International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 10s, 2025
https://theaspd.com/index.php

An Environmental Green Approach By Optimization Of Air Quality Index (AQI) Prediction Using Hybrid Machine Learning Combines With Swarm Intelligence Algorithm

Dr. Voore Subrahmanyam[1],Dr. Mulagundla Sridevi[2], R. Ravi Kumar[3], J. Yamini Devi[4], Indraneel K[5], Nagurla Mahender[6], Vedelli Naresh[7],

[1]*Associate Professor, Department of IT, Guru Nanak Institute of Technology, Hyderabad, Telangana., Subrahmanyam.voore@gmail.com*

[2]*Associate Professor, Department of CSE, CVR College Of Engineering, Hyderabad., sreetech99@gmail.com*

[3]*Assistant Professor, Department0020of CSE, CMR College of Engineering & Technology, Hyderabad,ravikumar.racha@gmail.com*

[4]*Assistant Professor, Department of CSE, Koneru Lakshmaiah Education Foundation, Hyderabad, India,yaminidevi@kl.edu.in*

[5]*Assistant Professor, Department of IT, Guru Nanak Institute of Technology, Hyderabad, Telangana, indraneelk@gmail.com*

[6]*Assistant Professor, Department of CS&AI, SR University, Warangal,mahi.nagurla@gmail.com*

[7]*Assistant Professor, Department of IT, Guru Nanak Institute of Technology, Hyderabad, Telangana, vedellinaresh@gmail.com*

Abstract: *The pollution in the air is a one of the challenge of green environmental in urban areas that effects the natural human living. The local authorities interact to real-time monitoring and analyze the pollution data caused as per the present status of traffic situation in the city. As per this, it is essential to make appropriate precautionary measures and decisions accordingly. The Internet of Things(IoT) based sensors play a major role in dynamic of predicting air quality. In the city, due to increase of vehicle's traffic, industrialization, urbanization are cause the air contamination that effects the air pollution that effect breathing issue of human and cause health condition. Air pollution is a major problem in the traffic areas of urban. Due of vehicles quantity and releasing various hazardous gases like CO, CO2, NO2, NO, SO2, NH3, PM2.5 cause environmental pollution that affects human health problems. Addition to this modern industrialization and heavy traffic contains harmful gas molecules are being spread and contaminated in the air. These particles in the size of PM10 and PM2.5 these are seriously effect the human health. In urban everyday releasing harmful gases and cause an un-curable diseases for human. In this paper, we performed pollution prediction by applying advanced regression methods actually predict air quality in the smart city environments. An optimization of Improved Grey Wolf Optimization(IGWO) combined with Decision Tree (DT) machine learning algorithm to find accurate values of prediction values of AQI in the urban areas in India. The dataset is for experimental purpose is available with Root Mean Square Error (RMSE for real-time data of various city of India for the experimental purpose. The air pollution monitoring is an essential to predict to take necessary actions to manage the ecological balance in the city. In this paper apply regression method to predict the air pollution in the major cities. This proposed research experimental approach by verify the repository metrics via R-Square, Mean Absolute Error (MAE), Mean Squared Error(MSE), Root Mean Square Error (RMSE) metrics. The machine learning algorithms like K-nearest neighbor, Random forest regression, are comparing with the proposed algorithm. The proposed research Hybrid IGWO-DT algorithm providing the best performance achievement results when compared with existing machine learning algorithms by gaining maximize the accuracy levels of various cities in India.*

Keywords *- Internet of Things(IoT), air pollution, machine learning algorithm, regression, smart city, Root Mean Error(RMSE), Mean Square Error(MSE).*

INTRODUCTION

The quality Air is most preferrable and one of the essential component for the human to live. Many precautionary measures are taken to control pollution. Still in urban areas increasing the air pollution because of industrialization, automobiles, chemicals, vehicles traffic. All these cause growth of air pollution. The air pollution increasing everyday due to introduce various vehicles. These vehicles releasing various harmful gases are contaminated the air that impact on human health and environment. The air pollution release harmful substances into the atmosphere that impact of

