Applied Computer Science, vol. 20, no. 3, pp. 132–152
doi: 10.35784/acs-2024-33

Submitted: 2024-05-21 | Revised: 2024-07-25 | Accepted: 2024-08-06

Keywords: mobile health, Machine Learning, design science, cardiovascular disease

Shahil SHARMA[*], *Rajnesh LAL* [0000-0002-3034-9751][*],
Bimal A Kumar [0000-0003-3622-7541][*]

DEVELOPING MOBILE MACHINE LEARNING APPLICATION FOR EARLY CARDIOVASCULAR DISEASE RISK DETECTION IN FIJI: A DESIGN SCIENCE APPROACH

Abstract

Cardiovascular disease (CVD) has become a significant contributor to premature deaths for many years in Fiji. CVD's late detection also significantly impacts annual deaths and casualties. Currently, Fiji lacks diagnosis tools to enable people to know their risk levels. In this paper, a machine learning mobile application was developed that can be easily accessible to the local population for early prediction of CVD risk. The design science approach was used to guide the development of the application. The design process involved identifying the problem and motivation, setting objectives, creating a machine-learning mobile application for medical record analysis, demonstrating the application to selected participants, evaluating its usability and the machine-learning model's performance, and communicating the findings. The results revealed that the proposed machine learning application achieved a high usability score of 87 on the System Usability Scale, indicating strong user-friendliness and adaptability. The machine learning model by random forest algorithm demonstrated the accuracy of 89% and was selected for implementation for CVD prediction in Fiji, as it outperformed other algorithms in the study: k-nearest neighbour, support vector machine, decision tree, and Naïve Bayes. The results highlight the effectiveness and user acceptance of the developed system in Fiji's medical facilities for CVD prediction.

1. INTRODUCTION

Cardiovascular disease (CVD) has been a significant contributor to premature deaths for many years in Fiji (Ministry of Health & Medical Services, 2015). According to Taylor et al. (2018), CVD develops gradually over time due to lifestyle choices or modifiable risk factors like inactivity, a poor diet, and the use of tobacco or alcohol. People rely solely on medical professionals to diagnose problems, give prognoses, and recommend appropriate preventative actions. Currently, Fiji lacks diagnostic tools such as software, websites or mobile apps to help people become aware of their CVD risk (Sharma et al., 2023).

[*] Fiji National University, School of Mathematical & Computing Sciences, Fiji, shahil.sharma@fnu.ac.fj, rajnesh.lal@fnu.ac.fj, bimal.kumar@fnu.ac.fj

Mobile health (mHealth) technology has improved health outcomes for communities worldwide over the past decade (Hoque et al., 2020; Kaium et al., 2020). In preventing communicable diseases like CVD, mHealth incorporates mobile information and communication technology in healthcare and has the potential to alter how healthcare is delivered on an individual level (Alaiad et al., 2019). Developing countries face a shortage of healthcare resources, including nurses and doctors; mhealth platforms offer a practical substitute for conventional, in-person healthcare interactions (Armaou et al., 2020; Gumede et al., 2023; Kruse et al., 2019).

Machine learning has revolutionised medical healthcare systems, where it has been successfully used for disease prediction (Kosarkar et al., 2022; Uddin et al., 2019). As a critical tool in healthcare, machine learning aids in disease detection at hospitals and significantly impacts healthcare decision-making processes, providing accurate advice that could potentially minimise casualties (Curigliano et al., 2020). This research aimed to develop a machine-learning application to predict CVD risk levels.

This research employs the Design Science Research (DSR) methodology developed by Peffers et al. (2007) to create cutting-edge artifacts that address issues related to human concerns (Kumar & Goundar, 2022; Hevner & Gregor, 2022). This study used DSR methodology to design and develop a machine learning-based mobile application for predicting CVD in developing countries like Fiji. Fiji, a small South Pacific Island nation, faces significant challenges in its healthcare due to limited access to advanced medical technologies and a shortage of healthcare professionals, including doctors and nurses (Razzaq et al., 2024). The choice of DSR methodology was suitable for this study for several reasons. Firstly, it enabled the development and evaluation of artifacts aligned with the research objective to develop a practical tool to address CVD prediction challenges in Fiji. This allowed the resulting system to be theoretical and applicable, where medical practitioners could use the system to understand evolving CVD patterns and perform better diagnosis and treatment.

Secondly, the agile nature of the DSR methodology enables continuous improvements of the proposed system, which was crucial for enabling the system to adapt to changing and evolving needs for CVD predictions in Fiji, ensuring the proposed system is effective in long-run implementation. Additionally, the methodology allowed the authors to contribute to the existing literature by providing a clear framework for system development, which could be replicated in similar contexts in a country with limited healthcare facilities for CVD diagnosis. By developing a functional system, this research provides valuable insights that could be utilised in Fiji’s medical setting and similar contexts, contributing to academic and practical domains.

The development of the machine learning artifact is mainly covered in this work, providing insights into the methodology for implementing a CVD prediction system in developing nations like Fiji. The structure of this paper is outlined as follows: Section 2 reviews the existing literature on advancements in mHealth applications. Section 3 describes the approach adopted for the artifact's development, which describes the design and development of the machine learning model and mobile artifact. Section 4 describes the demonstration of the artifact with the selected participants. The artifact's evaluation and results are elaborated on in section 5. The practical contributions and future research works are presented in Sections 6 and 7, respectively.

2. REALATED WORK

A thorough study of relevant available literature was conducted to evaluate the current state of mHealth applications intended to address various health concerns. This review of earlier research work served several functions: It provided insight into the approaches and initiatives already being used in mHealth applications and made it possible to identify gaps and opportunities for innovation for this paper.

The NusaHealth mobile health system developed by Sumarsono et al. (2023) was developed to address non-communicable diseases in Indonesia's rural areas. It is discussed in the article along with its design and development. The initiative, created in association with academic institutions, medical facilities, and local communities, takes a patient-centered approach and employs the DSR methodology in its creation. A network connecting mobile devices to hospital information systems is also a part of the system. The results show that a comprehensive strategy encompassing all pertinent partners and stakeholders is required for successful implementation. NusaHealth, initially implemented in Yogyakarta province's rural areas, calls for new initiatives and regional health regulations to further improve community health.

The growing problem of cardiovascular diseases, particularly in developing nations, was discussed in the study by Islam et al. (2023). The research creates and develops a system comprising a wearable gadget and a mobile application using a DSR methodology. The system accurately classifies users into different CVD risk levels using the Internet of Things and Machine Learning (Islam et al., 2023). The system has been evaluated for efficacy, usability, and efficiency, making it technologically cutting-edge and user-centric. It provides a substantial advance in biomedicine for the early identification and control of CVD by allowing users to track their CVD risk levels in real time (Islam et al., 2023).

Designing and evaluating a mHealth application created to promote healthy behaviors and lifestyles, which are essential in preventing cardiovascular illnesses, were covered in a study by Tundjungsari et al. (2018). The application underwent several stages of iterative development using the Human-centered design methodology, from an initial low-fidelity to a final high-fidelity prototype. The User Experience Questionnaire (UEQ) evaluation produced above-average results for several measures, including beauty, clarity, efficiency, and dependability. The program did an excellent job of encouraging user interaction. However, it received lower marks for dependability and accuracy, indicating a need for additional validation by medical experts to guarantee the app's reliability and effectiveness.

