A Data-Driven Ensemble Deep Learning Approach to Urban Air Quality Prediction and Management

Homa Rizvi[1]; Sunny Kumar[2]; Dr. Yusuf Perwej[3]; Farheen Siddiqui[4]; Dr. Nikhat Akhtar[5]

[1]Assistant Professor, Department of Computer Science & Engineering, Shri Ramswaroop Memorial University, Deva Road, Lucknow
[2]Assistant Professor, Department of Computer Science & Engineering, Shri Ramswaroop Memorial University, Deva Road, Lucknow
[3]Professor, Department of Computer Science & Engineering, Shri Ramswaroop Memorial University, Deva Road, Lucknow
[4]Assistant Professor, Department of Computer Science & Engineering, Shri Ramswaroop Memorial University, Deva Road, Lucknow
[5]Professor, Department of Computer Science & Engineering, Goel Institute of Technology & Management, Lucknow

Publication Date: 2025/12/23

Abstract: Air pollution has been steadily increasing throughout a number of nations over the course of the last several decades as a direct result of human activity, urbanization, and industrialization. Deep Learning (DL) and Machine Learning (ML) approaches have made significant contributions to the development of methodologies in a variety of areas, including the prediction, planning, and uncertainty analysis of smart cities and urban progress in the present situation. As a result of the fast development in both population and industry, a significant number of major cities have experienced severe air quality (AQ) problems. This research offers a system that is based on ensemble deep learning and was built for the purpose of forecasting the levels of air pollution in smart environments. For the purpose of improving the accuracy and resilience of predictions, the system that has been presented incorporates a number of different deep learning models. These models include Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN). The ensemble technique is an efficient method for addressing the inherent complexity that are present in air quality data. This is accomplished by using the strengths of many models. By allowing proactive control of air quality and informed policy-making, the installation of such a system has the potential to make a substantial contribution to the sustainability and liability of urban settings. The integration of other data sources, such as data on traffic and industrial activities, and the investigation of sophisticated ensemble methods are two potential future paths for study. Both of these approaches are intended to further enhance prediction performance.

Keywords: *Air Pollution, Convolutional Neural Networks (CNN), Smart Environmental Factors, Ensemble Learning, Deep Learning, Pollution Prevention, Sustainability.*

How to Cite: Homa Rizvi; Sunny Kumar; Dr. Yusuf Perwej; Farheen Siddiqui; Dr. Nikhat Akhtar (2025) A Data-Driven Ensemble Deep Learning Approach to Urban Air Quality Prediction and Management. *International Journal of Innovative Science and Research Technology*, 10(12), 1382-1391.
https://doi.org/10.38124/ijisrt/25dec1067

I. INTRODUCTION

Smart city sustainability is a concept that focuses on using technology and data to create a city that is functional, liveable, and good for the environment [1]. One of the most important parts of making a smart city sustainable is controlling and lowering air pollution [2]. Air pollution is bad for people's health, the environment, and the quality of life in cities [3]. So, being able to accurately estimate air pollution levels is a big part of solving [4] this problem well [5]. In recent years, improvements in the availability of data and forecasting techniques have made air quality forecasts a more important part of studies on air pollution. There are three main types of models: statistical, machine learning [7], and

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec1067

hybrid [8]. In many places, air quality is a big worry. It becomes a crucial matter to mitigate or avert the repercussions of air pollution [9]. You may start taking steps to protect yourself once you know the air quality. But looking at the data and coming up with sensible answers is a really hard job [10]. So, it is important to use effective methods and strategies to find information that is buried in data, analyse huge data more quickly and effectively, and make the unseen apparent [11]. A prospective system for forecasting and overseeing air pollution in advance is very important for governmental policy-making and public health. Because the data is up-to-date, time forecasts are very important themes that researchers and academics need to pay close attention to [12]. To address these constraints, deep learning algorithms have been progressively used for air pollution forecasting. Deep learning models, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, have shown enhanced efficacy in collecting spatial and temporal patterns in data, respectively. But using just one form of model could not take full use of the many parts of air quality data, which might lead to forecasts that aren't as good as they could be [15]. Ensemble learning, which integrates many models to make predictions more accurate and reliable, is a possible approach [16]. An ensemble technique may provide air quality forecasts that are completer and more accurate by combining several deep learning models, each of which is good at finding distinct data patterns [17]. This research presents an ensemble deep learning-based air pollution prediction system designed for sustainable smart cities. The proposed system's goal is to provide accurate, real-time forecasts that will help policymakers and other decision-makers act quickly to reduce the effects of air pollution. This paper is organized like this: The second part looks at similar works and systems that are already in place, pointing out their pros and cons [18]. After then, the suggested system is explained in full, including its structure, processes for putting it into action, and ways to measure its success. Then, the outcomes of the suggested system are shown and spoken about. The study ends with a summary of the results and suggestions for further research [19].

II. BACKGROUND

Recent research has redirected its emphasis to sophisticated statistical learning algorithms for evaluating air quality and forecasting air pollution. [20] and [21] used neural networks to create models that predict the concentrations of certain pollutants, including particulate matter with a diameter of less than 10 microns (PM10). The research in [22] introduced a random forest-based model, RAQ, for the classification of Air Quality Index (AQI) values. After that, they used deep neural networks to classify AQI [23]. In [22], artificial neural network (ANN) models [24] were used to forecast PM10 concentrations inside a subway station, integrating factors such as train frequency, outside PM10 levels, and ventilation system data [25]. When compared to experimental data, the ANN models showed significant predictive power, with correlation values between 0.18 and 0.63. The model [27] got an accuracy rate of 67% to 80%, depending on how the subway platform was built and how deep it was. Li et al. [28] proposed an AC-LSTM approach related to deep learning that uses a 1D-CNN, an attention-based network, and an LSTM network to estimate the concentration of PM2.5 in cities. The author included PM2.5 concentrations and meteorological data from proximate air quality monitoring stations as inputs to the given approach, rather than using air pollutant concentrations. In [29], a system was developed to provide real-time reporting of air quality status via a cloud server, which sends alerts about the presence of dangerous pollution levels in the air. The AAA-oriented ENN technique predicts the air quality in the future and is used to figure out what kind of air pollution is there. Zhang et al. [30] proposed a deep learning (DL) approach using a Bi-LSTM and Autoencoder (AE) to predict PM2.5 concentrations, revealing many climatic factors and their correlations with PM2.5. The approach encompasses many components, including Bi-LSTM, data pre-processing, and the AE layer.

