



COLLEGE OF SCIENCE AND TECHNOLOGY

Household Pollution Monitoring and Prediction System Using Internet of Things: A Case of Tanzania

By

Enatha Rweyemamu

College of Science and Technology

Master of Science in Internet of Things (Embedded Computing System)

2021





COLLEGE SCIENCE AND TECHNOLOGY

Household Pollution Monitoring and Prediction System Using Internet of Things

By

Enatha Rweyemamu

Reg. No.220013207

A dissertation submitted in partial fulfillment of the requirements of the degree of

MASTERS IN INTERNET OF THINGS (EMBEDDED COMPUTING SYSTEM)

In College of Science and Technology

Supervisor: Dr. Musabe Richard

Co-Supervisor: Dr. Kizito Nkurikiyeyezu

2021

Student Declaration

I declare that this Dissertation contains my own work except where specifically acknowledged.

Enatha Rweyemamu

220013207

Signature: …………

Date: 10 November 2021

Bonafide Certificate

This is to certify that the project entitled "Household Pollution Monitoring and Prediction System Using Internet of Things: A Case of Tanzania" is a record of original work done by Enatha Rweyemamu with registration number 220013207 in partial fulfillment of the requirement for the award of Masters of Science in Internet of Things in College of Science and Technology, University of Rwanda, Academic year 2019/2021.

This work has been submitted under the guidance of Dr. Musabe Richard and Dr. Kizito Nkurikiyeyezu.

Ass. Prof. Musabe Richard

Dr. Kizito Nkurikiyeyezu



Signature: 

Signature:

Date: 10/12/2021

Date: 10/12/2021

The Head of Masters and Trainings

Dr. James Rwigema

Signature: ……

Acknowledgements

This paper the study and research behind it would not have been possible without the exceptional support and guidance from different people to whom I am very thankful for.

My sincere gratitude goes to the Almighty God for everything through my research.

I thank Dr. Musabe Richard my supervisor and Dr. Kizito Nkurikiyeyezu his co-supervisor for their knowledge, inspiration and support that kept my work on track till the end.

I am most thankful to University of Rwanda for their financial support for the project. I also thank my colleagues at University of Rwanda for their support and unfailing patience to my numerous questions.

I acknowledge the contribution, encouragement and endless support from my family, especially my husband thanks for being patient with my busy schedules. I thank my parents for their prayers, support and for always believing in me.

Abstract

Indoor pollution has become a concern all over the world. The means of source of fire in low-income and middle-income countries use fire-wood, charcoal and liquefied petroleum gas (LPG) as for cooking daily. According to the world health organization an estimated 1.6 million people die annually due to indoor pollution. The Institute for Health Metrics and Evaluation (IHME) found out that 6% of deaths in low-income countries are caused by indoor air pollution. Indoor air pollution increases risks to contracting non communicable diseases such as heart diseases, stroke, lung cancer and chronic obstructive pulmonary disease (COPD), lung cancer and pneumonia among others. Among the main sources of household air pollution are harmful gasses from wood fuel and Liquefied Petroleum Gas (LPG). Even though, attempts have been made to come up with solutions using different technologies to monitor the pollution levels and give appropriate alerts. There is a need to develop a solution that will not only monitor but also predict pollution levels so that corrective measures may be taken early enough. This study aims at developing a prototype for a household air quality monitoring system and pollution prediction system using Artificial Intelligence (AI). The proposed solution will involve the installation of the system in the house with the use of a low cost Arduino based microcontroller that incorporates sensors to monitor carbon monoxide (CO) and particulate matter and environmental sensor to monitor temperature and humidity. The collected data will be sent to Cloud via Global System for Mobile (GSM) for data storage, processing and analysis. The alerting system will be developed along with the prediction system using Artificial Intelligence. Pollution early warning alerts generated appropriately. The expected results from the study are to collect data from a kitchen, alert people using the kitchen when the air quality is poor in real time, and also the prediction of air quality so as to minimize and later avoid the exposure of poor air quality in households. This work deployed sensors in the household connect them with ArduinoUNO and send the collected data to cloud platform known as Thingspeak with the use of GSM module, the data was further used on air quality prediction using Machine Learning algorithms, the air quality widget will clearly show the level of air pollution in colors. This system will help reduce the health risks relating to indoor air pollution and also provide data to be used by government, NGOs environmental bodies and agencies.

Key Words: *Indoor air pollution, Internet of Things (IoT), Machine Learning (ML), Artificial intelligence (AI), GSM, Cloud*

Abbreviations

IoT: Internet of Things

AI: Artificial Intelligence

IAQ: Indoor Air Quality

AQI: Air Quality Index

GSM: Global System for Mobile

LPG: Liquefied Petroleum Gas

WHO: World Health Organization

PM: Particulate Matter

CO: Carbon Monoxide

VOC: Volatile Organic Compound

ML: Machine Learning

List of Figures

Figure 1: African woman cooking in the kitchen

Figure 2: Block Diagram of Proposed System

Figure 3: MQ7 Circuit

Figure 4: Buzzer to ArduinoUno

Figure 5: Common Emitter Amplifier

Figure 6: Schematic system

Figure 7:MQ7 Gas Sensor

Figure 8: MQ7 Pinout

Figure 9: PM2.5 Sensor

Figure 10:PM2.5 Pinout

Figure 11:DHT11 Sensor

Figure 12: DHT11 Pinout

Figure 13: ArduinoUNO

Figure 14: ArduinoUNO Pinout

Figure 15: GSM800 Module

Figure 16: GSM800 Pin-out

Figure 17: LCD

Figure 18: LCD Pin-out

Figure 19: Simulation of the System

Figure 20: Deployment of the System in the kitchen

Figure 21: Tempe verse Time

Figure 22: Humidity verses Time

Figure 23: CO verses Time

Figure 24: PM2.5 verses Time

Figure 25: Graph of AQI Widget

Figure 26: Graph of Mean Square Error Plot

Figure 27: Graph of Gradient Plot

Figure 28: Graph of Error Histogram Plot Regression Plot

Figure 29: Graph of Autocorrelation Plot

List of Table

Table 1: Literature Review

Table 2: Dust Sensor connected to ArduinoUNO

Table 3: Components Specifications

Table 4: AQI levels with corresponding colors

Table 5: Indices Categories and corresponding concentration Breakpoint for each pollutant (ug/m3)

Table of Contents

Student Declaration i
Bonafide Certificate ii
Acknowledgements ii
Abstract v
Abbreviations vi
List of Figures vii
List of Table ix
Table 1: Literature Review ix
Table 2: Dust Sensor connected to ArduinoUNO ix
Table 3: Components Specifications ix
Table 4: AQI levels with corresponding colors ix
Chapter 1 1
1.1 Introduction 1
1.2 Problem Statement 2
1.3 Objective 3
1.4Research Questions 4
1.5Significance of the Study 4
1.6Scope of the Study 4
1.7 Motivation of the Study 4
1.8Organization of the Document 5
Chapter 2: Literature Review 6
2.1 IoT in Indoor Air Quality Monitoring 6
2.2 Indoor Air Pollution Prediction 7
2.4 Identified Gaps 12
Chapter 3: Methodology 13
3.1 Data Collection 13
3.2 Data Processing and Analysis 14
3.2.1 Sensing Unity 14
3.2.2 Microcontroller 16

3.2. 3 Connectivity 16
3.2.4 Monitoring 16
3.2.5 Prediction 17
3.2.6 Power Supply 18
3.3 Proposed Solution 20
Chapter 4: System Design and Implementation 21
4.1 System Design 21
4.1.1 The Sensing Layer 22
4.1.2 Processing Layer 24
4.1.3 Connectivity Layer 26
4.1.4 Monitoring, Storage and Analysis Layer 29
4.1.5 Air Quality Index 30
4.1.6 Prediction Layer 32
4.2 System Implementation 33
CHAPTER 5: Results and Data Analysis 35
5.1 Data Transmission, Storage and Monitoring 35
5.2 Air Quality Index Calculation 39
5.3 AQI Prediction using 41
CHAPTER 6: Conclusion 48
6.1 Limitations 48
6.2 How to Protect Ourselves from Indoor Air Pollution 48
6.3 Potential Impact 48
6.4 Future Work 49
References 50

Chapter 1

1.1 Introduction

People spend more time indoors exposing them to indoor air pollution repeatedly increasing associated risks. Indoor air pollution has been attributed to an estimate 4 million premature deaths yearly across the globe [1]. Indoor air pollution affects the productivity of individuals and is reported to be behind a loss between $20 to $ 200 billion per year [1]. The closed spaces may lead to a faster buildup of the pollutants making indoor air pollution to be at higher levels as compared to outdoor pollution levels [2]. The main source of indoor air pollutions is the use of solid fuels such as wood fuel, coal and biomass which when not fully combusted in poorly ventilated homes, can lead to increased level of toxic gasses such as carbon monoxide, nitrogen oxides benzene and particulate matter [3]. These can lead to chronic health problems when people are continually exposed to them. The rise of indoor pollution in urban settings may also be attributed to other sources such as the materials used for building, air conditioning systems, ventilation systems, heating systems, use of products reach in chemicals and other human related activities. A study in [4] is aimed at monitoring chemical air pollution in patients room. These are unique conditions that may not apply in household due to different source of pollutants, this also applies to a study in [19] which focuses on the pollution conditions in a hairdressing saloon where many chemicals are used, in this study a need for household's source of energy was considered and evaluated. . Dawei [1] proposed a web service to mobile devices to help users send photos to air pollution. The method includes GPS location data to retrieve the assessment of the quality of the air nearby air quality stations and they also applied convolution neural network on the photos uploaded to predict the air quality. The study has less errors but with less accuracy in stability of learning and not affordable to majority of people in Africa. Sharma [2] sequential modeling was proposed to monitor and predict air pollution from urbanization. The wireless sensor network and prediction was done using historical air quality that can be trick to understand to viewers hence need to have a system with an easy budget for all levels of people to understand and take precaution on time

[4].

The emerging concept of Internet of Things provides capabilities which have been exploited to develop indoor air pollution monitoring and alerts systems [5], [6],[1]. With IoT, sensors are used to detect levels of pollution in real time with the collected data being transmitted either to the cloud or processed locally. Different notification methods are used with most focusing on mobile user alerts. Getting real time alerts alone may not help effectively to solve the problem. Prediction of pollution prior to high levels of pollutants being exposed to people, is mostly needed. The use of prediction can be quicker, cheaper and also noninvasive when compared to measuring concentration of indoor pollutants. This can also help in planning by proving some long term expectations of household air quality in given areas [7].



Figure 1: African woman cooking in the kitchen

1.2 Problem Statement

According to the World Health Organization (WHO), countries that experiencing high levels of indoor air pollution and have exceeded the allowed limits for pollution from gases and particulate matter and Volatile Organic Compounds (VOCs) resulting in health hazards. Poor air quality has led to an increase in the risk of contracting life-threatening diseases such as heart diseases, stroke, bronchitis, and lung cancer to mention but a few. Indoor air pollution accounts for over 4 million deaths annually making it a major health risk [8]. There are economic risks as a result of air pollution. The estimated costs that are associated with health damage from air pollution at a global level is 5.7 trillion dollars, which is equivalent to 4.8 percent of global gross

domestic product [9]. A study in [6] notes that the use of IoT technologies in monitoring indoor air pollution is on the rise with most solutions processing data in the edge.

The proposed study clearly tender in providing first, a need to alert the users when the air quality is poor inside their homes, second to integrate cloud based technologies that provide powerful visualization and analytics to improve the efficiency and effectiveness of poor air quality by showing carbonmonoxide (CO) level, particulate matter (PM) level and the temperature and humidity of the kitchen in real time also the AQI color codes will be displayed too. Lastly, the existing systems are not tailored for the African household settings in terms of source of energy used in kitchen hence the system will cater for that also the use of real time monitoring system is considered invasive and costly unlike if AI is used. The proposed study is aimed at addressing this challenges and thereby mitigating the cases of indoor air pollution.

1.3 Objective

The aim of this study is to develop a prototype for indoor air pollution monitoring system and a prediction model using AI in households by considering a need to have a system to monitor air pollutants in homes, to alert and predict the pollution. This will reduce the breathing issues and deaths. The objective can be achieved by developing an affordable system, need to indicate at what time the indoors are not safe also to educate others on how to avoid pollution in homes. The system was carefully developed b considering the following objectives such as

i. To investigate the available open source technologies that can be used in monitoring indoor air pollution.
ii. To investigate and develop a real-time household pollution monitoring system using IoT that involves hardware devices.
iii. To configure the hardware and link it to the Cloud for data visualization, storage and analysis by deploying pollution monitoring systems to selected households.
iv. To propose and develop an AI based household pollution prediction algorithm and early warning system.

1.4Research Questions

i. How harmful is air quality in indoor African holdings?
ii. How is indoor air quality monitored in African holdings?
iii. How is data collected from African holdings stored and analyzed?
iv. How is indoor air quality predicted from African holdings?

1.5Significance of the Study

i. This study will bring awareness of indoor air pollution risks to people, where air quality authorities, agencies and individuals will make proper decisions based on data collected and predicted by this study.
ii. This study will highly reduce hospitalization risks caused by indoor air pollution after being aware of it.
iii. The NGOs and government will be able utilize the data obtained by this study to implement measures, rules and regulations.
iv. Researchers also will benefit from this study by accessing the data and being able to find a gap in this area for further studies.

1.6Scope of the Study

This study will focus on indoor air quality monitoring and prediction in African households in Dodoma-Tanzania. The CO sensor, PM2.5 sensor and environmental sensor of temperature and humidity will be connected to other IoT hardware and software.

1.7 Motivation of the Study

This research was conducted due to the fact that most African homes use non-combusted fuel as source of fire leading to health problem and deaths to women and kids who spend most of their time in the kitchen and in the house that is polluted.