Several research directions in designing and implementing mHealth apps to predict diseases in developing countries were discovered after examining the collection of existing prior literature. While several research used use cases and prototypes to convey their findings, these techniques are frequently needed to provide insights that can be implemented. These gaps in the existing body of work's methodology and evaluation frameworks are brought to light by these omissions, and our study intends to fill them to support the adoption of mHealth solutions in developing countries.

3. PROPOSED APPROACH

This research employed the design science methodology proposed by Peffers et al. (2007). DSR allows researchers to address critical human challenges by creating innovative artefacts, adding to the body of knowledge (Hevner & Gregor, 2022). The structure of the DSR is as follows: 1. Identify problem and motivation - a problem was identified to demonstrate the importance of developing the mobile application. 2. Define objectives of a solution – in this activity, the objectives to construct the mobile application were identified. 3. Design and development – this activity involved designing and developing a mobile application artifact. 4. Demonstration - demonstration of the mobile application artefact with a group of users to assess the usability. 5. Evaluation - monitor and evaluate the effectiveness of the mobile application in achieving the goal. 6. Communication - convey the issue and the significance of the artefact. Figure 1 shows the design science research methodology designed for this study.

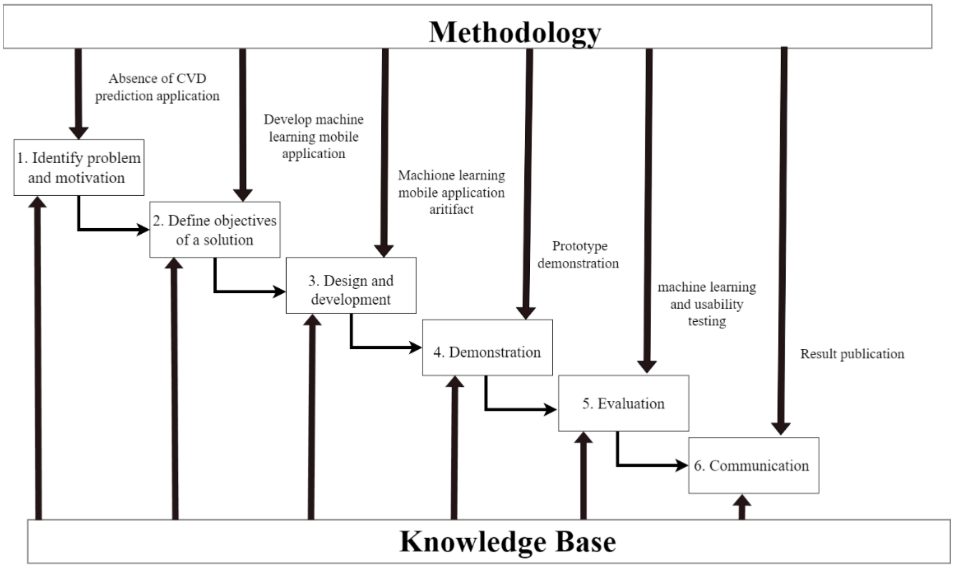

Fig. 1. DSR model to develop CVD prediction mobile application for Fiji

3.1. Activity 1: Identify problem and motivation

Unstructured interviews were undertaken to establish a research problem. Unstructured interviews were conducted to pinpoint existing gaps within Fiji's healthcare framework. The interview focused on three key components, which were: i) whether Fiji currently employs any diagnostic tools specifically for CVD prediction; ii) the extent to which existing diagnostic methods aid in CVD prediction; iii) the potential positive impact a dedicated CVD diagnosis tool might have on CVD prediction in the country. The conversations highlighted a critical issue: the late detection of CVD in Fiji, suggesting that machine learning technologies could be vital in mitigating this gap.

3.2. Activity 2: Define the objectives of a solution

Identifying the goals of the solution based on the defined problem and understanding of the practical and achievable aspects of the artifact (Peffers et al., 2007). The first objective was to Develop an Android mobile application that allows users to easily collect required medical data, analyse it, and disseminate the predicted CVD risk, with the machine learning model accessible to multiple users simultaneously. The second objective was to perform machine learning model accuracy testing and mobile application usability evaluation to determine the system's effectiveness for real-world implementation.

3.3. Activity 3: Design and development

The artefact was designed and developed in two stages: a machine-learning model and mobile application development. The development of the machine learning model was carried out to determine the optimal model for implementation. Afterwards, the mobile application was developed to implement the machine learning models. Figure 2 shows the framework adopted to create the mobile application.

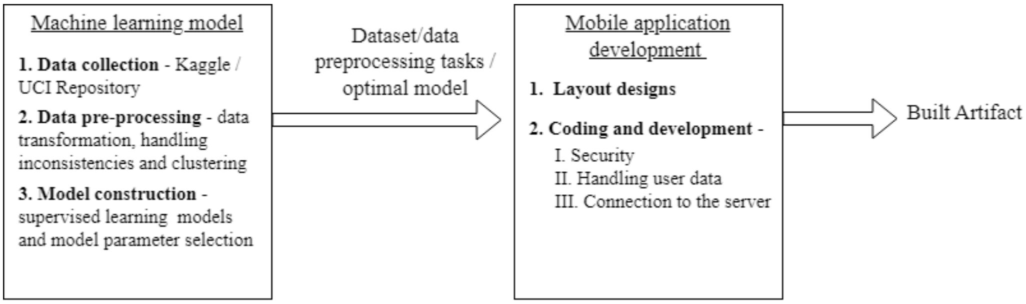

Fig. 2. Conceptual framework of the CVD mobile development process

3.3.1. Machine learning model

The machine learning model was constructed in three stages: (i) data collection, acquiring reliable datasets for building machine learning models; (ii) data pre-processing, correcting irregularities in the datasets and preparing the data for effective CVD prediction in Fiji; (iii) constructing the machine learning model.

The machine learning model was constructed using the Python programming language and libraries. The Python Pandas library was used to handle pre-processing data tasks. Pandas library has custom-made data pre-processing functions, making data pre-processing much simpler. The Python Sklearn library was used to construct machine-learning models. Six activities were conducted to construct the machine learning model, which was:

1. Data collection: datasets were collected from two reputable online repositories inspired by previous literature, which included UCI machine learning and Kaggle. A total of six datasets were chosen, which included the Heart Failure Prediction dataset, the Cleveland dataset, the Hungarian dataset, the Switzerland dataset, the Long Beach VA dataset with 200 instances, and the cardiovascular disease dataset. Two specific data selection criteria were considered when selecting the optimal datasets. The first factor was that the online datasets should share similar contributing risk factors to

those found in Fiji. The second factor was ensuring that both the quality and quantity of data were sufficient to test the model's scalability. The datasets selected were updated annually and consistently followed a standardised format, allowing the research team to combine the data into a single coherent source. The attributes used from these online datasets are detailed in Table 1, while Table 2 provides a detailed overview of the instances derived from various datasets.

