Nilansh Khurana, et. al. International Journal of Engineering Research and Applications
www.ijera.com
ISSN: 2248-9622, Vol. 10, Issue 9, (Series-III) September 2020, pp. 56-65

RESEARCH ARTICLE OPEN ACCESS

Prediction of Air Quality and Air Quality Level of India Using Machine Learning Algorithms

Nilansh Khurana*, Nandini Chauhan**
**(Department of Computer Science and Engineering, Manav Rachna University, Faridabad-43*
*** (Department of Computer Science and Engineering, Manav Rachna University, Faridabad-43*

ABSTRACT
As per the list of 2019, India ranks 5 among the most polluted countries across the globe and 21 out of 30 most polluted cities in the world are in India itself. This comes as a serious threat to human health and forces us to focus upon the air quality; timely and effective monitoring of which can contribute to pollution control and improving human health. In this paper, we use various machine learning techniques to forecast the AQI of a certain region. Publicly available datasets are used to determine the concentration levels of various factors affecting the quality of air like PM10, PM2.5, SO2, and so on and predict the AQI based on it. The model provides us with a satisfactory level of accuracy and very minimal errors. Through this paper, we also intend to imply that machine learning can come in handy while dealing with problems relating to the environment.
Keywords - Air pollutants, Machine Learning Techniques, Air Quality, AQI, Forecasting

---

Date of Submission: 29-08-2020 Date of Acceptance: 14-09-2020

I. INTRODUCTION

Air is in itself the major requirement, without it, nothing can ever exist; all living and non-living rely on it. But today the air around is not pure, it's polluted, so with every breath we take, an enormous amount of unwanted trash goes into causing a great impact on human health. Thus, the quality of air plays a major role, and to estimate it, various factors are put into consideration. To determine the air quality, AQI (Air Quality Index) is used. AQI is a concept used to determine the quality of air for a region and predict how worse it may become.[1] This helps in determining how far - fetched and worse the conditions are at a place, how that can be controlled, and what impact it is having on the health of the people there. As seen according to the Chinese Standard GB3095-2012[2], six major pollutants are considered while calculating AQI, that being: Ozone, Nitrogen Dioxide, Particulate matter, Carbon monoxide, Sulphur Dioxide and Inhalable particles.

For the calculation of AQI, generally, 7 measures are used: PM2.5, PM10, SO2, NOx, NH3, CO, and O3 [3]. The maximum of the last eight hours is used as a value for CO and O3 and the rest five, the average value in the last 24 hours is used**.** These measures are then converted to sub-index that are based on pre-defined groups. The maximum of these sub-indices is the final AQI with a constraint that three out of the seven measures are available and one of those three is either PM2.5 or PM10. The value of AQI lies within a range of 0-1000 (it's rare to find any value over 1000) with 0-50 having a minimal impact whereas above 400 causing respiratory impact even on the healthy ones [4]. The AQI bucket is predefined and is interpreted as follows:

Table 1.1: Air Quality Level Classification

0-50	Good	201-300	Poor
51-100	Satisfactory	301-400	Very Poor
101-200	Moderate	Above 400	Severe

Many researchers have used machine learning techniques both in short- term, as well as long-term prediction of the air quality, and many of them, have been successful in doing so. In Santiago, an hourly concentration of PM2.5 was predicted by Perez et al using a multilayer neural network [5]. In Xingtai, two hybrid models were proposed by Zhu et al to determine the AQI wherein he considered only one major indicator and ignored the rest of the pollutants (PM10, PM2.5, and others). The two models were: Empirical mode decomposition-Support vector regression (EMD- SVR) and Empirical mode decomposition- intrinsic mode functions (EMD- IMF). On comparison, the EMD-SRV had the highest accuracy of about 80% [6]. A model based on a recursive neural network was used by Bianconfiore et al to forecast the concentrations of PM10 on 1, 2, and 3 days. The model was accurate on 95% of the days; however, the

percentage of the false positive was up to 30% that demonstrated the limitations of the neural network model [7].

After analyzing the prior work, we observe that these models are based upon time series and require at least 8-10 hrs. to formulate and calculate AQI making it very time consuming and slow. AQI does not always depend on trends but also on the factors that affect it, thus time series forecasting may fail at some parts. Taking inspiration from the same, we aim to make it real-time using machine learning techniques so that the model can calculate the real-time AQI considering all the affecting factors and forecast the AQI for the next few days based upon it. We have performed EDA to analyze and visualize the relationship between features and predictors as well. We used RMSE, MSE, MAPE for calculation, and have also calculated R2 score to find the best fit.

The rest of the paper has been organized as per the following sections: Section 2 shows the work that was done previously in the field and the literature review; in section 3, we explain about machine learning and the models used. Starting in section 4, we introduce our model and present the working methodology. The conclusions are drawn in the final section accompanied by the future scope of the model.

II. RELATED WORK

A lot of work has been done in the field by various authors in the various parts of the earth with each concluding with their learnings and observations. This section reviews the analysis and the work of various authors across the globe in the field of AQI calculation and prediction using machine learning.