allergies, illness lung diseases and human deaths. The most dangerous environmental conditions in the city by Traffic. Huge number of vehicles every day releasing harmful gases in a city traffic that cause environmental disorder and cause diseases that even not cure in hospitals. It is essential to protect the environment green and clean city for the future generations. It is our response to address the escalating problem and recognize the need for an innovative solution by real-time monitoring, advanced technologies to monitoring of various harmful gases to use methods for safeguarding and sustainability of the urban areas[1,2].Internet of Things(IoT) with its ubiquitous nature involving to solve all types of problems. In the City duty to heavy traffic cause air pollution that effect environmental conditions that impact on harmful and unconditional human health to disrupt the quality of living environment. To predict air quality and to control air pollution for creating more healthier atmosphere for sustainable environment for better living of human. This research paper provides an intelligent analysis on air pollution prediction by regression methods. Regression is a famous of supervised ML algorithm for predicting quality of air within the city premises[3].By considering all aspects to smart traffic management system we are including the monitor & control environmental pollution in this proposal. If environmental pollution is high in a traffic signaling area, automatically the lights recognize status of aerial pollution content in the atmosphere dynamically the lights starts to blink(low, medium, high) as per status of releasing the harmful gases by vehicles. If pollution is more in signal area, the pedestrians as well as vehicles are being hurry to cross the signals to avoid from pollution effects. It will cause accident affects in traffic signal area. The Cyber physical system is an intelligent system in which an algorithms are to monitor and control to solve the transport related problems into well manageable informative solutions by adapting AI technology solutions. The Air Quality Index(AQI) is takes place an efficient role in predict the pure air in the environment. AQI indicates the quality of air and identify harmful gases as per predefined ranges of air pollution. An early precautions can preventing from potential dangers of occurs the pollution in the air of urban areas[4]. Because of industrialization as well as the vehicles in the city the air becomes polluted is cause the human related respiratory health problems. The solve this the air quality index(AQI) consider in the city premises to control the pollution in urban areas. One of the promising solution to monitor and control by IoT technologies by specific IoT devices and sensor to monitor air pollution levels the city. These devices to deployed in various places where traffic junctions while vehicles are releasing heavy contaminated gases. The sensor devices sense and access the data and continuously send this real-time data to cloud for further process[5].

Air Pollution in Smart City

The smart city interacted with Information and Communication Technologies(ICT) for providing better communication and utilizing smartness in services for public health, transportation, energy utilization for effectively uses available resources. The deployment of various sensor for data collection for the purpose of pollution monitor & control, environmental monitoring, street light, traffic control, waste management, energy conservation for providing smart city applications[6].

Air Quality Index(AQI)

The AQI used to measured quality in the air. It is issued by the Ministry of Environment, Forest and Climate Change. Mainly the pollutants like PM10, PM2.5, NO2, CO, NH3, O3 and Pb. For the air to be clean, these 8 pollutants need to be controlled and if not, it will harm us and cut short our lives. The AQI is categorized into 6 categories like AQI value 0-50 remarked as “Good” and it shows colour code in “Thick Green”. AQI value 51-100 remarked as “Satisfactory” and it shows colour code in “Light Green”. AQI value 101-200 remarked as “Moderate” and it shows colour code in “Yellow”. AQI value 201-300 remarked as “Poor” and it shows colour code in “thick Brown:. AQI value 301-400 remarked as “Very Poor” and it shows colour code in “Red”. AQI value 401-500 remarked as “Severe” and it shows colour code in “Maroon”.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 10s, 2025
https://theaspd.com/index.php

Calculation of Air Quality Index(AQI)

The range of AQI from 0 to 500 will be calculated by air quality standards. As per table 0-50 shows good air, 50-100 shows irrigation symptoms, 100-150 polluted areas, AQU > 150 shows un-healthy. The AQI is calculated as follows[7].

$$AQI = \frac{I_{high} - I_{low}}{C_{high} - C_{low}} + (C - I_{low}) + I_{low} \quad (1)$$

In above equation C denotes as pollutant concentration, Clow denotes as normal pollutant concentration. C_{high} denotes as above the normal pollutant concentration. I_{low} and I_{high} denotes as index breakpoint with Clow and C_{high}.

AQI category, pollutants and health breakpoints

Table taken for the reference from Natarajan, Suresh Kumar, et al. "Optimized machine learning model for air quality index prediction in major cities in India." *Scientific Reports* 14.1 (2024): 6795.

Table 1: AQI

AQI category(range)	PM_{10} 24-hr	$PM_{2.5}$ 24 hr	NO2 24 hr	O3 8-hr	CO 8-hr (mg/m^3)	SO2 24 hr	NH3 24-hr	Pb 24 -hr
Good(0-50)	0-50	0-30	0-40	0-50	0-10	0-40	0-200	0-0.5
Satisfactory (51-100)	51-100	31-60	41-80	51-100	1.1-2.0	41-80	201-400	0.5-1.0
Moderately polluted (101-200)	101-250	61-90	81-180	101-168	2.1-10	81-380	401-800	1.1-2.0
Poor (201-300)	251-350	91-250	181-280	169-208	10-17	381-800	801-1200	2.1-3.0
Very Poor (301-400)	351-430	121-250	281-400	209-748*	17-34	801-1600	1200-1800	3.1-3.5
Severe (400-500)	430+	250+	400+	748+*	34+	1600+	1800+	3.5+

