Acta Scientific COMPUTER SCIENCES

Volume 4 Issue 12 December 2022

Research Article

Creating a Classification Module to Analysis the Usage of Mobile Health Apps

Kuniko Azuma[1]*, Tareq Al Jaber[2] and Neil Gordon[2]

[1]*MSc Artificial Intelligence and Data Science, University of Hull, United Kingdom*

[2]*School of Computer Science, University of Hull, United Kingdom*

***Corresponding Author:** Kuniko Azuma, MSc Artificial Intelligence and Data Science, University of Hull, United Kingdom.

Received: October 18, 2022

Published: November 15, 2022

Abstract

With an ageing society becoming a major issue for many countries, health-related concerns are growing and mobile health applications (MHAs) are rapidly gaining users. The applications available range from those that promote exercise to maintain health, those that help to manage physical condition by recording weight and activity, and those that allow users to consult doctors and pharmacists. On the other hand, there are still many mobile users who do not use MHAs. In this case study from Japan, the range of diverse MHAs were classified into five categories by K-means clustering analysis and the results of a questionnaire on the use of MHAs were analyzed using a scientific approach to find out which types of users mainly use these applications. Based on the results of this analysis, a classifier was created using a Random Forest algorithm to extract MHAs that meet the needs of users based on their attributes and thoughts. With this Random Forest classification model, this paper recommends appropriate models for potential users who are not yet using MHAs.

Keywords: MHA (Mobile Health Application); m-Health; K-means; Clustering; Random Forest

Abbreviations

MHA: Mobile Health Application

Introduction

Context

In Japan, the ageing population is a major issue with a declining birth-rate and in three years' time, the number of people aged 65 and over is expected to reach 30% of the population, leaving a gap of around 10 years between average life expectancy and healthy life expectancy [1]. In order to close this 10-year gap, it is imperative to increase healthy life expectancy and enable older people to live independently. Meanwhile, smartphone ownership now exceeds 50% even among older people in their seventies [2]. This high ownership rate suggests that MHA could promote health at all ages and contribute to longer healthy life expectancy. Furthermore, MHAs have evolved rapidly in recent years, with the creation of significant applications for maintaining, promoting, and improving health with a variety of functions, exceeding 50,000 apps [3]. However, there are many dormant users and non-users. This is thought to have led to a mismatch of MHAs and uncertainty about which MHAs should be used. In order to solve this, this paper explains how MHAs were classified using a scientific approach based on the characteristics of the apps actually used, and details a classification model created to understand the characteristics of the user personas and propose applications that match the users' situation. Based on this analysis, we propose the types of apps for users who do not currently use MHAs, which could make a significant contribution to extending healthy life expectancy.

Previous work

Previous research has been conducted on the use of MHAs. In a survey conducted with health university students to identify barriers to MHA use, the purpose of MHA use was analyzed [4].

The study concluded that students were targeted and there were no differences in use by gender. However, it is thought that there may be differences when targeting Japan, which ranks 116th out of 146 countries in the global gender gap ranking [5]. A survey of MHAs among older people reported that one in ten said they used them to share health information with healthcare professionals [6]. The study shows that older users want to communicate with doctors and other experts when using MHAs. This is a unique use by a generation that often has health problems. This may therefore affect the usage depending on how they engage with MHAs. While there were no gender differences in the studies, another study found women tended to focus on the app's objectives, while men focused on the specific goals and features that the app offered [7]. The above previous studies have helped to understand that the evaluation of MHAs, or the purpose of the mobile health application, differs depending on gender, age and other factors. However, these studies were conducted using MHAs selected by the researchers. No analysis was conducted based on applications answered by users already using MHAs.

Aims of this Research

- Define the characteristics of the different types of MHAs, which are becoming increasingly large and diverse.
- Identify key user characteristics of classified MHAs
- Based on the characteristics of the identified user personas, create a Classification model to propose to users who are not yet MHA users.

