1. **INTRODUCTION**

A new car's price is established by the manufacturer, with the government incurring some additional costs in the form of taxes. Customers who buy a new car can rest assured that their money is being carefully spent. However, as new car prices rise and purchasers become unable to afford them, used car sales are on the rise all over the world. As a result, a method for estimating the value of a used car based on a variety of characteristics is urgently needed. The current approach incorporates a technique in which a vendor sets a price at random and the buyer has no idea how much the car is worth. The vendor has no idea how much the car is worth or how much he can get for it. We've designed an extremely effective model to address this problem. Because regression methods produce a continuous value rather than a categorized value as a put, they are used. As a result, instead of estimating a car's price range, it will be possible to estimate the car's actual price. In addition, a user interface has been developed that takes user input and displays the pricing of an automobile based on that input.

The investigation of training PCs to work all alone without being helped what to do is known as AI. Computer-based intelligence is a subset of ML, which has turned into the most well-known expression of the twenty-first hundred years. Man-made intelligence alludes to the activity of making PCs falsely astute with the goal that they can finish responsibilities all alone. These gadgets are incredibly exact and speedy at what they will do. We create and prepare AI Techniques by utilizing different ML approaches like Supervised Learning Unsupervised Learning, and Reinforcement Learning.

Design acknowledgment might be followed back to the starting points of Machine Learning. Furthermore, Machine learning utilizes different relapses and grouping ways to deal with train models so they can learn all alone.

**2. Plan of Automobile Price Prediction Project**

**2.1 Purpose of Project**

The automotive sector is an essential component of the global economy, and as it develops, it is critical to comprehend the present and upcoming market trends. Automobile price prediction, which involves examining historical data and current market conditions to forecast the anticipated price of a specific make and model of automobile in the future, is one of the primary topics of interest in the automotive industry. As it offers useful insights into the fair price of a car and assists them in making decisions about inventory management, pricing, and manufacturing, car price prediction can be beneficial for a variety of stakeholders, including consumers, car dealerships, and automotive manufacturers.

2.1.1. **Sales and marketing:** To maximize sales and marketing efforts, predictive models can assist automakers and dealerships in identifying client preferences, demand, and purchasing trends. Automakers can build the right models in the right amounts to satisfy demand and prevent overproduction by forecasting which vehicles are likely to be popular.

2.1.2. **Vehicle Identification:** Identification of the vehicle is essential for forecasting the price of a car because it identifies the precise year, make, and model of the car. The characteristics, mileage, condition, and history of the car, all key elements in establishing its value, can then be obtained using this information. For buyers and sellers who want to make sure they are obtaining a fair price for a vehicle or setting a price that appropriately reflects its condition and characteristics, this can be particularly crucial.

2.1.3. **Performance Evaluation:** Evaluation of performance is an essential component of automotive price forecasting. The performance and cost of a car are influenced by the engine's horsepower, torque, and fuel efficiency as well as the kind of gearbox and number of gears. The overall performance and appeal of an automobile, as well as the handling and braking system, also have an impact on its pricing.

2.1.4. **Research and Development:** Building models that can precisely forecast an automobile's price requires research and development in the field of car price prediction. This area of study is crucial for the auto industry since it aids in the pricing of vehicles by manufacturers and dealerships as well as the educated decision-making of consumers.

**2.2. Period of Project**

|  |  |  |
| --- | --- | --- |
| Date | Title | Hours |
| 25/11/22 | EDA(Exploratory Data Analysis) | 3 hr |
| 09/12/22 | Data Pre-processing on Automobile Dataset | 4 hr |
| 22/12/22 | Classification on Automobile Dataset | 4 hr |
| 16/01/23 | Regression on Automobile Dataset | 5 hr |
| 30/01/23 | Clustering on Automobile Dataset | 3 hr |
| 03/03/23 | Data Visualization on Automobile Dataset | 6 hr |
| 16/03/23 | Classification Evaluation Metrics | 2 hr |
| 09/03/23 | Regression Evaluation Metrics | 2 hr |

**2.3 Problem Statement Detailing**

Predicting a used car's sale or purchase price is a difficult task that takes into account a wide range of variables, including make, model, year, mileage, and condition, location, and market trends. Building a machine learning model that can precisely forecast the cost of a used car based on these variables is the goal of this problem statement. Features like the make, model, year, mileage, condition, location, and other pertinent characteristics of the car will be used as input to the model. The anticipated price of the car will equal the model's output.

We require a sizable dataset of used automobile sales data that contains details on the make, model, year, mileage, condition, location, and sale price of each vehicle in order to solve this issue. A machine learning model that can forecast the sale price of a car based on its attributes will be trained using this dataset. Different techniques, including linear regression, decision trees, random forests, and neural networks, can be used to construct the machine learning model. Metrics like mean absolute error (MAE), mean squared error (MSE), and R-squared can be used to assess the model's performance.

Once trained and tested, the model can be used to forecast a car's sale price based on its qualities. Car dealerships, purchasers, and sellers that wish to better grasp the fair market value of a used car may find this to be helpful. In general, estimating the sale price of a used car is a challenging challenge that calls for advanced machine learning algorithms, a sizable dataset, and thorough evaluation to assure accurate prediction.

**3. Objective of the Project**

**3.1. General Objective:**

Price prediction uses an algorithm to analyze a product or service based on its characteristics, demand, and current market trends. Then the software sets a price at a level it predicts will both attract customers and maximize sales. In some circles, the practice is called price forecasting or predictive pricing.

The general objective of automobile car prediction can vary depending on the specific context and use case, but in general, it involves using data analysis and machine learning techniques to make predictions about various aspects of automobiles and their performance. Some common objectives of automobile car prediction might include:

* Predicting the resale value of a car based on its make, model, age, mileage, and other factors.
* Predicting the fuel efficiency of a car based on its engine size, weight, and other specifications.
* Predicting the likelihood of a car experiencing mechanical issues or breakdowns based on its maintenance history and other factors.
* Predicting the safety rating of a car based on crash test results and other factors.
* Predicting the demand for a particular car model in a given market based on consumer preferences and trends.

The general objective of automobile car prediction is to use data analysis and machine learning techniques to make predictions about various aspects of automobiles and their performance, with the aim of informing decision-making around car buying, selling, and maintenance, as well as helping automotive companies better understand market demand and consumer behaviour. This may include predicting resale value, fuel efficiency, likelihood of mechanical issues or breakdowns, safety rating, and demand for a particular car model in a given market.

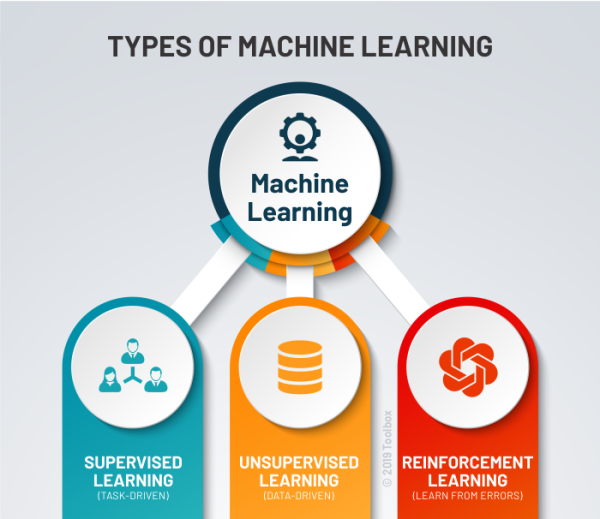
**4. Introduction**

**4.1. Introduction to Machine Learning:**

Machine learning is a type of computer technology that helps computers learn and make predictions or decisions based on data. Just like how people learn from their experiences and knowledge, machines can also learn from the data they are given.

Let's imagine you have a game where you need to sort different types of fruits. You show the computer pictures of bananas, apples, and oranges and tell the computer what each fruit is called. The computer then looks at the features of each fruit, such as its colour, size, and shape, and tries to figure out which fruit is which.

**Types of Learning :**

****

1. **Supervised Learning :**

Supervised learning is one of the most basic types of machine learning. In this type, the machine learning algorithm is trained on labeled data. Even though the data needs to be labeled accurately for this method to work, supervised learning is extremely powerful when used in the right circumstances.

It can solve the problems of **classification** and **regression**.

1. **Unsupervised Learning :**

unsupervised learning does not have labels to work off of, resulting in the creation of hidden structures. Relationships between data points are perceived by the algorithm in an abstract manner, with no input required from human beings.

It can solve **clustering** problem.

1. **Reinforcement Learning :**

[Reinforcement learning](https://it.toolbox.com/article/openais-robot-learns-to-solve-a-rubiks-cube-with-one-hand-peter-welinder-research-lead-openai-shares-insights) directly takes inspiration from how human beings learn from data in their lives. It features an algorithm that improves upon itself and learns from new situations using a trial-and-error method. Favourable outputs are encouraged or ‘reinforced’, and non-favourable outputs are discouraged or ‘punished’.

**4.2. Automobile Price Prediction Using ML:**

One of the main areas of research in machine learning is the prediction of the price of cars. It is based on finance and the marketing domain. It is a major research topic in machine learning because the price of a car depends on many factors.

The price of a car depends on a lot of factors like the goodwill of the brand of the car, features of the car, horsepower and the mileage it gives and many more. Car price prediction is one of the major research areas in machine learning. So if you want to learn how to train a car price prediction model then this article is for you. In this article, I will take you through how to train a car price prediction model with machine learning using Python.

**Some of the factors that contribute a lot to the price of a car are:**

1. Brand

2. Model

3. Horsepower

4. Mileage

5. Safety Features

6. GPS and many more.

**4.2.1. Technology is used :**

Accuracy, score of Decision Tree, Logistic Regression, Random Forest, Gradient Boosting Algorithms and Naive Bayes algorithms etc, for predicting the used car price.

**which variables are significant in predicting the price of a car?**

Price is highly (positively) correlated with wheelbase, carlength, carwidth, curbweight, enginesize, horsepower (notice how all of these variables represent the size/weight/engine power of the car

**4.2.2. Data Description :**

**Fuel type:**

Currently, the majority of motor vehicles worldwide are powered by gasoline or diesel. Other energy sources include ethanol, biodiesel, propane, compressed natural gas (CNG), electric batteries, and hydrogen (either using fuel cells or combustion). There are also cars that use a hybrid of different power sources.

**An aspiration :**

detection system is a highly sensitive form of fire detection that can detect a fire at an early stage. The system takes samples of the air and tests them for the presence of smoke. The air is drawn by a pump from the various test points.

**No. of doors :**

Passenger cars will typically have two-doors (such as coupes) or four-doors (such as sedans).A car door is a type of door opening, typically hinged on its front edge, but sometimes attached by other mechanisms such as tracks, for entering and exiting a vehicle. Carbody:

**Car body type :**

Body types are usually established by a car's 'body style', which can vary from a tall and boxy (SUV), low and sleek (coupe) or somewhere in-between (hatchback).India has a diverse range of cars available to own but most cars come in one of seven car body types: **Hatchback, Sedan, SUV, MUV, Coupe, Convertible, and Pickup Truck.**

**Drive wheel:**

A drive wheel is a wheel of a motor vehicle that transmits force, transforming torque into tractive force from the tires to the road, causing the vehicle to move. The powertrain delivers enough torque to the wheel to overcome stationary forces, resulting in the vehicle moving **forwards** or **backwards**.

**Engine type:**

a machine for converting any of various forms of energy into mechanical force and motion. also : a mechanism or object that serves as an energy source

**Types of engine:**

**Heat engine, Combustion engine, Internal combustion engine, External combustion engine.**

**Car cylinders Number:**

All engines these days are 4-stroke engines (intake, compression, power, exhaust). With a 4 cylinder engine, the overall balancing is perfect. Through every stroke made in a 4 cylinder engine, one cylinder is always in the power stroke and the rest are all at different positions than each other.

**Fuel system:**

The function of the vehicle fuel system is to store and supply fuel to the engine. The engine intake system is where the fuel is mixed with air, atomized, and vaporized. Then it can be compressed in the engine cylinder and ignited to produce energy or power.

**Wheelbase:**

The wheelbase of any vehicle is the distance between the centre of the front wheels and the centre of the rear wheels. the distance between the front and rear axles of an automotive vehicle.

**Curb weight:**

Curb weight is the weight of the vehicle including a full tank of fuel and all standard equipment. It does not include the weight of any passengers, cargo, or optional equipment. Curb weight is considered the closest weight to the actual weight of the vehicle.

**Engine size:**

Engine size, also known as 'engine capacity' or 'engine displacement' refers to the total volume of the cylinders in the engine **.** Car engine sizes are normally specified in litres, which is rounded up to the nearest tenth of a litre. For example, a car that has a 1390cc engine would be described as a 1.4 litre.

**Bore:**

Its main functions are to act as a seal against the combustion gases as well as lubricant ingression into the combustion chamber. Another key function is to conduct the heat away to the piston and the liner. in a piston engine, the bore (or cylinder bore) is the diameter of each cylinder. Engine displacement is calculated based on bore, stroke length and the number of cylinders: **displacement = π ( 12 × bore )2 × stroke × ncylinders**

**compression ration:**

The compression ratio is defined as the ratio between the volume of the cylinder with the piston in the bottom position, Vbottom (largest volume), and in the top position, Vtop (smallest volume). The higher this ratio, the greater will be the power output from a given engine. It is generally in the 6–10 range.

**Horse power:**

Horsepower is a measurement used to calculate how fast force is produced from an engine of a vehicle and focuses on how quickly it can accomplish the work. ”Right between 200 and 300 horsepower is the sweet spot for many drivers. Be cautious with models that approach 300 horsepower, unless the vehicle is a heavy truck or another large model.

**City mpg :**

ity MPG: the score a car will get on average in city conditions, with stopping and starting at lower speeds. And, combined MPG, a combined average of highway and city MPG.

**Highway mpg:**

It is the average a car will get while driving on an open stretch of road without stopping or starting, typically at a higher speed.Generally, a good rule of thumb for gas mileage nowadays is to make sure you are getting at least 23 MPG combined city and highway. However, there are all kinds of factors that will affect the mileage you get and how efficient your vehicle will be.

**Price:**

the sum you pay to get a vehicle from a dealer is known as the On-Road Price. It comprises the ex-showroom price, registration, road tax, and insurance fees, as well as any other fees..

**4.3. ML Algorithms used In Project:**

A machine learning (ML) algorithm is a set of rules or instructions used by a computer program to learn from data and improve its performance on a given task. ML algorithms can be used to build predictive models, identify patterns in data, and make decisions based on data. These algorithms can be supervised, unsupervised, or semi-supervised depending on the nature of the learning process.

