Data Preprocessing

This project demonstrates essential data preprocessing techniques required before building any machine learning model.
Use the file: Data_Preprocessing.ipynb to explore and run the project.

Overview

The notebook performs the following steps:

1. Handling Missing Data: Missing values are filled with the mean of the respective column.
2. Encoding Categorical Data: Categorical data is transformed using one-hot encoding and label encoding.
3. Splitting Dataset: The dataset is split into training and testing sets.
4. Feature Scaling: Standard scaling is applied to normalize the feature values.

Dataset

The dataset (Data.csv) is a sample tabular dataset with the following columns:

- **Country**: Categorical feature (France, Germany, Spain)
- **Age**: Numerical feature with missing values
- **Salary**: Numerical feature with missing values
- **Purchased**: Categorical target variable (Yes, No)

Notebook: Data_Preprocessing.ipynb

This notebook demonstrates how to:

1. Load the dataset
2. Handle missing data
3. Encode categorical variables
4. Split the dataset into training and test sets
5. Scale features (Standardization or Normalization)