

## **PREDICTION OF AIR POLLUTION BY USING MACHINE LEARNING ALGORITHM**

<sup>1</sup>N.TEJA SREE, <sup>2</sup>KARKALA NAVYA REDDY, <sup>3</sup>KOTAGIRI HARINI, <sup>4</sup>K.AKSHAYA

<sup>1</sup>Assistant Professor, Department of Information Technology, **MALLA REDDY ENGINEERING COLLEGE FOR WOMEN**, Maisammaguda, Dhulapally Kompally, Medchal Rd, M, Secunderabad, Telangana.

<sup>2,3,4</sup> Student, Department of Information Technology, **MALLA REDDY ENGINEERING COLLEGE FOR WOMEN**, Maisammaguda, Dhulapally Kompally, Medchal Rd, M, Secunderabad, Telangana.

### **ABSTRACT**

Controlling and defensive the higher air greatness has gotten one in everything about first imperative occasions in different creating and metropolitan districts at the present. The greatness of air is adversely contacting collectible to the different styles of tainting influenced through the transportation, power, powers consumptions, and so forth. In our country population is a big problem as day by day population is increasing, so the rapid increasing in population and economic upswing is leading environment problems in city like air pollution, water pollution etc. In some of air Pollution and air pollution is direct impact on human body. As we know that major pollutants are arising from Nitrogen Oxide, Carbon Monoxide & Particulate matter (PM), SO<sub>2</sub> etc. Carbon Monoxide is arising due to the deficient Oxidization of propellant like as petroleum, gas, etc. nitrogen oxide (NO) is arising due to the ignition of thermal fuel; Sculpture Dioxide (So<sub>2</sub>) is major spread in air, So<sub>2</sub> is a gas which is present more pollutants in air, it's affect more in human body. the predominance of air is overstated by multidimensional impacts containing spot, time and vague boundaries. The goal of this improvement is to take a gander at the AI basically based ways for air quality expectation. In this paper we will predict of air pollution by using machine learning algorithm.

### **LINTRODUCTION**

The Environment describe Pollution. If humidity is increasing more about the thing which is everything than automatically environment is going happening in encircles the Environment hotter. Major cause of increasing is polluted by human daily activities pollution is increasing day by day which include like air pollution, noise transport and industries there are 75 %

NO or other gas like CO, SO<sub>2</sub> and other particle is exist in environment.. The expanding scene, vehicles and creations square measure harming all the air at a feared rate.

Therefore, we have taken some attributes data like vehicles no., Pollutants attributes for prediction of pollution in specific zone of Delhi.

## II.LITERATURE SURVEY

The prediction of air pollution using machine learning (ML) algorithms has garnered considerable attention due to its potential to improve environmental management and public health. Goyal et al. (2006) introduced Key-Policy Attribute-Based Encryption (KP-ABE), a cryptosystem designed for fine-grained sharing of encrypted data, which lays foundational principles for data privacy in cloud computing [1]. This concept is pivotal when considering the secure handling of air quality data. Building on this, Chase and Chow (2009) advanced the field with improvements to attribute-based encryption by addressing privacy concerns associated with multi-authority systems, thus enhancing the secure sharing of environmental data [2]. Lin et al. (2010) further developed a threshold multi-authority attribute-based encryption scheme, eliminating the need

for a central authority and thereby increasing system robustness and privacy for data handling in various applications, including environmental data management [3].

In the context of air quality prediction, integrating diverse data sources and advanced techniques has been explored. Božović et al. (2011) proposed a multi-authority attribute-based encryption system with an honest-but-curious central authority, which offers insights into how to handle sensitive data with privacy considerations—a concept crucial for handling air quality information securely [4]. Hur (2014) extended this by addressing secure data sharing in smart grids, highlighting the need for privacy-preserving mechanisms when sharing sensitive data, such as air quality measurements, which can be adapted for pollution prediction systems [5].

## III.EXISTING SYSTEM

The Air Pollution Forecasting System: Air Quality Index (AQI) is a record that gives the public the degree of contamination related with its wellbeing impacts. The AQI centers on the different wellbeing impacts that individuals may encounter dependent fair and square and long stretches of

introduction to the poison concentration. The AQI values are not quite the same as nation to nation dependent on the air quality norm of the country.

The higher the AQI level more noteworthy is the danger of wellbeing related problems. The by and large point of this venture is to make a student calculation that will have the option to foresee the hourly contamination focus. Additionally, an Android application will be built up that will provide the clients about the constant contamination convergence of PM<sub>2.5</sub> alongside the hourly forecasted value of the toxin fixation from the student calculation. The Android application will also recommend data of the less dirtied[1].

