

Optimising Personalised Medical Insights by Introducing a Scalable Health Informatics Application for Sensor Data Extraction, Preprocessing, and Analysis

Chirath HETTIARACHCHI ^a, Robin VLIEGER ^{a,1}, Wenbo GE ^a, Deborah
APTHORP ^{a,b}, Elena DASKALAKI ^a, Anne BRÜSTLE ^a and Hanna SUOMINEN ^{a,c}

^aThe Australian National University, Australia

^bUniversity of New England, Australia

^cUniversity of Turku, Finland

ORCID ID: Chirath Hettiarachchi <https://orcid.org/0000-0002-7702-0718>,

Robin Vlieger <https://orcid.org/0000-0001-8492-6203>, Wenbo Ge
<https://orcid.org/0000-0002-0429-3061>, Deborah Apthorp <https://orcid.org/0000-0001-5785-024X>,
Elena Daskalaki <https://orcid.org/0000-0001-5416-3372>, Anne Brüstle
<https://orcid.org/0000-0002-3842-5683>, Hanna Suominen <https://orcid.org/0000-0002-4195-1641>

Abstract. Wearable sensors, among other informatics solutions, are readily accessible to enable noninvasive remote monitoring in healthcare. While providing a wealth of data, the wide variety of such sensing systems and the differing implementations of the same or similar sensors by different developers complicate comparisons of collected data. An online application as a platform technology that provides uniform methods for analysing balance data is presented as a case study. The development of balance problems is common in neurodegenerative conditions, leading to falls and a reduced quality of life. While balance can be assessed using, for example, perturbation tests, sensors offer a more quantitative and scalable way. Researchers can adjust the platform to integrate the sensors of their choice or upload data and then preprocess, featurise, analyse, and visualise them. This eases performing comparative analyses across the sensors and datasets through a reduction of heterogeneity and facilitates easy integration of machine learning and other advanced data analytics, thereby targeting personalising medical insights.

Keywords. health informatics, machine learning, signal processing, software tool

1. Introduction

The rapid development of wearable sensors and machine learning techniques has accelerated the development of personalised medicine, making cutting-edge technologies available to the general population [1]. These affordable solutions serve as valuable early screening tools, providing an accessible alternative to costly medical-grade options [2]. Everyday devices, such as smartphones and watches, effortlessly

¹ Corresponding Author: Robin Vlieger, robin.vlieger@anu.edu.au.

collect individual data, enabling the detection of diverse health conditions through sophisticated analysis. While providing a wealth of data, the wide variety of sensors and the differing implementation of the same or similar sensors by different developers makes it challenging to compare data collected using different devices. For example, smartwatches such as the Apple Watch 7 [3] and Samsung Galaxy Watch 4 [4] can collect similar data, but data formatting, preprocessing, and visualisation differ because they are products of different developers. Fortunately, many sensor technologies provide the option to export either raw or pseudo-raw data, often in the form of comma-separated values data files that can then be preprocessed, visualised, and analysed.

This manuscript introduces an online platform called SensAnalytics that provides uniform methods for data analysis and visualisation of data. By using methods of data preprocessing, feature extraction and visualisation, researchers can adjust the platform to integrate the sensors of their choice, upload data, and perform the same analysis on all of them for easier comparison through a reduction of heterogeneity. Using balance data as a case study, this study will (1) demonstrate the steps to use the SensAnalytics tool and (2) highlight its unique capabilities. In the scientific literature, balance has been assessed in various ways, sometimes using sensors. This is often undertaken in the context of, for example, neurodegenerative conditions, such as multiple sclerosis (MS) and Parkinson's Disease (PD), as people living with these conditions tend to develop balance problems as their disease progresses [5-8]. The gold standard is to use a balance plate (e.g. Bertec Force Plate) that measures variables such as anteroposterior and mediolateral sway at high sampling rates [9]. However, with the development of smartphones and wearable devices such as smartwatches (e.g. Apple Watch 7), it has also become possible to measure balance outside of clinical settings at a time convenient to the participant. While not producing the same accuracy as balance plates, they do provide useful information when used correctly. Similar developments can be observed in numerous domains where the focus has shifted towards consumer-grade sensors, creating new opportunities and challenges for research [1]. The capabilities of SensAnalytics aim to target these and are summarised as follows:

- SensAnalytics provides the capability to analyse both consumer- and research-grade sensors, allowing for analysis and comparison.
- A wide range of consumer sensors by multiple vendors is currently available. However, extracting data and performing analysis is challenging due to the different data schemas and technologies used. The SensAnalytics tool provides a standardised pipeline to analyse data from various sensors developed by other vendors.

