International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 6 June 2021, pp: 18-22 www.ijaem.net ISSN: 2395-5252

LSTM Based Air Quality Prediction Model In Smart Cities

P.Parkavi(1), S.Rathi (2)
[1]*Student, Government College of Technology, Coimbatore, Tamilnadu*
[2]*Professor, Government College of Technology, Coimbatore, Tamilnadu.*
Corresponding Author: P.Parkavi

Submitted: 25-05-2021 Revised: 31-05-2021 Accepted: 03-06-2021

ABSTRACT: The level of pollution is increasing due to factors like industries, vehicles use, dust particles which can affect the human health. Nowadays interests on measuring and predicting the quality of air is increased among researchers to improve the quality for people life to be pollutant-free area. The IoT also helps in a wide range to improve quality of air using the multiple sensors.

The IoT based air pollution mointering system is used to mointer the air quality with a web server using Internet. The sensors which like MQ2, MQ135, MQ9,and PM2.5 used to find the air quality. The work initiated with sensor and the quality of air values can be into the cloud server called 'THINKSPEAK'. These values helps to find the accuracy using LSTM algorithm and they compared with SVR algorithm for finding the better accuracy among them. According to EPA, the AQI values measures according to their concentration levels of air Quality.

KEYWORDS: Air Quality Prediction, Thinkspeak, Methane, Particulate matter2.5, carbon monoxide, Internet of Things.

I. INTRODUCTION

The term air quality refers to the degree to which the air in a particular place is free from pollutants. Air quality is influenced by a variety of factors and is a complex issue. Air pollution is a major problem that has been recognized throughout the world for hundreds of years. In the middle Ages, the burning of coal in cities released increasing amounts of smoke and sulphur dioxide to the atmosphere. In more recent times pollution from motor vehicles has become the most recognized air quality issue. Air pollutants are substances present in the atmosphere at concentrations above their normal background levels which can have a measurable effect on humans, animals and vegetation. Present pollution monitoring is revealing that if we do not think and act cautiously then vehicle pollution could harm the environment. People need to be encouraged to use public transport or share cars whenever possible so only the minimum amount of pollution is created. Poor air quality has negative effects on the environment in which we live. Air pollution from transport includes emissions of carbon monoxide, particulates, nitrogen oxides and hydrocarbons. Ozone is a secondary pollutant produced when many of these primary chemicals react in sunlight in the atmosphere. Such primary and secondary pollutants can impact on wildlife and vegetation, and human health.

The main objective of the project is to collect the dataset from the sensor (Mq-2, Mq-135,PM 2.5) to predict the air quality. The dataset can be saved automatically by connecting the sensor with Arduino. Arduino helps to get input from sensor and the data stored in the cloud. These dataset evaluated by algorithm.

The sensor used for predicting the air quality that fixed in a indoor of smart cities(sc). The sensors mounted in a place that will collect the data. The dataset of the sensors can be stored in the cloud called THINKSPEAK. ThingSpeak is an IoT analytics platform service that allows you to aggregate, visualize, and analyze live data stream in the cloud. We can send data to ThingSpeak from our devices, create instant visualization of live data. The collected dataset are taken and the dataset reports on weather and the level of pollution, also taken for predicting the air quality. These dataset used in DEEP LEARNING (DL) algorithm RECURRENT NEURAL NETWORK (RNN) specifically LONG SHORT TERM MEMORY(LSTM) and it compared with the machine learning algorithm called SUPPORT VECTOR REGRESSION.

II.LONG SHORT TERM MEMORY

LSTM is a special kind of recurrent neural network (RNN) and capable of learning long-term dependencies. It was introduced by Hochreiter and Schmidhuber in order to overcome vanishing gradient problem in 1997. In this neural network model, a memory block takes the place of each ordinary neuron in the hidden layer of standard recurrent neural network .

