**Project Report: Gender-Age-Salary Analysis for iPhone Purchase Prediction**

**Introduction**

The goal of this project is to analyse a dataset containing information about individuals' gender, age, salary, and likelihood to purchase an iPhone (binary classification: 1 for purchase, 0 for no purchase). The project involves various steps, including data preprocessing, exploratory data analysis (EDA), outlier detection and imputation, and building a predictive model using Decision Tree classification.

**Methodology**

**1. Data Import and Overview:**

The initial step involved importing necessary Python packages and loading the dataset. The dataset consists of the following columns: Gender, Age, Salary, and Purchase iPhone.

**2. Data Preprocessing:**

**2.1. Label Encoding for Categorical Data**

The 'Gender' column, being categorical, was converted into numerical values using label encoding to facilitate machine learning model training.

**2.2. Exploratory Data Analysis (EDA)**

EDA was performed to understand the relationships between different variables. Heatmap visualization was used for correlation analysis, and box plots were employed to detect outliers.

**3. Outlier Detection and Imputation:**

Outliers were detected using box plots, and a decision was made to impute these outliers with upper and lower threshold values to maintain data integrity.

**4. Model Planning and Training:**

**4.1. Decision Tree Classifier**

The decision was made to use the Decision Tree classification model for predicting iPhone purchases. 'Purchase iPhone' was considered the dependent variable, and the other columns were treated as independent variables.

**4.2. Model Fitting and Evaluation**

The dataset was split into training and testing sets (80:20 split), and the Decision Tree classifier was trained on the training set. Model accuracy was initially observed at 91.25%. To enhance accuracy, the entire process was iterated through a loop with 1000 different random state values, resulting in an increased accuracy of 98.75%.

**5. Model Evaluation:**

**5.1. Confusion Matrix**

A confusion matrix was applied to assess the performance of the model. Out of 80 test values, the model correctly predicted 79, demonstrating the effectiveness of the Decision Tree classifier in this context.

**Conclusion:**

The project successfully explored and analyzed a dataset related to gender, age, salary, and iPhone purchase behavior. By employing label encoding, outlier detection, and the Decision Tree classification model, the analysis yielded valuable insights into predicting iPhone purchases. The final model demonstrated a high accuracy rate, indicating its reliability in predicting purchase behavior based on given parameters.