**ABSTRACT**

Heavy-duty trucks contribute approximately 20% of fuel consumption in the United States of America (USA). The fuel economy of heavy-duty vehicles (HDV) is affected by several real-world parameters like road parameters, driver behavior, weather conditions, and vehicle parameters, etc. Although modern vehicles comply with emissions regulations, potential malfunction of the engine, regular wear and tear, or other factors could affect vehicle performance. Predicting fuel consumption per trip based on dynamic on-road data can help the automotive industry to reduce the cost and time for on-road testing. Data modeling can easily help to diagnose the reason behind fuel consumption with a knowledge of input parameters. In this paper, an artificial neural network (ANN) was implemented to model fuel consumption in modern heavy-duty trucks for predicting the total and instantaneous fuel consumption of a trip based on very few key parameters, such as engine load (%), engine speed (rpm), and vehicle speed (km/h). Instantaneous fuel consumption data can help to predict patterns in fuel consumption for optimized fleet operations. In this work, the data used for modeling was collected at a frequency of 1Hz during on-road testing of modern heavy-duty vehicles (HDV) at the West Virginia University Center for Alternative Fuels Engines and Emissions (WVU CAFEE) using the portable emissions monitoring system (PEMS). The performance of the artificial neural network was evaluated using mean absolute error (MAE) and root mean square error (RMSE). The model was further evaluated with data collected from a vehicle on-road trip. The study shows that artificial neural networks performed slightly better than other machine learning techniques such as linear regression (LR), and random forest (RF), with high R-squared (R 2 ) and lower root mean square error.

**Keywords:** heavy-duty vehicles; center for alternative fuels engines and emissions; average fuel consumption; machine learning; linear regression; random forest; artificial neural network

**TABLE OF CONTENTS:** -

**1. INTRODUCTION………………………………………………………1**

1.1OVERVIEW…………………………………………………...1

1.2 PURPOSE……………………………………………………...1

**2. LITERATURE SURVEY……………………………………………….2**

2.1 EXISTING PROBLEM………………………………………..2

2.2 PROPOSED SYSTEM………………………………………...2

**3. THEORITICAL ANALYSIS …………………………..…………….3-4**

3.1 BLOCK DIAGRAM…………………………………………...3

3.2 HARDWARE / SOFTWARE DESIGNING ……………….3-4

**4. EXPERIMENTAL INVESTIGATIONS ……………..………………..5**

**5. FLOW CHART** **…………………………………………….……………6**

**6. ADVANTAGES …………………….…….…………..…………………7**

**7. CONCLUSION ………………………………………………………….8**

**8.FUTURE SCOPE ………………………………………………………..9**

**9. BIBILOGRAPHY……………………………………………………….10**

**10.APPENDIX……………………………………………………………...11-22**

**11.RESULT ………………………………………………………………..23-24**

### **Trip Based Modelling of Fuel Consumption in Modern Fleet Vehicles**

### **using IBM Watson**

### 1.INTRODUCTION

**1.1OVERVIEW**

The fuel efficiency of fleet vehicles can be beneficial not only for the automotive and transportation industry but also for a country’s economy and the global environment . The cost of fuel consumed contributes to approximately 30% of a fleet vehicles life cycle cost. Reduction in fuel consumption by just a few percent can significantly reduce costs for the transportation industry . The effective and accurate estimation of fuel consumption (fuel consumed in L/km) can help to analyze emissions as well as prevent fuel-related fraud. As per Environmental Protection Agency (EPA) reports, 28% of total greenhouse gas emissions come from transportation (heavy-duty vehicles and passenger cars) . The United States Environmental Protection Agency (US EPA) has introduced Corporate Average Fuel Economy (CAFÉ) standards enforcing automotive manufacturers to be compliant with standards to regulate fuel consumption . US EPA regulations enacting fuel economy improvements in freights released in 2016 target truck fuel efficiency, which is predicted to improve by 11–14% by 2021 . Most states have now mandated that truck fleets update their vehicle inventory with modern vehicles due to air quality regulation.  
  
**1.2 PURPOSE**

Several studies have been presented in the past for evaluating the fuel efficiency of fleet vehicles using simulation-based models and data-driven models. A simulation model was developed based on engine capacity, fuel injection, fuel specification, aerodynamic drag, grade resistance, rolling resistance, and atmospheric conditions, with simulated dynamic driving conditions to predict fuel consumption . A statistical model which is fast and simple compared to the physical load-based approach was developed to predict vehicle emissions and fuel consumption . The impact of road infrastructure , traffic conditions , drivers’ behaviour , weather conditions , and the ambient temperature on fuel consumption were studied, and it was determined that fuel consumption can be reduced by 10% with eco-driving influences. The era of big data has enabled the modeling of huge volumes of data for companies to reduce emissions and fuel consumption. Machine learning techniques such as support vector machine (SVM) , random forest (RF) , and IBM watson are widely applied to turn data into meaningful insights and solve complex problems.