#### **Disadvantages**

- ❖ The system is not implemented Stepwise Multiple Linear Regression Method.
- ❖ The system is not implemented Instance-Linear Regression Model

#### **IV.PROPOSED SYSTEM**

1) Data assortment: There is a different method from which we collected data from various dependable sources like Delhi Gov. Site.

2) Exploratory examination: We research and explore examination with various parameter like ID of outliers,

consistency check, missing qualities, and so on, it's totally occurred in this period of the venture.

3) Data Manipulation control: In period of data control stage the required missing data need to insert in utilizing the mean estimations of that characteristic of information. [2]

4) Prediction of boundaries utilizing by gauge model: For appropriate data indirect relapse we have to keep future qualities for different boundaries just

5) Implementation of straight relapse: Whenever all the boundaries become in active mode or they are accessible mode, the direct relapse calculation would be used in anticipate the air quality index (AQI).

6) Data accuracy investigation: We have to analyze that used model is being fit for overall data or not so we have to cross check root mean error, absolute percentage error then after we have to assume this factor is good for accuracy or not

#### **Advantages**

- The proposed system implemented Linear Regression is basically use for predicting the real values data  $y$  using continuous parameter.
- Stepwise Multiple Linear Regression Method is used for

continuous data testing and training in effective way

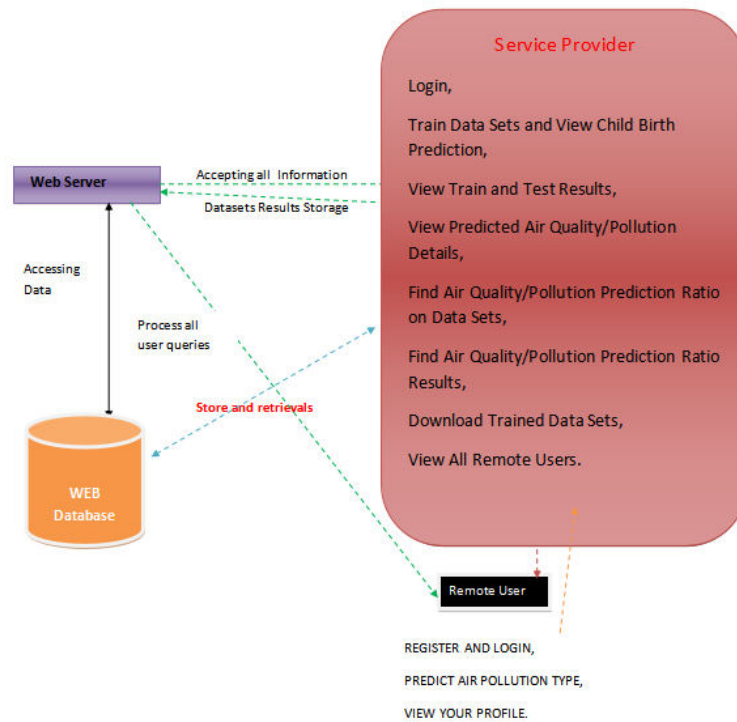
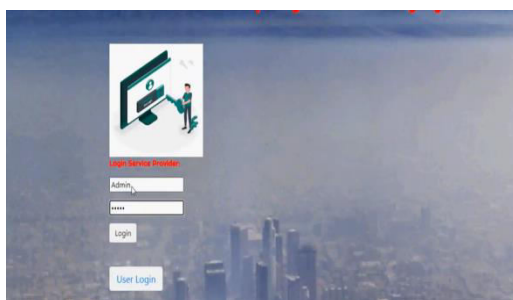


Fig1: System Architecture

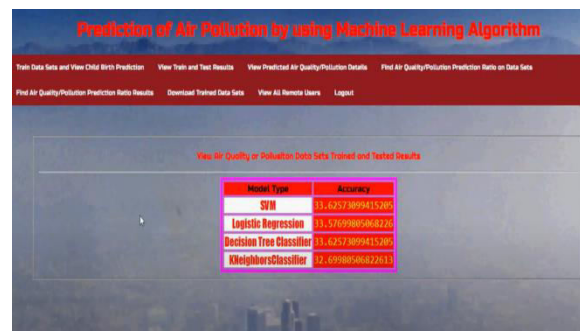
## V. MODULES

### Service Provider

In this module, the Service Provider has to login by using valid user name and password.



After login successful he can do some operations such as Login, Train Data Sets and View Child Birth Prediction, View Train and Test Results,



View Predicted Air Quality/Pollution Details, Find Air Quality/Pollution Prediction Ratio on Data Sets, Find Air Quality/Pollution Prediction Ratio

Results, Download Trained Data Sets, And View All Remote Users.

