Enhancing Patient-Centric Care and Doctor Convenience in Hospitals

Component – "**Automated Analysis of Medical Records and Predictive Modeling for Future Problems in Patients Using Machine Learning Algorithms**."

Project ID: TMP-2023-24-015

Project Proposal Report

Dewantha A.A.A.R.S

IT20618186

B.Sc. (Hons) Degree in Information Technology Specialized in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

August 2023

Enhancing Patient-Centric Care and Doctor Convenience in Hospitals

Component – "**Automated Analysis of Medical Records and Predictive Modeling for Future Problems in Patients Using Machine Learning Algorithms**."

Project ID: TMP-2023-24-015

Project Proposal Report

Dewantha A.A.A.R.S

IT20618186

B.Sc. (Hons) Degree in Information Technology Specialized in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

August 2023

Declaration page of the candidates & supervisor

We declare that this is our own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Group Member Name	Student ID	Signature
Dewantha A.A.A.R.S	**IT20618186**	

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Signature of the Supervisor: Date: 2023.08.25

(Dr. Kapila Dissanayaka)

Abstract

The number of people dying of heart disease in Sri Lanka is increasing day by day. Some heart conditions are manageable by the patient. But due to the inability to be aware of the risk situation in advance, many people are not interested in controlling certain factors. A lot of research has been done to diagnose heart conditions. Among them, much previous research has been conducted based on machine learning techniques to analyze the electronic health records of patients, including the most appropriate and accurate machine learning algorithms. My aim is to give prediction about the life risk that he or she will have to face in the future by identifying the current condition of the heart patient by using medical reports such as Electrocardiogram (ECG), Blood report, Magnetic resonance imaging (MRI) reports and comparing it with the previous condition. Here I hope to predict the level of risk in the form of a graph chart so that patients can understand. The main reason that our team suggested for creating a mobile application was that nowadays most people are tempted to use a smartphone. Mobile apps are easier to use and interact with people more quickly. Also, through this application, patients are anxious to identify their risk condition earlier and get proper treatment for it. Also, from the doctor's point of view, he can be aware of the risk conditions of the patient who comes to him and then check his sensitivity and give him treatment and advice.

Keyword: Electrocardiogram, Machine Learning, Electronic Health Record, Heart Disease, Predictive Modeling

Table of Contents

List of Figures

List of Table

01. Introduction

1.1 Background and Literature Survey

Heart disease is increasingly becoming prevalent in the world, and Sri Lanka marks no exception. According to data released by the World Health Organization (WHO) in 2020, over 22.6% of deaths in Sri Lanka are caused by cardiovascular diseases [1]. High blood pressure, high blood cholesterol and smoking have been identified as major factors that increase the risk of death from heart disease. Additionally, there are additional risk factors that can determine a person's likelihood of developing cardiovascular disease. Although variables no one can control, a person's age and ancestry play a significant role in assessing their susceptibility to heart disease [2].

