Project Proposal

TEAM 81

Team Members:

Ahmed Sameh.

Mohamed Yasser.

Abdullah Ahmed.

Rewan Adel.

Table of Contents

Table of Contents

Project Overview

The Healthcare Predictive Analytics Project aims to develop a predictive model to improve healthcare outcomes through advanced data-driven insights. By leveraging machine learning and statistical analysis, the project will assist healthcare professionals in patient risk prediction, health trend identification, and informed decision-making. The ultimate goal is to enhance patient care, optimize resource allocation, and support healthcare professionals with actionable insights.

The primary goals are:

- Reduce patient anxiety and healthcare system inefficiency by minimizing misdirected appointments.
- Empowering users with knowledge.

Objectives

- Collect, explore, and preprocess relevant healthcare data for predictive modeling.
- Conduct exploratory data analysis (EDA) to identify trends, patterns, and anomalies in patient data.
- Develop and optimize machine learning models for healthcare outcome prediction.
- Deploy the predictive model with MLOps integration to ensure scalability and continuous monitoring.
- Present findings, insights, and recommendations in a comprehensive report and presentation for stakeholders
- Ensuring they seek the right care at the right time.

DATA :

- Indicators of Heart Disease (2022 UPDATE)

- CAD Cardiac MRI Dataset

Milestones : 9/18/2025

Milestone 1: Project Design & Data Mapping (week 1-2)

Tasks:

- Define project objectives, scope, and success criteria.
- Design system architecture (data sources, storage, pipelines, model flow).
- Map required datasets (tabular health indicators, medical images, patient records).
- Identify target variables and expected outcomes.

Deliverables:

- Project Design Document (objectives, architecture, workflow).
- Data Mapping Schema (ERD, feature mapping, dataset description).
- Initial Project Plan (timeline, roles, tools).

Milestone 2: Data Collection, Exploration & Preprocessing (week 3-5)

Tasks:

- Acquire healthcare datasets (patient demographics, medical history, test results, outcomes).
- Perform EDA to understand data distribution, patterns, and missing values.
- Preprocess dataset: imputation, normalization, and categorical encoding.

Deliverables:

- Dataset Exploration Report.
- EDA Notebook with visualizations (histograms, boxplots, heatmaps).
- Cleaned Dataset.

Milestone 3: Data Analysis & Visualization (week 6-8)

Tasks:

- Perform advanced data cleaning and outlier handling.
- Analyze correlations and health metric relationships.
- Generate interactive dashboards and compelling visualizations using Plotly/Dash/Tableau.

Deliverables:

- Data Cleaning & Analysis Report.
- Visualizations of health trends (heatmaps, scatter plots, dashboards).

Milestone 4: Predictive Model Development & Optimization (week 6-8)

Tasks:

- Select suitable ML models (Logistic Regression, Random Forest, Gradient Boosting, Neural Networks).
- Train and evaluate models using cross-validation and performance metrics (Accuracy, Precision, Recall, F1, ROC-AUC).
- Perform hyperparameter tuning for optimization.

Deliverables:

- Predictive Model Evaluation Report.
- Python Code for model training and evaluation.
- Optimized Final Model.

Milestone 5: MLOps, Deployment & Monitoring (week 9-10)

Tasks:

- Implement experiment tracking using MLflow/Kubeflow.
- Deploy the model as an API or web application (Flask/FastAPI).
- Enable cloud deployment (AWS/Google Cloud/Heroku).
- Establish monitoring pipelines for model drift detection and retraining.

Deliverables:

- Deployed Predictive Model (API/Web App).
- MLOps Report.
- Monitoring Documentation.

Milestone 6: Final Documentation & Presentation (Week 11)

Tasks:

- Compile the full project report, including methodology, results, and challenges.
- Prepare a stakeholder-friendly presentation demonstrating the model's impact.
- Provide recommendations for healthcare integration and future improvements.

Deliverables:

- Final Project Report & Final presentation

Additional Ideas for Enhancement

1. CNN for Medical Image Analysis

To expand beyond structured patient data, we can integrate a **Convolutional Neural Network (CNN)** for analyzing medical images such as chest X-rays or MRIs. This feature will allow the system to:

- Detect diseases (e.g., pneumonia, tumors) automatically from uploaded images.
- Provide visual explanations using methods like **Grad-CAM** to show which part of the image influenced the prediction.
- Assist healthcare professionals by combining image-based predictions with structured data models for a more holistic assessment.

2. Healthcare Chatbot

We can develop a **chatbot assistant** to improve system usability and accessibility. The chatbot will:

- Guide patients or doctors in entering relevant data for predictions.
- Answer common healthcare-related questions in simple terms.
- Explain the predictive model's outputs in a user-friendly way (e.g., "This patient shows high risk due to high glucose and abnormal X-ray patterns").
- Be integrated into the web application frontend, making it an interactive tool for healthcare staff and patients.

Project Timeline

Phase	Week 1-2	Week 3-5	Week 6-8	Week 9-10	Week 11
1. Project Design & Data Mapping	✅				
2. Data Collection, Exploration & Preprocessing		✅			
3. Data Analysis & Visualization			✅		
4. Predictive Model Development & Optimization			✅		
5. MLOps, Deployment & Monitoring				✅	
6. Final Documentation & Presentation					✅

Team Roles & Responsibilities

- **Ahmed –** Data & ML: Structured data, preprocessing, predictive models.
- **Rewan** – Visualization & Chatbot: Dashboards, insights, and chatbot development.
- **Mohamed –** CNN & Imaging: X-ray preprocessing, CNN training, disease detection.
- **Abdullah –** MLOps & Deployment: Model deployment, monitoring, documentation.