The LSTM block shown in Fig. 1 has an input gate, a forget gate and an output gate which regulate the flow of information into and out of the cell. These gates, block input and block output as follows:

block input:$\boldsymbol{zt}$ = g **Wz**x! + **Rz**y!!! + b! (1)

input gate:$\boldsymbol{it}$ = σ **W**!x! + **R**!y!!! + p! ⊙c!!! + b! (2)

Forget gate:$\boldsymbol{ft}$ = σ **Wf**x! + **Rf**y!!! + p! ⊙c!!! + b! (3)

cell state:$\boldsymbol{ct}$ = $\boldsymbol{it}$ ⊙$\boldsymbol{zt}$ + $\boldsymbol{ft}$ ⊙c!!! (4)

output gate:$\boldsymbol{ot}$ = σ **Wo**x! + **Ro**y!!! + p! ⊙$\boldsymbol{ct}$ + b! (5)

block output:$\boldsymbol{yt}$ = $\boldsymbol{ot}$ ⊙$\boldsymbol{h}(\boldsymbol{ct})$ (6)

where $\boldsymbol{xt}$ is input vector at time t, $\boldsymbol{Wz}$,$\boldsymbol{Wi}$,$\boldsymbol{Wf}$,$\boldsymbol{Wo}$are the weights matrices connecting $\boldsymbol{xt}$ to the three gates and block input, $\boldsymbol{Rz}$,$\boldsymbol{Ri}$,$\boldsymbol{Rf}$,$\boldsymbol{Ro}$are recurrent weight matrices connecting $\boldsymbol{yt}$!**1** to the three gates and block input, b!, b!, b!, b! are the bias vectors. σ represents the logistic sigmoid function and h represents hyperbolic tangent function. σ is used for as activation of the gates and g is used as the block input and output activation function.

LSTM has been applied on a large variety of real world problems such as machine translation, speech recognition, neural language model and image recognition. In my project, a novel prediction model based on LSTM is proposed on IoT Smart city data.

III.SUPPORT VECTOR REGRESSSION

In machine learning, support-vector machines are supervised learning models with associated learning algorithms that analyze data for classification and regression analysis. It works really well with a clear margin of separation. It is effective in high dimensional spaces. But It doesn't perform well when we have large data set because the required training time is higher.

IV.LITERATURE SURVEY

1.PREDICTING AIR QUALITY USING MOVING SENSORS: The study was on the particular microfine dust (microdust) is known to cause serious health issues to the people. A typical way of measuring microdust using sensor in the fixed location. Modeling the air quality pattern in a given local area by using vehicles equipped with cheap IOT sensors. Inexpensive and disposable IOT sensor to collect the microdust level in the local area where sensors installed in the vehicles. Microdust sensor system, including temperature, humidity sensor, GPS, PM2.5 and PM10 microdust sensors, and Carbon Dioxide sensor, and Raspberry Pi 3B+,Arduino is used to build our microdust sensor box for moving vehicles. The sensor box are mounted in the top of a vehicle. During the experiment, the Geo tagged data is used, this consists of the sensors that predict longitude and latitude, and timestamp as well as temperature, humidity, and concentration of CO2, PM2.5. In addition, Machine learning algorithm to predict more accurate local microdust levels. The app is created to provide local microdust information to users. The algorithm used which are random forest regressor, support vector regressor and gradient boosting regressor as prediction models. The support vector regressor gives better accuracy.

2.AIR QUALITY MONITORING AND ASSESSMENT USING IOT.

The system interrelated with internet connected objects, which is used to collect and transfer data. The work based on monitoring the pollution using MQ135 sensor with GSM module. The MQ135 sensor is used in the air quality monitoring system. It can sense NH3, NOx, alcohol, Benzene, smoke, CO2, and some other gases. It gives the output in the form of voltage levels. Arduino for collecting the data from the MQ135 sensor and these data are stored. The MQ135 sensor has a copper inside it, so it sense restlessly since it does not have on and off keys. The sensor gives the value continuously but as the program that is given to Arduino displays the reading in the LCD and mobile only after storing five values to itself and display the average of the five readings. The values displayed are visible in LCD since it connected with LCD to the Arduino UNO with the help of breadboard. This GSM module gets the data from the microcontroller Arduino and sends it to mobile as a text message. Monitoring system helps us get the information without our presence which results in sending text messages to our mobile phone when needed to know the status of air in a particular place. The GSM module used to place the sim. When there is need to know the range of air, the text message sent to the sim it reflect the resultant message. when the requested message is collected by the antenna. Result obtained in the end of the experiment is the amount of polluted air in our environment which is displayed in ppm. The text message not only gives us the amount or the quality of the air it also says the status of air whether it is fresh or poor in quality.

