Integrative Approach to Cardiovascular Health Management

Roland Appuhamy
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
rsdew.appuhami99@gmail.com

Tharindu Lakshan
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
lakshanjat1@gmail.com

Oshadha Amaraweera
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
oshadageeth@gmail.com

Parami Jayasekara
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
paraminawodya500@gmail.com

Kapila Dissanayaka
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka kapila.d@sliit.lk

***Abstract*— Heart diseases remain a significant global health challenge, emphasizing the importance of early intervention and prevention. In addressing this issue, a mobile application was developed with four main components focusing on predicting the risk of becoming a heart patient, predicting the risk category of a heart patient, food item identification and dietary recommendation, and suggesting best rated nearby pharmacies to fulfill prescriptions. All these components include trained Artificial Intelligence models to assure a reliable performance within each component. In predicting the risk of becoming a heart patient, a trained logistic regression algorithm was used. As for predicting the risk category of a heart patient, a trained K-Nearest Neighbors algorithm was used. For the components of food item classification and suggesting best rated nearby pharmacies, YOLOv5 and TextBlob were utilized respectively. Results show that the accuracies of the trained models in the components are significant enough for mass scale utilization of all the four components of our mobile application in improving life standards of heart patients.**

Keywords—Cardiac health prediction, predictive modeling, logistic regression, machine learning algorithm, YOLOv5

I. INTRODUCTION

Heart diseases remain a leading cause of mortality with cardiac arrest posing a significant threat to public health. Despite remarkable advancements in medical science, accurately predicting the occurrence of a heart attack and its subsequent trajectory remains a critical challenge.

Early prediction of heart attacks and forecasting the likelihood of future occurrences are imperative for effective intervention and management. Furthermore, individuals who have suffered heart attacks require vigilant monitoring and access to appropriate medical resources to mitigate further risks and ensure optimal recovery. Therefore, it is necessary to develop predictive models capable of identifying individuals at high risk of heart attacks and forecasting the likelihood of subsequent occurrences.

The complexity of heart disease necessitates sophisticated approaches for accurate prediction and prevention. However, existing predictive models often lack the granularity and specificity required for effective risk assessment. Hence, it is better to suggest a comprehensive solution that empowers individuals to make informed dietary choices and access necessary medications promptly. By addressing these concerns positively, we can enhance the efficacy of

preventative measures and improve outcomes either for individuals who are at a high risk of having heart attacks or for individuals who are in the recovery phase after being a victim of a heart attack.

In addressing the above concerns, we developed an easy- to-use mobile app that can serve as a hub to handle the four functional components: predicting the risk of becoming a heart patient, predicting the risk category of a heart patient, providing food identification and dietary recommendation, and suggesting best rated nearby pharmacies to fulfill prescriptions. This approach guarantees a convenient and accurate way for individuals to assess cardiovascular health frequently while providing insights into mitigatory actions for not to become a victim of sudden heart failure.

II. LITERATUE REVIEW

Frequent cardiovascular health assessment is highly recommended for heart patients to take necessary actions before it becomes life threatening to them. In doing so, Artificial Intelligence (AI) or Machine Learning (ML) based predictive modeling plays a vital role since prediction accuracy can be tuned to trigger higher patient outcomes. In a study published in 2018 shows that AI is revolutionizing healthcare by automating diagnosis, improving data quality, reducing patient care variability, and lowering costs. They focus on AI predictive models for chronic disease diagnosis, enhancing early detection and patient outcomes. The majority of studies used regression-based ML models for diagnosing Celiac Disease, with 45% using Support Vector Machine (SVM), 23% using K-Nearest Neighbors (KNN) and Naïve Bayes (NB) models, 18% using Linear Regression, and 14% using Random Forest (RF) models. RF models were particularly effective in predicting liver fibrosis stages and identifying the liver fibrosis index degree. Bayesian network models were also used to analyze Chronic Obstructive Pulmonary Disease (COPD) patient data. SVM models showed better accuracy in predicting exacerbations and COPD detection [1]. AI can also revolutionize cardiovascular imaging since it can improve diagnostic accuracy, and physician productivity. However, challenges like interpretability and potential harmful recommendations remain [2]. In the study that comes under [3], they review the use of machine learning in predicting and early detection of cardiovascular issues in cancer survivors considering cardiac dysfunction and ECG use. It also highlights the prospects of ML in cardio-oncology. In another study, they explore the use of traditional machine learning algorithms and advanced technologies for predicting cardiovascular disease risk estimation. It compares various classifiers with J48 achieving about 70% accuracy and Keras achieving about 80% accuracy [4]. In their study [5], they highlight that heart diseases are a major global mortality cause, with ECGs being the most cost- effective diagnostic tool. However, challenges like limited experts and complex interpretations hinder accurate diagnosis. In this scenario, machine learning can be incorporated to unravel those complex interpretations along with the full explanations of the machine leaning models so that the practitioners can trust the results.

