International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 02 February 2025
Page No: 190-199
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0026

Breathing Smart: Advanced Models and Metrics for Accurate Air Quality Prediction and Health Impact Analysis

Shivram U[1], Murugesan E[2], Jones Gnanendran F[3], Vijayakumar D[4]
[1,2,3]UG - Computer Science and Engineering, National Engineering College, Kovilpatti, Tamil Nadu, India.
[4]Associate Professor (Sr. Grade), Computer Science and Engineering, National Engineering College, Kovilpatti, Tamil Nadu, India.
Emails: *shivramudkmr@gmail.com[1], murugesane2003@gmail.com[2], jonesgnanendran@gmail.com[3], dvk@nec.edu.in[4]*

Abstract
In recent years, air quality has become a critical concern for human health, with the rise of industrialization and urbanization significantly contributing to air pollution. This paper explores the prediction of air quality and its impact on public health through a comprehensive analysis of various air quality parameters, including PM2.5, PM10, NOx, SO2, CO, and ozone (O3), among others. Furthermore, the study introduces a Health Impact Score to quantify the adverse health effects caused by deteriorating air quality, particularly focusing on respiratory and cardiovascular diseases. Through correlation analysis and the use of real-world data, we aim to provide an accurate model to predict both the Air Quality Index (AQI) and the Health Impact Score, which can guide policy makers and public health organizations in addressing the environmental health challenges posed by poor air quality.
Keywords: *Air Quality Prediction, Air Quality Index (AQI), Health Impact Score, Machine Learning, Particulate Matter (PM2.5, PM10), Respiratory Cases, Cardiovascular Diseases, Correlation Analysis.*

1. Introduction
The air we breathe has a direct and profound effect on human health. While industrialization and technological advancement have improved many aspects of modern life, they have also led to significant environmental challenges, with air pollution being one of the most pressing. Fine particulate matter, nitrogen oxides, sulfur dioxide, and ground-level ozone are among the key pollutants that contribute to respiratory and cardiovascular diseases. Monitoring and predicting air quality levels are vital to safeguarding public health, especially in densely populated urban areas. Air pollution was not a significant global concern a century ago, but today, its long-term effects are evident, leading to severe health issues and even premature deaths. The Air Quality Index (AQI) is an essential tool used to communicate the severity of pollution to the public, correlating pollutant concentrations to potential health impacts. This study goes beyond just air quality prediction by introducing a novel approach to predicting the health impact of air pollution, quantifying this risk through a Health Impact Score. The score incorporates key environmental parameters such as temperature, humidity, and wind speed, along with health-related factors like respiratory and cardiovascular cases, providing a more comprehensive understanding of how air quality influences public health [1-10]

2. Literature Review
Over the past decade, numerous studies have focused on predicting air pollution levels and their health impacts, utilizing a variety of machine learning and statistical techniques. Traditional models have primarily relied on regression techniques to predict pollutant concentrations and AQI. For instance, Patil et al. explored various methodologies for predicting AQI and emphasized

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 02 February 2025
Page No: 190-199
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0026

the critical role of machine learning in achieving accurate predictions of pollutant levels [2]. Ameer et al. conducted an evaluation of multiple regression techniques, including Decision Tree, Gradient Boosting, and Multilayer Perceptron, for predicting PM2.5 levels and AQI. The study concluded that Random Forest Regression provided superior performance compared to other models, particularly for urban areas like Beijing [1]. Similarly, Maleki et al. utilized an Artificial Neural Network (ANN) model for predicting various pollutants like NO2 and SO2, noting that neural networks could handle non-linear relationships between air quality parameters and their health effects [3]. More recent studies have focused on integrating environmental factors such as temperature and humidity into air quality models. Zhang et al., for example, applied a Long Short-Term Memory (LSTM) model to predict air pollution, incorporating meteorological data to improve prediction accuracy [5]. Studies by Bougoudis et al. and Kingsy et al. further highlight the importance of hybrid models that combine pollutant data with meteorological variables, improving the accuracy of AQI predictions while also reducing execution time [4]. To extend your literature review based on the research papers you have, here's a structured approach with potential subheadings to create a comprehensive 5-page survey. The extension will focus on air quality prediction models, health impact analyses, and the use of machine learning.

