TURUN KAUPPAKORKEAKOULUN JULKAISUJA

PUBLICATIONS OF THE TURKU SCHOOL
OF ECONOMICS AND BUSINESS ADMINISTRATION

Eija Koskivaara

Artificial Neural Networks for Analytical Review in Auditing

Sarja / Series Ae-7:2004 (pdf) A-7:2004 (nid.)
ISBN: 951-564-185-3 (pdf) 951-564-184-5 (nid.)
ISSN: 1459-4870 (pdf) 0357-4652 (nid.)

ACKNOWLEDGEMENTS

Orienteering in the research world is challenging and sometimes lonely. For me this journey has been fun as well as educational. I have received a lot of support from various people and organisations.

My supervisor, Professor Barbro Back has supported me in numerous ways. Her interests in my research and recommendations have been of critical importance to me. Furthermore, I acknowledge all the valuable advice on society of science and the encouragement she gave me, not only as a supervisor but also as a real friend.

Rector, Professor Tapio Reponen has offered me a lot of tacit and explicit knowledge to find the right way whenever I got lost. It was an honour for me that he accepted the invitation to act as an esteemed custodian in the public examination of my thesis.

I sincerely thank the official examiners of this study, Professor Teija Laitinen and Professor Ari Visa, for their time and excellent comments that helped me a lot. I specially want to thank the Director of Rutgers Accounting Research Center, Professor Miklos Vasarhelyi, who accepted the invitation to act as an esteemed opponent of my thesis.

I thank the researchers and administrative staff of Information System Science at the Turku School of Economics and Business Administration for creating a positive and helpful working environment. Special thanks go to the Head of the Department of Management, Professor Hannu Salmela and Professor Reima Suomi for their continuous support to my research work and to Secretary Birgit Haanmäki for taking care of many practical things. The libraries of Turku School of Economics and Business Administration and Åbo Akademi deserve thanks for excellent service.

I have the privilege to have many good advisors around me. Professor Christer Carlsson was very kind to let me take part in the postgraduate seminars of the Institute for Advanced Management Systems Research. Christer has the special ability to see both a single tree and the whole forest simultaneously in the information systems research. Professor Kaisa Sere´s enthusiastic grip especially at the beginning of my study was essential.

The Turku Centre for Computer Science (TUCS) has financed my doctoral studies. I thank all the professors, colleagues, administrative staff, and friends at TUCS for their efforts to create a financially safe and inspiring research environment. In addition, the Foundation for Commerce Education in Turku, the Finnish Konkordia Fund, and the Turku School of Economics and Business Administration believed that my study was worth financial support, which was also mentally important to me. I thank the people at the Data Mining and Knowledge Management Laboratory. Collaboration with Iulian Nastac, Ph.D., has deepened my understanding of the ANNs. Special thanks are reserved to my roommates. My first roommate, late Eva Wilppu, Lic.Sc., shared the ups and downs of the research process at the beginning of the study. I miss the intellectual discussions with Eva, and the moments when we almost laughed our heads off. Eva has a special place in my scientific heaven. With my second roommate Jonna Järveläinen, M.Sc., I had stimulating discussions in e.g. reasoning.

I like to thank all the organisations that have participated in this study. Mayor Armas Lahoniitty, Chief Audit Executive Terttu Para, Audit Executive Ritva Rindell, Audit Executive Esko Herrala, Planning Officer Pekka Paatonen, Head of Audit, Authorised Public Accountant Pekka Pajamo, and Financial Adviser Heikki Rautalin helped me to find and use different kinds of interesting data to build and evaluate the artefact. I thank Marko Gröönroos, B.Sc., for programming the final ANNA-system, and my cousin Pasi Lehtonen, M.Sc., for programming a data sorter for me when I was drowning in the data. Senior Lecturer Satu-Päivi Kantola merits thanks for guiding me through the statistical jungle. Many thanks go to Eeva Lilley, MA, for the language checking.

During these years I have participated in many interesting workshops, conferences, and courses in, e.g. Angers, Bergen, and Cork. This has given me a lot of inspiration that was needed to conduct this research.

I started my orienteering career by accident thirty years ago, and I have enjoyed the special spirit of this sport of ”cunning running”. It is easy to learn how to orienteer, but the challenges the sport provides are endless. To achieve the world championship requires lots of work. Needless to say, research is similar. Orienteering and orienteers have provided me with a safety network that every researcher needs for sharing the ups and downs of the research work – as well as for all those unforgettable moments inside and outside work. I thank my orienteering friends around the world. Special thanks are reserved to my orienteering soul-mate Sanna Nymaln. I am proud to have had the privilege of growing up in the spirit of the orienteering world.

I thank my friends outside the world of orienteering for being around. Ladies who have created a research career in the pressure of work and family have provided me with the knowledge I have needed to materialise this research. Other mothers have shown me how to cope amidst it all. “Girls of the Ring” have kept my feet firmly on the ground. I feel lucky to share a friendship with all of you.

I owe my thanks to my late grandmother, Saima Varjonen, who always encouraged me to get a good life and education. My parents, farmers Pirjo and Niilo Koskivaara, have shown me the “learning by working” method, which is also a good method in science. My parents and my parents-in-law, Elvi and Olavi Heinonen, have unselfishly offered their time to me and my family helping us in numerous ways. Nothing could be more important for the smooth flow of my research than knowing that my children were in the best possible hands when I was in the grip of the research. I hope that I can give similar care to my own children. My brother Jouni and his family, especially his wife Tiina, have offered their home and farm as a real research field to our children.

Ever since I entered into this thesis the fertile discussions with my dear husband, Olli Heinonen, MD, Ph.D., have been very valuable. From him I have also got the sound confidence that I can do whatever I want in life. Besides, he has also offered me numerous lunches to keep the research going. Our lovely children, Eero and Elina have given me a lot of sunshine and happiness. There is nothing in my life to be compared with the importance of their existence.

May 7th 2004

Eija Koskivaara

…
East is east,
west is west,
and now you ask me,
what is best.
…

The answer is *yes* or *no*, depending on the interpretation.
Albert Einstein, in *Scientific American*, April 1950.

TABLE OF CONTENTS

LIST OF PUBLICATIONS

1. Koskivaara (2004a) Artificial neural networks in analytical review procedures. *Managerial Auditing Journal,* 19 (2): 191–223.
2. Koskivaara (2002) Design Science Approaches in Information Systems Research. In: *Proceedings of the Information Systems Foundations: Building the Theoretical Base Workshop*, ed. by Shirley Gregor – Dennis Hart. Canberra, Australia, October 1–2, 2002, 205–216.
3. Koskivaara (2000a) Artificial neural network models for predicting patterns in auditing monthly balances. *Journal of the Operational Research Society,* 51 (9): 1060–1069.
4. Koskivaara (2000b) Different Pre-Processing Models for Financial Accounts when Using Neural Networks for Auditing. In: *Proceedings of the European Conference on Information Systems*, ed. by Hans Robert Hansen – Martin Bichler – Harald Mahrer. Vienna, Austria, July 3–5, 2000, 326–332.
5. Koskivaara – Back (2003) An Artificial Neural Network Based Decision Support System for Budgeting. In: *Proceedings of the Fifth International Conference on Enterprise Information Systems*, ed. by Oliver Camp – Joaquim Filipe – Slimane Hammoudi – Piattini Mario. Angers, France, April 23–26, 2003, 108–114.
6. Koskivaara (2004c) Comparing Different Methods in Analytical Procedures Using Real World Data. In: *Proceedings of the IRMA 2004, Information Resources Management Association*, ed. by Mehdi Khosrow-Pour. New Orleans Louisiana, USA, May 23–26, 2004, 661–663.
7. Koskivaara (2004b) Visualization of patterns in accounting data with self-organizing maps. In: *Business Intelligence Techniques: A Perspective from Accounting and Finance*, ed. by Murugan Anandarajan – Asokan Anandarajan – Cadambi Srinivasan, 133–148. Springer-Verlag: New York.

LIST OF ABBREVIATIONS

ACL	audit common language
AF	audit fee
AMEX	American Stock Exchange
ANN	artificial neural network
ANNA	ANN-based auditing system
AQT	advanced quantitative technique
AR	analytical review
AVE	average of previous years
BPN	backpropagation network
CLN	categorical learning network
CPA	certified public accountant
CRA	control risk assessment
CTM	combined trivial prediction
DELTA	average delta prediction
ES	expert system
FD	financial distress
GAAP	generally accepted auditing principles
GC	going concern
GRG	generalised reduced gradient
HS	health care services
IFAC	International Federation of Accountants
IIA	Institute of Internal Auditors
ISA	International Standard of Auditing
IT	information technology
NQT	non-quantitative technique
NYSE	New York Stock Exchange
OFA	office facilities administration
OLS	ordinary least squares
PNN	probabilistic neural network
PT	port of town
PYS	previous year's value
RMSE	root mean square error
RPROP	resilient backpropagation
SAS	Statement of Auditing Standard
SEC	Securities and Exchange Commission
SIAS	Statement on Internal Auditing Standards
SOM	self-organising map
SQT	simple quantitative technique
SW	social welfare
ZERO	zero delta prediction

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

The aim of my study is to find out whether the artificial neural network (ANN) technique is suitable for supporting analytical review (AR) procedures in auditing. The development of tools for auditors is important with regard to the workload of auditors. Furthermore, the recent events in the business and auditing environment have highlighted problems in the auditing process.

Auditors may need to restore public confidence in the capital market system and accounting profession which might have been shaken by the collapses of Enron and Arthur Andersen and many others. A well-known fact is that many companies have problems with their accounting systems. Siebel Systems, Qwest, WorldCom, and Xerox are examples of companies that have been “cooking the books”. We can call the year 2002 “Annus Horribilis” from an accounting point of view. Unfortunately the year 2002 was not an exception. This manipulation is still going on. A fairly recent example is the Dutch retail trade company Royal Ahold whose subsidiary in the USA manipulated the operating profits. In Italy Parmalat and in Sweden Skandia have joined “the cooking the books club”. There are also examples from earlier years. Two of them come from the banking world where the English Baring's bank and the Japanese Daiwa's bank lost millions of dollars because they did not have effective control systems. The third one is a Finnish multinational company whose subsidiary in Italy overestimated the work in progress and recorded fictitious sales. The above examples reflect one characteristic and the demands of the audit process[1].

A common fact is that many parties, such as shareholders, investors, creditors, tax authorities, and managers are interested in the accuracy of the financial performances of organisations. Auditors are in a key position to monitor and control operations in organisations. The increasing use of information technology (IT) and computers in organisations requires auditors to obtain and evaluate evidence electronically. Technology has made the input of information for transactions and events easier and the evaluation of the relevant events more critical. Companies report their financial outcome quarterly and an increasing number of companies present their financial information on a public network. Sometimes the speed at which these reports

[1] Indeed, dishonesty in business is an eternal and universal problem, see e.g. ”Dishonest Accountant” in Luke 16:1-13 (Bible 1992, 126) or Kaptah’s accounting methods in The Egyptian (Waltari 1945).

are made makes one wonder whether all the relevant information is audited and reliable. Therefore, *continuous auditing* is given more attention today (Searcy & Woodroof 2003; Vasarhelyi, Kogan & Alles 2002; Woodroof & Searcy 2001).

Indeed, continuous process and utilisation of computers and information technology in auditing is not a new idea (Kunkel 1974). In addition, researchers advocated for continuous auditing already before the big scandals, but their calls fell mostly on deaf ears (Groomer & Murthy 1989; Kogan, Sudit & Vasarhelyi 1999). According to Vasarhelyi, Kogan, and Alles (2002) in continuous audit, software continuously monitors transactions and compares their characteristics to expected results. Any significant inconsistency in expectations sends alarms to the company's managers and auditors. One possible way to develop continuous auditing is to develop better analytical review (AR) tools for continuous monitoring.

This chapter provides an overview of the positioning of my study. Firstly, I present challenges to the concept of gathering audit evidence. Secondly, I discuss the AR in auditing through a presentation of developed AR techniques and trends. Thirdly, I give examples of ANN business applications in research and practice. Fourthly, I discuss the use of information technology (IT) in auditing in general. Finally, I present the research questions and the outline of the thesis.

1.1 Gathering audit evidence

Auditors need tools and methods that provide them with objective information about a client company. Surveys of Willet and Page (1996) and Lee (2002) support this demand. Willet and Page (1996) found that auditors are often tempted to speed up testing by irregular methods. They reported three major reasons for this: 1) budget pressure (61%), 2) boring work (30%), and 3) unimportant work (41%). Only 22% of the respondents claimed never to have succumbed to the temptation to speed up testing by irregular methods. Reported irregular audit practices were: rejecting problematic items from samples (54%), testing fewer items than reported (27%), or accepting doubtful audit evidence (24%). Although their questions in the survey were indirect because of the sensitive nature of the research, the answers give some idea of irregular auditing. Lee (2002) found that a correlation between commercial concerns, time pressures, and junior staff's "irregular auditing" resulted in audit failures. Furthermore, Lee (2002) found that commercial concerns often lead to the adoption of new audit methods.

It is well known that in many cases the relationships between accounting firms and their clients have been very tight. Therefore, according to Bazerman, Loewenstein, and Moore (2002) even the most honest and careful auditors can by accident distort a company's real financial status and therefore mislead investors, authorities, or shareholders. Also the subjective nature of accounting does not make the auditing tasks any easier.

Fischer (1996) has classified the literature on gathering and evaluating audit evidence into three general segments. Firstly, much of the literature has focused on the auditor's decision-making (Bonner 1990; Bonner & Walker 1994). For example, the mismatch between knowledge and task may hinder auditors' ability to draw on previous experience when making conditional probability judgements and when allocating audit hours (Nelson, Libby & Bonner 1995). Secondly, a sizeable amount of literature has focused on the development and evaluation of new audit methods. Thirdly, a smaller set of publications is concerned with the use of new audit technologies (Lieb & Gillease 1996). My study falls into the second segment and aims at getting further evidence of ANNs' feasibilities in the area of AR in auditing.

1.2 Overview on analytical review in auditing

The demands in the auditing environment have led to the publication of several auditing standards in different countries. The US standard, SAS 56, on AR[2,3] has often in the literature been considered an authoritative statement to analytical review in the literature. It states as follows (AICPA 1988):

> *Analytical procedures involve comparisons of recorded amounts, or ratios developed from recorded amounts, to expectations developed by the auditor. The auditor develops such expectations by identifying and using plausible relationships that are reasonably expected to exist based on the auditor's understanding of the client and of the industry in which the client operates.*

In the UK the guidance on the application of AR during an audit is included in the UK standard SAS 410 (APB 1995). In Finland the guidance is included in the recommendations of the Finnish Institute of Authorised Public

---

[2] Statement of Auditing Standard (SAS) No. 56 was issued by the ASB (= Auditing Standards Board) in 1988.

[3] AR first appeared in the authoritative literature of the AICPA (= American Institute of Certified Public Accountants) in 1972 (Kinney & William 1980).

Accountants', i.e. SAS 520 (KHT-yhdistys 2003, 182–187)[4]. The International Standard on Auditing, ISA 520, is becoming the universal standard for the analytical procedures (IFAC 2003a, 320–325). The guidance states as follows (KHT-yhdistys 2003, 632; IFAC 2003a, 321):

> *Analytical procedures consist of the analyses of significant ratios and trends including the resulting investigation of fluctuations and relationships that are inconsistent with other relevant information or which deviate from predictable amounts.*

The emphasis in the SAS 56 definition is on expectations developed by the auditor. The emphasis in the SAS 410 and SAS 520 guidances is on the analysis of ratios and trends.

Professional standards suggest that AR can be used for two purposes – attention directing and test reducing (Lin, Fraser & Hatherly 2003). For attention directing, auditors use AR as a diagnostic tool to identify the likelihood of material error. For test reducing, AR is used as a substantive test to justify the increasing or decreasing of other audit procedures.

In the literature different terms are commonly used to describe analytical review in auditing such as analytical auditing, analytical review, analytical review procedures, analytical procedures and analytical evidence. In my thesis, I have used the analytical review as an umbrella term for different kinds of analytical procedures. The AR produces analytical evidence to auditors. An analysis of AR in auditing is presented in chapter three.

Various techniques or methods may be used in performing the analytical procedures. These techniques range from simple comparison to complex analyses (e.g. Leitch & Chen 2003; Blocher, Krull, Tashman & Yates 2002; Fleming 2004). For example, I use here an ANN technique for estimating account values in order to direct auditors' attention. The AR techniques and trends are discussed in the following sections.

1.2.1 Analytical review techniques

Auditing researchers have developed a variety of procedures to assist in the AR process. Blocher and Patterson (1996) have identified three types of AR techniques, these are *trend analysis, ratio analysis,* and *model-based* technique. Fraser, Hatherly, and Lin (1997) have provided a slightly broader

[4] AR appeared in the recommendations of the Finnish Institute of Authorised Public Accountants 1.7.2000 (KHT-yhdistys 2003, 7).

classification perspective for AR techniques: non-quantitative (NQT) or *judgemental,* such as scanning; simple quantitative (SQT), such as trend, ratio, and reasonableness tests; and advanced quantitative (AQT), such as regression analysis[5]. These techniques differ significantly in their ability to identify a potential misstatement. Judgemental techniques include the auditor's subjective evaluations based on client knowledge and past experience. Trend analysis assesses whether there is a functional relationship between the variables over time. Ratio analysis incorporates the relationships between two or more variables. For example, turnover ratios are useful because there is typically a stable relationship between sales and other financial statement accounts, such as receivables and inventory. Ratios are easy to compute, and therefore they are tempting, but their interpretation is problematic, especially when two or more ratios provide conflicting signals. Indeed, ratio analysis is often criticised on the grounds of subjectivity, i.e. the auditor must pick and choose ratios in order to assess the overall performance of a client. In a reasonableness test, the expected value is determined with the data partly or wholly independent of the accounting information system, and for that reason, evidence obtained through such a test may be more reliable than evidence gathered using only an accounting information system. For example, the reasonableness of the total annual revenue of a freight company may be estimated by calculating the total tons carried during the year and the average freight rate per ton. In the regression analysis model the auditor may predict financial and operating data with the help of e.g. economic and environmental data.

Sad to say, many of the AQTs have not found their way into practice. Most of the AR techniques used in practice for signalling errors throughout the audit are relatively simple models and are not based on any statistical methods (Ameen & Strawser 1994; Cho & Lew 2000; Fraser, Hatherly & Lin 1997; Lin, Fraser & Hatherly 2003; Turpiainen 1994). It is possible that practitioners consider methods in terms of cost effectiveness: the more simple AR techniques are the most cost effective (Fraser, Hatherly & Lin 1997). Another obstacle is the required expertise knowledge to use AQT methods. This difficulty might be related to the age of today's auditors. The fact is that the older auditors are not so used and trained to use IT. Tiittanen (1998, 126) found that younger auditors seem to have more positive attitudes towards the use of IT. At the same time, for example, in Finland the average age of CPA-auditors has increased. The average age of CPA-auditors was 42.9 in 1980,

---

[5] Compare to Fischer's (1996) literature classification of gathering and evaluating the audit evidence. In the literature the judgemental based procedures are often studied under the behavioural aspects, see e.g. O'Donnell (2002).

45.4 in 1990, and 48.6 in 2000. The average age of approved accountants in Finland was 55 in 2000[6].

Needless to say, there is a need for better methods and I argue that ANNs are a technique to aid auditors in creating expectations and these expectations can then be compared to actual values automatically (cf. ISA 520). ANNs have many beneficial aspects in comparison to other techniques. They are statistical models of real world systems which are built by tuning parameters (Swingler 1996, 3). Therefore, they are adaptive tools for processing data. Swingler (1996, 4) states that ANNs are mainly used in two types of task: classification and continuous numeric functions. They are useful for noisy data and they are able to dynamically adapt to a changing environment (Medsker & Liebowitz 1994, 27, 170). Basically, ANNs learn from examples and then generalise the learning to new observations. Compared to regression analysis we do not need an a priori model because ANNs are data driven models. ANNs are, unlike traditional statistical techniques, capable of identifying and simulating nonlinear relationships in the data without any a priori assumptions about the distribution properties of the data. This means that ANNs are assumption-free approaches for approximating functions from sample data. When you use an ANN there is a possibility that the network finds a model that is not relevant and learns according to that. I have overcome this problem by carefully testing my models with holdout samples.

One advantage of the ANN-systems could be that they provide additional information to the decision process. With the help of an ANN an auditor may find something from the data more effectively and efficiently than with conventional AR methods. Therefore ANNs are potentially suitable for many tasks within auditing. Furthermore, ANNs have been considered one of the emerging technologies (Halal, Kull & Leffmann, 1998). IT development and the processing capacities of PCs have made it possible to model ANN-based information systems for monitoring and controlling operations.

1.2.2 Research trends in advanced quantitative techniques

A sizeable amount of literature has focused on the development and evaluation of new audit methods. In the seventies and eighties researchers used different statistical methods such as trend-, ratio-, and regression-analysis, and univariate and bivariate Box-Jenkins time series analysis for AR purposes

[6] CPA = Certified Public Accountant. The CPA-auditors (in Finnish KHT-auditors) are authorised by the Central Chamber of Commerce and the Approved Accountants (in Finnish HTM-auditors) are authorised by the local Chambers of Commerce. The average ages were achieved from the Central Chamber of Commerce.

(Arrington, Hillison & Jensen 1984; Biggs, Mock & Watkins 1988; Duke, Neter & Leitch 1982; Frost & Tamura 1982; Garstka & Ohlson 1979; Gibbins 1982; Kaplan & Reckers 1989; Kinney 1978; Kinney & Salamon 1982; Kinney & Warren 1979; Knechel 1988; Knechel 1986; Lev 1980; Matsumura & Tsui 1982). These methods have appeared to perform rather well in identifying unusual fluctuations in financial statements that need detailed investigation. For example, a regression-based analytical review and the use of monthly data increased audit effectiveness, and regression-based models were very efficient in detecting potential material misstatements (Knechel 1988).

At the end of the eighties and in the nineties researchers focused on the development of expert systems (ESs) for auditors (Bharadwaj, Karan, Mahapatra, Murthy & Vinze 1994; Brown & Murphy 1990; Lieb & Gillease 1996; Marcella & Rauff 1995; McCarthy, Denna & Gal 1992; McKee 1992; Murphy & Brown 1992; Sutton, Young & McKenzie 1995; Van den Acker & Vanthienen 1995). Van den Acker and Vanthienen (1995) reported that in Belgium the diffusion of ES showed a rather low occurrence rate in audited companies. However, frequencies for firms preparing to use ESs were higher and frequencies of ES abandonment were the lowest. According to Van den Acker and Vanthienen (1995), the ES-active business sectors were banking and insurance. Lieb and Gillease (1996) reported on Du Pont's audit risk-based decision support system. Because of the use of that system the company said that they had been able to reduce their staff while maintaining fully effective audit coverage. The idea behind the system is that it selects portfolios based on the potential risk of not auditing a specific auditable unit. This risk is based on the collective judgement of the management. Ye and Johnson (1995) suggested that explanation facilities could make ES-generated advice more acceptable to users. Nonetheless, the construction of an expert system is time consuming and sometimes the knowledge of the human experts is not available.

Auditing ANN research started a little more than a decade ago. The main ANN-application areas in auditing are detecting material errors (Coakley & Brown 1991a; Coakley & Brown 1991b; Coakley & Brown 1993; Wu 1994; Coakley 1995; Busta & Weinberg 1998), detecting management fraud (Green & Choi 1997; Fanning & Cogger 1998; Feroz, Kwon, Pastena & Park 2000), and supporting going concern decision (Hansen, McDonald & Stice 1992; Lenard, Alam & Madey 1995; Koh & Tan 1999; Anandarajan & Anandarajan 1999; Etheridge, Sriram & Hsu 2000). ANNs have also been applied to internal control risk assessment (Davis, Massey & Lovell 1997; Ramamoorti, Bailey & Traver 1999), to determination of the audit fee (Curry & Peel 1998), and to financial distress problems (Fanning & Cogger 1994). Going concern and financial distress are very close or can even be included in bankruptcy studies. In my review I have focused on those ANN-applications that are conducted from an auditing perspective. My research is a continuation on the line of research that Coakley and Brown (1991a) started on. My research also includes the discussion of the visualisation of the accounting data. This part of the study resembles studies of Martín-del-Brío and Serrano-Cinca (1993), and Back, Sere, and Vanharanta (1998). Chapter five presents a thorough review of auditing ANN-applications.

1.3 ANN business applications in research and practice

In the light of overview articles, ANNs have been extensively studied in the last decade (Wong, Bodnovich & Selvi 1995; Wong & Selvi 1998; O'Leary 1998; Zhang, Patuwo & Hu 1998; Vellido, Lisboa & Vaughan 1999; Coakley & Brown 2000). ANNs have been applied for various corporate functional activities, such as production, operation, and finance.

Examples of research applications of ANNs for financial analysis are the studies by Martín-del-Brío and Serrano-Cinca (1993), Hill, O'Connor, and Remus (1996), and Back, Sere, and Vanharanta (1998). Martín-del-Brío and Serrano-Cinca (1993) studied the financial statements of the Spanish companies, and attempted to predict bankruptcies among Spanish banks during the 1977–85 banking crisis. Back, Sere, and Vanharanta (1998) compared the financial performance of 120 companies in the international pulp and paper industry. Hill, O'Connor, and Remus (1996) observed that neural networks did significantly better than traditional statistical and human judgemental methods when forecasting quarterly and monthly data in financial time-series.

A small number of companies are using or have used ANN applications to support their business. Credit-card companies use ANN technology to reveal fraudulent clients (Mulqueen 1996; Fryer 1996; Fisher 1999). Former KPMG Peat Marwick has developed an ANN for bankruptcy prediction (Etheridge & Brooks 1994). Probably because these applications contain business confidential information, the models are kept secret. However, KPMG Peat Marwick reported that their ANN models produced between 10% and 15% of its revenues in 1995 (Mulqueen 1996). Chapter five summarises the reports related to my study.

1.4 Information technology in auditing

Toiviainen (1991) has developed four stages of information technology utilisation in auditing[7]. Table 1.1 is modified from Toiviainen (1991, 83). At stage one, standard software applications are used. At stage two, some databases, email, and graphics are also adapted. At stage three, several different external and internal databases, audit software applications, and company models are in use. At stage four, expert systems, decision support systems, and special audit software for continuous auditing are utilised. ANN-based support tools fit into stage four or possibly into the next stage. At this – the fifth stage – the software applications are based on advanced methods like ANNs, fuzzy logic, and genetic algorithms. They could produce information for continuous monitoring and controlling. The utilisation stage of information technology varies among auditors and auditing firms. Tiittanen (1998, 128) investigated Finnish auditing firms in 1997 and found that many small auditing firms were at the first stage and big auditing firms were at the second or at the third stage and only very few were at the fourth stage according to Toiviainen's classification.

Table 1.1 The five stages in the development of IT utilisation

STAGE	SOFTWARE APPLICATION	UTILIZATION
I	Word processing, spreadsheets	Documentation, auditor's report, financial analysis, and calculations
II	Graphics, external databases, electronic mail	Audit planning, comparison of financial information, company analysis
III	Company models, audit databases, IS audit software applications	Testing of information systems, database inquires
IV	Expert systems, decision support systems	Expert analysis for finding important tasks for audit
V	Advanced methods: ANN-based systems, fuzzy systems, genetic algorithm based systems	Assurance services, continuous monitoring, and controlling, business intelligence information for users

Glower and Romney (1998) have provided an alternative classification perspective for the utilisation of information technology in auditing. They have categorised software packages into five groups according to audit areas: 1) data extraction and analysis, 2) fraud detection, 3) internal control evaluation, 4) electronic commerce control, 5) continuous monitoring. As mentioned earlier, ANN-based AR tools could work in a continuous monitoring context.

---

[7] International Auditing Practice Statement 1009 Computer-Assisted Audit Techniques (CAATs) provides guidance on the use of CAATs (IFAC 2003c).

According to the survey of Glower and Romney (1998), 24% of the participants indicated that they used continuous monitoring software. Figure 1.1 illustrates the products used for continuous monitoring. Almost one-third of the respondents used internally developed software that has been customised to specific firm needs. Respondents used continuous monitoring software primarily to monitor high risk transactions, produce exception reports, detect fraud, monitor inventory trends and turnover, monitor computer processing functions, such as confidential file access and password usage. The study of Glower and Romney (1998) does not reveal whether any ANN-based tools were in use. The difference between ACL[8] and similar tools and ANN is that ACL is a query-based tool for database analysis whereas ANNs learn from examples.

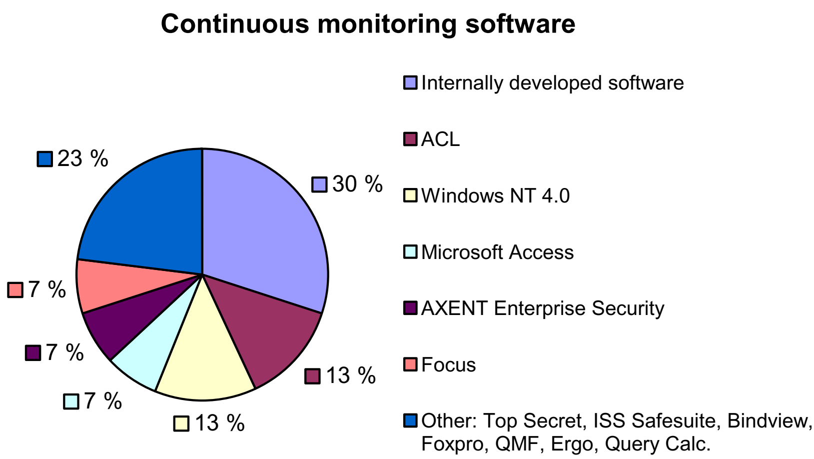


Figure 1.1 Proportion of continuous monitoring software (Glower & Romney 1998)

Glower and Romney (1998) say that a relatively large percentage of auditors have not adopted the support tools specialised in auditing. However, according to them the use of audit software is likely to become increasingly widespread among auditors as the use of advanced information systems becomes more prevalent among their clients. Therefore, IT is becoming an important enabler for the auditing process.

[8] ACL (audit common language) is a product of ACL Services Ltd., which provides business assurance solutions to financial executives and audit professionals.

1.5 Aim of the study

The focus in this thesis is on evaluating the feasibility of an ANN-based tool for supporting analytical review (AR) procedures in auditing. I will model, train, and test an ANN-based system that I have constructed (= an ANNA-system) for this problem domain.

The objective of the system is to form an expectation for monthly account values. Companies have hundreds of accounts and (in theory) they all should be audited. Therefore, a tool that supports the auditing of account values is welcomed.

The aim of this study can be broken down into the following research questions:

1. How to model an ANN-based monthly account predictor?
2. How to illustrate the monthly account models for auditing purposes?
3. How to construct the ANNA-system?
4. Which criteria should be used in evaluating the ANNA to model monthly account values?
5. How good is the ANNA-system based on these criteria?

To achieve the aim I selected and applied existing ANN techniques and data processing methods. This included five iterations of ANN-based models in the context of approximating function for account values. All the iterations were made with real world data.

1.6 Overview of the thesis

The introduction chapter positions my study to related publications. This means identification of the research scope and an overview of the research trends and techniques in the analytical review in auditing. It also outlines the research aim in the study.

The second chapter describes the research approach I have utilised in this research process[9]. The chapter presents research frameworks to design, build, and evaluate information systems. The design science approach is the glue, which ties up different parts of my thesis.

The third chapter gives an overview of AR in auditing. AR is a broad application domain for my research. AR is a way of thinking and behaving to

---

[9] It brings together some different views on design science approaches in the information systems research and is based on Koskivaara (2002).

achieve the audit evidence. In chapter three I propose the inclusion of AR tools into the continuous auditing environment.

The fourth chapter gives a brief description of ANNs. In particular, the chapter pays attention to learning algorithms and their feasibility to support my research problem. The ANN technique is the platform and the filter for the whole study. The ANN is the method I have applied in the AR contexts.

The fifth chapter provides a summary of related research in the field of ANN[10]. It goes through the applications domains of the researched articles. The review of the modelling issues pays attention to data and sample sets, ANN-architectures and learning parameters, and whether the model was evaluated with a holdout sample. The broad and extensive review of previous researches indicates that most of the authors state that the ANNs have the potential to improve AR and that there are some open ends in the field.

The sixth chapter presents an overview of the development of the ANN-based analytical review tool[11], ANNA. My research started by discovering the prediction model for the monthly account values. The proposed one-step-ahead model estimates the future revenues and expenses of the organisation based on past audited values. The monthly account values are regarded as a time-series and the target was to predict the current year's account values with the ANN. The prediction model was constructed with different input–output variables and learning parameters. My contribution here is that I show that an ANN-based system works in the monthly account value environment. Furthermore, the study revealed that the illustration of the results is a relevant problem[12]. Because the pre-processing has an effect on what the network learns and the successful data pre-processing might speed up ANNs learning, and give better and more general results, the prediction model was tested with different pre-processing models[13]. The contribution here is that it makes a difference as to how the data is pre-processed, i.e. scaling all data simultaneously gave the best results. Furthermore, the linear scaling has the advantage of preserving the relative position of each data point along the range.

The seventh chapter presents the results achieved with ANNA[14]. ANNA is a summary of all the earlier models. First, ANNA is presented as a decision

[10] The analysis is based on Koskivaara (2004a).
[11] The overview is based on Koskivaara, Back, and Sere (1996) and Koskivaara (1997, 2000a, 2000b, 2004b).
[12] The illustration of the results is also discussed in Koskivaara (2004b), where it is shown how Kohonen's self-organising map (SOM) can be used in the visualisation of complex accounting data. I used SOM for clustering ten years' monthly income statement values and showed that SOM can classify accounting data according to various accounts.
[13] See Koskivaara (2000b).
[14] The analysis is based on Koskivaara and Back (2003) and Koskivaara (2004c).

support system for budgeting, and a graph and a figure illustration of the results are given. The prediction values of ANNA are compared to the actual account values and to the account values budgeted by the organisation. The results of ANNA are evaluated with five conventional analytical review procedures. In each case ANNA predicts the best result with the lowest average error in money and the lowest standard deviation. Second, the evaluation of ANNA was continued with new data sets. It was tested whether there was a significant difference between the prediction accuracy in the monthly account values of ANNA and the conventional AR procedures. ANNA was significantly better than the other methods in most of the cases. The contribution here lies in that I have shown that ANN can be used for forming expectations for monthly account values. By illustrating the results and comparing them to the conventional AR procedures I have shown that in most of the cases the constructed ANNA-system made better and more consistent predictions than the comparison methods. While subject to limitations, such as ANNA was built for a specific situation of AR, training data sets were small, the best results were achieved when the management of the organisation selected the accounts, and every model needs to be trained individually, the results of the study have implications for the future. The chapter ends with the proposition for using this kind of system for directing attention.

The eighth chapter summarises the findings and outlines the recommendations for further research. Figure 1.2 illustrates the structure of the thesis and the relationships between the chapters and publications.

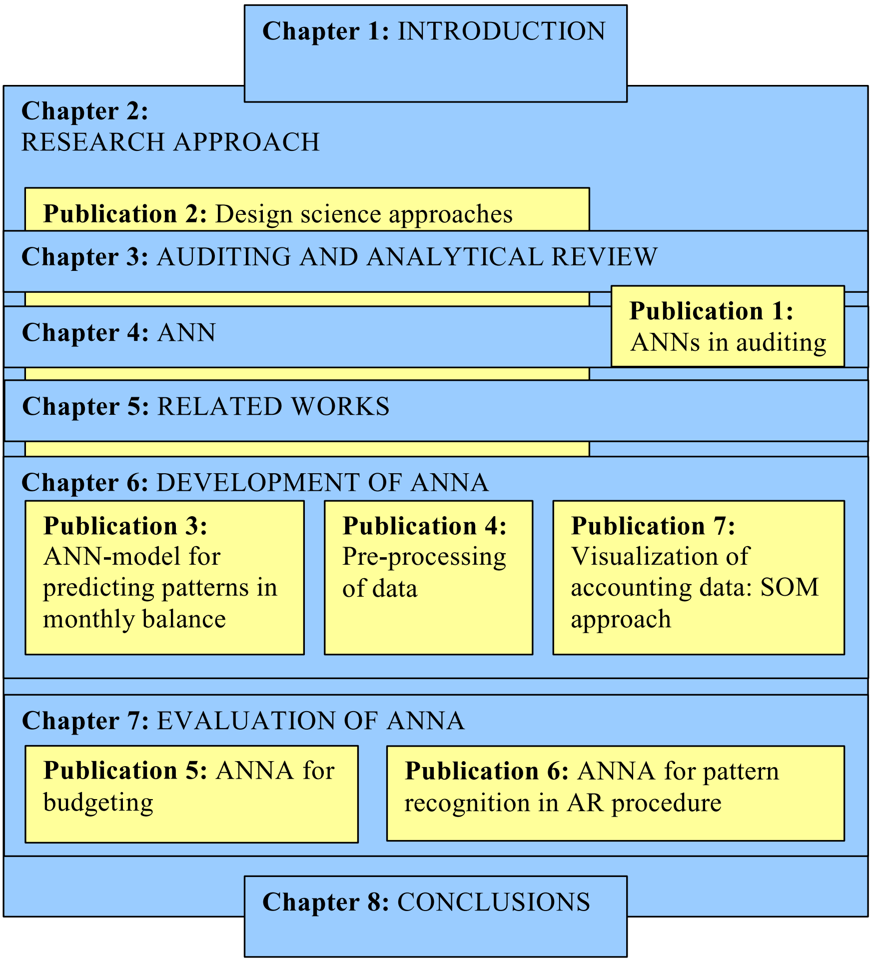


Figure 1.2 Interrelationships between the chapters and publications in the thesis

2 RESEARCH APPROACH

This chapter discusses and describes my research approach. I have used elements from different research approaches i.e. *the constructive research approach, the design science approach, the system development approach,* and *the development cycle approach*. I will describe them and state which parts I have used of each of them to achieve a coherent basis for my research. I consider my research approach an evolutionary research approach. It resembles the *spiral* model of Boehm (1988).

2.1 Approaches of designing artefacts

The constructive research approach has been applied e.g. in technical sciences, information systems, and operations research. Kasanen, Lukka, and Siitonen (1993) have introduced it into the field of management accounting research. According to Lukka (2002) the constructive research approach is a research procedure for producing innovative constructions, intended to solve problems faced in the real world. Lukka (2002) says that a construction can be such as models, diagrams, organisation structures, commercial products, and information systems designs.

March and Smith (1995) make a distinction between design science and natural science, and argue that both design science and natural science activities are needed to ensure that IS research is both relevant and effective. They refer to Simon's (1969, 55) work “The Sciences of the Artificial”, where he uses the expression “science of design” to refer to applied research. Science of design implies that this science is concerned with how to make artifacts that have desired properties and design skills. Thus design theory has two aspects: one dealing with the process of design and the other dealing with the product.

Nunamaker, Chen, and Purdin (1991) have introduced the systems development methodology which fits into the category of applied science and belongs to the engineering type of research. In the engineering approach, the artistry of design and the spirit of “making something work” are also essential. The constructive research approach has similarities to the system development methodology (Kasanen, Lukka & Siitonen 1993; Lukka 2002).

Medsker and Liebowitz (1994, 189) have presented a development cycle methodology for designing ANN systems. Boehm (1988) has introduced a spiral model of software development and improvement.

All these approaches are about one particular type of knowledge in IS, namely "theory of design". The research approach used in this study is a mixture of the above approaches. The spirit of making something work refers to pragmatism which is the philosophical background theory of this kind of research. Pragmatism emerged in the USA by Charles Sanders Pierce (1839–1914), William James (1848–1910), and John Dewey (1859–1952). Philosophical pragmatists deny the correspondence notion of truth proposing that truth essentially is what works in practice (see e.g. Niiniluoto 1986, 40, 52).

2.2 Constructive research approach

Lukka (2002) divides the constructive research process into the following phases:

1. *Find* a practically relevant problem. In my case the relevant problem is the workload of auditors in conjunction with the speed at which the financial reports are made and audited as well as the form of accounting data i.e. electronic. It is obvious that auditors need tools to support their work in this rapid business environment. They also need tools which either prevent or make the manipulation of accounting data more difficult by revealing the misstatements in the data efficiency. During the research process "cooking the books" -news in the newspaper became a more current theme than it was at the beginning of the research. Besides, the analytical review tools in auditing seemed an underanalyzed area at least in the Finnish auditing literature.

2. *Examine* the potential for long-term research co-operation. It was obvious from the beginning that I needed an organisation to co-operate with because neural networks are data driven models and in general companies and organisations own the data. I received the material to build and evaluate the ANN-based system from the organisation involved. At the beginning of the system building the data was received through personal relationships. Later on I had the benefit of working with personnel from a large organisation (City of Turku). In the evaluation phases the mayor gave permission for important data entries, which was recorded in the minutes of the city. This permission allowed access to data with the help of the internal auditors. I also received the right to publish the results of the project.

3. *Obtain* a deep understanding of the topic area. The application domain of the research, such as auditing and analytical review understanding, was mainly achieved from prior research, studies, and literature. The neural network understanding was achieved both from the literature and from building and testing different kinds of neural network models and applications.

4. *Innovate* a solution idea and develop a problem solving construction. The solution idea consists of building an ANN-system based on data which already exist in companies, and which is not properly used in the auditing context. Many companies collect reports on monthly basis to control their operations, however, many times the data in these reports is not properly used in the auditing process. The solution idea involved finding a suitable prediction model that could form the inner core in the system. A suitable data prediction model to create training sets for a chronological prediction ANN-system was found from the ANN-literature, and the idea of the ANNA-system was born. The neural networks would learn the relationships between the account values and the system could be used for predicting future values. If the prediction value and the actual value are within a previously stated threshold value, it is an indication for the auditor that those account values seem to be correct. It also involved studying how the data should be pre-processed before inputting it into the model's transforming process. Therefore, the prediction model was also applied to financial statements to study the effects of different pre-processing methods on the ANN results. Furthermore, it involved studying how to illustrate the output of the ANN-model. The prediction model was constructed on a five different input variable selection basis: literature, CPA-auditor, the most stable accounts compared to previous year's values, internal auditors, and management of the organisation. Another solution attempt was to let the SOM cluster the data in order to reveal patterns in the accounting data.

5. *Implement* and *test* the solution. The final ANNA-model is a prototype system. This is in line with the opinion of Kasanen, Lukka, and Siitonen (1993) and Järvinen (1999, 60). They pointed out that the output of a constructive research project may not necessarily be a fully developed system, but it could be a pilot system which shows how a system should be built. ANNA was tested with real operating data from the organisation. The empirical feasibility was tested by comparing the results of ANNA with conventional analytical review methods and by showing the prototype for potential users. ANNA predicted in each case the best results with the lowest average error in money and the lowest standard deviation. The difference was significant between the ANN and the other methods in most of the cases where the management of the organisation selected the accounts.

6. *Ponder* the scope of applicability of the solution. The results achieved indicated that this kind of system could offer a promising supplement approach to AR procedures used nowadays. Firstly, this kind of system can provide auditors with objective information about a client company. Therefore, it can also prove to be a useful tool when an auditor discusses problems with the clients and recommends corrections in the financial account values. Secondly, the auditor could focus substantive testing with the help of this kind of system. Thirdly, this kind of an ANN-system could be implemented in continuous monitoring and controlling. It could give a report, for example, once a month about those accounts that are within a certain threshold value. Then the auditor could decide whether and what kind of further audit with these accounts would be needed.

7. *Identify* and *analyse* the theoretical contribution. In this research an ANN-based prototype for analytical review of monthly account values was built and evaluated. In the development of the ANN-system I used different research approaches to achieve a sound basis for the development process. The evaluation criteria for the ANN-system were derived from the analytical procedures. The results of my study are in line with related research: although ANN models cannot entirely replace professional judgement, they offer a promising alternative approach to AR procedures used today. Indeed, the best results were achieved when the management of the organisation selected the input variables. This is not surprising. The management should know which accounts follow the trends best and are related to each other. However, from an auditor's point of view this result is doubtful because auditors have the owners' mandate to control also the management of the organisation.

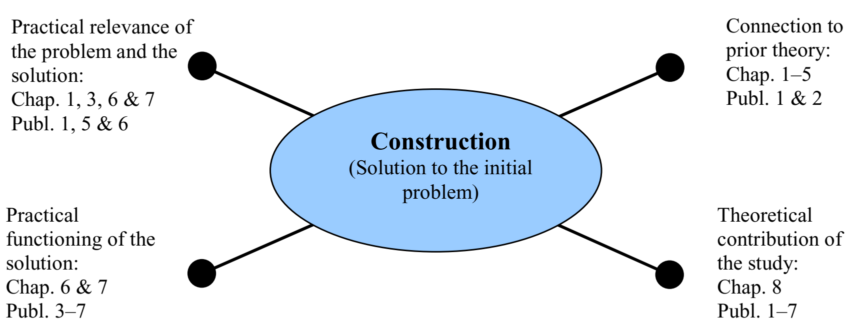


Figure 2.1 The interrelationships between the central elements of the constructive research approach and the thesis (adapted from Lukka 2002)

The key elements and the phases of the constructive research process turned out to be good general guidelines during the whole research process. Figure

2.1 illustrates the interrelationships between the key elements of the constructive research approach and the chapters and publications[15] of the thesis.

2.3 Design science research approach

According to March and Smith (1995), the design science[16] is knowledge using activity which produces and applies knowledge of tasks or situations in order to create an effective artifact. Design science research in information technology aims at improving performance. It is a technology oriented attempt to create things that serve human goals. Its products are assessed against criteria of value or utility: does it work and is it an improvement? In the designing of ANN models this means that the existing technology is applied to various situations to research whether it works and whether any improvement occurs.

March and Smith (1995) make the distinction between research activities and research outputs (Figure 2.2). The vertical dimension in their research framework is based on design science outputs or artifacts. The horizontal dimension in the framework distinguishes different research activities. In Figure 2.2 I classify different parts of my research into the framework of March and Smith (1995).

		Research Activities			
		Build	Evaluate	Theorise	Justify
	Constructs	Chap. 1			
Research	Model	Chap. 6 Publ. 3, 4, 7	Chap. 7 Publ. 5, 6	Chap. 8	
Outputs	Method	Chap. 3–5 Publ. 3, 4, 7	Chap. 5, 7 Publ. 1	Chap. 2, 5 Publ. 1, 2	
	Instantiations				

Figure 2.2 Relationship between the research framework and the chapters (Chap.) and the publications (Publ.)[15] in the thesis (adapted from March and Smith 1995)

[15] Publ. 1 = Koskivaara (2004a), Publ. 2 = Koskivaara (2002), Publ. 3 = Koskivaara (2000a), Publ. 4 = Koskivaara (2000b), Publ. 5 = Koskivaara and Back (2003), Publ. 6 = Koskivaara (2004c), Publ. 7 = Koskivaara (2004b), or see the list of the publications at the beginning of the thesis.

[16] Design science in information systems research is described also in Hevner, March, and Park (2004).

According to March and Smith (1995), there are four types of research outputs: constructs, models, methods, and instantiations. Constructs help to describe a research problem and to specify its solutions. I have described the research problem in chapter one. Models can be viewed as a description of how things are. In chapter six I have developed the model for the system[17]. Methods are sets of steps used to perform a task. In my research I have combined two different kinds of methods, namely the ANN and AR methods to solve a problem. The ANN method and how it could be used and is proposed to be used in AR is discussed more thoroughly in chapters three, four, and five. My contribution of using ANN in this field is discussed in chapters six and seven[18]. Instantiations demonstrate the feasibility and effectiveness of the models and methods they contain. Instantiations concern the efficiency and effectiveness of the artifact and its impacts on the environment and its users. The system I have created is not an instantiation that is in use; rather it is an instantiation in the focus of a research prototype. Therefore, instantiations are beyond my research scope.

March and Smith (1995) say that building and evaluating IT artifacts have design science intent, and theorising and justifying have natural science intent. Build refers to the construction of the artifact, demonstrating that such an artifact can be constructed. Chapter six in this thesis fits in the build column[19]. Evaluate refers to the development of criteria and the assessment of artifact performance against those criteria. Chapter seven in this thesis fits in the evaluation column. Chapter six introduces the assessment criteria for my prototype and in chapter seven the ANN results are compared with these conventional methods to see whether any improvement has occurred. The evaluation criteria are based on analytical procedures and this topic is discussed more thoroughly in sections 3.2.1 and 6.7.4. In the evaluation phase I also presented the prototype to possible users in order to get their feedback. Once the research output is built and evaluated there is a space for theorising and justifying. Chapter two brings together some different views on design science approaches in the IS research, and therefore its contribution suits in the theorise column[20]. Chapter five evaluates and theoretises the ANN-research done in the auditing field and is therefore suited to the evaluate and theorise columns. Chapter eight identifies and analyses the theoretical

[17] Chapter six is based on Koskivaara (2000a, 2000b, 2004b) = Publ. 3, 4, 7.
[18] This discussion in based on Koskivaara (2000a, 2000b, 2004c) = Publ. 3, 4, 6, and Koskivaara and Back (2003) = Publ. 5.
[19] As in Koskivaara (2000a, 2000b, 2004c) all present both model and method used in the empirical works, they are located in both of these boxes in the build column.
[20] Chapter two is based on Koskivaara (2002) = Publ. 2.

contribution of the thesis and therefore it also fits in the theorise column in the framework.

2.4 Systems development research approach

Nunamaker, Chen, and Purdin (1991) propose a multimethodological approach to IS research that consists of four research strategies: *theory building, experimentation, observation,* and *systems development*. Theory building includes the development of new ideas and concepts, frameworks, new methods, or models. Experimentation includes research strategies such as laboratory and field experiments, and computer simulations. Observation includes research methodologies such as case studies, field studies, and surveys. Systems development provides the exploration and synthesis of available technologies that produce the artifact (system) that is central to this process. My research approach mainly fits into the systems development research strategy, although my empirical studies also included building of the model and performing tests. This observation is in line with Järvinen and Järvinen (2000, 104) argument, in which they say that all the above research strategies can be applied to study systems development, but we cannot categorise them as subapproaches for the system development as Nunamaker, Chen, and Purdin (1991) suggest. The following discussion is based on the different stages of the system development research strategy.

According to Nunamaker, Chen, and Purdin (1991), a system development research consists of five sequential stages: *concept design, constructing the architecture of the system, prototyping, product development,* and *technology transfer*. Concept design means adaptation of technological and theoretic advances into potentially practical applications. System architecture provides a road map for the systems building process. The concept design and architecture construction stages resemble phases one, three, and four in the constructive research approach (see section 2.2 or Lukka 2002). Prototyping is used as a proof-of-concept to demonstrate feasibility. The prototyping includes stages 5–7 of the constructive research approach. The product development and the technology transfer stages are beyond my research. These latter phases of the system development approach might be hard to carry out in a public research community because these systems might contain business confidential information.

Figure 2.3 outlines the accumulative systems development research process and research issues together with the interconnection to the thesis. The development research starts by making assumptions about the research domain and the technical environment for developing systems (e.g. see section 6.1). In this research the ANN-based support system is proposed for analytical review and for continuous monitoring and controlling context. Design involves the understanding of the domain under study, the application of relevant scientific and technical knowledge, the creation of various alternatives, and the synthesis and evaluation of proposed alternative solutions. In designing the ANN building process the developers have to be aware of the learning paradigms, needed data, and how to evaluate the ANN's output. In this study different ANN-models were designed and built. One model was selected for prototyping and for further evaluation. This selection was based on learning through the design process. The design also included the selection of the comparison methods, which allowed me to test the prototype's feasibility.

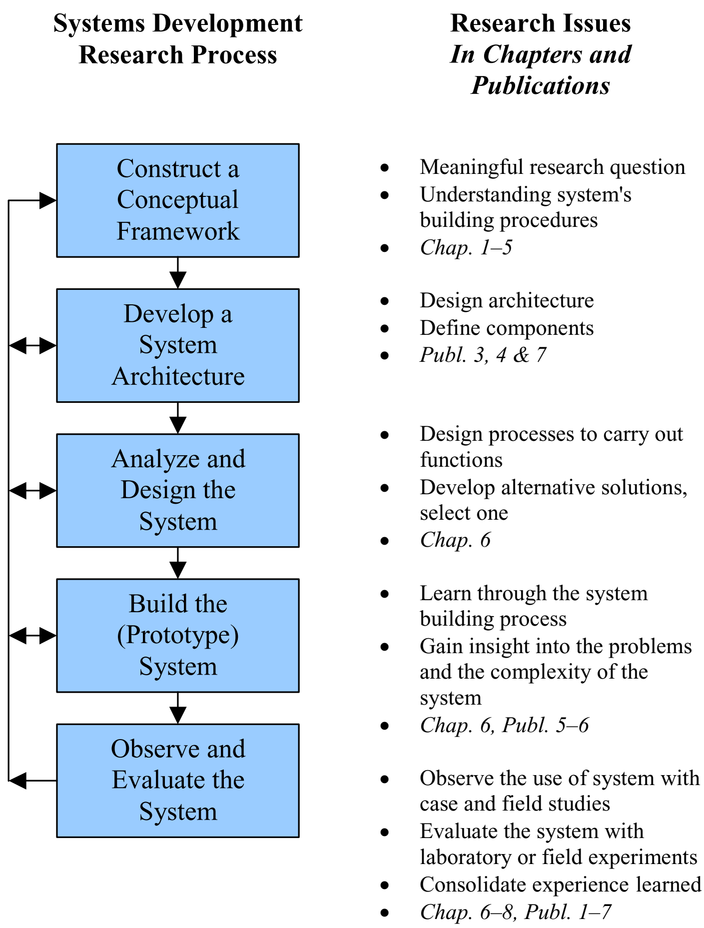


Figure 2.3 The interconnection between the process of systems development research and the thesis (adapted from Nunamaker, Chen & Purdin 1991)

The system development research approach provided me with good general elements and stages in the development of IS. Furthermore, like the design science research approach, it links this research to the IS research family.

2.5 Development cycle research approach

Medsker and Liebowitz (1994, 190–191) suggest that the development cycle of an ANN includes four major phases: *concept*, *design*, *implementation*, and *maintenance*. The phases included in my study are in *italics* in Figure 2.4.

The concept phase focuses on an application and paradigm selection. The point here is to focus on the applicability of neural networks as an appropriate technology to solve the problem at hand. Chapters one, three, four, five, and Koskivaara (2004a) focus on the demands of this applicability. Section 4.3 discusses more thoroughly the selection of learning the paradigm, i.e. unsupervised, supervised, or reinforcement learning.

At the design phase, the designs of a system, the gathering of data, and the selection of the development environment are decided. ANN system design means taking into account such things as the selection of a learning algorithm and activation function, and the designing of an ANN architecture. These things are discussed more thoroughly in chapter four. The last part of the design phase involves matching the designed model to an appropriate development environment. The design phase is discussed in chapter six[21].

In the implementation phase the ANN takes advantage of the data gathered for training and testing. The training and testing set size depends on the problem domain and on data. The training sample is used to determine values for the weights and parameters. The test sample is adopted for evaluating the model. Sometimes a third one called the validation sample is utilised to avoid the over fitting problem or to determine the termination point of the training process. Occasionally, only one set is used for both validation and testing purposes particularly with small data sets. This is true also in my case. I used only one set both for validation and testing purposes because my data sets were relatively small. During the iterative training process the ANN modifies the weights for the entire neural network. As the weights are usually difficult to interpret, the black-box testing is the primary approach for verifying and validating that inputs produce the appropriate outputs[22]. There are no clear guidelines on the selection of the performance measure of ANNs. The most

[21] Chapter six is based on Koskivaara (2000a, 2000b, 2004b).

[22] In traditional software development verification refers to the demonstration of consistency, completeness, and correctness of the system, and validation refers to the determination of the correctness of the final product with respect to the user needs and requirements.

commonly used approach is trial-and-error testing (Medsker & Liebowitz 1994, 206). The goal is that, when the training is over, the network has learnt an appropriate set of weights which allows the network to carry out the desired task (Rumelhart, Widrow & Lehr 1994). The implementation phase of the prototype is discussed in chapter seven[23].

The maintenance phase starts when the implementation is finished and continues during the application's entire lifetime. It evaluates periodically a system's performance and modifies it when necessary. Ongoing monitoring and feedback to the developers are required for system improvements and long-term success. The maintenance phase is outside my research.

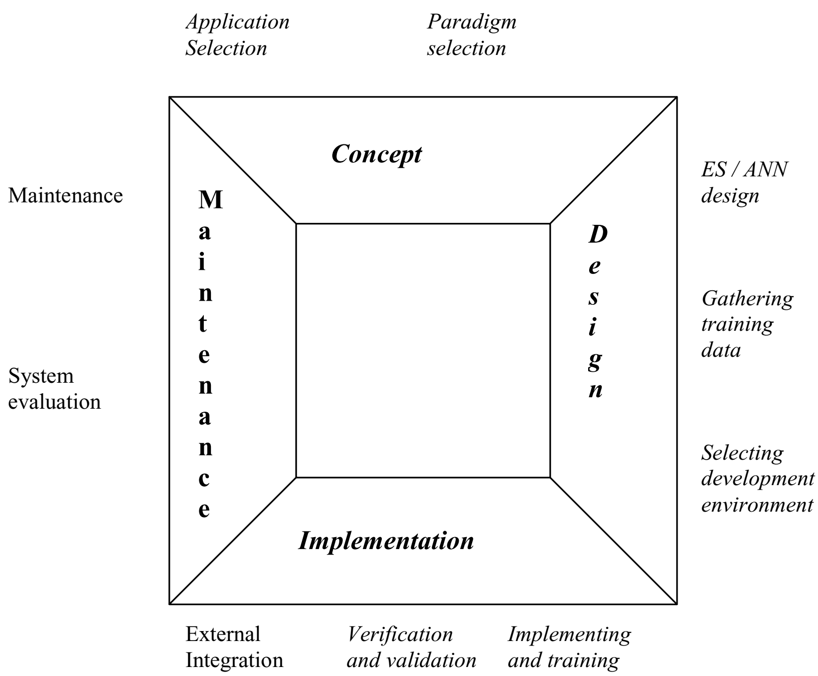


Figure 2.4 ANN development cycle (adapted from Medsker and Liebowitz 1994, 189)

Although Medsker and Liebowitz (1994) have included some unique steps and considerations in the development cycle of ANNs it still resembles a sequential process of conventional information systems design. One

[23] Chapter seven is based on Koskivaara and Back (2003) and Koskivaara (2004c).

explanation for this might be that they have tried to develop one joint design and development cycle for both the expert system (ES) and artificial neural network (ANN). This is based on their argument that the processes for developing expert systems and neural networks have important parallels. However, the idea of training a neural network system to carry out an information processing function is something very special and unique in comparison to i.e. the programming of rule and knowledge bases for expert systems.

2.6 Evolutionary research approach

All these approaches mentioned above have many good guidelines to facilitate the design and development of ANN systems. The constructive research approach gave me the general guidelines with its seven phases. The design science research approach provided me with the basic for research questions and research activities such as build and evaluate or train and test in the ANN environment. The systems development approach presented me the accumulative research process of the information system. The development cycle approaches gave me the phases of the ANN development process. The disadvantage of many development approaches is that they suppose an incremental or a sequential and quite enormous developments process whereas my development process tends to run through the several iterations. To describe my research process I use the spiral model[24] (Boehm 1988) in which designing, data acquiring, data processing, training, and testing cycles are repeated until the system meets an adequate solutions (Figure 2.5).

