Using Machine Learning to Predict Ranking of Webpages in the Gift Industry: Factors for Search-Engine Optimization

Joni Salminen
Qatar Computing Research Institute, HBKU; and Turku School of Economics
Doha, Qatar
jsalminen@hbku.edu.qa

Roope Marttila
Turku University of Applied Sciences
Turku, Finland
roope.marttila@edu.turkuamk.fi

Bernard J. Jansen
Qatar Computing Research Institute, HBKU
Doha, Qatar
bjansen@hbku.edu.qa

Juan Corporan
Banco Santa Cruz RD
Santo Domingo, Dominican Republic
juan.nunez.corp@gmail.com

Tommi Salenius
Parcero Marketing Partners
Turku, Finland
tommi.salenius@parcero.fi

ABSTRACT

We use machine learning to predict the search engine rank of webpages. We use a list of keywords for 30 content blogs of an e-commerce company in the gift industry to retrieve 733 content pages occupying the first-page Google rankings and predict their rank using 30 ranking factors. We test two models, Light Gradient Boosting Machine (LightGBM) and Extreme Gradient Boosted Decision Trees (XGBoost), finding that XGBoost performs better for predicting actual search rankings, with an average accuracy of 0.86. The feature analysis shows the most impactful features are (a) *internal and external links*, (b) *security of the web domain*, and (c) *length of H3 headings*, and the least impactful features are (a) *keyword mentioned in domain address*, (b) *keyword mentioned in the H1 headings*, and (c) *overall number of keyword mentions in the text*. The results highlight the persistent importance of links in search-engine optimization. We provide actionable insights for online marketers and content creators.

KEYWORDS

Search-Engine Optimization; e-Commerce; Online Marketing; Content Marketing; Machine Learning; Rank Prediction

1 INTRODUCTION

Search engines (SEs) are a common entry point to the content on the World Wide Web. SEs, such as Google, Baidu, and Yandex, enable users to find content with the most relevant information for their queries [8, 10, 12]. For this, SEs navigate through an immense amount of content [3]. SEs crawl text content of websites via links, storing it in their database ('index') for further analyses [3, 11]. Although exact details are unknown, SEs are speculated to use advanced computational techniques to evaluate the relevance of web pages for search queries [15].

Because of the power that SEs command in attracting online users [23], websites compete for top positions in search results. In general, higher positions generate more visits and revenue. Therefore, website owners are keenly interested in improving their search results ranking, a process known as *Search Engine Optimization* (SEO) [7] that aims to accommodate the webpage with the SEs' presumed ranking factors (i.e., variables the SE considers when ranking webpages for a specific search query).

Ranking high on organic search engine results is crucial for online businesses, especially in the e-commerce sector where a major part of the visitors typically originates from SEs. Often, e-commerce stores and other websites use blogs to increase the breadth and depth of their content for higher search rankings. Content is widely regarded as influential search ranking criteria among online marketing practitioners [14, 21].

Much of the previous work on SEO has focused on the importance of links in search rankings. This is because link quantity and quality are central to PageRank, Google's core algorithm [3]. However, considerably fewer works study the impact of content and textual features on search rankings in a manner that is practical for evaluation by site creators. In this research, we address that gap by asking: *How do different content features predict search ranking of online content websites?*

ICIST '19, March 24–26, 2019, Cairo, Egypt

ACM ISBN 978-1-4503-6292-4/19/03...$15.00
https://doi.org/10.1145/3361570.3361578

To address this question, we utilize predictive machine learning. We use a dataset of search queries and their search engine rankings to understand which factors elevate a website's position in search results. The data contains information on several websites' content and link profiles. With this information, we aim to pinpoint what features contribute to a page's position in the search results for a relevant query.

2 RELATED LITERATURE

2.1 Collection of Articles

We searched for empirical works on SEO using keywords such as "search engine optimization", "google rankings" and so on. We found 19 articles that were deemed relevant based on reading abstracts. The small number of articles shows that the research topic is scarcely studied in academia. Out of the found articles, 9 (47%) are empirical and 10 (53%) are conceptual works, the earliest is from 2005 and the most recent from 2017.

