International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Detection of Formaldehyde on Fruits and Vegetables using Machine Learning

[1]Mahanth S, [2]Hemalatha G, [3]Swathi H S, [4] Mrs. Madhu G

Department of Computer Science and Engineering, Jyothy Institute of Technology, Bengaluru, India

Abstract:

The food we take should be pure and unalloyed. In our paper we designed an IoT based hcho detection model. We used supervised machine - learning approaches to sort the data sets picked up by the HCHO gas detector connected with Arduino. Supervised machine literacy algorithms are incorporated into the system to directly prognosticate the correct attention of formalin. It's suitable to rightly classify the difference between instinctively added and naturally formed formalin. For the food products to remain rich in texture and appearance, a conception of contamination are extensively rehearsed. Substantially the contamination in fruits and vegetables are caused using a dangerous chemical substance called Formalin. Formalin is a colourless, waterless result of formaldehyde to save natural samples. This chemical is extremely lethal and just a few milli litters of formalin containing 37% of formaldehyde can easily kill an adult. Formalin is used as a preservative by the distributers to ameliorate the appearance of fruits and vegetables and to sustain for longer shelf life.

INTRODUCTION

This paper depicts detection of HCHO by using algorithms based on machine learning approaches. The process of manual detection method doesn't provide accurate levels of hcho. Our model provides a reliable formaldehyde detection technique based on IoT and M.L techniques. Detection of food contamination with the help using machine learning is one of the simple techniques which produces good accurate results by measuring the resistance difference from the fumes radiated by the food. Without the use of predefine model the detection of naturally formed formalin will be highly inaccurate. So the system will trace the artificially added formaldehyde as a preservative. The use of machine learning algorithms such as Logistic regression, SVM, K-NN Classifier are applied to the testing dataset to build a predictive model.

LITERATURE SURVEY

Currently a handheld formalin sensor called ZX300 used by mobile labs which have a narrow attention range of 1-30 ppm only and it is physically importable and impractical for everyday use and still provides frequent incorrect readings and costs a whopping $1175. Scientific and Industrial Researches have come up with a solution for 3USD model which still provides only a basic result whether formalin is present or not on the surface of the food. They designed system within the budget of $26 and integrated with their mobile application. Still their system designed using a new detector and a different package also with supervised machine learning algorithms still produced similar results.

In some cases, this harmful chemical is used for preserving specimens from going bad, but now being used to preserve edible items. The excessive levels of formaldehyde in air cause unfavourable outcomes like watery eyes, burning perceptions in one's eyes, nose and skin irritations. The statement from NIOSH states that 20 ppm of formalin is very lethal for humans. Due to lack of tight regulations and preventive measures, excessive concentration of hcho has been discovered in most fruits in the local markets. Due to these reasons, fruit identification is necessary for separating natural occurring formalin from artificially added formalin.

These smart device with trained model could be a better solution for regular people using these devices. Arduino based food adulteration detection is an effective method. The detector can detect the presence of formaldehyde in the surrounding air. These applications can be applied in food and dairy industry. Some of the paper proposes a microcontroller or microprocessor based formalin detection model along with an mobile application to produce the results of sample food on demand.

A related study that proposed on detecting the presence of formaldehyde in air in ppm unit and adding SD technology. A simple testing methods to detect the presence of formaldehyde in aqueous samples were also proposed with varying results. A test for detecting HCHO in sea foods and poultry animals were conducted. To the surprise the technique of use of HCHO is used on meat too. Problem is that the formalin naturally present in foods, similar as meat, poultry, fish, fruits and vegetables, mushrooms are also similar. It could intrude with instinctively added formalin when they're going to be detected. Arduino grounded food discovery is a simple system of detecting sample food by measuring the change in the concentration on them. The sensor can pick up the traces of the formalin by reading the fumes radiated in presence of formalin in the surrounding air.

