**ARTIFICIAL INTELLIGENCE**

****

**Session 2022 – 2026**

**Submitted by:**

Tayyaba Afzal 2022-CS-134

Raveeha Mohsin 2022-CS-149

**Supervised by:**

Samyan Qayyum Wahla

Department of Computer Science

**University of Engineering and Technology**

**Lahore Pakistan**

**Table Of Contents**

[1 Overview 5](#_Toc180415661)

[2 Domain 5](#_Toc180415662)

[3 SDG Focus: Good Health and Well-Being (SDG 3) 5](#_Toc180415663)

[1. Enhanced Access to Healthcare: 6](#_Toc180415664)

[4 Technology Stack 6](#_Toc180415665)

[4.1 MySQL: 6](#_Toc180415666)

[4.2 Express.js: 6](#_Toc180415667)

[4.3 React.js: 6](#_Toc180415668)

[4.4 Node.js: 7](#_Toc180415669)

[4.5 AI Integration Tools: 7](#_Toc180415670)

[4.6 Data Visualization Libraries: 7](#_Toc180415671)

[4.7 Authentication: 7](#_Toc180415672)

[4.8 Deployment: 7](#_Toc180415673)

[5 Datasets 7](#_Toc180415674)

[5.1 Heart Disease Presence Detection 7](#_Toc180415675)

[5.1.1 Attributes 7](#_Toc180415676)

[5.1.2 Categories 8](#_Toc180415677)

[5.2 Myocardial Infarction Type Classification 8](#_Toc180415678)

[5.2.1 Attributes 8](#_Toc180415679)

[5.2.2 Categories 9](#_Toc180415680)

[5.3 Artery Narrowing Analysis 9](#_Toc180415681)

[5.3.1 Attributes 9](#_Toc180415682)

[5.3.2 Categories 10](#_Toc180415683)

[6 Algorithms 10](#_Toc180415684)

[6.1 Dataset Analysis 10](#_Toc180415685)

[6.2 Algorithm Selection 11](#_Toc180415686)

[7 Data Preprocessing 11](#_Toc180415687)

[7.1 Data Cleaning 11](#_Toc180415688)

[7.2 Data Transformation 12](#_Toc180415689)

[7.3 Splitting the Data 12](#_Toc180415690)

[7.4 Feature Selection 12](#_Toc180415691)

[8 User Roles and Functionalities 12](#_Toc180415692)

[8.1 Admin 12](#_Toc180415693)

[8.2 Patient 13](#_Toc180415694)

[8.2.1 Authentication Flow 13](#_Toc180415695)

[8.2.2 Heart Disease Detection 13](#_Toc180415696)

[8.2.3 Chatbot (AI Health Assistant) 14](#_Toc180415697)

[8.2.4 Profile and Health Records Management 14](#_Toc180415698)

[8.2.5 Report Generation and Download 14](#_Toc180415699)

[8.2.6 Notifications and Reminders 15](#_Toc180415700)

[9 Complete Flow and UI Interfaces 15](#_Toc180415701)

[9.1 Admin 15](#_Toc180415702)

[9.1.1 Authentication 15](#_Toc180415703)

[9.1.2 Dashboard 15](#_Toc180415704)

[9.1.3 User Management: 15](#_Toc180415705)

[9.1.4 View Diagnostics: 16](#_Toc180415706)

[9.1.5 Performance Analytics: 16](#_Toc180415707)

[9.1.6 System Settings: 16](#_Toc180415708)

[9.2 For Patients 16](#_Toc180415709)

[9.2.1 Authentication 16](#_Toc180415710)

[9.2.2 Dashboard 16](#_Toc180415711)

[9.2.3 Heart Disease Detection 17](#_Toc180415712)

[9.2.4 Analysis and Predictions 17](#_Toc180415713)

[9.2.5 Report Generation 17](#_Toc180415714)

[9.2.6 Notifications and Reminders 17](#_Toc180415715)

[9.2.7 User Profile Management 18](#_Toc180415716)

[10 Key Performance Indicators (KPIs) 18](#_Toc180415717)

[11 Evaluation Criteria 19](#_Toc180415718)

