**📊 Bank Marketing Prediction using Decision Tree (Machine Learning Project)**

**🔍 Project Overview**

This project focuses on building a Decision Tree Classifier to predict whether a customer will subscribe to a term deposit (response: yes or no) based on data from a bank marketing campaign. The model uses structured preprocessing through a pipeline and visualizes the decision-making process via a decision tree.

**🧠 Objective**

To predict the outcome (y) of a marketing call using relevant customer features, and visualize how the decision tree arrives at its conclusions.

**📁 Dataset Used**

* Source: UCI Bank Marketing Dataset
* File: bank-full.csv
* Rows: ~45,000
* Separator: ;
* Target variable: y (Yes/No)

**🛠️ Tools & Libraries**

* **Python** (pandas, matplotlib, sklearn)
* **scikit-learn**: DecisionTreeClassifier, train\_test\_split, Pipeline, ColumnTransformer, OneHotEncoder
* **Jupyter/Colab** for development

**🧹 Preprocessing Steps**

1. **Drop Columns**:
   * duration: Leaks information about the result
   * day, default, contact, job: Not informative or too complex
2. **Handle Categorical Columns**:
   * Used OneHotEncoder for object-type columns (e.g., education, housing, loan)
   * Retained numeric columns (e.g., age, balance, campaign)

**🧪 Modeling Steps**

* **Pipeline** created for streamlined preprocessing + model training
* **Train-test split**: 80% training, 20% testing
* **Classifier**: DecisionTreeClassifier (max\_depth=4)

**📈 Visualization**

* Extracted the trained tree from the pipeline
* Visualized using plot\_tree() with labeled, colored decision paths

**✅ Outcome**

* The model provides clear insight into what influences customer responses
* Suitable for marketing teams to make better targeting decisions

**🔗 Next Steps**

* Evaluate performance metrics: Accuracy, Precision, Recall
* Try other models (Random Forest, Logistic Regression)
* Export to GitHub / LinkedIn with explanation and visuals