Style Guidelines for Final Year Project ReportsTitle of the Project

Final Year Project Proposal

Session 2021-2025

A 4th Year Student

A project submitted in partial fulfilment of the

COMSATS University Degree

of

BSc. (Hons.)BS in Computer Science (CUI)



Department of Computer Science

COMSATS University Islamabad, Lahore Campus

09 March 2024

**Project Registration**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Project ID (for office use) | | |  | | | | |
| Type of project | | | [✓] Traditional [ ] Industrial [ ] Continuing | | | | |
| Nature of project | | | [ ] **D**evelopment [ ] **R**esearch [✓] **R**&**D** | | | | |
| Area of specialisation | | | Machine Learning | | | | |
| **Project Group Members** | | | | | | | |
| Sr.# | Reg. # | Student Name | | CGPA | Email ID | Phone # | Signature |
| (i) | SP21-BCS-007 | Asad Ali | | 3.02 | sp21-bcs-007 @cuilahore.edu.pk | 0307 4315952 |  |
| (ii) | SP21-BCS-003 | Asad ur Rehman | | 3.02 | sp21-bcs-003 @cuilahore.edu.pk | 0335 4550064 |  |
| (iii) | SP21-BCS-017 | Muhammad Haroon Shahzad | | 2.99 | sp21-bcs-017 @cuilahore.edu.pk | 0318 4559527 |  |
| **Declaration:** FYP group members have cleared all prerequisites courses For FYP-I as per their degree requirements.  For BS (Computer Science)  (CSC241 Object Oriented Programming, CSC291 Software Engineering Concepts, CSC371 Database Systems-I, HUM102 Report Writing Skills)  For BS (Software Engineering)  (CSC241 Object Oriented Programming, CSE291 Introduction to Software Engineering, CSC371 Database Systems-I, HUM102 Report Writing Skills) | | | | | | | |

# Plagiarism Free Certificate

This is to certify that, I am Asad Ali S/D/o Ghaffar Ahmad, group leader of FYP under registration no CIIT/SP21-BCS-007/LHR at Computer Science Department, COMSATS Institute of Information Technology, Lahore. I declare that my FYP proposal is checked by my supervisor and the similarity index is \_\_\_\_\_\_\_\_% that is less than 20%, an acceptable limit by HEC. Report is attached herewith as Appendix A.

A black background with letters

Description automatically generatedDate: 11-03-2024 Name of Group Leader: Asad Ali Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Name of Supervisor: Muhammad Aksam Iftikhar Co-Supervisor (if any): \_\_\_\_\_\_\_\_\_\_\_\_\_

Designation: Associate Professor Designation: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Abstract**

Heart diseases is one of the most pressing issues globally. Effective and correct heart disease prediction methods are imperative for timely intervention. This project presents a comprehensive strategy and techniques for predicting heart disease through detailed analysis of electrocardiogram signals and image repots, employing advanced machine learning techniques. By combining demographic information with ECG data and exploring ECG signals, the project aims to enhance prediction accuracy and provide additional avenues for accurate predictions. The project primarily focuses on integrating demographic information such as age, gender, and medical history with heart rhythm recordings through a data pipeline and feeding it to machine learning model for training and making it able to predict for new patient. Additionally, the project deals with image-based ECG reports, establishing a separate pipeline to convert ECG images reports into signal data. After training, these machine learning models are deployed on a cloud-based service and can be integrated into a user-friendly interface that includes a mobile app and website. The web interface facilitates doctors to easily upload ECG data in bulk and process it for predictions and get medical prescriptions and diet plans based on predicted disease, also provides essential features like keeps track of patient information and history of the patient. Also, the addition of an assisting medical chatbot within the web interface help users in interpreting results and offers valuable medical guidance. Furthermore, a mobile app allows users to upload images and get predictions. In essence, this project represents a significant advancement in heart disease prediction methodologies, offering a helpful and user-friendly solution.

# Introduction

Heart disease is a common and serious health problem that affects many people worldwide. Predicting it accurately is important so that people can seek help in a timely manner. Electrocardiogram signals provide valuable insights about heart health and can be used to predict heart diseases. This project, called the "ECG Analyzer Toolkit," uses machine learning techniques and models to achieve precise predictions.

The project focuses on creating a robust data pipeline capable of integrating demographic information and ECG signals in such a way that they can be fed to a machine learning model. The pipeline will source raw data containing both demographic information and ECG signal information, converting it into a cohesive dataset. Leveraging PhysioNet's Wave Form Database, the fetched data will be converted into CSV files, which are easier to work with for analysis and machine learning model training.

