**An Immersive Model for Detecting Air Quality-Machine Learning Approach**

## Abstract

Air pollution has become a significant public health concern, particularly in urban areas. In this study, we propose a machine-learning approach for air quality detection in cities. The proposed model utilizes data from air quality sensors placed throughout the city and combines it with meteorological data and geographic information. The immersive nature of the model provides a more comprehensive view of the air quality development. The proposed model has the potential to be implemented in various settings, including smart cities, industrial zones, and traffic monitoring systems, to improve air quality monitoring and management. Many studies are being conducted in this area; however, the results are still not reliable. Datasets are grouped into two categories: training and testing and are available from air quality monitoring sites like Kaggle.

**Keywords**: LogisticRegression, GaussianNavieBayes, DecisionTreeClassifier, RandomForestClassifier, SupportVectorClassifier, KNeighborsClassifier

**Introduction**

Air pollution is a significant environmental and public health issue, particularly in urban areas where high concentrations of pollutants can have a detrimental impact on human health, the environment, and economic growth. In response to this growing problem, governments and organizations are increasingly investing in air quality monitoring systems to collect data on pollutants level and identify sources of pollution. In this study, we provide a machine-learning method for detecting urban air quality. The proposed model utilizes data from air quality sensors placed throughout the city and combines it with meteorological data and geographic information. The proposed approach provides a powerful tool for air quality detection and management in cities, enabling policy-makers, city planners, and citizens to make informed decisions about the air quality situation in their cities.

**Literature Review**

Air pollution is a significant environmental and public health concern in urban areas, and many researchers have proposed various approaches to detect and forecast air quality using machine learning algorithms.

Dixian Zhu et al. [1] proposed a regularization and optimization-based machine learning approach to forecast the air pollution concentration on an hourly basis. Their model outperformed standard regression models but faced the challenge of multi-task learning.

Pooja Bhalgat et al. [2] presented an ARIMA-based model to identify cities that require immediate attention due to high pollution levels. However, their model was limited by the data's lack of sequential order according to date columns.

Tanisha Madan et al. [3] used decision tree classifiers to predict air quality, which is a straightforward and easy-to-implement approach. However, the decision tree approach resulted in lower accuracy compared to other algorithms.

Ditsuhi Iskandaryan et al. [4] proposed a machine-learning model that utilized sensor data to predict air quality and improve prediction accuracy. However, comparing analysis results was challenging with their model.

Gaganjot Kaur Kang et al. [5] utilized big data technologies to predict air quality but faced issues at different levels of monitoring and analyzing dynamic air quality.

Lan Gao et al. [6] predicted air quality using satellite data and machine learning. However, their model did not provide the expected outcomes compared to real-time data.

In conclusion, these studies have shown that machine learning algorithms can be an effective tool for air quality management and decision-making in cities. However, challenges such as multi-task learning, data sequencing, and accuracy must be addressed to improve the models' performance.

**Methodology**

**Existing System**

The current approach for predicting air quality is based on information gathered from air quality monitoring stations and conventional statistical models. This approach does not account for real-time changes in the environment that might alter air quality and instead predicts air quality using past data. The method cannot give a complete picture of the air quality over a region since it depends on a set number of air quality monitoring stations that are positioned in particular places.

**Disadvantages:**

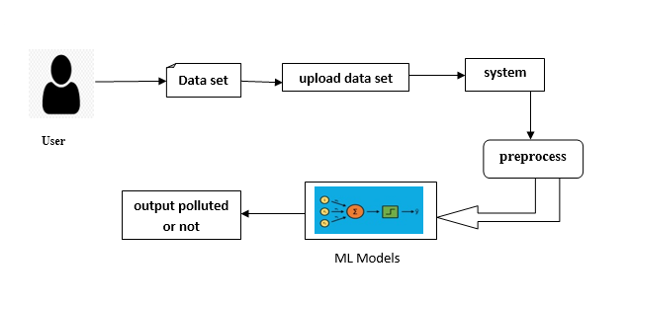
* They only provide a certain level of accuracy since they cannot foretell the maximum and minimum levels of pollution.
* They use inefficient approaches for better output prediction
* And complex mathematical calculations

**Proposed** **System**

Based on real-time data from several sources, the suggested system for air quality prediction employs machine learning techniques to forecast air quality. This system combines information from a variety of sources, including weather stations, satellite images, and air quality sensors, to give a complete picture of the air quality in an area. In order to anticipate air quality, the system uses machine learning techniques such as decision trees and neural networks, and random forests. These algorithms may be taught to take into account new data as it becomes available.