[11.1 Data Sourcing and Quality: 19](#_Toc180415719)

[11.2 Data Collection Methodology: 19](#_Toc180415720)

[11.3 Model Training and Testing: 19](#_Toc180415721)

[11.4 Performance Metrics: 19](#_Toc180415722)

[11.5 Real-World Applicability: 20](#_Toc180415723)

[11.6 User Feedback and Iteration: 20](#_Toc180415724)

[12 References 20](#_Toc180415725)

[12.1 Dataset References: 20](#_Toc180415726)

[12.2 Random Forest Algorithm: 20](#_Toc180415727)

[12.3 Support Vector Machines: 20](#_Toc180415728)

[12.4 Machine Learning in Healthcare: 20](#_Toc180415729)

[12.5 Previous Projects: 20](#_Toc180415730)

**CardioCare AI: Intelligent Heart Disease Prediction and Diagnosis System**

# Overview

The **CardioCare AI: Intelligent Heart Disease Prediction and Diagnosis System** is a comprehensive healthcare platform designed to enhance the early detection and diagnosis of heart diseases. By leveraging advanced machine learning algorithms and multi-level datasets, the system offers an intuitive interface for patients to input health data and receive personalized health insights.

The platform incorporates four distinct datasets that cover various aspects of heart health, including the presence of heart disease, identification of STEMI and NON-STEMI conditions, and narrowing arteries. Patients can engage with the system through a user-friendly dashboard that supports symptom entry, and AI-driven predictions.

Additionally, the system features an AI chatbot for health queries, personalized health reports, and automated notifications for appointments and medication reminders. Admin functionalities enable the management of patient records and the overall system performance, ensuring smooth operation and data integrity. Ultimately, the project aims to improve access to healthcare services, promote preventative care, and contribute to Sustainable Development **Goal 3 (Good Health and Well-Being)** by providing equitable health diagnostics.

# Domain

The domain of this project is **online healthcare services** powered by **artificial intelligence (AI)**. The platform allows patients to access health diagnostics, monitor symptoms, and check for diseases like heart conditions from anywhere. With AI tools for quick diagnosis and report generation, the system helps patients manage their health easily. The project focuses on making healthcare more accessible and efficient by offering users an online platform where they can get basic health advice and recommendations without needing in-person visits.

# SDG Focus: Good Health and Well-Being (SDG 3)

The "**CardioCare AI**" supports Sustainable Development Goal 3 by addressing critical healthcare challenges:

1. Enhanced Access to Healthcare:  
   The platform provides web-based heart disease diagnostics accessible to anyone with internet connectivity, particularly benefiting individuals in remote or underserved areas.
2. **Promoting Early Disease Detection**:  
   By utilizing AI for symptom analysis, the system facilitates early detection of cardiovascular issues, enabling timely medical intervention and reducing the risk of complications.
3. **Improvement in Healthcare Quality**:  
   AI-driven diagnostics enhance the accuracy of heart disease detection, offering personalized health reports that support informed decision-making by healthcare providers.
4. **Encouragement of Preventative Care**:  
   Real-time health insights empower patients to monitor their cardiovascular health regularly and take proactive steps to prevent serious conditions.
5. **Reduction of Healthcare Inequality**:  
   The platform aims to minimize disparities in access to healthcare by providing affordable and user-friendly diagnostic solutions, promoting equitable healthcare for all.

# Technology Stack

## MySQL:

* + Relational database for storing patient records, health data, and user profiles.
  + Ensures data integrity and supports complex queries.

## Express.js:

* + Web application framework for Node.js to build RESTful APIs.
  + Facilitates smooth communication between the front end and back end.

## React.js:

* + Frontend library for building user interfaces.
  + Provides a responsive and interactive experience for users, enabling dynamic data rendering.

## Node.js:

* + JavaScript runtime environment for server-side programming.
  + Handles client requests and serves APIs, ensuring efficient performance and scalability.

## AI Integration Tools:

* + Libraries like TensorFlow.js or Scikit-learn for implementing AI models for symptom checking and heart disease detection.

