International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8070
***Article Received:** 23 August 2023 **Revised:** 31 August 2023 **Accepted:** 04 September 2023*

Implementation of Feature Engineering in Prediction of AQI in India using Machine Learning

Reema Gupta[1],Dr. Priti Singla[2]
[1]Research Scholar,Department of Computer Science and Engineering
[2]Faculty of Engineering, Department of Computer Science and Engineering
[1,2]Baba Mastnath University,Rohtak,India
reema2405@gmail.com

Abstract—Prediction of Air Quality Index (AQI) is the necessity of today's era but for the prediction, analysis of different preprocessing techniques that can be applied, needs to be considered. In this study, first of all we explored various feature engineering techniques such as Data Imputation, Scaling, Extraction, Selection, and Data Split that can be used before applying machine learning algorithm for better results. Second, we used MLR and SVR (Linear, Gaussian) to build the prediction models. Finally, we used root mean square error (RMSE), R^2, Mean Squared Error (MSE) and Mean Absolute Error (MAE) to evaluate the performance of the regression models in collaboration with the feature engineering techniques. The results shows that the performance of Linear SVR is better when coupled with imputation and robust scaler (R^2=0.7557834846394744) as compared to the others, the performance of Gaussian SVR is better when coupled with the imputation only as compared to the others. In case of MLR, results (R^2=0.7769187383819041) are almost same in all the 4 cases and performance degraded when PCA was applied.

Keywords-Imputation,AQI,Standard Scaler,Modeling,Evaluation.

I. INTRODUCTION

Air Pollution Index (AQI) was created to make it easier for people to comprehend the impact of local air quality on their health. It is a health protection alarming tool made to assist people in making decisions about how to safeguard their health by changing their activity levels and reducing their short-term exposure when air pollution levels are high. Jind is one of the most polluted cities in the state of Haryana, India. As per the Information provided by IQAir, India was in the top 5 polluted country among 118 countries in 2021.

Traffic related pollution is one of the major sources of air pollution. There is correlation between air quality and traffic CO intensity because pollution increases in the traffic peak hours i.e. in the morning and the evening [1]. In case of travelling from one place to the other, Air quality varies location to location. IOT based technology [2] can predict the air quality of the entire route and the destination place. Dust that is created during the demolition of old buildings and the following construction of new ones is also one of the main reasons of deteriorating the air quality.

Air quality prediction for a particular location can be done with the help of image capturing through daily used devices [3]. There is a need to look towards the harmful pollutants that affects the respiratory system and causes breathing problem and other respiratory disease. [4]Air is the basic need of every organism because polluted inhalation creates several health problems. Currently, laws are only put in place by the government when air quality reaches dangerous levels. If it is possible to predict when the air quality will reach dangerous levels, the Government can enact these rules quickly, possibly halting further deterioration of the air quality. This research tries to create a model that can review historical air quality data and predicts air quality index and amounts of various pollutants.

This paper is divided into sections. Section II includes the Literature Review, Section III discusses the Materials and Methodology, Section IV contains Results that displays the outcomes obtained after applying techniques and Section V is conclusion followed by the references.

II. LITERATURE REVIEW

For differentiating between the seasonal Air Quality, principal components analysis (PCA) was performed along-with construction of Ensemble models [5]**.** The study emphasized on identifying the air pollution sources in a span of five years in the city of Lucknow, India. The major source of pollution found were fuel combustion and emissions emitted from vehicles. Another research used Principal component regression (PCR) technique for forecasting daily AQI in Delhi with the help of previous day AQI and meteorological variables in four seasons in the years 2000-2006 [9]**.** They performed many statistical parameters which produced the same result but the performance of the PCR model was way better in the winter season as compared to any other season. They used the covariance of the input data matrix as a principal component.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8070
***Article Received:** 23 August 2023 **Revised:** 31 August 2023 **Accepted:** 04 September 2023*

An alternative method was proposed to the traditional sensor network. They suggest Machine learning and the Internet of Things (IoT) be applicable that use cloud-centric IoT middleware. It receives data from both- air pollution sensors as well as weather sensors. It is cost-efficient and more reliable. To monitor and predict air pollution, Artificial Neural Network (ANN) results in a reliable and suitable candidature [6]. Artificial neural network (ANN) used to predict air quality index (AQI) and air quality health index (AQHI) in span of one year in Ahvaz, Iran. Predictions were made based on hourly criteria of air pollutant concentrations and [10] concluded that ANN is reliable and can be used by practitioners to estimate air quality indices and spatial-temporal profile of pollutants.

An optimal hybrid model was proposed for forecasting of AQI by combining AI method and secondary decomposition (SD), optimization algorithm [11]. Their proposed idea successfully solved problems like considering influential factors based on decomposition technology. For their case study, they took two daily AQI series from Beijing and Guilin, China from December 2016- December 2018 and comprehended that their optimal-hybrid model has success-rate of forecasting AQI.

There are multiple issues discussed by [12] in predictions of AQI and accordingly the future needs to face such challenges. They also compared it with current research work on AQI which uses various models like machine learning and big data analysis. LightGBM model was proposed to predict the PM 2.5 concentration on the basis of 35 air quality stations that are monitoring in Beijing for 24 hours [13]. They compared the predicted data and the actual data. For their data source, they used the forecasting data. By using the lightGBM model they also resolved problems such as processing large scale and high-dimensional data. Their proposed model resulted as better alternative. The integration of predicted data improves stability of model and understands data adequately.

In a study, support vector regression (SVR) and random forest regression (RFR) models which are based on machine learning algorithms were used to predict the AQI of Beijing and the Italian city [8]**.** The results suggested that the RFR model is more reliable and time-efficient to perform complex and large samples. They established various models to improve the accuracy of the prediction of air indicators. Support vector regression (SVR) with Radial basis function (RBF) kernel was used to predict the concentration of various pollutants, ground level ozone and Air quality Index using already available data at US embassy and central pollution control board in New Delhi [14]. They tested various alternatives and found that radial basis function (RBF) helps in predicting hourly concentrations most accurately in the state of California [15].

Four different machine learning models namely Artificial neural network, statistical multilevel regression, deep learning long-short-term memory and neuro-fuzzy were applied on meteorological parametric dataset of 5 years [16]**.** The work concluded that Dl-LSTM is highly correlated with the dataset with low error levels and thus best suited. To increase the quality prediction performance, [7] suggested a deep learning algorithm. They used pre-processed data to find AQI. They demonstrated the utilization of data mining methods in natural resource research and environmental monitoring. Principal component analysis was applied to the deep learning models namely recurrent neural network (RNN), long short-term memory (LSTM) and bidirectional LSTM to forecast fine particulate matter (PM 2.5) in eight Korean cities for 5 years. The study [17] resulted in the conclusion that models with application of PCA produced better results. The PCA applied model is accurate for improving the performance of the model.

Numerous regression techniques and machine learning fused with big data analytics and IoT for the prediction of AQI namely SDG regression, Support vector regression, gradient boosting, linear regression, decision tree regression, adaptive boosting, random forest regression and artificial neural networks. Major air pollutants were used as a source to analyze the techniques. In the results [18] SVR and neural networks were found as best suited techniques.

III. MATERIALS AND METHOD

The flowchart of complete step by step process for deriving the output from given input data is presented in Fig1

A. RAW DATA

Dataset for the prediction of air quality is taken from the publicly available website of Central Pollution Control Board. The dataset used has 12 attributes and 851 instances of the Jind city in Haryana considered from 1st March, 2020 to 29th June, 2022. Each instance consist of the concentration of various pollutants, Ozone and other parameters such as Temperature, Wind Speed, Wind Direction, Relative humidity, etc. that are required as an input parameters and AQI as the target variable that helps to analyze the air quality.

B. DATA IMPUTATION

This step deals with the irregularities within the dataset, for example missing values or data having values ‘None’ or ‘NA’ are replaced with the mean of the corresponding column. Dataset is represented as D1.

```
D1 = D1.replace(to_replace='NA',value=np.nan)

D1 = D1.replace(to_replace='None ',value=np.nan)

D1['PM2.5'] = D1['PM2.5'].astype('float64')
```

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8070
***Article Received:** 23 August 2023 **Revised:** 31 August 2023 **Accepted:** 04 September 2023*

D1['PM2.5'].replace({np.nan: D1['PM2.5'].mean()},inplace = True)

This task is carried out on all components i.e. on various pollutants, Ozone and Meteorological parameters i.e. RH, WS, WD and AQI that are used in the prediction of the air quality index.