**2.LITERATURE SURVEY**

**2.1 EXISTING PROBLEM**

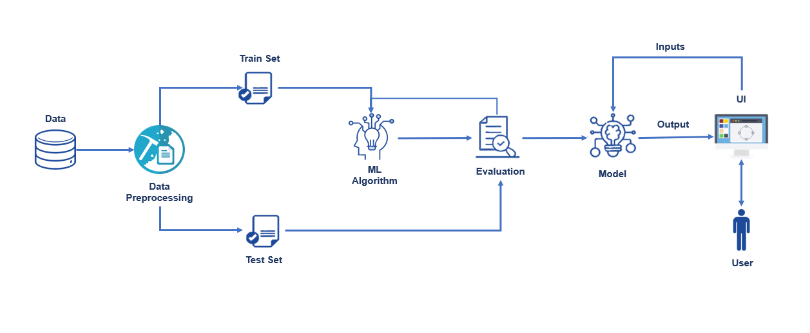
In recent years, deep learning has been used in various applications including the classification of ship targets in inland waterways for enhancing intelligent transport systems. Various researchers introduced different classification algorithms, but they still face the problems of low accuracy and misclassification of other target objects. Hence, there is still a need to do more research on solving the above problems to prevent collisions in inland waterways.

**2.2 PROPOSED SYSTEM**

In order to solve the problems for the accuracy of the classification system, we proposed a new classification model. First, based on the pretrained models, the models were fine-tuned with the public dataset we used. Based on their performance, the best model was selected in order to further adjust the performance for high accuracy in classifying ships in inland river waterways. After selecting the best model, the model was adjusted, and classification was conducted based on the modification of the network.

1. **THEORETICAL ANALYSIS**

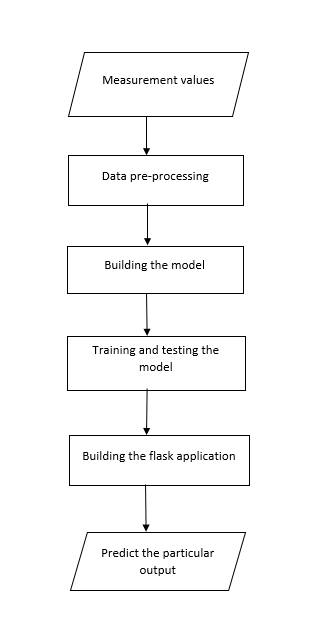
**3.1 BLOCK DIAGRAM**



**3.2 HARDWARE AND SOFTWARE DESIGNING  
  
Software requirements:  
  
Python**Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. It was created by Guido van Rossum , and first released on February 20, 1991. Its high-level built in data structures, combined with dynamic typing and dynamic binding , make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

**Anaconda navigator**  
Anaconda Navigator is a free and open-source distribution of the Python and R programming languages for data science and machine learning related applications. It can be installed on Windows, Linux, and mac OS.Conda is an open-source, cross platform, package management system. Anaconda comes with so very nice tools like JupyterLab, Jupyter Notebook, QtConsole, Spyder, Glueviz, Orange, Rstudio, Visual Studio Code. For this project, we will be using Jupyter notebook and Spyder.  
  
**Jupyter Notebook**The Jupyter Notebook is an open source web application that you can use to create and share documents that contain live code, equations, visualizations, and text. Jupyter Notebook is maintained by the people at Project Jupyter. Jupyter Notebooks are a spin-off project from the IPython project, which used to have an IPython Notebook project itself. The name, Jupyter, comes from the core supported programming languages that it supports: Julia, Python, and R. Jupyter ships with the IPython kernel, which allows you to write your programs in Python, but there are currently over 100 other kernels that you can also use.  
  
**Spyder**Spyder, the Scientific Python Development Environment, is a free integrated development environment (IDE) that is included with Anaconda. It includes editing, interactive testing, debugging, and introspection features. Initially created and developed by Pierre Raybaut in 2009, since 2012 Spyder has been maintained and continuously improved by a team of scientific Python developers and the community. Spyder is extensible with first-party and third party plugins includes support for interactive tools for data inspection and embeds Python specific code. Spyder is also pre-installed in Anaconda Navigator, which is included in Anaconda.  
  
**Flask**Web framework used for building. It is a web application framework written in python which will be running in local browser with a user interface. In this application, whenever the user interacts with UI and selects emoji, it will suggest the best and top movies of that genre to the user.  
  
