Customer Purchase Prediction using Decision Tree

Using the Bank Marketing Dataset (UCI Repository)

Dhavalsinh Parmar [ProdigyInfotech]

Objective

- To build a Decision Tree Classifier that predicts whether a customer will purchase a term deposit based on demographic and behavioral data.

Dataset Overview

- - Source: UCI Machine Learning Repository
- - Records: 4,521 (bank.csv)
- - Target Variable: y (Subscribed: yes/no)
- - Features: Age, job, marital status, contact type, duration, campaign outcomes, etc.

Steps Followed

- 1. Data Loading and Exploration
- 2. Data Preprocessing (Label Encoding, One-Hot Encoding)
- 3. Train/Test Split
- 4. Decision Tree Model Training (max_depth=5)
- 5. Evaluation using Accuracy and Classification Report
- 6. Visualization of Tree and Feature Importance

Key Insights

- - Duration of last contact is the most influential predictor.
- - Previous campaign outcome and contact method are also significant.
- - Customers with longer call durations and prior success were more likely to subscribe.

Conclusion

- - Decision Tree offers interpretable and effective predictions.
- - Helps marketing teams better target potential customers.
- - Model accuracy achieved: ~88% on test data.

Next Steps

- - Explore ensemble models like Random Forest or XGBoost.
- - Perform hyperparameter tuning.
- - Integrate with CRM for real-time predictions.