Once the data is organized and pre-processed, the machine learning model will be used to predict heart disease based on ECG signal data. The model will be thoroughly tested to ensure accurate predictions and efficient results. The project will also explore image-based ECG report processing, developing a separate pipeline to convert ECG images into analysable signals. These image-based signals will undergo the same predictive modelling process as conventional ECG signal data, enriching the predictive capabilities of the models by providing another option for disease predictions using ECG image-based reports.

After successful training and validation, the trained machine learning model will be deployed on a cloud service for accessibility and scalability. It will also be integrated into a website for use. This way, doctors and regular individuals can upload their heart rhythm data to obtain predictions. The website will even have a chatbot to assist and explain the results clearly. Additionally, the project includes a mobile app for those who can simply upload their ECG image to obtain predictions quickly.

# Success Criterion

* Achieve a minimum accuracy threshold (e.g., 85% or higher) in predicting heart disease when using both ECG signals and demographic data.
* Compare this with models using ECG signals alone to determine if incorporating demographic information improves performance.
* The model on the cloud can handle lot of ECG tests being uploaded at once without having any issues.
* The user finds the interface intuitive, effective and should find it easy to use and understand the results, prescriptions, and diet plans that it gives to them.
* The chatbot explains the result correctly and give advice based on medical guidelines.
* The system may be able to find heart disease earlier than normal tests doctors use now.
* Compare when the system predicts heart disease versus when a doctor normally finds it in the same group of patients. Track if early prediction leads to better outcomes.

# Related Work (mandatory)

# Project Rationale

Heart disease poses a critical global health threat, necessitating advanced methods for early detection and intervention. Traditional diagnostics often lead to delays in diagnosis, limiting treatment options and impacting patient outcomes. This project aims to develop a comprehensive heart disease prediction system leveraging electrocardiogram (ECG) signals and demographic data. This combination has the potential to improve prediction accuracy compared to existing methods. The development of image-based ECG processing pipelines will broaden input sources and enhance accessibility for diverse users. The motivation behind this process is to develop a more accurate and user accessible solution.

Through conducting this research and developing this solution, our main learning will be in depth understanding of how to train test and deploy machine learning models along with the integration process for web and mobile app. On of our mains learnings will be understanding data preprocessing, data pipeline designing and working with the imaging data. Also, we will learn how to deploy and maintain a machine learning model to keep it working for the user on our website and mobile app.

This project is driven by the need for accessible and user-friendly prediction solutions. A cloud-based machine learning model that integrate with web interface and a dedicated mobile application will prioritize widespread access. The integration of a medical chatbot will empower users by offering real-time interpretation of results and guidance on next steps. Ultimately, this project aims to introduce early detection tools for heart disease management. The potential impact includes improving patient outcomes and help them to seek better treatment early.

## Aims and Objectives

The primary aim of this project to develop an advanced system for heart disease prediction that integrates ECG signals analysis and image ECG image-based report processing and providing user-friendly interface to get accurate predictions about heart diseases.

Making it helpful for the doctors to get predicted disease labels along with ECG report for several patients in very sophisticated and rebuts way. It also aims to help a patient having any ECG report to get know about his disease, get prescriptions, and diet plans accordingly also integrating a chat bot ask questions about the disease or answer any concern of the patient.

### Objectives

* Desing and construct the data pipeline that combine ECG signals and demographic data from the PhysioNet WFDB database.
* Make the pipeline that converting ECG images into the data (signals) that model can understand.
* Train test and optimize the machine learning model to accurately and precisely predict heart disease based on ECG signals and image-based reports.
* Training a model to give prescriptions based on the predicted disease and have diet plan recommendations.
* Deploy the optimized prediction model on a cloud platform to ensure the scalability and accessibility.
* Build the website and app where the user can upload their heart data, safely log in manage their record, and see medical history.
* Include the medical chatbot (only for website). that can explain result answer the question and give them basic guidelines.