3.REAL TIME LOCALIZED AIR QUALITY MOINTORING AND PREDICTION THROUGH MOBILE AND FIXED IOT SENSING NETWORK.

The Internet of Things (IoT) has been widely used in different domains to improve the quality life for people by connecting multiple sensors in different places, it also makes the air pollution monitoring more easier than before. The research focuses on modeling the air quality pattern in a given region by adopting both fixed and moving IoT sensors, which are placed on vehicles patrolling around the region. With these approach, a full spectrum of how air quality varies in nearby regions can be analyzed. The approach demonstrate the feasibility in effectively measuring and predicting air quality using different machine learning algorithms with real world data. This evaluation shows a promising result for effective air quality monitoring and prediction for a smart city application.

4. INDOOR AIR QUALITY ANALYSIS USING DEEP LEARNING WITH SENSOR DATA.

Indoor air quality analysis is of interest to understand the abnormal atmospheric phenomena and external factors that affect air quality. By recording and analyzing quality measurements, they are able to observe patterns in the measurements and predict the air quality of future. They designed a microchip made out of sensors that is capable of periodically recording measurements, and proposed a model that estimates atmospheric changes using deep learning. In addition, they developed an efficient algorithm to determine the optimal observation period for accurate air quality prediction. Experimental results with real-world data demonstrate the feasibility of approach.

5. A DEEP LEARING MODEL FOR AIR QUALITY PREDICTION IN SMART CITIES.

The dataset is collected for past two years to predict the future air quality index. The dataset contains 8 features including ozone, particulate matter, carbon monoxide, sulfur dioxide, nitrogen dioxide, longitude, latitude and timestamp was used for experiment. The dataset has 17568 samples that are collected at five- minute intervals. Each sample value is given in the form of EPA's AQI standard.

In prediction context, deep learning techniques have been used for several forecasting problems in big data. In this paper, a novel deep learning model is proposed for analyzing IoT smart city data. They propose a novel model based on Long Short Term Memory (LSTM) networks to predict future values of air quality in a smart city. Evaluation results of the proposed model are found to be promising and they show that the model can be used in other smart city prediction problems as well.

6. PERSONAL POLLUTION MONITORING: MOBILE REAL-TIME AIR QUALITY IN DAILY LIFE.

The CitiSense system gives individuals, the real time tools they need to be able to identify when and where they are exposed to poor air. The CitiSense system is comprised of four main components: a wearable sensor board that pairs with an Android phone, a server-supported, web-based personalized daily pollution map, and a social component supported through Facebook and Twitter integration.

The mobile component of the CitiSense system consists of an Android mobile phone running custom application and a mobile air-quality monitoring unit that sends sensor data to the phone via Bluetooth. The air-quality monitoring unit contains the following 6 sensors attached to a custom board; Carbon Monoxide (CO ppm), Nitrogen Dioxide(NO_2 ppb), Ozone (O_3 ppb), Temperature (F°), Barometric Pressure (MBAR), Humidity (reported as percentage). They developed a modified version of the Environmental Protection Agency's (EPA) Air Quality Index (AQI) number and color mapping to help our users easily and quickly interpret sensor data. While the EPA's AQI values represent an average pollutant level at a location over time, CitiSense provides an instantaneous report of the same value. A personal map page was maintained for throughout work. These pages were generated in real-time, and feature a daily exposure map, and a chart displaying pollution exposure by time of day. This webpage was designed to allow users to dig deeper into their data and see trends in their exposure. The visual nature of the time chart and map allow users to quickly locate the time and place of peak exposures. The goal of the CitiSense project is to provide individuals with a system that makes the invisible visible.

