International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 3 Issue 6, October 2019 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

An Analytical Survey on Prediction of Air Quality Index

Suraj Kapse, Akshay Kurumkar, Vighnesh Manthapurvar, Prof. Rajesh Tak

Department of Information Technology, Dhole Patil College of Engineering, Pune, Maharashtra, India

ABSTRACT

A drastic increase of modernization gives birth to many industries and automobiles, which intern becomes the very common reason for the environmental issues like Air and water pollutions. Air pollution is the immediate affecting factors in our life, which contaminates the air that we breathe to cause serious health hazards. So it is very important to predict the Air quality index for the future coming days so that proper prompt action can be taken by the concern authorities to curb the same. The air quality reading for the different gases can be collect through the physical sensors and these readings can be used to predict the future Air quality index. Machine learning is acting as the catalyst in this prediction scenario to predict the accurate Air quality index for the future instance. Most of the learning systems need a huge amount of the data for the learning purpose and it is not possible to provide this every time. So it is a need to predict the air quality index by using considerable less amount of past instance data, This paper mainly concentrates on analyzing the past work in prediction of air quality index using machine learning and try to evaluate their flaws and to estimate the new possible way of prediction using machine learning.

KEYWORDS: *Air quality Index, Machine Learning, Gas Sensors*

How to cite this paper: Suraj Kapse | Akshay Kurumkar | Vighnesh Manthapurvar | Prof. Rajesh Tak "An Analytical Survey on Prediction of Air Quality Index" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-6, October 2019, pp.353-355, URL: https://www.ijtsrd.com/papers/ijtsrd28072.pdf

IJTSRD28072

I. INTRODUCTION

With economic development and population rise in cities, environmental pollution problems involving air pollution, water pollution, noise and the shortage of land resources have attracted increasing attention. Exposure to pollutants has an adverse effect on human health. The main reason of air pollution is energy production from power plants, residential heating, industries, fuel-burning vehicles, natural disasters, etc. Human health concern is one of the important consequences of air pollution, especially in urban areas. The global warming from anthropogenic greenhouse gas emissions is a long-term consequence of air pollution. Accurate air quality forecasting can reduce the effect of a pollution peak on the surrounding population and ecosystem, hence improving air quality forecasting is an important goal for society.

Air pollution is not only toxic to humans but also plays a major role in the overall integrity of the planet. The increase in air pollution has also been steadily increasing the overall temperature of the earth. This effect is known as Global Warming which can lead to devastating effects if not kept in check. Global Warming can cause a slew ofcatastrophic effects such as melting of the IceCaps on the poles of the earth, which has already been documented to be decreasing in size. Therefore, there is an absolute necessity to contain air pollution and monitor it to ensure that global warming does not escalate to a level that can be highly fatal to our planet. Many Air Quality monitoring mechanisms are implemented around the world and they have been highly successful in determining the changing quality of the air in real-time. Due to the importance of the Air Quality, WHO (World Health Organization) has issued warnings that stated that Air Quality is one of the most crucial aspects of Health care and it needs to be maintained at any cost.

The Air Quality Health Index (AQHI) may be public information tool designed in Canada to assist perceive the impact of air quality on health. The AQHI is outlined as AN index or rating scale vary from one to 10+ supported mortality study to indicate the extent of health risk related to native air quality. The larger the number, the greater the health risk and the need to take precautions. The formulation of Canadian national AQHI relies on three-hour average concentrations of ground-level ozone (O3), nitrogen dioxide (NO2), and fine particulate matter (PM2.5). The AQHI is calculated on a community basis, every community might have one or additional observation stations and also the average concentration of three substances is calculated at every station at intervals a community for the three preceding hours. AQHI is a purposeful index protective resident on a day to day from the negative effects of air pollution.

The existing systems detect the air quality of a particular metropolis into a different family like a commodity, satisfactory, centrist, poor, very poor, severe based on AQI (Air Quality Index). The data is displayed on a monthly, weekly or daily basis. Also, once the values are forecasted, the values do not vary with deference to the sudden change in the atmospheric changes or unexpected increase in traffic.

The values are detected for the whole metropolis, and cannot be verified for the accuracy of the forecasted values afterward. Some applications display the real number -time PM2.5 spirit level, while some show the forecast of a particular day. However, PM2.5 levels for dates after a week is not forecasted.

Most current air quality forecasting uses straightforward approaches like box models, Gaussian models, and linear statistical models. Those models are easy to implement and allow for the rapid calculation of forecasts. However, they typically don't describe the interactions and non-linear relationship that management the transport and behavior of pollutants within the atmosphere. With these challenges, machine learning ways originating from the sector of computer science became fashionable in air quality statement and alternative region issues. Broadly, machine learning is assessed into two categories; supervised and unsupervised. In unsupervised learning, the dataset is sculptured supported grouping and clustering to trace hidden data from it wherever target variables aren't clear. Supervised learning is the methodology wherever determined instances are labeled, or the target is known. Classification and regression are subtypes of supervised machine learning. If the target variable is continuous regression is employed, and for separate variables classification techniques are used.