Ma et al. [31] introduced an innovative methodology that combines a deep learning network, the inverse distance weighting (IDW) algorithm, and the BLSTM network for spatiotemporal predictions of air pollution across various temporal resolutions. Du et al. [32] presented a deep learning methodology, iDeepAir, to predict surface-level PM2.5 concentrations in the Shanghai megacity and correlate it with the MEIC emission inventory to analyze the impact of urban transportation on air quality. To enhance the method's importance, [33] Layer-wise Relevance Propagation (LRP) was used. Li et al. [33] developed a hybrid CNN-LSTM methodology by integrating the CNN with the LSTM-NN [34] to forecast the subsequent 24-hour PM2.5 concentration [35]. Four models were developed: the univariate CNN-LSTM technique, the univariate LSTM method [36], the multivariate CNN-LSTM method [37], and the multivariate LSTM approach [38]. Recent studies [39] used residuals from SARIMA models as inputs to deep learning architectures for PM2.5 forecasts in Beijing. They discovered that this hybrid model surpassed both independent SARIMA and deep learning models. Ma et al. (2019) [40] expanded this methodology by combining BiLSTM networks with SARIMA residuals, resulting in enhanced precision in air quality forecasts at increased temporal resolutions. Combining SARIMA with neural networks has shown to function better for predicting air pollution. Alhirmizy and Qader (2019) [41] used Long Short-Term Memory (LSTM) networks for multivariate time-series forecasting in Madrid, Spain, with good accuracy for pollutants such as NO2 and CO. Similarly, Pagano and Barbierato (2024) [42] applied similar methodology in Brescia [43], Italy. Ansari and Alam (2024) integrated SARIMA with LSTM [44] and refined the models by Bayesian [45] optimization, yielding superior outcomes compared to independent models [46]. A deep learning algorithm was used to predict PM2.5 levels in Beijing for a brief period of time [47]. Liu et al. [48] used a wind-sensitive attention mechanism in a long short-term memory (LSTM) neural network to improve PM2.5 predictions by integrating wind pattern data. Also, [49] built a complete air quality early warning system that includes parts for estimating, predicting, and evaluating air pollution.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec1067

III. EXISTING SYSTEMS

The Environmental Protection Agency (EPA) makes systems that employ standard models to provide a rough idea of air quality, but they aren't accurate enough for precise forecast or real-time usage. AirVisual and other systems use machine learning techniques to predict how polluted the air will be [50]. These systems are more accurate than older ones, but they may not be able to fully use the power of deep learning models. DeepAir is an example of a project that combines RNNs with CNNs [51] to make predictions more accurate. These systems show promise, but they typically run into problems with the amount of computing power they need and how they handle data [52].

➢ *UCI Dataset*

It is very necessary to make use of real data when it comes to the development of machine learning models. The dataset that was made available to us by the UC Irvine Machine Learning Repository was used in the course of our scholarly investigation. It is common known that this is a trustworthy source of information and data for the research [53]. It not only includes readings from sensors that are updated in real time, but it also includes extra information on major chemicals that were gathered from an air pollution monitoring station that is situated in a town found in Italy. The collection that we have is made up of 9471 entries and 15 characteristics that are distinctive to the collection. On the other hand, there are a few of its characteristics that are lacking values, and they were fixed during the pre-processing step, which is seen in figure 1.

