Food and Formalin Detector Using Machine Learning Approach

Afsana Azad Memi

ID: 2013–2–63–004

Nasrin Sultana

ID: 2014–1–60–060

Kanij Tabassum

ID: 2014–2–60–099

A thesis submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science in Computer Science and Engineering

Department of Computer Science and Engineering
East West University
Dhaka-1212, Bangladesh

September, 2018

Declaration

We, hereby, declare that the work presented in this thesis is the outcome of the investigation on "Formalin detection" performed by us under the supervision of Dr. Ahmed Wasif Reza, Associate Professor and Chairperson, Department of Computer Science and engineering, East West University. We also declare that no part of this thesis has been or is being submitted elsewhere for the award of any degree or diploma.

Countersigned   Signature

. . . .

Dr. Ahmed Wasif Reza   Afsana Azad Memi
Associate Professor & Chairperson   ID: 2013-2-63-004
Department of Computer Science & Engineering

                                                             Signature

                                                 . . . .

                                                 Nasrin Sultana
                                                 ID: 2014-1-60-060

                                                             Signature

                                                 . . . .

                                                 Kanij Tabassum
                                                 ID: 2014-2-60-099

Abstract

Unethical use of formalin, in the preservation of food items posing threat to public health. Without chemical experts accurately Formalin detection is a time consuming and complicated task. Moreover, the presence of naturally occurring formalin in food items may interfere in detecting artificially added formalin. Purpose of the study was to develop a simple cost-effective and reliable detection technique that can detect contaminated food. Therefore, quantifying artificially added formalin and naturally formed extent it is important to dynamically detect food for the comparison. With this view in mind, we have applied different machine learning algorithms like Naïve Bayes, Logistic regression, Support Vector Machine, K-NN Classifier on fruit's feature dataset to build a predictive model. We found that the K-NN algorithm works best in terms of accuracy. Finally using food conductance to electricity Rules have been developed and uploaded to the microcontroller unit. Combining with Arduino and the VOC HCHO gas sensor our own android application is able to detect 1-50 ppm of formalin. Several Tests are conducted and polynomial regression has been applied to predict the concentration of formalin in a given sample.

Acknowledgements

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Food is indispensable for surviving in the world. People consume food for the nutritional support of the human body. But now these foods especially fruits are intoxicated with the most dangerous chemical named Formalin. The extensive use of formalin, in the preservation of fish, fruit and other food items is posing a threat to public health. Formalin is an aqueous solution that is a mixture of water and 40% of formaldehyde [1]. The chemical used as a solution in water keeps fish fresh and makes fruits attractive [2]. A common use of this chemical is used to stop dead bodies from rotting, now being used to preserve edible items. No Smokey or harmful elements in formalin. But the presence of formalin in the air can be conceived very dangerous [4]. The exceeding level of formalin in air cause unfavorable outcomes like watery eyes; burning sensations in the eyes, nose, and throat; coughing; wheezing; nausea; and skin irritation [4]. The statement from NIOSH, 20 ppm of formalin is very harmful to human and life [2]. Considering the harmful condition of the human, OSHA set many standards to use formalin [4]. In the workplace like lab, office permissible exposure limit (PEL) for formalin is 0.75 parts per million parts of air (0.75 ppm) [4]. Another standard is the short-term exposure limit (STEL) that's range is 2ppm.

Now the concentration of formalin can be detected various way like e-nose technology, sensing formalin chemically [4]. To know the concentration of formalin a detector can be used. The detector can detect the formalin by sensing the presence of formalin in the air. This volatile chemical compound detector helps a person to identify the fruit which he/she doesn't know Contaminated or not [3]. This detector also helps the person to know the concentration of formalin which is safe or unsafe for his/her health. However, Formalin can be naturally found in foods, such as meats, fish, fruits and vegetables, dried mushroom and crustacean, as a typical metabolic by-product [5]. Presence of formalin content conjointly varies with the food conditions and items. This doesn't imply that formaldehyde isn't toxic. Any intense compound is probably going to be harmful in large doses; however, the span of the dosage, how it is devoured and who is presented to it makes a difference. As formalin naturally presents in food items, it could interfere with artificially added formalin when they are going to be detected. It is important to quantify the artificially added formalin from naturally occurring formalin in food items. We have applied various algorithms to the dataset we are working on. The accuracy of detecting sample food of different algorithms is different. We measured the accuracy of the following algorithms:

i.   Naïve Bayes

ii.   Logistic Regression

iii.   Support Vector Machine

iv.   K-NN classification

v.   Ruled based classification

The performance measure was different for each of the algorithms. We compared these algorithms on the same train and test dataset so that the decision of which algorithm works better on these algorithms can be taken. We found that K-NN Classifier provides the best accuracy among the algorithms we applied. The other algorithm also did well in other performance measures. Later in this paper, these performance measures are compared.

1.2 Motivation

The extensive admixture of up to date fruits, vegetables and fish with malignant concentrations of formalin has become a crucial outcome in People's Republic of Bangladesh which causes health risk of the human [5]. Due to lack of adequate observation and raids, an Excessive concentration of solution has been found in most fruits of the native market. Most of our citizens frequently shop for fruits and vegetables. Half of them buy these from the edge hawkers. These places are not monitored by mobile courts. For this to detect the moderate level of formalin in food especially in fruits and vegetables, we have to develop the idea of formalin detection system using machine learning algorithm like polynomial regression, rule base classification.

1.3 Problem Statement

Most of the fruits and vegetables naturally contain formalin in them and the amount varies with the food items. For this reason fruit identification is necessary for separating natural occurring formalin from artificially added formalin. Only sensing raw formalin without a predefine model of naturally formed formalin result could be misleading. On the other hand, sensing formalin chemically in contaminated food is not appropriate for general use. A digitalized device with pre-trained model could be a better solution for non-technical person using their own sell phone.

1.4 Objectives

The main objective of the project under the scrutiny is:

1. To develop an IoT based designed model to detect the presence of formalin in fruits and vegetables using machine learning approaches.

1.5 Outline of this Report

The layout of rest of this report has been organized as takes after.

Chapter 2 presents the literature review on the topic of Fruit and Formalin Detection.

Chapter 3 clarifies hardware specification of the components used in the project.

Chapter 4 explains methodology of the project that how the project being implemented.

Chapter 5 presents the real implementation of the project.

Chapter 6 outlines conclusion of this work precisely and describes the scope of future work for possible

extended application of the project.

Chapter 2

Literature Review

The surveys from various papers are taken and considered. Some of them are given below:

Fruits and vegetables can be detected by a machine in many ways. Arduino based food detection is a simple circuit using resistor can be found in and they detect sample food by measuring the resistance in them. They check resistance range of different foods using Serial Output and use those resistances to make create a range to detect food and display the food name on the LCD screen.

Some paper is mainly focused on the implementation and testing of formalin detector like. The Intense level of using formalin almost all kinds of raw foods, vegetables and dairy products have increased day by day alarmingly and terrifyingly in Bangladesh. Many researches worked to detect the extent of formalin use in fish available in Dhaka city. They collected five species of fishes from five several local markets of Dhaka city and measure the amount of formalin on their samples. Electronic-nose technologies and applications have come through advances in sensor design, software innovations, material improvements and progress in micro circuitry design and systems integration in the present world is proposed in [6]. Different types of sensory instruments and using e-nose technology in agriculture, forestry can be found in [5] and their applications applied in food and dairy industry can be found in [7], pharmaceutical industries and food application using e-noses tongues [8] are also reviewed of this papers. [9]. Another approach is to sense formalin chemically in contaminated food [10]. Simple spot test quantification methods to determine formaldehyde in aqueous samples were also proposed [11]. There are also several other methods proposed for formaldehyde detection, including (pulsed amperometric detection) PAD method [12], solid phase microextraction (SPME) extraction and gas chromatography-mass spectroscopy (GC–MS) analysis [13], gas chromatography-mass spectrometry with selected ion monitor (GC-MS-SIM) [14], by enhanced chemiluminescence [15], using cd-doped TiO2-SnO2 sensor [16] and highly sensitive colorimetric detection method for gaseous formaldehyde [21]. In [17], information is presented on five different local markets of Dhaka city in Bangladesh, where the presence of formalin was detected using the formalin detection kit for fish developed by Bangladesh Council of Scientific and Industrial Research (BCSIR).

Some paper proposes a microcontroller based electronic formalin detection system along with an Android smartphone app to show the results and status of sample food [5], [2]. They applied all methods in a specific environment and different temperatures. However in real life, it may tough to detect the formalin

accuracy at any fixed environment and temperature. We know that formalin is a volatile 40% aqueous solution of formaldehyde, this discussed in [21]. Grove sensor is able to detect the presence of formalin. The voltage output of the HCHO sensor is exponentially proportional to the concentration of formaldehyde gas [11]. In paper [5], their approach was mainly based on Formalin detector system incorporating supervised learning. They implemented a microcontroller based hardware using Grove VOC HCHO gas sensor and developed an android application and detect formalin in different temperatures. They took some sample data from their experimental sample and applied supervised machine learning approach like logistic regression and SVM (Support Vector machine). They used reference as formalin kit which is available in the market. Some Research was developed [3] on the detector that can sense the presence of formaldehyde in air measured in parts per million (ppm) and used SD technology. They used MQ138 which the semiconductor gas sensor used for formaldehyde detection and also monitored the air quality. In computer science, machine learning itself a term intended to identify a singular model that can be applied to any instance in order to make a prediction from sample dataset.

2.1 Machine Learning Algorithms

In this section we briefly discuss Naïve Bayes, Logistic regression, Support Vector Machine, K-NN, ruled based Classifier and Polynomial regression, that we applied for detecting the sample fruits and vegetables and formaldehyde in them.

2.1.1 Naïve Bayes

Naive Bayes is a first processing algorithm relative to other machine learning algorithms. It works on Bayes theorem of probability to predict the class of unknown data set. Naive Bayes model is easy to build and particularly useful for very large data sets. It also performs well in multiclass prediction. It is an eager learning algorithm. It is able to make predictions in real time.

2.1.2 Support Vector Machine

Support Vector Machine a supervised machine learning algorithm which can be used for both classification and regression challenges. It is a classification technique that seeks to find a hyperplane that partitions the data by their class label and at the same time avoid over-filtering the data. The learning of the hyperplane in linear SVM is done by transforming the problem by using linear algebra. And it makes a binary classification based on separating hyperplane on a remapped instance space. In SVM, it is easy to have a linear hyperplane between these two classes.

2.1.3 K-NN

KNN means K-Nearest Neighbors algorithm. KNN algorithm is one of the simplest classification algorithms and it is one of the most used learning algorithms. Its purpose is to use a database in which the data points are separated into several classes to predict the classification of a new sample point. KNN

is also a lazy algorithm as opposed to an eager algorithm. K-NN keeps all the training data. For more accuracy, K-NN requires all the training data for the testing phase.

2.1.4 Logistic Regression

Logistic regression is the suitable regression analysis to conduct when the dependent variable is dichotomous. Logistic regression is an analytical analysis. Logistic regression is used to define data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables.

2.1.5 Rule-base Modeling Approach

Rules base machine learning model approach uses a set of rules that specify a mathematical model thus can be translated into Markov chain or different equations. It is especially effective where set of rules is easier than the model it implies, that means the model repeats the demonstration of a limited range of patterns. A biochemical model of living organisms is an important domain for its appliance. Rule base classification can process both categorical and numerical data. Basically rule base classification is a straightforward way to classify data generating if-then rules to cover all cases.

2.1.6 Confusion matrix

Confusion matrix is a convenient way of performance measure of any predictive model. We have calculated Accuracy of the above mentioned algorithms. Accuracy show that how often the classifier is correct. It is mostly used figure to rate any classifier. To measure the performance metrics, the use of confusion matrix is useful and convenient. A confusion matrix is introduced to represent the relation between the actual and predicted class.

Table 2.1: Confusion Matrix

		Prediction	
		Positive	**Negative**
Actual	**Total=** TP+TN+FP+FN		
	Positive	**TP**	**FN**
	Negative	**FP**	**TN**

From this confusion matrix, accuracy is calculated by the equation,

$$Accuracy \ = \ \frac{TP + TN}{Total} \qquad (2.1)$$

We can calculate various performance measures from the confusion matrix of any given algorithm along with its accuracy.

2.1.7 Polynomial regression

Polynomial regression is a popular machine learning algorithm. It can be consider as a special type of linear regression. It has two axis X axis stands for the data value and Y axis is used for the Target value. Polynomial regression might better fit such cases where data may be scattered not linear. Below equation shows the Quadratic Equation of polynomial of degree 3 and where ω is constant.

$$Y = \omega_0 + \omega_1 * x + \omega_2 * x^2 + \omega_3 * x^3 \qquad (2.2)$$

After reviewed this papers, we observed that using e-nose technologies are easy and cost-effective. There are many proposals which based on how formaldehyde can be detected in contaminated foods. We chose the approach which ensures quality measurement of contaminated fruits and vegetables by quality assessment using sensor network.

