

# **Bank Marketing Customer Purchase Prediction**

## **(INSIGHTS)**

Project predicting customer purchase behavior using the Bank Marketing dataset, with EDA and a Decision Tree model. Here are the key insights:

### **1) Dataset Overview:**

- a) Contains 45,211 records and 17 features, with both numeric and categorical data.
- b) No missing values in the dataset.

### **2) Exploratory Data Analysis (EDA):**

- a) Initial exploration included summary statistics, data type checks, and null value assessments.
- b) Class distribution analysis showed an imbalance in the target variable ("y").
- c) Correlation heatmap highlighted the relationships between features, guiding feature selection.

### **3) Data Preprocessing:**

- a) Categorical variables (e.g., job, marital, education) were encoded using Label Encoding.
- b) Features like "duration" and "campaign" displayed high variance, potentially impacting predictions.

### **4) Model Building:**

- a) Split data into training and test sets (70%-30% split).
- b) Trained a Decision Tree Classifier with a max depth of 5 to prevent overfitting.

### **5) Model Evaluation:**

- a) Accuracy: 89.5% on the test set.
- b) Confusion Matrix: Showed the model performed well in identifying "No" (did not purchase) cases but struggled with "Yes" (purchased) cases.
- c) Classification Report:
  - i) Precision: 92% for "No", 58% for "Yes".
  - ii) Recall: 96% for "No", 40% for "Yes".
  - iii) F1-Score: 94% for "No", 47% for "Yes".
- d) The imbalance in target classes impacted recall for the "Yes" class.

### **6) Decision Tree Visualization:**

- a) The decision tree plot showed key splits, with features like "duration" and "pdays" playing major roles in the decision-making process.

**7) Insights:**

- a) Imbalance in the dataset suggests the need for techniques like SMOTE or cost-sensitive learning.
- b) Duration was a strong predictor of purchase behavior.
- c) Model performance could improve with hyperparameter tuning or ensemble methods like Random Forest or Gradient Boosting.