**KNN :**

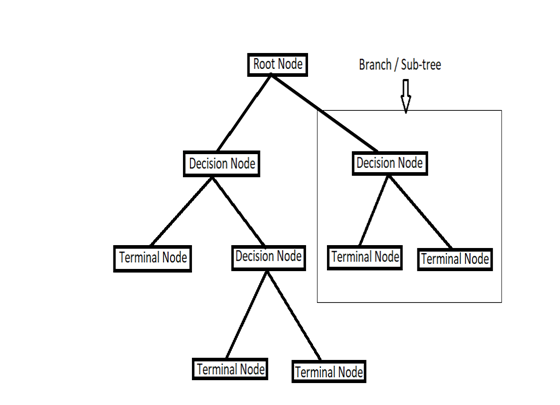
The k-nearest neighbours classifier (kNN) is a non-parametric supervised machine learning algorithm. It’s distance-based: it classifies objects based on their proximate neighbours’ classes. kNN is most often used for classification, but can be applied to regression problems as well.

**KNN Regression :**

KNN regression is a non-parametric method that, in an intuitive manner, approximates the association between independent variables and the continuous outcome by averaging the observations in the same neighbourhood.

**Decision Tree :**

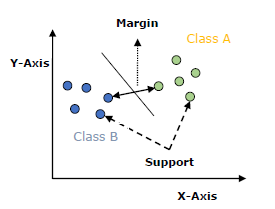
A decision tree is a non-parametric supervised learning algorithm, which is utilized for both classification and regression tasks. It has a hierarchical, tree structure, which consists of a root node, branches, internal nodes and leaf nodes.

****

A decision tree is a decision support tool that uses a tree-like model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements.

**Support Vector Machine(SVM) :**

Support Vector Machines (SVM) is a popular supervised machine learning algorithm used for classification and regression tasks. It is a type of discriminative classifier that works by finding the optimal decision boundary (or hyperplane) that separates the different classes in the input data.



**The followings are important concepts in SVM −**

•**Support Vectors** − Datapoints that are closest to the hyperplane is called support vectors. Separating line will be defined with the help of these data points.

•**Hyperplane** − As we can see in the above diagram, it is a decision plane or space which is divided between a set of objects having different classes.

•**Margin** − It may be defined as the gap between two lines on the closet data points of different classes. It can be calculated as the perpendicular distance from the line to the support vectors. Large margin is considered as a good margin and small margin is considered as a bad margin.

**Types of SVM :**

**SVM can be of two types:**

* **Linear SVM:**

Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is used called as Linear SVM classifier.

* **Non-linear SVM:** Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.

**Linear Regression and Logistic Regression :**

**Linear Regression**

Is a statistical approach to modeling the relationship between a dependent variable and one or

more independent variables. In simple linear regression, there is only one independent variable,

while in multiple linear regression, there are two or more independent variables.

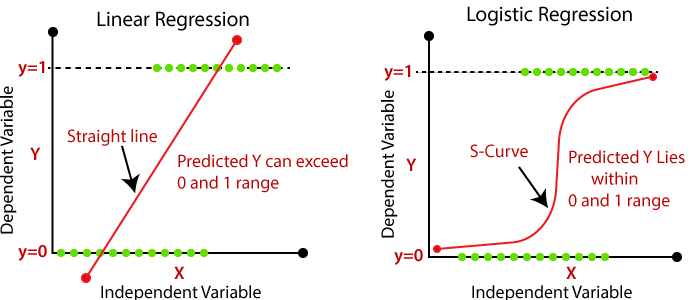
**y = β0 + β1x1 + β2x2 + … + βnxn + ε**

where y is the dependent variable, x1, x2, ..., xn are the independent variables, β0 is the intercept, β1, β2, ..., βn are the coefficients, and ε is the error term.

**Logistic Regression**

Is a statistical approach used to model the relationship between a binary dependent variable and one or more independent variables. The dependent variable can only take on two values, typically labeled as 0 and 1, representing the two classes.

**p = 1 / (1 + e^(-z))**

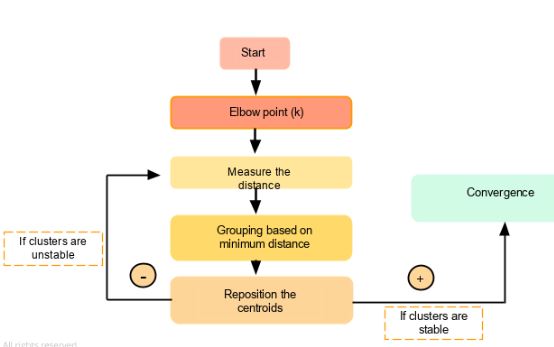


**K-Means Clustering :**

K-means clustering is a popular unsupervised machine learning algorithm used for clustering or grouping data points. The goal of K-means clustering is to divide a set of data points into K clusters, where each data point belongs to the cluster with the nearest mean. The value of K, the number of clusters, is determined by the user.

**The algorithm works as follows**

* Select K random points from the dataset to act as the initial centroids for each cluster.
* Assign each data point to the closest centroid, forming K clusters.
* Recalculate the mean of each cluster, moving the centroid to the center of the cluster.
* Repeat steps 2 and 3 until convergence, meaning the centroids no longer move or the maximum number of iterations is reached.



K-means clustering is an iterative algorithm that aims to minimize the sum of squared distances between each data point and its assigned centroid.

5. SDLC

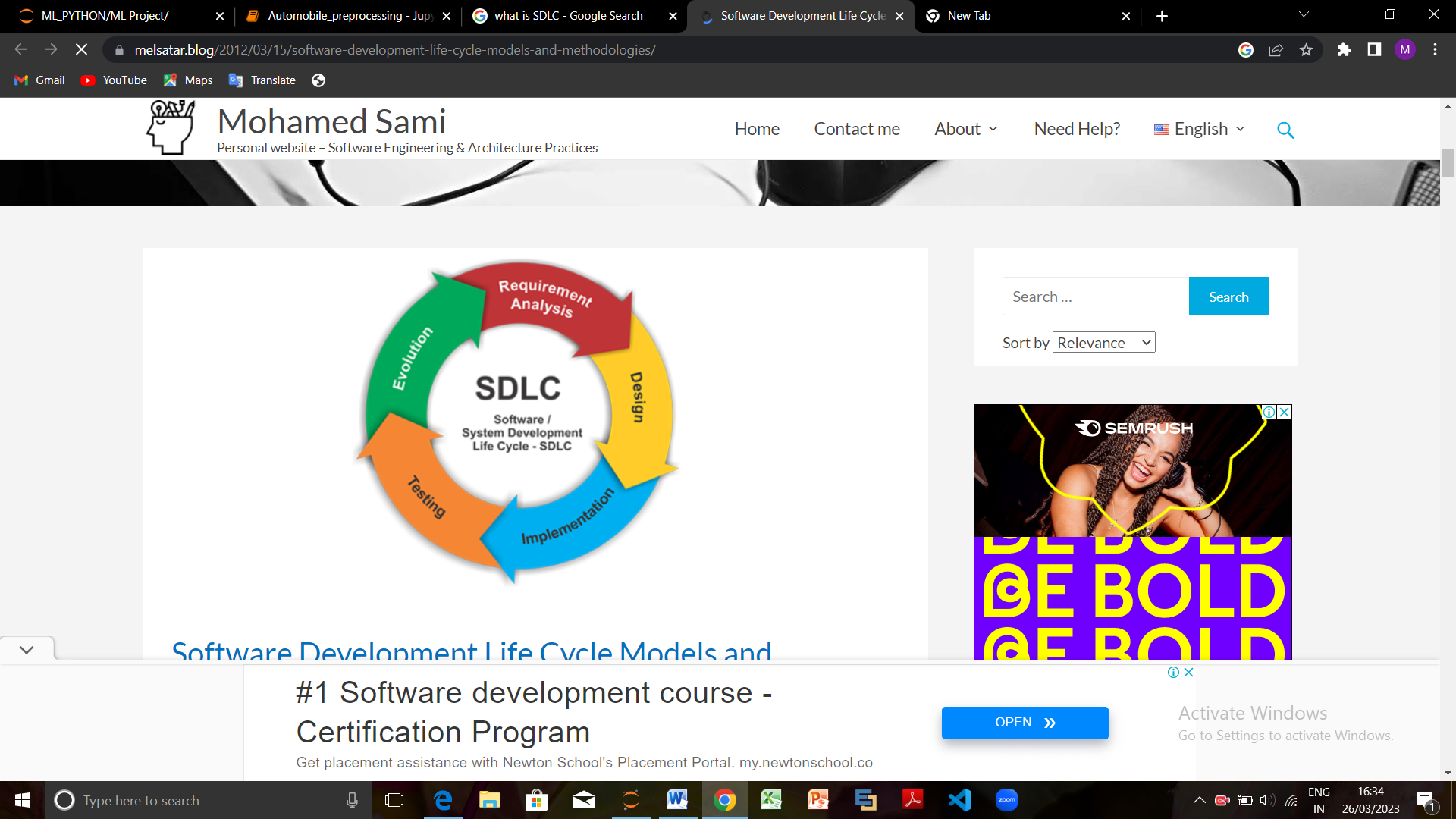
**Software development life cycle:**

A software life cycle model (also termed process model) is a pictorial and diagrammatic representation of the software life cycle. A life cycle model represents all the methods required to make a software product transit through its life cycle stages. It also captures the structure in which these methods are to be undertaken.

The Software Development Life Cycle (SDLC) is a structured process that enables the production of high-quality, low-cost software, in the shortest possible production time. The goal of the SDLC is to produce superior software that meets and exceeds all customer expectations and demands. The SDLC defines and outlines a detailed plan with stages, or phases, that each encompass their own process and deliverables. Adherence to the SDLC enhances development speed and minimizes project risks and costs associated with alternative methods of production.

**Why is the SDLC important?**

* It provides a standardized framework that defines activities and deliverables
* It aids in project planning, estimating, and scheduling
* It makes project tracking and control easier
* It increases visibility on all aspects of the life cycle to all stakeholders involved in the development process
* It increases the speed of development
* It improves client relations
* It decreases project risks
* It decreases project management expenses and the overall cost of production



**5.1. Evolution:**

**Data preprocessing:**

Data preprocessing transforms the data into a format that is more easily and effectively processed in data mining, machine learning and other data science tasks. The techniques are generally used at the earliest stages of the machine learning and AI development pipeline to ensure accurate results.

**Importing libraries for data preprocessing:**

**Import pandas as pd:**

Pandas is an open-source, BSD-licensed library writtenin Python Language. Pandas provide high performance, fast, easy-to-use data structures, and data analysis tools for manipulating numeric data and time series. Pandas is built on the numpy library and written in languages like Python, Cython, and C. In pandas, we can import data from various file formats like JSON, SQL, Microsoft Excel, etc.

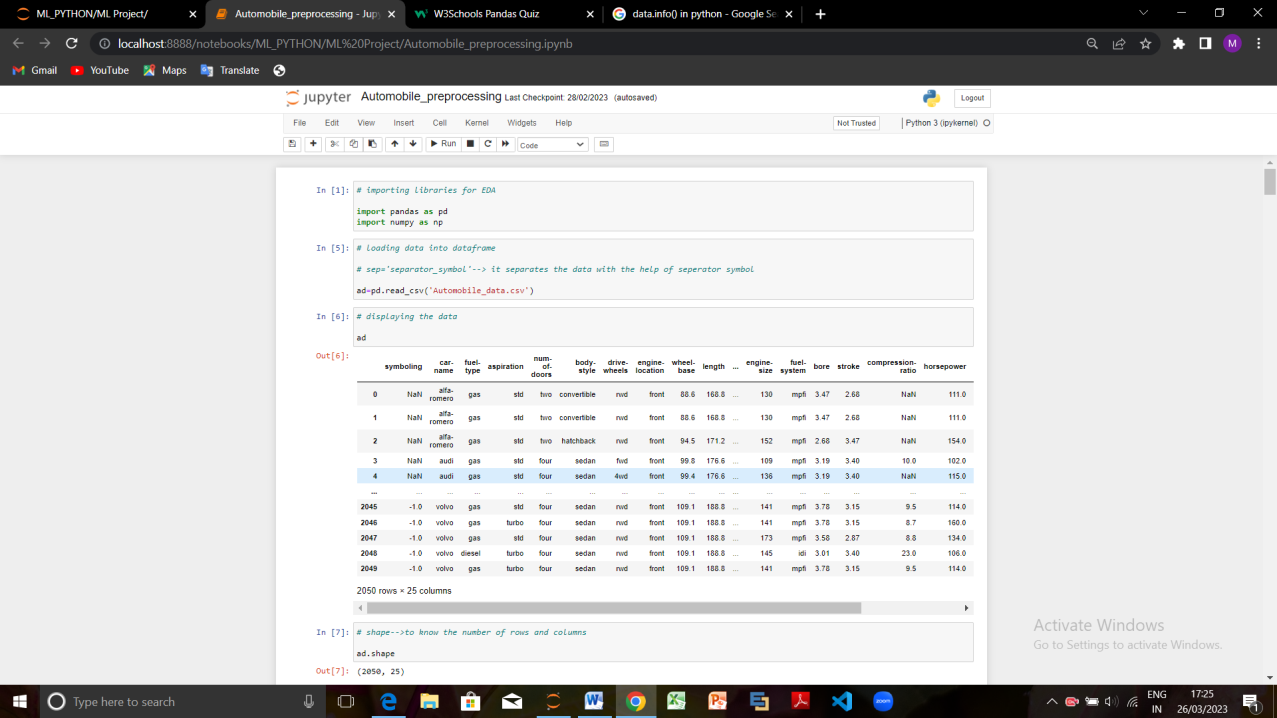
**Import numpy as np**

It is the fundamental library of python, used to perform scientific computing. It provides high-performance multidimensional arrays and tools to deal with them. A numpy array is a grid of values (of the same type) that are indexed by a tuple of positive integers, numpy arrays are fast, easy to understand, and give users the right to perform calculations across arrays

**Data=pd.read\_csv(‘automobile\_data.csv’):** loding data into dataframe

**Data.shape():**

the shape() method is used to fetch the dimensions of Pandas and NumPy type objects in python. Every value represented by the tuple corresponds to the actual dimension in terms of array or row/columns.

.

