IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed journal / Refereed journal ⁜ Vol. 12, Issue 10, October 2023

DOI: 10.17148/IJARCCE.2023.121015

Predicting Air Quality Index using most suitable ML model

Aishwarya Sajjan[1], M. Farida Begam[2], Ayush Dubey[3]

Department of Information Science and Engineering, CMR Institute of Technology, Bengaluru, India[1-3]

Abstract: Air quality fundamentally affects human wellbeing. Deterioration in air quality prompts an extensive variety of medical problems, mainly in young ones. The capacity to anticipate air quality allows the authority and also other related associations to take essential measures to protect the susceptible, while being prone to the air standards. Conventional ways of managing this problem have exceptionally limited success for the reason that an absence of access to such practices to adequate panel data. In this paper, we utilize a Support Vector Regression (SVR) model to predict the degrees of different contaminations and the Air Quality Index (AQI), utilizing chronicle pollution information which is made freely available by the Central Pollution Control Board and the US Embassy in New Delhi. Amidst the tried strategies, a Radial Basis Function (RBF) kernel generated the best outcomes with the SVR. In accordance with the analysis, utilizing the entire variety of accessible variables led to improved outcomes than utilizing characteristics chosen by principal component analysis. This model estimates levels of different pollutants, namely, nitrogen dioxide, sulfur dioxide, carbon monoxide, ground-level ozone, and particulate matter 2.5, as well as the Air Quality Index (AQI).

Keywords: Air Quality Index, Support Vector Regression, Radial Basis Function (RBF), Principal component analysis, Central Pollution Control Board

I. INTRODUCTION

The rapid increase in air pollution lately, because of agricultural and industrial activities, also an increased number of vehicles using gas powered engines, has caught the attention of the scientific community. Air pollution essentially affects human wellbeing and may cause long-term health issues in young ones. The significant increase in air pollution in New Delhi is attributed to increased vehicular emissions, burning of non-renewable energy sources at power plants, and other nearby industries and burning of fields by farmers in bordering states.

Air quality has been observed in New Delhi for around twenty years. This has allowed a deeper awareness of the changes in air contamination in response to particular activities and government regulations, but the air pollution in New Delhi remains a problem.

Air pollution is accountable for 30 percent of lower respiratory tract infections and is connected with 91 percent of premature deaths, from heart disease, acute respiratory infections, lung cancer, stroke and chronic obstructive pulmonary disease. It contributes to 20 percent of infant mortality worldwide and causes numerous short- and long term illnesses in children. Exposure of the mother to elevated levels of air pollution can negatively influence immune status, brain development, respiratory systems, and cardiometabolic health of the child. Air contamination has likewise been connected to low birth weight and stunted growth in children.

Air pollution is estimated to be liable for one in ten deaths of children under five years of age. In elderly people, air pollution causes high rates of asthma, with decreased cognitive performance. As India is the largest developing industrial country, it is producing a record measure of pollutants, explicitly CO2, PM2.5 etc and other harmful aerial contaminants. Air quality of a specific state or a nation is a measure on the impact of pollutants on the respective regions, according to the Indian air quality standard pollutants are recorded in terms of their scale, these air quality indexes demonstrate the levels of vital pollutants in the atmosphere. There are different atmospheric gases which cause contamination in our environment. Each pollution has an individual index and scales at various levels. The major pollutants such as (NO2, SO2, RSPM, SPM) indexes AQI are acquired, with this individual AQI, the data can be categorized on the basis of the limits.

Currently, the government implements regulations after the air quality has arrived at perilous levels. If there is a technique to foresee the air contamination reaching hazardous levels, the government can apply such regulations early, potentially preventing further deterioration of air contamination and being able to protect those from getting exposed to such air quality.

DOI: 10.17148/IJARCCE.2023.121015

This study aims to create a model that can look at previously recorded air contamination data and predict levels of various pollutants and also the Air Quality Index. For this we use a variation of Support Vector Machines (SVM), called Support Vector Regression (SVR).

