International Journal For Technological Research In Engineering
Volume 6, Issue 7, March-2019 *ISSN (Online): 2347 - 4718*

AIR POLLUTION MONITORING SYSTEM USING IOT AND DATA ANALYTICS

Ramya D[1], Rohini V S[2], Umaa Mahesshwari G[3], Vinitha Shree D[4], Mr. Hariharan B[5]
[5]Asst Prof
Department of Computer Science and Engineering, RMK College of Engineering and Technology

Abstract: An IoT based Air pollution monitoring system includes a MQ Series sensor interfaced to a NodeMCU equipped with a ESP8266 WLAN adaptor to send the sensor reading to a ThingSpeak cloud. Further scope of this work includes a suitable machine learning model to predict the air pollution level and a forecasting model, which is basically a subset of predictive modeling. We will be using out IoT device as a prototype to collect the data, and for expanding our model we used an authorized open source dataset provided by US Govt. The paper is mainly to monitor, visualize the pollution data and its forecasting. Specifically three machine learning (ML) algorithms were implemented to find out the best predictive model and a forecasting model for calculating AQI of four different gases: Carbon Monoxide (CO), Nitrogen Dioxide (NO2), Sulfur Dioxide (SO2) and Ozone (O3). The ML algorithms used over here are Linear Regression, Random Forest and XGBoost for predictive modeling and ARIMA model for time-series forecasting. The performance metrics was based on Root Mean Square Error (RMSE) and Mean Absolute Error (MAE).

It was observed that Random Forest had the best performance. From this paper, the model can thus be deployed in real-world in areas with high-pollution An IoT based Air pollution monitoring system includes a MQ Series sensor interfaced to a NodeMCU equipped with a ESP8266 WLAN adaptor to send the sensor reading to a ThingSpeak cloud. Further scope of this work includes a suitable machine learning model to predict the air pollution level and a forecasting model, which is basically a subset of predictive modeling. We will be using out IoT device as a prototype to collect the data, and for expanding our model we used an authorized open source dataset provided by US Govt. The paper is mainly to monitor, visualize the pollution data and its forecasting. Specifically three machine learning (ML) algorithms were implemented to find out the best predictive model and a forecasting model for calculating AQI of four different gases: Carbon Monoxide (CO), Nitrogen Dioxide (NO2), Sulfur Dioxide (SO2) and Ozone (O3).

The ML algorithms used over here are Linear Regression, Random Forest and XGBoost for predictive modeling and ARIMA model for time-series forecasting. The performance metrics was based on Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). It was observed that Random Forest had the best performance. From this paper, the model can thus be deployed in real-world in areas with high-pollution An IoT based Air pollution monitoring system includes a MQ Series sensor interfaced to a NodeMCU equipped with a ESP8266 WLAN adaptor to send the sensor reading to a ThingSpeak cloud.

Further scope of this work includes a suitable machine learning model to predict the air pollution level and a forecasting model, which is basically a subset of predictive modeling. We will be using out IoT device as a prototype to collect the data, and for expanding our model we used an authorized open source dataset provided by US Govt. The paper is mainly to monitor, visualize the pollution data and its forecasting. Specifically three machine learning (ML) algorithms were implemented to find out the best predictive model and a forecasting model for calculating AQI of four different gases: Carbon Monoxide (CO), Nitrogen Dioxide (NO2), Sulfur Dioxide (SO2) and Ozone (O3). The ML algorithms used over here are Linear Regression, Random Forest and XGBoost for predictive modeling and ARIMA model for time-series forecasting. The performance metrics was based on Root Mean Square Error (RMSE) and Mean Absolute Error (MAE).

It was observed that Random Forest had the best performance. From this paper, the model can thus be deployed in real-world in areas with high-pollution An IoT based Air pollution monitoring system includes a MQ Series sensor interfaced to a NodeMCU equipped with a ESP8266 WLAN adaptor to send the sensor reading to a ThingSpeak cloud. Further scope of this work includes a suitable machine learning model to predict the air pollution level and a forecasting model, which is basically a subset of predictive modeling. We will be using out IoT device as a prototype to collect the data, and for expanding our model we used an authorized open source dataset provided by US Govt. The paper is mainly to monitor, visualize the pollution data and its forecasting. Specifically three machine learning (ML) algorithms were implemented to find out the best predictive model and a forecasting model for calculating AQI of four different gases: Carbon Monoxide (CO), Nitrogen Dioxide (NO2), Sulfur Dioxide (SO2) and Ozone (O3). The ML algorithms used over here are Linear Regression, Random Forest and XGBoost for predictive modeling and ARIMA model for time-series forecasting. The performance metrics was based on Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). It was observed that Random Forest had the best performance. From this paper, the model can thus be deployed in real-world in areas with high-pollution An IoT

