Energy Prediction Model

Prediction of Turbine Energy Yield of Gas Turbine

INSTITUTE FOR ADVANCED COMPUTING AND SOFTWARE DEVELOPMENT AKURDI, PUNE

Documentation on

"Energy Prediction Model"

PG-DBDA MAR 22

Submitted By:

Group No: 12

Ankita Jain 223309

Gauri Wale 223320

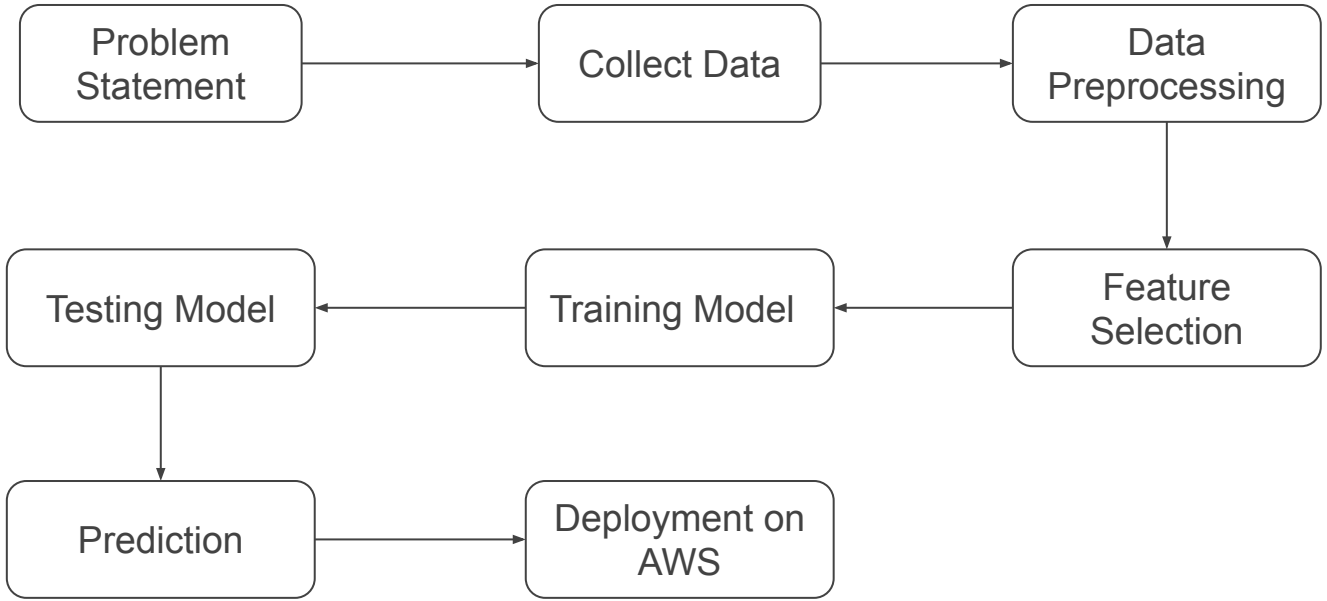
Mr. Prashant Karhale
Centre Coordinator

Mr. Akshay Tilekar
Project Guide

Introduction

Gas turbines are widely used in the energy production. In the present scenario, the quantity of the operating machines requires a special attention for prediction of power production in the energy marketing sector. Thus, the aim of this project is to support the sector by making the prediction of energy yield more computable. By using the data from an operating power plant, correlation and regression analysis are performed and model is developed to calculate energy yield.

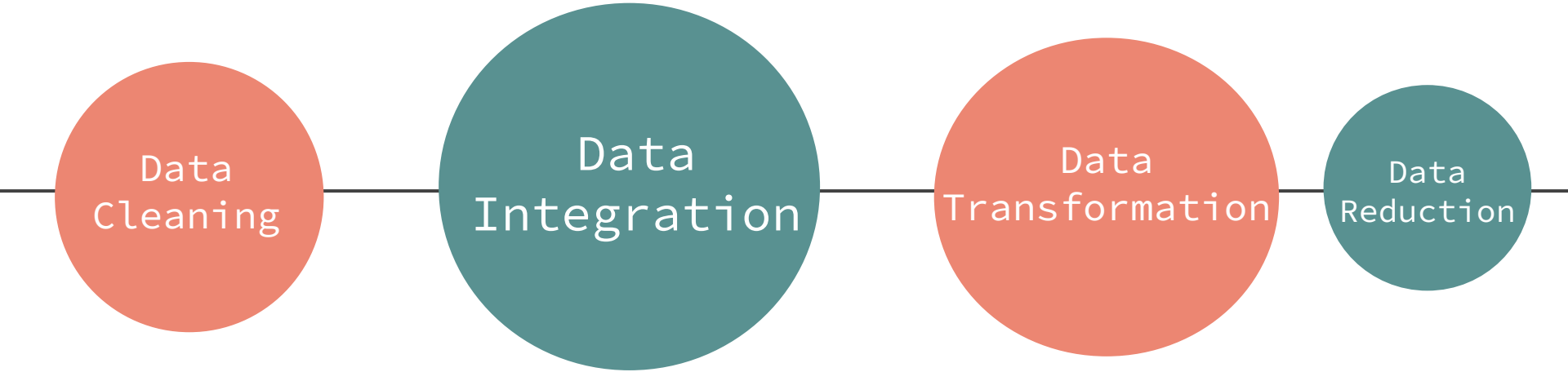
Lifecycle


Problem Statement
Collect Data
Data Preprocessing
Feature Selection
Training Model
Testing Model
Prediction
Deployment on AWS

Collect Data

- Data collected from UCI repository.
- It contains instances of 11 sensor measures aggregated over one hour (by means of average or sum) from a gas turbine located in Turkey's north western region for the purpose of studying flue gas emissions.
- Dataset consists of 11 features.