According to the study in [6], AI models are transforming medicine, enhancing efficiency and allowing medical professionals to focus on complex tasks. AI models can be used for disease prediction, diagnosis, and decision-making, improving data quality, reducing costs, and forecasting chronic diseases. Model accuracy in various diseases is a critical factor in predicting outcomes. They highlight that the accuracy of their models on predicting diabetes can vary in the range of 73.1-91.6%. The models that predict cardiac diseases show an accuracy within the range of 84-91%. NB, RF, KNN, and SVM models have the highest accuracy in liver disease prediction. LR models have been effective in identifying depression reasons, while KNN models are better for identifying disease patterns. The study in [7] proposes a data-driven approach using machine learning and electronic medical records to reduce preventable hospital readmissions for heart failure, stroke, and pneumonia. The Naïve Bayes algorithm outperformed existing models, achieving an accuracy of 83.19% and an Area Under Curve (AUC) of 0.78 with the possibility to enforce further refinement and validation [7]. Correct and effective assessment of diabetes is crucial to having a healthier heart. To address this fact, a study has been conducted to compare six machine learning algorithms for early diabetes prediction on a medical dataset, aiming to identify the most effective model for improving diagnostic accuracy and aiding doctors and healthcare practitioners in making timely decisions [8]. In 2019, a study has been conducted on a cloud-based healthcare application for predicting heart disease using the University of California Irvine ML Repository dataset and healthcare sensors. Various classification algorithms, including J48, were tested, with the J48 classifier showing the highest accuracy, precision, recall, F-score, and kappa value [9]. Also,

in another study, it highlights the potential of machine learning in cancer diagnosis and prediction, but challenges remain in managing large datasets and unstructured information. Efficient machine learning approaches could revolutionize healthcare by providing precise insights into symptoms, diagnoses, procedures, and treatments [10].

It is clearly noted that all the research mentioned above show the way that AI can be useful in predicting or assessing various diseases including cardiovascular diseases. However, an integrative approach in predicting, suggesting mitigatory actions, and aiding in dietary recommendations and medications are normally not addressed in single research itself. Hence, in this research, we suggest an AI base integrative approach to manage cardiovascular health by providing the features mentioned above with high accuracy.

III. METHODLOGY

In our research, we designed a mobile app for the convenience of heart patients and doctors focusing on risk assessment and control of cardiac failures. Here, the risk of a person becoming a heart patient can be found earlier. And a person who is at risk of heart disease should be advised to follow a dietary plan. Another unique feature here is the ability to automatically analyze the ECG reports and predict the risky situations that a person may face in the future with the help of previously obtained reports. Also, a patient can find the nearest pharmacy to get the prescription through this mobile application as well. The overall system diagram is shown in Fig. 1 and a brief description of the four main functional components is mentioned subsequently.