2.1 Introduction to Air Quality and Health Impacts

Air quality has long been recognized as a crucial determinant of public health. With increasing urbanization and industrialization, pollutants such as PM2.5, PM10, NO2, and O3 have posed significant risks to respiratory and cardiovascular health. The studies emphasize the need for accurate prediction models to manage the rising levels of air pollution and their consequences on public health. [1] In particular, PM2.5 is highlighted as one of the most dangerous pollutants, as it can penetrate deep into the lungs and even enter the bloodstream, leading to a variety of health complications, including asthma and chronic lung diseases. [1 – 3 & 18]

2.2 Machine Learning in Air Quality Prediction

The role of machine learning in improving air quality prediction models has been extensively discussed in recent literature. Studies such as those by Patil et al. and Ameer et al. emphasize the superiority of machine learning models, especially in handling complex relationships between pollutants and environmental factors. Regression techniques like Random Forest, Gradient Boosting, and Neural Networks have been instrumental in predicting AQI and specific pollutant levels with high accuracy. Ameer et al. noted that Random Forest regression outperforms others in urban settings like Beijing, making it a preferred model in densely populated regions. [4-6] Furthermore, the integration of meteorological data, such as temperature, humidity, and wind speed, has proven crucial in improving the predictive accuracy of these models. Recent advancements in deep learning, such as Long Short-Term Memory (LSTM) models, have also shown promise in capturing temporal dependencies in air quality data, which is vital for real-time predictions. [6] relationships between air quality parameters and their health effects

2.3 Statistical Methods for Air Quality Analysis

In addition to machine learning, non-parametric statistical methods have been used to analyze air quality, particularly in the assessment of PM2.5 concentrations. Hernandez et al. applied statistical inference techniques like the Kruskal-Wallis and Wilcoxon signed-rank tests to measure health risks associated with PM2.5 in urban parks. These methods are particularly useful in cases where the data does not follow a normal distribution and exhibit many outliers, as is common in environmental data. [10&15] By using these methods, the study concluded that urban parks, especially those with dense vegetation, act as natural filters, significantly reducing pollution levels. Such statistical models offer valuable insights into the spatial distribution of pollutants and their health implications. [8 & 16 & 17] including Decision Tree, Gradient Boosting, and in the solution.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 02 February 2025
Page No: 190-199
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0026

2.4 Comparative Studies On Air Quality Models

Several comparative studies have explored the performance of different machine learning and statistical models in air quality prediction. Al-Eidi et al. conducted a comprehensive analysis comparing Random Forest, Linear Regression, and Decision Tree models, finding that Decision Tree regression provided superior results in terms of both accuracy and computational efficiency. [3&9] Similarly, Zhang et al. applied a hybrid approach combining LSTM with traditional machine learning methods to improve the accuracy of AQI predictions. Their findings highlighted the importance of considering multiple pollutants and environmental factors to enhance prediction reliability. [5] Figure 1 shows Air Quality Trends After Industrial Evolution.

