International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 06 | June - 2021 ISSN: 2582-3930

Vaayu for Aayu (AQI) – A Study to Assess the Air Quality

Samraddh Saxena[1], Suyash Gupta[2], Sundram Gupta[3]

[1,2,3] *Computer Science Engineering, Acropolis Institute of Technology and Research, Indore, India*

---***---

Abstract – *Air pollution has become an important environmental issue in recent decades. Forecasts of air quality play an important role in warning people about and controlling air pollution. Air quality index are commonly used to indicate the level of severity of air pollution to the public. It is infeasible and perhaps impossible to formulate a universal technique for determining air quality index, one that considers all pollutants and that is appropriate for all situations. The intended use of the air quality index is to identify the vulnerable zone. Every index has its own characteristic strengths and weaknesses that affect its suitability for particular applications. In this paper, we will be assessing the quality of air by using machine learning as our prime ingredient. Machine learning is an area where system which implements artificial intelligence gathers data from an environment and learns how to act. One of the reason why we choose machine learning to predict air quality index, was this ability of adapting of machine learning (ML) algorithms. Here, we will be using various regression models to predict about the quality of air by visualizing a report (graphs) and a value as an end product. Also with the help of the AQI value the end user will be in a position to judge about the quality of air.*

Index Terms—Air Quality Index.

1. INTRODUCTION

Apart from land and water, air is the prime resource for the sustenance of life. Air is an integral and essential necessity in everyday life. Whether it is agriculture, or pollination of various crops, or even basic survival of numerous living species, everything everywhere is dependent on air. The importance of air cannot be overemphasized and hence the rising level of air pollution is a matter of serious concern.

Air pollution is nowadays a paramount concern worldwide. Because of this many people facing the problems generated by air pollution, like, Lung Diseases and Respiratory Problems, Cardiovascular Problem, chronic obstructive pulmonary disease (COPD), etc. Real-time air quality monitoring is a major concern in today's Internet era. Further, Air is an essential element for all the living species. Air Quality Index (AQI), is a value that tells about air quality. The Air Quality Index is an unit-less quantity, that determines whether the air quality is good or not for living beings. AQI can be calculated according to the required frequency, like, hourly AQI, daily AQI, weekly AQI, monthly AQI, or yearly AQI. Among all the previously stated frequencies, daily AQI is most widely used measurement of air quality monitoring. Real-time AQI monitoring of a city, make the people more concern about the environment, in which they are living, and thus make them avoid or minimize the use of pollutants for air, like burning of coal, waste, etc. Not only the outdoor environment, but indoor climate status produce a great impact on overall health of the human being.

Air Pollution is the inadequate change in physical, chemical or biological characteristics of air which hampers life as well as leads to potential health problems. Air pollution majorly affects the eyes, lungs, nose, and throat by causing irritation. It also creates respiratory problems and exacerbates existing conditions such as asthma and emphysema. The risk of cardiovascular diseases become much higher when humans are continually exposed to air pollution. In India, air pollution is the third highest cause of death among health risks and because of this, life expectancy has gone down by 2.6 years. Hence, it has become increasingly necessary to not only control the contamination but also enlighten the people affecting the quality of air effectively in a bid to maintain a healthy standard.

A reasonable way to analyze the amount of pollution is by determining the standard of air. With technological advancements, a vast amount of data on ambient air quality is generated which is used to establish the quality of air in different areas. The large monitoring data results have astronomical volumes of information that neither provides any useful insights to a decision-maker nor is intelligible to a common man who simply wants to understand how good or bad the air is. One way to describe air quality is to report the concentrations of all pollutants with acceptable levels (standards).

Air quality index (AQI) is used worldwide to inform the public about levels of air pollution (degradation or improvement) and associated to different biological effects. Due to pollution, the ambient air quality in major cities in India is very poor. According to WHO surveys, India is one of the most polluted countries in the world. Concentrations of air pollutants affect Air Quality Index. The higher the AQI value, the greater the level of air pollution and greater the health concern. Thus, we need a strong managerial system which can predict the adversity of air in advance and thus we can take immediate actions on it.

