

**ROYAL UNIVERSITY OF BHUTAN**

**Gyalpozhing College of Information Technology**

**Laptop Price Prognostication**

**CSA203 – ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**Group Member:**

DEKI LHAZOM(12210048)

KINZANG WANGDI(12210020)

CHIMMI RINZIN(12210002)

DAWA(12210046)

PEMA YANGCHEN(12210025)

**Module Tutor:**

NIMA DEMA

# Introduction

The laptop has become an important item in everyday life, but it can be difficult to sell and choose. Laptops come in all sorts of prices, features and speciﬁcations. Price estimation and prediction is an important part of consumer strategy. Deciding on the correct price of a product is very important for the market success of a product. Laptop price prediction model is needed to predict the price based on user input such as brand, processor, memory and storage.

# Aim

To provide more accurate and reliable forecasts of future laptop prices, improve decision making and enhance the efficiency and effectiveness of the laptop market.

1. **Goal**

To develop a machine learning model and accurately predict the price of the laptop.

# Problem Statement

Laptops prices tend to differ due to the rising demand from workers, students, and gamers. Laptop prices are subject to significant variations based on various factors such as their specifications, brand, and other .variables. Customers may face challenges in assessing the reasonable market value of a laptop or determining if they are receiving a reasonable deal. So, our platform will offer the solution to purchase a laptop at a reasonable price.

# Scope

# System scope:

System will predict the price of the laptop.

# User scope:

Price prognosticator will be targeted to all the customers who wants to purchase a laptop.

# Literature Review

1. **Keepa:**

Keepa is a web application that provides price history charts and price drop alerts for millions of products on Amazon. The Keepa website and browser extension offer a range of features that can help you track prices and make informed purchasing decisions. With Keepa, you can view the price history of a product over a specific period of time, ranging from a few days to several years. The price history chart shows the lowest and highest prices, as well as the average price, for the product during the selected time period. You can also view the price history for different Amazon marketplaces, such as Amazon US, Amazon UK, and Amazon Germany. Keepa also offers price drop alerts. Keepa offers a free version with limited features, as well as a paid version with additional features such as advanced price history charts and alerts for out-of-stock products. Keepa is a popular tool among Amazon shoppers and can be a useful resource for predicting the future price of a laptop or other products on Amazon.

1. **PriceSpy**:

PriceSpy is a website and app that allows you to compare prices of products from various online retailers. The website was founded in 2000 in Sweden, and it has since expanded to other countries, including the UK, Australia, and New Zealand. PriceSpy is popular among shoppers looking to find the best deals on products such as laptops, smartphones, and home appliances. The website has a large database of products from thousands of retailers, making Page 3 it easy to find the best prices and deals. One of the key features of PriceSpy is the price history graph, which shows you how the price of a product has changed over time. This information can be helpful in determining whether the current price is a good deal or if you should wait for a price drop. In addition to comparing prices, PriceSpy also provides user reviews, product specifications, and ratings to help you make an informed decision. You can also create price alerts for specific products, which will notify you when the price drops to a certain level.