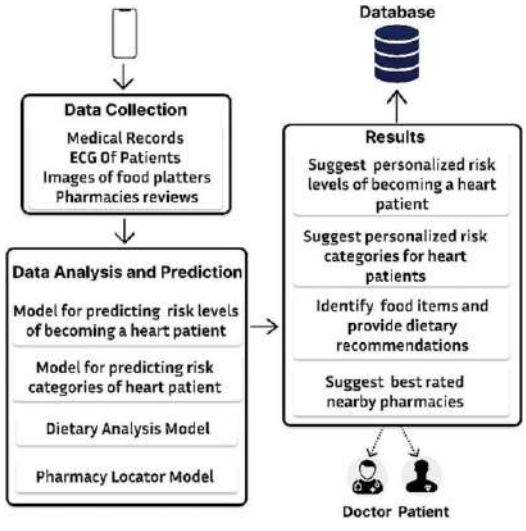


Fig. 1. Overall system diagram for cardiovascular health management.

A. Predicting the Risk Level of Becoming a Heart Patient

The methodology for developing a system for predicting the probability of becoming a heart patient involves several key steps. Firstly, a comprehensive survey was conducted among doctors from selected Sri Lankan government hospitals including Maharagama Apeksha Hospital, Ragama Hospital, Monaragala Sirigala Hospital to understand current diagnostic challenges and delays. Subsequently, data collection encompassed various demographic and health- related factors, including age, sex, cholesterol levels, blood pressure, heart rate, diabetes status, family history, smoking habits, obesity, alcohol consumption, exercise frequency, dietary habits, previous heart problems, medication use, stress levels, sedentary behavior, income, BMI, triglyceride levels, physical activity, and sleep duration. Data from multiple sources, including medical records and surveys, were collected and organized for analysis. This dataset is then preprocessed, including handling missing values, normalization, and feature engineering. Next, a logistic regression model was trained using the collected data, with feature selection and tuning performed to optimize predictive accuracy. The model's performance was evaluated through cross-validation and external validation on independent datasets to ensure robustness and generalizability. The flowchart for this component is shown in Fig. 2.



Fig. 2. Flowchart for risk levels of becoming a heart patient.

Subsequently, the model outputs were stratified into five risk levels (High Risk, Elevated Risk, Moderate Risk, Low Risk, and No Risk) according to WHO National Guidelines, enabling clear communication of individual risk status. Instructions for risk management and prevention were developed based on these risk levels, aligning with established guidelines to ensure clinical relevance and effectiveness. Finally, a user-friendly interface was designed to present the results in an easily understandable format, enhancing accessibility for both patients and healthcare providers. Iterative feedback and testing were conducted to refine the system's usability and efficacy, culminating in a comprehensive tool for predicting the possibility of becoming a heart patient and then providing actionable instructions for risk management.

B. Automated Analysis of ECG and Predictive Modeling for Future Heart Attack Risk

This component was developed using machine learning with expert knowledge to analyze ECG records. The required data was obtained from Maharagama Apeksha Hospital. Here, ECG is classified under four categories. They are Normal Person Heartbeat, Abnormal Heartbeat, History of Myocardial Infraction, and Myocardial Infraction. The first step is extracting 12 different leads of the ECG signal from the original ECG image. Each lead is extracted by defining a rectangular region of interest on the image using the slicing notation. Each lead image is then converted to a grayscale image, and the image is then smoothed to remove high-frequency noise. Those twelve leads are combined to create the 13th lead (see Fig. 3) and extract the signal from the Lead- 13 and save it as an image file, which is then ready to be used for further analysis, such as feature extraction and filtering. Then, a comma-separated values (CSV) file is created using the x and y coordinates of the waves in the thirteenth lead. A model was trained using K-Nearest Neighbors machine learning technique using the same CSV file. Here, x and y coordinates are used as input, and the ECG categorical group

as mentioned above is identified as output. The four categories mentioned at the beginning are classified under three risk levels, and a graph shows variation in them. The flowchart for this component is shown under Fig. 4.