2. Data pre-processing: data pre-processing was essential for solving inconsistencies in the data since the dataset was adopted from different sources. Missing values, duplicated data, data scaling, handling outliers, and transforming to suit the medical requirements of Fiji were problems in the dataset. There were five tasks related to data pre-processing, which included:

- Converting categorical data into numerical. Transforming categorical data into numerical form was essential, mainly because some machine learning algorithms cannot handle categorical data, which can impact model accuracy (Seger, 2018; McGinnis et al., 2018). In the datasets, the "Sex" attribute was converted from "M" for males and "F" for females to 1 and 0.
- Handling records with missing values. Duplicated data entries are standard and can lead to overfitting in machine learning models if used for training. The datasets had missing values. The chosen solution was to remove rows with these missing values.
- Removing outliers. Outliers are data points that significantly differ from the normal distribution and result in incorrect CVD diagnoses. Present in any feature attribute; they can impact the correlation between risk factors, leading to model biases (Sikder & Batarseh, 2023; Boukerche et al., 2020). The interquartile range (IQR) method removed outliers from the dataset. The IQR method calculated the range between the first and third quartiles to identify and remove outliers from a dataset. This helped the research team understand the data's central tendency and spread.
- Data scaling. The numerical attributes in the dataset vary in range and magnitude. For example, age values range from 28 to 77, blood pressure values range from 153 to 200, and cholesterol values range from 85 to 200. These differing magnitudes can impact the model's accuracy in predicting CVD risk. The robust scaling technique was employed to scale numerical data using the median and interquartile range (IQR), thereby transforming the numerical values to a consistent scale.
- Clustering CVD risks - the goal was to create a machine-learning model capable of predicting five CVD risk levels; however, the dataset used in the study did not contain risk levels. The optimal method for classifying risk levels in the existing records within the data frames is to utilise an unsupervised learning algorithm. This approach generated a labelled data frame, which was used to develop and test supervised learning algorithms. To handle this problem, the K-means unsupervised learning algorithm was utilised to cluster risk levels.
- Supervised model construction - six supervised machine-learning models were developed using the labelled data from the K-means models. The six supervised learning consisted of Decision trees, Support vector machines, Naïve Bayes, Random Forest, and K-nearest neighbour. The supervised model development process involved the selection of the best model parameters.

Tab. 1. Attributes obtained from online datasets

Attribute name	Data type	Description
Age	continuous	states that the patient's age ranges from 28 to 77
Sex	categorical	describes gender where 1 = Male 0 = Female
RestingBP	continuous	the resting blood pressure measured in mmHg when a patient has been admitted ranges from 0 to 200 mmHg
Cholesterol	continuous	cholesterol measured in mm/dl ranging from 0 to 603 mg/dl.
FastingBS	categorical	fasting blood sugar measured where fbs>120 mg/dl; 1 = true indicates the presence of diabetes, 0 = false indicates the absence of diabetes.
heart disease	categorical	target class determining cardiovascular disease :1 = present, 0 = absent.

Tab. 2. Dataset attributes detailed information

Datasets	Total Number of instances	Number of attributes	Number of duplicated instances	Instances of heart disease	Instances without heart disease
Cleveland	303	12	16	198	105
Heart failure	1911	14	22	1031	880
Hungarian	294	14	0	167	127
Switzerland	124	12	0	71	53
Long Beach VA	200	12	18	127	73
Cardiovascular disease	3888	14	24	2044	1844

3.3.2. Mobile application development

Predicting CVD is the main feature of the mobile application. Android Studio IDE, using the Java programming language, was used for this study. Once the technology was selected, three tasks were conducted to build the mobile application which was:

1. Layout design - user interface elements such as buttons, text fields, and menus were used to achieve maximum user interaction and more straightforward navigation. Figure 2 shows screenshots of the layouts developed for the mobile application.
2. Coding and development - the research team thoroughly inspected and refined the application during the development phase, ensuring its functionality and effectiveness in CVD prediction. Three main areas focused on the development phase were: 1. Security: user authentication and registration were critical components in developing a secure mobile application. The registration process allowed new users to set up an account in the mobile application, collecting the necessary information to identify them and establish a user profile. 2. Handling user data – mobile application users use a prediction form to input their medical data to send data to the model for prediction. Data validation was considered when retrieving user data from the prediction form, which checks for correct data formats, required fields, and consistency and manages any inconsistencies in data input. 3. Server development - The server was designed to prioritise scalability and security. The server was designed to perform three essential

tasks: i) Data pre-processing: correct inconsistencies when the user uploads new data. ii) Prediction: Construct a machine-learning model and make predictions using the medical data the user sends through the mobile application. iii) Saving medical data: A database was designed that enabled the server to store and retrieve medical data to find new emerging patterns of CVD. Figure 3 shows the machine learning model architecture.

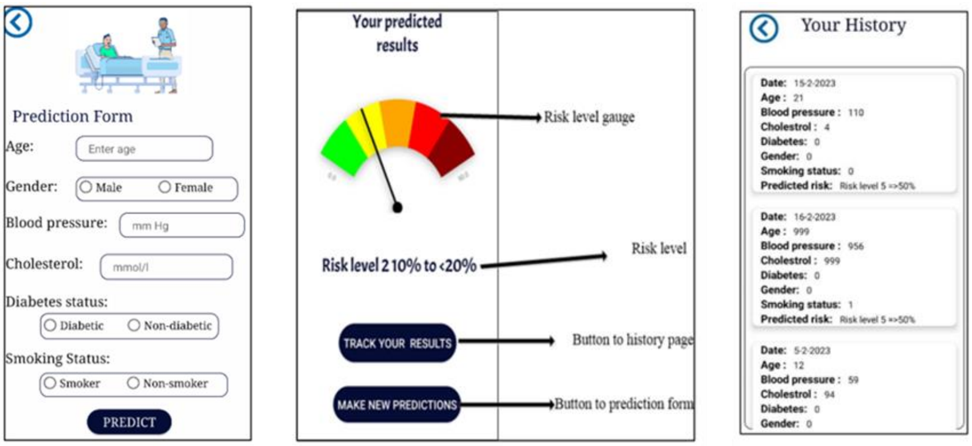

Fig. 3. Screenshots of the developed mobile application

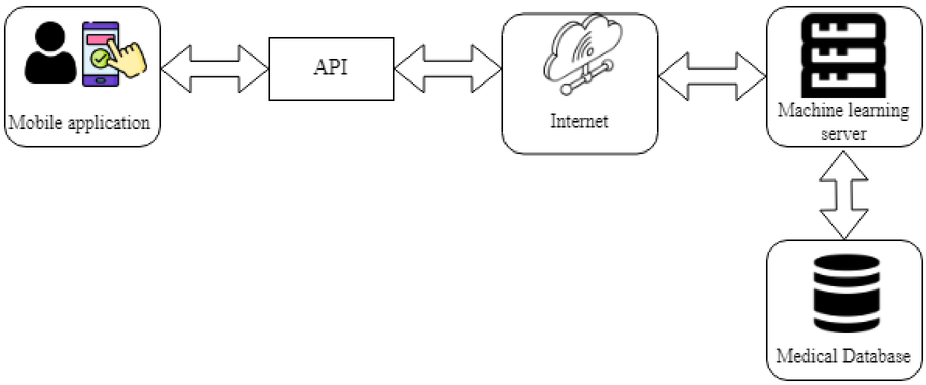

Fig. 4. Machine Learning model architecture

3.4. Activity 4: Demonstration

A total of 26 participants were chosen to evaluate the usability of the mobile application. The participants were divided into two primary categories. The first category comprised 80% of the participants and involved public members with varying experience with mobile applications. This group included individuals at risk of CVD and those without any identified CVD risk of different age groups. Specifically, ten individuals aged 50-70, 8 aged 30-40, and 5 aged 20-30.