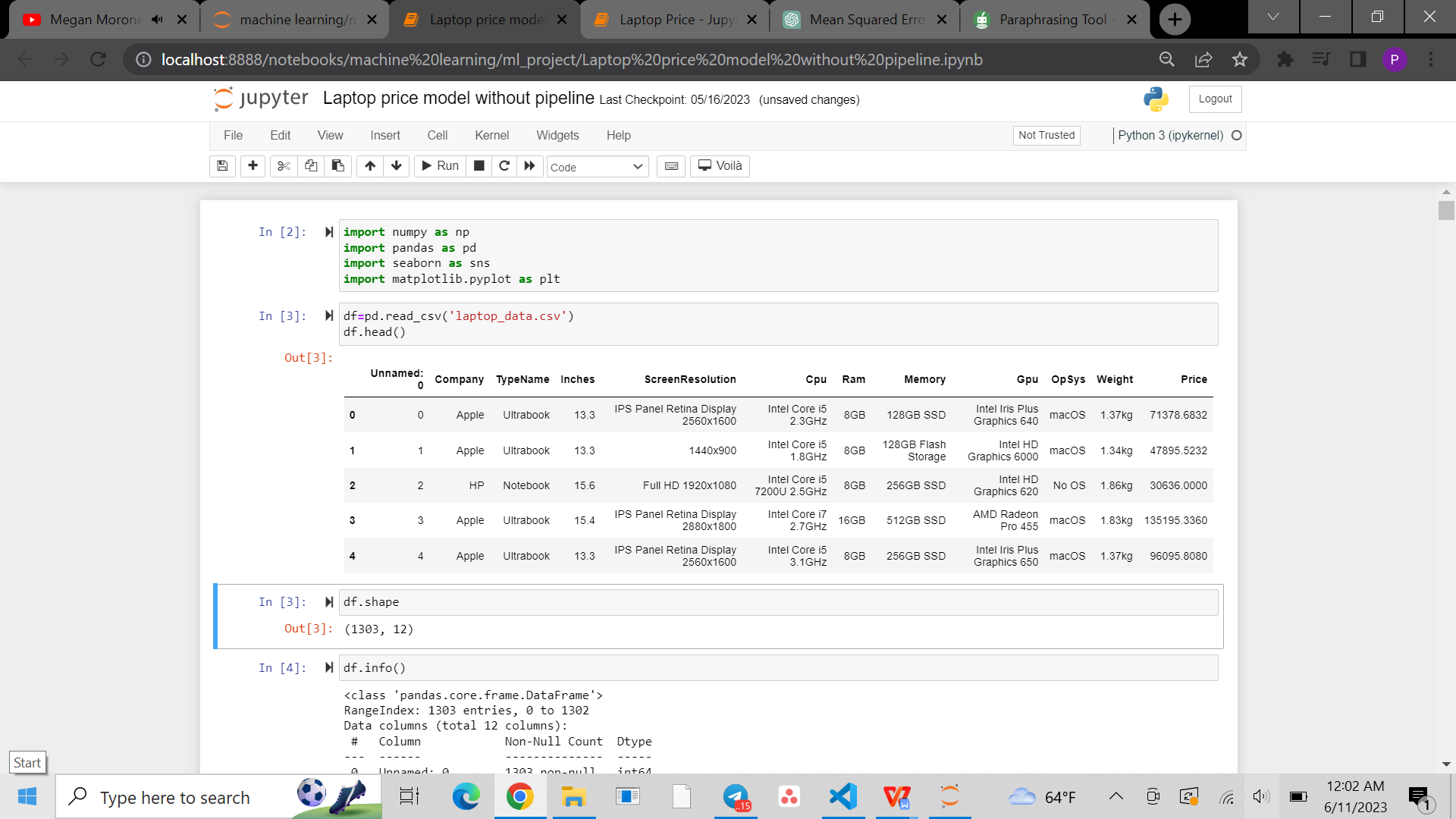
PriceSpy is free to use, and it earns revenue through affiliate links. When you click on a link to a retailer and make a purchase, PriceSpy earns a commission. Overall, PriceSpy is a useful tool for anyone looking to save money on their online purchases.

1. **Buyvia:**

Buyvia is a mobile app and website that offers price comparison and alerts for a wide range of products, including laptops. The site provides users with access to real-time pricing information from multiple retailers, allowing them to easily compare prices and find the best deal.