**Data.describe():**

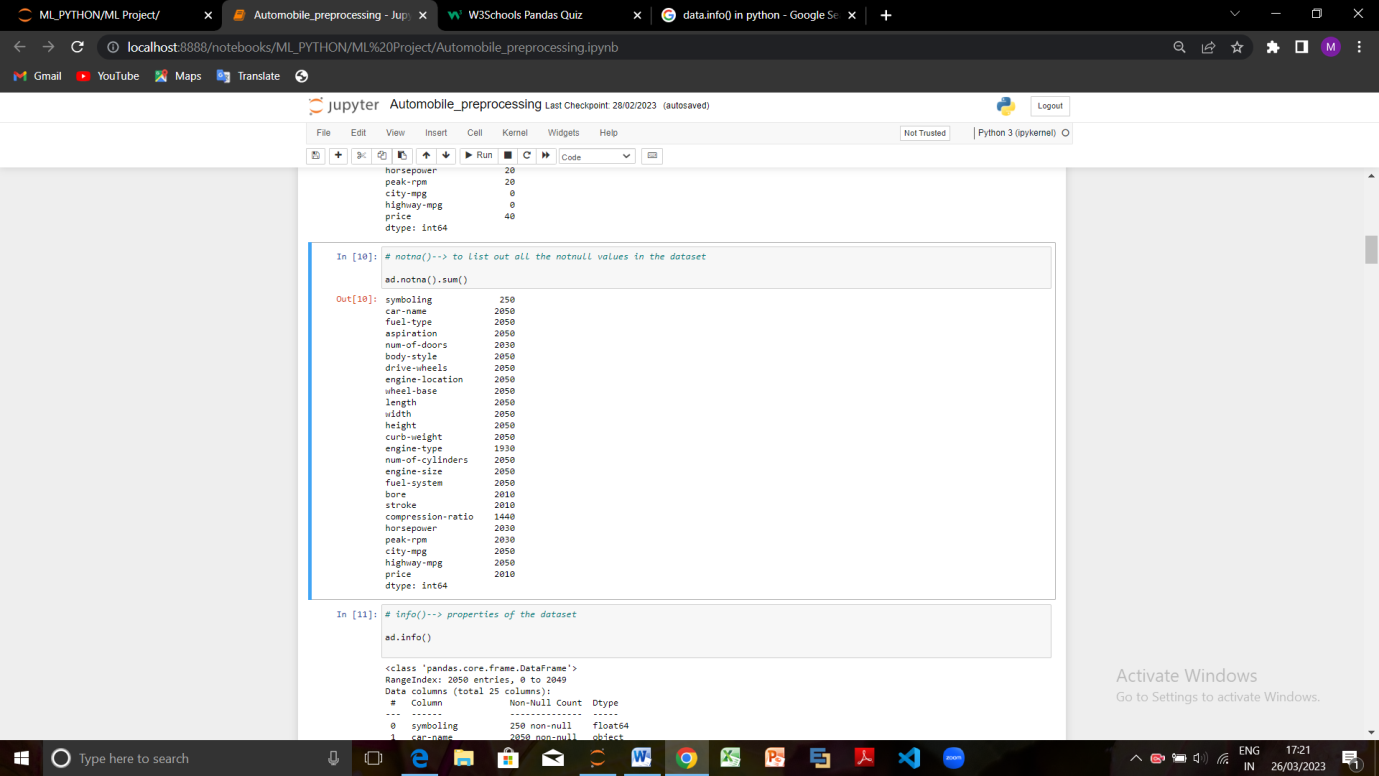
The describe() method returns description of the data in the DataFrame. column: count - The number of not-empty values. mean - The average (mean) value.

**Data.isna().sum():**

This returns the columns in our Pandas dataframe along with the number of missing values detected in each one, so 0 means there are no missing values, and 1 means there is a single missing value.

**Data.notna().sum():**

The notna() method returns a DataFrame object where all the values are replaced with a Boolean value True for NOT NA (not-a -number) values, and otherwise False.



**5.2.Requirement Analysis:**

**5.2.1.Software requirement:**

* **Anaconda:**

Anaconda is the Open Source data science distribution for the Python and R programming languages scientific computing, such as data science, machine learning applications, predictive analysis, large-scale data processing, etc. The Anaconda can be used to simplify package management and deployment. It includes more than 300 data science packages that are suitable for Windows, Linux, and MacOS.

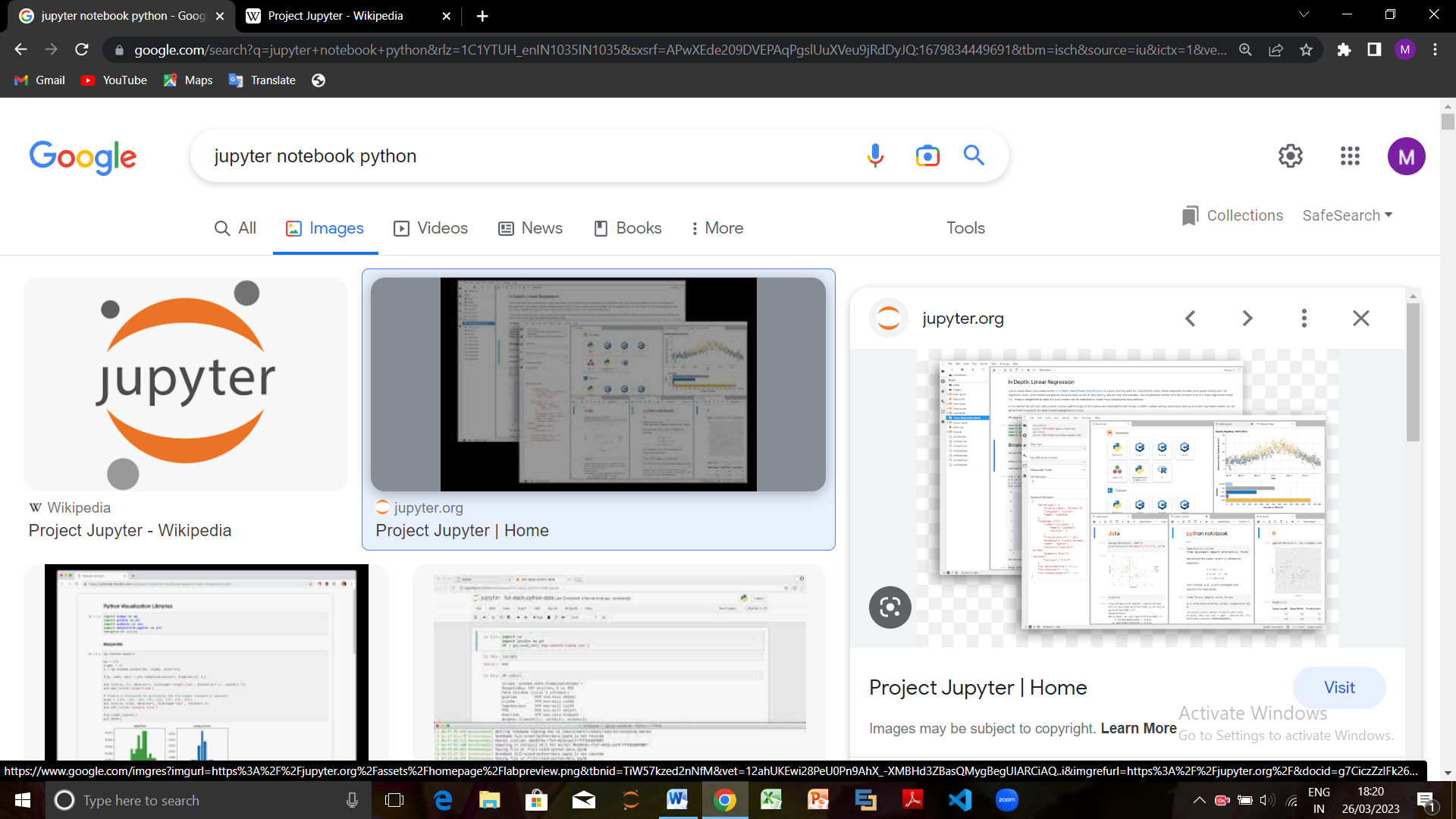


**Anaconda will be available in three different editions, which are:**

* **Individual Edition:** This edition was developed for Solo practitioners.
* **Team Edition:** The team edition has been developed to work with the team on the same page.
* **Enterprise Edition:** The Enterprise edition has been developed to use data science and machine learning to make a better decision.
* **Jupyter Notebook:**

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

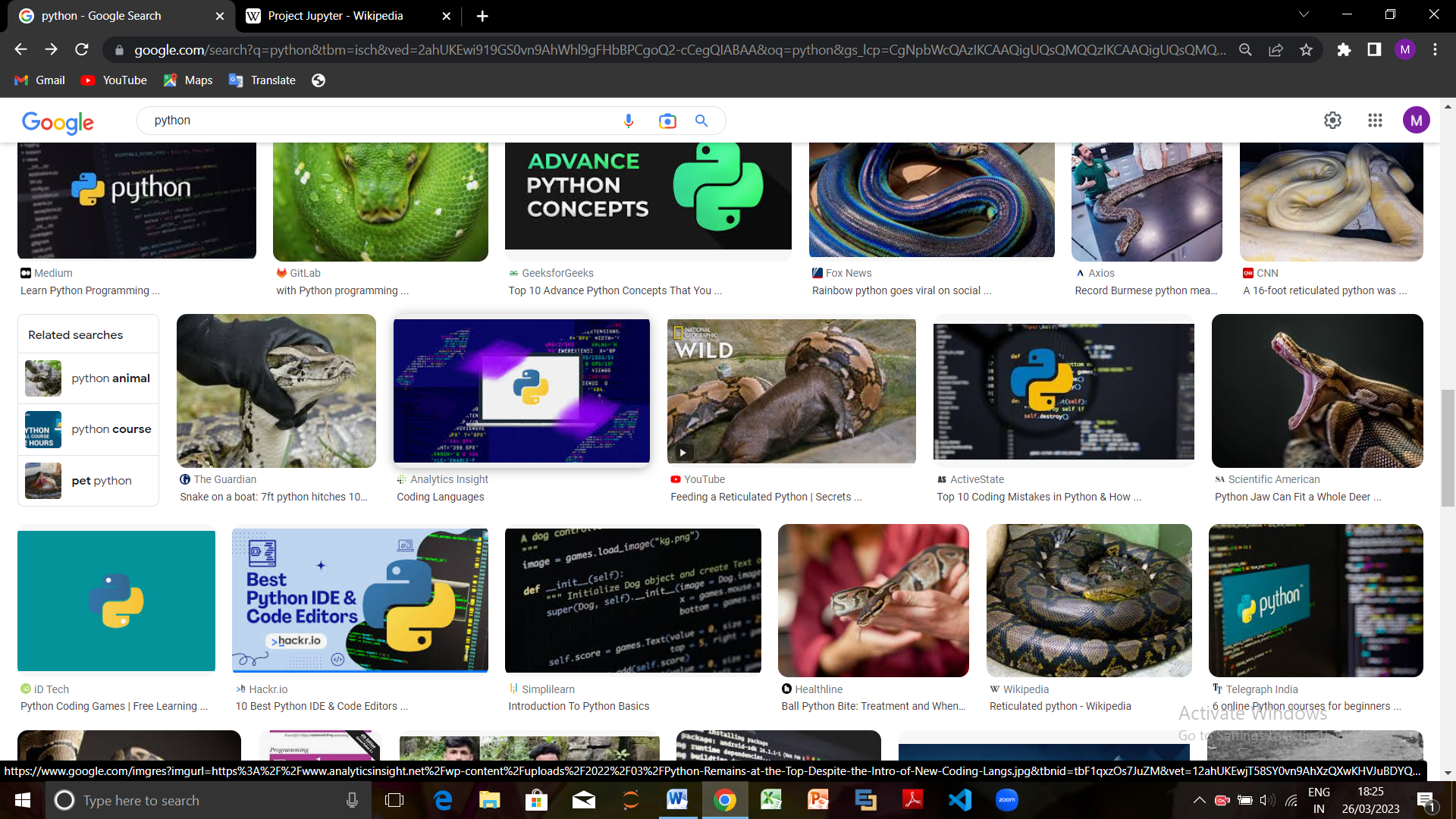
Jupyter has support for over 40 different programming languages and Python is one of them. Python is a requirement (Python 3.3 or greater, or Python 2.7) for installing the Jupyter Notebook itself.



* **Programming Language:**

Python is the most popular and widely-used programming language for machine learning. An open-source, Python is a dynamic, high-level, interactive, multi-paradigm, scripting, object-oriented, high-level, procedural language.

Python is the most used language for Machine Learning (which lives under the umbrella of AI). One of the main reasons Python is so popular within AI development is that it was created as a powerful data analysis tool and has always been popular within the field of big data.

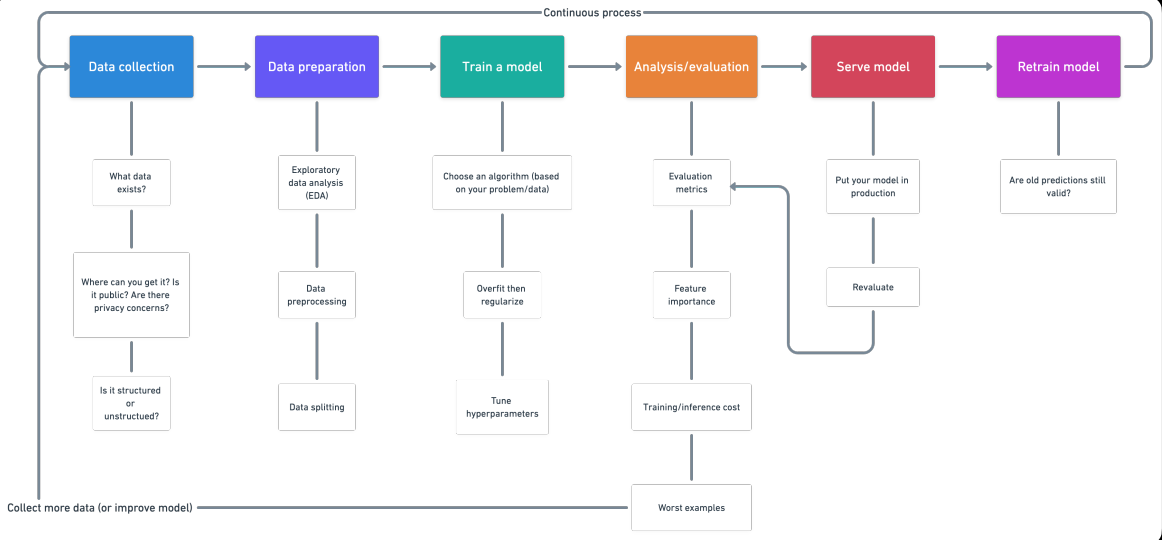


**5.2.2.Hardware Requirement:**

* **Processor (CPU):** A high-performance CPU is recommended for most ML projects, especially if the data set is large or the model is complex. A multicore CPU with a clock speed of at least 2 GHz is a good starting point.

