**Iphone Purchase**

**Advance Project-8**

**Report**

**iPhone Purchase Prediction using Demographic Data**

1. **Objective**

To develop a machine learning model that predicts whether a customer will purchase an iPhone based on their gender, age, and salary. This helps retail companies target customers more effectively, reduce marketing costs, and increase conversion rates.

**2. Solution Overview**

* **Data Source:**  
  Customer demographic data (Gender, Age, Salary, Purchase decision)
* **Tech Stack:**  
  Python, Jupyter Notebook, Pandas, Seaborn, Scikit-learn
* **Key Activities:**
  + Data loading and cleaning
  + Exploratory Data Analysis (EDA) with insights
  + Preprocessing: encoding, scaling
  + Model training and validation
  + Performance evaluation
  + Inference testing

**3. Solution Architecture**

Customer Data (CSV/Excel)

↓

Data Cleaning & EDA (Seaborn, Pandas)

↓

Preprocessing (Encoding, Scaling)

↓

Model Training (Logistic Regression)

↓

Model Evaluation (Accuracy, F1-Score, Confusion Matrix)

↓

Inference on New Customer Profile

**4. Methodology**

* **Step 1:** Define problem and business objective
* **Step 2:** Perform Exploratory Data Analysis
* **Step 3:** Encode categorical features (Gender)
* **Step 4:** Scale numerical features (Age, Salary)
* **Step 5:** Train classification model
* **Step 6:** Evaluate using classification report and confusion matrix
* **Step 7:** Run inference on sample input
* **Step 8:** Export results and package model

**5. Time Taken**

| **Task** | **Time Taken** |
| --- | --- |
| Data Understanding + Cleaning | 1 hour |
| Exploratory Data Analysis (EDA) | 1.5 hours |
| Feature Engineering | 0.5 hour |
| Model Building & Training | 1 hour |
| Evaluation + Inference | 0.5 hour |
| Documentation | 1 hour |
| **Total** | **5.5 hours** |

**6. Challenges Faced**

* **Small feature set:** Only 3 input features limited the complexity of the model
* **Overfitting risk:** Limited data and binary features may lead to overfitting
* **Gender imbalance:** Purchase patterns may vary significantly by gender, requiring careful analysis

**7. Complexity**

| **Component** | **Complexity Level** |
| --- | --- |
| Data Size | Small |
| Feature Engineering | Low |
| Model Complexity | Low |
| Deployment Potential | High |

**8. Business Impact**

* **Higher Targeting Accuracy:** Focus marketing on the most likely buyers
* **Reduced Cost Per Acquisition:** Fewer irrelevant leads = lower spend
* **Increased ROI:** Smart segmentation and outreach increases conversion
* **Scalable Framework:** The process can be applied to other purchase predictions

**9. Template Reusability for Future Projects**

Use this document structure for any classification project  
Replace only:

* Title
* Feature/target variables
* Business context
* Model used
* Metrics and insights