2.2 Findings from Previous Research

Zhang and Dimitroff [31] investigated the impact of metadata implementation on webpage visibility in search engine results. They modified metadata of 46 web pages and submitted the modified versions to 19 search engines (among them Google); over several weeks, the rankings were improved in eight of the search engines. Similar to Sen [25], Zhang and Dimitroff [31] found that pages with metadata had higher visibility than pages without metadata, especially when the metadata was mentioned also in the page's text content. Regarding content characteristics, Zhang and Dimitroff [32] found that using duplicate keywords in the title increases the ranking but only until three repetitions, after which there is a decrease in visibility. The same effect was found for the body text, except without a decrease from repetitive use of the keyword. Moreover, using keywords in both title and body text resulted in better performance than using the keywords in just either of the two, while changes in font color, case, size, or use of plural and adjectives did not affect ranking [32].

Evans [8] analyzed Google rankings of 50 optimized and 50 non-optimized web pages. The researcher found the following factors influential for higher rankings: a) PageRank score, b) the number of inbound links, c) age of the domain name, and d) listings in *Yahoo!* and *DMoz* directories. Malaga [19] conducted an experimental study to increase the search rankings of an ecommerce site by using on-site optimization and link building tactics, finding that links from reputable websites had a major impact. Wang et al. [27] collected data from 118 websites to measure the effect of ranking factors, including a number of inbound links, title length, and keyword density. They found link popularity as the most important criteria. Other recommendations included limiting website title to 80 characters, page size to smaller than 150 kilobytes, the hierarchical order in a directory listing less than 4, and keyword density between 2–8% [27].

Gasparotto [9] suggests that higher website rank is correlated with higher site visit numbers, resulting in an effect where big websites are able to maintain their positions. Lee et al. [16] conducted a case study to analyze the SEO techniques applied to the *LG Science Land* content, finding the following factors influential: (a) simplified URL structure, (b) internal redirect in the case of page removal, (c) XML sitemap to help search engines index the site, (d) descriptive title and meta-tags, (e) use of canonical URLs, and (f) removal of expired links and content [16].

Zhang and Cabage [33] compared the effect of link building and social sharing on search rankings. They analyzed three content-rich websites with similar content, site and page structure, the volume of traffic, and search rankings. SEO efforts, including content creation, link building, and social sharing were then applied to treatment websites while the control sites were left as they were. The findings showed that links had the strongest impact on ranking over 18 months. While social sharing had a rapid impact on traffic, the increase was proved having only a temporary effect on the search ranking [33].

Table 1: Evidence for SEO Factors in the Literature

SEO factor	Impact on ranking		
	Positive	Negative	Neutral
Meta tags	[25] [31] [16]		[27]
Age of domain	[8]		
Internal links	[16]		
External links	[8] [19] [27] [33]		
Number of pages			[8]
Use of keywords in body text	[32] [27]	[32]*	
Font color			[32]
Use of directories	[8] [27]		
Page file size	[27]		
Social sharing			[33]
Website traffic	[9]		

**when used excessively*

Overall, there has been surprisingly little empirical work on the effect of search-engine optimization techniques despite the impact of search engines on the revenue of companies. The earliest work [8, 31, 32] emphasizes the use of meta-tags, after which the focus shifts on content, and particularly giving more exact prescriptions on the length and densities of content elements in relation to keywords [27]. The changes in the emphasis of the earlier studies are descriptive of the ever-changing nature of search-engine algorithms. Links have remained as essential part of SEO research, mostly due to Google's PageRank algorithm emphasizing reputable inbound links [3].

However, based on the literature review, we notice several gaps. First, negative effects are rarely reported. Second, internal links are examined very rarely in comparison to external (outbound) links that garner the most attention. Third, most studies are over a decade old, and it is therefore uncertain if the findings are applicable to modern search engines. Overall, we aim to address these gaps through this research focused on the practical problem of search rankings for online content websites.

3 METHODOLOGY

3.1 Research Context

This research context focuses on gift-related search queries. We collect data on keywords relevant to an e-commerce store that sells experience gifts and is based in Finland. Experience gifts are gift-cards that include an experience service from a specified service provider. They are part of the trend for immaterial consumption, providing alternatives to material gifts [5, 6].

This online company has a SEO strategy based on periphery blogs, meaning content sites that are focused on some specific gift theme (e.g., Valentine's day) and are hosted under separate domains. The company has some 50 of these blogs, out of which we selected 30 for this study. The selected blogs reflect the diversity of products and gift occasions, for example, gift cards, business gifts, room escape, and so on. Each blog has a distinct theme relating to either a gift category or product category. A product-themed blog would focus on a product category, e.g. Room Escape, whereas gift-themed blog focused on a gift keyword (e.g., 'gift card') or gift occasion ('christmas gifts'). The purpose of these blogs is to provide relevant content for searchers and links to the company's main domain.