V.IMPLEMENTATION

The implementation process of the project is the dataset collected using sensors. The input stored in the cloud server. These dataset used for predicting the air quality in deep learning method and machine learning method.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 6 June 2021, pp: 18-22 www.ijaem.net ISSN: 2395-5252

SYSTEM ANALYSIS:
HARDWARE REQUIREMENTS:

• **ARDUINO UNO:**
Arduino UNO is developed by Arduino.cc which is an open source microcontroller board which is completely based on the microchip AT mega328P. It can detect the surroundings of the input.

• **MQ2 GAS SENSOR:**
MQ2 detects combustible gasses and smoke. The MQ-2 Gas sensor can detect carbon monoxide gas.

• **MQ135 GAS SENSOR:**
MQ135 detects combustible gasses and smoke. The MQ-135 Gas sensor can detect nitrogen oxides gases.

• **PM2.5:**
PM2.5 refers to the particles that have diameter less than 2.5 micrometer and suspended for long. It affects the lung and heart.

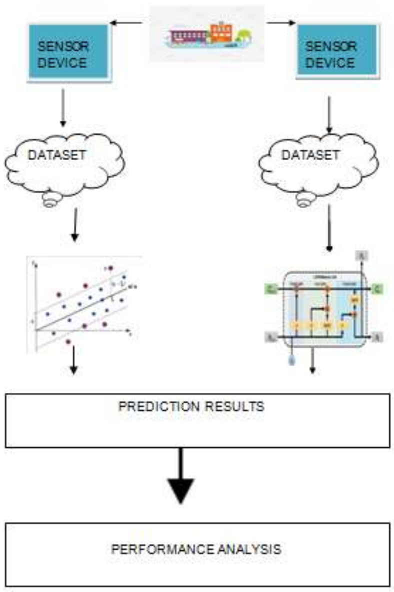
• **WIFI ESP8266:**
ESP8266 is small module allows microcontrollers to connect to a Wi-Fi network and make simple TCP/IP connections.

• **TRANSFORMER:**
A Transformer is a static electrical machine which transfers AC electrical power from one circuit to the other circuit at the constant frequency.show dynamic behavior under various operating and environmental conditions and demonstrate advantages of adaptive control over the non-adaptive type.

• **THINKSPEAK SERVER:**
ThingSpeak is an open-source Internet of Things (IoT) application and API to store and retrieve data from things using the HTTP and MQTT protocol over the Internet or via a Local Area Network.

• **PYTHON:**
Python is an interpreted, high-level and general-purpose programming language. TensorFlow provides a collection of workflows to develop and train models using Python. Keras is the high-level API of TensorFlow 2.0.

1.1 FRAMEWORK FOR PROPOSED SYSTEM:



i)DATASET COLLECTION:
The dataset can be collected through the sensor. The sensors which can able to collect data of carbon monoxide, nitrogen dioxide, particulate matter 2.5, Methane.

1.2SENSORKIT
The kit used for collecting the datas are as shown above.
The Thinkspeak cloud server helps to collect the data from the kit. The stored in cloud which can access by ourselfes using cloudserver website.

ii) THINKSPEAK SERVER:
ThingSpeak is an open-source Internet of Things (IoT) application and API to store and retrieve data from things using the HTTP and MQTT protocol over the Internet or via a Local Area Network. The arduino uno is coded to extract the gases from the air and the data updated to the server Thinkspeak.

iii)PREDICTION ACCURACY:

The evaluation prediction yi and the model's performance according to Root Mean Square Error (RMSE), is adopted as an error criteria.

VI. CONCLUSION

In this project, the algorithm of machine learning and deep learning compared with the dataset and the accuracy is found out with the prediction result that deep learning algorithm gives much accuracy than the machine learning algorithm.The long short term memory of deep learning algorithm and support vector regression of machine learning are implemented to predict the air pollution with various environmental consideration. The algorithm showed the prediction accuracy of air quality with 30000 samples data given as input to test and train. In Future the AQI values can be calculated for better results.

