



GOVERNMENT COLLEGE OF ENGINEERING BARGUR
(AUTONOMOUS)

Project : Cloud Application Development

Project Statement: Machine Learning Model Deployment
with IBM Cloud Watson Studio

Team members:

ANJALI B

AKSHAYA N

MUTHULAKSHMI R

MONIKA Y

Phase 2: Innovation

Consider experimenting with ensemble methods or hyperparameter tuning to optimize the model's performance.

1. Introduction

1.1 Project Overview

In this project, we aim to predict the presence of diabetes in patients using machine learning. We'll utilize the Diabetes Database, apply ensemble methods (Gradient Boosting), and perform hyperparameter tuning to optimize our predictive model.

1.2 Dataset Description

The Diabetes Database contains medical information about women, including features such as age, BMI, blood pressure, and pregnancy history. The goal is to predict whether a patient has diabetes based on these attributes.

1.3 Objectives

Our objectives are as follows:

Explore and preprocess the Diabetes Database.

Build and train a machine learning model to predict the presence of diabetes.

Optimize the model using hyperparameter tuning, feature engineering, and ensemble methods (Gradient Boosting).

Analyze the results and gain insights into factors influencing diabetes.

2. Data Exploration

2.1 Data Overview

We will provide an overview of the dataset, including statistical summaries and handling any missing values.

2.2 Data Preprocessing

Data preprocessing steps may include handling missing data, scaling features, and encoding categorical variables.

2.3 Data Visualization

We will create visualizations to better understand the relationships between features and the presence of diabetes.

3. Model Building

3.1 Selecting Machine Learning Algorithms

The Gradient Boosting algorithm will be selected and implemented for the classification task.

3.2 Model Training

We will train the Gradient Boosting model on the Diabetes Database.

3.3 Model Evaluation

The model's performance will be evaluated using appropriate metrics, including accuracy, precision, recall, and F1-score.

4. Model Optimization

4.1 Hyperparameter Tuning

Hyperparameter tuning will be employed to optimize the Gradient Boosting model's performance.

4.2 Feature Engineering

Feature engineering techniques will be applied to enhance model accuracy.

4.3 Ensemble Methods (Gradient Boosting)

The Gradient Boosting ensemble method will be used to improve model robustness and predictive accuracy.

5. Results and Analysis

We will present the results of our model, including performance metrics and insights gained from the analysis.

6. Conclusion

6.1 Summary of Achievements

We will summarize the key achievements of the project, including improvements achieved through hyperparameter tuning and Gradient Boosting.