## Data Visualization Libraries:

* + Tools like Chart.js or D3.js for creating graphical representations of patient data and analytics on the admin dashboard.

## Authentication:

* + Tools like JWT (JSON Web Tokens) for secure user authentication and session management.

## Deployment:

* + Hosting solutions such as Heroku, AWS, Netlify or Vercel for deploying the application and managing server resources.

# Datasets

The CardioCare AI system will utilize four distinct datasets, each corresponding to different levels of patient evaluation. These datasets are integral for accurately diagnosing heart disease and providing comprehensive health insights.

## Heart Disease Presence Detection

This dataset is used to determine whether a patient has heart disease based on various health attributes.

### Attributes

* **age**: Age of the patient (real)
* **sex**: Gender of the patient (real; 1 for male, 0 for female)
* **chest pain type**: Type of chest pain experienced (real)
* **resting blood pressure**: Blood pressure at rest (real)
* **serum cholesterol**: Serum cholesterol level (real)
* **fasting blood sugar**: Fasting blood sugar level (real)
* **resting electrocardiographic results**: Results from resting ECG (real)
* **maximum heart rate achieved**: Highest heart rate during exercise (real)
* **exercise induced angina**: Presence of angina during exercise (real)
* **oldpeak**: ST depression induced by exercise (real)
* **slope**: Slope of the peak exercise ST segment (real)
* **number of major vessels**: Number of major vessels colored by fluoroscopy (real)
* **thal**: Thalassemia (real)
* **class**: Diagnosis outcome (categories: **absent** or **present**)

### Categories

* **Absent**: Indicates that the patient does not have heart disease.
* **Present**: Indicates that the patient has been diagnosed with heart disease.

## Myocardial Infarction Type Classification

This dataset classifies patients based on the type of myocardial infarction (AMI) they experience, helping to differentiate between critical heart conditions.

### Attributes

* **Group**: Overall classification (0 for control, 1 for AMI)
* **Sub-type for AMI**: Type of AMI (categories: **STEMI: 0**, **NON-STEMI: 1**, **Control: 2**)
* **age**: Patient’s age
* **gender**: Patient’s gender (0 for female, 1 for male)
* **WBC**: White blood cell count
* **RBC**: Red blood cell count
* **HGB**: Haemoglobin level
* **Other blood parameters**: Including BA, EO, LY, MO, NEU, NEU/LY, PLT/LY, MPV/LY, LY/MO, MPV, PLT, HCT, MCV, RDW-SD, RDW-CV, MCH, MCHC, PDW, PCT

### Categories

* **Control (0)**: Indicates healthy patients without any myocardial infarction.
* **AMI (1)**: Indicates patients experiencing a myocardial infarction.
  + **STEMI (0)**: A severe type of heart attack.
  + **NON-STEMI (1)**: A less severe but still critical condition.

## Artery Narrowing Analysis

This dataset assesses artery narrowing, a significant indicator of heart disease.

### Attributes

* **age**: Age of the patient (real)
* **sex**: Gender of the patient (categories: **female**, **male**)
* **cp**: Chest pain type (categories: **typical angina**, **asymptomatic**, **non-anginal**, **atypical angina**)
* **trestbps**: Resting blood pressure (real)
* **chol**: Serum cholesterol (real)
* **fbs**: Fasting blood sugar (categories: **t** for true, **f** for false)
* **restecg**: Resting electrocardiographic results (categories **left ventricular hypertrophy**, **normal**, **ST-T wave abnormality**)
* **thalach**: Maximum heart rate achieved (real)
* **exang**: Exercise induced angina (categories: **no**, **yes**)
* **oldpeak**: ST depression induced by exercise (real)
* **slope**: Slope of peak exercise ST segment (categories: **up**, **flat**, **down**)
* **ca**: Number of major vessels (real)
* **thal**: Thalassemia status (categories: **fixed defect**, **normal**, **reversible defect**)
* **num**: Degree of artery narrowing (categories: **'<50'**, **'>50\_1'**, **'>50\_2'**, **'>50\_3'**, **'>50\_4'**)

### Categories

* **Chest Pain Type**:
  + **Typical Angina**: Classic symptom of heart disease.
  + **Asymptomatic**: No symptoms present.
  + **Non-Anginal**: Pain not related to heart disease.
  + **Atypical Angina**: Symptoms that are not typical for angina.
* **Degree of Artery Narrowing**:
  + **'<50'**: Narrowing of less than 50%.
  + **'>50\_1'**: Narrowing of 50% to 60%.
  + **'>50\_2'**: Narrowing of 60% to 70%.
  + **'>50\_3'**: Narrowing of 70% to 80%.
  + **'>50\_4'**: Narrowing of more than 80%.

