1. INTRODUCTION

a. Overview

Carbon emissions and environmental protection issues have brought pressure from the international community during Chinese economic development. Recently, Chinese Government announced that carbon emissions per unit of GDP would fall by 60–65% compared with 2005 and non-fossil fuel energy would account for 20% of primary energy consumption by 2030.The Beijing-Tianjin-Hebei region is an important regional energy consumption center in China, and its energy structure is typically coal-based which is similar to the whole country.Therefore, forecasting energy consumption related to carbon emissions is of great significance to emissions reduction and upgrading of energy supply in the Beijing-Tianjin-Hebei region. Thus, this study thoroughly analyzed the main energy sources of carbon emissions including coal, petrol, natural gas, and coal power in this region.

Carbon footprint has become a widely used term and concept in the public debate on responsibility and abatement action against the threat of global climate change. It had a tremendous increase in public appearance over the last few months and years and is now a buzzword widely used across the media, the government and in the business world. It is a measure of the total amount of carbon dioxide released into the atmosphere in the given time frame that is directly or indirectly caused by an activity to provide service or product . A carbon footprint is a measure of the amount of carbon dioxide emitted through the combustion of fossil fuels. In the case of a business organization, it is the amount of CO_2 emitted either directly or indirectly as a result of its everyday operations. It also might reflect the fossil energy represented in a product or commodity reaching the market. The carbon footprint of U.S. households is About 5 times greater than the global avera

ge, which is approximately 10 tons CO_2 per household per year. For most U.S. households, the single most important action to reduce their carbon footprint is driving less or switching to a more efficient vehicle

b. Purpose

ML depends heavily on data, without data, it is impossible for an "AI" to learn. It is the most crucial aspect that makes algorithm training possible. In Machine Learning projects, we need a training data set. It is the actual data set used to train the model for performing various actions.

There are many features which are responsible for CO-2 Emission of Countries, e.g. Country Name, Country Code, Indicator Name etc. For better prediction of the CO_2 Emission of Countries, we should consider as many relevant features as possible.

2. LITERATURE SURVEY

a. Existing problem

Research on CO_2 emissions is highly renewable, qualitative and quantitative researches that have been done until now is still in the process of the discussion, it makes the theory and approaching method in order to calculate the emitter of CO_2 emissions has not become one unit . Mitigation of Carbon Dioxide emission is the challenge of the future in order to stabilize global warming . CO_2 prediction using a computational intelligence approach . An adaptive neuro-fuzzy interference system (ANFIS) and multi-layer perceptron artificial neural network (MLP-ANN) have been developed to estimate CO_2. The proposed model of ANFIS and MLP-ANN demonstrates that both methods can solve CO_2 prediction problem. There's another method for prediction namely, support vector machine (SVM). Recently, several applications of SVM can be found both for classification and for regression problems . SVM is more accurate than semi empirical equations for predict solubility of different solutes in supercritical carbon dioxide . On the other hand, SVM was implements the structural risk minimization principle. SVM has successfully solved a prediction of CO_2 exchange rate . In their study, different tests were performed along the North Atlantic oceanic region with data obtained during

