

## **Project 1**

# Machine Learning Project 1 (Beginner to Intermediate Level)



### Project Title:

"Customer Purchase Prediction Using Classification Algorithms"

---



### Objective:

To build a basic classification model that predicts whether a customer will purchase a product or service based on their demographic and behavioral data.



### Dataset Example:

- Online Customer Dataset  
**Learn Fast, Understand Better.**
- Contains features like:
  - Age
  - Gender
  - Annual Income
  - Spending Score
  - Previous purchase behavior
  - Engagement metrics

You may also provide any similar dataset.

---

## Project Tasks / Workflow:

### 1. Data Loading & Understanding:

- Load the dataset using pandas.
  - Display first few rows, dataset shape, and column details.
  - Identify the target variable (Purchase / Not Purchase).
- 

### 2. Data Preprocessing:

- Handle missing values (imputation or removal).
- Encode categorical variables (Label Encoding / One-Hot Encoding).
- Scale numerical features using StandardScaler or MinMaxScaler.
- Split data into training and testing sets.

Learn Fast, Understand Better.

---

### 3. Exploratory Data Analysis (EDA):

- Plot distributions of key features.
  - Visualize correlations using heatmaps.
  - Identify customer behavior patterns (e.g., higher income correlates with purchases).
- 

### 4. Model Building:

Implement at least 2 classification algorithms:

- Logistic Regression
- Decision Tree
- Random Forest
- K-Nearest Neighbors (KNN)

Train each model and generate predictions.

---

## 5. Model Evaluation:

Compare models using:

- Accuracy
- Precision
- Recall
- F1-Score
- Confusion Matrix



Select the best-performing model.

---

## 6. Final Deliverables:

- Clean and documented notebook
  - Visualizations
  - Final model comparison summary
  - Explanation of project insights
-

## Learning Outcome:

By completing this project, the intern will:

- Understand how classification models work
- Learn how to preprocess and analyze real data
- Generate and interpret evaluation metrics
- Build confidence in ML model implementation