# Algorithms

In the CardioCare AI system, a variety of machine learning algorithms will be employed to analyze the datasets and provide accurate heart disease predictions. The choice of algorithm will depend on the nature of the datasets and the specific classification tasks.

## Dataset Analysis

Initially, we will explore each dataset using different visualization techniques to gain insights into the data characteristics. This will include:

* **Heatmaps**: To visualize correlations between different attributes, helping identify which factors are most influential in predicting heart disease.
* **Box Plots**: To detect outliers and understand the distribution of attributes like cholesterol levels and blood pressure.
* **Pair Plots**: To observe the relationships between multiple variables and how they cluster according to heart disease presence or absence.

## Algorithm Selection

After a thorough analysis of the datasets, we will research and evaluate several machine learning algorithms, including:

* **K-Nearest Neighbors (KNN)**: A simple, yet effective classification algorithm that can be used for predicting whether a patient has heart disease based on similarity to other cases in the dataset.
* **Random Forest**: An ensemble learning method that constructs multiple decision trees during training and outputs the mode of their predictions, providing robustness against overfitting.
* **Support Vector Machine (SVM)**: A powerful classification algorithm that finds the optimal hyperplane to separate classes, suitable for both linear and non-linear data.
* **Logistic Regression**: A statistical method for binary classification that models the probability of a class label based on one or more predictor variables.
* **Gradient Boosting**: An ensemble technique that builds models sequentially, optimizing the errors made by previous models, often leading to higher predictive accuracy.

We will compare the performance of these algorithms through various metrics such as accuracy, precision, recall, and F1-score. The best-performing algorithm will then be selected for implementation in the final model.

# Data Preprocessing

Data preprocessing is a crucial step in preparing the datasets for machine learning algorithms. Proper preprocessing ensures that the data is clean, formatted, and ready for analysis, enhancing the model's performance. The preprocessing steps include:

## Data Cleaning

* **Handling Missing Values**: Identify and impute or remove missing values to ensure a complete dataset. Techniques may include mean/mode imputation or utilizing algorithms that can handle missing data.
* **Removing Duplicates**: Check for and eliminate any duplicate records to maintain data integrity.

## Data Transformation

* **Normalization/Standardization**: Scale numerical attributes to ensure that they contribute equally to the distance calculations in algorithms like KNN and SVM. This involves scaling data to a range (0 to 1) or transforming it to have a mean of 0 and a standard deviation of 1.
* **Encoding Categorical Variables**: Convert categorical attributes (e.g., sex, chest pain type) into numerical formats using techniques like one-hot encoding or label encoding, enabling the algorithms to interpret them correctly.

## Splitting the Data

* **Train-Test Split**: Divide the dataset into training and testing sets, typically using an 80-20 or 70-30 ratio. The training set is used to train the model, while the testing set evaluates its performance.

## Feature Selection

* **Identifying Important Features**: Utilize techniques like recursive feature elimination, feature importance from tree-based models, or correlation matrices to select the most relevant features that contribute to heart disease predictions.

# User Roles and Functionalities

The CardioCare AI system supports two primary user roles: **Admin** and **Patient**. Each role has specific functionalities designed to enhance user experience and streamline system operations. Below are the detailed responsibilities and features available to each user role.

## Admin

The Admin role encompasses comprehensive control over the CardioCare AI system, facilitating efficient management and oversight of platform operations.