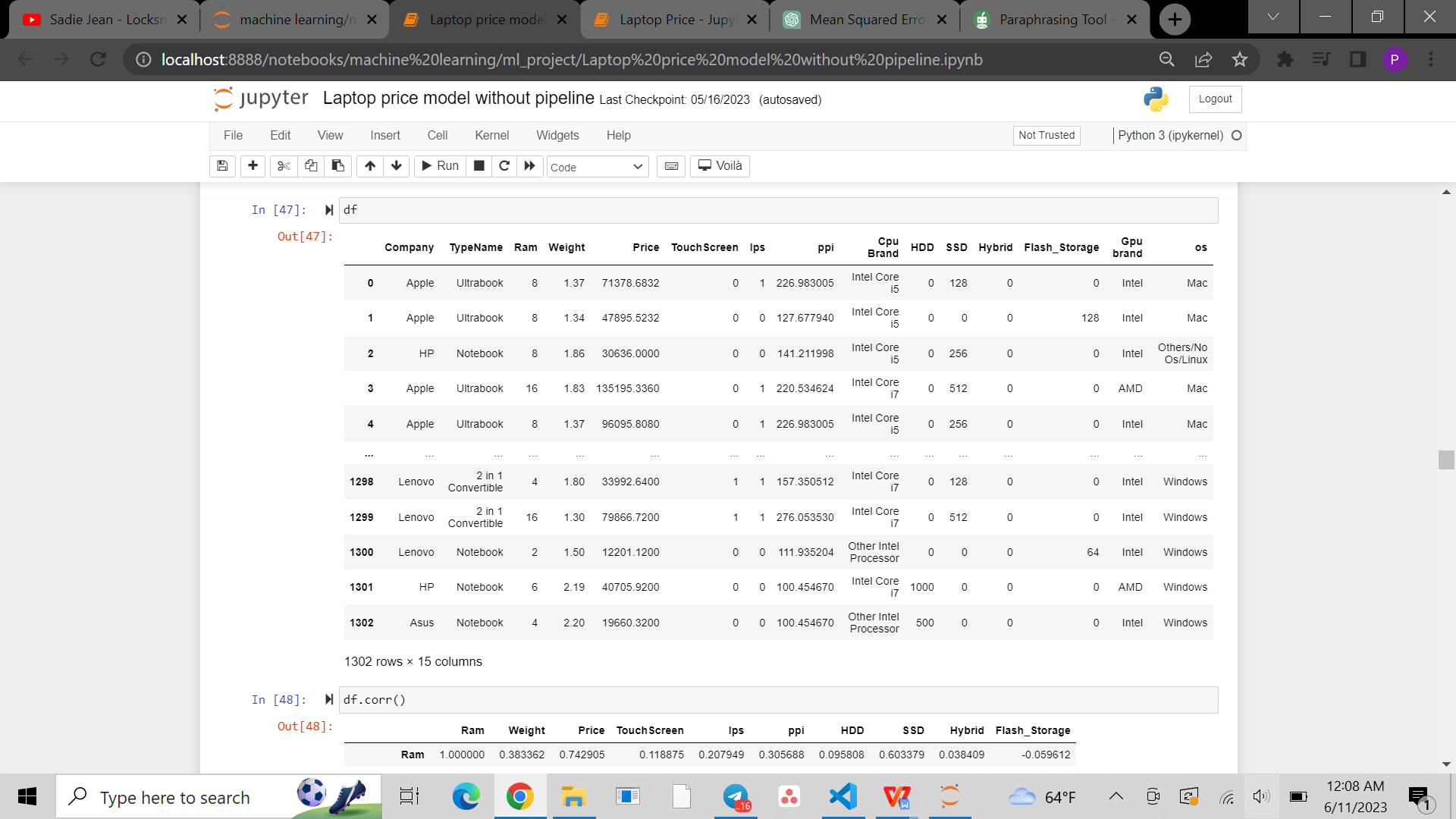
# Dataset

The dataset we collected has 1303 rows and 12 columns. It has features like Unnamed, Company, TypeName, Inches, ScreenResolution, Cpu, Ram, Memory, Gpu, OpSys, Weight and the target column/feature Price.



During preprocessing steps, we extracted important information from the columns into new columns and drop the columns after extracting the information. After the preprocessing steps were done, we are left with 15 features including the target feature.

