International Journal for Modern Trends in Science and Technology, 9(03): 175-179, 2023

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0903026
Available online at: http://www.ijmtst.com/vol9issue03.html

Air Quality Monitoring and Prediction using IoT and Machine Learning Approach

B Srinivas Raja, G.V.Padma Sai Sruthi, M.Y.Lokesh, A.Grace, S.Lakshmi Reddy

Dept of ECE, Godavari Institute of Engineering and Technology(A), Rajahmundry, A.P

To Cite this Article
B Srinivas Raja, G.V.Padma Sai Sruthi, M.Y.Lokesh, A.Grace and S.Lakshmi Reddy. Air Quality Monitoring and Prediction using IoT and Machine Learning Approach. International Journal for Modern Trends in Science and Technology 2023, 9(03), pp. 175-179. https://doi.org/10.46501/IJMTST0903026.

Article Info
Received: 26 February 2023; Accepted: 15 March 2023; Published: 18 March 2023.

ABSTRACT

The air quality index number established by the government has been exceeded, indicating dangerously high levels of pollution in today's major cities. A person's health is significantly impacted by it. Using iot and machine learning techniques, we created a model to track and foretell the air quality index in a given area. In order to keep air quality at a high standard, a system is in place to gather data on pollutants from a designated area utilising IOT (microcontrollers, sensors, and the like). Data is gathered with the help of an Arduino Uno microcontroller. Meanwhile, the dataset is sent through server using a device called Node MCU that has Wi-Fi capabilities. The dangerous contaminants in the air are detected using a variety of sensors, including the MQ135, MQ5, MQ2, and DHT11 sensor. While trying to determine the quality of the air in a particular area, Machine Learning (ML) combines statistical analysis and computational modelling to provide the most accurate predictions possible. Among the components of an Internet of Things–based air pollution monitoring system is a MQ Series sensor interfaced to a Node MCU fitted with an adapter to transmit the sensor reading to a Thing Talk cloud. With the use of gathered data on pollutant concentration, our model will be able to effectively monitor and forecast the air quality index of any enclosed zone.

Indicating the pollution levels during a certain time period, the Air Quality Index of India is a commonly used metric. For individuals to take preventative steps against air pollution, this technology provides forecasts of pollution levels.

KEYWORDS: Air Quality, IoT, Node MCU, WIFI

1. INTRODUCTION

Air pollution is one of the most pressing issues of our day. India, the fastest-growing industrial country, is responsible for a staggering increase in the global total of airborne pollution. The fast expansion of the global economy has led to a surge in industrial production, which in turn has increased air pollution. Pollution from industry is a major contributor to the worldwide environmental crisis that threatens all forms of life, including human beings. Air pollution may be caused by both gases and solid particles including dust, pollen, and spores. Combustion of natural gas, coal, and wood, as well as industries, automobiles, and other sources, releases a variety of air pollutants such as carbon monoxide, carbon dioxide, nitrogen dioxide, sulphur oxide, chlorofluorocarbons, and particulate matter, all of which contribute to air pollution. Breathing in polluted air causes several health issues, especially in the lungs and respiratory system, if exposed to it for lengthy periods of time.

If you want to know how much of a toll pollution is taking on a certain area, you may look to the air quality indices that have been developed in accordance with the Indian Air Quality Standard. There are a variety of

scales and indices to measure the effects of pollution on the environment. Using a variety of sensors, including MQ135, MQ5, MQ2, and DHT11, we gather data on a certain area. To transmit sensor data to the Thing Talk cloud, an IoT-based air pollution monitoring system uses a MQ Series sensor interfaced to a Node MCU using an adapter. Thingspeak will be used to export the data gathered by the sensors to an Excel file. We have developed a model that can anticipate air quality in any given location by calculating the air quality index for each data point in the dataset. Forecasting the air quality index allows us to trace the source of the pollution and the area's most severely impacted by it. This predictive model employs a number of methods for extracting information from data in order to identify the area's most likely to be adversely impacted by a certain event (cluster). This sheds light on the origin and chronology of the contaminants.