The paper is structured as follows. We state the motivations of this work and manage the related work that is accomplished in this field and, what are the results obtained, in section 2, stating the possible impact of being able to successfully predict the air quality. In section 3, we discuss the theoretical background, existing system and proposed system. Section 4 contains the methodology adopted for predicting air quality, and in section 5 we elaborated the results and performance analysis. In Section 6, we conclude the paper and discuss the results.

II. RELATED WORK

Air pollution resulted in the introduction of detrimental or imprudent quantities of substances into the air. It has been resolved that there is a relationship between exposure to such pollutants for a brief time and health issues like inflamed respiratory track in healthy individuals, elevated respiratory symptoms in individuals with asthma, trouble meeting high oxygen requirements while exercising and critical respiratory circumstances, particularly in children and the venerable. Hence, it is essential that we have a system to forecast rise in air pollution levels, so that government organizations might have the option to counter further increment through on-demand pollution control systems or an emergency response. This would make AQI more controllable to suit the overall demands of the populace.

The best statistical technique for anticipating time series data is the Autoregressive Integrated Moving Average model (ARIMA) depicted by G. E. Box and D. A. Pierce, “Distribution of residual autocorrelations in autoregressive-integrated moving average time series models”, 1970. It has several advantages in respect to its statistical properties , potential for an enormous number of applications and extendibility. With the expansion in importance for predicting air quality levels, ARIMA was applied to this task as well. It was demonstrated to reach accuracies around 95% for forecasting AQI monthly values.

J. Zhu, R. Zhang, B. Fu, and R. Jin, “Comparison of ARIMA model and exponential smoothing model on 2014 air quality index in yanqing county, Beijing, China.”, 2015 compared the interpretation of ARIMA with a Holt exponential smoothing model and demonstrated the superiority of the ARIMA model for forecasting AQI day to day values.

I. Goodfellow, Y. Bengio, and A. Courville, “Machine learning basics,” 2016 says availability of enormous amounts of archive data made it advantageous to utilize Machine Learning models for the time series prediction of AQI. ML models can automatically see a great deal of data and select significant features, thus reducing the requirement for human intervention. ML models are able to accomplish higher accuracies with enormous datasets, than classic statistical methods. Such models have for quite a while been utilized for AQI forecasting tasks.

ML models are nonlinear, nonparametric in nature and subsequently are better able to deal with the complexity of nonlinear elements like pollutant levels in the air. Thus, they outperform statistical methods like ARIMA, Winter exponential smoothing, and multivariate regression, which function well only with linear systems. ML models, for instance, Genetic programming, Support Vector Machine (SVM) and Artificial Neural Networks (ANN) are able to find hidden patterns in immense amounts of data.

Recent research concerns on latest statistical learning algorithms for air standards estimation and air contamination forecast. Raimondo et al., Park et al. and Garcia et al. have utilized neural networks that are used to enhance models for anticipating the extensiveness of single pollutants, for instance, particulate matter estimating under 10microns (PM10). Raimondo et al. used a SVM and ANN to train models. Best ANN model accomplished nearly 79 percent for accuracy with just a 0.82 percent false-positive rate, altogether their finest SVM model at a accuracy of 80 percent with a false-positive rate of just 0.13 percent. Yu et al. came up with a random forest proposal, named as RAQ, for AQI group prediction.

Later, Yi et al. enforced deep neural networks for AQI classification prediction. Veljanovska and Dimoski applied various settings to exceed the performance of K-Nearest Neighbor, SVM and decision tree for anticipating the AQI levels. The ANN accomplished a precision of 92.3 percent, outperforming all available tested algorithms.

E. Kalapanidas and N. Avouris says that external air quality performs an essential role in human well-being. Air pollution causes huge increases in scientific expenses, illness and also is evaluated to be almost 800,000 yearly rushed demises worldwide.