**Hardware Requirements:**  
o Operating system: window7 and above with 64bit   
o Processor Type -Intel Core i3-3220  
o RAM: 4Gb and above   
o Hard disk: min 100gb  
  
   
  
  
  
  
  
  
  
  
  
  
  
 **4.EXPERIMENTAL INVESTIGATION**   
The text data need to be organized before proceeding with the project. The original dataset has a single folder. We will be using the measurements.csv file to fetch the text data of training data. The data need to be unique and all fields need to be filled. The dataset images are to be pre-processed before giving to the model. We will create a function that uses the pre-trained model for predicting custom outputs. Then we have to test and train the model. After the model is build, we will be integrating I to a web application.

**5.FLOWCHART**



**6. ADVANTAGES**

* Data modeling can help to identify the trend in instantaneous fuel consumption and to calculate the total fuel consumed by the vehicle for each trip, which can further help in diagnosing vehicle performance in the case of abnormalities.
* The effective and accurate estimation of fuel consumption (fuel consumed in L/km) can help to analyze emissions as well as prevent fuel-related fraud

1. **CONCLUSION**

In conclusion, the study demonstrates the modeling of fuel consumption in modern fleet vehicles with an IBM Watson using very few technical parameters. An attempt was made to develop a model using very few parameters collected under different conditions. Data from modern fleet vehicles with the same make and model, driven by different persons on various routes under different external conditions, were used for training the IBM Watson.

### FUTURE SCOPE

* In future works, the proposed method will be improved in order to classify the people in different countries with extra features using more advanced technology.

### BIBILOGRAPHY

* **<https://researchrepository.wvu.edu/faculty_publications/3072/?utm_source=researchrepository.wvu.edu%2Ffaculty_publications%2F3072&utm_medium=PDF&utm_campaign=PDFCoverPages>**

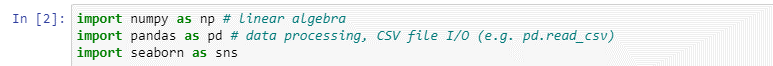
**10. APPENDIX**

### 

### SOURCE CODE:

### Import Required Libraries

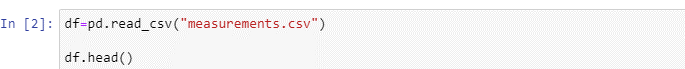
Go to the project folder which you have created copy the project path and open anaconda prompt from the menu and go to the location of your project folder in anaconda prompt and type jupyter notebook. Now jupyter notebook will be opened and create a python file and start the programming.



### Read The Datasets

The Our dataset format might be in .csv, excel files, .txt, .json, etc. We can read the dataset with the help of pandas.

In pandas, we have a function called read\_excel() to read the dataset. As a parameter, we have to give the directory of xlsx file.



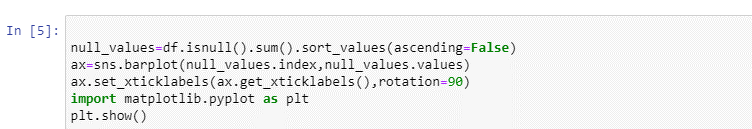
### 

### Check Null Values

For checking the null values, df.isnull() function is used. To sum those null values we use .sum() function to it. To visualize the null values heatmap() and barplot() from seaborn package is used.

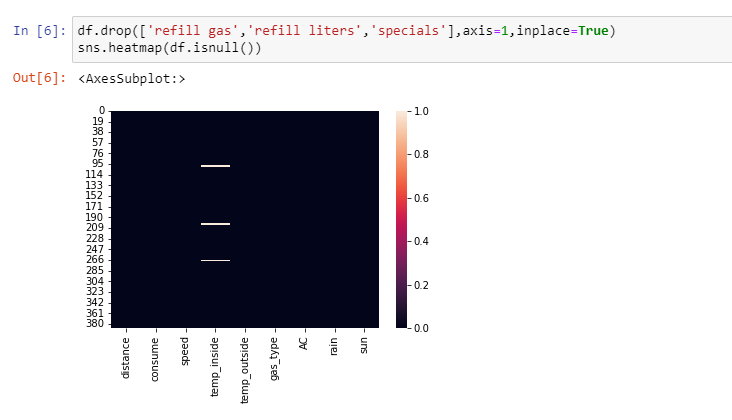
### 

* Plotting the variables which consist of maximum no of null values.



