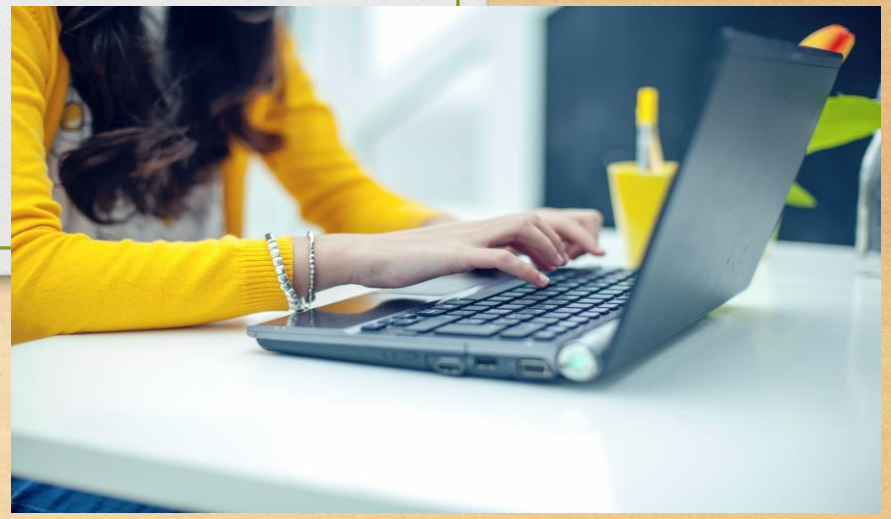
Laptop price predictor

- Ashwini patil

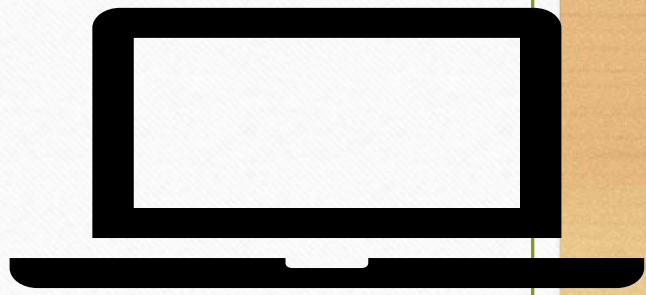


CONTENTS

INTRODUCTION

- Laptop price prediction is a challenging task due to the dynamic nature of the market and the many factors that can affect pricing. The goal of this model is to accurately predict the prices of laptops based on various features

- Laptop price prediction is important for both buyers and sellers buyers can use the model to predict the prices of laptops and make informed purchasing decisions while sellers can the model to set appropriate prices for their products



Problem statement

- We will make a project for laptop price prediction the problem statement is that if any user wants to buy a laptop then our application should be compatible to provide a tentative price of laptop according to the user configurations . Although it looks like a simple project or just developing a model , the dataset we have is noisy and needs lost of feature engineering , and preprocessing that will drive your interest in developing this project .

Data Exploration

- Most of the columns in a dataset are noisy and contain lost of information . But with feature engineering you do , you will get more good results . The only problem is we are having less data but we will obtain a good accuracy over it . The only good thing is it is better to have a large data . We will develop a website that could predict a tentative price of a laptop based on user configuration .