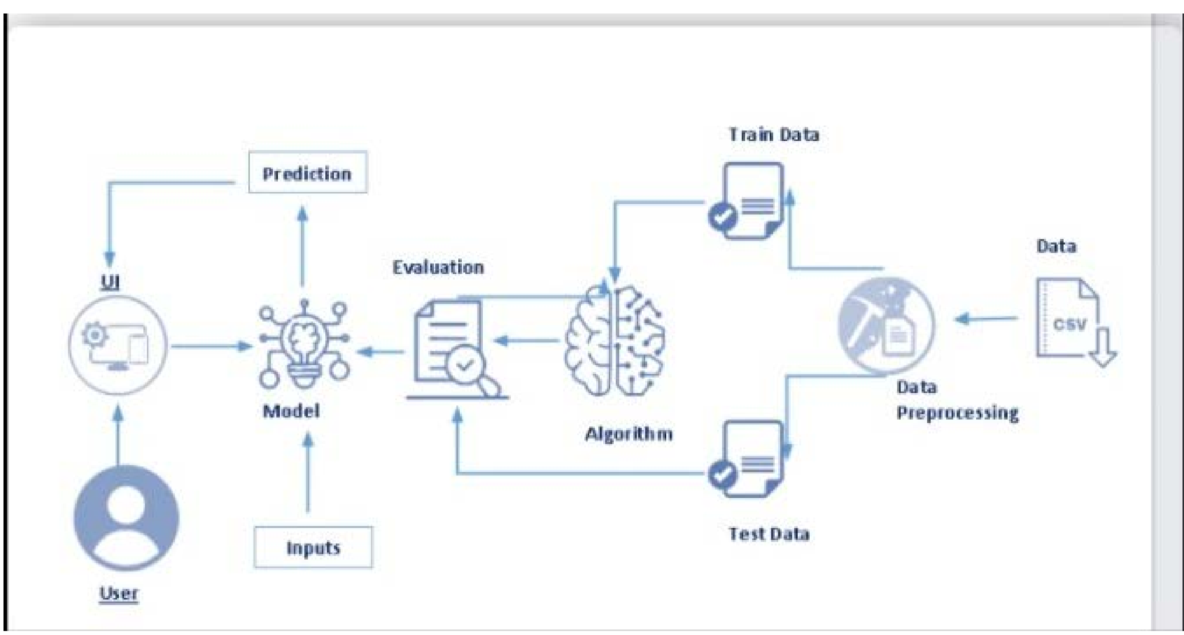
2009 and the proposed model of SVM demonstrates that SVM can solve CO_2 prediction problem. SVM is powerful machine learning tool that can be used for time-series prediction .

b. Proposed solution

The project "Prediction of CO2 emmisions bt Country using IBM Watson" aim to building a machine learning model to predict the amount of CO_2 emitted by country, Due to the increasingly deteriorating environment, it is time for the government to upgrade the Energy Consumption structure by making use of Machine Learning prediction to analyze and control the CO_2 emissions in future.

1. THEORITICAL ANALYSIS

a. Block diagram



b. Hardware / Software designing

Hardware Requirements:

Processor : Intel Core I3

RAM : 4.00 GB

Operating system : Windows/Linux/MAC

Software Requirements:

Anaconda

Jupyter Notebook

Spyder

4. EXPERIMENTAL INVESTIGATIONS

i) Dataset Collection

Collect the dataset or create the dataset.

ML depends heavily on data, without data, it is impossible for an "AI" to learn. It is the most crucial aspect that makes algorithm training possible. In Machine Learning projects, we need a training data set. It is the actual data set used to train the model for performing various actions.

There are many features which are responsible for CO-2 Emission of Countries, e.g. Country Name, Country Code, Indicator Name etc. For better prediction of the CO2 Emission of Countries, we should consider as many relevant features as possible.

We can collect dataset from different open sources like kaggle.com, data.gov, UCI machine learning repository etc.

The kaggle repository link is : https://www.kaggle.com/ashukr/exploring-co2-emission?select=Indicators.csv ii) Data Preprocessing :

Here, we are reading the dataset(.csv) from the system using pandas and storing it in a variable 'df'. It's time to begin building your text classifier! The data has been loaded into a DataFrame called df. The .head() method is particularly informative.

We might have your data in .csv files, .excel files or .tsv files or something else. But the goal is the same in all cases. If you want to analyse that data using pandas, the first step will be to read it into a data structure that's compatible with pandas.

load a .csv data file into pandas. There is a function for it, called read_csv().We will need to locate the directory of the CSV file at first (it's more efficient to keep the dataset in the same directory as your program).

iii) Check Unique Values In Dataset

Often, a DataFrame will contain columns that are having some unique values from which we can find out the unique records which are present in the dataset. find the min year and max year for count and also,find how many years of data we have.

iv) CO-2 Emission Of The Countries.

As our dataset is very huge, we are dealing with a few countries like the USA,SGP,IND,BRB,ARB. Here we are selecting the indicator name which is Co2-Emission (metric tons per capita) and also take the country code to find the co2 emission over the time.