This paper dedicates section 2 for analysis of past work as literature survey and section 3 concludes the paper with feasible statemement of the literature study.

II. LITERATURE SURVEY

This section of the literature survey eventually reveals some facts based on thought analysis of many authors work as follows.

Frances C. Moore [1], review the existing literature related to health, climatic impacts and economic, of black carbon emissions and tropospheric ozone in collaboration with mitigation options. The local character of many of the effects, merge with their little atmospheric lifetime and the presence of cost-effective decline technologies that are already broadly deployed in developed nations means to lessen these emissions provides a hugely climatically-effective alleviation option that is also appropriate to the development procedure of industrializing countries.

Markey Johnson, V. Isakov, J.S. Touma, S. Mukerjee, and H. Ozkaynak [2] presented the method for prediction of air quality concentrations of NOx, PM2.5 and benzene using hybrid modeling methods based on AERMOD and CMAQ model results. PM2.5 is hugely made in the atmosphere due to secondary reactions and appointed as criteria pollutant that is mostly regional. NOx was chosen as strongly affected by local burning sources and Benzene as illustrative of air toxics pollutants that are largely emitted from mobile sources. They made a model of 20 by 20 km area having air pollution impacts from emission sources. They utilized these modeled applications to developed and evaluate LUR models. They varying the number of training sites utilized to build the models to evaluate the LUR models on fit and performance.

Hsun-Ping Hsieh, Shou-De Lin and Yu Zheng [3], proposed a recommendation model which tell the most appropriate location of the building in which latest air quality monitoring stations can lead to the biggest accuracy enhancement in air quality inference. A framework that jointly infers air quality and recommends new locations is developed. They believed that the proposed framework is enough to be applied to the inference and deployment of other kinds of sensors. Several reasons lead to the success of the proposed model. First, the Affinity Graph seamlessly integrates spatial and temporal correlations. Second, the weights are learned to apprehend the correlation between AQI and features. It also lessens the uncertainty of the model. Finally, the proposed entropy-minimization greedy tries to identify a set of nodes that are uncorrelated with the more confident (i.e. low entropy) ones most of the time as the recommended locations for deployment.

Chamindi Malalgoda, Dilanthi Amaratunga, and Richard Haigh [4] elaborates the method of developing the abstract framework of a groundwork aimed toward developing a framework to empower native governments in creating a disaster-resilient engineered atmosphere among cities. The method includes distinctive key ideas, their inter-relationship and therefore the boundary of the study. The abstract framework is developed supported the literature review and any refined supported the findings from three knowledgeable opinions gathered as a part of the study. Consequently, the abstract framework illustrates the method for empowering native governments to create disaster-resilient engineered environments among cities.

Ulla Arthur Hvidtfeldt, Matthias Ketzel, Mette Sorensen, Ole Hertel, Jibran Khan, Jørgen Brandt, Ole Raaschou-Nielsen [5] evaluated calculations of PM2.5 and PM10 by AirGIS against concentrations measured at two fixed-site observation stations representing urban background and street, severally, and numerous address points within the Copenhagen space from two measurement campaigns B.C was evaluated against measured PM2.5 absorbance and PM10 absorbance from the two campaigns. Overall the concentrations sculptural by AirGIS correlated to well with the measured concentrations in relevance reproducing each temporal and spatial variation.

Wei Ying Yi, Kin Ming Lo, Terrence Mak, Kwong Sak Leung, Yee Leung and Mei Ling Meng [6] introduced the concept of TNGAPMS (The Next Generation Air Pollution Monitoring System) by utilizing the advance sensing technologies, Wireless Sensor Network (WSN), and Micro Electro Mechanical Systems (MEMS). Several of progressive pollution watching systems are enforced and tested. All of those systems proof that AN pollution monitoring system with a high spatio-temporal resolution, value, and energy potency, deployment, and maintenance practical feasibility, convenient accessing ability for the general public or skilled users are achievable.

Yu-Fei Xing, Yue-Hua Xu, Min-Hua Shi, Yi-Xin Lian [7] reviewed the effect of PM2.5 on respiratory disease and assists in preventing and designation the corresponding health problems and therefore the evolution of more practical strategies and technologies for the control and treatment of PM2.5 induced diseases.

Dixian Zhu, Changjie Cai, Tianbao Yang and Xun Zhou [8] proposed the efficient machine learning technique for air pollutant prediction. They developed the problem as regularized MTL and utilized advanced improvement algorithms for determination of totally different formulations. They have focused on alleviating model complexity by reducing the number of model parameters and on improving the performance by using a structured regularizer. They proved that the proposed system achieves far better performance than the other two model formulations which the regularization by imposing prediction models for two consecutive hours to be shut can even boost the performance of predictions. They also showed that advanced optimization techniques are important for improving the convergence of optimization and that they speed up the training process for big data.