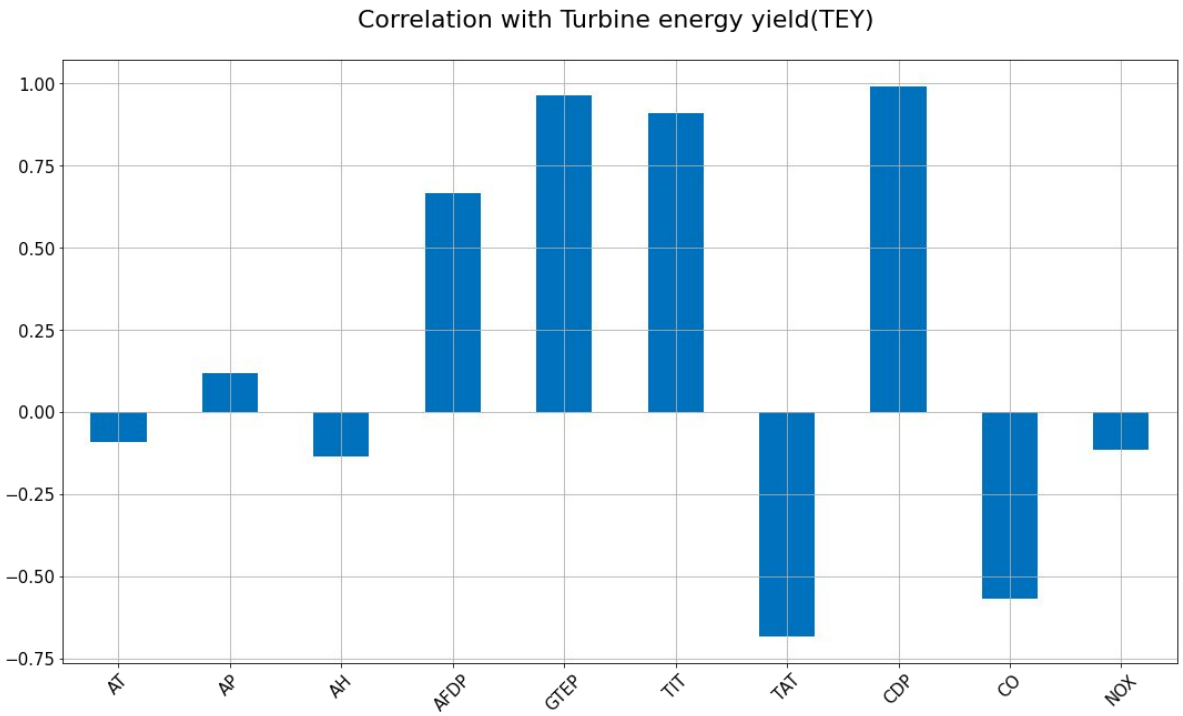
Data Preprocessing



Heat Map

- Graphical representation of multivariate data that is structured as a matrix of columns and rows.
- Useful in describing correlation among several numerical variables, visualizing patterns and anomalies.

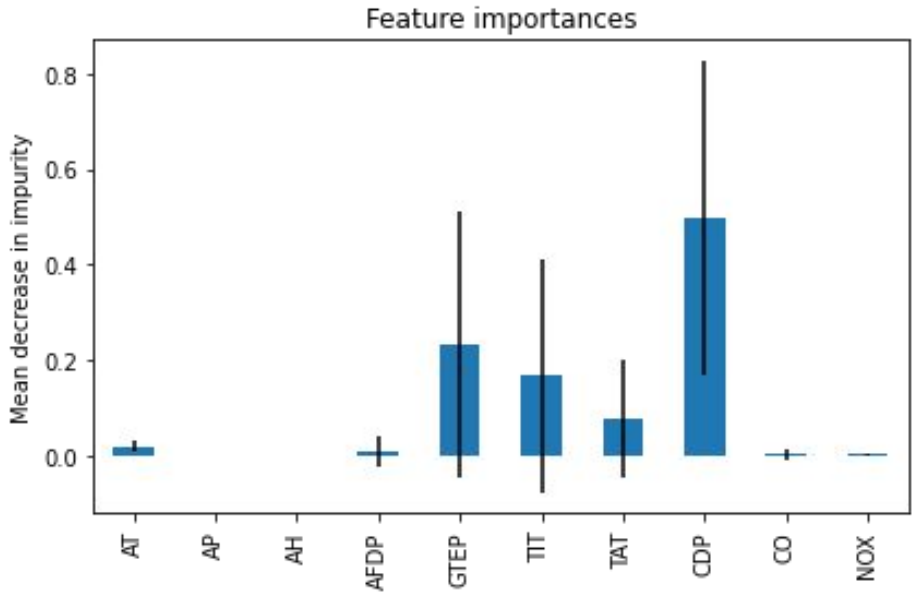
Correlation



Represents positive, negative relation between each feature and target.

- Negative Correlation
  -Inverse correlation
- Positive Correlation
  -Directly correlated

Feature selection



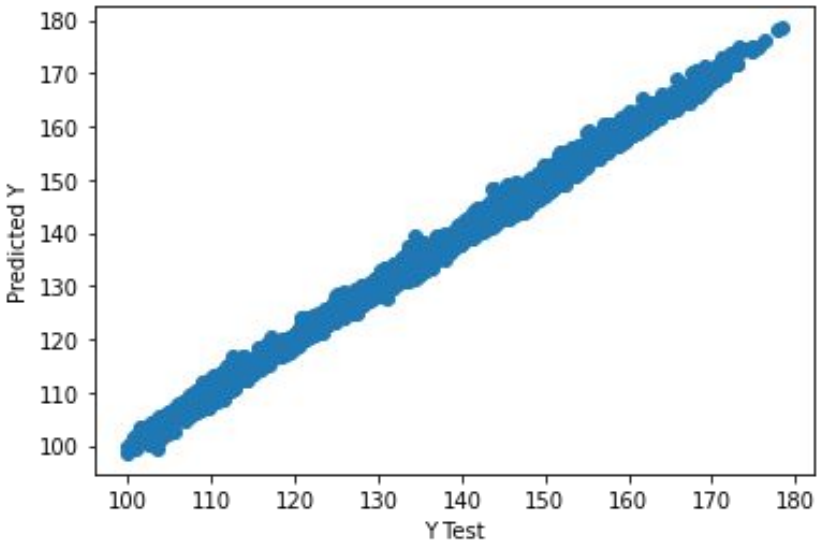
Feature importances are provided by the fitted attribute feature_importances_ and they are computed as the mean and standard deviation of accumulation of the impurity decrease within each tree.