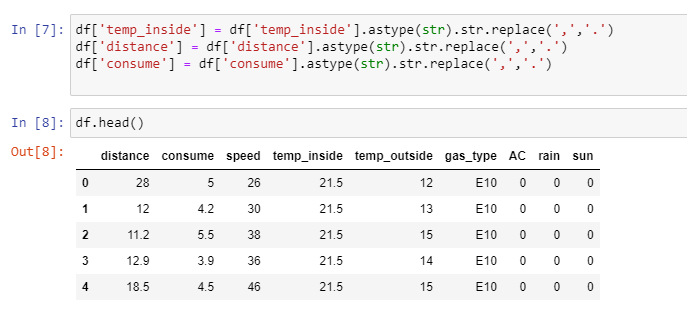
### 

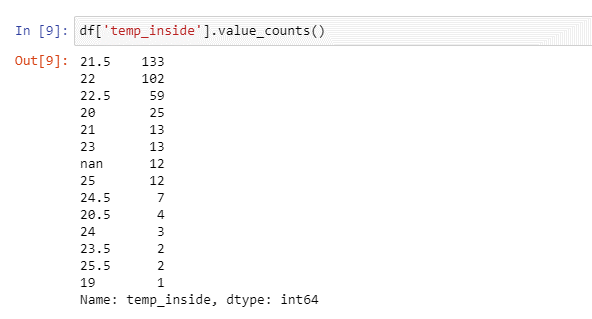
### Removing Null Values

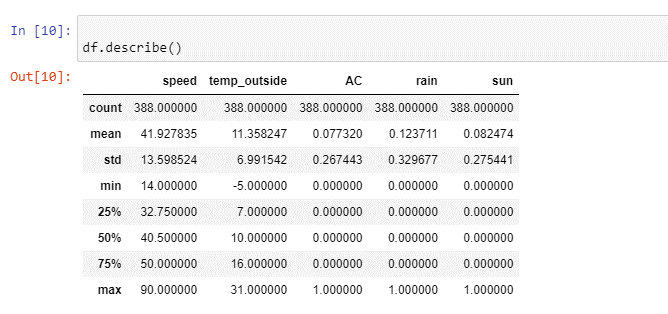
Refill gas, Refill liters, and specials columns are dropped using the drop() method from pandas. From the above image, we found these columns have many null values so it is dropped. Axis should be given as a parameter on the drop method.

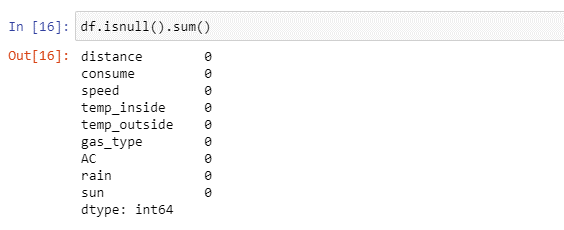
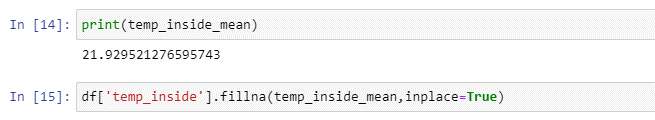
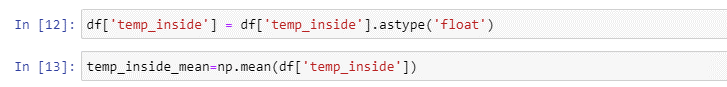
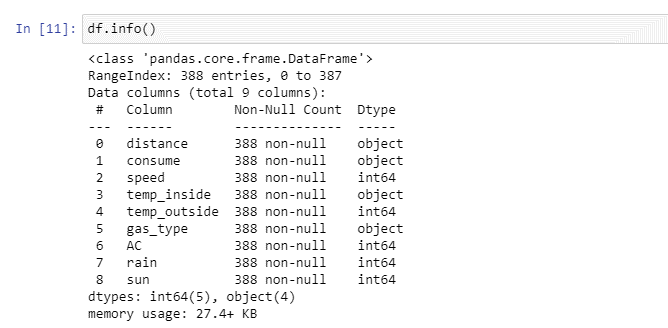
### Handling Null Values

Here we are going to handle null values. From activity 3 we found we have null values in the ‘temp\_inside’ column. So we are replacing the null value with its mean. Fillna() method from pandas is used to replace null values with their mean.