Select CO2 emissions for the Arab Country and stage is just those indicators matching the ARB for country code and CO2 emissions over time.

v) Observing Target,Numerical And Categorical Columns

vi) Data Visualization :

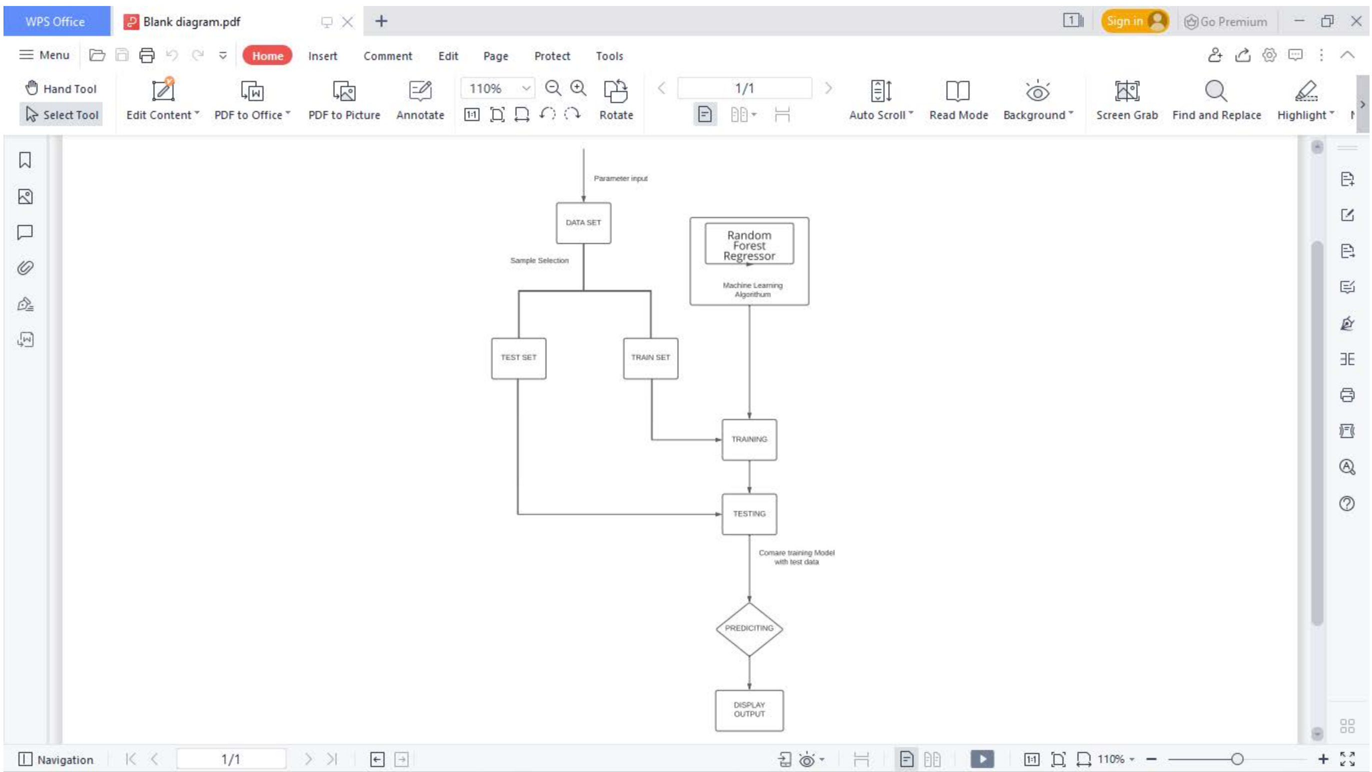
Data visualization is where a given data set is presented in a graphical format. It helps the detection of patterns, trends and correlations that might go undetected in text-based data.Understanding your data and the relationship present within it is just as important as any algorithm used to train your machine learning model. In fact, even the most sophisticated machine learning models will perform poorly on data that wasn't visualized and understood properly.

To visualize the dataset we need libraries called Matplotlib and Seaborn.The Matplotlib library is a Python 2D plotting library which allows you to generate plots, scatter plots, histograms, bar charts etc.