DOI: 10.17148/IJARCCE.2023.121015

However, most people spend almost every hour indoors of their homes, external air exceptions might upset the air that is inside to a huge measure. Moreover, numerous victims correlating to patients with allergies and chemical sensitivities, asthmatics, heart victims, stroke victims, COPD patients, pregnant ladies, diabetics, the elderly and young ones are mainly at risk. The extremity of the analysis work is on perceiving and estimating.

V. M. Niharika and P. S. Rao has shown viability of Recurrent Neural Networks (RNN) in managing time-related data. Yet, data from the future which might arise next illustrates that the current time is required for forecasting. RNNs can partially obtain this through postponing the outcome result with guide of a particular quantity of time-frames to integrate future understanding. Theoretically, a tremendous length can be applied yet in perception, it's found that forecast results drip if the extent is excessively enormous. To defeat these impediments, they came up with Bidirectional Recurrent Neural Community (BRNN) that might be capable of using all on hand input knowledge beforehand and future of a certain time period.

D. J. Nowak, J. C. Stevens and D. E. Crane says that evaluating air contamination in urban regions is a crucial errand. Source records no longer have anything to do with data and are not available. Thus, the predicted outcomes of a numerical model may not be exact and adequate. To overcome this, this study proposes a total analysis system framework to make the predicted outcomes sufficient and accurate. Trials on discrete characteristic organizations have given discoveries which are especially helpful. The AQI is utilized to calculate the quantity of contamination in air in a specific urban region.

III. THEORETICAL BACKGROUND

Perceiving and preserving high quality air has become a critical obstacle in metropolitan regions that have many organizations, industries and populace. As there is growth in the populace, there is an increase in the transportation, use of fuels and electricity.

A lot of waste is unloaded on the land that we are widely aware of. Air is likewise exceptionally contaminated which makes it a serious threat to a wide collection of living organisms on earth. And this brings about the necessity to monitor and assess the standard of air and appropriately the government authority should be made aware of taking necessary moves.

The goal of the “Predicting Air Quality Index using the most suitable ML model” is to discover the most appropriate machine learning model for the everyday prediction of the AQI for a particular area which is found on the historical data.

Data collection:

We gather the data from the Indian government database, that contains pollutant concentrations happening in the Safdarjung region of Delhi . We start by working out the individual index of the pollutant for each accessible data point and note their corresponding AQI for the area.

Preprocessing:

The data that we get is from various sources which might contain uncertain data, repeated data and missing values. To get appropriate prediction outcomes, the dataset should be cleaned, and missing values should be dealt with either by erasing or by loading up with average values or another method. Additionally redundant data should be taken out or disposed to try not to incline to the outcomes. Not many data set may have a few exceptional or outrageous values which additionally must be eliminated to get good expected accuracy. Clustering and classification algorithms, and other data mining techniques will function admirably provided that preprocessing is finished on the information.

Building the classification model:

● Firstly, we need to divide the data set into testing and training datasets. Predicting model is trained at first using the training data set. Next the testing set will be tested. Or else k-fold cross validation can be utilized.

● Subsequent to testing the model, the precision is evaluated by utilizing the parameters like MAE, MSE, RMSE.
We have outlined a model to forecast the AQI of every data point that is available in dataset. Our model is able to forecast the air quality of India in any given region. By forecasting the AQI, we can backtrack the significant pollution causing pollutant and the area that is seriously affected by the pollutant across India. With this prediction model, different information about the data is removed using various methods to get intensely impacted areas on a specific region(cluster).

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed journal / Refereed journal ⁜ Vol. 12, Issue 10, October 2023

DOI: 10.17148/IJARCCE.2023.121015

IV. METHODOLOGY

Machine Learning Prediction Models:

Linear Regression:

It is likely the technique where almost all academicians began their first machine learning understanding. Its foremost working guideline is situated behind the fitting of a single or more independent variable with dependent into a line in n proportions. n generally means the quantity of variables in a data set. The line is probably made as it limits all errors while attempting to fit every instance in line. In machine learning, this regression is enabled with the ability to advance constantly by enhancing the variables.

