Laptop Price Prediction in Machine Learning using Random Forest Classifier Technique

Katreddi.SAI SRINIVAS [1]
Kadali.TARUN SAI [2]
Kondaveeti.MOHAN SATYA SRIRAM [3]
P.Srinu Vasa Rao[4] ,
Swarnandhra college of Engineering and Technology.

Abstract:- Laptops are now became a most essential and mostly used gadgets around the world . It is widely used by the students, Employees and other working professionals. Considering the present technical advancements, the electronic gadgets like mobiles and laptops are getting costly and many more companies and models are evolving day by day. So it became difficult for the customers and also the sellers in the aspect of pricing.

To overcome this type of difficulties, we are building a model to predict the price of laptops based on the specifications they are made off. The price of the laptop is predicted by taking some input values from the user which were the specifications of the laptop such as Company, Model name, Category, screen type, screen resolution, CPU, RAM, Storage, GPU, Operating System, Operating system Version and Weight.

This laptop price prediction model is developed by using machine learning techniques. Leveraging historical pricing data, features and market trends are incorporated to train the model. The goal is to create an accurate system capable of forecasting laptop prices and helping the customers in making informed purchase decisions. It also aims to investigate the impact of external factors like technical advancements, economic conditions and consumer preferences on laptop prices.

To build this model, various machine learning algorithms such as linear regression, decision trees, and ensemble methods will be explored to identify the most suitable model for predicting laptop prices. Time-series analysis may be incorporated to capture temporal patterns in pricing fluctuations. The anticipated outcome is a well-tuned machine learning model capable of accurately predicting the laptop prices, offering a valuable tool for both the customers and retailers in understanding the pricing dynamics of the laptop market.

This model helps the customers to find the price of the laptop they are looking for according to their required specifications and also to the retailers to fix the price of the laptop according to the market trends and other aspects.

I. INTRODUCTION

Laptop price prediction is a challenging task that requires expert knowledge. The price of a laptop depends on various factors such as brand, model, RAM, ROM, GPU, CPU, etc. In India, the demand for laptops increased significantly highest in five years. This paper suggests that a laptop price prediction system was developed using the supervised machine learning techniques like Random Forest, Decision Tree. The system achieved 84% to 88% prediction precision by using factors like Laptop's model, RAM, ROM (HDD/SSD), GPU, CPU, IPS Display, and Touch Screen. However, the accuracy of the prediction model depends on the quality of the data used to train the model.

It is important to note that the price of laptops is subject to change due to various factors such as supply and demand, inflation, and currency exchange rates. Therefore, it is difficult to predict the exact price of a laptop in the future. However, you can use the current market trends and expert opinions to make an informed decision when purchasing a laptop.

In conclusion, predicting the price of a laptop is a complex task that involves many factors. This Analysis suggests that a laptop price prediction system was developed using the supervised machine learning technique. The system achieved 84% to 88% prediction precision by using factors like Laptop's model, RAM, ROM (HDD/SSD), GPU, CPU, IPS Display, and Touch Screen . However, the accuracy of the prediction model depends on the quality of the data used to train the model. Therefore, it is important to use the current market trends and expert opinions to make an informed decision when purchasing a laptop. By doing so, you can ensure that you get the best value for your money and avoid overpaying for a laptop that does not meet your needs.

II. LITERATURE SURVEY

[1] Applying rigorous analysis to aid investment decisions is gaining momentum in the United States and the United Kingdom, say Aminah Md Yusof and Ismail Syuhaida. But In Malaysia, the response from local

academics was very slow, and for doctors even slower. This article explains how multiple regression analysis (MRA) and its extension hedonic regression analysis can be used to explain price changes of selected houses in Malaysia. Theoretically, every attribute defined as a value judgment is a value, and the harmony of each attribute is specified. This article shows that statistical analysis can analyze real estate investment by taking into account various decisions. More rigorous consideration of various factors can lead to better investment decisions.

[2] It is necessary to increase the quality of education. Data mining technology plays an important role in data analysis. A decision tree algorithm related to data mining technology was used in this study to create a classification model that can predict student performance, especially in the field of engineering. Many factors can affect a student's performance. Here, some important factors are taken into consideration while creating a decision tree to classify students according to their characteristics (grades). This article compared four decision tree algorithms J48, NBtree, Reptree and Simple cart and found that the J48 decision tree algorithm is the most suitable algorithm for the design model. Cross-validation and percentage comparison methods are used to evaluate the performance of different algorithms. The traditional KDD method is used as the method. WEKA (Waikato Environment for Scientific Information) tools for analysis and prediction.