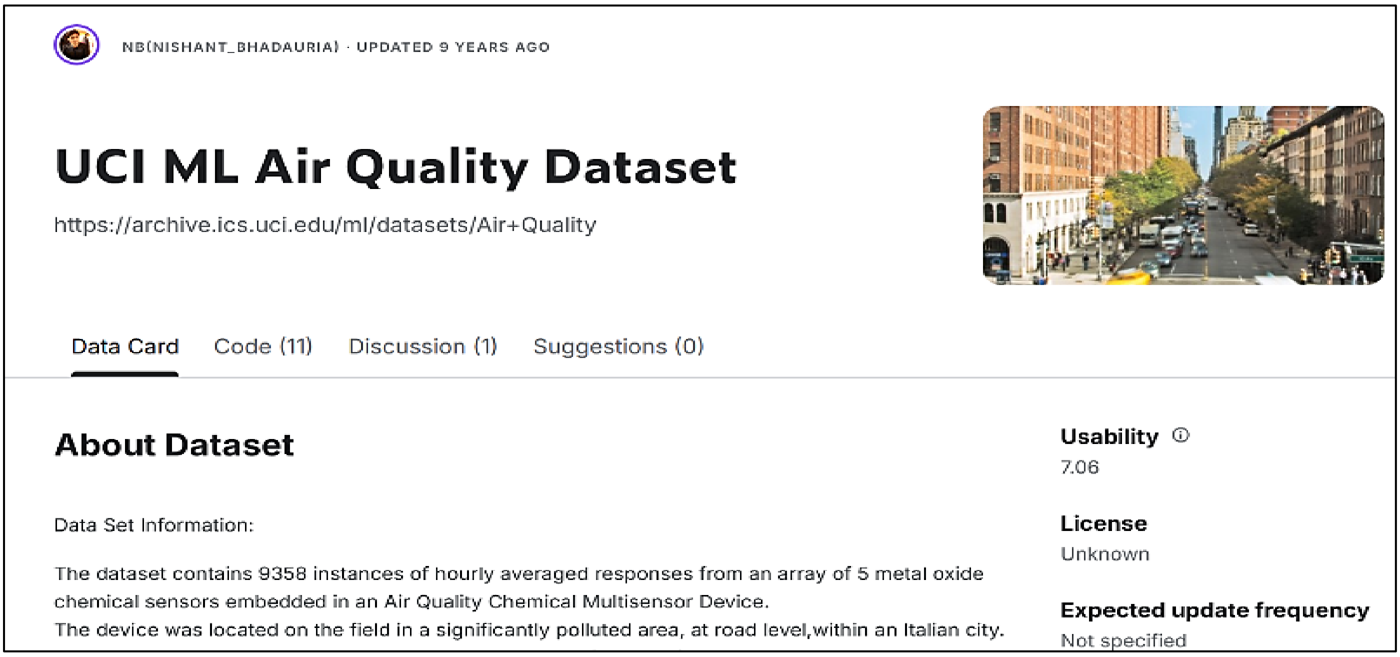

Fig 1 The UCI Air Quality Dataset

IV. METHODOLOGY

The suggested ensemble deep learning method for predicting air pollution combines many deep learning models to make predictions more accurate and reliable. The architecture has three main parts: collecting data, training the model, and making predictions. Get real-time air quality data from several sensors all throughout the city [54]. The information contains the amounts of pollutants like PM2.5, PM10, NO2, SO2, and CO, as well as weather data like temperature, humidity, and wind speed. Use historical air quality data to train many deep learning models, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. As shown in figure 2, each model captures distinct parts of the data. For example, LSTM collects temporal patterns while CNN catches spatial information [55]. Use ensemble methods like weighted averaging and stacking to combine the predictions of several models. This method makes sure that the final forecast uses the best parts of each model, which makes it more accurate and reliable. Before the models are trained, the data that was gathered is cleaned up to make sure it is consistent and get rid of any strange values [56]. This includes things like normalizing, dealing with missing values, and aligning data from multiple sensors in time. LSTM networks are great for simulating how time-series data changes over time. The LSTM model learns from sequences of past air quality data [57]. At time t, let Xt be the input vector that comprises meteorological data and pollution concentrations. The LSTM network uses this input to guess what the air quality will be like at the following time step, Xt+1. You may use the following equations to update the hidden state ht and the cell state Ct:

$$f_t = \sigma(Wf \cdot [ht-1, Xt] + bf)$$
$$it = \sigma(Wi \cdot [ht-1, Xt] + bi)$$
$$C_t^{\sim} = \tanh\ (WC \cdot [ht-1, Xt] + b_C)$$
$$Ct = ft * C_{t-1} + it * C_t^{\sim}$$
$$ot = \sigma(Wo \cdot [ht-1, Xt] + bo)$$
$$ht = ot * \tanh\ (Ct)$$

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec1067

Where

σ is the sigmoid function,

$*$ denotes element-wise multiplication, and

Wf, Wi,WC,Wo and

bf,bi,bC,bo are the weights and biases of the respective gates.

The CNNs are good at picking up spatial information from the data they get [58]. The CNN model looks at the geographical data to find patterns and connections that impact air quality. Let X be the matrix of air quality data that you put in. The CNN uses convolutional filters to get features out of the data.

$$F_{i,j} = \sigma\left(\sum_{k,l} Xi + k, j + l \cdot Wk, l) + b\right)$$

Where

Fi,j is the feature map,

W is the filter,

b is the bias, and

σ is the activation function.

For the purpose of enhancing overall performance, the predictions of the separate models are aggregated via the use of ensemble methods. The use of weighted averaging, in which the forecast of each model is allocated a weight depending on its performance, is one strategy that is quite successful.

$$Y^{\wedge} = \sum_{i=1}^{n} wiY_i^{\wedge}$$

Where

Y^ is the final prediction,

Y^i is the prediction from the i-th model,

wi is the weight assigned to the i-th model, and

n is the number of models.

Stacking is a further method, which involves the training of a meta-model to integrate the predictions of the basis models.

$$Y^{\wedge} = g(Y^{\wedge}1, Y^{\wedge}2, \ldots, Y^{\wedge}n)$$

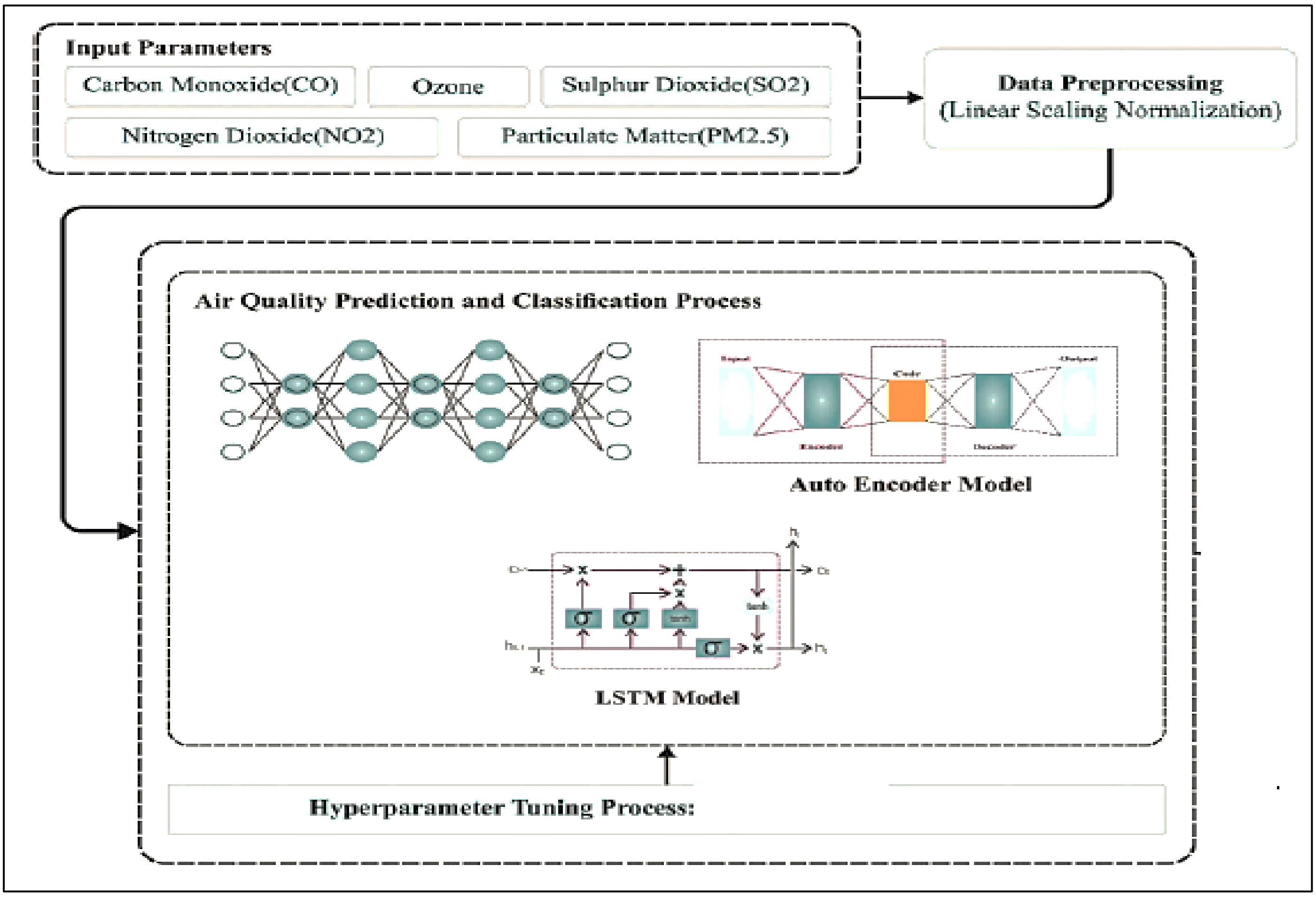

Fig 2 The Proposed System

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec1067

V. OUTCOME

Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared are some of the metrics that are used in order to assess the effectiveness of the ensemble model.