Oftentimes, optimization is done by a technique called gradient descent. It acts to a limited extent by obtaining the loss function and each of the parameters will be modified by deducting the prior value along with the derivative times at a particular learning rate. Rate can be tuned up by the uncomplicated method, that is a rule of thumb, trial and error, or advanced guideline, for instance, meta-heuristic. Other variable which is left for tuning is number of derivations that is added to this model. Regularisation is done as an attempt to reduce the odds of exceeding and expand the robustness. There are two kinds of regularisation that is utilized in linear regression, namely, lasso and ridge. Lasso regularisation will eradicate less significant features by allowing the attribute's coefficient to zero, and hold another more significant one. Further, Ridge regularization won't attempt to remove a feature, however it will try to reduce the extent of coefficients to have a lower conflict.

Lasso Regression:

Lasso is a kind of linear regression which utilizes the shrinkage. Shrinkage is a place where the data values are contracted to a middle point, similar to the mean. Lasso methodology supports basic, sparse models (that is models with fewer parameters). This specific type of regression is suitable for models showing elevated levels of multicollinearity or when you require to automate specific parts of model selection, identical to variable selection or parameter elimination. Lasso regression executes L1 regularization, that adds to a penalty which is proportionate to absolute value of magnitude of the coefficients. This kind of regularization may bring sparse models with less coefficients; Certain coefficients can turn into zero and eradicate from model. Huge penalties can derive coefficient values near to zero, that is optimal for making easier models. However, L2 regularization, for instance ridge regression, doesn't result in removal of sparse models or coefficients. This causes the lasso to be far more straightforward to interpret than the ridge.

Lasso Regression results are quadratic programming problems which are best tackled with software (for example Matlab).

XG Boost:

It is an open-source application software library that gives a regularising gradient boosting foundation for Java, Python, Scala, R, Perl, C++, and Julia. It works on Windows, Linux and also macOS. In the project description, it aims to give a "Scalable, Portable and Distributed Gradient Boosting (GBM, GBRT, GBDT) Library". It runs on one machine, also on distributed processing frameworks such as Apache Spark, Apache Hadoop, Apache Flink, and Dask. It has acquired a lot of recognition and attention lately as the algorithm of choice for the overwhelming majority of winning teams of ML competitions.

Salient features of XGBoost that makes it different from other gradient boosting algorithms includes:

· Clever penalization of trees
· A proportional shrinking of leaf nodes
· Newton Boosting
· Extra randomization parameter
· Automatic Feature selection

Random Forest Regression:

A supervised learning algorithm which uses ensemble learning strategies for regression. Ensemble learning technique is a way that combines forecasts from numerous machine learning algorithms to compute a more precise prediction than a single model.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed journal / Refereed journal ⁜ Vol. 12, Issue 10, October 2023

DOI: 10.17148/IJARCCE.2023.121015

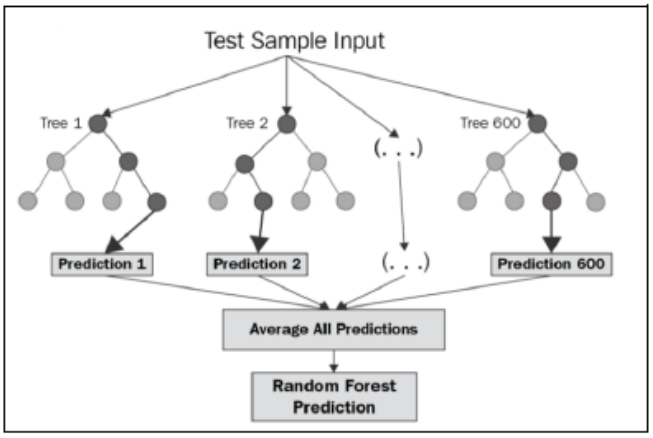


Fig. 1

The Fig. 1 above shows the framework of a Random Forest. You can see that the trees run in parallel with no interaction amongst themselves. A Random Forest works by developing several decision trees in training time and outputting the mean of the classes as the forecast of all the trees. To get greater understanding of Random Forest algorithm, let's walk through the steps:

- Pick at random k data points in the training set.
- Build a decision tree that is associated with k data points.
- Choose the number N of trees that you need to build and then repeat the above two steps.
- For other data point, make each and every one of your N-tree trees anticipate the value of y for the data in question and then allocate the other data point to average across the predicted y values.