3.2 Data Collection

For each chosen blog, we manually selected 4–10 keywords and phrases ('keyword' henceforth), using the following selection criteria: (a) *business value* – the business wants to rank high with the keyword; (b) *topical relevance* – the keywords relate to the theme of each blog; and (c) *search volume* – the keywords are sought in Google by many potential customers, verified using the Keyword Planner tool in Google Ads. This resulted in a total of 121 keywords, examples shown in Table 2.

Table 2: Examples of Keywords and Themes Corresponding to the Periphery Blogs.

Theme	Keywords (Translated from Finnish)
Valentine's day	valentine's day gift for boyfriend, valentine's day gift for girlfriend, valentine's gift for man, valentine's day gift for woman, valentine's day gift
Wedding gifts	gift ideas, wedding gift list, gift tips, gift for wedding couple, wedding gift
Christmas gifts	christmas gift ideas, christmas gifts, christmas gift for man, christmas gift for mom, christmas gift for dad, christmas gift for boyfriend, christmas gift for girlfriend, christmas gift online, christmas gift tips, christmas gift for child

After defining the list of keywords, we searched in Google for each keyword and recorded the addresses (URLs) of the first ten pages. This was done in privacy mode of Google Chrome to mitigate the impact of personalization on search results. Even though search results possibly vary by time and location, this cross-sectional sample represents the situation at a given point in time. We collected the URLs by manually performing searches and collecting the ten highest ranking URLs for each keyword.

From these searches, we had a list of 1,210 URLs, representing the highest-ranking web pages for each search term that we had defined. We skipped all duplicate URLs and keywords. The final count after skipping duplicate URLs was 750. There were also a few links that were invalid, some of which were combined, and some which were PDF. We fixed the combined ones manually, but the rest are skipped. There were 3 URLs that gave "404-Not Found" status and were skipped, resulting in 733 pages for analysis.

We then developed a script in Python to retrieve the hypertext (HTML) content from each of those pages. From this raw HTML text, we computed the content features, explained in Table 3. The content features were based on identifying the mentions of the chosen keywords in different HTML elements, as well as computing other common terms per document.

The counts were done on the raw HTML source code of the retrieved pages, in specific content elements. The content elements are *Paragraph* (<p>), *Heading 1* (H1), *Heading 2* (H2), *Heading 3* (H3), and the anchor text inside the link element (<a>), not including the *href* value of the link element. These content elements are a customary notation in HTML, see W3 guidelines[1].

We excluded links that relate to JavaScript functions or open an external application. We used the broad match of keywords [13], including both singular and plural forms of the keywords, and we did not separate between lower and upper case.

We also retrieved information about page loading times [2] , size and inbound links, as these are mentioned in prior research [29, 33]. Because retrieving this information for the 733 pages would be too time-consuming manually, we used *Netpeak Checker*. This tool was chosen for two reasons: (1) it collects the data automatically, and (2) it enables bulk upload of URLs and download of data. In summary, the extra information includes:

- **Content Download Time:** The time taken for all website content to download.
- **Content Size:** Size of the website taking just the content into account.
- **HTML Size:** Size of the website incl. HTML tags.
- **Related Pages in Google SERP:** Related pages of the domain on Google's search-engine results page.
- **Response Time:** The time is taken for the website to respond to a request.
- **Sites Linking in Alexa:** Websites that link to this page, based on Amazon Alexa results.

Finally, we computed one additional variable, **Secure URL**, that checks if the web page in question utilizes secure domain protocol, a choice that Google has recommended [24]. To merge this additional information with the original dataset, we used the URL as a unique identifier.

[1] https://www.w3.org/standards/

Table 3: Features Extracted from Webpages.

Feature	Definition
Amount of text	Count the number of characters in paragraph and titles (<p> and <h> elements)
H1 count of titles	Count the number of H1 titles on page
H1 length	Count the average length of H1 titles on page
H2 count of titles	Count the number of H2 titles on page
H2 length	Count the average length of H2 titles on page
H3 count of titles	Count the number of H3 titles on page
H3 length	Count the average length of H3 titles on page
Header total	Count of all the headers on page
Image count	Count the number of images
Internal links count	Count the number of internal links (internal = linking to a page in the same domain)
Keyword count H1	Count how many times the keyword mentioned in all the H1s
Keyword count H2	Count how many times the keyword mentioned in all the H2s
Keyword count H3	Count how many times the keyword mentioned in all the H3s
Keyword count p	Count how many times the keyword mentioned in all the paragraphs
Keyword in anchor text	0 if keyword not in anchor text of any link, 1 if keyword in anchor text of any link
Keyword in footer	0 if keyword not in footer, 1 if keyword in footer
Keyword in URL	0 if keyword not in URL, 1 if keyword in URL
Keywords in image alt	Count the number of times keyword mentioned in alt tag of images
Meta desc length	Count the length of the meta description. If no meta description, length = 0
Meta keywords count	Count the number of meta keywords used
Outbound links count	Count the number of outbound links (outbound = linking to a page not in the same domain)
Page title used	0 if no page title tag used, 1 if page title tag used
Total number of links	The number of total links on page