Chapter 3

Familiarization of Hardware

In this chapter, we depict the whole inserted arrangement of our venture where the entire fundamental parts name, their hugeness in the undertaking are portrayed intricately.

3.1 Arduino UNO

Founded on the Atmega328 datasheet Arduino UNO is made. It is a microcontroller board. It consists fourteen(14) digital input/output pins among these six(6) will be used as PWM outputs, six(6) analog inputs, a sixteen(16) megahertz quartz oscillator, a USB connection, a power jack, an ICSP header, and a reset button. Everything has in it for the assistance of the microcontroller; that helps to connect it to a PC with a USB cable or power it with an AC-to-DC adapter or battery to start. **Fig 3.1** refers Arduino Uno R3.
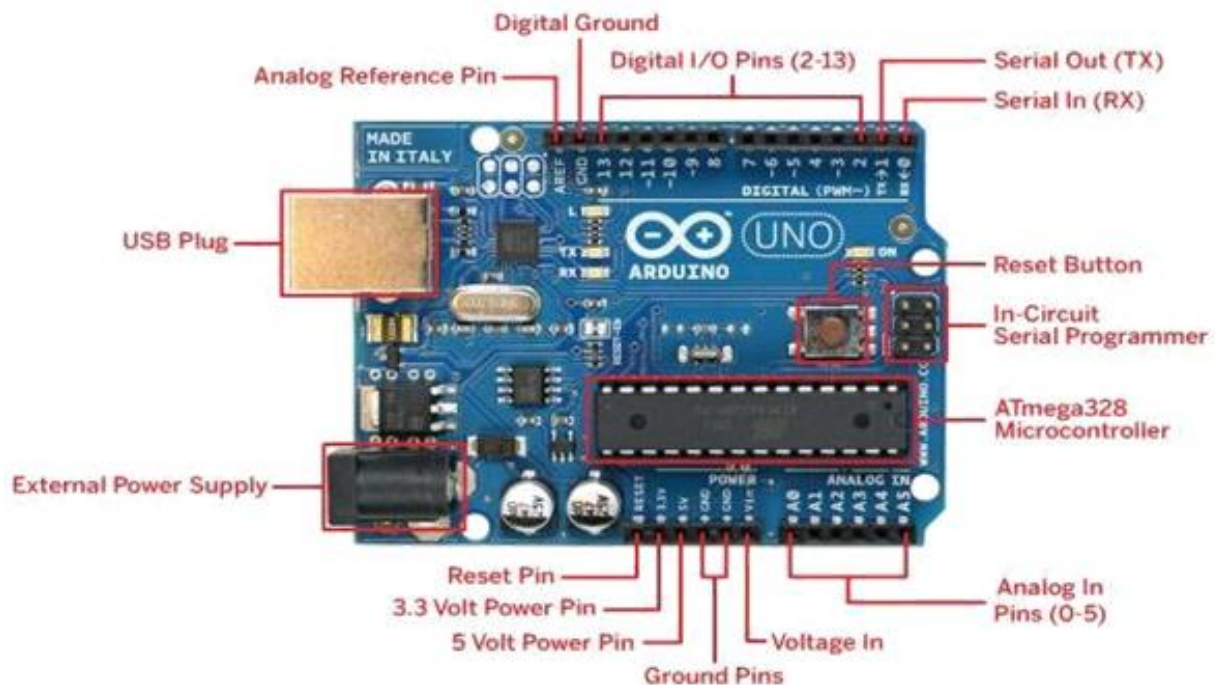
Fig 3.1: Arduino Uno R3

3.1.1 Features Of Arduino Uno:

- ❖ Microcontroller   : ATmega328
- ❖ Operating Voltage   : 5V

- ❖ Input Voltage (recommended)   : 7-12V
- ❖ Input Voltage (limits)   : 6-20V
- ❖ Digital I/O Pins   : 14 (of which 6 provide PWM output)
- ❖ Analog Input Pins   : 6
- ❖ DC Current per I/O Pin   : 40 mA
- ❖ DC Current for 3.3V Pin   : 50 mA
- ❖ Flash Memory   : 32 KB of which 0.5 KB used for boot loader
- ❖ SRAM   : 2 KB (ATmega328)
- ❖ EEPROM   : 1 KB (ATmega328)
- ❖ Clock Speed   : 16 MHz

3.1.2 Arduino Power

With the help of USB cable or an outer power supply Arduino UNO is powered. The supply source is chosen instinctively. From an AC-to-DC adapter or battery, the outer power will come back. A 2.1 mm center-positive plug is plugged into the board's power jack and the adapter is associated with it. At the top of the power connector the GND and Vin pins are implanted to conduct the battery.6 to 20 volts is the outer power supply to conduct the board. When the supply power is less than 7 volts, the board could be changeable. Again if the supply power is more than 12 volt, so voltage controller will be overheated. Then the board is spoiled. So the suggested range of the power is 7 to 12 volts. **Fig 3.2** shows Arduino Schematic I/O pins.

Fig 3.2: Arduino Schematic I/O pins

3.1.3 Arduino Power PINS

- VIN: The Arduino board uses the input voltage (VIN) when it is associated with the outer power supply.
- 5V: To provide power in the microcontroller and different parts on the board, a regulated power supply is used. This power comes back through the board regulator using input voltage (VIN) or by USB or by 5V power supply.
- 3V3: From the on-board regulator a 3.3 V power supply is produced. Flow of the current is maximum 50 mA.
- GND: GND is indicated the Ground pins.

3.1.4 Memory

For the code saving, The Atmega328 contains 32 KB of flash memory. In this 0.5 KB is employed for the bootloader. 2 KB of SRAM and 1 KB of EEPROM is also in it.

3.1.5 Arduino Inputs and Outputs

Arduino has fourteen digital input/output pins. By using pinMode(), digitalRead() and digitalWrite() functions,this pins are used.At 5V,every pins of arduino is operated. It contains an interior pull-up resistor which resistance is 20-50 kohms. It is disconnected by default. Among these fourteen pins, some pins have specific functions that are given below:

- **Serial Pins 0 (Rx) and 1 (Tx)**:

  Rx pins receive the TTL serial information and Tx pins transmit the TTL serial information. They're associated with same ATmega328P USB to TTL serial chip.
- **External Interrupt Pins 2 and 3:**

  For triggering interrupt or low value, these pins are organized.
- **PWM Pins 3, 5, 6, 9 and 11:**

  Using the analog Write () function, these pins produce an 8-bit PWM output.
- **SPI Pins 10 (SS), 11 (MOSI), 12 (MISO) and 13 (SCK):**

  For the communication of SPI, these pins are used.
- **In-built LED Pin 13:**

  With a built-in Light Emitting Diode, the pins are associated. LED is on; if pin 13 is at high voltage and LED is off, if pin 13 is at low voltage.

Arduino has some other pins like AREF and Reset Pin.

3.2 Grove HCHO Sensor

The Grove HCHO is a gas sensor. It follows semi conductivity rules to sense an element presence in the air. So it is also a semiconductor gas sensor. In this WSP2110 VOC gas sensor is used for sensing the

element. Its conductivity changes with the concentration of gas exist in the air. Grove HCHO sensor is very sensitive and stable sensor because it's ability to sense a gas which concentration is up to 1 ppm. Now this sensor is used for detecting formaldehyde combining with arduino. The output voltage of the sensor is exponentially proportional with the concentration of formalin gas [18]. Grove HCHO sensor is showing on **Fig. 3.3**.

Fig 3.3: Grove HCHO sensor

3.2.1 Features of Grove HCHO Sensor:

➢ Operating Voltage: 5.0V ±0.3V

➢ Target Gases: HCHO, Benzene, Alcohol

➢ Concentration Range: 1~50 ppm

➢ Sensor Resistance Value(Rs): 10KΩ-100KΩ(in 10ppm HCHO)

➢ Sensitivity: Rs (in air)/ Rs (10ppm HCHO)≥5

➢ Operating Voltage: 5.0V ±0.3V

3.3 Bluetooth Module hc-05

HC-05 Bluetooth module is a simple module for established a wireless serial connection. It is communicated with the microcontroller or PC through serial communication. The HC-05 Bluetooth module contains two functional modes. First one is data mode. It can send and receive data from any other Bluetooth features. The last one is AT command mode. In this mode the settings of the device can be modified by default**.** Both side of Bluetooth Module hc-05 is shows in **Fig 3.4.**
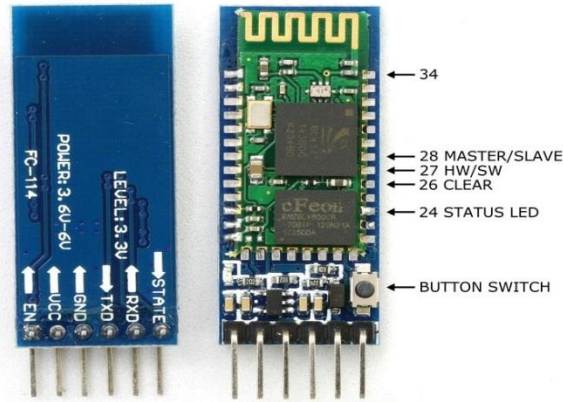
Fig 3.4: Bluetooth Module hc-05

3.3.1 Features of Bluetooth module:

➢ Model: HC-05

➢ Input Voltage: DC 5V

➢ Communication Method: Serial Communication

3.4 Resistors

Resistors are electronic elements that are used in the circuits to decrease the flow of the current and at the same time to divide the voltages. It also modifies the signal levels and terminates the lines of transmission in the circuits and microcontrollers. For the setup of circuit elements, resistors sometimes may be employed as sensing devices for light, heat and chemical action. The resistance of a resistor is measured in ohms. Resistors are made of carbon, metal or metal-oxide film. A skinny film of conductive material is wrapped like a helix in the resistors. The resistors are coated by a material which is insulating. **Fig 3.5** shows Resistors.

Fig 3.5: Resistors

Above mentioned equipment's are the main components for our developed device. Connection diagrams were discussed in next chapter.

Chapter 4

Design and Methodology

4.1 Decisions of design specification and hardware

The system detects the sample fruit by measuring the different range of resistances using serial output of Arduino. System uses the voltage drop to calculate the resistance associated with the food items. Rule based machine learning decision model is used to detect the given fruit by comparing the range of resistances in different fruits along with other features. The HCHO sensor senses the concentration of the formaldehyde from the detected fruit by placing near it. Considering naturally occurring formalin [22] in a fruits, our system identifies the level of artificially added formalin in a sample food. HC-05 Bluetooth module is used to interface the device with our self- made Android application. Apps display the food status along with concentration of formalin. A LED is used as an indicator for the presentence of high level of formalin concentration. All the parts are briefly discussed in this section. Following block Diagram shows the design of the food and formalin detection system.
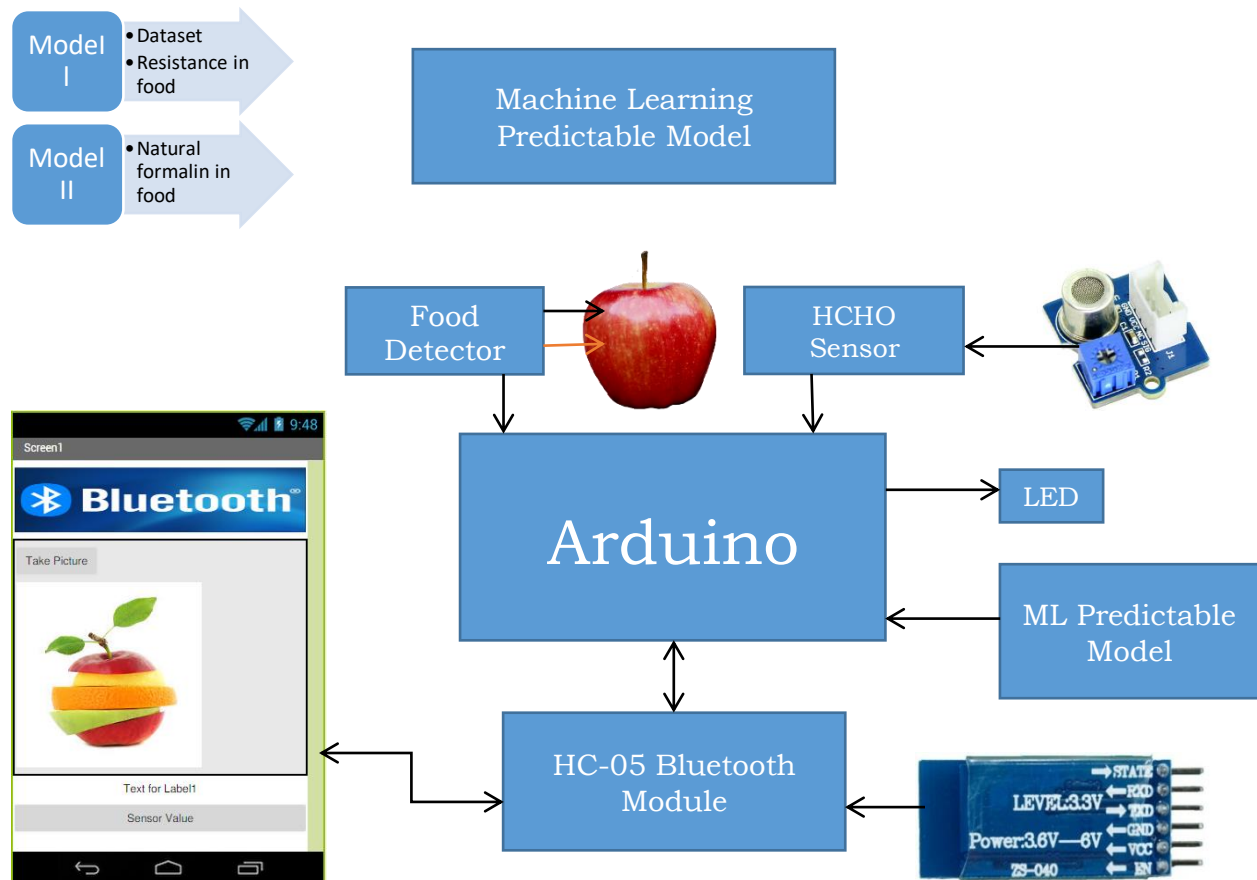

Fig 4.1: Block Diagram of the Food and Formalin Detector Device.