1.8Organization of the Document

This document is organized as follows: Chapter 1 comprises of Introduction, Problem Statement, Objectives, Research Questions, Significance and Scope of the study. Chapter 2: Literature review involves related works in series. Chapter 3: Methodology comprising the general introduction Chapter 4: System Design and Implementation involving description of components and platforms used, experiment setup and deployment of the prototype. Chapter 5: Results and Discussion comprises of analysis of collected data, visualization and interpretation of results and Chapter 6 Concluded the study.

Chapter 2: Literature Review

Air quality from indoor and outdoor has been studied, in this section the studies that have been done that relate to the proposed study have been analyzed beginning with the most recent studies available in online databases. Internet of things based air quality monitoring solution are first reviewed and thereafter AI technologies for indoor air pollution prediction.

2.1 IoT in Indoor Air Quality Monitoring

To begin with, Smart Air pollution system was proposed in [6]. In this solution sensors were used to measure the concentration of carbon dioxide, aerosol, Volatile Organic Compound (VOC), carbon monoxide, temperature and humidity to monitor indoor air quality. The data is transmitted to the cloud using LTE technologies where user can then view the data from a web portal. This solution is expensive hence the need of affordable system for all levels of people to use, the e-nose [10] a proposed air quality monitoring in real time. The system the capability of monitoring selected air parameters like carbon dioxide, carbon monoxide, particulate matter, nitrogen dioxide, temperature and humidity. The ESP32 Wi-Fi microcontroller is used with data being sent to the blink cloud platform. The blink platform limits ability to get alerts to only users within the same network with the system.

In a systematic literature review done in [1] it was noted that there is a need for more studies that focus on real time automatic alerts for indoor air quality conditions and the need to integrate automatic control of ventilation systems. The study also recommends the need to use latest technology of calibrated sensors in designing the real time monitoring application so as to enhance accuracy and efficiency. A study in [11] also lacks an automatic real time alert. It is based on an Arduino microcontroller to monitor the level of carbon monoxide indoors using an infrared gas module and display on a screen. In [5] a prototype is proposed that uses an Arduino microcontroller and an MQ135 sensor to measure a variety of toxic gasses in a room and displays on a screen. This study lack automatic alert functionality and does also not take into consideration other causes on indoor air pollution.

A study in [4] is aimed at monitoring chemical air pollution in patients room. These are unique conditions that may not apply in household due to different source of pollutants, this also applies

to a study in [19] which focuses on the pollution conditions in a hairdressing saloon where many chemicals are used. Elena [3] had a study on emission of volatile organic compounds, formaldehyde and ozone from a laser printing devices and the estimated of same substances through a paper filters operating through a mechanism of filtration surface with interstitial and penetration of particles into matrix filter on agglomeration. The system had a gap of not having an alerting system. Sofya [4] had a study on monitoring the content of carbondioxide in indoor air based on Arduino Uno microcontroller board,MH-219 NDIR infrared gas module, temperature and humidity module and LCD1602 module that was interfaced to i2C daughter board, same as [3] there was no alerting system in real time. Goncalo [5] had a study on reviewing air quality systems based on Internet of Things and wireless sensor networks in 2014-2019 by focusing on architecture, microcontrollers, connectivity and sensors used by the systems. It concluded that 57% of systems use Internet of Things (IoT), 33% use wireless architectures, Carbondioxide (CO2) and Particulate Matter (PM) were mostly used parameters with 67% and 29% respectively. Rui [6] had a study describing the low cost indoor air quality monitoring wireless sensor network system. It was developed using Arduino, Xbee modules and micro sensors such as temperature, humidity, carbonmonoxide, carbondioxide and luminosity and web portal was created for accessing the data in real time. The study used Xbee technology that has low data speed and low network stability. Sagar [7] had a study of prototyping an environmental air pollution monitoring system that monitored major air pollutant gases such as carbondioxide (CO2), carbonmonoxide(CO), sulphurdioxide (SO2), nitrogen dioxide (NO2) using semiconductor sensors. Data from sensors was sent to raspberry pi that acted as a base station and a mean stack was created for displaying data. The study worked on environment hence need for indoor system too to see how it affects human health.

2.2 Indoor Air Pollution Prediction

Vema [3] explained Bidirectional Long-Short Memory (BiLSTM) model by showing its benefit in forecast of air pollution with prediction of models of the long term, short term and critical consequence of PM2.5 severity levels at 6h,12h and 24h,from results the 12h results were consistent compared to those from 6h and 24h. Dawei [1] proposed a web service to mobile devices to help users send photos to air pollution. The method includes GPS location data to retrieve the assessment of the quality of the air nearby air quality stations and they also applied

convolution neural network on the photos uploaded to predict the air quality. The study has less errors but with less accuracy in stability of learning and not affordable to majority of people in developing countries such as Africa. Sharma [2] sequential modeling was proposed to monitor and predict air pollution from urbanization. The wireless sensor network and prediction was done using historical air quality that can be trick to understand to viewers hence need to have a system hence need for a system with both historical, colors and number displays and take precaution on time. Joovan [9] had a study on predicting the condition of indoor air quality through artificial neural network by using historical data from five days measurement consecutively in which the measurements took eight hours per day using Yes Plus LGA meter to capture the existing CO2. The results were further used for prediction using artificial neural network approaches getting the results as strong relation between input and target data using Levenbrg-algorithm hence need to use other kinds of algorithm like the one used in this study which was Bayesian algorithm. Ziyue [10] had a study on different machine learning approaches to determine the best in solving air prediction. The pollution data of Particulate Matter of 2.5 micrometer was collected form web-based resources and the data was analyzed using different learning models including linear regression, artificial neural network, long short term memory recurrent neural networks with their advantages and disadvantages. The study did not have the real time access to PM for real time prediction of pollution.

2.3 Air Monitoring using GSM

Mohanraj [11] developed a system for vehicle monitoring using GSM module. Threshold values were set and when high the motor was stopped and alarm starts. Level of pollution was displayed on LCD, messages were sent to RTO and the service center. PIC microcontroller was used. Afrah [12] detected gases and parameters and displayed them on LCD. They used the SIM900 GSM for connectivity and PIC 16f877A as a microcontroller. Fadi [13] used a mobile DAQ unit to collect air pollution levels convert them in frames that were sent to GPRS modem and to the pollution server using public mobile network as a outdoor system.

Table 1: Literature Review

STUDY	PROBLEMS	POLLUTANT	TECHNOLOGY
A System Based on the Internet of Things for Real-Time Particle Monitoring in Buildings-2018	contributing to the risk of developing cardiovascular and respiratory diseases, as well as of lung cancer.	Particulate matter	WEMOS D1 mini microcontroller and a PMS5003 sensor Wi-Fi connectivity
An Indoor Monitoring System for Ambient Assisted Living Based on Internet of Things Architecture-2016		air temperature, moistness, carbon monoxide, carbon dioxide, and glow	Arduino, ESP8266, and XBee technologies sensors SHT10, MQ7, T6615, LDR
Real-Time Monitoring of Indoor Air Quality with Internet of Things-Based E-Nose -2019	Air pollution can cause many serious health problems such as respiratory, cardiovascular and skin diseases in humans	CO2, CO, PM10, NO2 temperature and humidity	GP2Y1010AU, MH-Z14, MICS-4514 and DHT22 sensor array can be monitored via the 32-bit ESP32 Wi-Fi controller
Indoor Air Quality Monitoring Systems Based on Internet of Things: A Systematic Review-2020	*potential causes of several chronic health problems such as lung cancer, cardiovascular disease, and respiratory infections*	*70%, 65%, and 27.5% of studies focused on monitoring thermal comfort parameters, CO2, and PM levels with few on VOCs*	*there are 37.5% and 35% of systems based on Arduino and Raspberry Pi controllers 28 used WiFi with 1 using GSM/GPRS*

Air Quality Arduino Based Monitoring System-2019	Health problems	VOC,CO2	Arduino, Bluetooth
Air Quality Monitoring: The Use of Arduino and Android-2016	Health problems	CO, CO2 dust	GP2Y1010AU0F, MQ-9, MQ-135 Arduino Bluetooth
Development of an Arduino-based CO2 Monitoring Device-2019	Health issues	CO2, Temperature and Humidity	Arduino Uno microcontroller board, MH-Z19 NDIR infrared gas module, temperature and humidity module AM2302 and the LCD1602 module
Development of an IoT-Based Indoor Air Quality Monitoring Platform -2020	Health Issues	aerosol, VOC, CO, CO2, and temperature-humidity to monitor the air quality	microcontroller, pollutant detection sensors, and LTE modem
IoT Enabled Proactive Indoor Air Quality Monitoring System for Sustainable Health Management-2019	Health	tracking the ozone concentrations near a photocopy machine	ArduinoBT, Raspberry PI, WiFi Connectvity
Monitoring Indoor Air Quality by using	Study hygrothermal comfort-Health	Temperature, Humidity, Pressure,	Raspberry Pi 3, CoAP

IoT Technology-August 2018		Particulate Material, TVOC and eCO2	
Research on an Iot Based Air Pollution Monitoring System-July 2019	Health	Suspended Particulate Matter, NOx, and SO2	Node MCU, MQ-2, Wireless/ISP,Machine Learning,
Monitoring Indoor Air Quality to Improve Occupational Health-2016	HVAC-Health	temperature, relative humidity, (CO), (CO2) and luminosity.	Wireless/ZigBee, Arduino, SHT10 , MQ7-CO, T6615-CO2, LDR-Light,
An IoT Based Low Cost Air Pollution Monitoring System-October,2017	Health	CO,CO2,NH3	(Raspberry pi 3, Wi-Fi(TCP/UDP stack) ,MQ7, MQ135(CO2,NH3),
Indoor Pollution in Work Office: VOCs, Formaldehyde and Ozone by Printer-2014	Ventilation-Health	VOC,Ozone,Formaldehyde	-
IoT Enabled Proactive Indoor Air Quality Monitoring System for Sustainable Health Manag.	Health	Ozone	Arduino BT, Raspberry Pi-3, semiconductor O3 sensor
Assessment of Indoor Pollution in a School Environment through both Passive and	Health	VOC, CO2, CO, Temperature, Humidity	Photoionization detector (PID), Non-dispersive infrared, Electrochemical, Thermistor,

Continuous Samplings-July 2015			Capacitive
Arduino-Based Real Time Air Quality and Pollution Monitoring System-July 2018	Health	O3, CO, CO2, NO2	Arduino UNO, MQ-135 , LCD display
A Modular IoT Platform for Real-Time Indoor Air Quality Monitoring-February 2018	Health	SO2,NO2,O3,CO, CO2,CL2,T and RH	Arduino Uno, ZigBee,4-SO2-20,4-NO2-20,OX-A4-31,INE20-CO2P-NCVSP,4-C0-50,4-CL2-50,BME280

2.4 Identified Gaps

Different papers were read and discovered that many of the systems use the Internet of Things (IoT) technology in the different ways according to where the system is applied. The systems for outdoor, indoor, industries, offices, hospitals, salons and vehicles were developed. Many systems were developed from India and very few from Africa. Recent technologies such as 5G and LPWAN, Zigbee, Sigfox and Xbee were used.

The gaps that this proposed system addresses are the use of the technologies that were mostly used in reviewed systems by developing a system with technologies that are commonly used in Africa such as cellular.

An affordable system is developed to be used by all or majority people, this addresses the expensive systems.

Indoor system monitoring the pollution and hazardous gases that African women gat exposed too was developed to evaluate on how high they are exposed to them.

Prediction system on neural network with high accuracy of predicting had to be developed too.

Chapter 3: Methodology

The proposed system involved the qualitative and quantitative research was used with the main goal of identifying the existing studies and figure out the gaps so as to have a better and affordable system. The qualitative approach was done by examining the existing systems find out what to improve from them in terms of performance. The quantitative approach was done by collecting data, analyze them and use the data collected for prediction.

The system was developed with sensing unit, microcontroller unit, connectivity unit, monitoring unity and prediction unit as shown below

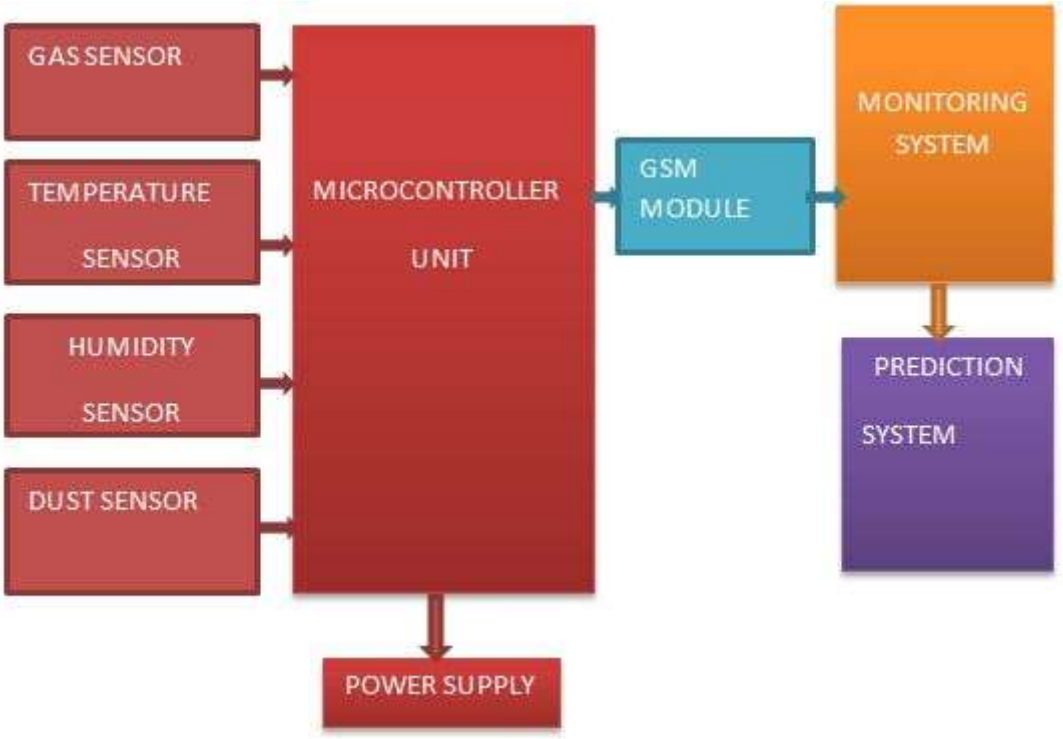


Figure 2: Block Diagram of a proposed system

3.1 Data Collection

The data collection was completed by searching for available, good quality and reasonable prices. On sensing unit different sensors were evaluated and the gas sensors, dust temperature and humidity sensors were clearly investigated and the best ones were used for the proposed system. On the microcontroller unit as the heart of the proposed system, several microcontrollers were evaluated and the affordable with high performance microcontroller was used, which helped to connect the sensing unit to other units. The connectivity system was clearly evaluated and came up with the best module foe sending collected data to the cloud. The monitoring

system was developed with the help of an application were collected data was stored and later used for prediction and the prediction was done on different algorithms and the best algorithm with high accuracy was used. The entire system was powered by the power supply for it to function well.