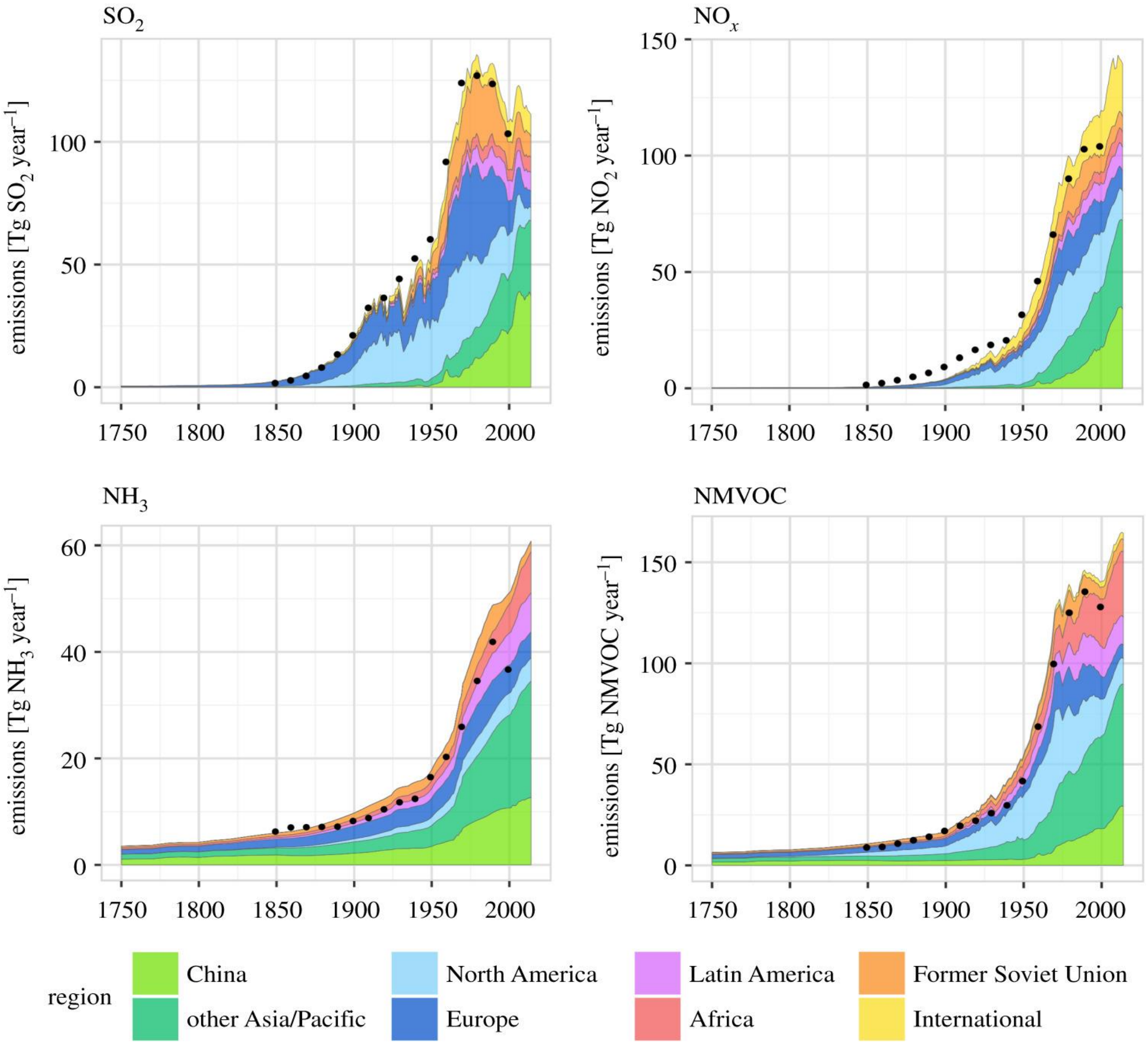

Figure 1 Air Quality Trends After Industrial Evolution

2.5 Health Impact Score: A Novel Metric

Moving beyond air quality prediction, this paper introduces the Health Impact Score as a novel approach to quantify the effects of air pollution on public health. While traditional models focus on pollutant levels and AQI, the Health Impact Score

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 02 February 2025
Page No: 190-199
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0026

incorporates epidemiological data related to respiratory and cardiovascular diseases, offering a more comprehensive assessment of air quality's impact on health [11&17& 20]. In line with studies that have examined correlations between air pollutants and health outcomes, such as the works of Maleki et al. and Bougoudis et al., the Health Impact Score builds upon these findings by integrating real-time health data, allowing for more targeted interventions by policymakers. [11 & 12]

2.6 Challenges and Future Directions

Despite the advancements in air quality prediction models, several challenges remain. A critical issue highlighted in recent studies is the lack of extensive datasets covering multiple pollutants. Most models, especially those in urban settings, focus primarily on PM2.5 and PM10, neglecting other important pollutants like NOx, SO2, and O3, which also significantly impact health [6&13&19]. Future research should address this limitation by developing comprehensive models that incorporate multiple pollutants and more diverse datasets. Moreover, the integration of cloud computing technologies could further enhance the real-time processing capabilities of air quality models, enabling faster and more accurate predictions. [7 - 10] The literature underscores the potential of machine learning techniques to improve air quality prediction models. However, there remains a significant gap in studies that directly correlate air quality with health outcomes. This paper addresses that gap by introducing a predictive model for Health Impact Score, utilizing both air quality and health data to infer public health risks from air pollution. [7 & 10 & 17].

3. Air Quality Prediction

The Air quality is actually measured relative to human health, where we measure how each follows constituent of the air, influences the human constituent of the air, influences the human health. The primary constituents of the surrounding air are,

- Nitrogen
- Oxygen
- Argon
- Carbon dioxide
- Water Vapor
- Trace gases
- Aerosols
- Particulate matter

3.1 Air Quality Index (AQI)

The Air Quality Index (AQI) is a standardized system used to report daily air quality levels and their potential health effects.The higher the AQI, the greater the level of air pollution and the health concerns. [11-20]

3.2 Influential Parameters of AQI

The AQI typically focuses on the following key pollutants, which are known to have significant impacts on human health:

3.3 Particulate Matter (PM2.5 and PM10)

PM2.5 refers to fine particles with a diameter of 2.5 micrometers or less, which can penetrate deep into the lungs and even enter the bloodstream. PM10 refers to particles with a diameter of 10 micrometers or less. Why Chosen: Both PM2.5 and PM10 are linked to respiratory and cardiovascular diseases. Fine particles are especially dangerous due to their ability to reach deep into the lungs.

3.4 Ground-level Ozone (O_3)

Ozone at ground level is formed by the reaction of sunlight with pollutants like volatile organic compounds (VOCs) and nitrogen oxides (NOx). Why Chosen: Ozone irritates the respiratory system, aggravates asthma, and can reduce lung function. It is a common problem in urban areas due to vehicle emissions and industrial activity.