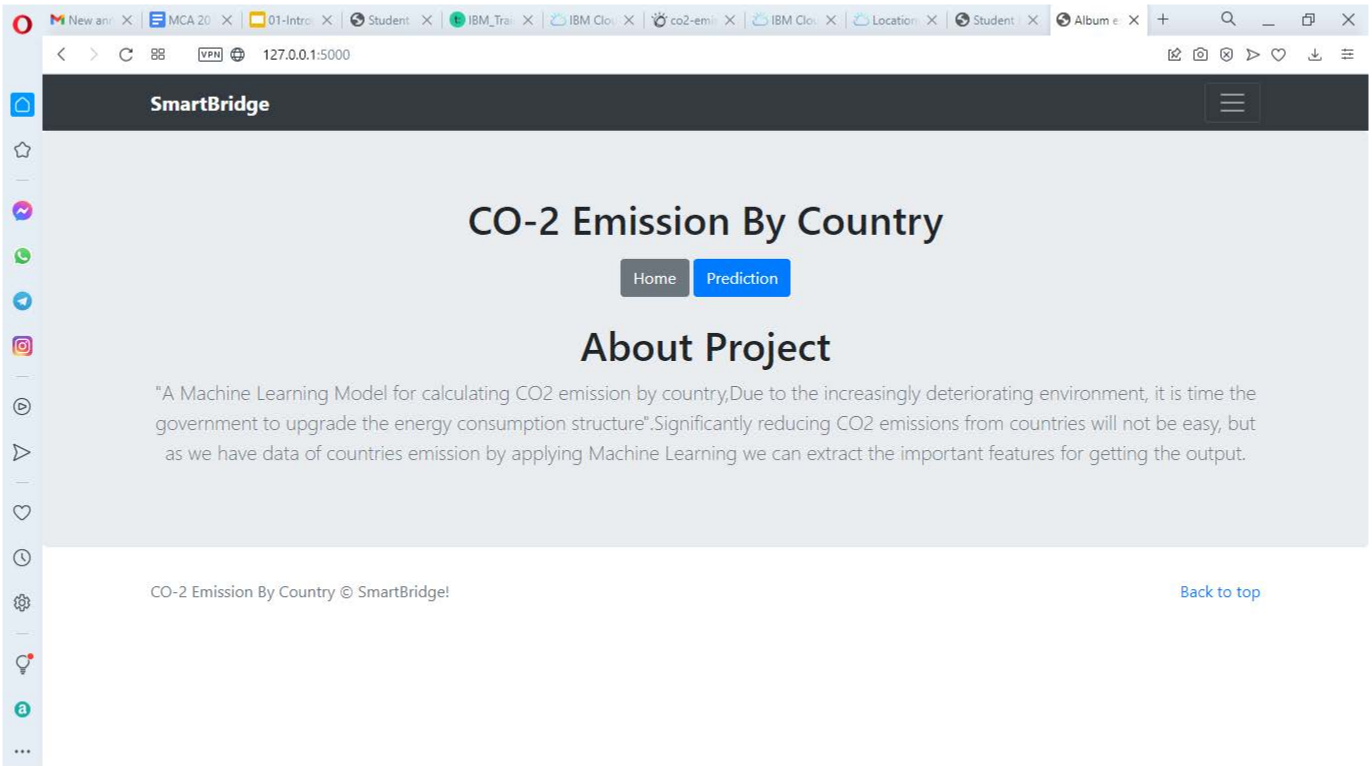
Let's visualize our data using Matplotlib and the Seaborn library.

