Mobile Price Range Prediction
Project Summary

The mobile phone industry is highly competitive, and the price of a mobile phone is determined by various factors, including battery power, Bluetooth, camera quality, screen size, and more. In this context, a study was conducted to understand the factors influencing the price range of mobile phones. The study used a dataset containing around 21 variables to predict the price range of mobile phones, categorized as low, medium, high, and very high.

The first step in the analysis involved data wrangling, where missing values were handled and unique values were checked. The study identified that 180 phones had pixel resolution height as 0 and two phones had screen width in cm as 0. The minimum value of px_height and sc_w should not be 0, as it does not make sense for a phone screen width or pixel height to be 0. Therefore, the study replaced these 0 values with the mean values, ensuring that no missing values were left in the dataset.

Next, the study performed exploratory data analysis (EDA), which revealed that all category phones were distributed with equal price range. The analysis also indicated that battery capacity was positively correlated with the price range of mobile phones, and the distribution of battery capacity gradually increased with the price range. This suggested that consumers may be willing to pay more for a mobile phone with a higher battery capacity. The study found that almost half the devices had Bluetooth, and half did not.

The scatter plot showed a clear positive correlation between RAM and price range, with the majority of the data points clustering towards the upper right corner. This indicated that as the price range increased, the amount of RAM in the device generally increased as well. The study also found that the count of devices with dual sim was increasing for the very high price range. Additionally, the distribution of primary camera megapixels across different target categories was relatively consistent, indicating that this feature may not significantly influence the price range of mobile phones.

The analysis of the screen size distribution among different target categories indicated that there was not a significant difference in the distribution, suggesting that screen size may not be the sole driving factor in determining the target categories. However, this uniformity in distribution can be advantageous for predictive modeling, as it implies that screen size may not be a significant variable in differentiating between different target categories, allowing other features to play a more crucial role in determining the target categories. The study also found that mobile phones with higher price ranges tended to be lighter in weight compared to lower price range phones.

After the EDA, the study performed hypothesis testing on three statements and handled outliers. The study identified that RAM, battery power, and pixel quality were the most significant factors affecting the price range of mobile phones. The study then performed feature engineering and implemented machine learning models such as logistic regression, random forest, and XGBoost. Based on the experiments, the study concluded that logistic regression and XGBoost algorithms with

hyperparameter tuning yielded the best results in predicting the price range of mobile phones.

In conclusion, the study found that the mobile phones in the dataset were divided into four different price ranges, each having a similar number of elements. Additionally, the study found that approximately half of the devices had Bluetooth, while the other half did not. Furthermore, the study found that as the price range increased, there was a gradual increase in battery power, and RAM showed continuous growth from low-cost to very high-cost phones. Moreover, the study found that the costly phones tend to be lighter than the lower-priced ones.

The study identified that RAM, battery power, and pixel quality were the most significant factors affecting the price range of mobile phones. Finally, the study found that logistic regression and XGBoost algorithms, coupled with hyperparameter tuning, provided the best performance in predicting the price range of mobile phones.

Contributors Roles:

Sandeep Salunke

1. Data Loading:
2. Data handling
3. Handling missing values
4. Data exploration/Visualization
5. Outliers detection
6. Hypothesis Testing
7. Feature engineering
8. Model deployment

Github Repo link

https://github.com/Sandeep81299/Mobile-Price-Range-Prediction-

to each other. That means more the people more will be adr. is_repeated guest and previous bookings not canceled has strong correlation. may be repeated guests are not more likely to cancel their bookings.