Finding a network structure and learning parameters is a search problem where earlier research processes have influences on the later processes. The research process enriches the understanding of the research problem, the design process, and the construction of the ANN-system. Indeed, the whole research process and the results achieved have helped me to identify and analyse the contribution of the thesis. Chapters six and seven present the iterative development process more thoroughly.

[24] Many software development approaches such as product-life-cycle have turned out to be too time-consuming and therefore faster approaches such as waterfall and spiral models have been developed.

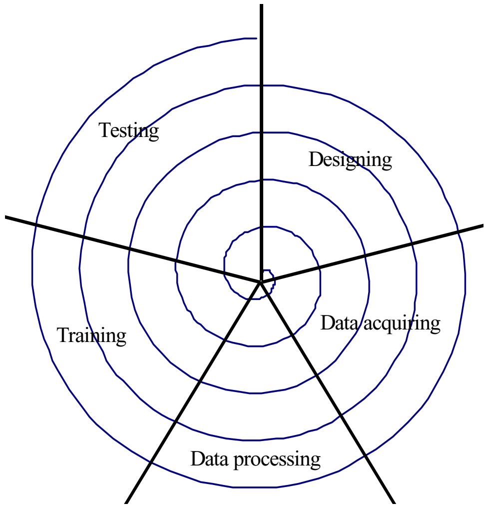


Figure 2.5 Spiral model

The ANNA-system has evolved via five steps or iterations. The first step was to model and test a suitable prediction model with the input variable based on literature. This is presented in section 6.2 (ANNA.01). The second step was to test the prediction model with the variables selected with the help of a CPA-auditor. This is discussed in section 6.3 (ANNA.02). The third step was to apply and test new features of the ANN to the prediction model and to decide how to illustrate the monthly account values for an analytical review purpose. This is discussed in section 6.4 (ANNA.03). The fourth step was to find a suitable pre-processing method for the account values. This is discussed in section 6.5 (ANNA.04). The next step was a side step or a replication of the third step in the development of ANNA, because in this step I explore how the SOM-method could be used in an analytical review context. This is discussed in section 6.6 (SOM-model). The fifth step was to construct an ANNA-system and test it with budget and real data. In the fifth step the most stable accounts compared to previous year's values were selected, also internal auditors selected accounts to test the system. This is discussed in sections 6.7 and 7.1. The fifth step was repeated and ANNA was tested with accounts selected by the management. This is discussed in section 7.2.

3 AUDITING AND ANALYTICAL REVIEW

This chapter overviews the audit environment and the role of analytical review (AR) in it. The audit environment is discussed in the view of audit phases. The role of AR in auditing is to obtain audit evidence. AR is supported with analytical procedures in which different kinds of techniques might be embedded. The chapter concludes with proposing continuous auditing and monitoring as a possible application domain for ANN-based AR tool for gathering audit evidence.

3.1 Audit phases with analytical review

Auditing is a part of the control process in organisations. In Finland, the auditor's report covers the accounting, the complete set of financial statements, and the corporate governance of the company (Finnish Audit Act 16 § 2). One purpose of auditing is to examine and observe the reliability of the accounts and that they give a true and fair view of the company's result of operations and financial position. An auditor should be experienced and independent. An engagement letter documents and confirms the auditor's acceptance of the appointment, the objective and scope of the audit, the extent of the auditor's responsibilities to the client, and the form of any reports. The engagement letter is the basis for the approach for auditing, e.g. standard or specific audit. The engagement letter finishes the client acceptance/retention phase. The audit process can be divided into three phases: *planning*, *testing*, and *overall review*. Figure 3.1 illustrates the audit phases with AR.

The planning phase includes the collection of information from various information sources, although the primary information source is the company itself. Annual reports, minutes, and financial statements, and interviews with management are natural information sources. In the planning phase AR is one way of collecting information for auditing purposes. For example, auditors may identify specific audit tasks with the help of AR. The planning phase includes the investigation of internal control systems. Despite the variations of the internal controls in different situations, all the internal controls will hold one inviolable principle: the separation of functions. Auditors may identify internal control risks with the help of AR. If an auditor wishes to rely on any internal controls, they should make sure they evaluate those controls.

Besides assessing risks of internal controls auditors have to assess risks related to the nature of the business in the planning phase. Auditors need to understand the company's nature of business, its organisation, its method of operating, and the industry in which it is involved. Thereby, they are able to estimate which events and transactions have a significant effect on the financial statements or which accounts might be the most vulnerable to manipulation. For example, in KPMG's audit approach, understanding the client's business strategy is considered to be the first step in assessing the adequacy and effectiveness of internal controls and in forming expectations regarding financial statement balances (Bell, Mars, Solomon & Thomas 1997, 1). The planning phase produces a plan for fieldwork, i.e. the testing phase which also depends on the audit approach, i.e. letters of engagement.

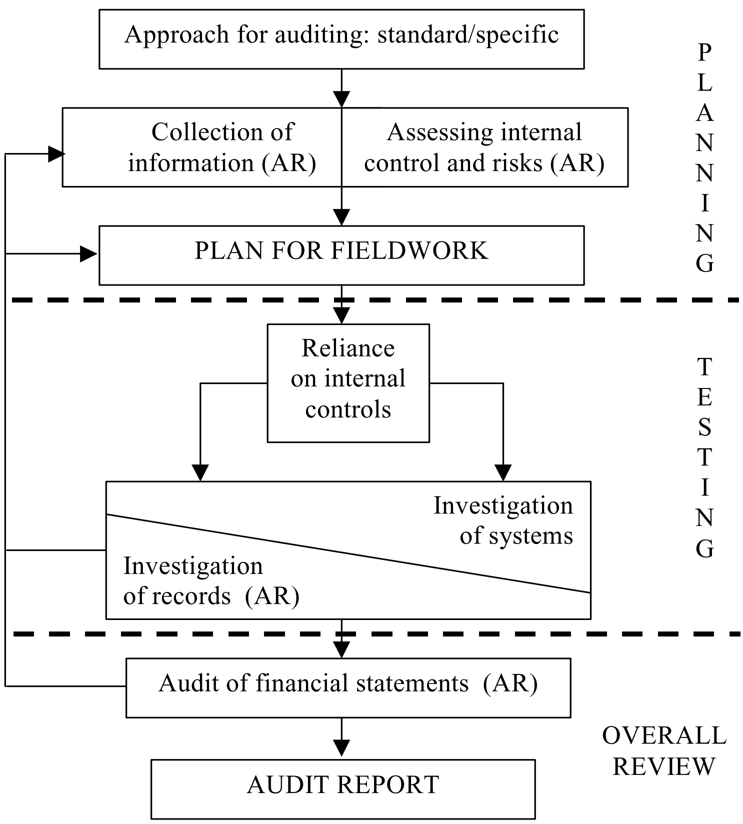


Figure 3.1 Audit phases with AR (modified from Riistama 2000, 74)

Reliance on a client's internal control system is the basis for the fieldwork. The bigger the reliance on the internal controls is, the more audit work may focus on information systems and internal control systems (Koch 1981). In a case where the internal controls are rudimentary, auditors will concentrate their efforts on investigation of the records and on examination of all available supportive audit evidence in the testing phase. They may obtain audit evidence as follows (IFAC 2003b): *investigation*, *observation*, *enquiry, computation, analytical review,* and *confirmation*. Investigation means examining records, documents, or tangible assets. Observation consists of looking at processes being performed by others. Enquiry involves seeking information of knowledgeable persons inside or outside the organisation. Computation consists of checking the arithmetical accuracy of source documents and accounting records or performing calculations. Analytical review consists of the analysis of significant ratios and trends including the investigation of fluctuations and relationships that are inconsistent with other relevant information or which deviate from expectations. Confirmation involves verifying the obtained evidence with e.g. the accounting records. The results of the testing phase are utilised to check whether the objectives of the audit plan are achieved.

The audit work ends at issuing the audit report. In forming the audit opinion for reporting purposes, i.e. whether the accounts present a true and fair view (e.g. Finnish Audit Act 19 § 2), the results of both formal and informal procedures must be once again carefully considered. In the overall review phase auditors may facilitate the review of financial statements with the help of AR. Obtaining the required knowledge of the business is a continuous and cumulative process of gathering and assessing the information. For example, although which information is to be gathered is decided at the planning phase, it is usually refined and added to at the later phases of the audit as the auditors and assistants learn more about the business.

3.2 Analytical review

AR focuses on the balance or summary of transactions rather than the components of the balance (Kinney 1978). AR involves estimation of expected amount for an account and comparison of that expectation to the actual book amount or AR involves comparing expected relationships among data items to actual observed relationships. Significant differences between the two figures are investigated. AR can be a relatively effective means of increasing the auditor's confidence in the validity and reasonableness of reported account values. AR can be used for directing attention and reducing tests (Lin, Fraser & Hatherly 2003). It assists the auditors in determining the nature, timing, and extent of their substantive testing (Weber & Goldstein 1999).

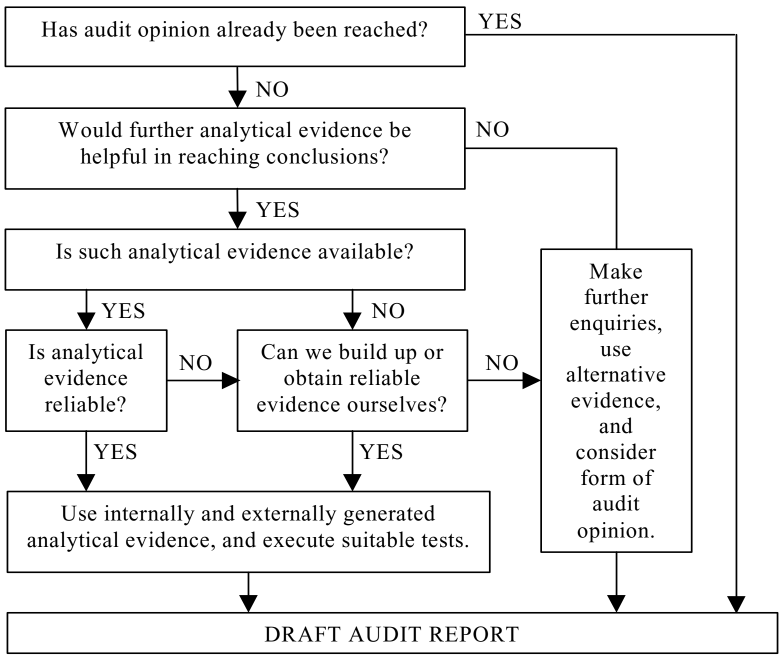


Figure 3.2 Flowchart framework of analytical review (modified from Woolf 1994, 268)

Figure 3.2 shows the AR as a decision-type flow diagram. The flowchart is modified from Woolf (1994, 268) and it suggests that the existence of available analytical evidence and the creation or procurements of analytical evidence by the auditors themselves may be embodied in the report. The flowchart suggests that the existence of available analytical evidence and the creation or procurement of analytical evidence may be viewed as alternatives. In practice, where both are suitable and usable, the auditor will make use of both sources of evidence. The flowchart makes a difference between internally and externally generated analytical evidence. The internal evidence is created within the client organisation. The auditor generates the external evidence. The exact form of such evidence will vary from one organisation to another.

Several different terms are commonly used to describe analytical review in auditing such as analytical auditing, analytical procedures, analytical review, analytical evidence, and analytical review procedures. I have used the analytical review as an umbrella term for different kinds of analytical procedures. AR may be performed for example:

- In the client acceptance/retention phase in order to settle the audit fee.
- In the planning phase to identify potential problem areas.
- In the testing phase to get evidence on account balances or transactions.
- In the overall review phase to gather evidence on the reliability of the financial statements with the auditor's knowledge of the business.

3.2.1 Analytical procedures

An auditor might use different kinds of analytical procedures to become convinced of the reliability of the audit evidence. These procedures may include (Waddington, Moreland & Lillie 2001; Gauntt & Gletzen 1997; SAS 56; ISA 520): comparison of current information with similar information for prior periods; comparison of current information with budget[25] or forecast or expectations of the auditor; study of relationships of financial information with the appropriate non-financial information; study of relationships among elements of information; comparison of information with similar information for other organisational units; comparison of information with similar information for the industry in which the organisation operates.

Analytical procedures involve the examination of the accuracy of account balances without considering the details of the individual transactions which make up the account balance. Analytical procedures may use methods and techniques to improve the efficiency of audits by determining the expectations and by comparing them to recorded amounts. These techniques range from simple comparisons to complex analyses using advanced statistical techniques or systems based on computational paradigms such as neural networks. AR techniques and research trends were discussed in sections 1.2.1 and 1.2.2.

Country specific business and accounting cultures influence the way in which auditors use analytical procedures in practice. For example, in the US the use of analytical procedures in the planning and overall review phases of an audit is required under generally accepted auditing principles (GAAP). Furthermore, SAS 96 contains amendments adding specific documentation requirements to the SAS 56, which at the moment requires auditors to document the factors they considered in developing the expectation for a substantive analytical procedure (AICPA 2002a). Besides, auditors have to document the expectation if it is not evident from other documentation.

[25] Budget-to-actual comparison of financial data in municipal audit is significant because municipal budgets have binding legal authorities (Johnson & Johnson 1995).

According to SAS 96, the auditors should also document (a) the results of their comparison of that expectation to the recorded amounts or ratios that they developed from recorded amounts, and (b) any additional auditing procedures they performed in response to significant unexpected differences arising from the analytical procedures, as well as the results of such additional procedures. In Finland the Finnish Institute of Authorised Public Accountants recommends the use of analytical procedures in the planning and overall review phases. Principally, analytical procedures could be performed at any phase of audit.

Recent research of Lin, Fraser, and Hatherly (2003) in Canada indicates that analytical procedures are extensively applied in practice, particularly by larger audit firms, and that their use dominates the overall review phase of audit regardless of the firm size. These results are comparable with the research conducted in the US (Ameen & Strawser 1994). One explanation for the greater use of analytical procedures by larger audit firms is the client size. Larger clients are more likely to have internal control systems that facilitate the reliance of accounting data and produce documents and data for AR purposes. Table 3.1 shows examples of the purposes of analytical procedures for each of the three audit phases (cf. application domains in chapter five). The X in the boxes in the matrix indicates that a certain purpose is applicable to a certain phase. The purposes vary in different phases of the audit.

Table 3.1 Timing and purposes of analytical procedures

	Planning	Testing	Overall Review
Indicating material error	X	X	X
Assessing going concern	X		X
Indicating management fraud	X	X	X
Reducing detailed test	X	X	
Assessing internal control risk	X		
Forecasting audit fee	X		X

Kinney and Felix (1980) present a summary table of the characteristics of AR techniques. In Table 3.2 I have kept the classification scheme but renamed the techniques according to Fraser, Hatherly, and Lin (1997). Auditors have to be aware of the characteristics of AR techniques in order to interpret the evidence they provide to be used in the audit process. However, when using SQTs or AQTs auditors should consider the possible effects of any uncorrected accounting errors in past data. Indeed, the range of deviations from what might be reasonable will still largely remain a subjective assessment. Ultimately the auditor's choice of procedures, techniques, and level of application is a matter of professional judgement (IFAC 2003a).

Table 3.2 Characteristics of AR techniques

AR technique	Information used	Predictions determined	Reliability of predictions determined
NQT	Any available information	Subjectively	Subjectively
SQT	Past audited values	Objectively	Objectively
AQT	Past audited values and quantifiable environmental information	Objectively	Objectively

Researchers have also stated that AR procedures are tools which management could use as a part of its responsibilities for controlling (Lee & Colbert 1997; Colbert 1994). A management accountant could effectively utilise the same benefits of AR procedures that auditors do. Accountants could apply the AR procedures to various accounts to search for trends and relationships that do not appear reasonable. If AR procedures are applied before the account values are integrated into the financial statements or prior to auditors' investigations, possible faults can be corrected in advance.

As mentioned earlier, ANNs offer a promising alternative technique to AR procedures. ANNs provide a non-linear dimension that captures the underlying representations within the data sets. In my research, I have used the past audited values to form expectations for the account values. The expectations are formed with the ANN-system. I have evaluated the ANN-system's output by comparing its results with the data from similar prior periods (cf. Gauntt and Gletzen, 1997; SAS 56; ISA 520). The detailed evaluation metrics are discussed more thoroughly in section 6.7.4. All my comparison metrics can be classified under the SQT category described in section 1.2.2. These comparison metrics are in line with, for example, the recent study of Lin, Fraser, and Hatherly (2003) which indicates that audit firms continue to emphasise judgement-based procedures where evaluation is based on knowledge and past experience, as compared to those which are more quantitatively based.

3.3 Continuous auditing

Although the idea of continuous auditing is not new (Kunkel 1974; Groomer & Murthy 1989; Vasarhelyi & Halper 1991; Kogan, Sudit & Vasarhelyi 1999), the development of continuous auditing tools is important to create a real on-line auditing environment. According to Kogan, Sudit, and Vasarhelyi (1999), continuous auditing is a type of auditing which produces audit results simultaneously with, or a short period of time after, the occurrence of relevant events. It is a systematic process of gathering electronic evidence under the paperless, real-time accounting system (Razaee, Elam & Sharbatoghlie 2001). Audit work will focus on monitoring transactions and comparing them to expected results on a continuous basis (Vasarhelyi, Kogan & Alles 2002). Conceptually, the continuous audit is an assurance service where the time between the occurrence of events underlying a particular subject matter and the issuance of an auditor's opinion on the reliability of a client's representation of the subject matter is eliminated (Woodroof & Searcy 2001).

Woodroof and Searcy (2001) have introduced the framework for continuous auditing with interconnected web servers, continuous auditing environment with agreements, characteristics of a reliable and secure system, and evergreen reports (see Figure 3.3). Evergreen or updated audit reports are dynamically dated according to continuous audit agreement. According to Woodroof and Searcey (2001), the continuous audit agreement is the contract between the parties participating in a continuous audit environment, i.e. the audit firm and the client. The third parties, such as shareholders, investors, tax authorities, and suppliers could have access to read these updated audit reports.

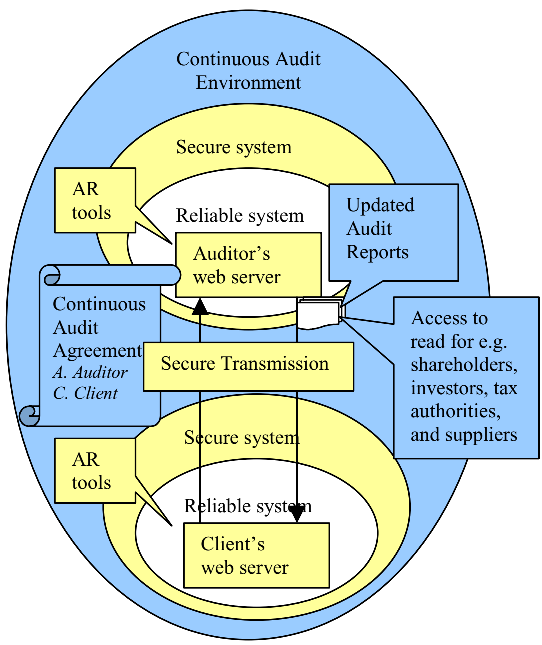


Figure 3.3 The framework of a continuous audit with AR (modified from Woodroof & Searcy 2001)

Woodroof and Searcy (2001) say that the continuous audit environment requires that the participating web servers are connected and given authority to communicate. The authority means that the client's web server allows the auditor controlled access to the client's database. The data flows through the client's system and is continuously monitored and analysed using for example AR tools integrated in the system. AR tools could be placed either in the auditor's systems or in the client's system. Woodroof and Searcy (2001) stress that the automated processes within the continuous audit environment must be highly reliable. According to them reliability encompasses SysTrust™ principles of *integrity*, *security*, *availability*, and *maintainability* (AICPA 2002b):

- Integrity means that the system is complete, accurate, timely, and authorized.
- Security means that the system is protected against unauthorised access.
- Availability means that the system is available for operation and use at the time of continuous audit agreement.
- Maintainability means that the system is updated when required and that the system's availability, security, and integrity are secured.

Already Kunkel (1974) argued that auditing by exception on a continuous basis could substantially increase the efficiency and effectiveness of the audit function. By utilising expectation made by AR tools on a continuous basis, the auditors may detect invalid data items or transactions and problem areas within the company. In the continuous audit environment, the differences in auditor-defined rules activate alarms via the Internet to the auditor regarding any potential anomaly in the system. The alarms require communication and possible investigation of the data. If there are no deviations in expectations at any of the levels of auditing (the accounting, the complete set of financial statements, and the corporate governance of the company) an unqualified updated audit report might be given.

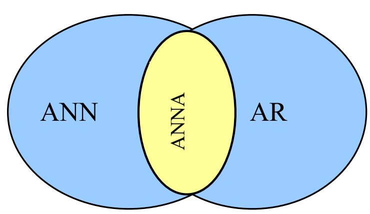


Figure 3.4 Research area

To sum up, AR tools should be embedded in the continuous auditing environment. The AR tools could utilise ANN-technology to create the expectations to be used in monitoring and controlling data. Therefore, the contribution of this research lies in the intersection of two areas: the analytical review and the artificial neural network (see Figure 3.4). The development of AR tools is important regarding the workload and the demand of auditors in today's business environment.

4 ARTIFICIAL NEURAL NETWORKS

This chapter starts by reviewing some notes of ANNs' history[26] and by giving the definition for an ANN. The discussion continues with some basics elements of an ANN such as a neuron, activation functions, and layers. The description of learning algorithms pays attention to their feasibility to my research problem, followed by presenting the selected methods.

4.1 History and definition of the ANN

Several researchers stand out in the pioneering work of ANNs. McCulloch and Pitts (1943) introduced the idea of ANNs as computing machines. Hebb (1949) postulated a self-organising learning rule, i.e. Hebbian learning. Rosenblatt (1958) suggested the perceptron model for learning with a teacher, i.e. supervised learning. Thereafter, there were some enthusiastic researches on ANNs in the fifties and sixties, but then ANNs were almost forgotten for more than ten years. One reason was technological; there were no personal computers or workstations for experimentation. The other reason was partly psychological partly financial; the book "Perceptrons" by Minsky and Papert (1969) did not encourage anyone to work on or to support the work on ANNs[27].

A new rise began in the eighties. Hopfield (1982) used the idea of an energy function to formulate a way of understanding the computation performed by recurrent networks with symmetric connections, i.e. Hopfield networks. Kohonen (1982a, 1982b) published the paper on self-organising maps (SOMs). Barto, Sutton, and Anderson (1983) demonstrated that a reinforcement learning system could learn in the absence of a helpful teacher. In 1986 the development of the backpropagation (BP) algorithm was reported by Rumelhart, Hinton, and Williams (1986). Broomhead and Lowe (1988) described a procedure for the design of layered feedforward networks using a radial basis function. Riedmiller and Braun (1993) proposed the *resilient*

[26] The history marks of ANN are mainly collected from the secondary references such as Hecht-Nielsen (1991, 14–19), Russell and Norvig (1995, 594–596), and Haykin (1999, 38–44).

[27] They overstated the limitations of the Rosenblatt's (1958) perceptron as being unable to solve nonlinear classification problems, although such a limitation was already known. Their campaign achieved its purpose: AI researchers got all of the ANN research money (Hecht-Nielsen 1991, 16–17).

backpropagation (RPROP) training algorithm to eliminate the effects of the magnitudes of the partial derivatives of the BP algorithm.

In the nineties the processing capabilities of the PCs improved enormously and made it possible to model ANN-based information systems on PCs. ANNs are nowadays a well established computational paradigm in the field of artificial intelligence (AI). ANNs have been applied extensively in solving many complex real-world problems. The attractiveness of ANNs comes from their information processing characteristics such as nonlinearity, parallelism, learning, and generalisation capabilities (Basheer & Hajmeer 2000). They are usually seen as a method of implementing complex non-linear mappings (functions) using parallel processing units (neurons) that are connected together with weighted, adaptable connections. By embedding a vast number of neurons in an interactive nervous system, it is possible to provide computational power for very sophisticated information processing. Haykin (1999, 2) provides the following definition of ANN:

> *An artificial neural network is a massively parallel distributed processor made up of simple processing units (neurons), which has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two aspects:*
>
> 1. *Knowledge is acquired by the network from its environment through a learning process.*
> 2. *Interneuron connection strengths, known as (synaptic) weights, are used to store the acquired knowledge.*

ANNs may be classified in many different ways. Chapter five classifies ANNs according to their application domains in auditing. This chapter classifies ANNs based on their features such as the supervision needed for training and the types of learning algorithm. Three different kinds of learning paradigms to train ANNs are *supervised*, *unsupervised*, and *reinforcement* learning. The learning algorithm is the function to modify the weights of the network in order to attain a desired objective. The learning paradigms are described in section 4.3 and the learning algorithms are discussed in section 4.4.

4.2 A neuron, activation functions, and layers

A *neuron* is an information-processing unit of an ANN (Figure 4.1a). Three basic elements of the neuron are (Haykin 1999, 10):

1. A set of *connecting links*, each of which is characterized by a weight of its own. A signal x_j at the input of neuron j connected to neuron k is multiplied by the weight w_{kj}.
2. An *adder* for summing ($\sum$) the input signals weighted by the respective connecting link of the neuron.
3. An *activation function* for limiting the amplitude of the output of a neuron.

There are multiple approaches depending on e.g. how the weights are connected and what the adder and activation functions are. These interconnections and functions can make ANNs mathematically complex and sophisticated. Each of these approaches has a unique mix of e.g. information-processing capabilities, domains of applicability, techniques for use, required training data, and training methods (Hecht-Nielsen 1991, 7). The different approaches do not compete against each other; rather, they represent various specialisations in solving different types of problems. Indeed, the approach strongly influences e.g. what the network can do (Hertz, Krogh & Palmer 1994, 9).

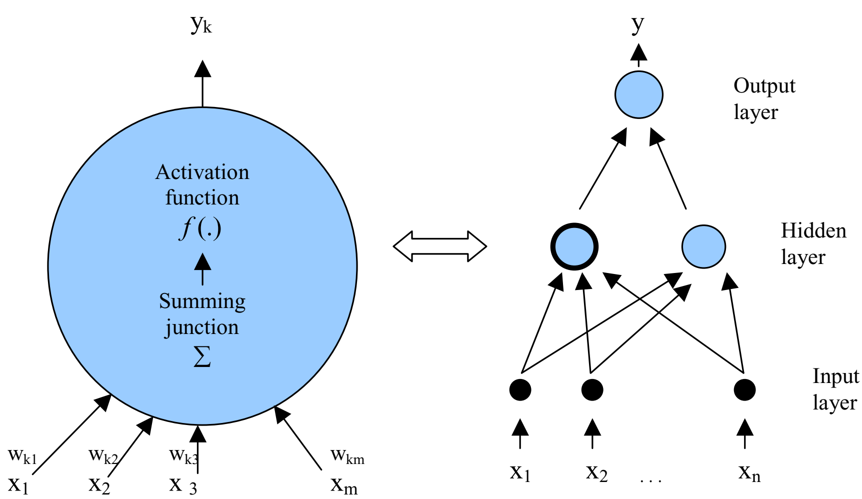


Figure 4.1a An artificial neuron Figure 4.1b A multilayer feedforward ANN

The activation function, denoted by $f(.)$ in Figure 4.1a, defines the output of a neuron. There are several activation functions available such as (Shih 1994, 83–87): *threshold (step) function, linear function,* and *sigmoid function.* An example of the threshold function is the early work of McCulloch and Pitts (1943) where the output of a neuron takes on the value of 0 if the induced local field of that neuron is negative, and 1 otherwise. The linear function generates the outputs, which are proportional to the inputs. The sigmoid function, whose graph is s-shaped, exhibits a graceful balance between linear and nonlinear behaviour. Examples of the sigmoid function are a *logistic function* and a *hyperbolic tangent function* (see e.g. Swingler 1996, 62; Mennon, Mehrotra, Mohan & Ranka 1996). Whereas a threshold function assumes the value of 0 or 1, the logistic function assumes a continuous range of values from 0 to 1. The hyperbolic function is shaped like the logistic function but it ranges from –1 to +1 rather than 0 to 1. The logistic function with the elegant balance between 0 and 1 was a natural choice for my prototype because different account values after scaling could get any positive continuous value in a nonlinear function. Figure 4.2 illustrates the activation functions. Very often, the neuron also includes a bias, which has the effect of increasing or lowering the net input of the activation function, depending on whether it is positive or negative, respectively.

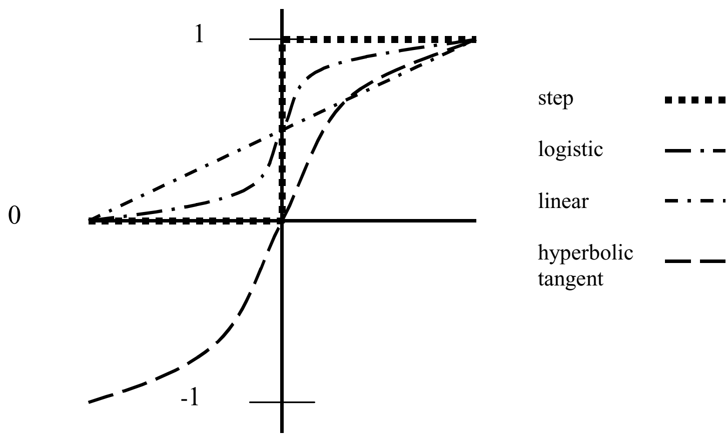


Figure 4.2 Activation functions

The neurons in an ANN are organised into *layers* (Figure 4.1b). The first layer is called the *input layer* and the last layer is the *output layer*. The inner layers are known as *hidden layers*. A hidden layer helps multilayer perceptron (MLP) network for mapping nonlinear input–output mapping, i.e. the universal approximation theorem, see e.g. Haykin (1999, 209). However, a single hidden layer might not be optimum in the sense of learning time, ease of implementation or generalisation[28]. Therefore, adding one or more hidden layers might help the network to find an unknown input–output mapping. The MLP network can also include multiple inputs and multiple outputs, i.e. MIMO-systems (Haykin 1999, 69). Sometimes the hidden layer might have some specific task (e.g. pattern classification and function approximation) like in the study of Etheridge, Sriram, and Hsu (2000)[29]. The ANN-*architecture* refers to the number and size of layers (Swingler 1996, 51).

The ANNs may be classified based on the direction of the information flow in the network. According to this classification, there are three different classes of ANNs: *single-layer feedforward networks, multilayer feedforward (perceptron) networks,* and *recurrent networks* (Haykin 1999, 21–23). The single-layered network has an input layer of source neurons and an output layer of neurons. This network is strictly a feedforward type and the input layer of source neurons is not counted because no computation is performed there. Figure 4.1b illustrates an example of multilayer feedforward network which consist of numbers of neurons that are interconnected and layered. A recurrent ANN distinguishes itself from a feedforward ANN in that it has at least one feedback loop (e.g. Hopfield network). For example, a recurrent network may consist of a single layer of neurons with each neuron feeding its output signal back to the inputs of all the other neurons.

The ANNs may be classified based on the degree of the connectivity of the neurons. The ANN is *fully connected* when every neuron in each layer of the network is connected to every other neuron in the adjacent forward layer. If some of the communication links are missing from the network, the network is *partially connected*.

4.3 Learning paradigms

As mentioned in section 4.1, the training of neural networks can be classified into three categories: *supervised*, *reinforcement (graded),* and *unsupervised (self-organisation) learning* (Hecht-Nielsen 1991, 7, 48). Reinforcement learning can be classified either into the supervised learning paradigm (Hertz,

---

[28] Generalisation refers to the ANN's capability to produce reasonable outputs for inputs not faced up during training (learning). According to Haykin (1999, 208), the generalisation is influenced by three factors: 1) the size of the training set, and how representative it is of the environment of interest, 2) the architecture of the ANN, and 3) the physical complexity of the problem at hand. General issues or rules of thumbs in ANN development may be found in the review articles or for example in Basheer and Hajmeer (2000).

[29] Etheridge, Sriram, and Hsu (2000) compared the performance of the: backpropagation (BPN), categorical learning network (CLN), and probabilistic neural network (PNN) as classification tools to assist and support the auditor's judgment about a client's continued financial viability in the future (see section 5.1.3).

Krogh & Palmer 1994, 10) or into unsupervised learning paradigm (Haykin 1999, 64). The distinction between supervised and unsupervised learning depends on information, e.g. pattern recognition is supervised if the training algorithm requires knowledge of the class membership of the training samples, unsupervised if it does not require it (Kosko 1990).

In supervised learning a teacher has some knowledge of the environment that is unknown to an ANN. The teacher expresses this knowledge with training examples which consist of input variables together with desired target values (Hecht-Nielsen 1991, 48). The network processes its output values from the input variables and compares them with the target output values. If an error, i.e. a difference between outputs and targets exists; the network adjusts the weights by a small amount in some direction in a step-by-step manner until the error is at an acceptable level. Therefore, supervised learning is an instructive feedback system.

In reinforcement or graded learning the training examples are given to a network without any desired outputs. In addition to the training data inputs, the network receives occasionally a grade, a performance score, from its environment. This grade tells how well the network has done overall since it was graded last time (Hecht-Nielsen 1991, 49). The reinforcement learning is on-line learning without a teacher. This paradigm is an evaluative feedback system since it evaluates the system's behaviour (Haykin 1999, 65).

In unsupervised learning nobody oversees the learning process. Therefore, the network is given only the training data inputs from which the network organises itself into some useful configuration (Hecht-Nielsen 1991, 49). The input vectors are classified according to their degree of similarity. The similar input vectors activate the same output cluster. The user is responsible for giving an interpretation to the clusters.

The hybrid use of learning paradigms can provide a better solution than one paradigm alone. For example, when similar input vectors produce similar outputs, it may be rational to categorise the inputs first with unsupervised learning paradigm and feed that information for supervised learning paradigm (Hertz, Krogh & Palmer 1994, 246). In my case I have used both supervised and unsupervised learning paradigms practically to the same kind of data to explore the applicability of the output data in the analytical review context.

However, I have assumed that the hybrid use of unsupervised and supervised learning paradigms could be useful in the AR context. Firstly, we could classify account values in the homogenous groups with the unsupervised learning algorithm and then we could use these groups as inputs for the supervised learning algorithm. However, the hybrid use of different learning algorithms is beyond the scope of my research.

4.4 Learning algorithm

A *learning algorithm* defines how the network weights are adjusted during the training process. There are four basic types of learning algorithms (Basheer & Hajmeer 2000; Haykin 1999, 51–61):

- Error-correction learning[30] tries to minimise an error between a desired and an actual output in accordance with some cost function (i.e. performance measure) such as the root-mean-square-error (RMSE). A correction occurs through weight modification.
- Bolzman learning is a device for modelling the underlying probability distribution of a given data set.
- Hebbian learning may be expressed as follows: if two neurons on either side of a connection are activated simultaneously, the strength of that connection is selectively increased. If these neurons are not activated simultaneously, the connection is selectively weakened or eliminated.
- In competitive learning, the output neurons of an ANN compete among themselves to become active. One of them will be a winner in accordance with a chosen metric and only it will be activated. Then, the winner's weight vector is updated to correspond more closely the input pattern.

The selection of the feasible learning algorithm depends on the problem and the data. Some networks are more suitable for classification problems, while others are more feasible for function approximation. A vast number of networks, new or modifications of existing ones, makes the selection of the appropriate ANN a demanding task. The selection could be based on an educated guess, such as:

> *The error-correction learning algorithm, such as backpropagation (BP), seems to be suited to indicating expected account values based on past audited values. The BP might approximate an unknown input–output mapping, i.e. the underlying rules relating the inputs to the outputs. The ANN might find the relationships between the account values and thereby give an estimation of a true and fair view of the account values.*
>
> *Or, allow the competitive learning algorithm to organise the nonlinear account values and then research the output results. The SOM might classify different account values in a meaningful manner.*

---

[30] In the literature, it might be also called a Widrow-Hoff rule according to its originators Widrow and Hoff (1960).

The selection could also be based on the popularity of the learning algorithm. For example, the BP algorithm has emerged as the standard for training multilayered feedforward networks (Sohl & Venkatachalam 1995; Koskivaara 2004a; Wong & Selvi 1998), whereas in the unsupervised learning paradigm the SOM has become the benchmark against which other innovations in the field are evaluated (Haykin 1999, 42). I started my ANN-development progress by selecting the most popular learning algorithm and by assuming that it might work in my problem domain.

4.4.1 Backpropagation algorithm

The BP algorithm operates in a multilayered feedforward (perceptron) network (see Figure 4.1b). BP has two kinds of signals (Haykin 1999, 159–160): *function signals* and *error signals*. A function signal is an input signal that comes in at the input layer of the network, propagates forward neuron by neuron through the network, and appears at the output layer of the network as an output signal. An error signal originates at an output neuron of the network, and propagates backward layer by layer through the network.

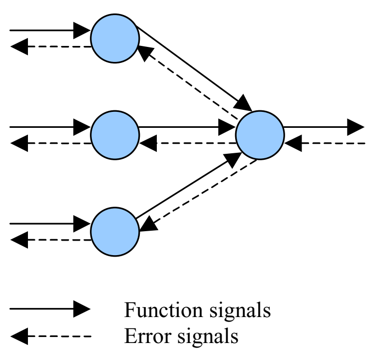


Figure 4.3 Illustration of function and error signals (Haykin 1999, 159).

Figure 4.3 illustrates the directions of two basic signal flows. Each hidden or output neuron of multilayer perceptron performs forward and backward computations. The forward computation of the functional signal, i.e. the output of a neuron, is a continuous nonlinear function of the input signals and weights associated with that neuron. The backward computation estimates the gradient vector (i.e. the gradients of the error surface with respect to the weights connected to the inputs of a neuron). BP can be used for both *batch training* in which the weights are updated after the processing of all the training examples that constitute an epoch and *incremental* (on-line) *training* in which the weights are updated after the processing of each training example.

In the standard BP algorithm the weight correction applies the delta rule which multiplies an input signal of neuron with a local gradient with a *learning rate* parameter (Haykin 1999, 163). The learning rate affects the speed at which the network settles on a solution by allowing us to regulate how much the error decreases in each iteration. Multilayer networks typically use sigmoid functions in the hidden layers. These functions are often called "squashing" functions, since they compress an infinite input range into a finite output range. Sigmoid functions are characterized by the fact that their slope must approach zero, as the input gets large. This causes a problem when using the steepest descent to train a multilayer network with sigmoid functions, since the gradient can have a very small size; and therefore, cause small changes in the weights and biases, even though the weights and biases are far from their optimal values. Therefore, parameters such as *adaptive learning rate* and *momentum* were developed. The *adaptive learning rate* accelerates the learning process by utilising the concept of the direction in which the error has been decreasing recently (Smith 1996, 88–90). It speeds up the training process when the ANN is far away from the correct weights and slows down when the ANN gets closer. A method of increasing the rate of learning is to modify the delta rule by including a *momentum* (Rumelhart, Hinton & Williams 1986). Momentum tends to keep the weight changes going in the same direction.

The supervised learning means iteration of the forward and backward computation by training epochs. When a neuron receives inputs, it computes its output value and sends it to the neurons on the next layer above. Consequently, the inputs of the network are fed forward through the entire network until they reach the output layer. Then, the difference, which is calculated with some cost function such as root mean square error (RMSE), between the output vector and the desired target vector is backpropagated through the ANN to modify the weights for the entire neural network. This iterative process is called training. After the BP network has been trained, it is tested against the records of a testing data set that have not been previously met with the network. For these records, the desired target output is known. The output generated for each record of these testing data is checked against the desired target output for that record. If there is a match, it is concluded that the trained network could recognise the record correctly.

4.4.2 Resilient backpropagation algorithm

A *resilient backpropagation* (RPROP) algorithm is an adaptive method for the BP. The purpose of RPROP[31] is to eliminate the influence of the size of the partial derivatives on the weights (Riedmiller 1994a; Riedmiller 1994b; Riedmiller & Braun 1993). In the RPROP the sign of the derivative is used to indicate the direction of the weight update. The size of the derivative has no effect on the weight update. The adaptation rule of RPROP works as follows. Every time the derivative of the weight changes its sign indicating that the last update was too big and the RPROP has jumped over a local minimum, the update value is decreased. If the derivative maintains its sign, the update value is increased in order to accelerate the learning process. The update process proceeds as follows: if the derivative is positive, the weight is decreased by its update value, if the derivative is negative, the weight is increased by its update value. The size of the weight change is determined by an update value. If the derivative is zero, then the update value remains the same. Whenever the weights are oscillating the weight change will be reduced.

There are two more reasons for selecting RPROP for the learning algorithm for the prototype. First, the performance of RPROP is not very sensitive to the settings of the training parameters. Second, RPROP uses a batch training algorithm and is therefore efficient and requires minimal storage. As mentioned in section 4.4.1, in the batch training weight updating is performed after the presentation of all the training examples that constitute an epoch. Besides, as the account values in my prototype are presented to the network in a time series manner, the use of holistic updating of weights makes the search of weights occasional by nature. This in turn makes it less likely for the network to be trapped in a local minimum.

4.4.3 Self-organising maps

The SOM is a clustering and visualisation method and the purpose is to show the data set in another representation form (Kohonen 1997, IX). The SOM facilitates the data compression. It has an input layer and an output layer or lattice. Each neuron in the lattice is fully connected to all the source neurons in the input layer. Three processes involved in the formation of the SOM are (Haykin 1999, 447–448):

[31] A complete description of the RPROP algorithm and formulas is given in Riedmiller and Braun (1993).

1. *Competition*. For each input pattern, the neurons in the network compute their respective values of a discriminant function. This discriminant function provides the basis for competition among the neurons. The particular neuron with the largest value of discriminant function is declared winner of the competition.
2. *Cooperation*. The winning neuron determines the location of a topological neighbourhood of neurons, thereby providing the basis for cooperation among neighbouring neurons.
3. *Adaptation* facilitates the neurons to increase their individual values of the discriminant function in relation to the input pattern through adjustments applied to their weighs. The adjustments are such that the response of the winning neuron to the following application of a similar input pattern is facilitated.

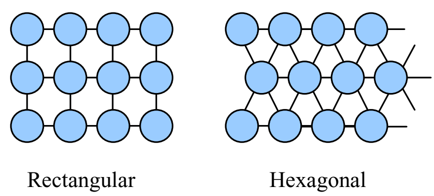


Figure 4.4 Forms of lattice

The SOM has six learning parameters, *topology, neighbourhood type, X- and Y-dimensions, training rate*, *training length*, and *network radius*. The network topology refers to the form of lattice. There are two types of lattices, rectangular and hexagonal (Figure 4.4). In a rectangular lattice each neuron is connected to four neighbours, except for the ones at the edge of the lattice. In a hexagonal lattice structure neurons are connected to exactly six neighbours, except for the ones at the edge of the lattice. Neighbourhood type refers to the neighbourhood function used and the options are Gaussian and bubble. X- and Y-dimensions refer to the size of the map. In maps that are too small differences between clusters are hard to identify and in maps too large clusters will appear to be flat. The training rate factor refers to how much the neuron in the neighbourhood of the winning neuron learns from the input data vector. The training length measures the processing time, i.e. the number of iterations through the training data. The network radius refers to how many neurons around the "winning" neuron are affected during the learning process. The training process of the network consists of two parts. In part one, the map is trained "roughly". In the second part, the network is fine-tuned.

The unsupervised learning means that all the output neurons compete against each other without any feedback from the environment. One of the neurons will be a winner and only it will be activated. The winner's weight vector is updated to correspond more closely to the input vectors. This means that two input items, which are close in the input space, are mapped into the same or neighbouring neurons on the map. Output neurons create groups, which together form a map of the input neurons. Normally, the SOM creates a two-dimensional map from n-dimensional input data. A one-dimensional lattice is a special case, which consists of a single column or row of neurons. This map resembles a landscape in which it is possible to identify borders that define different clusters (see e.g. Figure 6.2).

4.5 Benefits of ANNs for analytical review

An accounting system means the series of tasks and records of an entity by which transactions are processed as a means of maintaining financial records. In double accounting system every transaction is registered into two different accounts, and therefore there is a natural dynamics between different accounts, e.g. accounts receivable and incomes (Saario 1959, 24, 27). Accounting information systems identify, assemble, analyse, calculate, classify, record, summarise, and report transactions and other events. Their capacity to produce data and information in terms of volume and speed is vast. Indeed, accounting systems are also vulnerable to manipulation and fraud.

The auditor's work in each case is to confirm the reliability of the account values. The auditor should ascertain the entity's system of recording and processing transactions and assess its adequacy as a basis for the preparation of financial statements. Therefore, the auditor should carry out such a review of the financial accounts which is as sufficient, in conjunction with the conclusions drawn from the other audit evidence obtained as possible, to give them a reasonable basis for their opinion on the financial statements. As was mentioned in chapter three, this review i.e. analytical review, may include generating trends in accounting values or comparing accounting data items. It is supported with analytical procedures in which different kind of techniques might be embedded.

For analytical reviews ANNs offer an appealing choice. The research and development of audit tools is important, since the task of the auditor today is at once more onerous and more complex than ever before. As mentioned earlier, ANNs are good at handling data, and once trained, they can predict and classify new examples very quickly. Therefore, I argue that auditors may benefit from applying ANNs in revealing trends in accounting data or in the comparison of accounting records. For example, with the BP algorithm the auditor may generate evidence based on internal trends in accounting data and then compare ANN results with actual values. With the SOM algorithm the auditor may visualise clusters and reveal patterns in the accounting data. In brief, auditors may benefit from the ANN's ability to learn from data to support their experience and knowledge about a client company.

5 RELATED WORKS

This chapter provides an overview of ANN studies carried out in the field of auditing. The review takes a broad scope of auditing and encompasses both internal and external auditing. The review focuses on the application domains of the articles. I argue that these auditing ANN-applications could serve the analytical review process. In reviewing the modelling issues special attention is paid to the following issues: data and sample sets, ANN-architectures and learning parameters, and whether the model was evaluated with a holdout sample. The summary of the findings in Table 5.5 pays attention to whether authors state that ANNs have the potential to improve AR procedures.

5.1 ANN applications in auditing

The ANN-application areas in auditing are in detecting material errors (six studies), detecting management fraud (three studies), supporting going concern decisions (five studies), and determining financial distress problems (one study). ANNs have also been applied to internal control risk assessment (three studies) and audit fee forecasting (one study). Next, I outline the studies in chronological order by application area. Every application area ends with a summarisation table of modelling issues, as stated by the authors, in that particular area. Almost every model was evaluated with a holdout sample. The performance rates of the ANN models are reported based on the holdout samples.

5.1.1 Material errors

The major ANN-application area in auditing is material errors. Material error applications direct auditors' attention to those financial account values where the actual relationships are not consistent with the expected relationships. An auditor has to decide whether and what kind of further audit investigation is required to explain the unexpected results. Material error ANN-models either predict future values or classify data. Table 5.1 summarises the modelling issues of ANN literature pertaining to material error problems.

Coakley and Brown (1991a), Coakley and Brown (1993), and Coakley (1995) tested whether an ANN offered improved performance in recognising material misstatements. They used monthly data over four years of a medium-sized distributor. Their ANN-model was based on trend prediction. Data from three years were used for a training set and the fourth year of data was used as the forecast period to evaluate the performance of the ANN. They selected fifteen income statement and balance sheet accounts or aggregates to represent major balance sheet categories. The inclusion of all accounts values was not feasible due to the impact of the number of neurons on the time it takes to train an ANN. They compared a presumed lack of actual errors and seeded material errors to evaluate the ANN's performance. The results of the study were divided into findings based on: financial ratios, comparison of methods (financial ratio, regression, ANN), effect of error size, effect of statistical level of confidence, effect of source of material error, and applying methods to the base period. The results were compared to the results achieved with financial ratio and regression methods, and the ANN demonstrated a better predictive ability with less overall variation in the predicted values. However, the researchers argue that the fluctuating nature of the financial data within their studies limited the effectiveness of all the AR procedures. The fact is that higher unexpected fluctuations present in financial data sets cannot be effectively analysed by any known forecasting method.

Coakley and Brown (1991b) also tested ANN technology for recognising patterns in financial ratios of a medium-sized manufacturing firm. In this study they also predicted future values with an ANN. The financial accounts were selected so that they provided information about a company's solvency and the movement of accounts receivable and inventory. The model was trained using 36 months of data with an auto-association process, which means that the input pattern and the desired output pattern were the monthly account balances. Thus each pattern was associated with itself. They evaluated the effectiveness of the model by seeding errors in the data. Their preliminary results indicated that the use of ANNs for pattern recognition across related financial data sets might be viable. The research settings in my first studies, such as data sets, variables, and backpropagation learning algorithm, resemble this research.

Wu (1994) applied the ANN system to classify tax cases to ascertain whether a further audit was required or not. 180 sample companies were carefully gathered from an expert tax auditor’s audit case file, 90 of which required a further audit and 90 of which required no further audit. The input data (16 attributes) consisted of both reported values and ratios. Half of the cases were used as the training set and the other half was held out as the testing set to validate the implemented network. The cases consisted of information about a firm's business income tax behaviour. The classification accuracy for the neural network was 94% with a two-layer neural network and 95% with a three-layer neural network. The results are not compared to those of any other methods.

Table 5.1 Summary of modelling issues of material error applications

Researcher Country	Industry Data type	Train/test set	# inputs neurons	# hidden layers/ neurons	# outputs	Results[32]	Activation function	Learning algorithm	Performance measure
Coakley & Brown (1991a) USA	Distributor Monthly account values	48/48	11	2:	10	50–62%	Sigmoid	BP	Average error, standard deviation
Coakley & Brown (1991b) USA	Manufacturing firm Monthly account values, aggregates	36/36	42	1:15	15	30–82%	Modified sigmoid squashing	BP	MSE (mean square error)
Coakley & Brown (1993) USA	Distributor Monthly account values, aggregates	36/12 *	42	1:15	15	58–100 %	Modified sigmoid	BP	MSE
Wu (1994) Taiwan	Sample of firms Income tax behaviour data	90/90	16	1:8 0:0	1	95% 94%	Sigmoid	BP	Predictive accuracy
Coakley (1995) USA	Distributor Monthly financial ratios	36/12 *	5	2:11–11	3	60–90%	Hyperbol. Tangent activation	BP	SSE (sum of square error)
Busta & Weinberg (1998) USA	Simulated data Digits of numbers	800/ 800 *	34 24 15 5 1	1:4 1:6 1:6 1:6 1:4	1	67% 69% 53% 73% 75% **	Logistic	BP	accuracy %

* A holdout sample was employed. NA. = no answer. ** Level of contamination is 0%.

Busta and Weinberg (1998) used an ANN to distinguish between “normal” and “manipulated” financial data. They generated 800 data sets containing 200 two-digit numbers by simulating these data sets with the help of Benford’s law and non-Benford’s law[33]. The ANN analysed then input variables and the generated an estimation of the degree on contamination in the data sets. They

[32] Results are either as stated by the authors or my interpretation, e.g. if the average error is 20% I have interpreted that the result is 80%.

[33] Benford´s law says that the digits of naturally occurring numbers are distributed on a predictable and specific pattern. It determines the expected frequency for each digit in a position in a set of random numbers. This means that the chances for any number in a given database are mathematically predictable. Since the expected frequency for each number in the set is known, every item in excess of the expected frequency is deemed suspicious.

tested six ANN designs to determine the most effective model. In each design, the inputs to the ANN were the different subsets of the 34 variables. The results showed that if data have been contaminated at a 10% level or more, the ANN would detect this 68% of the time. If the data are not contaminated, the test indicated that the data are "clean" 67% of the time. The study uses only simulated data.

5.1.2 Management fraud

Auditors cannot make assumptions as to the honesty or dishonesty of the management. They should be aware of the possibility of management misrepresentation at the start of the audit and re-examine the likelihood of management misrepresentations as the audit progresses. Management fraud can be defined as deliberate fraud committed by the management that injures investors and creditors through materially misleading financial statements. Table 5.2 summarises the modelling issues of ANN literature pertaining to management fraud problems.

Green and Choi (1997) developed an ANN fraud classification model employing financial data. They used five ratios and three accounts as input variables. The selection of variables was determined by practical research. The fraud sample consisted of financial statements of various companies filed by SEC (Securities and Exchange Commission) that had been subsequently found to contain fraudulent account balances. The financial statements of the nonfraud sample received unqualified auditor opinions for the selected year. They were selected directly from COMPUSTAT and matched to the fraud sample on the basis of year, size, and industry (four digit SIC). Their training samples consisted of 44–49 companies and the holdout samples consisted of 42–46 companies respectively. The results showed that ANNs have significant potential as a fraud investigative and detection tool. All sums of the Type I and Type II error rates are significantly less than the random chance benchmark of 1.00[34]. Another contribution of their ANN models is the consistently low Type II error.

Fanning and Cogger (1998) used an ANN (AutoNet) to develop a model for detecting management fraud. They compared the results of an ANN with linear and quadratic discriminant analysis as well as logistic regression. Their sample consisted of 150 firms in the training sample and 54 firms in the holdout sample. The variables were selected by AutoNet and they were the

[34] Type I: There was fraud but ANN did not find it (= inefficient auditing). Type II: There was no fraud but ANN detected one (= ineffective auditing).

outsider director (not the owner), having a non-Big Six auditor, the geometric growth rate, accounts receivable to sales, net plan property and equipment to total assets, debt to equity and the trend variables for accounts receivable, and gross margin. The prediction accuracy of the ANN for the training sample was 75% and 63% for the holdout sample. The result of their models suggested there is potential in detecting fraudulent financial statements through analysis of public documents. They also showed that ANNs offer better ability than standard statistical methods in detecting fraud.

Feroz, Kwon, Pastena, and Park (2000) illustrated the application of the ANNs in order to test the ability of selected SAS 53 red flags to predict the targets of SEC investigations. They used both financial ratios and non-financial turnover red flags mentioned in SAS 53. They tested the ANN model with the sample of 42 firms from various industries. 28 firms were used for training the ANN and 14 firms were used for testing. The ANN models classified the membership in target (investigated) firms versus control (non-investigated) firms with an accuracy of 81%. The testing was biased because they only used those red flags that can be constructed from publicly available information. However, authors believed that the sampling choice was sound given the data constraints in that particular case.

Table 5.2 Summary of modelling issues of management fraud applications

Researcher Country	Industry Data type	Train/test set	# inputs	# hidden layers/ neurons	# outputs	Results[32]	Activation function	Learning algorithm	Performance measure
Green & Choi (1997) USA	Sample of firms Ratios and values of financial statements	45–49/ 42–46 *	8	1:4	1	31–60%	Sigmoid logistic	BP	Error rate %
Fanning & Cogger (1998) USA	Sample of firms Accounts and ratios of financial statements	150/54 *	8	NA.	1	63%	Simple quadratic, Quadratic	“AutoNet”	Prediction accuracy % Log./DA
Feroz *et al.* (2000) USA	Sample of firms 7 SAS No. 53 red flag data	28/14 *	7	1:14	1	81%	Binary sigmoid	BP	OER (overall error rate) MSE

* A holdout sample was employed. NA. = no answer

5.1.3 Going concern and financial distress

An auditor gives a going concern (GC) uncertainty opinion when the client company is at risk of failure or exhibits other signs of distress that threaten its ability to continue as a GC. In US, there are different degrees of financial distress (FD), and the auditors do have a choice between two types of going concern reports: modified audit report and disclaimer audit report. Bankruptcy is a situation where there is no ability to continue business (see for example the study of Tam & Kiang 1992). For example, in the US SAS 59 requires the auditor to evaluate whether there is a substantial doubt about a client's ability to continue for at least 1 year beyond the balance sheet data (Arens, Elder & Beasley 2003). GC and FD research as an application area of ANNs has been minimal, although it resembles bankruptcy studies which is one of the most popular ANN research areas in business sciences. Table 5.3 summarises the modelling issues of ANN literature pertaining to GC and FD problems.

Hansen, McDonald, and Stice (1992) samples consisted of 80 FD companies; 40 that received the GC audit report, and 40 that did not receive the GC audit report, and of 98 firms involved in litigation. The source of the sample was the Disclosure II Database that reported on all publicly traded companies whose fiscal year ended between March 31, 1981, and February 28, 1982. This database contains the financial statements of all firms on the New York Stock Exchange (NYSE) and the American Stock Exchange (AMEX). The holdout sample consisted of 50% of the total sample. They had two models with different variable settings. The audit opinion model had either twelve ratios from financial statements or other closing of the books information as variables. The litigation model had nine variables, which were client, auditor or engagement specific ones. The average error of the audit opinion model was 8.43% and 20.11% for the litigation model. The results indicated that in the case of predicting audit opinions, the qualitative-response models perform at a competitive level with the machine-learning models. Theoretical results inferred that this might be especially true when the training sets were relatively small. The authors stated that qualitative response models might be a desirable alternative when the training samples are relatively small and there is a need to incorporate additional parameters such as prior probabilities and error costs. The study uses small and old data sets.

Fanning and Cogger (1994) examined the efficiency of a generalised adaptive neural network algorithm (GANNA) processor in comparison to earlier model-based methods: a backpropagation ANN and logistic regression approaches to data classification. The research used the binary classification problem of discriminating between failing and non-failing firms to compare the methods. The sample consisted of 190 matched pairs of which the first 75 were selected for the training sample in chronological order and the last 115 matched pairs defined the holdout sample. All the ANN models had three inputs: the mean adjusted cash flow divided by its standard deviation, the firm's adjusted cash position divided by its standard deviation, and the number of years prior to failure for the failed year. No single approach studied was uniformly superior across all comparison and statistics. The reason for this might be the limited number of input variables. However, the results indicated the potential in time savings and the successful classification results available from GANNA and ANN processors. Another limitation to this study is that it was carried out with fairly old data; the data was from 1942 to 1965.

Lenard, Alam, and Madey (1995) studied the generalised reduced gradient (GRG2) optimiser for ANN learning, a backpropagation ANN, and a logit model to predict which firms would receive audit reports reflecting a GC uncertainty modification. GRG2 is a solver used in lieu of backpropagation's gradient decent algorithm for training feedforward supervised learning neural networks. The sample for the study was drawn from the 1988 Disclosure II Database. The training sample consisted of 70 firms and a holdout sample of 10 firms. The selection of variables was intended to determine whether the GC decision could be made from publicly available financial statement information. The variables consisted of ratios and account values. The ANN model formulated using GRG2 had the highest prediction accuracy of 95%. It performs best when tested with a small number of variables on a group of data sets, each containing 70 observations. The GRG2 based ANN was proposed as a robust alternative model for auditors to support their assessment of GC uncertainty affecting the client company. The authors state that a possible limitation to the model is whether all factors that affect the auditor's going concern uncertainty decision have been considered.

Anandarajan and Anandarajan (1999) compared ANN, expert system (ES), and multiple discriminant analysis models to facilitate the decision on the type of GC report that should be issued. The experimental sample of the study was drawn from the 1992 Disclosure database. The data consisted of 14 ratios calculated from the financial statements of 61 companies of which 27 were reserved for a holdout sample. The validity of the models was tested by comparing their predictive ability of the type of audit report, which should be issued to the client. The prediction accuracy of the ANN in the holdout sample for all reports was 85.8%. This was better than the prediction accuracy of the ES (69.1%) or the multiple discriminant (74.1%) models. The results of the study indicate that the ANN model predicts the best type of GC audit report that should be issued to the client when compared to the ES and the multiple discriminant analysis. The authors say that many qualitative variables, which could be useful in supporting the selection of the type of GC report, are not incorporated into the models.