The main focus of this paper is learning about modeling of data by supervised algorithms e.g. (Linear Regression (regression), Logistic Regression (classification) etc. The main focus of this particular kernel is AQI (Air Quality Index), and factors that affects AQI. Since in the dataset we have the concentration of pollutants and we need each pollutants index for calculating the air quality index, so that has been calculated further in the process and has been utilized in analysis . Also in the following paper there is a brief explanation of how combination of the independent variables (Interaction effect) has what impact on dependent variables.

2. LITERATURE SURVEY

Huixiang Liu (et al.2019) have taken two different cities Beijing and Italian city for the study purpose. They have forecasted the Air Quality Index (AQI) for the city Beijing and predicting the concentration of NOx in an Italian City depending on two different publicly available datasets. The first Dataset for the period of December 2013 to August 2018 having 1738 instances is made available from the Beijing Municipal Environmental Centre [5] which contains the fields like hourly averaged AQI and the concentrations of PM2.5, O3, SO2, PM10, and NO2 in Beijing. The second Dataset with 9358 instances is collected from Italian city for the period of March 2004 to February 2005. This dataset contains the attributes as Hourly averaged concentration of CO, Non methane Hydrocarbons, Benzene, NOx, NO2 [5]. But they focused majorly on NOx prediction as it is one of the important predictor for Air Quality evaluation.

Nidhi Sharma (et al.2018) had gone through the detailed data analysis of air pollutants from 2009-2017 and also proposed the critical observation of 2016-1017 air pollutants trend in Delhi, India [14]. They have predicted the future trends of various pollutants as Sulfur Dioxide (SO2), Nitrogen Dioxide (NO2), Suspended Particulate Matter (PM), Ozone (O3), Carbon Monoxide (CO) and Benzene. By using data analytics Time series Regression forecasting they have predicted the future values of the pollutants mentioned earlier on the of previous records. According to this study results the Anand Vihar and Shadipur monitoring stations of Delhi are under the study. The result shows that there is a drastic increase in PM10 concentration level, NO2 and PM2.5 are evidently increased showing the increased pollution in Delhi [14]. CO is predicted to reduce by 0.169mg/m3 , there is increase in NO2 concentration level for coming years by 16.77 µg/m3 , Ozone is predicted to increase by 6.11mg/m3 , Benzene reduce by 1.33mg/m3 and SO2 is forecasted to increase by 1.24µg/m.

Ziyue Guan and Richard O. Sinnot (2018) used the various machine learning algorithms to predict the PM2.5 concentration. Data were collected from the official website of Environment Protection Agency (EPA) for the city Melbourne that contains PM2.5 air parameter and they have also collected the unofficial data from Air beam which is the mobile device used to measure PM2.5 value [8]. The machine Learning Algorithms Artificial Neural Network (ANN), and Long Short Term Memory (LSTM) recurrent neural network were used for the PM2.5 prediction but out of these algorithms LSTM gives the best performance ad predict the high PM2.5 value with reasonable Accuracy.

S.Tikhe Shruti (et al.2013) the research employed two soft computing algorithms Artificial Neural Network (ANN) and Genetic Programming (GP) for the prediction of future concentration levels of air pollutants such as Oxides of Sulfur (SOx), Oxides of Nitrogen (NOx) and Respirable Suspended Particulate Matter (RSPM) over the year 2005 to 2011 for Pune city in Maharashtra which is at the second position I list of polluted cities in India. They have developed total six models (three of each algorithm ANN and GP) based on hourly average data values of pollutants concentration spanning greater than 7 years. Out of these two algorithms GP algorithms gives the better performance than ANN.