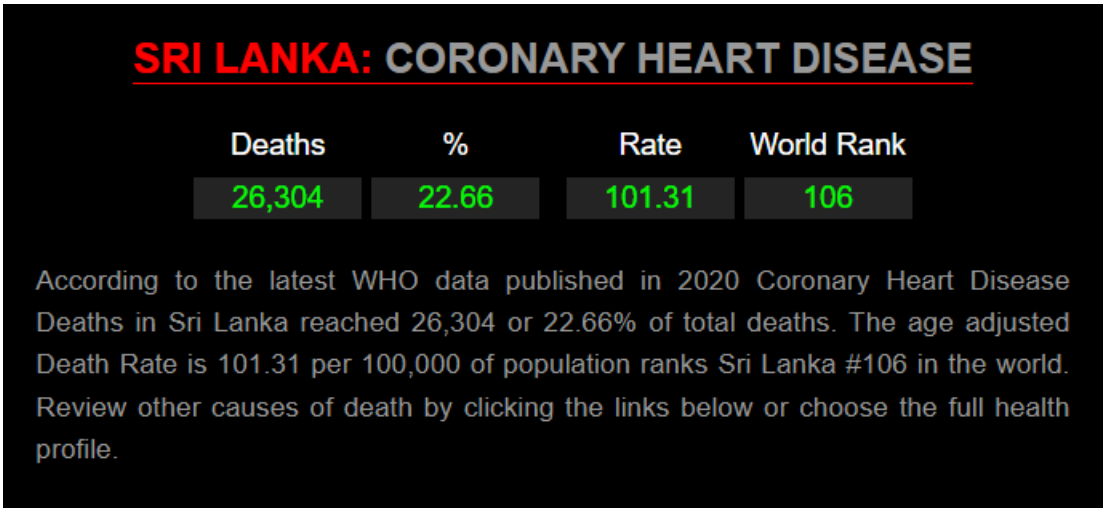

Figure 1 - coronary heart disease

Here, we consider helping people with preventable heart disease reduce their lifetime risk by providing advance notice of their risk status. Using the mobile application our team created for this project, people with heart illnesses can readily predict their doctor's recommendations for therapy. Through the AI model we have developed, the doctor can also predict about the patient's medical history who visits him in advance, as well as the likelihood of a heart attack and the risk that may materialize for a patient who is already experiencing a heart attack in the near future. Doctors no longer need to spend much time or effort comparing old and new records. The task I undertook in this section was to evaluate a patient's medical history, including an ECG or an EHR (Electronic Health Record), and determine the patient's future risk status.

Recent developments in automated analysis and predictive modeling for potential patient issues are made possible using machine learning algorithms in healthcare research. The application of machine learning techniques in the medical field has shown promise for transforming patient care, diagnosis, and prognosis. This research effort, which has a particular focus on cardiac patients, intends to build on previous work to develop a more thorough and precise predictive model by utilizing both medical records and electrocardiogram (ECG) data.

These initiatives represent the first steps towards creating predictive models. Previous research in the fields of automated analysis of medical records and predictive modeling for future problems in patients has yielded notable insights and achievements. Early studies primarily centered around applying traditional machine learning algorithms to electronic health records (EHRs) for identifying patterns and predicting medical conditions. Techniques such as decision trees, support vector machines, and logistic regression were used to extract relevant information from patient data.

Compared to other initiatives, medical and AI projects attracted more speculative funding in 2016 [3]. Artificial intelligence (AI) in medicine refers to the use of automated diagnosis procedures and the treatment of patients who need care. In 2017, “Predictive Modeling of Hospital Readmission Rates Using Electronic Medical Records-Wide Machine Learning: A Case-Study Using Mount Sinai Heart Failure Cohort” According too the research paper, they discuss their attempt to develop a data-driven, electronic medical record-wide (EMR-wide) feature selection approach and subsequent machine learning to predict readmission probabilities.

The environment of predictive modeling has changed recently because of the spike in interest in deep learning. Convolutional neural networks (CNNs) have proven to be incredibly accurate at analyzing image-based medical data, including interpreting X-rays and MRIs. For processing sequential data, such as ECG signals, recurrent neural networks (RNNs) and attention processes have proven invaluable in enabling the detection of anomalies and abnormal cardiac rhythms [4].

However, this study tries to fill in some gaps and take advantage of some opportunities. The combination of EHR information and ECG signals is an interesting feature. Prior studies have mostly focused on either specific data categories or medical records; however, merging both kinds of information can offer a more complete picture of a patient's health. This interdisciplinary method enables the discovery of nuanced relationships between EHR data, ECG abnormalities, and potential heart issues that may have previously gone unnoticed [5] [6].

Additionally, there is very little research on how to incorporate interpretability strategies into predictive models. Gaining the confidence of medical experts and increasing the use of machine learning solutions in clinical settings both depend on the capacity to explain model predictions. This research can improve model transparency and interpretability by including explainable AI approaches, such as attention maps that highlight ECG segments contributing to predictions.

1.2 Research Gap

My research topic is "**Automated Analysis of Medical Records and Predictive Modeling for Future Problems in Patients Using Machine Learning Algorithms**." As I hope in this research, what I will do is analyze the data and factors obtained from a heart attack patient's ECG reports and blood reports using a machine learning algorithm and estimate the patient's life risk. Through the mobile application, both the patient and the doctor are provided with a graphical chart to monitor their progress according to the treatment he or she has received, the patient's current condition, and the risks condition that may occur in the future.