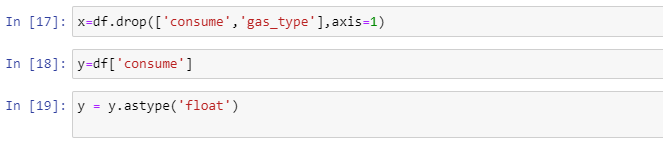
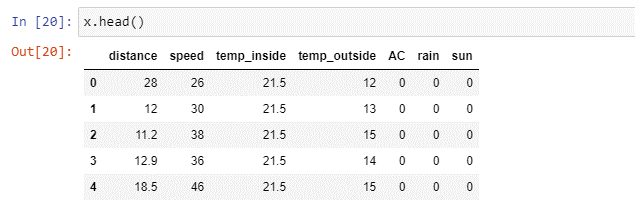


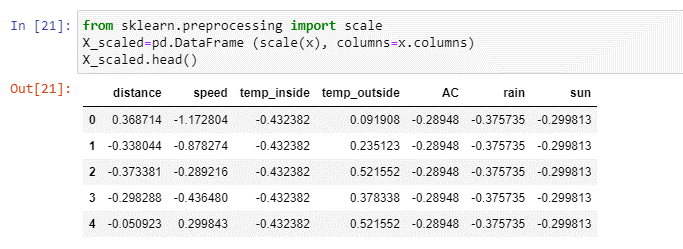
**Model building:**

### Seperating Independent And Dependent Variables

Now let’s split the Dataset into train and test sets. First split the dataset into x and y and then split the data set.

Here x and y variables are created. On x variable, df is passed with dropping the target variable. And on y target variable is passed.

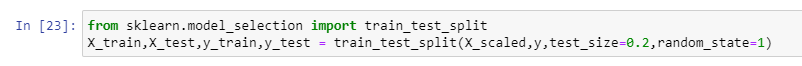




### 

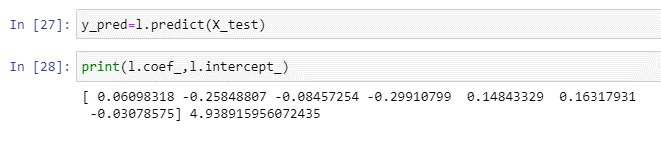
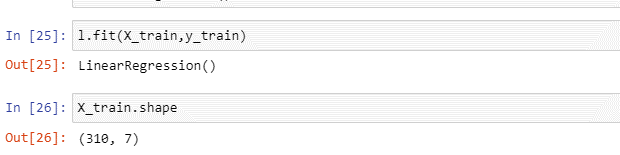
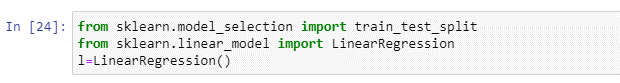
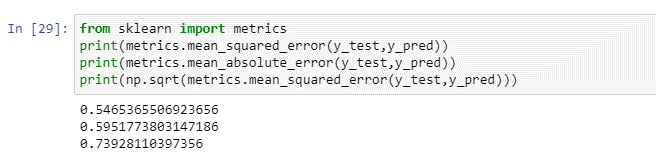
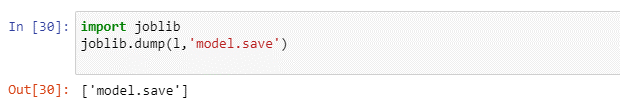
### Splitting Data Into Train And Test

For splitting training and testing data we are using train\_test\_split() function from sklearn. As parameters, we are passing x, y, test\_size, random\_state.