Models

- Linear Regression
- Decision Tree
- Random Forest
- XGboost
- Neural Network

Linear Regression

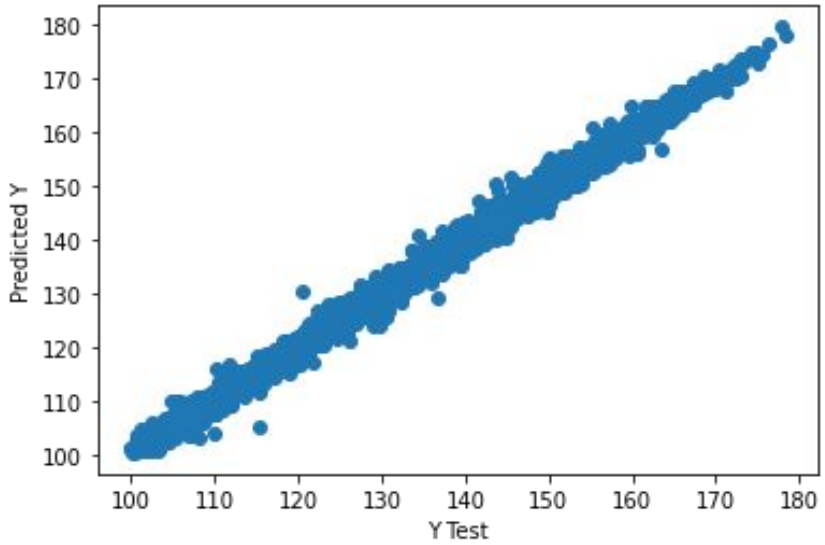
```
R2 score= 0.996334214930542
Root Mean Square Error (RMSE)= 0.9050421069525848
Mean Square Error (MSE)= 0.7436806483656044
Mean Absolute Error (MAE)= 0.951337010187549
```



Linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x). So, this regression technique finds out a linear relationship between x (input) and y(output). Hence, the name is Linear Regression.

Decision Tree Regression

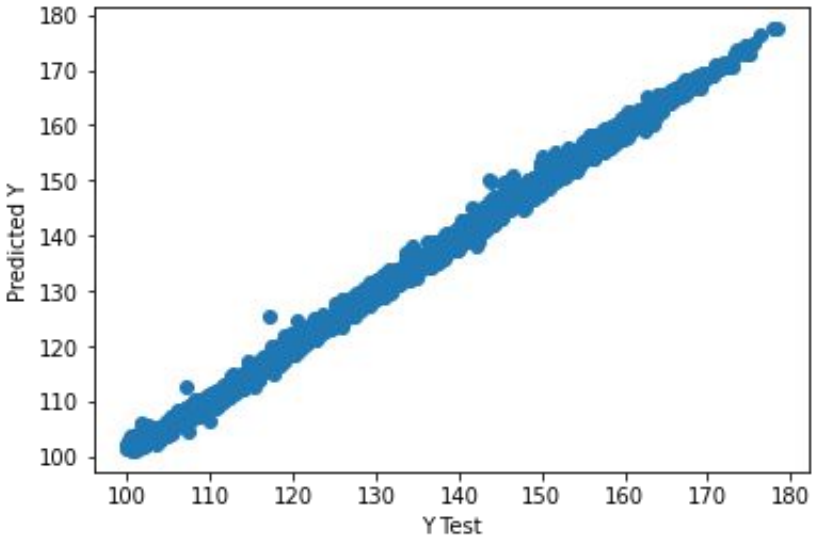
```
R2 score= 0.9964568305758154
Root Mean Square Error (RMSE)= 0.874769649664186
Mean Square Error (MSE)= 0.6430096206208024
Mean Absolute Error (MAE)= 0.9352912111552134
```



Decision tree regression observes features of an object and trains a model in the structure of a tree to predict data in the future to produce meaningful continuous output. Continuous output means that the output/result is not discrete, i.e., it is not represented just by a discrete, known set of numbers or values.

Random Forest

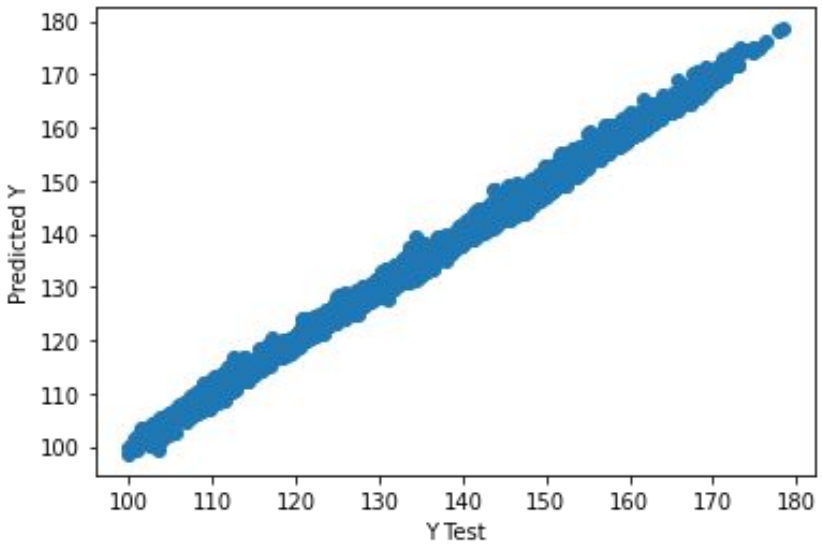
```
R2 score= 0.99837747562315
Root Mean Square Error (RMSE)= 0.4005834637826478
Mean Square Error (MSE)= 0.44277050444697186
Mean Absolute Error (MAE)= 0.6329166325691306
```