Research A – "Prediction of Diabetes Using Machine Learning Algorithms in Healthcare" According to this research paper [7], For experiment purpose, a dataset of patient's medical record is obtained, and six different machine learning algorithms are applied on the dataset. Performance and accuracy of the applied algorithms is discussed and compared. Comparison of the different machine learning techniques used in this study reveals which algorithm is best suited for prediction of diabetes. The performance and accuracy of the machine learning algorithms used here are discussed based on the data sets.

Research B – "IoT based heart disease prediction and diagnosis model for healthcare using machine learning models" According to this research paper [8], To avail good service to the user using the online healthcare services, a fresh Cloud as well as IoT based Healthcare application to monitor in addition to diagnose serious diseases is developed. In this study, an efficient framework is utilized for heart disease is created utilizing the UCI Repository dataset as well as the healthcare sensors to predict the public who suffer from heart disease. Here, using both techniques called IoT sensors and amazing cloud computing, a person's blood pressure and ECG are taken into consideration to highlight whether he is a heart patient or not. This is a somewhat successful experiment [9].

Research C – "Predicting Future Cardiovascular Events in Patients with Peripheral Artery Disease Using Electronic Health Record Data" According to this research paper [10], Patients with peripheral artery disease (PAD) are at risk of major adverse cardiac and cerebrovascular events. There are no readily available risk scores that can accurately identify which patients are most likely to sustain an event, making it difficult to identify those who might benefit from more aggressive intervention. Thus, they aimed to develop a novel predictive model—using machine learning methods on electronic health record data—to identify which PAD patients are most likely to develop major adverse cardiac and cerebrovascular events.

According to the research A, B and C considered above, I was able to find new profiles for the component to be developed. I hope to include features such as using a mobile application, generating information and graphical graphs depicting the risk level of heart patients, automatically analyzing disease and medical records, and especially predicting the current condition of a heart patient as well as possible future risk conditions in this research.

Features	A	B	C	Proposed System
Mobile App	✗	✓	✗	✓
Generates informative diagrams or charts	✗	✗	✗	✓
Automated Analysis of Medical Records	✓	✓	✓	✓
predict future problems and condition in heart patients	✗	✗	**Predict future cardiovascular events**	✓
Machine Learning Algorithms	K-Nearest Neighbors (KNN), Naive Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), Logistic Regression (LR) and Random Forest (RF)	logistic regression (LR), multilayer perception (MLP), support vector machine (SVM)	Decision Tree (DT), Logistic Regression (LR) and Random Forest (RF)	Convolutional Neural Networks (CNN), Support Vector Machines (SVM), Random Forest

Table 1 - Research Gap

1.3 Research Problem

Today, the number of deaths due to heart diseases has increased in Sri Lanka. The main reason for this is that patients suffering from heart diseases do not focus on controlling the disease. It is very important for people suffering from heart diseases to know the level of their disease tendency at present by following the treatment they receive, and the instructions prescribed by the doctors. It is important to be aware of it in advance in order not to allow your condition to become dangerous in the coming period. For this, what usually happens in Sri Lanka is that most people meet with heart disease specialists and show their medical records to gain awareness. Although this is a very successful method, if one wants to know the level of tendency in one's diseased nature now, it cannot be done. For that, those patients must bear their time, effort, and financial costs.

According to my research topic, the medical records of a person with a heart disease will be automatically analyzed through an image processing machine learning algorithm to thereby identify the current condition of that patient as well as the risk situations that may occur in the future. Although many tests have been conducted on related to early detection of heart failure, very little research has been done to detect this kind of life risky condition early for heart patient. And since this is used as a mobile application, the patient and the doctor can share information with each other, and the doctor can be aware of the current condition of his patient and possible future risks. If the doctor examines the dangerous condition and gives treatment for it first, the life risk of the patient can be reduced. I am very anxious to avoid some of the errors in them according to the research that has been done in the past.

02. Objective

2.1 Main Objective

- Develop an automated system that utilizes machine learning algorithms to analyze ECG medical records, predict future heart disease-related issues, and generate a graph chart for visual representation of the patient's current and predicted future status.

