MOBILE PRICE PREDICTION USING FEATURE SELECTION AND CLASSIFIER ALGORITHMS OF MACHINE LEARNING

Sireesha Mitta*[1], Dr. U. Arul*[2]

*[1]UG Student, Department Of Computer Science And Engineering, Madanapalle Institute Of Technology & Science, Angallu (V), Madanapalle-517325, Chittoor District, Andhra Pradesh, India.

*[2]Assistant Professor, Department Of Computer Science And Engineering, Madanapalle Institute Of Technology & Science, Angallu (V), Madanapalle-517325, Chittoor District, Andhra Pradesh, India.

ABSTRACT

The main aim of this research is to predict the mobile with available features are either economical or else expensive. We used various feature selection algorithms to identify and remove less important redundant features with minimum computational complexity. In addition our research used various classifiers to achieve high accuracy as much as possible. The obtained results by the experiments are compared to the existing systems and the comparison results proved that our proposed system achieves highest accuracy with minimum features selection. The conclusion is done based on the best feature selection algorithm and the best classifier for the given dataset. We can use this work in any type of marketing and business to get an idea while buying the products with minimum price and maximum features.

Keywords: Machine Learning, Prediction, Decision Tree, Naïve Bayes, Neural Networks, Linear Regression Algorithm, KNN.

I.   INTRODUCTION

The cost of a product is an important attribute in marketing as well as business. Every customer is primarily asking for the price of the item that they want to purchase. All the customers are first worried about whatever they purchase is worthy according to price and quality [1,2]. So the basic objective of every customer while buying the product is to estimate the price at home before they purchase. This paper moves the customer towards the first step of the above mentioned destination. The advancements of Artificial Intelligence in recent years has given an ability to computer science and automation industries which provided with robotic controlled tasks in most efficiently than humans however they require human intelligence and discernment. It makes machines capable to answer questions intelligently as well as technically in a very vast developing engineering field. Machine learning contains the best techniques and methods for artificial intelligence, for example, classification methods, regression techniques, supervised and unsupervised learning [3,4]. We can able to write machine learning algorithms by using different tools like python, MATLAB, WEKA, etc. In machine learning we found so many methodologies and classifiers, such as Decision tree, Naïve Bayes, and many more. Various types of feature selection algorithms are present in machine learning therefore we can select any of them to select the best features in order to minimize the dataset. It may reduce the computational complexity of the problem. Apart from it we can use optimization techniques to decline the dimensionality of the dataset. Mobile is the most important device nowadays almost all people maintain mobile, it is very difficult to imagine our lives without mobile. Since, it is the most selling and purchasing device in the market, every day new mobiles with the new version and more based on features are launched. So, in this paper we worked on the prediction of the mobile price and we can do a similar type of prediction on different products like cars, bikes, laptops, etc. In the analysis of mobile price prediction, our target variant is mobile price, since we are in the 21st century, while purchasing the mobile we have to consider the processor of the mobile, battery timing [5,6]. The size and thickness of the mobile are also important decision factors. Internal memory, Camera pixels, and video quality must be under consideration. The above mentioned are the main features while deciding the price of a mobile phone. Therefore, we have used many of the mentioned features to classify whether the mobile would be very economical and expensive.

II.   LITERATURE SURVEY

Using previous historical data we can predict the price of items by using some machine learning algorithms, best example for this is Sameer Chand-Pudaruth, the researchers worked to estimate the prices of second cars using algorithms such as multiple linear regression and k-nearest neighbors (KNN) [7,8]. Support Vector

Machine (SVM) concept is used the researcher Mariana Listiani[3,9] for the same work. The author predicted prices of leased cars using SVM techniques. It was found in this research that the SVM technique is far more better and accurate for price prediction as compared to other like multiple linear regression when a very large data set is available.

III.   METHODOLOGY

3.1 Data Collection

In this paper, we collected data from different websites which is related to mobiles like Kaggle, etc.

Features:

In order to predict the price of mobile we considered some feature for estimation such as battery power, bluetooth, width and height of mobile, browsing capacity, processor, primary camera mega pixel, touch screen, Wi-Fi, pixel resolution height, and pixel resolution width.

Use:

This type of estimation will help companies to predict the price of mobiles to give tough competition to another mobile manufacturer. It is also helpful for the customers who are investing money to buy mobile at the best price with the best features.
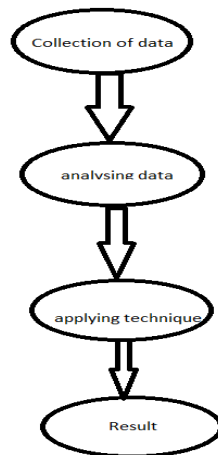

Fig 1: methodology of project

In order to achieve this project result we followed a procedure , first we collected the data from different websites and after the collection of data we analyse the patterns present in the data and apply the machine learning algorithms for classification and prediction purpose and finally we got the result with maximum accuracy.