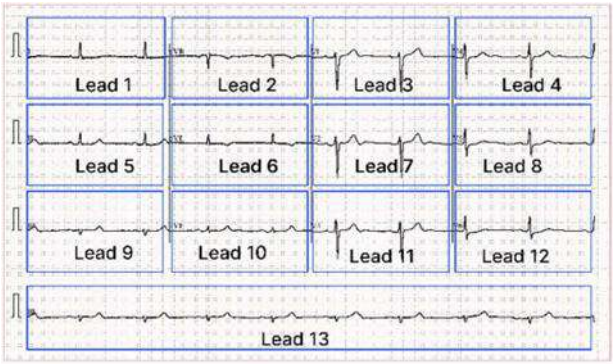


Fig. 3. Sample ECG with all thirteen leads.

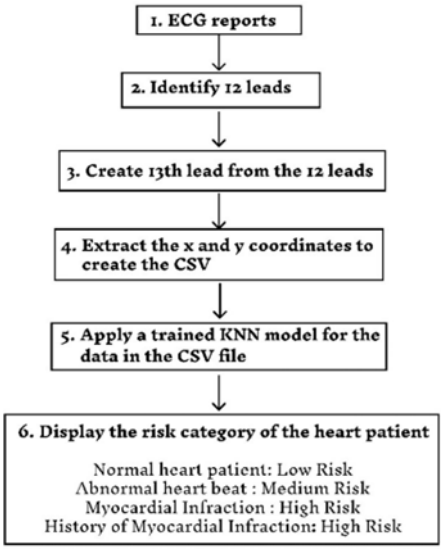


Fig. 4. Flowchart for automated analysis of ECG and predictive modeling.

C. Dietary Guidance System

The process of developing dietary guidance involves the following main steps. The first step is gathering a large amount of food images from various sources such as the internet and pictures manually captured from the camera, which forms the basis of the training dataset for the model. A test model is built and refined using approximately sixty images from reliable online repositories. After the model is validated, it is modified to account for the unique characteristics of Sri Lankan food, which requires a careful curation of one hundred images from the categories of meat, fish, rice, carrots, beetroot, eggs, and potatoes.

Every image in the dataset has been carefully annotated using expert annotation tools such as "MakeSense" to guarantee accurate labeling and efficient model comprehension. YOLOv5 architecture is used for model training, and Google Colab's processing power is utilized to maximize effectiveness and performance. The flowchart for this component is shown in Fig. 5.

After being integrated into the backend system, the trained model works in unison with the main features of the application, correctly recognizing the food items that are shown and offering relevant dietary recommendations based on the profiles of specific users. This essential feature enables users to make well-informed dietary choices that are in line with their individual nutritional needs and health goals.

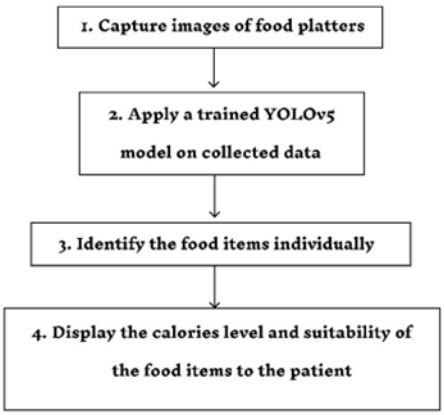


Fig. 5. Flowchart for dietary guidance.

D. Identify the Nearest Pharmacies with the Best Customer Recommendation

The overview of this model is described in Fig. 6, which facilitates different functions that are provided through the developed app. The input data for the model are user location, distance to nearby pharmacies obtained through Google Maps API, availability of required medicines and the ratings. Based on the distance, availability of medicines and the ratings, the pharmacies are listed upon the request of the user for an optimized result. Note that all these functions are performed within the main mobile application.