In 2000, Perez et al used a multilayer neural network to predict an hourly concentration of PM2.5 for the next 24 hours in Santiago. While experimenting, it was observed and hence concluded that the prediction errors in the model raised from about 30% in early hours to about 60% in late hours [5]. Kalapandias, in 2001, modeled a short-term air quality prediction model using a case-based classifier. He explained the effects on air pollution through the meteorological features only (such features include temperature, solar radiation, precipitation, humidity, wind [8]. In early 2002, Fuller, GW, and HW used an empirical approach to predict the daily mean PM10 concentrations in London and South East England [9]. In 2012, Jose Joan used fuzzy logic and auto-aggressive models to assess and predict the air quality of a certain region [10]. Earlier in 2010, Kurt and Oktay used neural networks by modeling geographic models and connections to predict the daily concentration levels of SO2, CO, and PM10 3 days in advance [11]. As the magnitude of the numeric data is ignored, the process of conversion of regression tasks to classification tasks is problematic and hence, is consistently inaccurate.

As many of the researches have been made to focus on implementing the machine learning algorithms and predicting the air quality, other researchers have worked on predicting the concentration levels of the pollutants. Corani, in 2005, tried to predict hourly concentration levels of O3 and PM10 based on the previous day data by training the neural networks in Milan. He mainly compared the performances of Pruned neural networks and feed-forward neural networks [12]. In Xingtai, 2017, two hybrid models were proposed by Zhu et al to determine the AQI wherein he considered only one major indicator and ignored the rest of the pollutants (PM10, PM2.5, and others). The two models were: Empirical mode decomposition- Support vector regression (EMD-SVR) and Empirical mode decomposition- intrinsic mode functions (EMD- IMF). On comparison, the EMD- SRV had the highest accuracy of about 80% [6]. Later in the same year, a model based on a recursive neural network was used by Bianconfiore et al to forecast the concentrations of PM10 on 1, 2, and 3 days. The model was accurate on 95% of the days; however, the percentage of the false positive was up to 30% that demonstrated the limitations of the neural network model [7].

The researches and the implementation of machine learning in these researches are a benefiting factor for mankind but there still exist some gaps that need focus and be filled by the researchers. All the models and researches mentioned above consider only a few of the features and the machine learning models may miss out on these features or even overweigh them. Since air quality depends on the meteorological features and also the parameters are wide enough, AQI needs consideration of both pollutant levels and meteorological parameters.

Most of the previous models use time series to predict the AQI and therefore take more time and hence, not reliable for a longer period. We have proposed a model that does not solely depend on time series but uses RMSE, MSE [13-14], MAPE [15] for calculation, and prediction. We have also performed EDA to analyze and visualize the relationship between features and predictors as well. R2 score has also been calculated to evaluate the best fit.

III. MACHINE LEARNING

Machine learning is a branch of science and an application of artificial intelligence that without programming, can improve and learn from experiences automatically [16]. It is the science of

making computers learn and act like humans by feeding them data in a manner that it seems like real-world interactions and observations.

Machine learning implementation is not only limited to computers and machines and its uses are not only for scientific studies but it is also used in various other sectors and fields. Banks and other businesses use it for financial services to have an insight into the data and prevent fraud. Government agencies use machine learning to maintain a huge amount of public data and mind the insights into the various sources to this data. Various other such fields are today very much dependent on machine learning like retail, health industry, transportation, and oil and gas.

Regression and Classification: Both of these algorithms are supervised machine learning algorithms. Even though they both are used for prediction in machine learning, they are used for different types of problems. A problem is said to be a regression problem when there exists a real or continuous value for the output variable such as "weight", "age" or "price" whereas a problem is considered as a classification problem when the output variable is a category such as "green" or "yellow", "true" or "false" or "right" or "left". A classification model uses the observed values and tries to draw some conclusions out of it. If one or more inputs are given, the classification model attempts to predict the value of one or more outcomes [16]

Regression involves the prediction of continuous values in an ordered manner using algorithms like linear regression and random forest whereas the latter involves the prediction of discrete values in an unordered manner using algorithms such as decision tree and logistic regression.

Models Used:

1. Lasso Regression: LASSO is a linear regression model and is short for the Least Absolute Shrinkage and Selection Operations. It performs L1 regularization and uses shrinkage. It is a regression model that encourages simple and sparse models i.e. models with fewer parameters or multicollinearity

2. Kernel Ridge Regression (KRR): This regression model is a non- parametric form of ridge regression that combines ridge regression with classification. It uses squared error loss combined with L2 regularization. KRR is a method of interpolation.

3. Elastic Net Model Regression (ENMR): It is a regularized method that uses both L1 and L2 regularization i.e. Lasso regression and Ridge regression. There exists a mix ratio 'r' that can be controlled. If the value of r is equal to 0, it is equivalent to ridge regression, and if the value of r is equal to 1, equivalent to Lasso regression.

4. Gradient Boosting Regressor (GBR): Boosting, also called additive model in machine learning, is a method to combine multiple simple models to make a single composite model. To minimize the loss, the algorithm uses gradient descent, and hence the term "gradient boosting". GBR is a regression and classification technique that produces a prediction model in the form of an ensemble of a weak prediction model. GBR also helps in providing higher flexibility, less pre-processing of data, and better accuracy as when compared to other techniques like linear regression.

5. XGBoost Regression: XGBoost stands for Extreme Gradient Boosting and is thus, the implementation of the Gradient Boosted decision tree. It improves model generalization capabilities using advanced regularization and is designed for computational speed and model performance.