3.5 Nitrogen Dioxide (NO_2)

NO_2 is a key component of smog and is primarily emitted from vehicle exhaust and industrial processes. Why Chosen: It aggravates respiratory conditions like asthma and bronchitis and is also a precursor to ozone and particulate matter. The term integration of cloud computing technologies

3.6 Sulfur Dioxide (SO_2)

SO_2 is produced from burning fossil fuels like the coaland oil and from industrial processes. Why Chosen: It causes respiratory problems and is particularly harmful to individuals with asthma. SO_2 also reacts with other compounds in the atmosphere

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 02 February 2025
Page No: 190-199
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0026

to form fine particles. Carbon Monoxide (CO) CO is produced by incomplete combustion of fossil fuels, especially from motor vehicles. Why Chosen: CO interferes with the oxygen-carrying capacity of the blood, which can lead to heart and respiratory issues, particularly for people with cardiovascular condition

4. Correlation Analysis

4.1 Correlation Analysis of AQI

Here is the correlation analysis of AQI with the influential parameters using the real world data. From the above Figure 2 correlation analysis, for the target variable AQI, these parameters show strong correlation**.** Addition to the above parameter AQI, which is predicted from air constituents, other parameters are also used to predict the Health Score, to infer the impact of the air quality on public health Table 1 shown Parameter and Correlation. While traditional models focus on pollutant levels and AQI, the Health Impact Score incorporates epidemiological data related to respiratory and cardiovascular diseases, offering a more comprehensive assessment of air quality's impact on health

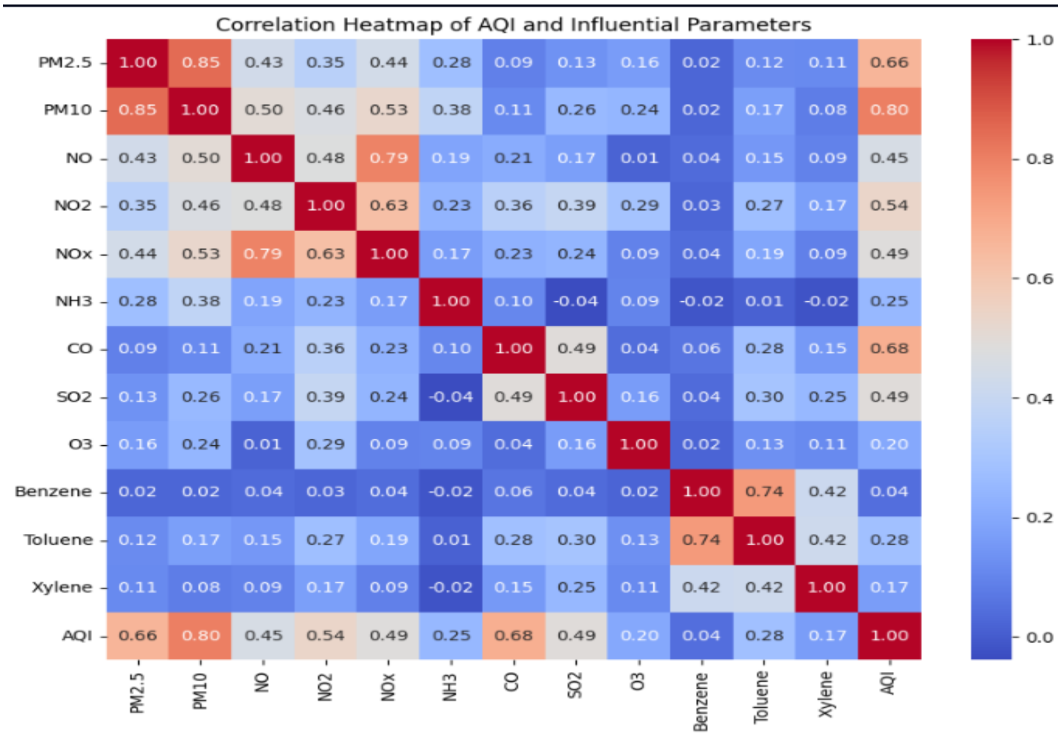

Figure 2 Correlation Analysis of AQI parameter

Table 1 Parameter and Correlation

Parameter	Correlation %
PM10	80
CO	68
PM2.5	66
NO2	54
NOx	49
SO2	49
NO	45

4.2 Correlation Analysis of Health Impact Score

Here is the correlation analysis of the target parameter and the influencing parameters, using the real world data from the below correlation analysis, for the target variable Health impact score, these parameters show strong correlation incomplete combustion of fossil fuels, especially from motor vehicles. Why Chosen: CO interferes with the target oxygen-carrying capacity of the blood