Koh and Tan (1999) predicted a firm's CG status from six financial ratios with an ANN model. Their data set contained a sample of firms divided into 165 with non-GCs and 165 matched GCs. 300 of the cases were used for training the ANN and the remaining 30 cases were reserved for testing. On an evenly distributed holdout sample, the trained network model correctly predicted all 30 test cases. They compared the GC results of the ANN to the probit model and the audit opinion. Their results suggested that the ANN was at least as good as both the auditors and the probit model for predicting the GC status of firms from financial ratios.

Table 5.3 Summary of modelling issues of GC and FD applications

Researcher Country	Industry Data type	Train/test set	# inputs	# hidden layers/ neurons	# outputs	Results[32]	Activation function	Learning algorithm	Performance measure
Hansen *et al.* (1992) USA	Sample of manufacturers Ratios, non-financial variables	40/40 *	12 9	NA.	1	80–91%	Hybrid of steepest gradient, Newton-Raphson	BP	MSE Average error Statistical models
Fanning & Cogger (1994) USA	Sample of firms Liquidity, cash-flow ratios	75/ 115 *	3	2:6–7	1	68–71%	Quadratic Sigmoid logistic	GANNA BP	% accuracy
Lenard *et al.* (1995) USA	Sample of firms Ratios and values of financial statements	70/10 *	8 (4)	1:5 (1:3)	1	78–90%	General. Reducent gradient optimizer	BP	% accuracy logit model
Anandarajan & Anandarajan (1999) USA	Sample of firms Ratios of financial statements	37/24 *	14	NA.	3	80–90%	Sigmoid	BP	Accuracy %
Koh & Tan (1999) USA	Sample of firms Financial ratios	300/ 30*	6	1:13	1	100 %	Sigmoid	BP	Accuracy %
Etheridge *et al.* (2000) USA	Sample of banks Financial ratios	863–892/ 215*	57	NA.	1	93–97% **	NA.	BP, CLN, PLN	OER Pearson's coefficient

* A holdout sample was employed. NA. = no answer. ** one year prior to failure.

Etheridge, Sriram, and Hsu (2000) compared the performance of three ANN approaches: backpropagation (BPN), categorical learning network (CLN), and probabilistic neural network (PNN) as classification tools to assist and support the auditor's judgment about a client's continued financial viability in the future (GC status). BPN has one or more hidden layers. CLN has normalisation and pattern layers. PNN has normalisation, pattern, and summation layers. The normalisation layer helps the pattern layer to develop correct classes for the input vectors. The pattern layer uses unsupervised learning to distinguish between categories. The summation layer helps in classifying data sets. The data was provided by a Big Six CPA-firm and consisted of 57 financial ratios for the years 1986–1988 for 1,139 banks in various regions of the U.S. The holdout sample consisted of 215 banks. They had three, two, and one year prior to failure models. When the overall error rate was considered, the probabilistic ANN (estimated error rate one year prior to failure 2.4%) was the most reliable in classification, followed by backpropagation (3.48%) and categorical learning ANN (7.15%). When the estimated relative costs of misclassification were considered, the categorical learning ANN was the least costly, followed by backpropagation and probabilistic ANN. The authors state that the results are only based on a single industry, banking, and this limits their generalizability to other industries.

5.1.4 Control risk assessment and audit fee

An auditor considers a huge amount of data when assessing the risk of the internal control structure of an entity failing to prevent or detect significant misstatements in financial statements. The relationships between internal control variables that must be identified, selected, and analysed often make assessing a control risk a difficult task. Therefore, control risk assessment (CRA) is a systematic process for integrating professional judgements about relevant risk factors, their relative significance and probable adverse conditions and/or events leading to identification of auditable activities (IIA, 1995, SIAS No. 9). Table 5.4 summarises the modelling issues of ANN literature pertaining to control risk assessment and audit fee problems.

Davis, Massey, and Lovell (1997) (see also Davis 1996) presented a construction of a prototype, which integrated an ES and an ANN. The rules were contained in the ES model basic CRA heuristics, thus allowing for efficient use of well-known control variable relationships. The 64 observations of auditors from Grant Thornton were used to develop and test an ANN model. The ANN training sample and testing sample each contained 32 observations. The ANN provided a way to recognise patterns in the large number of control variable inter-relationships that even experienced auditors could not express as a logical set of specific rules. The ANN was trained using actual case decisions of practising auditors. The input variables were judgement cues/variables from the general environment, computer processing, general computer, and accounting controls. The ANN model provided the auditor with information on how close a risk category border was. The testing resulted in an accuracy rate of 78%. The model is limited because of the number of cases and the fact that auditors were from the same firm.

Ramamoorti, Bailey, and Traver (1999) used quantitative (26 variables) and qualitative (19 variables) risk factors as input variables in the models. The risk was defined in an internal auditing context. The models were in the context of public state university departments. The sample consisted of 141 university departments and they used a training size, which covered 70% of the data and a holdout data of 30%. The quantitative data were downloaded from the University of Illinois financial and administration system. The qualitative risk factor values were elicited from audit staff using a pre-defined scale from 0 to 9. The eventual number of variables selected to construct the models were in the 7 to 18 range. The research project included a Delphi study and a comparison with statistical approaches and presented preliminary results which indicated that internal auditors could benefit from using ANN technology for assessing risk. The ANN models captured the top 25 risky departments at an accuracy rate between 72–84%. The results are based on a single industry, university, and this might limit the generalizability of the results to other industries as in the study of Etheridge, Sriram, and Hsu (2000).

Table 5.4 Summary of modelling issues of CRA and AF applications

Researcher Country	Industry Data type	Train/test set	# inputs	# hidden layers/ neurons	# outputs	Results[32]	Activation function	Learning algorithm	Performance measure
Davis (1996) USA	Sample of firms IC risk data, observations of auditors	37/27	107	1:5	1	56% 86%	NA.	BP	% absolute error
Davis *et al.* (1997) USA	Sample of firms IC risk data, observations of auditors	32/32 *	210	1:30	1	78%	Sigmoid	BP	RMSE Pearson's coeffic., accuracy %
Ramamoorti *et al.* (1999) USA	Sample of university departments Qualitative, quantitative risk factor data	100/ 41 *	10	NA.	1	72–84%	NA.	BP	R-squared %, Delphi overlap
Curry & Peel (1998) UK	Sample of electronic firms Financial, non-financial data	96/32 86/42 64/64 *	25	1:2 1:3 1:4	1	56% 69% 65%	Sigmoid	BP	MSE

* A holdout sample was employed. NA. = no answer

Curry and Peel (1998) provided an overview of the ANN modelling approach and the performance of ANNs, relative to conventional ordinary least squares (OLS) regression analysis, in predicting the cross-sectional variation in corporate *audit fees* (AF). The data was derived from a sample of 128 unquoted UK companies operating in the electronic industrial sector. The data was collected from the years 1986–1988 from Kompass 1990 and Macmilla's Unquoted Companies 1990. The audit fee, the dependent variable in the study, must be disclosed (under the UK company law) in a note to a company's annual statements. The input variables were related to auditee size, audit complexity, audit risk, auditee profitability, and auditor size. They tested the ANN's ability to generalise with three separate training and holdout samples: 96 and 32, 86 and 42, and 64 and 64 companies in the samples. The estimation accuracy for the holdouts was 56.2%, 68.6%, and 64.5%. These results were achieved with four hidden neurons. The ANN models also exhibited better forecasting accuracy than their OLS counterparts (31.9%, 59.7%, and 52.9%). The authors state that, although the study demonstrated that the ANN outperformed conventional linear techniques in forecasting audit fee data, there is a need for further research relating to the optimal architecture of ANNs, namely the number of hidden layers and neurons.

5.2 Issues in ANN modelling for analytical review

Auditing ANN research started in the beginning of the nineties. The applications have been developed mainly in the United States of America (16 studies). Some development work has been done also in Taiwan (1 study), and the United Kingdom (1 study). Many of the cited articles applied the ANNs as an extension of a previous statistical model-based study, and used publicly available stock market data. Table 5.5 summarises the findings of analytical review ANN studies in a chronological order by application areas. Most of the authors state that ANNs have the potential to improve AR procedures.

5.2.1 Data

Both quantitative and qualitative *data* were used as input variables in the applications. Most of the data is quantitative and gathered from public databases. Financial statement values and ratios and monthly account values were used mostly as quantitative input variables. Opinions and observations of auditors or red flag data defined by SAS in the US were included in quantitative input variables. In ten cases the sample data was gathered from various industries. Seven cases focused on one industry, such as the bank, manufacturing, university, or electronic industry. Focusing on one particular industry is one way of conducting analytical procedures (see section 3.2.1).

Many of the studies applied the ANNs as an extension of a previous statistical study. Therefore, in some cases the data was fairly old. Also to use only publicly available information might simplify the auditor's task too much and does not reveal all problems in the client company.

5.2.2 Samples

All but two applications reviewed in this survey had relatively small data samples to train and test the models with. ANNs can be more appropriate for large data sets. The material error ANN-model of Busta and Weinberg (1998) had bigger data sets than other studies in the review, but their data set was simulated. A Big Six CPA firm provided Etheridge, Sriram, and Hsu (2000) with the data set of financial ratios from 1,139 banks for a GC-application.

It seems that it is common to use small data sets in the early developing phase of ANN models. One explanation for the small data sets might be the high supply expenses of the data. Small data sets also limit the number of possible inputs into the models. However, it is critical to have both the training and test set representative of the population or underlying mechanism. The selection of the training and test set may affect the performance of an ANN. There are no mathematical rules for determining the sample sizes, only some rules of thumb derived from experience, see e.g. Basheer and Hajmeer (2000). Most authors select them based on the rule of 90% vs. 10%, 80% vs. 20%, or 70% vs. 30%, etc. 50% vs. 50% was used in four cases. Some choose them based on matched pairs. The median training samples and testing sample sizes are 80 and 44, respectively.

5.2.3 Architecture

The determining of the numbers of *input neurons*, *hidden layers* and *neurons,* and *output neurons* is problem dependent. The ANN-architecture must be sufficient for the correct representation of the problem, but sufficiently low to allow generalisations (Swingler 1996, 31). For example, all the income statement account values could be selected as input variables in many applications. However, some selection of variables might have helped an ANN to get better results. For example, in my studies I used different selection criteria for the input variables, i.e. literature, CPA-auditor, the most stable accounts compared to previous year's values, internal auditors, and management of the organisation.

The hidden neurons in a hidden layer allow the ANN to detect the feature, to capture the pattern in the data, and to perform complicated non-linear mapping between input and output variables. Linear relationships between the variable are very often a simplification of the natural financial data. The number of output neurons is relatively easy to specify as it is directly related to the problem under study. The median audit ANN architecture consists of 14 inputs with 6 neurons in one hidden layer and one output neuron.

The majority of the auditing ANN-applications under study had one output neuron. This is common in classification models whereas prediction models might have one or more neurons in the output layer. The choice network architecture is mostly subject to trial and error without guarantee of reaching an optimal solution. However, properly configured and trained ANNs appear to make consistently good classifications or generalisations.

5.2.4 Learning parameters and performance measurement

There are no clear guidelines on the selection of the performance and activation function. All but one study used the supervised *learning* paradigm. Fanning and Cogger (1998) used the self-organising learning paradigm in a management fraud application. Most studies used the backpropagation learning algorithm or some variation of it. A sigmoid[35] or a modified sigmoid *activation function* was used in nine studies; a logistic function was used in three studies. Seven applications used some other activation function.

Almost every model was evaluated with a holdout sample. The performance results of the ANN models in the previous review are based on holdout samples. The GC and FD applications got the best results with the prediction accuracy between 68–100%. All in all the performance accuracy varied between 30–100%. The *performance measurement* of the models varied a lot. Root mean square error, prediction accuracy, and average error were mostly used. In some cases also a comparison to the traditional statistical methods was used. It is not always clear whether the performance is calculated from actual data or scaled data. Researchers should emphasise the measurement of the performance from the auditing point of view. Visualisation of the results might help auditors to see the benefit of the ANNs.

[35] The author did not tell what kind of sigmoid function they used in the study.

Table 5.5 Findings of analytical review ANN studies

Authors	Summary of findings
	Material errors
Coakley & Brown (1991a)	Results tentatively suggest that the ANN recognised patterns across financial ratios more effectively than financial ratio and regression methods.
Coakley & Brown (1991b)	The use of ANNs for pattern recognition across related financial data sets may be viable.
Coakley & Brown (1993)	ANNs applied as a forecasting tool seem useful for identifying patterns that can indicate potential investigations of a firm's unaudited financial data in the current year.
Wu (1994)	The results strongly suggest that ANNs can be used to identify firms requiring further auditing investigation, and also suggest future implications for intelligent auditing machines.
Coakley (1995)	Results suggest that the use of an ANN to analyse patterns of related fluctuations across numerous financial ratios provides a more reliable indication of the presence of material errors than traditional analytic procedures or pattern analysis, as well as providing insight into the plausible causes of the error.
Busta & Weinberg (1998)	The results show that if data are contaminated at a 10% level or more ANNs will detect this 68% of the time. If the data are not contaminated, the test will indicate that the data are "clean" 67% of time.
	Management fraud
Green & Choi (1997)	The results show that ANNs have significant potential as a fraud investigative and detection tool.
Fanning & Cogger (1998)	The study showed that ANNs offer better ability than standard statistical methods in detecting fraud.
Feroz *et al.* (2000)	The ANN models classify the membership in target versus control firms with the accuracy of 81% across the board.
	Going concern and financial distress
Hansen *et al.* (1992)	The results indicate that the ANN model predicted more consistency than the other advanced statistical models used in the study.
Fanning & Cogger (1994)	The results indicate the potential in time savings and the successful classification results available from an ANN processor.
Lenard *et al.* (1995)	The ANN model is proposed as a robust alternative model for auditors to support their assessments of going concern uncertainty.
Anandarajan & Anandarajan (1999)	The results indicate that the ANN model has a superior predictive ability in determining the type of going concern audit report that should be issued to the client.
Koh & Tan (1999)	The results suggest that ANNs can be a promising avenue of research and application in the going concern area.
Etheridge *et al.* (2000)	When the overall error rate was considered, the probabilistic ANN was the most reliable in classification, followed by backpropagation and categorical learning ANNs. When the estimated relative costs of misclassification were considered, the categorical learning ANN was the least costly, followed by backpropagation and probabilistic ANN.
	Control risk assessment
Davis (1996)	The results show that the ANN models have been able to acquire, represent, and store the auditor's knowledge for selecting relevant information to assess control risk.
Davis *et al*. (1997)	The model was able to classify the auditors' control risk assessment in the testing sample with a 78% accuracy rate.
Ramamoorti *et al*. (1999)	The ANN, logistic regression, and stepwise multiple regression produced results that compared favourably together in terms of the variance in the rankings obtained.
	Audit fee
Curry & Peel (1998)	The ANN models exhibited better forecasting accuracy than their OLS counterparts, but this differential reduces when the models are tested out-of-sample.

5.3 Summary of ANNs in analytical review

This chapter presented the current state of the ANN-applications within the area of auditing. Application areas were detecting material errors and management fraud, supporting going concern decision, assessing internal control risk, forecasting audit fee, and determining financial distress problems. All the researches fit into analytical review (AR) procedures. Most of the authors state that the ANNs have the potential to improve AR procedures.

Many important qualitative variables such as management ability and future plans were not formally incorporated into the models. Bias in data samples might limit the generalizability of the results. Most of the reviewed applications used quite small data sets. Hence, a good performance generalisation in many cases is insufficient. In some cases the data was also quite old. However, in many cases the results serve the AR procedures such as comparison of the company data and the industry data or forming expectations to compare with company's data. Indeed, researchers should focus on developing models with data sets that are not publicly available in order to investigate client data more deeply. For example, in Koskivaara and Back (2003) and Koskivaara (2004c) ANNs operate with the operational monthly data. Therefore this study gives further information of the feasibility of ANNs in supporting the AR procedure of operational data.

The determining of the ANN architecture is problem dependent and is still part of the art in ANNs. Further research is needed to find out guidelines for optimal ANN architectures in auditing. Most studies used the BP algorithm while others employed some variants of it. There is clearly a need for models and networks that can handle the fluctuating nature of financial data. For instance, my prototype uses the supervised training method with the resilient backpropagation (RPROP) training algorithm with sigmoid function, which is one of the most efficient algorithms on pattern recognition problems (Demuth & Beale 2000).

The performance measure of applications varies a lot and it is not always clear whether the performance is calculated from actual data or scaled data. More emphasis should be put on the measurement of the performance of ANNs. Although the results are encouraging, further research is needed to study whether ANN models are better than traditional AR procedures in use. This is one aim of my study: to receive further evidence on whether the ANN is better than the conventional AR procedures. Besides, the visualisation of the results might give some added value to auditors. Therefore, another aim is to find an appropriate way of illustrating data to reveal unexpected patterns in it.

5.4 Limitations of ANNs

The presumptions are that ANNs learn in a semi-autonomous way and that ANNs allow us to bypass the need for careful thought and judgement (Curry & Peel 1998). In fact, nothing could be further from the truth. Users of ANNs are faced with modelling and parameter issues. To find a model for a monthly account predictor was one task in my study. Another task was to build a system based on that model.

A second key issue facing the users concerns the numbers of layers and neurons. There is no simple method for determining the numbers of hidden units a network requires, in practice the only serious alternative is to choose the number of layers and neurons experimentally (Coakley 1995). This trail and error development decreases the attractiveness of ANNs in practice.

The third obstacle with ANNs is that the internal structure makes it difficult to trace the process by which output is reached. This is why ANNs lack clear explanatory capabilities. The justifications for the results with the help of the connections' weights are difficult to obtain. Therefore, ANNs need a careful training and testing so that they do not find a pattern from the data, which is not relevant to the problem solving. In the case where the traditional statistical methods give as good results as ANNs, the traditional statistical methods are more effective because they require less computational power.

6 BUILDING ANNA

This and chapter seven give an overview of my ANN-based system (=ANNA) development process. Table 6.1 summarises the aims and the results and purposes of the different publications. The following discussion shows the research spiral I have followed: what are the reasons and lessons learnt from one phase that lead my research to the next phase.

6.1 Basics of ANNA

ANNA was developed in a step-by-step manner. The idea of ANNA is to take advantage of the data that already exist in the companies and use it in a monitoring and controlling context. A well-known fact is that companies are collecting monthly reports to support their operations. Therefore ANNA is built on the following basic assumptions:

- The broad and extensive review of previous researches indicates that most authors state that the ANNs have the potential to improve AR procedures, and that there are some open ends in the field.
- The ANN technology seemed an attractive method for forming expectations based on company data.
- The learning from the data could support auditors' experience and knowledge about a client company.
- The solution idea should be built based on data which already exist in companies and which could better be used in the auditing context.
- A suitable prediction model could form the inner core in the ANN-system.
- The selection of variables or accounts should be flexible.
- The results of ANN-system should be compared with the conventional AR procedures in use.

In brief, ANNA is an ANN-based tool for analysing monthly account values. ANN's predictions or account expectations are based on function approximation (modeling) capability from the previous years' monthly account values. It uses a flexible one-step-ahead prediction model described in section 6.2. It predicts comparison values with five conventional AR methods. The results are illustrated both by a graph on the computer screen and in the tables. ANNA belongs to the class that studied material errors in accounting data with ANNs. As mentioned in the introduction, according to Fischer's (1996) classification on gathering and evaluating audit evidence, my study belongs to the development of new tools and methods for auditors. According to Fraser, Hatherly, and Lin's (1997) classification of the AR techniques, the ANN-based AR tools belong to the AQTs.

Table 6.1 Purposes of the publications

Koskivaara (2000a) and Koskivaara (2004b)
How to model and illustrate ANN-based monthly account patterns for auditing purposes?
The one-step-ahead prediction model seems promising for recognising the dynamics and the relationships between monthly account values. The results should be illustrated both by a graph on the computer screen and in the tables. Also SOM has potential for clustering and visualisation data sets for analytical review purposes.
Koskivaara (2000b)
How to pre-process the data?
Linear scaling has the advantage of preserving the relative position of each data point along the range. The best result was achieved when all the data were scaled either linearly or linearly on a yearly basis.
Koskivaara and Back (2003) and Koskivaara (2004c)
How to construct and evaluate the ANNA-system in AR context, and how good is the ANNA-system based on these evaluation criteria?
To suit different kinds of situations ANNA needs to be flexible: sample size, number of variables, training options. A training algorithm has to suit a modeling (function approximation) problem. ANNA outperformed the second best method in 19 out of 20 cases. The results indicate that ANNA had the best and most consistent prediction accuracy compared the other methods used in the studies.

6.2 A model for ANN-based monthly account predictor (ANNA.01)

The practical study started by modelling an intelligent system based on ANNs for monitoring account values of monthly balances of a manufacturing firm. The account values of the monthly balances were regarded as a time-series and the aim was to recognise the dynamics and the relationships between accounts, and to predict the value of a certain account one-year in the future. For this purpose, it is possible to take a set of account values from a number of periods, for example three successive periods x_{t-2}, x_{t-1}, x_t, to be the input vector to a network and use the next value x_{t+1} as the target value as illustrated in Figure 6.1. Therefore, the data is presented to the network in chronological order. This is called a one-step-ahead prediction. By moving along the time axis it is possible to create a training data set consisting of many sets of input vector values with corresponding target values (Bishop 1995, 302–303). These assumptions mentioned above led me on to the world of supervised learning. The multilayered network can be applied directly to the problem described above (Bishop 1995, 302).

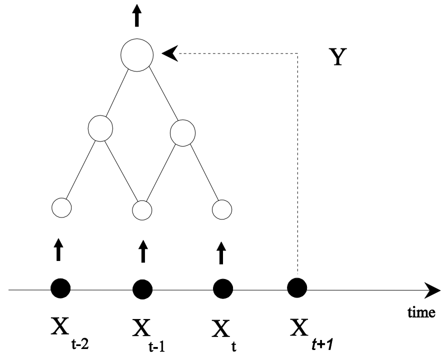


Figure 6.1 Generating data sets for multilayered network (modified from Bishop 1995, 303)

ANNA.01 was built using a commercial development package, Neuralyst v1.4, developed by Cheshire Engineering Corporation. Neuralyst was selected for the implementation platform because of its user-friendly features to start the design and development work. ANNA.01 uses the backpropagation algorithm described in section 4.4.1 for training the network. Neuralyst has several activation functions such as sigmoid, hyperbolic, linear, and gaussian function. Sigmoid function was selected for ANNA.01.

ANNA.01 used actual 66 monthly income statements from a manufacturing company. ANNA.01 was trained using the first 54 months' financial data. The rest of the data were held out for testing.

The one-step-ahead prediction model in a monthly account values environment was firstly tested with a subset of Coakley and Brown's (1991b) accounts (Koskivaara, Back & Sere 1996). Their study was the first attempt to apply an ANN to monthly financial statements. Following this account selection was a safe way to penetrate this application domain with a novel ANNA-model.

ANNA.01 was tested with two successive time periods to predict a third one. Therefore, ANNA.01 had 18 input data streams to represent the data from the previous two months for each of the nine account values. An additional 12 input neurons were added to indicate the month of the input data. This input value of the neuron was set to one if the neuron corresponded to the month of the data, and zero otherwise. The output layer had nine neurons which represented each account value.

Numerous experiments were conducted to find the network architecture that minimised the RMSE[36] between the target and the output account values for the 12-month holdout sample. The experiments included altering, for example,

[36] RMSE = root mean square error (see section 4.4.1).

the number of hidden layers or numbers of neurons per hidden layer. However, the training epoch that is the number of processing runs through the complete training data was limited to 10,000 to give an upper bound for the training time. The final configuration selected had one hidden layer with 16 neurons.

Normally, the input information to a network is pre-processed. Firstly, the input information has to code into numeric values. Secondly, the numeric input data may be pre-processed in a way that the network's learning task becomes easier. ANNA.01 used the same pre-processing method as in the study of Coakley and Brown (1991b): first each input vector was normalised by dividing it by that vector's length and then the account balances were scaled to a [0,1] range. The selected network worked best when its inputs ranged from 0 to 1. Sometimes the output data are also processed. Post-processing is the inverse of the pre-processing transformation. Output values of ANNA.01–04 were not post-processed.

The average difference between targets and output was 6% in the testing period. Therefore, the average prediction ability of a neural network to monitor the non-linear dynamics of monthly balances was good. The results were encouraging, but resulted also in new questions. Does this model work with other accounts? Are there any other pre-processing methods? How do we analyse the prediction capability in a practical way? For that purpose, ANNA was further designed and tested.

6.3 New account selection for ANNA-model (ANNA.02)

To get further evidence on ANNA's prediction capability it was tested with new account selection criteria (Koskivaara 1997). This time a Certified Public Accountant, CPA-auditor, assisted in choosing the accounts. The accounts that presented the major and the most interesting monthly balance categories were selected. Based on this selection, two different kinds of models, ANNA.021 and ANNA.022, were built with the same data set as in the first step.

ANNA.021 used two previous months to predict a third one. Therefore, ANNA.021 had 18 input data streams to represent the data from each of the nine financial accounts. ANNA.022 uses four previous months to predict a fifth one. ANNA.022 had 36 input data streams to represent the data from each of the nine financial accounts. The month indicator was the same as in the first step. To summarise, ANNA.021 had 30 input data streams and ANNA.022 had 48 input data streams in the input layer. Both models have nine neurons in the output layer, one for each financial account.

The data was pre-processed by a linear transformation to a [0,1] range. The linear transformation has the advantage of preserving the relative positions of each data point along the range (Swingler 1996, 31). I selected the linear scaling procedure because I did not want to lose any relative information between account values before the data is fed into the ANN-system.

The average difference between all the target values and all the output values was 12%/ANNA.021 and 10%/ANNA.022 in the testing period. The results indicated that the average prediction ability of the ANN models to monitor the non-linear dynamics of account values in the monthly balances was good. ANNA.022 was slightly better than ANNA.021. However, the results on average were slightly worse than in the first step. Therefore, the decision was made to expand the research to include more companies in the study, to investigate whether there are more sophisticated networks than the standard sigmoid based backpropagation network and try to get more monthly data for training purposes.

6.4 New parameters for ANNA-model (ANNA.03)

To take the advantage of the new training approaches of backpropagation network, which the Neuralyst v1.41 provided, once again ANNA was tested (Koskivaara 2000a). Especially, the *adaptive learning rate* seemed an attractive feature to support the ANN learning process.

Moreover, I managed to receive more monthly data to build and test the models. This time the ANN models used 72 monthly balances of a manufacturing firm. The last 12 months data were pertained for testing purposes.

I used the same account selection and the same pre-processing method as in step two (ANNA.02). Therefore, ANNA.031 and ANNA.032 were modified from ANNA.021 and ANNA.022. As mentioned earlier, the accounts were chosen with the help of a CPA-auditor in the way that they presented the major and the most interesting monthly balance categories. The following results were received.

The average difference between all the target values and all the output values was 6% for both models. The prediction capabilities of the models according to the accounts are presented in Table 6.2 (cf. Koskivaara 2000a: Table 4[37]).

Table 6.2 presents the prediction capabilities (average difference in per cents, difference range in per cents [min,max], and differences less than 10%)

[37] ANNA.031 corresponds to Model 1 and ANNA.032 corresponds to Model 2 in Koskivaara (2000a).

of the ANN models according to the accounts. Table 6.2 shows, for example, that the average difference for the output and the target values of receivables in ANNA.032 is 2%, and the difference range in per cents is [-1,5]. This means that all the values (100%) are inside the limit of 10% (≅ true and fair view). Overall 11% (14 data item out of 108) of the test data were not inside the 10% limit.

Table 6.2 The prediction capabilities of the models according to the accounts (Average difference/difference range/less than 10%)

Chosen accounts	ANNA.031	ANNA.032
Net sales	4/[-13,5]/92	5/[-10,12]/83
Materials + Change in inventory	3/[-7,12]/92	5/[-8,9]/100
Personnel costs	3/[-5,5]/100	2/[-4,4]/100
Gross margin	4/[-9,5] /100	5/[-11,8]/92
Administration	1/[-1,2] /100	1/[-1,4]/100
Total indirect	5/[-6,15]/83	6/[-7,15]/83
Operating profit	5/[-11,9]/92	5/[-7,12]/92
Receivables	2/[-3,4]/100	**2/[-1,5]/100**
Trade debts	28/[-23,88]/0	21/[-36,51]/33

Overall, we can conclude that ANNA.031 and ANNA.032 produce good results from the chosen accounts (materials + change in inventory, personnel costs, gross margin, administration, operating profit, and receivable accounts). The accounts net sales and total indirect attained by the ANN model are also on average within the true and fair view. The worst case was the account trade debt. The results of trade debt were fairly poor when comparing them to other account values. The reason for this might be that the trend of the trade debt does not follow the dynamics of the other account values and therefore the ANN-model was not able to predict the trade debt values as well as the other account values. For example, the average change of the trade debt values from the last training period to the test period is big, 20%. At the same time, the average change of all other accounts from the last training period to test period is 6%. Another reason for the poorer learning ability might be the fact that the trade debt account is the only big credit account in the data. Moreover, one explanation for this odd behaviour of the trade debt might be that the manufacturing company has bought raw material when it has been inexpensive or easily available. Also, when the company has been a little short of money it has delayed the payment of the trade debt. Those were some explanations why the trend of the trade debt was not as stable as the trend of other accounts. When an account behaves like this it is of course also a hint to the auditor to pay more attention to that account. ANNA.032 gave slightly better results with the smaller difference range. So with more months and with a more sophisticated learning method the results were better.

ANNA.032 was selected for further investigation. It was tested by seeding fictitious sales in May, which means that the net sales and the receivables should go up while others do not. ANNA.032 detected the error, when the amount of seeded error was 5% of the year's average net sales in two ways. In the case of fictitious sales the RMSE went up from 0.142205 to 0.162205 and the target values were unusually high and they were not inside the 10% limit in May. One might argue that the amount of seeded error is relatively high and the change in the RMRE is not so informative. Therefore the results were illustrated both by a graph on the computer screen and in the tables. This illustration is for monitoring and should tell whether further audit is required or not. More detailed results can be found in (Koskivaara 2000a).

The weaknesses of these ANNA versions were that their results were not compared to any other methods and that the results were counted from the scaled values. Therefore, the design of the comparison methods for ANNA was started. Before that I studied the effects of the different pre-processing methods.

6.5 Pre-processing data for ANNA-model (ANNA.04)

Successful data pre-processing speeds up neural networks learning and gives better and more general results. Swingler (1996, 48–50) suggests four pre-processing models for data: *linear scaling, linear normalisation, logarithmic scaling,* and *softmax scaling*. A linear scaling function moves the data from one interval to another interval. In a linear normalisation the data can be normalised, for example, using the average and the standard deviation (i.e. the inputs and targets have the zero average and unity standard deviation). In logarithmic scaling the data is first moved to the logarithmic values and then scaled from one interval to another interval. Softmax scaling squashes unevenly distributed data into an appropriate range: values, which are far from the average, are squashed most and a larger standard deviation requires a larger degree of scaling.

Linear scaling has the advantage of preserving the relative position of each data point along the range. This means that with the linear scaling the original and normalised values are one-to-one. Therefore, this scaling does not move any relevant information from the data before it is fed into the neural network. Hence this scaling was selected for further investigation. Especially, the four different alternative linear scaling methods were investigated: *all together, on a yearly basis, on a company basis,* and *on a yearly and company basis*.

ANNA.04 was built and tested using financial statements from 31 manufacturing companies over four years. Five of the companies were selected to the test set. From the AR point of view the model compares information of a firm with similar information for the industry in which the organisation operates. A large Finnish bank provided the sample data for the research. Sixteen income statement accounts, which represent the major financial statement categories and the average number of staff, were included in the study. The training data set consisted of 25 companies and their financial statement values and other variables from four years. The remaining six companies and their financial statement values and variables were divided into the training and the testing set so that the first three years were put into the training set, and the fourth year's variables were left for testing.

Firstly, *all* the data were scaled linearly. The reason for this was that the existing dynamics and relationships between all the variables should be kept totally one-to-one. In a case where there were any trends and relationships between the years and companies the ANN model might recognise them. In this case all data should always be checked and maybe pre-processed before entering new data into the support system.

Secondly, the data were pre-processed linearly *on a yearly basis*. The assumption in this case was that it is easier for the ANN model to learn the dynamics and relationships between the variables if the year effect was minimised on a yearly basis pre-processing. In this case when the year or other selected period is closed there is no need for data re-pre-processing.

Thirdly, the data were pre-processed linearly *on a company basis*. The assumption in this case was that it was easier for the ANN model to learn the dynamics and relationships between the variables if every company has its own sliding scale. All new data should be pre-processed when entered into the support system.

Fourthly, the data were pre-processed *on a yearly and company basis*. Correspondingly, the assumption in this case was that it is easier for the ANN model to learn the dynamics and relationships between the variables if every company had its own sliding scale on a yearly basis. The data from the period are ready after one pre-processing.

The results differed depending on what pre-processing method was used. The best results were achieved when all the data were scaled either all together or on a yearly basis. More detailed results can be found in Koskivaara (2000b). The weakness of the study was that the results are not compared to any other methods.

6.6 SOM-model

I also wanted to explore how an unsupervised neural network, especially Kohonen's self-organising map (SOM), can be used in the visualisation of accounting data for analytical review purposes. I used the SOM for clustering ten years' monthly income statements of a manufacturing firm. I used the same accounts as in ANNA.02.

I studied the account clusters and the yearly tendency in the data with the help of hexagonal lattice and the feature planes of the maps. I found that the data sets of various accounts and various years formed their own groups. In Figure 6.2, the SOM-map is on the left and on the right I have identified and named clusters according to the accounts these clusters contain. Thorough analysis of the accounts and the results are in Koskivaara (2004b). I labelled the last two years' accounts to see whether the accounts are close to each other and to name the clusters. I identified four main clusters: *revenues, margins, costs,* and *trade debts.* These clusters consist of the input variables with similar characteristics.

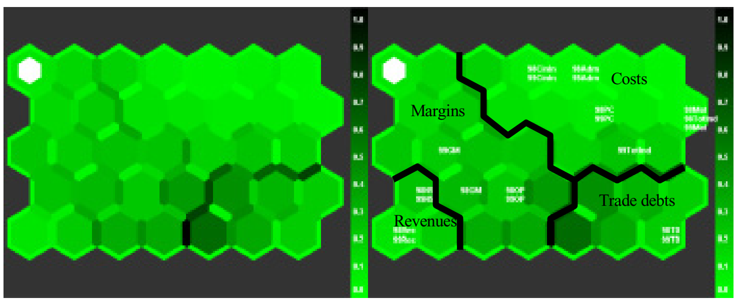


Figure 6.2 The SOM-map and the clusters on the map

This algorithm can discover features that may be used to classify a set of input vectors. Or, this algorithm may reveal something about the data that might be left hidden otherwise. Therefore, the SOM can be a visual aid for classifying and clustering accounting data sets, and it reveals if some cluster contains data that a priori should not be in it. Hence, it can be used for signalling unexpected fluctuations in data. SOM could also be a tool for studying the relationships among elements of information. Therefore, the SOM could be a possible technique embedded in the continuous monitoring and controlling tool. However, it could take some time before auditors accept this kind of approach to support their analytical procedures. As discussed earlier, the SOM could also be used for classifying the accounts into homogenous groups, which then can be inputs for a supervised ANN-system.

6.7 Construction of ANNA

The earlier version of ANNA seemed promising for supporting analytical review process. However, I needed a more flexible and effective environment to evaluate ANNA than Neuralyst was able to provide. Therefore ANNA was prototyped. The elements of ANNA are discussed in more detail in the following sections.

6.7.1 Learning algorithm

ANNA uses the supervised training method with the resilient backpropagation (RPROP) training algorithm with sigmoid function, which is one of the most efficient algorithms on pattern recognition problems (Demuth & Beale 2000). In my study I try to recognise patterns of different accounts by training an ANN with input–output data in order to approximate the underlying rules relating the input account values to the output account values. The selection of the RPROP learning algorithm is reasonable, since this algorithm was designed to overcome the difficulties caused by training with sigmoid functions, which have very small slopes when operating far from the centre point. Therefore, the RPROP has similarities to the adaptive learning rate described in section 4.4.1. Furthermore, the RPROP is not very sensitive to the settings of the training parameters, which might help to find some general values for the training parameters. The RPROP uses a batch training algorithm and therefore requires minimal storage. As mentioned in section 4.4.1, in the batch training weight updating is performed after the presentation of all the training examples that constitute an epoch. Moreover, as the account values in my prototype are presented to the network in a time series manner, the use of holistic updating of weights makes the search of weights occasional in nature. This in turn makes it less likely for the network to be trapped in a local minimum.

Training parameters included in the model are *training cycles*, *weight decay*, *delta0,* and *max delta.* Training cycles refer to the number of training runs needed to complete the task. Smaller values can give the prediction results faster which can sometimes be advantageous. The weight decay, which is used by default, is very useful in the training, as it reduces overlearning and therefore increases the generalisation ability. Delta0 and max delta are parameters specific to resilient backpropagation. They are the initial step size and the maximum step size, respectively. The optimal parameter values vary depending on the data, the amount of training cycles, and the network architecture. Figure 6.3 illustrates the options window in the system in which these parameters are determined before the ANN is trained.

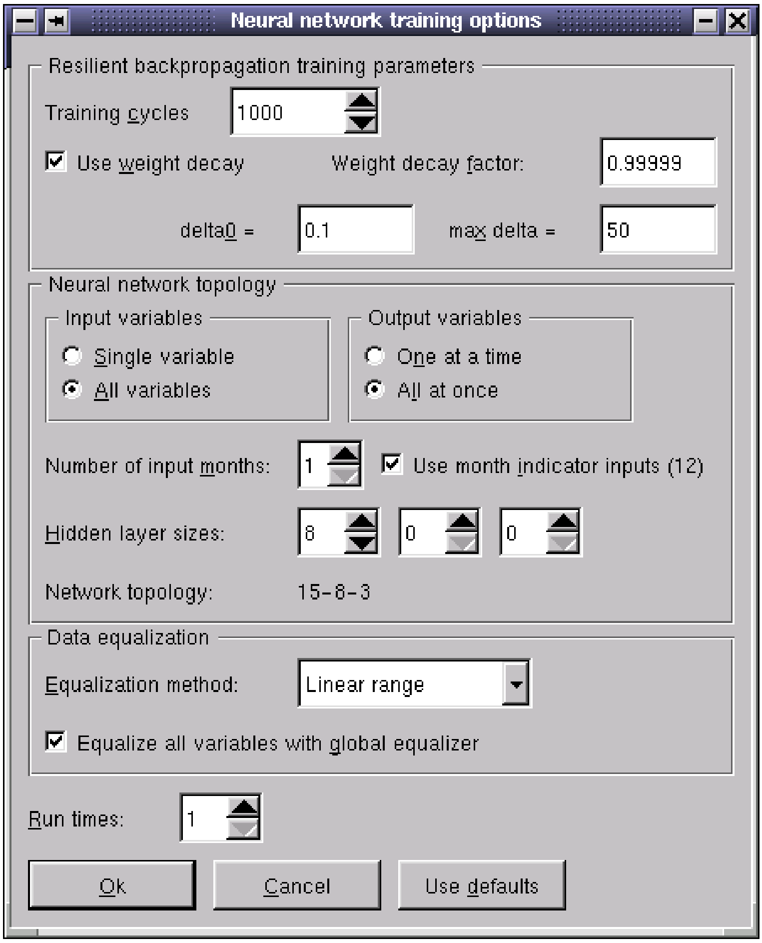


Figure 6.3 Option-window in the system (screen copy)

6.7.2 Choice of ANN architecture

The ANN architecture used in ANNA is a fully connected feedforward network with a maximum of three hidden layers. The specific network architecture is determined by a number of choices.

ANNA has two options regarding the choice of input variables given to the network: *the single variable option* (Figure 6.4) and *the multiple variables option* (Figure 6.5). With the single variable option, each monthly account value is predicted from the previous months' values of that account only. With this option, it is possible to predict only that single variable, and therefore there can be only one output variable. With the multiple variables option, each account is predicted from the values of all accounts in previous months. The output variables can be done either by one account at a time or all accounts at the same time. In the former case a separate ANN will be trained for each account. In the latter case, there is a single ANN with all the accounts in the input and output layers, i.e. multiple input–multiple output (MIMO) system.

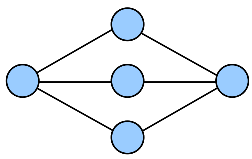


Figure 6.4 Single variable

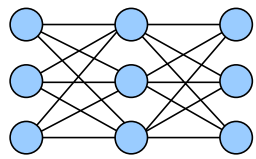


Figure 6.5 Multiple variables

The number of input months in ANNA may vary from one to twelve. Therefore, ANNA uses a flexible one-step-ahead prediction model. A month indicator shows the number of previous months given as inputs. This means that in ANNA the monthly account values of the organisation are regarded as a time-series and the aim is to recognise the models in the monthly account values. The goal is to predict the value of a certain account for one year in the future.

6.7.3 Data processing

In the ANNA-system the data is pre-processed by a linear transformation to a [0,1] range either one account at a time (locally) or all the accounts at the same time (globally). When only one account at the time is equalised, the value 0 would refer to the smallest value in the particular account, and 1 to the largest. When all the accounts at the same time are equalised, 0 and 1 would refer to the minimum and the maximum of the entire data. As mentioned earlier, the linear transformation has the advantage of preserving the relative positions of each data point along the range (Swingler 1996, 31). I selected this scaling procedure because I wanted to hold the relative information between account values when feeding them to the network in order to let the ANN approximate the unknown input–output mapping of the account values. The output data of ANNA is post-processed. The post-processing helps the comparison of results and the outputs of ANNA resemble actual values, which might be an advantage in analytical review.

6.7.4 Evaluation criteria of ANNA

As mentioned in the introduction, research has consistently indicated that auditors prefer simple scanning, reasonableness tests, and ratio analysis to a sophisticated statistical or mathematical model in AR (e.g. Ameen & Strawser 1994; Cho & Lew 2000; Fraser, Hatherly & Lin 1997; Turpiainen 1994; Lin, Fraser & Hatherly 2003; Lin & Fraser 2003). This is true also in the organisation that provided me with the data for building and evaluating ANNA. One explanation for the popularity of SQTs might be that they are quite straightforward and require only a few calculations. The AR procedures and techniques were discussed more thoroughly in sections 1.2 and 3.2.

There are three reasons for selecting the following evaluation metrics for the results achieved with the ANN-system. First, as SQTs play a fundamental role in the AR process in practice, they are selected as the basis for the comparison metrics for the ANN-system. Second, these comparison metrics are based on the nature of analytical procedures in auditing given in ISA 520. Third, the inner core of the ANNA is based on time trend, which could be classified either under the SQT or AQT depending on the utilized technique. The abbreviations in brackets and explanations of the comparison metrics used in this study (= traditional AR methods) are as follows:

- Previous year's value (PYS) is the same value from the previous year.
- Average of previous years (AVE) is the average of the same account values from all the previous years.
- Average delta prediction (DELTA) calculates an average of the monthly changes from previous years and makes predictions by adding the change to the account value of the previous month.
- Zero delta prediction (ZERO) values are the same as in the previous month.
- Combined trivial method (CTM) combines the above mentioned simple prediction models.

The predictions of these comparison metrics are made immediately when a data file has been loaded into the system. The predictions are displayed in the data analysis window. After the training of the ANN-system is finished it is possible to analyse the ANN’s prediction results together with these comparison metrics in the data analysis window. In effect, I have six populations, one population associated with each method (i.e. ANN, CTM, DELTA, AVE, PYS, ZERO).

6.7.5 System environment

ANNA has been developed and tested in a Linux environment. It works currently under Linux with the KDE1 desktop environment. It has four main windows: a file, a data analysis, an option, and a help window. Figure 6.6 depicts the main window of the system.

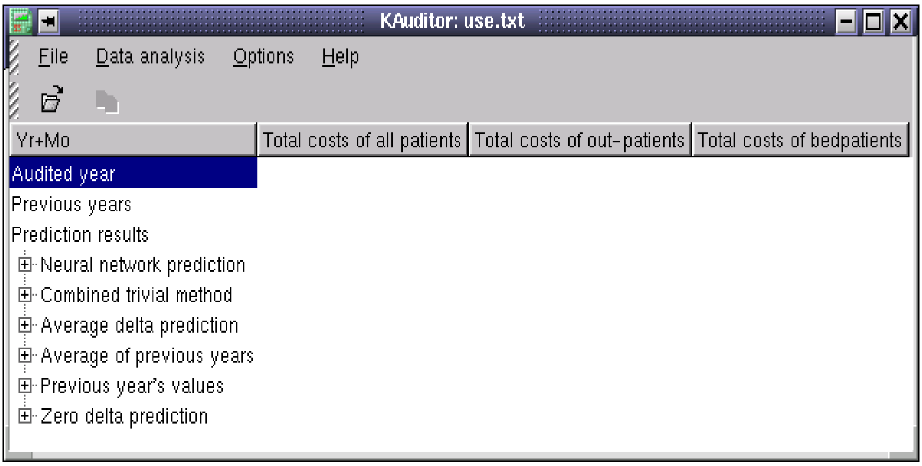


Figure 6.6 Main window of the system (screen copy)

It is possible to open the loaded data files from the file window. This data analysis window allows either viewing the prediction error statistics (see Figure 6.7) or plotting the prediction results together with the actual values (see Figure 6.8). It is also possible to illustrate the limits for the values achieved with ANNA, see dotted lines in Figure 6.8. The training settings for the ANN-system are in the option window (see Figure 6.3). Once the training settings for the model are determined it is possible to train the ANN model from the data analysis window. After the training is finished it is possible to view the ANN's prediction results in the data analysis window. When the network has been trained with a particular data set, it can be saved as a file. A previously saved ANN can be restored from the file window. From the help window one can find the manual of the system.

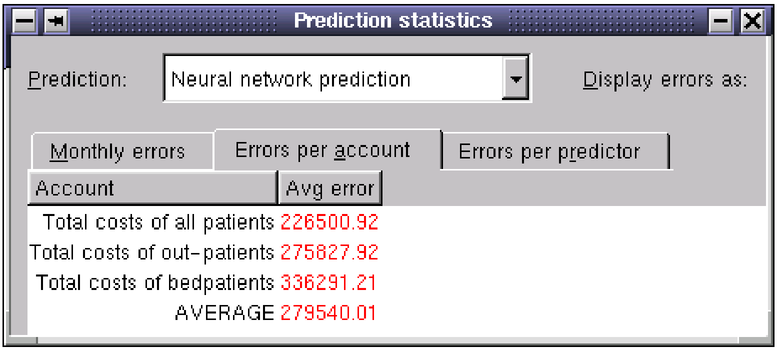


Figure 6.7 Prediction statistics of the system (screen copy)

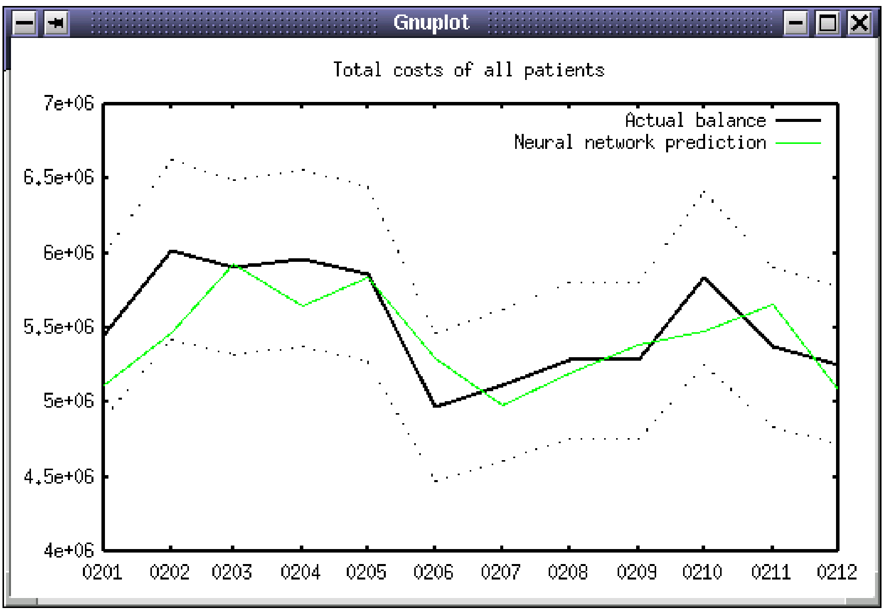


Figure 6.8 Actual monthly costs of all patients and the ANN predictions (screen copy)

7 EVALUATING ANNA

In this chapter I compared the expectations, i.e. monthly account values, formed with the ANN-method and with the conventional AR methods. The results of ANNA are also compared with budget values. Finally, the implementation framework for the model is proposed.

7.1 ANNA for different units

The ANN-system was evaluated with the monthly data of ten units from one big municipality. Internal auditors and managers of the municipality assisted in the evaluation process. The data was from the six health care services (HS) units, the social welfare (SW) unit, the office facilities administration (OFA) unit, and from the port of town (PT) unit. The data was five to seven years' operating monthly account values from these units. The year 2001 was selected as the testing period and all the previous years were used for training the ANNs. The performance was evaluated with the comparison methods described in section 6.7.4. Furthermore, ANNA was assessed with the budgeted yearly account values. Survey of Johnson and Johnson (1995) proposed budget-to-actual comparison of financial data in municipal audit.

In the raw data there were hundreds of accounts per unit. For the evaluation only those accounts that had entries in all the training and testing years were used. The numbers of accounts per unit ended up with a range of 29 to 157. For every unit two different kinds of models were built based on different account selection criteria. It was decided that the numbers of accounts per model were to be around ten in order to keep models user-friendly. First, those accounts per unit that were most stable compared to the previous year's values were selected. In most of the cases these accounts were the largest and most significant ones, such as salaries and salary-related accounts. These models were called A-models. Secondly, the internal auditors selected those accounts that were the most interesting from the auditing point of view. These models were called B-models. These selection criteria formed twenty models. Table 7.1 presents the average ANN-architecture of the ANNA. The more detailed architectures are found in Koskivaara and Back (2003). The average training cycle, 2100, is not high, but neither is the amount of the data. The weight decay is quite close to one and the reason for that is that the data fluctuates somewhat. This fluctuation is a reason for a low delta0 value. Multiple accounts were as input values and as output values 19 times, this indicates that the ANN utilises the relationships and dynamics between account values. The data are mostly equalised one account at a time; i.e. the values of the different accounts are in many cases far away from each other. On average there were no differences in the architectures between A- and B-models.

Table 7.1 Average ANN-architecture of ANNA

Average training parameters:
Training cycles: 2100 Weight decay: 0,96 Delta0: 0,17 Max delta: 50
ANN architecture (topology):
Inputs – Outputs: multiple – multiple accounts 19 times Inputs – Outputs: single – single accounts one time 15 times one previous month data 4 times 12 previous month data One time two previous month data
Data equalisation:
17 times one account at a time 3 times globally

Table 7.2 Average errors in money

Model	ANN-error	Next best error	Name of the method
HS1-A	**1138**	1269	CTM
HS1-B	**3339**	3527	CTM
HS2-A	**2122**	2170	CTM
HS2-B	**2747**	2944	PYS
HS3-A	**1061**	1201	CTM
HS3-B	**1001**	1027	CTM
HS4-A	**1434**	1471	CTM
HS4-B	277	**268**	**AVE**
HS5-A	6966	**6532**	**CTM**
HS5-B	**7452**	8640	AVE
HS6-A	**2504**	2581	PYS
HS6-B	**2085**	2103	AVE
SW-A	**43880**	46858	CTM
SW-B	**22844**	25417	CTM
OFA1-A	164569	**155803**	**DELTA**
OFA1-B	**159156**	165129	AVE
OFA2-A	**154646**	170119	CTM
OFA2-B	**79382**	80702	PYS
PT-A	**127367**	160502	PYS
PT-B	**49444**	94610	PYS

The ANN-method had better prediction accuracy with the lowest average errors in money than other AR methods used in the study (see Table 7.2 and

Figures 7.1 and 7.2). The ANN-method outperformed the second best method (CTM) in 19 out of 20 cases. The ANN-models also had the lowest standard deviations, which indicates that the ANN-method predicts more consistently. In Figure 7.3 the inner radars belong to the ANN-method and are smaller than the bigger radars, which belong to the CTM-method.

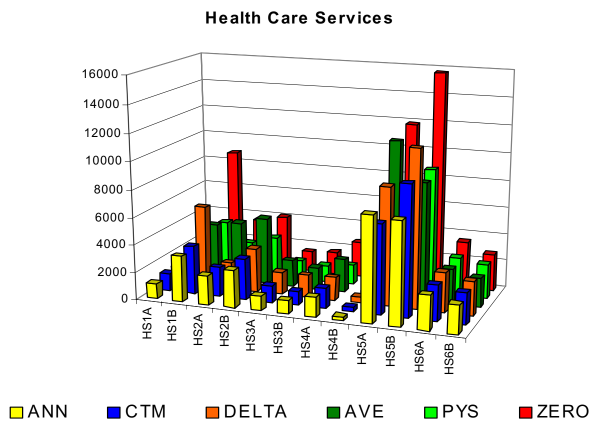


Figure 7.1 Average errors of AR methods in health care services

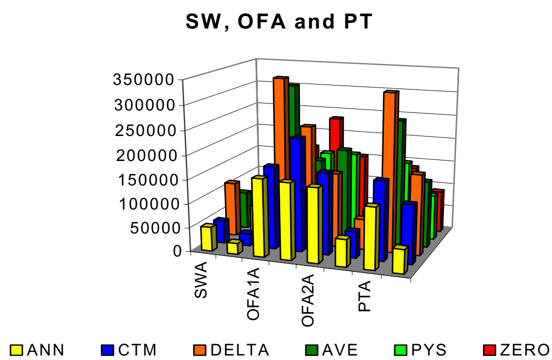


Figure 7.2 Average errors of AR methods in SW, OFA, and PT units

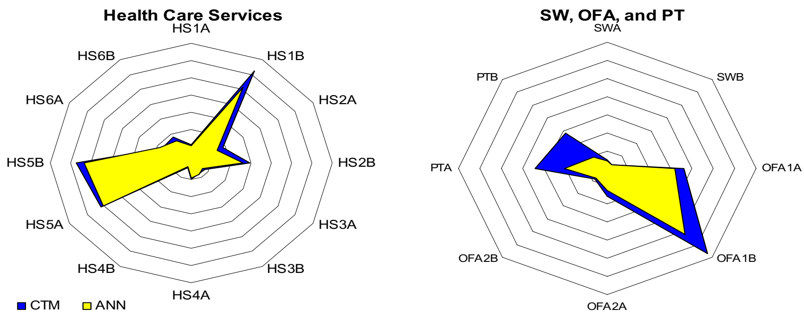


Figure 7.3 Standard deviations of ANN and CTM

At the top of Figures 7.4 and 7.5 are the yearly average errors of the ANN-method and the budget method in percentages. The average error of the ANN-method was 12%, and the average error of the budget was 19%. The biggest average errors % are in the HS3-A and HS4-B cases, although in money they are not the biggest error. The explanation might be that only one unexpected cost has turned the situation upside down. At the bottom of these figures are the yearly budget, actual, and ANN-method values. The deeper the pillar is the bigger the cost is. In four cases, HS2-A, HS2-B, HS5-A, the budget value is the biggest and the ANN-value is the smallest. These raise questions, why have the costs been cut down during the period, or has something failed to full? The ANN-method is very able to compete with the budget method. The ANN-method is especially competitive in cases where the budget values are missing for one or another reason.

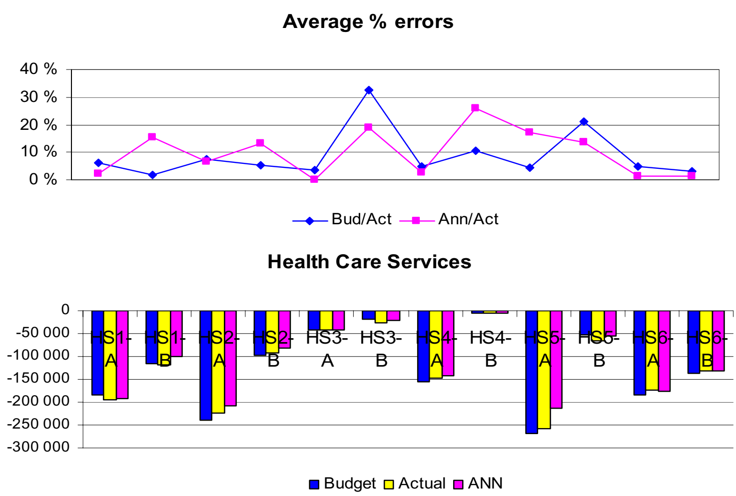


Figure 7.4 ANN vs. budget values in health care services

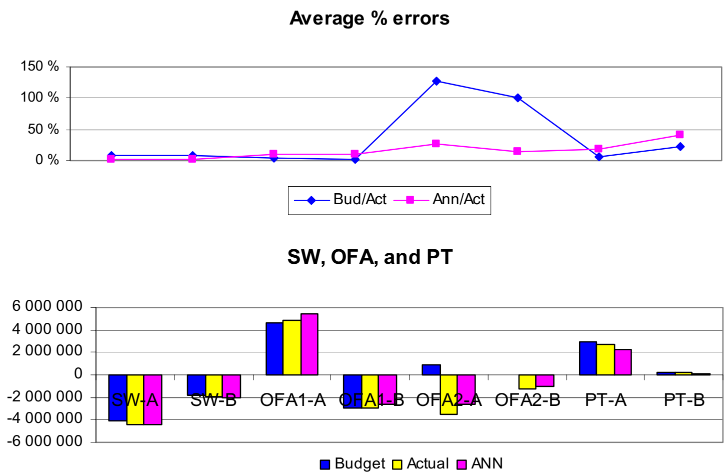


Figure 7.5 ANN vs. budget values in SW, OFA, and PT units

7.2 ANNA for modeling

The average results of ANNA were good compared to the related results, but there were no significant differences between the AR methods. Therefore, the evaluation of the ANN-method was duplicated with new data. Once again, I tested the ANN-method's modeling accuracy in the monthly account values and compared the results with the simple quantitative AR methods using real world data. The prediction accuracy of the ANN-method is compared to the results of the traditional AR methods (i.e. PYS, AVE, DELTA, DELTA, and CTM) described in section 6.7.4.

This time the unit from the municipality was selected with the help of their executive management. Eight years' monthly account values were collected from this unit. The data for year 2002 was held out for testing and all the data from the previous years were used for training the ANNs. Once again in the raw data there were hundreds of accounts. However, only those accounts that had entries in all the training and testing years were used. From these accounts the management of the municipality unit selected the accounts which they thought were the most significant.

Although it is doubtful from the auditing point of view to let the management of the organisation select those accounts that should be audited it is acceptable in a research setting. The management of the organisation should know which accounts follow the trends best and are related to each other. However, in principle, all the accounts should be audited in one way or another. Therefore, this selection basis might give the auditor an idea of which accounts would be the most suitable for the ANN-assisted auditing. Indeed, researchers have also proposed AR procedures to management accounting for controlling operations (Lee & Colbert 1997; Colbert 1994).