6. XGBoost Classifier: XGBoost classification model is known as XGBoost Classifier. Its uses include solving regression problems, problems of ranking, and user-defined prediction problems.

7. Decision Tree Classifier: In this, a decision tree is built to create a classification model; a test is specified on the attribute by each node in the tree. It is also considered to be one of the fastest ways of identifying the most significant variable and the relation between variables.

8. SVM Classification: SVMs are Support Vector Machines that are a discriminative and a non-probabilistic binary linear classifier. The SVM classifier provides great accuracy and can work in high dimensional space; it also uses very little memory.

9. Random Forest: Random forest, or well known as Random Decision Forest is flexible and easy to use an algorithm that produces a benefitting result. The term "Forest" is an ensemble of decision trees that are trained using the bagging method. Random Forest holds one of the major benefits and that is its versatility: it can be used

for both classification and regression tasks. They are simple, flexible, and hard to beat when it comes to performance but also take a long time to develop.
[17-20]

IV. DATA COLLECTION

Data collection is the first step of the Machine Learning/ Data Analysis process. Without data, we couldn't do anything as it is the major fuel. For this paper, the data is collected from the major data that has been made publicly available by the Central Pollution Control which is the official portal of Government of India [21]. The data contains factors that affect air quality; these factors are analyzed and used to predict the AQI Level.

The features are:
1. PM2.5 or Particulate Matter 2.5
2. PM10 or Particulate Matter 10
3. NO or Nitric Oxide
4. NO2 or Nitrogen Dioxide
5. NOx or Nitrogen Oxide
6. SO2 or Sulphur Dioxide
7. CO2 or Carbon Dioxide
8. NH3 Nitrogen Trihydrate (Ammonia)
9. O3 or Oxone
10. Benzene
11. Toluene

It also contains a predictor which is AQI.
The data is recorded daily by the Government of India from 25 major cities within the country [21], namely:

- Ahmedabad
- Aizawl
- Amaravati
- Amritsar
- Bengaluru
- Bhopal
- Brajrajnagar
- Chandigarh
- Chennai
- Kolkata
- Lucknow
- Mumbai
- Talcher
- Delhi
- Ernakulam
- Gurugram
- Guwahati
- Vishakhapatnam
- Hyderabad
- Jaipur
- Jorapokhar
- Kochi
- Patna
- Shillong
- Thiruvananthapuram