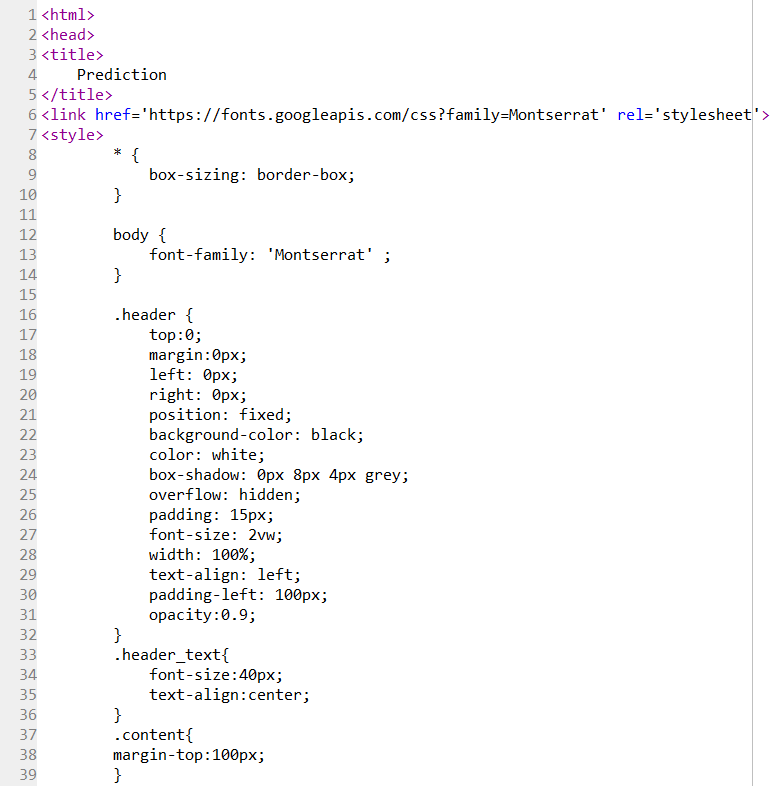
### Applying Linear Regression

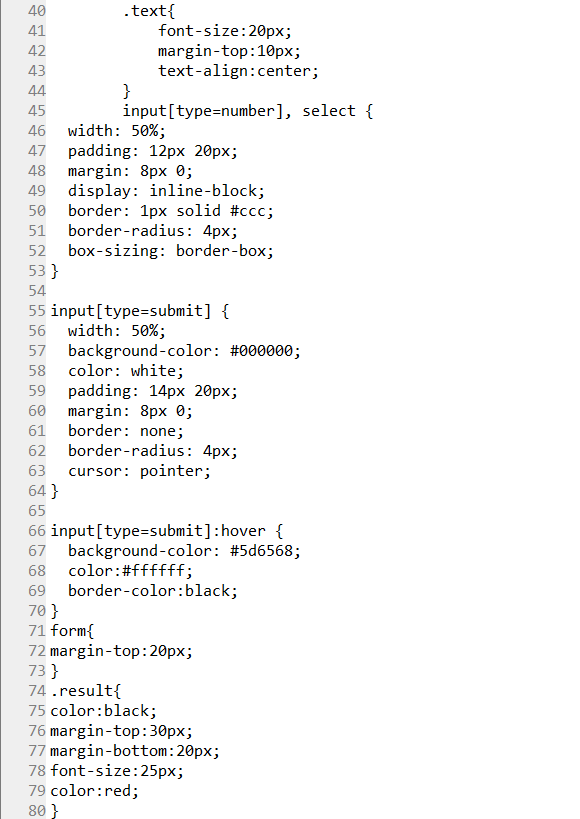
Now we are going to create our model with linear regression. As an initial step we have to initialize the linear model. Then train the model with fit() method. Now our model is trained and to test the model predict() method is used. To find the loss of linear regression model mean\_squared\_error and mean\_absolute\_error are used.

### Build An HTML Page:

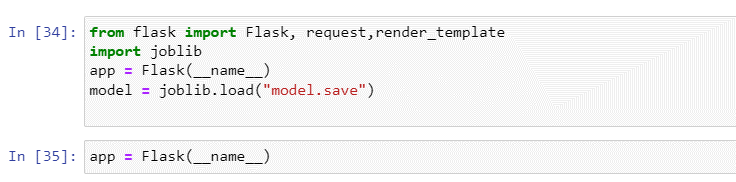
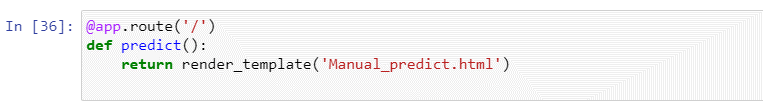
We Build an HTML page to take the values from the user in a form and upon clicking on the predict button we get the fuel consumption predicted.







### Build The Python Flask App

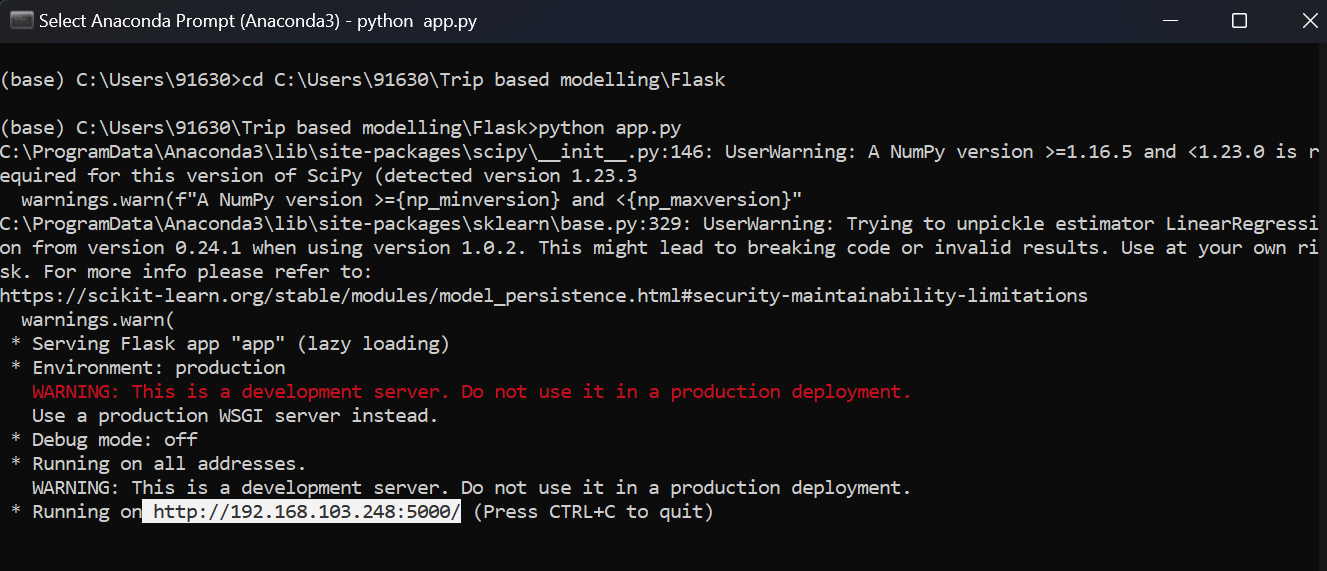
Load the saved model. Importing the flask module in the project is mandatory. An object of Flask class is our WSGI application. Flask constructor takes the name of the current module as argument.  Here we will be using declared constructor to route to the HTML page which we have created earlier.