4 MODEL DEVELOPMENT

4.1 Overview of Approach

Because of the nature of the data (search keywords and URL contents), and the objective of the problem (Rank per Query), we are faced with a ranking problem [22]. Even though the problem could also be interpreted as a classification or regression problem, the common algorithms for these types of problems do not work well with ranked responses. On the other hand, implementations of well-tested libraries for ranking problems are uncommon. Because of these reasons, two Python libraries (*LightGBM* and *XGBoost*) are evaluated using the LambdaRank algorithm [4].

Overall, a data cleaning process is performed on the dataset to eliminate missing values and then cross-validation used to find the best base model. After this, hyper-parameter random optimization is employed to find optimal parameters for the base model. When this model is created, we use a framework by Lundberg and Lee [18] to interpret the impact of each feature on the prediction.

4.2 Evaluation Metrics

To measure the quality of the ranking algorithm, we use a metric that takes ranking into account. This metric is called *Normalized Discounted Cumulative Gain* (NDCG). Ranking solutions are often evaluated using NDCG [28]. A key assumption for this measure is that highly relevant results are more useful when appearing higher in search engine results [20].

The premise of DCG is that highly relevant results appearing lower in a search result list should be penalized as the graded relevance value is reduced logarithmically proportional to the position of the result. This calculation is dependent on the size of the result list of the query. However, since not all queries return the same number of results, NDCG was formulated to consider what the perfect results would look like for a query. We can also limit how many results of the query to consider calculating the metric. When this step is taken, the metric is known as NDCG@K, where K is the number of results taken to calculate the NDCG. This metric allows us to evaluate the performance of the algorithm, since it should return higher ranked results first. In our case, since the maximum rank is 10, K is set to 10. This way, we can evaluate the model using the full information.

4.3 Model Development and Selection

As this is a ranking problem, we are limited in the kinds of models we can use on the data. The algorithms used in these kinds of problems are called *Learning to Rank* (LTR) algorithms. LTR is a class of techniques that apply supervised machine learning (ML) to solve ranking problems [30]. The main difference between LTR and traditional supervised ML is summarized in the following:

- **Traditional ML** solves a prediction problem (classification or regression) on a single instance at a time. For example, if in spam detection on email, the algorithms inspect all the features associated with that email and classify it as spam or not. The aim of traditional ML is to come up with a class (spam or no-spam) or a single numerical score for that instance.
- **Learning to Rank** solves a ranking problem on a list of items. The aim of LTR is to come up with the optimal

ordering of the items. As such, LTR does not consider the exact score that each item gets, but it cares more about the relative ordering of all the items.

On the algorithms we use, the ranking is transformed into a pairwise classification or regression problem. That means the algorithm looks at pairs of items at a time, chooses the optimal ordering for that pair of items, and we then use it to come up with the final ranking for all the results [17]. In our case, we take one keyword, and all the data related to it. We compare two URLs and their features, and their ranking and the model aims to determine what makes one URL higher ranked than the other. Two models will be compared that solve this problem:

- **Extreme Gradient Boosted Decision Trees (XGBoost).** XGBoost uses a combination of decision trees that split the data into smaller subsets and gradient boosting to construct successive models that learn from the previous models' mistakes. The models are added a penalty for growing too complex, thus helping the model generalize better to new data. XGBoost also includes a pairwise loss function, making it suitable for ranking problems.
- **Light Gradient Boosting Machine (LightGBM).** LightGBM is similar to XGBoost, but it uses a different mechanism for growing the decision trees. Instead of growing the trees in a spread manner, LightGBM focuses on specific leaves of the tree first, allowing the trees to be built faster. It also includes an implementation for LambdaRank, a pairwise loss function for ranking.