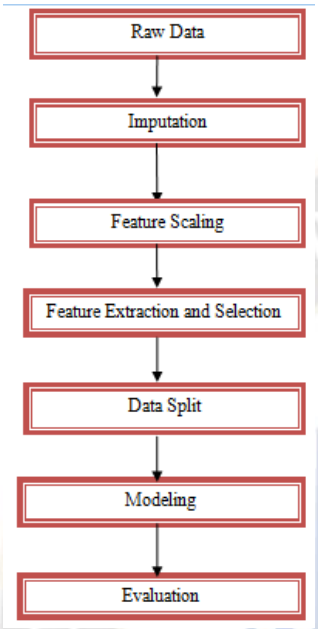

Fig. 1 Steps for Processing

C. FEATURE SCALING

This step standardizes the input variables in a fixed range and is a part of data preprocessing**.** For this, multiple techniques such as StandardScaler, Min max Scaler and Robust Scaler will be discussed in this paper.

1. **Standard Scaler-** This scaler transforms the each value in range of mean 0 and deviation 1.
2. **Minmax Scaler-** This scaler rescales each and every value to [0,1]
3. **Robust Scaler-** This scalers deals with handling of outliers

D. FEATURE EXTRACTION AND SELECTION

Feature Extraction is the process of reducing the dimensions of the dataset. It reduces the complexity and creates new ones which is linear combination of the existing ones.

Principal Component Analysis is one of the techniques that can be used for this purpose. PCA technique uses Covariance matrix, Eigen vectors and Eigen values.

Feature Selection is a method in which we have to select the best variables as input that can be used to build a prediction model. Output of the model is dependent on the quality of the data that are given as input to the model. Appropriate selection of the variables is the preprocessing step before applying the machine learning model.

E. DATA SPLIT

In this step given data set is splitted into train-test ratio. Splitting ratio should not undergo the over-fitting and under-fitting problems. In this study, train-test ratio considered is 80:20 respectively.

F. MODELING AND EVALUATION

For Modeling Machine learning algorithm, either classification or regression is applied on the problem and result obtained is compared with the actual value. If the difference between the actual value and the result is very low, this implies applied algorithm is working well in the environment. The performance of machine learning algorithms are evaluated through various metrics like accuracy, precision, Mean square error, Coefficient Metrics, etc.

The main objective is to predict the air quality index using machine learning technique. To create a prediction model two regression techniques are used i.e. MLR (Multiple linear regression) and SVR (Support Vector regression).

MLR (Multiple Linear Regression) – It is simply a regression algorithm in which the response variable is calculated based on multiple variables which serves as the input. Prediction of the response variable depends on how they are correlated with the independent variables.

SVR (Support Vector Regression) – SVR works on the same principle of SVM (Support Vector Machine) . The only difference is that, SVM deals with classification and regression problems while SVR deals with the regression problems.

IV. RESULTS

Python tool was used to obtain the results using machine learning algorithms. Numpy, pandas, Matplotlib, Scikit-learn, seaborn are the major libraries used in the implementation. Table I-III showing the performance of Linear SVR, Gaussian SVR and MultiLinear Regression algorithms are presented below wherein M1 represents Mean Absolute Error, M2 represents Mean Squared Error, M3 represents Root Mean Square Error and M4 represents Coefficient R^2. Different preprocessing techniques such as StandardScaler, Minmax Scaler, RobustScaler and PCA along with Imputation were applied for air quality prediction. Selection extracted the required components in the prediction from the dataset followed by the Data split in train-test ratio 80:20 respectively.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8070
***Article Received:** 23 August 2023 **Revised:** 31 August 2023 **Accepted:** 04 September 2023*

TABLE I
Performance Evaluation of Linear SVR while applying different Preprocessing techniques

Linear SVR				
Preprocessing Techniques	**M1**	**M2**	**M3**	**M4**
Imputation	33.04144965797549	2697.8424632036053	51.940759170458854	0.7330234447729286
Imputation and Standard Scaler	33.87663968082457	2472.5937100001697	49.72518184984515	0.7553138998382833
Imputation and Minmax Scaler	66.65289061446242	6406.411229474495	80.04006015411592	0.36602613950181295
Imputation and Robust Scaler	34.058417897389525	2467.848477537152	49.67744435392336	0.7557834846394744
Imputation and PCA	36.925125504159986	3087.234773408241	55.5628902542717	0.6944894610328636

TABLE II
Performance Evaluation of Gaussian SVR while applying different Preprocessing techniques

Gaussian SVR				
Preprocessing Techniques	**M1**	**M2**	**M3**	**M4**
Imputation	53.61567208810835	4741.171164960525	68.85616286840651	0.5308171019519038
Imputation and Standard Scaler	62.94170943020042	6346.393504078221	79.66425487053915	0.37196545056144925
Imputation and Minmax Scaler	65.13933618007908	6527.617399185145	80.79367177684861	0.3540316638780616
Imputation and Robust Scaler	58.46840816041663	5630.9255829013555	75.03949348777186	0.44276764289719117
Imputation and PCA	54.06578733773182	5237.580499642561	72.37113029131548	0.4816927902237098

TABLE III
Performance Evaluation of MLR while applying different Preprocessing techniques

MLR				
Preprocessing Techniques	**M1**	**M2**	**M3**	**M4**
Imputation	35.58786771686932	2254.2732257012267	47.47918728981391	0.7769187383819041
Imputation and Standard Scaler	35.587867716869305	2254.273225701229	47.479187289813936	0.7769187383819038
Imputation and Minmax Scaler	35.5878677168693	2254.2732257012285	47.47918728981393	0.7769187383819038
Imputation and Robust Scaler	35.5878677168693	2254.2732257012294	47.479187289813936	0.7769187383819038
Imputation and PCA	40.19107456509701	2785.609554106226	52.7788741269291	0.7243380764050298

A. Linear SVR Results

Fig 2-6 shows the actual and predicted AQI Values when using Linear SVR model and using imputation, imputation with scaling techniques and imputation with PCA as preprocessing technique to enhance the dataset before processing. The values shown in the below figures are of test dataset which includes actual values and predicted ones using the approach.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8070
Article Received: 23 August 2023 *Revised:* 31 August 2023 *Accepted:* 04 September 2023

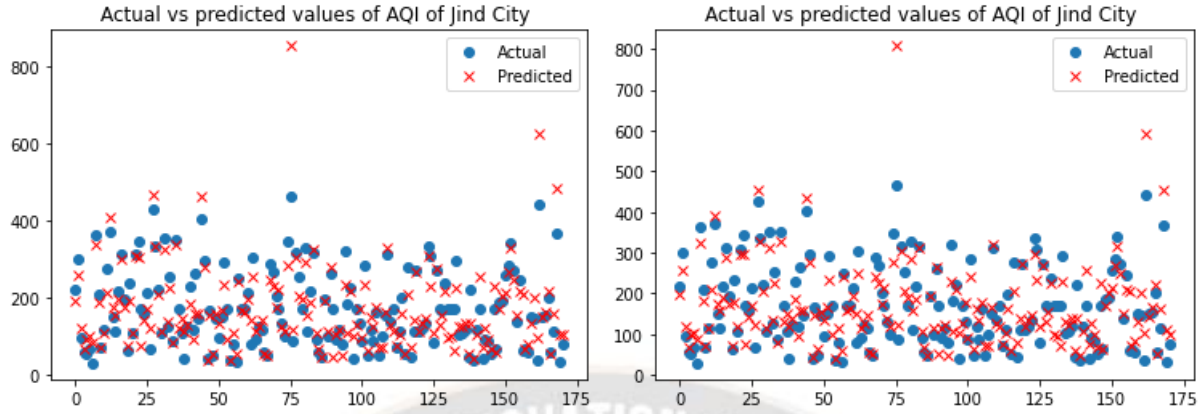

Fig 2 and 3 shows the actual and predicted values of AQI a) Imputation b) Imputation and Standard Scaler

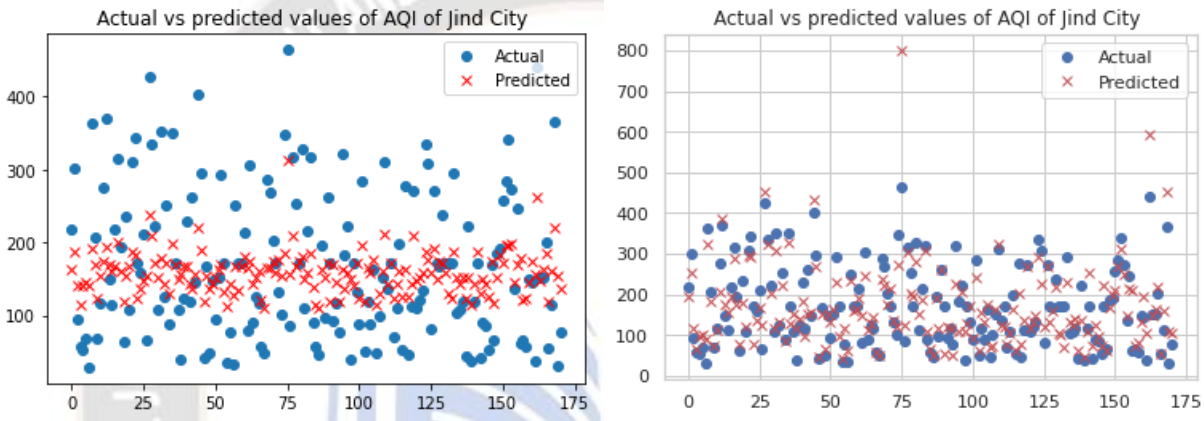

Fig 4 and 5 shows the actual and predicted values of AQI a) Imputation and minmax scaler b) Imputation and Robust Scaler