V. EXPLORATORY DATA ANALYSIS

The next step and the first step after collecting data is to perform EDA on it. It is one of the most important steps as it helps an analyst get some insights into the data such as the number of null values, data distribution over different categories, range of data, and different statistical values of data. In this research paper, we performed some EDA just after collecting the data. Our first step was to get an overview of the data to see its data type and count as well as to get an idea of null values in each of the features and also the predictor.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 28840 entries, 0 to 28839
Data columns (total 15 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   City        28840 non-null  object
 1   Date        28840 non-null  object
 2   PM2.5       24250 non-null  float64
 3   PM10        17795 non-null  float64
 4   NO          25313 non-null  float64
 5   NO2         25266 non-null  float64
 6   NOx         24659 non-null  float64
 7   NH3         18635 non-null  float64
 8   CO          26784 non-null  float64
 9   SO2         24989 non-null  float64
 10  O3          24821 non-null  float64
 11  Benzene     23220 non-null  float64
 12  Toluene     20802 non-null  float64
 13  AQI         24201 non-null  float64
 14  AQI_Bucket  24201 non-null  object
dtypes: float64(12), object(3)
```

Figure 5.1: Overview Information of Features and Predictor

The next step was to look for the exact number of null values in each column/feature.

PM2.5	4590
PM10	11045
NO	3527
NO2	3574
NOx	4181
NH3	10205
CO	2056
SO2	3851
O3	4019
Benzene	5620
Toluene	8038
AQI	4639
AQI_Bucket	4639

Figure 5.2: Number of Null Values in Each Column

Plotting those null values:

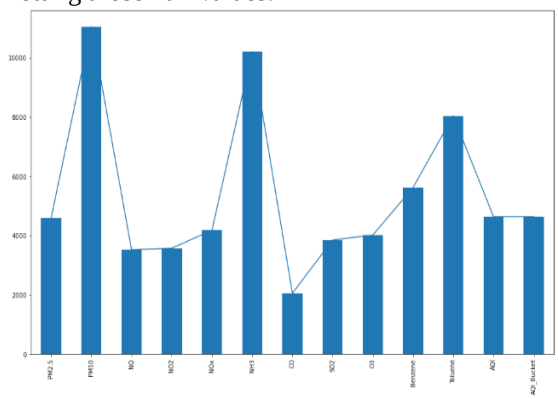


Figure 5.3: Visualization of Null Values

Our next step was to look for statistical values of each feature so, we performed statistical analysis on continuous features for their standard deviation, mean, median, and interquartile ranges.

	PM2.5	PM10	NO	NO2	NOx	NH3	CO	SO2	O3	Benzene	Toluene	AQI
count	24250.000000	17795.000000	25313.000000	25296.000000	24659.000000	18635.000000	26784.000000	24999.000000	24821.000000	23220.000000	21982.000000	24201.000000
mean	68.508440	120.465929	17.848399	28.722892	32.538442	23.733286	2.285140	14.716054	34.498114	3.324055	8.903265	168.846453
std	65.207394	90.977792	22.970994	24.621160	31.947611	25.949930	7.047618	18.341909	21.829299	16.013089	20.245367	141.737744
min	0.040000	0.010000	0.020000	0.010000	0.000000	0.010000	0.000000	0.010000	0.010000	0.000000	0.000000	13.000000
25%	29.472500	58.925000	5.710000	11.820000	12.710000	8.700000	0.510000	5.660000	18.770000	0.140000	0.700000	83.000000
