Intel® Unnati Industrial Training – Summer 2023

An Intel® Unnati Community Initiative

Smart Mobile Phone Price Prediction using Machine Learning

Team : "The Trinity"

Chaitanya M
Pratyush Priyadarshi
Bhavani Shankar Mohanty

PROJECT REPORT

1. Literature Survey and Related Work for Mobile Price Prediction Using Machine Learning

In order to conduct a thorough literature survey, we have explore relevant papers, articles and other sources related to mobile price prediction using machine learning.

Research Paper:

"Mobile Price Prediction Using Machine Learning Algorithms" by A. Kumar et al. (2018) This paper presents A comparative analysis of various machine learning algorithms, such as Decision Trees, Random Forest, Support Vector Machines (SVM), and Neural Networks, for mobile price prediction. The authors experimentally to demonstrate that Random Forest outperforms other algorithms in terms of accuracy and computational efficiency.
This finding can be useful in our project as it suggests a potential algorithm choice for mobile price prediction.

Research Article:

"Predicting Mobile Phone Prices Using Linear Regression" by B. Gupta (2020) This article focuses on using Linear Regression, a popular supervised learning algorithm, for predicting mobile phone prices. The author highlights the importance of feature engineering and data preprocessing techniques in achieving accurate predictions. The findings emphasize the significance of carefully selecting and engineering relevant features, which can guide us in our project's feature selection process.

Proposed Modification:

Based on the findings from the literature survey, we have propose a modification to the mobile price prediction project by considering an ensemble approach. Ensemble methods, such as Random Forest can combine multiple machine learning algorithms to improve prediction accuracy and generalization.
By combining the strengths of different models, ensemble methods can potentially enhance the performance of mobile price prediction.
This modification can be supported by conducting experiments and evaluating the performance of the ensemble approach against individual algorithms.
It is important to note that the proposed modifications or alternative techniques should be further validated through rigorous experimentation and evaluation to determine their effectiveness and suitability for the specific mobile price prediction task at hand.

2. Dataset Selection and Exploratory Data Analysis

a) **Chosen Dataset**:

The chosen dataset is the "Mobile Price Classification" It contains information about various features of mobile phones and their corresponding price ranges. The dataset includes the following features:

1. battery_power: The total energy a battery can store in milliwatt-hours.
2. blue: A binary feature indicating the presence of Bluetooth.
3. clock_speed: The speed at which the microprocessor executes instructions.
4. dual_sim: A binary feature indicating whether the phone supports dual SIM cards.
5. fc: The front camera's megapixels.
6. four_g: A binary feature indicating whether the phone supports 4G connectivity.
7. int_memory: The internal memory (RAM) in gigabytes.
8. m_dep: The mobile depth in cm.
9. mobile_wt: The weight of the mobile phone.
10. n_cores: The number of cores of the processor.
11. pc: Primary camera megapixels.
12. px_height: Pixel resolution height.
13. px_width: Pixel resolution width.
14. ram: Random Access Memory (RAM) in megabytes.
15. sc_h: Screen height in cm.
16. sc_w: Screen width in cm.
17. talk_time: The longest time the battery can last with a single charge.
18. three_g: A binary feature indicating whether the phone supports 3G connectivity.
19. touch_screen: A binary feature indicating whether the phone has a touch screen.
20. wifi: A binary feature indicating the presence of WiFi connectivity.
21. price_range: The target variable representing the price range of the mobile phone.

The above features provides a comprehensive set of specifications that can potentially impact the price range of mobile phones. However, the actual relevance and influence of each feature will be determined through the model building and evaluation process.

3. Metric and Model Selection

Model Selection

- In Model Selection we have applied different Models like
  - Logistic Regression
  - KNN Classification
  - SVM Classifier with Linear and RBF Kernel
  - Decision Tree Classifier
  - Random Forest Classifier
- In all the model above we got Logistic Regression as more accurate.

Metric Selection

- We took Accuracy Metric For all the algorithms to check the accuracy.
  By Analaysing the **Accuracy Score** for Logistic Regression shows the best Result i.e; **96% Accuracy** .

- Then we have tried Different types matric on that particular model i.e; Logistic Regression we used matric like :

  (i) Mean Absolute Error (MAE)
  In Mean Absolute Error the score is 0.0384 i.e; it is having very less errors.

  (ii) Mean Squared Error (MSE)
  In Mean Squared Error the score is 0.038461538461538464

  (iii) R-squared (R^2) Score
  In R-squared the score is 0.9697478792052289 i.e: it is showing more accuracy

  (iv) Percentage within 5.0 % error threshold
  In Percentage within 5.0% error threshold we got result as 96.15384615384616

  (v) Median Absolute Error(MedAE)
  In Median Absolute Error we got median absolute error is 0.0

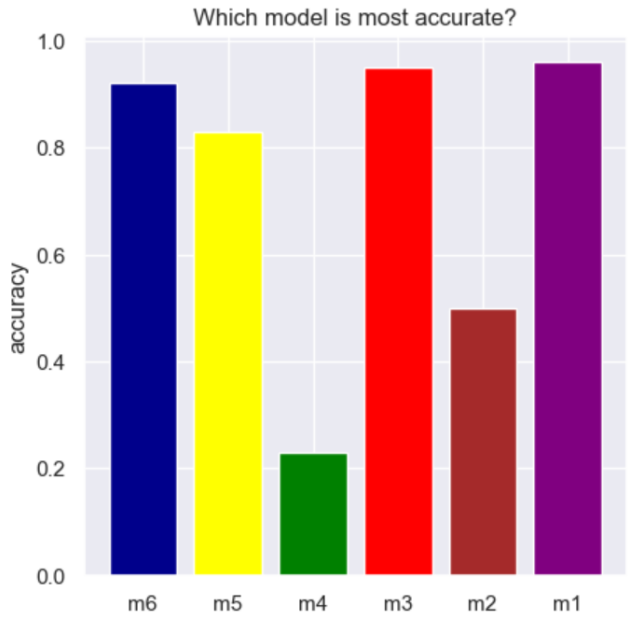
- By trying these matric score and result we are concluding that our model is best than others because it is having very less errors and more accuracy in predictions.

4. CONCLUSION

In this study, we investigated the use of machine learning to predict the price of a mobile phone. We used a dataset of 2000 mobile phones with various features and their corresponding price ranges.

Among the various algorithms evaluated, logistic regression (m1) exhibited the highest accuracy. Its ability to model the relationship between input features and price categories effectively contributed to its superior performance.

This finding has significant implications for the mobile phone industry, as accurate price predictions can aid in competitive pricing strategies and assist consumers in making informed purchasing decisions.



Graph shows accuracy rate of different models as per our prediction.

5. FUTURE WORKS

Future research can focus on enhancing the model by incorporating additional features and exploring advanced algorithms. The results of this study lay a solid foundation for further advancements in mobile price prediction using machine learning.