Software Requirements Specification (SRS)

Project: AI powered diabetes predictor

Date: 31/12/2024

Group Member:

1. Mahnoor Javed (FA21-BSI-024)
2. Aliha Batool (FA21-BSI-006)
3. Ayesha Dawood (FA20-BSI-019)

Purpose

The purpose of this document is to outline the requirements for the development of **Diabetes Predictor**, an AI-powered tool to predict the likelihood of diabetes in individuals based on health parameters. The tool will allow healthcare professionals to input patient data and receive predictions using a trained model.

Scope

Diabetes Predictor will be a tool used for:

- Predicting diabetes based on patient health data.

Product Description

Diabetes Predictor is a tool for machine learning, for diabetes risk assessment. It uses data that is publicly available at Kaggle for training and Random Forest for prediction.

Product Features

- **Data Input:** Upload individual patient records.

- **Prediction Engine:** Predict diabetes likelihood based on health metrics.

- **Error Handling:** Validate inputs for missing or incorrect data.

User Characteristics

- **Primary Users:** Healthcare professionals, researchers, and policymakers.

- **Technical Requirements:** Basic understanding of health metrics; no programming skills required.

Constraints

- The results depend upon the quality and completeness of the individual data.
- Accuracy is limited by the chosen Random Forest model and training dataset.

Hardware Requirements

- **Processor:** Intel Core i5 or equivalent.

- **Memory (RAM):** 8 GB minimum, 16 GB recommended.

- **Storage:** 10 GB free disk space.

Software Requirements

- **Programming Language:** Python 3.8 or latest.

- **Libraries:** scikit-learn, pandas, numpy, tkinter.

- **Environment:** Jupyter Notebook for development.

Network Requirements

- Internet access for downloading libraries and datasets.

Functional Requirements

Data Handling

- Accept health data through manual input or file upload.

- Validate input data and notify users of missing or invalid values.

Model Training and Prediction

- Train Random Forest model on Kaggle diabetes dataset.

- Provide predictions and confidence scores for diabetes likelihood..

Performance Requirements

- Process single-patient predictions in a couple of seconds.

- Achieve at least 85% accuracy on test datasets.

Schedule

Phases	Timeline
Data collection and cleaning	5 days
Modelling and Training	2 days
GUI Development	4 days
Testing and Deployment	5 days

Conclusion

This SRS defines the framework for building **DiabetesPredictor**, a simple yet robust tool for predicting diabetes in patients. By leveraging machine learning, the tool aims to enhance early detection and management of diabetes, improving patient outcomes.