2. Methods

The SensAnalytics tool (Figure 1) was developed using the Python framework [10] and publicly released under the MIT license at <https://github.com/chirathyh/SensAnalytics>. The tool's deployment requires no specific hardware configurations and can be accessed through desktop and mobile devices. The choice to build SensAnalytics as a cross-platform web-based application instead of an app for particular devices ensures easier access regardless of the device being used.

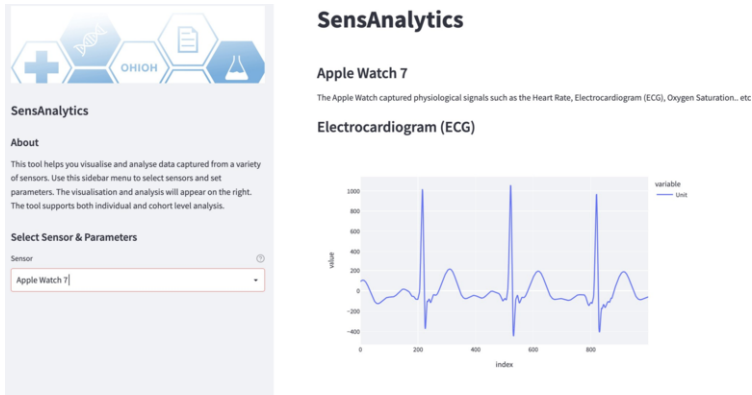


Figure 1. An example of plotted electrocardiogram data collected using an Apple Watch 7.

Five steps of using the SensAnalytics tool will now be described: *Step 1: Selecting a Sensor.* From a drop-down menu, the user is asked to select the sensor to analyse data from. SensAnalytics can accept data from the Bertec Force Plate BP5046, the Apple Watch 7, the Samsung Galaxy Watch 4, and the Zephyr Bioharness 3. *Step 2: Specifying an Analysis Type.* After selecting the desired sensor, the user is asked to specify if they wish to analyse data from an individual participant or perform an analysis on a cohort level. *Step 3: Data Upload.* SensAnalytics accepts comma-separated values data files as its import format. Individual files are required to have a maximum size of 200MB. *Step 4: Preprocessing, Feature Extraction, & Visualisation.* The user is asked to select the desired preprocessing steps (e.g. the filters and sampling frequency for the balance board [11]). After preprocessing, the user is provided with the choice of the extracted features from a drop-down menu and the choice of different plots to visualise the chosen feature. Currently, the feature extraction pipeline for the balance board proposed by [11] is implemented, and the choice of plots includes scatterplots, time series, and histograms, but this can be quickly extended to suit other preferences and research questions. *Step 5: Analysis.* The last step involves the performance of statistical testing. The researcher can quickly implement the statistical test of their preference to be performed on the data that was visualised in Step 4. These five steps support data science and, for example, subsequent machine learning.

Applying the above steps to dummy data, which would be collected using a Bertec Force Plate, provides a case study on balance sensing and analytics related to a balance board. In this scenario, founded on Ge et al. (2021) [11, 12]², data is collected from people standing on the balance board first with their eyes open and then with their eyes closed. The uploaded data consists of 1000 data points per second and six dimensions: time, force in newton, moment about the x-axis and y-axis, and the centre of pressure on the x-axis and y-axis.

² First, the systematic review [11] of 31 studies investigated postural sway in PD and found a total of 129 features extracted. Calculating effect sizes for the features, the review concluded that the most often extracted features (e.g., path length and sway area) were not the most discriminative and that higher sampling rates provided better outcomes. It recommended a sampling rate of 100 Hz and called for standardisation of experimental setup and statistical testing. Second, the experimental study [12] indicated that the sampling frequency of the data can be of influence on the utility of features when attempting to distinguish people with PD from controls.

Analysis of a person’s data through a platform like SensAnalytics can provide insight into potential balance problems a person might have. For instance, balance is often assessed in people living with neurodegenerative conditions, such as MS and PD [5-8]. People living with such neurodegenerative conditions report more falls than people without a neurodegenerative condition. In a study into older people with MS (PwMS), more than 50% had had a fall in the last six months [5], while a study into people with PD (PwPD) reported 46% as having had a fall in a period of three months [6]. Falls lead to injury, a fear of future falls and their consequences, and a reduced quality of life [4, 8].