The system shows status based on the fruit it detected and the sensor value, output voltage and concentration of formalin in ppm value. Comparing with naturally occurring formalin in food it considers either the food is "safe" or "Dangerous" for health.

4.2 Working Procedures of the Device

The working principle of the Food and formalin detection system is elaborated in the **Fig.** 4.2.
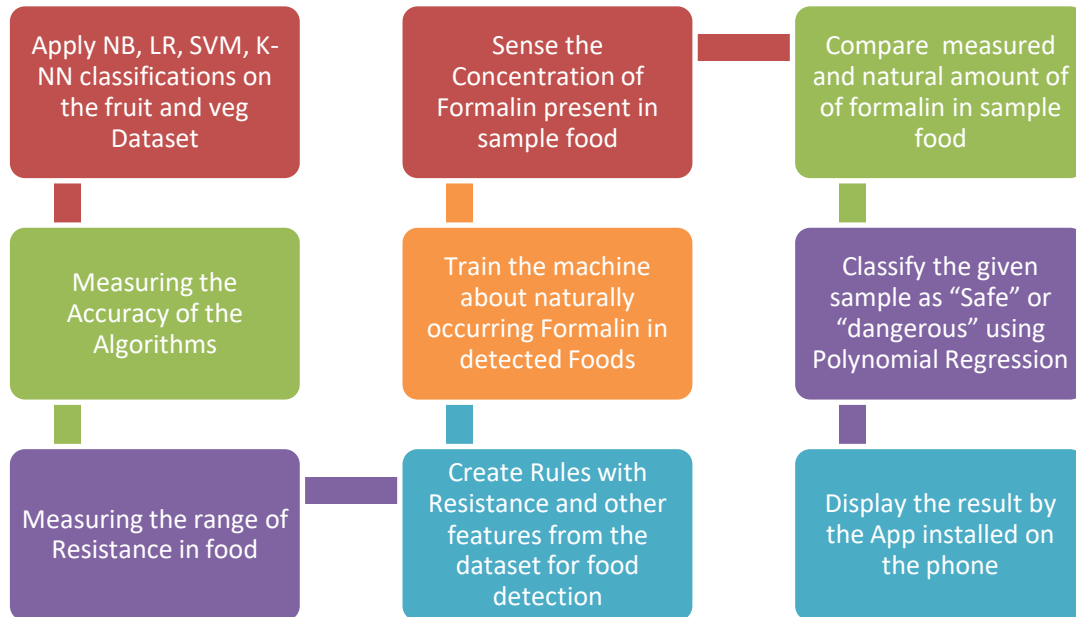
Fig 4.2: Working procedure of the Food and Formalin detector System.

Our first target was to detect sample food by the machine with the help of machine leaning approaches. Second, target was to detect the formalin in the detected fruits. Finally display the result by the Andriod App. Each steps of working procedures of our whole system are described completely in the next chapter.

4.3 Preparing Dataset

Dataset we used to build the predictive model was generated by Dr. Iain Murray from University of Edinburgh. The fruits dataset was also slightly formatted by professors at University of Michigan. The dataset has seven columns with fruit's label, fruit's name, fruit's subtype, mass, width, height and color score. There are four types of fruits in the dataset: 'apple' 'mandarin' 'orange' 'lemon'. As our working requirements we add two vegetables: carrot and cucumber with the original dataset. Then the dataset was preprocessed to predict different fruits and vegetables as different class. Then we used this preprocessed dataset to train the predictive model to detect given sample fruits and vegetables. The dataset is quite balanced and Fig shows the graphical representation of the class level. Feature scaling was used to the dataset for scaling all the features at the same range.
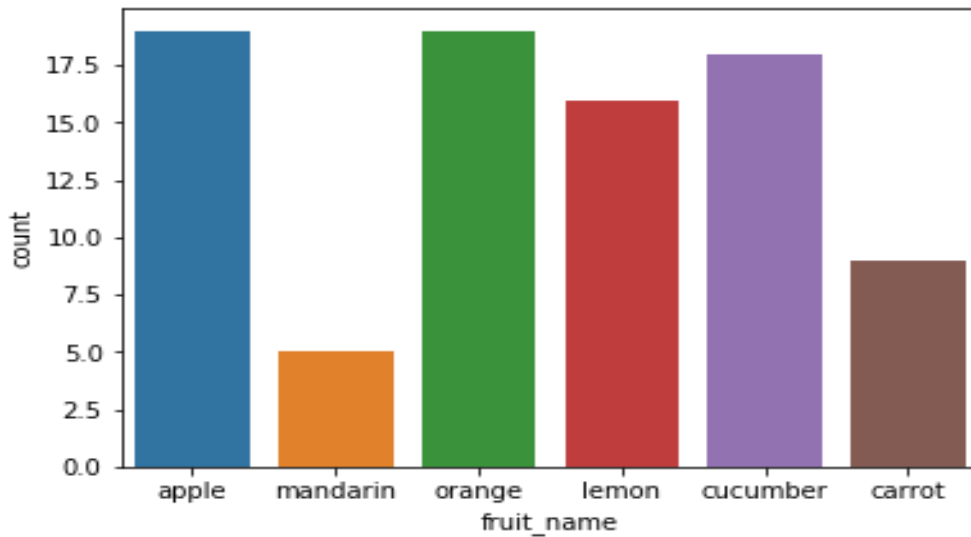
Fig 4.3: Six categories of class levels
Apple (19), mandarin (5), orange (19), lemon (16), cucumber (17) or carrot (8).

4.4 Hardware Design and Implementations

A Development research method is applied in our study to come up with a solution to find out formaldehyde detection problem in our daily life. This included hardware assembly and machine learning approaches for the development. The Arduino Uno R3 microcontroller board is attached with other input and output components: Grove HCHO sensor, HC-05 Bluetooth module, 10k ohm resistors and LED. ATmega328 AVR microcontroller is a dual-inline-package interface with CH340G chip. This works as the brain of the whole system and empowers the system to facilitate the process of formalin detection. The system is comprised of code using Arduino programing language, testing, debugging and simulating with Arduino IDE.

4.4.1 Hardwire Design of Food Detector

Food detector is a very simple low cost device which is basically follows the principles of physics and chemistry. Different resistances offered by the several food items are measured by the Arduino's analog inputs. For developing the system we used-

- Arduino uno
- Resistor 10k Ohm
- Breadboard
- Jumper Wires
- Computer with Adruino IDE
- USB cable to connect the Arduino to the laptop
- I2C 16x2 LCD Display
- Sample fruits

We know vegetables and fruits could be good source of conductor of electricity because of the presence of juice in them which forms mild acid. Acids are able to conduct electricity just like galvanic cell [19] where conductivity depends on the amount of chemical and water in them. More watery and chemical compound will allow more electrical energy to pass through. Acidic fruits are good conductor of electricity.
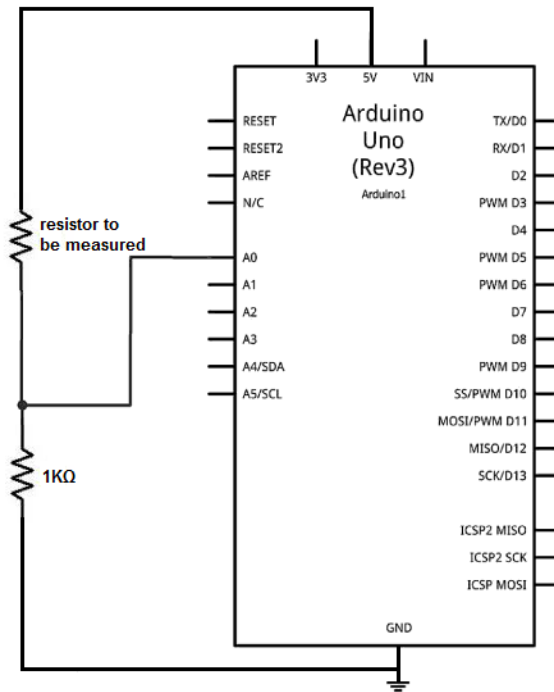
Fig 4.4: Schematic of Arduino-ohmmeter

Our system works based on a very simple version of ohm-meter using Arduino. The resistance of fruit or vegetable is unknown. Analog pin (A0) is used to measure the voltage drop across the resistance in the piece of fruit or vegetable. Larger readings offered lesser resistance in the food and this contains more conductivity.

4.4.1.1 Hardware connection

First we place a 10k ohm resistor on the breadboard in parallel and connect one point of the resistor with Arduino A0 analog pin. Then connect another point resistor with GND from Arduino as shown in the **Fig. 4.5**. A connection wire directly inserted into sample food and it's another terminal comes from Arduino 5 volt pin. The foods we want to test make a close circuit with Arduino. We use those resistances in the food and make range of different foods. Arduino's analog input is using here to measure this resistance offered by the foods. Figure shows the connection diagram of the food detector system where the basic idea is based on ohm meter finding the value of unknown resistor with a known resistor with the help of Arduino.
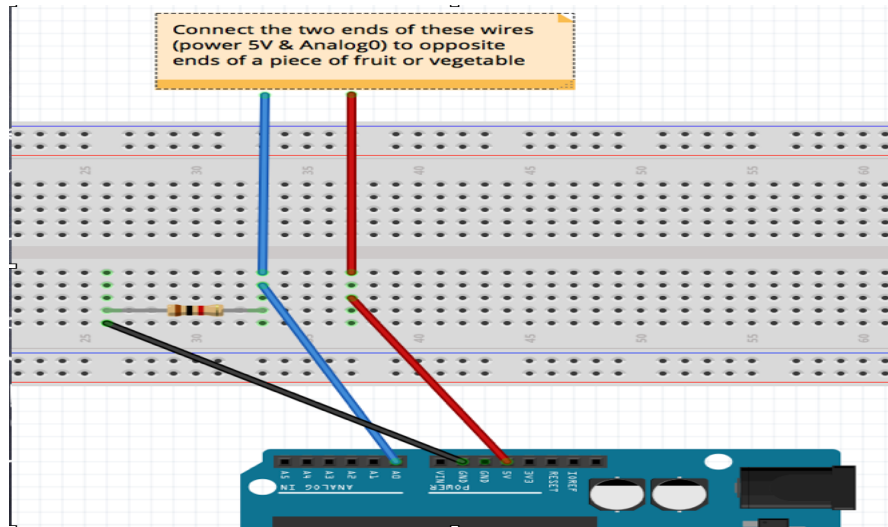
Fig 4.5: Connection Diagram of Food detector.