50%	49.620000	97.880000	10.080000	21.880000	23.730000	16.040000	0.900000	9.240000	30.830000	1.110000	3.090000	120.000000
75%	81.940000	152.175000	20.380000	37.820000	40.460000	30.210000	1.460000	15.510000	45.580000	3.142500	9.330000	212.000000
max	949.990000	1000.000000	390.680000	362.210000	467.630000	352.890000	175.810000	193.860000	257.730000	455.030000	454.850000	2049.000000

Figure 5.4: Statistical Description of all Data

Since we have pollutants as features in our dataset, considering that, the first thought that came into our mind was to find and plot which city has a greater number of which pollutants. So, we visualized that data of PM2.5, PM10, SO2, NO2 & NO.

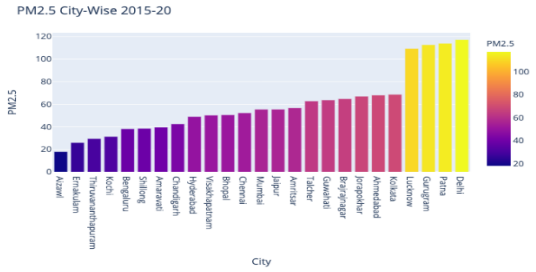


Figure 5.5: Visualization of amount of PM2.5 in each City.

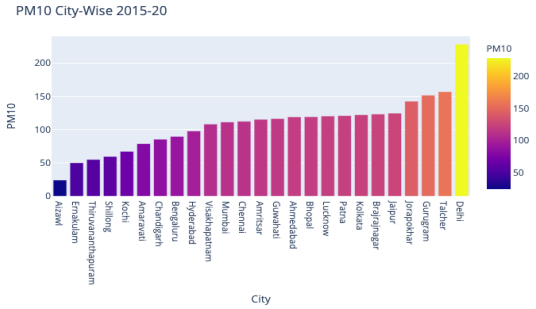


Figure 5.6: Visualization of amount of PM10 in each City.

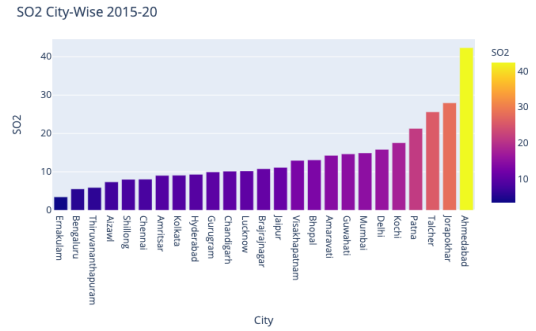


Figure 5.7: Visualization of amount of SO2 in each City.

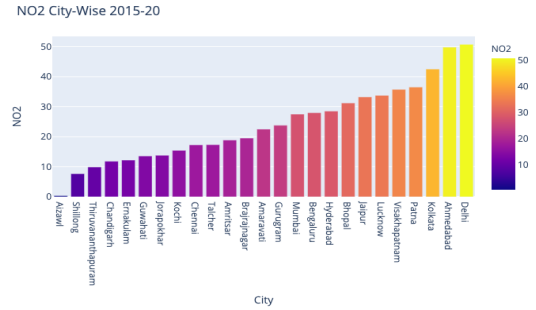


Figure 5.8: Visualization of amount of NO2 in each City.

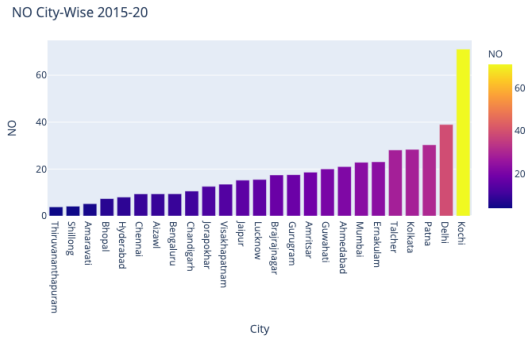


Figure 5.9: Visualization of amount of NO in each City.

For the next step, we performed univariate and bivariate analysis using pair plotting:

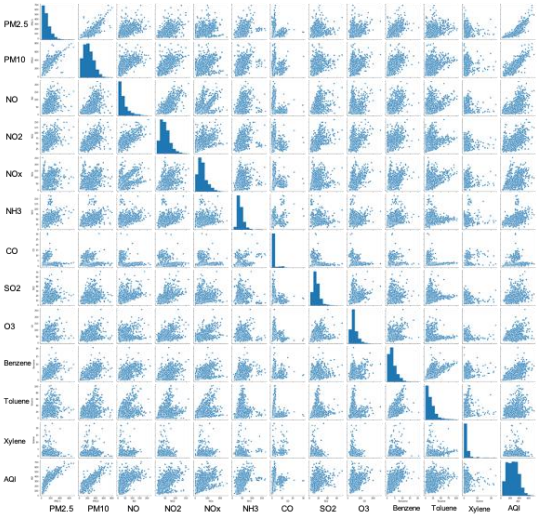


Figure 5.10: Pair Plotting

The next step is to check the correlation between features and predictors for easiness of feature selection and also to check multicollinearity.

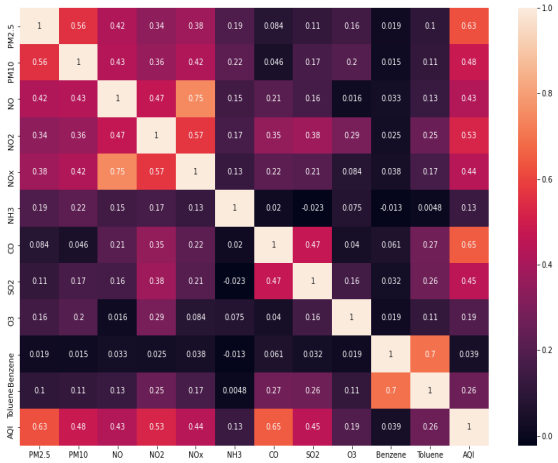


Figure 5.11: Correlation Heatmap

This ends our EDA and takes us onto the next step which is data preprocessing.

VI. DATA PREPROCESSING

The data needs to be preprocessed according to the EDA performed. Thus, the first step after performing EDA is to **impute null values.** We have used the **Mean** method for **Continuous values** and **Mode** for **Categorical (AQI Bucket)**.

After imputation, the non-numerical data is encoded. We only have 1 column with us i.e. AQI Bucket with no numerical data and we performed label encoding on it as it is Ordinal data.
Since each feature has its different units, the range of each feature differs. The next step is to normalize/standardize the data. We performed BoxCox Transformation [22].

Lambda value = 0.15

We performed box cox transformation only on data whose absolute skewness was greater than 0.75 to make stabilize the variance.

The next major step is to transform our predictor column. The original distribution of AQI (predictor) is shown below.

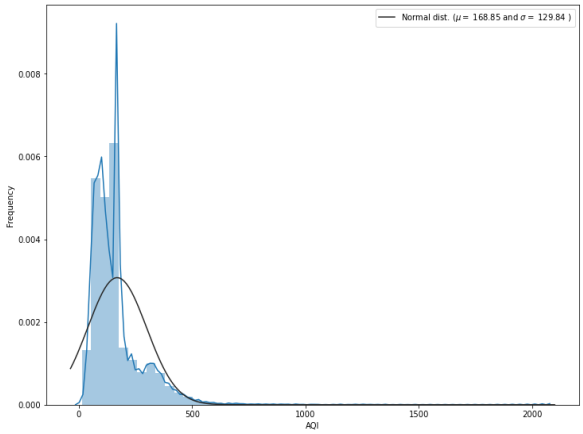


Figure 6.1: Distribution Plot AQI

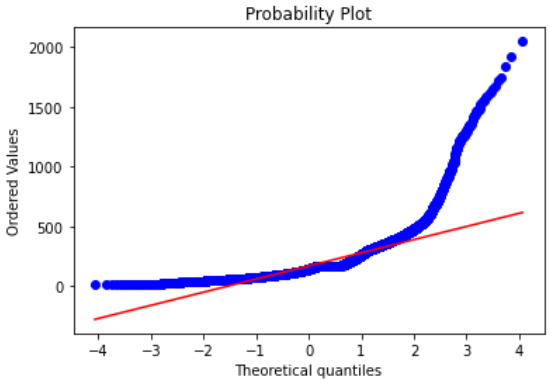


Figure 6.2: Probability Plot AQI

