

Intelligent System for Coffee Bean Roast Level Classification Using Electronic Nose and Artificial Neural Network

Justam¹, Sharto Batti², Erlita³, Luqman Fanani mz⁴, Milda Sibiti⁵
^{1,2,3,4,5}Universitas Mega Buana Palopo

ABSTRACT

Roasted coffee beans release gaseous compounds, primarily carbon dioxide (CO₂). Coffee comes in various types, including Robusta, Arabica, Excelsa, Tubruk, Latte, and Luwak. However, this study focuses only on Robusta and Arabica coffee. Each roast level of coffee beans has its own distinct aroma, necessitating a fast and accurate method to differentiate them. Therefore, this research aims to classify coffee bean roast levels based on their aroma profiles. The dataset for classifying Robusta and Arabica coffee roast levels was obtained from data collection using a miniature Electronic Nose system. A total of 900 data samples were collected, with 720 samples used for training and 180 samples used for testing. This study employs an Artificial Neural Network (ANN) with an Electronic Nose for classification. The True Positive (TP) results obtained for each coffee roast level are 44 for Light roast, 55 for Medium roast, and 57 for Dark roast. The classification accuracy achieved in determining the roast level of coffee beans is 86%.



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Keywords:

E-Nose,
Artificial Neural Network,
Coffee,
Microcontroller

Corresponding Author:

Justam

Universitas Mega Buana Palopo

Email: justam@umegabuana.ac.id

INTRODUCTION

Coffee is one of the most economically valuable agricultural commodities, playing a crucial role as a source of national revenue. The quality of coffee beans highly depends on proper post-harvest handling at every stage of processing. Coffee that is harvested when fully ripe ensures high quality, while prematurely harvested coffee results in inferior aroma and flavor due to incomplete maturation [1], [2].

The aroma of green coffee beans changes significantly after roasting, a complex process involving various chemical transformations [3]. Roasting is a crucial stage in coffee processing that induces chemical, physical, and structural changes, influenced by multiple factors such as plant variety, growing region, processing method, grinding size, and brewing method. During heating, the physical and chemical changes in coffee beans are highly intricate and follow specific guidelines throughout the roasting process.

Roasting is the process of developing the flavor and aroma of coffee beans. If the beans exhibit uniformity in size, specific gravity, texture, moisture content, and chemical structure, the roasting process becomes relatively easier to control. However, coffee beans vary significantly in these properties, making roasting an art that requires skill and experience to meet consumer preferences. Post-harvest handling and coffee bean processing must consider various aspects to maintain coffee quality, with roasting being one of the most critical factors. Coffee quality can be improved if roasting is performed at the appropriate temperature and duration, ensuring optimal moisture content and acidity levels as per SNI 01-2983-1992 (Indonesian National Standard, 1992) and SNI 01-3542-2004 (Indonesian National Standard, 2004) [4].

During the roasting process, coffee beans undergo temperature changes for different durations, depending on the desired final product characteristics. For instance, the Italian roasting style maintains temperatures between 200°C and 240°C [5], [6]. Several criteria are commonly used to evaluate coffee bean quality and roasting degree, including moisture content, weight loss, density, color, and taste. Among these, taste plays the most significant role in influencing consumer preferences and determining the final product quality [7].

Artificial Neural Networks (ANNs) offer various network architectures and training methods. The choice of network architecture and training enables ANNs to learn and analyze historical data patterns more accurately, minimizing errors and improving output precision. Backpropagation architecture is one of the ANN architectures that can effectively learn and analyze past data patterns, leading to more accurate predictions with minimal error.

METHODS

In this study, the system implements the Artificial Neural Network (ANN) algorithm to classify roasted coffee beans. The stages of the process are illustrated in Figure 1.