We create different ranges based on the measured resistances. We write a code for detecting apple, carrot, cucumber and red plum. The code is added at Appendix. However, the resistance values kept changing anomalously from time to time. The reason behind this could be the chemical changes inside the fruit due to air were changing the resistance. Following block Diagram shows the design of the food detection system.
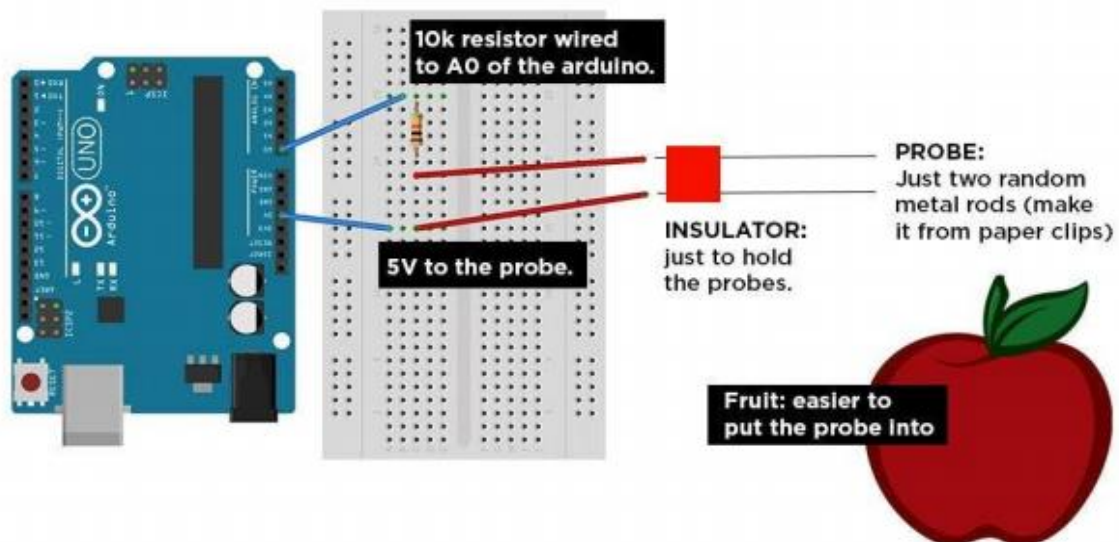
Fig 4.6: Block Diagram of the Food Detector Device.

4.4.2 Design of formalin detector sensor

Grove HCHO was the semiconductor VOC sensor used for detecting formaldehyde combining with arduino. This sensor featured to detect gas concentration up to 1 ppm. As formaldehyde is self-vaporized

solution, its presence can be detected by volatile organic compound sensor. The output voltage of the sensor is exponentially proportional with the concentration of formalin gas [18].

4.4.2.1 HCHO Sensor with Arduino

For Connecting HCHO sensor with Arduino required equipment's are:

- 1 Arduino Board (Uno or Mega)
- 1 Seed HCHO Sensor
- USB cable according to Arduino
- Computer with Adruino IDE
- Connecting wires.

Hardware Connection:

HCHO Sensor	Arduino Uno
GND	GND
VCC	5 Volt
Sig	Analog Pin (A_0)

Connection diagram of Arduino and the sensor is show in the below **Fig 4.7**. Then we plug the assemble Arduino into our laptop and load up Arduino IDE. The code is added to the Appendix B.
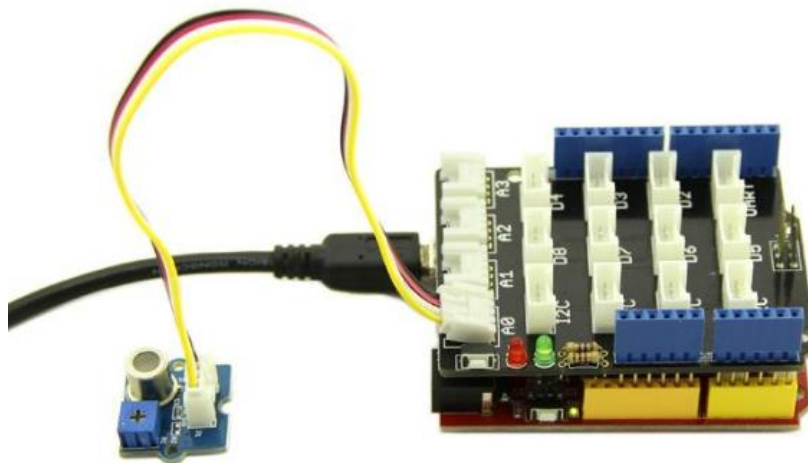
Fig 4.7: Grove HCHO Sensor connected with the Arduino.

4.4.2.2 Design of HC-05 Bluetooth module connection

In our system we interface Arduino UNO with our adriod app via HC-05 Bluetooth Module. Here we used Communication Mode. The hardware connection between HC-05 Bluetooth Module and Arduino Uno can be shown by the diagram given below:
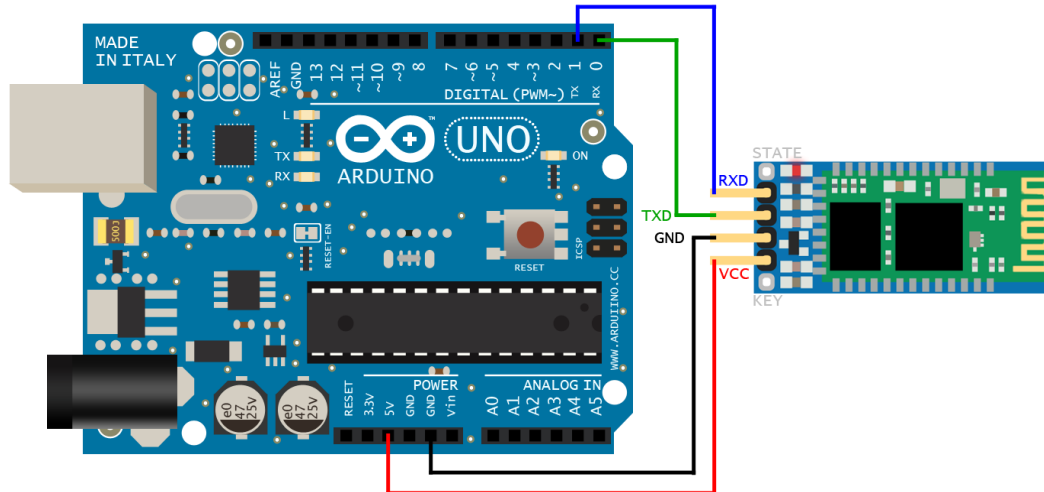
Fig 4.8: HC-05 Bluetooth module connected with Arduino.

Hardware Connection:

VCC connects to Arduino 5V

GND connects to Arduino GND

TXD connects to Arduino Pin RX

RXD connects to Arduino Pin TX

KEY connects to the air for communication mode.

## 4.5 Android App Development

Our Android application is developed by using MIT App Inventor 2. MIT App Inventor is a cloud based integrated Development Environment (IDE). This is a web based tool for building mobile apps on Android platform by drag and drop option. Graphical interface is used here which is very similar to Scratch. Block-based programming language is used to build the Application which is also known as visual programming [20]. The Designer, Block Editor and Android Emulator are the three section of the app inventor where the work takes place.

4.5.1 App Interface Design

For developing our Apps we follow the below steps:

1. User Interface-Our user interface palette is very simple. One Label and one ListPicker is used which is visible to the user. Label1 is used to display all the result we want to see in our phone Screen. It displays the instruction for the food we detected and sensor value, output voltage and concentration of formalin in ppm value in a sample food received from our device. It also displays the   Temperature and humidity of the environment. ListPicker button is used for whose text set as 'Get'. Two Button is used in our app. Button1 is TextToSpeech1 for reading the results from Label1. Button2 is for taking the picture of the food we want to test which is saved in our phone memory. The image will be displayed in the phone screen.

2. Connectivity Palette-Four nonvisible components are used like BluetoothClient1, Clock1, TextToSpeech1 and TinyDB1. BluetoothClient1 and Clock1 required for connecting the device and the app with the Bluetooth module. TextToSpeech1 is used to convert the text into speech. TinyDB1 is required for saving the picture in the SD Card.

4.5.2 App Blocks Design

Blocks are program the way we want our app to behavior. These Build-in blocks control things like logic, math and the text that want to add with each component. Bluetooth Connection: First we select Before picking from ListPicker and add set the element of list to show the available Bluetooth devices.

Fig 4.9: Before picking list to show the available Bluetooth devices.

Again we select After picking from ListPicker and add set status label for connecting with our device to Bluetooth App. When we click on the connect button, connection between HC-05 module and app will be established.
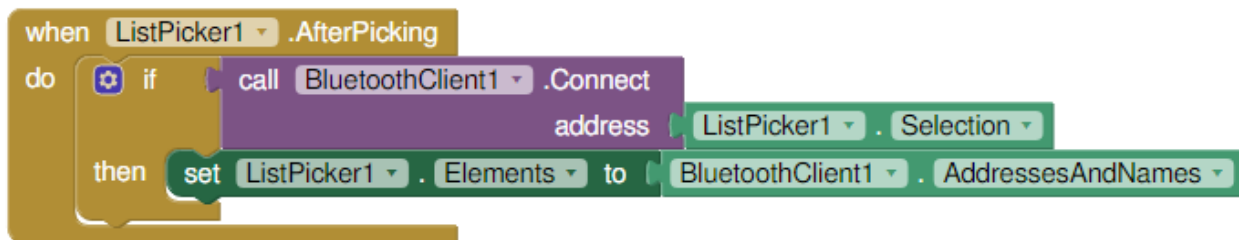
Fig 4.10: After picking list to set status label for connecting with the device.

When user press on Bluetooth image its elements set as paired Bluetooth address and name so it will show the paired Bluetooth address which is 48 bit long unique code. The screen will be look like **Fig. 4.11**.
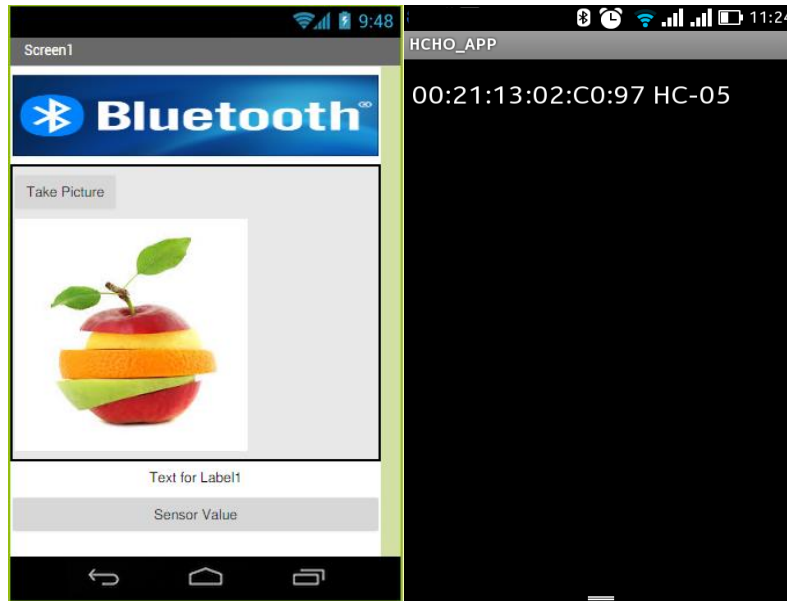
Fig 4.11: User interface (Left-side) and Available Bluetooth Address (Right-side)

Take image and display: We use button2 for clicking photo of our sample photo to be tested. When the button is pressed it called the camera's TakePicture method. Once camera finish taking picture, AfterPicture event occurs which fetch the image. It assigns the image to the picture property of the image control. This makes copy of the image that was just taken and displays it on our phone screen. TinyBD1 is used to store the taken picture in local directory automatically.
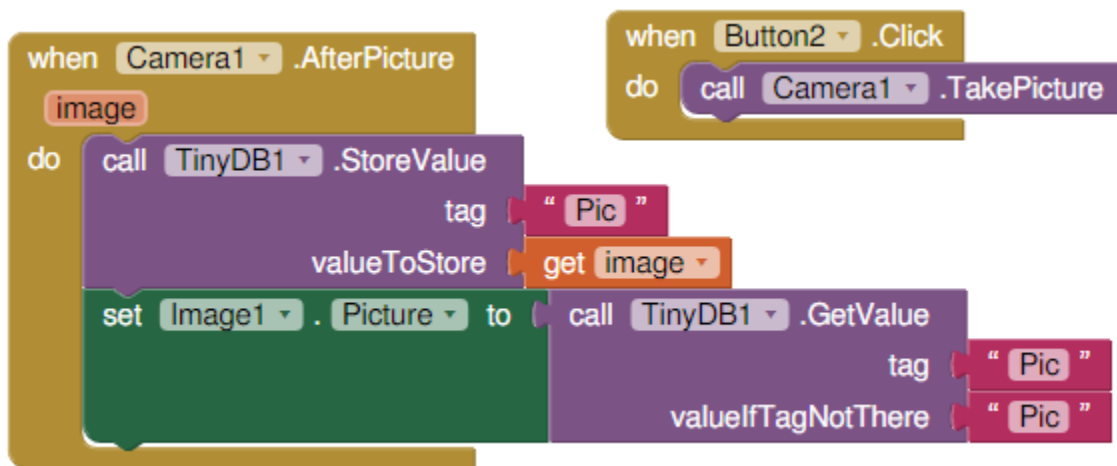
Fig 4.12: Block of Take picture and save it in local directory.