- Random Forest is an ensemble technique that uses multiple decision trees and a technique called as Bootstrap and Aggregation, commonly known as bagging. This helps to combine multiple decision trees in determining the final output rather than relying on individual decision trees.
- It has multiple decision trees as base learning models. We randomly perform row sampling and feature sampling from the dataset forming sample datasets for every model.

XGboost

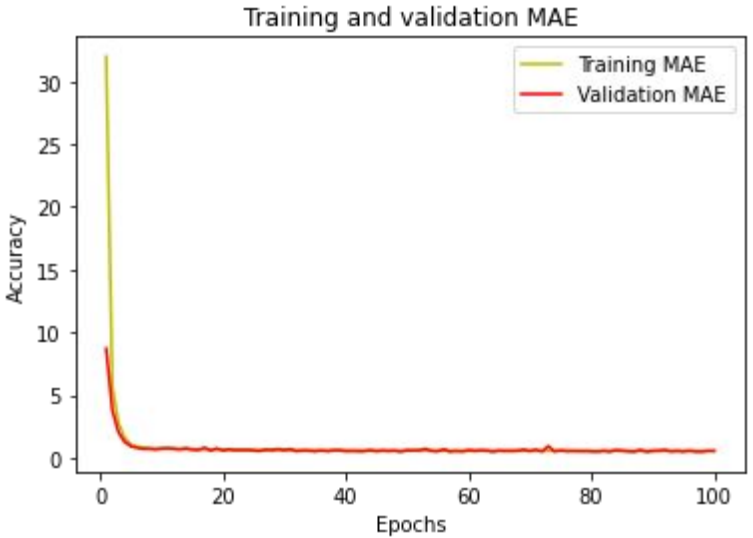
```
R2 score= 0.9987905619281734
Root Mean Square Error (RMSE)= 0.2985969880979626
Mean Square Error (MSE)= 0.3858000046276062
Mean Absolute Error (MAE)= 0.546440287769819
```

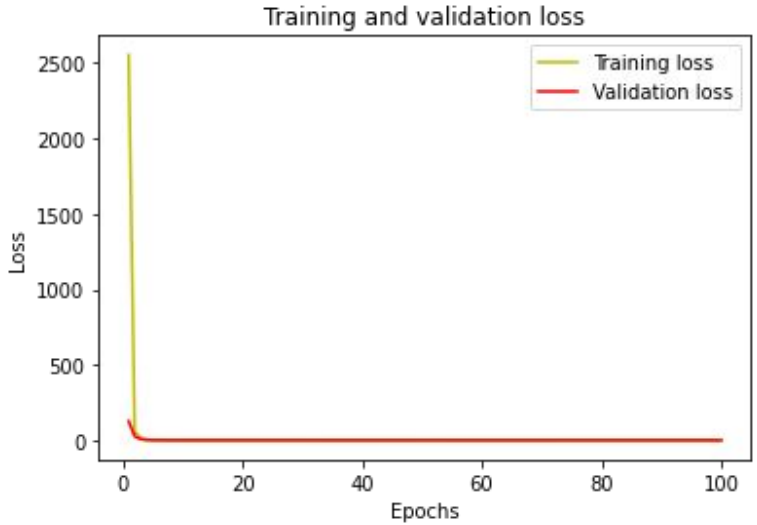


The objective function contains loss function and a regularization term. It tells about the difference between actual values and predicted values, i.e how far the model results are from the real values. Ensemble learning involves training and combining individual models to get a single prediction, and XGBoost is one of the ensemble learning methods. XGBoost expects to have the base learners which are uniformly bad at the remainder so that when all the predictions are combined, bad predictions cancels out and better one sums up to form final good predictions.

Neural Network

```
Mean squared error (MSE):  0.36022970528820664
Mean absolute error (MAE):  0.44616572637263746
R2 Score:  0.998533686560002
```





Result

Model	R2 Score	MAE	MSE
Linear Regression	0.9963342149	0.7436806484	0.905042107
Decision Tree	0.9964568306	0.6430096206	0.8747696497
Random Forest	0.9983774756	0.4427705044	0.4005834638
XGboost	0.9987905619	0.3858000046	0.2985969881
Neural Network	0.9985336866	0.4461657264	0.3602297053