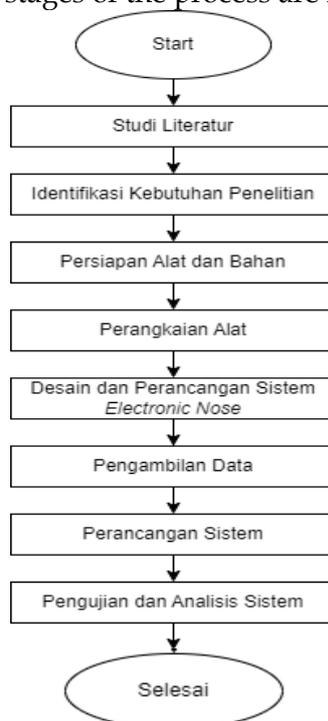


Figure 1. Research Workflow

The stages shown in Figure 1 can be broadly explained as follows:

1. Literature Review – This is the initial phase of the research. It involves gathering related studies on the use and development of an Electronic Nose with an Artificial

Neural Network (ANN). This stage also includes collecting references from national and international journals, books, articles, and research reports.

2. Research Requirement Identification – In this phase, various research requirements are determined to support the assembly and design of the Electronic Nose system.
3. Equipment and Material Preparation – The necessary tools and materials for the research are prepared, including Arduino Mega 2560, sensor modules MQ-2, MQ-5, MQ-135, and other essential components.
4. Electronic Nose Assembly – Once all components are available, the Electronic Nose system is assembled, including the integration of sensors with the microcontroller.
5. System Design & Programming – This step involves creating a flowchart to outline the workflow of the Electronic Nose system. The sensor data acquisition program is then developed using the microcontroller.
6. Data Collection – Data acquisition is performed using the assembled Electronic Nose system, capturing sensor readings for classification purposes.
7. System Development – A flowchart is created to guide the classification system for coffee beans. The system is built following the designed flowchart.
8. System Testing – Once all research requirements are met, the system is tested. This includes evaluating components like Arduino Mega 2560 as the microcontroller and analyzing sensor data to classify coffee beans based on their roasting level.

This study will use research instruments, as presented in Tables 1 and 2.

Table 1. *Software*

Software	Version
Windows 11	64 Bit
Arduino ID	1.8.19
Fritzing	0.9.3
Google Colab	-

Table 2. *Hardware*

Hardware
Laptop Lenovo V14 AMD Ryzen 5 5500U with Radeon Graphics RAM 16GB
Arduino mega 2560
Sensor MQ-2
Sensor MQ-5
Sensor MQ-135

The system developed in this study utilizes a total of 900 data samples, which are distributed as follows: 150 samples for light roast Robusta, 150 samples for medium roast Robusta, 150 samples for dark roast Robusta, 150 samples for light roast Arabica, and 150 samples for dark roast Arabica. The system design for this research is illustrated in Figure 2.

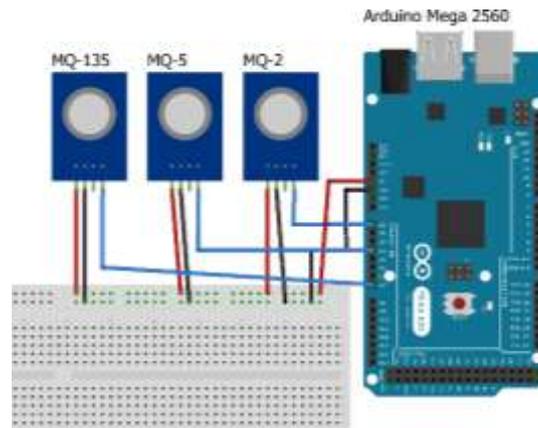


Figure 2. System Design

RESULT AND DISCUSSION

After completing the design and component assembly, the next step is the integration of system components into a unified Electronic Nose system. The assembled system can be seen in Figure 15



Figure 3. Purwarupa Electronic Nose

Figure 3 shows the prototype of the Electronic Nose, which serves as a container or housing for the sensor array used in the system, ensuring proper functionality.