Mohamed Shakir and N.Rakesh (2018) have analyzed the proportion of various air pollutants (NO, NO2, CO, PM10 and SO2) with respect to the time of the day and the day of the week and estimated the effect of environmental parameters as temperature, wind speed and humidity on the air pollutants mentioned above with the help of WEKA tool [15]. The data was collected from pollution control board of Karnataka. By using ZeroR algorithm in WEKA tool the study come up with the results that shows that the concentration levels of air pollutants increase during the working days and especially during the peak hours of the day and decrease during week-ends or holidays [15]. Using Simple K-means Clustering algorithms the study shows the relationship or dependencies between the environmental factors like Temperature, wind speed and humidity and the air pollutants like NO, NO2, PM10, CO and SO2.

Yusef Omidi Khaniabadi (et al.2016) the main aim of this study is to discover the relation or association between health impacts such as mortality rate of cardiovascular diseases and the air pollutants as PM10, NO2 and O3 over the years 2014 and 2015 for Kermanshah city in Iran. They used the AirQ software proposed by WHO for this purpose. The number of premature deaths for cardiovascular diseases is of 188 related to PM10, 33 related to NO2 and 83 related to O3[17].The results of this study indicates that if there is 10μ/m3 increase in PM10, NO2 and O3 concentration level the mortality risk will increase by 1.066, 1.012 and 1.020 respectively.

R. Gunasekaran (et al.2012) the main objective of this study is to monitor the air quality of Salem Swadeswari College, Tamil Nadu area for the period of April 2011 to March 2011 and it has been shown that this area has no serious pollution issues related to the pollutants as Sulfur Dioxide, Oxides of Nitrogen and Suspended Particulate Matter because their annual average concentration are within the range of national standards. But the annual average concentration of the pollutant PM10 is slightly higher than the levels of national standard. Also the monthly 24-hour average concentration of PM10 in the same year were crossed the national standard level except during July to October [18].

Kazem Naddaf (et al.2012) used the AirQ software proposed by WHO that provides the quantitative data on the impact of PM10, SO2, NO2 and O3 on the health of the people in Tehran City of Iran which is the most populated city in Iran [16]. Health impacts under the consideration were all cause mortality, cardiovascular diseases and the respiratory diseases. "The results of the study shows that the air pollutant PM10 had the highest heath impact on the 8,700,000 inhabitants of Tehran City and also caused an excess of total mortality of 2194 deaths out of 47284 in a year" [16]. The total number of excess cases of mortality due to SO2, NO2 and Ozone are 1458, 1050 and 819 respectively. These results shows that Tehran suffered from critical problem of air pollution and for Tehran there is a need to reduce the health burden of air pollution.

3. METHODOLOGY

Nowadays, general citizens have become more aware of the surroundings. They are well informed about their adverse health effects of poor air quality so it is important for people to know about the air they breathe. This information can be apprehended by society by assessing the daily air pollution levels. AQI is one such tool to canvass the air we breathe which is used to calculate the overall results based on standards and policies followed by the country. In India, these standards are set by the Central Board of Pollution Control (CPCB) under the law Air (Prevention and Control of Pollution) Act, 1981[4].

Now a day, it is important to the society to look for Awareness of daily levels of air pollution. AQI is a tool which is used to report the overall air quality status and trends based on a specific standard. In India we are using CPCB Standard for calculating air quality index or environment pollution index. This index gives an idea about the environmental status as air quality. And also tells the general public to understand how clean or pollute air is breathe daily.

Overall this index can be used to give meaningful evaluation of air pollution to the common man. It also helps to identify the air pollution control policies or control equipment can reduce level of dominating pollutant. AQI is representing the cumulative effect of all the pollutant to show overall air quality status in better way. The AQI of specific pollutant is derived mainly from the physical measurement of pollutant like PM10, PM2.5, NO2 and SO2 etc. In the present study, six different methods were used to calculate ambient air quality index.

Method I:

Air quality Index (AQI) is calculated based on the arithmetic mean of the ratio of concentration of pollutants to the standard value of that pollutant such as PM10, PM2.5, NO2 and SO2. The average is then multiplied by 100 to get the AQI index. AQI was then compared with rating scale (Kaushik et al., 2006). For individual pollutant AQI was calculated by the following formula

$$AQI = (C/C_s) * 100$$

Where,
AQI = Air Quality Index C= the observed value of the air quality parameters pollutant (PM10, PM2.5, NO2 and SO2) Cs= CPCB standard for residential Area (CPCB, 2009).

