Asian Journal of Convergence in Technology **Volume IV, Issue II**
ISSN No.:2350-1146, I.F-5.11

Analysis of Air Quality Estimation based on Air Pollutants Parameters

Tejaswini Rajendra Patil
School of Computer Science and Engineering
Dr.Vishwanath Karad MIT World Peace University
Pune, India
tejaswipatil395@gmail.com

Dr. Siddhivinayak Kulkarni
School of Computer Science and Engineering
Dr.Vishwanath Karad MIT World Peace University
Pune, Indiasiddhivinayak.kulkarni@mitcoe.edu.in

***Abstract*-Air quality is the degree that tells us how pure or polluted the air is. It is important to know air quality in our surrounding as it negatively impacts human health and environment. Modernization and industrialization have given birth to air pollution which has become hidden killer. Making cities more respirable starts by analyzing and seizing the air pollution data. Long-established air quality prediction model gave less accurate and unsatisfactory result. There is a need of reliable data-analytics based solutions which will optimally predict air quality thereby enhancing our quality life. We evaluated various studies in this domain and summated the important researches done. This work will give future directions to the upcoming researchers.**

Keywords— Air Quality Prediction, Deep Learning, Neural Networks, meteorological factors, PM values.

I. INTRODUCTION

Air Quality notifys how good or bad the air is to breathe. Bad quality air causes pollution which is one of the key causes of death in the world. A massive nine in 10 people on Earth breathe extremely polluted air, and more than 80% of urban resident have to tolerate outdoor pollution that exceeds health standards, according to the WHO's World Global Ambient Air Quality Database.Dismally,India is one of the countries with maximum number of most polluted cities in the world.Though Delhi air pollution is at the peak there are other 70 cities all over India that indicated unhealthy air quality.In 2017, air pollution estimated for 12.4 lakh deaths in India, in which 6.7 lakh deaths where due to outdoor particulate matter emission and 4.8 lakh deaths due to household air pollution.

To restrict the effect of air pollution and to save the people from its effect, real-time air quality information is required. Air quality prediction is essential to take important decision for environmental management. Along with the power of big data and analytics policymaker can provide solutions that whenever used optimally will reduce air pollution helping to enhance healthy life.

II. LITERATURE SURVEY

Over the past few years an immense work is done in the study and analysis of air pollution. Many research efforts have focused on approaches to predict air pollutant concentration.

In [1] deep learning model is applied on AirNet [2] dataset to predict the $PM_{10\ values}$. Framework used RNN, LSTM and GRU networks to forecast prediction.[3] adopted a distributed fusion architecture to fuse urban data using DNN-based approach. Proposed system is compared with previous online system and the result is 2.4%,12.2%,63.2% improved accuracy on long-short term prediction.

Cloud-based ETL framework is developed for Air quality analysisin [4]. Data mining algorithms are applied to store, reprocess and analyse the data. Results of air quality prediction are visualized through browser.[5] focused on implementing data mining techniques to predict patterns in pollutants' studio and tableau platform are used for time series regression forecasting.

[6] developed a heuristic recurrent air quality predictor (RAQP) to infer air quality. RAQP method recurrently applies the one-hour prediction model, which learns the current records of meteorology and pollution-related factors to measure the future air quality one hour later, to then estimate the air quality after several hour.[7]presents Long-Short Term Memory model to predict air pollutant concentration based on historical data. To look over the performance $PM_{2.5}$ concentration in Beijing City is collected. Different models like Spatiotemporal deep learning, time delay neural network, autoregressive moving average model and support vector regression model are compared. Among the all LSTM performs well.

[8] used data driven models, ST-DNN, to predict PM2.5 over 48 hours. The model shows that including an LSTM module improved first hour predictions, involvement of CNN module being more useful for longer time frame predictions, since CNN can extract the temporal delay factor from surrounding target features by learning spatial information. Result of proposed model is daily report shown on the Facebook chatbot to send the PM2.5 forecasts to users who subscribed the daily report of specific stations.