Table 3. Sensor Reading Results for Light Roast Coffee

No.	MQ2	MQ5	MQ135
1	413	193	157
2	414	192	143
3	415	192	145
..
148	429	194	126
149	432	199	143
150	424	192	177

Table 5 presents the gas sensor reading results for light roast coffee. The data collection was conducted on January 15, 2023.

Table 4. Sensor Reading Results for Medium Roast Coffee

No.	MQ2	MQ5	MQ135
1	490	174	256
2	490	174	464
3	490	174	279
..
148	479	163	529
149	479	162	326
150	480	165	309

Table 4 presents the gas sensor reading results for medium roast coffee. The data collection was conducted on January 15, 2023.

Table 5. Sensor Reading Results for Dark Roast Coffee

No.	MQ2	MQ5	MQ135
1	529	181	243
2	533	186	247
3	536	190	246
..
148	464	166	337
149	467	174	345
150	482	190	150

Table 5 presents the gas sensor reading results for dark roast coffee. The data collection was conducted on January 15, 2023. After the data collection process, data labeling is performed based on the roast level of the coffee beans. An example of the labeled data can be seen in Table 6

Table 6. Labeling of Coffee Bean Roast Levels

No.	MQ2	MQ5	MQ135	Label
1	413	193	157	Light
2	414	192	143	Light
3	415	192	145	Light
4	490	174	256	Medium
5	490	174	464	Medium
6	490	174	279	Medium
7	529	181	243	Dark
8	533	186	247	Dark
9	536	190	246	Dark

The labeled data is then processed through encoding to convert string-type data into a format that can be accepted by the Artificial Neural Network (ANN). This encoding process begins by converting labels into integer values, ensuring that the data is in a numerical format suitable for machine learning algorithms.

Light →1
Medium →2
Dark →3

After converting the labels into integers, the next step is the encoding process, which transforms the integer-type labels into a format that can be processed by the Artificial Neural Network (ANN). This step ensures that the data is properly structured for model training and classification.

1 →000
2 →010
3 →100

Next, the data splitting process is carried out to divide the dataset into training data and testing data. This step ensures that the model can learn patterns from the training set and be evaluated on unseen data from the testing set. After splitting the data, the training process is conducted using the designed system with the following parameters:

Batch

This parameter defines the batch size used in the training process. Batch size refers to the number of data samples processed at a time and propagated through the neural network during training. In the system developed, the batch size is set to 64, meaning that 64 data samples are used in one iteration to update the weights of the neural network. This setting helps balance computational efficiency and model convergence speed..

Epochs

An epoch is completed when the entire training dataset has been processed through the neural network once. One epoch is achieved when all batches have been passed through the network. The number of epochs can be one or more, depending on the training configuration.

Aktivasi ReLU

ReLU (Rectified Linear Unit) activation is an activation layer in a CNN model that applies the function $f(x) = \max(0, x)$. This means the function performs thresholding at zero on the pixel values of the input data. This activation ensures that all pixel values less than zero are set to 0.

Softmax

Softmax transforms the input into a probability distribution.

This study presents the classification system testing results for coffee bean roast level using an Electronic Nose and Artificial Neural Network (ANN). The main step in model development is the training process, which enables the model to learn and recognize patterns from the dataset. Once training is complete, the trained model is evaluated using testing data.

A total of 900 numerical data samples were collected, with 720 samples used for training and 180 samples used for testing. In this study, experiments were conducted using batch size = 64 and epochs = 2000.

To evaluate classification performance, a confusion matrix was used, as shown in Figure 3. The obtained results demonstrate an accuracy of 86%.