Random Forest Regression is powerful and precise. It typically performs well on numerous problems that include features with non-linear relationships. Disadvantages, on the other hand, include the following: interpretability is absent, overfitting may easily occur, and we must select the total number of trees to be included.

K-Nearest Neighbors Algorithm (KNN):

It is a non-parametric classification method first created by Evelyn Fix and Joseph Hodges in 1951, and later extended by Thomas Cover. It is utilized for regression and classification. In the two cases, the input comprises the k nearest training instances in a data set. The output relies upon whether k-NN is utilized for classification or regression:

- In kNN classification, the result is a class membership. An object is classified by a majority voting of its neighbors, with the object being allocated to class generally among its k nearest neighbors (k is a positive integer, usually small). If k = 1, then the object is basically given to the class of that single nearest neighbour.
- In kNN regression, the result is the property value for the object. This value is the median of the values of k nearest neighbors. It is a kind of classification where the purpose is just being close to nearby and all the calculations will be deferred up till function evaluation. As this algorithm depends on distance for classification, if the properties appear to be dissimilar physical units or come in immeasurably various scales the normalizing the training information can enhance its precision decisively. Both for regression and classification, an effective method can be to dole out the weights to the contributions of neighbors, in order that the nearer neighbours give more to the mean than the more far off ones. For instance, a typical weighting scheme comprises providing each neighbor a weight of 1/d, where d is the distance to the neighbor.

The neighbors are grasped from a collection of objects for which the class (for kNN classification) or the object property estimation (for kNN regression) is known. This can be intended as the training set for the algorithm, however no explicit training step is needed.

Discussion of Metrics Used for Prediction:
We evaluate the performance using the standard metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE).

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed journal / Refereed journal ⁜ Vol. 12, Issue 10, October 2023

DOI: 10.17148/IJARCCE.2023.121015

MEAN ABSOLUTE ERROR (MAE)

The average magnitude of the errors in a data values set (predictions) is measured using Mean Absolute Error, without considering the direction. The mean of the absolute differences between the actual and prediction observations in a test sample is termed as MAE. It is determined as in Equation:

$$MAE = (1/n)\ \Sigma(i=1 \text{ to } n)\ |y_i - \hat{y}_i|$$

where n = Number of Observations, y_i = Actual Value, $\hat{y}_i$ = Predicted Value

ROOT MEAN SQUARED ERROR (RMSE)

The average magnitude of the error can be calculated using RMSE also. It may be calculated by taking the median of squared differences among the actual and predicted values and taking square root of final result. It is determined as:

$$RMSE = \sqrt{(1/n)\ \Sigma(i=1 \text{ to } n)\ (y_i - \hat{y}_i)^2}$$

where n = Number of Observations, y_i = Actual Value, $\hat{y}_i$ = Predicted Value

With the aim to analyze the data sets we standardize the RMSE as follows:

$$\text{Normalized RMSE} = \text{RMSE}\ y_{max} - y_{min}$$

where y_{max} = Maximum value of dataset and y_{min} = Minimum value of dataset.

MEAN SQUARED ERROR (MSE)

An estimator (who calculates an unobserved quantity or value) calculates the average of the squares of the errors — that is to say, the average squared difference between estimated values and actual values. It is a peril function, correlating to the anticipated value of squared error loss. The verity that MSE is nearly inevitably positive (and not zero) is for the reason of randomness or as the estimator does not account for details that could make a more precise calculation.