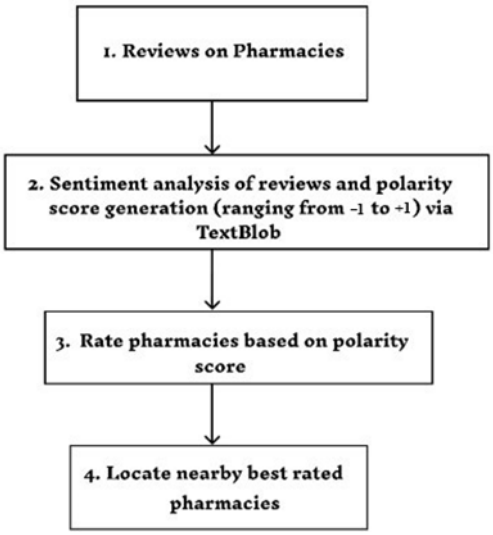


Fig. 6. Flowchart for generating the best rated nearby pharmacies.

The app creates a special token ID that is saved in Firebase and allows the device's token ID to be used to retrieve local pharmacies using Google Places API. Using Google APIs, GeoFire finds pharmacies within a given radius and sorts them automatically. While sentiment analysis uses rules and requires a pre-defined list of classified words, commonly supplied from databases like NLTK, TextBlob uses a machine learning approach to predict scores from user reviews. Sentiment analysis produces more accurate results by identifying sentiments in the input text based on word frequency and semantic correlations. TextBlob, a pre-trained Python library, is used to analyze the sentiment of patients' reviews. It determines whether a review is positive or negative by calculating its polarity score. This score ranges from □1.0 (indicating a highly negative sentiment) to +1.0 (indicating a highly positive sentiment). Based on this polarity score, reviews are classified as positive or negative. Pharmacies can then be ranked or positioned according to the average polarity score of their associated reviews. Based on this score, a corresponding sentiment rating will be assigned to each review. To display the higher rated Pharmacies,

Multinomial Naïve Bayes algorithm is applied after vectorization of the reviews.

IV. RESULTS AND DISCUSSION

The introduction of the proposed system shows several significant challenges faced by doctors and patients when managing health. The probability of having a heart attack was predicted through logistic regression model with training accuracy of 94% and testing accuracy of 93%. Stratification of predicted probabilities into five risk levels adheres to WHO National Guidelines, facilitating clear risk communication. Utilizing WHO National Guidelines, the model stratified predicted probabilities into five distinct risk levels: Red (High Risk), Orange (Elevated Risk), Yellow (Moderate Risk), Green (Low Risk), and White (No Risk). These risk levels facilitate clear communication of the individual's heart health status. Additionally, the system provides relevant instructions based on risk levels by considering WHO National Guidelines, ensuring appropriate actions and interventions are recommended for each risk category. The user-friendly interface enhances accessibility, aiding both patients and healthcare providers in understanding and navigating the results effectively. The integration of WHO National Guidelines into the logistic regression model represents a significant step forward in personalized cardiovascular risk assessment. By combining accuracy, standardization, and practicality, this approach holds promise for improving the early identification and management of individuals at risk of heart attacks, ultimately contributing to better heart health outcomes on a global scale.

Users of the mobile application can upload images of their meals to accurately identify the food items with an accuracy of around 85%. This has been tested only for a particular set of common foods used by Sri Lankans such as rice, potato, beetroot, fish, eggs, etc. By utilizing advanced algorithms, the system offers comprehensive details about every item, such as its calories content and suitability for specific patient conditions. The tailored approach makes it easier for users to make educated dietary decisions that meet their nutritional needs and overall health goals.

The KNN model, which was used in predicting the risk level of having a future heart attack, has a training accuracy of 96%. However, the testing accuracy drops to 88%. This could be because the ECG reports are unique from patient to patient. Note that the model effectively analyzes the ECG images to derive the ECG category and the risk level from that category. The normal person category was grouped as low risk (represented by 0), the abnormal heartbeat category as medium risk (represented by 1), and the myocardial infraction and history of myocardial categories were selected as high risk (represented by 2). Accordingly, a graph generated by the mobile application is shown below (see Fig. 7) with Risk Level in y-axis (vertical) and Dates in x- axis.