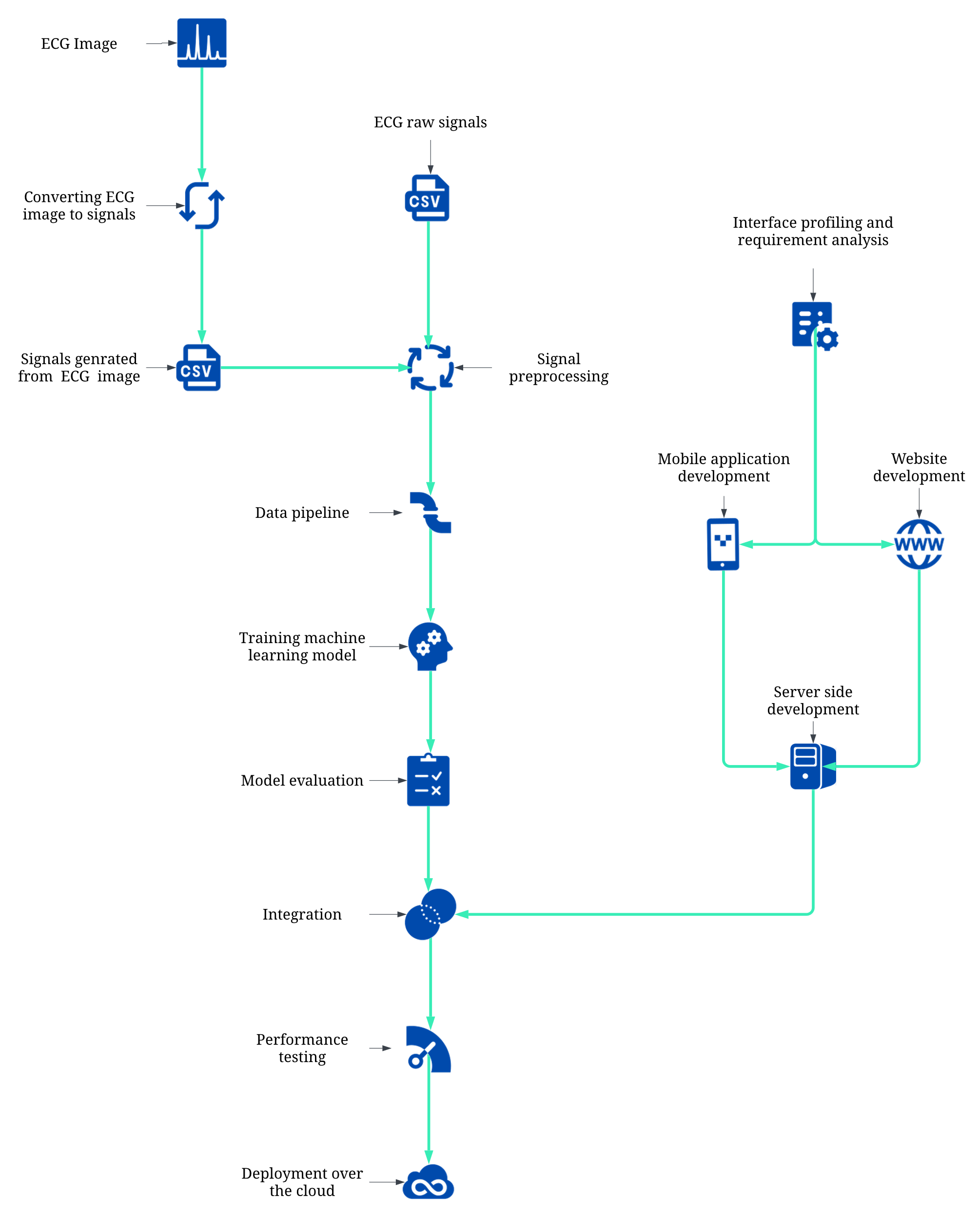
## Scope of the Project

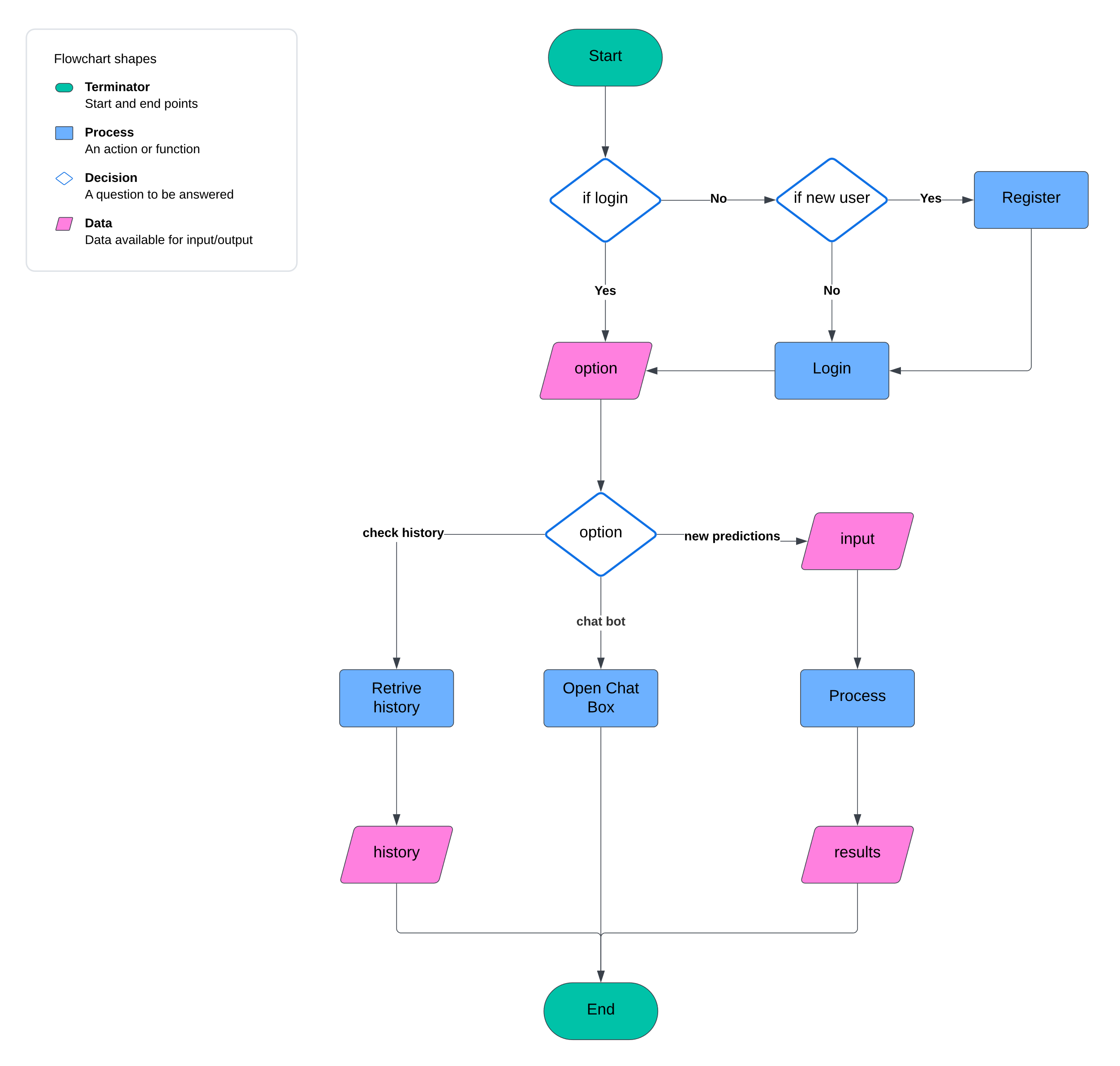
The scop of this project includes:

* Designing data pipeline and developing a preprocessing module to acquire data and convert it into a suitable format for applying machine learning techniques.
* Develop a machine learning based heart disease prediction system to accurately predict disease from ECG signal data or ECG image-based report.
* A fully functional cloud powered application solution accessible for the doctors and patients to process ECGs and get predictions, prescriptions, and assistance from the integrated medical chatbot fine-tuned for cardio related diseases.

# Proposed Methodology and Architecture

This section provides insight into what methodology you will employ in the development of the envisioned system. It is the systematic, theoretical analysis of the methods applied to your study. It can comprise step-by-step procedures, flowcharts, block diagrams or algorithms of the proposed system.





