

## Data Preprocessing Pipeline Report

This report describes the construction of a data preprocessing pipeline using Python, Pandas, and Scikit-learn.

### Key Steps:

- Data Loading using Pandas
- Missing value imputation for numeric and categorical columns
- Standardization of numeric features
- One-Hot Encoding for categorical features
- Model training using RandomForestClassifier
- Evaluation with accuracy, precision, recall, F1-score, and confusion matrix

This pipeline is modular and reusable for various machine learning tasks.

### Model Performance:

Accuracy: 55%

F1-Score (Purchased): 0.61

Generated with simulated sample data.

Prepared by: Rangesh Pandian PT