Prodigy Infotech
Name: Shruti Sirsat
TASK 3 : Decision Tree Classifier

❖ Customer Purchase Prediction using Decision Tree

1. Dataset Description

The **Bank Marketing dataset** was used for this task. It contains information about customers such as age, job, marital status, education, loan details, campaign information, and previous contact history. The dataset also includes a target column y, which indicates whether the customer purchased the product (yes) or not (no). This dataset helps in predicting customer purchase behavior based on demographic and behavioral data.

2. Data Preprocessing

The dataset was first uploaded into Google Colab and read using the Pandas library. While reading the CSV file, the separator sep=";" was used because the dataset values were separated using semicolons. The dataset was checked for missing values, and no missing values were found.

The data was then divided into input features (X) and the target variable (y). The target column was converted from text values (yes and no) into numerical values (1 and 0). Since the feature columns contained categorical text data, one-hot encoding was applied to convert them into numerical format so that the machine learning model could process the data correctly.

3. Steps Performed

The following steps were performed to build the Decision Tree model:

• The dataset was uploaded and read correctly using Pandas.
• Missing values were checked to ensure data quality.
• Features (X) and target variable (y) were separated.
• The target column was converted into numerical form.
• Categorical feature columns were converted using one-hot encoding.
• The dataset was split into training and testing sets using an 80:20 ratio.
• A Decision Tree Classifier was created and trained using training data.
• Predictions were made on the test dataset.
• Model performance was evaluated using accuracy score.

4. Model Building and Evaluation

After preprocessing, the dataset was split into training and testing sets. The training data was used to train the **Decision Tree Classifier**, and the testing data was used to evaluate the model's performance. The accuracy score was calculated to measure how correctly the model predicted whether a customer would purchase the product or not. The model achieved good accuracy, showing effective prediction performance.

5. Conclusion

The Decision Tree model was successfully built to predict customer purchase behavior using demographic and behavioral data. This task demonstrated the complete process of building a machine learning classification model, including data preprocessing, model training, prediction, and evaluation. It provided a clear understanding of how decision tree classifiers can be applied to real-world datasets.

🔗 Code & Notebook Reference

The complete implementation of this task was performed using **Python in Google Colab**, and the notebook includes all steps from data preprocessing to model evaluation.

:-
https://colab.research.google.com/drive/1q01Zm9BgZtdqFyp4BIqdg2Xu3V6YG5wb?usp=sharing