The number of accounts ended up as 25. They were divided into three models according to the level of account category and specialities. ANNA.USE includes three accounts from the upper levels. The accounts include all expenses of outpatients and bed patients as well as the total costs of all patients. These account values had the biggest entries in this data sample. ANNA.ONE includes eleven accounts from the outpatients' ward. The specialities are internal diseases, neurology, pulmonary diseases, surgery, gynaecological diseases, eye diseases, otolaryngology, psychiatry, adolescent psychiatry, and child psychiatry. ANNA.TWO includes the accounts from the same eleven specialities from the bed patient ward. These account values in ANNA.ONE and ANNA.TWO are from a lower account category level than in ANNA.USE.

7.2.1 Training options for ANNA

Table 7.3 presents the optimal network training parameters achieved for the ANN-models. The values of the training parameters were similar to the previous study. However, the training parameters were the same for all the ANN-models. The training cycles per model were 1000. This is not a high value, but neither is the amount of the data per models. The value of the weight decay was 0.99999 and the delta0 values were 0.1. These indicate the fluctuating nature of the data. Max delta was 50 in all the cases.

Table 7.3 Training parameters

ANNA	Training Cycles	Weight Decay	Delta0	Max Delta
USE	1000	0.99999	0.1	50
ONE	1000	0.99999	0.1	50
TWO	1000	0.99999	0.1	50

Table 7.4 presents the network architectures per models. The models have multiple variables as inputs and outputs. The models have the biggest difference in the number of the month indicator input. Either one, four or five previous months' data is given as input. All the models have one hidden layer. The data in ANNA.USE and ANNA.TWO were equalised globally, while the data in ANNA.ONE was equalised one account at a time.

Table 7.4 Network architectures

ANNA	Inputs	Outputs	Input months	Data Equalisation	Neurons in Layers
USE	Multiple	Multiple	1	Global	15–8–3
ONE	Multiple	Multiple	5	Local	55–11–11
TWO	Multiple	Multiple	4	Global	44–11–11

7.2.2 Comparison of the AR methods

A good measure of the accuracy of the methods is to compare the average errors and their standard deviations in the holdout group. Table 7.5 presents these values of all the methods used in this study. The ANN-method has in each case the lowest average error in money and the lowest standard deviation. This indicates that the ANN-method has the best average prediction accuracy. The lowest standard deviation illustrates that the values of the ANN method are less dispersed from the average value. So, the ANN-method predicts monthly account values more consistently than the conventional AR methods in the study. Based on the average errors and the standard deviations in money it seems that the ANN-based system's expectation values for monthly account values are the closest to the actual monthly account values. But is there a significant difference between the ANN-based expectations and the expectations achieved with the traditional AR methods[38].

Table 7.5 Average errors and standard deviations of the AR methods

ANNA	ANN	CTM	DELTA	AVE	PYS	ZERO
USE: average error	**279540**	396297	352831	752253	358826	365795
USE: standard deviation	**214154**	271171	273852	413668	294126	301408
ONE: average error	**23383**	29404	33580	46568	32832	33726
ONE: standard deviation	**5395**	7727	9785	7937	7573	9745
TWO: average error	**34261**	63304	66100	61884	54474	68347
TWO: standard deviation	**5340**	10243	10528	9358	6879	10766

As mentioned earlier, in effect, I have six populations, one associated with each method (i.e. ANN, ATM, DELTA, AVE, PYS, ZERO). I will test the following hypothesis:

[38] The statistical differences between the populations are studied with the help of SPSS for Windows 11.5 computer program.

H_0: *There is no significant difference between the expectations in the monthly account values of the ANN methods and the conventional AR methods.*

If H_0 cannot be rejected, I will not have evidence to conclude that the ANN-based system differs from the traditional AR methods. However, if H_0 can be rejected, I will conclude that the ANN-based system differs from the traditional AR methods.

The methodology for the parametric matched-sample analysis, the *t* test on paired differences, requires quantitative data and the assumption that the population of differences between the pairs of observation is normally distributed (Anderson, Sweeney & Williams 1999, 813). The skewness and kurtosis of the pair observations are in Table 7.6. In general, a skewness value greater than one indicates a distribution that differs significantly from a normal, symmetric distribution, i.e. the differences in ANNA.USE are normally distributed. With this assumption, the *t* distribution can be used to test the null hypothesis of no differences between the population averages in ANNA.USE. The differences in ANNA.ONE and ANNA.TWO are not normally distributed. Therefore in the cases of ANNA.ONE and ANNA.TWO the nonparametric Wilcoxon signed-test is used (Anderson, Sweeney & Williams 1999, 813).

Table 7.6 Skewness and kurtosis of the pair observations

	Skewness Statistic	Std. Error	Kurtosis Statistic	Std. Error
Difference: USE: ANN–CTM	**-0,694**	0,393	1,808	0,768
Difference: USE: ANN–DELTA	**-0,306**	0,393	-0,176	0,768
Difference: USE: ANN–AVE	**0,146**	0,393	0,630	0,768
Difference: USE: ANN–PYS	**-0,471**	0,393	0,456	0,768
Difference: USE: ANN–ZERO	**-0,213**	0,393	0,643	0,768
Difference: ONE: ANN–CTM	-5,345	0,211	58,662	0,419
Difference: ONE: ANN–DELTA	-7,192	0,211	65,210	0,419
Difference: ONE: ANN–AVE	-4,213	0,211	35,661	0,419
Difference: ONE: ANN–PYS	-4,679	0,211	48,903	0,419
Difference: ONE: ANN–ZERO	-6,932	0,211	61,956	0,419
Difference: TWO: ANN–CTM	-6,467	0,211	48,301	0,419
Difference: TWO: ANN–DELTA	-4,958	0,211	31,845	0,419
Difference: TWO: ANN–AVE	-3,295	0,211	13,788	0,419
Difference: TWO: ANN–PYS	-3,662	0,211	22,586	0,419
Difference: TWO: ANN–ZERO	-5,797	0,211	41,168	0,419

As shown in Table 7.7, there is a significant difference between the ANN-method and the traditional AR methods at the 0.05 level in four (in bold) out of five cases in ANNA.USE. Therefore, H_0 can be rejected in these cases.

There was no significant difference between the account values achieved with the ANN-system and with the PYS-method.

Table 7.7 Prediction accuracy of ANN-methods vs. AR methods in ANNA.USE

Paired Samples Test	t-value	df	Sig. (2-tailed)
ANN – CTM	-2,448	35	**0,020**
ANN – DELTA	-2,125	35	**0,041**
ANN – AVE	-6,416	35	**0,000**
ANN – PYS	-1,651	35	0,108
ANN – ZERO	-2,293	35	**0,028**

As shown in Table 7.8, there is a significant difference between the ANN-method and the traditional AR methods at the 0.05 level in eight (in bold) out of ten cases in ANNA.ONE and ANNA.TWO. Therefore, H_0 can be rejected in these cases. There was no significant difference between the account values achieved with the ANN-system and with the CTM-method or with the ZERO-method in ANNA.ONE case.

Table 7.8 Prediction accuracy of ANN-methods vs. AR methods

Test Statistics of ANNA.ONE and ANNA.TWO Wilcoxon Signed Ranks Test	Z	Asymp. Sig. (2-tailed)
ONE: CTM – ONE: ANN	-1,835	0,066
ONE: DELTA – ONE: ANN	-1,978	**0,048**
ONE: AVE – ONE: ANN	-7,370	**0,000**
ONE: PYS – ONE: ANN	-3,012	**0,003**
ONE: ZERO – ONE: ANN	-1,606	0,108
TWO: CTM – TWO: ANN	-5,583	**0,000**
TWO: DELTA – TWO: ANN	-4,731	**0,000**
TWO: AVE – TWO: ANN	-4,717	**0,000**
TWO: PYS – TWO: ANN	-4,954	**0,000**
TWO: ZERO – TWO: ANN	-5,764	**0,000**
Z=Based on negative ranks.		

7.3 Weak market test of ANNA

Kasanen, Lukka, and Siitonen (1993) present a three-level market test as a validation method of constructions. The weak market test answers the question whether any manager responsible for the financial results of his or her business unit has been willing to apply the construction in question in his or her actual decision-making. The semi-strong market-test answers the question whether the construction has become widely adopted by companies. The strong market test answers the question whether the business units applying the construction systematically produced better financial results than those which are not using it. The weak market test suits the validation of prototypes. The semi-strong and strong market tests are more appropriate in the product development and technology transfer phases.

My first intention was to test ANNA in real auditing situation, however, I did not get enough companies to commit themselves to the project. My second intention was to test ANNA in a firm of accountants. I contacted several accountancy companies. The smaller companies were ready to cooperate, but they did not have enough data. The bigger ones were not ready to cooperate, and they appealed to the confidential nature of their business relationships. Therefore, ANNA was evaluated with the data achieved from the public sector.

Within the evaluation process, the ANNA-system and its results were shown four different times to the internal auditors of the municipality. Two times the meeting was held together with the unit executives of the municipality. The main feedback from these discussions was that this kind of system could support the analytical review process. However, it should work quite automatically and monitor unusual expectations in the records and only give alarms when something is wrong.

7.4 Implementation framework for the model

Figure 7.6 depicts the framework of the ANNA-system in accounting data for continuous auditing purpose, which was discussed in section 3.3. This system could work in users' workstations or on the web. And the system would pick up the necessary data from the accounting data warehouses. In general, this accounting data should illustrate the business process. ANNA takes advantage of the data that already exists in accounting systems. It utilises the data and constructs expectations which can be either values or visualised pictures of the models in the data. The system could give new kind of information to analysts and therefore make the interaction between a human actor and information system more valuable.

Generally, the human actors may refresh ANNA in three ways. Firstly, they can make changes in the business processes. For example, allocate the use of material in the accounting system more timely. Secondly, they can correct, for instance, the faulty accounting records. Thirdly, a human actor may refresh the ANN-system either by changing the parameters and the network architecture in the model or by changing the variables of the model.

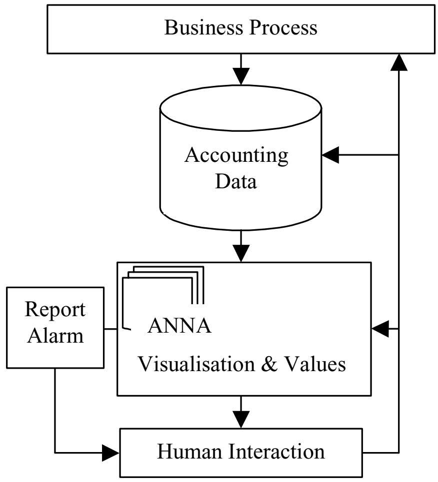


Figure 7.6 Framework of ANNA

ANNA could serve a continuous monitoring and controlling purpose as follows. It could automatically give, for example once a month, a report of those accounts which follow the trend (i.e. are inside a certain percentage or money limits) based on the ANN-method and two conventional AR methods. Furthermore, it could raise the alarm to those accounts which occasionally either start or stop to follow the trend. Then the user may plot those accounts on the screen to see and decide whether further investigation or action is needed. The investigation should include statistical sampling of the data in order to cover all the data statistically.

In any case users have to remember that the adaptability of an ANN-system implies that we should not allow an ANN to operate as a static tool. The ANN-system has to be refreshed continually when new data is entered into the system. For example, data gathered from the completed real audits with reasonably well know results could be used to develop and refine the ANN-system.

7.5 Limitations of the research

Since ANNA was built for a specific situation, it had a limited view of an AR procedure. The developed prototype focused on the monitoring models in the monthly data. Therefore the prototype had a narrow view of all the material control in auditing. Hence, other control indicators could produce some other auditing tasks. In spite of these limitations, the prototype is appropriate for my research aims since it highlights the building and evaluation process of an ANN-based system for AR. Furthermore, the development process assisted in highlighting the possible benefits, problem areas, and limits of this type of ANN-system.

Although, ANNA had the best and most consistent prediction accuracy and some general guidelines for modelling were found, some limitations occurred. It seems that every ANN-model needs to be trained individually. ANNA also used a limited number of accounts and a limited amount of data. However, this is also a common feature in other ANN-studies in auditing. Moreover, companies do not store the data forever, and even if the data are available very old data might not be so relevant.

The best results were achieved when the management of the organisation selected the accounts. This is not surprising: the management of the organisation should know which accounts follow the trends best and are related to each other. However, from an auditor's point of view this result is doubtful because auditors have a mandate from the owners to also control the management of the organisation. Nevertheless, researchers have stated that AR procedures are tools which also the management could use as a part of its control (Lee & Colbert 1997; Colbert 1994). A management accountant could effectively utilise the same benefits of AR procedures that auditors do. If AR procedures are applied before the amounts are integrated into the financial statements and prior to auditors' investigations, possible faults can be corrected in time.

8 CONCLUSIONS

The purpose of this thesis was to research whether artificial neural networks (ANNs) are feasible for supporting analytical review (AR) in auditing. In order to realise the purpose of the research, an ANN-based system was built as a computer program with monthly accounting data. The ANNA-system was developed using an iterative approach to answer the following research questions: 1) How to model an ANN-based monthly account predictor? 2) How to illustrate the monthly account models for auditing purpose? 3) How to construct the ANNA-system? 4) Which criteria should be used in evaluating the ANNA's capability to model monthly account values? 5) How good is the ANNA-system based on these criteria? The ANNA-system was iteratively constructed with five different data selection basis: literature, CPA-auditor, the biggest and most stable accounts compared to the previous year's values, internal auditors, and the management of the organisation.

The research started by considering the demands of today's business environment set for auditors. Auditors are faced with a dilemma. On one hand they are asked to prevent manipulation of the accounting data. On the other hand hangs the reality that auditing is a competitive business, subject to the same demands for profitability and the return on capital as other businesses. These economic demands create a conflict for auditors by constraining their ability to prevent data manipulation. In fact, Enron and its aftermath hit the headlines and put the reputations of auditors to shame. Although AR is one way to increase the effectiveness and efficiency of auditing, research has consistently indicated that auditors prefer simple scanning, reasonableness tests, and ratio analysis to sophisticated statistical or mathematical models in AR (Lin, Fraser & Hatherly 2003; Lin & Fraser 2003). Auditing by expectations on a continuous basis is proposed as a solution for monitoring actual values (Woodroof & Searcy 2001). Furthermore, Vasarhelyi, Kogan, and Alles (2002) argue that while many believe that a well-performed traditional audit could have detected many of Enron's operational problems, a well-performed continuous audit would have brought them to light much sooner. Besides, the development of IT makes the use of advanced methods easier and more cost effective. The current research argued that one possible way to develop continuous auditing is to develop AR tools and that they could be applied in the continuous monitoring context and that the ANN-technique is feasible for developing expectations for accounting data.

In order to build and evaluate the ANN-based system different kinds of research approaches were examined to achieve a coherent basis for the research process. These approaches were as follows: the constructive research approach, the design science research approach, the systems development approach, the development cycle approaches, and the spiral model approach. Different parts of these approaches were utilised to build and evaluate the ANN-based system. The constructive research approach gave the general guidelines with its seven phases. The design science research approach provided a basis for the research questions and activities. The systems development approach presented the accumulative research process of the information system. The development cycle approaches gave the main phases of the ANN development process. Altogether, the development process has been an evolutionary spiral approach with several iterations.

The ANN technology was selected as a technical platform in order to form expectations due to the following reasons. ANNs are data driven and they learn from examples. They are good at handling data, and once trained, they can predict and classify new examples very quickly. ANNs provide a non-linear dimension that captures the underlying representations within the data sets. They might reveal something in the data that is left hidden otherwise or they might find a function between data items that is hard to find with other techniques. Therefore, auditors may benefit from applying ANNs in revealing trends in accounting data or in the comparison of accounting records. In brief, auditors may benefit from the ANN's ability to learn from data to support the auditors' experience and knowledge about a client company.

Parallel with the development of the ANNA-system, both the research domain, i.e. analytical review in auditing and the ANN-research in auditing were analysed. The AR is a way of thinking and behaving to obtain audit evidence with the help of analytical procedures, in which again some special techniques, such as ANNs, might be embodied. The timing and purposes of analytical procedures as well as the characteristics of different kinds of AR techniques differ. The understanding of the AR procedures and techniques helps auditors to choose the best and possible AR tools and the level of application. Continuous auditing was proposed as a possible audit domain for AR tools.

The review of ANN-research in the auditing gave a summary of what has been done in this field. The ANN-application areas in auditing were detecting material errors (six studies), detecting management fraud (three studies), supporting going concern decision (five studies), and determining financial distress problems (one study). ANNs have also been applied to internal control risk assessment (three studies) and audit fee forecasting (one study). All the researches fit into AR procedures. Most of the authors state that the ANNs have the potential to improve the AR procedure. Indeed, the analysis revealed the open ends in the field and unveiled what kinds of ANN-models have been applied to auditing purposes. This research tended to fill the following open ends.

No ANN-research in the auditing field has previously compared ANN-based results to the budgeted values, although comparison of the current information with the budget or forecast is mentioned as one possible AR procedure. In Koskivaara and Back (2003) the results of the ANN-based system were compared with the budgeted values of the management.

Comparison of current information with ANN-based information based on similar information for prior periods was very common in the auditing ANN-studies. My research is a continuation of the line of research that Coakley and Brown (1991a) started on. They used monthly aggregated data based on financial statements whereas my research used the real operational monthly data (Koskivaara & Back 2003; Koskivaara 2004c).

My studies were the first studies to apply the resilient backpropagation (RPROP) training algorithm (Koskivaara & Back 2003; Koskivaara 2004c) and the SOM algorithm to analysing monthly account data sets for auditing purposes (Koskivaara 2004b).

While the pre-processing has an effect on what the network learns and the successful data pre-processing might speed up ANNs' learning and give better and more general results, the effect of pre-processing has not been studied in an auditing context. The ANNA-model was tested with financial statement values and different pre-processing models were compared (Koskivaara 2000b).

The performance measure of auditing ANN-applications varied a lot and it was not always clear whether the performance was calculated from actual data or scaled data. Therefore, one aim of the current research was to receive further evidence on whether the ANN-method is better than the conventional AR methods with actual values. Furthermore, the visualisation of the results might give some added value to auditors. Hence, the aim of finding an appropriate way of illustrating data to reveal unexpected values was included in my research.

8.1 Contributions of the research

For example by Simon (2002) stressed that new technology can be used as a research instrument. He states that we would look on the new technology as better ways of understanding problems and finding solutions to the problems. My research focuses on auditors' work in the sense that it tackles an important and inherent aspect of auditors' work: forming expectations to compare with client data. The expectations developed by the auditor are in the heart of the definition of analytical procedures. ANNs are an appealing technology embedded in the analytical procedures for forming expectations for auditing purposes. The precise application domain in AR was detecting material errors in the accounting data. Especially, the model attempted to monitor real operating data in order to model the monthly accounting data.

The contribution of my research is in the intersection between two areas: AR and ANN, i.e. in using and testing a technique in an area where new techniques are urgently needed. The *analytical review contribution* includes: (1) an analytical review theory synthesis (view of auditing with AR, synthesis from analytical procedures and from analytical review techniques, and their relationships for continuous auditing domain) and (2) a demonstration of the ANN-method vs. conventional AR methods for forming expectations for AR purposes. The *ANN-technology in auditing contribution* includes: (1) a comprehensive overview of ANN-studies done in the field of auditing and (2) building the model and choosing parameters for the ANN-based system for analytical review purposes, and building and evaluating the ANN-based system with several iterations. The contribution is discussed more thoroughly by answering research questions.

According to March and Smith (1995), the research contribution lies in the novelty and the effectiveness of the artifact. The ANNA-system is a novel construction because of the following reasons. First, the research presented the first attempt to apply the flexible one-step-ahead prediction model in the auditing domain. The architecture of ANNA is a fully connected feedforward network with a maximum of three hidden layers. ANNA includes both the single variable options and the multiple variable options for the choice of input and output variables. Second, this research is a first attempt to apply the RPROP learning algorithm in the analytical review context. The prediction capabilities of ANNA were based on its function approximation capability from previous years. Third, ANNA was developed based on real world operating data with several iterations and with five different variable selection basis.

The scaling of the data has effects on ANN's performance, i.e. it makes a difference how the data is pre-processed. In this study the best results were achieved in most of the cases by scaling the monthly data on an account basis. Furthermore, the linear scaling has the advantage of preserving the relative position of each data point along the range. The results of ANNA are presented both in tables and in figures on computer screen. It is also possible to illustrate the limits of ANNA (see dotted lines in Figure 6.8).

The evaluation criteria for the ANN-method were founded on the analytical procedures and AR techniques in use. The ANN-method had better and more consistent prediction accuracy than the conventional AR methods used in the study. The ANN-method had the lowest average errors and standard deviations. In some cases the difference was even significant in the favour of ANN-method. Indeed, the ANN-method was better than the organisation's budget method.

However, the auditor must never trust on only one method. Therefore, ANNA could serve the continuous monitoring and controlling purpose as follows. It could automatically give, for example once a month, a report of those accounts which follow the trend (i.e. are inside a certain threshold value) based on ANN-method and two conventional AR methods. Furthermore, it could give an alarm of those accounts which occasionally either start or stop to follow the trend. Then the user may plot those accounts on the screen to see and decide whether further investigation or action is needed. To sum up, the ANN-method was able to compete with the conventional AR methods used in this study. Based on the results of the study and the literature review, *it seems that ANNs are feasible techniques for supporting AR*.

8.2 Managerial implications

One advantage of this kind of ANN-system is that it can provide auditors with objective information about a client company. With the help of ANNA an auditor may focus substantive testing on the right places.

ANNA can also prove to be a persuasive analytical tool when an auditor discusses problems with the clients and recommends changes in operations. Thus, the prototype illustrates an ANN's usefulness for effective communication which is important not only for every auditing system but also for the planning part of management.

This kind of system could work on the web or on the user’s workstation using the data in the accounting databases. It could automatically report to the user about those accounts that follow the rules and alert about those accounts which irregularly begin or end to follow the trend. Therefore, ANNA is a possible future tool for continuous auditing and controlling.

Furthermore, management accountants could apply ANNA to different accounts to search for figures that do not appear reasonable. If ANNA is applied before the amounts are integrated into the financial statements and prior to auditors’ examinations, possible faults can be corrected before they become problems.

8.3 Future work

The future of ANNs in auditing is open and will be brighter as more and more research efforts are devoted to this area. There are many research questions and problems in this area. Are there any new application areas, such as authority checking and analysing minutes that would be suitable for ANNs? Budgets, forecasts, on-line analytical procedures, and data mining are areas in organisations where ANNs embedded in the systems may provide auditors with the business intelligent information they need to make decisions.

How do we systematically build an appropriate ANN-based model for auditing problems? Although some guidelines from the literature have been derived, we are still generalising across problem areas. Replication of the studies using different architectures and data sets is needed to refine these guidelines within auditing application areas. Research is also needed to examine the circumstances under which ANNs perform better than other techniques. Are some training methods and algorithms more suitable than others for auditing problems? Generalisation from these limited experiments is difficult. Again, replication of the experiments and further research is needed.

So far, only a few attempts have been made to combine ANNs with other advanced methods. Davis, Massey, and Lovell (1997) presented a construction of a prototype, which integrated an ES and an ANN in the control risk assessment domain. Additional research is clearly needed to evaluate the efficacy of hybrid models that combine the advantages of statistical methods, ANNs, fuzzy logic, and genetic algorithm based systems. Therefore, the future of ANNs in the auditing area may see increased integration with other existing or developing technologies and statistical techniques. Although ANN-based systems or other advanced methods cannot entirely replace professional judgement, they offer a promising alternative tool to AR procedures.

As advancements are made in information technologies and computer-based systems, there should be new opportunities to apply information technology to auditing. This would encourage and motivate academics and practitioners to collaborate in further exploration of the potential of IT in auditing. This is important, because a big challenge is to get practitioners to update (Fleming 2004; Blocher, Krull, Tashman & Yates 2002) and adopt new AR tools in today’s rapid and demanding business environment. To say the least, there is a place for innovations[39] in AR tools in the continuous auditing environment.

[39] See ”The Discipline of Innovation” by Drucker (1985, 1998, 2002).

REFERENCES

AICPA (1988) *Statement on Auditing Standard No. 56, Analytical Procedures.* American Institute of Certified Public Accountants. <http://www.aicpa.us/pubs/cpaltr/may98/suppl/pubacc.htm>, retrieved 22.11.2002.

AICPA (2002a) *AICPA Audit and Attest Standards, Statements on Auditing Standards (SAS) No. 96, Audit Documentation*. <http://www.aicpa.org/members/div/auditstd/riasai/sas96.htm>, retrieved 13.1.2004.

AICPA (2002b) *SysTrustTM Principles and Criteria*. American Institute of Certified Public Accountants, Inc. and Canadian Institute of Chartered Accountants. <http://www.aicpa.org/assurance/systrust/ princip.htm>, retrieved 23.1.2004.

Ameen, Elsie C. – Strawser, Jerry R. (1994) Investigating the Use of Analytical Procedures: An Update and Extension. *Auditing: A Journal of Practice & Theory,* 13 (2): 69–76.

Anandarajan, Murugan – Anandarajan, Asokan (1999) A comparison of machine learning techniques with a qualitative response model for auditor's going concern reporting. *Expert Systems with Applications,* 16 (4): 385–392.

Anderson, David R. – Sweeney, Dennis J. – Williams, Thomas A. (1999) *Statistics for business and economics 7e*. South-Western College Publishing: Cincinnati, Ohio.

APB (1995) *Analytical Procedures, Statement on Auditing Standards 410*. Auditing Practices Board. <http://www.frc.org.uk/images/ uploaded/documents/APB_410.pdf>, retrieved 16.2.2004.

Arens, Alvin A. – Elder, Randal J. – Beasley, Mark S. (2003) *Auditing and Assurance Services An Integrated Approach*. Prentice Hall: Upper Saddle River.

Arrington, C. Edward – Hillison, William – Jensen, Robert E. (1984) An Application of Analytical Hierarchy Process to Model Expert Judgements on Analytical Review Procedures. *Journal of Accounting Research,* 22 (1): 298–312.

Back, Barbro – Sere, Kaisa – Vanharanta, Hannu (1998) Managing Complexity in Large Data Bases Using Self-Organizing Maps. *Accounting, Management and Information Technologies,* 8 (4): 191–210.

Barto, Andrew G. – Sutton, Richard S. – Anderson, Charles W. (1983) Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, and Cybernetics,* SMC-13 (5): 834–846.

Basheer, Imad A. – Hajmeer, Maha N. (2000) Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods,* 43 (1): 3–31.

Bazerman, Max H. – Loewenstein, George – Moore, Don A. (2002) Why good accountants do bad audits. *Harvard Business Review,* 80 (11): 97–102.

Bell, Timothy – Mars, Frank – Solomon, Ira – Thomas, Howard (1997) *Auditing Organizations Through a Strategic-Systems Lens*. KPMG Peat Marwick LLP.

Bharadwaj, Anandhi – Karan, Vijay – Mahapatra, Radha K. – Murthy, Uday S. – Vinze, Ajay S. (1994) APX: An Integrated Knowledge-Based System to Support Audit Planning. *International Journal of Intelligent Systems in Accounting, Finance and Management,* 3 (3): 149–164.

Bible (1992) *Raamattu*. 3. Suomennoksen tekijän oikeudet: Kirkon Keskusrahasto, Sisälähetysseuran kirjapaino Raamattutalo: Pieksämäki.

Biggs, Stanley F. – Mock, Theodore J. – Watkins, Paul R. (1988) Auditor's Use of Analytical Review in Audit Program Design. *The Accounting Review,* 63 (1): 148–161.

Bishop, Christopher M. (1995) *Neural Networks for Pattern Recognition*. Clarendon Press: Oxford.

Blocher, Edward – Krull, Jr., George W. – Tashman, Leonard J. – Yates, Stephen V. N: (2002) Updating Analytical Procedures. *The CPA Journal,* 72 (11): 36–41.

Blocher, Edward – Patterson, Jr., George F. (1996) The Use of Analytical Procedures: The Importance of Expectation and Precision. *Journal of Accountancy,* 181 (2): 53–55.

Boehm, Barry (1988) A Spiral Model of Software Development and Enhancement. *IEEE Computer,* 21 (5): 61–72.

Bonner, Sarah E. (1990) Experience Effects in Auditing: The Role of Task-Specific Knowledge. *The Accounting Review,* 65 (1): 72–92.

Bonner, Sarah E. – Walker, Paul L. (1994) The Effects of Instruction and Experience on the Acquisition of Audit Knowledge. *The Accounting Review,* 69 (1): 157–178.

Broomhead, D.S. – Lowe, D. (1988) Multivariable functional interpolation and adaptive networks. *Complex Systems,* 2 (3): 321–355.

Brown, Carol E. – Murphy, David S. (1990) The Use of Auditing Expert Systems in Public Accounting. *Journal of Information Systems,* 4 (1): 63–73.

Busta, Bruce – Weinberg, Randy (1998) Using Benford's law and neural networks as a review procedure. *Managerial Auditing Journal,* 13 (6): 356–366.

Coakley, James R. (1995) Using Pattern Analysis Methods to Supplement Attention-Directing Analytical Procedures. *Expert Systems with Applications,* 9 (4): 513–528.

Coakley, James R. – Brown, Carol E. (1991a) Neural Networks Applied to Ratio Analysis in the Analytical Review Process. In: *Proceedings of the Fourth International Symposium on Expert Systemsin Accounting, Finance and Management*, University of Southern California. Pasadena California, USA, October 30 – November 1, 1991, 1–36.

Coakley, James R. – Brown, Carol E. (1991b) Neural Networks for Financial Ratio Analysis. In: *Proceedings of The World Congress on Expert Systems*, ed. by Jay Liebowitz. Orlando Florida, USA, December 16–19, 1991, 132–139.

Coakley, James R. – Brown, Carol E. (1993) Artificial Neural Networks Applied to Ratio Analysis in the Analytical Review Process. *International Journal of Intelligent Systems in Accounting, Finance and Management,* 2 (1): 19–39.

Coakley, James R. – Brown, Carol E. (2000) Artificial Neural Networks in Accounting and Finance: Modeling Issues. *International Journal of Intelligent Systems in Accounting, Finance & Management,* 9 (2): 119–144.

Colbert, Janet L. (1994) Analytical Procedures for Management Accountants and Auditors. *Managerial Auditing Journal,* 9 (5): 3–7.

Curry, Bruce – Peel, Michael J. (1998) Neural Networks and Business Forecasting: An Application to Cross-Sectional Audit Fee Data. *International Journal of Commerce and Management,* 8 (2): 94–120.

Davis, Jefferson T. (1996) Experience and Auditors' Selection of Relevant Information for Preliminary Control Risk Assessments. *Auditing: A Journal of Practice & Theory,* 15 (1): 16–37.

Davis, Jefferson T. – Massey, Anne P. – Lovell, Ronald E. R. II (1997) Supporting a complex audit judgement task: An expert network approach. *European Journal of Operational Research,* 103 (2): 350–372.

Demuth, Howard – Beale, Mark (2000) *Neural Network Toolbox*. The MathWorks, Inc. <http://www.mathworks.com/access/helpdesk/help/toolbox/nnet/nnet.shtml>, retrieved 16.2.2003.

Duke, Gordon L. – Neter, John – Leitch, Robert A. (1982) Power Characteristics of Test Statistics in Auditing Environment: An Empirical Study. *Journal of Accounting Research,* 20 (1): 42–67.

Drucker, Peter (1985) The Diciplines of Innovation. *Harvard Business Review,* 63 (3): 67–72.

Drucker, Peter (1998) The Diciplines of Innovation. *Harvard Business Review,* 76 (6): 149–155.

Drucker, Peter (2002) The Diciplines of Innovation. *Harvard Business Review,* 80 (8): 95–101.

Etheridge, Harlan L – Brooks, Richard C (1994) Neural networks: A new technology. *The CPA Journal,* 64 (3): 36–55.

Etheridge, Harlan L. – Sriram, Ram S. – Hsu, H. Y. Kathy (2000) A Comparison of Selected Artificial Neural Networks that Help Auditors Evaluate Client Financial Viability. *Decision Science,* 31 (2): 531–550.

Fanning, Kurt M. – Cogger, Kenneth O. (1994) A Comparative Analysis of Artificial Neural Networks Using Financial Distress Prediction. *International Journal of Intelligent Systems in Accounting, Finance and Management,* 3 (4): 241–252.

Fanning, Kurt M. – Cogger, Kenneth O. (1998) Neural Network Detection of Management Fraud Using Published Financial Data. *International Journal of Intelligent Systems in Accounting, Finance & Management,* 7 (1): 21–41.

Feroz, Ehsan Habib – Kwon, Taek Mu – Pastena, Victor S. – Park, Kyungjoo (2000) The Efficacy of Red Flags in Predicting the SEC's Targets: An Artificial Neural Networks Approach. *International Journal of Intelligent Systems in Accounting, Finance & Management,* 9 (3): 145–157.

Fischer, Michael J. (1996) "Real-izing" the Benefits of New Technologies as a Source of Audit Evidence: An Interpretive Field Study. *Accounting, Organizations and Society,* 21 (2/3): 219–242.

Fisher, Brian (1999) Mellon creates fraud watch to predict and manage risk using neural technology. *Journal of Retail Banking Services,* 21 (1): 15–17.

Fleming, John M. (2004) Analytical Procedures are a Helpful Tool. *Pennsylvania CPA Journal,* 74 (4): 3–4.

Fraser, Ian A. M. – Hatherly, David J. – Lin, Kenny Z. (1997) An Empirical Investigation of the Use of Analytical Review by External Auditors. *The British Accounting Review,* 29 (1): 35–47.

Freeman, James A. – Skapura, David M. (1991) *Neural Networks Algorithms, Applications, and Programming Techniques*. Addison-Wesley Publishing Company: Menlo Park, California.

Frost, Peter A. – Tamura, Hirokunui (1982) Jackknifed Ratio Estimation in Statistical Auditing. *Journal of Accounting Research,* 20 (1): 103–120.

Fryer, Bronwyn (1996) Visa cracks down on fraud. *InformationWeek,* August 26 (594): 87–88.

Garstka, Stanley J. – Ohlson, Philip A. (1979) Ratio Estimation in Accounting Population with Probabilities of Sample Selection Proportional to Size of Book Values. *Journal of Accounting Research,* 17 (1): 23–59.

Gauntt, Jr., James E. – Gletzen, G. William (1997) Analytical Auditing Procedures. *Internal Auditor,* 54 (1, February): 56–60.

Gibbins, Michael (1982) Regression and Other Statistical Implications for Research on Judgement Using Intercorrelated Data Sources. *Journal of Accounting Research,* 20 (1): 121–138.

Glower, Steven M. – Romney, Marshall B. (1998) The Next Generation. *Internal Auditor,* 55 (5): 47–53.

Green, Brain Patrick – Choi, Jae Hwa (1997) Assessing the Risk of Management Fraud Through Neural Network Technology. *Auditing: A Journal of Practice & Theory,* 16 (1): 14–28.

Groomer, S. Michael – Murthy, Uday S. (1989) Continuous auditing of database applications: an embedded audit module approach. *Journal of Information Systems,* 3 (1, Spring): 53–69.

Halal, William E. – Kull, Michael D. – Leffmann, Ann (1998) The George Washington University Forecast of Emerging Technologies: A Continuous Assessment of the Technology Revolution. *Technological Forecasting and Social Change,* 59 (1): 89–110.

Hansen, James V. – McDonald, James B. – Stice, James D. (1992) Artificial Intelligence and Generalized Qualitative-Response Models: An Empirical Test on Two Audit Decision-Making Domains. *Decision Science,* 23 (3): 708–723.

Haykin, Simon (1999) *Neural networks: a comprehensive foundation*. 2nd ed. Prentice-Hall, Inc.: New Jersey.

Hebb, Donald O. (1949) *The Organization of Behaviour*: A Neuropsychological Theory. Wiley: New York.

Hecht-Nielsen, Robert (1991) *Neurocomputing*. Addison-Wesley Publishing Company, Inc.: San Diego.

Hevner, Alan R. – March, Salvatore T. – Park, Jinsoo (2004) Design Science in Information Systems Research. *MIS Quarterly,* 28 (1): 75–105.

Hill, Tim – O'Connor, Marcus – Remus, William (1996) Neural Network Models for Time Series Forecasts. *Management Science,* 42 (7): 1082–1092.

Hopfield, John J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, PNAS*, 79: 2554–2558.

IFAC (2003a) *Handbook of International Auditing, Assurance, and Ethics Pronouncements 2003 Edition: International Auditing Standard on Auditing 520 Analytical Procedures*. International Federation of Accountants. <http://www.ifac.org/Members/DownLoads/2003Handbook.pdf>, retrieved 27.1.2004.

IFAC (2003b) *International Standard of Auditing 500 Audit Evidence*. International Auditing and Assurance Standards Board. <http://www.ifac.org/Members/DownLoads/AuditRisk-ISA500.pdf>, retrieved 28.1.2004.

IFAC (2003c) *Handbook of International Auditing, Assurance, and Ethics Pronouncements 2003 Edition: International Auditing Practice Statement 1009 Computer-Assisted Audit Techniques*. International Federation of Accountants. <http://www.ifac.org/Members/DownLoads/2003Handbook.pdf>, retrieved 28.1.2004.

Johnson, Eric N. – Johnson, Laurence E. (1995) Differences in municipal vs. commercial audit planning analytical procedures. *The CPA Journal,* 65 (12): 50–51.

Järvinen, Pertti (1999) *On Research Methods*. Opinpaja Oy: Tampere.

Järvinen, Pertti – Järvinen, Annikki (2000) *Tutkimustyön metodeista*. Opinpajan kirja: Tampere.

Kaplan, Steven E. – Reckers, Philip M. (1989) An examination of information search during initial audit planning. *Accounting, Organizations and Society,* 14 (5/6): 539–550.

Kasanen, Eero – Lukka, Kari – Siitonen, Arto (1993) The Constructive Approach in Management Accounting Research. *Journal of Management Accounting Research,* 5: 243–264.

KHT-yhdistys (2003) *Tilintarkastusalan suositukset 2003*. KHT-yhdistys: Helsinki.

Kinney, Jr., William R. (1978) ARIMA and Regression in Analytical Review: an Empirical Test. *The Accounting Review,* 53 (1): 48–60.

Kinney, Jr., William R. – Felix, Jr., William L. (1980) Analytical review procedures. *Journal of Accountancy,* 150 (4, Oct.): 98–102.

Kinney, Jr., William R. – Salamon, Gerald L. (1982) Regression Analysis in Auditing: A Comparison of Alternative Investigation Rules. *Journal of Accounting Research,* 20 (2): 350–366.

Kinney, Jr., William R. – Warren, Carl S. (1979) The Decision-Theory Approach to Audit Sampling: An Extension and Application to Receivables Confirmation. *Journal of Accounting Research,* 17 (1): 275–285.

Knechel, W. Robert (1986) A simulation study of the relative effectiveness of alternative analytical review procedures. *Decision Science,* 17 (3): 376–394.

Knechel, W. Robert (1988) The Effectiveness of Statistical Analytical Review as a Substantive Auditing Procedure: A Simulation Analysis. *The Accounting Review,* 63 (1): 74–95.

Koch, Harvey S. (1981) Online computer auditing through continuous and intermittent simulation. *MIS Quarterly,* 1 (5, March): 29–41.

Kogan, Alexander – Sudit, Ephraim F. – Vasarhelyi, Miklos A. (1999) Continuous online auditing: a program of research. *Journal of Information Systems,* 13 (2): 87–103.

Koh, Hian Chye – Tan, Sen Suan (1999) A neural network approach to the prediction of going concern status. *Accounting and Business Research,* 29 (3): 211–216.

Kohonen, Teuvo (1982a) Analysis of a simple self-organizing process. *Biological Cybernetics,* 44 (2): 135–140.

Kohonen, Teuvo (1982b) Self-organized formation of topologically correct feature maps. *Biological Cybernetics,* 43 (1): 59–69.

Kohonen, Teuvo (1997) *Self-Organizing Maps*. Springler-Verlag: Berlin.

Koskivaara, Eija (1997) Artificial neural network models for predicting patterns in auditing monthly balances. In: *Proceedings of the 5th European Conference on Information Systems*, ed. by Robert Galliers – Steven Carlsson – Claudia Loebbecke – Ciaran Murphy – Hans Robert Hansen – Ramon O'Callaghan. Cork, Ireland, June 19–21, 1997, 196–209.

Koskivaara, Eija (2000a) Artificial neural network models for predicting patterns in auditing monthly balances. *Journal of the Operational Research Society,* 51 (9): 1060–1069.

Koskivaara, Eija (2000b) Different Pre-Processing Models for Financial Accounts when Using Neural Networks for Auditing. In: *Proceedings of the European Conference on Information Systems*, ed. by Hans Robert Hansen – Martin Bichler – Harald Mahrer. Vienna, Austria, July 3–5, 2000, 326–332.

Koskivaara, Eija (2002) Design Science Approaches in Information Systems Research. In: *Proceedings of the Information Systems Foundations: Building the Theoretical Base Workshop*, ed. by Shirley Gregor – Dennis Hart. Canberra, Australia, October 1–2, 2002, 205–216.

Koskivaara, Eija (2004a) Artificial neural networks in analytical review procedures. *Managerial Auditing Journal,* 19 (2): 191–223.

Koskivaara, Eija (2004b) Visualization of patterns in accounting data with self-organizing maps. In: *Business Intelligence Techniques: A Perspective from Accounting and Finance*, ed. by Murugan Anandarajan – Asokan Anandarajan – Cadambi Srinivasan, 133–148. Springer-Verlag: New York.

Koskivaara, Eija (2004c) Comparing Different Methods in Analytical Procedures Using Real World Data. In: *Proceedings of the IRMA 2004, Information Resources Management Association*, ed. by Mehdi Khosrow-Pour. New Orleans Louisiana, USA, May 23–26, 2004, 661–663.

Koskivaara, Eija – Back, Barbro (2003) An Artificial Neural Network Based Decision Support System for Budgeting. In: *Proceedings of the Fifth International Conference on Enterprise Information Systems*, ed. by Oliver Camp – Joaquim Filipe – Slimane Hammoudi – Piattini Mario. Angers, France, April 23–26, 2003, 108–114.

Koskivaara, Eija – Back, Barbro – Sere, Kaisa (1996) Modelling Intelligent Information Systems for Auditing. In: *Proceedings of the Intelligent Systems in Accounting and Finance*, ed. by Guillermo J. Sierra – Enrique Bonsón, Punta Umbria Huelva, Spain, September 27–28, 233–247.

Kosko, Bart (1990) Unsupervised Learning in Noise. *IEEE Transactions on Neural Networks,* 1 (1): 44–57.

Kunkel, J. Gregory (1974) Continuous auditing by exception. *Management Accounting,* 56 (1): 45–48.

Lee, Bill (2002) Professional socialisation, commercial pressures and junior staff's time-pressured irregular auditing – a contextual interpretation. *British Accounting Review,* 34 (4): 315–333.

Lee, Minwoo – Colbert, Janet L. (1997) Analytical procedures: management tools for monitoring controls. *Management Decision,* 35 (5): 392–397.

Leitch, Robert A. – Chen, Yining (2003) The Effectiveness of Expectations Models in Recognizing Error Patterns and Generating and Eliminating Hypotheses While Conducting Analytical Procedures. *Auditing: A Journal of Practice & Theory,* 22 (2): 147–170.

Lenard, Mary Jane – Alam, Pervaiz – Madey, Gregory R. (1995) The Application of Neural Networks and a Qualitative Response Model to the Auditor's Going Concern Uncertainty Decision. *Decision Science,* 26 (2): 209–227.

Lev, Baruch (1980) On the Use of Index Models in Analytical Reviews by Auditors. *Journal of Accounting Research,* 18 (2): 524–550.

Lieb, Eugene B. – Gillease, Joanne K. (1996) Du Pont Uses a Decision Support System to Select Its Audit Portfolio. *Interfaces,* 26 (3): 10–21.

Lin, Kenny Z. – Fraser, Ian A. M. (2003) The use of analytical procedures by external auditors in Canada. *Journal of International Accounting, Auditing & Taxation,* 12 (2): 153–168.

Lin, Kenny Z. – Fraser, Ian A. M. – Hatherly, David J. (2003) Auditor analytical review judgement: a performance evaluation. *The British Accounting Review,* 35 (1): 19–34.

Lukka, Kari (2002) *The constructive research approach*. Metodix.com. <http://www.metodix.com/showres.dll/en/enindex>, retrieved 17.6.2002.

Marcella, Jr., Albert J. – Rauff, James (1995) Utilizing Expert Systems to Evaluate Disaster Recovery Planning. *Journal of Applied Business Research,* 11 (1): 30–39.

March, Salvatore T. – Smith, Gerald F. (1995) Design and natural science research on information technology. *Decision Support Systems,* 15 (4): 251–266.

Martín-del-Brío, Bonifacio – Serrano-Cinca, Carlos (1993) Self-organizing Neural Networks for the Analysis and Representation of Data: Some Financial Cases. *Neural Computing and Applications,* 1 (3): 193–206.

Matsumura, Ella Mae – Tsui, Kam-Wah (1982) Stein-Type Poisson Estimators in Audit Sampling. *Journal of Accounting Research,* 20 (1): 162–170.

McCarthy, William E. – Denna, Eric – Gal, Graham (1992) Expert Systems and AI-based Decision Support in Auditing: Progress and Perspectives. *International Journal of Intelligent Systems in Accounting, Finance and Management,* 1 (1): 53–63.

McCulloch, Warren S. – Pitts, Walter (1943) A logical calculus of the ideas imminent in nervous activity. *Bulletin of Mathematical Biophysics,* 5 (1): 115–133.

McKee, Thomas E. (1992) An Audit Framework for Expert Systems. *International Journal of Intelligent Systems in Accounting, Finance and Management,* 1 (4): 261–273.

Medsker, Larry – Liebowitz, Jay (1994) *Design and Development of Expert Systems and Neural Networks*. Macmillan College Publishing Company, Inc.: New York.

Mennon, Anil – Mehrotra, Kishan – Mohan, Chilukuri K. – Ranka, Sanjay (1996) Characterization of a class of sigmoid functions with applications to neural networks. *Neural Networks,* 9 (5): 819–835.

Minsky, Marvin – Papert, Seymour (1969) *Perceptrons: An Introduction to Computational Geometry*. The MIT Press: Cambridge, MA.

Mulqueen, John T. (1996) Neural nets block fraud. *CommunicationsWeek,* 625, 8/26: 95–96.

Murphy, David S. – Brown, Carol E. (1992) The Use of Advanced Information Technology in Audit Planning. *International Journal of Intelligent Systems in Accounting, Finance and Management,* 1 (3): 187–192.

Nelson, Mark W. – Libby, Robert – Bonner, Sarah E. (1995) Knowledge Structure and the Estimation of Conditional Probabilities in Audit Planning. *The Accounting Review,* 70 (1): 27–47.

Niiniluoto, Ilkka (1986) Pragmatismi. In: *Vuosisatamme filosofia*, ed. by Ilkka Niiniluoto – Esa Saarinen, 40–73. WSOY: Helsinki.

Nunamaker, Jr., Jay F. – Chen, Minder – Purdin, Titus D. M. (1991) Systems Development in Information Systems Research. *Journal of Management Information Systems,* 7 (3): 89–106.

O'Donnell, Ed (2002) Evidence of an Association between Error-Specific Experience and Auditing Performance during Analytical Procedures. *Behavioral Research in Accounting,* 14: 179–195.

Oikeusministeriö (1994) *Finnish Audit Act*. <http://www.Finlex1.edita.fi>, retrieved 7.7.2003.

O'Leary, Daniel E. (1998) Using Neural Networks to Predict Corporate Failure. *International Journal of Intelligent Systems in Accounting, Finance & Management,* 7 (3): 187–197.

Ramamoorti, Sridhar – Bailey, Andrew D. Jr – Traver, Richard O. (1999) Risk Assessment in Internal Auditing: A Neural Network Approach. *International Journal of Intelligent Systems in Accounting, Finance & Management,* 8 (3): 159–180.

Razaee, Zabihollah – Elam, Rick – Sharbatoghlie, Ahmad (2001) Continuous auditing: the audit of the future. *Managerial Auditing Journal,* 16 (3): 150–158.

Riedmiller, Martin (1994a) Advanced supervised learning in multi-layer perceptrons – From backpropagation to adaptive learning algorithms. *Journal of Computer Standards and Interfaces,* 16 (3): 265–278.

Riedmiller, Martin (1994b) Supervised Learning in Multilayer Perceptrons - from Backpropagation to Adaptive Learning Techniques. <http://amy.informatik.uos.de/riedmiller/publications.html>, retrieved 17.6.2002.

Riedmiller, Martin – Braun, Heinrich (1993) A direct adaptive method for faster backpropagation learning: the Rprop algorithm. In: *Proceedings of the IEEE International Conference on Neural Networks, ICNN. San Francisco California, USA, March 28 – April 1, 1993*. <http://amy.informatik.uos.de/riedmiller/publications.html>, retrieved 17.6.2002.

Rosenblatt, Frank (1958) The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review,* 65: 386–408.

Rumelhart, David E. – Hinton, Geoffrey E. – Williams, Ronald J. (1986) Learning internal representations by error propagation. In: *Parallel Distributed Processing*, ed. by D.E. Rumelhart – J.L. McClelland, 318-362. MIT Press, Cambridge, MA.

Rumelhart, David E – Widrow, Bernard – Lehr, Michael A (1994) The Basic Ideas in Neural Networks. *Communications of the ACM,* 37 (3): 87–92.

Russell, Stuart J. – Norvig, Peter (1995) *Artificial Intelligence A Modern Approach*. Prentice-Hall, Inc.: Upper Saddle River.

Saario, Martti (1959) *Kirjanpidon meno-tulo-teoria*. Otava: Porvoo.

Searcy, DeWaynen L. – Woodroof, Jon B. (2003) Continuous auditing: leveraging technology. *The CPA Journal,* 75 (5): 46–48.

Simon, Herbert (1969) *The Sciences of the Artificial*. MIT press: Cambridge.

Simon, Herbert (2002) Forecasting the future or shaping it? *Industrial and corporate Change,* 11 (3): 601–605.

Smith, Murray (1996) *Neural Networks for Statistical Modeling*. International Thompson Computer Press: Boston.

Sohl, Jeffrey E. – Venkatachalam, Angadipuram R. (1995) A neural network approach to forecasting model selection. *Information & Management,* 29 (6): 297–303.

Sutton, Steve G. – Young, Ronald - McKenzie, Phyllis (1995) An Analysis of Potential Legal Liability Incurred Through Audit Expert Systems. *International Journal of Intelligent Systems in Accounting, Finance and Management,* 4 (3): 191–204.

Swingler, Kevin (1996) *Applying Neural Networks A Practical Guide*. Academic Press Limited: London.

Tam, Kar Yan – Kiang, Melody Y. (1992) Managerial Applications of Neural Networks: The Case of Bank. *Management Science,* 38 (7): 926–947.

Tiittanen, Anne (1998) *The Role of Information Technology and IS/IT User Support Services in Modern Auditing*. Licentiate thesis. Swedish School of Economics and Business Administration: Helsinki.

Toiviainen, Kari (1991) *Tietotekniikan hyväksikäyttö tilintarkastuksessa*. Helsinki School of Economics and Business Administration: Helsinki.

Turpiainen, Janne (1994) Atk-avusteisen tilintarkastuksen työvälineet. Assistant tools for auditing. *Tilintarkastus,* 2: 143–146.

Waltari, Mika (1945) *Sinuhe egyptiläinen*. WSOY: Juva.

Van den Acker, Carine – Vanthienen, Jan (1995) The Definition and Diffusion of Expert Systems: An Empirical Survey of Bergian External Audit Firms. *International Journal of Intelligent Systems in Accounting, Finance and Management,* 4 (3): 205–220.

Vasarhelyi, Miklos A. – Halper, Fern B. (1991) The continuous audit of online systems. *Auditing: A Journal of Practice & Theory,* 10 (1): 110–125.

Vasarhelyi, Miklos A. – Kogan, Alexander – Alles, Michael (2002) Would continuous audit have prevented the Enron mess? *The CPA Journal,* 72 (7): 80.

Weber, John C. – Goldstein, Ellen J. (1999) *Public Accounting – Practice Alert 98-1*. CPA Letter Supplements. <http://www.aicpa.org/pubs/cpaltr/may98/suppl/pubacc.htm>, retrieved 26.4.2002.

Vellido, Alfredo – Lisboa, Paolo J.G. – Vaughan, John (1999) Neural networks in business: a survey of applications (1992–1998). *Expert Systems with Applications,* 17 (1): 51–70.

Widrow, Bernard – Hoff, Jr., Marcian E. (1960) Adaptive switching circuits. In *Neurocomputing:Foundations of Research*, 96–104. IRE WESCON Convention Record, New York.

Willet, Caroline – Page, Michael (1996) A Survey of Time Budgets and Irregular Auditing Practices among Newly Qualified UK Chartered Accountants. *The British Accounting Review,* 28 (2): 101–120.

Wong, Bo K. – Bodnovich, Thomas A. – Selvi, Yakup (1995) A bibliography of neural network business applications research: 1988-September 1994. *Expert Systems,* 12 (3): 253–262.

Wong, Bo K. – Selvi, Yakup (1998) Neural network applications in Finance: A review and analysis of literature (1990–1996). *Information & Management,* 34 (3):129–139.

Woodroof, Jon – Searcy, DeWayne (2001) Continuous audit: model development and implementation within a debt covenant compliance domain. *International Journal of Accounting Information Systems,* 2 (3): 169–191.

Woolf, Emile (1994) *Auditing Today*. Prentice Hall: Hertfordshire.

Wu, Rebecca Chung-Fern (1994) Integrating Neurocomputing and Auditing Expertise. *Managerial Auditing Journal,* 9 (3): 20–26.

Ye, L. Richard – Johnson, Paul E. (1995) The Impact of Explanation Facilities on User Acceptance on Expert Systems Advise. *MIS Quarterly,* 19 (2): 157–183.

Zhang, Guoqiang – Patuwo, B. Eddy – Hu, Michael Y. (1998) Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting,* 14 (1): 35–62.

PUBLICATION 1

Koskivaara (2004a) Artificial neural networks in analytical review procedures. *Managerial Auditing Journal,* 19 (2): 191–223. Reprinted by permission of Emerald Group Publishing Limited.

The Emerald Research Register for this journal is available at
www.emeraldinsight.com/researchregister

The current issue and full text archive of this journal is available at
www.emeraldinsight.com/0268-6902.htm

Artificial neural networks in analytical review procedures

Eija Koskivaara
Turku Centre for Computer Science and Turku School of Economics and Business Administration, Turku, Finland

Keywords *Auditing, Neural nets, Financial performance*

Abstract *This article gives an overview of artificial neural network (ANN) studies conducted in the auditing field. The review pays attention to application domains, data and sample sets, ANN-architectures and learning parameters. The article argues that these auditing ANN-applications could serve the analytical review (AR) process. The summary of the findings pays attention to whether authors state that ANNs have potential to improve analytical review (AR) procedures. Furthermore, the article evaluates which are the most influential contributions and which are open ends in the field. The article makes some practical suggestions to motivate academics and practitioners to collaborate in further exploration of the potential of ANNs.*

1. Introduction

The article gives an overview of artificial neural network (ANN) studies conducted in the auditing field. The development of tools for auditors is important with regard to the workload and demands of auditors. Furthermore, the recent events in the business and auditing environment have highlighted problems in the auditing process.

Auditors may need to take the necessary steps to restore public confidence in the capital market system and accounting profession that might have been shaken by the collapse of Enron and Arthur Andersen and many others. A well-known fact is that too many companies have problems with their bookkeeping. Siebel Systems, Qwest, WorldCom, and Xerox are examples of companies that have been “cooking the books”. We can call the year 2002 “the horrible year” from a bookkeeping point of view. Unfortunately the year 2002 was not an exception. This manipulation is still going on. A fairly recent example is the Dutch retail trade company Royal Ahold, whose subsidiary in the USA manipulated the operating profit. There are also examples from earlier years. Two of them come from the banking world, where the UK Baring’s Bank and the Japanese Daiwa’s Bank lost millions of dollars because they did not have effective control systems. The third one is a Finnish multinational company, whose subsidiary in Italy overestimated the work in progress and

The author appreciates comments received from anonymous referees and would also like to thank Barbro Back from the Turku Centre for Computer Science (TUCS) for her beneficial comments. The author wishes to thank the Foundation for Economic Education for providing financial support for this project.

Managerial Auditing Journal
Vol. 19 No. 2, 2004
pp. 191-223

0268-6902
DOI 10.1108/02686900410517821

recorded fictitious sales. Although the above examples are the tip of the iceberg, they reflect the changing nature and demands in the audit process.

A common fact is that many parties, such as shareholders, investors, creditors, tax authorities, and managers are interested in the accuracy of organisations' financial performance. Auditors are in a key position to monitor and control operations in organisations. The increasing use of information technology (IT) and computers in organisations requires auditors to obtain and evaluate evidence electronically. Technology has made the input of information for transactions and events easier and the evaluation of the relevant events more critical. Companies report their financial outcome quarterly and an increasing number of companies present their financial information on a public network. Sometimes the speed at which these reports are made makes one wonder whether all the relevant information is audited and reliable.

Needless to say, auditors need tools and methods that provide them with objective information about a client company (Willet and Page, 1996; Lee, 2002; Bazerman *et al.*, 2002). Sophisticated auditing tools could be a possible solution for preventing companies from manipulating account value and for helping auditors to answer the demands of today's business environment. One possible way to develop auditing tools is to embed ANNs in analytical review (AR) procedures.

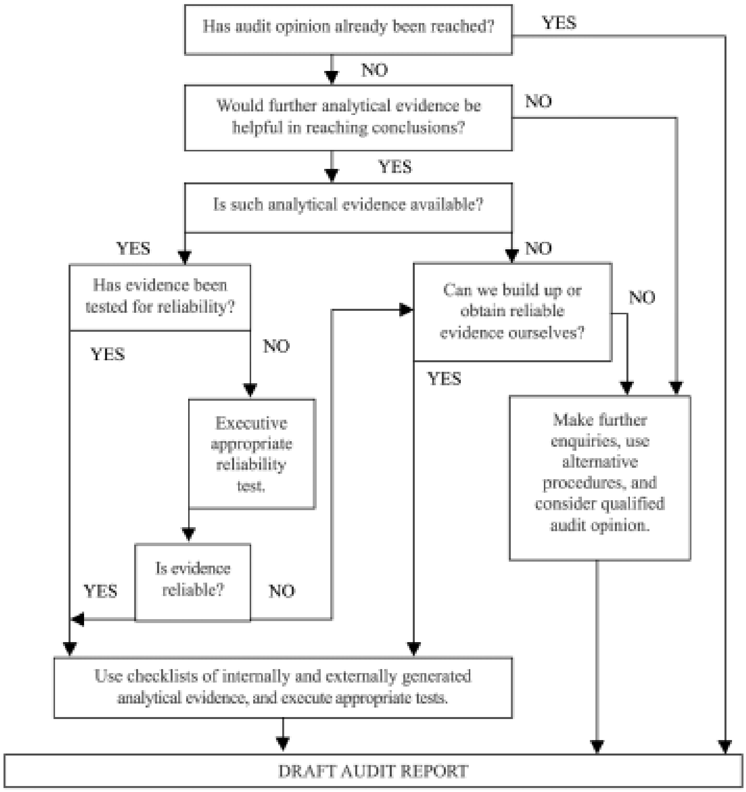
The article proceeds as follows. Section 2 provides an overview of the AR in auditing. Section 3 gives a brief description of ANNs. Section 4 summarises ANNs studies conducted in the field of auditing. The discussion section discusses the findings.