3.2 Data Processing and Analysis

3.2.1 Sensing Unity

The sensing unit had sensors which are temperature sensors, humidity sensors, gas sensors and dust sensors. On temperature sensors different sensors such as LMT84LP with 0.4^{o}C accuracy, ranges from -50^{o}C to 150^{o}C and use voltage from 1.5V to 5.5V which had low accuracy and was best for industrial usage, LM35DZ had 1.5^{o}C accuracy, and DS18B20 had accuracy of 0.5^{o}C which was also low, DHT11 was of low cost used 3V to 5V, 2.5 mA with 0-50^{o}C temperature and +-2^{o}C accuracy and 20-80% humidity with 5% accuracy and DHT22/AM2302 (wired version) had 0-100% humidity with 2-5% accuracy and -40C to 80^{o}C temperature with 0.5^{o}C accuracy DHT22 being the best followed by DHT11 and were both capacitive and digital sensors. In this proposed work the DHT11 was used due to its accessibility.

Gas sensors were evaluated and a series of MQ sensors was found, MQ2 sensor was found that and it has high sensitivity in Methane, Butane, LPG and smoke. MQ3 sensor has high sensitivity in alcohol, Ethanol and smoke. MQ4 has high sensitivity in Methane, CNG gas. MQ5 has high sensitivity in natural gas and LPG, MQ6 has high sensitivity in LPG and Butane gas. MQ7 has high sensitivity in CO gas. MQ8 sensor has high sensitivity in Hydrogen gas. MQ9 sensor has high sensitivity in CO and flammable gasses. MQ131 sensor has high sensitivity in ozone. MQ135 sensor has high sensitivity in CO2, Ammonia, Benzene, alcohol and smoke, MQ136 sensor has high sensitivity in Hydrogen Sulfide gas. MQ137 sensor has high sensitivity in Ammonia, MQ138sensor has high sensitivity in Benzene, Toluene alcohol, acetone, propane, formaldehyde gas, hydrogen gas. MQ214 sensor has high sensitivity in Methane and natural gas. As it shows from our study on gas sensors MQ9 and MQ7 both detect CO bur MQ9 can detect Carbonmonoxide (CO) gas ranging from 10-500ppm while the MQ7 sensor detects CO gas ranging from 1-1000ppm making it ideal.

MQ7 sensor standard working condition is Temperature between -20^{o}C +-2^{o}C and Relative Humidity of 65% +-5% and Road resistance of 10Kn+-5% with operating conditions OF -20^{o}C to +50^{o}C and <- 95%H it circuit comprises of two parts which are heating circuit and signal circuit. The heating circuit has time control function represented by High Voltage (VH) and Low Voltage Circularly (VC) as shown in the figure below

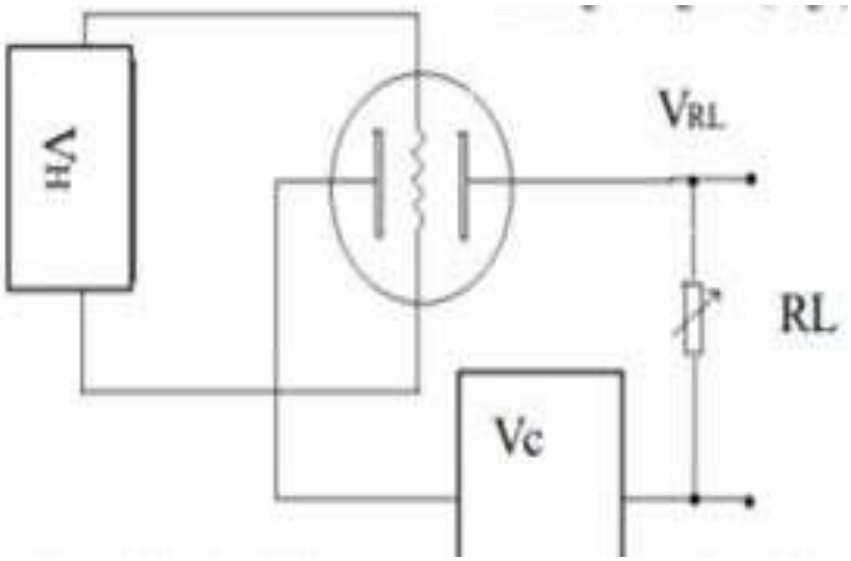


Figure 3: Circuit Diagram of MQ7

The second part deals with the signal output circuit which responds to changes of resistance of the sensor surface accurately. It has an operating principle of surface resistance of the sensor Rs obtained through effected voltage signal output of the load resistance RL.

$$Rs\backslash RL = (Vc - VRL) / VRL$$

The sensitivity of MQ7 sensor had to be adjusted by connecting the sensor to the application circuit and preheat it over 48 hours and the load resistance was adjusted till the CO need was obtained at both high voltage 5V and low voltage 1.4V, at the end the load resistance of 20Kn was fixed

The dust sensors were evaluated and several sensors were found. SM-PWM-01C can detect dust within the range of 1 micrometer to 20 micrometer, SM-UART-04L has high consumption of current of 100 micrometer and GP2Y1010AU0F sensor was selected due to its capability of sensing dust ranging from 0.8 micrometer to 10 micrometer, it consumes 20 mA, with sensitivity of 0.5V/100microgram/m3, having measuring range of 500microgram/m3, operating power ranging 2.5V-5.5V with 5 years of life time. It gives output as analog voltage which is proportional to the measured dust density hence using PWM pins when connected to the microcontroller. When it is connected to the microcontroller (MCU) here are the connections

Dust sensor	MCU
Vcc	2.5V-5.0V

GND	GND
A(OUT)	MCU.IO(Analog Output)
I(LED)	MCU.IO(Module driving pin)

Table 3: Dust Sensor connection with MCU

3.2.2 Microcontroller

Microcontrollers were analyzed Teensy 4.0 is one of the fastest microcontrollers but expensive, particle boron board is a standalone cellular endpoint or LTE enabled gateway but it has disadvantage of loss of mount ability due to moving from particle electron to particle boron. ArduinoUNO is an open source easy to learn, low costs, with wide array of sensors and several third-party libraries and large number of free online resources. The ArduinoUno was compiled using the Arduino IDE (Integrated Development Environment). The Arduino IDE was used to write, compile and upload the code to the Arduinouno board since it can work for al operating systems such as Linux, MAC and Windows. It runs on Java platform with commands helping in debugging, editing and compiling the code with the functions inbuilt. It supports both C and C ++ languages.

3.2. 3 Connectivity

The transferring of collected data from the sensors was done using a communication model. Different communication means were found such as Ethernet which is insensitive to interference with wire connectivity but limited with wire length making it difficult to troubleshoot. Mesh-Based IoT such as Zigbee and Thread were also studied and discovered that they use low power but has low bandwidth, limited range and complex to be set up for average customer. Cellular was used in this proposed work which has long range of connectivity with great bandwidth, low latency. Common standards for cellular are 3G and 4G (LTE) and machine to machine (M2M) standards developed for IoT are LTE-M and NB-IoT. It has nearly global coverage, high speed and band width reliability.

The GSM module was used to have a connection between the system and the cloud. The GSM of GSM900 was first used but it brought some difficulties in programming it hence we switched to GSM800 that gave a reliable connection and data was successfully sent to the cloud.

3.2.4 Monitoring

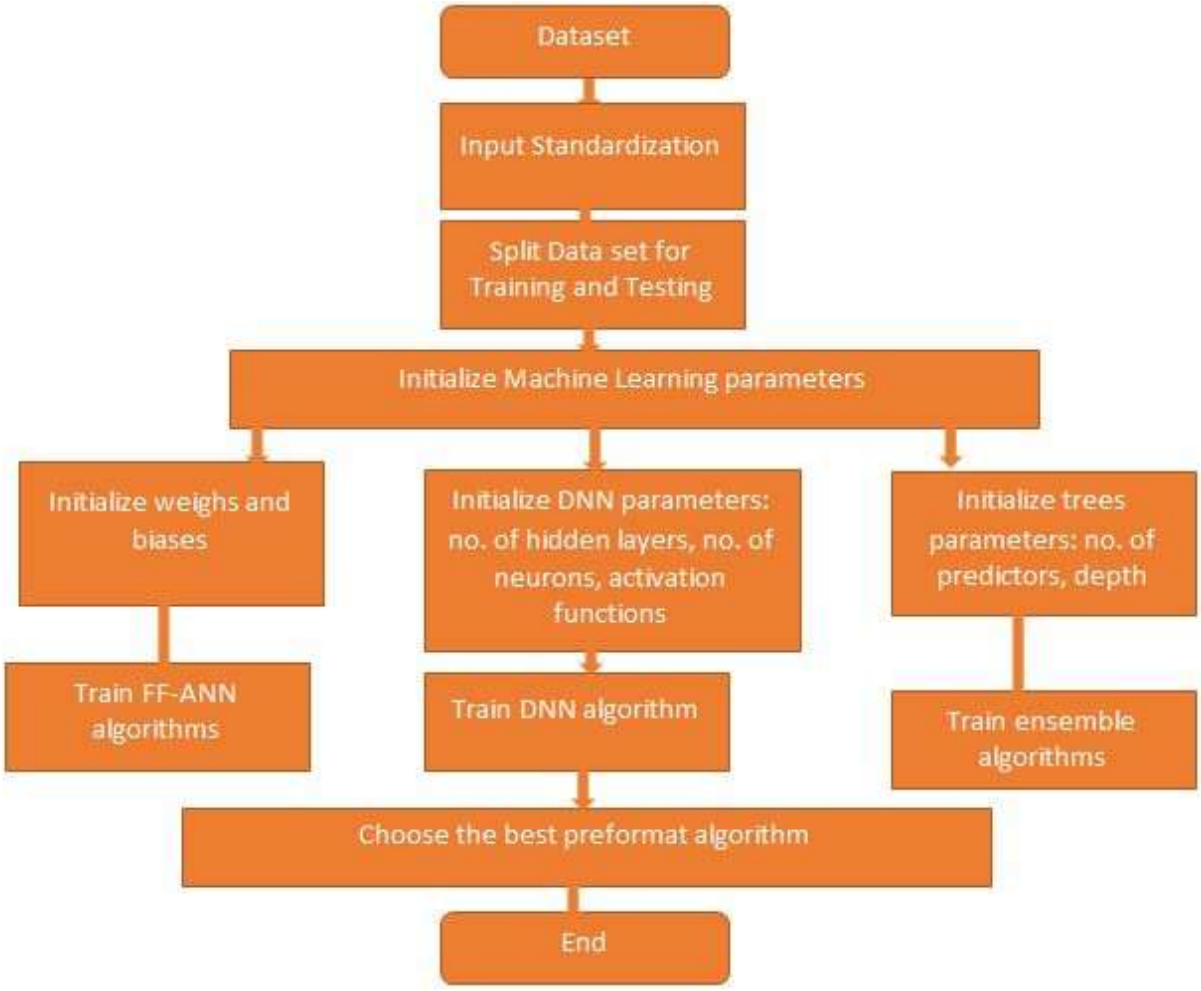
Cloud platforms have become helpful in having real time accessing of data, time saving and minimizing cost. Several plat forms were analyzed to choose the one suitable for our project. Cloud Linux was seen to involve own IT infrastructure, limiting the third-party provider and having a need for having Linux based operating systems. IBM Clod platform focuses on Infrastructure as a Service (IaaS), Software as a Service (SaaS) and Platform as a Service (PaaS), but very expensive. Google Cloud is security oriented and user friendly but most components are google based with limited programming languages choices. Azure IoT suite provides productivity and profitability with solution that are pre-built connected but it requires management, it is expensive with no support for bugs. Thingspeak IoT platform is an open-

source platform that helps in collecting sensor data to the cloud by providing app for analyzing and visualization of collected data in Matlab. It has advantage of free hosting for channels, easy to visualize with additional features for Ruby, Python and Node.js.

The system on monitoring, it comprised of an alerting system. The alerting system was selected depending on what kind of society the system aimed to help, hence the buzzer was seen as the best solution for that since it did need any smart phone, computer or any other installed system for that. The buzzer was used to convert the audio signals from the system into sound signals whenever the limited thresholds values of carbonmonoxide CO, particulate matter PM Temperature and Humidity values were reached.

3.2.5 Prediction

Methods for prediction that were used long time ago so tiresome and complex since the probability and statistics were done manually to achieve the prediction process. The machine learning was introduced and it simplified the whole process by using the software. The predictive models are of two types, the classification model and regression model. The classification model is used for predicting class membership while the regression model is used to predict a number. Both models are created by algorithms also known as classifiers in software format. Common and widely used prediction models are clustering algorithms used for outlier detection, factor analysis is used for finding independent variables and time series found under neural network is used for forecasting continuous vales over time. Artificial Neural Network (ANN) solves complex recognition pattern problems and can handle large dataset it is best when solving unknown variables with the flow chart as shown below



3.2.6 Power Supply

The power supply was used to power the entire system. It fist converted the main grid AC electricity into a DC one. It had to power Arduinouno, GSM module and Buzzer. The Buzzer was connected to the Arduino through a transistor that worked as an amplifier since the buzzer used in the proposed system used a different voltage as the one used by the Arduinouno. Arduinouno used 5V while the buzzer used 12V meaning there Arduino was going to be over drawn. A transistor NPN was used to connect the Arduinouno and the buzzer by allowing the buzzer to be powered from different voltage and the buzzer current drawn by the buzzer was up to 200 milliAmpere.