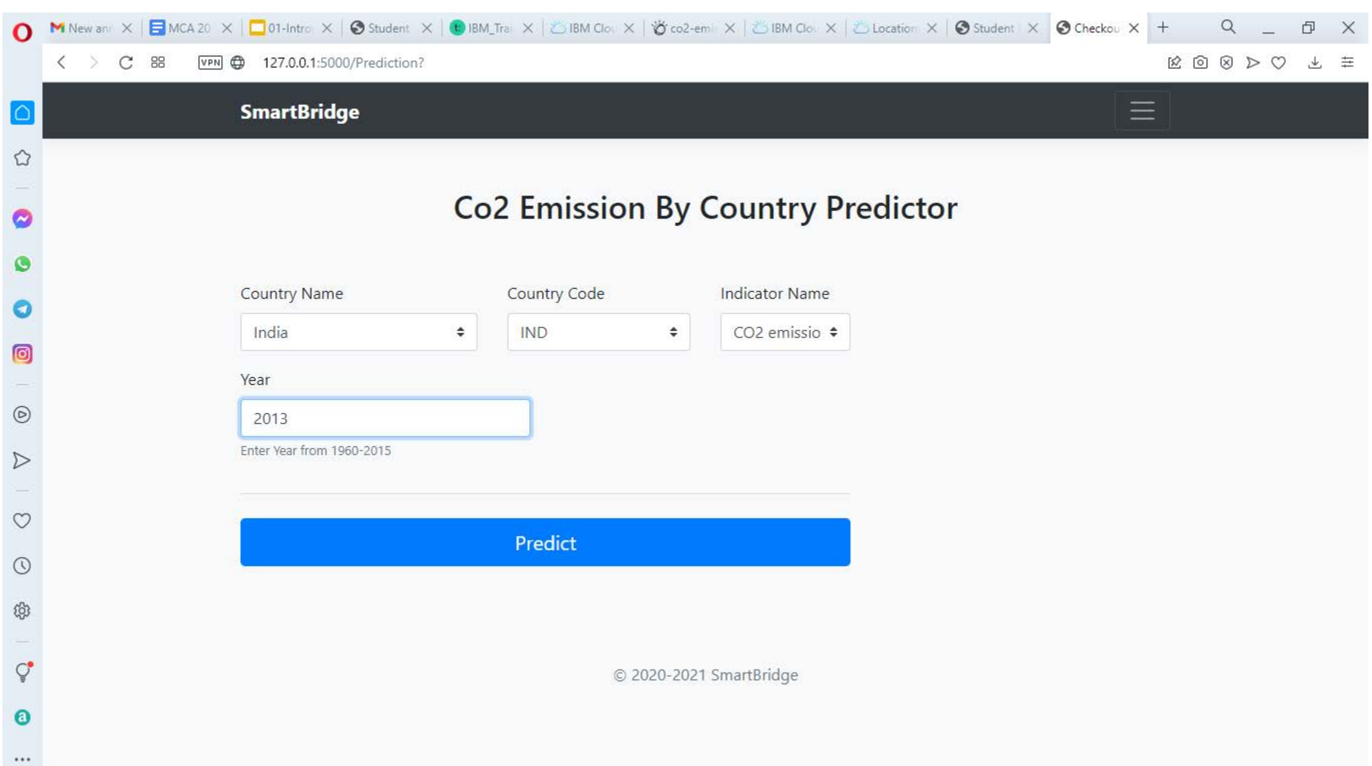
Get the years and co2 emission value in the particular year.