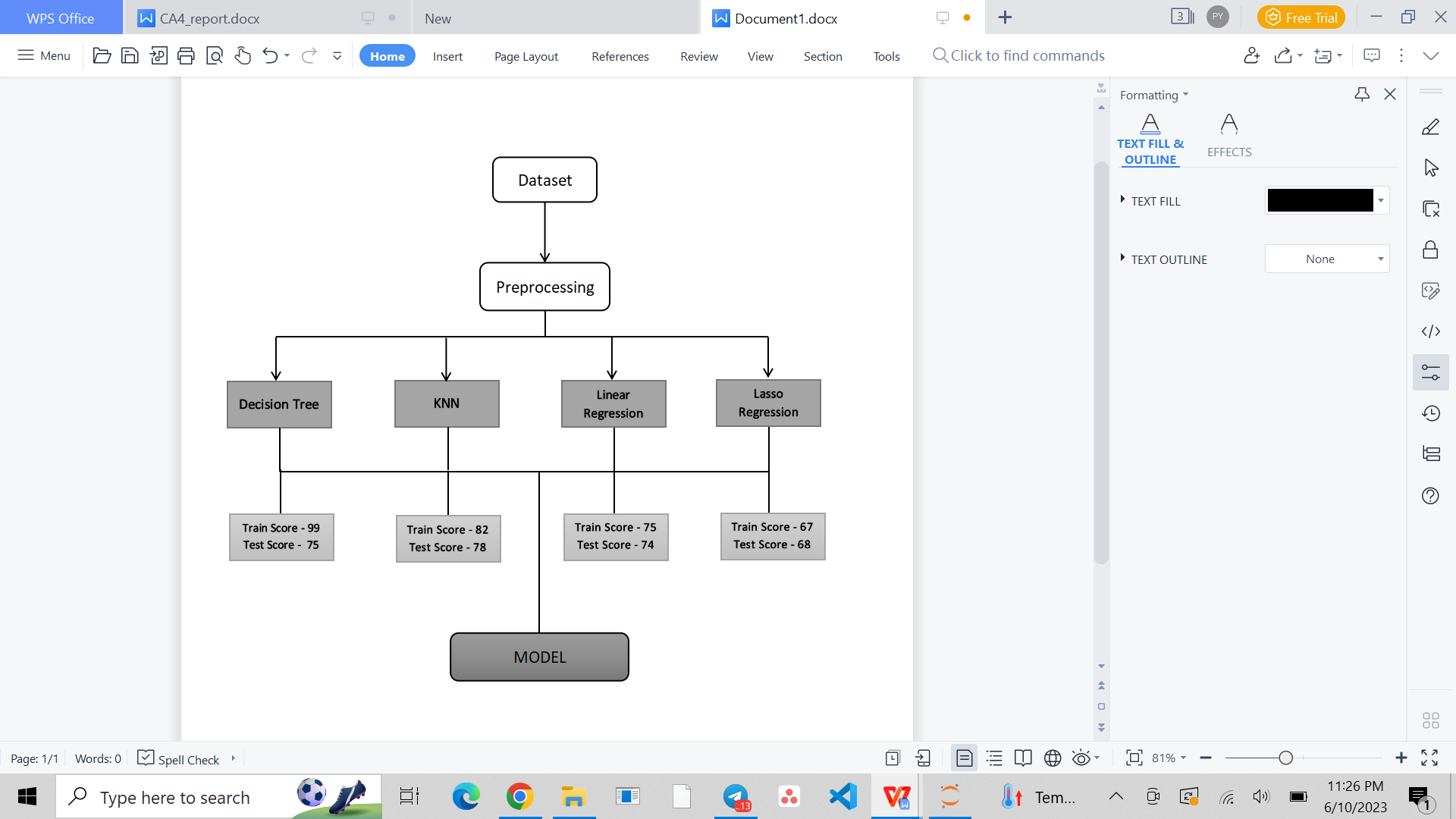
Our final dataset:



