International Journal of Innovative Research in Engineering
Volume 4, Issue 3(May-June 2023), PP: 488-491.
www.theijire.com ***ISSN No: 2582-8746***

Laptop Price Prediction Using Machine Learning

J. S. Pawar[1], Rahul Jeswani[2], Satyam Sadhotra[3], Pranjal Dube[4]
[1] *Professor, Department of Information Technology, Sinhgad College of Engineering, Pune, India.*
[2,3,4] *Department of Information Technology, Sinhgad College of Engineering, Pune, India.*

How to cite this paper:
J. S. Pawar[1], Rahul Jeswani[2], Satyam Sadhotra[3], Pranjal Dube[4], "Laptop Price Prediction Using Machine Learning", IJIRE-V4I03-488-491.

Abstract: *With the use of several features, including processor speed, memory, storage, screen size, brand, and other pertinent elements, this research intends to create a predictive model for laptop prices. Many e-commerce websites and other pertinent sources provided the data for this study. The purpose of the study is to construct a prediction model for estimating laptop pricing by identifying the major features that influence laptop prices using a variety of statistical and machine learning techniques. The findings of this study can help buyers and retailers forecast a laptop's pricing based on its features and other important considerations.*

Key Word: *Machine Learning, Random Forest.*

I.INTRODUCTION

After 40 years of development since the laptop's conception in 1984, laptops are now used for a wide range of activities, including business, education, and leisure. It goes without saying that people these days live mobile lifestyles both at home and at work. Customers need and require the flexibility to access their computers from any location, including the kitchen table and the office, which necessitates portable devices. Sales of desktop PCs are declining, while those of laptops are rising and are expected to keep doing so. According to a Statista survey, the laptop market segment was on a revenue growth trajectory of continuously rising even before the outbreak, rising from $132 billion in 2018 to $140 billion in 2020.

Yet, during the past three years, laptop sales have significantly increased as a result of remote working, learning, and expanding gaming demand. When comparing pricing between online and brick-and-mortar stores, where the same laptop model may be bought at some astonishingly different costs, the corresponding price has also varied dramatically. When they are uninformed of the price shift, consumers continue to overspend on subpar products. The team concentrated on this problem using a Kaggle dataset, and they were successful in developing models for random forest to predict laptop costs.

II.RELATED WORK

The task of predicting laptop prices for one of the Kaggle competitions has so many approaches in the current system that they combine common machine learning techniques with their unique concepts, such as neural network machines, logit transforms, and residual regression. Yet, data analysis findings indicate that the laptop price variation forecast results are insufficiently precise. Because of the tiny sample size, the results' Standard Deviation can occasionally be extremely high.

III.RESEARCH METHODOLOGY

Machine learning is a subfield of artificial intelligence that deals with developing software that can predict the future based on historical data. This article will assist you in starting from scratch to develop your own end-to-end machine learning project if you are a data science enthusiast or practitioner. Building a machine learning project involves a number of processes, although not all of them must be used in one project; it depends on the dataset.

Getting a dataset of laptop pricing

```
import numpy as np
import pandas as pd
from sklearn.impute import SimpleImputer
```

```
#reading dataset
dataset = pd.read_csv('laptop_price.csv', encoding='latin-1')
```

```
dataset.head()
```

	laptop_ID	Company	Product	TypeName	Inches	ScreenResolution	Cpu	Ram	Storage	Gpu	OpSys	Weight	Warranty	Price_euros
0	1	Apple	MacBook Pro	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 2.3GHz	8GB	1 TB HDD	Intel Iris Plus Graphics 640	macOS	1.37kg	1 Year Onsite Warranty	1339.69
1	2	Apple	Macbook Air	Ultrabook	13.3	1440x900	Intel Core i5 1.8GHz	8GB	256 GB SSD	Intel HD Graphics 6000	macOS	1.34kg	1 Year Onsite Warranty	898.94
2	3	HP	250 G6	Notebook	15.6	Full HD 1920x1080	Intel Core i5 7200U 2.5GHz	8GB	256 GB SSD	Intel HD Graphics 620	No OS	1.86kg	1 Year Onsite Warranty	575.00
3	4	Apple	MacBook Pro	Ultrabook	15.4	IPS Panel Retina Display 2880x1800	Intel Core i7 2.7GHz	16GB	256 GB SSD	AMD Radeon Pro 455	macOS	1.83kg	1 Year Onsite Warranty	2537.45
4	5	Apple	MacBook Pro	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 3.1GHz	8GB	1 TB HDD	Intel Iris Plus Graphics 650	macOS	1.37kg	1 Year Onsite Warranty	1803.60

Basic Data Exploration

After loading the dataset via Pandas, we can see a list of laptops and specs that are associated with each laptop. Looking at the dataset, we can see that some columns such as Screen Resolution and CPU have alphanumeric data while other features consist of purely numerical or alphabetical values. These data would need to be filtered and engineered later. To avoid any complications and error-prone predictions, useless features such as "Company" and "Product" will be removed from the dataset.