Materials and Methods

Methodology

In this study, data collection was conducted using a web-based questionnaire form called Surveroid from Marketing Applications, Inc [8]. The questionnaire asked basic user information and questions about the use of MHAs. Details are provided in the Appendices. Responded data was split into MHA and user information.

MHA analysis

For the MHA analysis, we first created a unique dataset on the applications used by users. This dataset was created by examining category information from the App Store and extracting the availability of features from the application's landing page. In addition, the source of the applications was researched. Furthermore, as the question on the names of MHA was asked in an open-ended format, 40 responses with abbreviations or aliases were seen. These have been corrected to the correct notation.

Application layers

Layer 1: The first layer was the default application, which can be used without users having to consciously install it in "Health" on iOS and "Google Fit" on Android.

Layer 2: This layer was assumed to be applications distributed by manufacturers that sell wearable devices, such as Garmin Connect and Fit Bit. These were created to be used alongside wearable devices.

Layer 3: The third layer is applications other than Layer 1 and Layer 2, which need to be deliberately selected and installed by the user.

Nine functions of apps from landing pages

- AI
- Doctors/Experts
- Food analysis
- Gamification
- Fitness
- Sleep
- Financial merits
- Medication management
- Female menstrual cycle records.

Based on the aspects of these applications, a clustering analysis was used to classify them into different types of applications. The K-means method was employed in this study for the following reasons. K-means clustering has been used since the algorithm was proposed by Stuart Lloyd in 1957 and first called "K-means" by MacQueen [9]. It is still a popular clustering algorithm. The K-means function of "scikit learn", a library commonly used by data scientists, also still uses algorithms from a method proposed in 1957 [10]. As an efficient model, a further improved algorithm has been proposed by Hartigan [11], which pointed out that Lloyd was computationally expensive and that the quality of the clustering results was highly dependent on the choice of initial centroid [12]. However, the amount of data in the current dataset is small and computational cost is therefore not an issue. Furthermore, the expected clustering classification values are also small, and the initial values calculated by the elbow method can be visually checked, so problems caused by initial value dependence can be avoided.

User analysis

For the analysis of users, we analyzed gender, age, employment status, geographical distribution, main purpose of use, overall satisfaction with the MHA being used, billing status and willingness to recommend it to others in order to extract the characteristics of users for each cluster. In addition, the relationship between the number of functions an MHA has and the level of satisfaction per cluster group was analyzed, as well as the quality of content, performance, credibility, confidentiality, entertainment value, visual appeal, ease of use, and ease of navigation of functions as a detailed satisfaction survey of MHAs.

Classification model

A classification model was created for the application cluster group, using the user characteristics of each cluster as input values. Random Forest [13] was used to create the model. It is said to be high performance compared to classic models such Support Vector Machine (SVM) [14] and logistic regression models [15]. In using machine learning models in this study, there was a large variation in the amount of data for each target MHA cluster, and the data set itself was unbalanced with a small number of samples. Unbalanced datasets have an impact on the quality of machine learning. The small number of data categories was therefore solved with the Synthetic Minority Over-sampling Technique (SMOTE) [16], which has proven to perform better than Random Over-sampling (ROS) [17]. To validate the model, project members were asked to complete a questionnaire on factors affecting the predictions. At the end of the questionnaire, they were asked to select which applications they actually use and, if not, which they would like to use.

Result and Discussion

Result

This survey is based on the results of a questionnaire conducted via a webpage among 841 randomly selected men and women aged 18 years and older. The results showed that 74.44% (626 respondents) stated that they were using an installed MHA, while 215 respondents stated that they did not have an installed MHA. Of the 626 respondents, 300 were male and 326 were female. As shown in Figure 1, the proportion of non-users is relatively low among younger generations, both men and women.