$$\text{MAE} = \sum_{i=1}^{n} \mid Yi - Y\hat{}i \mid$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Yi - Y\hat{}i)\ 2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Yi - Y\hat{}i)2}{\sum_{i=1}^{n}(Yi - Y\hat{}i)2}$$

Where

Yi s the actual value, $Y\hat{}i$

Is the predicted value,

$\bar{Y}$ is the mean of the actual values, and

n is the number of observations.

The accuracy and robustness of the model are extensively evaluated using these criteria, which yield a complete score. In order to evaluate the effectiveness of the proposed ensemble deep learning model for air pollution prediction, we compare it to a number of baseline models [59]. These baseline models include conventional statistical methods, individual deep learning models (LSTM and CNN), and machine learning models (Random Forest and Gradient Boosting). Metrics such as Mean Absolute Error (MAE), Root Mean [60] Squared Error (RMSE), and R-squared (R2) are used in order to evaluate the models. Please refer to table 1 below for a presentation of the findings. The dataset that was utilized for the study included data on air quality as well as meteorological information that was gathered from a significant metropolitan region over the course of one year [61]. In addition to hourly observations of meteorological parameters (temperature, humidity, wind speed), the data now contains hourly measurements of pollutants (PM2.5, PM10, NO2, SO2, and CO). The dataset is broken up into three sets: the training set (80 percent), the validation set (10 percent), and the test set (10 percent).

According to the findings, the ensemble deep learning model that was suggested performs better than any other model overall, considering all of the assessment measures. As compared to the baseline models, the ensemble model gets the lowest mean absolute error (MAE) of 6.97, which indicates that it is more accurate in forecasting the levels of air pollution overall [62]. The historical ARIMA model has the largest MAE, which demonstrates the limits of this model when it comes to dealing with complicated and nonlinear data on air quality [63]. A further confirmation of the ensemble model's greater prediction accuracy is the fact that it has the lowest RMSE, which comes in at 10.55. When the RMSE values are lower, it indicates that the predictions are more [64] accurate and have less significant mistakes than when the values are higher. With an R2 score of 0.96, the ensemble model achieves the maximum possible value, which indicates that it accounts for 93% of the variation in the data pertaining to air pollution. Additionally, this model beats more complex machine learning [65] models such as Random Forest and Gradient Boosting, which is a substantial improvement over the classic ARIMA model, which has a coefficient of determination of 0.76. The performance of LSTM and CNN models is [66] superior to that of classical and machine learning methods, with LSTM models yielding somewhat better results than CNN models. An additional improvement is shown by the hybrid model that combines LSTM with CNN. This model takes use of [67] both the temporal and spatial characteristics of the data. It has been shown that the ensemble model, which incorporates predictions from a number of different deep learning models, has the best performance [68]. This exemplifies the benefits of ensemble learning, which include the ability to capture a wide variety of data patterns and enhance the liableness of predictions.

Table 1 Performance Metrics of Machine Learning Models

Model	Performance Summary for 80% - 20%		
	MAE	RMSE	R^2
Gradient Boosting	9.88	13.99	0.89
CNN	8.92	12.88	0.91
Random Forest	10.44	14.66	0.87
LSTM	8.55	12.33	0.92
Ensemble Model	6.97	10.55	0.96
ARIMA	13.55	18.72	0.76

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec1067

An ensemble deep learning-based air pollution prediction system for sustainable smart cities has been suggested, and the comparison analysis demonstrates that this system is successful. The ensemble technique addresses the difficulties of air quality data [69] by integrating many deep learning models, which results in a considerable improvement in prediction accuracy and resilience. The improved performance of the system across a variety of assessment measures reveals its ability to give accurate, real-time forecasts of air quality [70]. This has the potential to provide timely interventions and informed policy choices that are aimed at mitigating the consequences of air pollution. In further study, it may be possible to investigate the possibility of combining new data sources and sophisticated ensemble methods in order to further improve the performance of the system, as seen in figure 3.