Air quality indices are calculated for every pollutant concentration and high values classify as per present locality AQI for the particular time period. The technology concern the IoT is gaining significant role for connectivity and minimize costs for sensing and monitoring the pollution data in the city. The Sensor devices gather the pollution data to measure air quality. In this paper shows the technically performing of pollution prediction using regression methods. As per evaluation criteria in this paper represents the comparative analysis of these methods. The MAE and RSME used to identify the suitable predictive model for estimation of the cause of pollution status in the areas of city.The pollution AQI data is large amount of data that require better performance algorithms. In existing algorithms for accurate prediction of air pollution including the statistical linear method used to predict air quality. Due data variations in practical approach gain poor estimation results. The inefficiency of issues in of statistical linear regression methods, the machine learning algorithms are taken place. Due to large amount of data these regression models shows poor performance. The complication can be minimized by selecting the optimality from the available dataset. These prediction techniques need required continuous learning model ultimately falls for local minima. The merging speed is showing low when comparison with SI algorithms. The swarm intelligence algorithms can solve these limitations of machine learning algorithms. In this proposed model, the optimization with regression model using combine with improved grey wolf optimization algorithm with decision tree algorithm. Using grey wolf optimization algorithm the optimal features from datasets are extracted. The next step by decision tree regression model used to already selected optimal features are finally classified. The benchmark models mean absolute error, root mean square error and mean square error are performed exact results. The comparative analysis is proposed model consider to show the performance hybrid model to find AQI.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 10s, 2025
https://theaspd.com/index.php

LITERATURE REVIEW

The various authors proposed their research for environmental pure air. The table is about existing literature and their research analysis and comparison.

Table 2: Existing literature review and research analysis and comparison

Problem statement	Technique	Strength	Limitations
Analyze air quality by ML[8]	Optimize	Minimize error rate	Data is small (available data sets). Processing time is not discussed
ML technique for classification of air quality[9]	Naïve Bayes Algorithm Decision Tree	Accurate for Decision Tree	Short data
low cost sensors to predict pollution[10]	DNN method	decreases sensor cost, prediction accuracy	datasets available too short period of 10 days.
ML techniques for AQI in Canada[11]	Neural networks with non-linear regression	its reduces error rate	data is very limited
Air Quality Index Level Prediction using Random Forest[12]	Random forest	Perform experiments for train datasets on the basis of clusters. Experiment found more efficient and accurate	random forest will perform only on classification

PROPOSED ARCHITECTURESensing Layer – This layer includes sensor devices, IoT based communication devices, pollution monitor sensor in the city, transport road and industry.[20,21,22].

Network Communication Layer - This network layer is responsible for data collection, and transferring to other layers. In ths layer network communication devices will be present including 5G,Wi-Fi, ZigBee, Zwave, Routers etc. IoT sensor devices accessing the pollution data and send the data to further process to cloud.

Data Management Layer- The bottom layer gather data from heterogenous sources in urban environment. The various air pollutants like sulphur dioxide, nitrogen dioxide, ozone etc are monitor by sensor devices at various release in this city. The heavy real-time data will be gathered and aggregation is takes place. The pre-processing is taken place when unwanted data will be removed. The dataset taken to consider for this proposed model from the Central Pollution Control Board(CPCB), India. The collected dataset includes air quality index (AQI) data from different corner of cities in India. The selected cities are Hyderabad, Calcutta, Delhi etc. Air Pollution level in Hyderabad location PM2.5, PM10, NO, NO2, NOx, NH3, CO, SO2, O3. The AQI is categorize into good, satisfactory, moderate, poor, very poor. In pre-processing step, the datasets contain null values. Premier the first step is data cleaning process, from dataset the null values are separate the remain datasets are used for experimental purpose. The datasets are imbalanced after removal of null values. To balance these data synthetic minority oversampling (SMOTE) apply for these datasets SMOTE process the un-necessary data add to extra data. SMOTE provides better balancing by generating synthetic data points and does not provide duplicate data.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 10s, 2025
https://theaspd.com/index.php