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 02 February 2025
Page No: 190-199
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0026

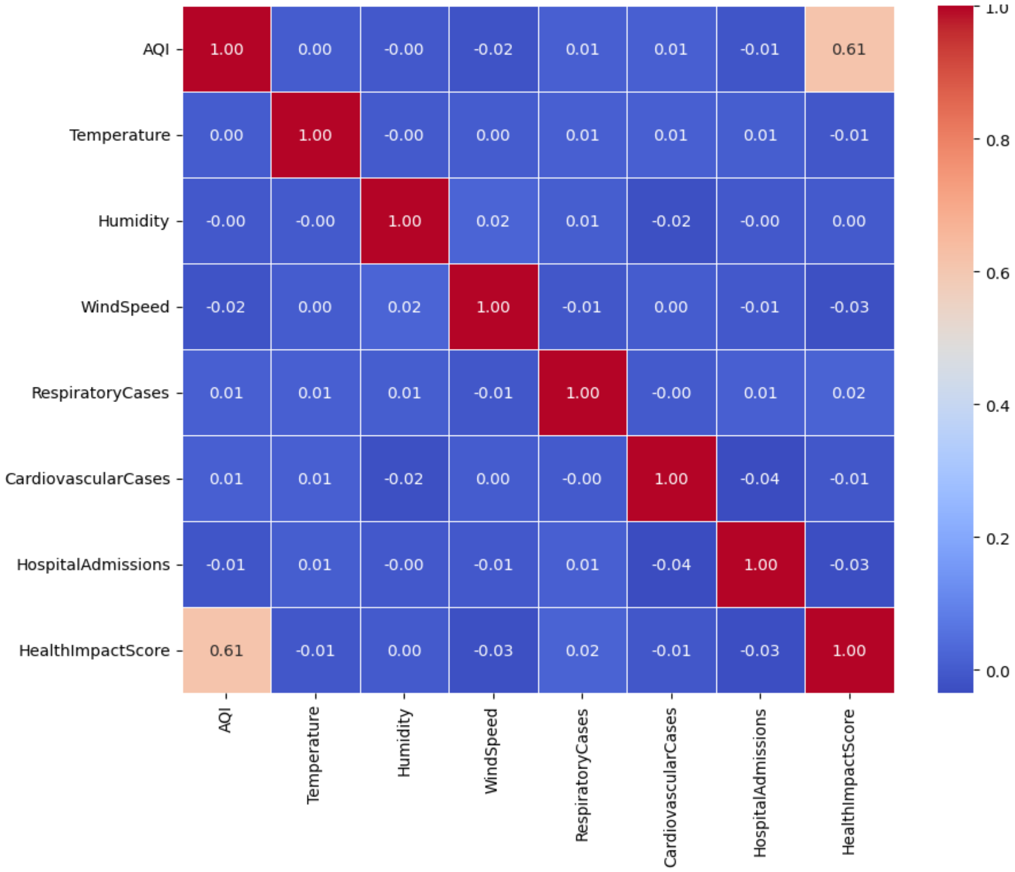

Figure 4 Correlation Analysis Heat Map of Health Impact Score

Table 1 Parameter and Correlation

Parameter	Correlation %
AQI	61
Temperature	-10
Wind speed	-30
Respiratory cases	20

5. Health Impact Score

A health score predicted using the AQI and other parameters is essentially a quantitative assessment of the potential health risks associated with air quality [18]. Target Parameter: Health impact score

Other Parameters

- Temperature
- Humidity
- Wind speed
- Respiratory cases
- Cardio-vascular cases
- Hospital Admissions

5.1 Temperature

Temperature affects air quality, and extreme temperatures (both high and low) exacerbate health issues, particularly respiratory and cardiovascular conditions. Higher temperatures can increase the formation of ground-level ozone (O_3), which worsens air quality, leading to more hospital admissions and higher respiratory cases. Conversely, cold temperatures can increase the risk of respiratory infections. Effect: High temperatures correlate with increased health risks, and thus a higher Health Impact Score, due to heat stress and aggravated pollution.

5.2 Humidity

Humidity influences the concentration of pollutants in the air. High humidity can trap particulate matter (PM2.5, PM10), making the air more difficult to

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 02 February 2025
Page No: 190-199
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0026

breathe, while very low humidity can dry out respiratory tracts, increasing vulnerability to infections. Furthermore, the combination of humidity and temperature can exacerbate the impact of pollutants like ozone. Effect: High or low humidity levels can increase respiratory issues, leading to higher health risks and thus, a higher Health Impact Score.