AQI is positively skewed as shown clearly in the distribution plot and it seems to have outliers and is deviated from a normal distribution which is clearly shown in probability plot. Probability plot ----- (data values Vs Z-score).

Keeping the above consideration in mind, we performed Log-Transformation and the results were perfectly satisfactory and are shown below.

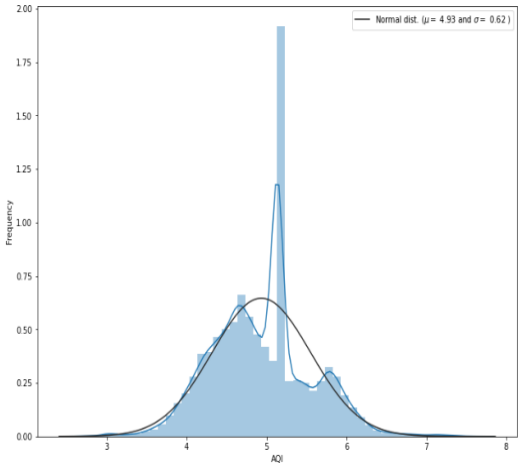


Figure 6.3: Distribution Plot AQI after Transformation

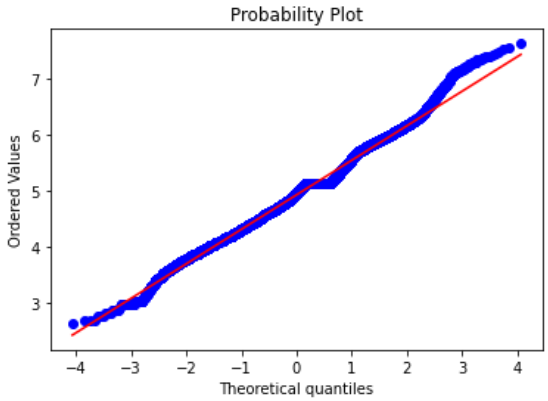


Figure 6.4: Probability Plot AQI after Transformation

The histogram can be seen to have a bell-like structure and the probability plot is also perfect.

VII. MODEL BUILDING AND EVALUATION

Model Building:

Many regression and classification models are used in this research with different hyperparameters to find the best output.

Starting with Regression:

Lasso: It is pipelined with Robust Scaler, with an alpha value of 0. 0005 and having a random state of 10.

ENet: It is also pipelined with robust Scaler having an alpha value of 0.0005 and l1_ration=0.9 and a random state of 10.

Kernel Ridge Regression: having alpha of 0.6, kernel type is polynomial, degree of 2, and the coefficient is 0.05.

XG Boosting: having the number of estimators=3200, gamma value=0.0468, with learning rate =0.05, and max depth of 3

Classification:

Decision Tree Classifier: It is kept simple without any hyperparameter.

XGB Classifier: with number of estimators=3000, gamma value=0, learning rate =0.05, max depth=5.

SVC: Here, one vs rest decision function shape is used.

Random Forest Classifier: Only max depth =5 is given as a parameter.

Model Evaluation:

We have used different Evaluation methods in regression and classification.

Regression:

1. Mean Absolute Percentage error: MAPE for each model is calculated as:
   Lasso= 13.28
   KRR= 3.7
   XGBoost=13.62
   ENet=13.29

2. Mean Square Error: MSE for each model is:
   Lasso=0.655
   KRR=0.06
   XGB=0.68

ENet=0.655

3. Root Mean Square Error: RMSE for each model is:
   Lasso=0.81
   KRR=0.25
   XGB=0.82
   ENet=0.81

4. R2 Score: R2 score for each model is also taken and its value in % is:
   Lasso=79%
   KRR=83%
   XGB=89%
   ENet=80%

Table 7.1: Error in Each Model

Model	**MSE**	**RMSE**	**MAPE**
Lasso	0.66	0.81	13.28
KRR	0.06	0.25	3.71
ENet	0.69	0.83	13.62
XGboost	0.66	0.81	13.30

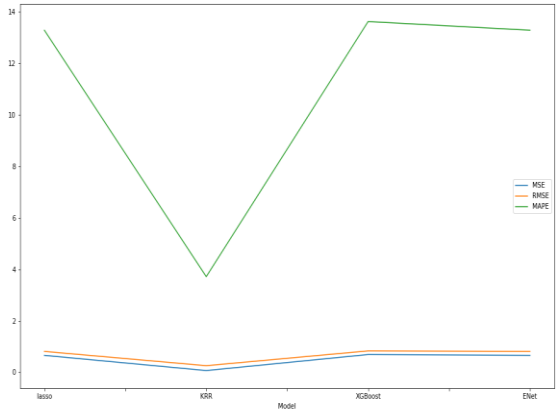
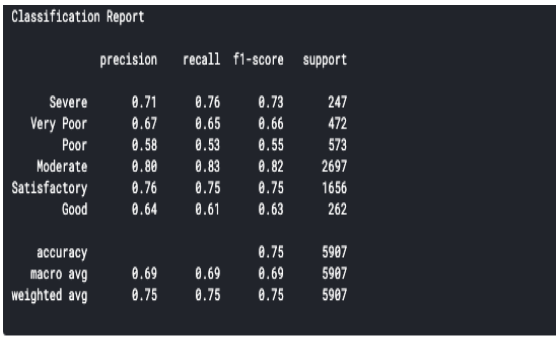


Figure 7.1: Visualizing Errors in Each Model

Classification: Accuracy, Precision, Recall, and F1 Score is calculated for each model with its confusion matrix as well as with classification report.

1. Decision Tree:
   Accuracy: 0.750126968004063
   Precision: 0.7476424238327595
   Recall: 0.750126968004063



Classification Report