To evaluate the alternative models, we use the NDCG@10 metric that was explained earlier. As a base, we build an XGBoost model for ranking. This model expects data to be in a specific DMatrix format, and each query to have its own group. We then create a model that uses each keyword, page information and ranking to learn to rank the sites and evaluate the performance of the model on test keywords. Essentially, the model generates predicted numbered ranks for each site of the data. The output is initially numerical float rankings that are converted into discrete rankings (see Table 4 for examples).

Table 4: Example Results with First-Run of XGBoost Model. Keywords Translated from Finnish.

Keyword	NDCG score
'Mother's Day gift online'	0.996
'Mother's Day gift'	0.986
'birth day gifts'	0.994

Overall, NDCG compares the order of the predictions against the perfect order possible. For example, say one has 5 items ordered [1, 2, 3, 4, 5]. If the model predicts [1, 2, 3, 4, 5] then the NDCG score is 1. If the model predicts [5, 4, 3, 2, 1], then the NDCG score is 0. However, if the model predicts [1, 3, 2, 4, 5], the NDCG score is still high; while a prediction like [1,4,3,2,5] would have a slightly lower NDCG score. Thus, NDGC is calculated based on how similar the predictions are to the optimal ranking order. This way, even if the exact ranking does not match, as long as the order is close to the prediction, the NDCG score is high.

Cross-validation divides the data into training data, that will be used to create the model, and validation data, that will be used to assess its performance. Then, the validation data is added to the training data, and another subset of the data is used to validate the model. The process is repeated until all data has been used to train and validate the model. However, for ranking, cross-validation works in a slightly different way. Instead of using a subset of the data, we use a subset of the queries. This way, we test on truly unseen queries and sites that the model ranks. After preparing all the functions, we evaluate both models to pick the one that generalizes better.

Table 5 shows the average NDCG scores for the two models.

Table 5: Average NDCG Scores of Cross-Validated Models.

Model	Avg. CV NDCG Score
XGBoost	0.852
LightGBM	0.848

Based on the cross-validation scores, we choose the XGBoost for final prediction and model interpretation.

4.5 Model Optimization and Interpretation

XGBoost involves parameters that can be tuned to fit better for a given problem. The possible combinations of values that these parameters can take are infinite. However, we can limit the number of possible values to pick from and simplify the combinations of parameters to be used. To do this, we first define a list of possible values of the parameters. We then define a function to pick a value for each parameter at random. Finally, we combine cross-validation with random parameters, known as Randomized Search Optimization [1]. Using this approach, we choose a combination of parameters that optimizes the performance of the XGBoost model, and obtain an NDCG of 0.858.

XGBoost models form a series of decision trees. While decision trees by themselves are easy to interpret, multiple decision trees used to train a model are not. One of the ways that gradient boosted decision trees or GBDT can be interpreted is by using feature importance. The importance is calculated by assessing how good is the model performance when that feature is absent. Features that are highly important to the model hurt its performance when they are missing. The measure that quantifies a feature's value to the model is known as *F Score*.

The F Score values in Figure 1 show that link features have the highest importance. A secure domain is also important. Features concerning headers, specifically *H3*, also have high importance. Other features like external links, response time, outgoing links, related pages, and the size of the page, provide some information to the model. Content features did not garner as much importance from the model, except if the keyword is visible in the footer.

