Capstone Project Submission

Name:- Tarun
Email:- jangratarun1020@gmail.com
Contribution:- Individual

Github Link:- https://github.com/tarun422/Mobile-Price-Range-Prediction

The Mobile Price Range Prediction dataset consist of 2000 observation with 21 features and it is a classification problem because its target column is contain categorical values and the main goal is to predict the category range of the mobile using dataset

After loading the dataset, first performed data preprocessing and checking data types, missing values, duplicate values and data description. In this dataset there are neither null values nor duplicate values. After that changed the date type to int16 which was int64 & float 64 . Their are some inappropriate values comes in which is imputed by knn imputer

After that checking class imbalance and lucky found their are all classes are symmetrically balanced

After that Exploratory Data Analysis is performed to obtain the insights of our dependent Price Range. Various graphs are constructed to comparing the Price Range column with other columns. It contain Univariat Analysis, Bivariate Analysis and Multivariate Analysis . RAM(Random Access memory) play an important role in price range after that battery power comes in for price observation

After that feature engineering comes in to remove less important features. For scaling independent features, StandardScaling is used.

After that modeling part begins and 7 Classification algorithms is used which are Logistic Regression , Decision Tree Classifier, Random Forest Classifiers , Gradient Boosting Classifier , XGBoost Classifier K-Neighbors Classifier and Support Vector Classifier . After fitting the models and evaluating metrics (Classification Report , Accuracy) and also hyperparameter tuning. In this observation Logistic Regression gives accuracy is 97.56 on training data and 96.67 on test data. So Logistic Regression model is the best for predicting the Mobile Price Range Prediction