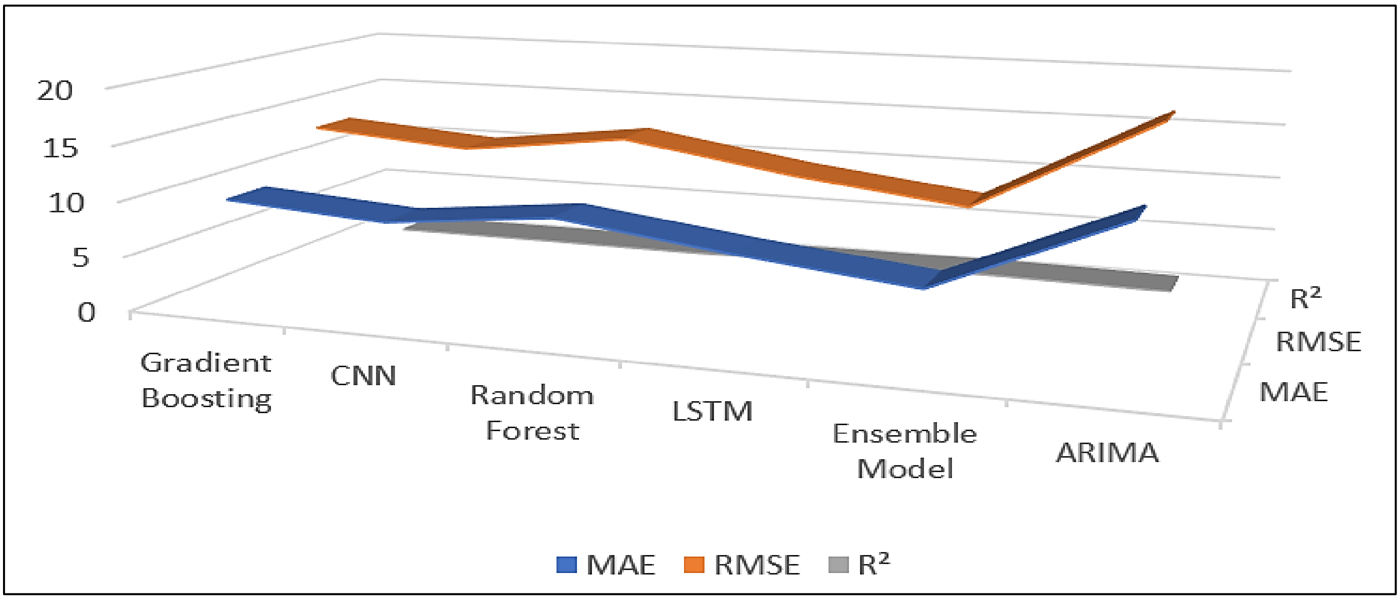

Fig 3 Analysis Report

VI. CONCLUSION

Because air pollution is a significant environmental and health problem on a worldwide scale, it is necessary to provide accurate and timely forecasts of pollutant levels in order to reduce the negative impacts of this pollution. Recent developments in machine learning and deep learning have made it possible to make accurate predictions about air quality; nevertheless, it is still difficult to implement these models in environments where resources are limited for real-world applications. The purpose of this study is to provide an ensemble deep learning-based system for predicting air pollution that is specifically designed for metropolitan areas. The system that has been developed displays greater accuracy and resilience in the management of the complexity of air quality data. This is accomplished via the integration of numerous deep learning models. A dependable instrument for predicting real-time air quality, the ensemble model outperforms conventional statistical approaches, individual machine learning models, and standalone deep learning models, according to the comparison study. This demonstrates that the ensemble model is superior to more traditional statistical methods. The findings of this research shed light on the potential of ensemble deep learning algorithms in the field of environmental management and monitoring. There is a possibility that future study may concentrate on increasing the data sources that are included into the prediction models. For example, traffic patterns, data on industrial activities, and information from social media might be incorporated in order to capture a wider variety of elements that influence air quality. In addition, the investigation of more sophisticated ensemble methods, such as adaptive boosting or deep stacking, has the potential to substantially improve prediction performance. Ultimately, the ensemble deep learning-based air pollution prediction system that has been developed provides a potential solution for sustainable smart cities. It lays the groundwork for proactive air quality management and contributes to the creation of urban settings that are both healthier and more sustainable.

REFERENCES

[1]. S. Myeong and K. Shahzad, "Integrating data-based strategies and advanced technologies with efficient air pollution management in smart cities," Sustainability, vol. 13, no. 13, pp. 7168, 2021
[2]. Chen, T., & Zhang, C. (2020). Ensemble learning for air pollution prediction. Environmental Modelling & Software, 124, 104602. https://doi.org/10.1016/j.envsoft.2019.104602
[3]. Gao, N., Fu, L., Tang, F., Zhang, Z., Zhang, X., & Gong, X. (2021). Air quality forecasting using a hybrid model based on deep learning and ensemble learning. Environmental Science and Pollution Research, 28(37), 51655-51667. https://doi.org/10.1007/s11356-021-14473-3.
[4]. Y. Perwej, "Unsupervised Feature Learning for Text Pattern Analysis with Emotional Data Collection: A Novel System for Big Data Analytics", IEEE

International Conference on Advanced computing Technologies & Applications (ICACTA'22), SCOPUS, IEEE No: #54488 ISBN No Xplore: 978-1-6654-9515-8, Coimbatore, India, 4-5 March 2022, DOI: 10.1109/ICACTA54488.2022.9753501

[5]. Y. Perwej, Firoj Parwej, "A Neuroplasticity (Brain Plasticity) Approach to Use in Artificial Neural Network", International Journal of Scientific & Engineering Research (IJSER), France, ISSN 2229 – 5518, Volume 3, Issue 6, Pages 1- 9, 2012, DOI: 10.13140/2.1.1693.2808

[6]. Y. Perwej, "An Evaluation of Deep Learning Miniature Concerning in Soft Computing", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), ISSN (Online): 2278-1021, ISSN (Print): 2319-5940, Volume 4, Issue 2, Pages 10 - 16, 2015, DOI: 10.17148/IJARCCE.2015.4203

[7]. Y. Perwej, "The Bidirectional Long-Short-Term Memory Neural Network based Word Retrieval for Arabic Documents", Transactions on Machine Learning and Artificial Intelligence (TMLAI), which is published by Society for Science and Education, United Kingdom (UK), ISSN 2054-7390, Volume 3, Issue 1, Pages 16 - 27, 2015, DOI: 10.14738/tmlai.31.863

[8]. Vaishali Singh, Soumya Verma, Ayush Srivastava, Abhishek Dubey, Dr. Nikhat Akhtar, "Eco-Sensing System for Water Pollution and Microplastic Detection", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN: 2456-3307, Volume 11, Issue 3, Pages 679-690, May 2025, DOI: 10.32628/CSEIT25113333

[9]. Alhirmizy, S.; Qader, B. Multivariate Time Series Forecasting with LSTM for Madrid, Spain Pollution. In Proceedings of the 2019 International Conference on Computing and Information Science and Technology and Their Applications (ICCISTA), Baghdad, Iraq, 16–17 March 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–5.

[10]. X. B. Jin, Z. Y. Wang, W. T. Gong, J. L. Kong, Y. T. Bai et al., "Variational Bayesian network with information interpretability filtering for air quality forecasting," Mathematics, vol. 11, no. 4, pp. 837, 2023.

[11]. C.Magazzino, M. Mele and N. Schneider, "The relationship between air pollution and COVID-19-related deaths: An application to three French cities," Applied Energy, vol. 279, pp. 115835, 2020.