REFERENCES

[1]. D. Zhang and S. S. Woo, ``Predicting air quality using moving sensors(poster)," in Proc. 17th Annu. Int. Conf. Mobile Syst., Appl., Services,Jun. 2019, pp. 604-605.

[2]. DAN ZHANG 1,2 AND SIMON S. WOO3,4 “Real Time Localized Air Quality Monitoring and Prediction Through Mobile and Fixed IoT Sensing”., Stony Brook, NY 11794, USA

[3]. J. Ahn, D. Shin, K. Kim, and J. Yang, ``Indoor air quality analysis using deep learning with sensor data," Sensors, vol. 17, no. 11, p. 2476, 2017.

[4]. Kok, M. U. Simsek, and S. Ozdemir, ``A deep learning model for airquality prediction in smart cities," in Proc. IEEE Int. Conf. Big Data (BigData), Dec. 2017, pp. 19831990.

[5]. Gulshan Taj Mohammed Navi Anwar et al Air quality monitoring and assessment using IoT; 2020 IOP Conf. Ser.: Mater. Sci. Eng. 955 012006

[6]. Elizabeth Bales, Nima Nikzad, Celal Ziftci, Nichole Quick, William Griswold “Personal Pollution Monitoring: Mobile Real-Time Air-Quality in Daily Life”,University of California, San Diego La Jolla, CA 92093-0404ess

[7]. M. Kampa and E. Castanas, ``Human health effects of air pollution,"Environ. Pollution., vol. 151, no. 2, pp. 362367, Jan. 2008.

[8]. Y. Jiang, L. Shang, K. Li, L. Tian, R. Piedrahita, X. Yun, O. Mansata, Q. Lv, R.P. Dick, and M. Hannigan, ``MAQS: A personalized mobile sensing system for indoor air quality monitoring," in Proc. 13th Int. Conf.

[9]. F. Xia, L. T. Yang, L. Wang, and A. Vinel, "Internet ofThings," International Journal of Communication Systems, vol. 25, pp. 1101-1102, 2012.

[10]. E. Boldo, S. Medina, A. Le Tertre, F. Hurley, H.-G. Mücke, F. Ballester,and I. Aguilera, ``Apheis: Health impact assessment of long-term exposureto PM2.5 in 23 European cities," Eur. J. Epidemiology, vol. 21, no. 6,pp. 449-458, Jun. 2006.

[11]. S. Devarakonda, P. Sevusu, H. Liu, R. Liu, L. Iftode, and B. Nath, ``Realtime air quality monitoring through mobile sensing in metropolitan areas,"in Proc. 2nd ACM SIGKDD Int. Workshop Urban Comput.UrbComp,2013,p.15

[12]. B. Maag, Z. Zhou, and L. Thiele, ``W-Air: Enabling personal air pollutionmonitoring on wearables," Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol., vol. 2, no. 1, p. 24, 2018

[13]. J. Lin, A. Zhang, W. Chen, and M. Lin, ``Estimates of daily PM2.5exposure in beijing using spatio-temporal kriging model," Sustainability,vol. 10, no. 8, p. 2772, 2018.

[14]. S. Fotouhi, M. H. Shirali-Shahreza, and A. Mohammadpour, `Concentration prediction of air pollutants in tehran," in Proc. Int. Conf. Smart Cities Internet Things SCIOT, 2018.

[15]. A. C. Rai, P. Kumar, F. Pilla, A. N. Skouloudis, S. Di Sabatino, C. Ratti,A. Yasar, and D. Rickerby, ``End-user perspective of low-cost sensorsfor outdoor air pollution monitoring," Sci. Total Environ., vols. 607-608,pp. 691-705, Dec. 2017.

[16]. Johnston, F.R.; Boyland, J.E.; Meadows, M.; Shale E. Some properties of a simple moving average when applied to forecasting a time series. J. Oper. Res. Soc. 1999, 50, 1267–1271.

[17]. Zhu, J.; Zhang, R.; Fu, B.; Jin, R. Comparison of ARIMA model and exponential smoothing model on 2014 air quality index in Yanqing county, Beijing, China. Appl. Comput. Math. 2015, 4, 456–461.