The second category of participants, made up of the remaining 20%, were medical practitioners from large hospitals and remote health clinics. The intention behind this diverse mix of participants was to gather a broad spectrum of perspectives and feedback regarding the mobile applications' usability. The mobile application was installed on the participant's smartphone, and then the participants were required to perform a series of tasks to evaluate the mobile application's usability. The participants from both groups received training on using the mobile application and were monitored for two months to observe their usage.

3.5. Activity 5: Evaluation

3.5.1. Usability testing

Usability was evaluated using two methodologies: observing participants through evaluation tasks and using the System Usability Scale (SUS). The evaluation tasks included observations on how participants interact with the proposed system and collection of feedback to identify any issues with usability, user experience challenges, or possible improvements to enhance user experience. The observation usability testing was conducted as follows:

1. Selecting participants. A total of 26 participants were selected to assess the proposed mobile application, which aimed to assess users from different backgrounds. The participants were categorised into two groups. The first group included 20 public participants. This group featured individuals at risk of CVD and those with no identified CVD risk, covering various age ranges (e.g., the second category consisted of 6 medical practitioners employed at large hospitals and remote health clinics, specifically general practitioners who conduct routine medical checkups. Table 3 presents the demographics of the surveyed subjects, including age, gender, professional background, and CVD risk status.
2. Confidentiality and consent – participants were informed about the purpose of the study and the importance of the data collected through them. A detailed consent form from the author's institution was provided, which outlined their participation rights and measures taken to ensure confidentiality and voluntary participation. Each participant signed the confidentiality consent form before conducting a usability test. Data was gathered through a survey form, which included participants' background information such as age, gender, area, occupation, and health status. The Fiji National University ethics committee provided ethical approval to ensure all data collected followed ethical standards and protect participants' rights and privacy. Participants were encouraged to use the synthetic medical data provided to them. Since our proposed system required an early-phase usability test, the main motive was to capture user experience in our initial research phase; hence, synthetic data was used in the study.
3. Demonstration/testing environments – a live demonstration was carried out with a few participants in a computer laboratory of the author's institute, and video recordings were emailed to remote participants. The usability testing was conducted in two different environments, urban and rural, to ensure diverse testing conditions that replicate Fiji's health context.
4. Evaluation tasks - the participants were required to perform a series of tasks, and feedback was collected based on the tasks through a questionnaire. Table 4 shows a list of tasks conducted in numerical order.

Tab. 3. Demographics of selected participants

Demographic Characteristic	Category	Number of Participants
Age	50 - 70	10
	30 - 40	11
	20 - 30	5
Gender	Male	15
	Female	11
Tech Savviness	Intermediate	15
	Advanced	11
Occupation	Healthcare	6
	Other	20
Health Status	Diagnosed with CVD	18
	No known conditions	8
Education	High school and below	16
	Undergraduate	8
	Postgraduate	2

Tab. 4. Mobile application evaluation tasks

Task	Description
Registration and Setup	Download and install the mobile application and create a new account by providing personal information such as email and password.
Data input	Enter medical data into the application. The users were required to find the prediction form and fill in the details. After completing the form, click the predict button at the bottom of the page, which displays predicted CVD results.
Interpreting predicted results	This task assessed how participants interpret their predicted risk level of CVD. The risk was shown on a gauge indicating various levels and colour representing different CVD risks.
Interpreting previously predicted results	This task assessed participants' understanding of their previously predicted CVD risk levels. The mobile application records each user's past CVD risk predictions in a database. Participants were required to go to the history page, which displays all past predictions chronologically, and check if the predictions list was understandable.

Afterwards, a post-evaluation was conducted using SUS, a dependable, cost-effective usability scale suitable for the worldwide evaluation of system usability artifacts (Thamilarasan et al., 2023; Zulzalil et al., 2023). This study followed a traditional test-and-measure methodology. Tasks, structured as questions, were created to evoke a form of interaction between participants and the proposed system for investigation, which was inspired by the works of (Del Mar-Raave et al., 2021). The questionnaire was divided into two major sections: 1) Questions associated with overall usage of the mobile application based on the evaluation tasks to get feedback on user perceptions, and 2) SUS, which consisted of ten questions. Participants selected the option for each statement that most accurately reflects their feelings and reactions when using the mobile application, and participants answered the questions using a scale ranging from 1 (Strongly Agree) to 5 (Strongly Disagree), with three being Neutral.

3.5.2. Machine Learning model testing

With a dataset of 6,000 records, the study outlined entailed carefully evaluating machine learning models. Five hundred records made up the original subset at the beginning. Up to 2000 records were added to the dataset sequentially at 500 records per stage, with each subset having three different random iterations to create complexity. Four were produced. As a result, they were totalling 6,000 records. The method allowed for cross-evaluations in the five selected machine-learning models while investigating various data circumstances. The procedure improved comprehension of machine learning scalability by revealing how the models responded to changes in the data. The performance metrics for each model were calculated, giving a view of how each model performs across unique datasets of the same size. To optimize the performance of the selected machine learning models, a grid-based search with 10-fold cross-validation was employed to identify the best parameter combinations. Each model was trained and tested ten times on different portions of the dataset for 10-fold cross-validation, giving an in-depth review of its performance. The benefit of using the whole dataset for training and testing, providing a reliable assessment of model performance, and reducing the risk of overfitting the training data made this approach ideal for testing the model's accuracy.

Because there are multiple CVD risk levels, comparing Precision, Recall, and F1-Score for each level could be complex and might not clearly show how well the models perform. Instead, the accuracy of each model was compared to that of the other to determine the best model for implementation. The model's accuracy was assessed in the study, and essential performance indicators, including Precision, Recall, and F1-Score, were carefully studied. These metrics offer a more thorough understanding of the model's classification abilities. Precision is the percentage of actual positive instances among those the model has classified as positive. In this paper, precision was calculated using Equation 1:

$$precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (1)$$

Instances where the model correctly predicted the positive class are called true positives. It is classified as a True Positive for that risk level, for instance, if the model predicts Level 5 danger for a sample and the ground truth indicates Level 5. When the model misclassifies a negative instance as positive, False Positives happen. For example, A False Positive for Level 5 in this study would occur if a sample belonged to Level 2 risk, but the model projected that it belonged to Level 5. Recall, called Sensitivity, measures how well a model can recognise all real positive cases. In this paper, the recall was calculated using Equation 2:

$$recall = \frac{True\ Positives}{True\ Positives + False\ Negative} \quad (2)$$

3.6. Activity 6: Communication

This paper focuses on the technical aspects of creating a prototype mobile application suitable to predict CVD in Fiji and its essential design requirements. The primary goal of this study was to develop an open-source CVD prediction system designed to be replicated in other developing countries facing challenges in healthcare technology. The development of the artifacts was structured following the DSR methodology, ensuring that the research is detailed and interpretable by other researchers in the field.

