**iPhone Purchase Prediction Project using KNN**

**Table of Contents**

1. Introduction
2. Objective
3. Solution Architecture
4. Methodology

* Data Collection
* Data Pre-processing
* Exploratory Data Analysis (EDA)
* Model Building
* Model Evaluation

1. Development Time
2. Challenges Faced
3. Complexity
4. Conclusion

**1. Introduction**

The iPhone Purchase Prediction Project aims to predict whether a customer will purchase an iPhone based on their gender, age, and salary. By leveraging machine learning techniques, businesses can enhance their marketing strategies, optimize inventory management, and gain deeper insights into customer behaviour.

## 2. Objective

The primary objective is to build a predictive model to determine if a customer will purchase an iPhone. The model should provide accurate predictions based on the provided demographic data.

## 3. Solution Architecture

The solution architecture for the project is divided into several key stages:

1. **Data Collection**: Gathering customer data from various sources.
2. **Data Pre-processing**: Cleaning and transforming the data for analysis.
3. **Exploratory Data Analysis (EDA)**: Understanding the data through visualization.
4. **Model Building**: Developing a machine learning model to predict iPhone purchases.
5. **Model Evaluation**: Assessing the performance of the model.

## 4. Methodology

### **Data Collection**

* Collected customer data including gender, age, and salary.
* The dataset contains 400 entries with no missing values.

### **Data Pre-processing**

* **Encoding Categorical Data**: Converted the 'Gender' column to numerical values using Label Encoding.
* **Standardizing Features**: Scaled the 'Age' and 'Salary' columns using StandardScaler.
* **Data Splitting**: Split the dataset into training (80%) and testing (20%) sets.

### **Exploratory Data Analysis (EDA)**

* **Age Distribution**: Analysed the distribution of customer ages.
* **Salary Distribution**: Examined the salary distribution.
* **Gender Distribution**: Visualized the gender distribution of customers.
* **Correlation Analysis**: Investigated the relationship between features and the target variable.

### **Model Building**

* **Model Selection**: Chose Logistic Regression for its simplicity and interpretability.
* **Training**: Trained the model on the training data.

### **Model Evaluation**

* **Accuracy**: Evaluated the model's accuracy on the test data.
* **Confusion Matrix**: Analysed the true positives, true negatives, false positives, and false negatives.
* **Classification Report**: Generated precision, recall, and F1-score for each class.

## 5. Development Time

The development of the iPhone Purchase Prediction model took approximately 2 weeks. The timeline is as follows:

* **Week 1**: Data collection, pre-processing, and exploratory data analysis.
* **Week 2**: Model building, evaluation, and documentation.

## 6. Challenges Faced

* **Data Imbalance**: The dataset had more males than females, which could potentially bias the model. Mitigation strategies included using balanced sampling techniques.
* **Feature Engineering**: Ensuring that the features selected were relevant and contributed positively to the model’s accuracy.
* **Model Selection**: Choosing the appropriate model that balances accuracy and interpretability.

## 7. Complexity

The complexity of this use case is moderate. While the data pre-processing and EDA steps are straightforward, selecting and fine-tuning the machine learning model requires a solid understanding of classification algorithms and performance metrics.

## 8. Conclusion

The iPhone Purchase Prediction Project successfully developed a predictive model with an accuracy of 87%. This model can significantly enhance business strategies by enabling targeted marketing, optimizing inventory management, and providing valuable customer insights. Continuous updates and improvements to the model will ensure its accuracy and relevance over time.