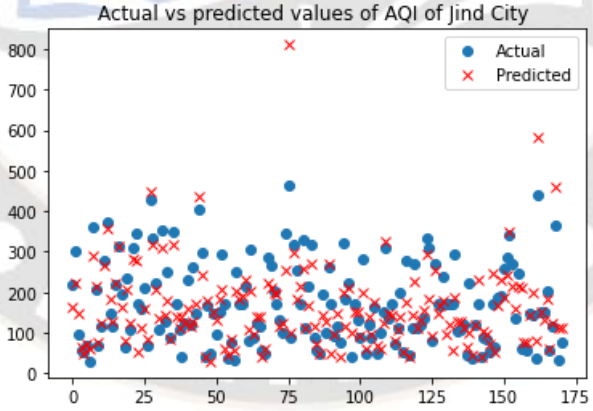

Fig 6 shows the actual and predicted values of AQI (Imputation and PCA)

B. Gaussian SVR Results

Fig 7-11 shows the actual and predicted AQI Values when using Gaussian SVR model and using imputation, imputation with scaling techniques and imputation with PCA as preprocessing technique to enhance the dataset before processing. The values shown in figures are of test dataset which includes actual values and predicted ones using different approaches.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8070
***Article Received:** 23 August 2023 **Revised:** 31 August 2023 **Accepted:** 04 September 2023*

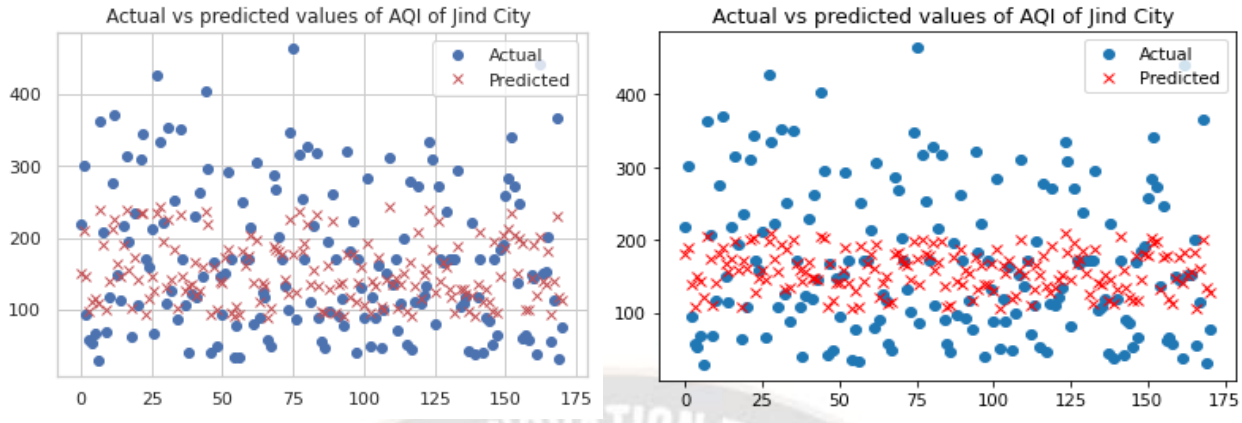

Fig 7 and 8 shows the actual and predicted values of AQI a) Imputation b) Imputation and Standard Scaler

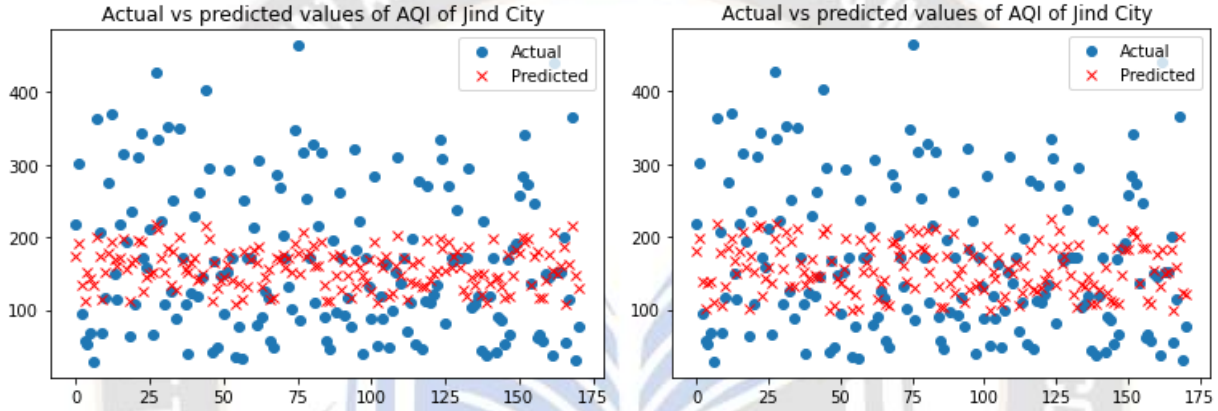

Fig 9 and 10 shows the actual and predicted values of AQI a) Imputation and minmax scaler b) Imputation and Robust Scaler

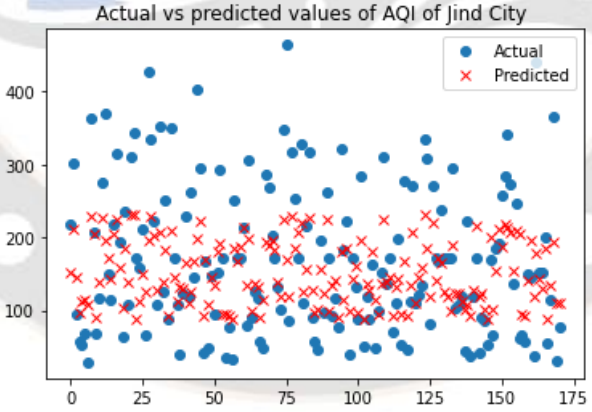

Fig 11 shows the actual and predicted values of AQI (Imputation and PCA)

C. MLR Results

Fig 12-16 shows the actual and predicted AQI Values when using MLR model and using imputation, imputation with scaling techniques and imputation with PCA as preprocessing technique to enhance the dataset before processing.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8070
***Article Received:** 23 August 2023 **Revised:** 31 August 2023 **Accepted:** 04 September 2023*

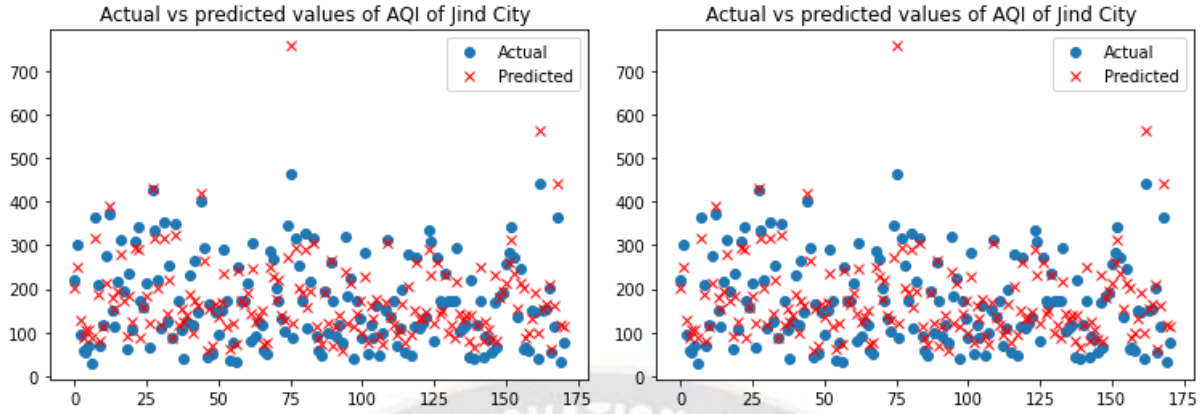

Fig 12 and 13 shows the actual and predicted values of AQI a) Imputation b) Imputation and Standard Scaler