Prediction

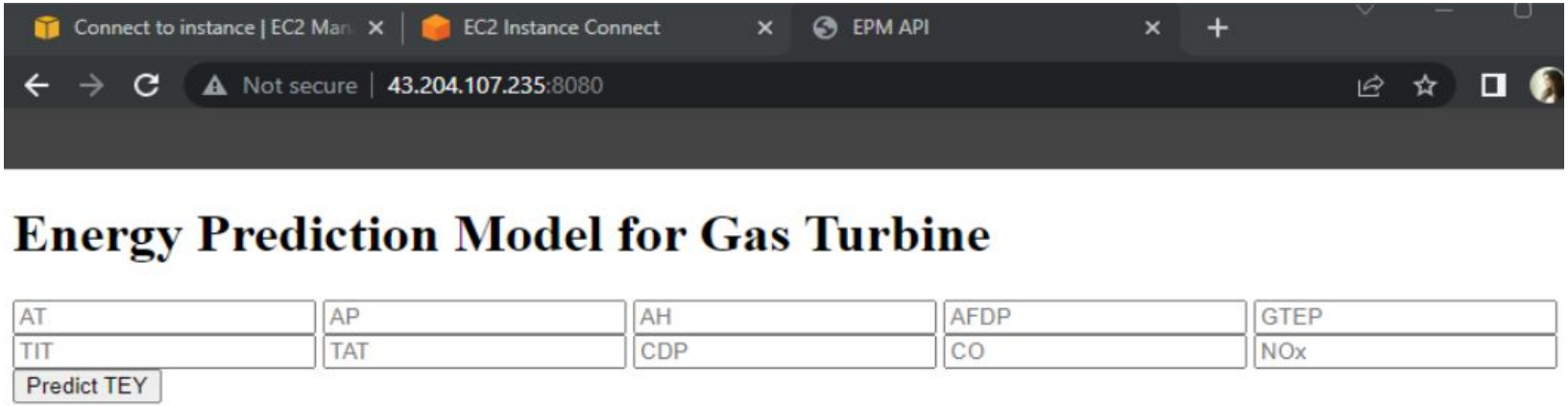
XGBoost Model

```
Mean Absolute Error (MAE): 0.3858000046276062
Mean Squared Error (MSE): 0.2985969880979626
Root Mean Squared Error (RMSE): 0.5464402877698l9
R2 Score: 0.9987905619281734
```

	Actual	Predicted	% Error
30175	129.98	130.158910	-0.137644
26404	127.89	127.852027	0.029692
11679	143.80	143.527853	0.189254
21622	155.10	155.600847	-0.322919
15549	133.15	133.433399	-0.212842
34083	110.06	109.869574	0.173020
26999	120.58	121.114361	-0.443159
31496	134.20	134.318019	-0.087943
28201	130.97	131.030923	-0.046517
17453	132.55	132.708656	-0.119695

Deployment on AWS

AWS


Connect to instance | EC2 Man
EC2 Instance Connect
EPM API
Not secure
43.204.107.235:8080
Energy Prediction Model for Gas Turbine
AT
AP
AH
AFDP
GTEP
TIT
TAT
CDP
CO
NOx
Predict TEY