2. Analytical review in auditing

The role of AR in auditing is to obtain audit evidence. The AR is an umbrella term for different kinds of analytical procedures. Various methods or techniques may be used in performing the analytical procedures.

In Woolf's (1994) book *Auditing Today* the AR is shown as a decision-type flow diagram (Figure 1). The flowchart suggests that the existence of available analytical evidence, and the creation or procurement of evidence by the auditor herself/himself may be embodied in the report.

The flowchart suggests that the existence of available analytical evidence and the creation or procurement of evidence may be viewed as alternatives. In practice, where both appropriate and practicable, the auditor will make use of both sources of evidence. The flowchart makes reference to "check-lists of internally and externally generated analytical evidence". The terms "internally" and "externally" should be taken to relate respectively to evidence created within the client organisation and that which has been independently generated by the auditor. The precise form of such a check-list will vary considerably from one firm to another.



Source: Adapted from Woolf (1994)

Figure 1.
Flowchart frameworks of analytical review

2.1 Analytical procedures

Analytical procedures use comparison and relationships to assess whether account balances or other data appear reasonable (Arens *et al.*, 2003). Analytical procedures are performed principally at any of the three phases (planning, testing, and completion) during an audit engagement. For example, in the USA, the use of analytical procedures in the planning and completion phases of an audit is required under generally accepted auditing principles (GAAP). As the statement on auditing standards (SAS) No. 56 states:

> Analytical procedures involve comparisons of recorded amounts, or ratios developed from recorded amounts, to expectations developed by the auditor. The auditor develops such expectations by identifying and using plausible relationships that are reasonably expected to

> exist based on the auditor's understanding of the client and of the industry in which the client operates . . . (AICPA, 1988).

The emphasis in the SAS 56 definition is on expectations developed by the auditor. Analytical procedures use methods and techniques to improve the efficiency of audits by developing these expectations and by comparing them with recorded amounts. The use of analytical procedures entails the examination of the accuracy of account balances without considering the details of the individual transactions, which make up the account balance. They play an important role in assisting the auditor in determining the nature, timing and extent of his or her substantive testing, and in forming an overall opinion as to the reasonableness of recorded account values (Weber and Goldstein, 1999). Basically, analytical procedures compare expected relationships among data items with actual observed relationships.

Examples of the purposes of analytical procedures for each of the three phases are shown in Table I. The X in the boxes in the matrix indicates that a certain purpose is applicable to a certain phase. The purposes vary for different phases of the audit. Analytical procedures are performed during the planning for all four purposes, whereas the other two phases are used primarily to determine appropriate audit evidence and to reach conclusions about the fair presentation of financial statements.

An important part of using analytical procedures is selecting the most appropriate procedures. According to Arens *et al.* (2003), there are five major types of analytical procedure:

(1) compare client and industry data;

(2) compare client data with similar prior-period data;

(3) compare client data with client-determined expected results;

(4) compare client data with auditor-determined expected results; and

(5) compare client data with expected results using non-financial data.

These procedures are facilitated by many of the firm's software tools, which provide direct linkages to client data. In practice the auditor decides which of the available analytical procedures to use, enters the data to make the calculations, and evaluates the results.

Table I.
Timing and purposes of analytical procedures

	Phase		
Purpose	Planning	Testing	Completion
Understand client's industry and business	×		
Assess going concern	×		×
Indicate possible misstatements	×	×	×
Reduced detailed test	×	×	

Source: Adapted from Arens *et al.* (2003)

2.2 Analytical review techniques

Methods and techniques embedded in analytical procedures range from simple comparisons to complex analyses using advanced statistical techniques (Woolf, 1994) or systems based on computational paradigms such as neural networks. Auditing researchers have developed and used a variety of models and techniques to assist in the AR process. Blocher and Patterson (1996) have identified three types of AR techniques: trend analysis, ratio analysis, and model-based procedures. Fraser *et al.* (1997) have provided a slightly broader classification perspective for AR techniques: non-quantitative (NQT) or judgemental, such as scanning; simple quantitative (SQT), such as trend, ratio and reasonableness tests; and advanced quantitative (AQT), such as regression analysis and ANNs.

These procedures differ significantly in their ability to identify potential misstatement. Judgemental techniques constitute subjective evaluations based on client knowledge and past experience. Trend analysis assesses whether there is a functional relationship between the variables over time. Ratio analysis incorporates directly the expected relationships between two or more accounts. For example, turnover ratios are useful because there is typically a stable relationship between sales and other financial statement accounts, especially receivables and inventory. Although ratios are easy to compute, which in part explains their wide appeal, their interpretation is problematic, especially when two or more ratios provide conflicting signals. Indeed, ratio analysis is often criticised on the grounds of subjectivity, i.e. the auditor must pick and choose ratios in order to assess the overall performance of a client. In a reasonableness test, the expected value is determined by reference to data partly or wholly independent of the accounting information system and, for that reason, evidence obtained through the application of such a test may be more reliable than evidence gathered using other analytical procedures. For example, the reasonableness of the total annual revenue of a freight company may be estimated by calculating the product of the total tons carried during the year and the average freight rate per ton. Regression analysis predicts financial and operating data by incorporating economic and environmental factors into formal models.

This article reviews the ANN studies carried out in the auditing field. The article argues that ANNs are a technique that aids auditors in creating expectations and these expectations can then be compared with actual values automatically (cf. SAS 56). ANNs have many beneficial aspects in comparison with other techniques. They are adaptive tools for processing data. They can learn, remember, and compare complex patterns (Medsker and Liebowitz, 1994). They are claimed to be able to recognise patterns in data, even when the data are noisy, ambiguous, distorted, or variable (Dutta, 1993). Basically, ANNs learn from examples and then generalise the learning to new observations. Compared with regression analysis we do not need an a priori model because

ANNs are data-driven. ANNs are, unlike traditional statistical techniques, capable of identifying and simulating non-linear relationships in the data without any a priori assumptions about the distribution properties of the data. Therefore, one advantage of ANN-systems could be that they provide additional information to the decision process. With the help of an ANN an auditor may find something from the data that might be left hidden otherwise and therefore ANNs are potentially suitable for many tasks within auditing. Furthermore, ANNs have been considered one of the emerging technologies (Halal *et al.*, 1998). Information technology development and the processing capacities of PCs have made it possible to model ANN-based information systems for monitoring and controlling operations.

3. Artificial neural networks

An ANN consists of a set of neurons (nodes) (Figure 2a) that interact through a dense web of interconnections. Each neuron has a value y_j, which is propagated through connections to other neurons in the network. Each connection has a weight denoted by w_{ij}, which dictates the effect of the jth neuron to the ith neuron. The neuron fires if the combined signal strength of the weights exceeds a certain threshold, θ_j. A net value, u_j, is yielded by an activation function. The standard form of the activation function is: $u_j(w, x) = \sum(w_{ij}, x_j)$. There are also other activation functions, like the logistic function and the tanh function (Swingler, 1996).

The net value is immediately transformed by a transfer function of the neuron. There are many possible transfer functions, i.e. step, sigmoid, hyperbolic, linear, bubble, Gaussian, and Mexican-hat function. The sigmoid transfer function, $f(u_j) = 1/(1 + e_j^{-u})$, is often used. It ranges from 0 to 1. The

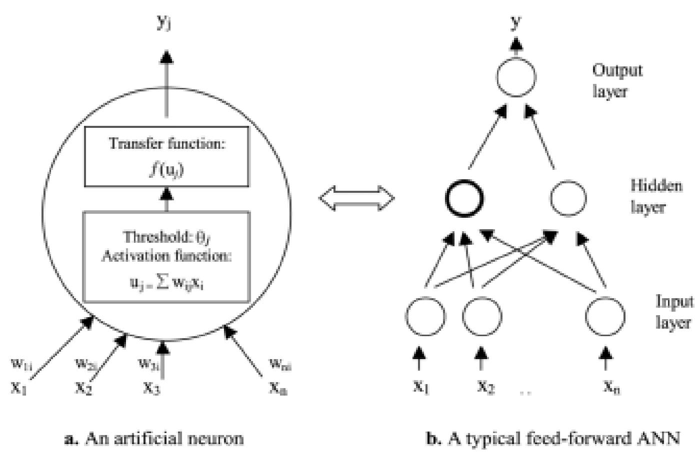


Figure 2.
Part (a) is an artificial neuron; part (b) is a typical feed-forward ANN

final output y can usually be expressed as a function of the input and the weights.

The neurons in an ANN are organised into layers (Figure 2b). The first layer is called the input layer and the last layer is the output layer. The inner layers are known as hidden layers. The numbers and also the names of the layers differ between different neural networks. A self-organising map (SOM) has no hidden layer at all (see Section 3.2.2). A backpropagation network (BPN) has one or more hidden layers. A categorical learning network (CLN) has normalization and pattern layers. The normalization layer helps the pattern layer to develop correct classes for the input vectors. The pattern layer uses unsupervised learning to distinguish between categories. A probabilistic neural network (PNN) has normalization, pattern, and summation layers. The summation layer helps in classifying data sets. Etheridge *et al.* (2000) compared the performance of the BPN, CLN, and PNN architectures (see Section 4.2.3).

Two layers of neurons communicate via a weight connection network. There are three types of connections. Feed-forward connections mean that data from neurons of a lower layer are propagated forward to neurons of an upper layer via feed-forward weight connection networks. Feedback connections networks bring data from neurons of an upper layer back to neurons of a lower layer. An example of a lateral connections network is the winner-takes-all circuit, which serves the important role of selecting the winner.

There are multiple approaches depending on, for example, how the weights are connected and what the basic and activation functions are. These interconnections and functions can make ANNs mathematically very complex and sophisticated. Each of these approaches has a unique mix of, for example, information-processing capabilities, domains of applicability, techniques for use, required training data, and training methods (Hecht-Nielsen, 1990). The different approaches do not compete against one another; rather, they represent various specialisations in solving different types of problems. Indeed, the approach strongly influences, for example, what the network can do (Hertz *et al.*, 1991). Section 3.1 discusses an important and attractive feature of an ANN. This is its learning capability.

3.1 Training and testing

As mentioned earlier, ANNs are data-driven models. This means that ANNs are assumption-free approaches for estimating functions from sample data. To build an ANN model one typically needs training and a test sample. The training sample is used to determine the weights and parameters that define the ANN model. The test sample is adopted for evaluating the model. The training and testing set size depends on the problem domain and on available data. Sometimes a third one called the validation sample is utilised to avoid the overfitting problem or to determine the stopping-point of the training process.

Occasionally, only one set is used for both validation and testing purposes, particularly with small data sets.

The training of neural networks can be classified into three categories:

(1) supervised;

(2) graded (reinforcement); and

(3) unsupervised (self-organisation) (Hecht-Nielsen, 1990).

Reinforcement learning is a special case of supervised learning (Hertz *et al.*, 1991). Thus, a learning mode is simply supervised or unsupervised. The distinction between supervised and unsupervised learning depends on information, e.g. pattern recognition is supervised if the training algorithm requires knowledge of the class membership of the training samples, unsupervised if it does not require it (Kosko, 1990). In the following we have taken a closer look at the different learning modes.

3.2 Learning modes

Learning occurs by training and, as mentioned above, there are two ways to train a network, supervised and unsupervised. In supervised learning a teacher has some knowledge of the environment that is unknown to an ANN. The teacher expresses this knowledge with training examples, which consist of input variables together with desired target values (Hecht-Nielsen, 1990). The network processes its output values from the input variables and compares them with the target output values. If an error, i.e. a difference between outputs and targets, exists, the network adjusts the weights by a small amount in some direction in a step-by-step manner until the error is at an acceptable level. Therefore, supervised learning is an instructive feedback system.

In unsupervised learning nobody oversees the learning process. Therefore, the network is given only the training data inputs from which the network organises itself into some useful configuration (Hecht-Nielsen, 1990). The input vectors are classified according to their degree of similarity. The similar input vectors activate the same output cluster. The user is responsible for giving an interpretation to the clusters.

The hybrid use of learning paradigms can provide a better solution than one paradigm alone. For example, when similar input vectors produce similar outputs, it may be rational to categorise the inputs first with unsupervised learning and use that information for supervised learning (Hertz *et al.*, 1991).

A learning algorithm is a set of well-defined rules for the solution of a learning problem. Several alternative learning algorithms, e.g. scale conjugate gradient (SCG), Levenberg-Marquardt (LM), backpropagation (BP) and conjugate gradient with Powel-Beale restarts (PBR), and competitive learning algorithm, exist and they all have their own specific advantages. The differences between them are based on various weight adjustments (Haykin, 1994). Detailed descriptions of learning algorithms can be found in the

ANN-books of Hecht-Nielsen (1990), Hertz *et al.* (1991), Freeman and Skapura (1991), Smith (1996), Swingler (1996), or Demuth and Beale (2000).

3.2.1 Backpropagation algorithm. The BP algorithm is used in supervised learning mode and it operates on a multi-layered feed-forward network (see Figure 2b). When the relationship between the input and output variables is non-linear, a hidden layer helps in extracting higher level features and facilitates the generalisation of outputs. This kind of network could include more hidden layers but it has been proven that one hidden layer can approximate even complex functions quite well. The network can also include several output neurons. Each neuron in the hidden or output layer is connected to all of the neurons in the layer below it and a weight is associated with each of the incoming connections of the neuron.

For a given input vector, the BP generates the output vector by a forward pass. When a neuron receives inputs, it computes its output value and sends it to the neurons on the next layer above. Thus, the inputs are fed forward through the entire network until they reach the output layer. Then, the difference between the output vector and the desired target vector, the root mean square error (RMSE), is backpropagated through the ANN to modify the weights for the entire neural network. This iterative process is called training. After the BP network has been trained, it is tested against the records of a testing data set that have not been previously met with the network. For these records, the desired target output is known. The output generated for each record of these testing data is checked against the desired target output for that record. If there is a match, it is concluded that the trained network could recognise the record correctly.

The BP algorithm has two important parameters: learning rate and momentum. The learning rate affects the speed with which the network settles on a solution by allowing us to regulate how much the error decreases in each iteration. The adaptive learning rate accelerates the learning process by utilising the concept of the direction in which the error has been decreasing recently (Smith, 1996). It speeds up the training process when the ANN is far away from the correct weights and slows down when the ANN gets closer. Momentum is another way to increase the speed of convergence, when calculating the weight change at each iteration; a fraction of the previous direction is added. This additional term tends to keep the weight changes going in the same direction.

The BP algorithm has become the most popular one for prediction and classification problems (Sohl and Venkatachalam, 1995). In business applications, BP is the most commonly used algorithm (Wong *et al.*, 1995; Wong and Selvi, 1998). Classification and prediction tasks, which require some modelling, are especially suitable for this type of network (Klimasauskas, 1991; Lehtokangas *et al.*, 1994). Hence, a multi-layer perceptron ANN can be considered, e.g. for forecasting where statistical methods like Box-Jenkins are used (Klimasauskas, 1991). This algorithm could also be possible for online process controls or adaptive controllers. It turned out to be the most popular in the auditing applications, as the review shows.

3.2.2 Self-organising maps. The SOM is an example of competitive learning (unsupervised learning mode). It is a clustering and visualisation method and the purpose is to show the data set in another representation form (Kohonen, 1997). A SOM has an input layer and an output layer. It is an example of an unsupervised learning algorithm where all the output neurons compete against one another (Haykin, 1994). One of them will be a winner in accordance with a chosen metric and only it will be activated. The winner's weight vector is updated to correspond more closely with the input vectors. This means that two input items, which are close in the input space, are mapped into the same or neighbouring neurons on the map. Output neurons create groups, which together form a map of the input neurons. It creates a two-dimensional map from n-dimensional input data. This map resembles a landscape in which it is possible to identify borders that define different clusters (see Figure 3). The SOM-map is on the left in Figure 3 and on the right we have identified and named clusters according to the accounts these clusters contain. These clusters consist of the input variables with similar characteristics. This algorithm can discover features that may be used to classify a set of input vectors. Or, this algorithm may reveal something from the data that might be left hidden otherwise.

The SOM has six learning parameters: topology, neighbourhood type, X- and Y-dimensions, training rate, training length, and network radius. The network topology refers to the form of lattice. There are two commonly used lattices: rectangular and hexagonal (Figure 4). In a rectangular lattice each neuron is connected to four neighbours, except for the ones at the edge of the lattice. In a hexagonal lattice structure neurons are connected to exactly six neighbours, except for the ones at the edge of the lattice. The SOM-map in Figure 3 is built using a hexagonal lattice. Neighbourhood type refers to the neighbourhood function used and the options are Gaussian and bubble. X- and

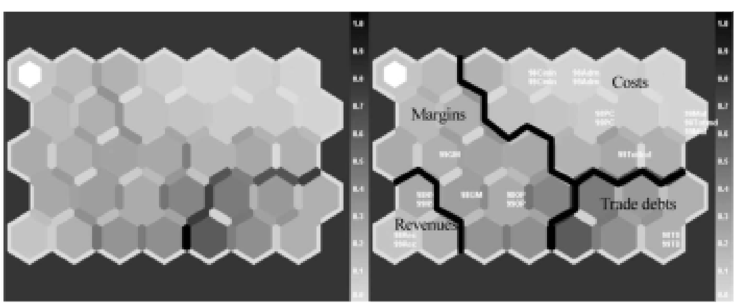


Figure 3.
The SOM-map and the clusters on the map

Y-dimensions refer to the size of the map. In too small maps differences between clusters are hard to identify and in too large maps clusters will appear to be flat. The training rate factor refers to how much the neuron in the neighbourhood of the winning neuron learns from the input data vector. The training length measures the processing time, i.e. the number of iterations through the training data. The network radius refers to how many nodes around the "winning" neuron are affected during the learning process. The training process of the network consists of two parts. In part one, the map is trained "roughly". In the second part, the network is fine-tuned.

4. ANNS in auditing

The following review takes a broad scope of auditing and encompasses both internal and external auditing. The scope and criteria for literature search are presented in Section 4.1. The review focuses on the application domains of the articles found. We argue that these auditing ANN-applications could serve the analytical review process. In reviewing the modelling issues we have paid special attention to the following issues: data and sample sets, ANN-architectures and learning parameters, and whether the model was evaluated with a hold-out sample. In the summary of the findings in Table VI we have paid attention to whether authors state that ANNs have potential to improve AR procedures.

4.1 Scope and criteria for literature search

For finding the studies on ANNs in auditing we used the ABI Inform/Proquest, Ebscohost, Emerald, JTORS, and Elsevier with search words "audit and neural". We also reviewed the electronic tables of contents of *Intelligent Systems in Accounting, Finance and Management, Expert Systems with Application & Auditing: A Journal of Practice and Theory for Years 1990-2003*, and a small number of proceedings from our university library. Furthermore, we reviewed special survey articles on ANNs applied to business situations. Wong *et al.* (1995) and Wong and Selvi (1998) surveyed articles from 1988-1996 and classified the articles among others by application areas. Only one article was categorised as belonging to the auditing discipline. O'Leary (1998) analysed 15 articles that applied ANNs to predict corporate failure. He provided information on data, ANN models, software, and architecture. Zhang *et al.*

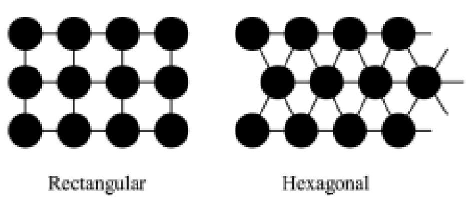


Figure 4.
Forms of lattice

(1998) surveyed 21 articles that addressed modelling issues when applying ANNs for forecasting. They compared the relative performance of ANNs with traditional methods in 24 cases. None of these articles pertained to auditing problems; however, the paper provides insights into modelling issues by summarising suggestions. Vellido *et al.* (1999) surveyed 123 articles from 1992 to 1998, of which six articles pertained to auditing problems. Besides modelling issues they summarise the most frequently cited advantages and disadvantages of the ANN models. Coakley and Brown (2000) surveyed accounting and finance ANN applications and classified them by research question, type of output (continuous versus discrete), and parametric model. None of these articles was categorized into audit discipline.

Altogether 22 articles, either focusing on or connected with the auditing environment, were found. This is seven auditing ANN articles more than in the Calderon and Cheh (2002) study, where they focus on ANNs as an enabler of the business risk auditing framework. They had also included in their study 12 bankruptcy studies, which we have excluded. Our review focuses only on those applications that are conducted from the auditing perspective. We focus on the application areas and the auditing point of view in applications. We also found that the ANN applications fit in different application areas in analytical auditing.

4.2 ANN applications in auditing

The main ANN-application areas in auditing are material errors (nine studies), management fraud (three studies), and support for going concern decision (five studies), and for financial distress problems (one study) ANNs have also been applied to internal control risk assessment (three studies), and audit fee forecasting (one study). Next, we outline the studies in chronological order by application areas. Every application area review ends with the summarisation table of modelling issues in that particular area. Almost every model was evaluated with a hold-out sample. The performance rate of ANN models is based on the hold-out samples in the following review.

4.2.1 Material errors. The major ANN-application area in auditing is material errors. Material error applications direct auditors' attention to those financial account values where the actual relationships are not consistent with the expected relationships. An auditor has to decide whether and what kind of further audit investigation is required to explain the unexpected results. Material error ANN-models either predict future values or classify data. Table II summarises the modelling issues of ANN literature pertaining to material error problems.

Coakley and Brown (1991a, 1993) and Coakley (1995) tested whether an ANN offered improved performance in recognising material misstatements. They used four years' monthly data of a medium-sized distributor. Their ANN-model was based on trend prediction. Three years' data were used for a training set

Researcher country	Industry data type	Train/test set	No. inputs nodes	No. hidden layers/nodes	No. outputs	Transfer function	Learning algorithm	Performance measure
Coakley and Brown (1991a) USA	Distributor Monthly account values	48/48	11	2:	10	Sigmoid	BP	Average error, standard deviation
Coakley and Brown (1991b) USA	Manufacturing firm Monthly account values, aggregates	36/36	42	1:15	15	Modified sigmoid squashing	BP	MSE (mean square error)
Coakley and Brown (1993) USA	Distributor Monthly account values, aggregates	36/12[a]	42	1:15	15	Modified sigmoid	BP	MSE
Wu (1994) Taiwan	Sample of firms income tax behaviour data	90/90	16	1:8 0:0	1	Sigmoid	BP	Predictive accuracy
Coakley (1995) USA	Distributor Monthly financial ratios	36/12[a]	5	2:11-11	3	Hyperbol. tangent activation	BP	SSE (sum of square error)
Koskivaara *et al.* (1996) Finland	Manufacturing firm monthly income statement values	54/12[a]	30	1:16	9	Sigmoid	BP	RMSE
Busta and Weinberg (1998) USA	Simulated data Digits of numbers	800/800[a]	34 24 15 5 1 1	1:4 1:6 1:6 1:6 1:4 0:	1	Logistic	BP	Accuracy per cent
Koskivaara (2000) Finland	Sample of manufacturers Financial statements' account values Non-financial values	25/6[a]	37	3:32-26-16; 1:27; 4:33-29-25-26; 1:27	16	Sigmoid	BP	RMSE QR
Koskivaara (2000b) Finland	Manufacturing firm Monthly balance account values	60/12[a]	30 48	2:27-18 4:40-32-26-18	9	Sigmoid	BP	RMSE

Note: [a] A hold-out sample was employed

Table II. Summary of modelling issues of material error applications and the fourth year of data was used as the forecast period to evaluate the performance of the ANN. They selected 15 income statement and balance-sheet accounts or aggregates to represent the major balance-sheet categories. The inclusion of all accounts values was not feasible due to the impact of the number of neurons on the time it takes to train an ANN. Coakley and Brown compared a presumed lack of actual errors and seeded material errors to evaluate the ANN's performance. The results of the study were divided into findings based on: financial ratios, comparison of methods (financial ratio, regression, ANN), effect of error size, effect of statistical level of confidence, effect of source of material error and applying methods to base period. The results were compared with the results achieved with financial ratio and regression methods, and the ANN demonstrated better predictive ability with less overall variation in the predicted values. However, the researchers argue that the fluctuating nature of the financial data within their studies limited the effectiveness of all the AR procedures. The fact is that bigger unexpected fluctuations present in financial data sets cannot be effectively analysed by any known forecasting method.

Coakley and Brown (1991b) also tested ANN technology for recognising patterns in financial ratios of a medium-sized manufacturing firm. In this study they also predicted future values with an ANN. The financial accounts were selected so that they provided information about a company's solvency and the movement of accounts receivable and inventory. The ANN was trained using 36 months of data with an auto-association process, which means that the input pattern and the desired output pattern were the monthly account balances. Thus each pattern was associated with itself. Coakley and Brown evaluated the effectiveness of the model by seeding errors in the data. Their preliminary results indicated that the use of ANNs for pattern recognition across related financial data sets might be viable. The weakness in the Coakley and Brown models is that, although the results are carefully reported from the auditing point of view, the results and the power of the ANN are not so clearly seen.

Wu (1994) applied the ANN system to classify tax cases to ascertain whether further audit was required or not. The 180 sample companies were carefully gathered from an expert tax auditors' audit case file, 90 of which required a further audit and 90 of which required no further audit. The input data (16 attributes) consisted of both reported values and ratios. Half of the cases were used as the training set and the other half were held out as the testing set to validate the implemented network. The cases consisted of information about a firm's business income tax behaviour. The classification accuracy for the neural network was 94 per cent with a two-layer neural network and 95 per cent with a three-layer neural network. The weakness of this model is that the results are not compared with those of any other methods.

Busta and Weinberg (1998) used an ANN to distinguish between "normal" and "manipulated" financial data. They generated 800 data sets containing 200 two-digit numbers by simulating these data sets with the help of Benford's law and non-Benford's law[1]. The ANN analysed the input variables and generated an estimation of the degree on contamination in the data sets. After that they tested six ANN designs to determine the most effective model. In each design, the inputs to the ANN were the different subsets of the 34 variables. The results showed that, if data have been contaminated at a 10 per cent level or more, the ANN will detect this 68 per cent of the time. If the data are not contaminated, the test indicated that the data are "clean" 67 per cent of the time. A key limitation of their study is that it uses simulated data.

Koskivaara (2000b) illustrated a business line ANN-model to compare information of a firm with similar information from the industry in which the organisation operates. Furthermore, the four different alternative models investigated the effect of the year and company on the ANN's performance. The ANN model used in this study was built by using the financial statements of 31 manufacturing companies over four years. Five of the companies were selected to the test set. One bank provided the sample data for the research. The values of the accounts were regarded as a time-series. Koskivaara selected 16 income statement accounts to represent the major financial statement categories; the average number of staff was also included in the model. The account values included financial assets and short-term liabilities, which were taken into the model in order to calculate quick ratio. The data were pre-processed linearly in four different ways: all together, on a yearly basis, on a company basis, and on a yearly and company basis. The results differed depending on what pre-processing method was used. The best results were achieved when all the data were scaled either all together or on a yearly basis. The weakness of these models is that the results are not compared with those of any other methods.

Koskivaara *et al.* (1996) modelled intelligent systems based on ANNs for auditing with the subset of Coakley and Brown (1991b) accounts. They used 66 actual monthly income statements from a manufacturing company. The model was trained using the first 54 months' financial data. The rest of the data were held out for testing. They introduced a one-step-ahead prediction model to observe the non-linear dynamics and the relationships between accounts based on monthly income statements and in that way monitored whether there were unusual fluctuations. Koskivaara (2000a) illustrated how an auditor can use an ANN model to support the planning of auditing monthly balances by a graph on the computer screen that signals either that "no further audit is required" or that "further audit is required". The ANN models used 72 monthly balances of a manufacturing firm. The last 12 months' data were retained for testing purposes. The accounts were chosen with the help of a CPA-auditor in the way that they presented the major and the most interesting monthly balance categories. This study has similarities to Wu's (1994) study. Two models were presented in the study. Model 1 operates inside the quartile and Model 2 has previous quartile data as input variables. The latter model gave slightly better results. The weakness in these models is that the results of the ANN are not compared with the results of any other methods.

4.2.2 Management fraud. Auditors cannot assume that the management is honest or dishonest. They should take a hard, cold look at the possibility of management misrepresentation at the start of the audit and re-examine the likelihood of management misrepresentations as the audit progresses. Management fraud (MF) can be defined as deliberate fraud committed by the management that injures investors and creditors through materially misleading financial statements. Table III summarises the modelling issues of ANN literature pertaining to management fraud problems.

Green and Choi (1997) developed an ANN fraud classification model employing financial data. They used five ratios and three accounts as input variables. The selection of variables was determined by both practical and empirical research. The fraud sample consisted of Securities and Exchange Commission (SEC)-filed financial statements of various companies that had been subsequently found to contain fraudulent account balances. The financial statements of the non-fraud sample received unqualified auditor opinions for the year of selection. They were selected directly from COMPUSTAT and matched the fraud sample on the basis of year, size, and industry (four digit SIC). The training samples consisted of 44-49 companies and the hold-out samples consisted of 42-46 companies respectively. The results showed that ANNs have significant potential as a fraud investigative and detection tool. All sums of the Type I and Type II error rates are significantly less than the random chance benchmark of 1.00[2]. Another contribution of their ANN models is the consistently low Type II error. As the sample is limited to SEC's data file, generalising may not be possible.

Fanning and Cogger (1998) used an ANN (AutoNet) to develop a model for detecting management fraud. They compared the results of an ANN with linear and quadratic discriminant analysis as well as logistic regression. The sample consisted of 150 firms in the training sample and 54 firms in the hold-out sample. The variables were selected by AutoNet and they were the outsider director, having a non-"Big Six" auditor, the geometric growth rate, accounts receivable to sales, net plan property and equipment to total assets, debt to equity and the trend variables for accounts receivable and gross margin. The prediction accuracy of the ANN for the training sample was 75 per cent and 63 per cent for the hold-out sample. The result of their models suggested that there is potential in detecting fraudulent financial statements through analysis of public documents. Fanning and Cogger also showed that ANNs offer better ability than standard statistical methods in detecting fraud. A key limitation of this study was that it was carried out with a sample that covered the period 1972 to 1984. The business environment has dramatically changed since then.

Researcher country	Industry data type	Train/test set	No. inputs	No. hidden layers/nodes	No. outputs	Transfer function	Learning algorithm	Performance measure
Green and Choi (1997) USA	Sample of firms' financial statements' ratios and values	45-49/42-46[a]	8	1:4	1	Sigmoid logistic	BP	Error rate per cent
Fanning and Cogger (1998) USA	Sample of firms' financial statements' accounts, ratios	150/54[a]	8	NA	1	Simple quadratic, quadratic	"AutoNet"	Prediction accuracy per cent Log./DA
Feroz *et al.* (2000) USA	Sample of firms' 7 SAS No. 53 red flag data	28/14[a]	7	1:14	1	Binary sigmoid	BP	OER (overall error rate) MSE

Note: [a] A hold-out sample was employed

Table III.
Summary of modelling issues of management fraud applications

Feroz *et al.* (2000) illustrated the application of the ANNs in order to test the ability of selected statements of auditing standards (SAS) No. 53 red flags to predict the targets of SEC investigations. They used both financial ratios and non-financial turnover red flags mentioned in SAS No. 53. They tested the ANN model with the sample of 42 firms from various industries. A total of 28 firms were used for training the ANN and 14 firms were used for testing. The ANN models classified the membership in target (investigated) firms versus control (non-investigated) firms with an accuracy of 81 per cent. The testing was biased because they only used those red flags that can be constructed from publicly available information. However, authors believed that the sampling choice was sound, given the data constraints in that particular case.

4.2.3 Going concern and financial distress. SAS 59 requires the auditor to evaluate whether there is a substantial doubt about a client's ability to continue as a GC for at least one year beyond the balance-sheet data (Arens *et al.*, 2003). Going concern (GC) and financial distress (FD) research as an application area of ANNs has been minimal, although it resembles bankruptcy studies, which is one of the most popular ANN research areas in business. An auditor gives a GC an uncertainty opinion, when the client company is at risk of failure or exhibits other signs of distress that threaten its ability to continue as a GC. In fact, there are different degrees of financial distress and the auditors do have a choice between two types of going concern report: modified audit report and disclaimer audit report. The latter is a financial distress situation (FD). Bankruptcy is a situation where there is no ability to continue business and the auditor gives an unqualified audit report (see, for example, Tam and Kiang (1992)). The decision to issue a GC opinion is an unstructured task that requires the use of the auditor's judgement. Table IV summarises the modelling issues of ANN literature pertaining to going concern and financial distress problems.

Hansen *et al.* (1992) samples consisted of 80 FD companies (40 that received the GC audit report, 40 that did not receive the GC audit report) and of 98 firms involved in litigation. The source of the sample was the Disclosure II Database that reported on all publicly traded companies whose fiscal year ended between 31 March 1981 and 28 February 1982. This database contains the financial statements of all firms on the New York Stock Exchange (NYSE) and the American Stock Exchange (AMEX). The hold-out sample consisted of 50 per cent of the total sample. Hansen *et al.* (1992) had two models with different variable settings. The audit opinion model had either 12 ratios from financial statements or other closing of the books information as variables. The litigation model had nine variables, which were client-, auditor- or engagement-specific. The average error of the audit opinion model was 8.43 per cent and 20.11 per cent for the litigation model. The results indicated that, in the case of predicting audit opinions, the qualitative-response models perform at a competitive level with the machine-learning models. Theoretical results inferred that this might be especially true when the training sets were relatively small. The authors

Table IV. Summary of modelling issues of going concern and financial distress applications

Researcher country	Industry data type	Train/test set	No. inputs	No. hidden layers/nodes	No. outputs	Transfer function	Learning algorithm	Performance measure
Hansen *et al.* (1992) USA	Sample of manufacturers' ratios, non-financial variables	40/40[a]	12 9	NA	1	Hybrid of steepest gradient, Newton-Raphson	BP	MSE Average error Statistical models
Fanning and Cogger (1994) USA	Sample of firms' Liquidity, cash-flow ratios	75/115[a]	3	2:6-7	1	Quadratic sigmoid logistic	GANNA BP	Per cent accuracy
Lenard *et al.* (1995) USA	Sample of firms' financial statements' ratios and values	70/10[a]	8(4)	1:5 (1:3)	1	General. Reducent gradient optimizer	BP	Per cent accuracy logit model
Anandarajan and Anandarajan (1999) USA	Sample of firms' financial statements' ratios	37/24[a]	14	NA	3	Sigmoid	BP	Accuracy per cent
Koh and Tan (1999) USA	Sample of firms' financial ratios	300/30[a]	6	1:13	1	Sigmoid	BP	Accuracy per cent
Etheridge *et al.* (2000) USA	Sample of banks' financial ratios	863-892/215[a]	57	NA	1	NA	BP, CLN, PLN	OER Pearson's coefficient

Note: [a] A hold-out sample was employed stated that qualitative response models might be a desirable alternative when the training samples are relatively small and there is a need to incorporate additional parameters such as prior probabilities and error costs. Still, the small training set and old data are the weaknesses of this study.

Fanning and Cogger (1994) examined the efficiency of a generalised adaptive neural network algorithm (GANNA) processor in comparison with earlier model-based methods:

- a backpropagation ANN; and
- logistic regression approaches to data classification.

The research used the binary classification problem of discriminating between failing and non-failing firms to compare the methods. The sample consisted of 190 matched pairs, of which the first 75 were selected for the training sample in chronological order and the last 115 matched pairs defined the hold-out sample. All the ANN models had three inputs: the mean adjusted cash flow divided by its standard deviation, the firm's adjusted cash position divided by its standard deviation, and the number of years prior to failure for the failed year. No single approach studied was uniformly superior across all comparison and statistics. The reason for this might be the limited number of input variables. However, the results indicated the potential in time-savings and the successful classification results available from GANNA and ANN processors. Another limitation to this study is that it was carried out with fairly old data (1942 to 1965).

Lenard *et al.* (1995) studied the generalised reduced gradient (GRG2) optimiser for ANN learning, a backpropagation ANN, and a logit model to predict which firms would receive audit reports reflecting a GC uncertainty modification. GRG2 is a solver used in lieu of backpropagation's gradient decent algorithm for training feedforward supervised learning neural networks. The sample for the study was drawn from the 1988 Disclosure II Database. The training sample consisted of 70 firms and a hold-out sample of ten firms. The selection of variables was intended to determine whether the GC decision could be made from publicly available financial statement information. The variables consisted of ratios and account values. The ANN model formulated using GRG2 had the highest prediction accuracy of 95 per cent. It performs best when tested with a small number of variables on a group of data sets, each containing 70 observations. The GRG2-based ANN was proposed as a robust alternative model for auditors to support their assessment of GC uncertainty affecting the client company. Lenard *et al.* (1995) state that a possible limitation to the model is whether all factors have been considered that affect the auditor's going concern uncertainty decision.

Anandarajan and Anandarajan (1999) compared ANN, expert system (ES) and multiple discriminant analysis models to facilitate the decision on the type of GC report that should be issued. The experimental sample of the study was drawn from the 1992 Disclosure database. The data consisted of 14 ratios calculated from the financial statements of 61 companies, of which 27 were reserved for a hold-out sample. The validity of the models was tested by comparing their predictive ability of the type of audit report that should be issued to the client. The prediction accuracy of the ANN in the hold-out sample for all reports was 85.8 per cent. This was better than the prediction accuracy of the ES (69.1 per cent) or the multiple discriminant (74.1 per cent) models. The results of the study indicate that the ANN model predicts the type of GC audit report that should be issued to the client best when compared with the ES and the multiple discriminant analysis. Anandarajan and Anandarajan state that many qualitative variables are not incorporated into the models.

Koh and Tan (1999) predicted a firm's GC status from six financial ratios with an ANN model. Their data set contained a sample of firms divided into 165 non-GCs and 165 matched GCs. Of the cases 300 were used for training the ANN and the remaining 30 cases were reserved for testing. On an evenly distributed hold-out sample, the trained network model correctly predicted all 30 test cases. Koh and Tan compared the GC results of the ANN with the probit model and the audit opinion. Their results suggested that the ANN was at least as good as both the auditors and the probit model for predicting the GC status of firms from financial ratios. In this study the ANN was tested only with six ratios.

Etheridge *et al.* (2000) compared the performance of three ANN approaches, backpropagation (BPN), categorical learning network (CLN), and probabilistic neural network (PNN) as classification tools to assist and support the auditor's judgement about a client's continued financial viability in the future (GC status). The data were provided by a "Big Six" CPA firm and consisted of 57 financial ratios for the years 1986-1988 for 1,139 banks in various regions of the USA. The hold-out sample consisted of 215 banks. They had three, two, and one year prior to failure models. When the overall error rate was considered, the probabilistic ANN (estimated error rate one year prior to failure 2.4 per cent) was the most reliable in classification, followed by backpropagation (3.48 per cent) and categorical learning ANN (7.15 per cent). When the estimated relative costs of misclassification were considered, the categorical learning ANN was the least costly, followed by backpropagation and probabilistic ANN. The authors state that the results are only based on a single industry, banking, and this limits the generalizability of the results to other industries.

4.2.4 Control risk assessment and audit fee. An auditor considers a huge amount of data when assessing the risk of the internal control (IC) structure of an entity failing to prevent or detect significant misstatements in financial statements. The relationships between IC variables that must be identified, selected, and analysed often make assessing a control risk a difficult task. Therefore, control risk assessment (CRA) is a systematic process for integrating professional judgements about relevant risk factors, their relative significance and probable adverse conditions and/or events leading to identification of auditable activities (IIA, 1995, SIAS No. 9). Table V summarises the modelling issues of ANN literature pertaining to control risk assessment and audit fee problems.

Davis *et al.* (1997) (see also Davis (1996)) presented a construction of a prototype, which integrated an ES and an ANN. The rules were contained in the ES model basic CRA heuristics, thus allowing for efficient use of well-known control variable relationships. The 64 observations of auditors from Grant Thornton were used to develop and test an ANN model. The ANN training sample and testing sample each contained 32 observations. The ANN provided a way to recognise patterns in the large number of control variable inter-relationships that even experienced auditors could not express as a logical set of specific rules. The ANN was trained using actual case decisions of practising auditors. The input variables were judgement cues/variables from the general environment, computer processing, general computer and accounting controls. The ANN model provided the auditor with information on how close a risk category border was. The testing resulted in an accuracy rate of 78 per cent. The model is limited because of the number of cases and the fact that auditors were from the same firm.

Ramamoorti *et al.* (1999) used both quantitative (26 variables) and qualitative (19 variables) risk factors as input variables in the models. The risk was defined in an internal auditing context. The models were in the context of public state university departments. The sample consisted of 141 university departments and they used a training size, which covered 70 per cent of the data and hold-out data of 30 per cent. The quantitative data were downloaded from the University of Illinois financial and administration system. The qualitative risk factor values were elicited from audit staff using a pre-defined scale from 0 to 9. The eventual number of variables selected to construct the models were in the seven to 18 range. The research project included a Delphi study and a comparison with statistical approaches, and presented preliminary results, which indicated that internal auditors could benefit from using ANN technology for assessing risk. The ANN models captured the top 25 risky departments at an accuracy rate of 72-84 per cent. The results are based on a single industry, university, and this might limit the generalizability of the results to other industries (as in the Etheridge *et al.* (2000) study).

Curry and Peel (1998) provided an overview of the ANN modelling approach and the performance of ANNs, relative to conventional ordinary least squares (OLS) regression analysis, in predicting the cross-sectional variation in corporate audit fees (AF). The data were derived from a sample of 128 unquoted UK companies operating in the electronic industrial sector. The data were collected from years 1986-1988 from *Kompass 1990* and Macmillan's *Unquoted Companies 1990*. The audit fee, the dependent variable in the study, must be disclosed (under UK company law) in a note to a company's annual

Researcher country	Industry data type	Train/test set	No. inputs	No. hidden layers/nodes	No. outputs	Transfer function	Learning algorithm	Performance measure
Davis (1996) USA	Sample of firms' IC risk data, observations of auditors	37/27	107	1:5	1	NA	BP	Per cent absolute error
Davis *et al.* (1997) USA	Sample of firms' IC risk data, observations of auditors	32/32[a]	210	1:30	1	Sigmoid	BP	RMSE Pearson's coeffic., accuracy per cent
Ramamoorti *et al.* (1999) USA	Sample of university departments' qualitative, quantitative risk factor data	100/41[a]	10	NA	1	NA	BP	*R*-squared per cent, Delphi overlap
Curry and Peel (1998) UK	Sample of electronic firms' financial, non-financial data	96/32 86/42 64/64[a]	25	1:2 1:3 1:4	1	Sigmoid	BP	MSE

Note: [a] A hold-out sample was employed

Table V.
Summary of modelling issues of internal control risk assessment and audit fee applications

statements. The input variables were related to auditee size, audit complexity, audit risk, auditee profitability, and auditor size. They tested the ANN's ability to generalise with three separate training and hold-out samples: 96 and 32, 86 and 42, and 64 and 64 companies in the samples. The estimation accuracy for the hold-outs was 56.2 per cent, 68.6 per cent, and 64.5 per cent respectively. These results were achieved with four hidden neurons. The ANN models also exhibited better forecasting accuracy than their OLS counterparts (31.9 per cent, 59.7 per cent, and 52.9 per cent respectively). Curry and Peel state that, although the study demonstrated that the ANN outperformed conventional linear techniques in forecasting audit fee data, there is a need for further research relating to the optimal architecture of ANNs, namely the number of hidden layers and neurons.

4.3 Issues in ANN modelling for analytical review

Auditing ANN research started at the beginning of the 1990s. The applications have mainly been developed in the USA (16 studies). Some development work has been done also in Finland (three studies), Taiwan (one study), and the UK (one study). Many of the cited articles applied the ANNs as an extension of a previous statistical model-based study, and used publicly available stock market data. Table VI summarises the findings of ANN studies in chronological order by application areas. All the authors state that ANNs have potential to improve AR procedures.

4.3.1 Data. Both quantitative and qualitative data were used as input variables in the applications. Most of the data are quantitative and gathered from public databases. Financial statement values and ratios and monthly account values were mostly used as quantitative input variables. Opinions and observations of auditors or red flag data defined by SAS in the USA were included in quantitative input variables. In ten cases the sample data were gathered from various industries. Seven cases focused on one particular industry, for example, the manufacturing, bank, university or electronic industry. Focusing on one particular industry is one way of conducting analytical procedures.

Many of the studies applied the ANNs as an extension of a previous statistical study. Therefore, in some cases the data were fairly old. Also to use only publicly available information might simplify the auditor's task too much and does not reveal any significant problems in the client company.

4.3.2 Samples. All but two applications reviewed in this survey had relatively small data samples with which to train and test the models. ANNs can be more appropriate for large data sets. The material error ANN-model of Busta and Weinberg (1998) had bigger data sets than other studies in the review, but their data set was simulated. The GC application of Etheridge *et al.* (2000) had a data set of financial ratios from 1,139 banks, which was provided by a "Big Six" CPA firm.

Authors	Summary of findings
Material errors	
Coakley and Brown (1991a)	Results tentatively suggest that the ANN recognised patterns across financial ratios more effectively than financial ratio and regression methods
Coakley and Brown (1991b)	The use of ANNs for pattern recognition across related financial data sets may be viable
Coakley and Brown (1993)	ANNs applied as a forecasting tool seem useful for identifying patterns that can indicate potential investigations of a firm's unaudited financial data in the current year
Wu (1994)	The results strongly suggest that ANNs can be used to identify firms requiring further auditing investigation, and also suggest future implications for intelligent auditing machines
Coakley (1995)	Results suggest that the use of an ANN to analyse patterns of related fluctuations across numerous financial ratios provides a more reliable indication of the presence of material errors than either traditional analytic procedures or pattern analysis, as well as providing insight into the plausible causes of the error
Koskivaara *et al.* (1996)	ANNs seem a promising technology for predicting monthly financial statement values
Busta amd Weinberg (1998)	The results show that, if data are contaminated at a 10 per cent level or more ANNs will detect this 68 per cent of the time. If the data are not contaminated, the test will indicate that the data are "clean" 67 per cent of the time
Koskivaara (2000a)	The best result was achieved when all the data were scaled either linearly or linearly on a yearly basis
Koskivaara (2000b)	The results achieved indicate that ANNs seem promising for recognising the dynamics and the relationships between financial accounts
Management fraud	
Green and Choi (1997)	The results show that ANNs have significant potential as a fraud investigative and detection tool
Fanning and Cogger (1998)	The study showed that ANNs offer better ability than standard statistical methods in detecting fraud
Feroz *et al.* (2000)	The ANN models classify the membership in target versus control firms with the accuracy of 81 per cent across the board
Going concern and financial distress	
Hansen *et al.* (1992)	The results indicate that the ANN model predicted more consistency than the other advanced statistical models used in the study
Fanning and Cogger (1994)	The results indicate the potential in time-savings and the successful classification results available from an ANN processor
Lenard *et al.* (1995)	The ANN model is proposed as a robust alternative model for auditors to support their assessments of going concern uncertainty

(*continued*)

Table VI.
Findings of analytical review ANN studies

Authors	Summary of findings
Anandarajan and Anandarajan (1999)	The results of the study indicate that the ANN model has a superior predictive ability in determining the type of going concern audit report that should be issued to the client
Koh and Tan (1999)	The results suggest that ANNs can be a promising avenue of research and application in the going concern area
Etheridge *et al.* (2000)	When the overall error rate was considered, the probabilistic ANN was the most reliable in classification, followed by backpropagation and categorical learning ANNs. When the estimated relative costs of misclassification were considered, the categorical learning ANN was the least costly, followed by backpropagation and probabilistic ANN
Control risk assessment	
Davis (1996)	The results show that the ANN models have been able to acquire, represent and store the auditors' knowledge for selecting relevant information to assess control risk
Davis *et al.* (1997)	The model was able to classify the auditors' control risk assessment in the testing sample with a 78 per cent accuracy rate
Ramamoorti *et al.* (1999)	The ANN, logistic regression and stepwise multiple regression produced results that compared favourably together in terms of the variance in the rankings obtained
Audit fee	
Curry and Peel (1998)	The ANN models exhibited better forecasting accuracy than their OLS counterparts, but this differential reduces when the models are tested out-of-sample

Table VI.

It seems that it is common to use small data sets in the early developing phase of ANN models. Small data sets also limit the number of possible inputs into the models. However, it is critical to have both the training and test set representative of the population or underlying mechanism. The selection of the training and test set may affect the performance of an ANN. The literature offers little guidance in selecting the training and the test sample. Most authors select them based on the rule of 90 per cent vs 10 per cent, 80 per cent vs 20 per cent, or 70 per cent vs 30 per cent, etc.; 50 per cent vs 50 per cent was used in four cases. Some choose them based on matched pairs. The median training samples and testing sample sizes are 64 and 36, respectively.

4.3.3 Architecture. The determining of the numbers of input nodes, hidden layers and nodes and output nodes is problem-dependent and is still part of the art in ANNs. For example, all the income statement account values could be selected as input variables in many applications. However, some selection of variables might have helped an ANN to get better results. For example, the selection of input variables for the prediction task was a difficult task in many studies. The hidden nodes in a hidden layer allow the ANN to detect the feature, to capture the pattern in the data, and to perform complicated non-linear mapping between input and output variables. Linear relationships between the variables are very often a simplification of the natural financial data. The number of output nodes is relatively easy to specify as it is directly related to the problem under study. The median audit ANN architecture consists of 24 inputs with 14 nodes in one hidden layer and one output node. The majority of the auditing ANN-applications under study had one output node. This is common in classification models, whereas prediction models might have one or more nodes in the output layer. The choice network architecture is mostly subject to trial and error without guarantee of reaching an optimal solution. However, properly configured and trained, ANNs appear to make consistently good classifications or generalisations.

4.3.4 Learning parameters. There are no clear guidelines on the selection of the performance and transfer function. All but one study used the supervised learning mode. Fanning and Cogger (1998) used the self-organising learning mode in a management fraud application. Most studies used the backpropagation learning algorithm or some variation of it. A sigmoid or a modified sigmoid transfer function was used in nine studies, a logistic function in three studies. Seven applications used some other transfer function.

4.3.5 Performance. Almost every model was evaluated with a hold-out sample. The performance results of the ANN models in the previous review are based on hold-out samples.

The performance measurement of the models varied a lot. Root mean square error, prediction accuracy, and average error were mostly used. In some cases also a comparison with the traditional statistical methods was used. It is not always clear whether the performance is calculated from actual data or scaled data.

Researchers should emphasise the measurement of the performance from the auditing point of view. Visualisation of the results might help auditors to see the benefit of the ANNs.

4.4 Limitations of ANNs

An inherent weakness of ANNs is that the internal structure makes it difficult to trace the process by which output is reached. This is why ANNs lack clear explanatory capabilities. Therefore, justifications for the results are difficult to obtain because the connections' weights do not usually have obvious interpretations. However, if the results are obvious from the auditing point of view, this "black box" problem is not a problem for auditors.

5. Discussion and conclusions

An accounting system's capacity to produce data and information in terms of volume and speed is vast. An accounting system means the series of tasks and records of an entity by which transactions are processed as a means of maintaining financial records. Such systems identify, assemble, analyse, calculate, classify, record, summarise and report transactions and other events. Accounting systems are also vulnerable to manipulation and fraud.

The auditor's work in each case is to confirm the reliability of the accounts. The auditor should ascertain the entity's system of recording and processing transactions and assess its adequacy as a basis for the preparation of financial statements. Therefore, the auditor should carry out such a review of the financial accounts which is as sufficient, in conjunction with the conclusions drawn from the other audit evidence obtained, as possible, to give her/him a reasonable basis for her/his opinion on the financial statements. This review may include generating trends in accounting values or comparing accounting data items. This is called AR and it is supported by analytical procedures, in which different kinds of techniques might be embedded.

The research and development of audit tools are important, since the task of the auditor today is both more onerous and more complex than ever before. Today's business environment is very commercial and time pressure is hard. Also, tight relationships between auditing firms and their clients do not make the auditing tasks any easier. The unveiling of the manipulation of accounting data could have serious consequences both for the companies and their stock value and for auditors and their reputations. The development of information technology support systems makes the use of advanced methods easier and more cost-effective.

As information technological changes occur at an increasing rate, auditors must keep pace with these emerging changes and their impact on their client's information-processing systems as well as on their own audit procedures. This article started by providing an overview of the AR in auditing and a brief description of ANNs. ANNs are an appealing technology embedded in the analytical procedures for forming expectations for auditing purposes. They are good at handling data and, once trained, they can predict and classify new examples very quickly. For example, with the BP algorithm the auditor may generate evidence based on internal trends in accounting data and then compare ANN results with actual values. With the SOM algorithm the auditor may cluster the data and reveal unexpected fluctuations in the data. In brief, auditors may benefit from the ANN's learning from data ability to support auditors' experience and knowledge about a client company. Therefore, we argue that auditors may benefit from applying ANNs in revealing trends and classes in accounting data or in the comparison of accounting records.

In this article we analysed a number of ANN studies carried out in the area. Indeed, the analysis revealed the open ends in the field. The main findings are summarised as follows.

All the studies fit into analytical review (AR) procedures. All the authors stated that ANNs have potential to improve AR procedures. ANNs outperformed the statistical methods in six cases. In many cases the results really do serve the AR procedures, such as comparison of client and the industry data or forming expectations to compare client's data.

The application areas were material errors, management fraud, and support for going concern decision, internal control, risk assessment, audit fee, and financial distress problems. Since the ANNs were built for specific situations, they have a limited view of AR procedures. Therefore, a question arises. Are there any new application areas, such as authority checking and analysing minutes that would be suitable for ANNs? Budgets, forecasts, online analytical procedures, and data mining are areas in organisations where ANNs embedded in the systems may provide auditors with such business-intelligent information they need to make decisions.

Data samples used in the studies might limit the generalisability of the results. Most of the reviewed applications used quite small data sets. In some case the data were also quite old. Researchers should focus on developing models with data sets that are not publicly available in order to investigate client data more deeply. For example, researchers should study ANNs with the operational data. Many of the analytical procedures are based on the client company's data and therefore this is an important research area. Indeed, many important qualitative variables such as management ability and future plans were not formally incorporated into the models.

The determining of the ANN architecture is problem-dependent and is still part of the art in ANNs. Further research is needed to find out optimal ANN architectures.

Most studies used the backpropagation algorithm, while others employed some variants of it. There is clearly a need for methods and networks that can handle the fluctuating financial data. For example, the supervised training method with the resilient backpropagation (Rprop) training algorithm with sigmoid function, which is one of the most efficient algorithms on pattern recognition problems (Demuth and Beale, 2000), may be possible, or the SOM algorithm for classifying accounting data sets.

There are no clear guidelines on the selection of the performance measure. The performance measure of applications varies a lot and it is not always clear whether the performance is calculated from actual data or scaled data. More emphasis should be put on the measurement of the performance of ANNs. Although the results are encouraging, further research is needed to examine whether ANN models are better than AR procedures in use. Furthermore, the visualisation of the results might give some added value to auditors.

One advantage of an ANN-based system is that it can provide auditors with objective information about a client company. With the help of an ANN-based system an auditor may, for example, focus substantive testing on the right places. The ANN-based system could work on the Web or on the user workstation using the data in the accounting databases. It could automatically alert the user when certain amounts of the data are not inside the agreed limits.

Therefore, the ANN is a possible future technique for analytical auditing, and suitable for continuous auditing and controlling.

Although ANN-based systems cannot entirely replace professional judgement, they offer a promising alternative approach to AR procedures. It should also be emphasised that the number of variables, which can be inputs and outputs of the models, is limited. However, ANNs provide a non-linear dimension that captures the underlying representations within the data sets. Therefore, the future of ANNs in auditing is open and will be even brighter as more and more research efforts are devoted to this area.

So far, only a few attempts have been made to combine ANNs with other advanced methods. Davis *et al.* (1997) presented a construction of a prototype, which integrated an ES and an ANN in the control risk assessment domain. Additional research is clearly needed to evaluate the efficacy of hybrid models that combine the advantages of ANNs, fuzzy logic- and genetic algorithms-based systems.

Furthermore, the future of ANNs in the auditing area may see increased integration with other existing or developing technologies and statistical techniques. As advancements are made in AI technologies and computer-based systems, there should be new opportunities to apply ANNs to auditing. This would encourage or motivate academics and practitioners to collaborate in further exploration of the potential of ANNs.

Notes

1. Benford's law says that the digits of naturally occurring numbers are distributed on a predictable and specific pattern.
2. Type I: There was fraud but ANN did not find it (= inefficient auditing). Type II: There was no fraud but ANN detected one (= ineffective auditing).

References

AICPA (1988), *Statement on Auditing Standards #56, Analytical Procedures*, American Institute of Certified Public Accountants, New York, NY.

Anandarajan, M. and Anandarajan, A. (1999), "A comparison of machine learning techniques with a qualitative response model for auditors' going concern reporting", *Expert Systems with Applications*, Vol. 16, pp. 385-92.

Arens, A.A., Elder, R.J. and Beasley, M.S. (2003), *Auditing and Assurance Services: An Integrated Approach*, Prentice-Hall, Upper Saddle River, NJ.

Bazerman, M.H., Loewenstein, G. and Moore, D.A. (2002), "Why good accountants do bad audits", *Harvard Business Review*, Vol. 80 No. 11, pp. 97-102.

Blocher, E. and Patterson, G.F. Jr (1996), "The use of analytical procedures: the importance of expectation and precision", *Journal of Accountancy*, February, p. 53.

Busta, B. and Weinberg, R. (1998), "Using Benford's law and neural networks as a review procedure", *Managerial Auditing Journal*, Vol. 13 No. 6, pp. 356-66.

Calderon, T.G. and Cheh, J.J. (2002), "A road-map for future neural networks research in auditing and risk assessment", *International Journal of Accounting Information Systems*, Vol. 3 No. 4, pp. 203-36.

Coakley, J.R. (1995), "Using pattern analysis methods to supplement attention-directing analytical procedures", *Expert Systems with Applications*, Vol. 9 No. 4, pp. 513-28.

Coakley, J.R. and Brown, C.E. (1991a), "Neural networks applied to ratio analysis in the analytical review process", paper presented at Expert Systems Symposium, Pasadena, CA.

Coakley, J.R. and Brown, C.E. (1991b), "Neural networks for financial ratio analysis", paper presented at The World Congress on Expert Systems, Orlando, FL.

Coakley, J.R. and Brown, C.E. (1993), "Artificial neural networks applied to ratio analysis in the analytical review process", *Intelligent Systems in Accounting, Finance and Management*, Vol. 2, pp. 19-39.

Coakley, J.R. and Brown, C.E. (2000), "Artificial neural networks in accounting and finance: modeling issues", *International Journal of Intelligent Systems in Accounting, Finance and Management*, Vol. 9, pp. 119-44.

Curry, B. and Peel, M.J. (1998), "Neural networks and business forecasting: an application to cross-sectional audit fee data", *International Journal of Commerce and Management*, Vol. 8 No. 2, pp. 94-120.

Davis, J.T. (1996), "Experience and auditors' selection of relevant information for preliminary control risk assessments", *Auditing: A Journal of Practice & Theory*, Vol. 15 No. 1, pp. 16-37.

Davis, J.T., Massey, A.P. and Lovell, R.E.R. II (1997), "Supporting a complex audit judgement task: an expert network approach", *European Journal of Operational Research*, Vol. 103 No. 2, pp. 350-72.

Demuth, H. and Beale, M. (2000), *Neural Network Toolbox*, The Math Works, Inc., Natick, MA.