4. RESULTS EVALUATION

4.1. Usability testing results

The usability score was analysed using two basics: the first focused on the user response to the mobile evaluation task, and the second on the SUS score.

The first analysis used the findings as a strategic blueprint for future enhancements to develop a more user-focused machine learning mobile application. A detailed analysis was conducted on each question in the first section of the usability questionnaire. The individual responses from all 26 participants were examined to gain insights into various user experiences and perspectives. Figure 4 shows the summary of the responses gathered by the participants based on the mobile application evaluation tasks. Table 5 elaborates on the responses gathered by users based on each task used for the next development phase of the mobile application.

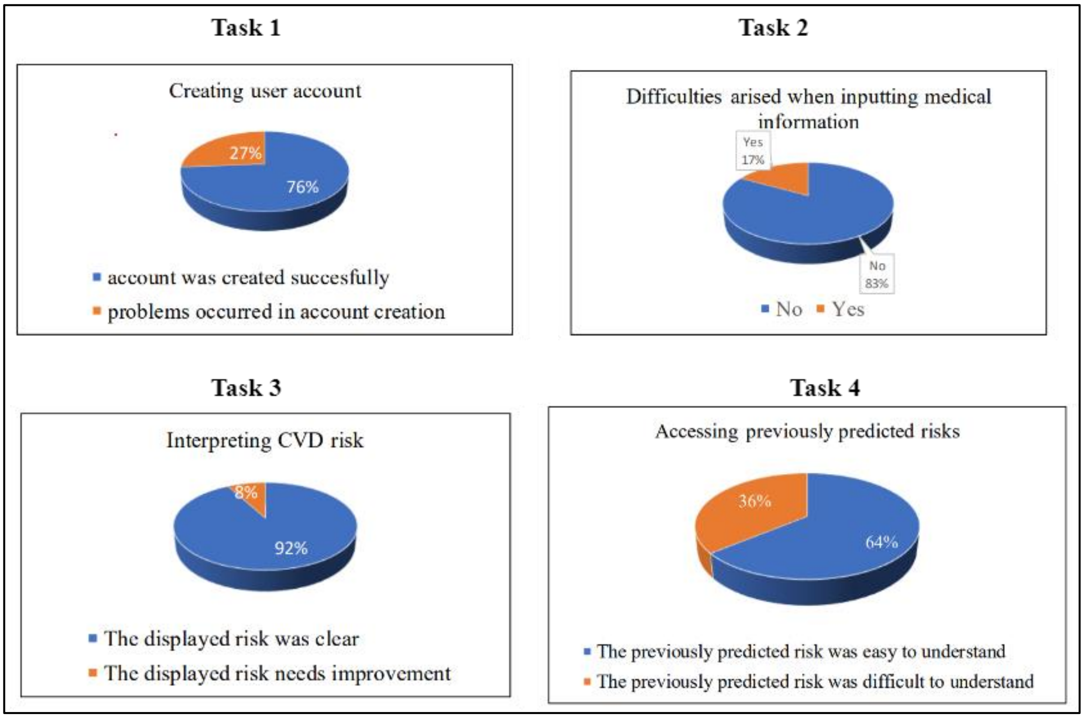

Fig. 5. Usability results based on the mobile application evaluation tasks

Tab. 5. Responses collected from users based on each mobile application evaluation task

Task	Response
Registration and Setup	Participants stated that they faced difficulties with password creation and that the password criteria were not explained. Participants have also stated that a button is needed to reset their password.
Data input	The data format required by the systolic blood pressure and cholesterol input fields was in mmHg and mmol/l formats, but certain users were unsure of the format. The participants also highlighted visibility issues with input fields and fonts; both input fields and fonts require a bigger size for clear visibility.
Interpreting predicted results	The participants have suggested that the precise percentage of the CVD risk be calculated instead of showing the approximate range of risk.
Interpreting previously predicted results	The participants who faced difficulties stated that how past prediction results are displayed is confusing and that the previous prediction should be displayed in ascending order, showing the latest predictions first.

For the SUS, each question was rated from 0 to 4. To calculate the score, the research team subtracted one from the responses to odd-numbered questions and subtracted the responses to even-numbered questions from 5. This way, the research team handled the survey's positive and negative statements. The overall SUS score was calculated by summing these adjusted scores and multiplying the total by 2.5, effectively scaling it from 0 to 100. Scores higher than 68 are typically seen as above average, while scores below 68 are regarded as below average (Blattgerst et al., 2022; Dasmen et al., 2021; Del Mar-Raave et al., 2021).

While the SUS score provided a general measure of system usability, it lacked specific context and was challenging to interpret. Therefore, some studies recommended using an adjective rating scale to contextualise the SUS scores. In this study, the SUS results included adjective-adjusted scores adopted from the works of (Blattgerste et al., 2022; Dasmen et al., 2021). Figure 12 displays the usability scores for the proposed system, including overall usability, learnability, and System Usability Scale (SUS) scores across 26 participants from P1 to P26. The usability and learnability scores vary among participants but generally remain high and share a positive relationship, often above 80%, indicating an intuitive interface design, leading to higher user satisfaction, efficiency, and adoption in Fiji's health context. The SUS scores consistently stayed above 75, suggesting overall solid usability. The average SUS score was approximately 85, significantly higher than the baseline of 68, demonstrating the proposed system's more excellent usability than the standard benchmark. Despite some variability among individual participants, the close alignment between usability and learnability scores highlighted the proposed system's robust usability and learnability. The maximum usability score recorded was 96, while the minimum score was 73. However, the high scores across all participants suggested that users found the application easy to use and quick to learn, essential for user satisfaction and efficiency.

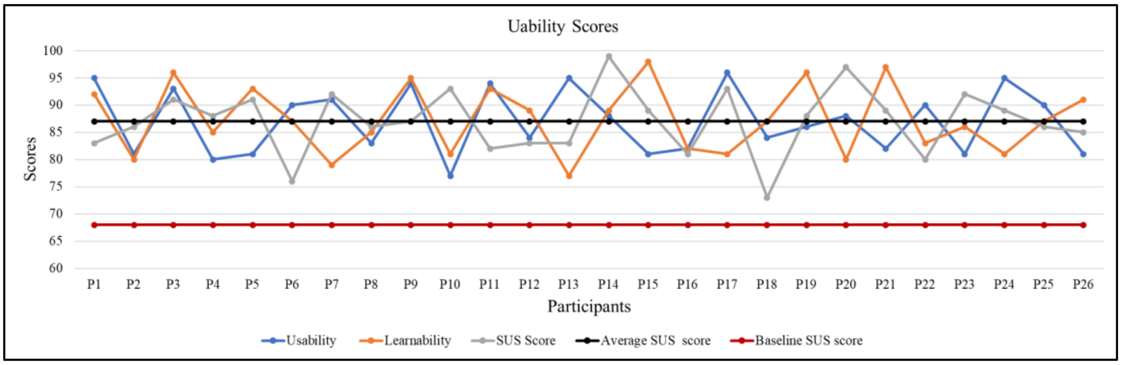

Fig. 6. SUS score results

4.2. Machine Learning model testing results

The accuracy of various machine learning models was assessed based on the best parameters, allowing a performance comparison. Using three distinct datasets, the performance metrics for each model were calculated, giving a view of how each model performs across unique datasets of the same size. The process ensured that the evaluation of the models was both thorough and fair, considering the inconsistency in the datasets. Table 6 shows the best parameters used to construct each model. Figure 8 displays each model's average performance metrics results using the three subsets. The five models can be easily compared using this methodology in their most optimal parameter settings, providing a clear understanding of each model's potential.