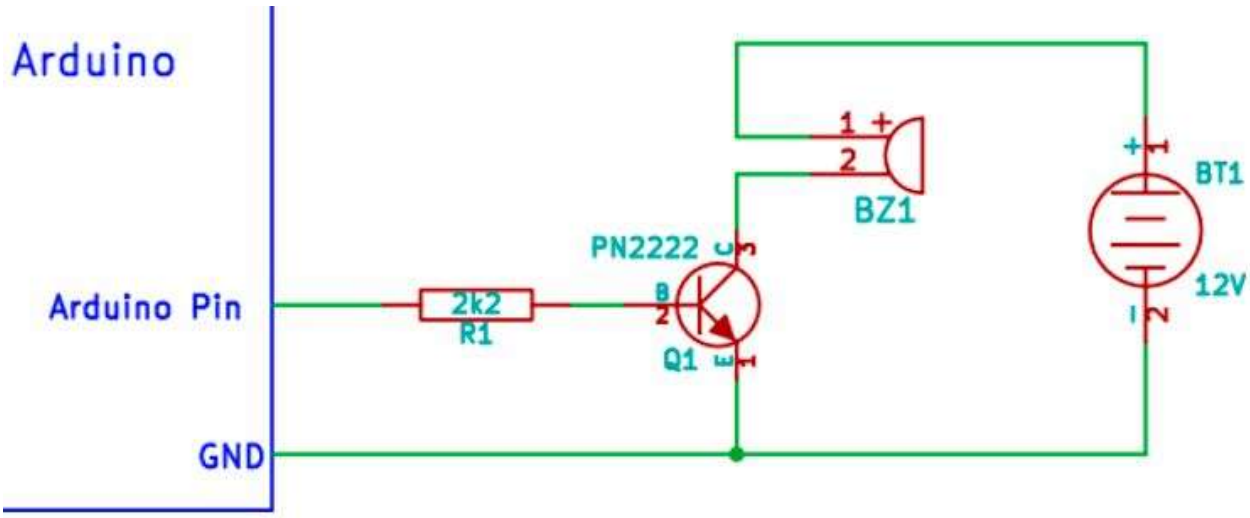


Figure 4: Buzzer connected to ArduinoUno through Transistor

The common emitter transistor was used as shown below, it is also known as voltage divider biasing with the formula of voltage divider as

$$V_B = \frac{V_{CC} R_2}{R_1 + R_2}$$

The voltage gain as

$$\text{Voltage Gain} = \frac{\text{Vout}}{\text{Vin}} = \frac{\Delta V_L}{\Delta V_B} = -\frac{R_L}{R_E}$$

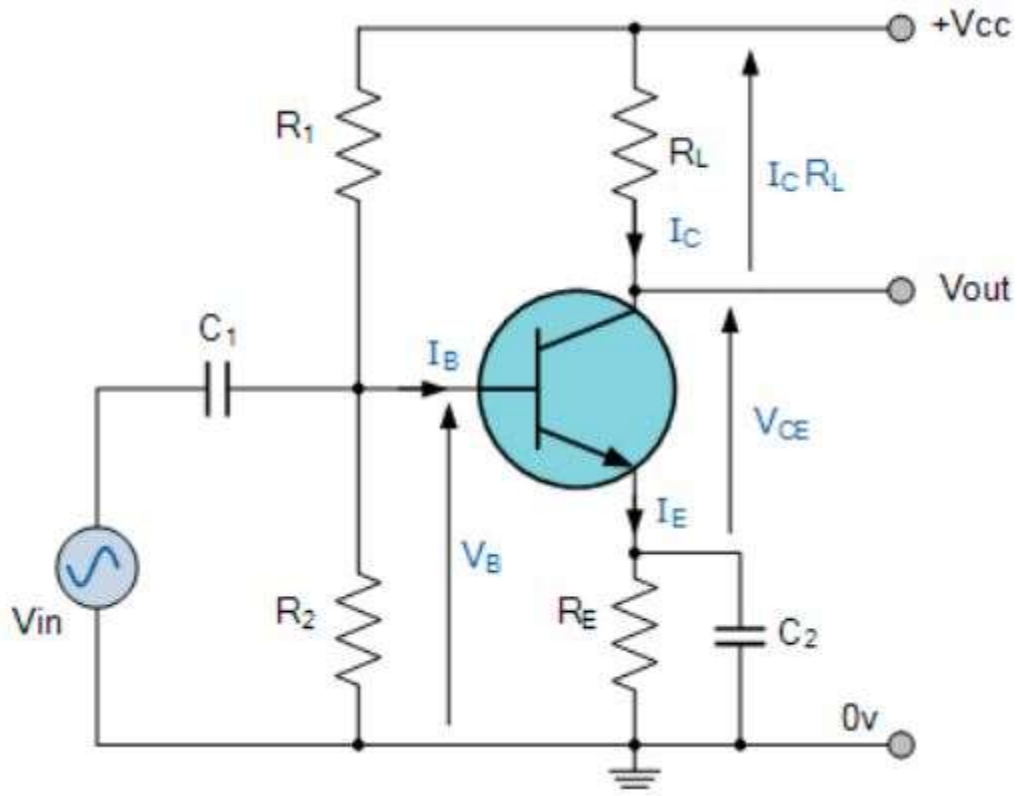


Figure 5: Common Emitter Amplifier

3.3 Proposed Solution

The proposed system was established and placed in one household for 8 weeks. Data was collected for 24 hours and it was sent to the cloud every after 5 minutes. The alerting was done through beeping produced by the buzzer and the prediction was achieved through the machine learning.

Chapter 4: System Design and Implementation

This section outlines the design and implementation for indoor air quality monitoring system and predicting, the system design comprise the sensing, processing, communication, power supply, monitoring and prediction parts with their respective hardware and software used. The implementation consists of the system developed.

4.1 System Design

A MQ7 carbon monoxide (CO) sensor, PM2.5 particulate matter sensor, DHT11 temperature and humidity sensor were all connected to a microcontroller known as an Arduino-UNO so as to collect specified gas, matter and environmental conditions. Data was sent to the cloud platform for IoT projects known as Thingspeak over the internet with the help of GSM module for storage and further analysis. Artificial Intelligence (AI) technique known as ANN was used to predict pollution. Figure one gives the system architecture

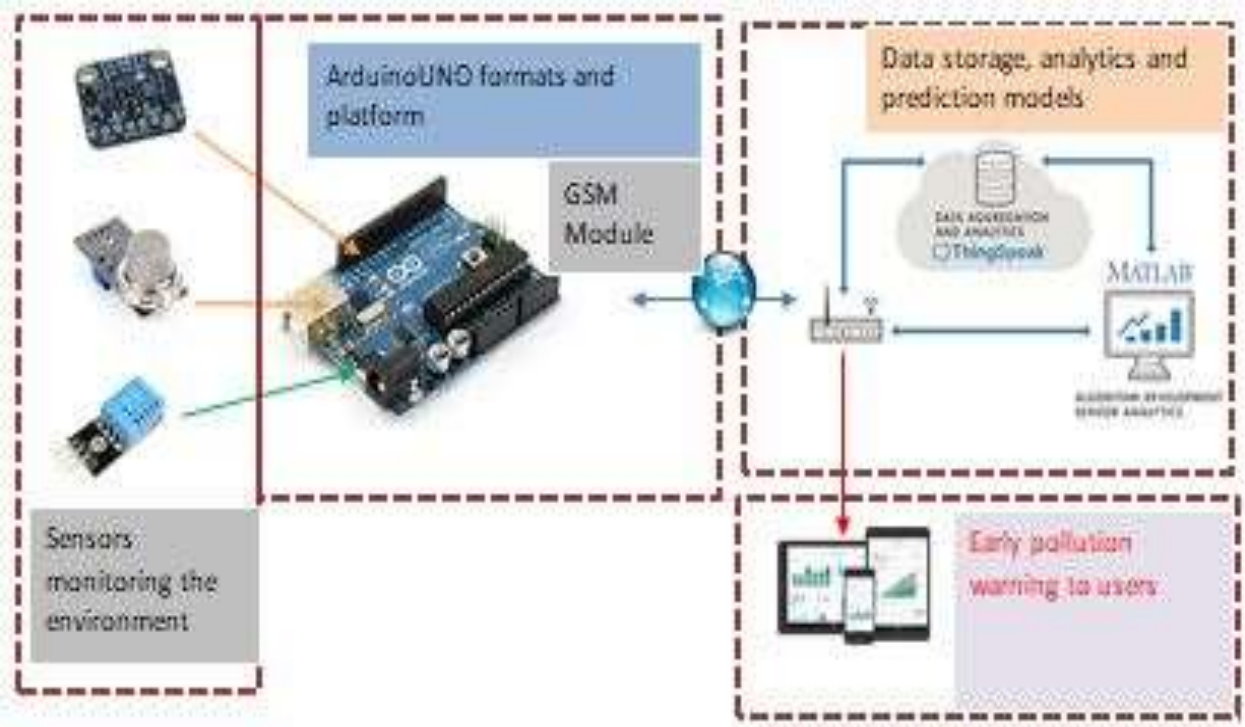


Figure 6: Schematic system Circuit

4.1.1 The Sensing Layer

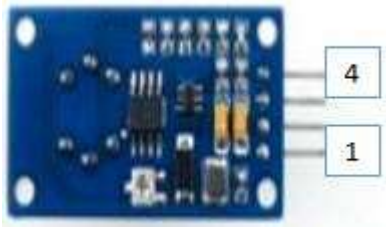
This layer comprises of three sensors, MQ7 gas sensor, PM2.5 particulate matter sensor and DHT11 a temperature sensor. Each component is well described below with their specifications as follows

4.1.1.1 MQ7 Gas Sensor

MQ7 CO gas sensor has adjustable and high sensitivity for CO, has fast response, recovery and signal output indicator, stable and long life that detects a wide range of CO gas of 20-2000 in parts per million (ppm). It uses power of 2.5V to 5.0V. The MQ7 sensor is composed by micro Al2O3 ceramic tube, Tin Dioxide (SnO2) a sensitive layer, heater fixed into a crust made by plastic and stainless steel and measuring electrode [4]



Figure 7: MQ7 Gas Sensor



Pin Number	Pin Name
1	AO
2	DO
3	GND
4	VCC

Figure 8: MQ7 Pin-out

4.1.1.2 Particulate Matter Sensor

The PM2.5 and PM10 refer to particulate matter which are among the most dangerous air pollutants with particulate diameter up to 2.5 microns and 10 microns respectively. PM2.5 particles can travel deep into the human lung due to their small size and cause a severe health issues compared to PM10 particles. PM2.5 sensor of GP2Y1010AUOF is a digital and universal particle concentration sensor that can be used to obtain the number of suspended particles in the

air and the output is in the form of digital interface. It can detect house duct, cigarette and also designed for automatic running of application like air purifier and air conditioner.

PM 2.5 uses light scattering working principle to detects and counts particles, where the particle passes through the light source (laser beam) and causes scattering of the incoming light where the scattering light is received and detected by a photodiode and converted into an electrical signal that calculates the particle concentration.



Figure 9: PM2.5 Sensor

Pin number	Pin name
1	GND
2	Vcc
3	NC
4	NC
5	RxD
6	TxD

Figure 10: PM2.5 Pin-Out

4.1.1.3 DHT11 Sensor

The DHT11 is temperature and humidity sensor. It measures temperature using NTC and uses an 8 bit microcontroller to output the serial data values of temperature and humidity hence easy to be set up. The DHT11 sensor can measure temperature from 0 C to 50 C with accuracy of +-1% and measures humidity from 20% to 90% with accuracy of +_1%. [5]

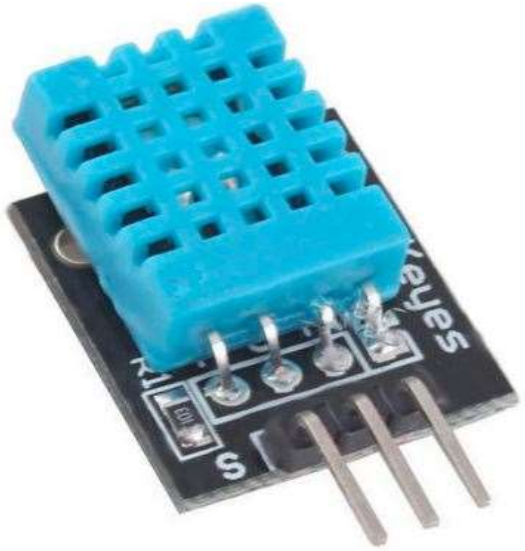


Figure 11: DHT11 Sensor

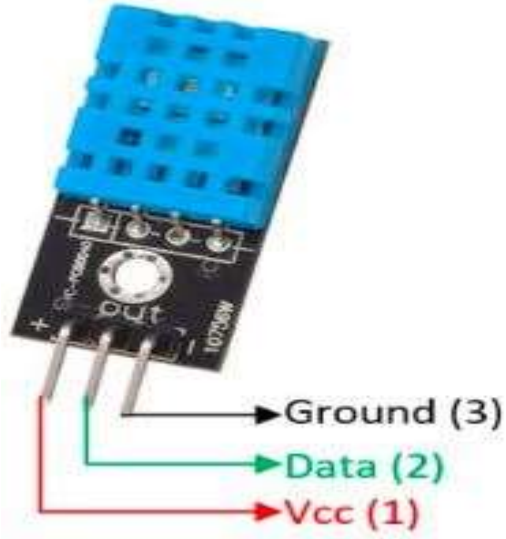


Figure 12: DHT11 Pin-out

4.1.2 Processing Layer

This layer is made up of a powerful microcontroller named Arduino-UNO. An Arduino-Uno is a microcontroller board based on ATmega328P. It is low power, high performance and 8-bit microcontroller. It has 6 analog inputs, 14 digital input/output of which 6 can be used as PWM output pins, a 16 MHZ ceramic resonator, power jack, an ICSP header, USB connection and a

reset button and can be easily connected to a computer with a USB cable or can be powered by battery or an AC to DC adapter [6].