Feature Engineering

Our next step would be to extract and restructure the data in order to gain a deeper understanding of the underlying factors that contribute to laptop pricing. Upon analyzing the Screen Resolution column, we observe that certain laptops possess touch screen capabilities. As touch screen laptops are typically more expensive than their non-touch counterparts, we will add a Touch screen feature to identify laptops with this capability.

Exploratory Data Analysis (EDA)

Through Exploratory Data Analysis (EDA), we can now generate visualizations and tables to examine the relationship between each feature and the variation in laptop prices. By utilizing the barplot function from Matplotlib, we can evaluate and confirm our initial assumptions about how certain features impact laptop pricing.
To illustrate, we can generate a barplot for the Type Name feature, which denotes the type of laptop.

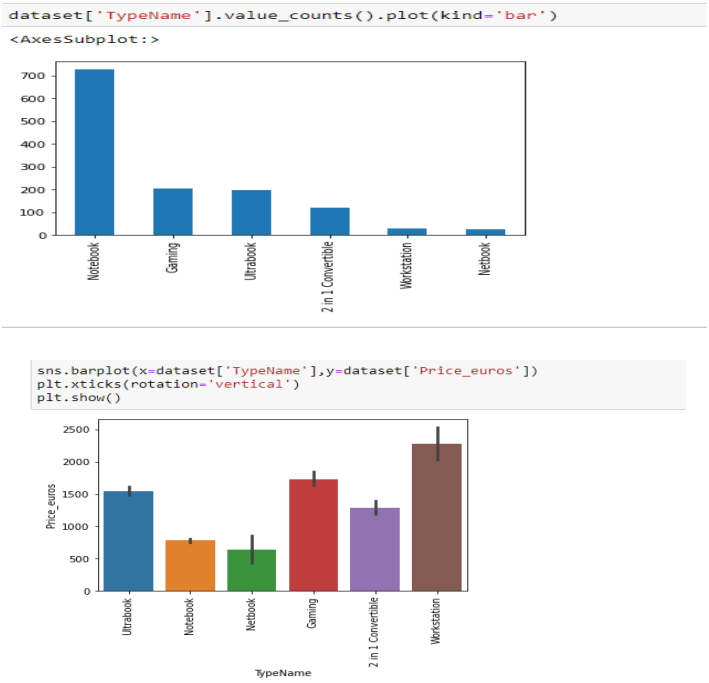


Figure: Data Visualization using Different Plots

Based on the above barplot, we can infer and conclude that workstation and gaming laptops have higher average prices compared to other laptop types. This can be expected since these laptops typically have superior specifications such as better CPUs and higher memory capacity, which are necessary for meeting the demands of professional work environments. Notebooks and net books have lower prices due to their lower-powered configurations.

When plotting bar graphs for the CPU features, some interesting findings emerge. In general, processors with higher power should have higher prices than those with lower power. Prices for Intel processors generally follow this pattern: Xeon > i7 > i5 > i3. The same principles also apply to AMD CPUs, with Ryzen > AMD A series > E series.

Data Preprocessing

The following section will involve renaming and converting categorical features into numerical ones, which is necessary for training machine learning models. This is because ML models only accept numerical inputs.

To begin, we will identify non-numerical features (those with Object type) and calculate their cardinalities, which refer to the number of categories present in each feature.

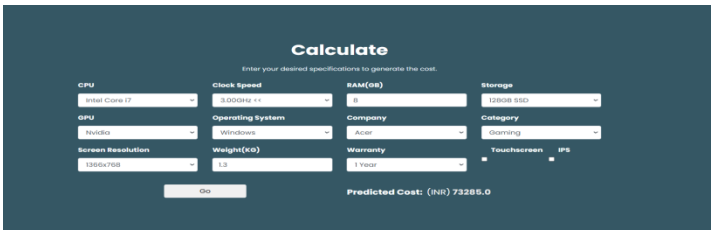
Modeling

After importing the preprocessed .csv dataset, we proceed to identify our dependent variable (Price) and create a separate data frame for it. Machine Learning involves detecting patterns in data, and one can use either supervised or unsupervised learning approaches. Common ML tasks include regression, classification, forecasting, and clustering.

At this stage of the process, we must apply mathematical, computer science, and business knowledge to train a Machine Learning algorithm that can generate predictions based on the given data. This is a critical step that will determine the accuracy and quality of future predictions in new scenarios. Additionally, ML algorithms can assist in identifying key features with high predictive value.