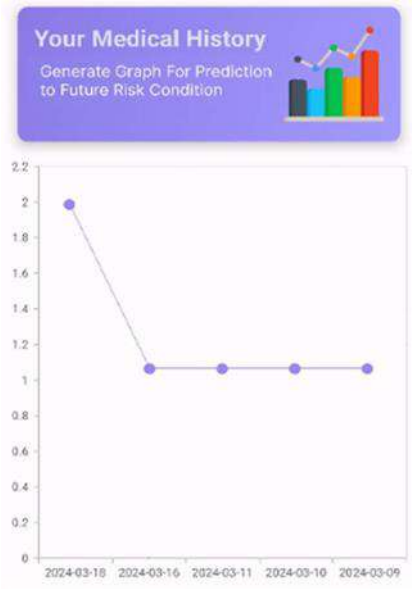


Fig. 7. Sample risk category levels of having a heart attack.

Identifying the nearest pharmacies with the best customer recommendations was implemented successfully. A sample suggesting the best rated pharmacies—within a given radius via google map—on user request is displayed in Fig. 8.



Fig. 8. Sample list of best rated nearby pharmacies.

V. CONCLUSION

The integrative approach presented in this paper offers a significant step forward in the early prediction and management of cardiovascular health. By integrating the four functional components: predicting the risk of becoming a heart patient, categorizing the risk level for heart patients, providing personalized dietary recommendations, and locating top-rated nearby pharmacies, the mobile application developed enhances both preventive care and post-heart attack recovery. The user-friendly interface allows individuals to assess their cardiovascular health frequently and take proactive steps towards reducing their risk of having a sudden heart failure. Developing a commercial-grade version of this application holds great potential for serving a large community to provide effective cardiovascular health management and to enhance the quality of life of those individuals who are either at risk of having heart attacks or who are recovering from such situations.

REFERENCES

[1] W. J. Kipp et al., "Artificial Intelligence in Cardiology," Journal of the American College of Cardiology, vol. 71, no. 23, pp. 2668-2679, 2018.

[2] L. P. Badano, D. M. Keller, D. Muraru, C. Torlasco, and G. Parati, "Artificial intelligence and cardiovascular imaging: A win-win combination," (in eng), Anatol J Cardiol, vol. 24, no. 4, pp. 214-223, Oct 2020.

[3] M. Daniel Sierra-Lara et al., "Artificial intelligence opportunities in cardio-oncology: Overview with spotlight on electrocardiography," American Heart Journal Plus: Cardiology Research and Practice, vol. 15, p. 100129, 2022.

[4] M. Ashraf et al., "Prediction of Cardiovascular Disease Through Cutting-Edge Deep Learning Technologies: An Empirical Study Based on TENSORFLOW, PYTORCH and KERAS," Singapore, 2021, pp. 239-255: Springer Singapore.

[5] Y. M. Ayano, F. Schwenker, B. Dufera, and T. Debelee, "Interpretable Machine Learning Techniques in ECG-Based Heart Disease Classification: A Systematic Review," Diagnostics, vol. 13, p. 111, 2022.

[6] G. Battineni, G. G. Sagaro, N. Chinatalapudi, and F. Amenta, "Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis," (in eng), J Pers Med, vol. 10, no. 2, Mar 31 2020.

[7] K. Shameer et al., "Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: A case-study using mount sinai heart failure cohort," Pacific Symposium on Biocomputing, vol. 0, no. 212679, pp. 276-287, 2017.

[8] M. A. Sarwar, N. Kamal, W. Hamid, and M. Shah, "Prediction of Diabetes Using Machine Learning Algorithms in Healthcare," pp. 1-6, 2018.

[9] M. Ganesan and N. Sivakumar, "IoT based heart disease prediction and diagnosis model for healthcare using machine learning models," in 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), 2019, pp. 1-5.

[10] N. B and I. Velchamy, "Predictive analytics in health care using machine learning tools and techniques," pp. 492-499, 2017.