Figure 13: ArduinoUNO

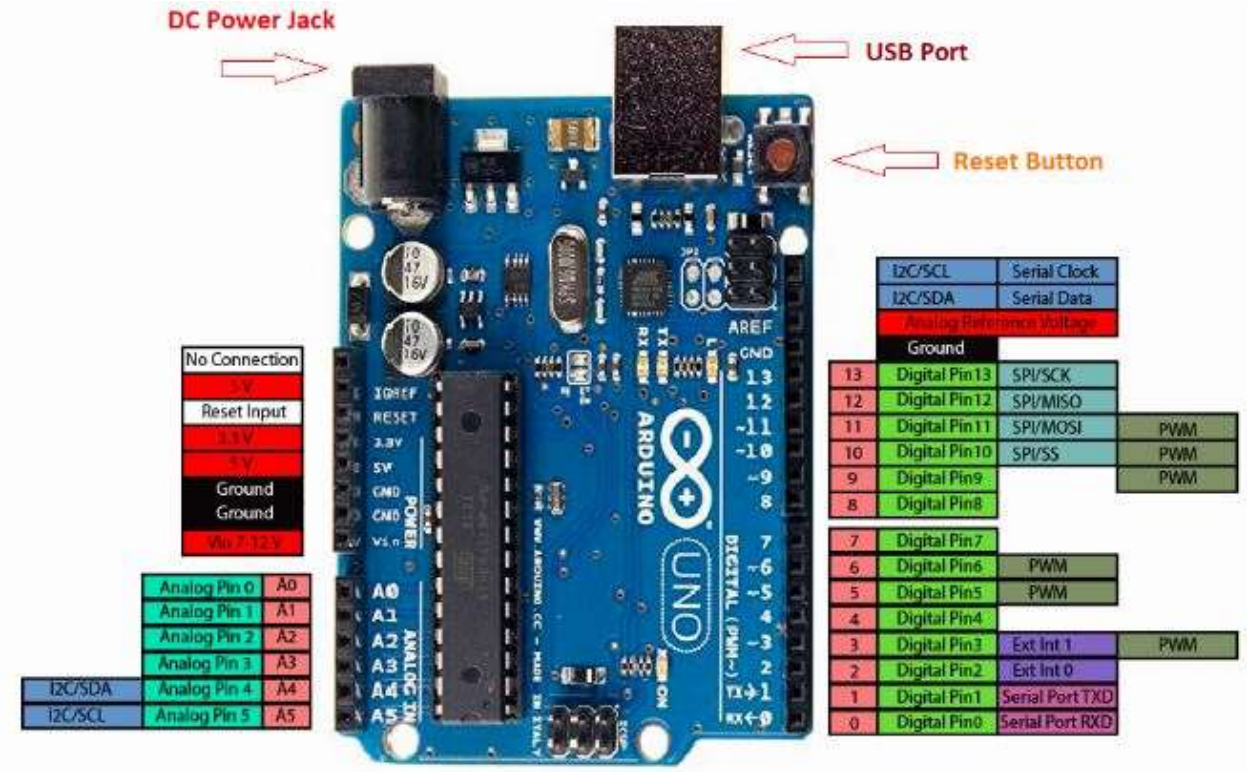


Figure 14: ArduinoUNO Pin-out

4.1.3 Connectivity Layer

SIM800 is a quad-band GSM/GPRS module that works on frequencies of 850MHz GSM, 900MHZ EGSM, 1800MHz DCS, and 1900MHZ PCS. It has one UART port for updating and debugging it. It also has one SIM card interface to be integrated to TCP/IP protocol and it also features GPRS MULTI-SLOTS [7]. GSM SIM800 can be used to send/receive message, make calls, send/ receive data over the internet as used in this study hence making it useful for other IoT applications.



Figure 15: GSM800 Module

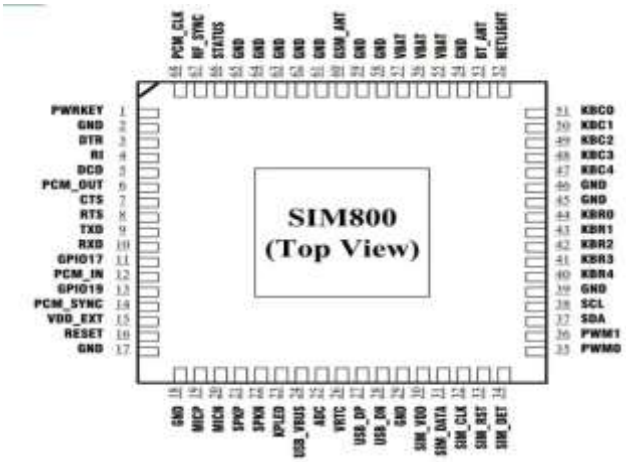


Figure 16: GSM800 Pin-out

Table 3: Components Specifications

All the three layers' components were further specified as shown below

Product Name	Sensing Quality		Power Specifications			Performance(Speed)		Design	Unit Cost per Sample($)
	Accuracy/Resolution	Capability/Working Conditions	Active Current	Sleep Current	Power Supply Range	Input Reference Clock	Response Time	Dimension	
Temperature/Humidity Sensor(DHT11)	+2 ^{0}C +5%RH	0-50^{0}C 20-80%	0.3mA	60uA	3.5-5.5V	1Hz	80usec	(l*w*h) 12.6mm *5.83mm*16mm	30
MQ7	20*2000(resolution)	-20℃±2℃ 65%±5% RL:10KΩ ±5%			2.5V-5.0V		<1sec	(l*h) 40.0mm*21.0 mm	35
PM 2.5	+10%	0-300ug/m^3		<200 uA	4.95-		<10s	(w*d*h)mm 46*30*17.6	35

					5.05 V				
Embedded Processor	Memory		Power Specifications			Performance(Speed)		Design and Output	Unit Cost
	Flash	RAM	Active Current	Sleep Current	Power Supply Range	Input Reference Clock	Response Time	Dimension	US dollars
Arduino-Uno AtMEGA328p	32KB	2KB	50mA	20 mA	3.3-5V	16 MHz		(l*w)mm 68.6*58.4	50
Display Module	Sensing Quality		Power Specifications			Performance		Design	Weight	Unity Cost per sample
	Accuracy/ Resolution	Working Condition	Active Current(mA)	Sleep Current	Power Supply Range(V)	Data hold time(ns)	Data Output delay (ns)	mm)	g	$

LCD 24*4	95%	25C	1.5	25	5	20	120	125*39*14(w*h*t) 24char.*4lines	150	25

Communication Module	Memory		Power Specification			Performance Speed		Dimension	Weight	Unit Cost
	FLASH	RAM	Active Current	Sleep Current	Power Supply Range	Uplink	Downlink	(l*w*h) mm	grams	US dollars
SIM800C	24Mbit	32Mbit		0.88mA	3.4V - 4.4V	85.6 kbps	85.6 kbps	17.6*15.7*2.3	1.3g	50

Other components that were used in the proposed system were the jump wires, a buzzer and a case used to enclose the entire system.

4.1.4 Monitoring, Storage and Analysis Layer

This layer consists of how the data collected by the system was stored and analyzed for further usage, it involved the LCD for displaying the results and the cloud platform known as Thingspeak.

4.1.4.1 Liquid Crystal Display (LCD)

The LCD of 24character and 4 lines was used to display the readings from the sensors where the readings of temperature, humidity, CO and PM from the kitchen were displayed in real time. Is

comprises up of two matters, the solid and liquid and composed of several layers including electrodes and two polarized panel filters. It has low power consumption compared to LED or cathode ray displays [8]. It works on principle of blocking light where an electrical current applied to the liquid crystal molecule, the molecule tends to untwist causing the angle of light passing through the molecule of polarized glass and causing a change in an angle of top polarizing filter. Hence a little light is allowed to pass the polarized glass through a particular area of the LCD. LCD uses a light to produce a visible image [9].



Figure 16: LCD

PIN CONFIGURATION

PIN NO.	1	2	3	4	5
SYMBOL	Vss	V_{DD}	Vo	RS	R/$\overline{W}$
PIN NO.	6	7~14	15	16	
SYMBOL	E	DB0~DB7	A	K	

Figure 17: LCD Pin-out

4.1.4.2 Thingspeak

This layer consists of Cloud Platform for IoT innovations known as Thingspeak platform. Thingspeak is an open source of "Internet of Things" application and API that store and retrieve data from things using http over the internet or via a local area network. It allows you to aggregate, visualize and analyze live data streaming in the cloud. Sensor logging applications, location tracking applications and social network of things with status updates and send alerts.

4.1.5 Air Quality Index

Air is mainly made up of two gases which are nitrogen and oxygen, and other smaller amounts of many other gases and particles that include mainly five pollutants including ground level ozone, carbon monoxide, sulfur dioxide, nitrogen dioxide and airborne particles/ aerosols, the most dangerous ones being ground level ozone and aerosols. [10] The Air Quality is based on measurement of particulate matter (PM2.5 and PM10), and other gases such as CO. Air Quality Index (AQI) is a key tool used by local air quality agencies to provide simplified local air quality

information. The AQI works like a thermometer that runs from 0-500 degrees, by showing changes in the amount of pollution in the air. The color code is used for easy understanding of AQI by community. [10] This information enables people to know how unhealthy the local air quality is, how unhealthy air might affect them and to take appropriate measures to protect themselves from polluted air.

Table 4: AQI levels with corresponding colors

Air Quality Index (AQI) value	Levels of Health Concern	Colors
0 to 50	Good	Green
51 to 100	Moderate	Yellow
101 to 150	Unhealthy for Sensitive Groups	Orange
151 to 200	Unhealthy	Red
201 to 300	Very Unhealthy	Purple
301 to 500	Hazardous	Maroon

The color-coded Air Quality Index (AQI) describes air quality levels in a standard way across the world, where each color corresponds to a different level of healthy concern and warnings. The higher the AQI value the greater the level of air pollution and the greater the health danger. For example from the table above an AI value ranging from 0 to 50 represents “Good Air Quality” which is satisfactory and poses little or no health risk. An AQI value from 51 to 100 represents “Moderate Air Quality” with acceptable air quality but may pose a moderate health concern to people who are sensitive to particulate pollution, an AQI value ranging from 101 to 150 represents “Unhealthy for Sensitive Groups” the general public are unaffected but the sensitive ones may experience health effects and an AQI value ranging from 151 to 200 represents “Unhealthy” which is dangerous to both sensitive and to non-sensitive ones. [11] The air is polluted by different mixture of small particles with a diameter less than 10μm and this mixture is referred as particulate matter. Particulate matter is a mixture of airborne solid particles

and liquid droplets that can be inhaled to human body due to its small size which can penetrate to different parts of human body such as lungs. After a long period of time it can cause lung diseases such as asthma and cardiovascular diseases. The World Health Organization (WHO) reports airborne particulate matter as Group 1carcinogen with the highest risk to both health and environment and with responsibility for one death in every nine deaths annually. [12] Carbon monoxide (CO) is a colorless and odorless gas, which is composed of carbon (C) one atom of oxygen (O). The CO gas comes from incomplete combustion of fuels. The CO gas causes poison inside human bodies by displacing oxygen in them. The CO poisoning can be prevented with simple actions by installing CO alarm and maintaining fuel burning appliances to avoid illness and deaths since the gas is odorless, colorless and tasteless hence cannot be detected by our senses [13]. This study used MQ7 CO sensor that has adjustable and high sensitivity for CO, has fast response, recovery and signal output indicator, that detects a wide range of CO gas of 20-2000 in parts per million (ppm).

4.1.6 Prediction Layer

The sensing device was used to collect data for a period of two weeks for 24 hours from a household, it is assumed that after this period a pattern will have been established. In a related study in [23] data was collected for a period of one week for 8 hours daily. The prediction layer was performed with the help of Machine Learning technology. MATLAB was used to develop the prediction model using logic rules implemented in It uses neural network of dynamic time series app of a nonlinear autoregressive with external (exogenous) input (NARX) with predict series y(t) given d past values of y(t) and other series of x(t). It has three different networks used to create algorithm which are Levenberg-Marquardt, Bayesian Regulation and Scaled Conjugate Gradient. The Scaled Conjugate Gradient network uses less memory and the training automatically stops when generalization stops improving as indicated by an increase in the mean square error of the validation samples, the Levenberg-Marquardt network uses more memory but less time and the training automatically stops when generalization stops improving as indicated by an increase in the mean square error of the validation samples and the Bayesian Regulation network requires more time but result in good generalization for difficult small or noisy dataset and the training stops according to adaptive eight minimization (regularization). In this work the Bayesian Regulation network was used to generate the algorithm.

4.2 System Implementation

The system was developed and simulated using Proteus software.. The prototype was later developed through connecting sensors to the microcontroller and with the help of GSM module the data from the sensors were sent to thingspeak cloud platform for storage and analysis. The MQ7 gas sensor was connected to analog pin of UNO through pin A2, PM2.5 sensor was connected to analog pin through pin A3, LCD module was connected at analog pin A4 and A5 of a microcontroller while DHT11 sensor was connected to digital pin of microcontroller through pin 2, GSM module was connected to digital pin of UNO through pin 3 and 4, and LEDs were connected to digital pin of the Arduino-UNO.