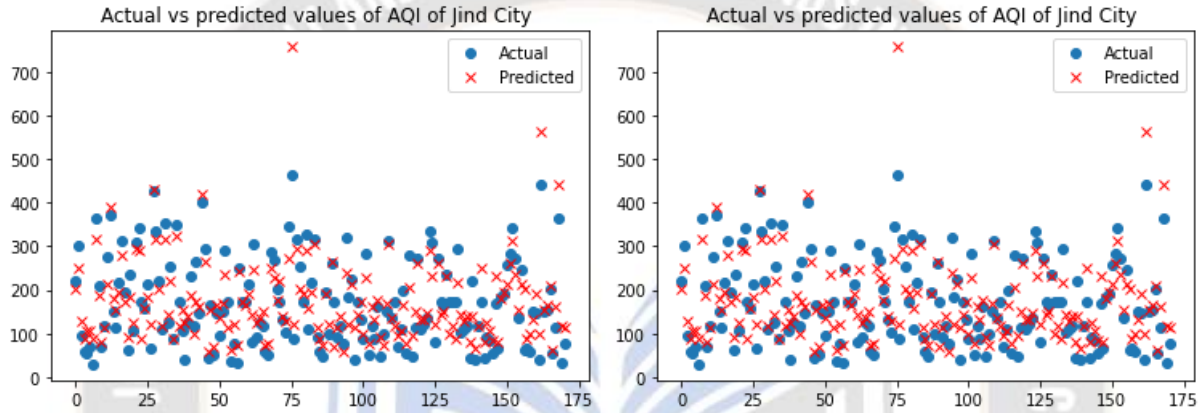

Fig 14 and 15 shows the actual and predicted values of AQI a) Imputation and minmax scaler b) Imputation and Robust Scaler

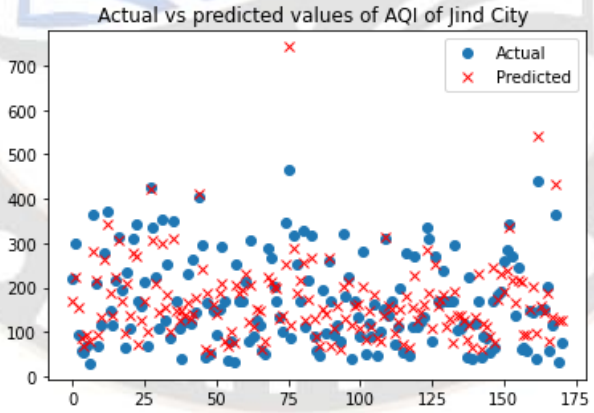

Fig 16 shows the actual and predicted values of AQI (Imputation and PCA)

V. COMPARATIVE ANALYSIS

Imputation with Robustscaler technique with MAE 34.058417897389525, RMSE 49.67744435392336, R^2 0.7557834846394744 outperforms as compared to other preprocessing technique with Linear SVR. In case of Gaussian SVR performance is best when only imputation technique with MAE 53.61567208810835, Coefficient R^2 0.5308171019519038 is used. In case of MLR, performance is same when the imputation technique is applied and when imputation is used along with the scaling techniques with MAE, RMSE and R^2 35.58786771686932, 47.47918728981391, 0.7769187383819041 respectively. These best approaches of each regression technique are compared based on evaluation metrics is shown in Fig 17.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8070
***Article Received:** 23 August 2023 **Revised:** 31 August 2023 **Accepted:** 04 September 2023*

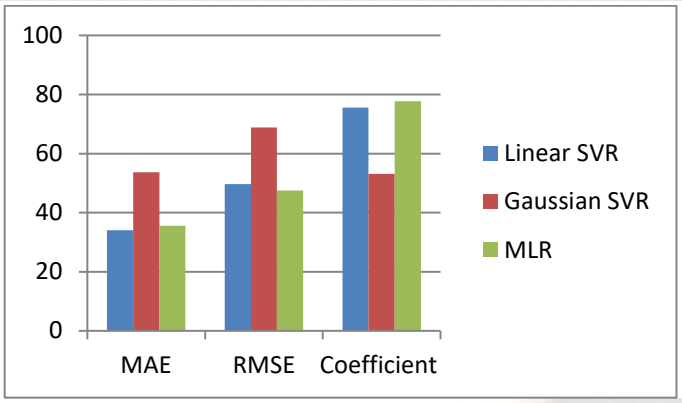

Fig 17 Comparison based on Evaluation Metrics

VI. CONCLUSION

Air quality index is predicted using Prediction models based on MLR, SVR. Dataset contained null values and outliers that need to be handled before applying model for the prediction. This paper concludes that the performance of the algorithm improves if outliers and irregular values in the dataset can be handled using above discussed techniques. Transformation of raw data into the fruitful data is one of the needs of the machine learning model. In this study PCA degraded the performance of the existing model in the prediction. RMSE, R^2, MSE and MAE have been used to evaluate the performance of the regression models when different preprocessing techniques were applied to improve the performance. In case of Linear SVR, imputation and robustscaler when applied gave better result as compared to the others while results produced using Gaussian SVR is better when applied imputation only as compared to the others. In case of MLR, performance results are almost same in all the 4 cases and performance degraded when PCA was applied. Overall the performance of MLR is best as compared to the others in terms of various evaluation metrics used in the paper.

REFERENCES

[1] L. Pan, E. Yao and Y. Yang, "Impact Analysis of Traffic-Related Air Pollution based on Real time Traffic and Basic Meteorological Information," Journal of Environmental Management, vol. 183, no. 3, pp. 510-520, 2016.

[2] S. Dhingra, R. B. Madda, A. H. Gandomi, R. Patan and M. Daneshmand, "Internet of Things Mobile-Air Pollution Monitoring System (IoT-Mobair)," IEEE Internet of Things Journal, vol. XX, no. XX, p. 8, 2019.

[3] J. Ma, K. Li, Y. Han and J. Yang, "Image based Air Pollution Estimation Using Hybrid Convolutional Neural Network," in 24th International Conference on Pattern Recognition, Beijing, China, 2018.

[4] R. Brugha and J. Grigg, "Urban Air Pollution and Respiratory Infections," Paediatric Respiratory Reviews, vol. 15, no. 2, p. 6, 2014.

[5] K. P. Singh, S. Gupta and P. Rai, "Indentifying Pollution Sources and Predicting Urban Air Quality using Ensemble Learning Methods," Atmospheric Environment, vol. 80, pp. 426-437, 2013.

[6] I. U. Samee and M. T. Jilani, "An Application of IOT and Machine Learning to Air Pollution Monitoring in Smart Cities," IEEE, p. 6, 2019.

[7] S. K, A. K, G. M, G. R and M. A, "Air Quality Prediction Using Classification Techniques," Annals of R.S.C.B, vol. 25, no. 4, pp. 3794-3805, 2021.

[8] H. Liu, Q. Li, D. Yu and Y. Gu, "Air Quality Index and Air Pollutant Concentration Prediction based on Machine Learning Algorithms," MDPI, no. 4069, p. 9, 2019.

[9] A. Kumar and P. Goyal, "Forecasting of air quality in Delhi using principal component regression technique," Atmospheric Pollution Reasearch 2, pp. 436-444, 2011.

[10] H. Maleki, A. Sorooshian, G. Goudarzi, Z. Baboli, Y. T. Birgani and M. Rahmati, "Air Pollution Prediction using an Artificial Neural Network Model," Clean Technologies and Environmental Policy, vol. 21, pp. 1341-1352, 2019.

[11] Q. Wu and H. Lin, "A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors," Science of the total environment, vol. 683, pp. 808-821, 2019.

[12] G. K. Kang, J. Z. Gao, S. Chiao, S. Lu and G. Xie, "Air Quality Prediction: Big Data and Machine Learning Approaches," International Journal of Environmental Science and Developement, vol. 9, no. 1, pp. 8-16, 2018.

[13] Y. Zhang, Y. Wang, M. Gao, Q. Ma, J. Zhao, R. Zhang, Q. Wang and L. Huang, "A Predictive Data Feature Exploration - Based Air Quality Prediction Approach," IEEE Access, vol. 7, pp. 30732-30743, 2019.

[14] S. Bhattacharya and S. Shahnawaz, "Using Machine Learning to Predict Air Quality Index in New Delhi," arXiv:2112.05753, p. 7, 2021.

[15] M. Castelli, F. M. Clemente, A. Popovik, S. Silva and L. Vanneschi, "A Machine Learning Approach to Predict Air Quality in California," Wiley, vol. 2020, p. 23, 2020.

[16] C. AmuthaDevi, D. S. Vijayan and V. Ramachandran, "Development of Air Quality Mornitoring (AQM) Models using different Machine Leaning Approaches," Journal of Ambient Intelligence and Humanized Computing, p. 13, 2021.

[17] S. W. Choi and B. H. Kim, "Sustainability," Applying PCA to Deep Learing Forecasting Models for Predicting PM 2.5, vol. 13, no. 7, p. 30, 2021.

