ISSN 2395-1621

DATA MONITORING USING IOT AND ML

Vikrant Ghode, Akshay Borawake, Pratik Ithape, Sahib Singh Makhija,

Prof. K. Thoke

kiranthoke16@gmail.com

Department of Computer Engineering, KJCOEMR, Pune,India, 411048

ABSTRACT

Using empirical analysis, conventional air automatic monitoring system has high precision, but large bulk, high cost, and single datum class make it impossible for large-scale installation. Based on introducing embedded system into the field of environmental protection, this project puts forward a kind of real-time air pollution monitoring and forecasting system. By using embedded, this system can reduce the hardware cost into 1/10 as before. The system can be laid out in a large number in monitoring area to form monitoring sensor network. Besides the functions of conventional air automatic monitoring system, it also exhibits the function of forecasting development trend of air pollution within a certain time range by analyzing the data obtained by front-end perception system according to neural network technology. Targeted emergency disposal measures can be taken to minimize losses in practical application.

Keywords — nodeMCU, MQ135, DHT11, MQ5

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

With the rapid development of economy, chemical industrial park construction and production activity are increasingly frequent, leading to increasing probability of environmental pollution accidents, especially air pollution accident. Affected by meteorological and geographical conditions, air pollution will be highly clustered in a short time after happening, causing great harm or even extreme destruction to both human and environment. So, it is particularly important to set up a real-time air pollution monitoring system. Using laboratory analysis, conventional air automatic monitoring system has relatively complex equipment technology, large bulk, unstable operation and high cost. High cost and large bulk make it impossible for large-scale installation. This system can only be installed in key monitoring locations of some key enterprises; thus, system data is unavailable to predict overall pollution situation. To overcome defects of traditional monitoring system and detection methods and reduce test cost, this paper proposes a method combining embedded technology with environment monitoring. By replacing monitoring equipment in traditional empirical analysis with sensor network in IOC technology, through which inexpensive sensors can be laid out flexibly in the whole area to monitor omni-directionally to provide data support for prediction.

II. LITERATURE SURVEY

Wearable sensors have also been used to monitor air quality, for instance in the Common Sense [5] and CitiSense [6] solutions. Both rely on small, battery978-1-5090-13890/16/$31.00 ©2016 IEEE 443 2016 39th International Spring Seminar on Electronics Technology (ISSE) powered sensor nodes that measure the concentrations of polluting gases and send the data to users' smartphones via Bluetooth [1].

This data and the GPS coordinates are then shared with other users through a dedicated website. The Common Sense system explored outdoor sensors in a number of contexts including sensors attached to street sweeper vehicles, and hand held sensors that could be used to sample interesting outdoor locations. The utilization of street sweeper mounted sensors aimed to enlarge the existing sensor infrastructure in the city. The CitiSense system provides desktop based, reflection supporting visualizations and "in-themoment" visualizations that support real-time analysis [2]. A similar approach is also followed in where, a small and portable ozone measurement system GasMobile is proposed. It uses a particular sensor equipped with a transmitter board that allows easy connection of the sensor to the user's smartphone through the USB port. The system is extremely compact and employs an Android application to calibrate the sensor and upload the acquired data to a server. But the readings are very limited and refer only to a single gas. Ching-Biau Tzeng et al [3]. It have described an indoor air quality (IAQ) monitoring system based on ZigBee wireless sensor network implemented with the TI CC2430 chip. In the system they propose, each sensor node measures temperature, relative humidity and carbon dioxide. The combination of amperometric gas sensors for pollutant and oxygen measurement with sensors for precise measurement of basic physical parameters, such as atmospheric pressure, temperature and humidity, make it possible to perform higher precision measurements of gas concentrations [4].

III. PROPOSED SYSTEM

The main working of this project has two steps: The hardware and the machine learning and prediction. The hardware collects data from various sensors and makes a dataset out of this data.This data is then used for analysis [1].

The machine learning working is separated into three main stages: Initial, Middle, Last stage [2]. The Initial stage is identified with Data Exploration, Data Cleaning and Data Transformation [3]. The centre stage comprises of data modelling [4]. The final stage comprises of data analysis using four models viz. Random Forest, SVM and Decision Tree [5]. Data exploration is similar to initial data analysis, visual exploration to understand what is in a dataset and the characteristics of the data, rather than through traditional data management systems [6]. Data Cleaning is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data [7]. Data transformation is the process of converting data from one format to another, typically from the format of a source system into the required format of a destination system [8]. Once the first stage is cleared then we move to data modelling. Data modelling is the process of producing a descriptive diagram of relationships between various types of information that are to be stored in a database. One of the goals of data modelling is to create the most efficient method of storing information while still providing for complete access and reporting [9]. After this the data is processed using algorithms and results are obtained [10]. This result are the test results generated by training the models on the train dataset [11]. Once the dataset is processed then we can make use of the actual dataset to predict air quality [12].

IV. MATHEMATICAL ALGORITHMS

[1] Random Forest :

A Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap Aggregation, commonly known as bagging. First, we pass the features(X) and the dependent(y) variable values of the data set, to the method created for the random forest regression model. We then use the grid search cross validation method (refer to this article for more information) from the sklearn library to determine the optimal values to be used for the hyperparameters of our model from a specified range of values. [2] SVM :

Support vector machines so called as SVM is a supervised learning algorithm which can be used for classification and regression problems as support vector classification (SVC) and support vector regression (SVR). It is used for smaller dataset as it takes too long to process. In this set, we will be focusing on SVC. The points closest to the hyperplane are called as the support vector points and the distance of the vectors from the hyperplane are called the margins.