	precision	recall	f1-score	support
Severe	0.71	0.76	0.73	247
Very Poor	0.67	0.65	0.66	472
Poor	0.58	0.53	0.55	573
Moderate	0.80	0.83	0.82	2697
Satisfactory	0.76	0.75	0.75	1656
Good	0.64	0.61	0.63	262
accuracy			0.75	5907
macro avg	0.69	0.69	0.69	5907
weighted avg	0.75	0.75	0.75	5907

Figure 7.2: Classification Report Decision Tree

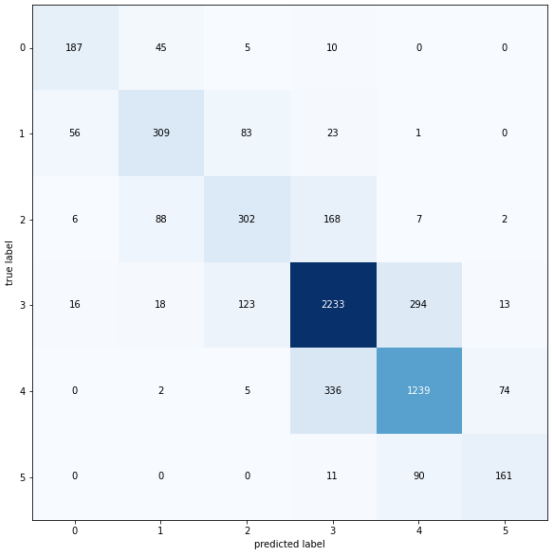
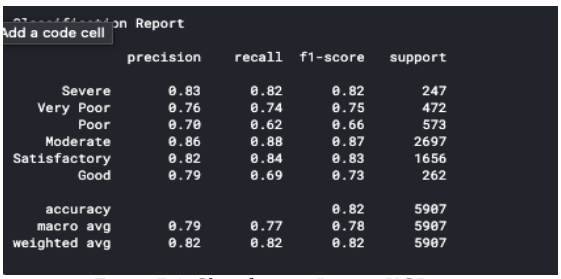


Figure 7.3: Confusion Matrix Decision Tree

2. XGBoost Classifier:
   Accuracy: 0.8219062129676655
   Precision: 0.819878418735058
   Recall: 0.8219062129676655



Add a code cell

Report

	precision	recall	f1-score	support
Severe	0.83	0.82	0.82	247
Very Poor	0.76	0.74	0.75	472
Poor	0.70	0.62	0.66	573
Moderate	0.86	0.88	0.87	2697
Satisfactory	0.82	0.84	0.83	1656
Good	0.79	0.69	0.73	262
accuracy			0.82	5907
macro avg	0.79	0.77	0.78	5907
weighted avg	0.82	0.82	0.82	5907

Figure 7.4: Classification Report XGBoost

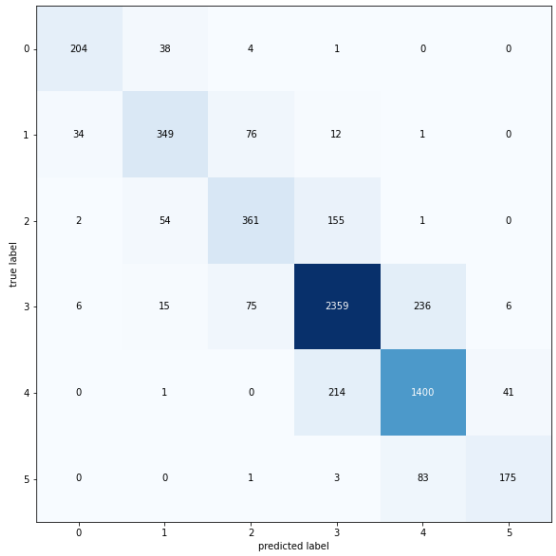
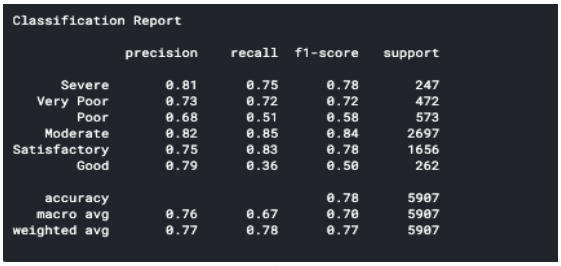


Figure 7.5: Confusion Matrix XGBoost

3. SVC:
   Accuracy: 0.776705603521246
   Precision: 0.7748304408289808
   Recall: 0.776705603521246



Classification Report

	precision	recall	f1-score	support
Severe	0.81	0.75	0.78	247
Very Poor	0.73	0.72	0.72	472
Poor	0.68	0.51	0.58	573
Moderate	0.82	0.85	0.84	2697
Satisfactory	0.75	0.83	0.78	1656
Good	0.79	0.36	0.50	262
accuracy			0.78	5907
macro avg	0.76	0.67	0.70	5907
weighted avg	0.77	0.78	0.77	5907

Figure 7.6: Classification Report SVC

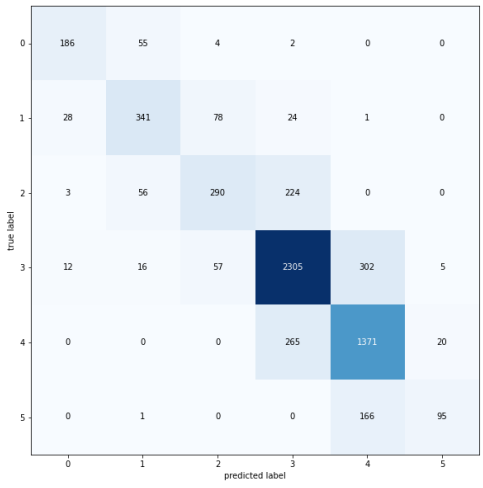
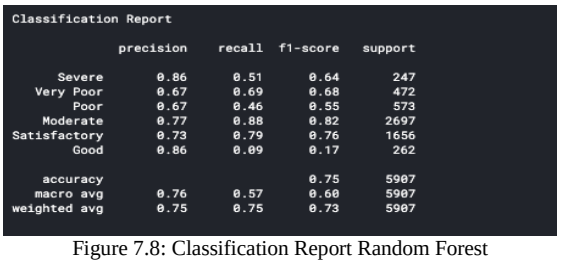


Figure 7.7: Confusion Matrix SVC

4. Random Forest:
   Accuracy: 0.7469104452344676
   Precision: 0.7498485970022315
   Recall: 0.7469104452344676



Classification Report

	precision	recall	f1-score	support
Severe	0.86	0.51	0.64	247
Very Poor	0.67	0.69	0.68	472
Poor	0.67	0.46	0.55	573
Moderate	0.77	0.88	0.82	2697
Satisfactory	0.73	0.79	0.76	1656
Good	0.86	0.09	0.17	262
accuracy			0.75	5907
macro avg	0.76	0.57	0.60	5907
weighted avg	0.75	0.75	0.73	5907

Figure 7.8: Classification Report Random Forest

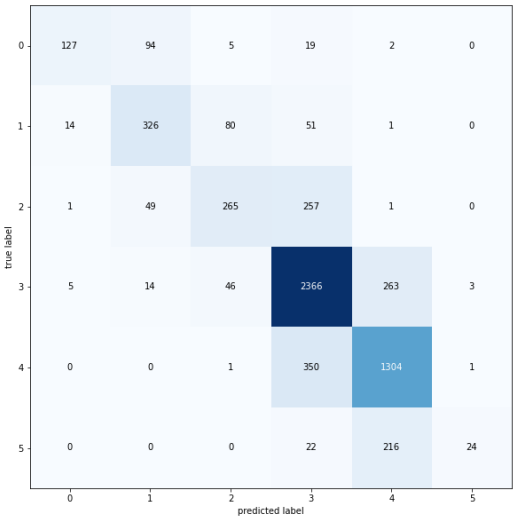


Figure 7.9: Confusion Matrix Random Forest

VIII. CONCLUSION

After building the model and evaluating based on various parameters, it was observed that for regression while predicting AQI, Kernel Ridge Regression performs the best. It was also observed that XGBoost performs the best while Classification of AQI Levels.

References

[1]. 1.National Air Quality Index, Central Pollution Control Board, (https://cpcb.nic.in/)
[2]. 2. Ying Yan, Yuangang Li, Maohua Sun and Zhenhua Wu, Primary Pollutants and Air Quality Analysis for Urban Air in China: Evidence from Shanghai, 2019, 11, 2319