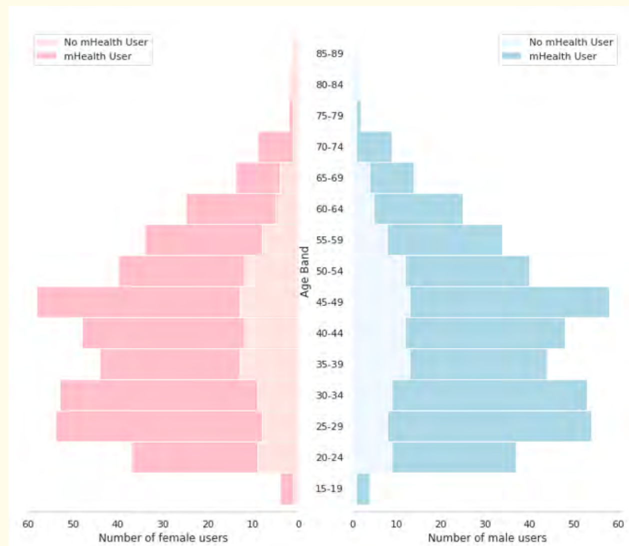


Figure 1: Usage of MHA by gender and age band.

MHA analysis

The MHA App Store categories used by users included 49 "Health and Fitness" and four "Medical" categories, as well as various other categories, as table 1 shows. Although this is not the main intention of the distributors, it shows that there are a number of applications available to improve health. However, in order to focus solely on MHAs, the survey was limited to "Health and fitness" and "Medical" applications. In addition, in surveying the applications, the applications were categorized into three layers.

App Store Category		**Layer**		**Provider**	
Category	n	Layer	n	Type	n
Health and fitness	49	Layer 1	2	Health-tech	11
Medical	4	Layer 2	12	Non-health-tech	63
Lifestyle	7	Layer 3	60		

Navigation	2				
Photo and video	2				
Food and drink	1				
Social networking	1				
Books	1				
News	1				
Weather	1				
Utility	1				
Education	1				
Game	1				
No Category	2				

Table 1: Breakdown of MHAs.

The heatmap in figure 2 shows the results of examining the correlation between these features. As seen in this heatmap, there was a relatively large association between "AI" and "Food Analysis", "Games" and "Fitness", "Points" and "Fitness", "Fitness" and "Games", "Menstrual Cycle Record" and "Food Analysis", "Menstrual Cycle Record" and "Medicine Management". From this it can be inferred that the applications tend to combine several functions. The applications with the most functions had six of the above functions. The most common application had three functions, followed by applications with none of the functions. The median was 2 functions.

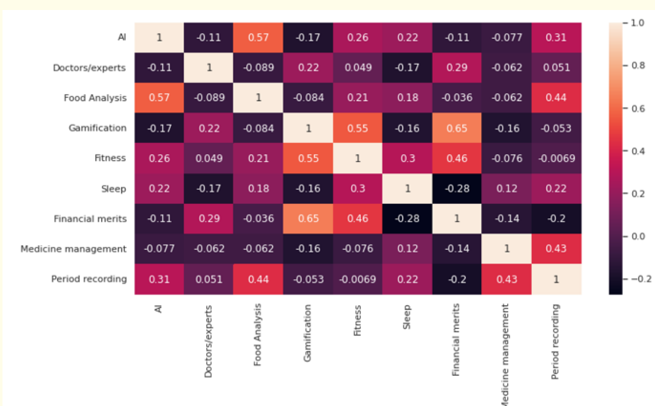


Figure 2: Correlation of 9 functions.

In addition, the study investigated whether the application was delivered by a health tech company, which has seen impressive growth in recent years [18]. 13 apps in this survey were created by health tech companies. The applications were analysed using a clustering method based on the above application information. The results are as follows of figure 3.

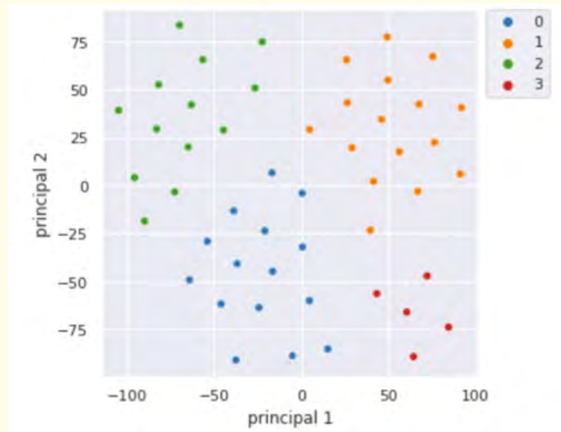


Figure 3: App clustering by K-means.