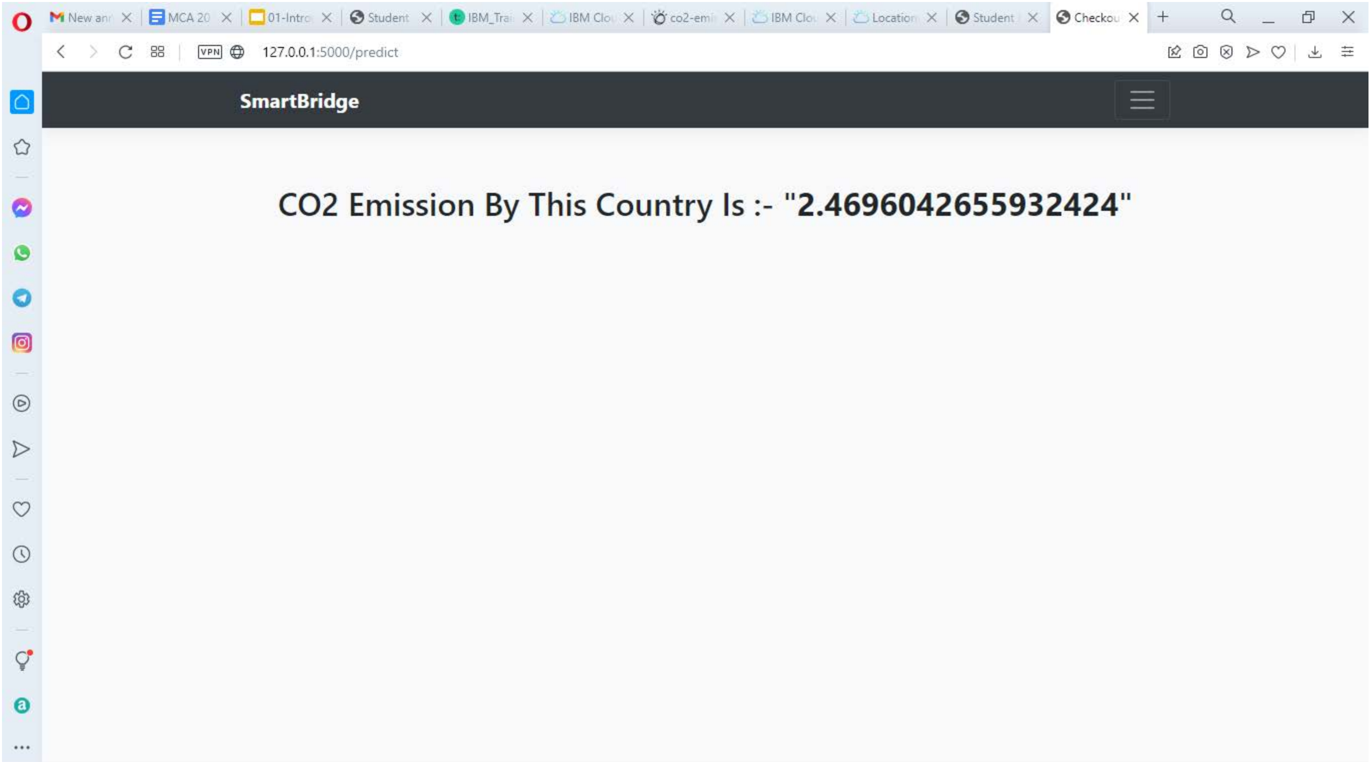
5. FLOWCHART



6. RESULT


127.0.0.1:5000
SmartBridge
CO-2 Emission By Country
Home
Prediction
About Project
"A Machine Learning Model for calculating CO2 emission by country,Due to the increasingly deteriorating environment, it is time the government to upgrade the energy consumption structure".Significantly reducing CO2 emissions from countries will not be easy, but as we have data of countries emission by applying Machine Learning we can extract the important features for getting the output.
CO-2 Emission By Country © SmartBridge!
Back to top


127.0.0.1:5000/Prediction?
SmartBridge
Co2 Emission By Country Predictor
Country Name
India
Country Code
IND
Indicator Name
CO2 emissio
Year
2013
Enter Year from 1960-2015
Predict
© 2020-2021 SmartBridge



ADVANTAGES & DISADVANTAGES

Advantages

- Fast convergence speed
- High training accuracy
- No manual tuning

Disadvantages

- .Slow learning speed
- Need of numerous training sample
- Over fitting

8. APPLICATIONS

1. Machine Learning models such as SVM, tuned SVM, linear model etc are used in Estimating CO_2 emissions from Electricity generation

2. banking and Financial services industry uses Machine Learning to detect and reduce fraud, measure market riska dn identify oppurtunities.

3. Machine Learning play a key part in security as they typically use predictive analysis to improve services and performance, but also to detect anamolies, fraud, understand consumer behavior and enhance data security.

4. ML Algorithms are also known for it's recommendation algorithms like in Facebook or YouTube where similar content will be suggested to engage the user

9. CONCLUSION

The Support Vector Machine model can be applied to predict CO_2 emissions from energy consumption which can give us more accuracy . This model is used to monitor electrical energy and burning coal which affect the amount of CO_2 emitted. Trial and error approach was applied in order to obtain a better prediction model with a lower error. The results obtained show that the lower error (RMSE) value was 0.004 with optimal parameters for the SVM model of 0.1 for the C parameter and 0 for Epsilon. Prediction with high accuracy can give information concerning about CO_2 emissions Furthermore, the main objective in this work is to achieve the lower RMSE when designing the model prediction. It can be concluded that with the high accuracy of the prediction model, then the lower RMSE value must be obtained ..

10 . FUTURE SCOPE

With increasing investment in adaptation, one important challenge for the future will be to establish the right sets of indicators for identifying priorities, as well as monitoring and evaluation frameworks for adaptation. While progress on mitigation can be interpreted from trends in national GHG emissions,comparable measurable outcomes do not yet exist for adaptation. The difficulties in monitoring and evaluating adapftation range from the ambiguous definition of adaptation to the identification of targets and the choice of indicators used to monitor performance. Consequently, while international discussions on adaptation have focused on implementation of adaptation and the associated costs, systematic evaluation of

how much progress is being made in this direction is generally lacking and needs further development.