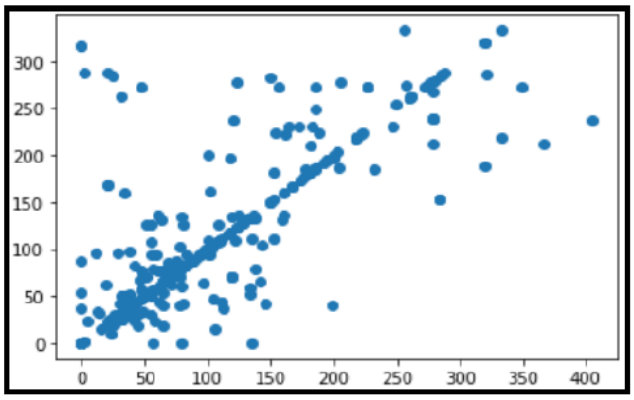
It is the amount of the quality of an estimator. It is obtained from a square of Euclidean distance, so it is always a positive value that reduces as the error proceeds towards zero.

V. RESULTS & PERFORMANCE ANALYSIS

OUTPUT GRAPHS:

In this present section, snapshots showing the functioning of final code that is generated is shown. For each model, graphs are shown where X-axis has the original values and Y-axis has the predicted values.

K-Nearest Neighbour:



Mean Absolute Error : 26.46491361788618
Mean Squared Error : 3282.779510713076
Root Mean Squared Error : 57.295545295538254

Fig. 2

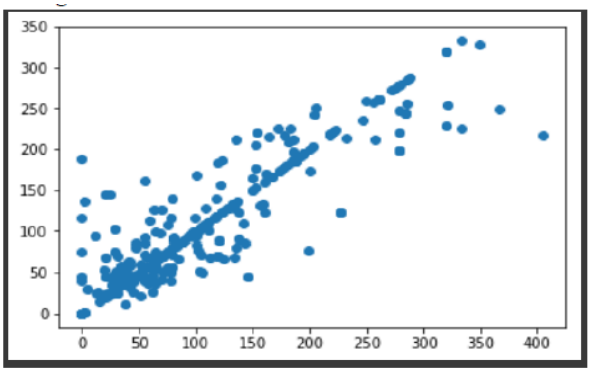
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed journal / Refereed journal ⁜ Vol. 12, Issue 10, October 2023

DOI: 10.17148/IJARCCE.2023.121015

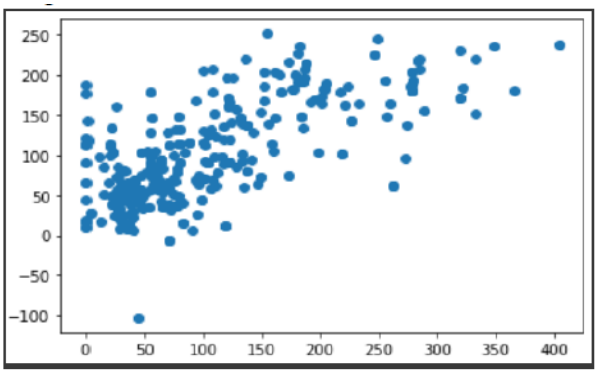
XGBoost Regressor:



Mean Absolute Error : 25.245581808943093
Mean Squared Error : 1681.8142776645072
Root Mean Squared Error : 41.009929013161035

Fig. 3

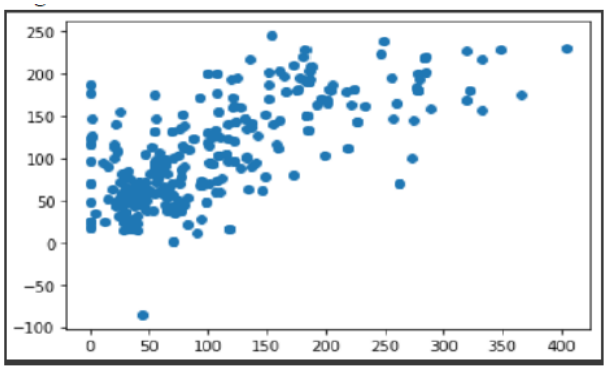
Linear Regression:



Mean Absolute Error : 25.245581808943093
Mean Squared Error : 1681.8142776645072
Root Mean Squared Error : 41.009929013161035