The elbow method showed that the applications could be classified into four clusters. Therefore, clusters were created using the k-means method and drawn as two-dimensional using t-Distributed Stochastic Neighbor Embedding (T-SNE), which is good at preserving the feature structure of local data when converting high-dimensional data to low-dimensional data, and the four clusters could also be visually identified [19]. It is said that T-SNE provides a better visualization of clusters than Principal Component Analysis (PCA) [20]. Then, by examining the applications in detail, the following four application type characteristics could be extracted:

- **Cluster 0**: This cluster did not have many functions but was an application that focused on monitoring changes in a woman's physical condition, such as recording menstrual cycle, diet and sleep.
- **Cluster 1:** The group of applications in this cluster had a relatively high number of functions, were games and fitness-oriented, and focused on non-health benefits, mainly enjoying health-related activities rather than improving serious health conditions.
- **Cluster 2:** For cluster 2, the default applications for Health, Google fit and for use with Layer 2 devices were classified. This is the base group of applications for activity recording.

- **Cluster 3:** The applications classified here were highly functional applications created by health tech companies, using AI and other technologies.

Although the classification of applications by the K-means clustering method allowed for four categories of applications, a study of the purpose of use of users for each cluster showed that there were two types of characteristics within Cluster 2. It was found that users of Health and Google Fit, which are pre-installed with the OS, have a large number of users who answered “Other” as their purpose of use, whereas the overwhelming majority of Layer 2 users’ purpose of use is “Monitoring”. Due to the different user personas, it was decided to survey users of five different applications in Cluster 2, with Health and Google Fit as Cluster 4.

MHA user analysis

Next, the number of MHAs and users of the five types of applications classified above were then shown in table 2 below.

Cluster	Number of apps	Number of users
0	15	35
1	16	86
2	11	19
3	5	45
4	2	356

Table 2: Number of MHAs and the users each cluster.

Gender and age band

The graph in figure 4 shows the percentage of each cluster by gender and age group. Although most respondents use pre-installed models, a slightly higher proportion of women use apps that are not pre-installed, indicating that they are more particular about their choice of apps. Women are relatively more likely to use Cluster 0, perhaps because they are more concerned about changes in their health status when they are younger. However, from middle age onwards, use of high-performance apps and fitness and gaming apps increases. Cluster 2 of device use was also higher than among women.

User satisfaction

The next part of the survey is in terms of application satisfaction by users in each cluster. Only in cluster 3, the high-performance application tier created by health tech companies, nearly 40% of users said that they already pay for the application or that they use the free version but would use it even if the application was paid for. Overall satisfaction was low in all clusters, with a low percentage of “Very Good” or “Good” responses, especially in Cluster 4, where less than 20% of respondents answered, “Very Good” or “Good”. Cluster 2 has a low overall satisfaction level, but a very large proportion of respondents would recommend it to others.

Figure 4: Comparison of users each app cluster by age band and gender.

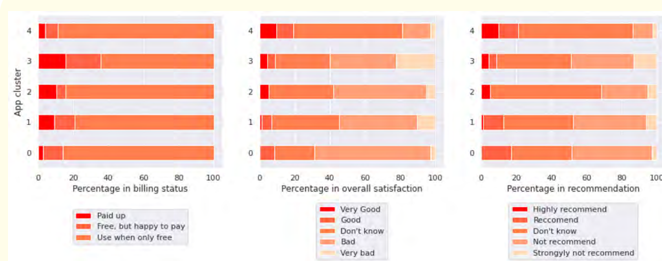


Figure 5: Comparison of user satisfactions in each app cluster.

Further investigation into the details of satisfaction by cluster revealed that Cluster 3 users were highly satisfied in all areas. Whereas, Cluster 4, the pre-installed software, tended to be rated lower overall.