Electronic- nose technology for seeing formalin studied in. Their operations applied in food and dairy assiduity considered in, pharmaceutical diligence and food operation using - noses speeches are also reviewed in the paper. In paper their approach was to descry Formalin incorporating supervised literacy. An affiliated study proposed on detecting the presence of formaldehyde in air using s d technology. There were many more methodology proposed for formaldehyde discovery, including P.A.D system, S.P.M.E and gas spectroscopy gas chromatography, mass spectrometry by using SnO2 detector. The use of various machine learning algorithms was used to classify the features of the fruits and vegetables on their features. The grove hcho voc detector connected with were used to prize the attention. Latterly, rule- grounded bracket and polynomial retrogression algorithms were applied to prognosticate the attention of the formalin from the uprooted affair voltage.

PROPOSED SOLUTION

We used CH2O sensor also known as grove HCHO VOC sensor and Arduino uno and i2c 0.9in oled display for the hardware. Compare data sets from food safety board, prepare a dataset for each food item then compare the results with our recorded data from the gas sensor. Use of Support Vector Machine [SVM], K-NN Classifier and few other regression algorithms to compare the taken dataset to build a predictive model.

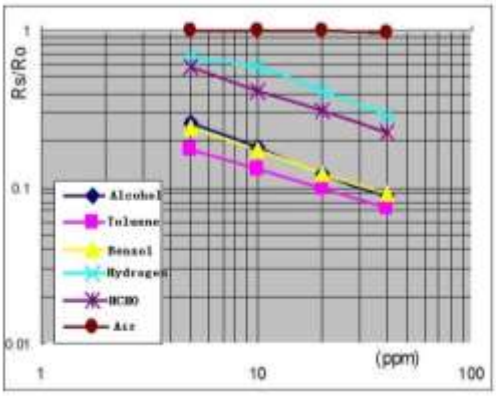

Figure 1. From the sensitivity curve we could know that the detection range is 1-50ppm.

The grove hcho sensor is able to detect the presence of formaldehyde on the sample food by measuring the different range of concentration radiated from the surface of the fruit. The analog data from the sensor is then sent to the Arduino uno. Arduino reads the sensor value from one of the analog pins A0 on the board. In the Arduino a series of command lines are used for the calibration of the grove hcho sensor. The process of calibration improves the accuracy range of the recorded formaldehyde values. The formula for the calibration is mentioned. Another set of code is instructed for the detection on hcho vapours in ppm. The formula is further explained below. Now the recorded values from the Arduino IDE is entered into the software. The software is comprised of supervised machine learning algorithms, which then process the recorded values from the hardware. The ml model consists of a prediction model which then compares the test values with the actual values and produces an accurate level of concentration present on the food sample. The hardware model is an isolated system designed to sense the accurate levels of formaldehyde on the surface of the food. This helps in isolating the formaldehyde gas from other gases present in the atmosphere. The use of separate dataset for each food helps in the accuracy of the prediction model. The entire setup costs about 2500 INR. This model is partially portable and cost effective to manufacture. The entire model is designed to be fool proof, specifically tuned to detect the formaldehyde vapours at the range from 1 – 50 ppm. Which is the chemical range of formaldehyde. The model can safely differentiate the composition of naturally occurring formaldehyde from the adulterated levels. The prediction model uses the test datasets to classify the values recorded and give an accurate detected value of formaldehyde. The working of our model is represented below by the data flow diagram. The level 2 data flow diagram shows the working of the model.

