Machine Learning
Project

Bachelor of Computer Application

Team Members

Rajas Bagga	500126194
Tanisha Saini	500119718
Akshay Kakkar	500119191
Japtez Singh	500121370

Submitted to
Mr. Pankaj Kumar

University of Petroleum & Energy Studies
Bidholi, Via Prem Nagar, Dehradun, Uttarakhand

Jan-May – 2025

Diabetes Detection

Diabetes Detection and Classification Using AI/ML

Introduction:
Diabetes is a chronic disease that affects millions worldwide. Early detection and classification of diabetes can help in better management and treatment. This project aims to develop an AI/ML-based model to predict whether a person has diabetes based on key health indicators.

Dataset:
The model is trained on a simple CSV file containing medical attributes such as glucose level, BMI, insulin, blood pressure, and other relevant features. The dataset is preprocessed to handle missing values, normalize data, and enhance feature selection.

Methodology:
Multiple machine learning algorithms, including Logistic Regression, Decision Trees, Random Forest, Support Vector Machines (SVM), will be tested to determine the most effective model.

Technologies Used:

- Python (NumPy, Pandas, Scikit-Learn)
- Jupyter Notebook / Google Colab for experimentation
- Matplotlib & Seaborn for data visualization

Expected Outcome:
The project will identify the best-performing algorithm for diabetes classification based on predictive accuracy and generalization capability.

Dataset Used:

https://www.kaggle.com/datasets/iammustafatz/diabetes-prediction-dataset