In the above example, ‘/’ URL is bound with Manual\_predict.html function. Hence, when the home page of the web server is opened in browser, the html page will be rendered. Whenever you enter the values from the predict html page the values can be retrieved using POST Method.

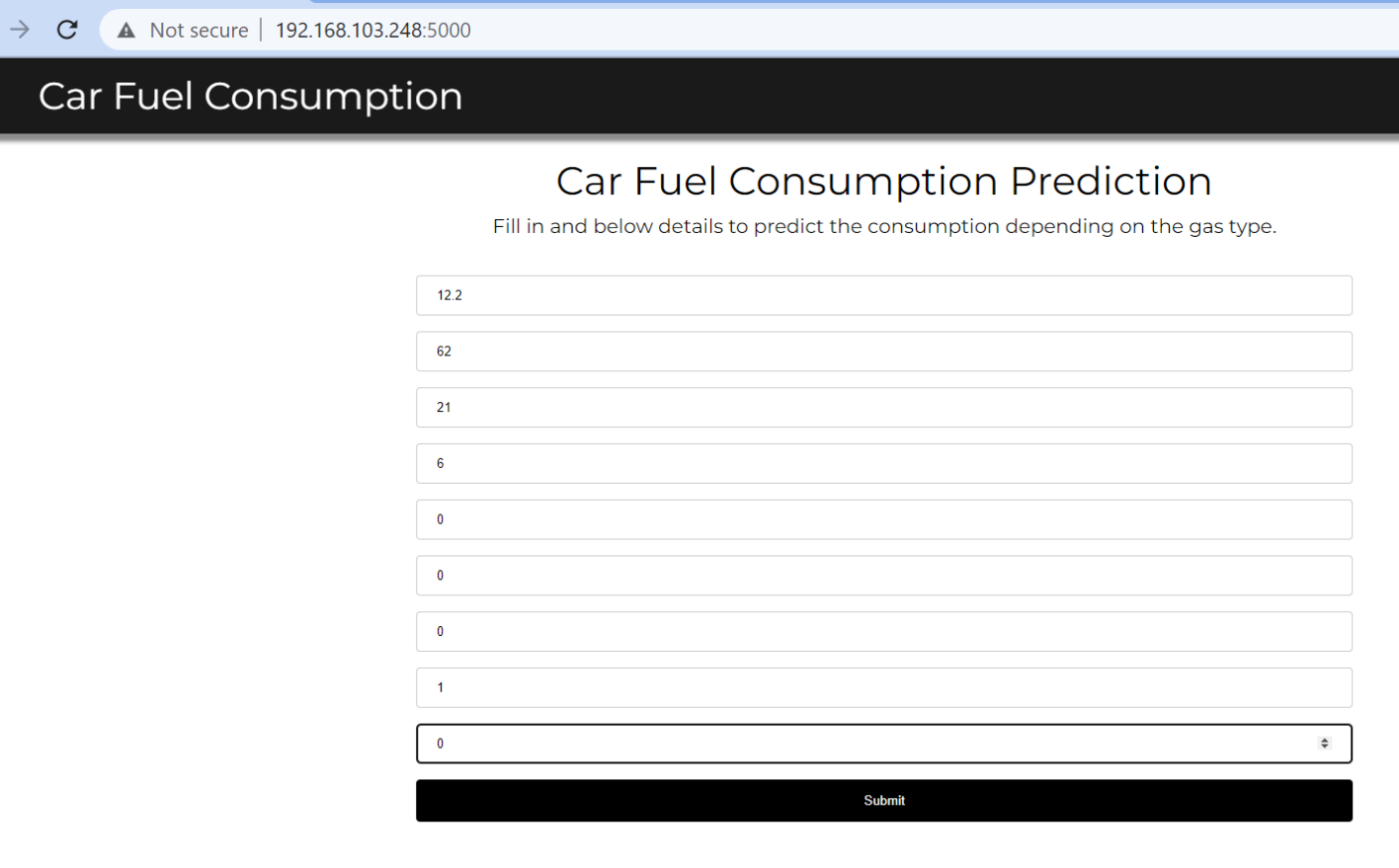