In this study, the models were designed to predict five levels of CVD risks: 1, 2, 3, 4, and 5. Precision refers to the model's accuracy in predicting each specific level of risk. Precision measures the proportion of correct optimistic predictions out of all the predictions that the model labelled as positive for a particular risk level. Recall refers to the model's ability to identify all actual cases of each specific risk level correctly. Recall measures the proportion of actual positives that the model correctly identifies. F1-score is a metric that averages precision and recall. Because there are multiple risk levels, comparing Precision, Recall, and F1-Score for each level could be complex and might not clearly show how well the models perform. Instead, the accuracy of each model was compared to that of the other to determine the best model for implementation.

Tab. 6. Optimal model parameters obtained from grid-based search

Model	Parameters
KNN	Neighbours = 2 for smaller dataset and 3 for larger dataset, distance metric = Euclidean distance
Decision tree model	Criterion = Gini (CART)
Random Forest model	number of trees = 100, criterion = Gini (CART)
Naïve Bayes	Gaussian
SVM	kernel functions = Polynomial

Model	Dataset 1 (500 instances)				Dataset 2 (1000 instances)				Dataset 3 (1500 instances)			
	Precision	**Recall**	**F1-score**	**Accuracy**	**Precision**	**Recall**	**F1-score**	**Accuracy**	**Precision**	**Recall**	**F1-score**	**Accuracy**
KNN	0.73	0.75	0.75	75%	0.74	0.76	0.76	78%	0.85	0.84	0.85	83%
Naïve Bayes	0.73	0.71	0.72	75%	0.68	0.71	0.69	73%	0.63	0.60	0.61	72%
Random Forrest	0.82	0.78	0.80	86%	0.89	0.84	0.86	90%	0.82	0.86	0.84	89%
Decision Tree	0.88	0.86	0.86	86%	0.89	0.88	0.89	87%	0.86	0.87	0.86	89%
SVM	0.74	0.76	0.76	75%	0.63	0.65	0.65	69%	0.54	0.51	0.51	66%

Fig. 7. Model evaluation results

This approach gives a more complete and fair evaluation of each model's performance by considering how often each risk level appears in the dataset and how vital each level is. During the testing phase with 500 records, the KNN model achieved a 75% average accuracy, likely due to its effectiveness with smaller datasets and using the Euclidean distance metric. The most successful model was the Random Forest, which averaged an impressive 86% accuracy, utilising 100 decision trees and bagging techniques to reduce variance. The Naive Bayes model also performed well, reaching 75% accuracy with the Gaussian Naive Bayes approach, potentially due to the independence of features. Using a polynomial kernel, the SVM model attained a 75% average accuracy, effectively handling linearly inseparable CVD risks. Lastly, the Decision Tree model equalled Random Forest's 86% accuracy, benefiting from the clarity of the Gini criterion and its suitability for smaller datasets.

When testing with a larger dataset of 1000 records, the KNN model's accuracy increased from 75% to 78%, likely due to more reliable majority voting with closer neighbours. The Random Forest model also showed improvement, rising from 86% to 90% accuracy, continuing to use 100 decision trees with the Gini criterion. On the other hand, the Naive Bayes model decreased average accuracy to 73%, potentially because of higher prediction errors with the larger dataset. The SVM model's accuracy dropped to 69%, whereas the Decision Tree model experienced a slight increase to 87% accuracy. In the last testing phase with 1500 records, the KNN model's accuracy increased significantly from 78% to 83%, probably due to a more diverse range of nearest neighbours. The Random Forest model's accuracy improved from 90% to 91%, likely benefiting from the larger dataset to create more diverse decision trees. In contrast, the Naive Bayes model saw a slight drop in accuracy to 72%. The SVM model's accuracy decreased from 69% to 66%, indicating potential challenges with high-dimensional data. The Decision Tree model showed a slight improvement to 89%.

After comparing accuracy across five machine learning models, the most suitable model for implementation was the Random Forrest model, created using the Gini criterion parameter and 100 trees. Random Forrest outperformed other machine learning models in terms of accuracy using different datasets. Figure 9 shows the accuracy comparison between the machine-learning models.

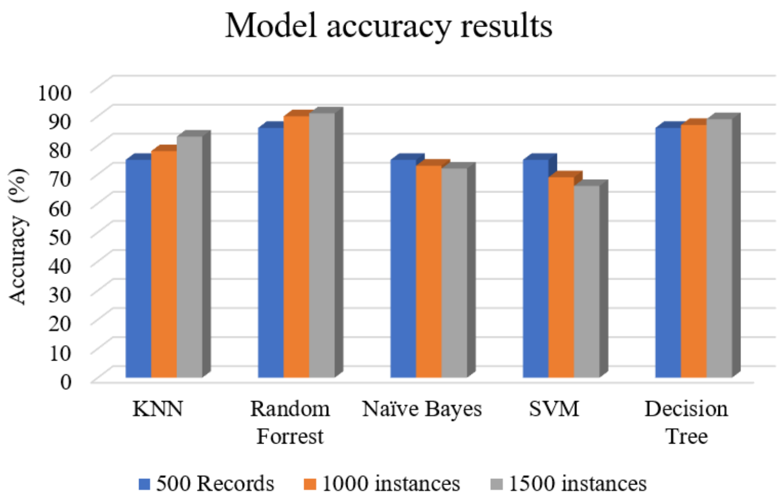

Fig. 8.Model accuracy comparison

4.3. Limitations of the proposed system

1. **1.** Batch model training. The CART Random Forest algorithm was selected to create a machine learning model in this study; however, this algorithm performs batch training as the model has to be completely retrained every time new data is added. This is unsuitable for implementing machine learning applications in the real world. As the model needs to be retrained completely, it impacts memory and computational power, which could be costly for handling large medical datasets in the future.
2. **2.** Security. Security concerns need to be addressed due to the sensitive nature of healthcare information and its accessibility through the Internet and mobile technology systems, which is causing significant concerns (Sumarsono et al., 2023). The prototype artifact developed in this study uses the default firewall provided by the cloud service provider, which could cause network breaches and increased vulnerability to cyber-attacks.
3. 3. Network connection issue. The mobile application requires a constant and stable internet connection to function correctly. During the usability testing phase, participants from remote areas have stated that the mobile application takes a while to respond due to poor internet connection. Since some remote areas in Fiji have poor reception, this could cause problems with the CVD prediction. Edge computing could be utilised to solve this issue.