AWS

Energy Prediction Model for Gas Turbine

4.5878	1018.7	83.675	3.5758	23.979
1086.2	549.83	11.898	0.32663	81.952

Predict TEY

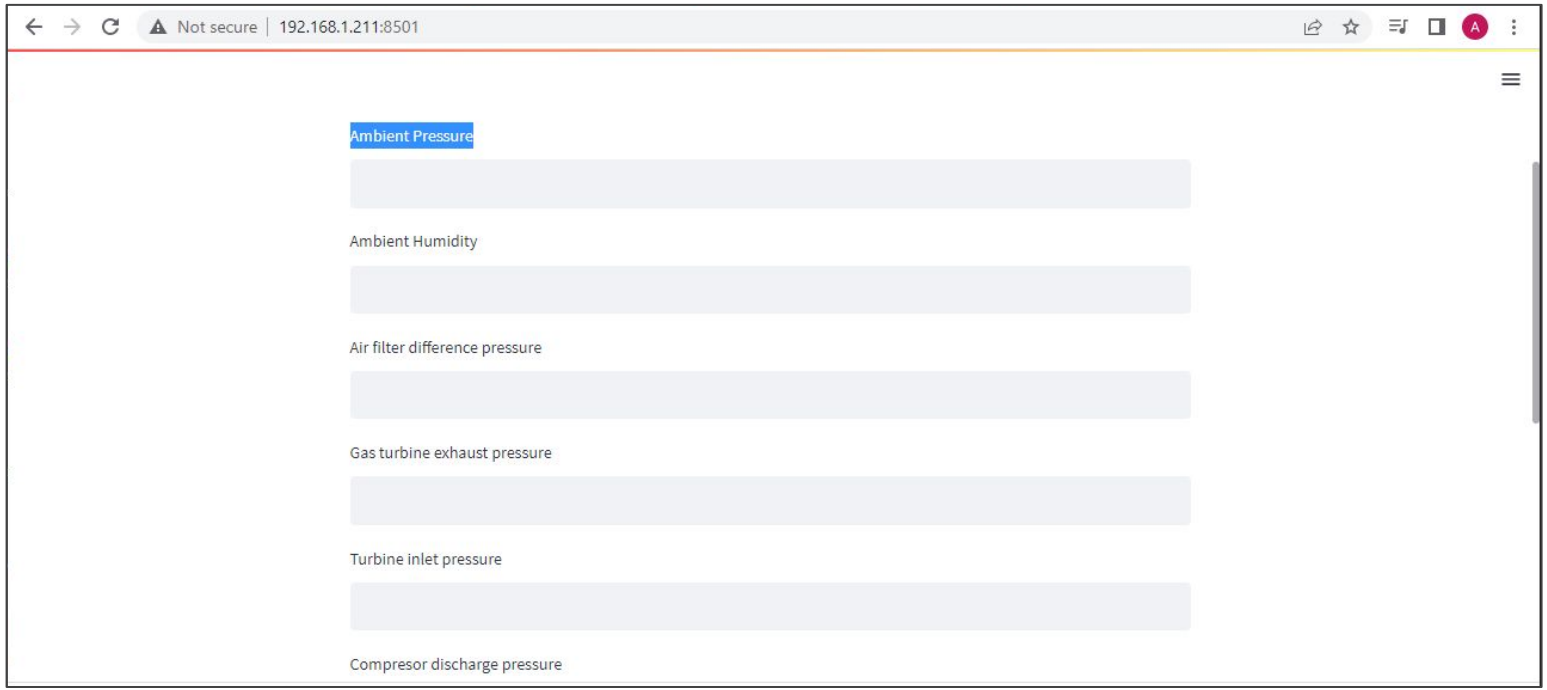
Energy Prediction Model for Gas Turbine

4.5878	1018.7	83.675	3.5758	23.979
1086.2	549.83	11.898	0.32663	81.952

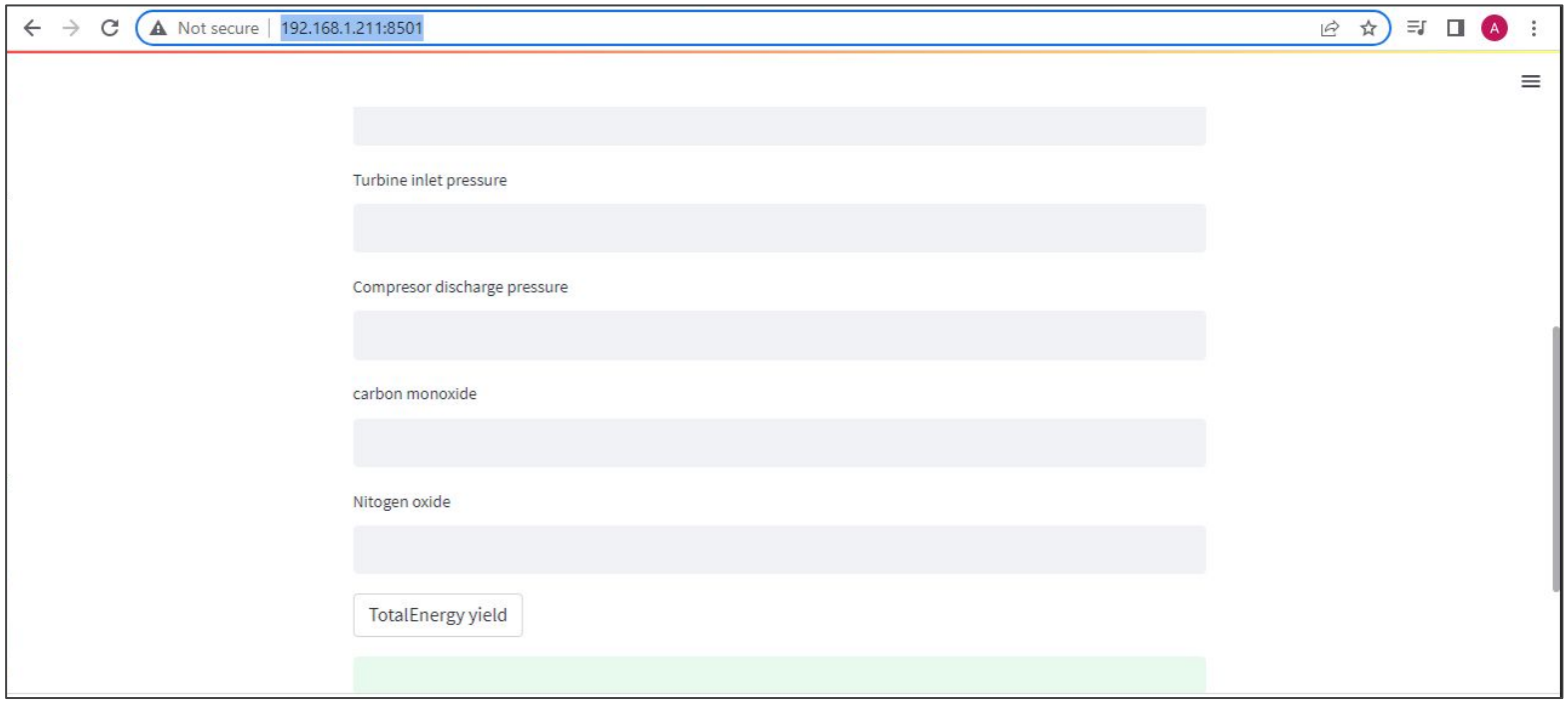
Predict TEY

Turbine Energy Yield 134.64

Streamlit


Not secure | 192.168.1.211:8501
Ambient Pressure
Ambient Humidity
Air filter difference pressure
Gas turbine exhaust pressure
Turbine inlet pressure
Compresor discharge pressure

Streamlit


Not secure
192.168.1.211:8501
Turbine inlet pressure
Compresor discharge pressure
carbon monoxide
Nitogen oxide
TotalEnergy yield