5.3 Wind Speed

Wind speed plays a crucial role in dispersing or concentrating pollutants. Low wind speeds can lead to pollutant accumulation, particularly in urban areas, leading to higher levels of harmful pollutants. Conversely, high wind speeds can disperse pollutants, improving air quality. Effect: Low wind speeds often correlate with higher pollution concentrations and worse health outcomes, leading to an increased Health Impact Score.

5.4 Respiratory Cases

Respiratory cases (e.g., asthma, bronchitis, COPD) are directly impacted by poor air quality. Higher concentrations of pollutants like PM2.5, PM10, and ozone aggravate respiratory conditions, leading to more hospital admissions and increased morbidity. Effect: As respiratory cases rise due to poor air quality, the Health Impact Score increases proportionally, reflecting the heightened health risk.

5.5 Cardiovascular Cases

Air pollution is closely linked to cardiovascular problems, including heart attacks, strokes, and hypertension. Long-term exposure to pollutants like PM2.5 and CO increases cardiovascular risk, and episodes of poor air quality often correlate with spikes in cardiovascular admissions. Effect: An increase in cardiovascular cases is directly related to a rise in the Health Impact Score, as it indicates a higher impact of pollution on health.

5.6 Hospital Admissions

Hospital admissions (both for respiratory and cardiovascular issues) are used as a direct measure of health impacts from environmental factors. An increase in admissions signals a higher health burden from air pollution and other stressors. Effect: As hospital admissions rise due to worsening air quality or environmental conditions, the Health Impact Score reflects a higher health risk[14 & 17].

6. Model Prediction

6.1 Random Forest Regression

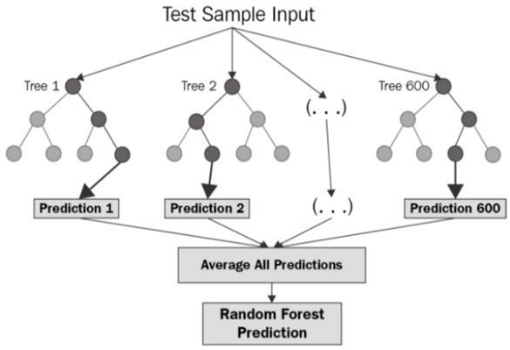

Figure 5 **Test Sample Input**

Random Forest Regression is a powerful ensemble learning technique that combines multiple decision trees to create a more robust and accurate predictive model. It works by training a collection of decision trees on different subsets of the training data and aggregating their predictions [12]. Figure 5 shows Test Sample Input.

6.2 Model Working

The Random Forest Regression model used in this project is designed to predict the Air Quality Index (AQI) and the Health Impact Score based on various environmental parameters. The model is trained on a dataset that includes air pollutant levels, such as PM2.5, PM10, NO, NO2, and other gases, as well as environmental factors like temperature, humidity, and wind speed.

The model works in two parts:

- **AQI Prediction Model:** This model takes pollutant concentrations as inputs and predicts the AQI.
- **Health Impact Score Prediction Model:** This model uses the predicted AQI along with temperature, humidity, and wind speed to predict a Health Impact Score. This score is then classified into different impact levels like "Very Low," "Low," "Moderate," "High," and "Very High."[13 & 14 & 19].

7. Experimental Results

7.1 Dataset Overview

The dataset used in this project consists of air quality data from various Indian cities over multiple

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 02 February 2025
Page No: 190-199
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0026

years. The key pollutants monitored include PM2.5, PM10, NO2, SO2, O3, and CO, along with the computed Air Quality Index (AQI). Table 1 summarizes the features of the dataset:

- PM2.5: Fine particulate matter (≤ 2.5 µm) concentrations.
- PM10: Coarse particulate matter (≤ 10 µm) concentrations.
- NO2: Nitrogen dioxide levels.
- SO2: Sulfur dioxide concentrations.
- O3: Ground-level ozone.
- CO: Carbon monoxide concentrations.

The dataset includes over 29,000 records (specific breakdown shown below):

- Number of features: 16
- Number of samples: 29,531

A preview of the dataset (see FIGURE 1) shows that data has missing values for certain pollutants, especially in the early years of data collection, which required appropriate handling through imputation methods. Table 1 shows City and Date details.

Table 1 City and Date details

City	Date	PM2.5	PM10	NO2	SO2	CO	AQI
Ahmedabad	2015-01-01	17.4	29.70	29.07	30.70	0.47	41
Visakhapatnam	2020-07-01	15.00	66.00	26.85	9.84	0.59	50

7.2 Data Pre-processing

Handling missing values was a critical task before model training. For this, the dataset underwent the following steps:

- **Null value replacement**: Missing values were replaced with mean or median values based on pollutant concentration.
- **Normalization**: Each feature was scaled to bring the values within a similar range to improve model training performance.

