

## Phase-1 Submission Template

**Student Name: P.M.Nirbhayaram**

**Register Number: 620123106074**

**Institution: AVS Engineering College**

**Department: Electronics and Communication**

**Engineering**

**Date of Submission: 30.04.2025**

---

### 1.Problem Statement

Traditional disease diagnosis often relies on observable symptoms and manual evaluation, which can result in late detection and inefficient care. With rising chronic disease prevalence and overburdened healthcare systems, there is an urgent need for predictive tools that utilize patient data to detect disease risks early and improve clinical decision-making.

### 2.Objectives of the Project

- To develop an AI-based system that predicts disease risks using comprehensive patient data.
- To assist healthcare providers in making early, informed, and personalized treatment decisions.
- To demonstrate the potential of machine learning in reducing diagnostic errors and improving patient outcomes.

### 3.Scope of the Project

- Focus on common chronic diseases (e.g., diabetes, cardiovascular diseases, respiratory conditions).

- **Use anonymized datasets from hospitals, open-source medical databases, or synthetic data for development.**
- **Deploy a web or mobile-based interface for healthcare professionals to input data and view predictions.**
- **Ensure ethical AI practices including bias reduction, data privacy, and model transparency**

#### **4.Data Sources**

- **Open-source datasets (e.g., UCI Machine Learning Repository, MIMIC-IV, Kaggle health datasets)**
- **Simulated or synthetic patient records for training and validation**
- **Optional integration with Electronic Health Records (HER) from healthcare institutions (subject to privacy agreements)**

## 5.High-Level Methodology

- Data Collection: Gather structured patient data (demographics, medical history, lab reports, lifestyle).
- Data Preprocessing: Clean, normalize, and prepare data; handle missing values and outliers.
- Model Development: Train machine learning models (e.g., Random Forest, XGBoost, Neural Networks).
- Model Evaluation: Validate models using metrics such as accuracy, precision, recall, and AUC.
- Deployment: Build a user interface and deploy the model for real-time prediction use.

## 6.Team Members and Roles

- Programming Languages: Python
- Libraries/Frameworks: Scikit-learn, TensorFlow, Keras, Pandas, NumPy, SHAP (for explainability)
- Development Tools: Jupyter Notebook, VS Code
- Deployment Platforms: Streamlit, Flask, or web hosting services (e.g., Heroku, AWS)
- Version Control: Git/GitHub

## 7.Team Members And Role

P.M.Nirbhayaram –Project Lead & Data Scientist

Oversees the overall project development and timeline.Designs and implements machine learning models.Performs data preprocessing, feature engineering, and model evaluation..Ensures model explainability, accuracy, and compliance with healthcare standards.

Naveenkumar.K –Software Developer & System Integrator

Develops the front-end and back-end for the prediction system (web or mobile interface).Integrates the trained AI model into the application.Manages deployment using cloud platforms (e.g., AWS, Heroku).Ensures UI/UX is user-

friendly for healthcare professionals.

### B.Salman-3. Data Engineer & Analyst

Collects, cleans, and prepares patient datasets from various sources. Handles database management and data pipelines. Conducts exploratory data analysis (EDA) and generates visual insights. Works with the data scientist to ensure high-quality training data.