based Air pollution monitoring system includes a MQ Series sensors for detecting the amount of industrial effluents in the air, interfaced to an PIC IC equipped with UART adaptor to send the sensor reading to the IOT Module (GPRS)which will in turn send the values to the php server. Further scope of this work includes a suitable machine learning model to predict the air pollution level and a forecasting model, which is basically a subset of predictive modeling. We will be using an IoT device as a prototype to collect the data, and for expanding our model we used an authorized open source dataset provided by TNPCB. The paper is mainly to monitor, visualize the pollution data and its forecasting. Specifically ARIMA algorithm was implemented to find out the best predictive model and a forecasting model for calculating AQI of four different gases: Carbon Monoxide (CO), Carbon Dioxide (CO2), Sulphur dioxide (SO2) and Temperature sensor. From this paper, the model can thus be deployed in real-world in areas with high-pollution.

Keywords: IoT, Mqseries sensors, PIC IC, Arima, Predictive modeling, Machine learning.

I. INTRODUCTION

With fast development of industrialization and urbanization pollution has become more common. Air pollution is presence of contaminants or pollutant substances that effect human health. If we know the quantity of pollutant, then proper precautions can be taken to minimize the pollution levels in air. Recent researches prove the high correlation between atmospheric pollutants and disease like asthma. The recent advancements in embedded electronics have led to the usage of wireless network technologies in monitoring sensor data and air pollution . The aim in this paper is to come out with prediction and forecasting model for certain air pollutants like CO,CO2,SO2 and temperature which are considered to be quite harmful. Two machine learning algorithm have been implemented which include Arima.These models have very good predictive capacity, generalization power and have a wide range of applications.

II. PROPOSED SYSTEM

A. IOT device Prototype:

IoT mainly deals with connecting smart devices (embedded electronics devices) to internet by harnessing the advantage of OSI layered Architecture. In the context of this work we propose a cluster of Air Quality Monitoring Sensor motes, which are used to measure the concentration of Air pollutants in the air. All the Air Sensors are interfaces with a tiny embedded platform equipped with network connectivity and are interconnected to internet making it a global network of connected things. We have used PIC 16F877A microcontroller which features 256 bytes of EEPROM data memory, self programming features which make it ideal for more advance level A/D applications in automotive, industrial, appliances and consumer applications.Mq-7(CO) , Mq-135(CO2),Mq-136(SO2) and temperature sensors are used to collect the gas concentration measurements. This sensor data would be captured and sent to php server for IoT based data acquisition.

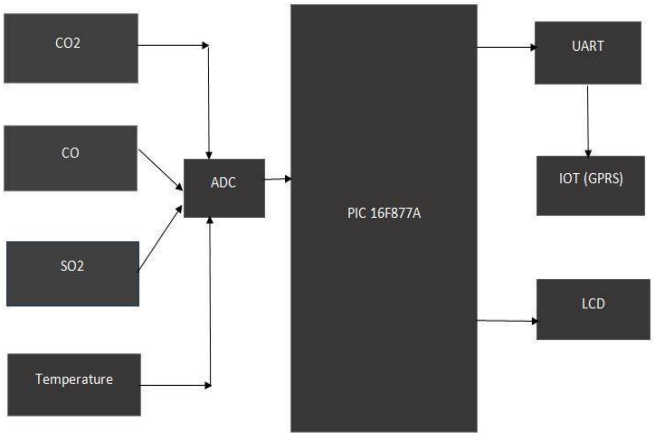

Fig.1: Block Diagram of proposed IoT based air pollution monitoring system

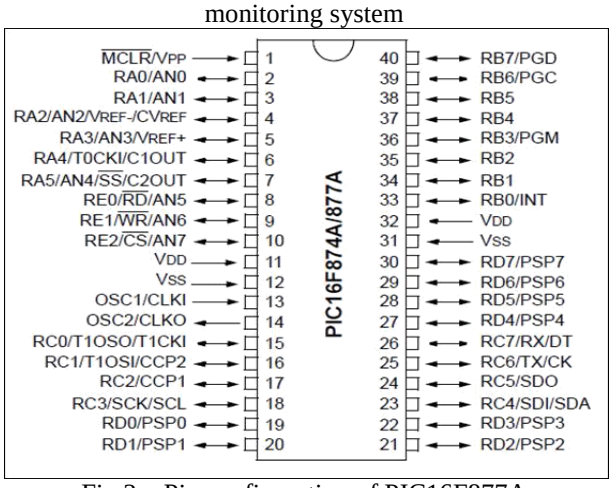

Fig.2 – Pin configuration of PIC16F877A

Fig.3- Mq-7 Gas Sensor

Further data processing can be done on the data obtained from the php server rather than on the device to reduce the computational over head of the tiny embedded devices as they are constrained low power devices and mostly powered by battery.

