International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Survey on Air Quality Prediction in Smart Cities.

Prof. Manjunath Patil [a], Shivanand Lamani [b], Raghavendra Kusoor [c], Suhas Pawar [d], Vivek Borgalle [e]

[a,b,c,d,e] *Department of Computer Science and Engineering, Angadi Institute of Technology and Management, Belagavi-590009, India*

A B S T R A C T

Air pollution, a perilous global challenge, introduces harmful toxins into the atmosphere, adversely impacting human health and the environment. Recognized as one of humanity's most significant threats, it causes widespread damage to animals, crops, and forests. Addressing this concern, the transport sector employs machine learning techniques to predict air quality, focusing on preprocessing an air quality dataset and utilizing supervised machine learning algorithms like Logistic Regression, Random Forest, K-Nearest Neighbors, Decision Tree, and Support Vector Machines. Decision Tree algorithm proves effective in air quality prediction, with potential applications in meteorological departments. This research optimally utilizes machine learning and artificial intelligence techniques for future advancements.

Keywords: Air Pollution, Machine Learning, Feature Engineering, Feature Scaling, Machine Learning Models.

1. Introduction

The passage highlights the severe health issues caused by air pollution, emphasizing its impact on respiratory and coronary health. It states that air pollution is recognized as a significant global threat and outlines the various diseases attributed to it, such as stroke, heart disease, lung cancer, and respiratory infections. The passage also notes the alarming number of deaths worldwide linked to air pollution, with a significant portion occurring in lower and middle-income countries.

The text then shifts focus to the situation in South Africa, providing statistics on mortality rates related to outdoor and indoor air pollution. It mentions the challenges in maintaining air quality monitoring devices and highlights the diversity of urban air-related data, emphasizing the lack of a clear understanding of the primary causes and distribution of air pollution.

The passage discusses broader environmental concerns, including temperature rise, global warming, and rising sea levels. It introduces international efforts, such as the Paris Climate Agreement, aimed at mitigating the impact of pollution. The text emphasizes the importance of predicting the carbon footprint of nations and implementing policies to address pollution and reduce fossil fuel usage.

In the latter part of the passage, the focus shifts to the development of a spatiotemporal prediction model to determine the Air Quality Index (AQI) and individual pollutant levels over time. The model aims to overcome challenges posed by large-scale data and high variability. The text mentions conducting experiments with different models and identifying a low-cost, low-complexity combination to accurately predict pollution levels and the AQI. In summary, the passage addresses the significant health risks posed by air pollution globally, with a specific focus on the situation in South Africa. It discusses challenges in monitoring air quality, highlights international efforts to combat pollution, and introduces a spatiotemporal prediction model as a solution to accurately determine pollution levels and the AQI.

2. Methodology

2.1 Preprocessing:

- Preprocessing in air quality data involves handling missing values, temporal and spatial aggregation, normalization, and feature engineering. Quality control, outlier identification, and integration of meteorological data enhance the reliability and accuracy of predictive models for air quality analysis.

2.2 Training dataset:

- In the realm of air quality prediction for smart cities using machine learning, a training dataset comprises historical data encompassing air quality parameters, weather conditions, and geographical features. This dataset forms the basis for training machine learning algorithms, enabling them to discern patterns, correlations, and trends. Through iterative learning, algorithms adapt their parameters to create predictive models, capable of forecasting future air quality conditions. This foundational process facilitates the algorithms in making informed predictions based on new input data, contributing to the advancement of accurate and efficient air quality prediction systems in smart city environments.

2.3 Testing dataset:

- The testing dataset plays a pivotal role in assessing the performance and generalization capabilities of air quality prediction models in smart cities. This dataset, comprising new and unseen data, is input into the model to evaluate its ability to predict air quality values. The testing process scrutinizes the model's performance under real-world complexities and variations. By measuring accuracy and assessing predictive capabilities, the testing dataset ensures the reliability and applicability of the model. This evaluation phase is essential for validating the model's effectiveness, providing insights into its robustness and suitability for accurately predicting air quality in dynamic smart city environments.

2.4 Supervised Machine Learning:

- Supervised machine learning entails training algorithms using labeled datasets, where input data is paired with corresponding output labels. The goal is to learn a mapping function from input to output, allowing the algorithm to make accurate predictions with unseen data. Tasks like classification categorize data, while regression predicts numerical values. This approach leverages the known outcomes during training to enable the algorithm to generalize and infer correct outputs for new, unseen data, making it a fundamental technique in various applications, including categorization, prediction, and decision-making.

2.5 Prediction of air quality:

- Air quality prediction is a vital application employing algorithms and historical data to anticipate forthcoming pollution levels. This process integrates meteorological factors and pertinent features to enhance accuracy. Utilizing machine learning techniques, including regression or classification models, facilitates the development of forecasts for specific parameters like PM, NO2, or O3 concentrations. Precise predictions play a pivotal role in effective public health management and environmental planning, enabling proactive measures to mitigate potential health risks associated with poor air quality. By leveraging predictive models, decision-makers can make informed choices, contributing to a healthier environment and enhancing overall quality of life in communities, making air quality prediction an essential tool in environmental monitoring and management.