# Methodology

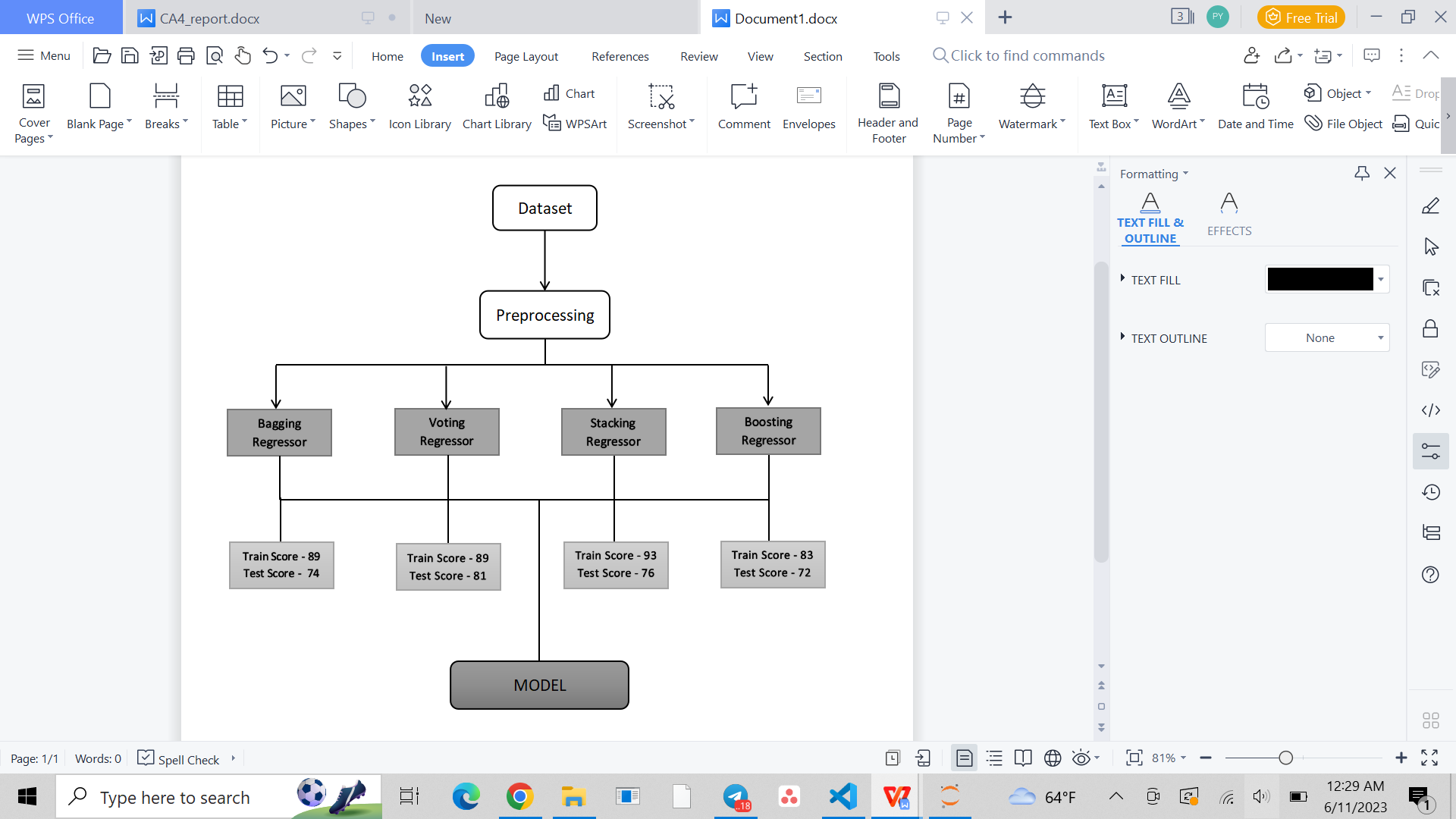
# Algorithm

We have implemented the following algorithms to our datasets and checked the accuracy for each.