# Individual Tasks

|  |  |  |
| --- | --- | --- |
| **Team Members** | **Activity** | **Tentative Date** |
| Asad Ali, Asad ur Rehman, Muhammad Haroon Shahzad | Research and Project Planning | 11 March 2024 – 07 April 2024 |
| Asad ur Rehman, Muhammad Haroon Shahzad | Literature Review | 15 March 2024 – 07 April 2024 |
| Asad Ali, Muhammad Haroon Shahzad | Dataset Collection | 20 March 2024 – 29 March 2024 |
| Muhammad Haroon Shahzad, Asad Ali | Data Preprocessing and Pipeline Designing | 01 April 2024 – 12 April 2024 |
| Asad Ali | Interface Profiling and Requirement Analysis | 11 March 2024 – 12 April 2024 |
| Asad Ali, Asad ur Rehman, Muhammad Haroon Shahzad | Model Development and Training | 15 April 2024 – 10 May 2024 |
| Muhammad Haroon Shahzad | Website Development | 15 April 2024 – 10 May 2024 |
| Asad Ali, Asad ur Rehman | Model Evaluation and Validation | 13 May 2024 – 24 May 2024 |
| Asad Ali, Asad ur Rehman, Muhammad Haroon Shahzad | Integration and Testing | 13 May 2024 – 24 May 2024 |
| Asad Ali | Mobile App Development | 24 June 2024 – 09 Aug 2024 |
| Muhammad Haroon Shahzad, Asad Ali | Model Optimization and Enhancement | 24 July 2024 – 06 Sep 2024 |
| Asad Ali, Asad ur Rehman, Muhammad Haroon Shahzad | Interface Refinements | 11 Aug 2024 – 06 Sep 2024 |
| Muhammad Haroon Shahzad, Asad Ali | Backend Development and Integration | 11 Aug 2024 - 09 Nov 2024 |
| Asad Ali, Muhammad Haroon Shahzad | Deployment Process | 11 Nov 2024 – 30 Dec 2024 |
| Asad Ali, Asad ur Rehman, Muhammad Haroon Shahzad | Documentation and Reporting | 11 March 2024 – 30 Dec 2024 |

# Gantt Chart (Mandatory)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Activity** | |  |  |  | | --- | --- | --- | | Semester 7 | Summer Break | Semester8 | |
| |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec | |
| Research and Project Planning | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Literature Review | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Dataset Collection | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Data Preprocessing and Pipeline Designing | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Interface Profiling and Requirement Analysis | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Model Development and Training | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Website Development | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Model Evaluation and Validation | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Integration and Testing | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Mobile App Development | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Model Optimization and Enhancement | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Interface Refinements | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Backend Development and Integration | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Deployment Process | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |
| Documentation and Reporting | |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  | |

# Tools and Technologies

We have carefully selected different tools and technologies to make sure that we successfully achieve our goals and develop and implement a complete solution to our chosen problem.

## Technologies

### Machine Learning Model

* **Python:** Primary language for implementing ML models and web frameworks.
* **TensorFlow:** Building and training neural networks and deep learning models.
* **Keras:** Used with TensorFlow for rapid prototyping of high level ANNs.
* **scikit-learn:** Library for handling, classifications, and clustering problems.
* **OpenCV:** Library providing tools for image processing and manipulation.
* **NumPy:** Providing supports to other required libraries.
* **Pandas:** EDA and data pipeline designing of feeding data to ML model.
* **Matplotlib:** Visualizing ML model’s performance and confusion matrix.



### Interface Design and Development

* **HTML 5:** Assisting in structuring our web pages.
* **CSS 3:** Style sheets to give a robust look to the user interfaces.
* **JavaScript:** Making applications interactive to enhance user experience.
* **Bootstrap 5:** Ensure responsive design and quick prototyping.
* **Flask:** Quickly develops a web server in python dealing with ML model.
* **Pillow:** Processing images on the Python server for ML model.
* **React Native:** Building mobile app for engaging more users.
* **React:** Making fully functional web application for the project.
* **Django:** Hosts and manages the complete server for web and mobile.

 A blue and black logo

Description automatically generated A blue and black logo

Description automatically generated

### Databases

* **MS SQL Server:** Storing our model training data after pipeline and preprocessing.
* **MongoDB:** Keeping user demographics and medical history for applications.

## Tools

### Integrated Development Environments

* **Google Collab:** Cloud based notebooks for Model development and collaborations.
* **Visual Studio Code:** Light weight and versatile IDE for app and web development.
* **Jupyter Notebook:** Interactive code writing environment for EDA and prototyping.
* **Anaconda:** Facilitates Python virtual environments and robust library management.

 A logo with orange and grey circles

Description automatically generatedA green circle with a white border

Description automatically generated

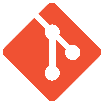
### Deployment

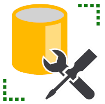
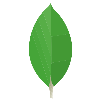
* **Microsoft Azure:** Facilitates ML model deployment and server less app hosting.
* **Docker:** Streamlines application management across different environments.

### Others

* **Git and GitHub:** Enables code management, sharing and team to collaborate.
* **SSMS:** SQL Server Management Studio to manage SQL database server.
* **MongoDB Compass:** Efficiently managing MongoDB from the desktop.

 A black cat with a tail

Description automatically generated  

# References (Mandatory)

You must provide references (IEEE style) when appropriate to justify your study.

Appendix A