Project Report: Smartphone Price Prediction

Submitted by: Isha Yadav, 21i190002 (MSc-PhD IEOR)

Introduction:

The Smartphone Price Prediction project aimed to develop a robust predictive model to estimate the price range of smartphones based on their features. The proliferation of smartphones and the varying features that influence their prices have driven the need for accurate price estimation. This report outlines the project's objectives, methodology, findings, and conclusions.

Motivation and Objectives:

The rapid evolution of the smartphone market necessitates accurate pricing to facilitate informed purchase decisions. The primary objective of this project was to leverage various classification techniques, fine-tune models, and identify the best model in terms of prediction accuracy and computational efficiency. The project aimed to provide a tool for both consumers and sellers to approximate the value of smartphones.

Data and Methodology:

The dataset used for this project was obtained from Kaggle and comprised 2000 samples with 20 features. The features ranged from battery capacity to camera specifications. The dataset provided four price range classes: Very Low, Low, High, and Very High. The project followed a structured approach:

- *Data Preprocessing*: The data underwent preprocessing steps such as removing duplicates, handling missing values, and encoding categorical features.

- *Exploratory Data Analysis (EDA)*: Pandas and Seaborn were used for EDA, revealing the distribution of features, correlation among variables, and potential relationships.

- *Model Selection and Implementation*: Various classification models were employed, including Logistic Regression, Decision Trees, Random Forest, K-Nearest Neighbors (KNN), and a Fully Connected Neural Network (FCNN). Each model was trained and evaluated for accuracy and efficiency.

Model Evaluation and Results:

- *Logistic Regression:* The Logistic Regression model achieved an accuracy of 68.83%. Its performance was visualized through accuracy vs. epochs and a confusion matrix.

- *Decision Trees:* The Decision Trees model yielded an accuracy of 81%. A classification report and confusion matrix illustrated its performance.

- *Random Forest:* Employing 20 trees, the Random Forest model reached an accuracy of 88%. Its ensemble nature demonstrated favorable performance.

- *K-Nearest Neighbors (KNN):* Selecting an optimal K value of 9, the KNN model achieved an accuracy of 92.25%. The project navigated overfitting concerns.

- *Fully Connected Neural Network (FCNN):* The FCNN, a more complex model, achieved the highest accuracy of 95%. By tuning hyperparameters, including hidden layers and epochs, computational efficiency was balanced.

Conclusion:

The Smartphone Price Prediction project emphasized meticulous dataset analysis, model selection, and optimization to achieve accurate price range estimation. Through the exploration of various classification techniques, it became evident that the FCNN model was best suited for this task, yielding a 95% accuracy rate. This project not only fulfilled its objectives but also highlighted the significance of robust prediction models in addressing practical challenges in the technology sector.

In conclusion, this project demonstrated the viability of machine learning techniques in predicting smartphone prices based on features. The outcomes have implications for both consumers seeking to make informed purchasing decisions and businesses aiming to optimize pricing strategies.