Dataset

```
In [10]:
1 df=pd.read_csv('laptop_data.csv')
2 df
```

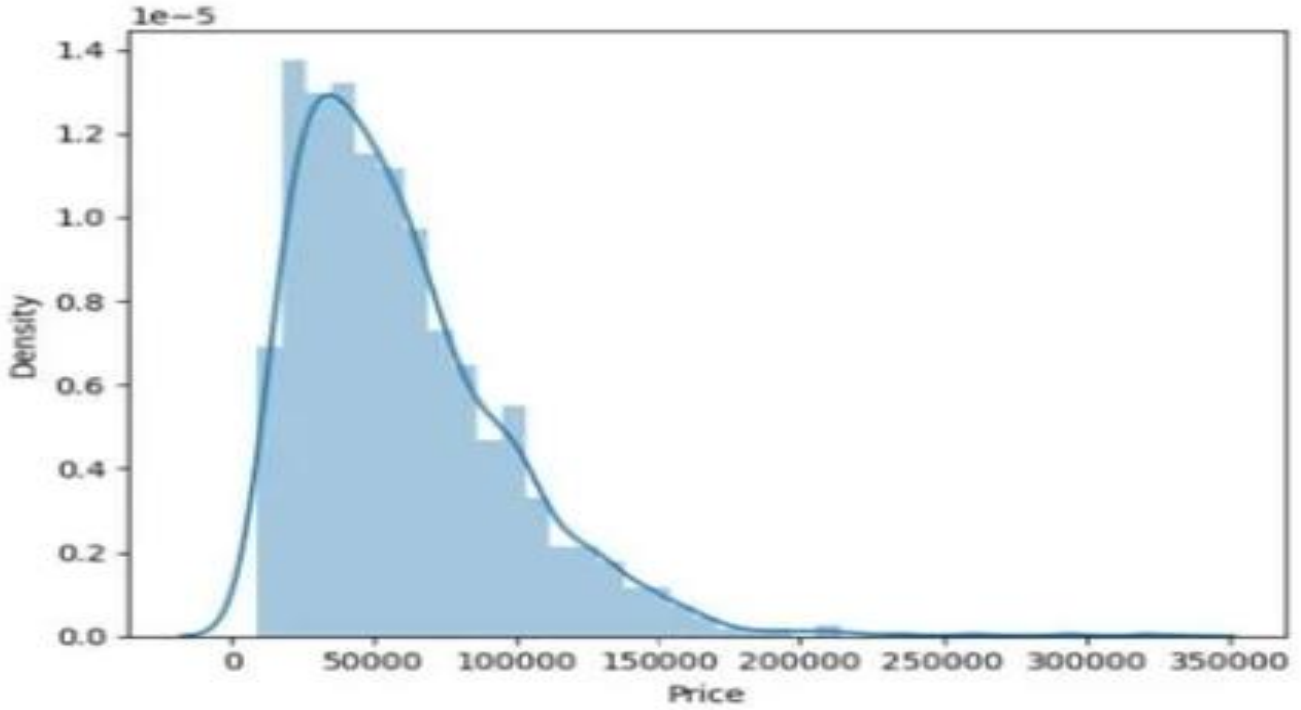
1298	1298	Lenovo	2 in 1 Convertible	14.0	IPS Panel Full HD / Touchscreen 1920x1080	Intel Core i7 6500U 2.5GHz	4GB	128GB SSD	Intel HD Graphics 520	Windows 10	1.8kg	3399
1299	1299	Lenovo	2 in 1 Convertible	13.3	IPS Panel Quad HD+ / Touchscreen 3200x1800	Intel Core i7 6500U 2.5GHz	16GB	512GB SSD	Intel HD Graphics 520	Windows 10	1.3kg	7986
1300	1300	Lenovo	Notebook	14.0	1366x768	Intel Celeron Dual Core N3050 1.6GHz	2GB	64GB Flash Storage	Intel HD Graphics	Windows 10	1.5kg	1220
1301	1301	HP	Notebook	15.6	1366x768	Intel Core i7 6500U 2.5GHz	6GB	1TB HDD	AMD Radeon R5 M330	Windows 10	2.19kg	4070
						Intel						

EXPLORATORY DATA ANALYSIS

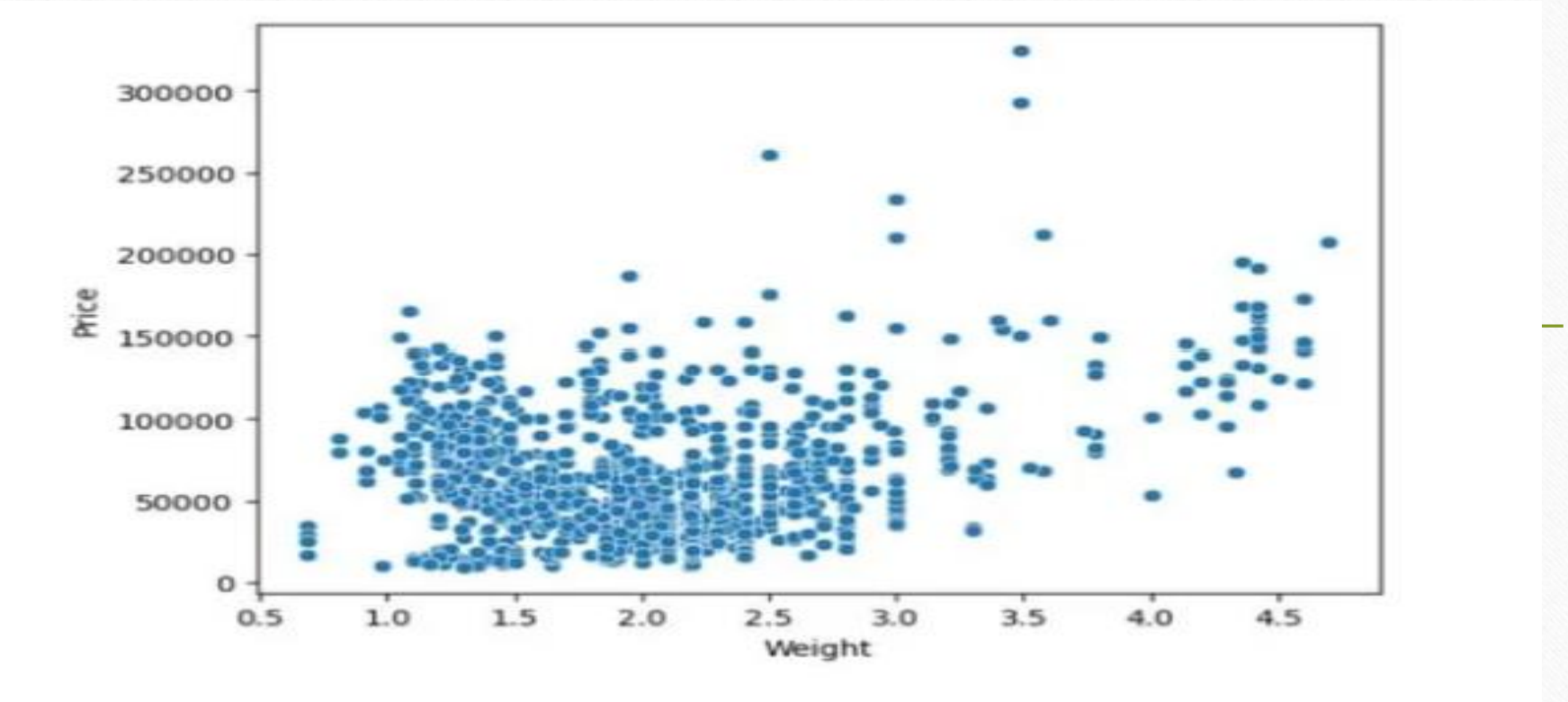
Exploratory analysis is a process to explore and understand the data and data relationship in a complete depth so that it makes feature engineering and machine learning modeling steps smooth and streamlined for prediction. EDA involves Univariate, Bivariate, or Multivariate analysis. EDA helps to prove our assumptions true or false. In other words, it helps to perform hypothesis testing. We will start from the first column and explore each column and understand what impact it creates on the target column. At the required step, we will also perform preprocessing and feature engineering tasks. our aim in performing in-depth EDA is to prepare and clean data for better machine learning modeling to achieve high performance and generalized models. so let's get started with analyzing and preparing the dataset for prediction.