3. Results

As a further system demonstration, Figure 2 visualises the balance board data from the case study using SensAnalytics as a scatter plot with different colours. After selecting the Bertec Force Plate as the sensor of choice (Step 1 of SensAnalytics), data is analysed at a cohort level (Step 2), and the files containing the data are uploaded (Step 3). After selecting the desired filter frequency (the options being 10, 20, 40, and 100 Hz), the SensAnalytics tool preprocesses the data and extracts the 129 features in accordance with Ge et al. (2021) [11, 12] (Step 4).

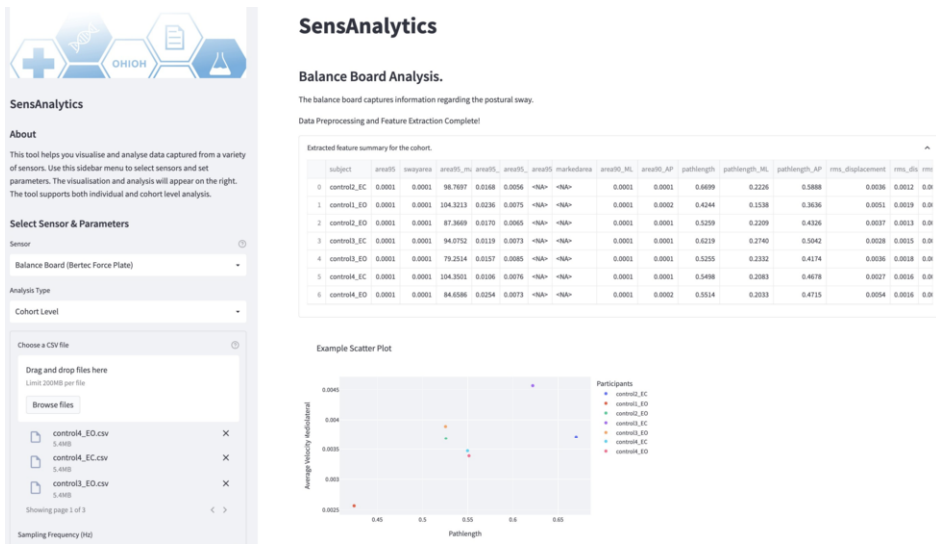


Figure 2. An example of a cohort analysis of Bertec balance data.

SensAnalytics now presents a table containing each participant’s extracted features, ready for the user to choose the type of visualisation and data analysis (Step 4 & 5). The chosen plot will be produced based on the current choice of scatterplots, time series, and histograms.

As SensAnalytics allows for the analysis of sensor data from different sources and different modalities, the platform is set up in such a manner that a researcher can amend the code to visualise data by group or in any other way they desire. Furthermore, the analysis part of SensAnalytics is designed to be customisable by the researcher. For

example, in the case of eyes open and eyes closed data, a conventional statistical test can be carried out to look at the difference between the two conditions for any of the 129 features. With the same ease, it is also possible to implement a machine learning method to provide a personalised analysis of a participant's data compared to the data of others.

4. Discussion

The SensAnalytics application presented in this manuscript is a platform that enables users, both in academia and outside it, to easily visualise and analyse data collected using different sensors. In this way, SensAnalytics provides a solution to the problem of rich data being collected from various technologies from other developers that use the same or similar sensors. By streamlining the whole pipeline from data uploading to visualisation and analysis, shown here using the Bertec Force Plate BP5046 as an example, SensAnalytics enables users to quickly make sense of postural data.

The SensAnalytics platform allows for personalised medical insights through its capability to preprocess, analyse, and visualise data from the individual, and through the capacity to implement machine learning algorithms that would allow a user to place individual results in the context of the larger cohort. While other applications may be able to provide some of this functionality — for example, Apple Health, Samsung Health, or Polar Beat — these applications allow only for collection and analysis of data with their proprietary devices, while SensAnalytics is device-agnostic. With its potential for advanced data analysis, SensAnalytics follows up on the recommendation made in a 2022 review into sensor technologies for the management of Type 1 diabetes that, while sensor technologies have the potential in disease management, the development of analysis tools was lacking and should be developed [13].

In this manuscript, the focus was on balance data, but SensAnalytics can be easily adapted to perform the same or a similar analysis on data from other sensors, such as balance data collected from the Apple Watch 7 or the Samsung Galaxy Watch 4. Additionally, the framework of SensAnalytics can be easily adapted to perform analyses on data of other origins than balance data, such as heart rate, electrocardiographic data, and blood oxygen level.