1.1 PROBLEM DEFINITION

Countries throughout the globe are increasingly worried about air pollution. It poses risks to human and environmental health but is difficult to detect since it is often not apparent to the naked eye. To choose the right course of action, one must first ascertain the extent of their focus at any particular moment. Although tools like the recurrent air quality predictor might help us anticipate when pollution levels in the air could rise or decrease, their performance degrades after a few hours. As a result, this work employs an ensemble of iot and ML algorithms to develop a system that can foretell the concentration of air pollutants in the future.

2. PROPOSED WORK

Two controllers, an Arduino and a NodeMCU, are used in the suggested technique. We are using Arduino to interface air quality sensors such as the MQ5, MQ2, and MQ135. The DHT11 sensor measures ambient temperatures and relative humidity. These sensor readings will be sent to the server via NodeMCU. Because NodeMCU has a WIFI module, we can connect it to our network to retrieve the data it has gathered and send it on to our server. The LCD will show off these details. On the server, we may access Excel files containing the data. Using this information, machine learning processes may be carried out.

While writing new code, we always use the most recent version of py.charm.

The findings will be shown in a whole new format. The code may then be deployed by using the appropriate run command. After the run command has been executed, the resulting processing will provide a link to the template. To get access to the information in the template, one must first register. After that, the Excel sheet with the information is sent out to be processed. We see the accuracy of prediction made by many algorithms. Logistic Regression is the most accurate approach for predicting air quality.