Data Distribution

```
Out[15]: <AxesSubplot:xlabel='Price', ylabel='Density'>
```



Price corr with weight

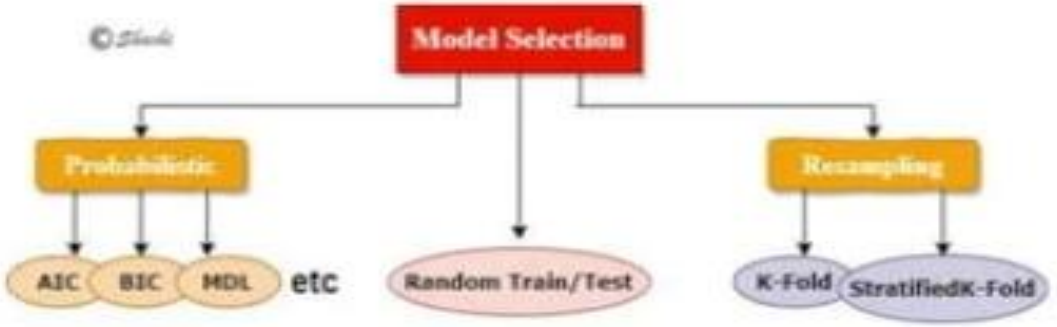


Model Selection

In the first step for categorical encoding, we passed the index of columns to encode, and pass-through means pass the other numeric columns as it is. The best accuracy I got is with all-time favorite Random Forest. But you can use this code again by changing the algorithm and its parameters. I am showing Random forest. you can do Hyper parameter tuning using Grid search CV or Random Search CV. we can also do feature scaling but it does not create any impact on Random Forest.

Different ML models use for Model Selection:

- Linear Regression
- Ridge Regression
- Lasso Regression
- KNN
- Decision Trees
- Random Forest
- SVM
- Ada Boost
- Gradient Boost



conclusion

- **Summary of key findings:** Summarize the key findings of the model development process, including the performance of the final model and any insights into the factors that affect laptop prices.
- **Model performance:** Provide a summary of the model's performance, such as the accuracy, recall, precision, F1-score, and AUC-ROC, and compare it with the benchmark models or with other models that were considered.
- **Model interpretability:** Discuss how the model provides insights into the factors that affect laptop prices, and how it can be used to make more informed decisions.
- **Potential Impact:** Discuss the potential impact of the model on the domain, such as how it can be used to improve pricing strategies and make better decisions for buyers and sellers.
- **Limitations:** Summarize the limitations of the model, such as the availability of data, the quality of data, or the complexity of the model, and discuss how they may have affected the results.
- **Future work:** Discuss any future work that could be done to improve the model, such as collecting more data, incorporating additional features, or testing different model architectures.
- **Conclusion:** Sum up the main takeaways from the model development and results and its potential impact on the domain.


Thank You