2.2 Specific Objectives

- To study and understand what the conditions and factors heart patients are.
- To study and understand analyzing ECG and Other electronic medical records.
- Collecting a diverse and representative dataset of ECG records.
- Preprocessing the ECG data to remove noise and artifacts using machine learning algorithms.
- Develop a user-friendly interface for generating graph chart and to the doctor and the patient, to take care of patient’s risk.
- Create a No-SQL database for store patient’s medical reports and ECG Record Images

03. Methodology

- Data Preparation: Collecting and preprocessing a diverse dataset of ECG records.
- Feature Extraction: Extracting relevant features from the ECG signals.
- Model Training: Training machine learning models to analyze ECG records and predict future health problems.
- Graph Chart Generation: Designing a graph chart that visually represents the patient's current and predicted future status based on the model's output.
- Evaluation: Assessing the performance and accuracy of the automated system using appropriate metrics.
- Validation: Validating the system on independent datasets to ensure generalizability.

3.1 Requirement Gathering

- As the initial step of requirement analysis, I talked with Dr. Kapila Dissanayaka (Our Supervisor) about our research topic and individual components. After we decided to meet up with the co-supervisor and external supervisor about our research ideas. First got a meeting with our co-supervisor and solved some issues. We contacted a doctor in apeksha hospital maharagama to get the data and medical records we need.

3.2 Feasibility Study

3.2.1 Technical Feasibility Study

- project members have the knowledge of developing mobile applications and web applications. Therefore, this project is technically feasible to develop. Because our team is learning about machine learning algorithms using the internet, and because we have the necessary facilities to study related to it in our college, and because we have the ability to get relevant medical information from hospitals, we are very anxious to do this research.

3.2.2 Schedule Feasibility Study

- Each phase of the proposed system should be completed with credible outputs while keeping in line with the timeline and finalized. The final product should then be submitted on the predefined due date.