4.4. Technical challenges addressed

During the development phase of the artifact, key technical challenges were addressed. The research team aims to elaborate on these issues to assist other researchers and practitioners working on similar studies. Listed below are a few difficulties encountered:

1. Model parameter selection. Parameter values influence the model's behaviour significantly influence the overall performance of the model's outcomes, and are impacted by the number of training data. Compared to previous studies, this research

aimed to develop a system that utilizes any data size for model training to make it more practical. Selecting optimal parameters for different training data sizes was critical in this research. Hence, data division and hypermeter tuning were used to solve the issue. The dataset was divided into three sizes, 500, 1000, and 1500 records, allowing for a simulation of the real-world scenario. Afterwards, hyperparameter tuning was combined with 10-fold cross-validation; this methodology provided a practical approach for hyperparameter tuning. 10-fold cross-validation, each model was trained and evaluated ten times on different parts of the dataset, offering a comprehensive assessment of different parameter performances.

2. Machine learning hosting services. Locating a compatible server hosting service was difficult as limited technologies support customised machine learning model hosting with mobile application support. To solve this issue, the research team developed a custom server environment using the services provided by Linode Cloud Hosting. Linode was utilised for server deployment, setting up, managing, and scaling virtual machines (VMs) within its cloud infrastructure. VMs were utilised to host the server components and communicate with the mobile application. The VMs allow concurrent operations of multiple operating system instances on a single system. VMs were established using the Linux-based Ubuntu Server operating system.

5. CONTRIBUTION TO THEORY AND PRACTICE

This study presented a novel artefact tailored for the early detection of CVD in the healthcare contexts in Fiji, attempting to fill a crucial void in healthcare technology by employing machine learning algorithms and predictive models specifically adapted to Fiji's unique demographic and health environment. This research aimed to assist healthcare practitioners in Fiji by improving personalise awareness and risk prediction for CVD, facilitating a more efficient diagnostic process, optimising palliative care, customising follow-up timelines, and utilising projected outcomes for better-informed clinical decisions. The mobile application created in this research holds the promise for extensive utilisation, offering substantial benefits to healthcare institutions and individual users by significantly contributing to enhancing preventive care practices through risk prediction.

6. CONCLUSION AND FUTURE WORK

While this study achieved significant progress in developing a machine-learning mobile application for CVD risk prediction in Fiji, there remains room for further enhancements. The system is currently in its prototype stage, and its reliability in real-world applications has not been established. This is primarily due to the need for long-term empirical clinical trials to verify the system's effectiveness, which would involve comparing mobile application’s predicted CVD risks with actual incidences. Due to the limited timeframe of the study, conducting extensive, long-term clinical studies to evaluate the model's reliability in real-world scenarios was not feasible. This approach will assist in identifying the most effective machine-learning model for practical application. In future works, the model will undergo enhancements for clinical application, aiming to transform the prototype into a dependable tool that can be deployed for early detection of CVD.

Artificial Intelligence (AI) is set to influence the future of mobile health applications significantly. AI’s ability to enhance the accuracy of medical diagnoses and treatment for various conditions signifies a transformative shift in patient care (Poalelungi et al., 2023; Wen & Huang, 2022). By enabling personalised medicine, AI's integration into mHealth promises to make healthcare more efficient and patient-centered, highlighting its essential role in advancing medical practices and outcomes. Mobile health initiatives often start as small pilot projects within the broader scope of electronic health (Ma et al., 2023).

The development of the mobile application in the research project has shown it can be effectively replicated and scaled up within the community, integrating well with existing health practices. The developed mobile application, with its high usability score and scalable machine learning models, underscores the effectiveness and user-friendliness of applying design science research principles in health app development for enhancing app utility and user experience.

However, it is essential to acknowledge the limitations of the proposed system and concerns relating to online model training, security, and network connection. Therefore, an additional future is essential for addressing the limitations and improving the reliability of the developed artifact in large-scale implementation in Fiji. Future work could lie in the following:

1. Online training model for batch training: methods such as deep learning algorithms, like neural networks that enable online training, are much more suitable for this approach. Online training allows a model to learn and adapt to new training data without causing significant changes to the model (Mehmood et al., 2021). With time, increasing prediction accuracy results from this adaptability, which is particularly helpful for dealing with shifting trends or patterns in the data (Thanga et al., 2021). Building a neural network from scratch was impractical due to the limited time and resources available for the study. Hence, there is a need to test and implement deep learning algorithms for future work.
2. Advanced security measures: sophisticated security technology, such as virtual private networks and intrusion detection and prevention systems, must be implemented to handle cyber-attacks for future work.
3. Edge computing to solve network connection issues: a connecting point within the Internet of Technology framework, forming a link between the physical world and digital spaces. It tackles issues such as quick response times, extending battery life, reducing bandwidth expenses, and ensuring data security and privacy (Hu et al., 2019). Edge computing minimises latency by handling data processing on the mobile application rather than depending on a remote cloud server.
4. Adopting the proposed system to Fiji’s existing healthcare: the proposed system can utilise machine learning algorithms to predict CVD effectively; the next step would be to conduct a large-scale empirical study to identify additional healthcare needs and newer medical data for creating and fine-tuning a model which is suitable to cope with the evolving patterns of CVD in Fiji. The research time aims to collaborate with local healthcare practitioners, policymakers, the Ministry of Health Fiji, and government institutions to customise CVD risk factors and diagnoses for other regional and remote clinics in the South Pacific. Integrating the proposed system with the existing system is also essential for practical implementation. The machine learning models require quality and quantity data to train and find evolving patterns of CVD. Hence, it is

necessary to develop a customised electronic health record system allowing numerous users seamless data collection.

Author Contributions

The first author has done research data collection, methodology, model, and system development. The second author has done the result analysis and drafted the manuscript. The third author has done a manuscript review and editing.

Funding

This study was supported by a research project through the Higher Degree Research Scholarship provided by Fiji National University to the first author.

Conflicts of Interest

The authors declare that they have no financial or non-financial interests related to the topic, including honoraria, educational grants, speakers' bureaus, memberships, employment, consultancies, stock ownership, expert testimony, patent-licensing agreements, personal or professional relationships, affiliations, knowledge, or beliefs.