SensAnalytics, being an application that does not require any specific hardware configurations and can be accessed through desktop machines and mobile devices, provides the capacity for quick and standardised data analysis of rich data on both the individual and cohort levels. By enabling the integration of consumer wearables and research devices, this technology platform is valuable for the general public and researchers.

5. Conclusions

The work presented here showcases SensAnalytics as an application that can be of value to both researchers and people outside of academia, allowing easy preprocessing, visualisation, and analysis of data collected from various sensors. The ease with which the platform can be adjusted to the requirements of the user shows it to be a promising solution for personalised medical insights.

Acknowledgments

This research was funded by and has been delivered in partnership with Our Health in Our Hands (OHIOH), a strategic initiative of the Australian National University (ANU), which aims to transform healthcare by developing new personalised health technologies and solutions in collaboration with patients, clinicians, and healthcare providers. We gratefully acknowledge the funding from the ANU School of Computing and the Australian Government Research Training Program (AG RTP) for the first three authors' PhD studies.

References

- [1] Dunn J, Runge R, Snyder M. Wearables and the medical revolution. *Pers Med*. 2018 Sep;15(5):429-48, doi: 10.2217/pme-2018-0044.
- [2] Hsieh KL, Roach KL, Wajda DA, Sosnoff JJ. Smartphone technology can measure postural stability and discriminate fall risk in older adults. *Gait Posture*. 2019 Jan;67:160-5, doi: 10.1016/j.gaitpost.2018.10.005.
- [3] Apple Watch Series 7 — Technical Specifications [Internet]. Apple Support. [cited 2024 Mar 21]. Available from: <https://support.apple.com/en-us/111909>.
- [4] Samsung Galaxy Watch 4 [Internet]. The Official Samsung Galaxy Site. [cited 2024 Mar 21]. Available from: <https://www.samsung.com/global/galaxy/galaxy-watch4/specs>.
- [5] Sosnoff JJ, Motl RW, Morrison S. Multiple sclerosis and falls - an evolving tale. *US Neurology*. 2013 Sep;9(1):30-34, doi: 10.17925/USN.2013.09.01.30.
- [6] Boonstra TA, Van Der Kooij H, Munneke M, Bloem BR. Gait disorders and balance disturbances in Parkinson's disease: clinical update and pathophysiology. *Curr Opin Neurol*. 2008 Aug;21(4):461-71, doi: 10.1097/WCO.0b013e328305bdaf.
- [7] Sosnoff JJ, Socie MJ, Boes MK, Sandroff BM, Pula JH, Suh Y, Weikert M, Balantrapu S, Morrison S, Motl RW. Mobility, balance and falls in persons with multiple sclerosis. *PLoS One*. 2011 Nov;6(11):e28021, doi: 10.1371/journal.pone.0028021.
- [8] Apthorp D, Smith A, Ilschner S, Vlieger R, Das C, Lueck CJ, Looi JCL. Postural sway correlates with cognition and quality of life in Parkinson's disease. *BMJ Neurol Open*. 2020 Nov;2(2):e000086, doi: 10.1136/bmjno-2020-000086.
- [9] Zampogna A, Mileti I, Palermo E, Celletti C, Paoloni M, Manoni A, Mazzetta I, Dalla Costa G, Pérez-López C, Camerota F, Leocani L, Cabestany J, Irrera F, Suppa, A. Fifteen years of wireless sensors for balance assessment in neurological disorders. *Sensors*. 2020 Jun;20(11):3247, doi: 10.3390/s20113247.
- [10] Van Rossum G, Drake FL, editors. Python language reference manual. Amsterdam: Centrum voor Wiskunde en Informatica; 1995. 150 p, ISBN: 978-1906966140.
- [11] Ge W, Lueck CJ, Apthorp D, Suominen H. Which features of postural sway are effective in distinguishing Parkinson's disease from controls? A systematic review. *Brain Behav*. 2021 Jan;11(1):e01929, doi: 10.1002/brb3.1929.
- [12] Ge W, Apthorp D, Lueck C, Suominen H. Which features of postural sway are effective in distinguishing Parkinson's disease patients from controls? An experimental investigation. In: 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2021 Dec 9 (pp. 860-867). IEEE.
- [13] Daskalaki E, Parkinson A, Brew-Sam N, Hossain MZ, O'Neal D, Nolan CJ, Suominen H. The potential of current noninvasive wearable technology for the monitoring of physiological signals in the management of type 1 diabetes: literature survey. *J Med Internet Res*. 2022 Apr;24(4):e28901, doi: 10.2196/28901.