* **System Management**: The admin has full authority to manage the overall platform's functionality, ensuring a seamless experience for all users.
* **Patient Monitoring**: Admins can view a complete list of patients along with their diagnostic history. This allows for tracking patient engagement and ensuring timely follow-ups based on their health records.
* **User Account Management**: The admin is responsible for creating, editing, and deleting user accounts, both for patients and any additional administrative staff. This functionality ensures that the platform maintains a secure and organized user base.
* **Review AI-Generated Reports**: Admins have access to all AI-generated reports for patient diagnostics, allowing them to verify the accuracy of the predictions made by the AI models.
* **Data Oversight**: The admin can monitor all data being processed across the platform, including the performance metrics of various algorithms, to ensure optimal functioning of the system.
* **Visualization Tools**: The admin can view graphical data representations, such as heatmaps of patient health trends, statistics on disease prevalence, and overall platform usage metrics, aiding in strategic decision-making for service improvements.

## Patient

* Patients form the core user group of the CardioCare AI system. They will engage with the platform through an intuitive interface that provides access to health diagnostics, personalized insights, and essential health records.

### Authentication Flow

* **Sign-up/Login**: Patients can easily register or log in using their email or phone number. Enhanced security measures, such as OTP-based authentication, will be implemented to protect user accounts.

### Heart Disease Detection

* **Symptom Entry**: Patients can enter relevant heart-related symptoms, such as chest pain or shortness of breath, which will be analyzed against the first dataset (Heart Disease Presence).
* **AI Predictions**: The AI model will leverage textual input data to assess the likelihood of heart disease, categorizing results into actionable insights such as whether to seek further consultation with a cardiologist.
* **Real-Time Results**: Patients receive immediate feedback with a detailed breakdown of the analysis, including risk factors based on their provided symptoms and scan results.

### Chatbot (AI Health Assistant)

* **AI Chat Support**: An integrated AI-powered chatbot will assist patients with health-related inquiries. The chatbot will provide guidance based on the current health trends indicated in the datasets, such as general heart health tips or responses to specific symptom queries.
* **Natural Language Processing (NLP)**: Using NLP algorithms, the chatbot will understand patient inputs in natural language, allowing for meaningful interactions and relevant health information retrieval.

### Profile and Health Records Management

* **Health Dashboard**: Patients will have access to a personalized health dashboard that displays their vital statistics, including BMI, heart rate, and other relevant health metrics. Integration with wearable devices will enhance data accuracy and provide real-time monitoring.
* **Medical Records**: Patients can view their past medical records, diagnostic reports from the AI predictions based on the datasets, and treatment histories, ensuring they are well-informed about their health journey.
* **Record Updates**: Patients can easily update their personal details and track changes in their medical history, ensuring the information in the system is current and accurate.

### Report Generation and Download

* **Custom Reports**: After diagnostic sessions, patients will receive personalized health reports that summarize their results from the datasets, including the likelihood of heart disease based on symptom entry and diagnostic imagery.
* **Visual Insights**: Reports will include visual elements such as graphs, charts, and tables summarizing health trends and metrics, facilitating a better understanding of their health status.
* **Downloadable Reports**: Patients can download their reports in PDF format for personal use or for sharing with healthcare providers, enabling continuity of care.

### Notifications and Reminders

* **Health Alerts**: The system will send automated notifications about upcoming medical appointments, new health reports, or significant changes in health status based on the AI’s analysis of their input data.
* **Medication Reminders**: Patients can set reminders for their medications and follow-up appointments, enhancing adherence to prescribed healthcare regimens.

# Complete Flow and UI Interfaces

The **CardioCare AI** system provides a structured flow for patients to access heart disease diagnostics seamlessly, integrating four multilevel datasets. This section outlines the user experience from authentication to report generation, accompanied by detailed UI components.

## Admin

Admins have access to various features for managing the system and ensuring smooth operations:

### Authentication

* **Elements**: Input fields for email and password.
* **Buttons**:
  + - **Login**: Authenticates user credentials.
    - **Sign Up**: Redirects to the registration page.
    - **Forgot Password**: Sends a password reset link.

