ISSN 2395-1621

Air Quality Prediction Using IoT-Based Sensors and Machine Learning

Prajakta Yadav, Rujuta Rathod, Prof. Vineeta Phillip

prajaktayadav199@gmail.com
rujutarathod1305@gmail.com
vineeta.philip@aissmsioit.org

Department of E&TC, SPPU, AISSMS IOIT, Pune, India.

ABSTRACT

All the countries in the world are in a race against each other for making various technological advancements. Due to the mad race in technology man made blunders always cause harm to nature. Among which air pollution is one, which is a curse to mankind by the mother nature. According to World Health Organization, International agency for cancer research has identified that air pollution causes carcinogenic, a major cause for lung cancer in humans, putting air pollution in the same category as tobacco, UV rays etc. Urban air pollution causes 1.3 million deaths globally every year. A quarter of these deaths are due to lung cancer. So many measures are taken to curb air pollution, but due to immature decisions and negligence towards nature, things become worse day by day. It is necessary to measure the air pollution attribute and to predict its adverse effects on the environment so that some concrete measures can be taken to tackle it. To achieve this, the Internet of things (IoT) is the best technique.[1] When IoT blends with the machine learning algorithms it works perfectly for the reason of prediction. Internet of things (IoT) is getting popular day by day due to its flexibility and cost-effectiveness.[1] Therefore, to enhance the process of air pollution detection and prediction proposed model provides a way to predict the pollution from the collected data from sensors using Linear Regression analysis and Fuzzy Artificial Neural Network and the whole process is powered with Fuzzy C-Means Clustering.

Keywords— Air Quality, Pollution, Prediction, IoT-based sensors, Machine Learnin

ARTICLE INFO

Article History

Received: 8th March 2020

Received in revised form :

8th March 2020

Accepted: 10th March 2020

Published online :

11th March 2020

I. INTRODUCTION

Air pollution is the biggest problem in every country around the globe, be it developed or developing. Health problems have been growing at a faster rate especially in urban areas of developing countries where industrialization and the growing number of vehicles lead to the release of a lot of gaseous pollutants. Harmful effects of air pollution range from mild allergic reactions such as irritation of the throat, eyes and nose to some serious problems like bronchitis, heart diseases, pneumonia and asthma. According to a survey, 50,000 to 100,000 premature deaths occur per year in the U.S. alone due to air pollution. Whereas in the European Union the number reaches to 300,000 and over 3,000,000 worldwide. The main objective of the IoT Air Monitoring System is that air pollution is a growing issue these days. Air quality monitoring and control of pollution levels are necessary for a better future and healthy living for all. With urbanization and with the increase in the vehicles on the road the atmospheric conditions have considerably affected. The commercial meters available in the market are Fluke CO220 carbon monoxide meter for CO, Amprobe CO2 meter for CO2, Forbix Semicon LPG gas leakage sensor alarm for LPG leakage detection. Researchers in the field have proposed different air quality monitoring systems based on WSN, GSM, and GIS. Now each technology has limitations in usage according to the

intended application, as Zigbee is meant for users with Zigbee trans-receiver, Bluetooth. Systems based on GIS are designed, implemented and tested to monitor the points of air pollution of any area. This system comprises gas sensors, a micro-controller, a temporary memory buffer, a mobile unit and a web server with uninterrupted internet connectivity which collects real-time data from different locations along with location coordinate's information at a particular time of a day. The readings for a particular location are averaged for a specific time and space. The Global Positioning System (GPS) module is attached to a system to provide an accurate representation of pollution sources in an area. The recorded data is periodically transferred to a computer through a General Packet Radio Service (GPRS) connection and then the data will be displayed on the dedicated website with user acceptance. This, in turn, is beneficial for a large number of people.