Fig. 4

Lasso Regression:



Mean Absolute Error : 44.50831198875126
Mean Squared Error : 3627.8109390424697
Root Mean Squared Error : 60.2313119485411

Fig. 5

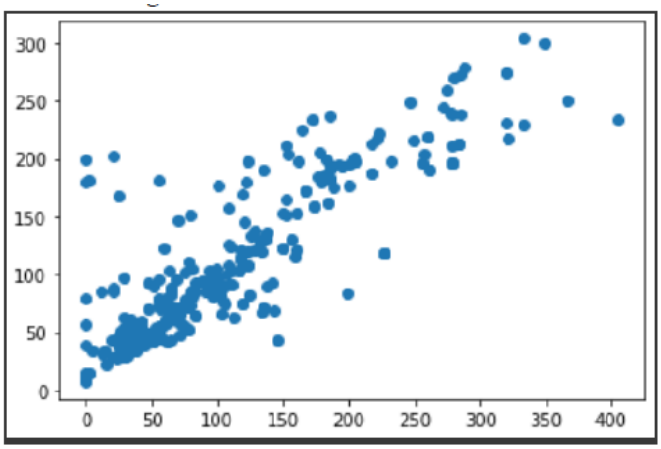
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed journal / Refereed journal ⁜ Vol. 12, Issue 10, October 2023

DOI: 10.17148/IJARCCE.2023.121015

Random Forest Regression:



Mean Absolute Error : 24.74410997787267
Mean Squared Error : 1675.4203418373647
Root Mean Squared Error : 40.931898830097836

Fig. 6

Summary:

The table shows each model's performance on forecasting levels of pollutants, using publicly available data for New Delhi, in terms of MAE, MSE, RMSE values.

ML Model	Mean Absolute Error (MAE)	Mean Squared Error (MSE)	Root Mean Squared Error (RMSE)
K-Nearest Neighbor	26.46	3282.78	57.30
XGBoost Regressor	25.24	1681.81	41.01
Linear Regression	44.83	3687.54	60.72
Lasso Regression	44.50	3627.81	60.23
Random Forest Regressor	24.74	1675.42	40.93

VI. CONCLUSION

The task of forecasting pollutant levels is innately hard since the nature of the information is volatile and dynamic and its irregularity in space and time. Nevertheless, the task of forecasting pollutant levels has been expanding significantly since the impact of pollution on the population and the ecosystem.

In this task, we have used various ML models for forecasting levels of pollutants like PM2.5 and Air Quality Index (AQI), using publicly accessible data for New Delhi and figured out that Random Forest Regressor gives the most accurate results with Mean Average Error of 24.75, Mean Squared Error of 1675.42 and Root Mean Squared Error of 40.93.

DOI: 10.17148/IJARCCE.2023.121015

REFERENCES

[1] U. A. Hvidtfeldt, M. Ketzel, M. Sørensen, O. Hertel, J. Khan, J. Brandt, and O. Raaschou-Nielsen, "Evaluation of the danish airgis air pollution modeling system against measured concentrations of pm2.5, pm10, and black carbon," Environmental Epidemiology, vol. 2, no. 2, p. e014, 2018.

[2] Y. Gonzalez, C. Carranza, M. Iniguez, M. Torres, R. Quintana, A. R. Osornio-Vargas, C. Gardner, S. Sarkar, and S. Schwander, "Inhaled air pollution particulate matter in alveolar macrophages alters local pro-inflammatory cytokine and peripheral ifn production in response to mycobacterium tuberculosis," in B17. MYCOBACTERIAL HOST DEFENSES. American Thoracic Society, 2017, pp. A2901–A2901.

[3] V. Kanawade, A. Srivastava, K. Ram, E. Asmi, V. Vakkari, V. Soni, V. Varaprasad, and C. Sarangi, "What caused severe air pollution episode of november 2016 in new delhi?" Atmospheric Environment, vol.222, p. 117125, 2020.

