Transforming healthcare with AI-powered disease prediction based on patient data.

Student Name: S. SUJI
Register Number:511023205004
Institution: Jei Mathaajee College of Engineering
Department: B.Tech(IT)
Date of Submission: 25-04-2025

1.Problem Statement:

Develop an AI-powered disease prediction system that leverages patient data to accurately predict the likelihood of disease onset, enabling early intervention and improving patient outcomes.

2.Objective:

Artificial Intelligence (AI) has revolutionized the field of healthcare by enhancing disease diagnosis and treatment. AI's ability to analyse vast amounts of data quickly and accurately has led to significant advancements in medical practice.

- ✓ **Develop an AI-powered disease prediction system:** Create a system that leverages machine learning algorithms to predict disease onset based on patient data.
- ✓ **Improve patient outcomes:** Enable early intervention and treatment, potentially improving patient outcomes.
- ✓ **Enhance patient care:** Provide personalized disease risk assessments and treatment plans.

3.Scope of the Project:

- ✓ **Data Collection:** Gathering patient data from various sources (e.g., EHRs, medical imaging, genomic data).
- ✓ **AI Model Development:** Developing and training machine learning models for disease prediction.
- ✓ **Model Evaluation:** Evaluating the performance of the AI models using relevant metrics.

- ✓ **System Development**: Developing a user-friendly system for healthcare professionals to access and utilize the disease prediction models.
- ✓ **Clinical Validation:** Validating the effectiveness of the AI-powered system in clinical settings.

4.Data Sources:

- ✓ **Electronic Health Records (EHRs):** Patient medical history, diagnoses, medications, and treatment plans.
- ✓ **Medical Imaging Data:** Images from X-rays, CT scans, MRI scans, and other medical imaging modalities.
- ✓ **Genomic Data:** Genetic information, including DNA sequencing and genotyping.
- ✓ **Wearable Device Data:** Data from wearable devices, such as fitness trackers and smartwatches.
- ✓ **Patient-Generated Data:** Data generated by patients, such as self-reported symptoms and health behaviours.

5. High-Level Methodology:

- ➢ **Data Collection**- Identify data sources: Gather patient data from various sources (E.g EHRs, medical imaging, wearable devices).
- ➢ **Data Cleaning –** Data cleaning helps remove errors, inconsistencies, and inaccuracies.
- ➢ **Exploratory Data Analysis (EDA)-** Analyze data distributions and statistics to understand the characteristics of the data.
- ➢ **Feature Engineering**- Feature engineering helps create relevant features that improve model accuracy and performance.
- ➢ **Model Building –** Use supervised learning algorithms, such as logistic regression, decision trees, or random forests, to predict disease onset.
- ➢ **Model Evaluation-** Measure the proportion of correctly predicted instances, Measure the model's ability to distinguish between positive and negative classes.
- ➢ **Visualization & Interpretation-** Visualize feature importance and correlations, contribution of each feature to the model's predictions.

- **Deployment-** Deploy models on cloud platforms, such as AWS or Google Cloud, for scalability and reliability.

6.Tools and Technologies:

- ✓ Programming Language – Python
- ✓ Notebook/IDE – Google Colab / Jupyter Notebook
- ✓ Libraries – pandas, numpy, seaborn, matplotlib, scikit-learn, xgboost, folium
- ✓ Optional Tools for Deployment – Streamlit, Flask.

7. Team Members and Roles

1. K. KOMALA
2. D. ANUSUYA
3. J. KUTTI