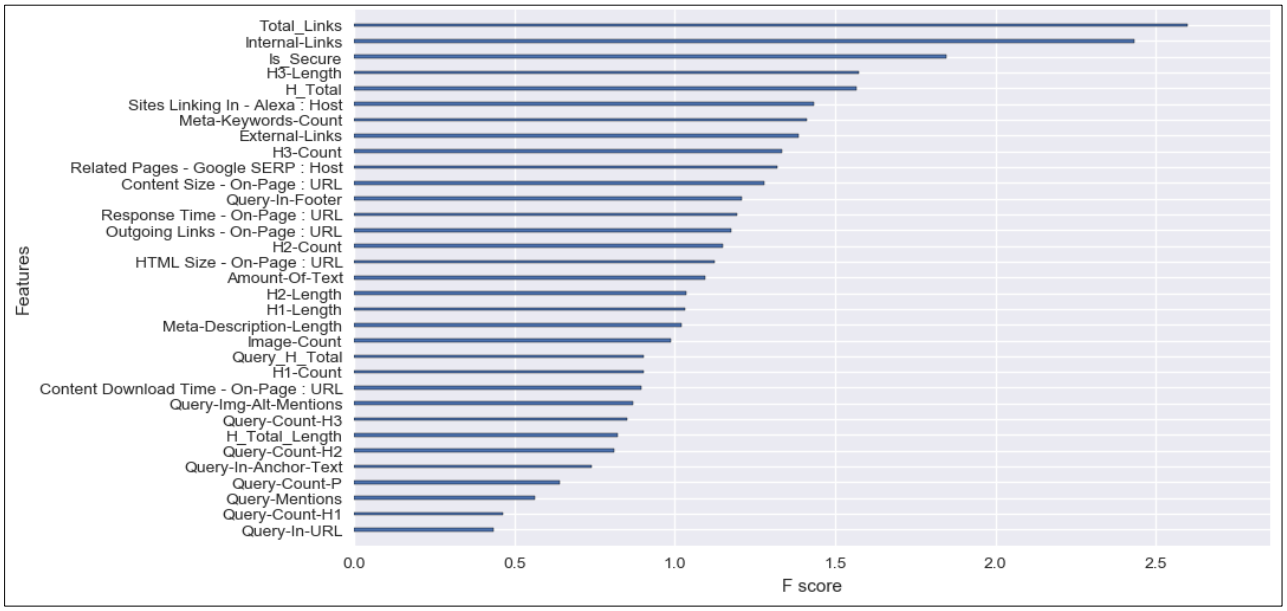

Figure 1: Feature Importance Values. Higher Values Indicate More Impact on the Predictions Made by the Model. The Analysis Shows that Link Features Are the Most Influential.

While feature importance indicates the feature's weights when predicting rank, it does not indicate the direction. For example, while *Total Links* has a high weight when determining rank, does it increase or decrease the rank?

To gauge the direction each feature takes the prediction into, we use the SHAP (SHapley Additive exPlanations) algorithm [26]. This algorithm takes each feature and assigns a weight to it when making a prediction, thereby capturing the directional impact of the feature. Figure 2 shows the SHAP values of each feature. The most common contributing features are at the top – red points indicate high values and blue points low values.

While the content features do not seem to be contributing much to the model, not reducing them would have led to a very complex model. We can see how content features impact the model by creating a summary plot (see Figure 3). While these features have a negligible impact on the model, higher values for the content features tend to yield higher rankings.

5 DISCUSSION

5.1 Positioning to Earlier Research

While most previous SEO research focuses on external links and their impact on search rankings [8, 27, 33], we are among the first ones to examine the impact of internal linking schemes. Apart from the study by Lee et al. [16] that considers the impact of site structure and accessibility on the search ranking, we could not locate other studies making this association. Our finding uses more variables than that of Lee et al. [16], yet corroborating their finding on internal linking schemes playing a role for search engine optimization. Moreover, our findings show that keyword mentions in the various HTML elements have a positive impact on rankings, corroborating earlier findings by Zhang and Dimitroff [32]. However, the excessive use of keywords (Query-Mentions in Figure 3) have a slight negative effect on the ranking, again reconfirming some earlier research on keyword density [32]. While the impact of these features is small on the overall model, as search algorithms are speculated to involve hundreds of individual features with varying weights [15], the use of keywords in content for SEO efforts cannot be neglected.

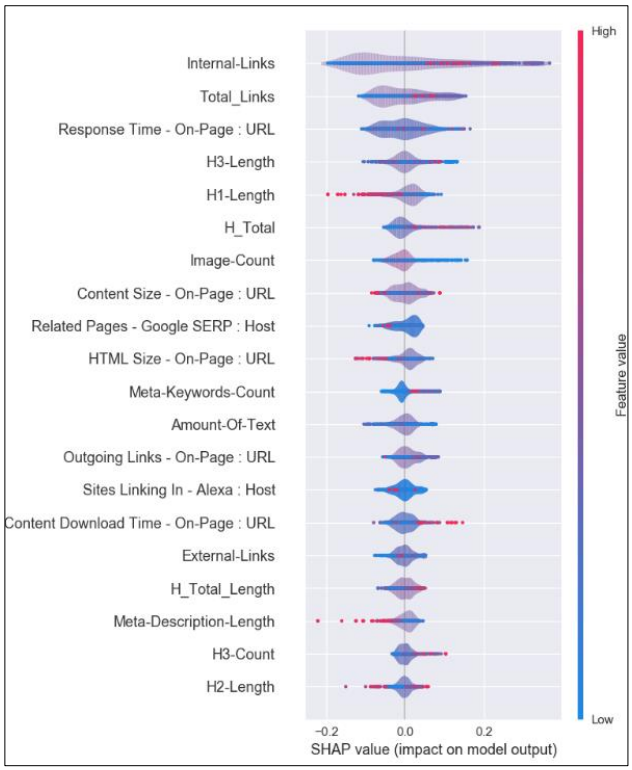

Figure 2: SHAP Values of the Features.