### **View and Authorize Users**

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

### **Remote User**

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like register and login, predict air pollution type, view your profile.

## **VI.CONCLUSION**

Precision of our model is very acceptable. The anticipated AQI has a precision of 96%. Future upgrades incorporate expanding the extent of district and to incorporate whatever number locales as could be allowed as of now this venture targets foreseeing the AQI estimations of various areas of close by New Delhi. Further, by utilizing information of various urban areas the extent of this venture can be

exhausted to anticipate AQI for different urban communities also. The integration of machine learning algorithms into air pollution prediction systems represents a significant advancement in environmental management and public health. By leveraging ML techniques, researchers can enhance the accuracy and reliability of air quality forecasts, providing valuable insights for mitigating the impacts of pollution on human health and the environment. The reviewed literature highlights the importance of secure data handling and privacy-preserving mechanisms, emphasizing the need for robust encryption and multi-authority systems to protect sensitive environmental data.

The application of advanced ML algorithms, such as deep learning models, combined with secure data management practices, offers a promising approach to addressing the challenges of air pollution forecasting. Future research should focus on integrating diverse data sources, such as satellite imagery and IoT sensors, to improve prediction accuracy and spatial resolution. Additionally, advancements in real-time decision support systems and personalized health recommendations can further enhance

the effectiveness of pollution management strategies.

## VII. REFERENCES

- [1] Ni, X.Y.; Huang, H.; Du, W.P. "Relevance analysis and short-term prediction of PM 2.5 concentrations in Beijing based on multi-source data." *Atmos. Environ.* 2017, 150, 146-161.
- [2] G. Corani and M. Scanagatta, "Air pollution prediction via multi-label classification," *Environ. Model. Softw.*, vol. 80, pp. 259-264, 2016.
- [3] Mrs. A. GnanaSoundariMtech, (Phd), Mrs. J. GnanaJeslin M.E, (Phd), Akshaya A.C. "Indian Air Quality Prediction And Analysis Using Machine Learning". *International Journal of Applied Engineering Research* ISSN 0973-4562 Volume 14, Number 11, 2019 (Special Issue).
- [4] Suhasini V. Kottur , Dr. S. S. Mantha. "An Integrated Model Using Artificial Neural Network
- [5] RuchiRaturi, Dr. J.R. Prasad . "Recognition Of Future Air Quality Index Using Artificial Neural Network". *International Research Journal of Engineering and Technology (IRJET)* .e ISSN: 2395-0056 p-ISSN: 2395-0072 Volume: 05 Issue: 03 Mar-2018
- [6] Aditya C R, Chandana R Deshmukh, Nayana D K, Praveen Gandhi Vidyavastu ." Detection and Prediction of Air Pollution using Machine Learning Models". *International Journal o f Engineering Trends and Technology (IJETT)* - volume 59 Issue 4 - May 2018
- [7] Gaganjot Kaur Kang, Jerry ZeyuGao, Sen Chiao, Shengqiang Lu, and Gang Xie." *Air Quality Prediction: Big Data and Machine Learning Approaches*". *International Journal o f Environmental Science and Development*, Vol. 9, No. 1, January 2018
- [8] PING-WEI SOH, JIA-WEI CHANG, AND JEN-WEI HUANG," Adaptive Deep Learning Based Air Quality Prediction Model Using the Most Relevant Spatial-Temporal Relations," *IEEE ACCESS* July 30, 2018. Digital Object Identifier 10.1109/ACCESS.2018.2849820.
- [9] GaganjotKaur Kang, Jerry Zeyu Gao, Sen Chiao, Shengqiang Lu, and Gang Xie," *Air Quality Prediction: Big Data and Machine Learning Approaches*," *International Journal of Environmental Science and Development*, Vol. 9, No. 1, January 2018.
- [10] Haripriya Ayyalasomayajula, Edgar Gabriel, Peggy Lindner and Daniel

Price, "Air Quality Simulations using Big Data Programming Models," IEEE Second International Conference on Big Data Computing Service and Applications, 2016.

She received B.Tech Degree (INFORMATION TECHNOLOGY) from RRS college of engineering technology in 2014. She received MASTER OF ENGINEERING (COMPUTER SCIENCE) Degree from Mallareddy Engineering college for Women in 2016. She is having Good Academic and Experience of more than

4 years. Now she is associated with MRECW.

Her current area of research includes Machine Learning and Network Technology. She is having 4 papers in reputed International Journals. She Attended Various Workshops and Faculty Development Programs.