3.3 System Analysis

3.3.1 Overall System Overview

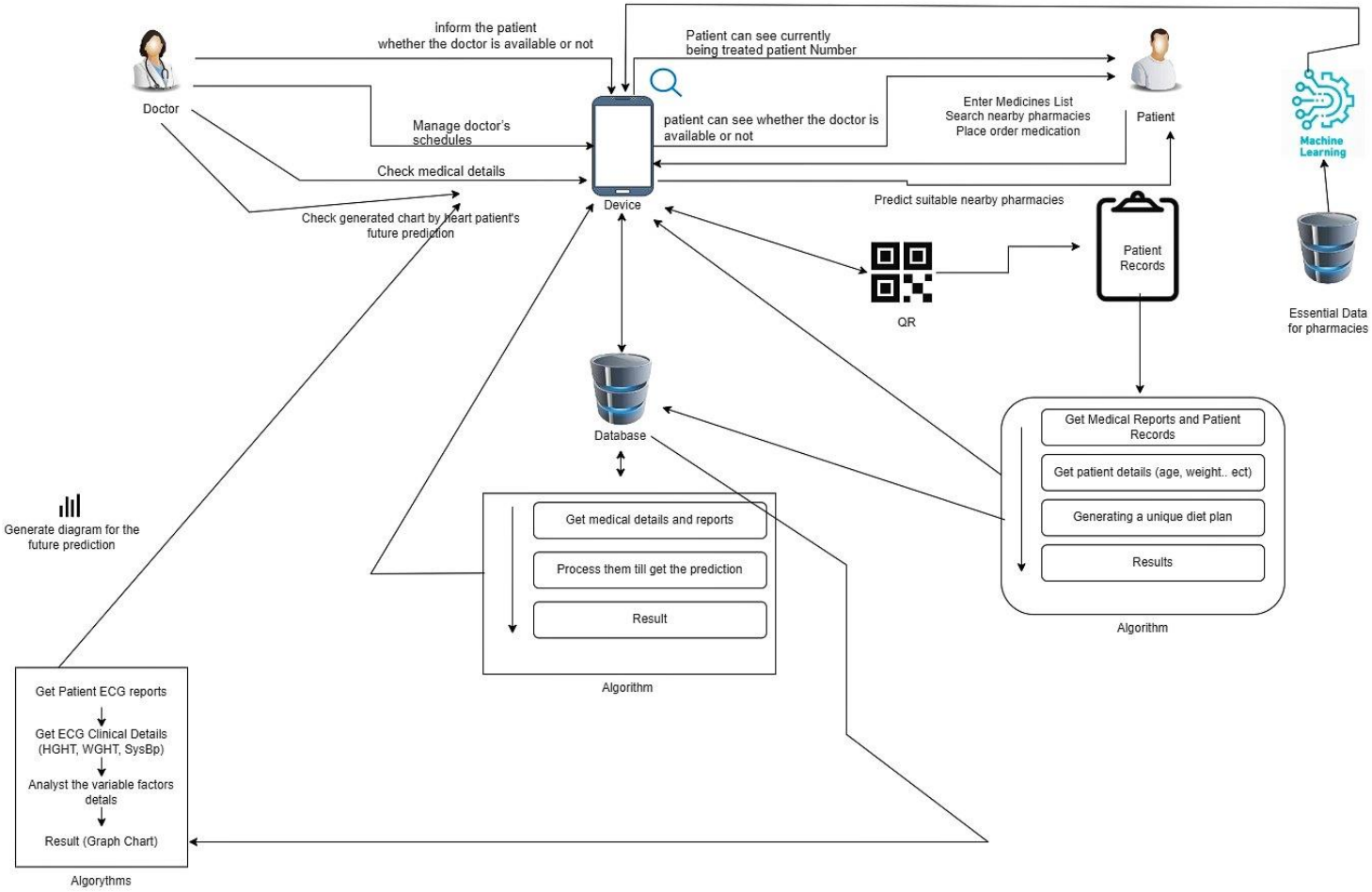

Figure 2 - Overall System Overview Diagram

3.3.2 System Overview for Individually Component Diagram

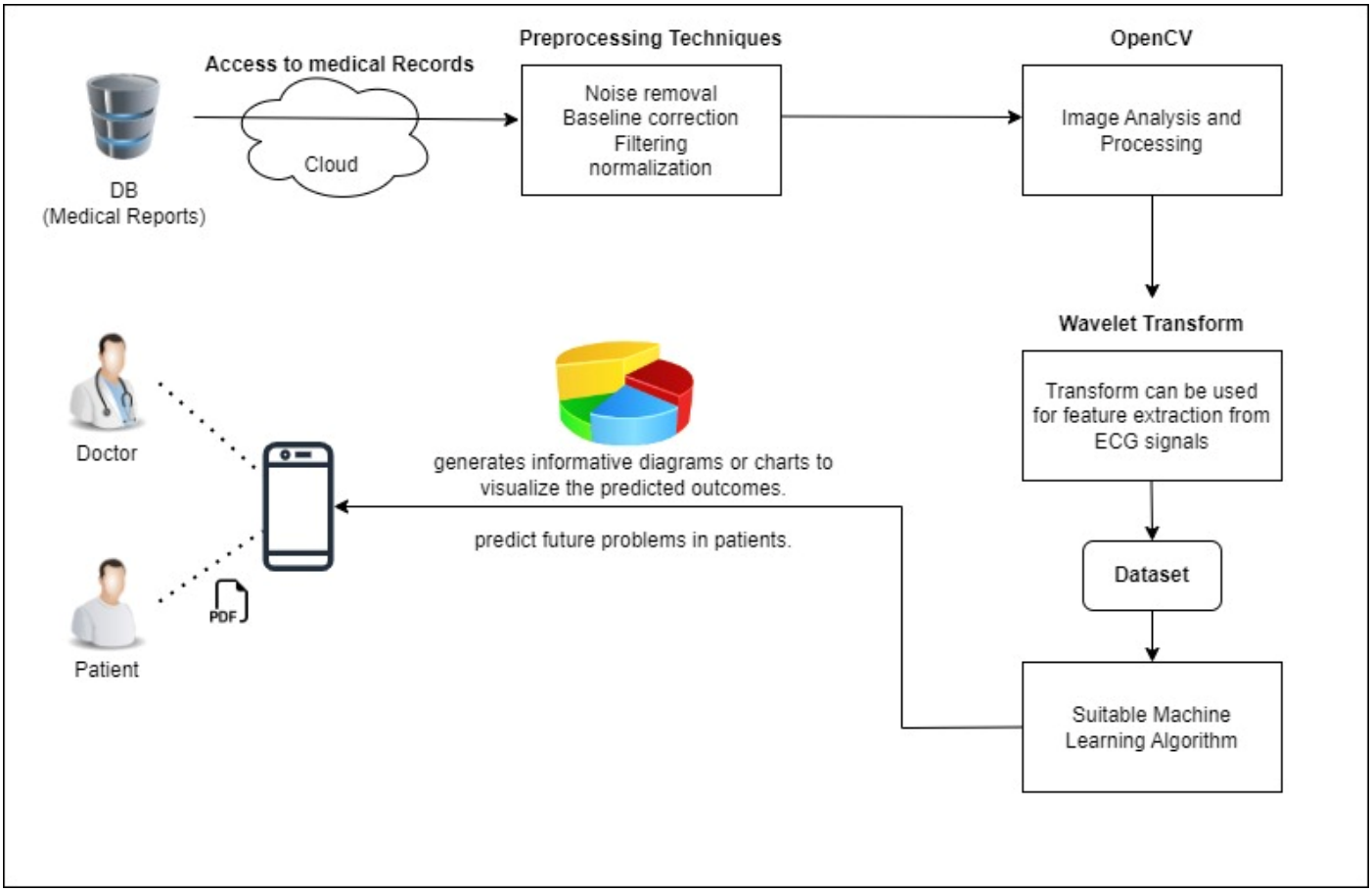

Figure 3 - System Overview for Individual Diagram

04. Project Requirements

4.1 User Requirements

- A doctor should be able to view own patient's heart failure risk prediction.
- User school login to the mobile Application and he want to navigated future risk prediction component.
- User should be able to view the prediction of future heart failure conditions.

4.2 Software Requirements

4.2.1 Technologies

- Image Processing Algorithms
- Artificial Intelligent
- Python Language
- React Native

4.2.2 Tools

- OpenCV
- Jupyter Notebook

4.3 Functional Requirements

- Data Collection and Preprocessing
- Automated ECG and EHR (Electronic Health Record) Analysis
- Predictive Modeling
- Graph Chart Generation
- User Interface

4.4 Non-functional Requirements

- Accuracy and Reliability
- Performance
- Scalability
- Security and Privacy
- Interoperability

4.5 Test Cases

This phase verifies that each function of this application has no errors. According to do that unit testing, integration testing, system testing, and user acceptance testing will be done. Furthermore, in this phase will focus on increasing the quality of the software.

1. Unit Testing
    - In this step, each function will check individually and verify that each function has no errors. In this phase, more focus on analyzing the source code. This type of testing is called white box testing.

2. Integration Testing
    - In this phase, integrate the collection of sub-modules and verify the interaction and how it works. After integrating individual components, there may be some failures that would happen. In this phase, fine-tune those failures.

3. System Testing
    - The purpose of this phase is to test the fully integrated application. This testing is done after the entire system is fully assembled.

4. User Acceptance Testing
    - In this phase comparing this developed application with initial requirements. This task will be performed by heart patients and doctors.