[9] aims to develop a decisive prior-warning system that consist of air quality prediction and assessment model. To mark the ambiguity in prediction, use of back propagation neural network algorithm, combined with a probabilistic parameter model and data pre-processing techniques are used. Second model aims to provide clear description of air quality condition using fuzzy theory and Analytics Hierarchy Process [10] using raw images CNN model is applied for air quality estimation model involves a negative log-log ordinal classifier to fix the ordinal output well. A new activation function is developed to

improve the discriminative property of model. [11] Threenonlinear Auto Regressive model based Artificial Neural Network are used to predict industrial air quality.

TABLE I. LITERATURE SURVEY

Sr. no	Paper Title	Predicted Parameter	Methodology	Accuracy
1	A Solution for Air Quality Monitoring & Forecasting [12]	NO2, temperature, humidity, pressure	Multiple linear regression model	Model performed well, having high coefficients of determination (R2), close to 1 and Standard Error of 3.3183
2	A Bayesian hierarchical model for urban air quality prediction under uncertainty [13]	SO2, NOx, and dust fall	Bayesian hierarchical model	Average relative errors between observed data and mean values from model for SO2, NOx, and dust fall is 6.81%, 6.79% and 3.52%.
3	Performance Analysis of SVM ensemble methods for Air Pollution Data [14]	SO_2	SVM, Bagging, AdaboostM1	Accuracy of single SVM method was 76.33% & of bagging algorithm it was 79.66%. AdaboostM1 algorithm proved to be best with accuracy 91.28%
4	Air Pollution Analysis Using Enhanced K-Means Clustering Algorithm for Real Time Sensor Data.[15]	SO_2, NO_x, PM_{10}, O_3 and CO	K-Means Clustering algorithm, Possibility Fuzzy C-Means (PFCM) clustering algorithm.	K-means algorithm gives 40% more time than PFCM Algorithm.
5	Evolving Different evolution method with random forest for prediction of Air Pollution [16]	C_6H_6, NO_2, O_3, SO_2, CO, $PM_{2.5}$ and PM_{10}	Random Forest Algorithm	Combined differential evolution strategy with random forest method outperforms the stand with existing independent classifier of Bayesian network &multi-label classifier technique.
6	A deep learning model for air quality prediction in smart cities [17]	NO_2, O_3	Long Short-Term Memory (LSTM) Support Vector Regressor (SVR)	The prediction accuracy of the LSTM and SVR based model were calculated as 95% and 92.9%, respectively.

Result shows model 3 has highest accuracy that include combined meteorological variable with pollutants as compared to other two models.
Thus, as given in Table I., related work in the area of air quality prediction involves different analysis and prediction algorithm.

III. DATA SOURCES

The data can be obtained Central Pollution Control Board (CPCB). Currently, CPCB tracks the Air Pollution levels across different dimension (variables). Day wise, hour wise (for some variables). Following are some dimensions that can be considered:
Carbon Dioxide (CO_2): Carbon dioxide is high density greenhouse gas than dry air. A high concentration can upset oxygen in the air. Major sources of $CO_{2\ are}$ industrial emission, fertilizers, deforestation as well as the burning of fossil fuels like coal, oil and natural gas. CO_2 contributes 65% of total air pollution. Human exposure to high level of CO_2 causes suffocation, unconsciousness, headache.
Ozone (O_3): Ground level Ozone is perilous air pollutant. Chemical reaction of pollutants discharged from vehicles, power plants, factories with sunlight gives rise to ozone gas. It harms lungs and causes diseases such as asthma, chronic bronchitis.
Carbon Monoxide (CO): Carbon Monoxide is very poisonous odourless gas. Itacts as silent killer. Automobileexhaust, gasoline, furnaces are sources of CO. It causes dizziness, vomiting, and nausea.
Sulphur Dioxide (SO_2): It is nonflammable gas with a penetrating odour. Volcanic eruptions, Fuel combustion, emission from locomotives, ships, industrial activities give rise to SO_2. Sulphur dioxide causes acid rains. Itcauses eye irrition, coughing, mucus secretion, lungs malfunctioning.
Nitrogen Dioxide (NO_2):It is one of a group of highly reactive gases known as nitrogen oxides (NO_x). Burning of fuel are main source of (NO_2). Excess inhalation of (NO_2) causes respiratory disease.
Particulate Matter : It is mixture of solid particles like dust,pollen,ash,smoke and liquid droplet suspended in air many of which are hazardous.There are two types Particulate Matter PM_{10} (particle size 10 micrometers or less in diameter) and $PM_{2.5}$ (particle size 2.5 micrometers or less in diameter).It causes lung cancer,asthma, cardiovascular problems.
It is important to measure the following factors as they can help us understand the chemical reactions that occur in the atmosphere:

- Relative Humidity (RH)
- Temperature
- Wind Speed (Wind speed S)
- Vertical Wind speed (Wind speed V)
- Wind Direction
- Solar Radiation

Not all monitoring stations track Air Pollution on all the above-mentioned parameters and for all days.

IV. PROPOSED SYSTEM ARCHITECTURE

Proposed System architecture for our current study is given in Figure 1. Air quality prediction model consist of three parts. In the first part, data is collected from Central Pollution Control Board. This dataset is pre-processed to remove outliers, noise and redundancy. In second part, thepre-processed dataset is given to prediction model as input. Based on the previous labels model is trained and values for air pollutant are predicted.

Data visualization part generate the report. According to changing parameter and standards of Environment Protection Agencies we need to update dashboard. Updation of dashboard include changes in values, Graphs, charts. Statistical report Air Quality Index will be generated.

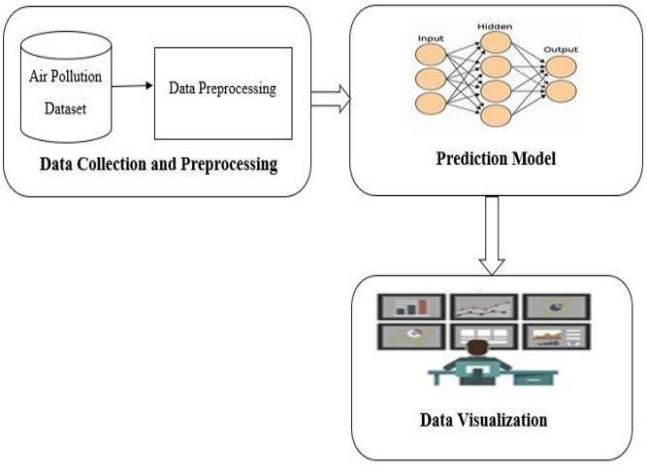

Fig. 1. Proposed System Architecture

Application:

- Industrial: Industries are giants emitting hazardous gases in environment. There is great scope of air quality prediction model in Industries. It will alert the industries with information needed to minimize pollution crest before they are likely to happen.

- Traffic Signals: Traffic Signals are major contributors of Vehicular Pollutions. Vehicles waiting at signals emits toxic gases that badly impact on humans and environment. Dashboard displaying future air quality will help in taking decisions to reduce pollution.

- GovernmentAgencies: Organizations like Environmental Protection Agency (EPA) and Pollution Control Boards can use this prediction model to forecast the daily air quality index. Model

will give authorities the required information that will help in preventing and controlling pollution.

V. CONCLUSION

Prediction algorithms are developing as a promising technique for forecasting non-linear time series information like meteorological and pollution data.It results to good performance for air quality prediction by extracting air quality features.In Future, we need to focus on long-term sudden changes in air quality level as it is very important. Large data set consisting various pollutants need to be considered for prediction. Forcasting model which is adaptable to dynamic atmosphere variable need to be build. The connection between different features can likewise be assessed hence enabling us to see whether there is any hidden parameter which will correlate the performance of features that appears to be different from the first peek.