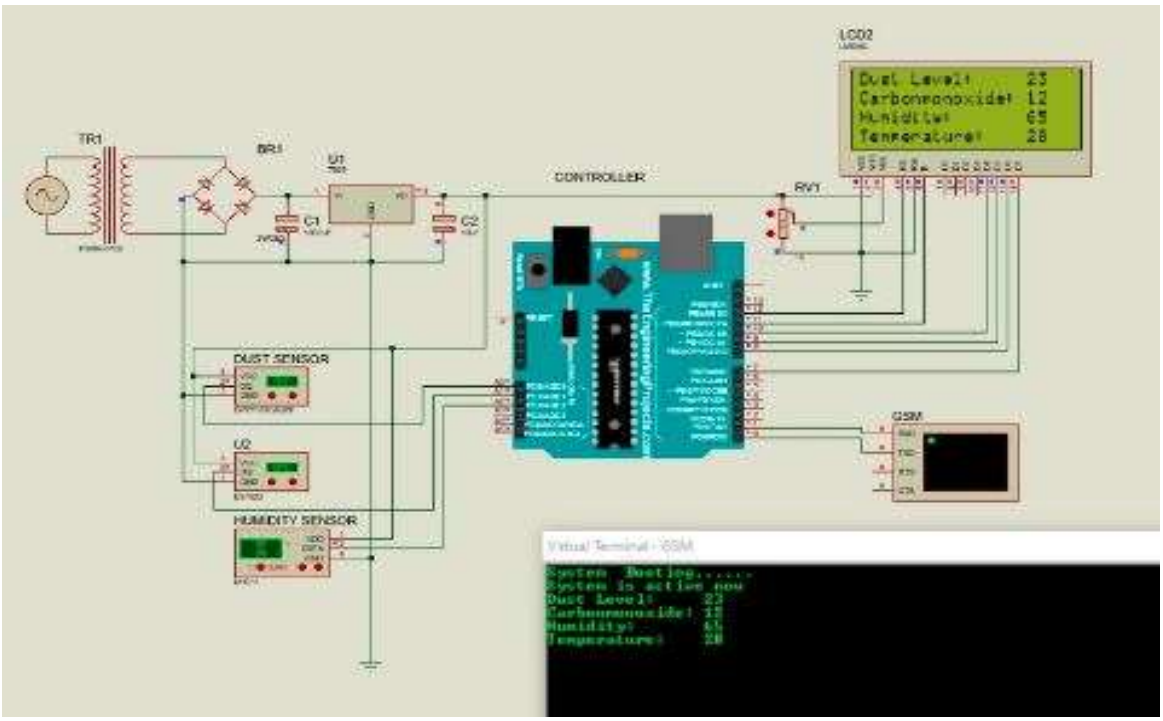


Figure 18: Circuit of the Proposed System

The system was deployed in the kitchen having one window and charcoal has source of fire. The data was collected continuously with-in two months and transmitted to cloud using GSM module. The system was powered by the main grid electricity since the system was inside the house as source of power. The figure below shows the complete circuit being deployed in the kitchen

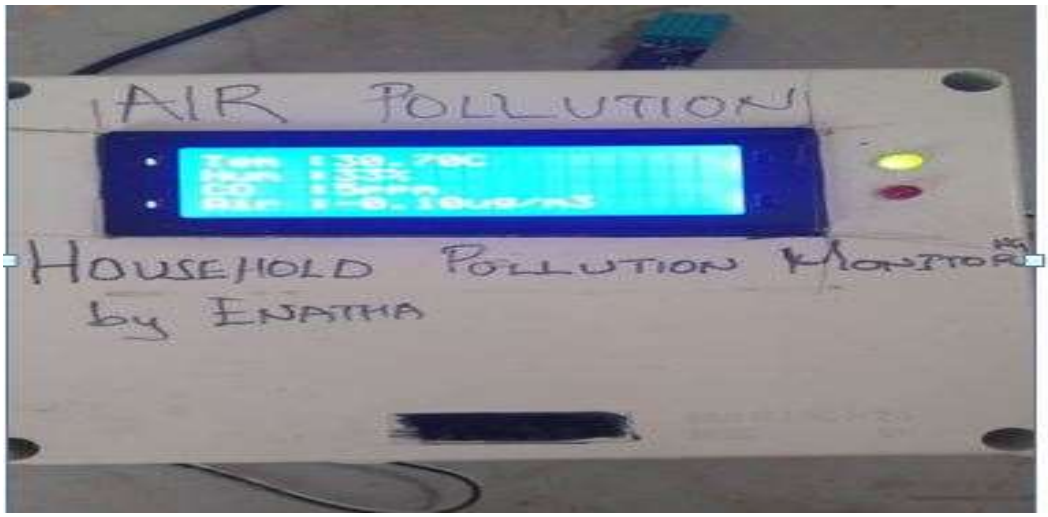


Figure 19: Deployment of the System in the kitchen

CHAPTER 5: Results and Data Analysis

5.1 Data Transmission, Storage and Monitoring

The system was deployed at a house in the kitchen and was collected for 24 hours in every 5minutes. SIM800 a GSM module was used in this project for transmission of data from the system to cloud. The data collected had noise to the extent of having negative values mostly from MQ7 gas sensor. The Thingspeak platform was used and different graphs were analyzed from the collected data as shown in the figures below



Figure 20: Graph of Temperature verse Time

The figure 18 above describes the temperature detected in Universal Time Coordinated (UTC) time zone which is equal three hours behind East African Time (EAT). The room temperature is normal when there are no activities in the kitchen but the room temperature increases during noon and in the evening when the stove is on and kitchen is busy.

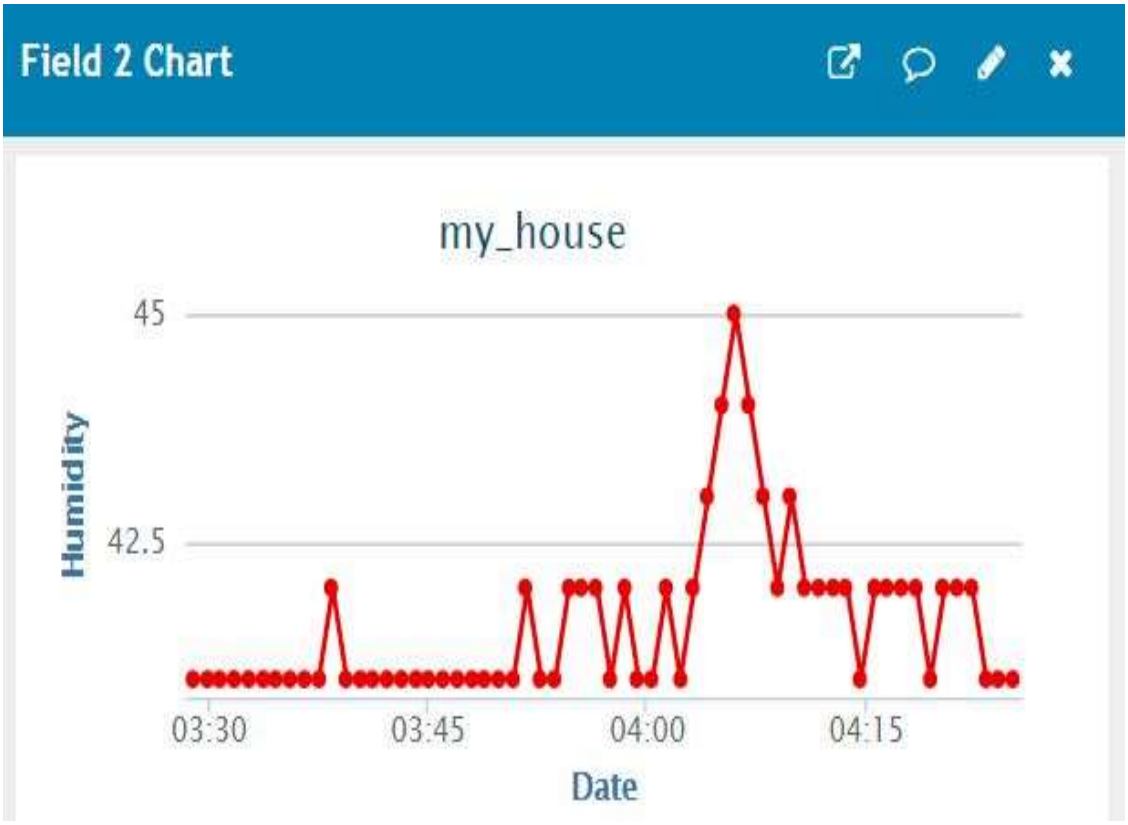


Figure 21: Graph of Humidity to Time

The graph of Humidity to Time from figure 21 humidity increases when a kitchen is busy too and then goes back to normal.

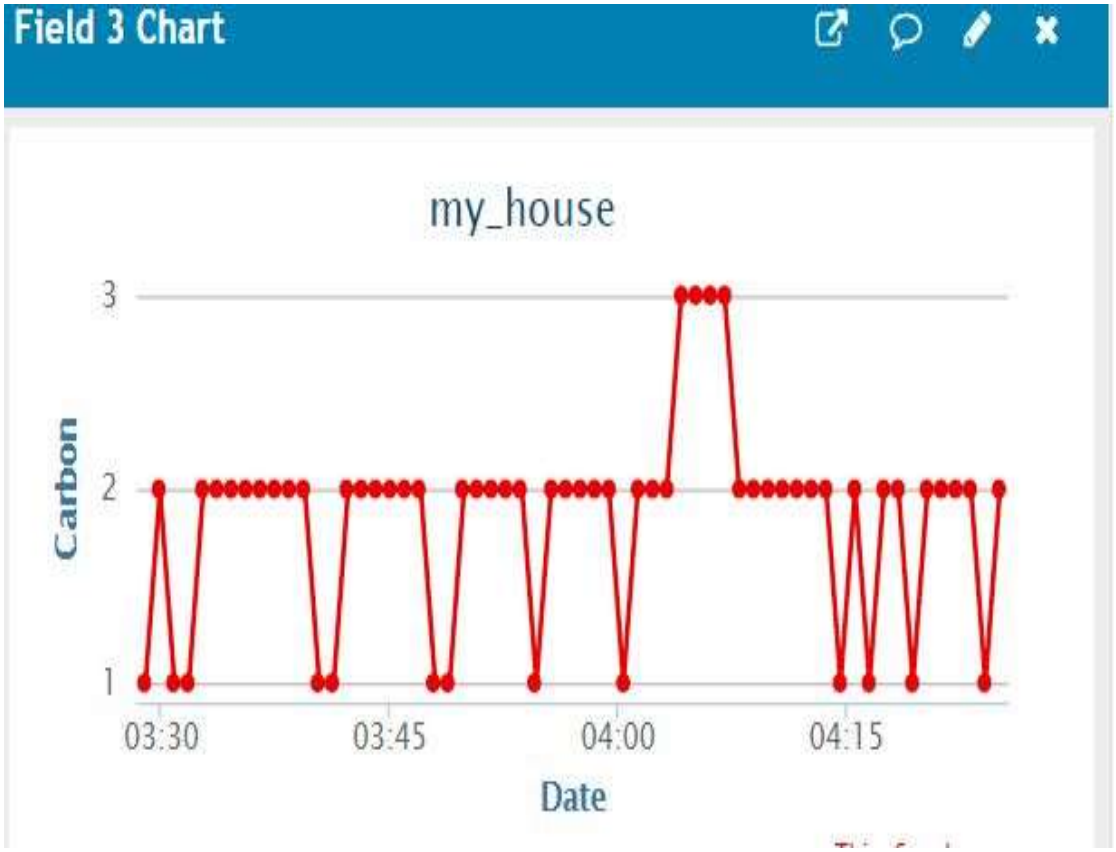


Figure 22: Graph of CO vs Time

Figure 22 shows the detection of CO from the kitchen,it was seen that CO was normal but it was detected as 13 which is poor according to the AQI chart and can cause health issues and death w to the people in the house since it was detected at night.This when the stove is used in the kitchen and it goes back to normal after the usage of the stove in the kitchen.



Figure 23: PM2.5 VS Time

Figure 23 shows the particulate matter being detected, with higher values at different points of time, from this we can clearly conclude that the particulate matter or the room air quality changes when ever there is any rise in particles in the air that is when ever the stove is on or when the number of people in the room increases and during morning, noon and evening the readings were higher and the highest values was 63 which is moderate and dangerous to people with breathing problems.

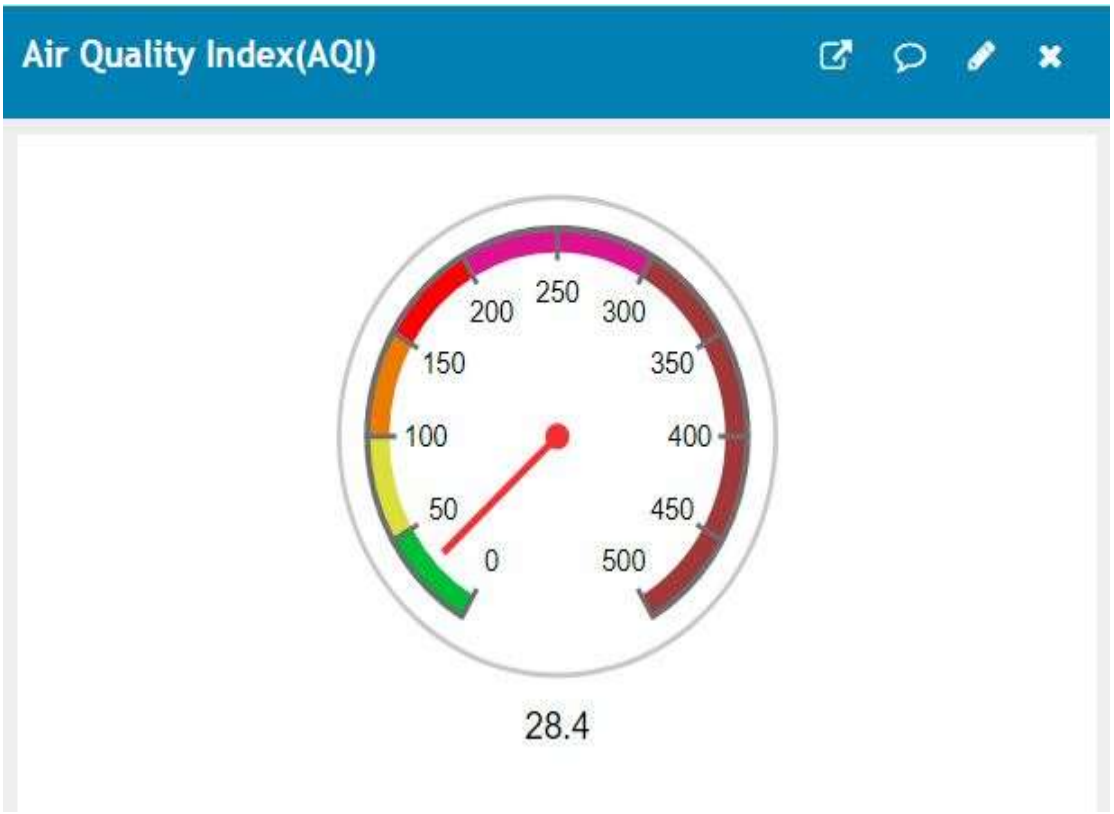


Figure 24: Graph of AQI Widget

The air quality widget from figure 21 shows the air quality by colors which can be easily understood by just looking at the color being pointed at.