[3] Machine learning (ML) can provide great help in making decisions and predictions based on the amount of facts generated. We also looked at machine learning techniques used in recent developments in the Internet of Things (IoT). Many studies have provided insight into the use of machine learning techniques to predict laptop-only prices. As demonstrated in this article, we present a new method that improves the prediction of desktop usage to better understand the context of avoiding analysis. process. The truth about laptop prices. This prediction model is used

by a unique combination of features and learning models from various popular computer programs. We use special computer programs to learn about models such as decision trees, multilayer networks, KNN, and random forests to test which design model is more accurate in predicting laptop speed.

[4] Mehtab, S., Sen, J. said, Predicting future trends of stock prices has always been a difficult task for researchers. While proponents of efficient market theory (EMH) argue that it is not possible to develop forecasting systems that can accurately predict market prices, pioneering studies in data have shown that the timing of stock prices can be predicted with accuracy. In this article, we propose a very powerful and accurate stock price prediction framework that includes a combination of statistical, machine learning and deep learning models. We use daily price data collected at five-minute intervals from a well-known company listed on the National Stock Exchange (NSE) of India. Granular data is collected in three periods throughout the day, and the aggregated data is used to develop a forecast model. Detailed results on the performance of this model are presented.

III. METHODOLOGY

This study aims to create a regression algorithm model using the Google Collaboratory tool and Jupyter Notebook. The steps of the modeling process are shown in Figure . The first step is to obtain the datasets from the Kaggle website; Then, the data is pre-processed, which involves Data Cleaning and Feature Engineering; Next, the Exploratory Data Analysis (EDA) is performed; To apply machine learning with the Decision Tree and Random Forest algorithms for some sample data is required. The table below shows some data about different laptops and their prices based on their specifications. The data is sourced from Kaggle.com

	Company	TypeName	Inches	ScreenResolution	Cpu	Ram	Memory	Gpu	OpSys	Weight	Price
0	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 2.3GHz	8GB	128GB SSD	Intel Iris Plus Graphics 640	macOS	1.37kg	71378.6832
1	Apple	Ultrabook	13.3	1440x900	Intel Core i5 1.8GHz	8GB	128GB Flash Storage	Intel HD Graphics 6000	macOS	1.34kg	47895.5232
2	HP	Notebook	15.6	Full HD 1920x1080	Intel Core i5 7200U 2.5GHz	8GB	256GB SSD	Intel HD Graphics 620	No OS	1.86kg	30636.0000
3	Apple	Ultrabook	15.4	IPS Panel Retina Display 2880x1800	Intel Core i7 2.7GHz	16GB	512GB SSD	AMD Radeon Pro 455	macOS	1.83kg	135195.3360
4	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 3.1GHz	8GB	256GB SSD	Intel Iris Plus Graphics 650	macOS	1.37kg	96095.8080
...	...	...	...	...	...	...	...	...	...	...	...
1298	Lenovo	2 in 1 Convertible	14.0	IPS Panel Full HD / Touchscreen 1920x1080	Intel Core i7 6500U 2.5GHz	4GB	128GB SSD	Intel HD Graphics 520	Windows 10	1.8kg	33992.6400
1299	Lenovo	2 in 1 Convertible	13.3	IPS Panel Quad HD+ / Touchscreen 3200x1800	Intel Core i7 6500U 2.5GHz	16GB	512GB SSD	Intel HD Graphics 520	Windows 10	1.3kg	79866.7200
1300	Lenovo	Notebook	14.0	1366x768	Intel Celeron Dual Core N3050 1.6GHz	2GB	64GB Flash Storage	Intel HD Graphics	Windows 10	1.5kg	12201.1200
1301	HP	Notebook	15.6	1366x768	Intel Core i7 6500U 2.5GHz	6GB	1TB HDD	AMD Radeon R5 M330	Windows 10	2.19kg	40705.9200
1302	Asus	Notebook	15.6	1366x768	Intel Celeron Dual Core N3050 1.6GHz	4GB	500GB HDD	Intel HD Graphics	Windows 10	2.2kg	19660.3200