Dutta, S. (1993), *Knowledge Processing & Applied Artificial Intelligence*, Butterworth-Heinemann, Oxford.

Etheridge, H.L., Sriram, R.S. and Hsu, H.Y.K. (2000), "A comparison of selected artificial neural networks that help auditors evaluate client financial viability", *Decision Science*, Vol. 31 No. 2, pp. 531-50.

Fanning, K.M. and Cogger, K.O. (1994), "A comparative analysis of artificial neural networks using financial distress prediction", *Intelligent Systems in Accounting, Finance and Management*, Vol. 3, pp. 241-52.

Fanning, K.M. and Cogger, K.O. (1998), "Neural network detection of management fraud using published financial data", *International Journal of Intelligent Systems in Accounting, Finance & Management*, Vol. 7 No. 1, pp. 21-41.

Feroz, E.H., Kwon, T.M., Pastena, V.S. and Park, K. (2000), "The efficacy of red flags in predicting the SEC's targets: an artificial neural networks approach", *International Journal of Intelligent Systems in Accounting, Finance & Management*, Vol. 9, pp. 145-57.

Fraser, I.A.M., Hatherly, D.J. and Lin, K.Z. (1997), "An empirical investigation of the use of analytical review by external auditors", *The British Accounting Review*, Vol. 29, pp. 35-47.

Freeman, J.A. and Skapura, D.M. (1991), *Neural Networks Algorithms, Applications, and Programming Techniques*, Addison-Wesley Publishing Company, Menlo Park, CA.

Green, B.P. and Choi, J.H. (1997), "Assessing the risk of management fraud through neural network technology", *Auditing: A Journal of Practice & Theory*, Vol. 16 No. 1, pp. 14-28.

Halal, W.E., Kull, M.D. and Leffmann, A. (1998), "The George Washington university forecast of emerging technologies: a continuous assessment of the technology revolution", *Technological Forecasting and Social Change*, Vol. 59, pp. 89-110.

Hansen, J.V., McDonald, J.B. and Stice, J.D. (1992), "Artificial intelligence and generalized qualitative-response models: an empirical test on two audit decision-making domains", *Decision Science*, Vol. 23 No. 3, pp. 708-23.

Haykin, S. (1994), *Neural Networks: A Comprehensive Foundation*, Macmillan, New York, NY.

Hecht-Nielsen, R. (1990), *Neurocomputing*, Addison-Wesley Publishing Company, Inc., San Diego, CA.

Hertz, J., Krogh, A. and Palmer, R.G. (1991), *Introduction to the Theory of Neurocomputing*, Addison-Wesley Publishing Company, Redwood City, CA.

Klimasauskas, C.C. (1991), "Applying neural networks. Part 1: an overview of the aeries. Part 2: a walk through the application process. Part 3: training a neural network. Part 4: improving performance. Part 5: integrating a trained network into an application. Part 6: special topics", *PC/AI Magazine*, No. 5.

Koh, H.C. and Tan, S.S. (1999), "A neural network approach to the prediction of going concern status", *Accounting and Business Research*, Vol. 29 No. 3, pp. 211-16.

Kohonen, T. (1997), *Self-Organizing Maps*, Springler-Verlag, Berlin.

Koskivaara, E. (2000a), "Artificial neural network models for predicting patterns in auditing monthly balances", *Journal of the Operational Research Society*, Vol. 51 No. 9, pp. 1060-9.

Koskivaara, E. (2000b), "Different pre-processing models for financial accounts when using neural networks for auditing", paper presented at European Conference on Information Systems, Vienna.

Koskivaara, E., Back, B. and Sere, K. (1996), "Modelling intelligent information systems for auditing", paper presented at the Intelligent Systems in Accounting and Finance Conference, Huelva.

Kosko, B. (1990), "Unsupervised learning in noise", *IEEE Transactions on Neural Networks*, Vol. 1 No. 1, pp. 44-57.

Lee, B. (2002), "Professional socialisation, commercial pressures and junior staff's time-pressured irregular auditing – a contextual interpretation", *British Accounting Review*, Vol. 34 No. 4, pp. 315-33.

Lehtokangas, M., Saarinen, J. and Kaski, K. (1994), "Aikasarja ja ennusteet", in Koikkalainen, P. (Ed.), *Neurolaskennan mahdollisuudet*, TEKES, Paino-Center Oy, Helsinki.

Lenard, M.J., Alam, P. and Madey, G.R. (1995), "The application of neural networks and a qualitative response model to the auditor's going concern uncertainty decision", *Decision Science*, Vol. 26 No. 2, pp. 209-27.

Medsker, L. and Liebowitz, J. (1994), *Design and Development of Expert Systems and Neural Networks*, Macmillan College Publishing Company, Inc., New York, NY.

O'Leary, D.E. (1998), "Using neural networks to predict corporate failure", *International Journal of Intelligent Systems in Accounting, Finance & Management*, Vol. 7, pp. 187-97.

Ramamoorti, S., Bailey, A.D. Jr and Traver, R.O. (1999), "Risk assessment in internal auditing: a neural network approach", *International Journal of Intelligent Systems in Accounting, Finance & Management*, Vol. 8 No. 3, pp. 159-80.

Smith, M. (1996), *Neural Networks for Statistical Modeling*, International Thompson Computer Press, London.

Sohl, J.E. and Venkatachalam, A.R. (1995), "A neural network approach to forecasting model selection", *Information & Management*, Vol. 29, pp. 297-303.

Swingler, K. (1996), *Applying Neural Networks: A Practical Guide*, Academic Press Limited, London.

Tam, K.Y. and Kiang, M.Y. (1992), "Managerial applications of neural networks: the case of bank", *Management Science*, Vol. 38 No. 7, pp. 926-47.

Vellido, A., Lisboa, P.J.G. and Vaughan, J. (1999), "Neural networks in business: a survey of applications (1992-1998)", *Expert Systems with Applications*, Vol. 51, pp. 51-70.

Weber, J.C. and Goldstein, E.J. (1999), *Public Accounting – Practice Alert: CPA Letter Supplements*, Vol. 98 No. 1.

Willet, C. and Page, M. (1996), "A survey of time budgets and irregular auditing practices among newly qualified UK chartered accountants", *The British Accounting Review*, Vol. 28, pp. 101-20.

Wong, B.K. and Selvi, Y. (1998), "Neural network applications in finance: a review and analysis of literature (1990-1996)", *Information & Management*, Vol. 34, pp. 129-39.

Wong, B.K., Bodnovich, T.A. and Selvi, Y. (1995), "A bibliography of neural network business applications research: 1988-September 1994", *Expert Systems*, Vol. 12 No. 3, pp. 253-62.

Woolf, E. (1994), *Auditing Today*, Prentice-Hall, Hemel Hempstead.

Wu, R.C-F. (1994), "Integrating neurocomputing and auditing expertise", *Managerial Auditing Journal*, Vol. 9 No. 3, pp. 20-6.

Zhang, G., Patuwo, B.E. and Hu, M.Y. (1998), "Forecasting with artificial neural networks: the state-of-the-art", *International Journal of Forecasting*, Vol. 14, pp. 35-62.

PUBLICATION 2

Koskivaara (2002) Design Science Approaches in Information Systems Research. In: *Proceedings of the Information Systems Foundations: Building the Theoretical Base Workshop*, ed. by Shirley Gregor – Dennis Hart. Canberra, Australia, October 1–2, 2002, 205–216. Reprinted by permission.

Design Science Approaches in Information Systems Research

Eija Koskivaara

TUCS Turku Centre for Computer Science, and
Turku School of Economics and Business Administration,
Finland
Eija.Koskivaara@tukkk.fi

Abstract

This paper outlines design science approaches in information systems (IS) research. Design science attempts to create effective artifacts that serve human purposes. Therefore, design science activities are an important part of knowledge representation in IS research. The paper brings together different views of design science approaches. Furthermore, it illustrates how these approaches are applied to designing artificial neural networks (ANNs).

Introduction

This paper focuses on the different design science approaches in information system (IS) research. Whereas natural sciences and social sciences try to understand reality, design science attempts to create effective artifacts that serve human purposes. If we think that science is an activity that produces guidelines, then design science is an important part of it (Simon, 1969). In this paper we illustrate how different approaches are brought together and apply them to designing and developing artificial neural network (ANN) systems.

We have found four approaches in the literature to design and develop information systems that are possible to follow in the ANN development process. March *et al.* (1995) make a distinction between *design science* and natural science and argue that both design science and natural science activities are needed to ensure that IS research is both relevant and effective. They refer to Herbert Simon's (1969) work "The Sciences of the Artificial", where he uses the expression "science of design" to refer to applied research. A science of design implies that this science is concerned with how to make artifacts that have desired properties and how to design. Thus a design theory has two aspects - one dealing with the process of design and one dealing with the product. Nunamaker *et al.* (1991) have introduced the *systems development* methodology, which fits into the category of applied science and belongs to the engineering type of research. The results of applied sciences should be relevant, simple and easy to use (Kasanen *et al.*, 1993). In the engineering approach, the artistry of design and the spirit of "making something work" are also essential. These refer to pragmatism, which is the philosophical background theory of this kind of research. Pragmatism emerged in the USA by Charles Sanders Pierce (1839-1914), William James (1848-1910), and John Dewey (1859-1952). Philosophical pragmatists deny the correspondence notion of truth, proposing that truth essentially is what works in practice (see e.g. Niiniluoto, 1986). Another research approach, which has similarities to the system development, comes from the field of business administration, where a *constructive research* approach has been developed (Kasanen *et al.*, 1993; Lukka, 2002). Medsker *et al.* (1994) have presented a *development cycle* methodology for designing ANN systems. All these approaches are about one particular type of knowledge in IS, namely "theory of design".

Although the design of an ANN system resembles conventional information system design, it includes some unique steps and considerations. According to our experience the approaches of information systems design and development give too narrow a picture of building and assessing ANN systems. Therefore, one purpose of this paper is to bring together these approaches and thereby give a wider approach for the design of ANN systems.

ANNs are selected for illustration purposes for several reasons. ANNs are nowadays a well-established computational paradigm in the field of artificial intelligence. They are usually seen as a method for implementing complex non-linear mappings (functions) using simple elementary units that are connected together with weighted, adaptable connections. The history of ANNs can be traced far back to the birth of the digital computers (McCulloch *et al.*, 1943; Hebb, 1949). There was some enthusiastic research on ANNs in the fifties and sixties (Rosenblatt, 1958; Minsky *et al.*, 1969), but subsequently ANNs were almost forgotten. The rise that began in the eighties is still going on. One reason for this is that information technology development and the processing capacities of PCs have made it possible to model ANN-based information systems.

ANNs are data driven systems that can be used for prediction, classifying, and clustering tasks. They can learn, remember, and compare complex patterns (Medsker *et al.*, 1994).They are also claimed to be able to recognize patterns in data even when the data is noisy, ambiguous, distorted, or variable (Dutta, 1993). Furthermore, they continue to perform well even with missing or incomplete data, and they are capable of discovering data relationships. One advantage of ANN-systems could be that they provide "intelligent real time" information to add value to the decision process. These features make ANNs suitable for many business decisions. ANNs have already been applied and have proven their usefulness in many different business areas (Wong *et al.*, 1995; Wong *et al.*, 1998; O'Leary, 1998; Zhang *et al.*, 1998; Vellido *et al.*, 1999; Coakley *et al.*, 2000). Budgets, forecasts, on-line analytical procedures, and data mining are areas in organizations where ANNs embedded in the systems may provide users such business intelligent information they need to make tactical and strategic decisions.

The rest of the paper is organized as follows. In section 2 we review the different design science approaches that at least partly suit the ANN development environment. Our aim is to find an effective and efficient way for ANN-based system design and development. In section 3 we combine and compare these approaches. Section 4 brings together some different views of the knowledge of these approaches in order to design ANN systems for organizations. This developed approach is partly based on our experience in designing several ANN models.

Approaches and methodologies in information system research design

Design Science Research Approach

Design science is knowledge using activity, which produces and applies knowledge of tasks or situations in order to create an effective artifact (March *et al.* 1995). Design science research in information technology aims at improving performance. It is a technology-oriented attempt to create things that serve human goals. Its products are assessed against criteria of value or utility - does it work? is it an improvement? In the designing of ANN models this means that we apply the existing technology to various situations to research whether any improvement occurs.

March *et al.* (1995) make the distinction between research outputs and research activities (Figure 1.). The first dimension in the framework is based on design science outputs or artifacts. The second dimension is based on broad types of research activities. Building and evaluating IT artifacts have design science intent. Theorizing and justifying have natural science intent.

Research Activities

		Build	*Evaluate*	Theorize	Justify
Research Outputs	***Constructs***				
	Model				
	Method				
	Instantiation				

Figure 1. A research framework

Design science consists of two basic activities: *build* and *evaluate*. They are aimed at improving performance. Build refers to the construction of the artifact, demonstrating that such an artifact can be constructed. Evaluate refers to the development of criteria and the assessment of artifact performance against those criteria. In the construction of an ANN system this means that we try to find appropriate learning paradigms and parameters, and network architecture to receive reasonable outputs. This research activity should be judged based on the value or utility to a community of users. This means that we have to be able to show the value of the ANN's output information to users. The research contribution lies in the novelty and the effectiveness of the artifact. In analyzing the result of an ANN model we have to be able to show that it either works more efficiently than the old method or it produces some added value to the user. In comparing the efficiency of the methods, ANN vs. other, the criteria has to be the same for both methods and hopefully not invented by the researcher. The development of the criteria might bias the research results.

According to March *et al.* (1995) design science products are of four types, *constructs*, *models*, *methods*, and *instantiations*. Constructs or concepts form the elements of the semantic field of a research domain. They help to specify the research problem and its solutions. Models are propositions and statements, which express the relationships between constructs. Methods are sets of steps used to perform a task. The desire to utilize a certain type of method can influence the constructs and models developed for a task. Instantiations are the actual realizations of the artifact in the environment. They demonstrate the feasibility and effectiveness of the models and methods. With the ANN's development process we may produce all these products depending on the final status of the developed artifact.

Systems Development Research Approach

Nunamaker *et al.* (1991) propose a multimethodological approach to IS research that consists of four research strategies: *theory building, experimentation, observation, and systems development*. Theory building includes development of new ideas and concepts, frameworks, new methods, or models. Experimentation includes research strategies such as laboratory and field experiments, and computer simulations. Observation includes research methodologies such as case studies, field studies, and surveys. According to Nunamaker *et al.* (1991) a system development research consists of five sequential stages: *concept design, constructing the architecture of the system, prototyping, product development,* and *technology transfer*. Concept design is the adaptation and amalgamation of technological and theoretic advances into potentially practical applications. A system architecture provides a road map for the systems building process. Prototyping is used as a proof-of-concept to demonstrate feasibility. Much of the system development research stops at this stage because it fails to meet initial expectations. In the ANN development process this mostly means that we have to redesign the ANN model to get it to work. Those research projects that are judged successful are expanded into fully articulated production systems. This allows a realistic evaluation of the impacts of the included information technologies and their potential for acceptance. The transfer of technology to organizations represents the final success of those theories, concepts, and systems that complete this race. These latter phases of the system development approach are hard to carry out because these applications may contain business secrets and therefore the models are kept secret. Difficulties in the systems development process can be used to modify the concepts and theories from which the application systems are derived.

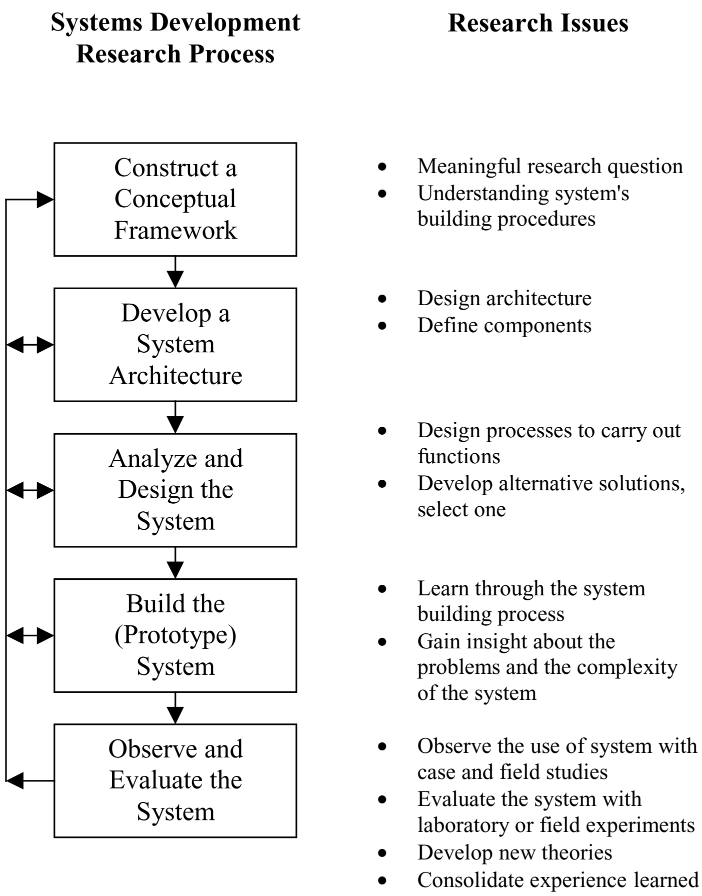


Figure 2. A Process for Systems Development Research

The systems development research process and research issues are outlined in Figure 2. In the development type of research, researchers do not usually formulate an explicit hypothesis, but they do make assumptions about the research domain and the technical environment for developing systems. In many ANN development processes researchers propose a new way of doing things. Design involves the understanding of the domain under study, the application of relevant scientific and technical knowledge, the creation of various alternatives, and the synthesis and evaluation of proposed alternative solutions. In designing a road map for the ANN building process the researcher has to be aware of the learning paradigms, data needed or available, and how to evaluate the ANN's output. Researchers state the system requirements under the constraints of these assumptions, and design and implement the system according to the requirements. Researchers in system development often conduct their research by building the prototype system. The process of implementing a working process can provide researchers with insights into the advantages and disadvantages of the concepts, the frameworks, and the chosen design alternatives. Once the system is built, the researcher can test its performance and usability as stated in the requirements definition phase. Experiences gained from the developing of the system usually lead to the further development of the system, or even to the discovery of a new theory to explain newly observed phenomena.

Constructive Research Approach

The aim of a constructive research approach is to solve a real-world problem with an implemented new innovative construction, which has practical and theoretical contribution. This kind of research result would satisfy all stakeholders of the research project. However, from the academic point of view, practically failing projects could also have theoretical implications. This resembles the systems development approach where the difficulties in the research process can be used to modify the theory. The constructive research approach has been developed in the field of business administration. However, its roots are in technical sciences, education, clinical medicine, information systems and operations research. The constructive type of studies have rarely been published because they have often been considered consulting and, therefore, non-scientific. Another reason for non-publishing could be the business secrets those studies contain. According to Lukka (2002) the constructive research approach is one of the options currently available for a case researcher. It adds an alternative and a strong, problem-solving type of intervention and an intensive attempt to draw theoretical conclusions based on empirical work. Figure 3 illustrates the key elements of the constructive research approach.

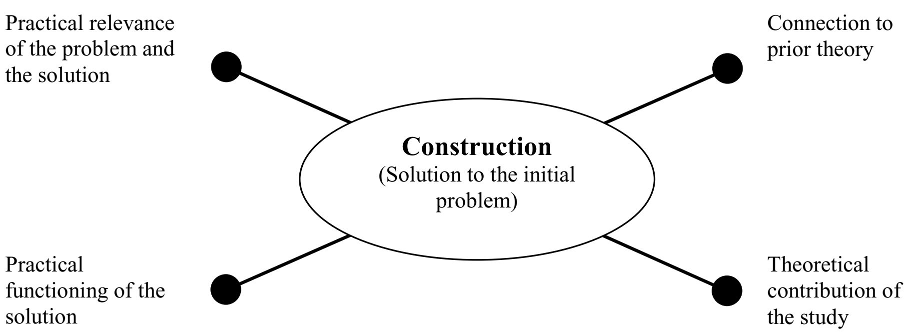


Figure 3. The central elements of the constructive research approach

Lukka (2002) divides the constructive research process into the following phases:

1 *Find* a practically relevant problem, which also has potential for theoretical contribution. An ideal topic has practical relevance, which is underanalyzed in prior literature.

2 *Examine* the potential for long-term research co-operation with the target organization(s). This involves making a formal research agreement, which deals with issues such as research funding, access to data, and particularly the conditions of publishing the results of the project.

3 *Obtain* a deep understanding of the topic area both practically and theoretically. This resembles a field study. With observations, interviews and analysis of archives the researcher gets insights into the state of affairs in the organizations. This analysis may reveal the explicit and implicit problems and should support the purposes of the research subject. The researcher should be aware of the potentially existing prior theories of the topic area. This is a basis for further development work based on prior knowledge, but also later on, to be able to identify and analyze the theoretical contribution of the study.

4 *Innovate* a solution idea and develop a problem solving construction, which also has potential for theoretical contribution. When an innovative construction cannot be designed, there is no point in going on with the project. This phase is creative and heuristic by nature. The implementation of an earlier designed construction should not be regarded as a constructive research approach.

5 *Implement* the solution and test how it works. The first level practical test, "market test", of the construction should be viewed as a key characteristic of the constructive research approach. This approach differs from a typical analytical model building, in which most sophisticated model designs are often just constructed while their empirical feasibility is not tested in any way.

6 *Ponder* the scope of applicability of the solution. In this phase the researcher has to be able to step back from the empirical work, control his or her level of commitment, and start reflecting the learning process. This means analyzing the results of the process and its preconditions. If the innovative construction passed the primary market test, i.e. it produced positive results, the next phase is to study whether the construction might be transferable to other organizations. Even if the market test failed, there will be room for theoretical analysis, i.e. whether the causes of the failure could be avoided in other projects.

7 *Identify* and *analyze* the theoretical contribution. The researcher has to be able to explain the theoretical contribution of the project, i.e. reflect the findings back to prior theories.

The constructive research approach differs from the other approaches in the strong emphasis on the formal co-operation agreement. One explanation for this slightly different emphasis might be that it has been developed in the field of business administration. But as its developers state, its potential application field is broad, also including information system science. This co-operation aspect is very important with the design of ANN systems because they are data driven systems and the data is normally owned by organizations.

Development Cycle Methodology

Medsker *et al.* (1994) suggest that the development cycle of an ANN includes four major phases (Figure 2).

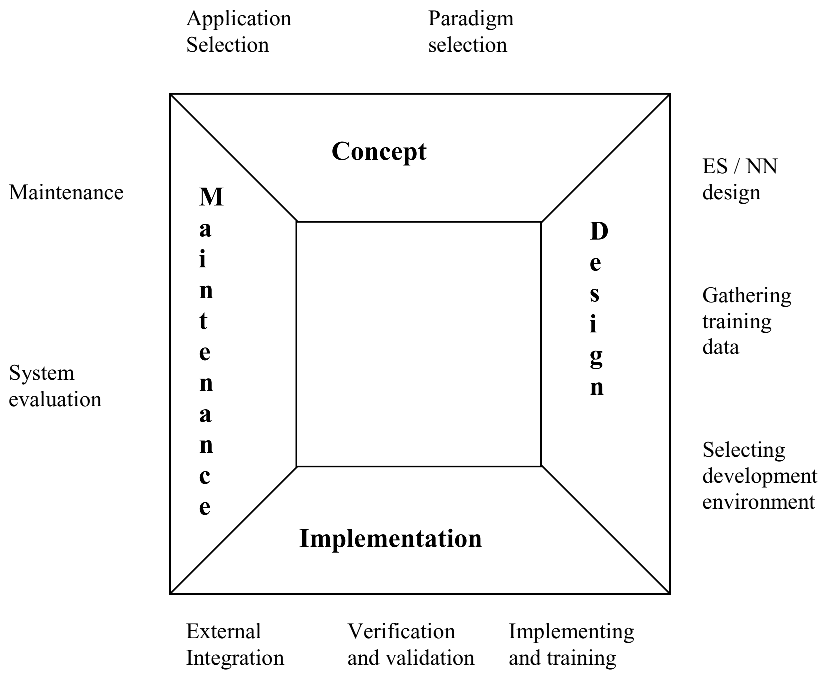


Figure 4. ANN development cycle

The *concept phase* focuses on an application assessment and selection of an appropriate ANN architecture to guarantee a successful application. The point here is to focus on the applicability of neurocomputing as an appropriate technology to solve the problem at hand. The available data, software, and the experience level of the developer are selection criteria for choosing the ANN paradigm for the problem. There are two basic learning paradigms in ANNs: *supervised* and *unsupervised*. In supervised learning the network learns from previous correct input-output values. In unsupervised learning the network is provided with no previous output values.

At the *design phase*, general and detailed designs of an individual system are created, the data required are gathered, and a development environment is chosen. ANN system creation means that a number of neurons, layers, and parameters like the transfer function are decided on. Adequate data must exist and be obtained and then divided into the training and testing sets. The developer must identify and clarify data relevant to the problem. Before the data is fed to the ANN it is somehow pre-processed. The last part of design phase involves matching the application's design to appropriate development environment.

At the *implementation phase*, a network is created, trained, and tested. Besides, external integration is included at this phase. The implementation phase involves the actual construction of the ANN and makes use of the data gathered for training and testing. During the iterative training process the ANN modifies the weights for the entire neural network. When the training is over the network has learnt an appropriate set of weights, which allows the network to carry out the desired task (Rumelhart *et al.*, 1994). Validation refers to a process of demonstrating that the correct system has been built. Verification involves proving that the system was constructed from the mechanical standpoint. Ongoing monitoring and feedback to the developers are required for system improvements and long-term success.

The *maintenance phase* starts when the building process is finished and continues during the application's entire lifetime. It evaluates periodically a system's performance and modifies it when necessary.

Although Medsker *et al.* (1994) have included some unique steps and considerations in the development cycle of ANNs it still resembles a sequential process of a conventional information systems design. One explanation for this might be that they have tried to develop one joint design and development cycle for both the expert system (ES) and neural network (NN). This is based on their argument that the processes for developing expert systems and neural networks have important parallels. However, the idea of training a neural network system to carry out an information processing function is something very special and unique in comparison to i.e. the programming of rule and knowledge bases for expert systems.

Comparison of approaches and methodologies

If we compare the approaches in Section 2 with the product and with the process of design aspects there are some differences. The design science (March *et al.* 1995) and the system development (Nunamaker *et al.* 1991) approaches combine both these aspects. The constructive research (Lukka 2002) approach puts emphasis on the process of the design aspect. The development cycle of ANN (Medsker *et al.* 1994) deals mainly with the design of the product aspect.

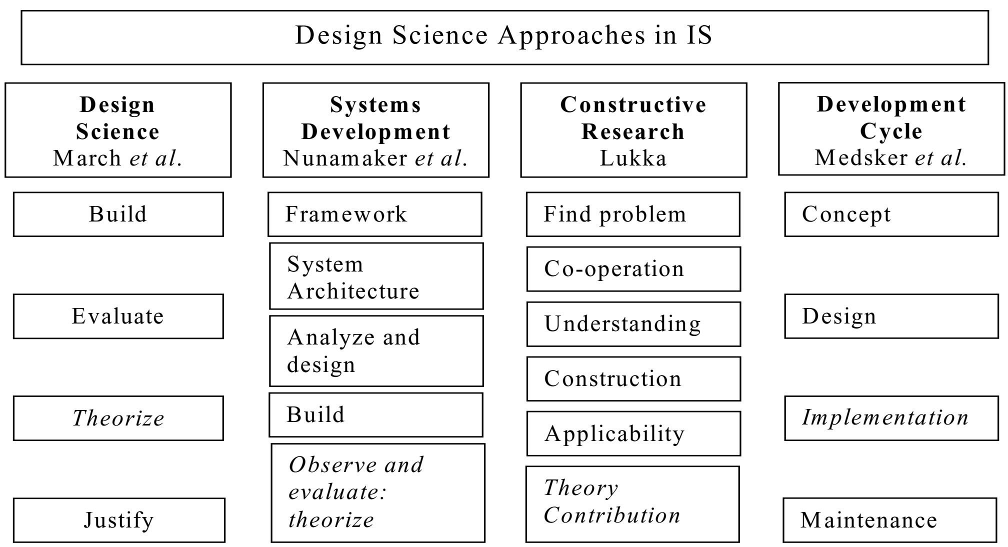


Figure 5. The comparison of approaches

In Figure 5 we have depicted the phases of these four approaches from top down. All the approaches have the same general idea that we have the problem and try to solve it with new artifacts. The new construction could be technology-oriented like ANNs. The aim is to get some improvement when comparing it to the traditional way of doing. However, the research projects, which fail to get improvement, could also have theoretical contributions. We have in *Italics* for those phases that have the issue of "theorize" in the approaches. With this "theorise" we mean the modification of the approach or framework or the explanation of the theoretical contribution.

In design science theorize activity involves explaining why and how the effects came about: why and how the constructs, models, methods, and instantiations work (March *et al.,* 1995). This activity attempts to unify the known data (observations of effects) into viable theory - explanations of how and why things happen. It may involve developing constructs with which to theorize about constructs, models, methods, and instantiations. The justify activity performs empirical and/or theoretical research to test the proposed theories. Such theories, once justified, can provide direction for the development of additional and better technologies.

According to the systems development methodology, once the system is built the researcher can test its performance and usability as stated in the requirement definition phase, as well as observe its impacts on individuals, groups and organizations (Nunamaker *et al.* 1991). The test results should be interpreted and evaluated based on the conceptual framework and the requirements of the systems defined at the earlier stages. Experiences gained from the developing of the system usually lead to the further development of the system, or the discovery of a new theory to explain newly observed phenomena. There are several ways of gaining experience e.g. experimentation and observation.

According to Lukka (2002) the constructive research approach has two major types of theoretical contributions: 1) The novel construction itself is a contribution. It should be regarded as a new means of achieving certain ends and therefore it is for further analysis. 2) The positive relationships behind the constructions are an arena for applying, testing, and developing the existing theoretical knowledge.

Medsker *et al.* (1994) do not emphasis the theory connection in the ANN development cycle at all. However, the implementation phase could have the theory contribution part in it because ANN development is an evolutionary process.

A new research approach for designing and developing ANNs

All these approaches mentioned above have many good guidelines to facilitate the design and development of ANN systems. However, none of them is comprehensive alone. Therefore, we present here a new approach based on these approaches and on our own experience. This approach combines three aspects of design - one dealing with the artifact, another with the process of design, and the third with the theory aspect. All these aspects should support each other all the time during the research process. Therefore our approach is broader, more concrete and detailed than the guidelines given before for designing and developing ANNs. This research approach is outlined in Figure 6.

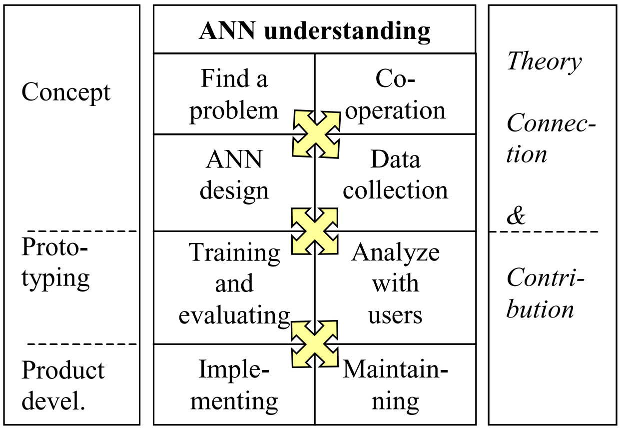


Figure 6. An approach for ANN systems design and development research

We argue that there are three aspects in the design and development of ANN systems. The left side illustrates the artifact aspect. It tells what the phase of the artifact is. It begins with the concept phase, continues with prototyping and with possible product development. The right side focuses on the theory aspect (*in Italics)*. This theory aspect should support both the designs of the artifact and the design process. In the beginning the prior theories have bigger influence on the research process than in the latter phases. At the end of the research process the emphasis can be opposite: the research process and the results may have influences on the theory. This aspect also illustrates the theoretical groundwork needed to carry out during the research process. This secures that the research is linked to prior knowledge. The experience learned during the research process is also easier to identify and analyze when the existing theories are understood. Our approach differs in this respect from the consulting approach.

In the middle is the core of ANN systems' design and development process. First of all, a researcher needs to obtain general understanding on how artificial neural networks work and what is possible to do with them. At the same time as the relevant problem is anticipated, one should start serious research co-operation discussions with the target organizations. This may also be needed to investigate whether the problem has practical concerns. Another reason is to get data for an ANN system. With the ANN system design and development this phase is extremely important because ANNs are data driven systems and data is needed to test whether the designed model works in the real world. Arrows in this process of design aspect illustrate the flow of data, information and knowledge needed to develop ANN systems for organizations.

The design of the ANN system starts with the selection of the learning paradigm for the problem. The available data and the assumption either the developer of the ANN system or the users in the target organization have about data may influence the choice of an appropriate leaning paradigm. If the output information is not known before, it is possible to first apply the unsupervised learning paradigm to classify the data and then use the supervised learning paradigm. The design of ANN architecture means that the number of neurons and layers is decided. During the training process the optimal numbers of hidden layers and neurons, the optimal values of parameters, and training times are determined. In order to test the ANN system the data should be divided into training and testing sets. Sometimes the data is somehow pre-processed before it is fed to the network. Very often it is useful to analyze the training and the testing results with possible users. The results of these analyses may lead to the redesigning of the ANN architecture and collection of more data to get better results. This analyze phase may also have such an effect that new problems are found that can be solved with the ANN technique. For example, the ANN system developed for auditing monthly account values is used for budgeting purposes. The careful analyses of the design process, the designed artifact, and the results of the assessment of the artifact are a point for the theory contribution.

At the implementation phase the constructed ANN system is implemented into the organization's processes. At the same time the maintenance phase starts. This means that ongoing monitoring of the ANN system and feedback to the responsible user or developer is required. This is needed to update the ANN system so that it works with new input data. Modification of the ANN system based on the new theories or knowledge might also be needed.

Acknowledgements

The author appreciates the comments received from two anonymous referees. The author also would like to thank Barbro Back for her comments and the Academy of Finland for financial support.

References

Coakley JR, Brown CE. 2000. Artificial Neural Networks in Accounting and Finance: Modeling Issues. International Journal of Intelligent Systems in Accounting, Finance & Management 9: 119-144

Dutta S. 1993. Knowledge Processing & Applied Artificial Intelligence. Butterworth-Heinemann: Oxford, England

Hebb D. 1949. The Organization of Behavior. Wiley: New York

Kasanen E, Lukka K, Siitonen A. 1993. The Constructive Approach in Management Accounting Research. Journal of Management Accounting Research (Fall): 243-264

Lukka K. 2002. The constructive research approach, Vol. 2002. Metodix.com

March ST, Smith GF. 1995. Design and natural science research on information technology. Decision Support Systems 15: 251-266

McCulloch W, Pitts W. 1943. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics (5): 115-133

Medsker L, Liebowitz J. 1994. Design and Development of Expert Systems and Neural Networks. Macmillan College Publishing Company, Inc.

Minsky M, Papert S. 1969. Perceptrons: An Introduction to Computational Geometry. The MIT Press: Cambridge, MA

Niiniluoto I. 1986. Pragmatismi. WSOY: Helsinki

Nunamaker Jr. JF, Chen M, Purdin TDM. 1991. Systems Development in Information Systems Research. Journal of Management Information Systems 7(3): 89-106

O'Leary DE. 1998. Using Neural Networks to Predict Corporate Failure. International Journal of Intelligent Systems in Accounting, Finance & Management 7: 187-197

Rosenblatt F. 1958. The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review 65: 386-408

Rumelhart DE, Widrow B, Lehr MA. 1994. The Basic Ideas in Neural Networks. Communications of the ACM 37(3): 87-92

Simon H. 1969. The Sciences of the Artificial. MIT press: Cambridge

Vellido A, Lisboa PJG, Vaughan J. 1999. Neural networks in business: a survey of applications (1992-1998). Expert Systems with Applications 51: 51-70

Wong BK, Bodnovich TA, Selvi Y. 1995. A bibliography of neural network business applications research: 1988-September 1994. Expert Systems 12(3): 253-262

Wong BK, Selvi Y. 1998. Neural network applications in Finance: A review and analysis of literature (1990-1996). Information & Management 34: 129-139

Zhang G, Patuwo BE, Hu MY. 1998. Forecasting with artificial neural networks: The state of the art. International journal of Forecasting 14: 35-62

PUBLICATION 3

Koskivaara (2000a) Artificial neural network models for predicting patterns in auditing monthly balances. *Journal of the Operational Research Society,* 51 (9): 1060–1069. Reprinted by permission of Palgrave Macmillan Limited.

Journal of the Operational Research Society (2000) 51, 1060–1069 0160-5682/00 $15.00

www.stockton-press.co.uk/jors

Artificial neural network models for predicting patterns in auditing monthly balances

E Koskivaara

Turku Centre for Computer Science and Turku School of Economics and Business Administration, Turku, Finland

The aim of this study is to investigate the potential of artificial neural network (ANN) models to recognise patterns when auditing monthly balances in financial accounts. ANNs have been used in many different disciplines as a basis for building intelligent information systems. This study examines the predictive ability of an ANN by building models using the 72 monthly balances of a manufacturing firm. The monthly balances are regarded as a time-series and the target is to recognise the dynamics and the relationships between different accounts. Furthermore, a certain seeded material error with signals from the ANN model is investigated. The results achieved indicate that neural networks seem promising for recognising the dynamics and the relationships between financial accounts.

Keywords: auditing; artificial neural networks; analytical review

Introduction

Background

A company has many accounts and they depend on each other. Transactions must be recorded in such a way that their relationships with the profit and loss account and the balance sheet can be ascertained without difficulty. Auditing is a part of the control process in organisations. One purpose of auditing is to examine and observe the validity of the accounts and financial reporting, and to ensure that they give a true and fair view of the company. This means that an auditor has independently and adequately to examine financial accounts or the related financial information of an entity. However, manual auditing can never cover all the evidence entirely and in detail, except in very small companies. An auditor needs, however, to determine whether accounts are free of material errors. This can be done by an analytical review process, which is an essential audit procedure used to examine the accuracy of account balances, without considering the details of the individual transactions which make up the account balance.[1] The major benefit of using analytical review techniques is the increased objectivity of the auditor's tests.[2] This means that the decision on whether and to what extent, to conduct further detailed tests can be made considerably more objective than intuitive auditing decisions. One example of an analytical review process is where the auditor compares an estimated expected value with the actual balance, to identify those accounts where further audit testing seems to be needed. Another example is where items in the profit and loss account and in the balance sheet are compared to other records in bookkeeping. For instance, where interest costs are outstanding compared to the average loan during the accounting period.

A new dimension in today's auditing is that a large amount of audit material, for example, receipts and accounting records, is increasingly displayed only electronically. Naturally, this kind of information also has to be audited. This is one reason why auditors need more, or maybe different kinds of, support systems.

In a recent survey, Willett and Page[3] found that auditors are often tempted to speed up testing by irregular methods. They reported three major reasons for this: (1) budget pressure (61%); (2) boring work (30%); and (3) unimportant work (41%). Only 22% of the respondents claimed never to have succumbed to the temptation to speed up testing by irregular methods. Reported irregular audit practices were: rejecting problematic items from samples (54%), testing fewer items than reported (27%), or accepting doubtful audit evidence (24%). Although the questions were indirect because of the sensitive nature of the research, the answers give some idea of irregular auditing. Furthermore, our daily newspapers report on problems in the auditing field every now and then. The following three examples are only the tip of the iceberg. Two of them come from the banking world, where the English Baring's bank and the Japanese Daiwa's bank lost millions because they did not have effective control systems. The third one is a Finnish multinational company, whose subsidiary in Italy overestimated the work in progress and recorded fictitious sales.

Correspondence: E Koskivaara, Turku Centre for Computer Science and Turku School of Economics and Business Administration, PO Box 110, FIN-20521 Turku, Finland.
E-mail: eija.koskivaara@tukkk.fi

Hansen and Messier[4] pointed out three factors illustrating why accounting data in computer systems must be controlled in a more timely fashion. Firstly, more individuals have access to computer systems. Secondly, more individuals are being trained to use computers, which means that the knowledge of how to use computer systems is no longer confined to small isolated groups. Thirdly, the amount of data in a computer system may be very large. These factors mean that effective monitoring tools are needed to ensure the protection of data against either fault or fraud. Vasarhelyi[5] has outlined a Continuous Process Auditing System (CPAS), where the data that flow through the system are monitored and analysed continuously using a set of auditors' defined rules.

Gathering and evaluating audit evidence

Fischer[6] has classified the literature on gathering and evaluating audit evidence into three general segments. Firstly, much of the literature has focused on the auditor's decision-making.[7,8] For example, the mismatch between knowledge and task may hinder auditors' ability to draw on previous experience when making conditional probability judgements and when allocating audit hours.[9]

Secondly, a sizeable amount of literature has focused on the development and evaluation of new audit methods. Researchers have, for example, used different statistical methods such as trend-, ratio-, and regression-analysis and univariate and bivariate Box-Jenkins time series analysis in auditing, for identifying unusual fluctuations which need more detailed investigation.[10,11] These methods have appeared to perform rather well in identifying unusual fluctuations in financial statements that need detailed investigation. For example, a regression-based analytical review and the use of monthly data increased audit effectiveness, and regression-based models were very efficient in detecting potential material misstatements.[12] However, these methods have not found their way into practice. The analytical review techniques used in practice for signalling errors throughout the audit are relatively simple models and are not based on any statistical methods.[13–15]

Thirdly, a smaller set of papers is concerned with the use of new audit technologies. For instance, Lieb and Gillease[16] reported Du Pont's audit risk-based decision support system, which has been used since 1986. Because of the use of this system the company has been able to reduce its staff while maintaining a fully effective audit coverage. The idea behind the system is that it selects portfolios based on the potential risk of not auditing a specific auditable unit. This risk is based on the collective judgement of the management.

ANNs as a new audit method

This study falls into the second segment and aims at getting further evidence of ANNs' capabilities in the area of auditing by building an ANN model using the monthly balances of a manufacturing firm. It follows that the focus is on the development and evaluation of ANNs as a new audit method. The following factors are the reasons for selecting the ANN as a new audit method.

Auditing requires prediction, control, and classification capabilities. ANNs have shown in earlier studies that they are suitable for a set of such applications in other fields, for example, in the capital markets. Refenes[17] points out two driving forces of ANNs. Firstly, ANNs are powerful tools for modelling and understanding human cognitive behaviour, and secondly, ANNs have powerful pattern recognition properties. They are able to recognize patterns even when the data are noisy, ambiguous, distorted, or variable.

Furthermore, because of their adaptive characters and their ability to handle non-linearity ANNs are an interesting alternative to existing methods.[18] ANNs continue to perform well even with missing or incomplete data, and ANNs are capable of discovering data relationships. Nor do a large number of input parameters pose a problem for ANNs. Besides, linear relationships between variables are very often a simplification of the nature of economic data. Moreover, Hill *et al*[19] observed that neural network models did significantly better than traditional statistical and human judgement methods when forecasting quarterly and monthly data in a financial time-series. Finally, an ANN has been considered one of the emerging technologies in this millennium.[20]

Although ANN techniques are being developed in a wide variety of business fields,[21,22] only a few works have applied ANNs to auditing problems. Coakley and Brown[10,11,23] compared financial ratio, regression and artificial neural network methods in their researches. Their results tentatively suggest that the artificial neural network method was able to extrapolate financial trends more effectively than the other statistical procedures due to its ability to recognise patterns within the financial accounts. Coakley[24] applied the ANN procedure to analysing patterns across financial ratios, and suggests that the use of the pattern analysis method as a supplement to traditional analytical procedures will offer improved performance in recognising material misstatements within the financial accounts. Wu[25] applied the neural network system in classifying tax cases to ascertain whether further audit is required or not. The 180 sample examples were gathered from the expert tax auditors' audit case file. The cases consist of information about a firm's business income tax behaviour. Fanning *et al*[26] and Brian, Green and Choi[27] have used ANNs to detect management fraud. Klersey and Dugan[28] have applied artificial neural networks to evaluate going concern situations. Tam and Kiang[29] have applied the neural network approach to predict bank failure. Koskivaara *et al*[30] modelled intelligent systems based on ANNs for auditing with the same accounts as Coakley and Brown. As far as we know ANNs are not yet used in practice for auditing.

Our aim is to build an ANN-based tool, which forecasts and recognises patterns in the financial accounts in monthly balances. We are visualising the account trends and seeding error in order to show how this kind of a system could assist the audit decision. The aim is that, if the ANN model works and is useful, the ANN model could be embedded into auditors' analytical review processes, and therefore be used in the audit-planning phase. Because the ANN model is embedded auditors do not have to be professionals in neural networks, they only need to understand the basic principles of the ANN. We will not compare our ANN results to any other statistical method for two reasons. Firstly, experts claim that the ANN technology is very competitive with standard statistical methods when applied to examining actual financial data,[23,24] and secondly, techniques such as regression analysis have attracted little use in the auditing practice.[13–15] Ameen and Strawser[14] have made the point that the more sophisticated analytical review techniques require large amounts of historical data, a stable trend, and familiarity with modelling. Moreover, interpretation of the regression results requires expertise, as does the understanding of data transformation, multi-collinearity, homoskedasticity and other technical issues associated with the use of regression.[15] According to Wallin[31] auditors are not familiar with these things. This is one reason why we are not promoting the use of regression analysis or any other statistical method in this study but a new method that might be more appealing for auditors. We expect good performance from an ANN in auditing because an ANN application:

- always works essentially with the same capacity
- does not complain if the work is too boring
- does not have an illusion of knowledge
- makes the balance between effectiveness and efficiency more appropriate by focusing the auditing on the right places

The remainder of the paper is organised as follows. Firstly, a brief description of the neural network method is given. Then, the models used in prediction and pattern recognition are specified. Finally, results and outlines for further research directions in this area are presented.

The basics of ANNs

The architecture of an ANN

A brief description of an ANN model is presented here. More details can be found in Hecht-Nielsen,[32] Hertz *et al*,[33] Freeman and Skapura,[34] Smith,[35] or Swingler.[36]

An ANN is composed of a set of *neurons*, which have many *input* vectors and one *output* vector. Figure 1 presents an artificial neuron with three inputs and one output. The *transfer function* transforms the inputs into the output. Within the transformation process the inputs are *weighted*.

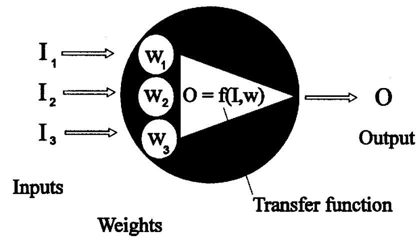


Figure 1 Simple structure of an artificial neuron.

Neurons are grouped into a number of *layers*. The first layer and the last layer within an ANN are called input and output layers, respectively. The inner layers (one or more) are known as *hidden* layers, which give the neural network its non-linear capability. The output of each neuron in a given layer, except the output layer, is fed as an input to every neuron of the next layer. Figure 2 shows a feed-forward ANN with three layers.

The learning method of an ANN

The ANN is based on a mathematical approach that uses differential equations, matrices, and linear algebra. Unlike traditional expert systems, where knowledge is made explicit in the form of rules, neural networks generate their own rules through training examples. Neural networks learn through a training procedure, in which the network is trained to perform certain tasks with a given set of examples. Learning is achieved by a *learning rule* that adapts or changes the weights of the network in response to example inputs and outputs.

Although there are a number of learning algorithms to train an ANN, the *backpropagation* paradigm has become the most popular one for prediction and classification problems.[37] It operates on a multilayered network like the one in Figure 2. When the relationship between the input and output variables is non-linear, a hidden layer helps in extracting higher level features and facilitates generalisation of outputs. For a given input vector, it generates the output

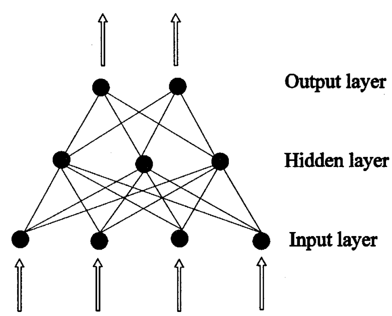


Figure 2 Structure of a multilayered neural network.

vector by a forward pass. Then, the difference between the output vector and the desired target output vector, the *root mean square error* (RMSE), is backpropagated through the ANN to modify the weights for the entire neural network. This iterative process is called *training*.

The backpropagation algorithm has two important parameters, *learning rate* and *momentum*. The learning rate affects the speed in which the network settles on a solution by allowing us to regulate how much the error decreases in each iteration. The *adaptive learning rate* accelerates the learning process by utilising the concept of the direction in which the error has been decreasing recently.[35] It speeds up the training process when the ANN is far away from the correct weights and slows down when the ANN gets closer. Momentum is another way to increase the speed of convergence: when calculating the weight change at each iteration, we add a fraction of the previous direction. This additional term tends to keep the weight changes going in the same direction.

Swingler[36] states that the training may be stopped when one or more of the following criteria have been satisfied: (1) The average training error has reached a predetermined target value; (2) The average training error no longer falls, or falls by an insignificant amount; (3) The average independent test error starts to rise, indicating the onset of overfitting. Early stopping and weight limitations are also approaches to reduce problems of overfitting. All these approaches have contributed considerably to the robustness of neural network training.

Normally, the raw input data to a network is somehow pre-processed. Firstly, the input data should be in such a form that a neural network can process it; this means numeric values. Secondly, the numeric input values may be pre-processed in a way that the network's learning task is easier. In many practical applications the choice of pre-processing will be one of the most significant factors in determining the performance of the ANN. Sometimes the output data are also processed. This is called post-processing, which is the inverse of the pre-processing transformation. The details of the backpropagation algorithm and pre- and post-processing of the data are described in standard textbooks on ANNs.[34–36]

After the ANN has been trained, it is tested against the records of a *testing* data set that have not been previously met with the network. For these records, the desired output is known. The output generated for each record of these testing data is checked against the desired output for that record. If there is a match, it is concluded that the trained network could recognise the record correctly.

The prediction model

The monthly balances of a firm are regarded as a time-series and the aim is to recognise the dynamics and the relationships between accounts. The goal is to predict the value of a certain account for a short time in the future. The feed-forward network can be applied directly to such problems.[38] We can take a set of account values from a number of periods, for example three successive periods x_{t-2}, x_{t-1}, x_t, to be the input vector to a network and use the next value x_{t+1} as the target value as illustrated in Figure 3. This is called *one-step ahead* prediction. In the *multi-step ahead* prediction the outputs of the neural network are cycled around to the inputs of the network, then predictions can be made for further points. By moving along the time axis we can create a training data set consisting of many sets of input vector values with corresponding target values.[38]

A one-step ahead prediction model was selected for this study because our aim is to forecast the account values for a short time in the future. Our opinion is that this forecasting is sufficient when auditing monthly balances. The reason for this is that all the input values are based on real data. In contrast, multi-step ahead prediction is at least partly based on predicted input values.

Monitoring the financial performance of a firm with ANNs

Next we will test the prediction model described above with an ANN. Our aim is to recognise the dynamics and the relationships between accounts, and, in that way, monitor if there are unusual fluctuations. In other words, we are applying the ANN method in an analytical review process. The annual accounts can be partitioned into twelve monthly ones. This test is based on the monthly balances of one company. A Certified Public Accountant (CPA)-auditor assisted in choosing the accounts. Next we will give further information about the selection criteria of the accounts.

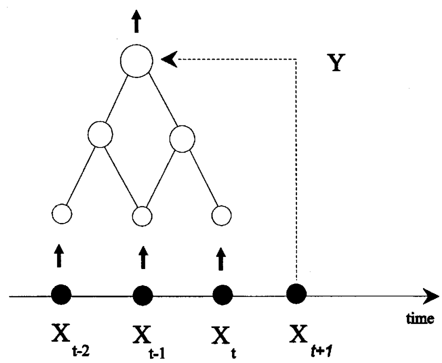


Figure 3 Generating a set of training data for a feed-forward ANN (*Source*: Based on Bishop[38]).

Description of the firm and the accounts

We used actual data comprising 72 monthly income statements—from January 1990 to December 1995—of a manufacturing company. The company was a medium-sized firm in Finland and its net sales amounted to approximately FIM 52 million per year. There are no changes in accounting methods in the five-year period. The accounts and their monthly averages in FIM thousand are presented in Table 1.

The reasons for selecting the above accounts for our models are as follows.

Net sales are a significant value to predict. In this particular case variation between months, especially between July and the other months, are very big. From the management's point of view it is better if the actual value is bigger than the budgeted value because then there are fewer disappointments. From the auditor's point of view this might raise doubts about whether all sales are recorded if the actual value is much below the prediction value.

Materials + Change in inventory together should tell the total use of material during a certain period. The value should be in alignment with net sales as this is a manufacturing company.

Personnel costs should be in alignment with production and the total use of material.

Gross margin is an important value at least from the prediction point of view as well as to see how much money is left to cover indirect costs and profit.

Administration is a good value to see the overall trend of the costs in the company and in the line of business.

Total indirect indicates all fixed costs. This value should be predicted in all cases because these costs do not depend on sales.

Operating profit is an interesting value at least from the prediction point of view. Furthermore, it is important to see that operation is profitable in the long run.

Receivables are an interesting and important value to follow in order to know how much of the company's money is 'outside'. Moreover, receivables have been taken into this model to illustrate how big the seeded fictitious sales must be before the model gives an alarm. In the case of fictitious sales, both net sales and receivables should go up while materials and personnel costs do not (= seeded material error).

Trade debt tells how much the company has to pay 'outside'.

Table 1 Financial accounts (in FIM 1000)

Net sales	4338
Materials + change in inventory	1096
Personnel costs	682
Gross margin	2560
Administration	351
Total indirect	1952
Operating profit	607
Receivables	8527
Trade debt	6777

Model 1 and Model 2

The implementation environment is a Pentium PC. A commercial development package, Neuralyst, developed by Cheshire Engineering Corporation was selected for prototyping because of its user-friendly features. A multilayer architecture was selected for our study. Our aim is to train the network to recognise a pattern in the material so that the network is able to predict future values based on prior values. The current study uses the backpropagation algorithm described earlier for training the networks. The data were pre-processed by a linear scaling to a [0, 1] range to make the network's learning task easier. This was done because the selected neural network works best when its inputs range from 0 to 1.[39] Another reason for this was that we wanted to keep the existing dynamics and relationships between the account values.

The training data set consists of the first 60 months of the total 72 months. The last 12 months were selected for testing. The ANN models were used to make a *one-step ahead* prediction described earlier for training and testing. Therefore, the training data were presented to the network in a chronological order. We tested the following two ANN models.

Model 1 uses two previous months to predict a third one. The idea is that Model 1 works inside a quarter. This required 18 input data streams to represent the data from each of the nine financial accounts. Model 2 uses four previous months to predict a fifth one. Model 2 is double the size of Model 1 when comparing input months. This required 36 input data streams to represent the data from each of the nine financial accounts. Both models have additional 12 input neurons to indicate the months of the input financial data. The input value of a neuron was set to one if the data corresponded to the month, and zero otherwise. To summarise, Model 1 has 30 input data streams and Model 2 has 48 input data streams in the input layer. Both models have nine neurons in the output layer, one for each financial account.

Training the networks

Numerous experiments were conducted to find the network architecture that minimised the RMSE between the target account balances and the predicted output balances. However, the training epoch, that is the number of processing runs through the complete training data, was limited to 10 000 to give an upper limit for the training time. This was also done because the average training error no longer fell by a significant amount. These experiments include altering, for example, the number of hidden layers or the number of neurons per hidden layer, the training tolerance between the target values and the output values in the training period, the learning rate, and momentum. Besides, an adaptive learning rate parameter was in use and Neuralyst provides this option. The adaptive learning rate enables a special mode which evaluates the RMSE after every training epoch and causes a revision of learning rate for the next training epoch. If the RMSE is high, then a high learning rate will be set, as the RMSE is reduced, the learning rate will be reduced correspondingly. The net effect is to speed up the training process when the neural network is far away from the correct weight, but to slow it down as it gets closer. The transfer function in both models is sigmoid. Table 2 summarises the models used in this study. The target value is the actual value and the output value is the value the ANN model gives.

The RMSEs for both models in the training period are close to each other. We may conclude that the learning ability of a neural network to monitor the non-linear dynamics of monthly balances is good. So, we managed to build such models that do not have significant differences in the training period. In the daily work of auditors this kind of analytical review method should be updated before use. The auditor may proceed to use an ANN-based analytical review method in cases where the difference between the target values and the output values is low. Another way to find out that an ANN model works is to visualise figures from the training period and compare the training results to the actual values by the account bases. Figure 4 shows the difference between the target values and the output values of receivables in the training period. It is very easy to see that the output receivables follow the target receivables very closely, and therefore, it is possible to apply this model to the test period.

Testing the networks

Both ANN models were tested after training. In theory, audited monthly balances should be free of material error. Therefore, this testing situation resembles a real auditing situation. In the planning phase an auditor may predict the account values with an ANN model and then compare them to the non-audited account values and therefore support the decision on how much further work is needed. The test data set consists of financial data from a 12-month period. We compared the ANN output values to the target values. Table 3 summarises the prediction capabilities of both models.

Table 2 Training statistics of Model 1 and Model 2

	Model 1	*Model 2*
Neurons in input layer	30	48
Hidden layers/neurons	2/27–18	4/40–32–26–18
Neurons in output layer	9	9
Set size (months)	54	54
RMSE	0.0261	0.021105
Learning rate	0.2	0.3
Momentum	0.2	0.3
Adaptive learning rate	True	True

Since the average difference between all the target values and all the output values is 6% for both models in the testing period, the average prediction ability of the ANN models to monitor the non-linear dynamics of monthly balances is fairly good. On average this kind of difference is, in my opinion, approximately within a true and fair view. However, an auditor has to go through all the figures in order to clarify those account values which differ the most. Especially, if the account value is not in line with other account values an auditor has to decide how to proceed. Model 2 is slightly better than Model 1. Figure 5 shows the difference between the target values of the receivables and the output values of the receivables in the testing period. Figure 5 is based on the values of Model 2. Figure 5 illustrates how the ANN model has recognised the non-linear dynamics and the relationship between account values. We see for example that in April the output receivable is about 5% higher than the target receivable. Similarly, in December the output receivable is about 1% lower than the target receivable. Similar pictures can be constructed from the other accounts. The output line of receivables clearly follows the target line of receivables, however in March, April, and May the output values are higher than the target values. One explanation for this is that the values in the training period are clearly higher than the values in the testing period.

Analysing the results

Table 4 presents the prediction capabilities (average difference in per cents, difference range in per cents [min, max], and differences less than 10%) of the ANN models according to the accounts. Table 4 shows, for example, that the average difference for the output values and the target values of receivables in Model 2 is 2%, and the difference range in percents is $[-1, 5]$. This means that all the values (100%) are inside the limit of 10% (≅ true and fair view). Overall 11% (14 data item out of 108) of the test data were not inside the 10% limit.