5.2 Air Quality Index Calculation

The AQI was calculated using a segmented linear equation. The dataset was developed from the calculated AQI and it was further used for AQI prediction. The Segmented Linear equation is shown below

$$I = \frac{I_{high} - I_{low}}{C_{high} - C_{low}}(C - C_{low}) + I_{low}$$

Where:

I=Air Quality

C=Pollutant Concentration

C(low)=The Concentration Breakpoint that is < C

C(high)=The Concentration Breakpoint that is > C

I(low)= The Index Breakpoint Corresponding to C(low)

I(high)= The Index Breakpoint Corresponding to C(high)

Table 5: Indices Categories and corresponding concentration Breakpoint for each pollutant (ug/m3)

AQI Category	Range	PM2.5 24hrs	CO 8hrs
NAAQ Standard	0-500	60	2.0
Good	0-50	0-30	0-1.0
Satisfactory	51-100	31-60	1.1-2.0
Moderate	101-200	61-90	2.1-10
Poor	201-300	91-120	10-17
Very Poor	301-400	121-250	17-34
Severe	401-500	250+	34+

5.3 AQI Prediction using

The data collected from the system was used to calculate the AQI as shown in the previous chapter and the dataset was developed. The AQI prediction was done with the help of MATLAB software that solves complicated issues such as prediction applications. It uses neural network of dynamic time series app of a nonlinear autoregressive with external (exogenous) input (NARX) with predict series y(t) given d past values of y(t) and other series of x(t). On network architecture that defines a NARX neural network was performed several times at different number of hidden neurons and an output of 9.2 performance achieved at 18 number of hidden neurons and 2 number of delays d. Algorithm was created using Bayesian Regularization by dividing the dataset into three groups, the training group the validation group and the testing group. The algorithm randomly divided the dataset into 19 target time steps where the training group had 13 target time steps with 70% of dataset and were adjusted according to its error, the validation group had 3 target time steps with 15% of dataset which helped to measure network generalization and to halt training when generalization stops improving lastly the testing network had 3 target time steps with 15% of the dataset having no effect on training and so provide an independent measure of network performance during and after training. After the training different plots were generated and analyzed where the mean squared error (MSE) which is the

amount of error in statistical models found at 18 numbers of hidden layers was 2 meaning the error obtained was very small. The lower the mean square error the higher the accuracy of prediction, when a model has no error the MSE is zero meaning the match between the actual and the predicted dataset is excellent as shown in figure 23

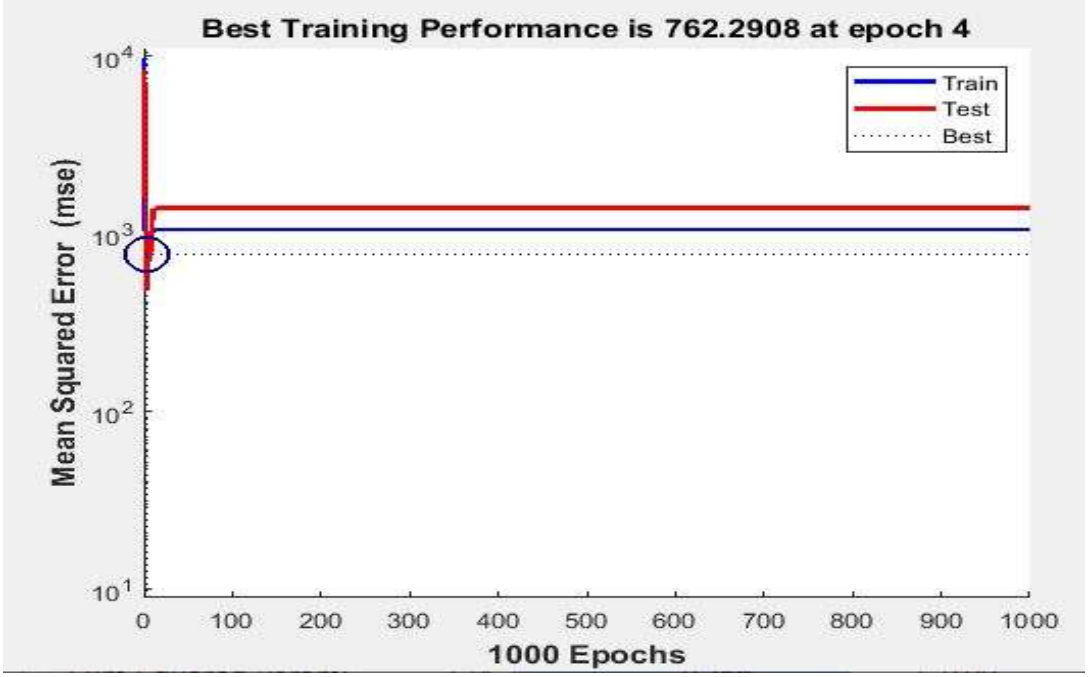


Figure 25: Graph of Mean Square

The gradient was 256.7232 at epoch 1000 with 3.447 number of parameters, 0.35591 as sum squared parameter giving zero as the validation check as shown in figure 23, since gradient is a first-order iterative optimization algorithm for finding a local minimum of a differentiable function by taking repeated steps in the opposite direction of the gradient of the function at the current point.

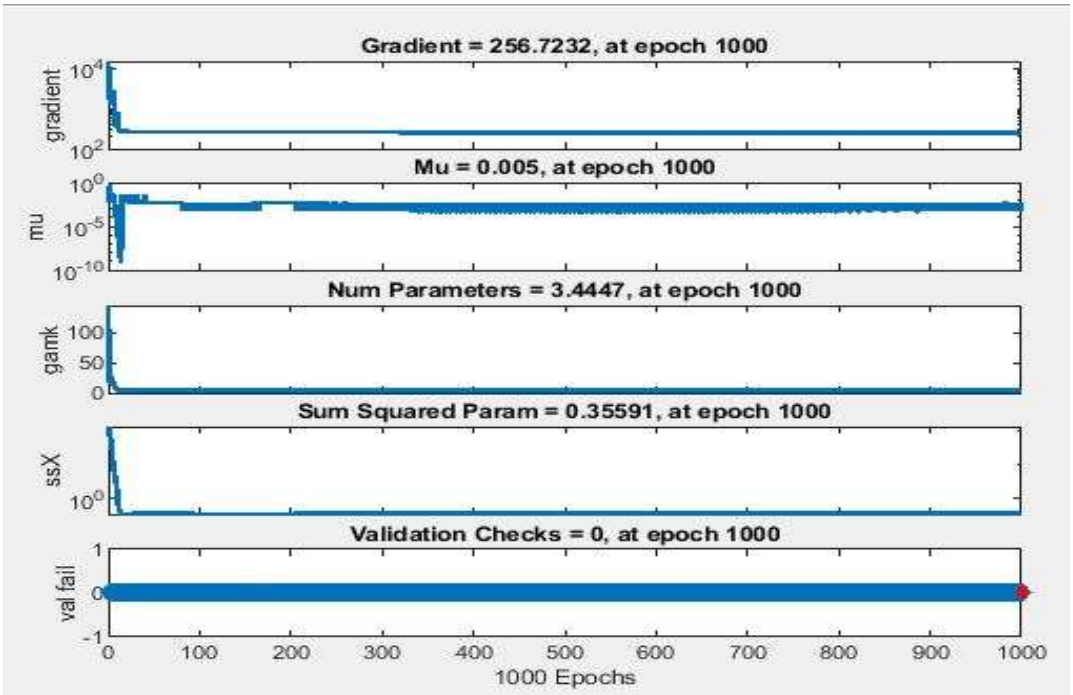


Figure 26: Graph of Gradient

The error histogram which is the histogram of the errors between target values and predicted values after training a feedforward neural network indicating how predicted values are differing from the target values with 20 bins that are numbers of vertical bars observed on the graph had zero error at 0.8517 and training between 0-4 with test lying between 4-8 meaning many samples from dataset used had an error range between 4 and 8as shown from figure 26,

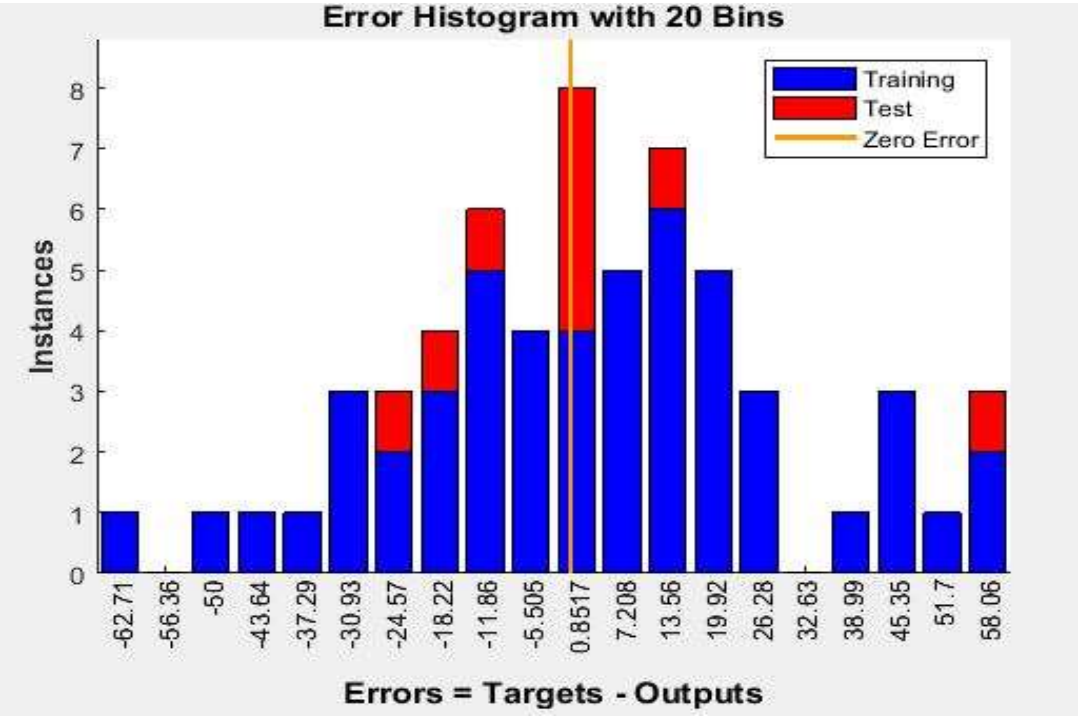


Figure 27: Graph of Error Histogram Plot

A linear regression value was obtained, where the regression plots a best fit line or a curve for relationship between variables and predicted data. During training the regression was 0.62586 for the testing the regression was 0.868 giving an overall regression value as 0.66807 meaning the relation was close as shown in figure 28. We can see the regression of testing has improved meaning there was no nose as compared to the training process.

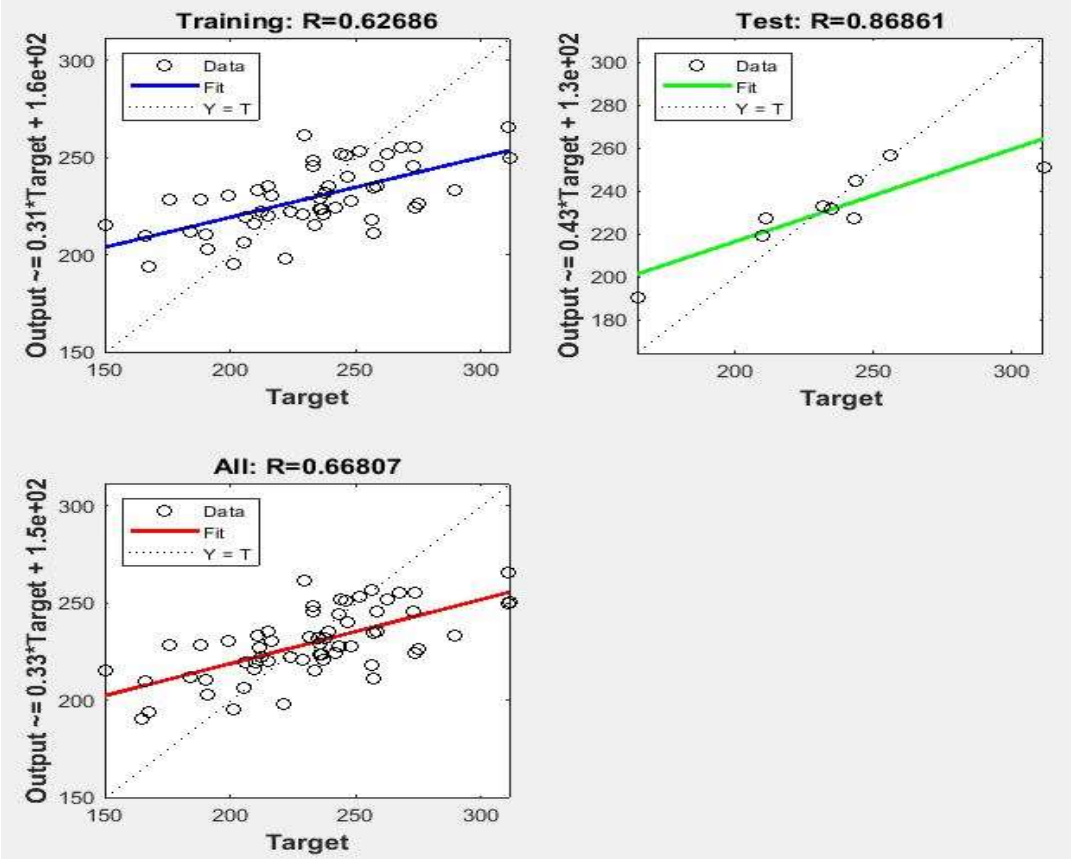


Figure 28: Graph of Regression Plot

The response of prediction was also obtained between 210 and 260 with few errors in between as shown below