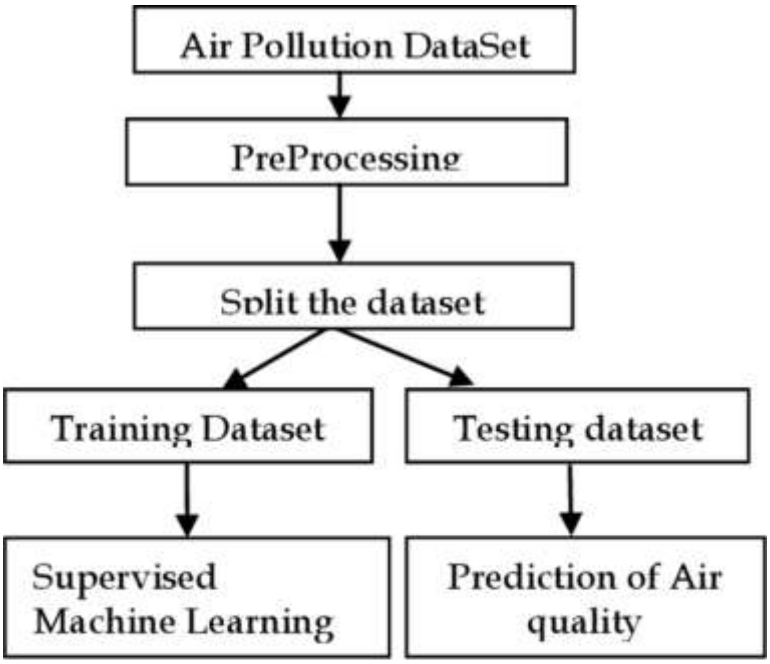

Fig. 1-Methedology Flow

3. Conclusion

Engaging in this project proved to be an exhilarating experience. The pivotal phases of planning, designing, and implementation, as gleaned from our textbooks, became clearer through practical application. This endeavor facilitated collaboration, providing a platform for expressing creativity. The project underscored the significance of cooperation and effective communication. Following extensive hours of labor, the successful completion involved multiple stages, including compilation, debugging, fault correction, and bug-free development. Additional machine learning features and interactions were incorporated, augmenting the project's dependability and utility. Employing Exploratory Data Analysis (EDA), we scrutinized feature distribution, conducted bivariate and multivariate analyses, and ensured a clean, evenly distributed dataset. The meticulous process included the deletion of Null values and the creation of dummy variables for constructing the model.

References

[1] Shah, J., Mishra, B.: IoT-enabled low power environment monitoring system for prediction of PM2. 5. Pervasive Mob. Comput. 67, 101175 (2020).

[2] L. Bai, J. Wang, X. Ma, and H. Lu, “Air pollution forecasts: an overview,” International Journal of Environmental Research and Public Health, vol. 15, no. 4, p. 780, 2018.

[3] Khaled Bashir Shaban, Abdullah Kadri, Eman Rezk,”Urban Air Pollution Monitoring System With Forecasting Models”,IEEE SENSORS JOURNAL, VOL. 16, NO. 8, APRIL 15, 2016

[4] Forecasting Criteria Air Pollutants Using Data Driven Approaches; An Indian Case Study Tikhe Shruti, Dr. Mrs. Khare , Dr. Londhe, IOSR-JESTFT (Mar. - Apr. 2013)

[5] Forecasting Criteria Air Pollutants Using Data Driven Approaches; An Indian Case Study Tikhe Shruti, Dr. Mrs. Khare , Dr. Londhe,IOSR-JESTFT (Mar. - Apr. 2013)

[6] S. Simu, V. Turkar, R. Martires, V. Asolkar, S. Monteiro, V. Fernandes,and V. Salgaoncary, ‘‘Air pollution prediction using machine learning,’’in Proc. IEEE Bombay Sect. Signature Conf. (IBSSC), Dec. 2020,pp. 231–236.

[7] Z. Zhang, H. Chen, and X. Huang, ‘‘Prediction of air quality combiningwavelet transform, DCCA correlation analysis and LSTM model,’’ ApplSci., vol. 13, no. 5, p. 2796, Feb. 2023.

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

[9] S.-C. Hu, Y.-C.Wang, C.-Y.Huang, and Y.-C. Tseng, “Measuring air quality in city areas by vehicular wireless sensor networks,” J. Syst. Softw., vol. 84, no. 11, pp. 2005–2012, 2011

[10] J.-Y. Kim, C.-H.Chu, and S.-M. Shin, “ISSAQ: An integrated sensing systems for real-time indoor air quality monitoring,” IEEE Sensors J., vol. 14, no. 12, pp. 4230–4244, Dec. 2014.