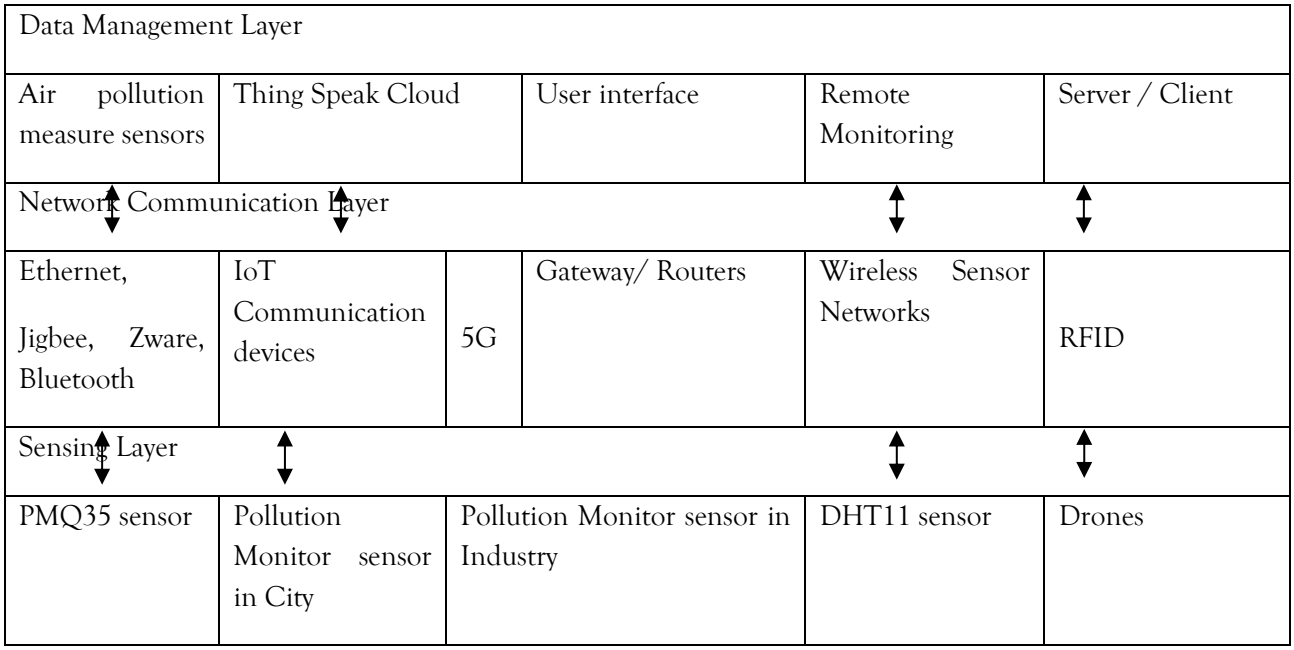

Figure1: Layered architecture for Pollution prediction in City

In this proposed model after the completion of data balancing, next step by optimize with nature inspired algorithm i.e. improved grey wolf optimization to select optimal features of dataset. GWO algorithm features are overcome effectively local search and explore global search features in available search space. As per improved grey wolf optimization algorithm providing solutions, the optimal features extracted from datasets to predict the air quality. The regression technique apply in proposed algorithm i.e. decision tree used to predict qualitative as well as quantitative variables. Using decision tree the data response can be predicted. The decision tree having tree like structures that provides classification as well as regression by direct and indirect relations between independent and target variables. The already predicted values are taken to consider an optimum solutions. The decision tree regression algorithm used to predict the air quality. In this proposed model consider regression method to obtain better performance results for evaluating region of interest. In figure shows step-by-step for the final prediction model by improved GWO combine with DT regression algorithm. In premier step performs the data clearing to remove the null value from dataset. Next step, the optimal features are extracted by grey wolf optimization algorithm. Next step, to predict AQI, the extracted optimal features are analyzed by decision tree. Finally, the test measure by metrics MSE, MAE, RMSE etc.

Figure 2: Proposed prediction model using Improved Gray Wolf Optimization algorithm combination with Decision Tree Regression model