Method II:

In this procedure AQI is calculated by taking the geometric mean of the ratio of concentration of pollutants to the standard value of that pollutant such as PM10, PM2.5, NO2 and SO2. AQI was then compared with rating scale. (Ravikumar et. al., 2014).

Method III:

Oak Ridge National Air Quality Index (ORNAQI) is used for the relative ranking of an overall air quality status. Over all AQI was estimated by the following mathematical equation developed by the Oak Ridge National Laboratory (ORNL), USA is given below.

$$AQI = [39.02 \sum C/C_s]^{0.967}$$

Air quality Index then measured and compared with relative ORAQI values (Bhuyan et al. 2010).

Method IV:

Air Quality Index was done for combining qualitative measures with qualitative concept of the environment. The individual air quality index here is calculated as follow:

$$AQI = (W * C/ C_s)$$

Where,
AQI = Air Quality Index
W= Weighted of Pollutant
C= the observed value of the air quality parameters pollutant (PM10, PM2.5, NO2 and SO2)
Cs= CPCB standard for residential Area (CPCB, 2009)

Method V:

Air Quality Index was done based on dose response relationships of pollutants to obtain break point concentration.(USEPA, 2006,CPCB 2014) The individual air quality index for a given pollutant concentration (Cs) as based on linear segmented principle.

4. RESULT AND DISCUSSION

Based on above methodologies we can determine the Air quality index value and that value can be used to accumulate the amount of quality of any particular area based on the standard AQI chart which helps to understand the value of AQI. The value can be understand as:

a. Good (0–50) - Minimal Impact

b. Satisfactory (51–100) - May cause minor breathing difficulties in sensitive people.

c. Moderately polluted (101–200) - May cause breathing difficulties in people with lung disease like asthma, and discomfort to people with heart disease, children and older adults.

d. Poor (201–300) - May cause breathing difficulties in people on prolonged exposure, and discomfort to people with heart disease

e. Very Poor (301–400) - May cause respiratory illness in people on prolonged exposure. Effect may be more pronounced in people with lung and heart diseases.

f. Severe (401-500) - May cause respiratory issues in healthy people, and serious health issues in people with lung/heart disease. Difficulties may be experienced even during light physical activity.

Based on this values, we can understand the adversity of air and can take necessary actions against it in advance.

All of this may sound complicated, but when broken down it really is quite simple. Basically, when the air quality is good the AQI is low (zero 50) and the color associated with it is greenv(Based on below given chart). As the air quality gets worse the number gets higher and the color associated with it becomes a darker shade of red.

Air Quality Index (AQI) Values	Levels of Health Concern	Colors
When the AQI is in this range:	*...air quality conditions are:*	*...as symbolized by this color:*
0-50	Good	Green
51-100	Moderate	Yellow
101-150	Unhealthy for Sensitive Groups	Orange
151 to 200	Unhealthy	Red
201 to 300	Very Unhealthy	Purple
301 to 500	Hazardous	Maroon

5. CONCLUSION

Accurate air quality forecasting has important theoretical and practical value for the public; without it, neither the government nor the public can effectively avoid the health damage caused by air pollution or improve the emergency response capability of heavy pollution days. In this study, we built regression models to predict air indicators based on various methodologies and algorithms.

Air quality Index can give clear view about ambient air and critical pollutant mainly responsible for the quality of air. Predicting the air quality is a complex task due to the dynamic nature, volatility, and high variability in space and time of pollutants and particulates. At the same time, being able to model, predict, and monitor air quality is becoming more and more important, especially in urban areas, due to the observed critical impacts of air pollution for populations and the environment. In this paper we presented various methodologies to study and correctly identify the AQI value. Thus, with the help of them we can get value of AQI and based on standard chart for AQI we can easily identify the adversity of an air and can take necessary immediate actions against it to protect our environment against air pollution.