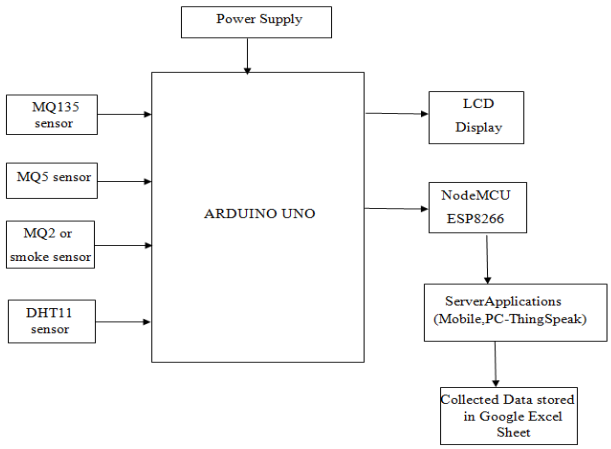


Fig1: Block Diagram of Hardware Implementation

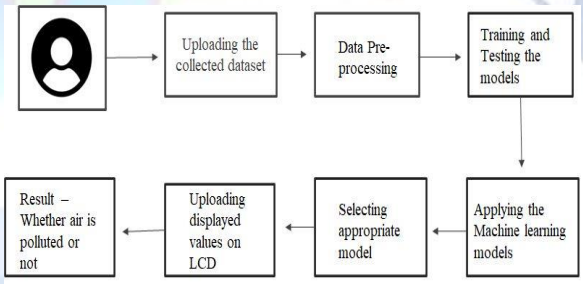


Fig2: Block Diagram of Software Implementation

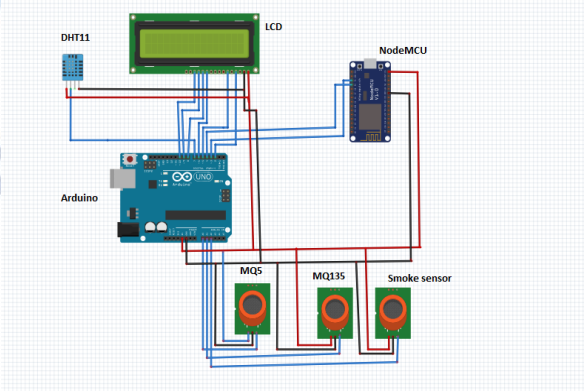


Fig3: Circuit Diagram

3. COMPARATIVE STUDY OF EXISTING AND PROPOSED SYSTEM

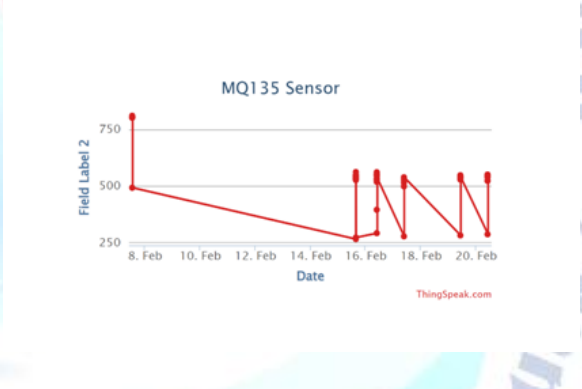
We compare the accuracy of many methods for predicting air quality in our project, including X Gradient Boost (XGB), Random Forest (RF), Decision Tree (DT), Convolution Neural Network (CNN), and Logistic Regression. The main shortcoming of the current method is its low accuracy, whereas the suggested approach provides high forecast accuracy.

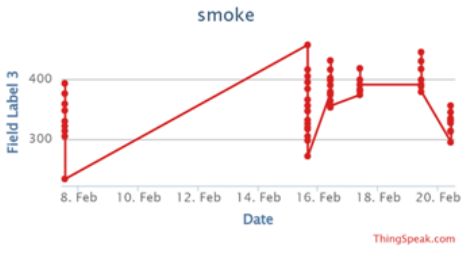
The supervised classification technique known as logistic regression (LR) is robust and widely used. LR is useful for calculating the likelihood that a novel instance belongs to a certain class. As this is a probability, the result may be anything from zero to one. If you want to use the LR as a binary classifier, you'll need to set a threshold to separate the two groups.One of the first and most well-known machine learning algorithms is the decision tree (DT). A decision tree is a graphical representation of the logical steps involved in categorising data, such as the tests to be performed and the corresponding results. Based on the result of the test, the classification algorithm branches to the most relevant child node, and so on, all the way up to the leaf node.

Similar to how a forest is a collection of numerous trees, a random forest (RF) is an ensemble classifier made up of several DTs. Deep DTs tend to overfit the training data, leading to a large amount of uncertainty in the classification result for even a little change in the input data. These models are too dependent on their training data, leading to inaccuracies when applied to a test set.A convolutional neural network (CNN) is a special form of network design for deep learning algorithms that is used in image recognition and other pixel-based applications. Although deep learning employs a wide variety of neural networks, convolutional neural networks (CNNs) are the preferred network design for tasks such as object recognition and identification.Although some machine learning models perform better than others, the best ones can only be determined by the distribution of training data.

ALGORITHM	ACCURACY
Random Forest	92.95%
Decision Tree	63.38%
XGBoost Classifier	97.18%
Logistic Regression	98.59%
Convolutional Neural Network	93.35%

4. RESULTS





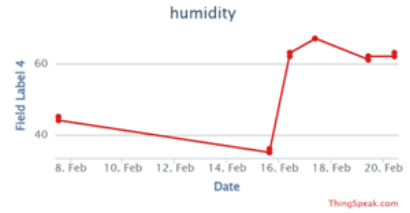


Table1: AQI RANGE

AQI	Descriptor	General Health Effects
0–50	Good	None
51–100	Moderate	Few or none for the general population
101–200	Unhealthy	Everyone may begin to experience health effects; members of sensitive groups may experience more serious health effects. To stay indoors.
201–300	Very unhealthy	Health warnings of emergency conditions. The entire population is more likely to be affected.
301+	Hazardous	Health alert: everyone may experience more serious health effects