[4] B. R. Gurjar, L. T. Molina, and C. S. P. Ojha, Air pollution: health and environmental impacts. CRC press, 2010.

[5] D. C. Payne-Sturges, M. A. Marty, F. Perera, M. D. Miller, M. Swanson, K. Ellickson, D. A. Cory-Slechta, B. Ritz, J. Balmes, L. Anderko etal., "Healthy air, healthy brains: advancing air pollution policy to protect children's health," American journal of public health, vol. 109, no. 4, pp. 550–554, 2019.

[6] M. Castelli, F. M. Clemente, A. Popovi˘c, S. Silva, and L. Vanneschi, "A machine learning approach to predict air quality in california," Complexity, vol. 2020, 2020.

[7] G. E. Box and D. A. Pierce, "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models," Journal of the American statistical Association, vol. 65, no. 332, pp. 1509–1526, 1970.

[8] J. Zhu, R. Zhang, B. Fu, and R. Jin, "Comparison of arima model and exponential smoothing model on 2014 air quality index in yanqing county, beijing, china," Appl. Comput. Math, vol. 4, pp. 456–461, 2015.

[9] I. Goodfellow, Y. Bengio, and A. Courville, "Machine learning basics," Deep learning, vol. 1, no. 7, pp. 98–164, 2016.

[10] R. Sharda and R. Patil, "Neural networks as forecasting experts: an empirical test," in Proceedings of International Joint Conference on Neural Networks, vol. 2. IEEE, 1990, pp. 491–494.

[11] I. Alon, M. Qi, and R. J. Sadowski, "Forecasting aggregate retail sales:: a comparison of artificial neural networks and traditional methods," Journal of retailing and consumer services, vol. 8, no. 3, pp. 147–156, 2001.

[12] Raimondo, G.; Montuori, A.; Moniaci, W.; Pasero, E.; Almkvist, E. A Machine Learning Tool to Forecast PM10 Level. In Proceedings of the Fifth Conference on Artificial Intelligence Applications to Environmental Science, San Antonio, TX, USA, 14–18 January 2007; pp. 1–9.

[13] Garcia, J.M.; Teodoro, F.; Cerdeira, R.; Coelho, R.M.; Kumar, P.; Carvalho, M.G. Developing a Methodology to Predict PM10 Concentrations in Urban Areas Using Generalized Linear Models. *Environ. Technol.* 2016, *37*, 2316–2325.

[14] Park, S.; Kim, M.; Kim, M.; Namgung, H.-G.; Kim, K.-T.; Cho, K.H.; H, K.; Kwon, S.-B. Predicting PM10 Concentration in Seoul Metropolitan Subway Stations Using Artificial Neural Network (ANN). *J. Hazard. Mater.* 2018, *341*, 75–82.

[15] Yu, R.; Yang, Y.; Yang, L.; Han, G.; Move, O.A. RAQ A Random Forest Approach for Predicting Air Quality in Urban Sensing Systems. *Sensors* 2016, *16*, 86.

[16] Yi, X.; Zhang, J.; Wang, Z.; Li, T.; Zheng, Y. Deep Distributed Fusion Network for Air Quality Prediction. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, London, UK, 19–23 August 2018; pp. 965–973.

[17] Veljanovska, K.; Dimoski, A. Air Quality Index Prediction Using Simple Machine Learning Algorithms. *Int. J. Emerg. Trends Technol. Comput. Sci.* 2018, *7*, 25–30.

[18] E. Kalapanidas and N. Avouris, "Applying machine learning techniques in air quality prediction," in Proc. ACAI, vol. 99, September 2017.

[19] V. M. Niharika and P. S. Rao, "A survey on air quality forecasting techniques,"International Journal of Computer Science and Information Technologies, vol. 5, no. 1, pp.103-107, 2014.

[20] D. J. Nowak, D. E. Crane, and J. C. Stevens, "Air pollution removal by urban trees and shrubs in the United States,"Urban Forestry &Urban Greening, vol. 4, no. 3, pp. 115-123, 2014`.