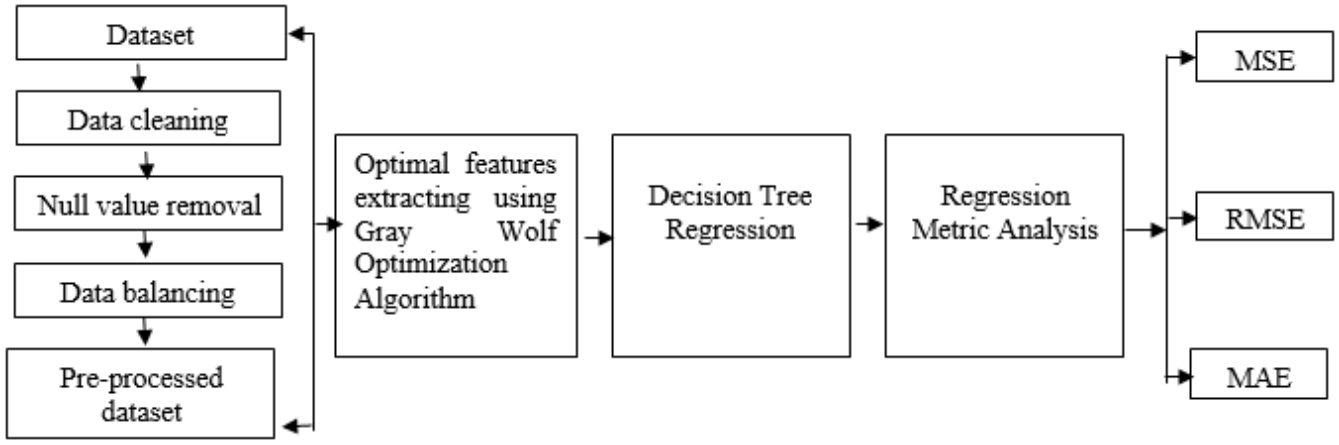

Table 3: Air Pollution level in Hyderabad location

LOCATIONS	Status	AQI-US	PM2.5	PM10	SO2	NO2	CO	Temp.	Humidity
Zoo Park	GOOD	50	12	14	17	10	16	29	69
Vittal Rao Nagar	GOOD	46	11	34	10	7	11	30	53
Somajiguda	MODERATE	74	23	62	15	11	17	29	64
Shirdi Sai Nagar	GOOD	38	9	26	16	10	15	30	54
Saidabad	MODERATE	55	14	37	19	22	42	30	55
Puppalguda	GOOD	42	10	30	14	8	11	30	54
New Malakpet	MODERATE	68	20	76	13	15	20	29	64
Manikonda	GOOD	50	12	36	17	18	25	30	53
Madhapur Village	GOOD	42	10	31	18	17	26	25	74
Kphb Phase Iii	MODERATE	53	13	30	19	16	29	25	74
Koti	MODERATE	55	14	42	15	10	21	25	74
Kokapet	MODERATE	57	15	41	14	11	18	28	71
Hyderabad Us Consulate	MODERATE	66	19	47	12	15	19	29	64
Central University	MODERATE	55	14	48	9	13	21	29	68
Banjara Hill	MODERATE	63	18	45	7	12	22	25	74

Table 4: Air Pollution level in New Delhi

LOCATIONS	Status	AQI-US	PM2.5	PM10	SO2	NO2	CO	Temp.	Humidity
Ashok Vihar Phase 1	POOR	168	88	185	15	11	14	33	34
Civil Lines	POOR	157	68	120	9	8	12	33	34
Defense Colony	GOOD	49	11	12	14	12	18	28	67

Delhi Cantt	GOOD	44	10	32	15	12	15	27	65
Delhi Institute Of Tool Engineering	POOR	176	104	242	18	21	19	33	35
Dwarka Sector 11	POOR	156	65	214	13	9	10	33	34
Dwarka Sector 12	POOR	154	62	209	15	16	18	33	34
Janakpuri	POOR	166	85	233	20	21	27	33	34
Jawaharlal Nehru Stadium	GOOD	48	10	11	17	18	10	27	66
Kashmiri Gate ISBT	POOR	155	64	110	18	17	30	33	34
New Delhi Us Embassy	POOR	165	83	148	16	17	18	30	61
New Friends Colony	GOOD	50	12	11	14	11	19	27	65
RK Puram	POOR	167	87	207	13	16	18	30	61
RK Puram North Block	POOR	161	74	170	9	14	20	33	34
Saket	MODERATE	63	18	45	9	10	28	25	74
Saket Block C	MODERATE	63	18	45	15	10	30	25	74

The air pollutants like PM2.5, PM10, SO2,NO2, CO, TEMP, HUM etc are effects the pollution levels in the air. The essential objective of these cities to deploy the sensor devices to monitor the status of air pollution in a particular areas where vehicles are jammed in a circles.. The sensors are also failure sometimes to predict these dataset. To solve these problems to monitor the quality of air is a promising task. The research objective to data analysis by various ML combine with Swarm Intelligence for forecasting the particular matter(PM2.5).[13,14,15]. The research to predict concentration of PM2.5 in the atmospheric air regression models were evaluated in the form of statistical measures like Root Mean Error(RMSE), Mean Square Error(MSE)[16,17,18,19].