[18] C. Srivastava, S. Singh and A. P. Singh, "Estimation of Air Pollution in Delhi using Machine Learning Techniques," in 2018 International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, 2018.

Forecasting Air Quality Index(AQI) For Cities In Haryana State Using Machine Learning

Reema Gupta,Research Scholar
and
Dr.Priti Singla,Faculty Of Engineering
Baba Mastnath University,Rohtak (India)

In India and throughout the world, air pollution is becoming a severe worry day by day. Governments and the general public have grown more concerned about how air pollution affects human health. Consequently, it is crucial to forecast the air quality with accuracy. In this paper, Machine learning methods SVR and RFR were used to build the hybrid forecast model to predict the concentrations of Air Quality Index in Haryana Cities. The forecast models were built using air pollutants and meteorological parameters from 2019 to 2021 and testing and validation was conducted on the air quality data for the year 2022 of Jind and Panipat city in the State of Haryana. Further, performance of hybrid forecast model was enhanced using scalar technique and performance was evaluated using various coefficient metrics and other parameters. First, the important factors affecting air quality are extracted and irregularities from the dataset are removed. Second, for forecasting AQI various approaches have been used and evaluation is carried out using performance metrics. The experimental results showed that the proposed hybrid model had a better forecast result than the standard Random forest Regression, Support Vector Regression and Multiple Linear Regression.

Keywords: Forecasting,Air Quality Index,Hybrid Model,Linear Regression,RFR,Support Vector Regression.

1. INTRODUCTION

The Air Quality Index (AQI) is a widely used metric that provides a measure of the air quality in a given region. Forecasting AQI values can aid in the effective management of air quality by enabling timely interventions and interventions. Most urban areas struggle with serious air pollution issues. People need to be conscious of the air they are breathing. The National Ambient Air Monitoring Network produces data that displays the amount of various air pollutants; however it is not an easy job to understand this data. Accordingly, Central Pollution Control Board (CPCB) created the national Air Quality Index (AQI) for Indian cities. Nowadays, the air in most cities is polluted, and other toxins have been introduced, making the air even more poisonous. Activities both natural and human-made can pollute the air. (Bhat, Manek, and Mishra, 2019) Human activities are adding pollutants to the air, including lead, mercury, chlorofluorocarbons (CFC), sulphide oxides, nitrogen oxides, carbon monoxide (CO) and carbon dioxide (CO2).Humans are known to suffer substantial health impacts from exposure to high levels of PM, particularly the tiny PM2.5 particles. Environmental consequences are obviously seen, in addition to health implications. (Boonphun, Jirat, Kaisornsawad, Chalet, Wongchaisuwat, and Papis, 2019) There has been extensive use of time series forecasting techniques in the field of air quality forecasting to predict future AQI values. Machine learning, in particular, has emerged as a promising technique for AQI forecasting due to its ability to handle complex and non-linear relationships in data.(Mahalingam, Usha, Elangovan, Kirthiga, Dobhal, Himanshu, Valliappa, Chocko, Shrestha, Sindhu, Kedam, and Giriprasad, 2019) Monitoring urban air quality has been difficult since industrialization took hold. Around the world, air pollution continues to be a significant problem for both the general public and the government. Both the environment and human health are significantly harmed by air pollution, which leads to acid rain, global warming, heart illnesses, and skin cancer in humans. In order to reduce pollution before it has a negative impact, this

research uses 2 machine learning algorithms, support vector machines and neural networks to tackle the problem of forecasting the Air Quality Index (AQI). This paper seeks to explore the application of machine learning algorithms for the time series forecasting of AQI values. The study utilizes a range of machine learning techniques to predict AQI values based on historical data, and the results are compared to traditional time series forecasting methods. The results of this study can support the creation of precise and trustworthy AQI forecasting models, ultimately resulting in improved air quality management procedures. This study uses data from the Haryana cities of Jind and Panipat to examine whether machine learning techniques are most suited for a prediction model under various conditions. In this work, a forecast model based on RFR and SVR and other existing ones are used to anticipate the values of air quality. The forecast model was constructed to forecast the air quality index for the year 2022 of Jind and Panipat city.

This paper is divided into sections. Section 2 includes the Related Works, Section 3 discusses the Proposed Methodology, Section 4 contains Results and discussions, Section 5 includes the performance analysis based on the parameters and Section 6 is conclusion followed by the references.

2. RELATED WORKS

Machine learning has gained popularity in recent years due to its ability to handle complex and non-linear relationships in data. For predicting air quality, a variety of machine learning methods have been applied, including Artificial Neural Networks (ANN), Support Vector Regression (SVR), Random Forest (RF), and Gradient Boosting Decision Trees (GBDT). The review also emphasizes that the selection of the appropriate machine learning algorithm (Patil, Dinde, Powar, and Ganeshkhind, 2020)is crucial for accurate and reliable forecasting. The paper further highlights that studies have shown the superiority of machine learning over traditional statistical methods such as linear regression and ARIMA. The literature review concludes by noting that although machine learning techniques have shown promising results, there is still a need for further research to investigate their effectiveness in different geographical locations and for different pollutants. The study by (Lei, Siu, Monjardino, Mendes, and Ferreira, 2022) aims to address this gap by applying various machine learning algorithms to predict the AQI values in Macao, providing a useful case study for further research in the field.

In the study by (Kumar and Pande, 2023), the application of machine learning methods for air pollution prediction in Indian cities is investigated. This paper provides an overview of previous research conducted on air pollution estimation using machine learning methods, with a particular focus on studies conducted in India. The review highlights that the issue of air deteriorating air quality has become a major cause of concern in India, and traditional methods of air pollution prediction have limitations in terms of accuracy and timeliness. The capacity of machine learning techniques to manage sizable and complicated datasets as well as to identify non-linear correlations between variables has led to an increase in their application in recent years. This paper also highlights the importance of input variables in air pollution prediction, including meteorological data, traffic volume, and industrial emissions. Several studies have investigated the impact of these variables on air pollution prediction accuracy, and the review notes that the selection of appropriate input variables is crucial for accurate and reliable air pollution prediction (Aarthi, Gayathri, Gomathi, Kalaiselvi, and Gomathi, 2020). The environment's carbon monoxide C.O. concentration is predicted using regression analysis techniques. Long-term exposure to carbon monoxide can result in irregular heartbeat, nonfatal heart attack, and even death from lung illness. Therefore, it is important to reduce the use of motor vehicles, power plants, and cigarettes. (Sankar Ganesh, Arulmozhivarman, and Tatavarti, 2017) proposed several ensemble models to forecast the air quality index of Houston and Los Angeles for the year 2010-2016.For two separate sets of base learners, the effectiveness of the ensemble approaches that were deployed has been assessed and contrasted. It has been concluded that for both sets of base learners, cas-

cade forward ensemble beat SVR ensemble. The capacity to more precisely forecast the AQI from highly nonlinear data is the current study's shortcoming.(Sanjeev, 2021) implemented RF, Support Vector Regression, Artificial Neural Network and to predict the Air quality and precision, F-score, Specificity parameters are used to evaluate the performance of the said techniques and it has been found that the performance of RF outperforms as compared to the others.(Zhu, Wu, Chen, Zhou, and Tao, 2018) proposed a hybrid EEMD-MM-CFM model to forecast the air quality of Hefei, China. Dataset of the year 2016-2018 is considered and from 2016- April 2018 data is considered to train the model and May 2018 data is considered for the test. The mirror method and EEMD are used to sort the original AQI time series; a mixed forecasting model is used to predict the IMFs and residue; and finally, the forecasts are added to provide the results. To assess the efficiency of the suggested model, 4 statistical indicators—MAE, MAPE, SSE, and RMSE—as well as mode correctness and the DM test were used. We can improve human health, which is of the utmost importance as PM2.5, by monitoring the amount of particulate matter in the air. (Zamani Joharestani, Cao, Ni, Bashir, and Talebiesfandarani, 2019) Using Taiwan AQM data sets from 2012 to 2017, machine learning predictive models for predicting the amount of particulate matter in the atmosphere are examined. When it comes to making predictions, these models perform better than the current industry standards. These models' efficacy was evaluated using statistics including the Mean Absolute Error (MAE), Mean Square Error (MSE) are all types of errors. For the purpose of predicting particulate matter (PM2.5), statistical computations utilizing measures such as MAE, MSE, RMSE, and R2 are employed to develop machine learning models. (Harishkumar, Yogesh, Gad, et al., 2020) The findings indicate that the suggested model values perform better than those of the prior models and that the actual and anticipated values are quite similar. Finally, we draw the conclusion that using the TAQMN data, the GBR model is superior for predicting air pollution. The combustion of fossil fuels, transportation patterns, and industrial variables like power plant emissions all has a substantial impact on air pollution. Particulate matter (PM 2.5) requires more attention among all the particulates that affect air quality. When it is present in large amounts in the air, it seriously affects people's health. Therefore, it's crucial to control it by regularly monitoring its level in the atmosphere.(Aditya, Deshmukh, Nayana, and Vidyavastu, 2018) In order to establish whether a data sample is contaminated, logistic regression is used. Auto-regression is used to forecast future PM2.5 values using prior PM2.5 observations.