REFERENCES

[1]Athira Va, Geetha Pb, Vinayakumar Rab, Soman K P*, "DeepAirNet: Applying Recurrent Networks for Air Quality Prediction", Elsevier 2018, Procedia Computer Science 132(2018) 1394-1403

[2] Songgang Zhao, Xingyuan Yuan, Da Xiao, Jianyuan Zhang, Zhouyuan Li. (2018) "AirNet: a machine learning dataset for air quality forecasting"

[3] XiamenYi, Junbo Zhang Zhaoyuan Wang, Tianrui Li Yu Zheng "Deep Distributed Fusion Network for Air Quality Prediction", KDD2018, August 1923,2018, London, United Kingdom.

[4]Yue Shan Chang, Kuan-Ming Lin, Yi-Ting Tsai, Yu-Ren Zeng, Cheng-Xiang Hung, "Big data platform for air quality analysis and prediction", IEEE 2018 27th Wireless and Optical Communication Conference (WOCC),30 April-1 May 2018, Hualien, Taiwan.

[5] NidhiSharma, ShwetaTaneja, VaishaliSagar, Arshita Bhatt, "Forecasting air pollution load in Delhi using data analysis tools", Elsevier2018, International Conference on Computational Intelligence and Data Science, Page no.1077–1085.

[6] Key Gu, June Qiao,,Weisi Lin , "Recurrent Air Quality Predictor Based on Meteorology- and Pollution-Related Factors",IEEE 2018, Volume: 14 , Issue: 9, Pages. 3946 - 3955

[7] Xiang Li, LingPeng, XiaojingYao, ShaolongCui, YuanHu, ChengzengYou, Tianhechi, "Long short-term memory neural network for air pollutant concentration predictions: Method development and evalution", Elsevier2017, Environmental Pollution 231(2017),997-1004.

[8]Ping-Wei Soh, Jia-Wei Chang, Jen-Wei Huang, "Adaptive Deep Learning-Based Air Quality Prediction Model using the most Relevant Spatial-Temporal Relations", IEEE 2018, Pages 38186 - 38199.

[9] Jianzhou Wang, XiaoboZhang, ZhenhaiGuo, Haiyan Lu, "Developing an earlywarning system for air quality prediction and assessment of cities in China", Elsevier 2017 Expert System with Applications 84(2017)102-116.

[10] ChaoZhang, JunchiYan, ChangshengLi, XiaoguangRui, LiangLiu, Rongfang Bie, "On Estimating Air Pollution from Photos Using Convolutional Neural Network", ACM2016, Amesterdam, Netherland.

[11]Nadjet Djebbri, Mounira_Rouainia, "Artificial Neural Networks Based Air Pollution Monitoring in Industrial Sites", IEEE ICET2017, Antalya, Turkey.

[12] Yasmine Gruicin, Marian-Emanuel Ionascu, Mircea Popa, "A solution for Air Quality Monitoring and Forecasting", IEEE 12th International Symposium on Applied Computational Intelligence and Informatics 2018, May 17-19, Timisoara, Romania.

[13] Young Liu, HuaichengGuangzhouMao, Pingjian Yang, "A Bayesian hierarchical model for urban air quality prediction under uncertainty",Elsevier 2008,Atmospheric Environment 42 (2008) 8464–8469

[14] Shahid Ails S Tirumala,"Performance Analysis of SVM ensemble methods for Air Pollution Data", ACM 2016,ICSPS2016, November 21 to 24, 2016, Auckland, New Zealand.

[15] Kings Grace. R, Animalia. R, Geetha Devising. M.S, Rajah. S, Usha. K, Radiatus Baseria. N,"Air Pollution Analysis Using Enhanced K-Means Clustering Algorithm for Real Time Sensor Detainee 2016.

[16] Rubal, Dinesh Kumar "Evolving Different evolution method with random forest for prediction of Air Pollution",Elsevier2018, International Conference on Computational Intelligence and Data Science (ICCIDS 2018).

[17]Ibrahim KOK, MehmetUlvi, SIMSEK, Suat OZDEMIR, "A deep learning model for air qualityprediction in smart cities",2017 IEEE International Conference on Big Data (BIGDATA), Pages: 1983 - 1990, Boston, MA, USA.