5. CONCLUSION

We can confidently anticipate the air quality index of any given place since our model can accurately forecast the existing data. We can predict the AQI and send an alert to the appropriate area of the nation using this model, and since it is a progressive learning model, we can use the collected time series data to pinpoint exactly where our focus should be directed. China's air quality monitoring and research stage provided the data used for this project, which includes daily averages for fine particulate matter (PM2.5), inhalable particulate matter (PM10), ozone (O3), carbon monoxide (CO), sulphur dioxide (SO2), nitrogen dioxide (NO2) fixation, and air quality record (AQI). When determining the poison concentration, it is crucial to take into account both the many factors that contribute to the concentration and the various sources from which the poison was originally derived. The CNN approach used for classification in this study was shown to be more effective than previously used algorithms in terms of both accuracy and precision.

FUTURE SCOPE

The CNN approach used in this study to execute classification proved to be more effective than previously used algorithms in terms of both accuracy and precision. In the near future, we will be able to apply this to deep learning strategies. The India Meteorological Department is looking to implement an automated system for determining whether or not air quality meets qualifying requirements (real time). With the forecast result shown in a web application or desktop application, we can streamline this procedure. In order to improve the efficiency of the task in an AI setting.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

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] Ayele, TemeseganWalelign, and RutvikMehta."Air pollution monitoring and prediction using IoT." In 2018 Second International Conference on Inventive Communication 6 Fig. 12. RH w.r.t Temperature Fig. 13. RH w.r.t CO and Computational Technologies (ICICCT), pp. 1741-1745. IEEE,2018.

[5] Djebbri, Nadjet, and MouniraRouainia. "Artificial neural networksbased air pollution monitoring inindustrial sites." In 2017 International Conference on Engineering and Technology (ICET), pp. 1-5. IEEE,2017.

[6] Kumar, Dinesh. "Evolving Differential evolution method with random forest for prediction of Air Pollution." Procedia computer science 132 (2018): 824-833.

[7] Jiang, Ningbo, and Matthew L. Riley. "Exploring the utility of the random forest method for forecasting ozone pollution in SYDNEY." Journal of Environment Protection and Sustainable Development 1.5 (2015): 245-254.

[8] Svetnik, Vladimir, et al. "Random forest: a classification and regression tool for compound classification and QSAR modeling." Journal of chemical information and computer sciences 43.6 (2003): 1947-1958.

[9] Biau, GA˜ Srard. "Analysis of a random forest model." ˇJournal of Machine Learning Research 13.Apr (2012): 1063- 1095.

[10] Biau, Gerard, and ErwanScornet. "A random forest ´ guided tour." Test 25.2 (2016): 197-227.

[11] Grimm, Rosina, et al. "Soil organic carbon concentrations and stocks on Barro Colorado Island— Digital soil mapping using Random Forests analysis." Geoderma 146.1- 2 (2008): 102-113.

[12] Strobl, Carolin, et al. "Conditional variable importance for random forests." BMC bioinformatics 9.1 (2008): 307.

[13] Svetnik, Vladimir, et al. "Random forest: a classification and regression tool for compound classification and QSAR modeling." Journal of chemical information and computer sciences 43.6 (2003): 1947-1958.

[14] Verikas, Antanas, AdasGelzinis, and MarijaBacauskiene. "Mining data with random forests: A survey and results of new tests." Pattern recognition 44.2 (2011): 330-349.

[15] Ramasamy Jayamurugan,1 B. Kumaravel,1 S. Palanivelraja,1 and M.P.Chockalingam2 International Journal of Atmospheric Sciences Volume 2013, Article ID 264046, 7 pages http://dx.doi.org/10.1155/2013/264046

[16] V. M. Niharika and P. S. Rao, "A survey on air quality forecasting techniques,"International Journal of Computer Science and Information Technologies, vol. 5, no. 1, pp.103-107, 2014. 2.

[17] NAAQS Table. (2015). [Online]. Available: https://www.epa.gov/criteria-air-pollutants/naaqs-table 3.