1303 rows × 11 columns

Fig1. Dataset

This data is cleaned and explorated using machine learning techniques and that mostly suitable for Decision tree and Random forest algorithms.
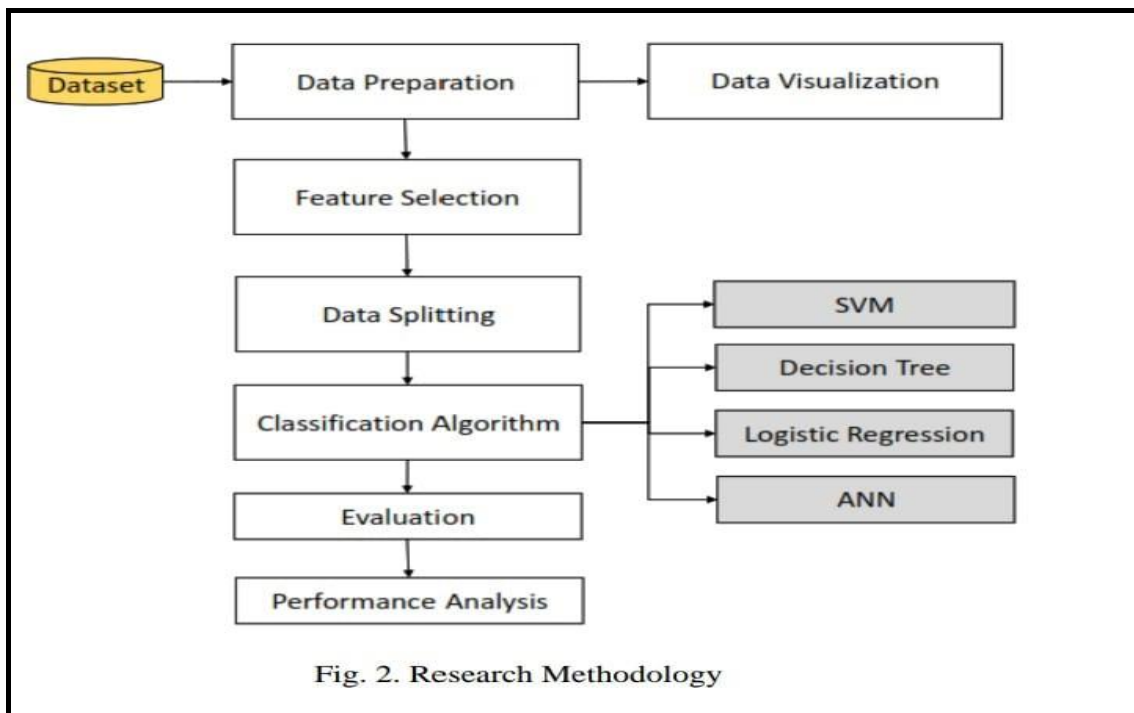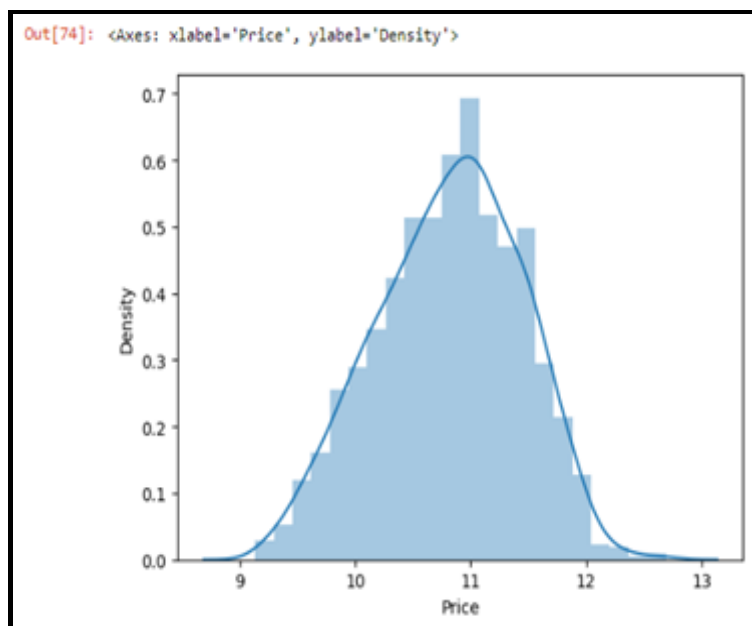

Fig2. The representation of data flow.