### Dashboard

* + **Elements**: Overview of platform metrics (active users, diagnostics performed).
  + **Buttons**:
    - **Manage Users**: View and edit patient accounts.
    - **View Diagnostics**: Access diagnostic history of all patients.
    - **Generate Reports**: Create system-wide performance reports.

### User Management:

* + **Elements**: List of registered patients with search and filter options.
  + **Buttons**:
    - **Edit User**: Modifies patient details.
    - **Delete User**: Removes a patient from the system.

### View Diagnostics:

* + **Elements**: Comprehensive view of patient diagnostic history.
  + **Buttons**:
    - **View Report**: Opens individual diagnostic reports.

### Performance Analytics:

* + **Elements**: Graphs and charts depicting platform usage, patient trends, and diagnostics.
  + **Buttons**:
    - **Download Report**: Exports performance data for offline analysis.

### System Settings:

* + **Elements**: Configuration options for system parameters.
  + **Buttons**:
    - **Update Settings**: Saves changes to system configurations.

## For Patients

### Authentication

* **Login/Sign-Up Screen**:
  + **Elements**: Input fields for email and password.
  + **Buttons**:
    - **Login**: Authenticates user credentials.
    - **Sign Up**: Redirects to the registration page.
    - **Forgot Password**: Sends a password reset link.

### Dashboard

After successful authentication, patients are directed to their dashboard.

* + **Elements**: Profile overview displaying health statistics (e.g., BMI, heart rate).
  + **Buttons**:
    - **Start Heart Disease Detection**: Initiates the diagnostic process.
    - **View Reports**: Access previously generated reports.
    - **Chat with AI Assistant**: Opens the AI chatbot for health inquiries.

### Heart Disease Detection

* **Symptom Entry Interface**:
  + **Elements**: Text fields for entering heart-related symptoms.
  + **Buttons**:
    - **Submit Symptoms**: Triggers analysis using the first dataset (Heart Disease Presence).

### Analysis and Predictions

* + **Elements**:
    - **Predicted Diagnosis**: Displays the likelihood of heart disease based on the analysis.
    - **Suggested Actions**: Recommendations for follow-up consultations with cardiologists.
  + **Buttons**:
    - **Generate Report**: Creates a personalized health report based on the analysis.

### Report Generation

* + **Elements**:
    - **Health Report**: Includes diagnostic results, AI predictions, and personalized health insights.
    - **Visualizations**: Graphs and charts illustrating health metrics.
  + **Buttons**:
    - **Download PDF**: Downloads the report for personal records.
    - **Email Report**: Sends the report to the registered email address.

### Notifications and Reminders

* + **Elements**: Organized list of notifications with timestamps.
  + **Buttons**:
    - **View Details**: Expands notifications for more information.

### User Profile Management

* **Profile Management Interface**:
  + **Elements**: Editable fields for personal information and medical history.
  + **Buttons**:
    - **Save Changes**: Updates the profile with new information.

# ****Key Performance Indicators (KPIs)****

The success of the **CardioCare AI** system will be measured through the following KPIs:

* **Diagnostic Accuracy**: Target 90% accuracy in heart disease predictions to ensure reliable results.
* **User Satisfaction Rate**: Aim for an 85% satisfaction rate based on user feedback to confirm that the platform meets patient needs.
* **Rate of Early Disease Detection**: Strive for a 75% detection rate for early-stage heart diseases to improve treatment outcomes.
* **Chatbot Engagement Rate**: Target a 60% engagement rate with the AI chatbot for health inquiries, indicating its usefulness.
* **Report Generation Time**: Ensure reports are generated in under 5 minutes for timely access to health insights.
* **System Uptime**: Maintain 99.5% uptime to ensure continuous access for users.
* **Number of Users Registered**: Aim for 1,000 registered users within the first year to reflect acceptance and reach.
* **Frequency of Use**: Encourage patients to use the system at least twice a month for ongoing health monitoring.
* **Compliance with Follow-Up Recommendations**: Target a 70% compliance rate for patients following up with healthcare providers after diagnostics.
* **Data Security Incidents**: Aim for zero security incidents to protect patient data and maintain trust.

These KPIs will help evaluate the system's effectiveness and guide continuous improvements in patient care.