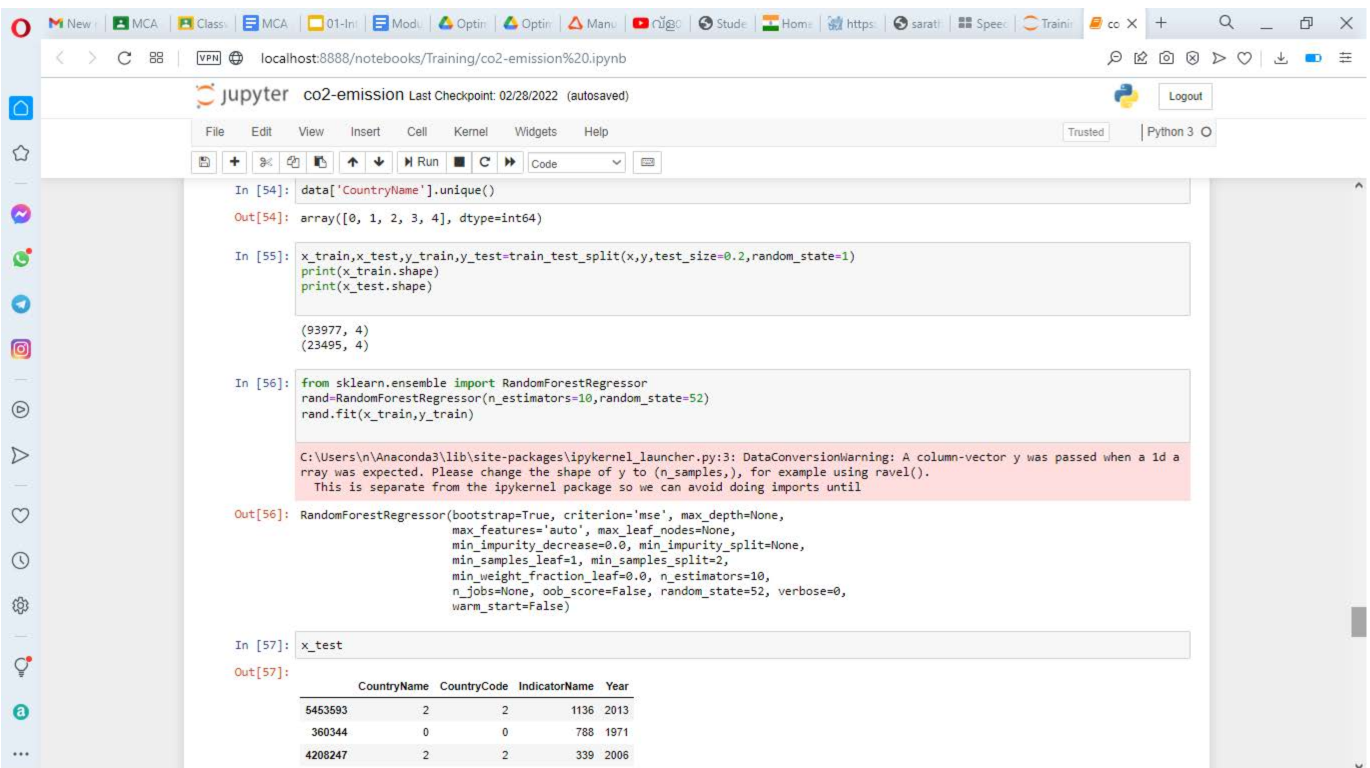
11 . BIBILOGRAPHY

https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html

https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeRegressor.html

APPENDIX

A. Source Code



localhost:8888/notebooks/Training/co2-emission%20.ipynb

jupyter co2-emission Last Checkpoint: 02/28/2022 (autosaved)

5018058	1.018949e+01
1089477	4.456807e+00
4806531	2.701080e+00
3245984	8.660000e+01
5506001	1.790000e+01
4873892	3.870000e+01
3705457	3.038204e+06
498576	4.454733e+01
2551074	8.507301e+00
5585649	1.417129e+09
1876725	1.392817e+00
821618	1.755728e+10
532921	1.005420e+10
2173208	8.390000e+01

23495 rows × 1 columns

In [62]:
```
rand.score(x_train,y_train)
```
Out[62]: 0.9953047142094279

In [63]:
```
data.head()
```
Out[63]:

	CountryName	CountryCode	IndicatorName	IndicatorCode	Year	Value
0	0	0	40	1181	1960	1.335609e+02
1	0	0	44	1204	1960	8.779760e+01