

**CS550/DSL501: Machine Learning (2024–25–M)**  
**Statement Of Purpose**

**Full name:** R C Yajour Kichenamourty

**ID:** M24MT006

## **Explainable AI Based Disease Prediction System**

### **1 Problem Statement**

People nowadays suffer from a variety of diseases because of environmental factors and their lifestyle choices. As a result, disease prediction at an earlier stage becomes a critical task. In the current healthcare landscape, disease prediction and diagnosis often rely on traditional methods, which may not always yield accurate or timely results, especially for complex diseases. With the growing availability of healthcare data from Electronic Health Records (EHRs), medical imaging, and wearable devices, the application of Machine Learning (ML) in healthcare has gained prominence. However, while ML models have shown great potential in disease prediction, they often operate as "black boxes," making it difficult for medical professionals and patients to trust and understand the predictions. This lack of transparency limits the adoption of ML system in critical healthcare environments, where explainability and trust are significant.

The problem lies in developing an ML-based healthcare application that can not only predict diseases with high accuracy but also provide clear explanations for its predictions. This will allow healthcare professionals to make informed decisions and gain trust in AI-powered medical tools.

### **2 Motivation**

The motivation behind this project is to improve healthcare outcomes by leveraging Machine Learning techniques to enable earlier and more accurate disease diagnosis. Many diseases, such as diabetes, heart disease and Alzheimer's disease have better prognoses when detected early. Machine learning models have shown great potential in analyzing large datasets and identifying patterns that are not immediately visible to human practitioners. However, their adoption in real-world clinical practice has been slow due to the lack of interpretability in decision-making processes.

Explainable AI (XAI) offers the opportunity to bridge this gap by ensuring that the predictions made by the ML models are not only accurate but also understandable and transparent to healthcare providers. By combining ML with explainable AI, the project aims to increase trust in AI-based healthcare systems, making them more practical for real-world clinical use, improving patient outcomes, and assisting doctors in making better decisions.

### 3 Objectives

The **primary objective** of the proposed system is to bridge the gap between medical expertise and the general public by providing a user-friendly and transparent platform for self-diagnosis and medication guidance. The system integrates Explainable AI (XAI) to enhance user understanding of the reasoning behind disease predictions and medication recommendations.

**The secondary objectives of this project are:**

**To improve accuracy and reliability:** Achieve high predictive accuracy while maintaining the interpretability of the model's decisions.

**To evaluate the system on real-world datasets:** The model will be tested and validated on public healthcare datasets to ensure its effectiveness in disease prediction.

**User-friendly interface:** Design a user-friendly interface for medical professionals and general public to interact with the model's predictions and explanations.

### 4 Relevant Study

A large body of research highlights the effectiveness of machine learning models in disease prediction across multiple healthcare areas. Algorithms such as decision trees, neural networks, support vector machines (SVMs), and deep learning have been applied to predict diseases such as cancer, diabetes, and cardiovascular diseases using patient data from EHRs, genetic data, and imaging data. Recent advancements in Explainable AI (XAI) have led to methods like LIME (Local Interpretable Model-Agnostic Explanations), SHAP (SHapley Additive exPlanations), and Grad-CAM (Gradient-weighted Class Activation Mapping) being used to interpret complex models and provide human-interpretable insights.

However, many machine learning systems used in healthcare still face challenges related to model transparency and user trust, particularly in high-stakes environments like healthcare. Explainable AI has emerged as a key solution to address this issue, offering methods that can explain the predictions made by complex models in ways that are accessible to non-experts. Studies have shown that implementing XAI increases trust in AI systems and enhances their adoption in critical sectors like healthcare.

Moreover, research suggests that combining predictive analytics with interpretability can improve the decision-making process for healthcare providers, leading to better patient outcomes.

Sl. No	Paper Name and Year	Journal/ Paper Details	Paper Link
1	Explainable Artificial Intelligence Based Framework for Non Communicable Diseases Prediction (2023)	This Journal proposes a Deep Shapley Additive Explanations (DeepSHAP) based deep neural network framework equipped with a feature selection technique for NCDs prediction and explanation among the population in the United States.	Link1
2	Integration of Explainable Artificial Intelligence (XAI) in the Development of Disease Prediction and Medicine Recommendation System (2024)	The paper proposes to bridge the gap between medical expertise and the general public by providing a user-friendly and transparent platform for self-diagnosis and medication guidance. The system integrates Explainable AI (XAI) to enhance user understanding of the reasoning behind disease predictions and medication recommendations.	Link2
3	Comparison and Analysis of Various Machine Learning Algorithms for Disease Prediction (2023)	This paper presents a system wherein data can help the medical filed experts to detect the fatal diseases early and thus the survival rate of the victims will be increased. The Disease Prediction system is based on predicting the name of the disease of the user examining the symptoms that user gives as an input to the system. The system also predicts the risk due to the occurrence of the disease compared to the general disease which is lower or higher.	Link3

## 5 Proposed Solution

The proposed solution is a comprehensive Machine Learning-Based Healthcare Application with Explainable AI for Disease Prediction. This system will consist of the following components:

**Data Collection and Preprocessing:** The system will gather patient data from publicly available healthcare datasets and preprocess it to handle missing values, outliers, and irrelevant features.

**Machine Learning Models for Disease Prediction:** A set of machine learning algorithms, including logistic regression, decision tree, random forest classifier and support vector machines will be trained to predict various diseases using patient medical history and clinical test results.

**Explainable AI Integration:** XAI techniques such as SHAP, LIME, and explainable Boosting Machines will be integrated into the ML models to generate human-understandable explanations for the predictions. For instance, for a patient diagnosed with potential heart disease, the system will explain how various factors, such as cholesterol level, blood pressure, and lifestyle choices, contributed to the prediction.

**Model Evaluation:** The system's performance will be evaluated using metrics such as accuracy, precision, recall, F1-score, and AUC-ROC curves.

**User-Friendly Interface:** A front-end dashboard will be developed for healthcare professionals and general public to interact with the system, view predictions, and understand the model's reasoning through visual and textual explanations.

The system aims to provide doctors, medical professionals and public with not only accurate predictions but also a high level of transparency and trust, thereby making the application practical and reliable in clinical environments.

Team Name: MT67

- 1) R. C. Yajour Kichenamourty - M24MT006
- 2) Sagar Sudhir Pathak - M24MT007