Overall, we can conclude that the ANN model produces good results with materials + change in inventory, personnel costs, gross margin, administration, operating profit, and receivable accounts. Net sales and total indirect values attained by the ANN model are also on average within a true and fair view, the worst case was the trade debt. The results of trade debt were fairly poor when comparing them to other account values. The reason for this might be that the trend of the trade debt does not follow the dynamics of the other account values and therefore the ANN-model was

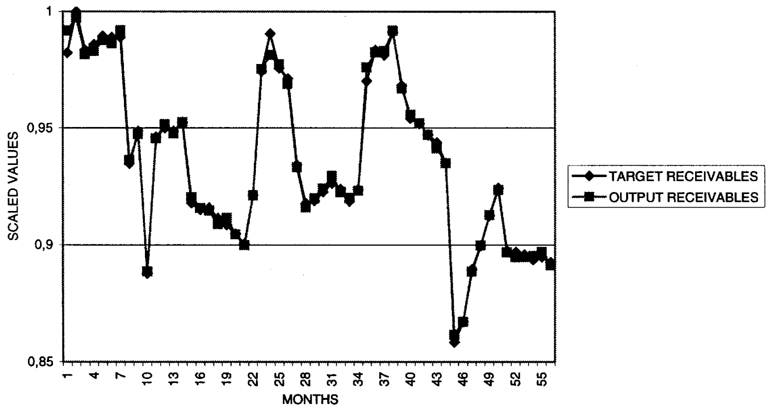


Figure 4 Difference between the target and output values, which are scaled to [0, 1], of receivables in the training period.

Table 3 The test statistics of the models used in this study

	Model 1	*Model 2*
Set size: accounts/months	9/12	9/12
RMSE	0.149692	0.142205
Outputs differ less than 5% from targets	24%	26%
Outputs differ less than 10% from targets	50%	51%
Outputs differ less than 30% from targets	94%	97%
Average difference between targets and outputs	6%	6%

Table 4 The prediction capabilities of the models according to the accounts (Average difference/difference range/less than 10%)

	Model 1	*Model 2*
Net sales	4/[−13,5]/92	5/[−10,12]/83
Materials + Change in inventory	3/[−7,12]/92	5/[−8,9]/100
Personnel costs	3/[−5,5]/100	2/[−4,4]/100
Gross margin	4/[−9,5]/100	5/[−11,8]/92
Administration	1/[−1,2]/100	1/[−1,4]/100
Total indirect	5/[−6,15]/83	6/[−7,15]/83
Operating profit	5/[−11,9]/92	5/[−7,12]/92
Receivables	2/[−3,4]/100	2/[−1,5]/100
Trade debts	28/[−23,88]/0	21/[−36,51]/33

not able to predict the trade debt values as well as the other account values. For example, the average change of the trade debt values from the last training period to the test period is big, 20%. At the same time, the average change of all other accounts from the last training period to test period is 6%. Another reason for the worse learning ability might be the fact that the trade debt account is the only big credit account in the data. Moreover, one explanation for this odd behaviour of the trade debt might be that the manufacturing

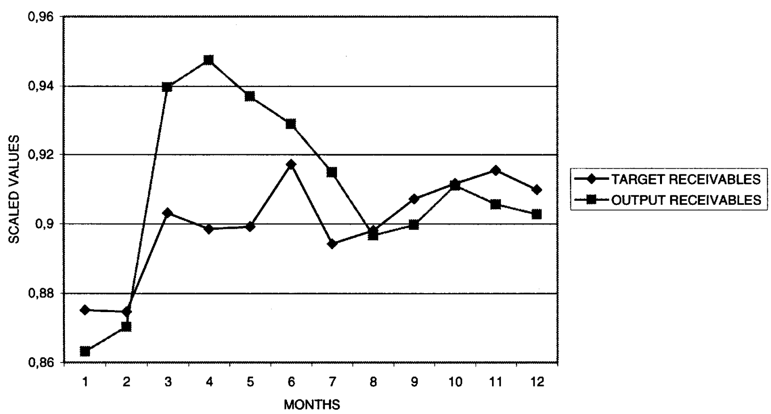


Figure 5 Difference between the target and output values, which are scaled to [0, 1], of receivables in the training period (= 12 months).

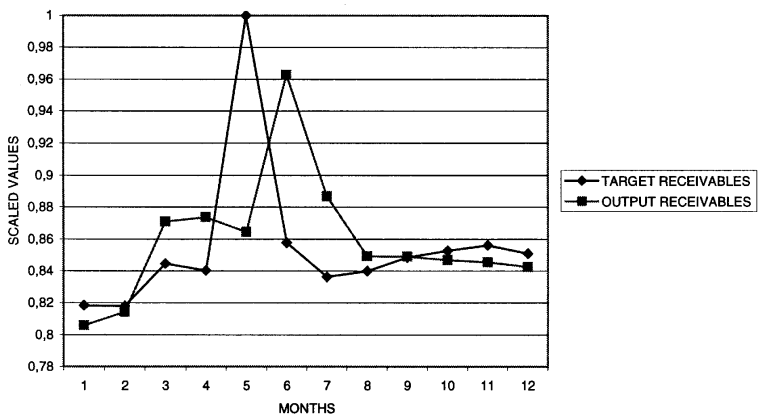


Figure 6 Error in receivables when seeding fictitious sales in May.

company has bought raw material when it has been inexpensive or easily available. Another explanation might be that when the company has been a little short of money it has delayed the payment of the trade debt. Those were some explanations why the trend of trade debt was not as stable as the trend of other accounts. When an account behaves like this it is of course also a hint to the auditor to pay more attention to that account.

Seeded error

The availability and affordability of technology appropriate to the task and problem are a necessity for applying a decision support system to audit tasks. One way to determine the appropriateness of the decision support system for auditing financial accounts is to find out if the system detects material errors, assessing materiality is not unambiguous. However, some quantitative thresholds, such as 5% as 'a rule of thumb', may provide the basis for a preliminary assumption in assessing materiality.[40]

We selected the better model (Model 2), which uses four prior months to predict the fifth one, for further investigation. We tested the model by seeding fictitious sales in May, which means that the net sales and the receivables should go up whiles others do not. The ANN model detected the error, when the amount of seeded error was 5% of the year's average net sales in two ways. In the case of fictitious sales the RMSE went up from 0.142205 to 0.162205 and the target values were unusually high and they were not inside the 10% limit in May. This seeded error also has effects on other account values because now overall 14% (19 data item out of 108) of the test data were

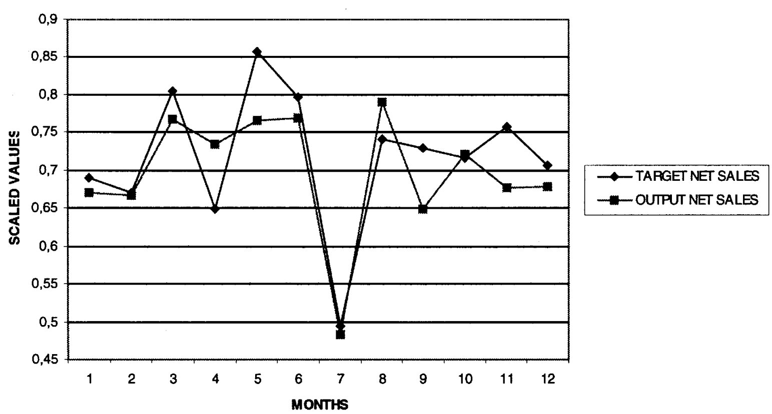


Figure 7 Error in scaled net sales when seeding fictitious sales in May.

not inside the 10% limit. Figure 6 and Figure 7 present the results of seeded fictitious sales in May and they show that both target values are clearly above the output values.

In situations like this, auditors should clarify why the actual target values are higher than the output values the net has predicted. This kind of support system works well in the beginning of an analysis of materiality. However, it cannot appropriately be used as a substitute for a full analysis of all relevant considerations.

Discussion and further research

The main purpose of this study was to build an ANN-based monitoring tool, which forecasts and recognises patterns when auditing financial accounts in monthly balances. We illustrated how an auditor may use an ANN model to support the planning of auditing by a graph on the computer screen that either signals that 'no further audit is required' or 'further audit is required'. This is different from the Coakley and Brown study.[23] The neural network application succeeded in predicting the future account values as well as in monitoring the seeded material error. Therefore, we think the neural network method could provide improvement in the efficiency and effectiveness in auditing financial records, however this has to be tested in further experiments. Further experiments will have to be made to test what the acceptable difference between the target value and output value might be without actions in the management or in the bookkeeping procedure. Another point to focus on is the pre-processing methods. In this study, we used a linear pre-processing method on the data before the network training procedure. Other pre-processing methods might give better results. We also plan to extend our research to include more companies and other accounts as variables. In addition to account values, there can be other variables, for example, orders in the 'queue', the failure rate of machines, and the number of staff turnover, which could better predict changes in accounts. We will also apply alternative prediction models and alternative network structures. If the ANN model is useful in practice, it will be embedded into the auditors' analytical review process, and thereby used in the audit-planning phase. Therefore, we plan to further apply the neural network technology to a real audit situation to get more results.

Acknowledgements—The author appreciates comments received from two anonymous referees, and is grateful to the company who provided the data for this study. The author would also like to thank Pekka Pajamo from KPMG Wideri and colleagues Barbro Back, Tapio Reponen, Kaisa Sere and Marko Grönroos from Turku Centre for Computer Science TUCS, for their beneficial comments. The author wishes to thank the Academy of Finland for providing financial support for this project.

References

1 Knechel WR (1986). A simulation study of the relative effectiveness of alternative analytical review procedures. *Decis Sci* **17**: 376–394.

2 Lev B (1980). On the use of index models in analytical reviews by auditors. *J Account Res* **18**: 524–550.

3 Willett C and Page M (1996). A survey of time budget pressure and irregular auditing practices among newly qualified UK chartered accountants. *Brit Account Rev* **28**: 101–120.

4 Hansen JV and Messier WF Jr (1983). Scheduling the monitoring of EDP controls in online systems. *Inter J Comput and Inform Sci* **12**: 35–46.

5 Vasarhelyi MA and Halper FB (1991). The continuous online systems. *Auditing* **10:** 110–125.

6 Fischer MJ (1996). 'Real-izing' the benefits of new technologies as a source of audit evidence an interpretative field study. *Account Organiz and Soc* **21**: 219–242.

7 Bonner SE (1990). Experience effects in auditing: the role of task-specific knowledge. *The Account Rev* **65**: 72–92.

8 Bonner SE and Walker PL (1994). The effects of instruction and experience on the acquisition of auditing knowledge. *The Account Rev* **69**: 157–178.

9 Nelson MW, Libby R and Bonner SE (1995). Knowledge structure and the estimation of conditional probabilities in auditing planning. *The Account Rev* **70**: 27–47.

10 Coakley JR and Brown CE (1991). Neural networks for financial ratio analysis. In: *Proceedings of The World Congress on Expert Systems I*. Pergamon Press: Oregon, pp 132–139.

11 Coakley JR and Brown CE (1991b). Neural networks applied to ratio analysis in the analytical review process. In: *Proceedings of the Fourth International Symposium on Expert Systems in Accounting, Finance and Management*, University of Southern California and The Peak Marwick Main Foundation, Pasadena, CA, pp 1–36.

12 Knechel WR (1988). The effectiveness of statistical analytical review as a substantive auditing procedure: a simulation analysis. *The Account Rev* **63**: 74–95.

13 Turpiainen J (1994). Assistant tools for auditing. *Tilintarkastus-Revision*, **2**: 143–146 (in Finnish).

14 Ameen EC and Strawser JR (1994). Investigating the use of analytical procedures: an update and extension. *Auditing: A Journal of Practice & Theory*, **13**: 69–76.

15 Fraser IAM, Hatherly DJ and Lin KZ (1997). An empirical investigation of the use of analytical review by external auditors. *Brit Account Rev* **29**: 35–47.

16 Lieb EB and Gillease JK (1996). Du Pont uses a decision support system to select its audit portfolio. *Interfaces* **26**: 10–21.

17 Refenes AP (ed.) (1995). *Neural Networks in the Capital Markets*. John Wiley & Sons Ltd: Chichester.

18 Scherer A (1996). Combining neural networks with other prediction techniques, an empirical study. In: *Proceedings of the 29th Annual Hawaii International Conference on Systems Sciences*. IEEE Computing Society Press, USA pp 437–446.

19 Hill T, O'Connor M and Remus W (1996). Neural network models for time series forecasts. *Mgmt Sci* **42**: 1082–1092.

20 Halal WE, Kull MD and Leffmann A (1998). The George Washington University forecast of emerging technologies: a continuous assessment of the technology revolution. *Technol Forecast and Social Change* **59**: 89–110.

21 Wong BK, Bodnovich TA and Selvi Y (1995). A bibliography of neural network business applications research: 1988–September 1994. *Expert Sys* **12**: 253–262.

22 Wong BK and Selvi Y (1998). Neural network applications in finance: a review and analysis of literature (1990–1996). *Inform Mgmt* **34**: 129–139.

23 Coakley JR and Brown CE (1993). Artificial neural networks applied to ratio analysis in the analytical review process. *Intelligent Systems in Accounting, Finance and Management* **2**: 19–39.

24 Coakley JR (1995). Using pattern analysis methods to supplement attention-directing analytical procedures. *Expert Systems with Applications* **9**: 513–528.
25 Wu RC-F (1994). Integrating neurocomputing and auditing expertise. *Managerial Audit J* **9**: 20–26.
26 Fanning KM, Cogger KO and Srivastava R (1995). Detection of management fraud: a neural network approach. *International Journal of Intelligent Systems in Accounting, Finance and Management* **4**: 113–126.
27 Brian P, Green J and Hwa C (1997). Assessing the risk of management fraud through neural network technology. *Auditing: A Journal of Practice and Theory* **16**: 14–28.
28 Klersey GF and Dugan MT (1995). Substantial doubt: using artificial neural networks to evaluate going concern. *Advances in Account Inform Syst* **3**: 137–159.
29 Tam K and Kiang M (1992). Managerial applications of neural networks: the case of bank failure predictions. *Mgmt Sci* **38**: 926–948.
30 Koskivaara E, Back B and Sere K (1996). Modelling intelligent systems for auditing. In: Sierra GJ and Bonsón E (eds). *Intelligent Systems in Accounting and Finance*. Papel Copy SL, Plaza de la Merced: Huelva, pp 233–252.
31 Wallin J (1998). Analytisk granskning med hjälp av Exel och regression analysis. *Tilintarkastus—Revision*, **1:** 37–40 (in Swedish).
32 Hecht-Nielsen R (1991). *Neurocomputing*. Addison-Wesley: Reading, MA, USA.
33 Hertz J, Krogh A and Palmer RG (1991). *Introduction to the Theory of Neural Computation*. Addison-Wesley: CA, USA.
34 Freeman JA and Skapura DM (1991). *Neural Networks: Algorithms, Applications and Programming Techniques.* Addison-Wesley Publishing Company, Inc., MA, USA.
35 Smith M (1996). *Neural Networks for Statistical Modeling* International Thompson Computer Press: Boston, USA.
36 Swingler K (1996). *Applying Neural Networks: A Practical Guide.* Academic Press: London.
37 Sohl JE and Venkatachalam AR (1995). A Neural network approach to forecasting model selection. *Inform & Mgmt* **29**: 297–303.
38 Bishop CM (1995). *Neural Networks for Pattern Recognition*. Clarendon Press: Oxford.
39 Shih Y (1994). *Neuralyst*™ *User's Guide*. Cheshire Engineering Corporation, USA.
40 See, eg, Securities and Exchange Commission Staff Accounting Bulletin: *No. 99—Materiality.* www.sec.gov/rules/acctreps/sab99.htm.

Received December 1998;
accepted May 2000 after two revisions

PUBLICATION 4

Koskivaara (2000b) Different Pre-Processing Models for Financial Accounts when Using Neural Networks for Auditing. In: *Proceedings of the European Conference on Information Systems*, ed. by Hans Robert Hansen – Martin Bichler – Harald Mahrer. Vienna, Austria, July 3–5, 2000, 326–332. Reprinted by permission.

Different Pre-Processing Models for Financial Accounts when Using Neural Networks for Auditing

Eija Koskivaara
Turku Centre for Computer Science TUCS, and
Turku School of Economics and Business Administration
P.O.Box 110, FIN-20521 Turku, Finland
Eija.Koskivaara@tukkk.fi

Abstract-**The aim of this study is to investigate the impact of various pre-processing models on the forecast capability of artificial neural network (ANN) when auditing financial accounts. Hence, the focus of this paper is on the pre-processing of the data. ANNs are selected for auditing purposes because they are capable of learning complex, non-linear underlying relationships. Therefore, they are used to model the dynamics and the relationships between account values in order to find unexpected fluctuations. This study uses a multi-layered neural network with the backpropagation algorithm. The artificial neural network model used in this study was built by using the financial statements of 31 manufacturing companies over four years. The values of the accounts were regarded as a time-series. The data were pre-processed in four different ways. Firstly, all the data were scaled linearly. Secondly, the data were pre-processed linearly on a yearly basis. Thirdly, the data were pre-processed linearly on a company basis. And fourthly, the data were pre-processed on a yearly and company basis. The best results were achieved when all the data were scaled either linearly or linearly on a yearly basis.**

I. Introduction

Auditing is a part of the control processes in organisations. One purpose of auditing is to examine and observe the creditability of the account values, and that they give a true and fair view of the company. This means that an auditor has to examine financial accounts or related financial information of an entity. However, manual auditing can never cover all the evidence entirely and in detail, except in very small companies. Auditing of financial accounts and other financial information is often made with the help of analytical review procedures. Analytical review procedures are defined as an evaluation of financial information studying plausible relationships among both financial and non-financial information. Analytical review techniques in the literature are generally classified as non-quantitative or judgmental such as scanning, simple quantitative such as trend, ratio and reasonableness test, and advanced quantitative such as regression analysis and neural networks. However, techniques such as regression analysis have attracted little use in the auditing practice [1][2]. Moreover, the research of auditor judgement and decision-making has focused on revealing the limitations of human auditors [3][4]. These research findings have concluded that developing tools to overcome these limitations might enhance auditors' effectiveness and/or efficiency [5]. Indeed, a growing dimension in today's auditing is that a large amount of audit material, e.g. receipts and accounting records, is increasingly displayed only electronically. Naturally this kind of information also has to be audited. Furthermore, the American Institute of Certified Public Accountants (AICPA) Committee believes that major changes are currently under way involving both the kinds of information with which auditors are involved and the nature of that involvement. They describe this as a shift from an old audit paradigm to a new audit paradigm. This shift is also referred to as a transformation from audit to assurance. It also means that auditors have more and more varying kind of work to do than they had earlier. The report about the future of the financial statement audit is found on the homepage of AICPA (http://www.aicpa.org). These are some reasons why auditors need more or maybe different kind of support systems.

We are investigating the possibility to develop a support system, which uses an artificial neural network (ANN), for auditing financial accounts. We think that an ANN is suitable for these kinds of problems and will be one of the emerging technologies in this millennium [6]. Many ANN models are being developed in a wide variety of business fields [7][8]. Auditing as an application area for ANNs is also emerging [9][10][11][12][13][14][15][16][17][18][19]. As far as we know, ANNs are not yet used in practice within auditing.

The aim of this study is to get further evidence of the capability of an ANN to forecast and recognise patterns when auditing financial accounts. Therefore, the focus is on the development and evaluation of an ANN as a new audit method. The ANN architecture category and the prediction model for predicting and recognising patterns in this study is similar to the one in Koskivaara's [18] study. The difference is that we have more accounts, and in addition we have the number of staff turnover as variables. Furthermore, we have more companies in the study, because our aim is to build an ANN-based support system which forecasts and recognises patterns in financial accounts within one business line. In particular, we are focusing on the impact of various pre-processing models on the forecast capability of an ANN when auditing financial accounts.

The rest of the paper is organised as follows. First, a brief description of the multi-layer feedforward ANN method will be given. Then, the models used in prediction and pattern recognition will be specified. Finally, results and outlines for further research directions in this area will be presented.

II. ANN For Forecasting

We will focus on a particular structure of an ANN, the multi-layer feed-forward network, which is the most popular and widely used network paradigm in forecasting. We will present a brief description of it here. More details and applications of ANNs can be found in standard textbooks on ANNs [20][21][22][23][24].

The basic element of an ANN is a *neuron*. A neuron has many *input* vectors and one *output* vector. Fig. 1 presents an artificial neuron with three inputs and one output. The *transfer function* transforms the inputs into the output. The inputs are *weighted* within the transformation process.

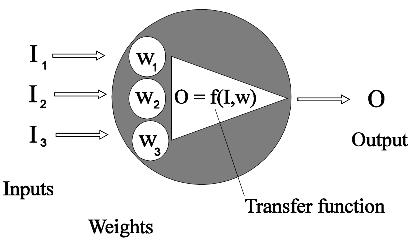


Fig. 1. Simple structure of an artificial neuron

An ANN is typically composed of layers of neurons. The first layer and the last layer within an ANN are called input and output layers, respectively. The inner layers (one or more) are known as *hidden* layers. The output of each neuron in a given layer, except the output layer, is fed as an input to every neuron of the next layer. Fig. 2 shows a feed-forward ANN with three layers.

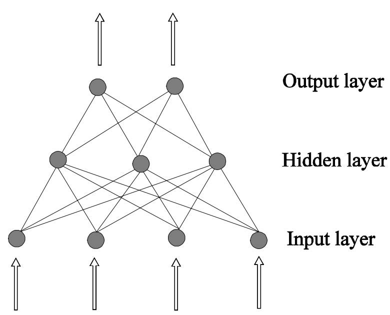


Fig. 2. Structure of a multi-layered neural network

In designing an ANN model, one must determine the following network architecture variables:

- *The number of input neurons*
- *The number of hidden layers and hidden neurons*
- *The number of output neurons*
- *Transfer or activation function*
- *Training algorithm*
- *Pre- and post-processing*
- *Training sample and test sample*
- *Performance measures*

The number of input neurons: The number of input neurons corresponds to the number of variables in the input *vector* used to forecast future values.

The number of hidden layers and hidden neurons: The hidden neurons in the hidden layer allow neural networks to detect the feature, to recognise the pattern in the data, and to perform complicated non-linear mapping between input and output variables. When the relationship between the input and output variables is non-linear, a hidden layer helps in extracting higher level features and facilitates the generalisation of outputs.

The number of output neurons: The number of output neurons is directly related to the problem under study.

Transfer or activation function: In practice, only a small number of activation functions is used [25]. These include the sigmoid (logistic) function, the hyperbolic tangent function, the sine or cosine function, and the linear function. The sigmoid transfer function is the most popular choice.

Training algorithm: During training, the weights of a network are iteratively modified to minimize the overall mean or total squared error between the desired and actual output values. Neural networks receive their *intelligence* through this training, which is also called *learning*. Although there are number of learning algorithms to train an ANN, the *backpropagation* paradigm has become the most popular one for prediction and classification problems [26]. It operates on a multilayered network like the one in Fig. 2. The backpropagation algorithm has two important parameters, the *learning rate* and *momentum*. The learning rate affects the speed in which the network settles on a solution by allowing us to regulate how much the error decreases at each iteration. Momentum is another way to increase the speed of convergence: when calculating the weight change at each iteration, we add a fraction of the previous direction. This additional term tends to keep the weight changes moving in the same direction.

Pre- and post-processing: The input data should be in such a form that a neural network can process it, this means numeric values. Furthermore, the numeric input values may be pre-processed in a way that the network's learning task is easier. In many practical applications, the choice of pre-processing will be one of the most significant factors in determining the performance of the ANN. One commonly used pre-processing method is data normalisation. Pre-processing is always performed before the training process begins. Sometimes the output data are also processed. This is called post-processing, which is the inverse of the pre-processing transformation.

Training sample and test sample: The data set used in network development is divided into a training sample and a testing sample. The training sample is used to build the network. Swingler [24] states that the training may be stopped when one or more of the following criteria have been satisfied: 1) The average training error has reached a predetermined target value. 2) The average training error no longer falls, or falls by an insignificant amount. 3) The average independent test error starts to rise, indicating the onset of overfitting. After training, the ANN model is tested against the records of a test sample that have not been previously met with the network. For these records, the desired output is known. The output generated for each record of this test sample is checked against the desired output for that record. There can be many performance measures for an ANN model, such as modelling time, training time or development costs. The most important performance measurement is the prediction accuracy it can achieve beyond the training data.

Performance measures: An accuracy measure is often defined in terms of a forecasting error which is the difference between the actual/desired and the predicted value. The most frequently used are the mean absolute deviation (MAD), the sum of squared error (SSE), the mean squared error (MSE), the root mean squared error (RMSE), and the mean absolute percentage error (MAPE) [25].

III. The Prediction Model Used in This Study

Our aim is to build an ANN-based support system, which will forecast and recognise patterns in financial accounts within one business line. The system could assist in the audit decision by giving a suggestion that could either be "no further audit required" or "further audit required". The accounts in financial statements are regarded as a time-series and the aim is to observe the dynamics and the relationships between the accounts. The goal is to predict the value of a certain account for a short time in the future. The feed-forward network can be applied directly to such problems [27].

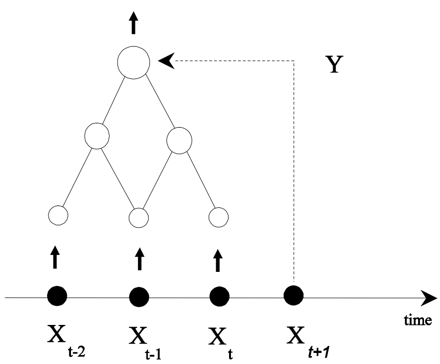


Fig. 3. Generating a set of training data for a feed-forward ANN. [27]

We can take a set of account values from a number of periods, for example two successive periods x_{t-2}, x_{t-1} and x_t to be the input vector of a network and use the next value x_{t+1} as the target value as illustrated in Fig. 3. This is called *one-step ahead* prediction. In the *multi-step ahead* prediction the outputs of the neural network are cycled around the inputs of the network, then predictions can be made for further points. By moving along the time axis we can create a training data set consisting of many sets of input vector values with corresponding target values [27].

The one-step ahead prediction model was selected for this study because our aim is to forecast the account values for a short time in the future. This also means that all the input values are based on real data.

IV. Monitoring The Financial Accounts of One Line of Business with A Neural Network

An ANN is a new possible method to be used in the analytical review process. An auditor needs to determine whether a financial statement is free of material errors and this can be done through an analytical review process [28]. For example, an auditor compares an estimated expected value with the actual value to identify those accounts where further audit testing seems to be needed. Our aim is to observe the dynamics and the relationships between accounts in one line of business and in that way monitor if there are unusual fluctuations. Furthermore, we are investigating four different ways to pre-process the data.

We will test the prediction model described above with a multi-layer feed-forward neural network. The study is based on the financial statements of 31 manufacturing companies over four years (1993-1998), all in the same line of business. The companies were picked from one bigger database. The limited companies with a turnover of over 3 million Finnish marks (~0.5 million euros) were accepted. Next we will give further information about the selection criteria of the variables. The variables and their averages and medians in thousand FIMs are presented in Table 1.

Variables

The reasons for selecting the above accounts for our models are as follows:

Net sales and other business earnings are significant values to predict. From the management's point of view it is better if the prediction value is lower than the actual value because then there are fewer disappointments. From the auditor's point of view this might raise doubts about whether all sales are recorded if the actual value is much below the prediction value.

Materials + change in inventory together should tell the total use of material during a certain period. The value should be in alignment with the net sales as these are manufacturing companies.

External service is quite a significant cost for some companies.

Personnel costs (manufacturing) and other direct costs should be in alignment with production and the total use of material.

Sales margin is an important value at least from the prediction point of view as well as to see how much money is left to cover indirect costs and profit.

Personnel costs (administration) and other indirect costs are good values to see the overall trend of the costs in the company and in the line of business. These values should be predicted always and everywhere because these costs do not depend on sales.

Profit (total) is an interesting value at least from the prediction point of view. Furthermore, it is important to see that operation is profitable in the long run.

Average staff number should be in alignment with personnel costs.

Receivables are an interesting and important value to follow in order to know how much of the company's money is "outside". Moreover, the receivables have been taken into the model to illustrate how big the seeded fictitious sales must be before the model gives an alarm. In the case of fictitious sales, both net sales and receivables should go up.

Accounts payable are interesting and important values to follow, and they tell how much the company has to pay "outside". It should be in alignment with net sales and the total use of material.

Financial assets and *short-term liabilities* are taken into the model in order to calculate quick ratio (QR), which shows the liquidity of the company.

TABLE I
Variables

Variable	Averages	Median
1. Net sales	19076	9389
2. Other business earnings	113	11
3. Materials + change in inventory	8966	3591
4. External service	686	5
5. Personnel costs (manufacturing)	3046	2386
6. Other direct costs	295	141
7. Sales margin	6437	2761
8. Personnel costs (administration)	1633	756
9. Other indirect costs	1979	626
10. Operating margin	2825	1387
11. Profit (total)	1433	606
12. Average staff (number)	26	19
13. Receivables	2862	895
14. Financial assets	5381	1709
15. Accounts payable	1331	535
16. Short-term liabilities	4171	1833

The implementation environment is a Pentium PC. A commercial development package, Neuralyst, developed by Cheshire Engineering Corporation was selected for prototyping because of its user-friendly features. A multilayer architecture was selected for our study. Our aim is to predict future values based on prior values. The current study used the back-propagation algorithm described in Section II for training the networks. The data were pre-processed to a [0,1] range to make the network's learning task easier. This was done because the selected neural network works best when its inputs range from 0 to 1 [29]. This pre-processing is made in four different ways.

Four different pre-processing ways

Model A: *All* the data are scaled linearly. The reason for this is that we wanted to keep the existing dynamics and relationships between the variables. In the case where there are any trends and relationships between the years and companies the ANN model might recognise them. In this case all data should always be checked and maybe pre-processed before entering new data into the support system.

Model Y: The data are pre-processed linearly *on a yearly basis*. An assumption in this case is that it is easier for the ANN model to learn the dynamics and relationships between the variables if the year effect is minimised on a yearly basis pre-processing. In this case when the year or other selected period is closed there is no need for data re-pre-processing.

Model C: The data are pre-processed linearly *on a company basis*. Respectively, an assumption in this case is that it is easier for the ANN model to learn the dynamics and relationships between the variables if every company has its own sliding scale. All of the company's data should pre-process when new data are put into the support system.

Model C&Y: The data are pre-processed *on* a *yearly and company basis*. Correspondingly, an assumption in this case is that it easier for the ANN model to learn the dynamics and relationships between the variables if every company has its own sliding scale on a yearly basis. The data from the period are ready after one pre-processing.

Training and testing sets and the prediction model

The training data set consists of 25 companies and their financial statement values and other variables from four years. The remaining 6 companies and their financial statement values and variables were divided into the training and the testing set so that the first three years were put into the training set and the fourth year's variables were left for testing.

The ANN models were used to make a *one-step ahead* prediction described in Section III, for training and testing. Therefore, the training data were presented to the network in a chronological order. Our models use two previous years to predict a third one. This required 32 input data streams to represent the data from each of the 16 account values. Four additional input neurons indicate the year of the input variable. The input value of this neuron was set to one if the neuron corresponded to the year of the data, and otherwise to zero. An additional input neuron was added to indicate the company. To summarise, the model has 37 input data streams in the input layer, and 16 neurons in the output layer, one for each account variable.

The following results are based on experiments to find (train) the network architecture that minimised the RMSE between the target account balances and the predicted output balances. The transfer function is the sigmoid function. We built (trained) the neural network model to monitor the dynamics of variables in such a way that it accepts 5 % difference between the target (= real, desired) and the output (= the value the networks give) values. We varied both layers and numbers per layer to find the best results. The training tolerance has no effect on the learning algorithm. However, when the Neuralyst finds 100% right, as defined by training tolerance, it will stop training. After training the ANN models were tested. The testing data set consists of the previous year's variables from six companies. The output values were compared to the target values. Table 2 summarises the best-achieved prediction capability of the models.

TABLE 2
Test statistics of the models

	Model A	Model Y	Model C	Model C&Y
Layers	5	3	6	4
Neurons per Layer	37-32-26-20-16	37-27-16	37-33-29-25-21-16	37-27-16
Training Epochs	3002	4091	3627	2593
RMSE	0.0048932	0.069367	0.132907	0.187114
Number of Items	96	96	96	96
Number Right	91	89	61	61
Number Wrong	5	7	35	35
Percent Right	**95 %**	93 %	64 %	64 %
Percent Wrong	5%	7 %	36 %	36%

The neural network output and the test target values are scored as "right" if they are within the testing tolerance, which was 10 %. From table 2 we can see that Model A received the best results with five layers. 91 data items out of 96 are within the 10% testing tolerance whilst others are not. Model Y was almost as good. Models C and C&Y attained much lower results.

Another way to compare the models is to calculate the quick ratio (QR) in the testing period. The QR shows the liquidity of the company and its financial assets divided by short-term liabilities. From table 3 we can see the calculated QRs in the testing period and that they support the test statistics of the models. Model A is best followed by Model Y.

TABLE 3
Quick Ratio in the test period

QR	**Right**	**A**	Y	C	C&Y
Company 26	**0.8**	**1.0**	**1.0**	0.5	0.1
Company 27	**1.1**	1.6	0.8	**1.0**	2.1
Company 28	**1.3**	**1.3**	**1.3**	2.2	2.7
Company 29	**1.5**	1.2	**1.5**	0.8	1.3
Company 30	**1.1**	**1.1**	1.2	2.6	1.8
Company 31	**1.2**	**1.1**	2.4	1.5	0.6

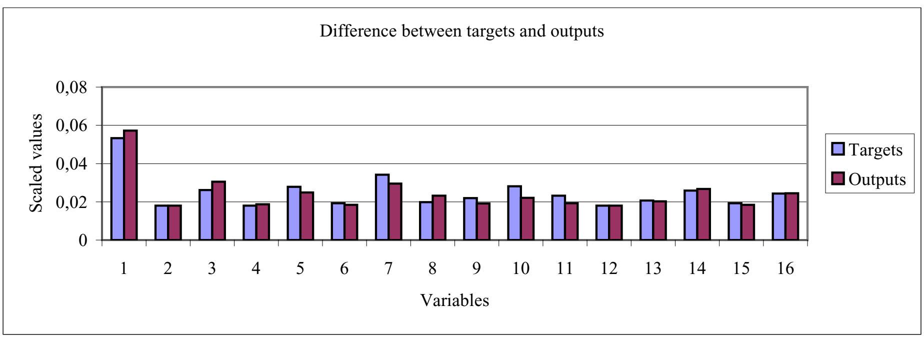


Fig. 4. Company 31, difference between target and output variables in the testing period

The columns in Fig. 4 show the difference between the target and the output values with Model A of Company 31's variables in the testing set. The difference values between targets and outputs are scaled and the number of the column equals the number of the variables in table 1.

The biggest differences are in:
(1) Net sales 7%
(3) Materials + change in inventory 17%
(5) Personnel costs (manufacturing) -11%
(7) Sales margin -14%
(8) Personnel costs (administration) 17%
(9) Other indirect costs -13%
(10) Operating margin -22%
(11) Profit -17%

The predicted output values of net sales (1) and materials + change in inventory (3) are bigger than the actual value.

However, these differences are in line. But if only the net sales were bigger an auditor might doubt whether all sales are recorded. The actual values of sales margin (7), operating margin (10), and profit (11) are bigger than the predicted output value. This indicates that the company attains better results than an average company in the same line of business. Personnel costs (manufacturing) (5) and other direct costs (6) also support this conclusion. However, personnel costs (administration) (8) do not support this conclusion.

We tested Model A by seeding fictitious sales into Company 31's data, which means that the net sales and receivables go up while the others remain at the initial level. The amount of the seeded error was 10 % of the average net sales. Fig. 5 shows that the target values of both net sales (1) and receivables (13) are clearly above the ANN output values.

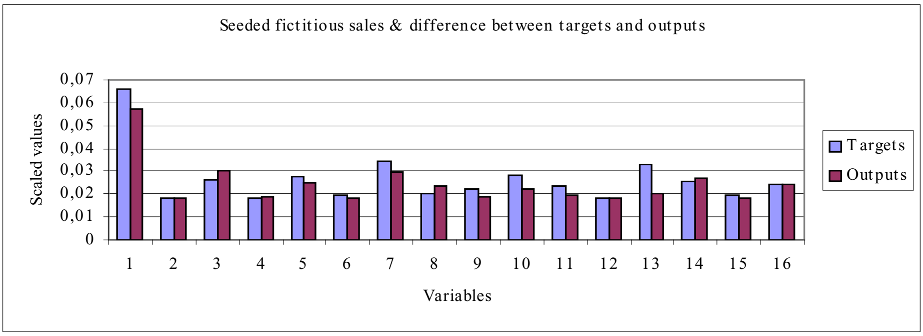


Fig. 5. Company 31, difference between target and output variables when seeding fictitious sales

V. Discussion and Further Research

We applied neural network technology to forecast and recognise patterns within one line of business when auditing financial accounts. The aim of this study is to get further evidence of the capability of an artificial neural network (ANN) to forecast and recognise patterns when auditing financial accounts. Therefore, the focus is on the development and evaluation of the ANN as a new audit method. We had both financial account values and other variables in our models. In particular, we focused on the impact of various pre-processing models on the forecast capability of an ANN when auditing financial accounts. The data were pre-processed in four different ways. Firstly, all the data were scaled linearly. Secondly, the data were pre-processed linearly on a yearly basis. Thirdly, the data were pre-processed linearly on a company basis. And fourthly, the data were pre-processed on a yearly and company basis. The best result was achieved when all the data were scaled linearly and linearly on a yearly basis. These two models give such encouraging results that it is worthwhile testing an ANN-based support system in a real audit situation. Before applying these kinds of methods in practice it has to be ensured that the results are stable enough. Comparing an ANN-model with other models such as regression analysis with the same data can do this.

However, pre-processing data on a company or on a yearly and company basis did not give as good results. One explanation for this might be that there were not enough data for learning the dynamics between variables. Another explanation might be that these pre-processing methods remove the existing dynamics and relationships between the variables within the business line. This has to be studied by using a larger data sample, i.e. data from more companies or by convincing that the sample data represent the whole population.

We think that neural network technology provides new opportunities for auditors and leads to improvements in the efficiency and effectiveness in auditing financial records, however, this has to be tested in further experiments. Moreover, the question arises how big the difference between the target value and output value might be without actions in the management or in the bookkeeping procedure. Furthermore, the question arises whether the neural network technology provides any improvement in the efficiency and effectiveness in auditing financial records. We plan to further apply the neural network technology to a real audit situation to get more results.

Acknowledgment

The author would like to thank Barbro Back for her comments and suggestions. Furthermore, the author would like to thank the company which provided the data for the study.

References

[1] I.A. Fraser, D.J. Hatherly, and K.Z. Lin, "An empirical investigation of the use of analytical review by external

auditors", *British Accounting Review,* vol. 29, pp. 35-47, 1997.
[2] E.C. Ameen and J.R. Strawser, "Investigating the use of analytical procedures: an update and extension", *Auditing: A Journal of Practice & Theory,* vol. 13, pp. 69-76, Fall 1994.
[3] S. Bonner, (1990), "Experience effects in auditing: The role of task-specific knowledge", *The Accounting Review,* vol. 65(1), pp. 72-92, 1990.
[4] S. Bonner and P.L. Walker, "The effects of instruction and experience on the acquisition of auditing knowledge", *The Accounting Review,* vol. 69(1), pp. 157-178, 1994.
[5] M.J. Fisher, ""Real-izing" the benefits of new technologies as a source of audit evidence an interpretative field study", *Accounting, Organizations and Society,* vol. 21(2/3), pp. 219-242, 1996.
[6] W.E. Halal, M.D. Kull, and A. Leffmann, "The George Washington University forecast of emerging technologies a continuous assessment of the technology revolution", *Technological Forecasting and Social Change,* vol. 59, pp. 89-110, 1998.
[7] B.K. Wong, T.A, Bodnovich, and Y. Selvi, "A bibliography of neural network business applications research: 1988 - September 1994", *Expert Systems,* vol. 12(3), pp. 253-262, 1995.
[8] B.K. Wong and Y. Selvi, "Neural network applications in finance: A review and analysis of literature (1990-1996)", *Information & Management*, vol. 34, pp. 129-139, 1998.
[9] J.R. Coakley and C.E. Brown, "Neural networks for financial ratio analysis", In *Proceedings of The World Congress on Expert Systems*, pp. 132-139, Pergamon Press, 1991.
[10] J.R. Coakley and C.E. Brown, "Neural networks applied to ratio analysis in the analytical review process", In *Proceedings of the Fourth International Symposium on Expert Systems in Accounting, Finance and Management*, pp. 1-36, Pasadena, CA, 1991.
[11] J.R. Coakley and C.E. Brown, "Artificial neural networks applied to ratio analysis in the analytical review process", *Intelligent Systems in Accounting, Finance and Management,* vol. 2, pp. 19-39, 1993.
[12] J.R. Coakley, "Using pattern analysis methods to supplement attention-directing analytical procedures", *Expert Systems with Applications,* vol. 9(4), pp. 513-528, 1995.
[13] R.C-F. Wu, "Integrating neurocomputing and auditing expertise", *Managerial Auditing Journal,* vol. 9(3), pp. 20-26, 1994.
[13] K.M. Fanning, K.O. Cogger, and R. Srivastava, "Detection of management fraud: A neural network approach", *International Journal of Intelligent Systems in Accounting, Finance and Management,* vol. 4(2), pp. 113-126, 1995.
[14] P. Brian, J. Green, and C. Hwa, "Assessing the risk of management fraud through neural network technology", *Auditing,* vol. 16(1), pp. 14-28, 1997.
[15] G.F. Klersey and M.T. Dugan, "Substantial doubt: using artificial neural networks to evaluate going concern", *Advances in Accounting Information Systems*, vol. 3, pp. 137-159, 1995.
[16] K. Tam and M. Kiang, "Managerial applications of neural networks: the case of bank failure predictions", *Management Science,* vol. 67(4), pp. 783-801, 1992.
[17] E. Koskivaara, B. Back, and K. Sere, "Modelling intelligent systems for auditing", *Intelligent Systems in Accounting and Finance,* ed. by Guillermo J. Sierra & Enrique Bonsón, Plaza de la Merced, Huelva, pp. 233-252, 1996.
[18] E. Koskivaara, "Artificial neural network models for predicting patterns in auditing monthly balances", In *Proceedings of the 5th European conference on information systems*, Cork-Ireland June 19-21, pp.196-209, 1997.
[19] S. Ramamoorti, A.D. Bailey Jr., and R.O. Traver, "Risk assessment in internal auditing: a neural network approach", *International Journal of Intelligent Systems in Accounting, Finance & Management,* vol. 8(3), pp. 159-180, 1999.
[20] R. Hecht-Nielsen, *Neurocomputing*, Addison-Wesley, Reading, MA, 1991.
[21] J. Hertz, A. Krogh, and R.G. Palmer, *Introduction to the theory of neurocomputing*, Addison-Wesley, Reading, MA, 1991.
[22] J.A. Freeman and D.M. Skapura, *Neural Networks: algorithms, applications and programming techniques.* Addison-Wesley Publishing Company, Inc., 1991.
[23] M. Smith, *Neural Networks for Statistical Modeling*, International Thompson Computer Press, Boston, 1996.
[24] K. Swingler, *Applying neural networks: A practical guide,* Academic Press, London, 1996.
[25] G. Zhang, B.E. Patuwo, and M.Y. Hu, "Forecasting with artificial neural networks: The state of the art", *International Journal of Forecasting,* vol. 14, pp. 35-62, 1998.
[26] J.E. Scholl and A.R. Venkatachalam, "A neural network approach to forecasting model selection", *Information & Management,* vol. 29, pp. 297-303, 1995.
[27] C. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford, 1995.
[28] W.R. Knechel, "A simulation study of the relative effectiveness of alternative analytical review procedures", *Decision Science,* vol. 17(3), pp. 376-394, 1986.
[29] Y. Shih, *Neuralyst™ User's Guide,* Cheshire Engineering Corporation, 1994.

PUBLICATION 5

Koskivaara – Back (2003) An Artificial Neural Network Based Decision Support System for Budgeting. In: *Proceedings of the Fifth International Conference on Enterprise Information Systems*, ed. by Oliver Camp – Joaquim Filipe – Slimane Hammoudi – Piattini Mario. Angers, France, April 23–26, 2003, 108–114. Reprinted by permission.

AN ARTIFICIAL NEURAL NETWORK BASED DECISION SUPPORT SYSTEM FOR BUDGETING

Eija Koskivaara
TUCS, Turku Centre for Computer Science and Turku School of Economics and Business Administration, Lemminkäisenkatu 14 A, 20520 Turku, Finland
Email: eija.koskivaara@tukkk.fi

Barbro Back
TUCS, Turku Centre for Computer Science and Åbo Academi University, Lemminkäisenkatu 14 B, 20520 Turku, Finland
Email: Barbro.Back@abo.fi

Keywords: Artificial neural networks, decision support system, budgeting

Abstract: This paper introduces an artificial neural network (ANN) based decision support system for budgeting. The proposed system estimates the future revenues and expenses of the organisation. We have built models based on four to six years' monthly account values of a big organisation. The monthly account values are regarded as a time-series and the target is to predict the following year's account values with the ANN. Thus, the ANN's output information is based on similar information on prior periods. The prediction results are compared to the actual account values and to the account values budgeted by the organisation. We also evaluated our method with four other methods and found that ANN outperformed the second best method in 19 out of 20 cases.

1 INTRODUCTION

Many parties are interested in the future financial performance of a company. Managers want to estimate the future revenues and expenses, in order to optimise the operations during a certain period. Auditors want to assess the accuracy of an organisation's financial statements. Creditors want to analyse the organisation's payment ability. All these parties may use budgets of the organisation to support their decisions and often they also want to evaluate the budgets. There are common ways to create a budget: "break down", "build up", and collaboration methods (Riistama and Jyrkkiö, 1975). In the "break down" method the management creates the budget. In the "build up" method all the responsible people prepare the budget. The collaboration method joins the best practices from the "break down" and the "build up" methods. In many organisations budgeting is a long, involved and inefficient process and consumes a huge amount of management time and effort. Despite careful preparation the budgets very often fail in today's rapidly changing business environment. Information technology has an important role to play in improving the efficiency of the budgeting process. Harvey (1999) has suggested enhancing managerial intelligence activities through information systems for budgeting. Mosmans *et al.* (2002) developed a multicriteria-based decision support system for a Belgian hospital. Their system allows the decision-makers several degrees of freedom in order to analyse the reality of consumption from a patient's point of view and to test a prospective scenario. From the theoretical point of view their system is based on fuzzy conjunctive aggregation operators (t-norms) weighted through fuzzy implications operators. Another possible enhancement would be an artificial neural network (ANN) approach, which according to our information has not been applied earlier to support the budgeting of monthly account values.

In this paper, we introduce an ANN-based decision support system for budgeting. With the proposed system one can either model the dynamics of the account values on a monthly basis or estimate the yearly account values. It also enables companies to adjust their budgets easily at any point during the year.

The rest of the paper is organised as follows: Section 2 describes the methodology and the model and the selection criteria for the accounts included in the study. Section 3 presents a detailed analysis of the results. Section 4 presents the general characteristics of the decision support system. The conclusions of the paper are presented in Section 5.

2 ANN DECISION SUPPORT MODEL

2.1 Prediction Model

We have used the supervised training method with the resilient backpropagation (Rprop) training algorithm. Backpropagation with sigmoid function has become the most popular one for prediction and classification problems (Sohl and Venkatachalam, 1995).

Learning parameters included in the model are *training cycles*, *weight decay*, *delta0* and *max delta* (Bishop, 1995; Sarle, 1995; Smith, 1996; SNNS, 1995; Swingler, 1996). Training cycles refer to the number of training runs needed to complete the task. Smaller values can give the prediction results faster, which can be advantageous sometimes. The weight decay, which is used by default, is very useful in the training, as it reduces overlearning and therefore increases the generalisation ability. Delta0 and max delta are parameters specific to resilient backpropagation. They are the initial step size and the maximum step size, respectively. The optimal parameters' values vary depending on the data, the amount of training cycles, and the network topology.

The network topology used in our model is a feed-forward multilayer perceptron topology with the sigmoid function. Our model can have a maximum of three hidden layers. The specific network topology is determined by a number of choices. Our model has two options regarding the choice of input variables given to the network. With the single variable options, each monthly account value is predicted from the previous months' values of only that account. With this option, it is possible to predict only that single variable, and therefore there can be only one output variable. With the all variables option as inputs, each account is predicted from the values of all accounts in previous months. The output variables can be done by either one account at a time or all accounts at the same time. In the former case a separate ANN will be trained for each account. In the latter case, there is a single ANN with all the accounts in the output layer. The number of input months may vary from one to twelve. A month indicator indicates the number of previous months given as inputs.

The monthly account values of the organisation were regarded as a time-series and the aim was to recognise the dynamics and the relationships between the accounts. The goal was to predict the value of a certain account for a short time in the future. The feed-forward network can be applied directly to such problems (Bishop, 1995).

We take a set of account values from a number of periods, for example four successive periods x_{t-2}, x_{t-1}, and x_t to be the input vector to a network and use the next value x_{t+1} as the target value as illustrated in Figure 1. This is called one-step-ahead prediction. In the multi-step ahead prediction the outputs of the neural network are cycled around to the inputs of the network, then predictions can be made for further points. By moving along the time axis we can create a training data set consisting of many sets of input vector values with corresponding target values (Bishop, 1995).

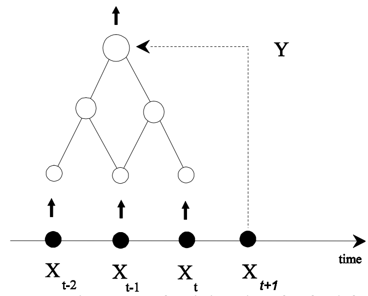


Figure1: Generating a set of training data for feed-forward ANN (adapted form Bishop, 1995).

2.2 Choice of sample and data

For the study, we have selected ten units from one big municipality with the help of internal auditors. We have six units from the Health Care Services (HS) and one from the Social Welfare (SW). We also have account values from the Office Facilities Administration (OFA) and from the Port of Town (PT). We have collected five to seven years' monthly account values from these units. The year 2001 was selected as the testing period and all the previous years were used for training the ANNs. In the raw data there were hundreds of accounts per unit. In this preliminary experiment we have not used all accounts. We only used those accounts that had entries in all the training and testing years. We first ended up with a range of 29 to 157 accounts per unit.

Table 1 - Number of accounts per models.

		Number of accounts in		
Model	Years	Total	Model	Budgeted
HS1-A	1995-2001	157	10	10
HS1-B			10	10
HS2-A	1995-2001	64	10	10
HS2-B			10	10
HS3-A	1995-2001	47	10	10
HS3-B			10	9
HS4-A	1995-2001	56	10	10
HS4-B			10	10
HS5-A	1995-2001	76	10	9
HS5-B			10	10
HS6-A	1995-2000	29	10	9
HS6-B			10	10
SW-A	1995-2001	133	9	9
SW-B			10	10
OFA1-A	1995-2001	48	10	10
OFA1-B			10	9
OFA2-A	1997-2001	136	10	3
OFA2-B			14	0
PT-A	1997-2001	94	10	10
PT-B			10	10

For every unit we then built two different kinds of models based on different account selection criteria. We decided to have about ten accounts per model in order to keep it user-friendly. First, we selected those accounts per unit that were most stable compared to previous year's values. In most of the cases these accounts were the largest and most significant ones like salaries and salary-related accounts. We called these models A-models. Secondly, we let the internal auditor select those accounts which were the most interesting from the auditing point of view. We call these models B-models. With these selection criteria we received twenty models.

Our data file contains a number of monthly account values from consecutive time periods. The last 12 months of the data are used as test data, and the rest is used as training data. The training data is used to predict the account values for the most recent period. The prediction results are compared to the actual account values and to the account values budgeted by the organisation.

The data in our model was pre-processed with linear equalisation, which maps the account values linearly to range [0,1]. It is possible to equalise the values either one account at a time or all accounts at the same time (globally). If we equalise only one account at a time, the value 0 will refer to the smallest value in the particular account, and 1 to the largest. If we equalise all the accounts at the same time, 0 and 1 will refer to the minimum and the maximum of the entire data.

In the evaluation of the models we have used four different kinds of comparison methods for ANN's prediction results: Previous year's value (p.year's) is the same value from the previous year. Average of previous years (ave of p.) is the average of the same account values from all the previous years. Average delta prediction (ave delta pre) calculates an average of the monthly changes from previous years, and makes predictions by adding the change to the account value of the previous month. Zero delta prediction (z delta pre) does not predict anything, the values are the same as in the previous month. Combined trivial prediction (ctm) combines the above mentioned simple prediction methods. These comparison metrics, as well as comparison with budgets and forecasts, are mainly based on analytical review procedures that are used in auditing (Gauntt and Gletzen, 1997).

For evaluating purposes we collected the account values of year 2001 budgeted by the organisation. We noticed that some budgeted values were missing in the system. Table 1 shows the years and number of accounts per units and models.

2.3 ANNs for organisation units

Table 2 - Error in currency.

Model	ANN error	Next best error	Name of the method
HS1-A	**1138**	1269	ctm
HS1-B	**3339**	3527	ctm
HS2-A	**2122**	2170	ctm
HS2-B	**2747**	2944	p. year's
HS3-A	**1061**	1201	ctm
HS3-B	**1001**	1027	ctm
HS4-A	**1434**	1471	ctm
HS4-B	277	**268**	**ave of p.**
HS5-A	6966	**6532**	**ctm**
HS5-B	**7452**	8640	ave of p.
HS6-A	**2504**	2581	p. year's
HS6-B	**2085**	2103	ave of p.
SW-A	**43880**	46858	ctm
SW-B	**22844**	25417	ctm
OFA1-A	164569	**155803**	**z delta pre**
OFA1-B	**159156**	165129	ave of p.
OFA2-A	**154646**	170119	ctm
OFA2-B	**79382**	80702	p. year's
PT-A	**127367**	160502	p. year's
PT-B	**49444**	94610	p. year's

We performed several tests in order to determine suitable parameters and topologies for every model. We compared the ANN's results with the previously mentioned comparison methods (previous year's value [p. year's], average of previous years [ave of p.], average delta prediction [ave delta pre], zero delta prediction [z delta pre], combined trivial prediction [ctm]). Only in three cases, models HS4-B, HS5-A, and OFA1-A, the ANN was the second best method as to errors in currency, otherwise ANN-models were the best. However, the results of these three ANN-models were very close to the best (ave of p., ctm, z delta pre) methods. In the other cases the average errors of the next best method were also close to the ANN error values. However, the second best method is not the same for every model, although ctm is the second best model most frequently - ten times. Table 2 presents the average error in euros of the ANN and the second best or best (in bold) method. We have presented the error of the methods in currency, because we think that it is more valuable for the decision-makers. The error per model shows the average error over the 12 predicted months for all the accounts in the model. The amount of the error depends on the values of accounts included in the model. The bigger the account values the bigger the errors. Therefore, one cannot compare the ANN models against each other by comparing these error amounts.

Table 3 - Training parameters.

Model	Training cycles	Weight decay	Delta0	Max delta
HS1-A	500	0.9991	0.1	50
HS1-B	1000	1	1	50
HS2-A	2000	0.98	0.09	50
HS2-B	1000	0.99	0.09	50
HS3-A	1000	0.9991	0.1	50
HS3-B	1000	0.99	0.1	50
HS4-A	5000	0.98	0.1	50
HS4-B	1000	0.9	0.1	50
HS5-A	1000	0.95	0.09	50
HS5-B	1000	0.99	0.09	50
HS6-A	500	0.9992	0.1	50
HS6-B	1000	1	0.11	50
SW-A	1000	0.9	0.2	50
SW-B	1000	0.9	0.2	50
OFA1-A	20000	0.95	0.1	50
OFA1-B	1000	0.95	0.2	50
OFA2-A	500	0.9992	0.1	50
OFA2-B	1000	0.9	0.2	50
PT-A	500	0.9991	0.1	50
PT-B	1000	0.84	0.3	50

The achieved optimal network training parameters per model are presented in Table 3. The training cycles per model differ from 500 to 20 000, however, the best result was most often achieved with 1000 training cycles. The value of the weight decay was 19 times between 0.9 and 1, it was below 0.9 i.e. 0.84 only once. The delta0 values were between 0.09 and 0.3. In general it seems that the smaller the weight decay the bigger the delta0. Max delta was 50 in all the cases.

The network topologies per models are presented in Table 4. In 19 out of 20 the models have all the variables as inputs and outputs. Only once, in HS3-A, the topology was based on the single variable option as input and output. This strengthens our assumption that the network was able to capture the dynamic and some relationships between the account values. There are two general models for the month indicator input, either one previous month's data is given as inputs (14/20) or twelve previous months' data is given as inputs (5/20). Only once, in HS3-A, the number of previous months is something else, namely two. In 19 out of 20 the models have one hidden layer. Three models, HS1-B, HS3-A and HS6-B, were equalised globally, while others were equalised one account at a time.

Table 4 - Neural networks topologies.

Model	Inputs	Outputs	Input months	Neurons in layers
HS1-A	all	All	1	22-10-10
HS1-B	all	All	1	22-10-10
HS2-A	all	All	1	22-20-20-20-10
HS2-B	all	All	1	22-10-10
HS3-A	single	One	2	14-6-1
HS3-B	all	All	1	22-10-10
HS4-A	all	All	1	22-10-10
HS4-B	all	All	1	22-10-10
HS5-A	all	All	12	132-10-10
HS5-B	all	All	1	22-10-10
HS6-A	all	All	1	22-10-10
HS6-B	all	All	1	22-10-10
SW-A	all	All	12	120-9-9
SW-B	all	All	12	132-10-10
OFA1-A	all	All	12	132-10-10
OFA1-B	all	All	1	22-10-10
OFA2-A	all	All	1	22-10-10
OFA2-B	all	All	1	26-26-10
PT-A	all	All	12	132-10-10
PT-B	all	All	1	22-44-10

3 ANALYSIS OF THE ANN OUTPUT DATA AND BUDGETING DATA WITH ACTUAL DATA

In this section we compare the results of the above mentioned ANN models and the account values budgeted by the organisation to the actual account values. We have collected the yearly budget values and counted the yearly actual values from the bookkeeping system. We counted the yearly ANN values of the accounts included in the models. These values are presented in Table 5. We also counted the average per cent and euro errors of the ANN and budget values. These values are presented in Table 6.

Table 5 - The yearly budget, actual and ANN values of the units.

Model	Budget	Actual	ANN
HS1-A	-184 090	-195 840	-191 866
HS1-B	-116 454	-118 403	-99 992
HS2-A	-240 230	-223 664	-208 545
HS2-B	-97 465	-92 761	-80 608
HS3-A	-41 283	-42 718	-42 662
HS3-B	-17 929	-26 496	-21 452
HS4-A	-154 095	-147 218	-143 134
HS4-B	-6 198	-5 607	-4 164
HS5-A	-269 553	-258 522	-213 753
HS5-B	-51 589	-65 291	-56 465
HS6-A	-183 491	-174 973	-176 986
HS6-B	-136 647	-132 842	-130 980
SW-A	-4 092 573	-4 411 907	-4 359 886
SW-B	-1 796 247	-1 941 962	-1 987 871
OFA1-A	4 659 310	4 878 824	5 396 956
OFA1-B	-2 948 503	-2 888 307	-2 590 945
OFA2-A	946 898	-3 489 975	-2 583 310
OFA2-B	0	-1 219 423	-1 036 741
PT-A	2 913 546	2 739 970	2 254 394
PT-B	176 967	226 548	131 724

On average the ANN method was better than the budgeting method. The average error of the ANN method was 12% and the average error of the organisation's budget was 19%. The ANN values were closer to the actual values eleven times whereas the organisation's budget values were closer to the actual values eight times. Once they were equally good. If we think that anywhere between plus or minus 10% may be sufficient for decision purposes (Wheelwright and Makridakis, 1977), then the values budgeted by the organisation were better. The account values budgeted by the units of the organisation were 14 times within this 10 % marginal whereas the ANN was nine times within this marginal (columns two and three in Table 6). Next we compared the average difference between the A-models and the B-models. With the A-models the ANN was better seven out of ten times and with the B-models ANN was better five times and the budget was better five times. So on average ANN was better in those cases where the account values were more stable. On average both models were equal in those cases which were more interesting from the auditing point of view.