Dashboard

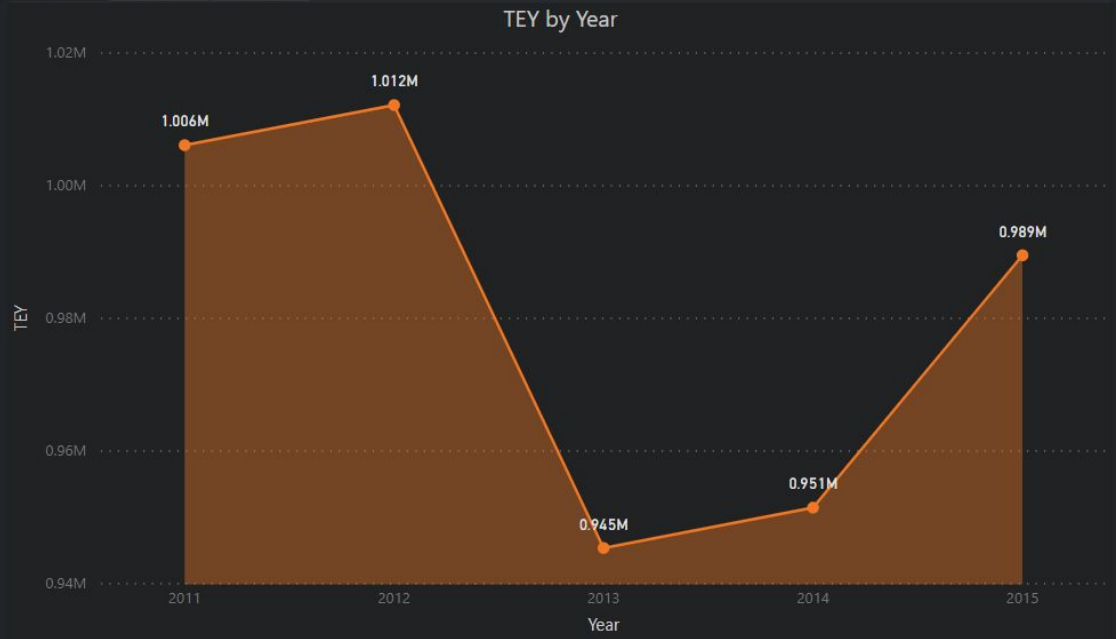
Gas Turbine Parameters Dashboard

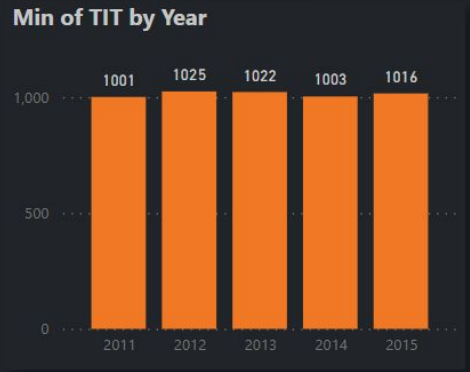
Overview | Details | Correlation | Trend