IMPLEMENTATION

Our model uses KNN, SVM Linear, SVM Gaussian and Logistical regression algorithm and produces a prediction model based on the provided datasets. Each fruit or vegetables are given its own datasets. These data sets consist of three parameters i.e., fruit_id, ppm value, state. The grove HCHO VOC gas sensor is used to detect the volume of the formaldehyde present on the surface of the food. These recorded ppm values of formaldehyde is read by the analog pin A0 by the Arduino uno. The Arduino uno consists of two sets of codes, One of the code is used to calibrated the sensor for increased accuracy. The formula for the accuracy model is represented by:

```
Vf = sensorValue * 4.95 / 1023; // Voltage through pin when in presence of VOCs (formaldehyde).
```

The second part of the model is used to calculate the concentration of the formaldehyde. This formula is represented by the formula:

```
con = 212.32619543773774 * exp( -( 7.653015806367451 * (4.95 - V) * Vstart ) / ( V * (4.95 - Vstart) ) );
```

The Arduino IDE produces the concentration of formaldehyde in ppm units. This is displayed on the serial monitor.

These ppm values are then entered to our software where the supervised algorithms are used with the datasets for particular food and with the prediction model the results are displayed on the screen.

The prediction model produces accurate results of the detected content of formaldehyde present on the surface of the food. The algorithm is capable of differentiating between naturally present formaldehyde and adulterated values. And thus, the prediction model is completed.

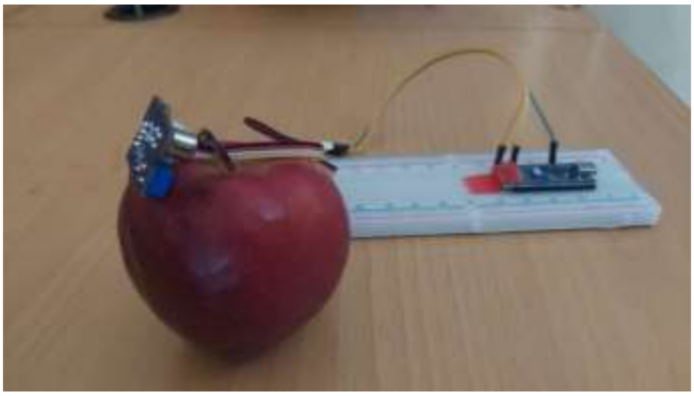

Figure 2. Detection of HCHO on apple.

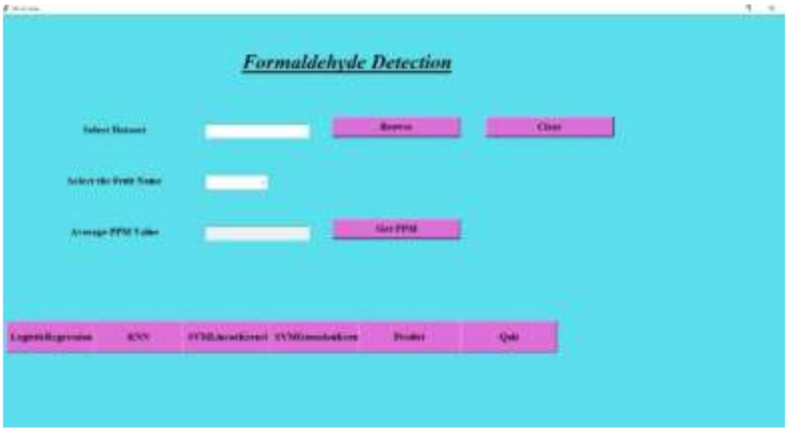

Figure 3. Prediction model interface.

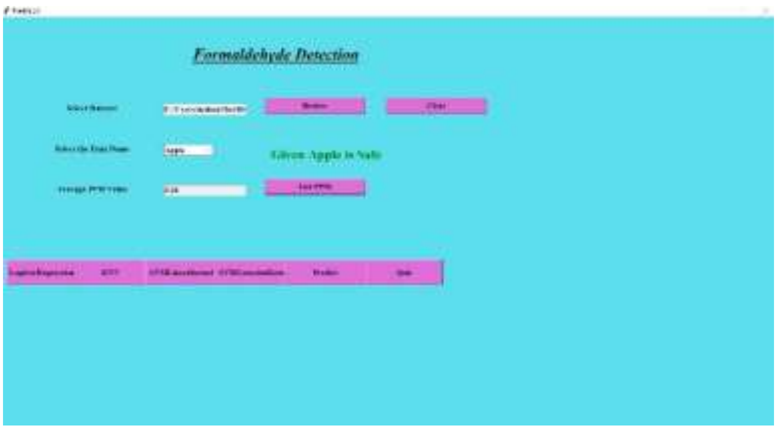

Figure 4. Prediction model for apple(safe).