05. Evaluation Criteria

5.1 Gantt Chart

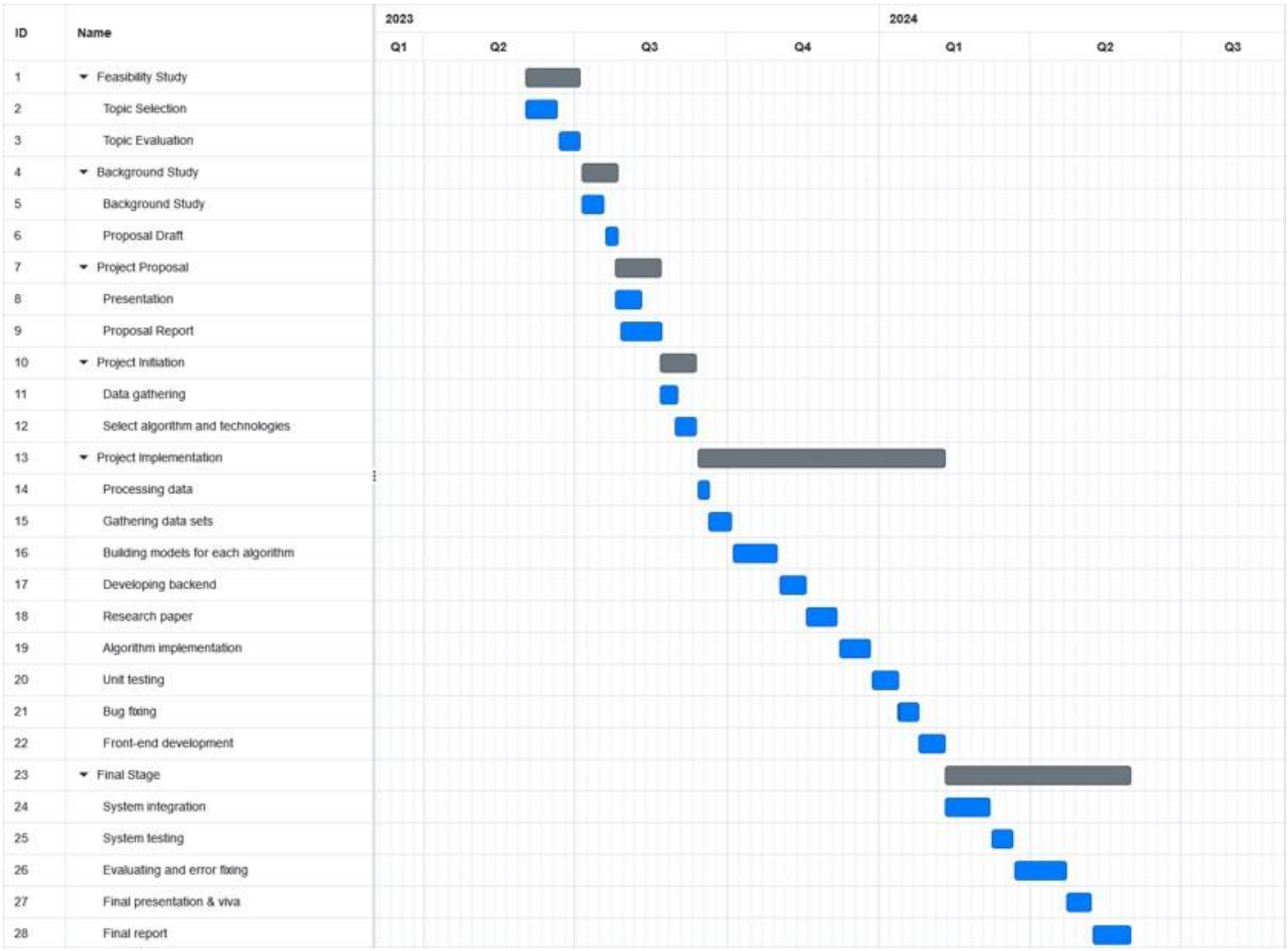

Figure 4 - Gantt Chart

5.2 Work Breakdown Structure

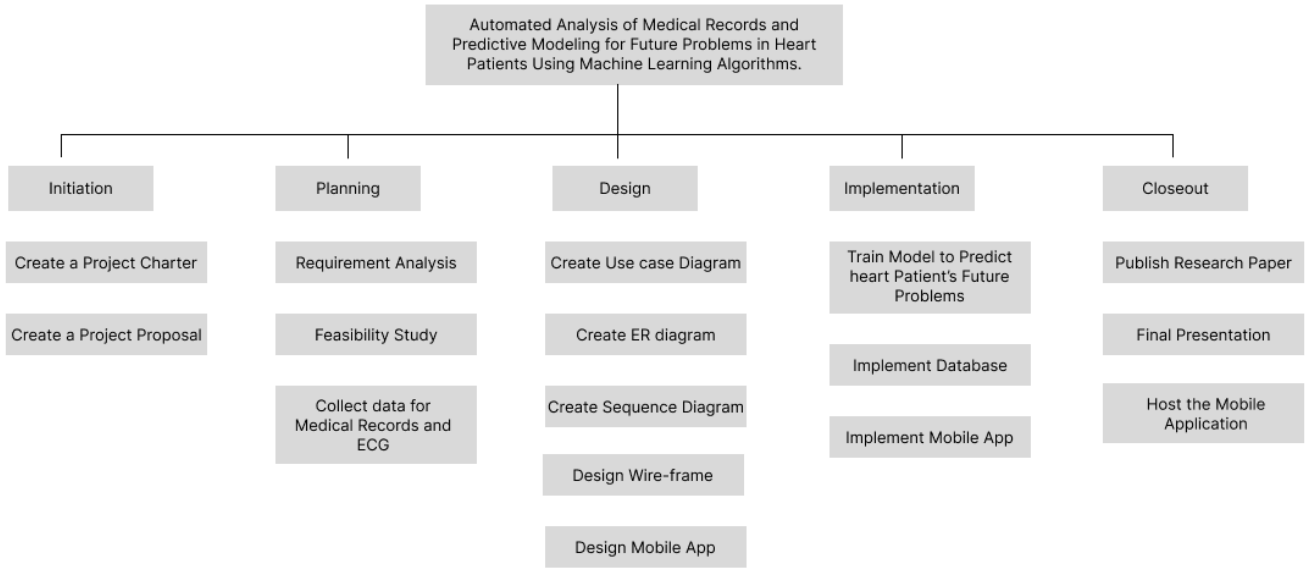

Figure 5 - Work breakdown Chart

06. Budget

Component	Cost (LKR)
Traveling Cost	Rs. 12,000.00
Deployment Cost	Rs. 18,000.00
Mobile App Hosting and Play Store	Rs. 20,000.00
Others	Rs.10,000.00
Total Cost	Rs. 60,000.00

Table 2 – Budget

07. References

[1] W. H. ORGANIZATION, "WORLD HEALTH RANKINGS," WORLD HEALTH ORGANIZATION, 2020. [Online]. Available: https://www.worldlifeexpectancy.com/sri-lanka-coronary-heart-disease.

[2] A. Insurance, "AIA.com," AIA Insurance, 28 October 2021. [Online]. Available: https://www.aialife.com.lk/en/life-challenges/what-matters/health-wellness/The-cardio-epidemic-growing-case-of-heart-disease-in-sri-lanka.html#:~:text=The%20Cardio%20Epidemic%3A%20Growing%20Case%20of%20Heart%20Disease%20in%20Sri%20Lanka,-28%20October%2020.

[3] I. A. M. M. Varun H Buch, "Artificial intelligence in medicine: current trends and future possibilities," *Br. J. Gen. Pract,* Vols. 143-144, p. 68, 2018.

[4] D. V. I. B. Nithya, "Predictive Analytics in Health Care Using Machine Learning Tools and Techniques," *International Conference on Intelligent Computing and Control Systems,* p. 8, 2017.