A study by (Espinosa, Palma, Jiménez, Kamińska, Sciavicco, and Lucena-Sánchez, 2021) took into account 3 years worth of data which show nitrogen oxide levels in the air on an hourly basis; excessive levels of presence of these contaminants have been linked to a number of respiratory, circulatory, and even mental illnesses. For each measurement, nitrogen oxide concentrations are combined with weather and traffic information. A technique based on exactness and robustness criteria was presented to analyze various models for the prediction of different sorts of contaminants. The regression models are examined using a variety of window sizes. Deep learning models include Support Vector Machines, Lasso Regression, Random Forest, DCNN, GRU, and LSTM. As a result, our most accurate models provide a highly accurate projection of the amount of contaminants in the air in the location under consideration 24 hours in advance. This prediction may be used to plan and implement a range of measures and initiatives to mitigate the population's consequences.

This study (Freeman, Taylor, Gharabaghi, and Thé, 2018) used deep learning to predict the 8-hour mean surface ozone (O3) concentrations using a Long Short-Term Memory (LSTM) and Recurrent Neural Network (RNN). With minimal inaccuracy, values up to 72 hours were trained and projected using hourly air quality and meteorological data. The length of continuous O3 exceedances could also be predicted using the LSTM.(Shaban, Kadri, and Rezk, 2016) research presents a system for tracking and predicting urban air pollution. The system makes use of competitively cheap, multi-sensor gaseous and meteorological air-quality monitoring motes. With a framework for intelligent sensing made up of numerous modules, these motes wirelessly inter-

act. The modules are in charge of collecting and storing the data, preprocessing it to create information that is useful, Pollutants are predicted based on previous data, and the information is then presented through a variety of media, including mobile applications, web portals, and short messaging services. (Boonphun et al., 2019)Humans are known to suffer substantial health impacts from exposure to high levels of PM, particularly the tiny PM2.5 particles. Environmental consequences are obviously seen, in addition to health implications. The purpose of this research is to calculate the likelihood that PM2.5 will exceed a set safety threshold. In this paper, many machine learning methods are investigated. In particular, categorization models are put into practice based on weather information and air pollution characteristics obtained at various heights above ground level. These features are pushed back to different time steps, producing time-lagged features that are more revealing. A feature selection technique is also used to describe the ideal group of crucial features.

(Bhat et al., 2019) This paper considered Data from two Bengaluru-area sources: a government website and static sensors made using an Arduino board. Three computer algorithms, Random Forest Regression (RFR), Decision Tree Regression (DTR), and Linear Regression (LR), are used to calculate the CO level. The findings demonstrate that RFR provides the least inaccuracy of the three, leading to greater accuracy.

(Rahman, Panchenko, and Safarov, 2017) This paper discusses the development of two distinct prediction models based on neural networks for calculating the air pollution index. These models are spatial (prediction of atmospheric pollution index at every area inside a city) and temporal (short-term prediction of the contaminants in air for the next few days). The suitability of the forecasting models is also assessed on basis of calculation of the correlation coefficient between the output and reference data.

This work employs variant machine learning algorithm rather than reliance on one algorithm to achieve a better model for prediction. This work considered large dataset (1453 instances of each city) with more parameters and variables which can support more predictive model's performance in forecasting.

3. PROPOSED METHODOLOGY

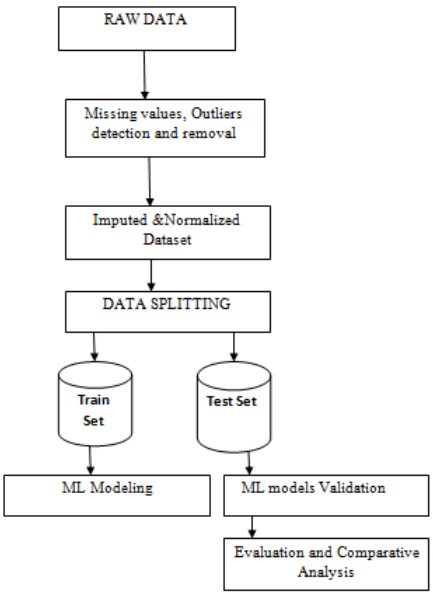

Figure 1. Showing the process for Forecasting AQI

Figure 1 shows flow-chart for the process of forecasting AQI.Dataset of Jind city and Panipat City is considered for the forecasting the value of AQI in which input dataset of approx two years from 2019 to 2021 is considered and result is validated and evaluated for the year 2022. Dataset is collected from the Central Pollution Control Board Website which contains numerous pollutant values and meteorological parameters of 1453 instances of each city. The pollutants like PM10, PM2.5, SO2, NO2 and Ozone were collected everyday from Jind city and Panipat City of Haryana and data is preprocessed and data is splitted into training set and testing set (80:20). Result is validated and performance is compared using performance metrics. Table1 provides information about the dataset that was used for the analysis.

Table-1 Details of Experimental Dataset

Parameters	**Value**
Dataset Year	2019-2022
Train:Test	80:20
Dataset Source	Central Pollution Control Board
Location	Jind and Panipat City of Haryana
Pollutants	NO2,PM2.5,PM10,SO2,O3
Weather Variables	Wind direction, Wind speed, Temperature and Relative humidity
Forecast Variable	AQI

Table I: Dataset Details

3.1 Approaches

a) **MLR**

Use multiple linear regression to determine the relationship between two or more manipulated variables and one measured variable. How closely two or more manipulated factors are related to one measured variable (for example, how added pollutants, temperature, humidity, and wind direction and speed affect air quality) can be determined using multiple linear regression. A multiple linear regression follows this formula:

$$z = a_0 + a_1 Y_1 + \ldots + a_n Y_n + \varepsilon \tag{1}$$

where:

$$\begin{aligned} a_0 &= \text{y-intercept} \\ a_1 Y_1 &= \text{regression coefficient of the first variable } (Y_1) \\ a_n Y_n &= \text{regression coefficient of the last manipulated variable} \\ n &= \text{model error} \\ z &= \text{the predicted value of the variable} \end{aligned}$$

b) **SVR**

Support Vector Machine (SVR)(Saeed, Hussain, Awan, and Idris, 2017), which forecasts real values rather than a class, uses the same basic idea as SVR. SVR acknowledges the non-linearity in the data while still providing an effective prediction model. When predicted variable is numerical rather than categorical, SVR is the modified form of SVM. The non-parametric nature of SVR is a significant advantage.

c) **RFR**

A random forest is a machine learning algorithm for tackling issues regarding regression and classification. It employs ensemble learning, a method for combining a number of classifiers to address complicated issues. There are too many possible decision trees in this algorithm. A "forest" is created by the random forest algorithm and trained using bagging or bootstrap

aggregation. This method determines the outcome using averages of outcomes from various decision trees.

d) **Hybrid Approach**
An approach is proposed which includes the RFR and SVR. Several simple algorithms complement and enhance one another. They can solve problems that they weren't meant to manage independently by working together.

3.2 Evaluation

Model performance MAE is calculated using Equation 2, MSE is calculated using Equation 3, RMSE is the square root of MSE which is calculated using Equation 4 and the performance indicator R square is calculated using Equation 5.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|a_i - \hat{a}_i| \tag{2}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(a_i - \hat{a}_i)^2 \tag{3}$$

$$RMSE = \sqrt{MSE} \tag{4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(a_i - \hat{a}_i)^2}{\sum_{i=1}^{n}(a_i - \bar{a})^2} \tag{5}$$

Here n is the sample size

a_i represents the actual value,

$\hat{a}_i$ represents the predicted value,

$\bar{a}$ denotes the mean of the values.

4. RESULTS AND DISCUSSION

Air Quality Index is forecasted using different approaches and data of the year 2019-2021 are used to train the model and data for the year 2022 is used for validation. Implementation is done using Python and its libraries.

—

[✓] Forecasting Using MLR Approach:

Fig 2-3 shows the forecasted values of AQI of Jind city and Panipat city for the year 2022 using Multiple Linear Regression approach. The applied method forecast the AQI value for the given dataset of Jind and Panipat City.