IV.   EXPLORATORY DATA ANALYSIS

In this section, we compared the price range versus mobile features by plotting different types of graphs.

```
sns.jointplot(x='ram',y='price_range',data=dataset,color='red',kind='kde');
```
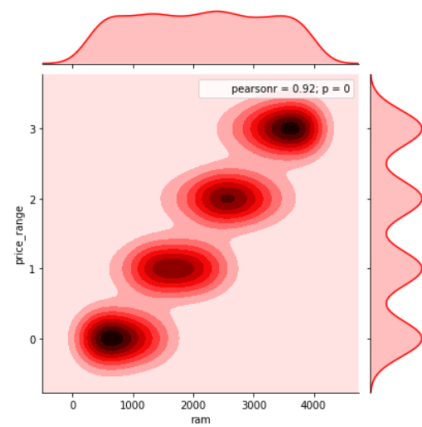
This is the graph which we got while comparing the ram of mobile with the price range, if the ram capacity increases then price_range is also increasing.

```
sns.pointplot(y="int_memory", x="price_range", data=dataset)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7ff5cea69c18>
```
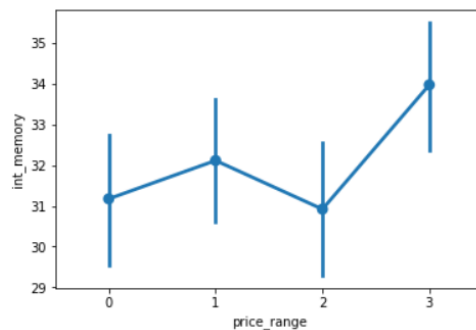

This is the graph which we plot to compare the internal memory capacity with the price_range.

4.1 Training the Model

Here we developed the training model by using some techniques such as decision trees, KNN (K-nearest neighbors) and logistic regression techniques.

```
from sklearn.tree import DecisionTreeClassifier
dt=DecisionTreeClassifier()
```

```
dt.fit(X_std,Y)
```

```
dt.predict(data_test_std)
```

```
data_test
```

KNN:

Here, by using KNN model we trained the data set by importing the modules like from sklearn and Kneighbors classifiers.

```
from sklearn.neighbors import KNeighborsClassifier
knn=KNeighborsClassifier()
```

```
knn.fit(X_std,Y)
```

```
KNeighborsClassifier()
```

```
knn.predict(data_test_std)
```

Logistic Regression:

Here, we used the technique called logistic regression to train the model.

```
from sklearn.linear_model import LogisticRegression
lr=LogisticRegression()

lr.fit(X_std,Y)

LogisticRegression()

lr.predict(data_test_std)
```

Accuracy Prediction:

Accuracy is also predicted which are helpful to find definite accuracy in order to take decision for purchasing the right mobile using techniques such as decision tree, KNN, Logistic regression.

By using decision tree techniques we got the accuracy on the data set which we had considered for the final analysis.

```
dt_ac=accuracy_score(Y_test,Y_pred)

dt_ac

0.845
```

By using KNN technique we got the accuracy but the KNN model produced less accuracy as depicted below.

```
knn_ac=accuracy_score(Y_test,Y_pred)

knn_ac

0.5225
```

By using logistic regression we got the highest accuracy of any other AI techniques.

```
: lr_ac=accuracy_score(Y_test,Y_pred)

: lr_ac

: 0.96
```

V. CONCLUSION

The conclusion of this work is the comparative result of feature selection and classifier algorithms. The main aim of this work is to get maximum accuracy by considering minimum features but the most appropriate feature. It is important to note that forwarding selection by considering irrelevant and redundant features to the dataset definitely decreases the efficiency of both classifiers, while in the case of backward selection if we

remove any important feature from the dataset, its efficiency decreases. One more important conclusion is derived from our analysis is while working conversion of a regression problem into a classification problem that introduces more error.

VI.   REFERENCES

[1]   Sameer Chand Pudaruth. "Predicting the Price of Used Cars using Machine Learning Techniques", International Journal of Information & Computation Technology. ISSN 0974-2239 Volume 4, Number 7 (2014), pp. 753764.

[2]   Shonda Kuiper, "Introduction to Multiple Regression: How Much Is Your Car Worth? " , Journal of Statistics Education · November 2008.

[3]   Mariana Listiani, 2009. "Support Vector Regression Analysis for Price Prediction in a Car Leasing Application". Master Thesis. Hamburg University of Technology.

[4]   U. Arul & S. Prakash, 'Toward Automatic Web Service Composition based on Multilevel Workflow Orchestration and Semantic Web Service Discovery', International Journal of Business Information Systems, Inderscience Publishers, Vol. 34, Issue 1, pp. 128–156, April 2020.

[5]   https://www.kaggle.com/vikramb/mobileprice-prediction

[6]   file:///C:/Users/Lenovo/Downloads/asim2018-ijca-9165551%20(1).pdfMuhammad   Asim   UET Lahore Pakistan, Zafar Khan UET Lahore Pakistan.

[7]   U. Arul & Dr. S. Prakash, 'Towards Fault Handling in B2B Collaboration using Orchestration based Web Services Composition', International Journal of Emerging Technology and Advanced Engineering (IJETAE), Vol. 3, Issue 1, pp. 388-394, 2013.

[8]   https://github.com/vikram-bhati, Classification classify mobile price range.

[9]   Introduction   to   dimensionality   reduction,   A   computer   science   portal   for   Geeks. https://www.geeksforgeeks.org/dimensionalityreduction.