* **Graphics Processing Unit (GPU):** A GPU can significantly speed up the training process for deep learning models. NVIDIA GPUs are most commonly used for this purpose, and a dedicated GPU with at least 4GB of VRAM is recommended.
* **Random Access Memory (RAM):** The amount of RAM needed depends on the size of the data set being used. As a general rule, you should have at least 8GB of RAM for small to medium-sized data sets, and 16GB or more for larger data sets.
* **Storage:** You'll need plenty of storage to store the data sets, code, and any intermediate results generated during the training process. An SSD is recommended for faster access times.
* **Operating System:** Most ML projects are developed on Linux, but Windows and macOS can also be used.

**5.3.Design**

****

* **Data Collection:** Collect the relevant data for your problem. This could involve scraping data from websites, using APIs to collect data, or manually entering data.
* **Data Cleaning and Preprocessing**: Once you have collected the data, you need to clean and preprocess it. This could involve removing missing values, outliers, and duplicates. You may also need to transform the data into a suitable format for machine learning algorithms.
* **Model Training:** Train the chosen ML model on the preprocessed data.
* **Model Evaluation:** Evaluate the performance of the trained model using appropriate metrics such as accuracy, precision, recall, and F1-score. You may also want to use techniques like cross-validation to ensure that your model is not overfitting.
* **Hyperparameter Tuning:** Once you have evaluated the performance of your model, you may want to fine-tune its hyperparameters to further improve its performance.

**5.4. Implementation**

Implementing a project involves turning the design and plan into a working product or solution. The following are some steps that can help you in implementing a project:

**5.4.1. Exploratory Data Analysis (EDA) :**

Exploratory Data Analysis, is the process of analyzing and understanding the structure and characteristics of a dataset before applying any statistical models or machine learning algorithms.

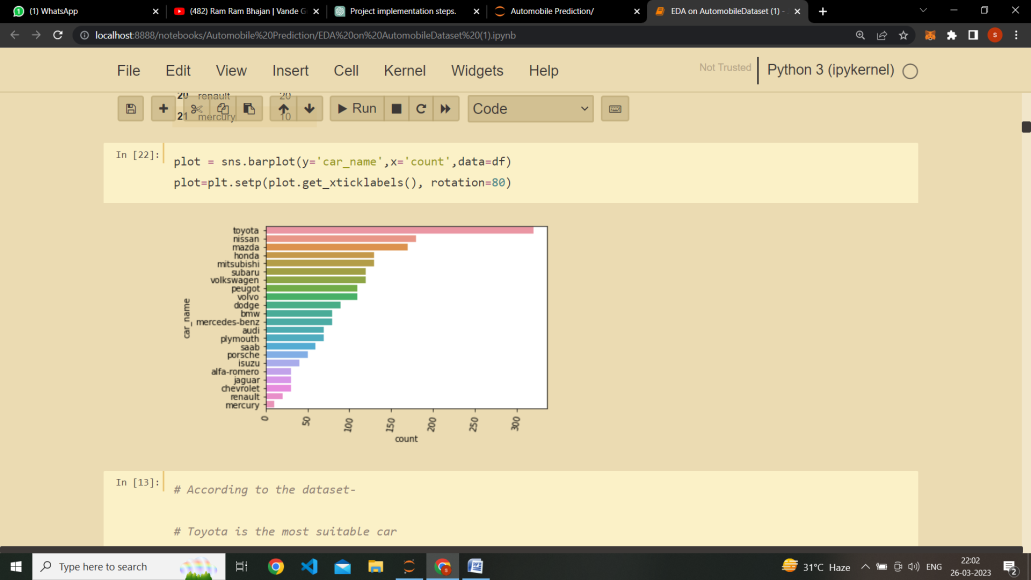
For EDA following libraries are important :