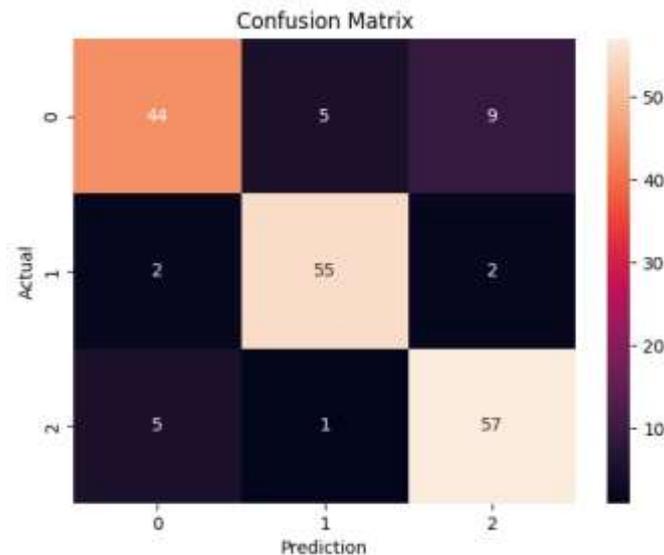


Figure 3. Confusion Matrix

$$Accuracy = \frac{TLight + TMed + TDark}{All\ Prediction\ Results} \times 100\%$$

$$Accuracy = \frac{44 + 55 + 57}{180} \times 100\% = 86\%$$

CONCLUSION

Based on the research conducted on the coffee bean roast level classification system, it can be concluded that after preparing the necessary tools and materials for the Electronic Nose, including MQ-2, MQ-5, and MQ-135 sensors, these components were connected to the Arduino Mega 2560 using jumper wires. The assembled device was then used for data collection, which was further processed for classification of coffee bean roast levels using an Artificial Neural Network (ANN). The model's processing results showed True Positive (TP) values of 44 for Light Roast, 55 for Medium Roast, and 57 for Dark Roast. The classification system successfully achieved an accuracy of 86% in determining the roast level of coffee beans.

REFERENCES

- [1] B. T. Edvan and R. Edison, "Pengaruh Jenis dan Lama Penyangraian pada Mutu Kopi Robusta (*Coffea robusta*)," vol. 4, no. 1, [Online]. Available: [URL jika tersedia].
- [2] K. Diantoro, "Implementasi Sensor MQ 4 dan Sensor DHT 22 pada Sistem Kompos Pintar Berbasis IoT (SIKOMPI)," **Electrician**, vol. 14, no. 3, pp. 84-94, 2020, doi: 10.23960/elc.v14n3.2157.
- [3] D. A. Putra, T. Rahmadani, A. D. Wicaksono, and A. Triwiyatno, "Sistem Pendeteksi Kadar Gas Methana (CH₄) Berbasis IoT Menggunakan NodeMCU ESP8266 dan Sensor Gas MQ-5," vol. 8, no. 2, 2019, [Online]. Available: [URL jika tersedia].
- [4] P. Rahardjo, "Sistem Penyiraman Otomatis Menggunakan Sensor Kelembaban Tanah Berbasis Mikrokontroler Arduino Mega 2560 Pada Tanaman Mangga Harum Manis Buleleng Bali," **Majalah Ilmiah Teknologi Elektro**, vol. 21, no. 1, p. 31, 2022, doi: 10.24843/MITE.2022.v21i01.P05.
- [5] B. Rahman, F. Fernando, and N. Indriawan, "Sistem Monitoring Kebocoran Gas dan Api Menggunakan Sensor MQ-2 dan Flame Sensor Berbasis Android," **Journal Sensi**, vol. 8, no. 2, pp. 209-222, 2022, doi: 10.33050/sensi.v8i2.2429.
- [6] A. A. Rosa, B. A. Simon, and K. S. Lieanto, "Sistem Pendeteksi Pencemar Udara Portabel Menggunakan Sensor MQ-7 dan MQ-135," [Online]. Available: [URL jika tersedia].

- [7] A. Safuan' and D. I. K. Swakarma, "Pengujian Respon Sensor MQ2 dan MQ8 dengan Metode Analisis Titik Pusat Klaster Berbasis Bulbus Olfactory Electronic (BOE)," [Online]. Available: [URL jika tersedia].
- [8] Zainuddin, Zahir. "Detection and Counting of the Number of Cocoa Fruits on Trees Using UAV." *2023 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*. IEEE, 2023.