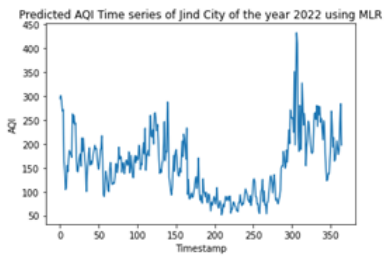

Figure 2. Forecasted AQI (Jind City) Time series using MLR

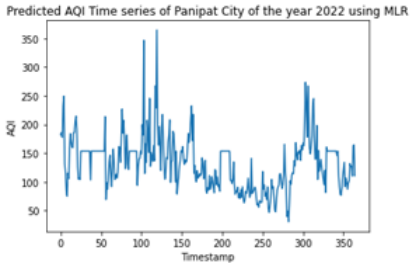

Figure 3. Forecasted AQI (Panipat City) Time series using MLR

Fig 4-5 shows the forecasted and actual values of AQI of Jind and Panipat city for the year 2022 and the comparison graph is plotted which shows the difference in the actual and forecasted values.

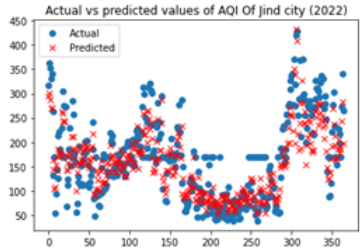

Figure 4. Actual and Forecasted AQI (Jind City) Time series using MLR

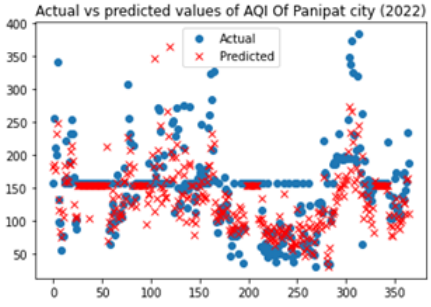

Figure 5. Actual and Forecasted AQI (Panipat City) Time series using MLR

[✓] Forecasting Using SVR Approach: Fig 6-7 shows the forecasted values of AQI of Jind and Panipat city for the year 2022 using Support Vector Regression approach. The applied methods forecast the AQI value for the given dataset of Jind and Panipat City.

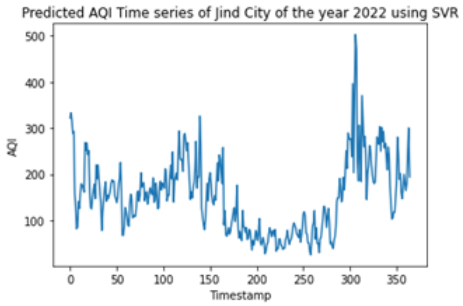

Figure 6. AQI forecasted values (Jind City) using SVR

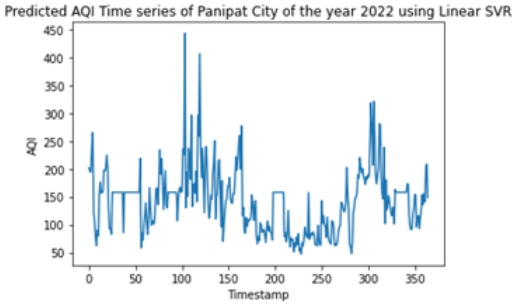

Figure 7. AQI forecasted values (Panipat City) using SVR

Fig 8-9 shows the forecasted and actual values of AQI of Jind city and Panipat city for the year 2022 and the comparison graph is plotted which shows the difference in the actual and forecasted values using SVR approach.

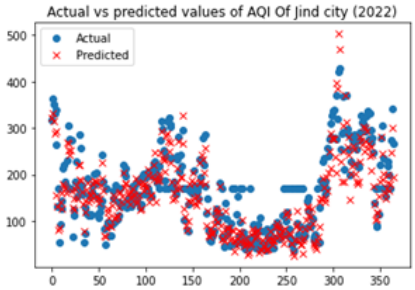

Figure 8. Actual and Forecasted AQI (Jind City) Time series using SVR

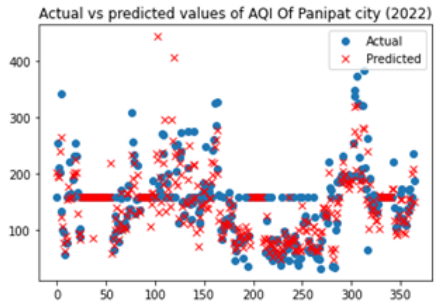

Figure 9. Actual and Forecasted AQI (Panipat City) Time series using SVR

[✓] Forecasting Using RFR Approach: Fig 10-11 shows the forecasted values of AQI of Jind city and Panipat city for the year 2022 using Random Forest Regression approach. The applied methods forecast the AQI value for the given dataset of Jind city and Panipat City.

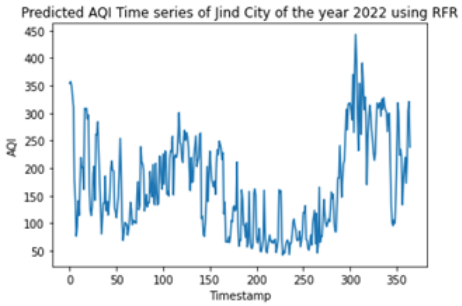

Figure 10. AQI forecasted values (Jind City) using RFR

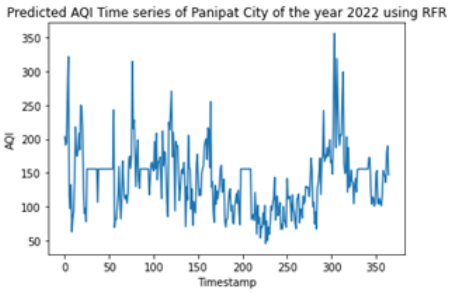

Figure 11. AQI forecasted values (Panipat City) using RFR

Fig 12-13 shows the forecasted and actual values of AQI of Jind city and Panipat city for the year 2022 and the comparison graph is plotted which shows the difference in the actual and forecasted values using RFR approach.

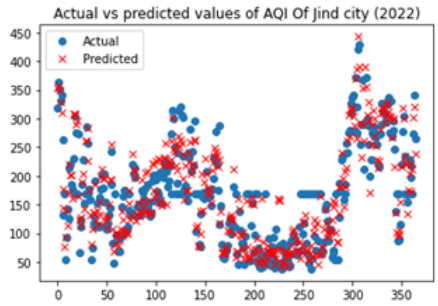

Figure 12. Actual vs AQI forecasted values (Jind City) using RFR

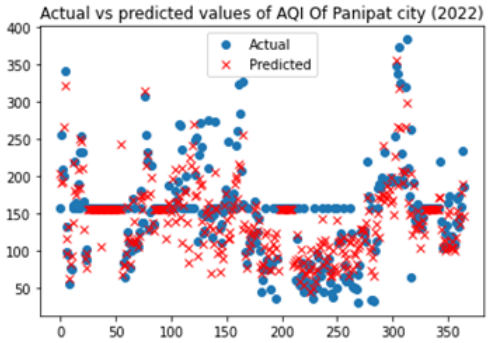

Figure 13. Actual vs AQI forecasted values (Panipat City) using RFR

[✓] Forecasting Using Hybrid Approach

Fig 14-15 shows the forecasted values of AQI of Jind city and Panipat city for the year 2022 using hybrid approach which combines SVR and RFR. The applied methods forecast the AQI value for the given dataset of Jind city and Panipat City.

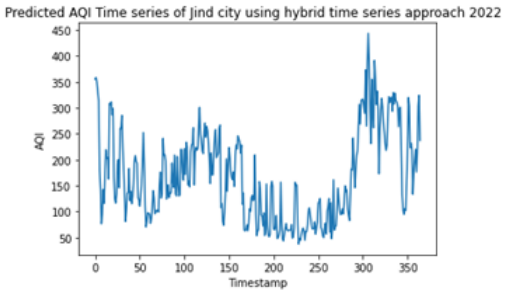

Figure 14. AQI forecasted values (Jind City) using Hybrid Approach

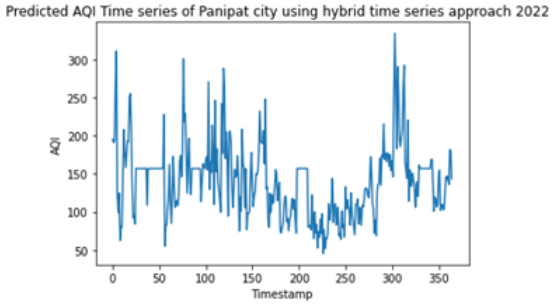

Figure 15. AQI forecasted values (Panipat City) using Hybrid Approach

In Fig 16-17 the actual value is shown in blue color circle for the AQI of Jind city and Panipat City, the forecasted value of the AQI is shown in red color cross using Hybrid approach. The model is observed to do a decent job of fitting both the training and test data.

