

Name: Adhija Bachhav

Roll No: 281053

Prn: 22311656

Batch: A-3

ML Assignment 4

Problem Statement:

Apply appropriate ML algorithm on a dataset collected in a cosmetics shop showing details of customers to predict customer response for special offer. Create confusion matrix based on above data and find

- a) Accuracy
- b) Precision
- c) Recall
- d) F-1 score

Introduction

Machine Learning (ML) is widely used for customer data analysis, helping businesses make informed decisions. In this project, we apply an appropriate ML algorithm on a dataset containing customer details to classify customer behaviors or preferences. After training the model, we evaluate its performance using a confusion matrix and calculate key metrics: Accuracy, Precision, Recall, and F1-score.

Objectives

- Apply a suitable ML algorithm to classify customer data.
- Construct a confusion matrix to evaluate model performance.
- Compute key evaluation metrics: Accuracy, Precision, Recall, and F1-score.
- Interpret the results and derive meaningful insights.

Theory: Confusion Matrix

A confusion matrix is a table used to describe the performance of a classification model on a set of test data for which the true values are known. It consists of four elements:

- **True Positives (TP):** Correctly predicted positive instances.
- **True Negatives (TN):** Correctly predicted negative instances.
- **False Positives (FP):** Incorrectly predicted positive instances (Type I error).
- **False Negatives (FN):** Incorrectly predicted negative instances (Type II error).

a) Accuracy: Accuracy measures the overall correctness of the model and is given by:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

b) Precision: Precision measures the correctness of positive predictions and is given by:

$$Precision = \frac{TP}{TP + FP}$$

c) Recall: Recall (or Sensitivity) measures the ability of the model to detect positive instances and is given by:

$$Recall = \frac{TP}{TP + FN}$$

d) F1-score: F1-score is the harmonic mean of Precision and Recall, balancing both metrics:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Conclusion

In this assignment, we successfully applied a machine learning model to classify customer data. Using the confusion matrix, we evaluated model performance by calculating Accuracy, Precision, Recall, and F1-score. These metrics help in assessing how well the model performs and indicate areas for improvement. The results provide valuable insights into customer classification and can guide further optimization of the ML model.