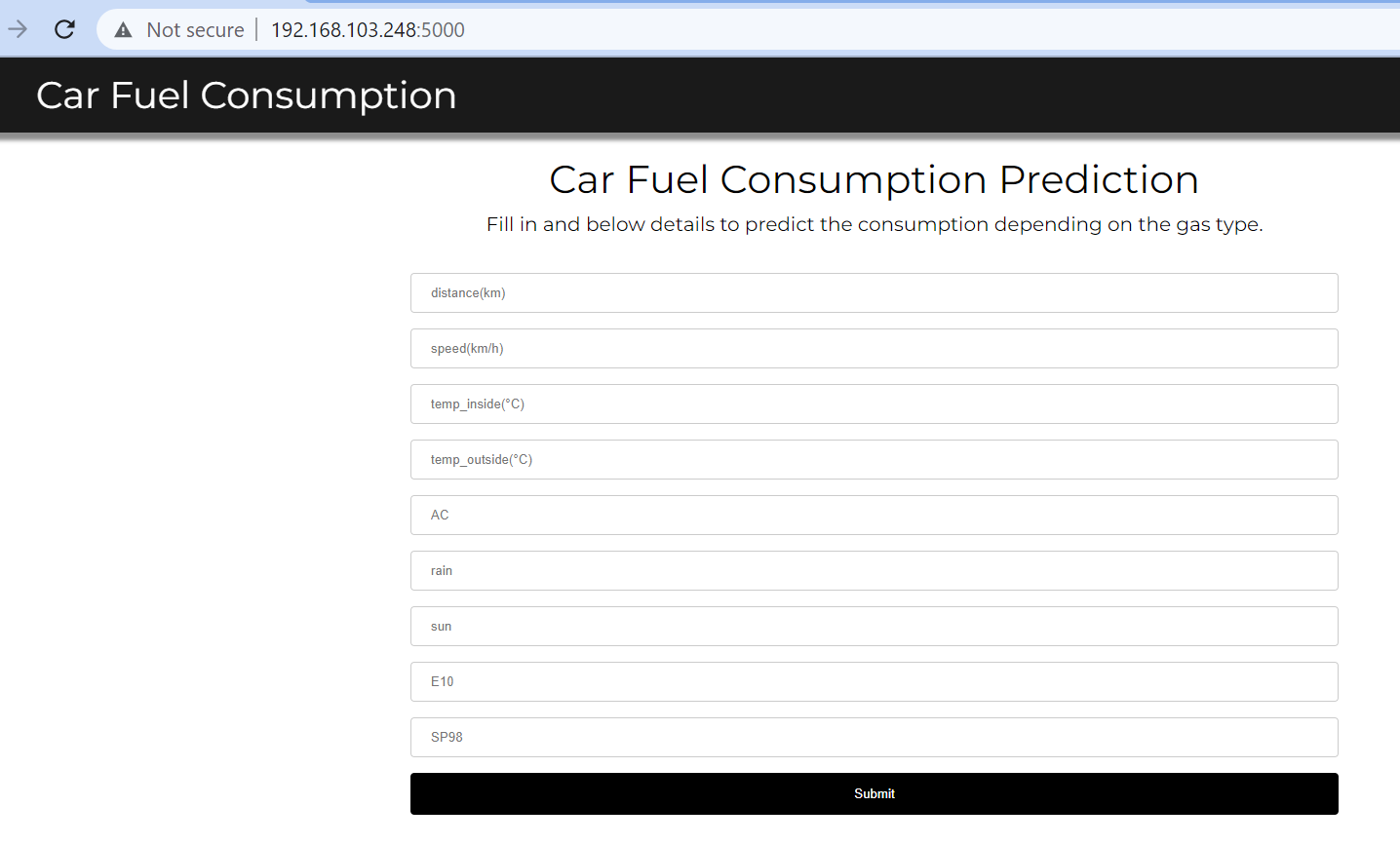
**RETRIVES THE VALUES FROM UI:**

### 

### Run The Application:

**Step 1:** Open anaconda prompt go to project folder and in that go to flask folder and run the python file by using the command “**python app.py”**

**11. RESULT**

**OUTPUT:**