Table 6 - The average % errors and euro errors

Model	Bud/Act	Ann/Act	Bud/Act	Ann/Act
HS1-A	6 %	2 %	-11 750	-3 974
HS1-B	2 %	16 %	-1 949	-18 412
HS2-A	7 %	7 %	16 566	-15 119
HS2-B	5 %	13 %	4 704	-12 154
HS3-A	3 %	0 %	-1 434	-56
HS3-B	32 %	19 %	-8 567	-5 044
HS4-A	5 %	3 %	6 877	-4 084
HS4-B	11 %	26 %	591	-1 443
HS5-A	4 %	17 %	11 031	-44 770
HS5-B	21 %	14 %	-13 702	-8 826
HS6-A	5 %	1 %	8 517	2 013
HS6-B	3 %	1 %	3 806	-1 862
SW-A	7 %	1 %	-319 334	-52 021
SW-B	8 %	2 %	-145 715	45 909
OFA1-A	4 %	11 %	219 514	-518 132
OFA1-B	2 %	10 %	60 196	-297 362
OFA2-A	127 %	26 %	-4 436 873	-906 664
OFA2-B	100 %	15 %	-1 219 423	-182 682
PT-A	6 %	18 %	-173 576	485 576
PT-B	22 %	42 %	49 581	94 824
Average	19 %	12 %	-297 547	-72 214

For further analysis we compared the values of individual accounts of the units. We counted how many individual account values of the ANN and the budget were inside a 10% marginal. The budgeted values were 96 times inside this marginal whereas the ANN values were 82 times inside this marginal. When comparing the A-models and B-models the budget account values were inside the 10% marginal 55 times and the ANN account values were 52 times within this marginal. If we look at the individual accounts the budget by the organisation was slightly better. With the B-models this difference was 41 to 30 in favour of the budget by the organisation. So in the individual account level the budget by the organisation was better than the ANN method. However, these figures were a little bit biased because not all accounts had a budget value at all. 25 out of 203 accounts did not have a budget value in the system. Table 7 shows the number of the individual account per units, whose values are inside the 10% marginal.

Table 7 - Number of accounts per models in 10% marginal.

Model	Budget	ANN
HS1-A	10	10
HS1-B	3	1
HS2-A	5	9
HS2-B	2	3
HS3-A	7	6
HS3-B	5	3
HS4-A	4	7
HS4-B	3	0
HS5-A	9	9
HS5-B	7	6
HS6-A	6	5
HS6-B	5	3
SW- A	5	5
SW-B	8	9
OFA1-A	3	1
OFA1-B	3	1
OFA2-A	0	0
OFA2-B	0	0
PT-A	6	0
PT-B	5	4
Total	96	82

The worst units with the ANN are OFA and PT. With the models from OFA and PT the ANN did not learn the dynamics and the trends as well as in the HS as SW units. There are some explanations for the ANN's bad performance in these cases. The training data of OFA2- and PT-models consisted of four-year data whereas in the other models there were either five or six years' training data sets. Also the organisational changes in the OFA-units might have influences in the data. The data in the PT-model was entered in the system on a monthly basis only in the last training year, which is all too short a time for the network to learn the dynamics and relationships in the data. We conclude that there are not enough training data for the PT- and the OFA-units. If we skip these units both the budgeted and the ANN values are on average equally close to the actual values.

The extra power of the ANN model described in this study lies in it's ability to model the monthly dynamics of account values. The ANN model works best with accounts with a trend and with some causality to each other. These kinds of accounts in our models are salaries and salary-related accounts, as well as the use of materials in some cases. The ANN model does not work very well with those accounts which are mainly based on occasional decisions. However, sometimes the yearly ANN value of an account like this might be quite close to the actual value. In Figure 2 we show the dynamics the ANN model has learnt. This is a salary account from the HS1-A unit. The actual yearly value of this account is 127 416,26 euros and the budgeted value is 121 095,31 euros, the ANN´s yearly value is 125 663,80 euros. Both budgeted and ANN values are within the 10% marginal. However, the ANN model also predicts the monthly variation of the salary expenses. For accounts that have a trend or causality this kind of model adds values in enabling to allocate money into the right places at the right time.

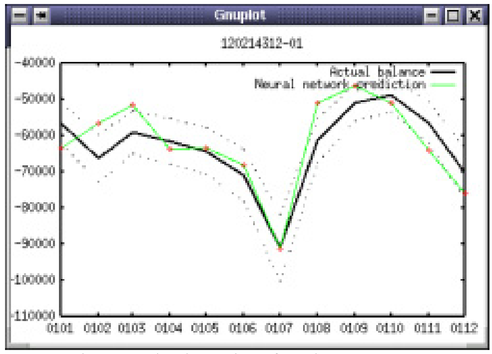


Figure 2:The dynamics of a salary account.

4 THE NEW DECISION SUPPORT SYSTEM

Our DSS has been developed and tested in Linux environment. It works currently under Linux with KDE1 desktop environment. It has four main windows: a file, a data analysis, an option, and a help window. It is possible to open the loaded data files from the file window. The predictions with the trivial methods are made immediately when a data file has been loaded. These predictions are displayed in the data analysis window. This data analysis window allows either viewing the error statistics or plotting the results. In the option window are the settings for the ANN training. Once the setting for the model is determined it is possible to train the ANN model from the data analysis window. After the training is finished it is possible to view the ANN's prediction results together with the trivial methods in the data analysis window. After the network has been trained with a particular data set, it can be saved as a file. A previously saved ANN can be restored from the file window. From the help window one can find the manual for the DSS.

5 CONCLUSION

In this paper we have proposed an ANN-based decision support system for budgeting. The proposed system estimates the future revenues and expenses of an organisation. The ANN-model is trained and tested with the monthly account values of a big organisation. These monthly account values are regarded as a time-series and the target was to predict the account values for the most recent period with the ANN. Thus, the ANN's output information was based on similar information from prior periods. Our model uses the one-step-ahead prediction model. The network topology used in the model was a feed-forward multilayer perceptron with sigmoid function. In the training of the models we compared ANN's prediction results with the previous year's values, average values of previous years, average delta prediction, zero delta prediction, and with combined trivial prediction. On average the ANN was the best method. With the training parameters it seems that in general the smaller the weight decay the bigger the delta0. In 19 out of 20 models we have all the variables as inputs and outputs, which indicates that there is a dynamic and relationships between the account values. We found two general models for the month indicator input, either one previous month is given as inputs or twelve previous months are given as inputs. In most cases the accounts were equalised one account at a time.

We compared our prediction results and the organisation's budgeted account values to the actual account values. On average it seems that ANN was better than the organisation's method. The average error of the ANN method was 12% and the average error of the budget was 19%.

The development of a decision support system (DSS) arises from the necessity to manage an organisation's internal operations better. For instance, budgets are usually adjusted either annually or quarterly based on income, expenses, and other factors. The ANN- model can be used for either predicting the yearly account values or even more specifically for modeling the dynamics of the account values on a monthly basis. This feature makes it possible for companies to adjust their budgets at any point during the year. Furthermore, it helps managers to estimate the future revenues and expenses in order to optimise the operations during the following periods. At the moment the system is at the prototype stage. It is worth noticing that a team of internal auditors and possible future users of municipal organisation supervise the work.

ACKNOWLEDGEMENTS

The authors would like to thank Marko Grönroos for programming the DSS. The authors are grateful for the municipality that provided the data and the supervision of internal auditors and users for the study. The authors wish to thank the Academy for Finland for providing financial support for this project.

REFERENCES

Bishop, C. M., 1995. *Neural Networks for Pattern Recognition*. Oxford: Clarendon Press.

Gauntt, J. E. Jr. and Gletzen, G. M., 1997. "Analytical Auditing Procedures." *Internal Auditor*,:February, pp. 56-60.

Harvey, J., 1999. Supporting managerial intelligence tactics through information technology. *International Journal of Technology Management*, 17:1,2, pp. 56-72.

Mosmans, A., Praet, J-C, and Dumont, C., 2002. A decision support system for the budgeting of the Belgian health care system. *European Journal of Operational Research*, 139:2, June 1, pp. 449-60.

Riistama, V. and Jyrkkiö, E., 1975. *Operatiivinen laskentatoimi*. Helsinki: Weilin+Göös.

Sarle, W. A., 1995. Stopped training and other remedies for overfitting. *27th Symposium on the Interface of Computing Science and Statistics*: 352-60.

Smith, M., 1996. *Neural Networks for Statistical Modeling*: International Thompson Computer Press.

SNNS. 1995. SNNS, Stuttgart Neural Network Simulator, User Manual, Version 4.1. Institute for Parallel and Distributed High Performance Systems (IPVR), University of Stuttgart: Stuttgart.

Sohl, J. E., and Venkatachalam, A. R., 1995. A neural network approach to forecasting model selection. *Information & Management*, 29, pp. 297-303.

Swingler, K., 1996. *Applying Neural Networks A Practical Guide*. London: Academic Press Limited.

Wheelwright, S. C. and Makridakis, S., 1977. *Forecasting methods for Management*. New York: Wiley-Interscience.

PUBLICATION 6

Koskivaara (2004c) Comparing Different Methods in Analytical Procedures Using Real World Data. In: *Proceedings of the IRMA 2004, Information Resources Management Association*, ed. by Mehdi Khosrow-Pour. New Orleans Louisiana, USA, May 23–26, 2004, 661–663.

Comparing Different Methods in Analytical Procedures Using Real World Data

Eija Koskivaara
Turku School of Economics and Business Administration, , Information System Science, Rehtorinpellonkatu 3, 20500 Turku, Finland,
Telephone: +358-2-48 14 457, Fax: +358-2-48 14 453, Email: eija.koskivaara@tukkk.fi

ABSTRACT

Analytical procedures play an important role in assisting the auditor in determining the nature, timing and extent of their substantive testing, and in forming an overall opinion as to the reasonableness of recorded account values. The present study compares the artificial neural network (ANN) system and traditional analytical procedures on pattern recognition in monthly account values. The results of the study indicate that the ANN-system has a better predictive ability on pattern recognition in monthly account values than the traditional analytical procedures used in this study.

1 INTRODUCTION

The demands in the auditing environment have led to the publication of several standards on analytical procedures (APs) in different countries (e.g. AICPA, 1988; APB, 1995; KHT-yhditys, 2003). The US standard (SAS 56) on AP[1] has generally, in the literature, been considered as an authoritative pronouncement to AP (AICPA, 1988). The emphasis in the SAS 56 definition is on expectations developed by the auditor. It states as follows:

Analytical procedures involve comparing of recorded amounts, or ratios developed from recorded amounts, to expectations developed by the auditor. The auditor develops such expectations by identifying and using plausible relationships that are reasonably expected to exist based on the auditor's understanding of the client and of the industry in which the client operates.

SAS 96 contains amendments adding specific documentation requirements to the SAS no. 56, which at the moment requires auditors to document the factors they considered in developing the expectation for a substantive analytical procedure (AICPA, 2002). Besides, they have to document the expectation if it is not apparent from the documentation of the work that they performed. The auditors also should document (a) the results of their comparison of that expectation to the recorded amounts or ratios that they developed from recorded amounts, and (b) any additional auditing procedures they performed in response to significant unexpected differences arising from the APs, as well as the results of such additional procedures.

APs may be performed:

- in the *client acceptance/retention stage* in order to assist in obtaining a better understanding of the client's business
- in the *audit planning stage* to identify possible problem areas
- in the *substantive testing stage* as a means of gathering substantive evidence in relation to one or more account balances or classes of transactions
- in the *opinion formulation stage*, as a means of gathering evidence as to the consistency of the financial statements with the auditor's knowledge of the business

Auditing researchers have developed a variety of models to assist in the APs Techniques included in these models range from simple comparisons to complex analyses (e.g. Leitch and Chen, 2003; Blocher, et al. 2002; Fleming, 2004). Fraser, Hatherly, and Lin (1997) have identified three types of AP techniques: non-quantitative (NQT) or *judgmental,* such as scanning; simple quantitative (SQT) such as trend, ratio and reasonableness tests; and advanced quantitative (AQT), such as regression analysis. These techniques differ significantly in their ability to identify potential misstatement. Judgmental techniques include auditor's subjective evaluations based on client knowledge and past experience. Trend analysis assesses whether there is a functional relationship between the variables over time. Ratio analysis incorporates the expected relationships between two or more accounts directly. For example, turnover ratios are useful because there is typically a stable relationship between sales and other financial statement accounts, especially receivables and inventory. Although ratios are easy to compute, which in part explains their wide appeal, their interpretation is problematic, especially when two or more ratios provide conflicting signals. Indeed, ratio analysis is often criticized on the grounds of subjectivity, i.e. the auditor must pick and choose ratios in order to assess the overall performance of a client. In a reasonableness test, the expected value is determined by reference to data partly or wholly independent of the accounting information system, and for that reason, evidence obtained through the application of such a test may be more reliable than evidence gathered using only an accounting information system. For example, the reasonableness of the total annual revenue of a freight company may be estimated by calculating the total tons carried during the year and the average freight rate per ton. With a regression analysis model the auditor may predict financial and operating data by incorporating e.g. economic and environmental factors into the model. Sad to say, many of the AQT-based models have not found their way into practice. Most of the AP models used in practice for signaling errors throughout the audit include relatively simple techniques and are not based on any statistical methods (Ameen and Strawser, 1994; Cho and Lew, 2000; Fraser, Hatherly, and Lin, 1997; Lin, Fraser, and Hatherly, 2003).

Needless to say, there is a need for better AP tools and this study argues that ANNs (artificial neural networks) could be a feasible technique to aid auditors in creating expectations, and these expectations can then be compared to actual values automatically (cf. SAS 56). ANNs have many beneficial aspects in comparison to other techniques. They are adaptive tools for processing data. They can learn, remember, and compare complex patterns (Medsker and Liebowitz, 1994). They are useful for recognition of patterns from noisy data and they are able to dynamically adapt to a changing environment (Dutta, 1993). Basically, ANNs learn from examples and then generalize the learning to new observations. Compared to regression analysis we do not need an a priori model because ANNs are data driven. In addition, ANNs are, unlike traditional statistical techniques, capable of identifying and simulating non-linear relationships in the data without any a priori assumptions about the distribution properties of the data. One advantage of the ANN-systems could be that they provide additional information to the decision process. With the help of an ANN an auditor may find something from the data more effectively and efficiently than with conventional APs.

Auditing ANN research started a little more than a decade ago (Koskivaara, 2004). The main ANN-application areas in auditing are detecting material errors, detecting management fraud, and supporting going concern decision. ANNs have also been applied to internal control risk assessment, to determination of the audit fee, and to financial distress problems. Going concern and financial distress are very close or can even be included in bankruptcy studies. All these above mentioned researches fit into APs. Most of the researchers state that the ANNs have the potential to improve APs. This research can be classified under detecting material error, especially illustrating monthly account values, applications. The paper proceeds as follows: Section 2 identifies the research settings; Section 3 presents the results; Section 4 discusses the findings.

2 RESEARCH SETTINGS

2.1 Sample and data selection

For the study, we have selected three units from a big organization with the help of their executive management. Although it is doubtful from the auditing point of view to let the management of the organization select those accounts that should be audited it is acceptable in a research setting. The management of the organization should know which accounts follow the trends best and are related to each other. Besides, in principle, all the accounts should be audited in one way or another. Therefore, this selection basis might give an idea to the auditor of which accounts would be the most suitable for ANN-assisted auditing. Indeed, researchers have also proposed APs to management accounting for controlling operations (Lee and Colbert, 1997; Colbert, 1994).

We collected eight years' total monthly costs of theses units. The data for year 2002 was held out for testing and all the previous years' data were used for training the ANN. All the values had been audited, and therefore, in theory, they should have been correct.

2.2. ANN- model

The ANN-model uses the supervised training method with the *resilient backpropagation* (RPROP) training algorithm with sigmoid function, which is one of the most efficient algorithms for pattern recognition problems (Demuth and Beale, 2000). The forming expectations for account values can be classified into the pattern recognition problem. The purpose of the RPROP training algorithm is to eliminate the effects of the magnitudes of the partial derivatives (Riedmiller, 1994; Riedmiller and Braun, 1993). Only the sign of the derivative is used to determine the direction of the weight update; the magnitude of the derivative has no effect on the weight update. The size of the weight change is determined by a separate update value. The update value for each weight and bias is increased whenever the derivative of the performance function with respect to that weight has the same sign for two successive iterations. The update value is decreased whenever the derivative with respect to that weight changes sign from the previous iteration. If the derivative is zero, then the update value remains the same. Whenever the weights are oscillating the weight change will be reduced. If the weight continues to change in the same direction for several iterations, then the magnitude of the weight change will be increased. There are two more reasons for selecting RPROP for the learning algorithm for the prototype. First, the performance of RPROP is not very sensitive to the settings of the training parameters. Second, RPROP uses a batch training algorithm and is therefore efficient and requires minimal storage. In the batch mode of learning weight updating is performed after the presentation of all the training examples that constitute an epoch. Besides, as the patterns in the system are presented to the network in a time series manner the use of holistic updating of weights makes the search in weight space stochastic in nature. This in turn makes it less likely for the network to be trapped in a local minimum.

Training parameters included in the system are *training cycles*, *weight decay*, *delta0* and *max delta*. Training cycles refer to the number of training runs needed to complete the task. Smaller values can give the prediction results faster, which can sometimes be advantageous. The weight decay, which is used by default, is very useful in the training, as it reduces overlearning and therefore increases the generalization ability. Delta0 and max delta are parameters specific to RPROP. They are the initial step size and the maximum step size, respectively. The optimal parameter values vary depending on the data, the amount of training cycles, and the network architecture.

The achieved optimal network training parameters for the ANN-system are as follows. The training cycles of the ANN-system were 1000, this is not a high value, but neither is the amount of the data in the model. The value of the weight decay was 0.99999 and the delta0 values were 0.1. These indicate the fluctuating nature of the data. Max delta was 50 in all the cases. The construction of the ANN-system is presented more thoroughly in Koskivaara and Back (2003).

The network architecture of the ANN-system of the present study is as follows. The ANN-system has multiple variables as inputs and outputs (i.e. MIMO-system). Additional 12 input neurons indicate the month of the input data. One previous month was given as input. The ANN-architecture has one hidden layer with eight neurons. To summarize, the ANN-system has 15 neurons in the input layer, eight neurons in the hidden layer, and three neurons in the output layer. The data in the ANN-system was equalized with linear scaling of all accounts at a time. Linear scaling has the advantage of preserving the relative position of each data point along the range. This means that with the linear scaling the original and normalized values are one-to-one. Therefore, this scaling does not move any relevant information from the data before it is fed to the ANN.

2.3 Development of comparison metrics and hypothesis

To study (with *t* test on paired differences) whether the ANN-system recognize patterns in monthly account values more accurately than traditional APs in the real data environment, requires the selection of the traditional APs and an assumption that the population of differences between the pairs of observations is normally distributed. This was done.

As mentioned in the introduction, research has consistently indicated that auditors prefer simple scanning, reasonableness tests, and ratio analysis to sophisticated statistical or mathematical models in APs. This is true also in the organization that provided us the data for testing the ANN-based system. One explanation for the popularity of simple techniques might be that they are quite straightforward and require only few calculations. As simple techniques play a fundamental role in the APs in practice, they are selected as the basis for the comparison metrics for the ANN-based system. These comparison metrics are based on the guideline for APs in auditing given by Gauntt and Gletzen (1997). The abbreviations in brackets and explanations of the comparison metrics (= traditional APs) used in this study are as follows:

- Previous year's value (PYS) is the same value from the previous year.
- Average of previous years (AVE) is the average of the same account values from all the previous years.
- Average delta prediction (DELTA) calculates an average of the monthly changes from previous years, and makes predictions by adding the change to the account value of the previous month.
- Zero delta prediction (ZERO) values are the same as in the previous month.
- Combined trivial prediction (CTM) combines the above mentioned simple prediction models.

The predictions of these comparison metrics are made immediately when a data file has been loaded in the system. In effect, we have five populations, one population associated with each method (i.e. ANN, CTM, DELTA, AVE, PYS, ZERO). The following hypothesis will be tested:

H_0: *There is no significant difference between the prediction accuracy in the monthly account values of the ANN-system and the traditional APs.*

If H_0 cannot be rejected, we will not have evidence to conclude that the ANN-system differs from the traditional APs. However, if H_0 can be rejected, we will conclude that the ANN-system differs from the traditional APs.

3 RESULTS

First we measured the accuracy of the ANN-system and the traditional APs in their expectations forming capability by comparing

Table 1. Means and standard deviations of the comparison methods

	ANN	CTM	DELTA	AVE	PYS	ZERO
MEAN	**279540**	396297	352831	752253	358826	365795
STANDARD DEVIATION	**214154**	271171	273852	413668	294126	301408

Table 2. Prediction accuracy of ANN-system vs APs

Paired Samples Test	t-value	df	Sig. (2-tailed)
ANN – CTM	-2,448	35	**0,020**
ANN – DELTA	-2,125	35	**0,041**
ANN – AVE	-6,416	35	**0,000**
ANN – PYS	-1,651	35	0,108
ANN – ZERO	-2,293	35	**0,028**

their mean errors and their standard deviations in the holdout sample. Table 1 presents these values of all the methods used in this study. The ANN-system has the lowest average error in money (= the average difference between the actual account value and the account values achieved with some method such as ANN) and the lowest standard deviation of these average errors. The lowest average currency error indicates that the ANN-system has the best average prediction accuracy. The lowest standard deviation illustrates that the ANN-values are less dispersed from the average value. So, the ANN-system forms expectations for monthly account values more consistently than the other methods in the study.

Paired samples t-tests were conducted to assess the significance of the differences of the methods, this was possible, because the population between the differences between the pairs of observations was normally distributed. As shown in Table 2, there is a significant difference between the ANN-model and the traditional APs at the 0.05 level in four (in bold) out of five cases. Therefore, H_0 can be rejected in these cases. There was no significant difference between the account values achieved with the ANN-system and with the PYS-method. Although the results are encouraging, it should also be emphasized that the number of variables, which were inputs and outputs in the system, were limited.

4 DISCUSSION

The Enron incident and other such scandals also highlight the use and efficiency of APs. Recent research of Lin, Fraser, and Hatherly (2003) conducted in Canada indicates that APs are extensively applied in practice, particularly by larger audit firms, and that their use dominates the completion phase of audit regardless of the firm size. Their results are comparable with those from research conducted in the US (Ameen and Strawser, 1994). One explanation for the greater use of APs by larger audit firms is the client size. Larger clients are more likely to have strong internal controls that facilitate the reliance of accounting data and support documents and data for using APs. An important part of using APs is to select the most appropriate procedures.

Despite investments by larger firms in technology and in audit automation, audit firms of all sizes continue to emphasize judgment-based procedures as compared with those which are more quantitative based. However, the development of IT supports systems makes the use of advanced methods easier and more cost effective. Therefore, auditors must keep pace with the emerging IT changes and their impact on their client's information processing systems, as well as on their own audit procedures. This also means the development of APs.

In this study, an ANN-based system was trained and tested with the real world operating monthly data. The predictive ability of the ANN-system was compared with the predictive ability of the five other AP methods. The results indicate that the ANN-system has a better prediction accuracy than the other AP methods used in the study. One advantage of the ANN-based system such as used in this study is that it can provide auditors with objective information about a client company. Therefore, it can prove to be a persuasive analytical tool when an auditor discusses problems with the clients and recommends changes in the financial account values. The ANN-system could be adapted to the changes in the environment by retraining it with the new audited data. In our opinion the ANN-based system could serve a continuous monitoring and controlling purpose. For example, it could automatically trace, once a month, those accounts that follow a trend and are inside a certain threshold limit. Then the auditor could decide whether and what kind of further audit with these accounts would be needed.

Although the ANN-based systems cannot entirely replace professional judgment, they offer a promising alternative approach to APs. Indeed, ANNs provide a non-linear dimension that captures the underlying representations within the data sets. Therefore, the future of ANNs in auditing is open and challenging, but it will be brighter as more and more research efforts are devoted to this area. Another challenge is to get practitioners to adopt ANN-based or other feasible statistical methods.

[1] Statement of Auditing Standards (SAS) No. 56 of the AICPA (American Institute of Certified Public Accountants). AP first appeared in the authoritative literature of the AICPA in 1972 (Kinney, Jr. and William Jr., 1980).

REFERENCES

AICPA. 2002. AICPA Audit and Attest Standards, Statements on Auditing Standards (SAS) No. 96, Audit Documentation.

AICPA. 1988. *Statement on Auditing Standards #56, Analytical Procedures.* New York: American Institute of Certified Public Accountants.

Ameen, Elsie C. & Jerry R Strawser. 1994. Investigating the Use of Analytical Procedures: An Update and Extension. *Auditing: A Journal of Practice & Theory*, 13(2): 69-76.

APB. 1995. *Analytical Procedures, Statement on Auditing Standards 410*. London: Auditing Practices Board.

Blocher, Edward, Jr. Krull, George W., Leonard J. Tashman, & Stephen V. N: Yates. 2002. Updating Analytical Procedures. *The CPA Journal*(November): 36-41.

Colbert, Janet L. 1994. Analytical Procedures for Management Accountants and Auditors. *Managerial Auditing Journal*, 9(5): 3-7.

Demuth, Howard & Mark Beale. 2000. *Neural Network Toolbox*. Natick: The Math Works, Inc.

Dutta, S. 1993. *Knowledge Processing & Applied Artificial Intelligence*. Oxford, England: Butterworth-Heinemann.

Fleming, John M. 2004. Analytical Procedures are a Helpful Tool. *Pennsylvania CPA Journal*, 74(4): 3-4.

Fraser, Ian A. M., David J. Hatherly, & Kenny Z. Lin. 1997. An Empirical Investigation of the Use of Analytical Review by External Auditors. *The British Accounting Review*, 29: 35-47.

Gauntt, James E. Jr. & G. William Gletzen. 1997. Analytical Auditing Procedures. *Internal Auditor*(February): 56-60.

Kinney, William R. Jr. & William L. Jr. Felix. 1980. Analytical review procedures. *Journal of Accountancy*(Oct.): 98-102.

Koskivaara, Eija. 2004a. Artificial neural networks in analytical review procedures. *Managerial Auditing Journal*, 19(2).

Koskivaara, Eija & Barbro Back. 2003. An Artificial Neural Network Based Decision Support System for Budgeting. Paper presented at Fifth International Conference on Enterprise Information Systems, Angers, France.

Lee, Minwoo & Janet L. Colbert. 1997. Analytical procedures: management tools for monitoring controls. *Management Decision*, 35(5): 392-97.

Leitch, Robert A. & Yining Chen. 2003. The Effectiveness of Expectations Models in Recognizing Error Patterns and Generating and Eliminating Hypotheses While Conducting Analytical Procedures. *Auditing: A Journal of Practice & Theory*, 22(2): 147-70.

Lin, Kenny Z., Ian A. M. Fraser, & David J. Hatherly. 2003. Auditor analytical review judgement: a performance evaluation. *The British Accounting Review*, 35(1): 19-34.

Medsker, L. & J. Liebowitz. 1994. *Design and Development of Expert Systems and Neural Networks*: Macmillan College Publishing Company, Inc.

Riedmiller, Martin. 1994. Supervised Learning in Multilayer Perceptrons - from Backpropagation to Adaptive Learning Techniques. *Journal of Computer Standards and Interfaces*, 16.

Riedmiller, Martin & Heinrich Braun. 1993. A direct adaptive method for faster backpropagation learning: the Rprop algorithm. Paper presented at Proceedings of the ICNN, San Francisco.

PUBLICATION 7

Koskivaara (2004b) Visualization of patterns in accounting data with self-organizing maps. In: *Business Intelligence Techniques: A Perspective from Accounting and Finance*, ed. by Murugan Anandarajan – Asokan Anandarajan – Cadambi Srinivasan, 133–148. Springer-Verlag: New York.

8 Visualization of patterns in accounting data with self-organizing maps

Eija Koskivaara

TUCS Turku Centre for Computer Science, Turku School of Economics and Business Adminstration, Finland

Abstract. Neural networks are data driven methods. They could provide such additional information to the decision process as might be left hidden otherwise. Neural networks have already been applied in many different business areas and they can be used for prediction, classifying, and clustering. They can learn, remember, and compare complex patterns. This chapter shows how a neural network, especially Kohonen's self-organizing map (SOM), can be used in visualization of complex accounting data. The SOM is used for clustering ten years' monthly income statements of a manufacturing firm. The purpose is to show how the data sets of various accounts and various years form their own groups. We found that the SOM can be a visual aid for classifying and clustering data sets, and that it reveals if some cluster contains data that a priori should not be in it. Hence, it can be used for signaling unexpected fluctuations in data. Furthermore, the SOM is a possible technique embedded in the continuous monitoring and controlling tool.

8.1. Introduction

Many parties, such as investors, creditors, and managers, are interested in the accuracy of organizations' financial account values. Managers want to estimate the future revenues and expenses, in order to optimize the operations during a certain period. Auditors want to assess the accuracy of an organization's financial statements. Creditors want to analyze the organization's payment ability. All these parties may benefit from a tool that monitors and visualizes complex accounting data.

One possible technique embedded in the monitoring tool could be an artificial neural network (ANN). ANNs are not a priori statistical formal techniques but instead they are data driven techniques. ANNs have already been applied in many different business areas and they can be used for prediction, classifying, and clustering tasks (Vellido *et al.* 1999, Zhang *et al.* 1998). They can learn, remember, and compare complex patterns (Medsker and Liebowitz 1994). They are claimed to be able to recognize patterns in data even when the data are noisy, ambiguous, distorted, or variable (Dutta 1993). They are capable of discovering data relationships. It is well known, for example, that, unlike traditional statistical techniques, ANNs are capable of identifying and simulating non-linear relationships in the data without any a priori assumptions about the distribution properties of the data. ANNs learn from examples and then generalize them to new observations. They could provide such an additional information to the decision process as might be left hidden otherwise. These features make ANNs potentially suitable for many tasks within accounting. Information technology development and processing capacities of PCs have made it possible to model ANN-based information systems for monitoring and controlling operations.

Researchers have developed a variety of ANN-models to assist in the monitoring and controlling of operations such as detecting material errors in the data (Coakley and Brown 1991a, Coakley and Brown 1991b, Coakley and Brown 1993, Wu 1994, Coakley 1995, Busta and Weinberg 1998, Koskivaara 2000), detecting management fraud (Green and Choi 1997, Fanning and Cogger 1998, Feroz *et al.* 2000), and support for going concern decision (Hansen *et al*. 1992, Lenard *et al*. 1995, Koh and Tan 1999, Anandarajan and Anadarajan 1999, Etheridge *et al.* 2000). ANNs have also been applied to internal control risk assessment (Davis *et al.* 1997, Ramamoorti *et al.* 1999), and financial distress problems (Fanning and Cogger 1994). Material error applications direct users' attention to those financial account values where the actual relationships are not consistent with the expected relationships. A user has to decide whether and what kind of further investigation is required to explain the unexpected results. Material error ANN-models either predict future values or classify data. The results of material error ANN-models seem promising, at least as a supplement to traditional analytical auditing procedures, and offer improved performance in recognizing material misstatements within the financial accounts.

Companies are using or have used ANN applications to support their business. Credit-card companies use ANN technology to reveal fraudulent clients (Mulqueen 1996, Fryer 1996, Fisher 1999). KPMG Peat Marwick has already developed an ANN for bankruptcy prediction (Etheridge *et al.*

1994). Pohjola insurance company has applied an ANN solution to their direct marketing (Sinkkonen and Lahtinen 1998).

In this chapter we show the feasibility of an ANN, especially Kohonen's self-organizing map (SOM), in monitoring account values in monthly income statements of a manufacturing firm. We study whether the SOM is capable of revealing monthly variations and clusters in the data sets in the meaningful matter. The SOM has proven to be suitable for data analysis tasks (Kohonen 1997). Although many papers on self-organizing maps, since its invention in 1981, have been published, very few studies have dealt with the use of self-organizing maps for financial analysis. The studies by Martín-del-Brío and Serrano-Cinca (1993) and Back *et al.* (1998) are examples of the application of the SOM for financial analysis. Martín-del-Brío and Serrano-Cinca (1993) studied the financial statements of Spanish companies, and attempted to predict bankruptcies among Spanish banks during the 1977-85 banking crisis. Back *et al.* (1998) compared 120 companies in the international pulp and paper industry. In this study, the aim is to receive evidence of this method's suitability to analyze monthly account data. We anticipate that the SOM recognizes the relationships between different account values and reveals the time variation of the data sets. According to our information the SOM has not earlier been applied to analyzing monthly account data sets.

The remainder of the chapter is organized as follows: Section 8.2 describes the research methodology. This includes a brief description of SOM and the choice of the financial account values used in this experiment. Section 8.3 presents the construction of the self-organizing maps and Section 8.4 presents a detailed analysis of the maps and gives guidelines to implement this kind of model. The conclusions of this paper are presented in Section 8.5.

8.2. Methodology

8.2.1. Self-organizing maps

The SOM is a clustering and visualization method and the purpose is to show the data set in another representation form (Kohonen 1997). It creates a two-dimensional map from n-dimensional input data. This map resembles a landscape in which it is possible to identify borders that define different clusters. These clusters consist of the input variables with similar characteristics, i.e. in this report, of account values' monthly variation. The self-organizing training trials continue until two input items which are close in the input space are mapped into the same or neighboring neurons on the map. Output neurons create groups which together form a map of the input neurons.

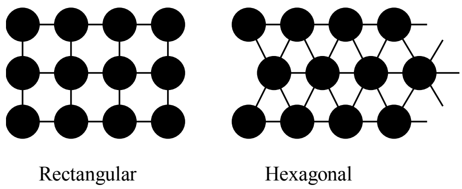


Fig. 8.1. Forms of lattice

The SOM has six learning parameters, *topology, neighborhood type, X- and Y-dimensions, training rate*, *training length*, and *network radius*. The network topology refers to the form of lattice. There are two commonly used lattices, rectangular and hexagonal (Fig. 8.1). In a rectangular lattice each neuron is connected to four neighbors, expect for the ones at the edge of the lattice. In the entire network we used, the output neurons are arranged in a hexagonal lattice structure. This means that every neuron is connected to exactly six neighbors, expect for the ones at the edge of the lattice. This choice was made following the guidelines of Kohonen (1997), since we expected the SOM to provide some benefit for the monitoring due to its visualization capability. Neighborhood type refers to the neighborhood function used and the options are Gaussian and bubble. X- and Y-dimensions refer to the size of the map. In too small maps differences between clusters are hard to identify and in too large maps clusters will appear to be flat. The training rate factor refers to how much the neuron in the neighborhood of the winning neuron learns from the input data vector. The training length measures the processing time, i.e. the number of iterations through the training data. The network radius refers to how many nodes around the "winning" neuron are affected during the learning process. The training process of the network is split into two parts. In part one, the map is "roughly" trained. In the second part, the network is fine-tuned.

8.3. Data

We used actual data comprising ten years' monthly income statements of a manufacturing company. The company was a medium-sized firm in Fin land and its net sales amounted to approximately EUR 11 million per year. The accounts were chosen with the help of a certified public accountant (CPA)-auditor in the way that the accounts represented the major and the most interesting monthly income statement categories. The accounts and their monthly averages in thousand euros are presented in Table 8.1.

Table 8.1. Financial accounts (in EUR 1000)

	90-99	1998	1999
Net sales	916	1325	1250
Materials + Change in inventory	215	297	259
Personnel costs	125	165	161
Gross margin	571	864	830
Administration	58	57	58
Total indirect	340	360	383
Operating profit	215	462	396
Receivables	1450	1630	1591
Trade debt	1468	2563	1965

The reasons for selecting the above accounts for our models are as follows.

- *Net sales* (NS) are a significant value to monitor. In this particular case variation between July and the other months is big because the company is closed in July. From the management's point of view it is better if the actual value is bigger than the budgeted value because then there are fewer disappointments. From the auditor's point of view this might raise doubts about whether all sales are recorded if the actual value is much below the prediction value. On the other hand, if the actual value is much higher than the prediction values the company might have recorded some fictitious sales.
- *Materials* (Mat) + *Change in inventory* (CinIn) together should tell the total use of material during a certain period. The value should be in alignment with net sales as this is a manufacturing company.
- *Personnel costs* (PC) should be in alignment with production and the total use of material.
- *Gross margin* (CM) is an important value at least from the prediction point of view as well as in seeing how much money is left to cover indirect costs and profit.
- *Administration* (Adm) is a good value in seeing the overall trend of the costs in the company and in the line of business.

- *Total indirect* (TotInd) indicates all fixed costs. This value should be predicted in all cases because these costs do not depend on sales.
- *Operating profit* (OP) is an interesting value at least from the prediction point of view. Furthermore, it is important to see that the operation is profitable in the long run.
- *Receivables* (Rec) are an interesting and important value to follow in order to know how much of the company's money is "outside".
- *Trade debt* (TD) tells how much the company has to pay "outside". Receivables and trade debts should be in alignment with the net sales.

8.4. Clustering with the SOMs

For the clustering purpose we used The Self-Organizing Map Program Package version 3.1 created by The SOM Programming Team of the Helsinki University of Technology in the network building (Kohonen *et al.* 1995). For the visualization of the results of the SOM we used Nenet - Demo version 1.1a created by The Nenet Team (Elomaa *et al.* 1999). Nenet is a user-friendly program designed to illustrate the use of SOMs, and it also provides individual parameter level maps, feature planes. This property suits our purposes perfectly, because we want to compare different accounts and months with each other.

We constructed two different kinds of maps with different input vectors. Firstly, we constructed a map so that in a vector there were the monthly data of the account per year as vector items. With this A-map we wanted to see how different accounts are situated in comparison to each other and to the previous years' values. Secondly, we constructed a map with the values of a certain month's data as vector values and presented them in a chronological order for the neural network. With this B-map we wanted to see whether there are any yearly tendencies in the data sets. These map approaches resemble analytical auditing procedures such as a comparison of current information with similar information for prior periods, and a study of relationships among the elements of information (Gauntt *et al.* 1997).

There are some rules of thumb when creating maps. The map ought to be rectangular, rather than square, in order to achieve a stable orientation in the data space. Normally, the x-axis should be about 30 per cent greater than the y-axis, thus forming a rectangular output map. Another recommendation is that the training length of the second part should be at least 500 times the number of the network units, in order to reach statistical accuracy (Kohonen 1997). We chose one where the layer consisted of 35 neurons arranged in a 5*7 hexagonal grid. As mentioned earlier hexagonal lattices are good for visualization purposes. The neighborhood function was the bubble. The training length and training rate in the first phase were 1750 and 0.5 and in the second phase 17500 and 0.05. The neighborhood radius in the first phase was 9 and in the second phase, 1.

To visualize the final self-organizing map we used the *unified distance matrix (U-matrix)*. This U-matrix method can be used to discover otherwise invisible relationships in a high-dimensional data space. It also makes it possible to classify data sets into clusters of similar values. The simplest U-matrix method is to calculate the distances between neighboring neurons, and store them in the matrix, i.e. the output map, which can then be interpreted. If there are "walls" between neurons, the neighboring ones are distant, i.e. the values differ significantly. The distance values are also displayed in colors when the U-matrix is visualized. On the maps we define the clusters by looking at the color shades of the borders between the hexagons. The dark colors in the walls represent great distances while brighter colors indicate similarities amongst the neurons. The colored borders between the hexagons are of great value when trying to determine and interpret clusters.

By viewing the individual feature planes it is possible to visualize the values of a single vector column, i.e. in this study, the maps for one month (A-maps) or for one account (B-maps). These feature planes can be visualized in order to discover how the company has been doing according to different months or different accounts. Because we selected accounts that depend on each other the feature planes of the months should be more or less similar.

8.5. Visualization of the data

Next we show how the outputs of the SOM can be used as a visual aid for classifying and clustering the monthly income statements values over ten years time period. The user analyses the clusters of accounts and variations of the monthly feature planes in order to find whether the clusters are close enough to each other or whether there are any significant differences between the monthly feature planes (A-map). The user may also analyze whether the account values are close enough to the previous year's values (B-map). If the difference is significant the user has to decide how much and what kind of further investigation is needed.

8.5.1 Studying the account cluster

Studying the underlying monthly feature planes of the A-map (Fig. 8.2) and the final A-map (Fig. 8.3), a number of clusters of accounts, and the characteristics of these clusters were identified (Fig. 8.4).
The feature planes in Fig. 8.2 show a map for each month in this study where the red color in the bottom left corners represents high values, which in our case implies revenue accounts. Dark colors in the bottom right corners show negative values, which in our case implies a trade debt account. From these feature planes we see that there is only a little variation between the months. For example, in the feature planes of June and July there are a little lighter neurons in the middle than in the other months' planes. However, the feature planes of the months are so similar that none of them gives any reason for auditing implications. This means that the relationships between the accounts included in this study are quite stable during the year. If the feature plane of the month differs much from the other feature planes it is a hint for a user to make some more investigations.

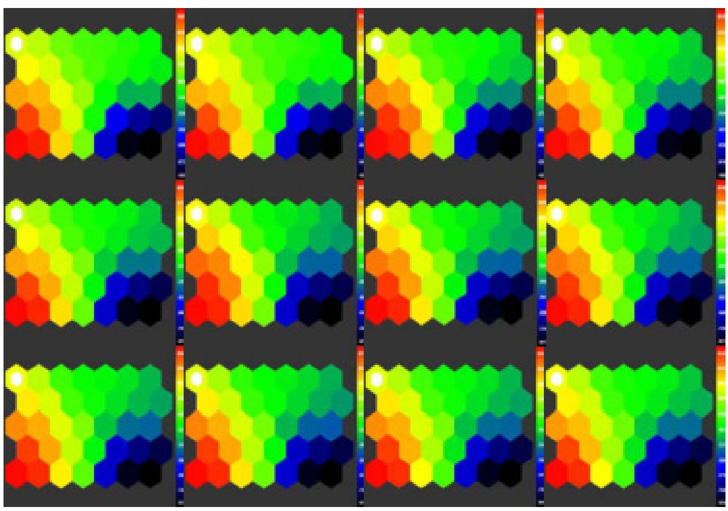


Fig. 8.2 The feature planes maps for the months: January, February, March, and April at the top, May, June, July, and August in the middle, and September, October, November, and December at the bottom

In Fig. 8.4, we have named identified clusters according to the accounts these clusters contain. We labeled the last two years accounts to see whether the accounts are close to each other and to name the clusters. We identified four main clusters: *revenues, margins, costs, and trade debts.* Revenue and trade debt clusters were easy to identify based on the feature planes of the different months. Although the trade debts of the last two years are in the same neuron, the cluster itself is much bigger because the earlier years' trade debt values were more spread out on the map. The revenue cluster could be bigger based on the feature planes on March, June, July, and August, however, the labeling of gross margin and operating profit reveals that these neurons belong to the margins cluster. All the cost accounts in our study are situated in the upper right corner and therefore we named it the cost cluster. Receivables (Rec), net sales (NS), operating profit (OP), personnel costs (PC), change in inventory (CinIn), administration (Adm), and materials (Mat) of two last years are in the same neuron. The gross margin (GM) and total indirect (TotInd) of 1998 and 1999 are in different neurons. This indicates that the relative monthly account value's variation of gross margin and total indirect is bigger than that of the other accounts in these years.

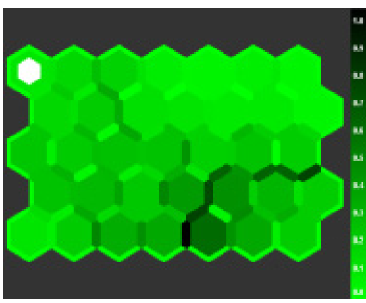


Fig. 8.3. The final A-map

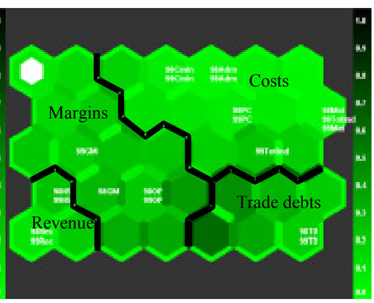


Fig. 8.4. Clusters on the A-map

8.5.2 Studying the yearly tendency

We also let the SOM cluster the account based on the month. With this B-map we wanted to see whether the months are close to each other and whether the different years are close to each other. We analyzed the B-map with account feature planes (Fig. 8.5) and by labeling all the months on the map. This way we found six clusters in a map. The feature planes of net sales, gross margin, operating profit, and receivables are at the top of the

Fig. 8.5. These accounts present in our sample the income accounts where the red color illustrates high values, which are situated on the left side of the feature planes. The overall outlook of the net sales, gross margin and operating profit seems very similar. The feature planes of materials, change in inventories, personnel costs, and administration are in the middle. Materials and personnel costs have the same general outlook as net sales although the colors are opposite. Its is good because they should be in alignment. The feature planes of total indirect and trade debts are on the bottom line. In these feature planes the darker the color gets the bigger the cost is.

In Fig. 8.6. we have counted how many monthly data of the year belong to one cluster in the B-map. We also see a tendency starting from the bottom right corner, where the early nineties data is situated, up towards the top of the map and then to the bottom left corner and once again up towards the top of the map. It seems that the best performance is in the bottom left corner where the monthly data from the year 1998 is located. All the July data are in the bottom line neurons. The first seven years' July data are in the ultimate bottom right corner of the map. The July data of 1997-1999 are also grouped at the bottom right neurons of these years' clusters.

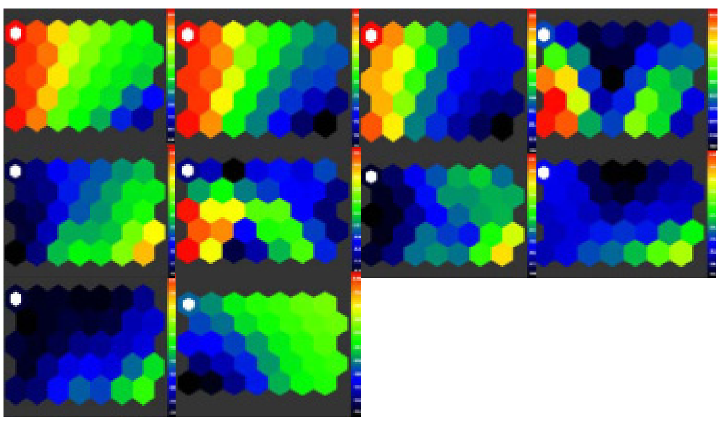


Fig. 8.5. Feature planes of accounts: net sales, gross margin, operating profit, and receivables at the top, materials, change in inventories, personnel costs, and administration in the middle, and total indirect and trade debts at the bottom

We also illustrate with a black arrow the monthly movement in 1998 and with a white arrow the monthly movement in 1999 (Fig. 8.7). From these arrows we see that the movement of the 1999 monthly data on the B-map is much broader. The reason for the more compact movement of the 1998 data might be that the account values of that year are the biggest in the whole data set and therefore they have concentrated in one corner. The same but opposite reason applies to the July data, especially with the years 1990-1996. These account values are the smallest in the whole data set and therefore have concentrated in the bottom right corner. We also see from Fig. 8.7 that both arrows start from and end at adjacent neurons.

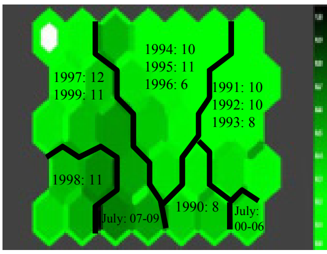


Fig. 8.6. Yearly clusters of B-map

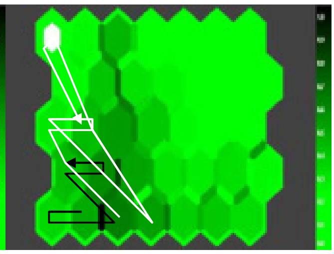


Fig. 8.7. Movements of the 1998 (Black) and 1999 (White) months

8.5.3 Account values

To ease the ANN's learning process and improve the quality of the map, the data is very often pre-processed in some manner. We did not use any pre-processing method because we wanted to calculate the 1999 average account values based on the vector values in the output maps. In Table 8.2, we compare the actual average monthly account values to values calculated from the vector values of the A-map and B-map. On average it seems that the vector order we have in the B-map is better than the vector order in the A-map if we compare the output vector values.

Table 8.2. Account values

	1999	A-map	B-map	A/1999	B/1999
Net sales	1250	1238	1239	-1 %	-1 %
Materials + Change in inventory	259	273	269	5 %	4 %
Personnel costs	161	123	161	-24 %	0 %
Gross margin	830	873	807	5 %	-3 %
Administration	58	59	56	2 %	-2 %
Total indirect	383	471	361	23 %	-6 %
Operating profit	396	467	394	18 %	0 %
Receivables	1591	1408	1532	-12 %	-4 %
Trade debts	1965	1626	1978	-17 %	1 %

8.5.4 Seeded error

We seeded an extra use of material in the data in order to see whether the SOM recognizes any difference. We manipulated the data by doubling the use of material in December 1999 in order to see its effects on maps and feature planes. This is a very tiny effect considering that in one map we have all the data from the ten years visible at the same time. In Fig. 8.8 we show how the feature planes of the maps change because of this manipulation. On the left side of the figure we have the feature planes of the original data, and on the right side we have the feature planes of the manipulated data. The feature planes at the top of the figure are based on months (A-map) and the feature planes at the bottom of the figure are based on accounts (B-map). The white neurons on the left show the right place for the vector. The white neurons on the right show where the manipulated data vectors are. We have also circled the effects the manipulation has on the whole map. The monthly use of material is situated in the same neuron in both cases (see Fig. 8.8 upper feature planes). However, the color of the adjacent neuron has changed dramatically. On the account feature planes the change is more radical. The whole feature plane looks very different. The manipulation has turned the whole feature plane inside out. The neuron has changed its place and the colors of the adjacent neurons have changed.

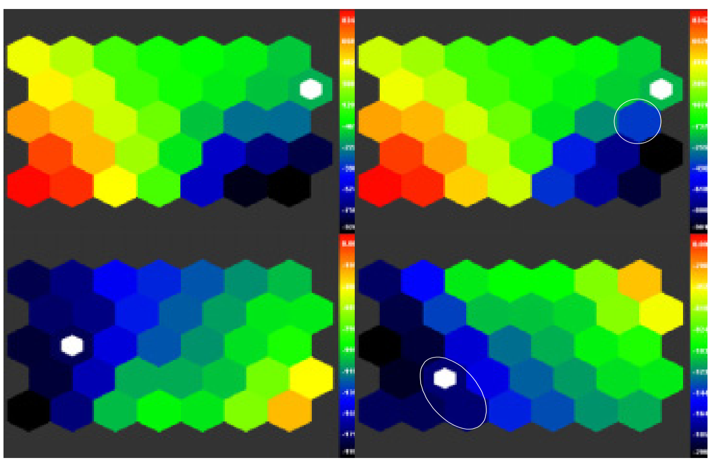


Fig. 8.8 Seeded use of material in December 1999

8.5.5 Implementing the model

Fig. 8.9 depicts the framework of our SOM-based pattern analysis model in accounting data. In general, this accounting data should illustrate the business process. The SOM takes advantage of the data that already exists in the accounting systems. The SOM utilizes the data and gives a compact and excellent visualization of it. Furthermore, it clusters the data in a meaningful manner. The SOM gives new kind of information to human analysts and therefore makes the interaction between a human actor and information system possible. Generally, the human actor may refresh the model in three ways. Firstly, they can make changes in the business processes. For example, allocate the use of material in the accounting system more timely. Secondly, they can correct the faulty data. For instance, take away all the fictitious sales from the bookkeeping records. Thirdly, a human actor may refresh the model either by changing the parameters in the model or by chancing the variables of the model.

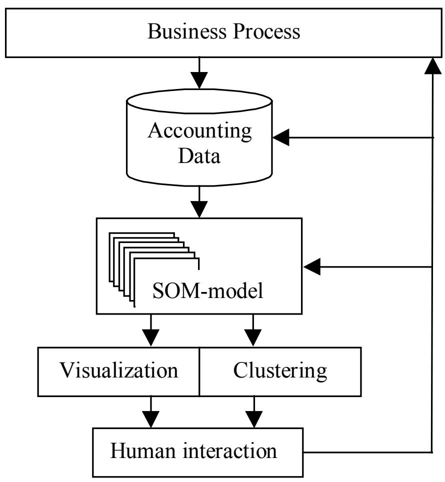


Fig. 8.9 Framework of the SOM pattern analysis model

8.6 Conclusions

In this chapter we showed how the SOM could be used in the visualization of the monthly income statement values. The SOM was used for the clustering of the data sets and the purpose was to show account values in an other representation form. We let the SOM cluster a manufacturing company's monthly account data from ten years. We found that the SOM is a tool for classifying these data sets, and that similar accounts form their clusters close to each other. We argue that the SOM can assist users by visualizing irregularities in the data and guiding the user to the heart of the problem. The SOM utilizes the entire data and finds homogenous groups in the data. With the SOM embedded in the monitoring system it is possible to plot a picture on screen from complex data sets, for example once a month, and give a visual aid for analyzers of the business data. We used monthly values of account but the SOM could also be used for analyzing all the transactions of certain accounts.

The development and assessment of advanced analysis methods like ANNs in an accounting context and for continuous monitoring and controlling is important in order to supply users with more efficient and effective means of monitoring account values. This is one way of restoring the public confidence in the capital market system and accounting profession that might have been shaken by the collapses of Enron and Arthur Andersen etc.

References

Anandarajan M, Anandarajan A. 1999. A comparison of machine learning techniques with a qualitative response model for auditor's going concern reporting. Expert Systems with Applications 16: 385-392

Back B, Sere K, Vanharanta H. 1998. Managing Complexity in Large Data Bases Using Self-Organizing Maps. Accounting, Management and Information Technologies 8: 191-210

Busta B, Weinberg R. 1998. Using Benford's law and neural networks as a review procedure. Managerial Auditing Journal 13(6): 356-366

Coakley JR. 1995. Using Pattern Analysis Methods to Supplement Attention-Directing Analytical Procedures. Expert Systems with Applications 9(4): 513-528

Coakley JR, Brown CE. 1991a. Neural Networks Applied to Ratio Analysis in the Analytical Review Process, Expert Systems Symposium. University of Southern California, School of Accounting: Pasadena, California

Coakley JR, Brown CE. 1991b. Neural Networks for Financial Ratio Analysis. In J Liebowitz (Ed.), The World Congress on Expert Systems, Vol. 1: 132-139. Pergamon Press: Orlando, Florida

Coakley JR, Brown CE. 1993. Artificial Neural Networks Applied to Ratio Analysis in the Analytical Review Process. Intelligent Systems in Accounting, Finance and Management 2: 19-39

Davis JT, Massey AP, Lovell RERI. 1997. Supporting a complex audit judgement task: An expert network approach. European Journal of Operational Research 103(2): 350-372

Dutta S. 1993. Knowledge Processing & Applied Artificial Intelligence. Butterworth-Heinemann: Oxford, England

Elomaa J, Halme J, Hassinen P, Hodju P, Rönkkö J. 1999. Nenet -Demo version 1.1a, Vol. 2001. The Nenet Team

Etheridge HL, Brooks RC. 1994. Neural networks: A new technology. The CPA Journal 64(3): 36-

Etheridge HL, Sriram RS, Hsu HYK. 2000. A Comparison of Selected Artificial Neural Networks that Help Auditors Evaluate Client Financial Viability. Decision Science 31(2): 531-550

Fanning KM, Cogger KO. 1994. A comparative Analysis of Artificial Neural Networks Using Financial Distress Prediction. Intelligent Systems in Accounting, Finance and Management 3: 241-252

Fanning KM, Cogger KO. 1998. Neural Network Detection of Management Fraud Using Published Financial Data. International Journal of Intelligent Systems in Accounting, Finance & Management 7(1): 21-41

Feroz EH, Kwon TM, Pastena VS, Park K. 2000. The Efficacy of Red Flags in Predicting the SEC's Targets: An Artificial Neural Networks Approach. International Journal of Intelligent Systems in Accounting, Finance & Management 9: 145-157

Fisher B. 1999. Mellon creates fraud watch to predict and manage risk using neural technology. Journal of Retail Banking Services 21(1): 15-17

Fryer B. 1996. Visa cracks down on fraud. InformationWeek(594): 87

Gauntt JEJ, Gletzen GW. 1997. Analytical Auditing Procedures. Internal Auditor (February): 56-60

Green BP, Choi JH. 1997. Assessing the Risk of Management Fraud Through Neural Network Technology. Auditing: A Journal of Practice & Theory 16(1): 14-28

Hansen JV, McDonald JB, Stice JD. 1992. Artificial Intelligence and Generalized Qualitative-Response Models: An Empirical Test on Two Audit Decision-Making Domains. Decision Science 23(3): 708-723

Koh HC, Tan SS. 1999. A neural network approach to the prediction of going concern status. Accounting and Business Research 29(3): 211-216

Kohonen T. 1997. Self-Organizing Maps. Springler-Verlag: Berlin

Kohonen T, Hynninen J, Kangas J, Laaksonen J. 1995. SOM_Pak v3.1, Vol. 2001. SOM Programming Team of the Helsinki University of Technology Laboratory of Computer and Information Science

Koskivaara E. 2000. Artificial neural network models for predicting patterns in auditing monthly balances. Journal of the Operational Research Society 51(9): 1060-1069

Lenard MJ, Alam P, Madey GR. 1995. The Application of Neural Networks and a Qualitative Response Model to the Auditor's Going Concern Uncertainty Decision. Decision Science 26(2): 209-227

Martín-del-Brío B, Serrano-Cinca C. 1993. Self-organizing Neural Networks for the Analysis and Representation of Data: Some Financial Cases. Neural Computing & Applications 1: 193-206

Medsker L, Liebowitz J. 1994. Design and Development of Expert Systems and Neural Networks. Macmillan College Publishing Company, Inc.

Mulqueen JT. 1996. Neural nets block fraud. CommunicationWeek(625): 95

Ramamoorti S, Bailey ADJ, Traver RO. 1999. Risk Assessment in Internal Auditing: A Neural Network Approach. International Journal of Intelligent Systems in Accounting, Finance & Management 8(3): 159-180

Sinkkonen J, Lahtinen I. 1998. Hermoverkot markkinoinin apuna. Pcmikro(4): 70-74

Vellido A, Lisboa PJG, Vaughan J. 1999. Neural networks in business: a survey of applications (1992-1998). Expert Systems with Applications 51: 51-70

Wu RC-F. 1994. Integrating Neurocomputing and Auditing Expertise. Managerial Auditing Journal 9(3): 20-26

Zhang G, Patuwo BE, Hu MY. 1998. Forecasting with artificial neural networks: The state of the art. International journal of Forecasting 14: 35-62