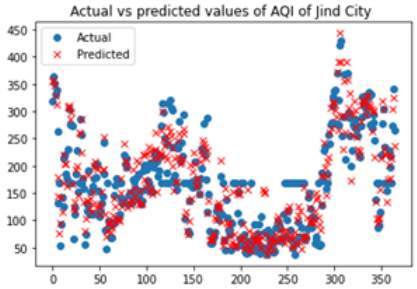

Figure 16. Actual vs AQI forecasted values (Jind City) using Hybrid Approach

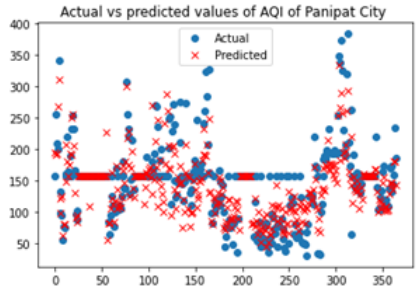

Figure 17. Actual vs AQI forecasted values (Panipat City) using Hybrid Approach

5. PERFORMANCE ANALYSIS

The performances of various approaches are analyzed on the basis of evaluation parameters such as Absolute Error, MSE, R2.The performance of the approaches is compared on the basis of the parameters. Values are shown in Table II.

It has been observed from the table II that the performance of hybrid approach is better in forecasting the Air Quality Index of Jind city as compared to the other approaches.

Table III shows that the performance of Hybrid approach is best and efficient in forecasting followed by RFR approach.

Algorithm	MAE	MSE	RMSE	R square
MLR	33.91502207582924	1846.9217466425782	42.97582746896886	0.7451440355966558
Linear SVR	30.650362858444247	1806.0676926619412	42.49785515366559	0.7507814695301447
Gaussian SVR	36.86712083903777	2233.676751725465	47.26178955271864	0.6917758731461358
RFR	27.882126505470257	1620.4026160249819	40.25422482206038	0.776401316309441
Proposed Hybrid	26.785454536451308	1549.8122177040416	39.36765446028048	0.7861420560426773

Table II: Performance Analysis of various approaches in Forecasting AQI (Jind City)

Algorithm	MAE	MSE	RMSE	R square
MLR	29.84153178724068	1795.0200481675258	42.36767692672712	0.5047064559803252
Linear SVR	22.547601957738973	1469.106676809112	38.32892741532317	0.5946345818017054
Gaussian SVR	24.674839065900507	1371.719234806852	37.03672818712328	0.6215063548170272
RFR	24.56880283380499	1314.6144371305365	36.25761212670432	0.6372630800137751
Proposed Hybrid	24.525400488778672	1217.8194662577373	34.89727018346474	0.6639713745622466

Table III: Performance Analysis of various approaches in Forecasting AQI (Panipat City)

6. CONCLUSION

Due to the dynamic environment and variety of pollutants in the air, it is quite difficult to predict or forecast the air quality. Machine learning and statistical methods are applied to forecast the air quality index of the year 2022 of Jind and Panipat City of Haryana using dataset for the year 2019-2021. Accurately predicting pollutant components or air quality is quite challenging.

In the current paper, 4 years worth of Haryana city air pollution statistics are examined. By filling in NAN, None, and NA values, addressing outliers, and normalizing data values, the dataset is initially cleaned and preprocessed. The methodologies for data analysis are put to use to discover numerous hidden patterns that are present in the dataset. The dataset is divided between train and test subgroups at a ratio of 80:20. Techniques are used to do ML-based AQI prediction, and a comparison analysis is then provided.

The result shows that the Hybrid approach is more efficient as compared to the others in forecasting the AQI as it has high R2 and low error rate.

References

Aarthi, A., Gayathri, P., Gomathi, N., Kalaiselvi, S., and Gomathi, V. 2020. Air quality prediction through regression model. *Int. J. Sci. Technol. Res Vol.9,* No.3.

Aditya, C., Deshmukh, C. R., Nayana, D., and Vidyavastu, P. G. 2018. Detection and prediction of air pollution using machine learning models. *International Journal of engineering trends and Technology (IJETT) Vol.59,* No. 4, pp.204–207.

Bhat, A., Manek, A. S., and Mishra, P. 2019. Machine learning-based prediction system for detecting air pollution. *Int J Eng Res Technol Vol.8*, pp.155–9.

Boonphun, Jirat, Kaisornsawad, Chalet, Wongchaisuwat, and Papis. 2019. Machine learning algorithms for predicting air pollutants. In *E3S Web of Conferences*. EDP Sciences, pp 03004.

Espinosa, R., Palma, J., Jiménez, F., Kamińska, J., Sciavicco, G., and Lucena-Sánchez, E. 2021. A time series forecasting based multi-criteria methodology for air quality prediction. *Applied Soft Computing Vol. 113*, pp. 107850.

Freeman, B. S., Taylor, G., Gharabaghi, B., and Thé, J. 2018. Forecasting air quality time series using deep learning. *Journal of the Air & Waste Management Association Vol. 68,* No. 8, pp. 866–886.

Harishkumar, K., Yogesh, K., Gad, I., et al. 2020. Forecasting air pollution particu-

late matter (pm2.5) using machine learning regression models. *Procedia Computer Science Vol.171*, pp. 2057–2066.

Kumar, K. and Pande, B. 2023. Air pollution prediction with machine learning: a case study of indian cities. *International Journal of Environmental Science and Technology vol.20,* No.5, pp.5333–5348.

Lei, T. M., Siu, S. W., Monjardino, J., Mendes, L., and Ferreira, F. 2022. Using machine learning methods to forecast air quality: A case study in macao. *Atmosphere vol.13,* No.9, pp. 1412.

Mahalingam, Usha, Elangovan, Kirthiga, Dobhal, Himanshu, Valliappa, Chocko, Shrestha, Sindhu, Kedam, and Giriprasad. 2019. A machine learning model for air quality prediction for smart cities. In *2019 International Conference on wireless communications signal processing and Networking (WiSPNET)*. IEEE, pp 452–457.

Patil, R. M., Dinde, H., Powar, S. K., and Ganeshkhind, P. M. 2020. A literature review on prediction of air quality index and forecasting ambient air pollutants using machine learning algorithms. *International Journal of Innovative Science and Research Technology Vol.5,* No.8.

Rahman, P., Panchenko, A., and Safarov, A. 2017. Using neural networks for prediction of air pollution index in industrial city. In *IOP Conference Series: Earth and Environmental Science*. Vol. Vol.87. IOP Publishing, pp.042016.

Saeed, S., Hussain, L., Awan, I. A., and Idris, A. 2017. Comparative analysis of different statistical methods for prediction of pm2. 5 and pm10 concentrations in advance for several hours. *IJCSNS International Journal of Computer Science and Network Security Vol.17,* No.11, pp.45–52.

Sanjeev, D. 2021. Implementation of machine learning algorithms for analysis and prediction of air quality. *International Journal of Engineering Research & Technology (IJERT) Vol.10,* No.3, pp.533–538.

Sankar Ganesh, S., Arulmozhivarman, P., and Tatavarti, R. 2017. Forecasting air quality index using an ensemble of artificial neural networks and regression models. *Journal of Intelligent Systems Vol.28,* No.5, pp.893–903.

Shaban, K. B., Kadri, A., and Rezk, E. 2016. Urban air pollution monitoring system with forecasting models. *IEEE Sensors Journal* No.8, pp. 2598–2606.

Zamani Joharestani, M., Cao, C., Ni, X., Bashir, B., and Talebiesfandarani, S. 2019. Pm2. 5 prediction based on random forest, xgboost, and deep learning using multisource remote sensing data. *Atmosphere vol.10,* No.7, pp.373.

Zhu, J., Wu, P., Chen, H., Zhou, L., and Tao, Z. 2018. A hybrid forecasting approach to air quality time series based on endpoint condition and combined forecasting model. *International Journal of environmental research and Public Health Vol.15,* No.9, pp. 1941.

Ms. Reema Gupta is the research scholar at the Baba Mastnath University, Rohtak, and working as an Assistant Professor in Goverment College Under Higher Education,Haryana.She is having 8 years of teaching experience. Ms. Gupta pursued her engineering education at Guru Jambheshwar University,Hisar and Master's under Kurukshetra University,Kurukshetra and qualified for GATE in 2012 and UGC-NET in 2014. She has presented various papers in national and international seminar,conferences and many research papers published in international journals.
Email: reema2405@gmail.com

Dr. Priti Singla is Professor in Computer Science and Engineering department at Baba Mastnath University,Rohtak. With over 21 years of teaching experience and a trail of accomplishments, she has established herself as a distinguished figure in academia. She is holding PhD, M.Tech, M.Sc, Dr. Singla's academic prowess is further highlighted by the remarkable achievement of being a Double Gold Medalist. Her illustrious career includes teaching positions at esteemed institutions such as Carleton University in Canada ,Georgia Institute of Technology,Bowling Green State University and Heriot-Watt University,where she taught a wide range of courses and made significant contributions to the field.
Email: pritisingla04@gmail.com