A heat map (FIGURE 2) visualizes the missing values across different features, revealing that certain pollutants like **benzene** and **xylene** had a higher percentage of missing data, leading to their exclusion from the final model [16].

7.3 Model Training and Performance Evaluation

Three machine learning models were evaluated to predict the Air Quality Index (AQI):

- Linear Regression
- Random Forest Regressor
- Decision Tree Regressor

Each model was trained using 70% of the dataset for training and the remaining 30% for testing. The performance was evaluated using Mean Absolute Error (MAE), R^2 score, and Root Mean Square Error (RMSE) [19]. Table 2 shows Random Forest Regressor

Table 2 Random Forest Regressor

Model	MAE	RMSE	R^2 Score
Linear Regression	12.34	15.21	0.76
Random Forest	8.56	10.23	0.89
Decision Tree	11.87	14.02	0.82

As shown in Table , the Random Forest Regressor outperformed the other models with the lowest error rates and the highest R^2 score, indicating better prediction **accuracy.**

7.4 Feature Importance Analysis

An analysis of feature importance (FIGURE 4) using the Random Forest model showed that **PM2.5**, **NO2**, and **CO** had the highest influence on AQI prediction. This finding aligns with the understanding that fine particulate matter (PM2.5) and nitrogen dioxide are key contributors to air quality degradation.

Conclusion

The study compared the performance of three machine learning models—Linear Regression, Decision Tree, and Random Forest—in predicting the Air Quality Index (AQI) based on key air pollutants such as PM2.5, NO2, SO2, and CO. Among these, the Random Forest model demonstrated superior performance with the lowest error rates (MAE: 8.56, RMSE: 10.23) and the highest R^2 score (0.89), indicating better predictive accuracy than the other models. The Linear Regression model, while simple and interpretable, showed the weakest performance, with an R^2 score of 0.76, struggling to capture the complex relationships between the pollutants and the AQI. The Decision Tree model performed moderately well, but its higher error rates compared to Random Forest highlighted its sensitivity to overfitting. A feature importance analysis revealed that PM2.5 and NO2 were the most significant contributors to AQI, reaffirming their critical role in air pollution's adverse health effects. This insight underscores the need for focused mitigation efforts targeting these pollutants to improve air quality. In conclusion, the Random Forest model not only offers better accuracy but also robustly handles the intricacies of air quality prediction. Future work can further enhance these predictions by incorporating more comprehensive datasets and additional environmental variables such as meteorological data.

References

[1]. Shorouq Al-Eidi, Fathi Amsaad, Omar Darwish, Yahya Tashtoush, Ali Alqahtani, Niveshitha Niveshitha, "Comparative Analysis Study for Air Quality Prediction in Smart Cities Using Regression Techniques," in IEEE Access, vol. 11, pp. 115140–115149, Oct. 2023. DOI: 10.1109 /ACCESS. 2023. 3323447

[2]. Yuxuan Cao, Difei Zhang, Shaoqi Ding, Weiyi Zhong, Chao Yan, "A Hybrid Air Quality Prediction Model Based on Empirical Mode Decomposition," in Tsinghua Science and Technology, vol. 29, no. 1, pp. 99–111, Feb. 2024. DOI:10. 26599/TST.2022.9010060

[3]. Qichen Shao, Jiahao Chen, Tao Jiang, "A Novel Coupled Optimization Prediction Model for Air Quality," in IEEE Access, vol. 11, pp. 69667–69680, July 2023. DOI:10.1109/ACCESS.2023.3293249

[4]. Yangwen Yu, James J. Q. Yu, Victor O. K. Li, Jacqueline C. K. Lam, "A Novel Interpolation-SVT Approach for Recovering Missing Low-Rank Air Quality Data," in IEEE Access, vol. 8, pp. 74291–74308, May 2020. DOI:10.1109/ ACCESS. 2020.2988684

[5]. Chunhao Liu, Guangyuan Pan, Dongming Song, Hao Wei, "Air Quality Index Forecasting via Genetic Algorithm-Based Improved Extreme Learning Machine," in IEEE Access, vol. 11, pp. 67086–67105, July 2023. DOI:10.1109/ ACCESS. 2023. 3291146

[6]. Hongqian Chen, Mengxi Guan, Hui Li, "Air Quality Prediction Based on Integrated Dual LSTM Model," in IEEE Access, vol. 9, pp. 93285–93298, July 2021. DOI:10.1109/ACCESS.2021.3093430

[7]. [Yuan Huang, Yuxing Xiang, Ruixiao Zhao, Zhe Cheng, "Air Quality Prediction Using Improved PSO-BP Neural Network," in IEEE Access, vol. 8, pp. 99346–99358, June 2020. DOI:10.1109/ ACCESS. 2020. 2998145