# Evaluation Criteria

The **CardioCare AI** system's effectiveness will be evaluated based on its performance with a dataset sourced from the **Punjab Institute of Cardiology (PIC)**. This dataset is critical for ensuring that the system is well-tailored to the healthcare needs of the Pakistani population. The following criteria will be applied during the evaluation phase:

## Data Sourcing and Quality:

* The dataset has been meticulously curated from the PIC, one of the leading cardiac care institutions in Pakistan. This ensures that the data reflects real-world conditions and patient demographics specific to the region.
* Reference consultants **Dr. Gulfam Hashmi** and **Dr. Umair Abid**, both of whom are affiliated with PIC, provided insights and guidance during the data collection process. Their expertise ensures the relevance and accuracy of the dataset.

## Data Collection Methodology:

* The dataset was acquired through physical visits to the PIC, allowing for direct interaction with healthcare professionals and access to comprehensive medical records. This hands-on approach enhances the reliability of the data.
* Patient records include various attributes such as age, gender, medical history, diagnostic results, and treatment outcomes, providing a well-rounded basis for analysis.

## Model Training and Testing:

* The dataset from the Punjab Institute of Cardiology (PIC) will be used to train and validate the AI models, focusing on accurately predicting the presence and type of heart disease.
* Algorithms such as K-Nearest Neighbors (KNN), Random Forest, and Support Vector Machines (SVM) will be evaluated based on their performance with the PIC data.

## Performance Metrics:

* Key metrics, including accuracy, precision, recall, and F1 score, will assess the model's performance in correctly classifying heart disease cases and minimizing false positives/negatives.

## Real-World Applicability:

* The system will be tested with actual patient inputs from PIC to evaluate its effectiveness in real-world settings. Feedback from healthcare professionals will ensure the validity of model predictions.

## User Feedback and Iteration:

* Post-deployment, feedback from patients and healthcare providers will be gathered to assess satisfaction and usability. This feedback will guide further refinements and updates to the system.

# References

## Dataset References:

* Heart Disease Data Set (Presence or Absence of Disease): [**View Dataset**](https://www.openml.org/data/download/53/dataset_53_heart-statlog.arff)
* STEMI and NON-STEMI Data Set: [**View Dataset**](https://prod-dcd-datasets-public-files-eu-west-1.s3.eu-west-1.amazonaws.com/0cbfa50b-a0ec-46c5-96a2-78bad67a857c)
* Narrowing Arteries Data Set: [**View Dataset**](https://www.openml.org/data/download/49/dataset_49_heart-c.arff)

## Random Forest Algorithm:

* + Breiman, L. (2001). "Random Forests." *Machine Learning*, 45(1), 5-32. [View Article](https://link.springer.com/article/10.1023/A:1010933404324)

## Support Vector Machines:

* + Cortes, C., & Vapnik, V. (1995). "Support-Vector Networks." *Machine Learning*, 20(3), 273-297. [View Article](https://link.springer.com/article/10.1007/BF00994018)

## Machine Learning in Healthcare:

* + Rajpurkar, P., et al. (2017). "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning." *arXiv preprint arXiv:1711.05225*. [View Paper](https://arxiv.org/abs/1711.05225)

## Previous Projects:

* [Heart Disease Prediction by nripstein](https://github.com/nripstein/Heart-Disease-Prediction/blob/main/heart%20disease%20prediction%20notebook.ipynb)
* [Heart Disease Prediction by chayandatta](https://github.com/chayandatta/Heart_disease_prediction/blob/master/HeartDisease.ipynb)
* [Heart Disease Prediction by kb22](https://github.com/kb22/Heart-Disease-Prediction/blob/master/Heart%20Disease%20Prediction.ipynb)
* [Heart Disease Prediction by suhasmaddali](https://github.com/suhasmaddali/Heart-Disease-Prediction/blob/main/Heart%20Disease%20Prediction%20Using%20Machine%20Learning.ipynb)
* [Heart Disease Prediction by g-shreekant](https://github.com/g-shreekant/Heart-Disease-Prediction-using-Machine-Learning/blob/master/Heart_disease_prediction.ipynb)