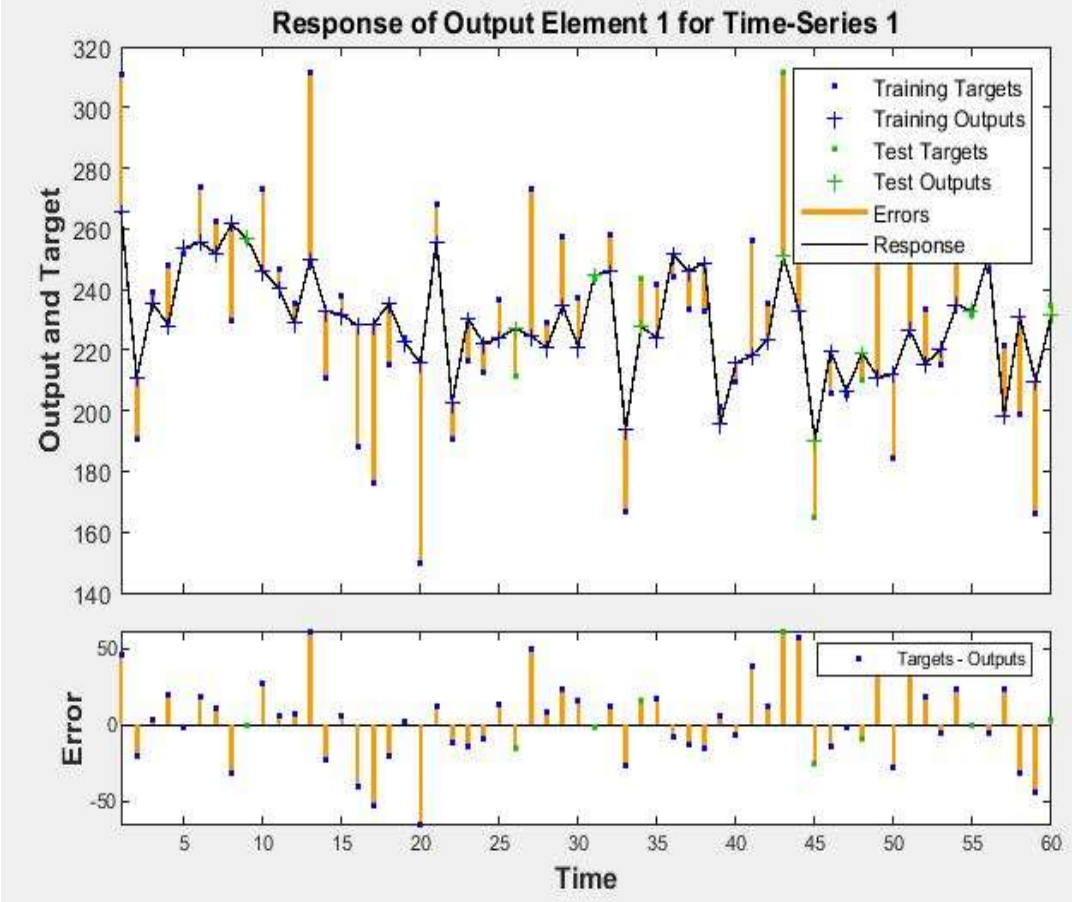


Figure 29: Response of Output Element

Lastly the autocorrelation was calculated by using the same time series twice, once in its original form and once lagged one. The autocorrelation of error obtained was 1 meaning the similarity between a given time series and a lagged version of itself over a successive time intervals was perfect as shown in figure below

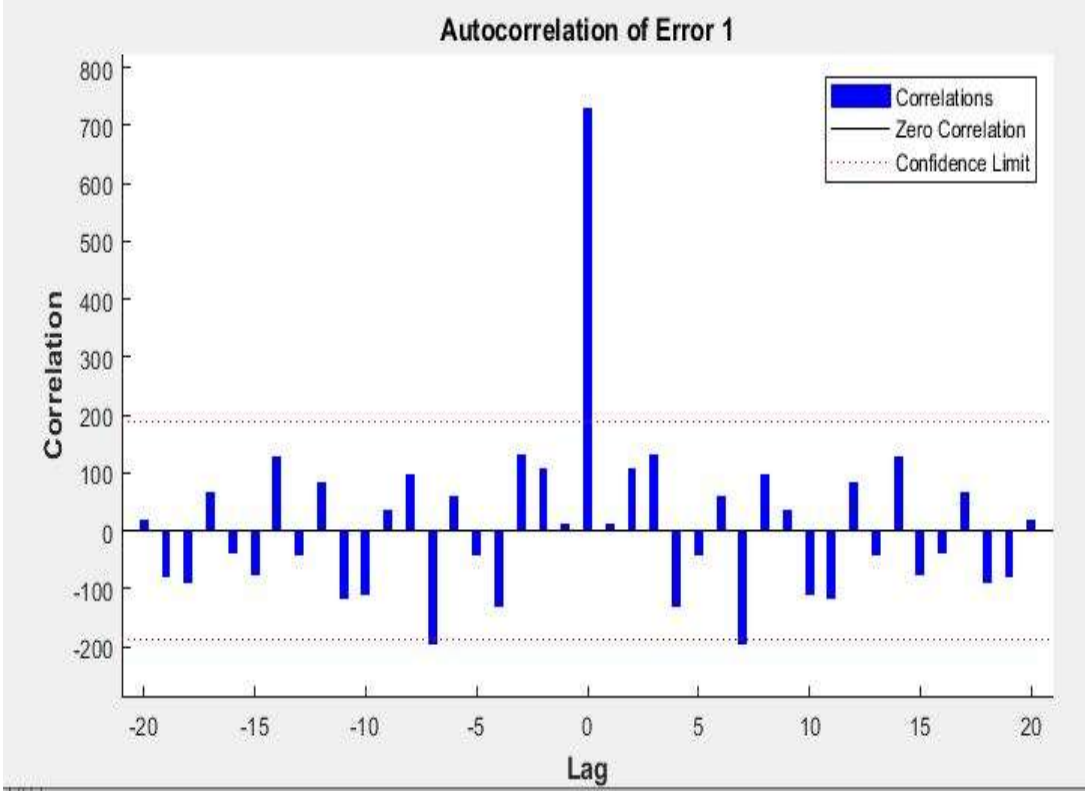


Figure 30: Autocorrelation

CHAPTER 6: Conclusion

The carbon monoxide gas (CO), the particulate matter, temperature and humidity were collected from the kitchen that uses charcoal as source of fire, and data was sent to Cloud platform with the help of GSM module and the alerting was done through a beep sound from a buzzer. The Thingspeak platform was used to monitor the data collected from the kitchen and it was observed that during day mostly at after noon the CO and particulate matter were at high values while at night around 12-8pm the readings were the worst. This gave the conclusion that when the charcoal fire is used during day when it is hot outside and the window is open the CO and PM2.5 readings are not as high as the ones at night when the temperature is low and the window is closed. It is better to have an outside kitchen so as to avoid being exposed to the harmful air mostly at night when everyone is at home hence reducing and eliminating the diseases and deaths caused by the poor air quality.

6.1 Limitations

The GSM module uses a SIM cad which brought some issues such as need to put it in the phone for recharging it with internet bundle, connectivity can be more achieved using LoRaWAN, BLE, Sigfox, and NB-IoT once they become popular in Africa.

The power went off for some hours for few days hence the long lasting batteries or solar power can be used to solve the problem.

6.2 How to Protect Ourselves from Indoor Air Pollution

i. Knowing the CO poisoning symptoms
ii. Properly vent and maintain fuel burning appliances
iii. Installing and maintaining Air Quality Monitoring and Prediction System

6.3 Potential Impact

i. The proposed system will be of benefits to the population at large by helping them to be aware of the risks they may be facing so that they may take corrective measures and precautions. Such actions will lead to a healthier and more productive citizens leading to economic growths.

ii. Hospitalizations cases resulting from air pollution will also be greatly reduced leading to less stress in the health sector.
iii. The collected data and prediction algorithm will be used by Governments and NGOs for proper planning.
iv. Researchers will also benefit from this study

6.4 Future Work

The system comprised of two harmful gases, if the more gases are used we can be able to determine how each gas affects the air quality in the household.

The prediction can be done using a different algorithm to compare the accuracy.

References

[1] Baoxian Liu,Junchi Yan,Zhang,Chao,Jinghai Yan,Lingjun Li,Dawei Zhang,Xiaogung Rui and Rongfang Bie, "Hybrid Measurement of Air Quality as an Image Based Approach," ISSN 853-856, 2017.

[2] S.Maurya,S.sHARMA,pYadav, "Internet of Things Based Air Pollution Penetrating System ," vol. 1, 2018.

[3] Vema,Rahul Ahuja, Hardik Meisheri,Ishan and Lapila Dey, "Air Pollutant Severity Prediction Using Bi-directional Network," 651-654, 2018.

[4] HANWEI ELECTRONICS CO., LTD, "MQ-7," CHINA,.

[5] Components 101. (2021, July) Components 101. [Online]. https://components101.com/sensors/dht11-temperature-sensor

[6] Atmel Corporation, *ATmega328P*. San Jose: Atmel Corporation, 2015.

[7] ElectronicWings. (2021) SIM800 GSM GPRS module.

[8] Electronics Projects Focus, "LCD Working Principle,".

[9] FocusLCDS.com, "LCD MADE EASY,".

[10] San Salvador, El Salvador, "Air Quality Index," USA, 2012.

[11] Environmental Protection Agency, "Air Quality Index," vol. III, no. 12, 2014.

[12] Livio Lattanzio. (2021) Particulate Matter Sensing for Air Quality Measurements. [Online]. https://www.sensirion.com/en/about-us/newsroom/sensirion-specialist-articles/particulate-matter-sensing-for-air-quality-measurements/

[13] Minnesota Department of Health, "Preventing Carbon monoxide at home," st.Paul, 4/2017 ID #56687, 2017.

[14] J.Saini, M.Dutta and G.Marques, "Indoor Air Quality Monitoring Syatems Based on Internet of Things:A System Review," vol. 17, 2020.

[15] P.Pierucci and D.nIEGO E.Benfenati, "A Case study of indoor pollution by Chinese cooking ," vol. 65, 1998.

[16] D.Dey AND a.Chattopadhyay, "Solid Fuel Use in Kitchen and ChildHelath in India," vol.

58, 2016.

[17] G.Settimo and S.Capolongo M.Gola, "Indoor air healing environments:Monitoring chemical pollution in inpatient rooms," vol. 37, no. 9-10, 2018.

[18] M.A.Al Ahasan,S.Roy,A.H.M Saim, R.Akter and M.ZHossain, "Arduino Based Real Time Air Quality and Pollution Monitoring Sytem," vol. 6, no. 4, 2018.

[19] J.Jo, B.Jo,J.Kim, S.Kim and W.Ham, "Development of an IoT-Based Indoor Air Quality Monitoring Platform," vol. 2020, no. 13-15, 2020.

[20] W.Wei' O.Ramalho, L.Malingre,S.Sirvanantham, "Machine learning and statistical models for predictioing indoor air quality," vol. 29, no. 5, 2019.

[21] House air pollution and health. [Online]. https://www.who.int/news-rooms-room/fact-sheets/detail/household-air-pollution-and-health

[22] Pollution. [Online]. https://www.worldbank.org/en/topic/pollution

[23] M.Tastam and H.Gkozan, "Real Time Monitoring of indoor air quality with internet of thngs-based e-nose," vol. 9, no. 16, 2019.

[24] S.P.Kurilova and A.A.Belitsky P.D.Lapshima, "Development of an Arduino-Based CO2 Monitoring Device," no. 595-597, 2019.

[25] A.J.Alabdullah,B.I.Farhat and S.Chtouron, "Air Quality Arduino Based Monitoring System," no. 1-5, 2019.

[26] G.Marques,C.Roquee Ferreira and R.Pitarn, "Asystem based on the internet of things for real-time particle monitoring in buildings," vol. 15, 2018.

[27] A.M.Husain, T.H.Rini, M.I.Haque and M.R.Alam, "Air Quality Monitoring:The use of Arduino and Android," vol. 4, no. 1, 2016.

[28] A.Roch, A.M.Corrieia, H Adei,L.P.Reis and M.M.Teixeira, "New advances in information systems and technologies," vol. 445, no. 13-21, 2016.

[29] I.Vassura,E.Venturini,E.Bernadi,F.Passarini and G.Settimo, "Assessment of indoor pollution in a school environment through both passive and continuous samplings," vol. 14, no. 7, 2015.

[30] G.R.Ana,A.S.Alli,D.C.Uhiara and D.G.Shendell, "Indoor air quality and reported health symtoms among hair dressers in salons inIbadan," vol. 26, no. 1, 209.

[31] M.F.M.Firdhous,B.H.Sudantha and P.M.Karunaratne, "IoT enabled proactive indoor air quality monitoring system for suistainable health management," 2017.

[32] Sanayama Gunjai GH,Madhuri VM,Savitha Kamalapurkar, "Air Pollution Prediction Using Machine Learning Supervised Learning Approach," vol. 9, no. 04, 2020.