# Disease Prediction Using Machine Learning

***Project Overview:***

This project aims to leverage machine learning techniques to predict the likelihood of a person having a particular disease based on various health parameters. Specifically, we focus on predicting diabetes using a dataset containing medical measurements and patient information. The application provides an interactive web interface using Streamlit, enabling users to input their health data and receive predictions in real-time.

***Objectives:***

* To develop a predictive model using machine learning algorithms that can accurately predict the presence of diseases based on input report.
* To create a user-friendly web application using Streamlit for easy interaction and real-time predictions.
* To implement data preprocessing steps, including feature scaling, to ensure the model performs optimally.
* To educate users on the potential of machine learning in healthcare and its applications in early disease detection.

***Technologies Used:***

1. Programming Language: Python
2. Libraries: scikit-learn, NumPy, Pandas , Streamlit
3. Tools: Jupyter Notebook for model training, Streamlit for web application

**Methodology:**

1. Data Collection: Gather a dataset containing health-related features such as blood pressure, BMI, age, etc.

2. Data Preprocessing: Clean the dataset and perform necessary preprocessing steps like handling missing values and feature scaling using StandardScaler.

3. Model Training: Train a machine learning model (e.g., Logistic Regression, Random Forest, etc.) using the processed data.

4. Model Evaluation: Evaluate the model's performance using metrics like accuracy, precision, recall, and F1-score to ensure reliability.

5. Web Application Development: Develop a Streamlit application that allows users to input their health parameters and receive predictions.

**Implementation Steps:**

1. Load and Preprocess Data:

* Use pandas to load the dataset.
* Apply StandardScaler for feature scaling.

2. Train the Model:

* + Split the data into training and testing sets.
  + Train a machine learning model using the training data.
  + Save the trained model and scaler using pandas.

3. Develop Streamlit Application:

* + Create input fields for users to enter their health data.
  + Load the saved scaler and model.
  + Preprocess user input data and make predictions.
  + Display the prediction results on the web interface.

***Key Features:***

- User Input: Interactive input fields for users to enter their health parameters.

- Real-time Prediction: Immediate prediction results based on the entered data.

- Scalability: The application can be extended to predict other diseases by training on relevant datasets.

**Expected Outcomes:**

- A functional web application capable of predicting diseases based on user input.

- Increased awareness and understanding of how machine learning can be applied to healthcare.

- A basis for further development and enhancement to include more diseases and advanced predictive models.

**Future Work:**

- Integrate additional health metrics and data sources for more comprehensive predictions.

- Implement more advanced machine learning techniques such as ensemble methods or deep learning for improved accuracy.

- Expand the application to predict multiple diseases and provide more detailed health insights.

**Conclusion:**

This project demonstrates the potential of machine learning in the field of healthcare, specifically for disease prediction. By developing an accessible web application, we aim to provide a tool that can assist in early detection and prompt medical consultation, ultimately contributing to better health outcomes.