Website

To Build our website Flask is used. Flask provides developers with tools and features for building web applications and APIs using Python. It is open-source and has a simple and intuitive API that makes it easy to use and extend. It is designed to be lightweight, flexible, and modular, so that developers can easily customize it to meet their specific needs.



IV.ARCHITECTURE

Dataset Selection

The first stage of the machine learning process involves selecting an appropriate dataset for training the algorithm. The choice of dataset depends on the specific problem at hand and the availability of relevant data.

Data Pre-Processing

Data pre-processing involves the collection and cleaning of data in order to minimize noise and enhance accuracy. To achieve this, algorithms can be employed to eliminate unknown values, outliers, or any other factors that may have an impact on the accuracy of predictions.

Feature Selection

Feature selection involves the identification of the most significant features that will aid in predicting the outcome variable in our model. This can be achieved through an exploratory process or by employing a pre-processing step such as principal component analysis. The objective is to choose features that are pertinent to the prediction of the outcome variable in the model.

Prediction Model

After converting your data into a machine learning compatible format, we can construct an accurate prediction model tailored to your needs. We offer a variety of models specialized for various purposes such as predicting high-end laptops, budget laptops, or gaming laptops. You can choose from our selection of models based on your specific needs and budget constraints.

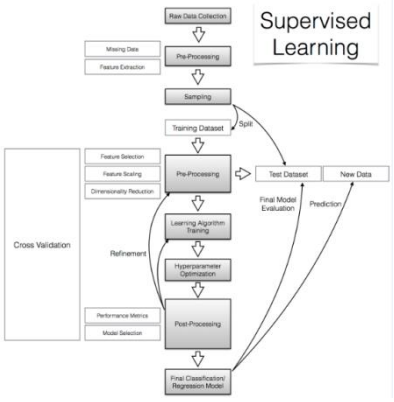


Fig. Working of Machine Learning Algorithm

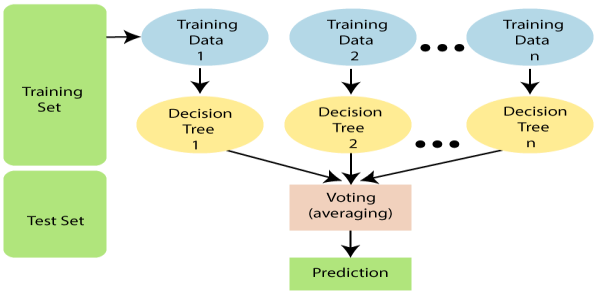


Fig. Working of Random Forest Algorithm

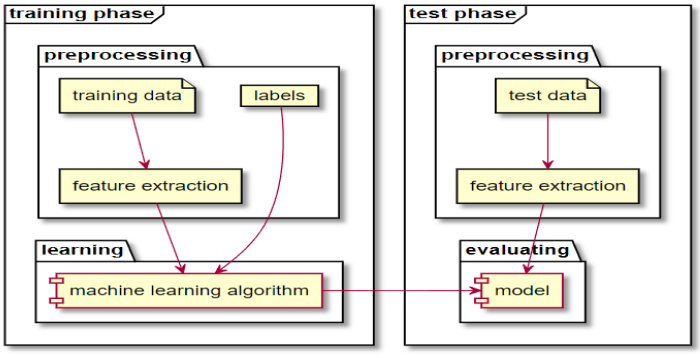


Fig. UML Diagram

V.RESULT ANALYSIS AND DISCUSSION

To create a laptop price prediction website, data on laptop specifications and prices must first be collected, cleaned, and processed. Utilizing machine learning algorithms, accurate prediction models can be generated for various laptop brands and types. The end result is a website that enables users to input laptop specifications and receive an estimated price range based on available data. This type of platform is useful for consumers who want to make informed purchasing decisions and for manufacturers and retailers looking to gain insights into pricing strategies and market trends. Ultimately, the development of a laptop price prediction website provides valuable information on the correlation between laptop specifications and pricing in the market.

VI.FUTURE SCOPE AND CONCLUSION

By adding more sophisticated algorithms, extending the database, and including user ratings, the laptop prediction website has room to grow. Buyers of laptops may find it to be a useful tool.

References

[1] International Journal of Computer Science and Mobile Computing. Laptop Price Prediction using Machine Learning.
[2] https://www.researchgate.net/publication/ 50946368_Exploratory_data_analysis_in_the_ context_of_data_mining_and_resampling.
[3] https://www.academia.edu/69591584/Laptop_Price_Prediction_using_Machine_Learning
[4]https://medium.com/analytics-vidhya/laptop-price-prediction-by-machine-learning-7e1211bb96d1
[5] https://www.digitalocean.com/community/tutorials/how-to-make-a-web-application-using-flask-in-python-3
[6] https://www.kaggle.com/datasets/muhammetvarl/laptop-price