PREDICTION Of Air Quality Index Using Machine Learning

A Mini Project Synopsis
Submitted in partial fulfilment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

in

INFORMATION TECHNOLOGY

by

Rohan Khilnani   **Sayyam Jain**   **Vasu Bansal**

00411503119   **02311503119**   **35311503119**

Guided by

Mr. Puneet Singh Lamba

DEPARTMENT OF INFORMATION TECHNOLOGY
BHARATI VIDYAPEETH'S COLLEGE OF ENGINEERING,
PASCHIM VIHAR, DELHI – 110063
(AFFILIATED TO GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, DELHI)
NOVEMBER, 2020

Introduction

The Nature is nothing but everything that encircles us. The environment is getting polluted due to human activities and one of the most common is air pollution. The concentration of air pollutants in ambient air is governed by the meteorological parameters such as atmospheric wind speed, wind direction, relative humidity, and temperature. The major pollutants are Nitrogen Oxide (NO), Carbon Monoxide (CO), Particulate matter (PM), SO2 etc. Carbon Monoxide is produced due to the deficient Oxidization of propellant such as petroleum, gas, etc. Nitrogen Oxide is produced due to the ignition of thermal fuel; Carbon monoxide causes headaches, vomiting; Benzene is produced due to smoking, it causes respiratory problems; Nitrogen oxides causes dizziness, nausea; Particulate matter with a diameter 2.5 micrometre or less than that affects more to human health. Air Quality Index (AQI), is used to measure the quality of air. Earlier it was difficult and unprecise to know the quality of air and was based on predictions but now due to advancement of technology, it is easy to fetch the data about the pollutants of air using sensors. Assessment of raw data to detect the pollutants needs vigorous analysis. Convolution Neural networks, Recursive Neural networks, Deep Learning, Machine learning algorithms assures in accomplishing the prediction of future AQI so that measures can be taken appropriately.

Machine learning which comes under artificial intelligence has three kinds of learning algorithms, they are the Supervised Learning, Unsupervised learning, Reinforcement learning. In the propose work we will use these all algorithms to bring more accuracy like supervised learning algorithms such as Linear Regression, Nearest Neighbour, SVM, kernel SVM, Naive Bayes and Random Forest.

Literature Survey

Several studies, experiments, and researches have been approved over the years to get precise and accurate results for prediction of air quality index using Machine Learning algorithms., here are the details of some foremost research papers, thoroughly examined.

Ishan et.al [1] described the benefits of the Bidirectional Long - Short Memory [BiLSTM] method to forecast the severity of air pollution. The proposed technique achieved better prediction which models the long term, short term, and critical consequence of PM2.5 severity levels. In the proposed method prediction is made at 6h, 12h, 24h. The results obtained for 12h is consistent, but the result obtained for 6h, and 24h are not consistent. Chao Zhang et.al

[2] proposed web service methodology to predict air quality. They provided service to the mobile device, the user to send photos of air pollution. The proposed method includes 2 modules a) GPS location data to retrieve the assessment of the quality of the air from nearby air quality stations. b) they have applied dictionary learning and convolution neural network on the photos uploaded by the user to predict the air quality. The proposed methodology has less error rate compared to other algorithms such as PAPLE, DL, PCALL but this method has a disadvantage in learning stability due to this the results are less accurate. Ruijun Yang et.al

[3] used the Bias network to find out the air quality and formed DAG from the data set of the town called as shanghai. The dataset is dived for the training and testing model. The disadvantage of this approach is they have not considered geographical and social environment characteristics, so the results may vary based on these factors. Madhuri VM et.al

[4] The concentration of air pollutants in ambient air is governed by the meteorological parameters such as atmospheric wind speed, wind direction, relative humidity, and temperature. Air Quality Index (AQI), is used to measure the quality of air. The proposed work is a supervised learning approach using different algorithms such as LR, SVM, DT and RF. The result show that AQI prediction obtained through RF are promising which are analysed with results.

Objectives of the Project

Objective of the project is to measure air quality index accurately, by predicting the air quality index, we can backtrack the major pollution causing pollutant and the location affected seriously by the pollutant. With the Machine Learning based forecasting model, various knowledge about the data is extracted using various techniques to obtain heavily affected regions on a particular

region(cluster). This gives more information and knowledge about the cause and seniority of the pollutants.

The air quality index is needed to provide a metric for warning citizens about the dangers of air pollution at varying levels of intensity

AQI informs the public about environmental conditions. It is especially useful for people suffering from illnesses aggravated or caused by air pollution. If AQI is not taken seriously it can cause damage to nature and human life. Some of the risk associated with High AQI to humans is lung diseases, such as asthma, chronic bronchitis, and emphysema.

Table 1: Air pollution attributes (values are expressed in $\mu g/m^3$, and CO is expressed in mg/m^3)

Attributes	Values	Meaning
SO2	1,2,3,4,5	Sulfur dioxide-(0-50,50-100,100-350,350-500,500+)
NO2	1,2,3,4,5	Nitrogen dioxide-(0-50,50-100,100-200,200-400,400+)
O3	1,2,3,4,5	Ozone-(0-60,60-120,120-180,180-240,240+)
CO	1,2,3,4,5	Carbon monoxide-(0-5,5-7.5,7.5-10,10-20,20+)
PM 2.5	1,2,3,4,5	Fine particles-(0-15,15-30,30-55,55-110,110+)
PM 10	1,2,3,4,5	Large particles-(0-25,25-50,50-90,90-180,180+)
AQI	0,1,2	Air Quality Index- (Low, Medium, High)

References

[1] Verma, Ishan, Rahul Ahuja, HardikMeisheri, andLipikaDey. "Air pollutant severity rediction using Bi-directional LSTM Network." In 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI), pp. 651-654. IEEE, 2018.

[2] Figures Zhang, Chao, Baoxian Liu, Junchi Yan, Jinghai Yan, Lingjun Li, Dawei Zhang, XiaoguangRui, and RongfangBie. "Hybrid Measurement of Air Quality as a 5 Fig. 8. RH w.r.t tin oxide Fig. 9. RH w.r.t C6H6 Mobile Service: An Image Based Approach." In 2017 IEEE International Conference on Web Services (ICWS), pp. 853- 856. IEEE,2017.

[3] Yang, Ruijun, Feng Yan, and Nan Zhao. "Urban air quality based on Bayesian network." In 2017 IEEE 9th Fig. 10. RH w.r.t NO Fig. 11. RH w.r.t NO2 International Conference on Communication Softwareand Networks (ICCSN),

pp. 1003-1006. IEEE,2017.

[4] Madhuri VM, Samyama Gunjal GH, Savitha Kamalapurkar "Air Pollution Prediction Using Machine Learning Supervised Learning Approach" In 2020 international journal of scientific and technology research volume 9 pp. ISSN 2277-8616

[5] Kumar, Anikender and P. Goyal, "Forcasting of daily air quality index in Delhi", Science of the Total Environment 409, no. 24(2011): 5517- 5523.