PROPOSED MODEL

The simulation analysis consider benchmark collect Air Quality Datasets from Central Control Room for Air Quality Management New Delhi. The major cities are selected as per the AQI given datasets range from good, moderate, poor, satisfactory and very poor. Initially the optimal features are extracted using Improved Gray Wolf Optimization algorithm after than classified using Decision tree regression model. To evaluate the performance levels of pollution status and prediction by applying R-Square, MAE and RMSE. All 5 ML regression algorithm to predict AQI. It is not classification problem The target variable is AQI which continuous and hence regression analysis.

R-Square

$$R - Square = SS_{regr} / SS_{tt} \quad (2)$$

In above equation SS_{regr} denotes the model for regression sum of squares SS_{tt} denotes sum of squares.

Mean Square Error (MSE).

This metrics used to measure how results are closely to the data points. The MSE is minimum. Whether MSE noted as equivalent to 0 is consider as real model. The MSE equation as follows.

$$MSE = \sum_{i=1}^{n} \frac{(x_i - \widehat{x_i})^2}{n} \quad (3)$$

In above equation x_i is denote as observed values. x_i denote as predicted values. n denote as number of observations.

Mean Absolute Error(MAE)

It is apply to measure to find errors to predict the data. MAE provide an accuracy of differentiate the actual data and prediction data values. The MAE is performance by following formula[23].

$$MAE = \frac{1}{n}\sum_{j=1}^{n}(y_j - \acute{y_J}) \quad (4)$$

In above equation n denote as no. of observations, y_j denote as actual value, $\grave{y_J}$ denote as predicted value.

Root Mean Squared Error(RMSE)

It is square root of the average of squared error being difference between the actual target and models prediction. Low the RMSE better the model performance. The model the RMSE zero then it's a perfect. RMSE is highly sensitive metrics. If the data does not contain any outliers and if our model is not expected to deal with these outliers, we can apply RMSE. The equation is...

$$RMSE = \sqrt{\frac{1}{n}(\sum_{j=1}^{n}(y_j - \acute{y_J})^2)} \quad (5)$$

Normalize the RMSE for comparing the datasets to apply this equation..

$$Normalized\ RMSE = \frac{RMSE}{y_{max} - y_{min}} \quad (6)$$

In above, Y_{max} = data set max. value, Y_{min} = dataset min. value

Accuracy:

The accuracy metric used to identifies the relations in the dataset. It is measured by mean absolute error formulation by following equation.

Accuracy = (1 −MAE)∗100 (7)

RESULSTS AND DISCUSSION

The proposed research model include various steps like data cleaning, removal of null values and balance of data etc. In this data cleaning process, at first the xylene was removed from data. Next process the blanks in other attributes are removed. For this process the datasets with all values of attributes available. The dataset AQI will be categorized into various classes like poor, very poor, Good, satisfactory and moderate. In this, the removal of null values from dataset then it become imbalanced. So these dataset is balanced by applying SMOTE i.e. Synthetic minority

oversampling technique apply for various major cities in India. In figure the balanced and un-balanced dataset shows accuracy analysis of the proposed research. It shows the accuracy of balanced dataset over un-balanced dataset. The accuracy availing of datasets of various cities New Delhi is 84.65% of imbalanced whereas balanced gain 90.10%. The accuracy availing of datasets of city Bangalore is 71.23% of imbalanced whereas balanced gain 92.50%, The accuracy availing of datasets of city Kolkata is 92.65% of imbalanced whereas balanced gain 96.68%, The accuracy availing of datasets of city Hyderabad is 92.68% of imbalanced whereas balanced gain 98.88%, The accuracy availing of datasets of city Chennai is 80.49% of imbalanced whereas balanced gain 97.44%, The accuracy availing of datasets of city Visakhapatnam is 92.84% of imbalanced whereas balanced gain 98.75%,