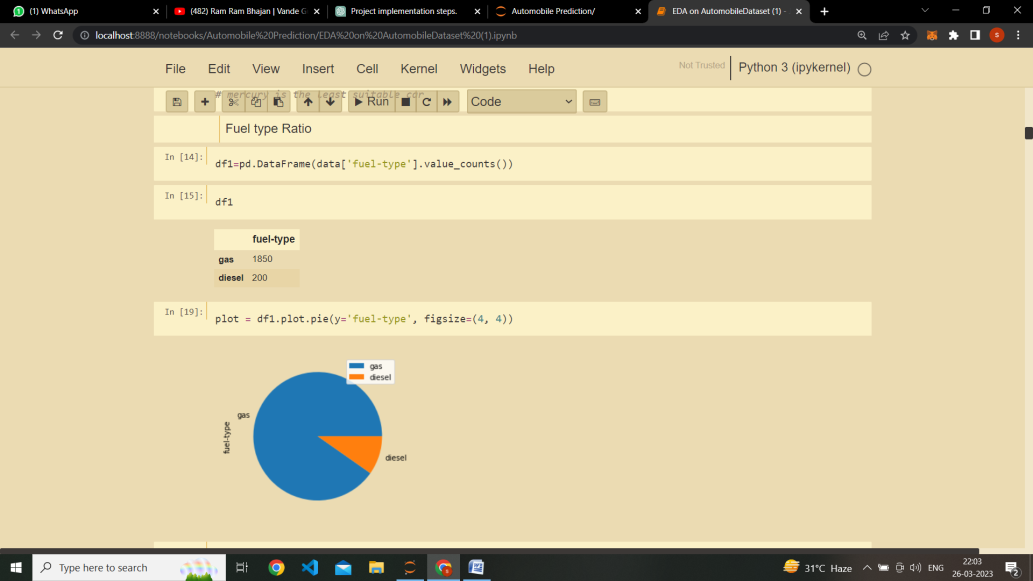
**import pandas as pd**

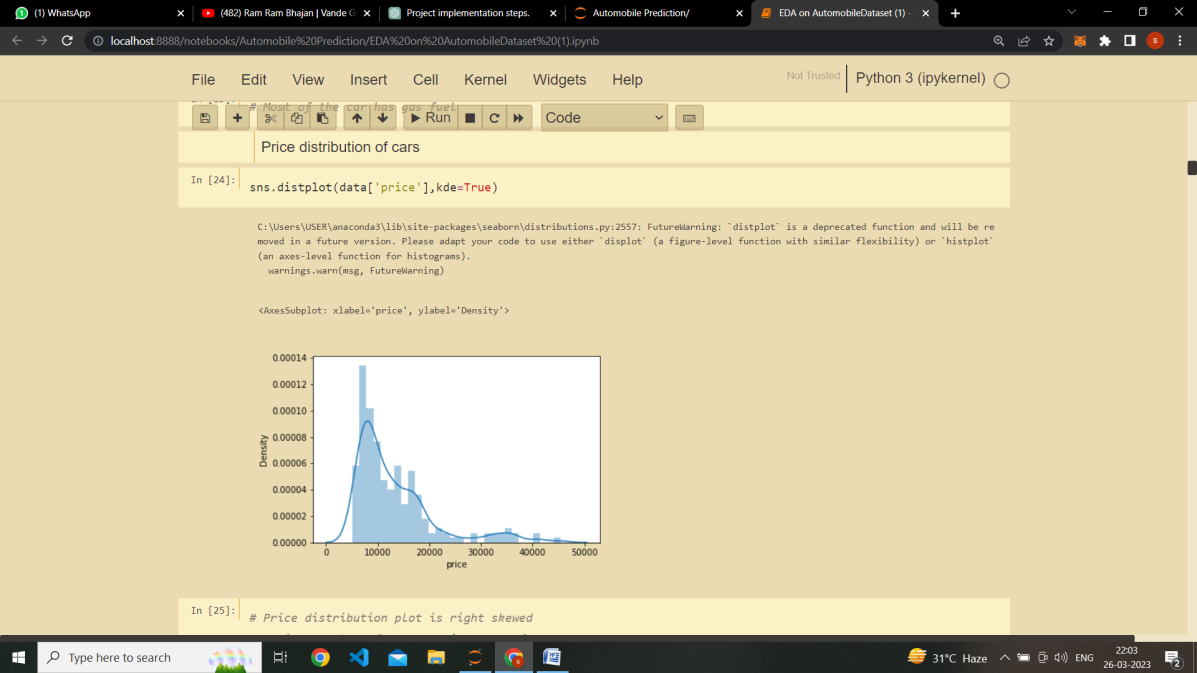
**import numpy as np**

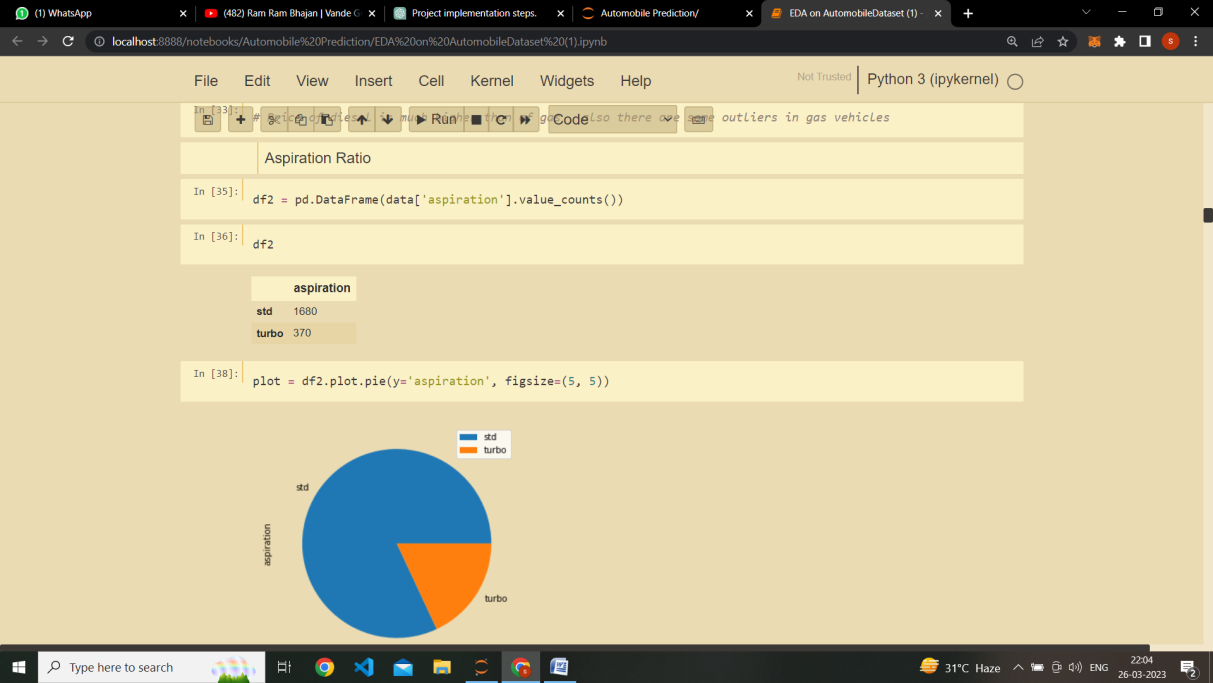
**import seaborn as sns**

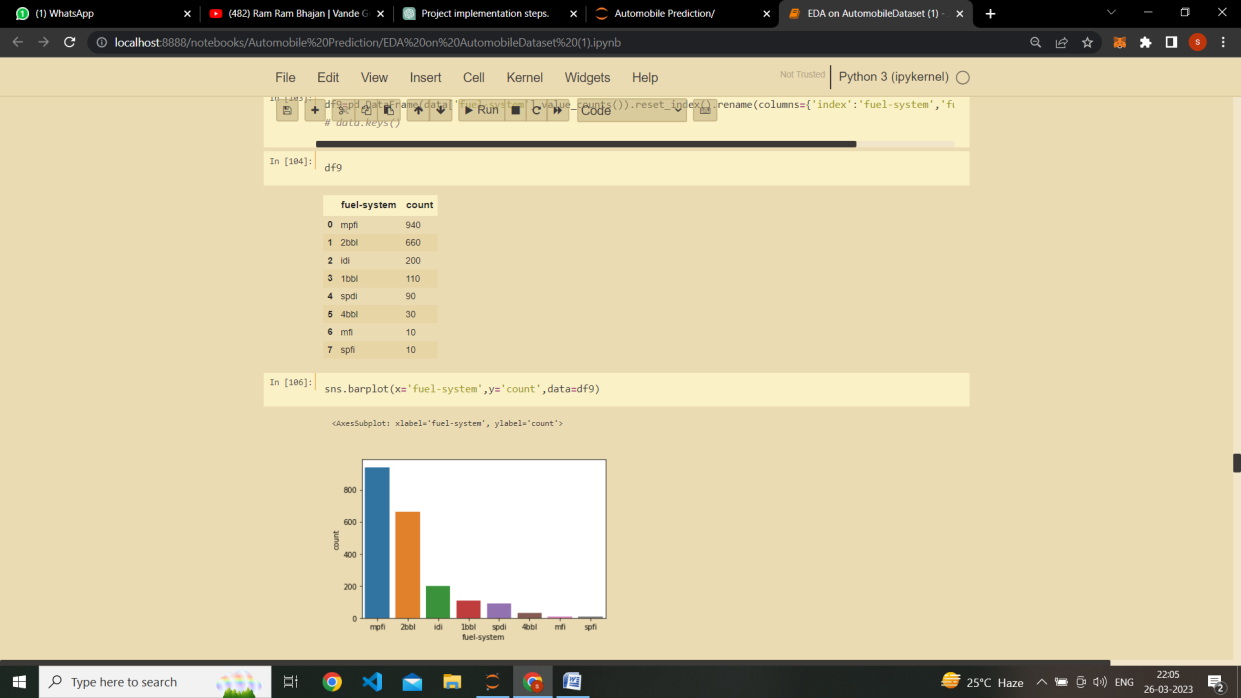
**import matplotlib.pyplot as plt**

****

****

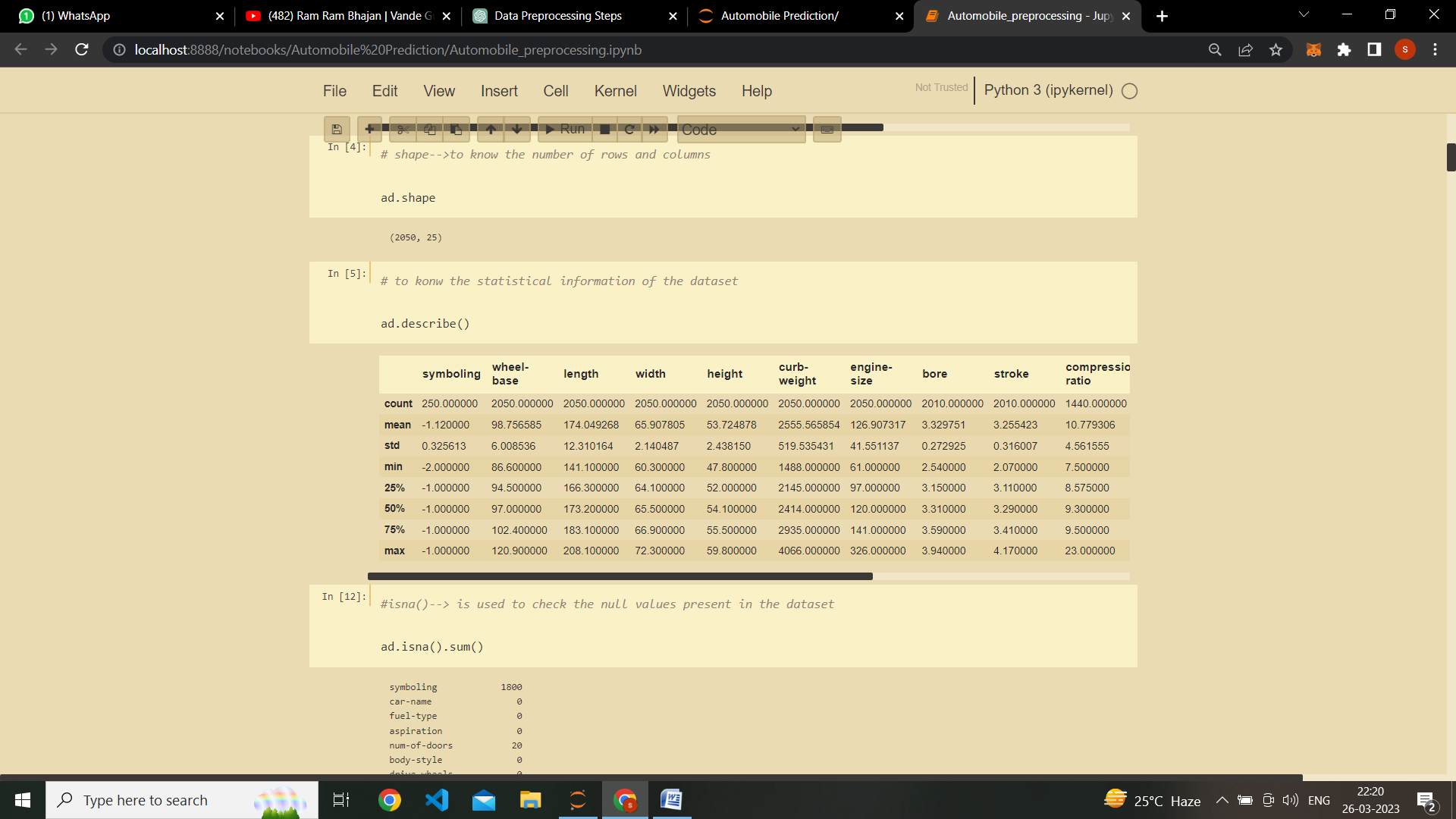
****

****

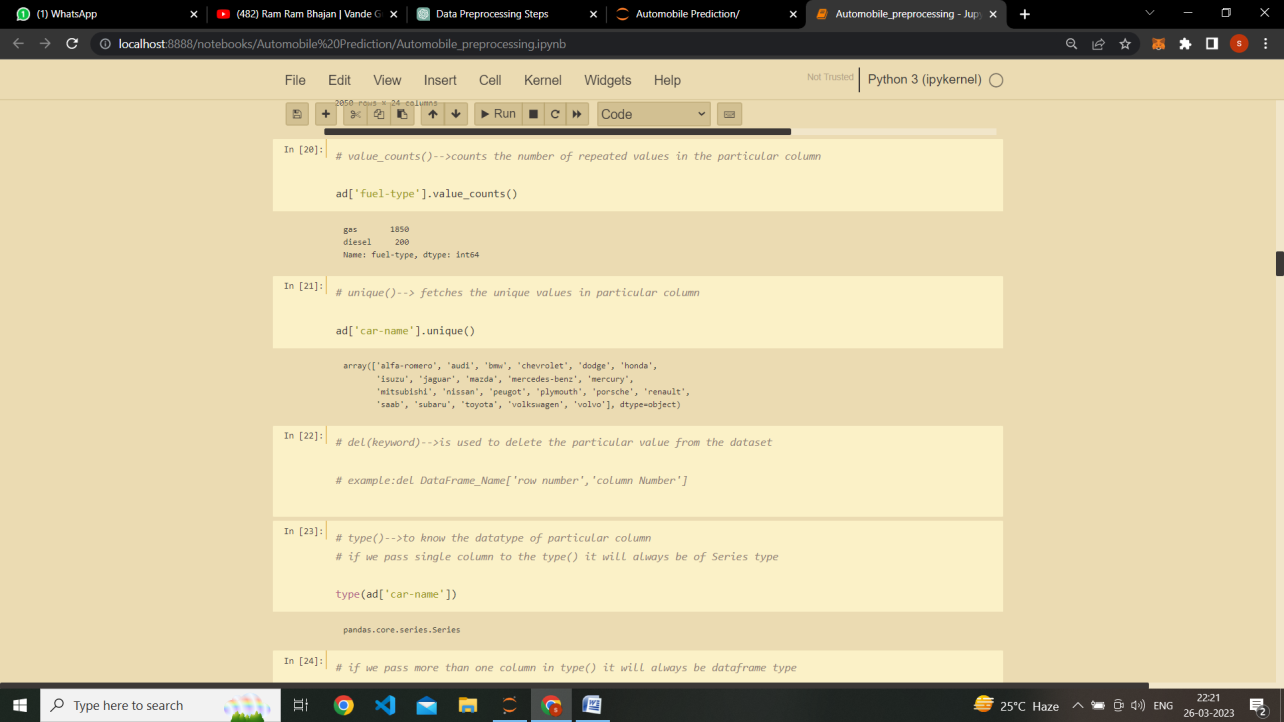
****

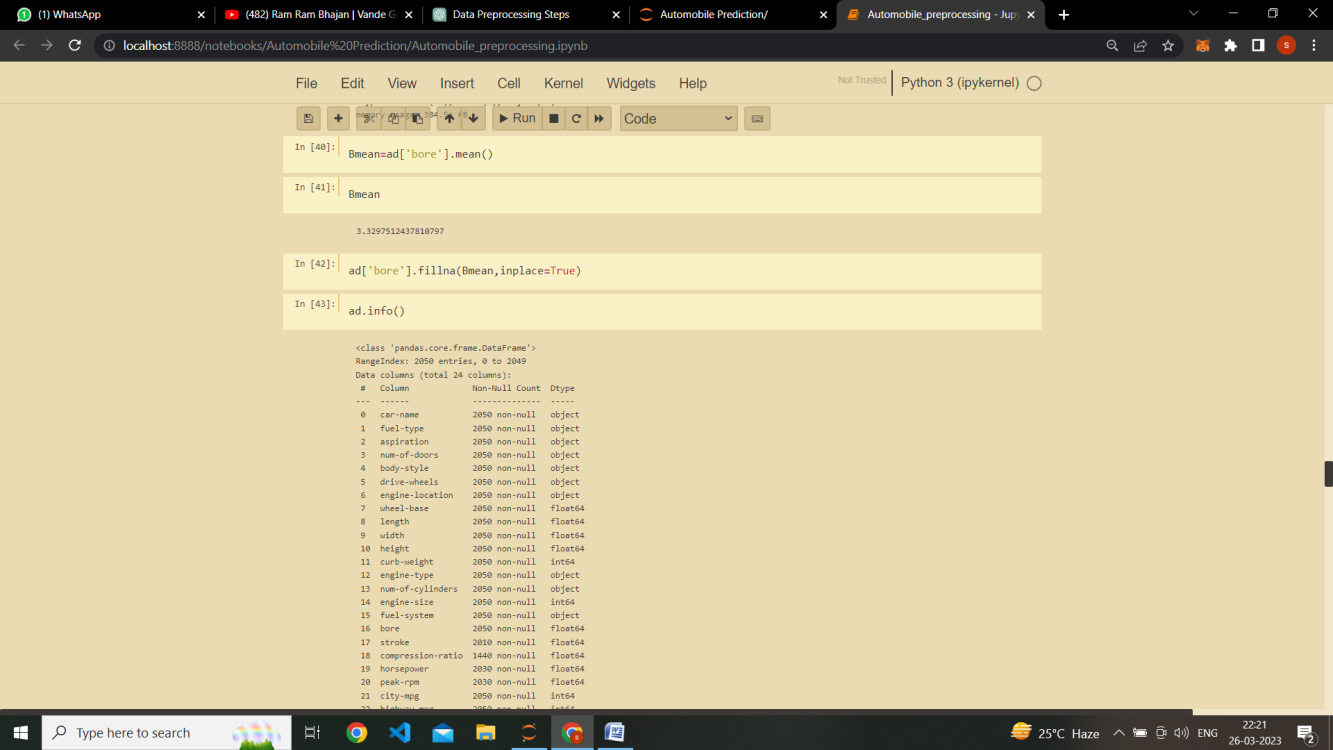
**5.4.2. Data Preprocessing :**

Data preprocessing refers to the process of cleaning, transforming, and preparing raw data for analysis. It is an important step in the data science pipeline as it helps to ensure that the data used for analysis is accurate, complete, and in the right format.









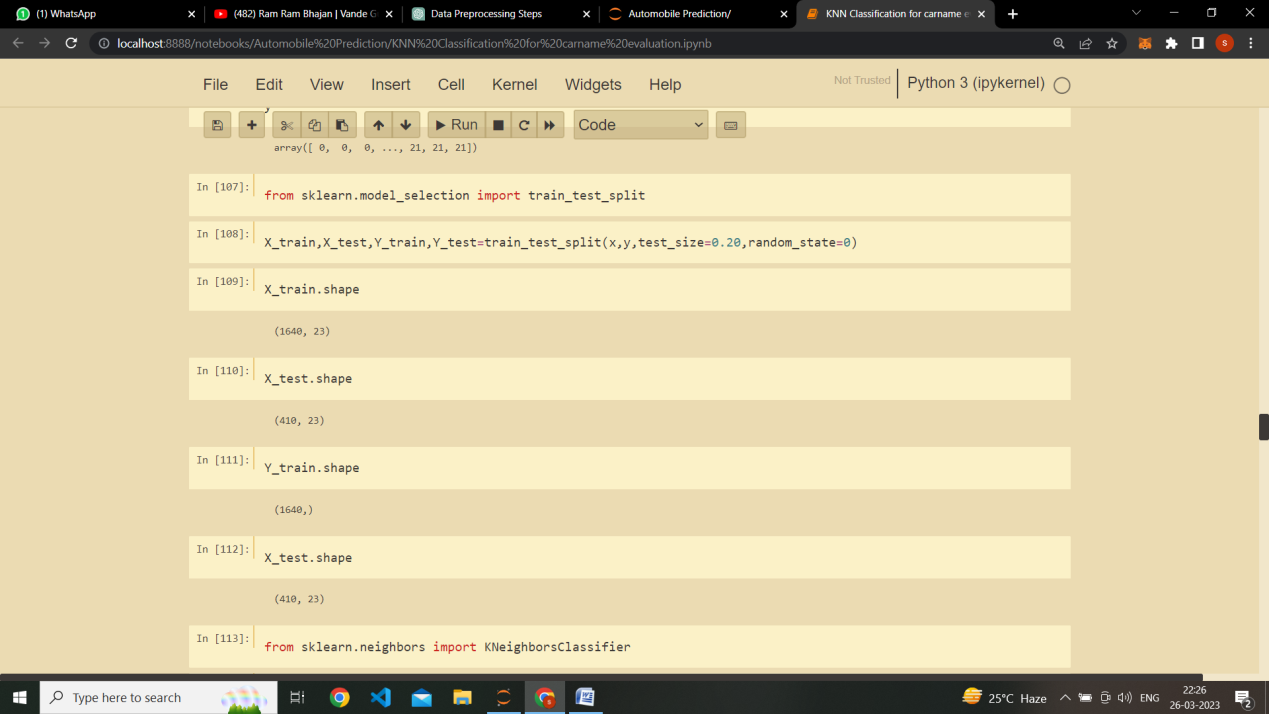
**5.4.3. Training the model :**

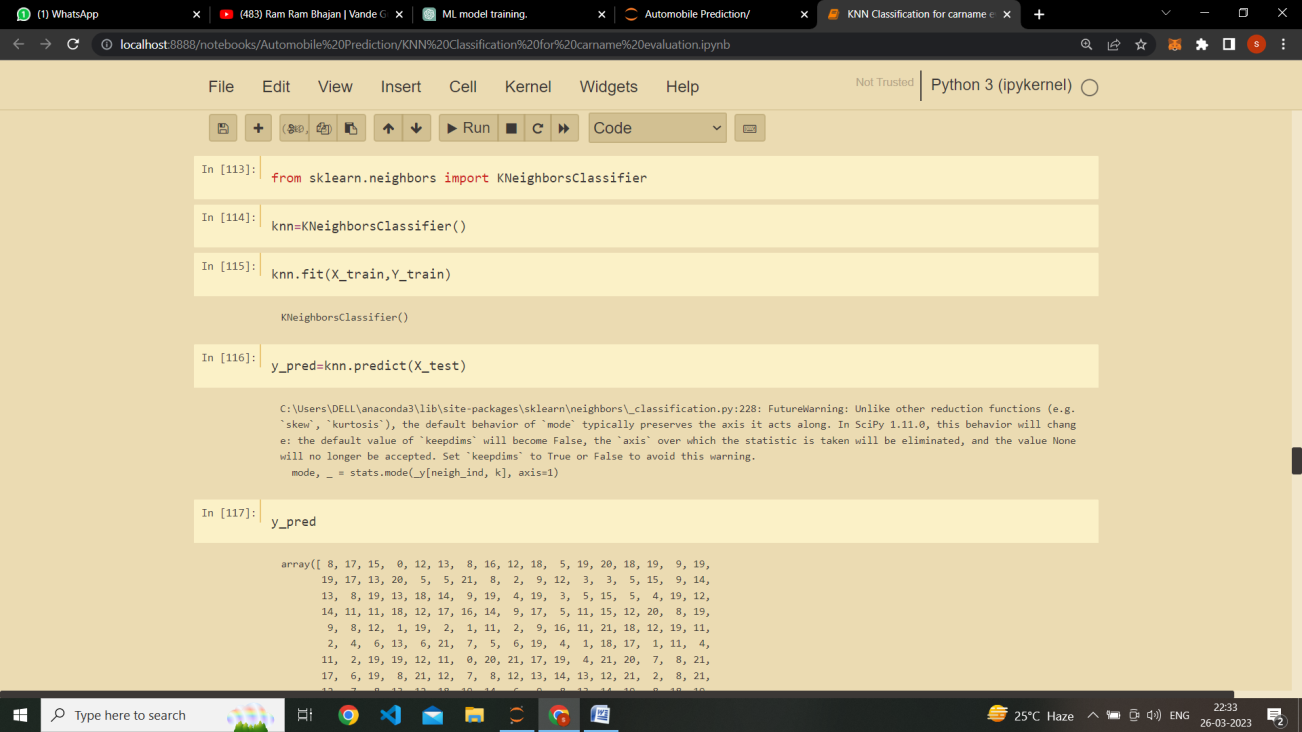
The training process typically involves breaking the data into two parts:

a **training set** and a **testing set.**

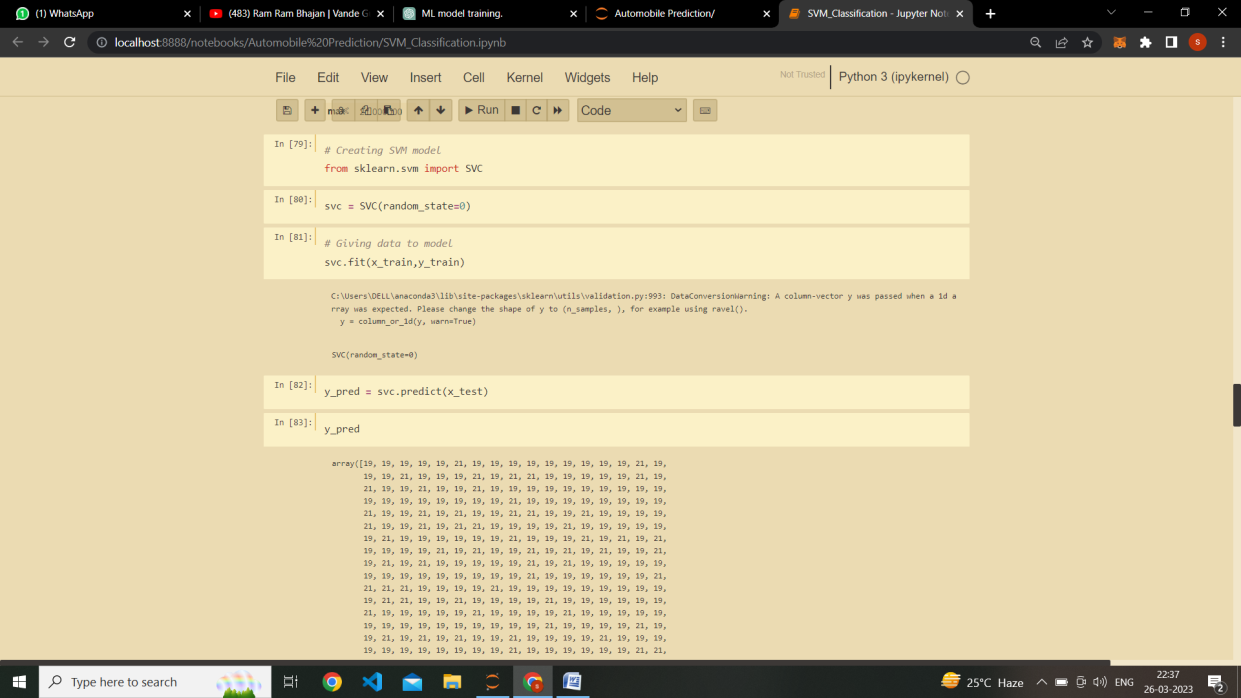
By using from **sklearn.model\_selection import train\_test\_split** python library.

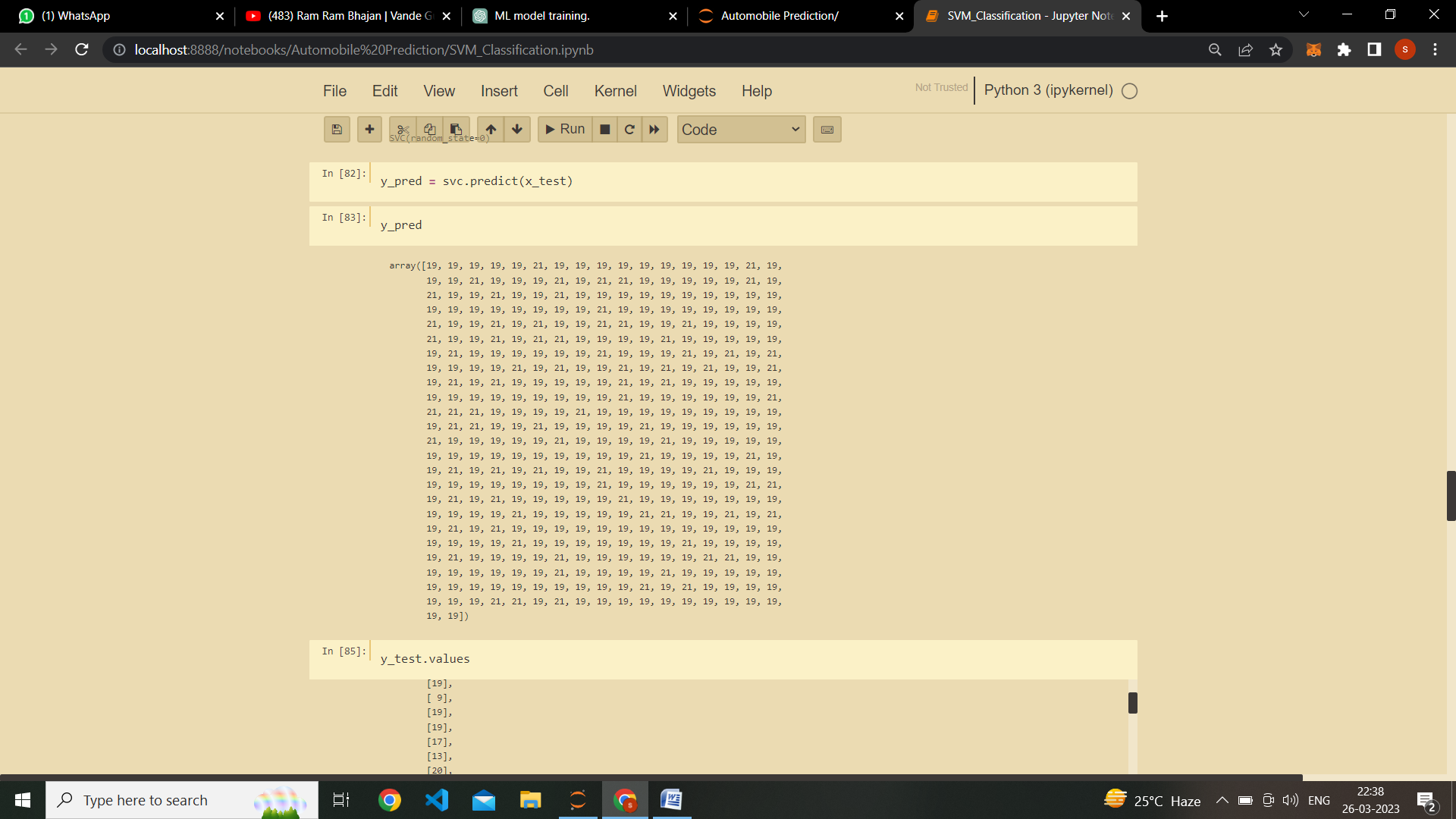
1. **KNN Classification Implementation :**

****

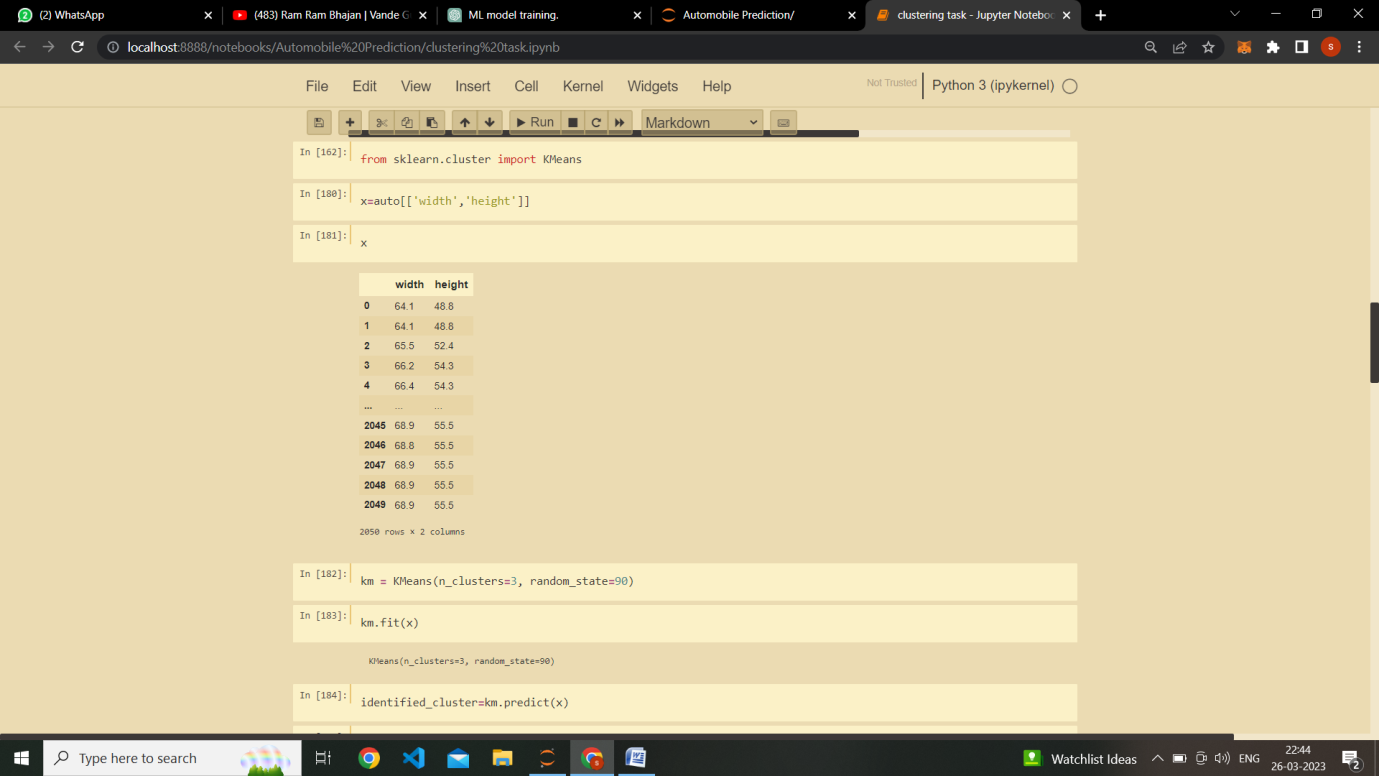
****

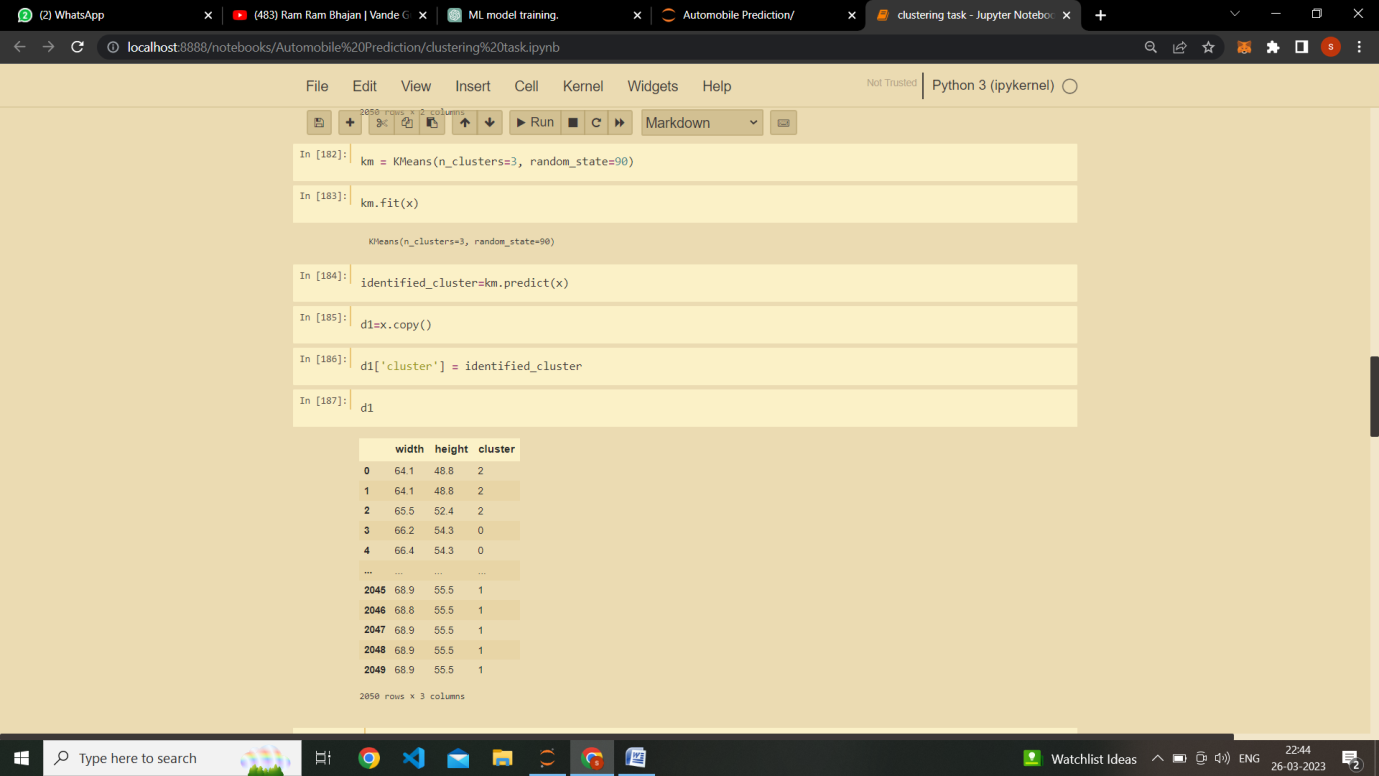
1. **SVM (Support Vector Machine) Implementation :**