****

1. Decision Tree’s goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. A tree can be seen as a piece wise constant approximation.
2. K-Nearest Neighbors It is also known as KNN or k-NN, is a non-parametric, supervised learning classifier that employs proximity to classify or anticipate how a certain data point will be grouped.
3. Linear Regression uses independent variables to model a goal prediction value. It mostly serves to determine how factors and forecasting interact.
4. The lasso regression allows you to shrink or regularize these coeﬃcients to avoid overﬁtting and make them work better on diﬀerent datasets.
5. **Ensemble Learning**

After that, we perform ensemble learning to improve the overall predictive performance of our model by combining the predictions of multiple model.



1. Bagging regressor: combines multiple regression model through the process of bagging (bootstrap aggregating).
2. Voting regressor: combines the predictions of multiple model to generate the final prediction.
3. Stacking regressor: combines multiple model in hierarchical manner to make prediction.
4. Boosting regressor: combines multiple weak model to create a strong predictive model.

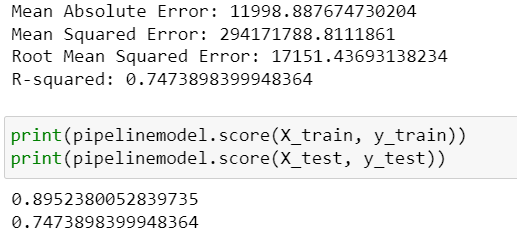
# Evaluation Metrics

Once the machine learning model is trained we should see its eﬃciency to make sure that our model is well trained to make good predictions on newly observed data.

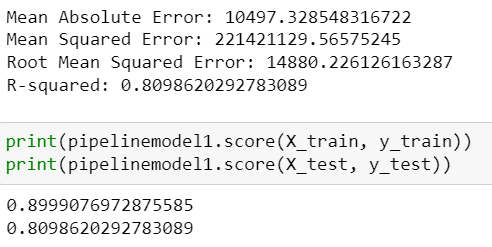
Once the model is trained, we checked the accuracy of the model.

* + - * R-squared is a goodness-of-ﬁt measure for linear regression models. This statistic indicates the percentage of the variance in the dependent variable that the independent variables explain collectively.
      * Mean Absolute Error (MAE): MAE measures the average magnitude of the errors in a set of predictions, without considering their direction. It's the average over the test sample of the absolute diﬀerences between prediction and actual observation where all individual diﬀerences have equal weight.
      * Mean Squared error (mse): MSE is a common metric used to measure the average squared difference between the predicted value and the actual value.

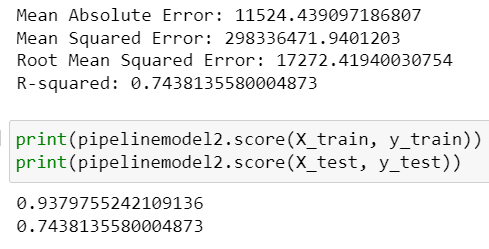
1. **Evaluation of model**
2. **Bagging ensemble learning**



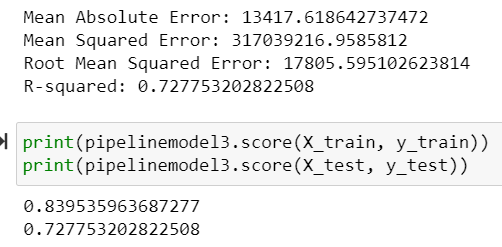
1. **Voting ensemble learning**



1. **Stacking ensemble learning**



1. **Boosting ensemble learning (AdaBoostRegressor)**



1. **Final model selection**

The final model selection was voting regressor ensemble learning with accuracy score of 80%, trains score of 89% and 81% of test score. We choose the voting regressor learning model because compared to other estimator with and without ensemble learning, it gives high accuracy score and have least difference between train and test score. Without ensemble learning, linear regression have best performance with 75% train score and 74% test score. After implementing voting ensemble learning, accuracy improve to 89% and 81% train and test score respectively.