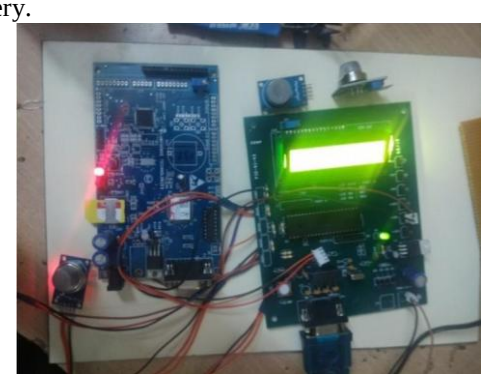

Fig.4 shows the implementation of proposed device to monitor air quality.

B. Predictive and Forecasting Modeling:

We aim to accurately predict concentrations of CO2, SO2, CO and temperature. Using the historical data of the gas sensors and their AQI value we try to obtain a predictive model that indicates the graphical representation of the AQI value by one-step ahead forecast and dynamic forecast implemented by ARIMA algorithm. The process of building the prediction models mainly deal with 3 steps:

- Data Pre-processing: The first step of building a prediction model is data pre-processing where data is cleaned, missing values are filled, outliers are removed and also data is arranged in a way to fit for the Machine Learning algorithms.
- Feature Engineering: Features are one of the major factors which increase the prediction accuracy such, day, month, time of the day, etc.
- Building Forecasting Model: Model is built to predict the future, i.e. on the unseen data based on the historical data.

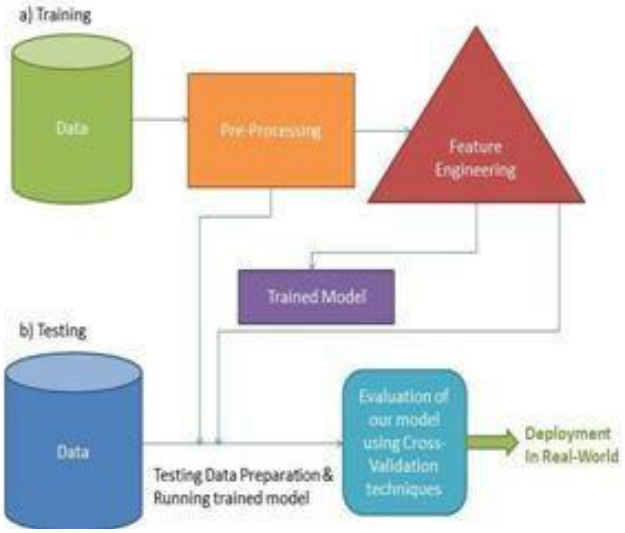

Fig.5 Architecture to build predictive model

```
In [42]: y_forecasted = pred.predicted_mean
         print(y_forecasted)
         y_truth = data['2018-10-23':]
         print(y_truth)

Date
2018-10-23    120.858173
2018-10-26    126.556226
2018-11-16    120.279137
2018-11-19    121.331396
2018-11-22    107.507738
2018-11-26    118.753330
2018-11-28    130.979785
2018-11-29    122.336238
2019-02-01    117.026381
2019-01-13    115.190405
2019-01-21    110.621976
2019-01-23    129.627457
2019-01-28    105.187684
2019-01-30    108.815296
dtype: float64
              AQI
Date
2018-10-23  131
2018-10-26  117
2018-11-16  118
2018-11-19   87
2018-11-22  111
2018-11-26  141
2018-11-28  123
2018-11-29  109
2019-02-01  104
2019-01-13   94
2019-01-21  137
2019-01-23   85
2019-01-28   90
```

Fig.6 Predicted data VS Actual data

Autoregressive integrated moving normal (ARIMA) is one of the well known linear models in time arrangement determining amid the previous three decades. Late research exercises in anticipating with artificial neural systems (ANNs) recommend that ANNs can be a promising contrasting option to the customary linear methods. ARIMA models and ANNs are regularly contrasted and blended conclusions regarding the prevalence in forecasting models. The proposed System architecture interprets the training and testing data set.