[5] G. Battineni, G. G. Sagaro, N. Chinatalapudi and F. Amenta, "Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis," *Journal of Personalized Medicine,* p. 11, 2020.

[6] K. SHAMEER, K. W. JOHNSON, A. YAHI, R. MIOTTO, D. RICKS, J. JEBAKARAN, P. KOVATCH, P. P. SENGUPTA, A. GELIJNS, A. MOSKOVITZ, B. DARROW, D. L. REICH, A. KASARSKIS, N. P. TATONETTI and S. PINNEY, "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,* p. 12, 2017.

[7] N. K. W. H. M. A. S. Muhammad Azeem Sarwar, "Prediction of Diabetes Using Machine Learning Algorithms in Healthcare," in *International Conference on Automation & Computing*, Newcastle University, 2018.

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

[9] B. Nithya and D. V. Ilango, "Predictive Analytics in Health Care Using Machine Learning Tools and Techniques," *International Conference on Intelligent Computing and Control Systems,* p. 8, 2017.

[10] E. G. Ross, K. Jung, J. T. Dudley, L. Li, N. J. Leeper and N. H. Shah, "Predicting Future Cardiovascular Events in Patients With Peripheral Artery Disease Using Electronic Health Record Data," *Circulation: Cardiovascular Quality and Outcomes,* p. 9, 2019.