The suggested system can offer precise and current air quality forecasts that people, organisations, and politicians may utilise to make knowledgeable decisions about managing air quality. To give the general public access to real-time air quality information, the system may potentially be coupled with mobile applications and other technologies.

**Block Diagram:**

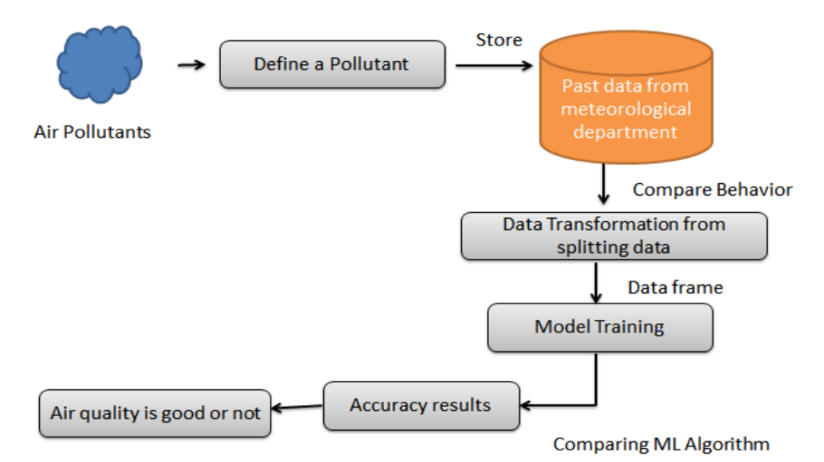


**Figure 1: Block Diagram**

**Advantages:**

* Easy to detect and predict pollution levels.
* An efficient approach for better output prediction.

