Project Report: Term Deposit Subscription Prediction

Overview:

- Goal: Predict customer subscription to a term deposit.
- Dataset: https://www.kaggle.com/datasets/janiobachmann/bank-marketing-dataset
- Workflow: cleaned, scaled and encoded dataset's numerical and categorical variables to train and evaluate models.

Key Findings:

- Class Imbalance: Must be handled to improve model performance to perform logistic regression.
- Logistic Regression: Doesn't fit that well according to the ROC & AOC score and f1 score as these came out to be pretty low.
- kNN: Fits very well to the dataset according to the ROC & AOC score and f1 score as these came out to be almost perfect. But it can be such that it's overfitted to the dataset
- Decision Trees: Fits very well to the dataset according to the ROC & AOC score and f1 score as these came out to be pretty high.
- Naïve Bayes: Doesn't fit that well according to the ROC & AOC score and f1 score as these came out to be pretty low.