# Result and Discussion

# Experimental Setup

1. Anaconda

An open-source distribution of Python and R for data research called Anaconda seeks to make package management and deployment easier. Packages for data science are available in the distribution that work with Windows, Linux, and macOS.

1. Jupyter Notebook

The objective of Project Jupyter is to create open-source software, open standards, and interactive computing services for a variety of programming languages. The first web application for producing and sharing computational documents was called a Jupyter Notebook. It provides a straightforward, efficient, document-focused experience.

1. Scikit Learn

It is a Python programming language machine learning package that is available for free. Support-vector machines, random forests, gradient boosting, k-means, and DBSCAN are just a few of the classification, regression, and clustering algorithms it offers. It is also built to work with Python's NumPy and SciPy scientific and numerical libraries.

1. Matplotlib and seaborn

For plotting, we use matplotlib and seaborn. On top of matplotlib, seaborn offers an API that provides reasonable options for plot style and color defaults, outlines straightforward high-level functions for popular statistical plot kinds, and connects with Pandas DataFrame's features.

# Programming language: Python

1. Python Libraries for Data Processing and Model

* Pandas: Is one of those libraries for data analysis that contains high-level data structures and tools to manipulate data in a simple way.

Sci-Kit Learn: It has supervised and unsupervised machine learning algorithms for production applications. Sci-Kit Learn focuses on code quality, documentation, ease of use, and performance as this library provides learning algorithms.

* NumPy: Is a Python library for numerical calculations and scientiﬁc computations. NumPy provides numerous features which Python enthusiasts and programmers can use to work with high-performing arrays and matrices.

1. **Web deployment**

We integrated model with web using flask and deployed in python anywhere platform as it consist of libraries we required.

### [**laptopprice.pythonanywhere.com**](http://laptopprice.pythonanywhere.com/)

# Conclusion

Based on a few particular specifications, the goal of this research was to estimate the cost of laptops from manufacturers like Apple, Dell, Lenovo, Acer, HP, etc.

We used a variety of supervised regression strategies to solve this problem and determined the most effective one. To determine which algorithm provides the greatest prediction, we tested many evaluation metrics such as mse, mae, accuracy\_score and rmse.

Among different algorithm, we chose voting ensemble learning with the accuracy score of 80% and train score of 89% and test score of 81%.

# Bibliography

Chapke, A., Barapatre, A., & Lohakare, S. (2022, January). Laptop Price Prediction using Machine Learning. *11*(1), 164–168. doi:10.47760/ijcsmc. 2022.v11i01.021

Journal, I. J. C. S. M. C. (2022). Laptop Price Prediction using Machine Learning. IJCSMC, (1), 164–168. <https://doi.org/10.47760/ijcsmc.2022.v11i01.021>

S., Subhiksha & Thota, Swathi & Sangeetha, J.. (2020). Prediction of Phone Prices Using

Machine Learning Techniques. 10.1007/978-981-15-1097-7\_65. Retrieved from <https://www.researchgate.net/publication/338471736_Prediction_of_Phone_Prices_Using_Machine_Learning_Techniques>

*BuyVia*. (n.d.-b). BuyVia. https://www.buyvia.com/ Kopischke, N. (2022, March 19). *How to Use Keepa (Amazon Sales Rank + Price History Charts)*. Online Selling Experiment.

https://onlinesellingexperiment.com/how-to-use-keepa/#:~:text=Keeps%20is%20a%20

tool%20that,area%20on%20a%20listing%20page

*PriceSpy Ltd.* (n.d.). PriceSpy UK. https://pricespy.co.uk/