

# Logistic Regression Project Report

## 1. Introduction

This project demonstrates the implementation of Logistic Regression, a widely used supervised machine learning algorithm for classification problems. The notebook includes steps such as exploratory data analysis (EDA), data preprocessing, model training, evaluation, and visualization of results.

## 2. Exploratory Data Analysis (EDA)

EDA was performed to understand the dataset, check for missing values, analyze data distributions, and identify relationships between variables. Visualization techniques such as histograms, count plots, and correlation heatmaps were used.

## 3. Data Preprocessing

The preprocessing steps included handling missing values, encoding categorical variables, feature scaling, and splitting the dataset into training and testing sets.

## 4. Model Training

A Logistic Regression model was trained using the Scikit-learn library. The model was fitted on the training data and then tested on the testing data.

## 5. Model Evaluation

The performance of the model was evaluated using accuracy score, confusion matrix, classification report, and ROC-AUC curve. These metrics helped assess the precision, recall, F1-score, and overall effectiveness of the model.

## 6. Visualization

Visualizations included confusion matrix heatmaps and ROC curves, which provided insights into the classification performance and trade-offs between sensitivity and specificity.

## 7. Conclusion

The Logistic Regression model achieved a good level of accuracy and demonstrated its effectiveness in solving binary classification problems. Further improvements could be achieved by hyperparameter tuning, feature engineering, or using advanced algorithms.