Receive Data: This section is added to display all the value we received from the Arduino. When device is connected with the app it displays the detected food name and the concentration of formalin in it. Based on that it shows either the food is safe or dangerous for taking.
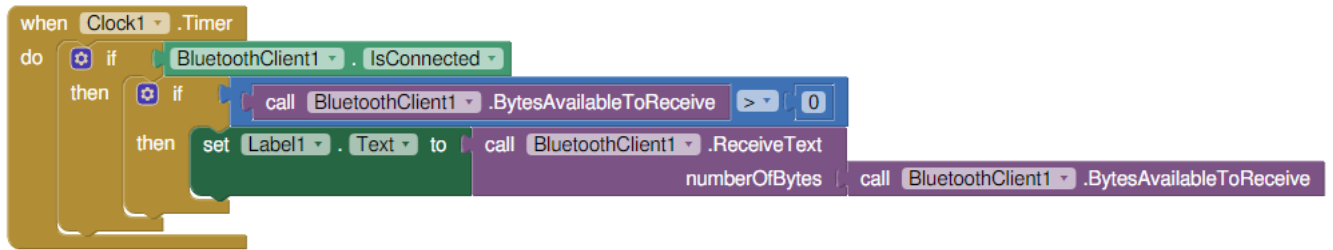
Fig 4.13: Block of Receiving Data from the connected device.

Text to speech: This Block is added to read the data we display in label1. When sensor value button is pressed it read all the available data.
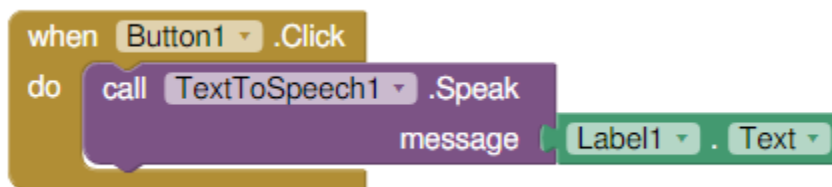
Fig 4.14: Block of Reading Data from the connected device.

4.6 Complete Device

The complete device of Food and Formalin detection is show in **Fig. 4.15** which is a comprehensive low-cost device. This package is composed with Arduino, sensor, Bluetooth module, LED and our self-made Android App.
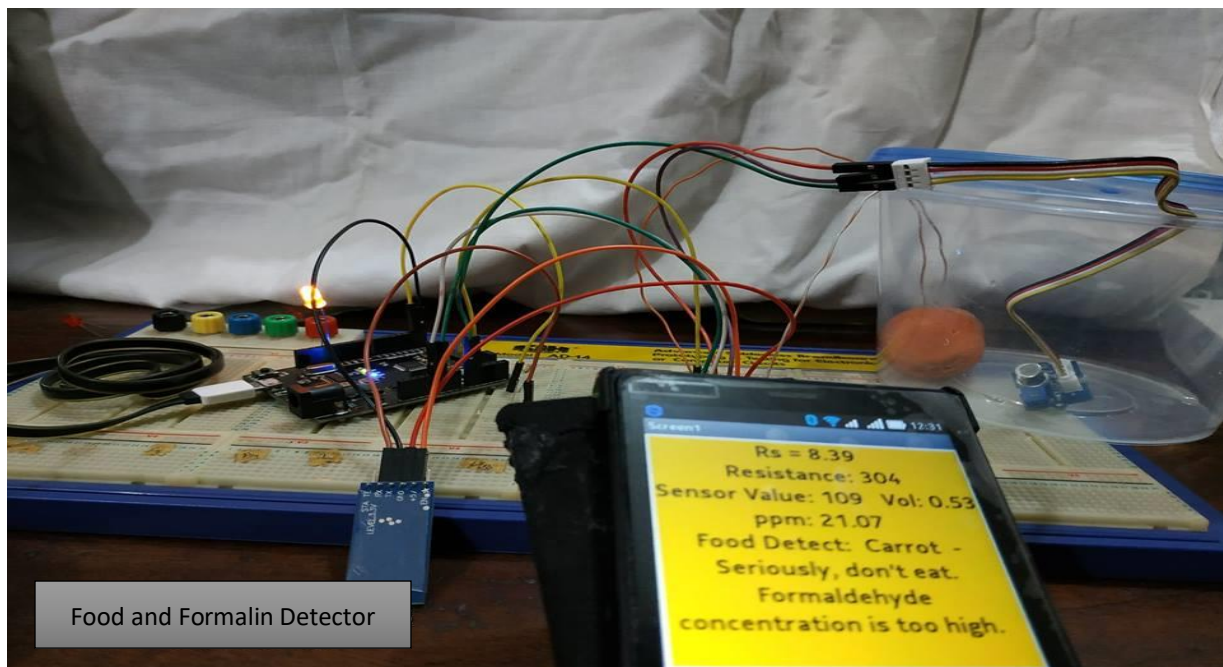
Fig 4.15: Final Device

4.7 Flowchart of the Program

After designing the working principle, the flow chart of the system is implemented where the code and the model is developed and tested. The flowchart of the complete system is shown in **Fig. 4.16**.
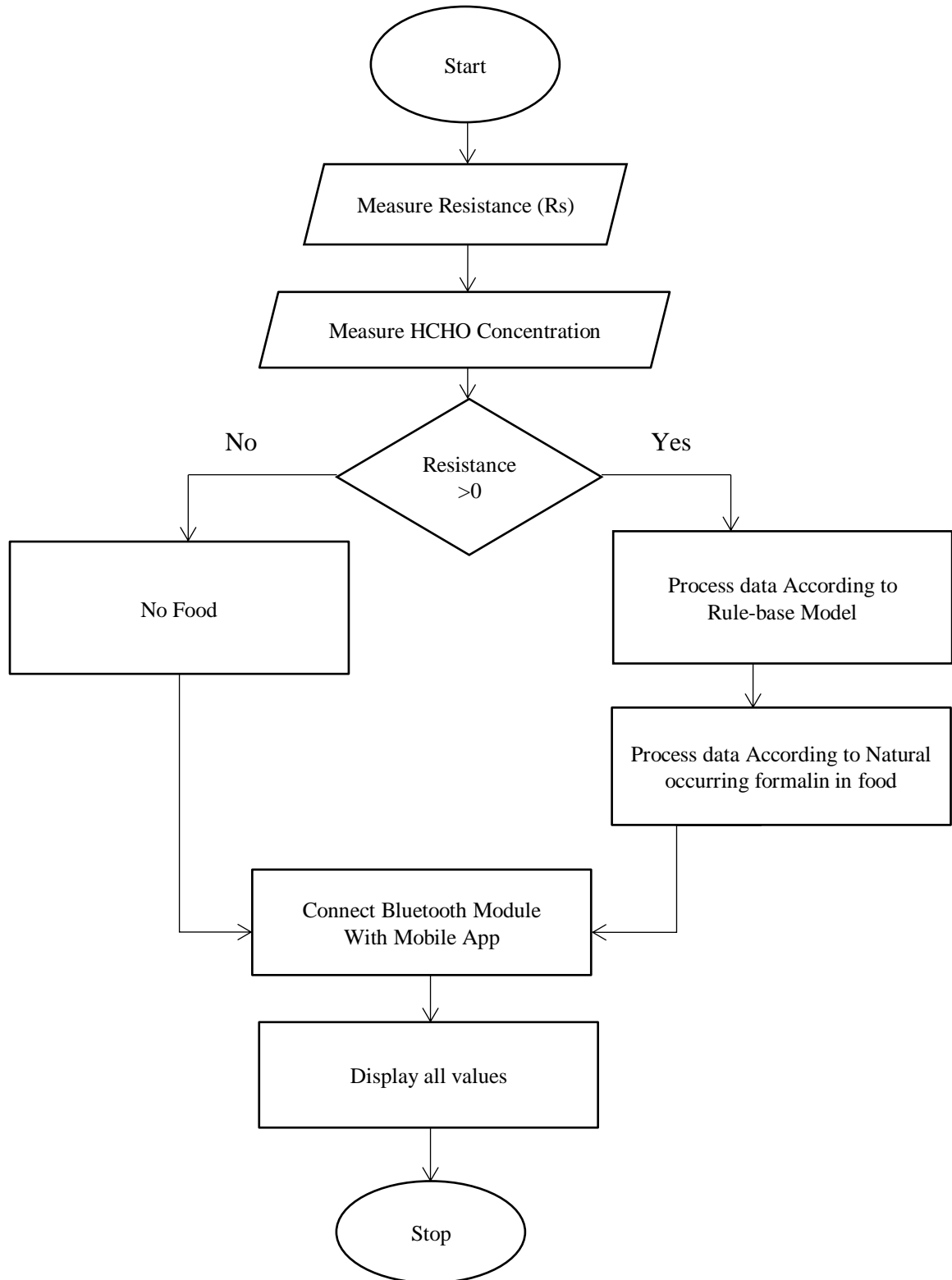
Fig 4.16: Flow Chart for the Complete Program

Chapter 5

Experiments and Result Analysis

5.1 Performance of Various algorithms

Different algorithms have provided different accuracies and their performances were measured. We calculated the accuracy of each of the algorithms from the confusion matrix. Now we represent the confusion matrices for all the algorithms we are interested in. Then from the confusion matrix we would calculate the accuracy of each algorithm and put them in comparison. For the aggregate view we put all the values of confusion matrix in one table. The following table shows different performance measures of different algorithms:

Table 5.1: Accuracy of the algorithms

Algorithm	Accuracy On train set y	Accuracy On test set y
Naïve Bayes	45 %	36%
Logistic Regression	64 %	55%
Support Vector machine	45%	36%
K-NN	97%	100%

The algorithms we applied show different performance measures. K-nearest neighbors provided the highest accuracy (100.00%) and the Naïve Bayes did the least performance in terms of accuracy (36%). For this precise dateset, we acquire the highest accuracy when k=5
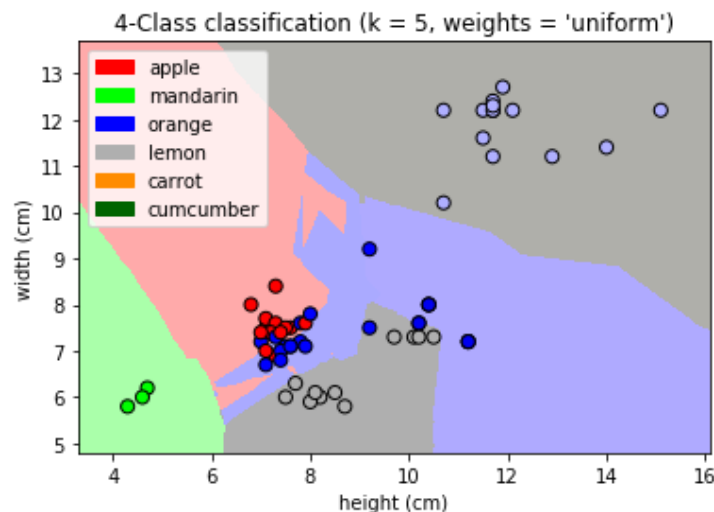
Fig 5.1: Plot of the Decision Boundary of the k-NN Classifier

5.2 Experimenting the Range of Resistance in food

Different foods contain different rage of resistance. Arduino's analog pin (A0) was used to measure the voltage drop across the resistance in the piece of fruit or vegetable. We take some sample food like apple, cucumber, carrot, etc. and measure the resistances to create a range using serial output of Arduino. The ranges are outlined in **Table 5.2**. Rules were established with other features like output voltage, color, width, height and mass of sample food for the analysis.

Table 5.2: Training set for detecting fruits and vegetables from Experiment B.