6. ACKNOWLEDGEMENT

We would like to extend our gratitude towards our mentor Mrs. Juhi Srivastava , Faculty at Department of Computer Science, Acropolis Institute of Technology and Research. She provided great insights and expertise to this research and was also an integral part of this research without whom this would have not been possible. We would also like to thank our families and friends who supported us in the course of writing this paper. Their contributions to our efforts are genuinely appreciated.

7. REFERENCES

1. Kanchan Prasad, Amit Gorai, and Pramila Goyal "A review on air quality indexing system", Asian Journal of Atmospheric Environment, Vol. 9-2, pp. 101-113, June 2015.

2. Central Pollution Control Board, Government of India, Ministry of Environment, Forest and Climate Change "National Air Quality Index", 2015 [Online] Available: https://www.cpcb.nic.in/

3. Air Now, "Air Quality Index Basics", June 2019. [Online],Available: https://www.airnow.gov.in

4. U. S. Environmental Protection Agency (USEPA). Guidelines for reporting of daily air quality- air quality index (AQI), Series EPA-454/B-06-001. Research Trangle Park , North Carolina. 2006.

5. Central Pollution Control Board (CPCB).Guidelines for National ambient air

quality monitoring, Series:NAAQM/25/2003-04.Parivesh Bhavan,Delhi; 2009.

6. Shivangi Nigam, B. P. S. Rao, Navneet Kumar,and V. A. Mhaisalkar, "Air Quality Index – A Comparative Study for Assessing the Status of Air Quality", Research Journal of Engineering and Technology, pp. 267-274, 2016.

7. R. Gunasekaran, K. Kumaraswamy, P.P. Chandrasekaran, R. Elanchezhian, "MONITORING OF AMBIENT AIR QUALITY IN SALEM CITY, TAMIL NADU", International Journal of Current Research, ISSN: 0975-833X, Vol. 4, Issue, 03, pp.275- 280, March, 2012.

8. S.Tikhe Shruti, K.C. Khare, S.N. Londhe, "Forecasting Criteria Air Pollutants Using Data Driven Approaches: An Indian Case Study", International Journal of Soft Computing 8 (4), 305-312, 2013, ISSN: 1816-9503.

9. Yusef Omidi Khaniabadi, Gholamreza Goudarzi, Seyed Mohammad Daryanoosh, Alessandro Borgini, Andrea Tittarelli, Alessandra De Marco, "Exposure to PM10, NO2, and O3 and impacts on human health", Environ SciPollut Res, 2016.

10. . Mohamed Shakir, N. Rakesh, "Investigation on Air Pollutant Data Sets using Data Mining Tool", IEEE Xplore Part Number:CFP18OZV-ART; ISBN:978-1- 5386-1442-6.

11. Nidhi Sharma , ShwetaTaneja , VaishaliSagar , Arshita Bhatt, "Forecasting air pollution load in Delhi using data analysis tools", ScienceDirect, 132 (2018) 1077– 1085.

12. .Pallavi Pant, Raj M. Lal, Sarath K. Guttikunda, Armistead G. Russell, Ajay S. Nagpure, AnuRamaswami, Richard E. Peltie, "Monitoring particulate matter in India: recent trends and future outlook", Air Quality, Atmosphere & Health, 2018.

13. WHO (2018) Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: global update. Summary of risk assessment. Geneva, World Health Organization.

14. U.S. Environmental Protection Agency (1999) Guideline for reporting of daily air quality - air quality index (AQI). EPA-454/R-99-010. Office of Air Quality Planning and Standards, Research Triangle Park, NC 27711.

15. WHO (1999) Monitoring ambient air quality for health impact assessment. Copenhagen, World Health Organization Regional Office for Europe, WHO regional publication, European Series, No. 85. Available online at http://www.euro.who.int/__data/assets/pdf_file/ 0010/119674/E67902.pdf (Last accessed on 12th September, 2014).