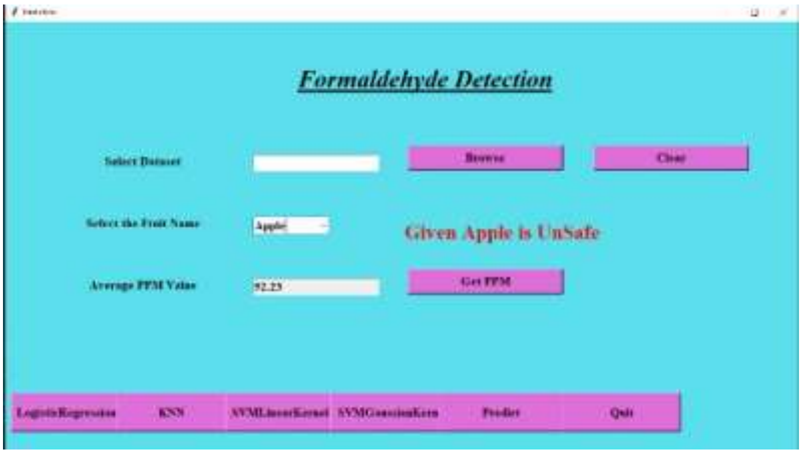

Figure 5. Prediction model for apple(unsafe).

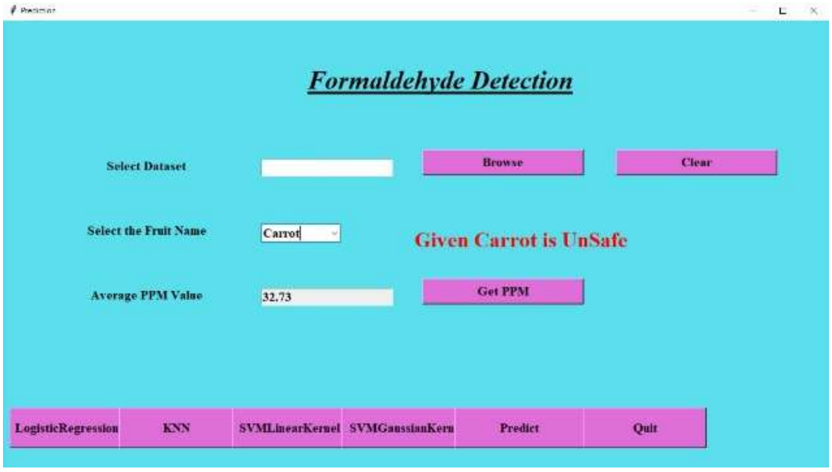

Figure 6. Prediction model for carrot(unsafe).

CONCLUSION

Adulteration is posing much more health risks as the use of more and more techniques are practiced to preserve the shelf life of produces. This is the reason our project is out to provide simple but effective solution in detecting adulteration. Our model ensures the cost effectiveness and reliability for the proposed solution. With the rapid growth of technology and improvement for quality of living it is much easier to provide solutions for the problems.

We aim to improve and iterate the proposed solution further incorporating technologies making it a worthwhile model.

References

[1]. https://en.wikipedia.org/wiki/Arduino_Nan

[2]. https://docs.arduino.cc/

[3]. https://wiki.seeedstudio.com/Grove-HCHO_Sensor/

[4]. https://files.seeedstudio.com/wiki/Grove-HCHO_Sensor/res/Wsp2110-1-.pdf

[5]. https://www.spyder-ide.org/

[6]. https://www.arduino.cc/en/donate/