II. RELATED WORK

It is found that air pollution is the fifth cause for death in India after high blood pressure, heart diseases, poor nutrition and tobacco smoking, about 6, 20,000 deaths occur due to air pollution every year. The traditional air monitoring stations available are quiet few and are very expensive. In paper we came across a dynamic and proficient method of scrutinize air pollution using wireless sensor network. The sensor materials monitor and control the air quality system. The nodes are distributed in area of interest to measure air quality index, compared with standard value and the collected information are transmitted to the nearest base station and to the control room for air pollution. For power management requirements the processor can be switched to sleep mode during idle conditions.[1]

In another experiment, the measurements acquired from gas sensors have been investigated in Bangkok metropolitan, and the results show air quality index (AQI) via Narrow-band Internet of Things (NBIoT).[2] One of the proposed works collected carbon dioxide and carbon monoxide levels in the air along with GPS location by using pollution detection sensors and uploaded it into Azure cloud services. Low cost embedded Beagle bone board along with gas sensors is used for data acquisition. Microsoft's Azure Machine learning service is used to predict the pollution metrics with the help of previous data. Processed data is fetched and represented by the Power BI tool. Calibrated gas sensor data is fetched from sensors and successfully uploaded into the cloud. Data stored in the cloud is utilized by different cloud services to make the data meaningful. The proposed system was implemented and useful to monitor and reduce pollution in a smart city by avoiding the causes.[4]

From the literature survey, it can be summarized that it is possible for determining the quality of Air of our surroundings. This is very essential because due to rapid urbanization and industrial growth in recent years, the amount of suspended particulate matter in major cities has increased exponentially. This could lead to an increased number of Lung ailments and suffocation for individuals suffering from asthma or other diseases. As the level of pollution increases, it also increases the chances of serious Lung Diseases for the people residing in those particular areas. Therefore, it is highly necessary to monitor Air Quality as it has an immense impact on a person's health.

III. METHODOLOGY

Data collected using gas sensors will be stored in the database. After a specific interval data will be monitored using fuzzy c-means algorithm. Fuzzy C-means is a clustering algorithm using fuzzy logic. It has 5 clusters namely very low, low, medium, high, very high. This data will then be analyzed using linear regression and entropy estimation. Entropy estimation consists of comparisons and distribution of most frequently occurring values. We are also going to use fuzzy ANN which is an artificial neural network. Using this method, we will then be able to predict the parameters for the next day or week.

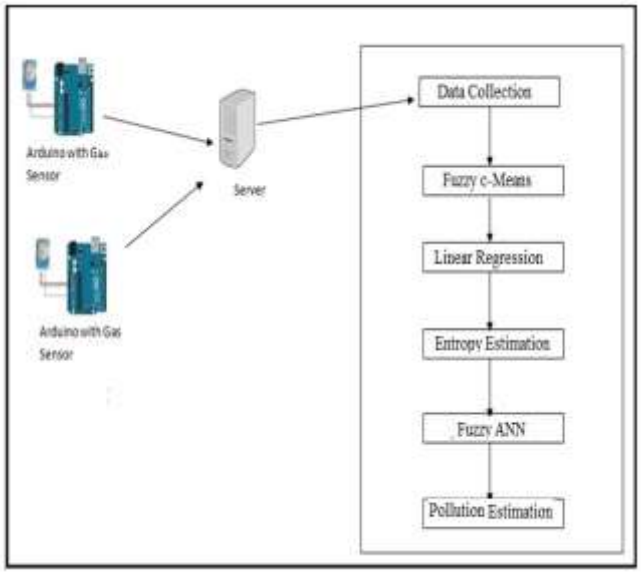

Fig.1. System architecture

Data Collection:
A gas sensor is interfaced with an Arduino board and this setup is used for data collection on regular intervals in a particular area. This data is stored continuously in the database using timestamps. The output of this step is a data vector.

Fuzzy C-Means:
The data vector is used as input here. Fuzzy C-means is a clustering algorithm. It forms clusters of the data received, more like grouping the data, and in the output giving cluster vectors.

Linear Regression:
This helps in estimating the threshold of prediction. A graph is plotted based on the values obtained from the cluster vectors. Roughly a straight line is drawn using the points which are approximately linear to each other.

Entropy Estimation:

Identifying distribution among the cluster based on threshold value of clusters.

Fuzzy ANN:
The entropy factors and the regression list are used as inputs here. This helps in estimating pollution index based on Artificial Neural Network and Fuzzy Crisp Value like very low, low, high, very high.

System Specifications

• Software Specifications:
Coding Language: Java.
Development Kit: JDK 1.8, JRE.
Front End: Java Swing
Development IDE: Netbeans 8.0,
Data Base**:** My SQL 5.0

•Minimum Hardware Specification:
Sensors: MQ135 for Co2, MQ7 for Co
Kit: Arduino
Transmission KIT: RF 433
Router: D Link

System Design

In the proposed system, we will be having two Arduino ATMEGA328p and two gas sensors one for Co2 namely MQ135 and another one for Co namely MQ7.
The hardware will be interfaced with PC using RF433 kit. The software part consists of Netbeans8.0 development IDE as it is very flexible to use.
Data collected from sensors will be stored in database using MY SQL 5.0 software. Along with Netbeans IDE we are also using JAVA DEVELOPMENT KIT because our main programming part will include Java programming language.