Fruit	Resistance Range	Color	Mass(g)	Correct 1 Incorrect 0
Apple	Res >140 and Res <300	Red	172	1
Cucumber	Res >400 and Res <700	Green	160	1
Mango	Res >410 and Res <700	Green	40	0
Tomato	Res >510 and Res <710	Red	149	1
Carrot	Res >300 and Res <399	Orange	320	1
Red plum	Res >100 and Res <279	Red	42.5	1
Banana	Res >200 and Res <500	Yellow	200	0

5.3 Developing models

In this phrase fruit is detection by knowledge-based comparison and decision making. Rule base machine learning algorithm pay important rule here to classification and decision making that we already discuss in chapter 2.Using the Rule-base Algorithm we set some rules to detect sample food, some of them are given below in **Fig 5.2** :

```
R1: IF fruit_subtype = mandarin and Color= "yellow" and height = 4.3
THEN Fruit = mandarin
R2: IF fruit_subtype = "unknown" and Color= "green" and (Res >400 and Res
<700)
THEN Fruit= Cucumber
R3: IF fruit_subtype = "braeburn" and width = 7.1 and Color= "red" and
(Res >140 and Res <300)
THEN Fruit= apple
R4: IF fruit_subtype = "estern" and Color= "orange" and (Res >300 and Res
<399)
THEN Fruit =Carrot
R5: IF fruit_subtype = "spanish belsan" and Color= "yellow" or "green" and
height =0.72
THEN Fruit= lemon
```

Fig 5.2: Rule-base Algorithm applied to detect some fruits and vegetables.

The result demonstrated that the rule based system can effectively detect foods like apple, cucumber, carrot, red plum comparing with other machine learning approaches. **Fig. 5.3** Show the expected outcomes of the device**:**

Fig 5.3: Apple and Cucumber detected by the device.

5.4 Train the machine with naturally Occurring forming

The aim of our study is to conduct a sport test to detect artificially added formalin in food items to prohibit its appliance in food. However, formaldehyde is a natural chemical compound which is ubiquitous in our environment and it's also present in many plant and animals for their normal metabolism. It can be also found naturally in different food items including vegetable, fruit, fish and meats [21]. Concentration of natural occurring formalin in food items varies according to conditions and types of food [22]. This natural occurring formalin may interfere in detection of artificially added formalin in food items. Therefore it is important to quantify the artificially added formalin from naturally occurring formalin in food items.

Several researches were conducted on analyzing the naturally occurring formalin in food items [21, 22]. Using Secondary data from The Centre for Food Safety (CFS) we create a model to identify the safe or dangerous level of artificially added formalin in food items. **Fig 5.4** shows the different level [mg/m3] of natural occurring formalin in fruits and vegetables. We train our model with this value for the further comparison with measured value by our detector. According to NIOSH, 1 ppm = 1.23 mg/m3.
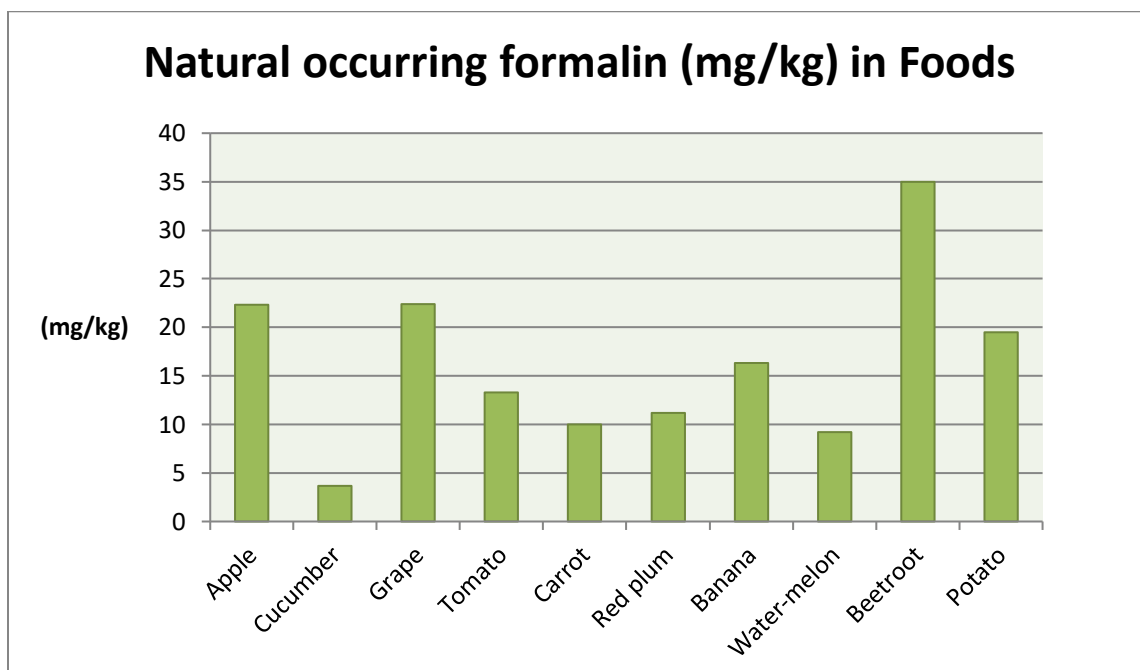
Fig 5.5: Foods Contain Naturally Occurring Formaldehyde

5.5 Experiment on detecting Artificially added Formalin from sample food

Authors applied a small scale experimental analysis to evaluate the capability of the device to detect formalin. We collect some fresh fruits and vegetables from Shwapno Super Shop and measure the concentration of formalin in electric signal output voltage and also in ppm (parts per million). For testing the accuracy of the device sample fruits i.e. carrot placed in 1 liter closed container and added 5 milliliters of formaldehyde with it. Our aim is to get targeted concentration (ppm) of gas greater than natural occurring formalin for carrot (10 mg/m^3). System prediction YES is considered as 1, as well as our targeted concentration is 1. The Experiment was conducted at measured temperature (300K). We leave the container for 3 hours to expose with formaldehyde. We take the readings of targeted gas to calculate the accuracy referred as (5.1) in percentage scale [14]. System follows basic K-NN machine learning principle. We repeat the experiment for different fruits and vegetables with 5 ml of formaldehyde and measure the accuracy.

$$Accuracy\ (\%) = 1 - \frac{|targeted\ concentration - observed\ value|}{targeted\ concentration} * 100\% \qquad (5.1)$$

After connecting all hardwires and applying our developed model on sample food our developed system detect concentration of formalin in sample food and send the result to our developed Android app which is installed in the mobile phone via Bluetooth module. Apps display the result we obtain from the experiments. Experimental set-up was shown in **Fig. 5.5** where test sample as carrot was placed in a 1 liter closed container which is exposed with Formaldehyde and HCHO sensor was placed inside the container.
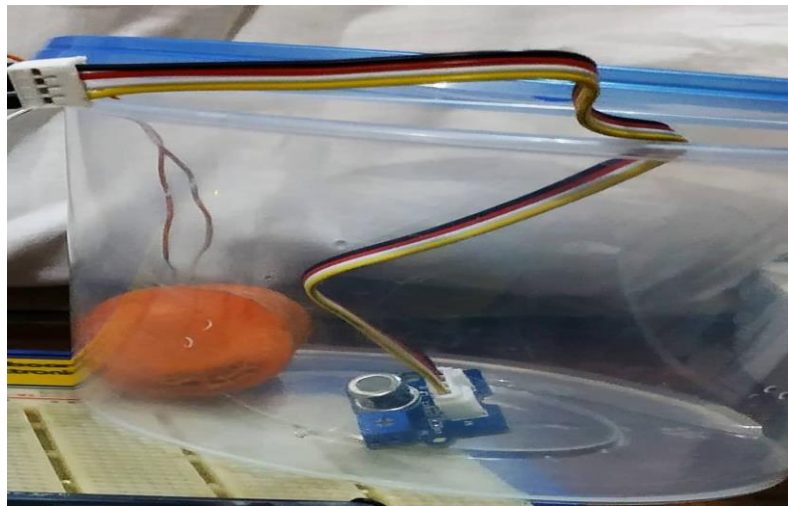
Fig 5.6: Experimental set-up where sensor and carrot is placed inside the container.

First we monitor the output Voltage (Vol) in fresh air. We also measure R0 which is Resistance of the sensor in fresh air using equation (5.2). Vc is Input voltage for the sensor. Then we adjust R1 which is the blue potentiometer attach with HCH0 sensor with a small screwier. This fixed R0 in the number of range 10-100. For our case we fix R1 at 1. Relationship between Vol and R0:

$$R0 = (Vc/Vol - 1) \times R1 \qquad (5.2)$$

To demonstrate the accuracy of the device expected result was to get the target gas. After waiting three hours we again measure the resistibility (Rs) of the sensor. Microcontroller converts the ratio of the resistances (y) to ppm value using equation (5.3). The scale of the measurement is log-log which means in a linear scale the characteristic of the concentration of gas with respect to the resistance ration is exponential [18].

$$x\ =\ 10\hat{}\{[log(y) - b]/m\} \qquad (5.3)$$

Where y = (R0/RS), b = 0.0827, m = -0.4807 [18]

Fig. 5.6 outlines the result of the output Voltage and Resistance in the presence of Formaldehyde and show the observed concentration of formalin in parts per million (ppm).From the result we can see our developed system is able detect formalin. We repeat the experiment with other samples.
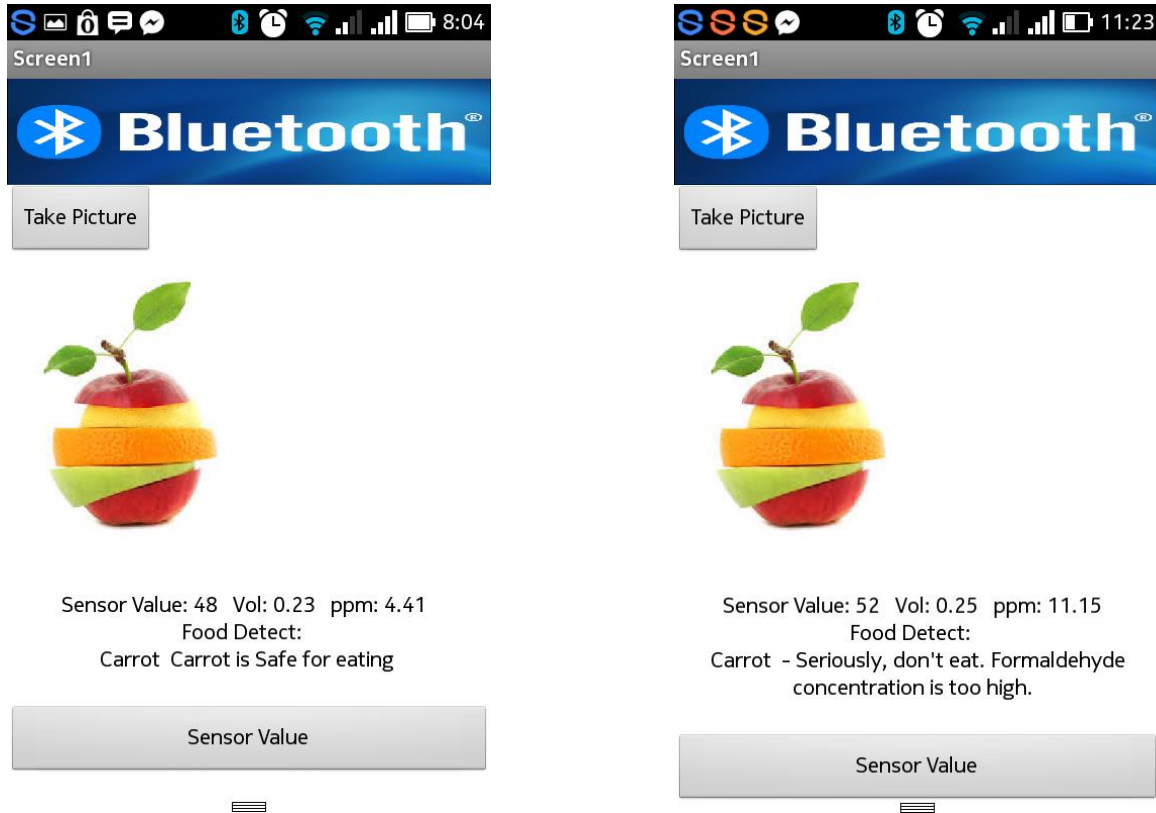
Fig 5.7: Concentration of formalin in ppm (parts per million) Displayed by App.

TABLE 5.3 outlines the concentration of formalin of various fruits and vegetables measured by our developed system. Concentration of formalin detected by our device, if lower than the natural occurring formalin in that particular food item, our system marks it as safe for eating. Otherwise for greater concentration system detect the food as dangerous for consumption. System predicted the presence of artificially added formalin as yes or no. Food with asterisk (*) mark represented the artificially added formalin with them.

Table 5.3: Naturally occurring and artificially added formalin present in foods detected by the system.

