Decision Tree – Bank Marketing Subscription Prediction

1. Objective

The objective of this task is to build a **Decision Tree classification model** to predict whether a customer will subscribe to a bank's term deposit based on customer and campaign-related features.

---

2. Dataset Description

- **Dataset Name:** Bank Marketing Dataset

- **Source:** UCI Machine Learning Repository

- **File Used:** bank.csv

- **Target Variable:** y

    - yes → customer subscribed

    - no → customer did not subscribe

The dataset contains customer demographic information, contact details, and previous campaign outcomes.

---

3. Tools & Technologies Used

- Python

- Pandas

- Scikit-learn

- Matplotlib

---

4. Data Preprocessing

- Loaded the dataset using Pandas

- Checked for missing values

- Encoded categorical variables using **Label Encoding**

- Separated features and target variable

- Split the dataset into **training (80%)** and **testing (20%)**

5. Model Building

- Algorithm Used: **Decision Tree Classifier**

- Criterion: **Gini Impurity**

- max_depth was limited to prevent overfitting

- Model trained on training dataset

---

6. Model Evaluation

- Evaluated model using:

  - Accuracy score

  - Classification report (Precision, Recall, F1-score)

- Compared training accuracy and testing accuracy to check overfitting

---

7. Decision Tree Visualization

- Visualized the trained Decision Tree using plot_tree()

- The tree shows feature splits and decision paths clearly

---

8. Key Decision Rules

1. If the call duration is high, the customer is more likely to subscribe.

2. Customers contacted via cellular with fewer campaigns show higher subscription probability.

3. Very short call duration usually results in no subscription.

---

9. Conclusion

The Decision Tree model successfully predicts customer subscription behavior.
Limiting tree depth helped reduce overfitting while maintaining good accuracy.
This task demonstrates the interpretability and effectiveness of Decision Trees in real-world classification problems.