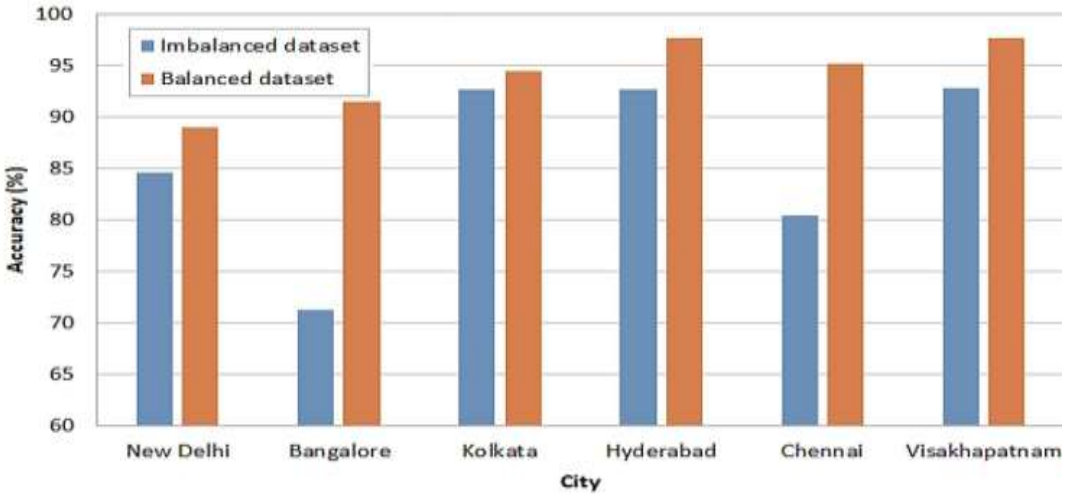

Figure 3: Accuray analysis

COMPARATIVE ANALYSIS

The comparative analysis comparison results of existing model with proposed model as follows. The existing models like random forest regression, support vector regression, k-nearest neighbor model etc. Obtaining results has been consider for summarization the proposed model Hybrid Random Forest combined with Improved Gray Wolf Optimization Algorithm perform high accuracy results comparisons with existing regression model. The comparison performance results shown in Table form. As per the feature selection process, the regression technique i.e. Decision Tree avail the maximize performance of prediction process and obtain results when compare with existing algorithms. As per results shown in figure it is observed that the proposed hybrid optimization algorithm to optimize with regression model gain high performance results compare with existing repression models.

Table 5: Proposed algorithm results comparison with existing ML Algorithms

S.No.	City	Accuracy				
		Support vector regression	K-nearest neighbor	Random forest regression	IGWO-DT	Proposed IGWO-DT
1	New Delhi	84.83	83.68	84.73	88.98	90.75
2	Bangalore	87.18	89.47	90.31	91.49	93.50
3	Kolkata	91.56	90.65	93.74	94.48	96.25
4	Hyderabad	93.57	93.68	97.61	97.66	98.75
5	Chennai	92.65	93.48	94.48	95.22	96.35
6	Visakhapatnam	92.24	92.11	95.65	97.68	98.45

CONCLUSION

The proposed research to optimize to predict the AQI by hybrid IGWO-DT algorithm for major cities in India. In this proposed model we consider the Decision tree regression model and Grey Wolf Optimization algorithm to find optimal features for prediction process. The data sets is used for this proposed model to evaluate various cities using testing by performance metrics like MSE, MAE, RMSE are applied better way. Comparison with existing regression methods taken to consider for this proposed research are k-nearest neighbor, support vector machine and random forest models, the proposed Hybrid IGWO-DT shows high performance results of New Delhi gain accuracy of 90.75%, New Bangalore gain accuracy of 93.50%, New Kolkata gain accuracy of 96.25%, New Hyderabad gain accuracy of 98.75%, New Chennai gain accuracy of 96.35%, New Visakhapatnam gain accuracy of 98.45%, All these experiments it is confirmed for better performance of proposed algorithm comparison with existing methods. In future the proposed model should extend by applying Hybrid Artificial Intelligence with recent Swarm Intelligence algorithms to acquire more better performance results for monitoring air quality.

ACKNOWLEDGEMENT

My sincere gratitude to our Institute Principal and Management for their kind support for providing computer centre for completion of the experimental work of this paper. I am very thankful my friends and colleagues for supporting this research paper.

AUTHOR CONTRIBUTIONS

Dr. V. Subrahmanyam is corresponding author of the research paper. His contribution of this paper includes Literature Survey, implementation, optimization of algorithm and results. Other co-authors Dr. Mulagundla Sridevi, R. Ravi Kumar contribution is to verify the results, J. Yamini Devi, Indraneel K contribution is to data gather, Nagurla Mahender and Vedelli Naresh verify the available resources.

.REFERENCES

[1] Liang, Lu, and Peng Gong. "Urban and air pollution: a multi-city study of long-term effects of urban landscape patterns on air quality trends." Scientific reports 10.1 (2020): 18618.

[2] Wang, Yiyi, et al. "High-resolution modeling for criteria air pollutants and the associated air quality index in a metropolitan city." Environment International 172 (2023): 107752.

[3] Yu, Jingze, et al. "Association between air pollution and cardiovascular disease hospitalizations in Lanzhou City, 2013–2020: A time series analysis." GeoHealth 8.1 (2024): e2022GH000780.