**Architecture:**

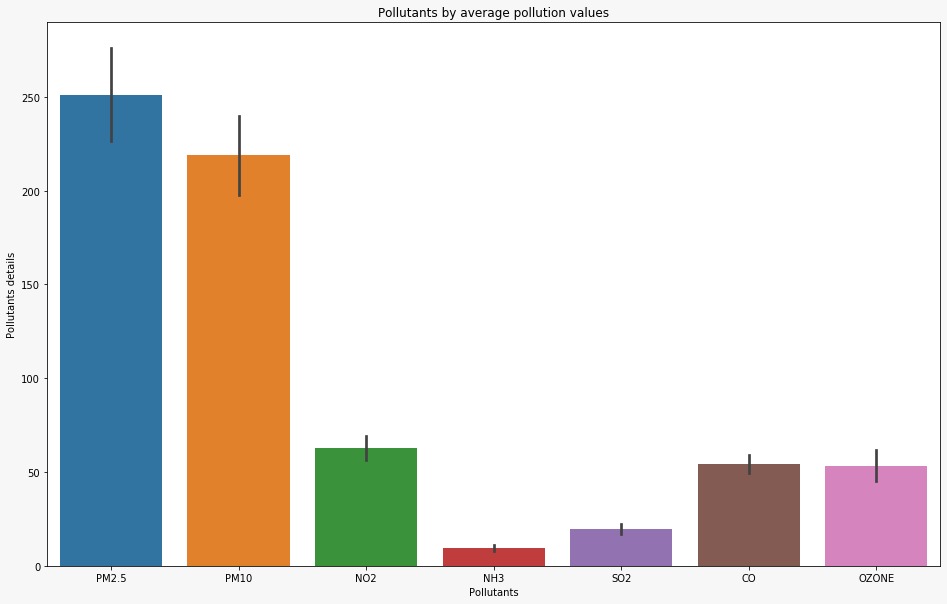


**Figure 2: Architecture of model**

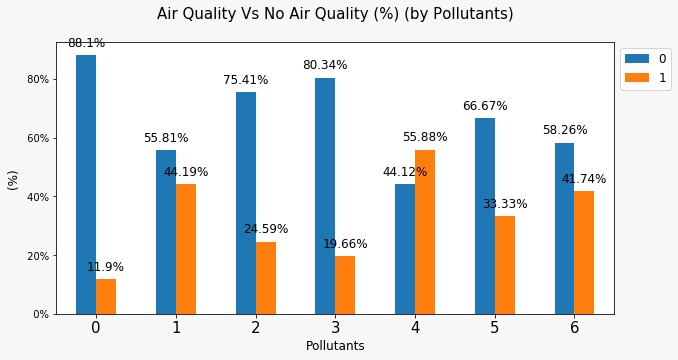
**Algorithm:**

1. Data Collection: Collect air quality data, including historical and real-time data, from various sources such as government monitoring stations, satellite images, and weather forecasts.
2. Data Preprocessing: To prepare the data for machine learning models, clean, preprocess, and modify it to eliminate noise, missing values, and outliers.
3. Feature Engineering: the preprocessed data's important characteristics should be extracted, including meteorological data, pollutant concentrations, and other relevant variables that could impact air quality.
4. Data Splitting: Divide the preprocessed data into training, validation, and testing datasets to train and test machine learning models.
5. Model Selection: Choose appropriate machine learning models, such as regression, decision tree, or neural network models, to predict air quality based on the extracted features.
6. Model Training: Train the selected machine learning models on the training dataset using appropriate algorithms and hyperparameters.
7. Model Evaluation: Use the validation dataset and metrics like mean squared error (MSE), mean absolute error (MAE), and coefficient of determination to assess the trained models (R-squared).
8. Model Tuning: Adjust the model hyperparameters and algorithms to improve model performance and generalization.
9. Model Testing: Test the final model on the testing dataset and evaluate its performance using similar metrics used in the evaluation step.
10. Deployment: Once the model is optimized and tested, deploy it for real-world use to predict air quality levels in real-time.

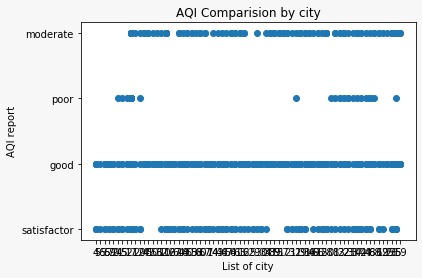
**Results and Discussions**



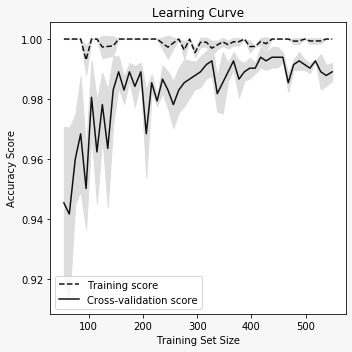
**Figure 3: Average pollution values**

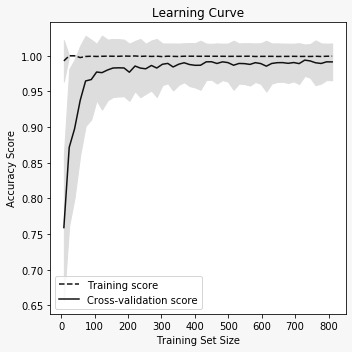


**Figure 4: Representing pollutants of air quality**



**Figure 5: AQI comparision by city**





**Figure 6.1&6.2: Accuracy score of training set**

**TABLE 1: Accuray of the proposed model**

|  |  |  |
| --- | --- | --- |
| **S.NO** | **ALGORITHMS** | **ACCURACY** |
| 1. | K Neighbors Classifier | 78.57% |
| 2. | Random Forest Classifier | 92.74% |
| 3. | Support Vector Machine | 76.92% |
| 4. | Decision Tree Classifier | 71.42% |
| 5. | Gassuain Naïve Bayes | 78.57% |
| 6. | Logistics Regression | 85.71% |