Year: All

179.50	119.91	44.10
Max of TEY	Max of NOX	Max of CO

TEY | CO | NOx





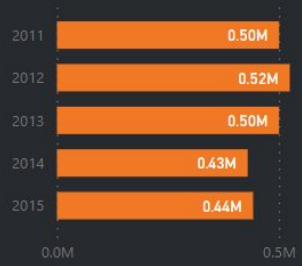
Gas Turbine Parameters Dashboard

Overview | Details | Correlation | Trend

TEY by Year

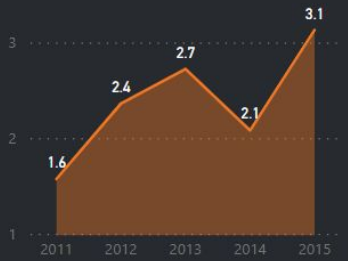
CO by Year

NOX by Year

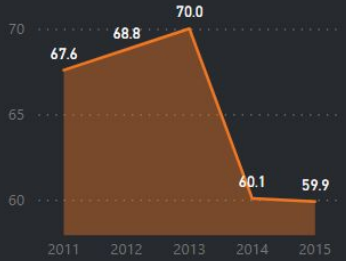


Average of TEY by Year

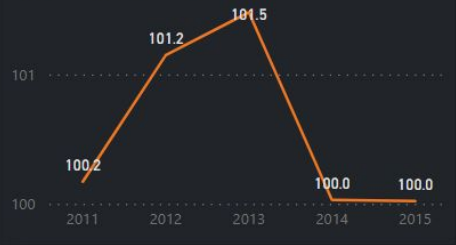
Average of CO by Year



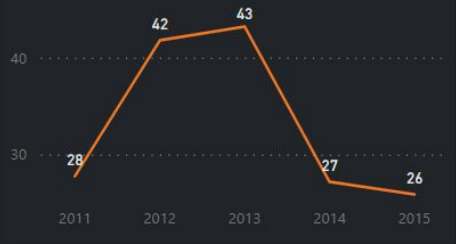
Average of NOX by Year



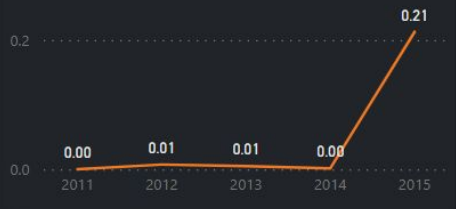
Min of TEY by Year



Min of NOX by Year



Min of CO by Year



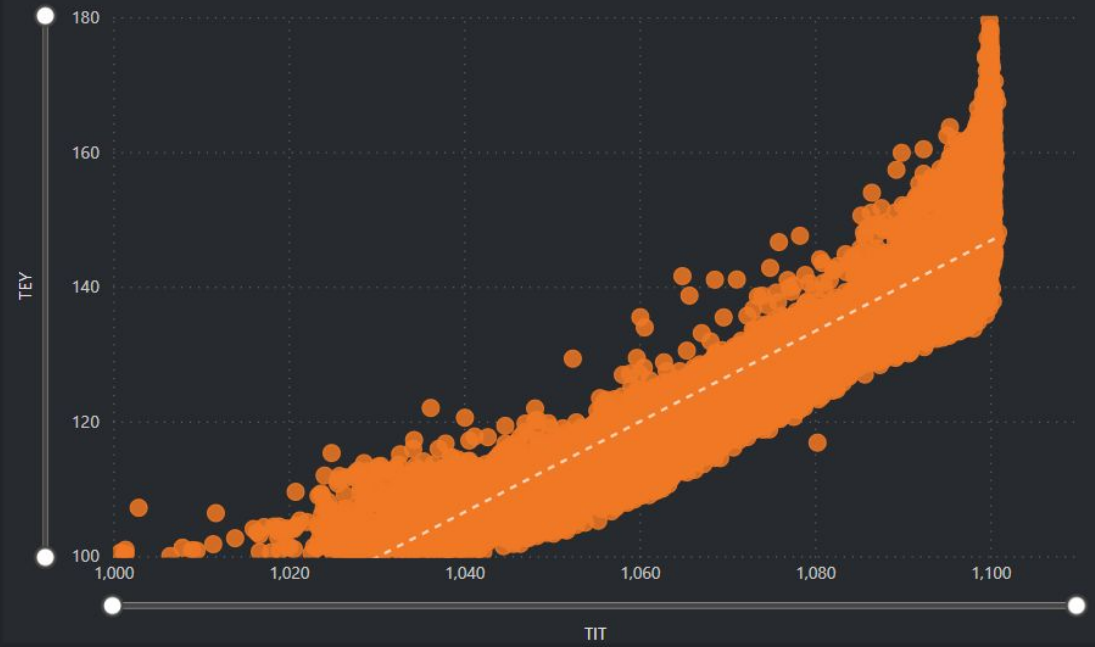
Gas Turbine Parameters Dashboard

Overview | Details | **Correlation** | Trend

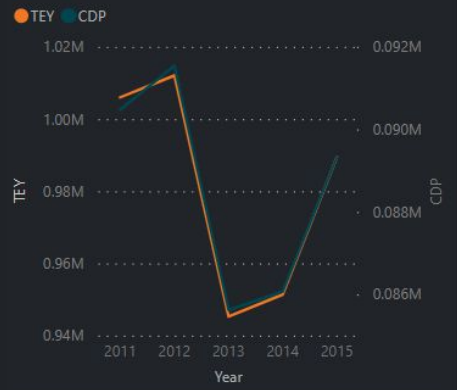
1st Parameter: TIT

2nd Parameter: TEY

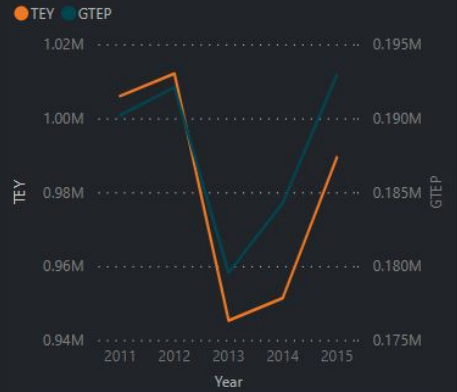
TIT and TEY



TEY and CDP by Year



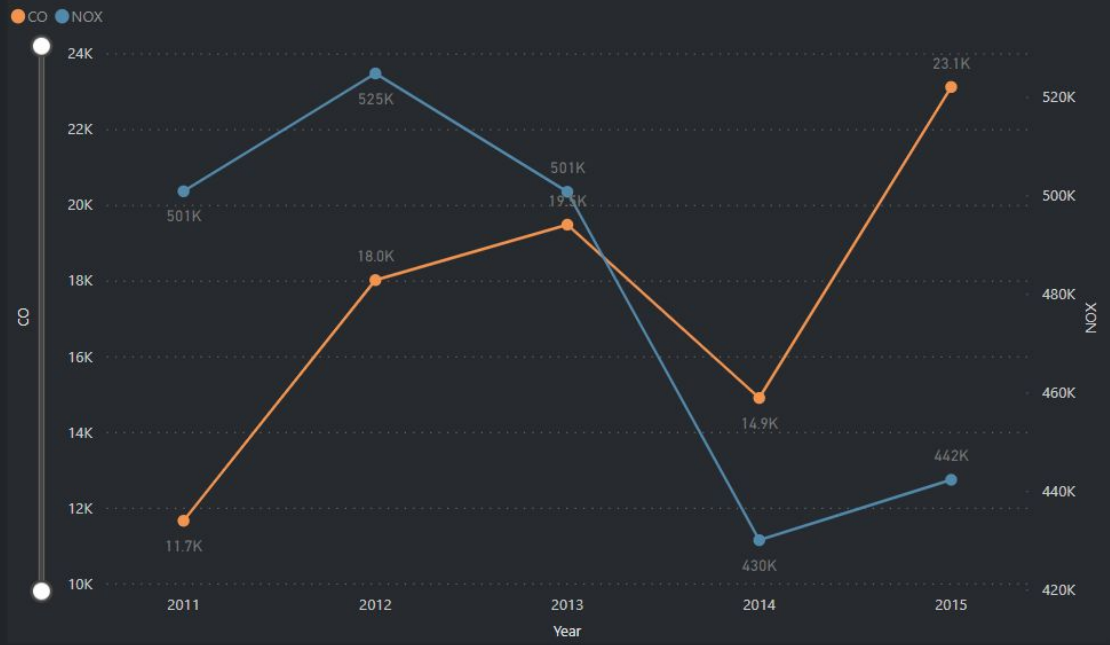
TEY and GTEP by Year



Gas Turbine Parameters Dashboard

Overview | Details | Correlation | **Trend**

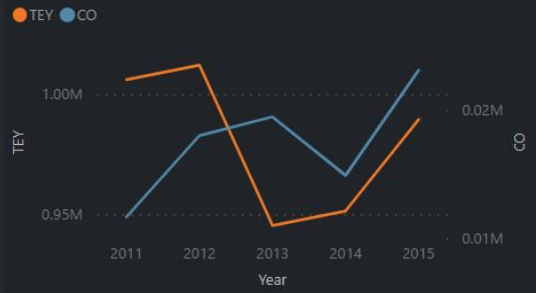
CO and NOX by Year



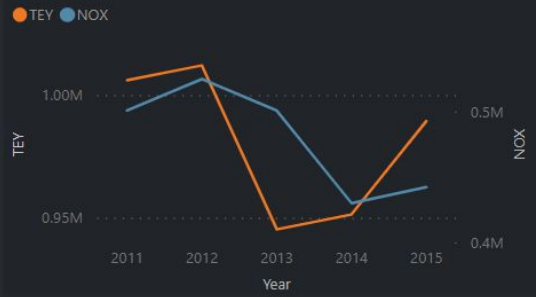
Max Limit of CO
150 mg/m3

Max Limit of NOx
50000 mg/m3
(50 ppm)

TEY and CO by Year



TEY and NOx by Year



Thank You

Any questions?