Figure 3: Impact of Content Features on Prediction.

5.2 Practical Implications for SEO

Key insights (seen from Figure 2) include:

- Having more internal links increases the ranking, but a very high number of links has a negative effect.
- Mid to low response times yield better results for ranking that high response times.
- Long H1 tags lead to lower rankings, while a high count of H tags leads to higher overall rankings.
- Low image counts lead to higher rankings.
- Lower amounts of related Google pages lead to marginally higher rankings.

Although having keywords in content is critical for the page's discoverability, their effect on ranking seems to be less impactful than internal linking, page-loading times, and external links. Several of these factors seem to exhibit non-linear behavior, meaning that adding keywords and internal links become redundant or even detrimental beyond a certain point. Therefore, we advise against keyword stuffing and link-farming that increase the numbers of such variables without considering their quality. Overall, a well-structured and sectioned page, with H1-H3 headers that are short and to the point, aids in obtaining better search engine rankings.

5.3 Limitations and Future Work

The primary limitation of this work is that the keywords and associated webpages came from only one company and industry. In addition, the chosen language (Finnish) might affect the results. In general, the gift industry can be considered as a highly competitive online industry with a lot of SEO activity taking place. Although the range of keywords was relatively large in the context of that company, as was the number of webpages, this research would need to be replicated using data on other companies, industries, and languages in order to claim generalizability of the findings.

Even though we mitigate the impact of potential personalization by using an anonymous browser, there are other factors that impact the search results, such as click logs, ranking information from past SERPs, and so on. These factors make search results structurally unstable and make it more difficult to replicate research in this domain. Moreover, as the ranking algorithms of the major search engines undergo periodic changes, any research in the SEO field is subject to expiration.

Even with the mentioned limitations, the results are indicative of the impact of content and link features on search rankings. Acquiring more data would allow for the use of more features (e.g., utilizing unsupervised methods such as topic modeling), and more learning examples to further improve the algorithm. In addition, more features about the actual content of the sites, would provide more distinct information about each site.

Apart from obtaining data from other contexts, future research could focus on specific website elements. In particular, the relatively high correlation of H3 and rankings is an interesting finding. One reason for this can be that the use of H3 tags is rarer than the use of H1 and H2 tags and, therefore, websites using H3 tags are applying more advanced SEO and content marketing strategies. This proposition should be explored in future research.

6 CONCLUSION

A good search engine ranking is instrumental in obtaining more website visitors, more clicks, and more revenue. In this research, we analyzed what factors drive this ranking up in order to better understand what factors a website owner should optimize to improve rankings. Our results show that webpages that contain a high but not too high amount of internal and external links, several keyword mentions in content, and low loading times and file sizes have higher rankings, while those without these characteristics have lower rankings.

REFERENCES

[1] Auger, A. and Doerr, B. 2011. *Theory of randomized search heuristics: Foundations and recent developments.* World Scientific.

[2] Bai, X. and Cambazoglu, B.B. 2019. Impact of response latency on sponsored search. *Information Processing & Management.* 56, 1 (2019), 110–129.

[3] Brin, S. and Page, L. 1998. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems.* 30, 1–7 (1998), 107–117.

[4] Burges, C.J. 2010. From ranknet to lambdarank to lambdamart: An overview. *Learning.* 11, 23–581 (2010), 81.

[5] Clarke, J. 2008. Experiences as gifts: from process to model. *European Journal of Marketing.* 42, 3/4 (Apr. 2008), 365–389. DOI:https://doi.org/10.1108/03090560810852986.

[6] Clarke, J.R. 2006. Different to 'dust collectors'? The giving and receiving of experience gifts. *Journal of Consumer Behaviour.* 5, 6 (Nov. 2006), 533–549. DOI:https://doi.org/10.1002/cb.201.

[7] Davis, H. 2006. *Search engine optimization.* O'Reilly Media, Inc.

[8] Evans, M.P. 2007. Analysing Google rankings through search engine optimization data. *Internet research.* 17, 1 (2007), 21–37.

[9] Gasparotto, M. 2014. Search Engine Optimization for the Research Librarian: A Case Study Using the Bibliography of U.S. Latina Lesbian History and Culture. *Practical Academic Librarianship: The International Journal of the SLA Academic Division.* 4, 1 (Jun. 2014), 15–34.