III. RESULT ANALYSIS

From the context of the work here are the insights and analytics we got from the data that we generated. The graphical representation is the AQI values predicted for one particular area.

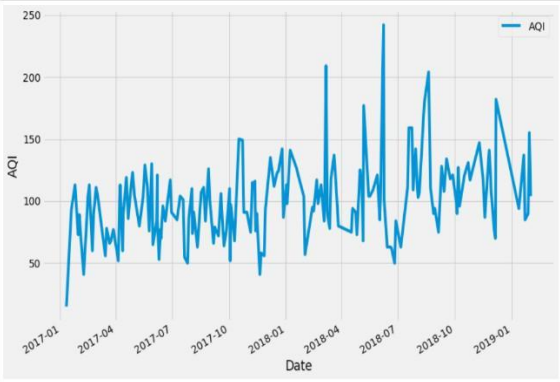

Fig.7(a) Result(Graph) generated by the given data set.

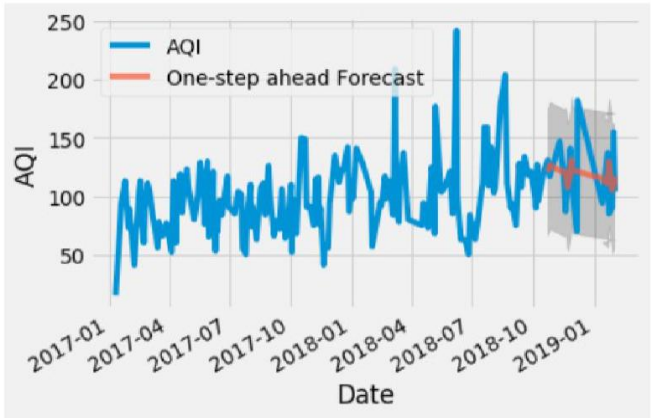

Fig.7(b) Result(Graph) generated by One step ahead forecast method using ARIMA algorithm (prediction).

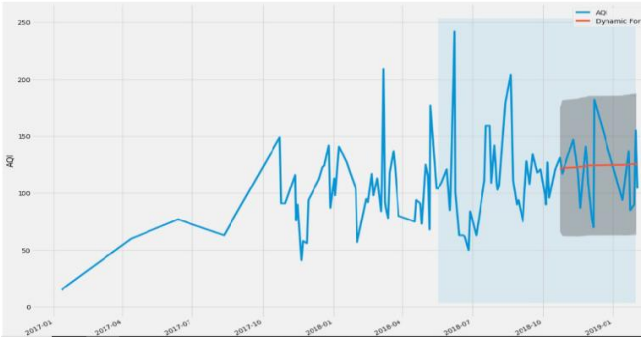

Fig.7(c) Result(Graph) generated by Dynamic Forecast model using ARIMA algorithm (prediction)

IV. CONCLUSION

Air quality is a critical issue that straightforwardly influences human wellbeing. Air quality information are gathered remotely from checking bits that are outfitted with a variety of vaporous also, meteorological sensors. This information are investigated and utilized as a part of anticipating fixation estimations of contaminations utilizing savvy machine to machine stage. The stage comprises of a ML-based calculations to construct the estimating models by training from the gathered information .ARIMA performs pretty well as a forecasting model, what can be used as to make a daily forecast just like regular weather forecasting.

REFERENCES

[1] Ayaskanta Mishra "Air pollution monitoring system based on IOT :Forecasting and predictive model using Machine learning". International Conference on Applied Electromagnetics ,Signal Processing and Communication.2018

[2] Bhavika Bathiya; Sanjay Srivastava, Biswajit Mishra; "Air Pollution Monitoring Using Wireless Sensor Network," IEEE International WIE Conference on Electrical and Computer Engineering, 2016.

[3] Snehal Sirsikar, Priya Karemore; "Review Paper on Air Pollution Monitoring system," International Journal of Advanced Research in Computer and Communication Engineering,2015.

[4] Abdullah Kadri ; Elias Yaacoub ; Mohammed Mushtaha ; Adnan Abu-Dayya; "Wireless Sensor Network For Real-Time Air Pollution Monitoring" 1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA) ,2013.

[5] Liu ; Yu-Fan Chen ; Tzu-Shiang Lin ; Da-Wei Lai ; Tzai-Hung Wen ; Chih-Hong Sun , "Developed urban air quality monitoring system based on wireless sensor networks," Fifth International Conference on Sensing Technology , 2011.

[6] Boyanka Marinova Nikolova, Marin Berov Marinov, Georgi Todorov Nikolov ; "Air Quality Monitoring System,"2006