In addition, high-performance applications have a relatively high number of functions. We checked what happens to the evaluation

Figure 6: Detailed evaluation of aspects of MHA by clusters.

of the application as the number of functions increases. In cluster 3, the overall evaluation was not that high, indicating that the number of functions is not directly related to the level of satisfaction.

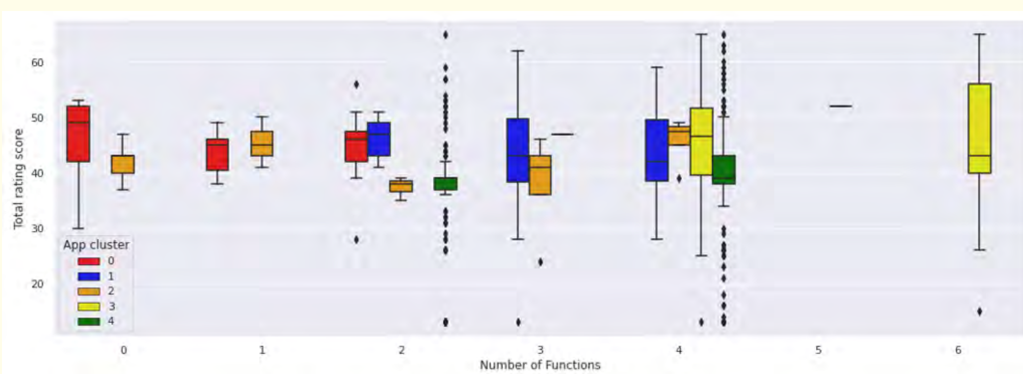


Figure 7: Impact of number of functions to user satisfaction by App cluster.

Employment status

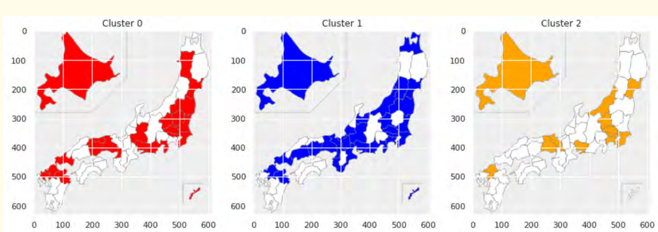
Next, analysis of whether there are differences in the choice of MHA type by employment status shows a higher proportion of permanent employees in the Cluster 2 group. In the current survey, the proportion of full-time employees was almost twice as high, with 92 women compared to 183 men. It can be assumed that this group of full-time employees with higher incomes are more likely to use wearable devices as they have to purchase the devices, which are generally not inexpensive. Conversely, cluster 1, the gaming and fitness cluster, had the lowest usage rate among full-time employees. Cluster 1 has a high utilisation rate among part-timers and housewives. Time availability and financial incentives for health-related activities may be motivating factors.

Use of purposes

An analysis of the purpose of use showed that the most common cluster 2 need for wearable devices was "Monitoring". They are trying to observe more detailed activity and sleep records through the device and to monitor their health in detail. The pre-installed model group in Cluster 4 had a large number of respondents who stated "Other" as their main purpose, while the other purposes were sparse, suggesting that their main purpose was not clear. In the cluster 3 group, the most common purposes were "Diagnosis" and "Education", and some respondents also answered "Treatment", indicating that they expect serious improvement or betterment of their health rather than enjoyment.

Geographical distribution of clusters

Cluster 4 had a high number of respondents with pre-installed software, indicating that it is used nationally. On the other hand, for Cluster 2, it is shown to be concentrated in urban areas with relatively large populations.



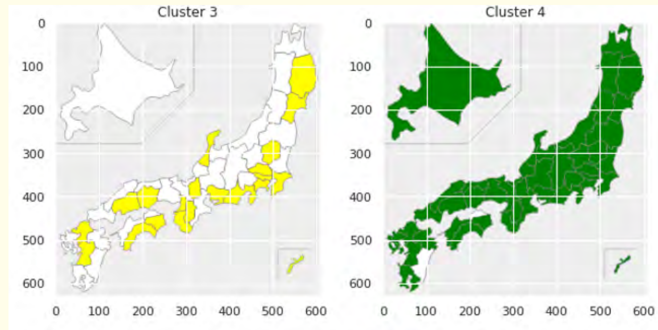


Figure 8: Geographical distribution of clusters.