[12]. M. Alghieth, R. Alawaji, S. H. Saleh and S. Alharbi, "Air pollution forecasting using deep learning," International Journal of Online & Biomedical Engineering, vol. 17, no. 14, pp. 50–64, 2021

[13]. Yusuf Perwej , Asif Perwej , "Forecasting of Indian Rupee (INR) / US Dollar (USD) Currency Exchange Rate Using Artificial Neural Network", International Journal of Computer Science, Engineering and Applications (IJCSEA), which is published by Academy & Industry Research Collaboration Center (AIRCC), USA , Volume 2, No. 2, Pages 41- 52, April 2012, DOI: 10.5121/ijcsea.2012.2204

[14]. Y. Perwej, "Unsupervised Feature Learning for Text Pattern Analysis with Emotional Data Collection: A Novel System for Big Data Analytics", IEEE International Conference on Advanced computing Technologies & Applications (ICACTA'22), SCOPUS, IEEE No: #54488 ISBN No Xplore: 978-1-6654-9515-8, Coimbatore, India, 2022, DOI: 10.1109/ICACTA54488.2022.9753501

[15]. Y. Perwej, S. A. Hann, N. Akhtar, "The State-of-the-Art Handwritten Recognition of Arabic Script Using Simplified Fuzzy ARTMAP and Hidden Markov Models", International Journal of Computer Science and Telecommunications (IJCST), Sysbase Solution (Ltd), UK, London, ISSN 2047-3338, Volume, Issue 8, Pages, 26 - 32, 2014

[16]. Y. Perwej, Asif Perwej, Firoj Parwej, "An Adaptive Watermarking Technique for the copyright of digital images and Digital Image Protection", International journal of Multimedia & Its Applications (IJMA), Academy & Industry Research Collaboration Center (AIRCC) , USA , Volume 4, No.2, Pages 21- 38, 2012, DOI: 10.5121/ijma.2012.4202

[17]. Y. Perwej , Firoj Parwej, Asif Perwej, "Copyright Protection of Digital Images Using Robust Watermarking Based on Joint DLT and DWT ", International Journal of Scientific & Engineering Research (IJSER), France, ISSN 2229-5518, Volume 3, Issue 6, Pages 1- 9, June 2012

[18]. Gilik, A.; Ogrenci, A.S.; Ozmen, A. Air quality prediction using CNN+ LSTM-based hybrid deep learning architecture. Environ. Sci. Pollut. Res. 2022, 29, 11920–11938

[19]. Han, S., Lee, S., & Hwang, M. (2022). A deep learning-based hybrid model for air pollution prediction. Journal of Environmental Management, 305, 114385. https://doi.org/10.1016/j.jenvman.2021.114385

[20]. J. M. Garcia, F. Teodoro, R. Cerdeira, L. M. Coelho, P. Kumar, and M. G. Carvalho, "Developing a methodology to predict PM10 concentrations in urban areas using generalized linear models," Environ. Technol., vol.37, no.18, pp.2316-2325, 2016. doi: 10.1080/09593330.2016.1149228.

[21]. S. Park, M. Kim, M. Kim, H. G. Namgung, K. T. Kim, K. H. Cho, and S. B. Kwon, "Predicting PM10 concentration in Seoul metropolitan subway stations using artificial neural network (ANN)," J. Hazard. Mater., vol.341, pp.75-82, 2018. doi: 10.1016/j.jhazmat.2017.07.050.

[22]. R. Yu, Y. Yang, L. Yang, G. Han, and O. A. Move, "RAQ-A Random Forest Approach for Predicting Air Quality in Urban Sensing Systems," Sensors (Basel, Switzerland), vol.16, no.1, pp.86, 2016, doi: 10.3390/s16010086.

[23]. Shobhit Kumar Ravi, Shivam Chaturvedi, Dr. Neeta Rastogi, N. Akhtar, Y. Perwej, "A Framework for Voting Behavior Prediction Using Spatial Data", International Journal of Innovative Research in Computer Science & Technology (IJIRCST), ISSN:

2347-5552, Volume 10, Issue 2, Pages 19-28, 2022, DOI: 10.55524/ijircst.2022.10.2.4
[24]. Y. Perwej, "An Optimal Approach to Edge Detection Using Fuzzy Rule and Sobel Method", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE), ISSN (Print) : 2320 – 3765, ISSN (Online): 2278 – 8875, Volume 4, Issue 11, Pages 9161-9179, 2015, DOI: 10.15662/IJAREEIE.2015.0411054
[25]. Shweta Pandey, Rohit Agarwal, Sachin Bhardwaj, Sanjay Kumar Singh, Y. Perwej, Niraj Kumar Singh, "A Review of Current Perspective and Propensity in Reinforcement Learning (RL) in an Orderly Manner" , the International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), Volume 9, Issue 1, Pages 206-227, 2023, DOI: 10.32628/CSEIT2390147
[26]. Mahmoud AbouGhaly, Yusuf Perwej, Mumdouh Mirghani Mohamed Hassan, Nikhat Akhtar, "Smart Sensors and Intelligent Systems: Applications in Engineering Monitoring" , International Journal of Intelligent Systems and Applications in Engineering, SCOPUS, ISSN: 2147- 6799, Volume 12, Issue 22s, Pages 720–727, July 2024
[27]. N. Akhtar, Nazia Tabassum, Dr. Asif Perwej, Y. Perwej," Data Analytics and Visualization Using Tableau Utilitarian for COVID-19 (Coronavirus)", Global Journal of Engineering and Technology Advances (GJETA), Volume 3, Issue 2, Pages 28-50, 2020, DOI: 10.30574/gjeta.2020.3.2.0029
[28]. S. Li, G. Xie, J. Ren, L. Guo, Y. Yang et al., "Urban PM2.5 concentration prediction via attention-based cnn–lstm," Applied Sciences, vol. 10, no. 6, pp. 1953, 2020.
[29]. P.Asha,L.Natrayan, B. T. Geetha, J. R. Beulah,R. Sumathy et al., "IoT enabled environmental toxicology for air pollution monitoring using AI techniques," Environmental Research, vol. 205, pp. 112574, 2022.
[30]. B. Zhang, H. Zhang, G. Zhao and J. Lian, "Constructing a PM2.5 concentration prediction model by combining auto-encoder with Bi-LSTM neural networks," Environmental Modelling Software, vol. 124, pp. 104600, 2020.
[31]. Farheen Siddiqui, Sunny Kumar, Dr. Yusuf Perwej, Homa Rizvi, Dr. Nikhat Akhtar, "Optimizing Air Quality Labelling with Advanced Machine Learning Techniques", Journal of Emerging Technologies and Innovative Research (JETIR), ISSN-2349-5162, Volume 12, Issue 12, Pages 688-694, December 2025, DOI: 10.6084/m9.jetir.JETIR2512381
[32]. W. Du, L. Chen, H. Wang, Z. Shan, Z. Zhou et al., "Deciphering urban traffic impacts on air quality by deep learning and emission inventory," Journal of Environmental Sciences, vol. 124, pp. 745–757, 2023.
[33]. T. Li, M. Hua and X. Wu, "A hybrid CNN-LSTM model for forecasting particulate matter (PM2.5)," IEEE Access, vol. 8, pp. 26933–26940, 2020.
[34]. Y. Perwej, Nikhat Akhtar, Devendra Agarwal, "The emerging technologies of Artificial Intelligence of Things (AIoT) current scenario, challenges, and opportunities", Book Title "Convergence of Artificial Intelligence and Internet of Things for Industrial Automation", SCOPUS, ISBN: 978-1-032-42844-4, CRC Press, Taylor & Francis Group, 2024, DOI: 10.1201/9781003509240-1
[35]. Patil V. H. , Dhanke J. ., Ghaly M. A. ., Patil, R. N. ., Yusuf Perwej, Shrivastava A., "A Novel Approach for Crop Selection and Water Management using Mamdani's Fuzzy Inference and IoT" ,International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), SCOPUS, ISSN: 2321-8169, Volume 11, Issue 9, Pages 62-68, 2023, DOI: 10.17762/ijritcc.v11i9.8320
[36]. Shubham Mishra, Mrs Versha Verma, Nikhat Akhtar, Shivam Chaturvedi, Yusuf Perwej, "An Intelligent Motion Detection Using OpenCV" , International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Print ISSN: 2395-1990 , Online ISSN : 2394-4099, Volume 9, Issue 2, Pages 51-63, 2022, DOI: 10.32628/IJSRSET22925
[37]. Venkata K. S. Maddala, Dr. Shantanu Shahi, Yusuf Perwej, H G Govardhana Reddy, "Machine Learning based IoT application to Improve the Quality and precision in Agricultural System", European Chemical Bulletin (ECB), ISSN: 2063-5346, SCOPUS, Hungary, Volume 12, Special Issue 6, Pages 1711 – 1722, May 2023, DOI: 10.31838/ecb/2023.12.si6.157
[38]. N. Akhtar, Hemlata Pant, Apoorva Dwivedi, Vivek Jain, Y. Perwej, "A Breast Cancer Diagnosis Framework Based on Machine Learning", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Print ISSN: 2395-1990, Volume 10, Issue 3, Pages 118-132, 2023, DOI: 10.32628/IJSRSET2310375
[39]. Pak, U.; Ma, J.; Ryu, U.; Ryom, K.; Juhyok, U.; Pak, K.; Pak, C. Deep Learning-Based PM2.5 Prediction Considering the Spatiotemporal Correlations: A Case Study of Beijing, China. Sci. Total Environ. 2020, 699, 133561.
[40]. Ma, J.; Cheng, J.C.; Lin, C.; Tan, Y.; Zhang, J. Improving Air Quality Prediction Accuracy at Larger Temporal Resolutions Using Deep Learning and Transfer Learning Techniques. Atmos. Environ. 2019, 214, 116885.
[41]. Alhirmizy, S.; Qader, B. Multivariate Time Series Forecasting with LSTM for Madrid, Spain Pollution. In Proceedings of the 2019 International Conference on Computing and Information Science and Technology and Their Applications (ICCISTA), Baghdad, Iraq, 16–17 March 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–5.
[42]. Pagano, E.; Barbierato, E.A. Time Series Approach to Smart City Transformation: The Problem of Air Pollution in Brescia. AI 2024, 5, 17–37.
[43]. KDV Prasad, Yusuf Perwej, E. Nageswara Rao, Himanshu Bhaidas Patel, "IoT Devices for Agricultural to Improve Food and Farming Technology", Journal of Survey in Fisheries Sciences (JSFS), ISSN: 2368-7487, SCOPUS, Volume 10, No.

1S (2023): Special Issue 1, Pages 4054-4069, Canada, 2023

[44]. Kajal, Neha Singh, Nikhat Akhtar, Ms. Sana Rabbani, Y. Perwej, Susheel Kumar, "Using Emerging Deep Convolutional Neural Networks (DCNN) Learning Techniques for Detecting Phony News", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN: 2456-3307, Volume 10, Issue 1, Pages 122-137, 2024, DOI: 10.32628/CSEIT2410113

[45]. Firoj Parwej, N. Akhtar, Y. Perwej, "A Close-Up View About Spark in Big Data Jurisdiction", International Journal of Engineering Research and Application (IJERA), ISSN: 2248-9622, Volume 8, Issue 1, (Part -I1), Pages 26-41, January 2018, DOI: 10.9790/9622-0801022641

[46]. Saurabh Sahu, Km Divya, Neeta Rastogi, Puneet Kumar Yadav, Yusuf Perwej, "Sentimental Analysis on Web Scraping Using Machine Learning Method" , Journal of Information and Computational Science (JOICS), ISSN: 1548-7741, Volume 12, Issue 8, Pages 24- 29, 2022, DOI: 10.12733/JICS.2022/V12I08.535569.67004

[47]. Q. Tao, F. Liu, Y. Li, and D. Sidorov, "Air pollution forecasting using a deep learning model based on 1D convnets and bidirectional GRU," IEEE Access, Vol.7, pp.76690-76698, 2019. doi: 10.1109/ACCESS.2019.2921578.

[48]. Y. Liu, P. Wang, Y. Li, L. Wen, and X. Deng, "Air quality prediction models based on meteorological factors and real-time data of Industrial Waste Gas," Sci. Rep., Vol.12, No.1, pp.8392, 2022. doi: 10.1038/s41598-022-13579-2.

[49]. P. Jiang, C. Li, R. Li, and H. Yang, "An innovative ensemble air pollution early-warning system based on pollutants forecasting and Extenics," Knowl.-Based Vol.164, pp.174-192, 2019. doi: 10.1016/j.knosys.2018.10.036.