V. ARCHITECTURAL REPRESENTATION

NodeMCU is an open source IoT platform. It includes firmware which runs on the ESP8266 WiFi SoC from Espressif Systems, and hardware which is based on the

ESP-12 module. The term "NodeMCU" by default refers to the firmware rather than the development kits. The firmware uses the Lua scripting language. It is based on the eLua project, and built on the Espressif Non-OS SDK for ESP8266. It uses many open source projects, such as luacjson and SPIFFS [1]. The DHT11 is a basic, ultra-low-cost digital temperature and humidity sensor. It uses a capacitive humidity sensor and a thermistor to measure the surrounding air, and spits out a digital signal on the data pin (no analog input pins needed). It's fairly simple to use, but requires careful timing to grab data. The only real downside of this sensor is you can only get new data from it once every 2 seconds, so when using our library, sensor readings cabe up to 2 seconds old [2].

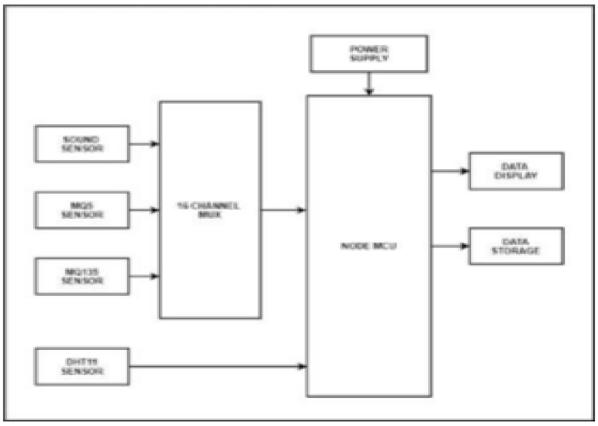

Fig.3.1 Basic Architecture (1)

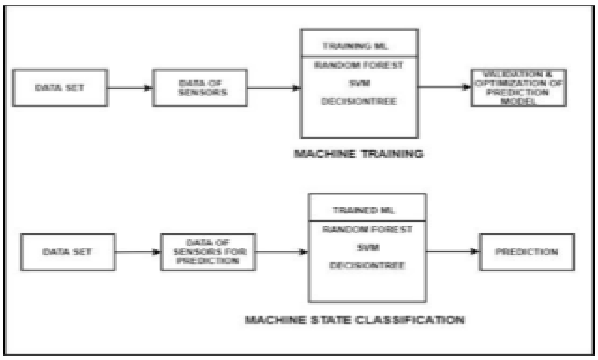

Fig.3.2 Basic Architecture (2)

Gas Sensor(MQ5) module is useful for gas leakage detection (in home and industry). It is suitable for detecting H2, LPG, CH4, CO, Alcohol. Due to its high sensitivity and fast response time, measurements can be taken as soon as possible [3].

VI. GRAPHICAL RESULTS AND ANALYSIS

VII. SYSTEM FEATURES

Since multiple models are used it increases the efficiency of the system [1]. As we can estimate the air pollution, we can take preventive measures for the same [2]. The model uses large number of parameters for prediction which makes the system accurate [3]. Reliable, Easy to use and Portable [4].

VIII. CONCLUSION

Air pollution is one of the vital factors that affects the quality of people's life in the urban environment. Existing monitoring systems play an important role in many smart city developments in urban areas for monitoring and governing air quality and the main pollutant concentrations. Using such information, we could revisit the concepts like when, where the emission decreases and why the dangerous exposures of atmospheric pollution have been overestimated from urban areas. Using empirical analysis, conventional air automatic monitoring system has high precision, but large bulk and high cost makes it impossible for large-scale installation, to create better and safer environment for human beings, animals, plants, there is a need to monitor and control the pollution. Current monitoring systems are precise but massive, expensive in nature and hence seldom.

IX. REFRENCES

[1] World Health Organisation, "7 million premature deaths annually linked to air pollution," 2012.

[2] U.S. Environmental Protection Agency, "Air Pollution Monitoring," [Online]. Available: http://www.epa.gov/airquality/montring.html. [Accessed 09 April 2015].

[3] "Smart City-Strategie Berlin," Senatsverwaltung für Stadtentwicklung und Umwelt - Berlin, 2015.

[4] "Impact Analysis," European Commission, 2013.

[5] Dutta, P.; P.M. Aoki, N. Kumar et al., "Common Sense: Participatory urban sensing using a network of handheld air quality monitors," in Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, Berkeley, CA, USA, 4–6 November, 2009.

[6] Nikzad, N., N. Verma, C. Ziftci, C. et al., "CitiSense: Improving geospatial environmental assessment of air quality using a wireless personal exposure monitoring system," in Proceedings of the ACM Conference on Wireless Health, San Diego, CA, USA, 23–25 October 2012.

[7] Alphasense Ltd., "Amperometric Electrochemical Gas Sensors," November 2015. [Online]. Available: http://www.alphasense.com/index.php/air/downloads/.

[8] Z. Kokolanski, C. Gavrovski, V. Dimcev, "Modified single point calibration with improved accuracy in direct sensor-to-microcontroller interface," Measurement, vol. 53, p. 22–29, 2014.

[9] "Executive Environment Agency - Air Quality Monitoring," [Online]. Available: http://eea.government.bg/airq/bulletin.jsp.

[10] R. Piedrahita, Y. Xiang, N. Masson, "The next generation of low-cost personal air quality sensors for quantitative exposure monitoring, Atmospheric Measurement Techniques," vol. 7, p. 3325–3336, 2014.

[11] Ching-Biau Tzeng, Tzuu-Shaang Wey, "Design and Implement a Cost Effective and Ubiquitous Indoor Air Quality Monitoring System Based on ZigBee Wireless Sensor Network," in Innovations in Bio-inspired Computing and Applications (IBICA), 2011 2nd Int. Conf., 16-18 Dec. 2011.