[4] Debbarma, Sohana, et al. "Impact of real-world traffic and super-emitters on vehicular emissions under inter-city driving conditions in Maharashtra, India." Atmospheric Pollution Research 15.4 (2024): 102058.

[5] Saikiran, Karlapudi, et al. "Prediction of Air Quality Index Using Supervised Machine Learning Algorithms." 2021 2nd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS). IEEE, 2021.

[6] Srivastava, Chavi, Shyamli Singh, and Amit Prakash Singh. "Estimation of air pollution in Delhi using machine learning techniques." 2018 International Conference on Computing, Power and Communication Technologies (GUCON). IEEE, 2018.

[7] Ameer, Saba, et al. "Comparative analysis of machine learning techniques for predicting air quality in smart cities." IEEE access 7 (2019): 128325-128338.

[8] K.-H. Kim, E. Kabir, and S. Kabir, "A review on the human health impact of airborne particulate matter," Environ. Int., vol. 74, pp. 136_143, Jan. 2015.

[9] 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," Environ. Epidemiol., vol. 2, no. 2, p. e014, 2018.

[10] D. Zhu, C. Cai, T. Yang, and X. Zhou, ``A machine learning approach for air quality prediction: Model regularization and optimization," Big Data Cogn. Comput., vol. 2, no. 1, p. 5, 2018.

[11] N. Zimmerman, A. A. Presto, S. P. N. Kumar, J. Gu, A. Hauryliuk, E. S. Robinson, A. L. Robinson, and R. Subramanian, ``Closing the gap on lower cost air quality monitoring: Machine learning calibration models to improve low-cost sensor performance," Atmos. Meas. Tech. Discuss, vol. 2017, pp. 1_36, 2017. doi: 10.5194/amt-2017-260.

[12] O. A. Ghoneim, H. Doreswamy, and B. R. Manjunatha, ``Forecasting of ozone concentration in city using deep learning," in Proc. Int. Conf.Adv. Comput., Commun. Inform. (ICACCI), 2017, pp. 1320_1326.

[13] Mohammadi, Farzaneh, et al. "Prediction of atmospheric PM2. 5 level by machine learning techniques in Isfahan, Iran." Scientific Reports 14.1 (2024): 2109.

[14] Ameer, Saba, et al. "Comparative analysis of machine learning techniques for predicting air quality in smart cities." IEEE Access 7 (2019): 128325-128338.

[15] Liu, Huixiang, et al. "Air quality index and air pollutant concentration prediction based on machine learning algorithms." Applied Sciences 9.19 (2019): 4069.

[16] Bhalgat, Pooja, Sejal Pitale, and Sachin Bhoite. "Air quality prediction using machine learning algorithms." International Journal of Computer Applications Technology and Research 8.9 (2019): 367-370.

[17] Harishkumar, K. S., K. M. Yogesh, and Ibrahim Gad. "Forecasting air pollution particulate matter (PM2. 5) using machine learning regression models." Procedia Computer Science 171 (2020): 2057-2066.

[18] Maheshwari, Khushi, and Sampada Lamba. "Air quality prediction using supervised regression model." 2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT). Vol. 1. IEEE, 2019.

[19] Aram, S. A., et al. "Machine learning-based prediction of air quality index and air quality grade: A comparative analysis." International Journal of Environmental Science and Technology 21.2 (2024): 1345-1360.

[20] Martín-Baos, José Ángel, et al. "IoT based monitoring of air quality and traffic using regression analysis." Applied Soft Computing 115 (2022): 108282.

[21] Zhu, Dixian, et al. "A machine learning approach for air quality prediction: Model regularization and optimization." Big data and cognitive computing 2.1 (2018): 5.

[22] Van, N. H., et al. "A new model of air quality prediction using lightweight machine learning." International Journal of Environmental Science and Technology 20.3 (2023): 2983-2994.

[23] Baba Ameer et.al. “Comparative Analysis of Machine Learning Techniques for Predicting Air Quality in Smart Cities” “IEE Xplore, publication June 26, 2019,

[24] Murugan, Rishanti, and Naveen Palanichamy. "Smart city air quality prediction using machine learning." 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2021.

[25] Wang, An, et al. "Potential of machine learning for prediction of traffic related air pollution." Transportation Research Part D: Transport and Environment 88 (2020): 102599.

[26] Rahimpour, Alireza, Jamil Amanollahi, and Chris G. Tzanis. "Air quality data series estimation based on machine learning approaches for urban environments." Air Quality, Atmosphere & Health 14 (2021): 191-201.

[27] Zimmerman, Naomi, et al. "A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring." Atmospheric Measurement Techniques 11.1 (2018): 291-313.