Fig3. Density Plot regarding Price for fig1 dataset

Random Forest algorithm can be applied on the the dataset for the classification(To get the optimal price from the dataset). By using the random forest algorithm works 88% accurately.
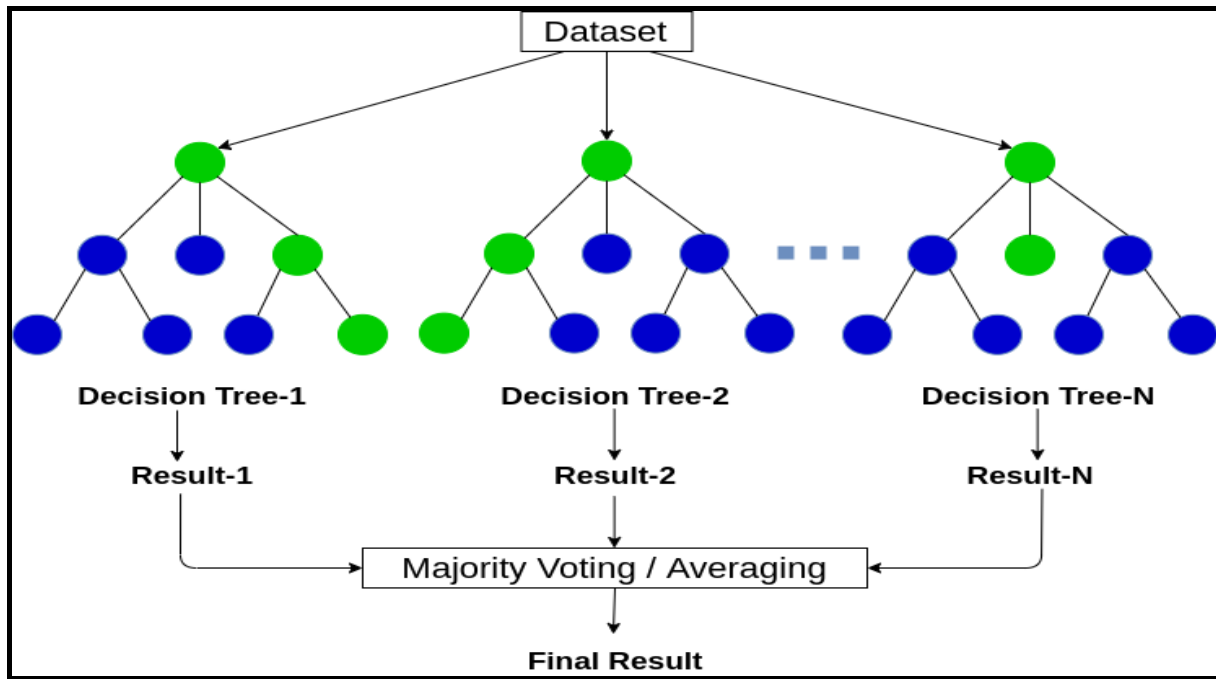
Fig4: Random Forest

Entropy: Entropy is the measurement of disorder or impurities in the information processed in machine learning. It determines how a decision tree chooses to split data.

$$S = -\sum_{i=1}^{n} p_i \log{_2} p_i$$

S: Case set
n: Number of S partition
Pi: Probability obtained from the total (Yes / No) divided by the total case

Applying Random Forest:

For this dataset we got the accuracy upto 88% by using random forest algorithm.

```python
step1 = ColumnTransformer(transformers=[
    ('col_tnf',OneHotEncoder(sparse=False,drop='first'),[0,1,7,10,11])
],remainder='passthrough')

step2 = RandomForestRegressor(n_estimators=100,
                              random_state=3,
                              max_samples=0.5,
                              max_features=0.75,
                              max_depth=15)

pipe = Pipeline([
    ('step1',step1),
    ('step2',step2)
])

pipe.fit(X_train,y_train)

y_pred = pipe.predict(X_test)

print('R2 score',r2_score(y_test,y_pred))
print('MAE',mean_absolute_error(y_test,y_pred))

R2 score 0.8873402378382488
MAE 0.15860130110457718
```

Applying Decision Tree:

For this dataset we got the accuracy upto 84% by using Decision making algorithm.

```python
step1 = ColumnTransformer(transformers=[
    ('col_tnf',OneHotEncoder(sparse=False,drop='first'),[0,1,7,10,11])
],remainder='passthrough')

step2 = DecisionTreeRegressor(max_depth=8)

pipe = Pipeline([
    ('step1',step1),
    ('step2',step2)
])

pipe.fit(X_train,y_train)

y_pred = pipe.predict(X_test)

print('R2 score',r2_score(y_test,y_pred))
print('MAE',mean_absolute_error(y_test,y_pred))
```

```
R2 score 0.8309669409734675
MAE 0.18631000131141304
```

IV. CONCLUSION

The data which we use in this analysis have taken from the Kaggle, which are based on the real world pricing. In this analysis we had done data cleaning, exploratory data analysis and data visualization. We here by conclude that this laptop price prediction can be give the 88% accurately work by doing the classification using random forest which is a machine learning technique.

REFERENCES

[1]. Aminah Md Yusof and Ismail Syuhaida, "Multiple Regressions in Analyzing House Price Variations", IBIMA Publishing Communications of the IBIMA, vol. 2012, pp. 9, 2012.

[2]. Pandey M, Sharma VK. A decision tree algorithm pertaining to the student performance analysis and prediction. International Journal of Computer Applications. 2013 Jan 1;61(13).

[3]. M. A. Shaik, M. Varshith, S. Sri Vyshnavi, N. Sanjana and R. Sujith, "Laptop Price Prediction using Machine Learning Algorithms," 2022 International Conference on Emerging Trends in Engineering and Medical Sciences (ICETEMS), Nagpur, India, 2022

[4]. Mehtab, S., Sen, J.: A time series analysis-based stock price prediction using machine learning and deep learning models, 1–46 (2020).