Fruit	concentration of formalin (V)	System Prediction	Status
Apple	0.25	**No**	Safe
*Apple	0.32	**Yes**	Dangerous
Cucumber	0.18	**No**	safe

*Cucumber	0.32	**Yes**	Dangerous
Grape	22.4	**No**	Safe
Tomato	0.25	**No**	Safe
*Tomato	0.48	**No(wrong prediction)**	Dangerous
Carrot	0.20	**No**	Safe
*Carrot	0.25	**yes**	Dangerous
Red plum	0.19	**No**	Safe
Banana	0.21	**No**	Safe

Based on the result our system obtains the accuracy of 100% in first iteration. After repeating several times with different samples, the average accuracy of our developed device becomes 99%. **Fig 5.7** shows the Natural and artificially added formalin in different food items in voltage response (V). From the experiment we also gain that system could appropriately detect 1-50 ppm of formalin.
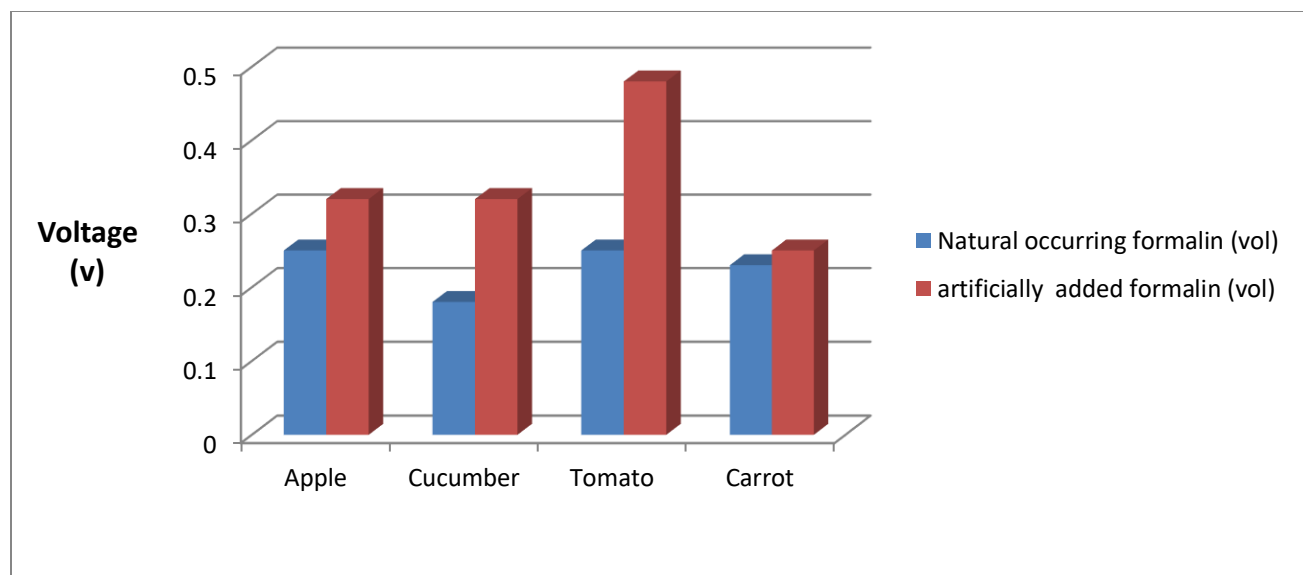
Fig 5.8: shows the Natural and artificially added formalin in voltage response (V).

5.6 Experiment on Raw formalin detection

Our second experiment is to find out the system response for various concentration of raw formalin. At first we measure the temperature and humidity of the environment. We take 14 different concentrations of constant level (1ml) of formalin and measured the response of the device. Three trials were conducted at three measured temperature and humidity for testing the consistencies of the data. The following **Fig. 5.8** shows the graph of the experiment's result of the target gas for different concentration.
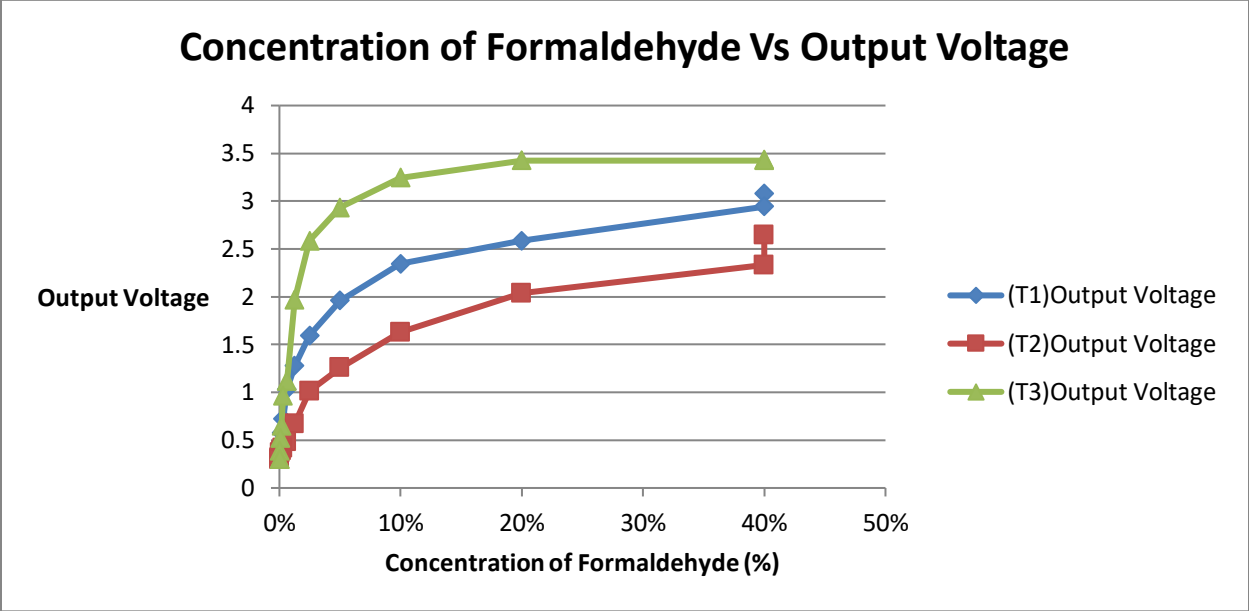
Fig 5.9: shows the concentration (%) of formalin versus system response (V).

The figure indicates that system's response is higher for high concentration of formalin for three different trials. Using polynomial regression which is a non-linear curve fitting algorithm curves were fitted as it provides us a good approximation. Relationship between output voltage and concentration of formalin can be mathematically represented by the polynomial degree of three equations. Using equation (2.2) we can predict the concentration of formalin in a given sample from sensor's output voltage. Here, in **Fig 5.9** X axis represent voltage and Y axis represent the same fixed concentration of formalin for three different trials.
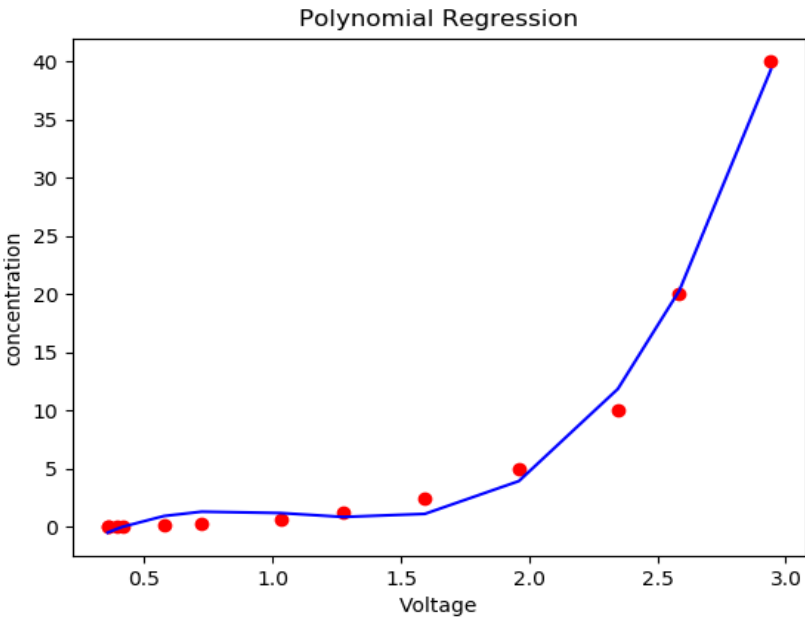
Fig 5.10: shows the Relationship between output voltage and concentration of formalin.

Our Food and Formalin detector device is capable to detect concentration of formalin present in food

items like fruits and vegetables. It is also able to detect various concentration of row formalin in any temperature. This device can be applied for other sample food items also. Accuracy of the device was measured.

There are some drawbacks of our developed system. When concentration of formalin is more than 50 ppm system fail to detect the accurate targeted concentration. Better sensor is required for better performance. Detecting fruits by only using resistance give sometimes wrong result and not appropriate for all fruits. Using image processing machine learning approaches this problem can be solved.

Chapter 6

6.1 Overall Conclusion

Formaldehyde is one of the commonest molecules in nature. A solution of 35 to 40 percent of water in formaldehyde is called formalin. Some foods and fruits are known to contain naturally occurring formaldehyde in some specific ranges. To keep fruits, some raw foods like fishes and meats for a long time, prevent from decomposing people are using a large amount of formalin which reflects a big impendence of our health issues. But Bangladesh Food Safety Authority claimed that there is no scope to adulterate vegetable and fruits by mixing formalin, a health hazardous chemical which is not fully true.

The rapid development of artificial senses technologies in the present world is increasing day by day. By using the machine learning algorithm is avowedly underpin and also increase the capacity of the system. , which is kind capable of differentiating between natural and artificial formalin at all temperatures in PPM value and all within a portable, lavish, accurate and user-friendly package. Our implemented device can detect the fruit and makes use of Grove VOC HCHO gas sensor which can feasibly detect given sample fruit and concentration of formaldehyde present in comestible products. The human sensing system cannot detect the amount of formaldehyde or even the presence of it in any comestible products. But our implemented device can replicate the human sensing system to detect this harmful element in comestible products. Our ultimate goal is to help the general community regarding health concerns of them. To ensure healthy eating we involve ourselves in this implementation to make our country a better place for living.

6.2 Discussion

By naturally fruits and vegetables contain different level of formalin in them. It is necessary to develop a system that behaves in such a way that can detect formalin based on their natural occurring formalin and will not interfere with artificially added formalin. It is therefore obvious to detect fruit or vegetable dynamically by the system. Considering this we develop a system which first detect a fruit or vegetable then compare the measured concentration present in them with their natural occurring formalin. Different machine learning algorithms where applied for detecting food items like fruits and vegetables where K-NN algorithm gives us the best result for the particular dataset. Different Rules were developed and applied on sample food for detection. Our first Experiment shows the accuracy of the device we developed. Targeted amount is measured in ppm value, finally comparing with naturally occurring

formalin in food items system classify either artificially formalin is added or not. Based on the classification our system make decision either it is safe or dangerous. Integrated Android application added extra feature with our developed system and make it more interesting. We conduct our second Experiment for analyzing the sensing capacity of the sensor and monitoring its behavior in different concentration of formalin. For low concentration linearity is obtained and for higher concentration non-linear behavior is found. We fit the curve using Polynomial regression.

6.3 Future work

Ours proposes to develop a smartphone-based system that offers a simple & rapid route to detect sample foods and sense harmful element formaldehyde in our daily consumed. The further development of our device will be to ensure the following properties which are to more utility, efficiency, and effectiveness in formaldehyde sensing operations in our daily consumed foods. The foods can be detected by using the application of image processing also. It can explore the possibility of implementing a neural network rather than using the different machine learning algorithms. Also can explore the possibility of commercially developing a photonic or MEMS-based formalin sensor and using it in our designed system [4] [21]. Java and python based program will be developed for further improvements of the Android application which will be able to show all past detection results with the concentration of formalin. It could be possible to Connect to the cloud or server may also be considered where all the previous data will be stored. Design the hardware device outer look can be better which will also easily portable to anywhere.

References

[1] M. Rahman, M. Sultan, M. Rahman and M. Rashid, "Food Adulteration: a serious public health concern in Bangladesh", *Bangladesh Pharmaceutical Journal*, vol. 18, no. 1, p. 1, Dec. 2015.

[2] A. Mamun, A. Al Mamun, A. Begum, M. Shafiul Alam and M. Asaduzzaman, *Microcontroller-based Formalin Detection: Hardware Design and Android App Development*. Dhaka, Bangladesh: Department of Electrical and Electronic Engineering, University of Dhaka, 2017.

[3]C. Cortez and J. Santos, "Evaluation of the Developed Formaldehyde Detector", *International Journal of Environmental Science and Development*, vol. 7, no. 6, pp. 449-452, 2016.

[4]C. Cortez, F. Christian Bato and T. Jake G. Bautista, "Development of Formaldehyde Detector", *International Journal of Information and Electronics Engineering*, vol. 5, no. 5, p. 385, 2015.

[5] S. Sayan Saha, M. Siraj and W. Bin Habib, "FoodAlytics: A formalin detection system incorporating a supervised learning approach", in *In Humanitarian Technology Conference (R10-HTC), 2017 IEEE Region 10*, Dhaka, Bangladesh, 2018, pp. 26-29.