[10] Jansen, B.J. 2009. Understanding user-web interactions via web analytics. *Synthesis Lectures on Information Concepts, Retrieval, and Services.* 1, 1 (2009), 1–102.

[11] Jansen, B.J. and Booth, D. 2010. Classifying Web Queries by Topic and User Intent. *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 2010), 4285–4290.

[12] Jansen, B.J. and Mullen, T. 2008. Sponsored search: an overview of the concept, history, and technology. *International Journal of Electronic Business.* 6, 2 (2008), 114–131.

[13] Jansen, J. 2011. *Understanding Sponsored Search: Core Elements of Keyword Advertising*. Cambridge University Press.

[14] Kent, P. 2012. *Search engine optimization for dummies*. John Wiley & Sons.

[15] Latent Dirichlet Allocation (LDA) and Google's Rankings are Remarkably Well Correlated: 2010. *https://moz.com/blog/lda-and-googles-rankings-well-correlated*. Accessed: 2018-04-23.

[16] Lee, S. et al. 2016. Search engine optimization: A case study using the bibliographies of LG Science Land in Korea. *Library Hi Tech*. 34, 2 (2016), 197–206.

[17] Lei, Y. et al. 2017. Alternating Pointwise-Pairwise Learning for Personalized Item Ranking. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (New York, NY, USA, 2017), 2155–2158.

[18] Lundberg, S.M. and Lee, S.-I. 2017. A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems 30*. I. Guyon et al., eds. Curran Associates, Inc. 4765–4774.

[19] Malaga, R. 2007. The Value of Search Engine Optimization: An Action Research Project at a New E-Commerce Site. *Journal of Electronic Commerce in Organizations*. 5, 3 (2007), 68–82.

[20] Meng, Z. et al. 2018. Search result diversification on attributed networks via nonnegative matrix factorization. *Information Processing & Management*. (2018).

[21] Mill, D. 2005. *Content is King: Writing and Editing Online*. Elsevier Butterworth-Heinemann.

[22] Rafailidis, D. and Crestani, F. 2017. A Collaborative Ranking Model for Cross-Domain Recommendations. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (New York, NY, USA, 2017), 2263–2266.

[23] Salminen, J. 2009. *Power of Google: A study on online advertising exchange*. Master's thesis. Turku School of Economics.

[24] Secure your site with HTTPS - Search Console Help: 2018. *https://support.google.com/webmasters/answer/6073543?hl=en*. Accessed: 2018-04-30.

[25] Sen, R. 2005. Optimal Search Engine Marketing Strategy. *International Journal of Electronic Commerce*. 10, 1 (Oct. 2005), 9–25. DOI:https://doi.org/10.1080/10864415.2005.11043964.

[26] shap: A unified approach to explain the output of any machine learning model: 2018. *https://github.com/slundberg/shap*. Accessed: 2018-08-16.

[27] Wang, F. et al. 2011. An empirical study on the search engine optimization technique and its outcomes. *2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)* (Aug. 2011), 2767–2770.

[28] Wang, Y. et al. 2013. A Theoretical Analysis of NDCG Ranking Measures. *JMLR: Workshop and Conference Proceedings* (2013), 1–30.

[29] Webmaster Guidelines - Search Console Help: 2018. *https://support.google.com/webmasters/answer/35769?hl=en*. Accessed: 2018-04-23.

[30] Xu, B. et al. 2017. Learning to Rank with Query-level Semi-supervised Autoencoders. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (New York, NY, USA, 2017), 2395–2398.

[31] Zhang, J. and Dimitroff, A. 2005. The impact of metadata implementation on webpage visibility in search engine results (Part II). *Information Processing & Management*. 41, 3 (May 2005), 691–715. DOI:https://doi.org/10.1016/j.ipm.2003.12.002.

[32] Zhang, J. and Dimitroff, A. 2005. The impact of webpage content characteristics on webpage visibility in search engine results (Part I). *Information Processing & Management*. 41, 3 (May 2005), 665–690. DOI:https://doi.org/10.1016/j.ipm.2003.12.001.

[33] Zhang, S. and Cabage, N. 2017. Search Engine Optimization: Comparison of Link Building and Social Sharing. *Journal of Computer Information Systems*. 57, 2 (Apr. 2017), 148–159. DOI:https://doi.org/10.1080/08874417.2016.1183447.