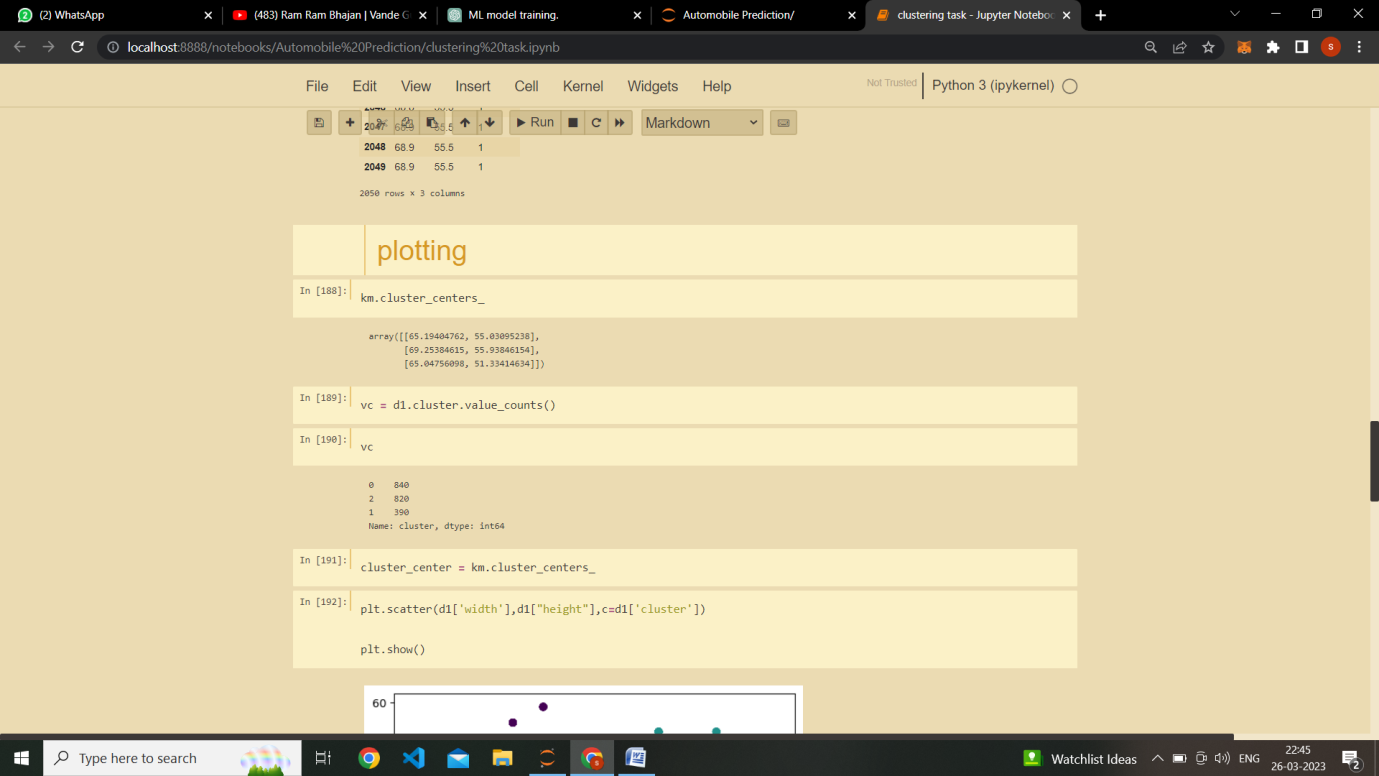
****

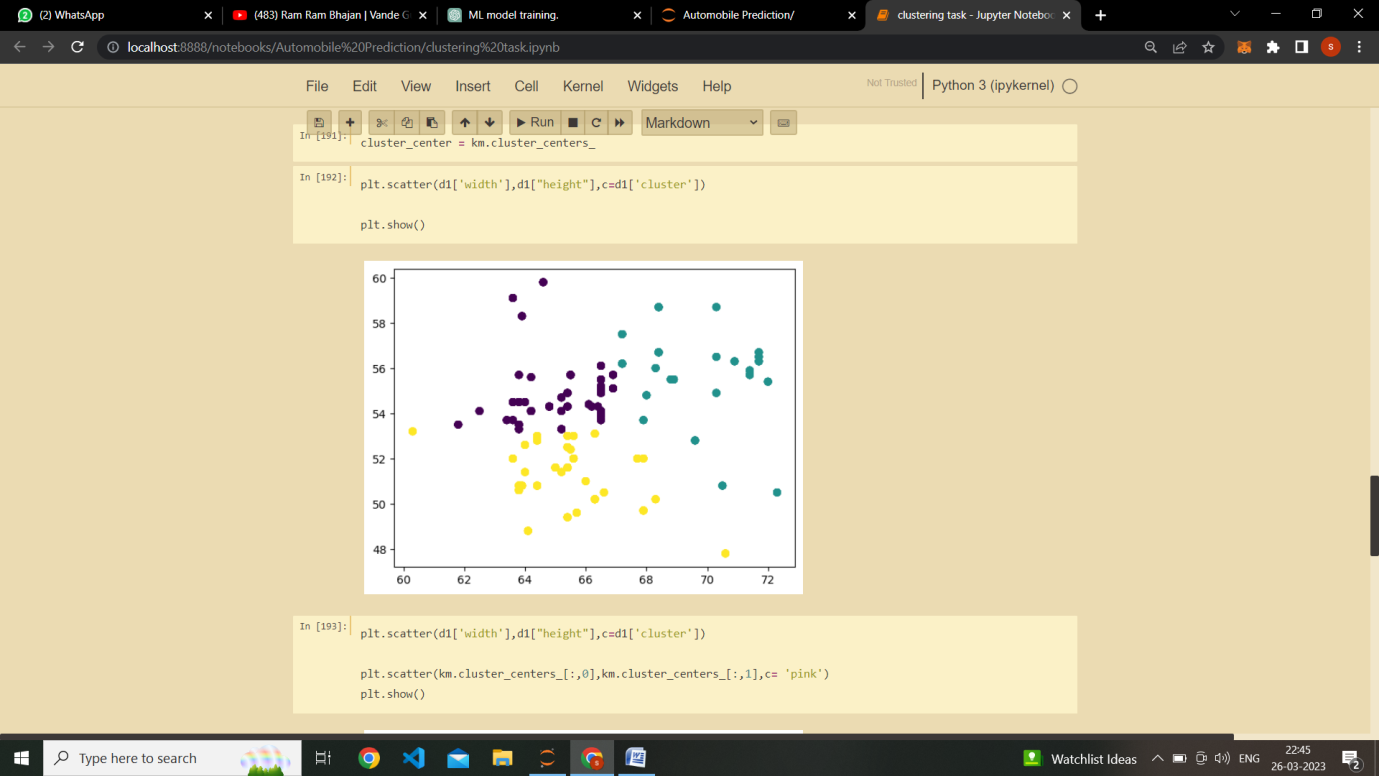
****

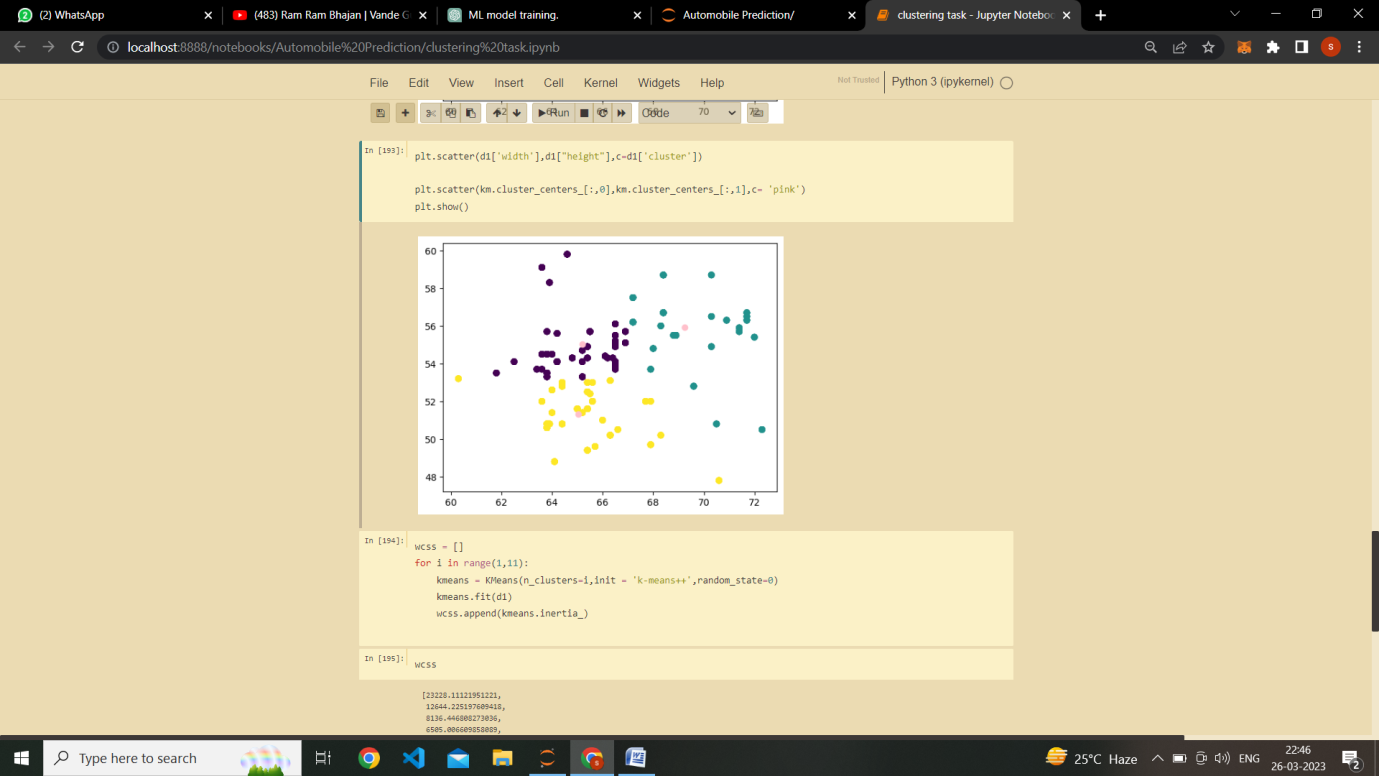
**3. K-Means Clustering Implementation :**