[3]. 3. Filonchyk, M.; Yan, H.; Li, X. Temporal and spatial variation of particulate matter and its correlation with other criteria of air pollutants in Lanzhou, China, in spring-summer periods. Atmos. Pollut. Res. 2018, 9, 1100–1110.
[4]. National Air Quality Index, Central Pollution Control Board, (https://cpcb.nic.in/)
[5]. Pérez, P.; Trier, A.; Reyes, J. Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile. Atmos. Environ. 2000, 34, 1189–1196.

[6]. Zhu, S.; Lian, X.; Liu, H.; Hu, J.; Wang, Y.; Che, J. Daily air quality index forecasting with hybrid models: A case in China. Environ. Pollut. 2017, 231, 1232–1244
[7]. Biancofiore, F.; Busilacchio, M.; Verdecchia, M.; Tomassetti, B.; Aruffo, E.; Bianco, S.; Di Tommaso, S.; Colangeli, C.; Rosatelli, G.; Di Carlo, P. Recursive neural network model for analysis and forecast of PM10 and PM2.5. Atmos. Pollut. Res. 2017, 8, 652–659.
[8]. Kalapanidas, E.; Avouris, N. Short-term air quality prediction using a case-based classifier. Environ. Model. Softw. 2001, 16, 263–272
[9]. Fuller, G.W.; Carslaw, D.C.; Lodge, H.W. An empirical approach for the prediction of daily mean PM10 concentrations. Atmos. Environ. 2002, 36, 1431–1441.
[10]. Carbajal-Hernández, José Juan "Assessment and prediction of air quality using fuzzy logic and autoregressive models." Atmospheric Environment 60 (2012): 37-50.
[11]. Kurt, A.; Oktay, A.B. Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. Expert Syst. Appl.2010, 37, 7986–7992.
[12]. Corani, G. Air quality prediction in Milan: Feed-forward neural networks, pruned neural networks and lazy learning. Ecol. Model. 2005, 185, 513–529.
[13]. T. Chai and R. R. Draxler: RMSE or MAE, Geosci. Model Dev., 7, 1247–1250, 2014
[14]. 14. NEIL J. SALKIND, ROOT MEAN SQUARE ERROR, 2010
[15]. Sungil Kim [a] Hee young Kim, A new metric of absolute percentage error for intermittent demand forecasts, 2016
[16]. Trends in extreme learning machines: a review, by Huang, G., Huang, G., Song, S., & You, K. (2015). Neural Networks, (cited 323 times, HIC: 0, CV: 0)
[17]. Shalabh (1999): "Improving the Predictions in Linear Regression Models", *Journal of Statistical Research*, Vol. 33, No. 1.
[18]. Toutenburg, H. and Shalabh (2000): "Improved Prediction in Linear Regression Model with Stochastic Linear Constraints", Biometrical Journal, 42, 1, 71-86
[19]. S. B. Kotsiantis, I. D. Zaharakis, Machine learning: A review of classification and combining techniques, 2007
[20]. A review of supervised machine learning applied to ageing research, Fabio Fabris, Joa Pedro de Magalha, Alex A. Freitas
[21]. Central Pollution Control Board (https://cpcb.nic.in/)
[22]. Minbo Kim, R Carter Hill, The BoxCox transformation of variables in regression, 1993.