**Conclusion**

Based on the information provided, it appears that "An Immersive Model for Detecting Air Quality-Machine Learning Approach" is a research paper or a project that proposes a machine learning-based approach for detecting air quality. The approach involves using various sensors to collect air quality data and using machine-learning algorithm to analyze and predict air quality levels. While the details of the proposed approach are not provided, such an approach has the potential to improve air quality monitoring and management, as accurate and timely air quality predictions can help inform decisions about activities and policies that may affect air quality. However, as with any machine learning-based system, the accuracy and effectiveness of the approach will depend on the quality and quantity of data that the algorithms utilise to be trained and tested.

The proposed approach is an interesting and innovative idea that could have significant practical applications in improving air quality and public health. Further research and development will be needed to refine and optimize the approach for real-world implementation.

**References**

1. Pandey, Gaurav, Bin Zhang, and Le Jian. " Predicting sub-micron air pollution indicators: a machine learning approach.”; Environmental Science: Processes & amp; Impacts 15.5 (2013): 996-1005.
2. Bhushan, Shashi, Mohammed Alshehri, Ismail Keshta, Ashish Kumar Chakraverti, JitendraRajpurohit, and Ahed Abugabah. "An experimental analysis of various machine learning algorithms for hand gesture recognition." *Electronics* 11, no. 6 (2022): 968.
3. Chung, Hung-Yuan, Yao-Liang Chung, and Wei-Feng Tsai. "An efficient hand gesture recognition system based on deep CNN." In *2019 IEEE International Conference on Industrial Technology (ICIT)*, pp. 853-858. IEEE, 2019.
4. Cormoş, Angel Ciprian, Răzvan Andrei Gheorghiu, Valentin Alexandra STAN, and Ion Spirea Dănăilă. "Use of TensorFlow and OpenCV to detect vehicles." In *2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pp. 1-4. IEEE, 2020.
5. Ertam, Fatih, and Galip Aydın. "Data classification with deep learning using Tensorflow."

In *2017 international conference on computer science and engineering (UBMK)*, pp. 755-758. IEEE, 2017.

1. Hasan, Mokhar M., and Pramod K. Mishra. "Features fitting using multivariate Gaussian

distribution for hand gesture recognition." *International Journal of Computer Science & Emerging Technologies IJCSET* 3, no. 2 (2012): 73-80.

1. Hasan, Mokhtar M., and Pramod K. Mishra. "Robust gesture recognition using Gaussian

distribution for features fitting." *International Journal of Machine Learning and Computing* 2, no. 3 (2012): 266.

1. Hasan, Mokhtar M., and Pramod K. Mishra. "HSV brightness factor matching for gesture

recognition system." *International Journal of Image Processing (IJIP)* 4, no. 5 (2010): 456-467.

1. Hussain, Soeb, Rupal Saxena, Xie Han, Jameel Ahmed Khan, and Hyunchul Shin. "Hand

gesture recognition using deep learning." In the *2017 International SoC Design Conference*

*(ISOCC)*, pp. 48-49. IEEE, 2017.

1. Oudah, Munir, Ali Al-Naji, and Javaan Chahl. "Hand gesture recognition based on computer vision: a review of techniques." *journal of Imaging* 6, no. 8 (2020): 73.
2. Sharma, Sakshi, and Sukhwinder Singh. "Vision-based hand gesture recognition using deep learning for the interpretation of sign language." *Expert Systems with Applications* 182 (2021): 115657.
3. Strezoski, Gjorgji, Dario Stojanovski, Ivica Dimitrovski, and Gjorgji Madjarov. "Hand gesture recognition using deep convolutional neural networks." In *ICT Innovations 2016: Cognitive Functions and Next Generation ICT Systems*, pp. 49-58. Springer International Publishing, 2018.
4. Sun, Jing-Hao, Ting-Ting Ji, Shu-Bin Zhang, Jia-Kui Yang, and Guang-Rong Ji. "Research on

the hand gesture recognition based on deep learning." In *2018 12th International Symposium on antennas, propagation, and EM theory (ISAPE)*, pp. 1-4. IEEE, 2018.