[50]. Qiu, Y., Zhang, W., Wang, L., & Liang, P. (2021). Multi-step air quality forecasting based on deep LSTM with spatiotemporal correlation. Science of Total Environment, 776

[51]. Shamim Ahmad, Farheen Siddiqui, Y. Perwej, Homa Rizvi, Dr. Nikhat Akhtar, "Modelling Crop Yield with Deep CNN-LSTM for Spatiotemporal Data Analysis", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN: 2456-3307, Volume 11, Issue 5, Pages 54-64, September 2025, DOI: 10.32628/CSEIT25111697

[52]. Y. Perwej ," The Hadoop Security in Big Data: A Technological Viewpoint and Analysis ", International Journal of Scientific Research in Computer Science and Engineering (IJSRCSE) , E-ISSN: 2320-7639, Volume 7, Issue 3, Pages 1- 14, June 2019, DOI: 10.26438/ijsrcse/v7i3.1014

[53]. Vito, S. (2008). Air Quality [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C59K5F

[54]. M. A. Hamza, H. Shaiba, R. Marzouk, A. Alhindi, M. M. Asiri et al., "Big data analytics with artificial intelligence enabled environmental air pollution monitoring framework," Computers, Materials & Continua, vol. 73, no. 2, pp. 3235–3250, 2022

[55]. Y. Perwej, Shaikh Abdul Hannan, Firoj Parwej, Nikhat Akhtar, "A Posteriori Perusal of Mobile Computing", International Journal of Computer Applications Technology and Research (IJCATR), which is published by ATS (Association of Technology and Science), India, ISSN 2319–8656 (Online), Volume 3, Issue 9, Pages 569 - 578, 2014, DOI: 10.7753/IJCATR0309.1008

[56]. Apoorva Dwivedi, Basant Ballabh Dumka, Nikhat Akhtar, Ms Farah Shan, Yusuf Perwej, "Tropical Convolutional Neural Networks (TCNNs) Based Methods for Breast Cancer Diagnosis", International Journal of Scientific Research in Science and Technology (IJSRST), Print ISSN: 2395-6011, Online ISSN: 2395-602X, Volume 10, Issue 3, Pages 1100 - 1116, May-June- 2023, DOI: 10.32628/IJSRST523103183

[57]. Du, S.; Li, T.; Yang, Y.; Horng, S.J. Deep air quality forecasting using hybrid deep learning framework. IEEE Trans. Knowl. Data Eng. 2019, 33, 2412–2424

[58]. Gilik, A.; Ogrenci, A.S.; Ozmen, A. Air quality prediction using CNN+ LSTM-based hybrid deep learning architecture. Environ. Sci. Pollut. Res. 2022, 29, 11920–11938

[59]. Neha Kulshrestha, Nikhat Akhtar, Yusuf Perwej, "Deep Learning Models for Object Recognition and Quality Surveillance", Accepted International Conference on Emerging Trends in IoT and Computing Technologies (ICEICT-2022), ISBN 978-10324-852-49, SCOPUS, Routledge, Taylor & Francis, CRC Press, Chapter 75, pages 508-518, Goel Institute of Technology & Management,Lucknow, May 2022

[60]. Firoj Parwej, N. Akhtar, Y. Perwej, "An Empirical Analysis of Web of Things (WoT)", International Journal of Advanced Research in Computer Science (IJARCS), ISSN: 0976-5697, Volume 10, No. 3, Pages 32-40, May 2019., DOI: 10.26483/ijarcs.v10i3.6434

[61]. Y. Perwej, Firoj Parwej, N. Akhtar, "An Intelligent Cardiac Ailment Prediction Using Efficient ROCK Algorithm and K- Means & C4.5 Algorithm", European Journal of Engineering Research and Science (EJERS), Bruxelles, Belgium, ISSN: 2506-8016, Vol. 3, No. 12, Pages 126 – 134, 2018, DOI:10.24018/ejers.2018.3.12.989

[62]. Aram, S.A.; Nketiah, E.A.; Saalidong, B.M.; Wang, H.; Afitiri, A.R.; Akoto, A.B.; Lartey, P.O. Machine Learning-Based Prediction of Air Quality Index and Air Quality Grade: A Comparative Analysis. Int. J. Environ. Sci. Technol. 2024, 21, 1345–1360

[63]. Guler, E.; Ozcan, B. PM2.5 Concentration Prediction Based on Winters' and Fourier Analysis with Least Squares Methods in Cerkezkoy district of Tekirda˘ g. Int. J. Environ. Pollut. Environ. Model. 2021, 4, 17–27

[64]. R. Priyadarshini, Naim Shaikh, Rakesh Kumar Godi, Yusuf Perwej, P.K. Dhal, Rajeev Sharma, "IoT-Based

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec1067

Power Control Systems Framework for Healthcare Applications" , Measurement: Sensors,ELSEVIER, ScienceDirect, SCIE, Web of Science, SCOPUS, ISSN 2665- 9174, Volume 25, Pages 1-6, January 2023 DOI: 10.1016/j.measen.2022.100660

[65]. Asif Perwej, K. P. Yadav, Vishal Sood, Yusuf Perwej, "An Evolutionary Approach to Bombay Stock Exchange Prediction with Deep Learning Technique", for published in the IOSR Journal of Business and Management (IOSR-JBM), e-ISSN: 2278-487X, p-ISSN: 2319-7668, USA, Volume 20, Issue 12, Ver. V, Pages 63-79, December. 2018., DOI: 10.9790/487X-2012056379

[66]. Y. Perwej, Nikhat Akhtar, Firoj Parwej, "The Kingdom of Saudi Arabia Vehicle License Plate Recognition using Learning Vector Quantization Artificial Neural Network", International Journal of Computer Applications (IJCA), USA, ISSN 0975 – 8887, Volume 98, No.11, Pages 32 – 38, http://www.ijcaonline.org., 2014, DOI: 10.5120/17230-7556

[67]. Asif Perwej, Y. Perwej, N. Akhtar, and Firoj Parwej, "A FLANN and RBF with PSO Viewpoint to Identify a Model for Competent Forecasting Bombay Stock Exchange COMPUSOFT, SCOPUS, An International Journal of Advanced Computer Technology, 4 (1), Volume-IV, Issue-I, Pages 1454-1461, January-2015, DOI: 10.6084/ijact.v4i1.60

[68]. Pagano, E.; Barbierato, E.A. Time Series Approach to Smart City Transformation: The Problem of Air Pollution in Brescia. AI 2024, 5, 17–37

[69]. Li, Y.; Guo, J.; Sun, S.; Li, J.; Wang, S.; Zhang, C. Air quality forecasting with artificial intelligence techniques: A scientometric and content analysis. Environ. Model. Softw. 2022, 149, 105329

[70]. Liang, Y.-C.; Maimury, Y.; Chen, A.H.-L.; Juarez, J.R.C. Machine learning-based prediction of air quality. Appl. Sci. 2020, 10, 9151