Cluster group main user definitions

From the results of the above analysis, the following characteristics of each cluster group can be concluded.

- **Cluster 0 users:** Women who are relatively young, or middle-aged men, who have minor concerns about their health and would rather not spend money on their health care. They believe that they would be more satisfied with mobile applications if the functions that each of them focusses on were available.
- **Cluster 1 users:** In addition to young people, many have low incomes, such as part-timers and housewives. This is a group of people who enjoy fitness and the other financial benefits it offers, as well as incorporating health-related gaming into their lives. Therefore, it can be considered that they do not have any serious health concerns.
- **Cluster 2 users:** Most men are middle-aged, married and have children. They also have full-time jobs, tend to live in urban areas and are considered relatively well-off. They have a high level of non-entertainment satisfaction with their applications and are willing to recommend them to others. They are more concerned about their own health and activity status and are in the minority overall.
- **Cluster 3 users:** This group is also clear about the purpose of using MHAs and is willing to implement health initiatives at a cost and is the most health-conscious compared to the other groups. In addition, nuclear family households are married but have no children.
- **Cluster 4 users:** This group is the largest majority. They are ambiguous about the purpose of their use, have low levels of satisfaction and do not want to spend much money. This

suggests that they have just started working on their health or are somewhat interested in health but do not know what to do.

Classification model

A classifier was created to suggest applications matched to users in a random forest. Although the number of data was small and the data was unbalanced with a lot of variation in each objective variable, by using SMOTE, the Accuracy, average of Precision and Recall went up to 86%. Cluster 2 and cluster 4 had the highest Precision, marking 89%. Cluster 3 had the highest Recall of 95%. Details are given in the table below. We also compared the results with SVM, SV Linear and Naive Bayes in practice, and Random Forest was still the best model.

	Random Forest	SVM	Linear SV	Naïve Bayes
Accuracy	0.86	0.72	0.63	0.38
Average Precision	0.86	0.72	0.62	0.4
Average Recall	0.86	0.71	0.62	0.38

Table 3: Performance score report of Random Forest.

To test this pilot model, seven project members were asked to complete a questionnaire focusing only on the items that were explanatory variables. The proposed application by Random Forest was then compared with the actual application used or the application they wanted to use. Unfortunately, the results were 43% accuracy, 58% average precision and 44% average recall.

Discussion

The analysis revealed user characteristics corresponding to objectives 1 and 2, the types of MHAs. Some studies have examined differences in the evaluation of application features based on user characteristics [21], but this focuses on features and does not examine application's functions' correlation. As the results of this study show, there is a correlation between the features that MHAs have, and the balance between them is the source of the classification of the application by K-means clustering. Some studies have investigated how much awareness there is of the existence of MHAs and why they are not being used [22]. In cases where the existence of such a wide variety of applications is not known, the ability to propose MHAs, such as this classifier, can be an opportunity to learn about their existence and generate interest.

Conclusion

From this study, a detailed application survey based on user responses revealed that MHAs can be categorised into five types, and for each group of applications, different user personas were detected. Validation with a Random Forest classifier using the characteristics of those personas and explanatory variables allowed an accuracy of over 85%. It is possible to suggest appropriate applications for potential users and dormant users due to application mismatches. It can contribute to raising the age of health by encouraging people who are not yet interested in and engaged in health-related activities to use the system. Whilst the data of the case study was based in Japan, it is relevant to other nations. Limitations of this work include the sample size of this survey, which was relatively small and not necessarily representative of the whole of Japan. In addition, the questionnaire used in this research was designed not only for this study but also for several similar projects so was not a strong fit to the specific research focus of this study. Future work would be a larger survey using a more bespoke questionnaire. A further limitation is that the MHAs analyzed have not been verified whether they are of the quality expected by users. Therefore, future research should also consider the quality of applications belonging to the clusters using indicators to measure expected effectiveness values using the proposed mHealth education-specific evaluation framework [23,24].