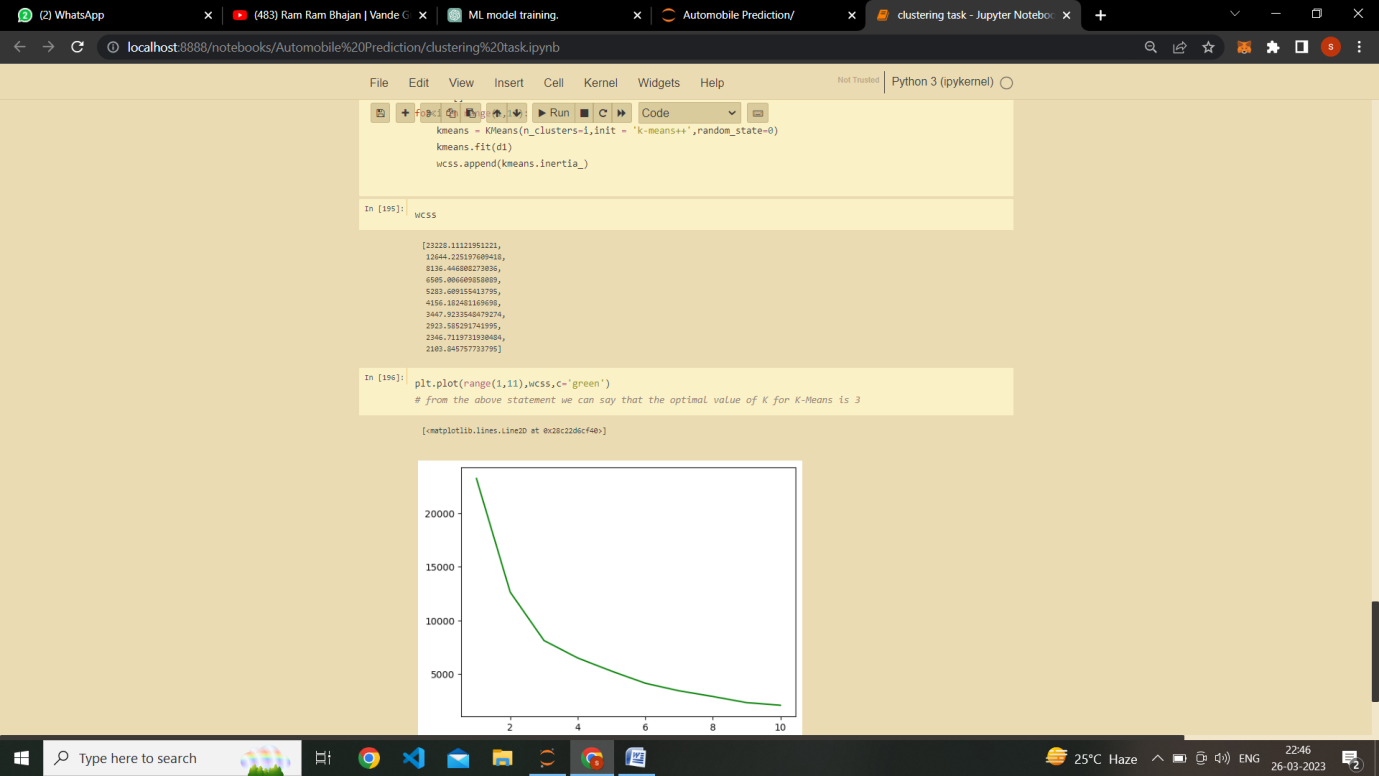
****

****

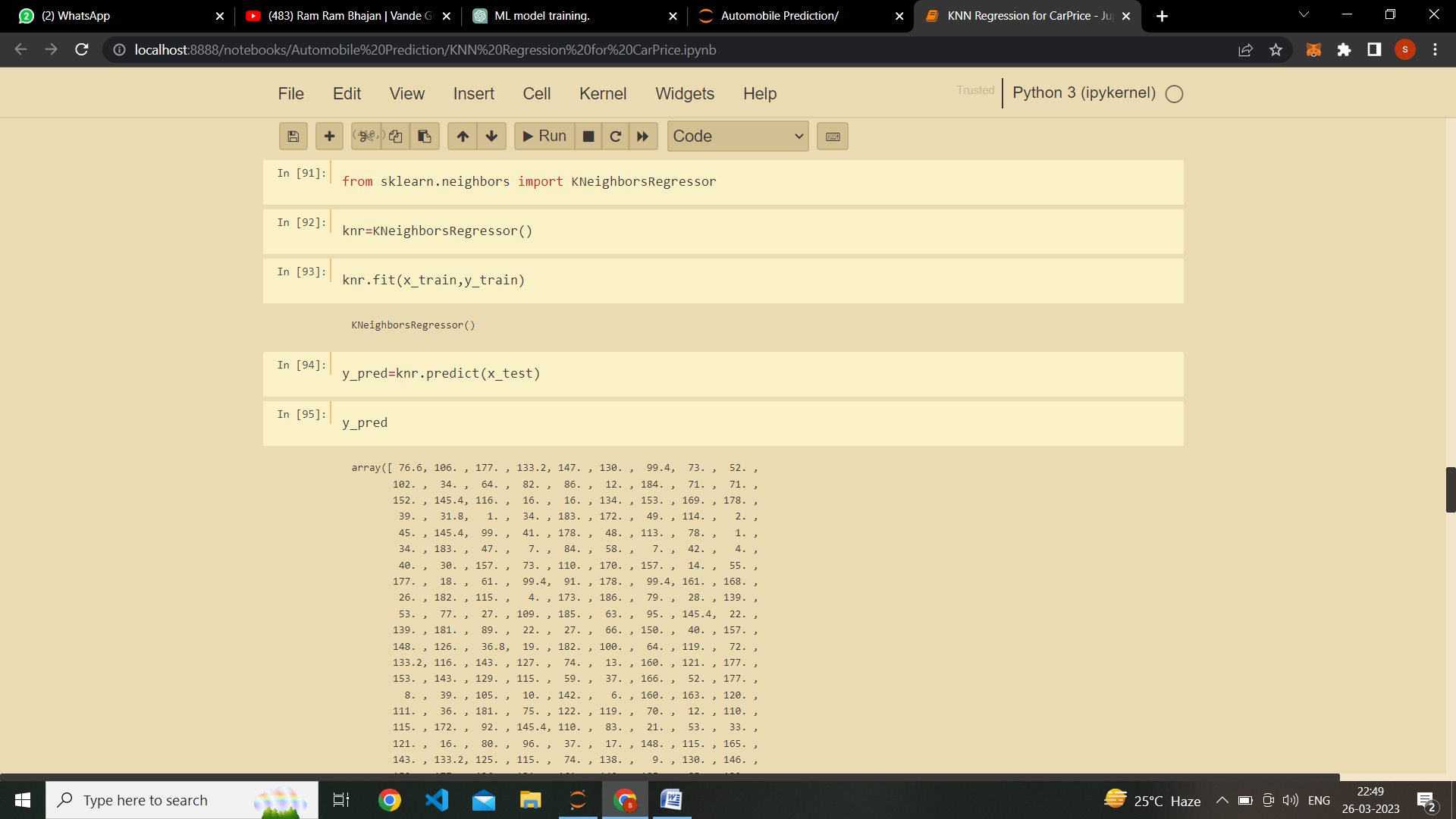
****

****

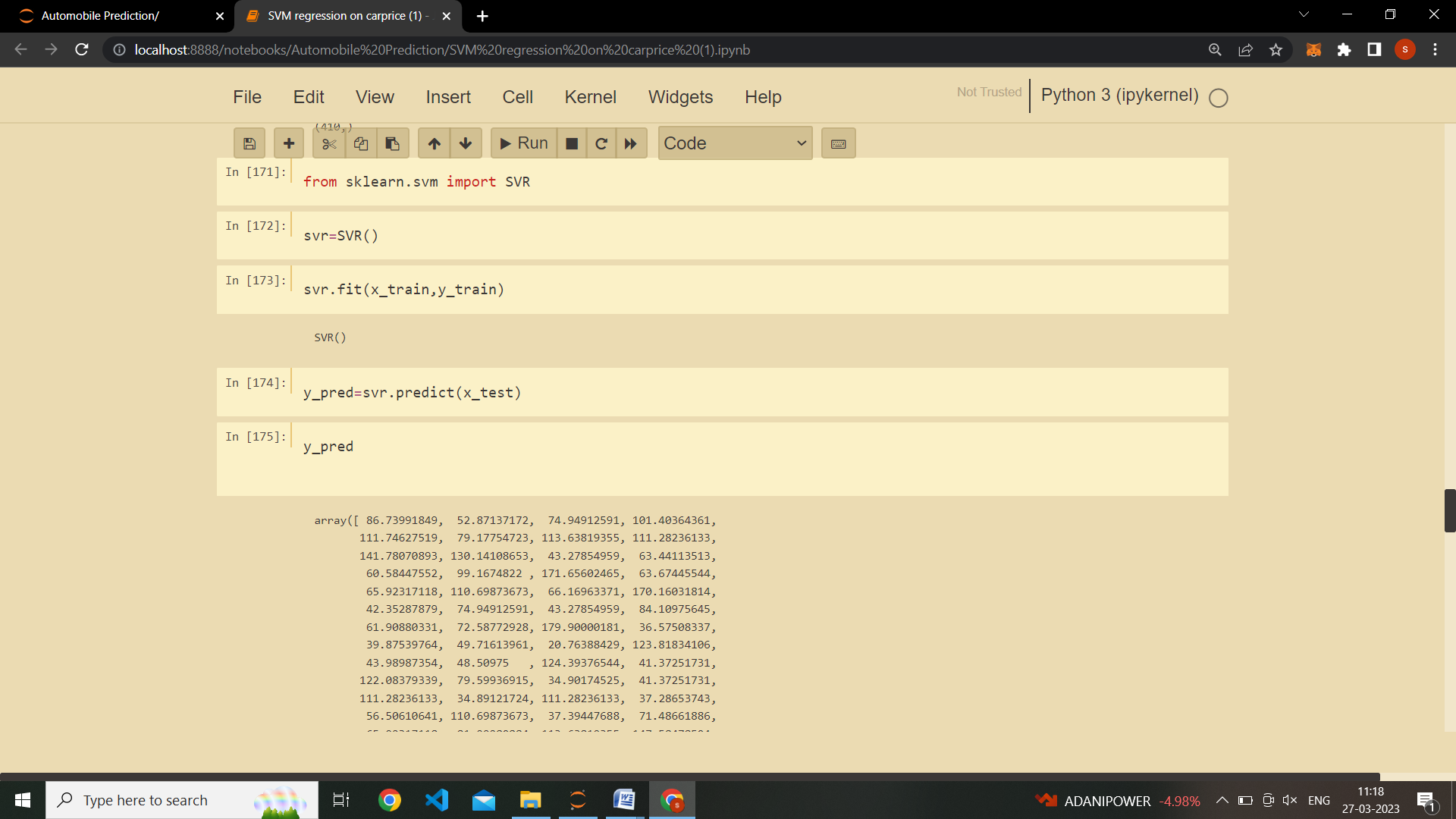
****

****

**4 . KNN Regression :**

****

**5. SVM Regression :**

****

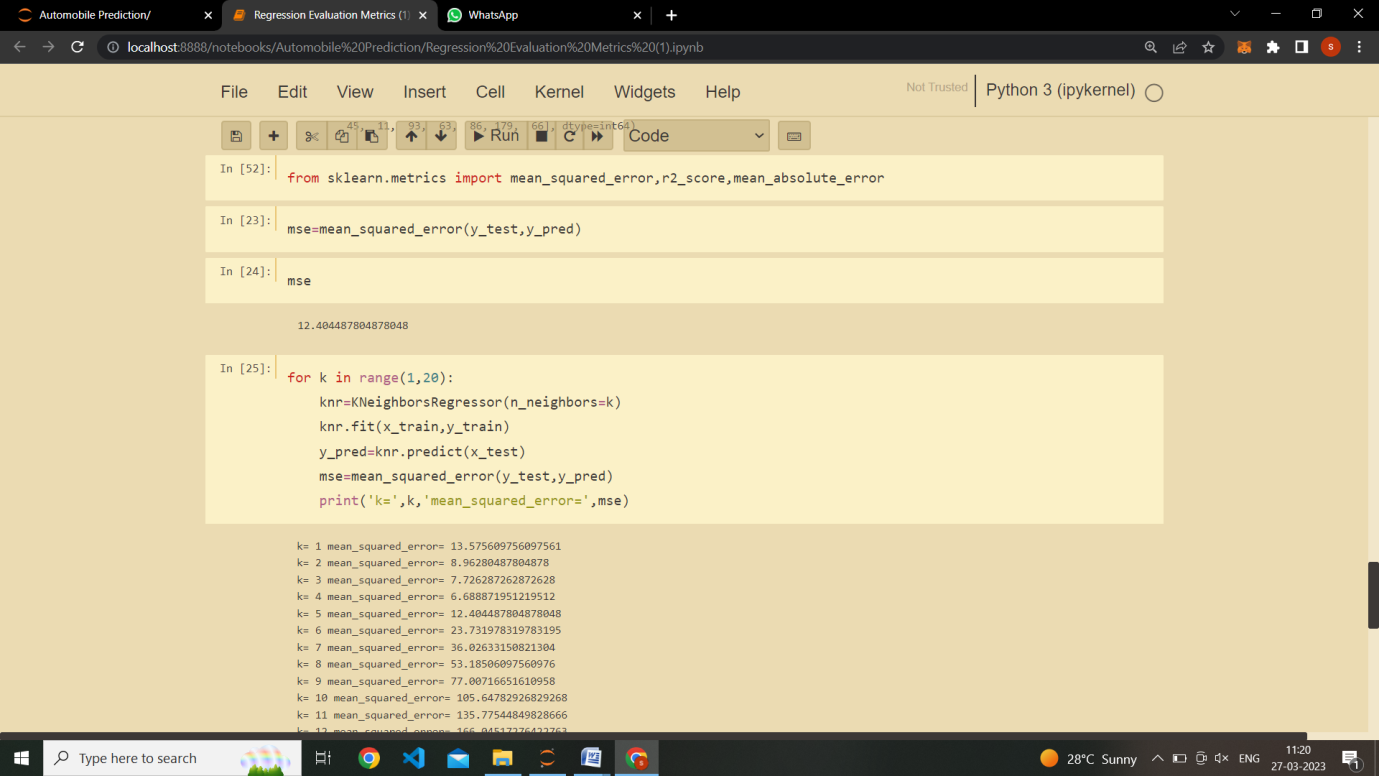
**5.5. Testing (Evaulation) :**

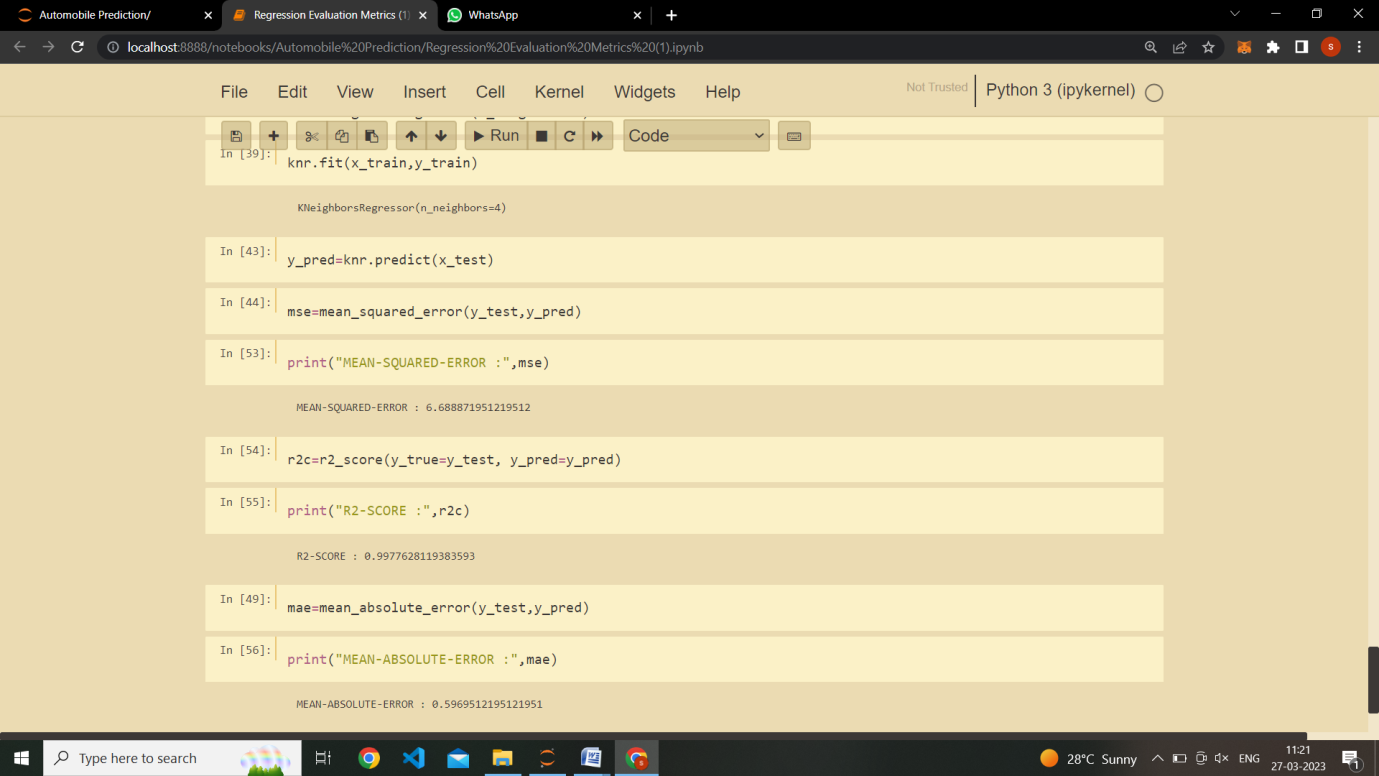
**1. Classification Evaulation metrices :**

Classification evaluation metrics are used to assess the performance of a classification model.

****

**2. Regression Evaulation metrices :**



****

**CONCLUSION**

Utilizing AI calculations and the Kaggle dataset, we endeavoured to expect the selling cost of pre-owned vehicles. This dataset was anticipated utilizing Random Forest and Extra Tress Regressor. The expectation of the model is contrasted with a test dataset made by haphazardly picking values from the first dataset, and the forecast is surveyed utilizing an assortment of techniques. We might presume that the expectation model is exceptionally precise after a far-reaching study and that Random Forest and Extra Tree Regression are among the best calculations for relapse issues. These two methodologies are amazingly precise and quick, no matter what the size of the dataset.

**ACKNOWLEDGEMENT**

We would like to express my sincere thanks to our guide Mrs. Prof. K.R.Ghule for her valuable guidance and support in completing my project. We would also like to express my gratitude towards our Head of Department of Computer Science & Engineering, for giving us this great opportunity to do a project on Crowdfunding Website. Without their support and suggestions, this project would not have been completed.