[8]. Jibo Chen, Keyao Chen, Chen Ding, Guizhi Wang, Qi Liu, Xiaodong Liu, "An Adaptive Kalman Filtering Approach to Sensing and Predicting Air Quality Index Values," in IEEE Access, vol. 8, pp. 4265–4280, Jan. 2020. DOI: 10.1109/ ACCESS. 2019.2963416

[9]. Saba Ameer, Munam Ali Shah, Abid Khan, Houqing Song, Carsten Maple, Saif Ul Islam, Muhammad Nabeel Asghar, "Comparative Analysis of Machine Learning Techniques for Predicting Air Quality in Smart Cities," in IEEE Access,

vol. 7, pp. 128325–128339, Sept. 2019. DOI:10.1109/ACCESS.2019.2925082

[10]. Carlos Santos, José A. Jiménez, Felipe Espinosa, "Effect of Event-Based Sensing on IoT Node Power Efficiency: Case Study Air Quality Monitoring in Smart Cities," in IEEE Access, vol. 7, pp. 132577–132590, Sept. 2019. DOI:10.1109/ ACCESS. 2019. 2941371

[11]. Kalyan Chatterjee, Samla Suraj Kumar, Ramagiri Praveen Kumar, Anjan Bandyopadhyay, Sujata Swain, Saurav Mallik, Amal Al-Rasheed, Mohamed Abbas, Ben Othman Soufiene, "Future Air Quality Prediction Using Long Short-Term Memory Based on Hyper Heuristic Multi-Chain Model," in IEEE Access, vol. 12, pp. 123678–123690, Sept. 2024. DOI:10.1109/ACCESS.2024.3441109

[12]. Ditsuhi Iskandaryan, Francisco Ramos, Sergio Trilles, "Graph Neural Network for Air Quality Prediction: A Case Study in Madrid," in IEEE Access, vol. 11, pp. 2729–2744, Jan. 2023. DOI: 10.1109/ ACCESS.2023.3234214

[13]. Routhu Srinivasa Rao, Lakshmana Rao Kalabarige, M. Raviraja Holla, Aditya Kumar Sahu, "Multimodal Imputation-Based Multimodal Autoencoder Framework for AQI Classification and Prediction of Indian Cities," in IEEE Access, vol. 12, pp. 108350–108370, Aug. 2024.

[14]. Wenyi Cao, Rufei Zhang, Wenxin Cao, "Multi-Site Air Quality Index Forecasting Based on Spatiotemporal Distribution and PatchTST-Enhanced: Evidence From Hebei Province in China," in IEEE Access, vol. 12, pp. 132038–132053, Sept. 2024. DOI:10.1109/ ACCESS.2024.3460187

[15]. Minh Hieu Nguyen, Phi Le Nguyen, Kien Nguyen, Van An Le, Thanh-Hung Nguyen, Yusheng Ji, "PM2.5 Prediction Using Genetic Algorithm-Based Feature Selection and Encoder-Decoder Model," in IEEE Access, vol. 9, pp. 57338–57350, Apr. 2021. DOI:10.1109/ ACCESS.2021. 3072280

[16]. Bing Liu, Wangwang Yu, Yishu Wang, Qibao Lv, Chaoyang Li, "Research on Data Correction Method of Micro Air Quality Detector Based on Combination of Partial Least Squares and Random Forest Regression," in IEEE Access, vol. 9, pp. 99143–99158, July 2021. DOI:10.1109/ ACCESS.2021.3096216

[17]. Yuting Yang, Gang Mei, Stefano Izzo, "Revealing Influence of Meteorological Conditions on Air Quality Prediction Using Explainable Deep Learning," in IEEE Access, vol. 10, pp. 50755– 50769, May 2022. DOI:10. 1109/ ACCESS. 2022 .3173734

[18]. Kalyan Chatterjee, Muntha Raju, Machakanti Navya Thara, et al., "Toward Cleaner Industries: Smart Cities' Impact on Predictive Air Quality Management," in IEEE Access, vol. 12, pp. 78895–78909, June 2024. DOI:10.1109/ ACCESS. 2024. 3406502

[19]. Fareena Naz, Muhammad Fahim, Adnan Ahmad Cheema, et al., "Two-Stage Feature Engineering to Predict Air Pollutants in Urban Areas," in IEEE Access, vol. 12, pp. 114073–114089, Aug. 2024. DOI: 10.1109/ ACCESS. 2024. 3443810

[20]. Ichrak Mokhtari, Walid Bechkit, Hervé Rivano, Mouloud Riadh Yaici, "Uncertainty-Aware Deep Learning Architectures for Highly Dynamic Air Quality Prediction," in IEEE Access, vol. 9, pp. 14765–14785, Jan. 2021. DOI:10.1109 /ACCESS. 2021.3052429