REFERENCES

Alaiad, A., Alsharo, M., & Alnsour, Y. (2019). The determinants of m-health adoption in developing countries: An empirical investigation. *Applied Clinical Informatics*, *10*(05), 820–840. https://doi.org/10.1055/s-0039-1697906

Armaou, M., Araviaki, E., & Musikanski, L. (2020). eHealth and mHealth interventions for ethnic minority and historically underserved populations in developed countries: An umbrella review. *International Journal of Community Well-Being*, *3,* 193-221. https://doi.org/10.1007/s42413-019-00055-5

Blattgerste, J., Behrends, J., & Pfeiffer, T. (2022). A web-based analysis toolkit for the system usability scale. *15th International Conference on PErvasive Technologies Related to Assistive Environments* (pp. 237-246). Association for Computing Machinery. https://doi.org/10.1145/3529190.3529216

Curigliano, G., Lenihan, D., Fradley, M., Ganatra, S., Barac, A., Blaes, A., Herrmann, J., Porter, C., Lyon, A. R., Lancellotti, P., Patel, A., DeCara, J., Mitchell, J., Harrison, E., Moslehi, J., Witteles, R., Calabro, M. G., Orecchia, R., De Azambuja, E., … Jordan, K. (2020). Management of cardiac disease in cancer patients throughout oncological treatment: ESMO consensus recommendations. *Annals of Oncology*, *31*(2), 171–190. https://doi.org/10.1016/j.annonc.2019.10.023

Dasmen, R. N., Fatoni, F., Wijaya, A., Tujni, B., & Nabila, S. (2021). Pelatihan uji kegunaan website menggunakan system usability scale (SUS). *ABSYARA: Jurnal Pengabdian Pada Masyarakat*, *2*(2), 146-158. https://doi.org/10.29408/ab.v2i2.4031

Del Mar-Raave, J. R., Bahşi, H., Mršić, L., & Hausknecht, K. (2021). A machine learning-based forensic tool for image classification design science approach. *Forensic Science International: Digital Investigation*, *38,* 301265. https://doi.org/10.1016/j.fsidi.2021.301265

Gumede, D. M., Taylor, M., & Kvalsvig, J. D. (2023). Causes and consequences of critical healthcare skills shortage in the Southern African development community. *Development Southern Africa*, *40*(6), 1174-1199. https://doi.org/10.1080/0376835X.2023.2203155

Hevner, A., & Gregor, S. (2022). Envisioning entrepreneurship and digital innovation through a design science research lens: A matrix approach. *Information & Management*, *59*(3), 103350. https://doi.org/10.1016/j.im.2020.103350

Hoque, M. R., Rahman, M. S., Nipa, N. J., & Hasan, M. R. (2020). Mobile health interventions in developing countries: A systematic review. *Health Informatics Journal*, *26*(4), 2792-2810. https://doi.org/10.1177/1460458220937102

Islam, M. N., Raiyan, K. R., Mitra, S., Mannan, M. R., Tasnim, T., Putul, A. O., & Mandol, A. B. (2023). Predictions: An IoT and machine learning-based system to predict the risk level of cardiovascular diseases. *BMC Health Services Research*, *23*, 171. https://doi.org/10.1186/s12913-023-09104-4

Kaium, M. A., Bao, Y., Alam, M. Z., & Hoque, M. R. (2020). Understanding continuance usage intention of mHealth in a developing country: An empirical investigation. *International Journal of Pharmaceutical and Healthcare Marketing*, *14*(2), 251-272. https://doi.org/10.1108/IJPHM-06-2019-0041

Kosarkar, N., Basuri, P., Karamore, P., Gawali, P., Badole, P., & Jumle, P. (2022). Disease prediction using machine learning. *10th International Conference on Emerging Trends in Engineering and Technology-Signal and Information Processing (ICETET-SIP-22)* (pp. 1-4). IEEE. https://doi.org/10.1109/ICETET-SIP-2254415.2022.9791739

Kruse, C., Betancourt, J., Ortiz, S., Valdes Luna, S. M., Bamrah, I. K., & Segovia, N. (2019). Barriers to the use of mobile health in improving health outcomes in developing countries: Systematic review. *Journal of Medical Internet Research*, *21*(10), e13263. https://doi.org/10.2196/13263

Kumar, B., & Goundar, M. S. (2022). Kid-learn: A mobile language learning application for pre-schoolers. *International Journal of Virtual and Personal Learning Environments*, *12*(1), 1-16. https://doi.org/10.4018/IJVPLE.314950

Ma, E.-Y., Kim, H., & Lee, U. (2023). Investigating causality in mobile health data through deep learning models. *2023 IEEE International Conference on Big Data and Smart Computing (BigComp)* (pp. 375-377). IEEE. https://doi.org/10.1109/BigComp57234.2023.00089

Ministry of health & medical services. (2015). *NCDs in Fiji*. https://www.health.gov.fj/ncds/ncds-in fiji/#:~:text=ncds%20in%20fiji&text=in%20recent%20decades%2c%20ncd's%20have,and%20those%20numbers%20are%20growing

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45-77. https://doi.org/10.2753/MIS0742-1222240302

Poalelungi, D. G., Musat, C. L., Fulga, A., Neagu, M., Neagu, A. I., Piraianu, A. I., & Fulga, I. (2023). Advancing patient care: How artificial intelligence is transforming healthcare. *Journal of Personalized Medicine*, *13*(8), 1214. https://doi.org/10.3390/jpm13081214

Razzaq, A., Travaglia, J., Raynes-Greenow, C., & Alam, N. A. (2024). Understanding Fijian health system challenges in the prevention of mother-to-child transmission of HIV services in the three tertiary hospitals in Fiji. *AIDS Care*, *36*(7), 954-963. https://doi.org/10.1080/09540121.2024.2331215

Sharma, S., Lal, R., & Kumar, B. A. (2023). Machine learning for early detection of cardiovascular disease in Fiji. *2023 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)* (pp. 1-6). IEEE. https://doi.org/10.1109/CSDE59766.2023.10487655

Sumarsono, S., Sakkinah, I. S., Permanasari, A. E., & Pranggono, B. (2023). Development of a mobile health infrastructure for non-communicable diseases using design science research method: A case study. *Journal of Ambient Intelligence and Humanized Computing*, *14*, 12563-12574. https://doi.org/10.1007/s12652-022-04322-w

Taylor, R., Lin, S., Linhart, C., & Morrell, S. (2018). Overview of trends in cardiovascular and diabetes risk factors in Fiji. *Annals of Human Biology*, *45*(3), 188-201. 10.1080/03014460.2018.1465122

Thamilarasan, Y., Ikram, R. R. R., Osman, M., Salahuddin, L., Bujeri, W. Y. W., & Kanchymalay, K. (2023). Enhanced system usability scale using the software quality standard approach. *Engineering, Technology & Applied Science Research*, *13*(5), 11779-11784. https://doi.org/10.48084/etasr.5971

Tundjungsari, V., Sofro, A. S. M., Yugaswara, H., & Putra, A. T. D. (2018). Development of mobile health application for cardiovascular disease prevention. *International Journal of Advanced Computer Science and Applications*, *9*(11). https://doi.org/10.14569/IJACSA.2018.091175

Uddin, S., Khan, A., Hossain, M. E., & Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making*, *19*, 281. https://doi.org/10.1186/s12911-019-1004-8

Wen, Z., & Huang, H. (2022). The potential for artificial intelligence in healthcare. *Journal of Commercial Biotechnology*, *27*(4). https://doi.org/10.5912/jcb1327

Zulzalil, H., Rahmat, H., Abd Ghani, A. A., & Kamaruddin, A. (2023). Expert review on usefulness of an integrated checklist-based mobile usability evaluation framework. *Journal of Computer Science Research*, *5*(3), 57-73. https://doi.org/10.30564/jcsr.v5i3.5816