Bibliography

1. Ministry of Health, Labour and Welfare. "About the Population of our Country". Ministry of Health, Labour and Welfare (2022).
2. Mobile Society Research Institute. "Over 50% Smartphone Ratio among People in their 70s". Mobile Society Research Institute. 12 June (2019).
3. Statista. "Number of mHealth Apps Available in the Google Play Store from 1st Quarter 2015 to 2nd Quarter 2022". Statista. July (2022).
4. Jabour Abdulrahman M., *et al*. "The adoption of mobile health applications among University students in health colleges". *Journal of Multidisciplinary Healthcare* 14 (2021): 1267.

5. World Economic Forum. "Global Gender Gap Report 2022". World Economic Forum. 13. July (2022).

6. Lee P., *et al*. "Mobile Health App use among Older Adults". Institute for healthcare policy and innovation National poll on healthy aging University of Michigan. February (2022).

7. Antezana Gaston., *et al*. "Do young men and women differ in well-being apps usage? Findings from a randomised trial". *Health Informatics Journal* 28.1 (2022): 14604582211064825.

8. Marketing Applications, Inc. "Surveroid". Surveroid (2022).

9. Macqueen WM. "Some observations about aims and methods in the teaching of building construction". *The Vocational Aspect of Secondary and Further Education* 1.2 (1949): 156-168.

10. Scikit-learn. "Sklearn.Cluster.KMeans". Scikit-learn Machine learning in Python (2022).

11. Hartigan John A and Manchek A Wong. "Algorithm AS 136: A k-means clustering algorithm". *Journal of the Royal Statistical Society. Series c (applied statistics)* 28.1 (1979): 100-108.

12. Nazeer KA Abdul and M P Sebastian. "Improving the Accuracy and Efficiency of the k-means Clustering Algorithm". Proceedings of the world congress on engineering. London, UK: Association of Engineers London 1 (2009).

13. Breiman Leo. "Random forests". *Machine Learning* 45.1 (2001): 5-32.

14. Sheykhmousa Mohammadreza., *et al*. "Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review". *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13 (2020): 6308-6325.

15. Kirasich Kaitlin., *et al*. "Random forest vs logistic regression: binary classification for heterogeneous datasets". *SMU Data Science Review* 1.3 (2018): 9.

16. Chawla Nitesh V., *et al*. "SMOTE: synthetic minority over-sampling technique". *Journal of Artificial Intelligence Research* 16 (2002): 321-357.

17. Viloria Amelec., *et al*. "Unbalanced data processing using oversampling: Machine learning". *Procedia Computer Science* 175 (2020): 108-113.

18. Statista. "Projected Global Digital Health Market Size from 2019 to 2025*". Statista. April (2019).

19. Van der Maaten Laurens and Geoffrey Hinton. "Visualizing data using t-SNE". *Journal of Machine Learning Research* 9.11 (2008).

20. Devassy Binu Melit and Sony George. "Dimensionality reduction and visualisation of hyperspectral ink data using t-SNE". *Forensic Science International* 311 (2020): 110194.

21. Gimpel, Henner, *et al*. "Understanding the evaluation of mHealth app features based on a cross-country Kano analysis". *Electronic Markets* 31.4 (2021): 765-794.

22. Kayyali Reem, *et al*. "Awareness and use of mHealth apps: a study from England". *Pharmacy* 5.2 (2017): 33.

23. Aljaber Tareq and Neil Gordon. "A Hybrid Evaluation Approach and Guidance for mHealth Education Applications". International Conference on Applied Human Factors and Ergonomics. Springer, Cham, (2017).

24. Aljaber Tareq., *et al*. "An evaluation framework for mobile health education software". 2015 Science and Information Conference (SAI). IEEE, (2015).