[6]A. Wilson, "Diverse Applications of Electronic-Nose Technologies in Agriculture and Forestry", *Sensors*, vol. 13, no. 2, pp. 2295-2348, 2013.

[7] P. R. Chaudhari and A. L. Judal, "Sensory Instruments and Its Application in Dairy and Food Industry", *In International Conference of Advance Research and Innovation, ICARI'15*, India, 2015.

[8] E. A. Baldwin, J. Bai, A. Plotto and S. Dea, "Electronic noses and tongues: Applications for the food and pharmaceutical industries", *Sensors*, vol. 11, no. 5, pp. 4744-66, 2011.

[9] Bandal, Amol, and M. Thirugnanam. "Quality Measurements of Fruits and Vegetables Using Sensor Network", *In Proceedings of the 3rd International Symposium on Big Data and Cloud Computing Challenges,* ISBCC'6, pp. 121-130. Feb. 2016.

[10] M. Adlim, Z. Fitri, Y. Amri, M. Sari and S. Mahya, "Preparations of chemical sensors for simple formalin detection in contaminated food (A model for final project works for pre-service teachers in chemistry subject)', *In Proceedings of The Annual International Conference, Syiah Kuala University-Life Sciences & Engineering Chapter*, vol. 1, no. 2, Dec. 2011.

[11]A. Dar, J. Amara, U. Shafique, J. Anwar, and A. Naseer. "A simple spot test quantification method to determine formaldehyde in aqueous samples." *Journal of Saudi chemical society*, vol. 20, pp. S352-56, 2016.

[12]S. Ngamchana and W. Surareungchai, "Sub-millimolar determination of formalin by pulsed amperometric detection", *Analytica Chimica Acta*, vol. 510, no. 2, pp. 195-201, 2004.

[13] F. Bianchi, M. Careri, M. Musci, and A. Mangia. "Fish and food safety: Determination of formaldehyde in 12 fish species by SPME extraction and GC–MS analysis," *Food Chemistry*, vol. 100, no. 3 pp. 1049-53, 2007.

[14]O. Lau and S. Wong, "Determination of plasticisers in food by gas chromatography-mass spectrometry with ion-trap mass detection", *Journal of Chromatography A*, vol. 737, no. 2, pp. 338-342, 1996.

[15] F. Bianchi, M. Careri, M. Musci, and A. Mangia. "Fish and food safety: Determination of formaldehyde in 12 fish species by SPME extraction and GC–MS analysis," *Food Chemistry*, vol. 100, no. 3 pp. 1049-53, 2007.

[16]W. Zeng, T. Liu, Z. Wang, S. Tsukimoto, M. Saito and Y. Ikuhara, "Selective Detection of Formaldehyde Gas Using a Cd-Doped TiO2-SnO2 Sensor", *Sensors*, vol. 9, no. 11, pp. 9029-9038, 2009.

[17]R. Uddin, M. Wahid, T. Jasmeen, N. Huda and K. Sutradhar, "Detection of Formalin in Fish Samples Collected from Dhaka City, Bangladesh", *Stamford Journal of Pharmaceutical Sciences*, vol. 4, no. 1, 2011.

[18] "Grove - HCHO Sensor - Seeed Wiki", wiki.seed.cc, 2017. [Online].Available: http://wiki.seeed.cc/Grove-HCHO_Sensor. (Accessed on June 9, 2018)

[19] Galvanic cell, Available in: https://en.wikipedia.org/wiki/Galvanic_cell [online] (Accessed on August 9, 2018)

[20] D. Walter, M. Sherman, "Learning MIT App Inventor a hands-on guide to building your own android apps",U S. River, NJ: A.Wesley, 2015.

[21]P. Wahed, M. Razzaq, S. Dharmapuri and M. Corrales, "Determination of formaldehyde in food and feed by an in-house validated HPLC method", *Food Chemistry*, vol. 202, pp. 476-483, 2016.

[22]F. Nowshad, M. Islam and M. Khan, "Concentration and formation behavior of naturally occurring formaldehyde in foods", *Agriculture & Food Security*, vol. 7, no. 1, p. 17, 2018.

[23] W. C. Seng and S. H. Mirisaee, "A new method for fruits recognition system", *In International Conference on Electrical Engineering and Informatics, ICEEI'09*, vol. 1, pp. 130-134, Aug. 2009.

Appendix A

List of Acronyms

IoT	Internet of things
HCHO	Formaldehyde
KNN	K-nearest neighbors
SVM	Support Vector Machines
VOC	Volatile Organic Compounds
PPM	Parts Per Million
STEL	Short-term exposure limit
NIOSH	National Institute for Occupational Safety and Health
OSHA	Occupational Safety and Health Administration
PEL	Permissible Exposure Limit
E-nose	Electronic nose
LCD	Liquid Crystal Display
SPME	Solid-phase Microextraction
PAD	Pulsed Amperometric Detection
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
GND	Ground
VIN	Input Voltage
V	Voltage
KB	Kilobyte
SRAM	Static Random-access Memory
MIT	Massachusetts Institute of Technology
App	Application
SD	Secure Digital
Mg/m3	Milligram Per Cubic Meter
CFS	Centre for Food Safety
Rs	Resistibility

Appendix B

Essential Source Code of This Project

/**************** Arduino Programing **************/

We create different ranges based on the measured resistances. We write a code for detecting apple, carrot, cucumber and red plum. The code is:

```
const int frootSense = 0;
int frootResistance, high = 0, low = 1023;
int frootDetect;
#include <LiquidCrystal.h>
LiquidCrystal lcd(12,11,5,4,3,2);
void setup(){
Serial.begin(9600);
lcd.begin(16, 2);
lcd.clear();
}
void loop()
{
lcd.clear();
lcd.setCursor(0,0);
lcd.print("Food Detect:");
lcd.setCursor(0,1);
frootResistance = analogRead(frootSense);
Serial.print("Resistance:");
Serial.print(frootResistance);
Serial.print("\n");
if (frootResistance>400 & frootResistance<700){
Serial.print("Cucumber \n");
lcd.print("Cucumber");      }
else if(frootResistance>140 & frootResistance<300){
Serial.print("Apple");
lcd.print("Apple");     }
else if(frootResistance>300 & frootResistance<380){
Serial.println("Food Detect:  ");
Serial.print("Carrot ");
lcd.print("Carrot");
else if(frootResistance> 100 & frootResistance< 279){
Serial.println("Food Detect:  ");
Serial.print("Red Plum ");
lcd.print("Red Plum ");
else {
Serial.print("No Food \n");
lcd.print("No Food");
}
delay(1000);
}
```

After connecting all devices with Arduino, we plug the assemble Arduino into our laptop and load up our developed code to Arduino IDE.

Following code need to upload to the board:

```
// demo of Grove - HCHO Sensor
#include <Wire.h>
#include <math.h>
#define Vc 4.95

void setup()
{
    Serial.begin(9600);
}

void loop()
{
    int sensorValue=analogRead(A0);

    float Vol=sensorValue*4.95/1023;
    Serial.print("Vol = ");
    Serial.println(Vol);

    float R0=(Vc/Vol-1);
    Serial.print("R0 = ");
    Serial.println(R0);
    delay(500);
}
```

After waiting two hours we again measure the resistibility (Rs) of the sensor.) . Again upload the following code and get the concentration of formalin in ppm.

```
// demo of Grove - HCHO Sensor

#include <Wire.h>
#include <math.h>
#define Vc 4.95
//the number of R0 you detected just now
#define R0 23.95

void setup()
{
    Serial.begin(9600);
}

void loop()

{
    float Vol=sensorValue*4.95/1023;
    Serial.print("Vol = ");
    Serial.println(Vol);

    float RS=(Vc/Vol-1);
```

```
    Serial.print("RS = ");
    Serial.println(RS);

    double ppm = pow(10.0,((log10(Rs/R0)-0.0827)/(-0.4807)));
    Serial.print("HCHO ppm = ");
    Serial.print(ppm);

//Get ppm value in linear scale according to the ratio value
//Convert ppm value to log scale

    delay(500);
}
```

Complete code of full developed System:

```
// demo of Grove - HCHO Sensor
#include <math.h>
//the number of R0 you detected just now
#define R0 20.00
#define Vc 4.95 // Constant voltage from pin

const int frootSense = 0;
int frootResistance, high = 0, low = 1023;
int frootDetect;
//#include <LiquidCrystal.h>              //remove this if u don't have
LCD, anyways it won't affect.
//LiquidCrystal lcd(12,11,5,4,3,2);        //remove this if u don't have
LCD, anyways it won't affect.
//#include <Wire.h>
//#include <LiquidCrystal_I2C.h>
//LiquidCrystal_I2C lcd(0x27, 2, 1, 0, 4, 5, 6, 7, 3, POSITIVE);  // Set the
LCD I2C address
float Vol, ppm;
int sensorValue;
int redLed = 13; // Red Led for indication
int c = 0;

void setup(){
Serial.begin(9600);
//lcd.begin(16, 2);                       //remove this if u don't have LCD,
anyways it won't affect.
//lcd.clear();                            //remove this if u don't have
LCD, anyways it won't affect.
pinMode(redLed,OUTPUT);
digitalWrite(redLed, LOW);
}


void loop()
{

//lcd.clear();                    //remove this if u don't have LCD, anyways
it won't affect.
```

```
//lcd.setCursor(0,0);          //remove this if u don't have LCD, anyways it
won't affect.
//lcd.print("Food Detect:");            //remove this if u don't have LCD,
anyways it won't affect.
//lcd.setCursor(0,1);                    //remove this if u don't have LCD, anyways
it won't affect.
frootResistance = analogRead(A1);
sensorValue=analogRead(A0);
Vol=sensorValue*4.95/1023;

float Rs=(Vc/Vol-1);
//Serial.print("Rs = ");
//Serial.println(Rs);
double ppm=pow(10.0,((log10(Rs/R0)-0.0827)/(-0.4807))); //Get ppm value in
linear scale according to the the ratio value
                                                    //Convert ppm value to
log scale
//Serial.print("Resistance: ");
//Serial.print(frootResistance);
Serial.print("\n");

Serial.print("Sensor Value: ");
    Serial.print(sensorValue);
    Serial.print("   Vol: ");
    Serial.print(Vol);
    Serial.print("   ppm: ");
    Serial.print(ppm);
    Serial.println(" ");

if (frootResistance>400 & frootResistance<700){
Serial.println("Food Detect:   ");
Serial.print("Cucumber \n");
//lcd.print("Cucumber");      //remove this if u don't have LCD, anyways it
won't affect.

 if (ppm*1.23 >3.7) {
       Serial.println(" - Seriously, don't eat.Formaldehyde concentration is
too high. ");
      digitalWrite(redLed, HIGH);
     }
     else {
     Serial.println(" Cucumber is Safe for eating");
     digitalWrite(redLed, LOW);

     }
}
else if(frootResistance>140 & frootResistance<300){
Serial.println("Food Detect:   ");
Serial.print("Apple ");
//lcd.print("Apple");     //remove this if u don't have LCD, anyways it wont
affect.
 if (ppm*1.23 >22.3) {
       Serial.println(" - Seriously, don't eat.Formaldehyde concentration is
too high. ");
```

```arduino
        digitalWrite(redLed, HIGH);
      }
else {
      Serial.println(" Apple is Safe for eating");
      digitalWrite(redLed, LOW);

      }
}
else if(frootResistance>300 & frootResistance<380){
Serial.println("Food Detect:   ");
Serial.print("Carrot ");
//lcd.print("Carrot");     //remove this if u don't have LCD, anyways it won't
affect.
 if (ppm*1.23 >10) {
        Serial.println(" - Seriously, don't eat. Formaldehyde concentration
is too high. ");
        digitalWrite(redLed, HIGH);
      }
      else {
      Serial.println(" Carrot is Safe for eating");
      digitalWrite(redLed, LOW);

      }
}
else {
Serial.print("No Food Detected \n");
//lcd.print("No Food");     //remove this if u don't have LCD, anyways it
won't affect.
}
delay(1000);
}
```

/**************** **Python Code for** **************/
Polynomial regression for raw HCHO curve fitting

```python
import pandas as pd
df=pd.read_csv("DataSet.csv")
x=df.iloc[:,0:3].values
y=df.iloc[:,3].values
from sklearn.preprocessing import PolynomialFeatures
poly=PolynomialFeatures(degree=3)
poly_x=poly.fit_transform(x)
from sklearn.linear_model import LinearRegression
regressor=LinearRegression()
regressor.fit(poly_x,y)
import matplotlib.pyplot as plt
plt.scatter(x,y,color='red')
plt.plot(x,regressor.predict(poly.fit_transform(x)),color='blue')
plt.show()
```