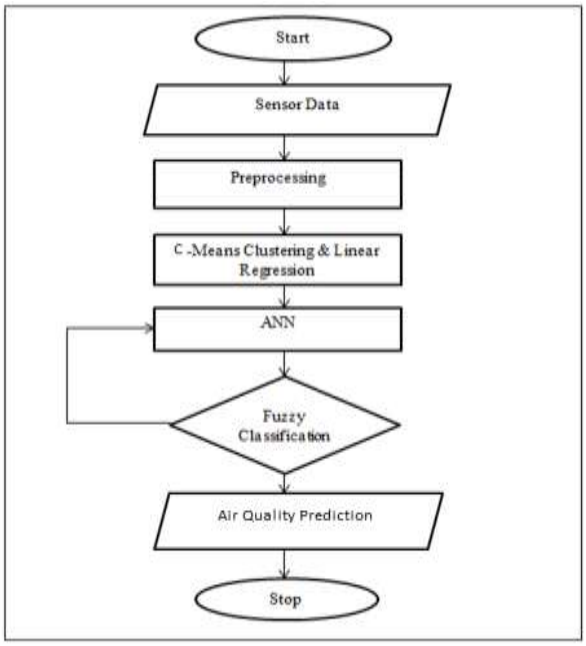

Fig.2. flow diagram

IV. RESULT

The data collected from the sensors will be collected in the fixed interval and this data using the machine learning algorithms will help predict the air quality for the next day or week. Increased accuracy will be achieved using machine learning. More number of sensors can be used to monitor various other gases too.

V. LIMITATIONS AND ASSUMPTIONS

The limitations of the proposed system are that limited number of sensors are used, hence, restricting the number of gases that can be measured. The system is assumed to work effectively for both dataset as well as real time data. Limited attributes are considered for air quality prediction. A small room is considered as a bigger area and lighting of an incense stick or matchstick is considered as the source.

VI. CONCLUSION

A conclusion section must be included and should indicate clearly the advantages, limitations, and possible applications of the paper. Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions.

REFRENCES

[1] Baowei Wang, Weiwen Kong, Hui Guan, Neal N. Xiong, “Air Quality Forecasting based on Gated Recurrent Long Short-Term Memory Model in Internet of Things”, 2017 IEEE Access Research Paper.

[2] Sarun Duangsuwan, Aekarong Takarn, Rachan Nujankaew, Punyawi Jamjareegulgarn, “A Study of Air Pollution Smart Sensor LPWAN via NB-IoT for Thailand Smart Cities 4.0”, IEEE Communication Surveys and Tutorials, vol. 19, 2017, pp. 855-873.

[3] Yi-Ting Tsai, Yu-Ren Zeng, Yue-Shan Chang, “Air Pollution forecasting using RNN with LSTM”,2018 IEEE 16th international conference on dependable, autonomic and secure comp.

[4] Changmin Li,Hongxia Wang,”A Rating Charge Model of Transfer Tax for China’s Interprovincial Air Pollution Control”2473-3547/16 $31.00 © 2016 IEEE DOI 10.1109/ISCID.2016.176.

[5] Yunping Chen, Shudong Wang, Weihong Han, Yajv Xiong, Wenhuan Wang, and Ling Tong,”A New Air Pollution Source Identification Method Based on Remotely Sensed Aerosol and Improved Glowworm Swarm Optimization, “IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING”, 2017 IEEE paper.

[6] Nitin Sadashiv Desai, John Sahaya Rani Alex, "IoT based air pollution monitoring and predictor system on Beagle Bone Black", 978-1-5090-5913-3/17/$31.00c 2017 IEEE.

[7] Norsuzila Ya'acob,AzieanM. Azize ,Noranizar M. Adnan , Azita Laily Yusof1, and Suzi Seroja Sarnin, "Haze Monitoring based on Air Pollution Index (API) and Geographic Information System (GIS)", 2016 IEEE Conference on Systems, Process and Control (ICSPC 2016), 16-18 December 2016, Melaka, Malaysia.