Ilias Bougoudis, Konstantinos Demertzis, Lazaros Iliadis [9] presented the design, implementation, and testing of an innovative hybrid model capable of forecasting the concentrations of air pollutants. They utilize the method of combined learning for forecasting homogenous data vector clusters. This technique keeps away bad local behaviors. The presented model has been developed by considering data vectors related to all involved factors obtained from various representative measuring stations with specific topographic and microclimate characteristics. Thus, it can be considered a rational modeling effort with a good level of convergence and with high practical merit. Testing (rather difficult forecasting and decision making task) has been performed with reliable and rational results.

Kingsy Grace. R, Manimegalai, Geetha Devasena. M.S, Rajathi. S, Usha. K, Raabiathul Baseria [10] proposed the technique for finding AIQ (Air Quality Index) by using both PFCM and enhanced K-Means algorithms are implemented for different Datasets. Real-time data sets are taken from different places. The improved k-mean cluster algorithm provides AQI worth in higher accuracy however less execution time compared to the PFCM cluster Algorithm. The projected increased k-means cluster algorithm provides four-hundredth additional potency in terms of Accuracy and Execution time than PFCM algorithmic program. A distributed version of the K-means Clustering algorithm can be implemented where data or computational power is distributed. Efficiency can also be improved by using variable clusters instead of constant 'K' number of clusters.

III. CONCLUSION

As we know the prediction of Air pollution is always a beneficial thing in the process of curbing the Air pollution. Most of the traditional prediction systems are working on the linear process rather than the occasional and instance techniques, which yields poor results in prediction of Air quality index. So this paper deals with the most of the past methodologies and try to evaluate their gaps. On precise study of the past work this paper comes to a conclusion that still there is a lot of scope is existed in Air quality index measurement process. So this paper introduces using of K nearest neighbor classification protocol along with the Hidden Markov model to estimate the next instance Air pollution Quality Index, which will be reflect in the coming edition of our research paper.

REFERENCES

[1] Frances C. Moore, "Climate Change and Air Pollution: Exploring the Synergies and Potential for Mitigation in Industrializing Countries", Open Access Sustainability, 1, 43-54; DOI: 10.3390/su1010043, 2009.

[2] Markey Johnson, V. Isakov, J. S. Touma, S. Mukerjee, and H. Ozkaynak, "Evaluation Of Land-Use Regression Models Used To Predict Air Quality Concentrations In An Urban Area", Atmospheric Environment 44 (2010) 3660e3668, DOI:10.1016/j.atmosenv.2010.06.041, ELSEVIER, 2010.

[3] Hsun-Ping Hsieh, Shou-De Lin, and Yu Zheng, "Inferring Air Quality for Station Location Recommendation Based on Urban Big Data", KDD '15, August 10 - 13, 2015, Sydney, NSW, Australia, ACM, 2015.

[4] Chamindi Malalgoda, Dilanthi Amaratunga, and Richard Haigh, "Local Governments and Disaster Risk Reduction: A Conceptual Framework", https://www.researchgate.net/publication/309396143 , Research gate, September 2016.

[5] Ulla Arthur Hvidtfeldt, Matthias Ketzel, Mette Sorensen, Ole Hertel, Jibran Khan, Jørgen Brandt, Ole Raaschou-Nielsen, "Evaluation Of The Danish AirGIS Air Pollution Modeling System Against Measured Concentrations Of PM2.5, PM10, And Black Carbon", Environmental Epidemiology (2018)2:e014, DOI: 10.1097/EE9.0000000000000014, 30 May. 2018.

[6] Wei Ying Yi, Kin Ming Lo, Terrence Mak, Kwong Sak Leung, Yee Leung, and Mei-Ling Meng, "A Survey of Wireless Sensor Network Based Air Pollution Monitoring Systems", www.mdpi.com/journal/sensors, DOI: 10.3390/s151229859, 2015.

[7] Yu-Fei Xing, Yue-Hua Xu, Min-Hua Shi, Yi-Xin Lian, "The Impact of PM2.5 on The Human Respiratory System", Journal of Thoracic Disease, *J Thorac Dis*; 8(1): E69-E74., 2016.

[8] Dixian Zhu, Changjie Cai, Tianbao Yang and Xun Zhou, "A Machine Learning Approach for Air Quality Prediction: Model Regularization and Optimization", www.mdpi.com/journal/bdcc, DOI: 10.3390/bdcc2010005, Big Data Cogn. Comput. 2018.

[9] Ilias Bougoudis, Konstantinos Demertzis, LazarosIliadis, "HISYCOLaHybrid Computational Intelligence System for Combined Machine Learning: The Case of Air Pollution Modeling In Athens", DOI 10.1007/s00521-015-1927-7, Neural Comput & Applic, Springer, 10 June 2015.

[10] Kingsy Grace. R, Manimegalai, Geetha Devasena. M.S, Rajathi. S, Usha. K, Raabiathul Baseria, "Air Pollution Analysis Using Enhanced K-Means Clustering Algorithm for Real-Time Sensor Data", DOI: 978-1-5090-2597-8/16, IEEE, 2016.