sciendo INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS

Article | DOI: 10.2478/ijssis-2024-0019 Issue 1 | Vol. 17 (2024)

Development of a portable electronic nose for the classification of tea quality based on tea dregs aroma

Adi Djoko Guritno[1,*], Agus Harjoko[2], Megita Ryanjani Tanuputri[1], Diyah Utami Kusumaning Putri[2] and Nur Achmad Sulistyo Putro[2]

[1]Department of Agroindustrial Technology, Faculty of Agricultural Technology, Universitas Gadjah Mada, Yogyakarta, Indonesia

[2]Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Yogyakarta, Indonesia

*E-mail: adidjoko@ugm.ac.id

Received for publication June 26, 2023.

Abstract

The current assessment of tea quality is considered subjective. This study aims to develop a portable electronic nose to assess the aroma of tea dregs objectively by relying on the aromatic capture process through sensors and using multilayer perceptron (MLP). A MLP with some hyperparameter variations is used and compared with five machine-learning classifiers. The classification using MLP model with ReLU activation function and 3 hidden layers with 100 hidden nodes resulted in the highest accuracy of 0.8750 ± 0.0241. The MLP model using ReLU activation function is better than Sigmoid while increasing the number of hidden layers and hidden nodes does not necessarily enhance its performance. In the future, this research can be improved by adding sensors to the portable electronic nose, increasing the number of datasets used, and using ensemble learning or deep learning models.

Keywords

electronic nose, tea dregs aroma, multilayer perceptron, machine learning

1. Introduction

Camellia sinensis is a species in the flowering plant family of Theaceae whose leaf buds and leaves are used to produce the popular beverage tea. Tea is famous worldwide, although in significantly varied amounts; tea is widely regarded as the most consumed beverage on earth after water [1]. The combination of the processing techniques and the characteristics of manufactured tea determines the tea quality. Tea has different quality characteristics, including aroma, taste, color, and appearance due to the manufacturing process. The different types of tea, including black tea, can be classified based on its processing method: CTC (crush, tear, and curl) and orthodox tea. Orthodox tea has increasingly become famous because of its antioxidant and antibacterial properties [2]. Moreover, in terms of its quality, global black tea can be categorized into four grades of tea: leaf grade, broken, fannings, and dust. Based on this global category, the tea industry subsequently classifies its black tea into the first grade (i.e., Broken Orange Pekoe [BOP], Broken Orange Pekoe Fanning [BOPF], Pekoe Fanning [PF], Fanning), second grade (i.e., PFII, FII, etc.), and off-grade (i.e., Bohea). This grading category defines the price of tea. Kumar et al. [3] revealed that the color of tea liquor, taste, and aroma predominantly contribute to the quality of tea. Aroma is one of the crucial aspects of tea quality, which can decide the acceptance or rejection of a tea [1]. Tea aroma depends upon certain volatile aromatic substances developed during the fermentation process [4]. During the fermentation process, the grassy smell of tea leaves is converted into a floral smell through a complex chain of biochemical reactions [5, 6]. Therefore, the fermentation process is the most crucial of all the processes in the formation of tea aroma. The determination of volatile compounds

Open Access. Published by Sciendo. CC BY-NC-ND
© 2024 Guritno et al. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License https://creativecommons.org/licenses/by-nc-nd/4.0/

that actively contribute to the aroma of tea has been heavily studied in many varieties of tea using various methods [7–11].

The assessment of tea quality is one of the essential points since it affects the price of the tea product and customer acceptance. The traditional evaluation of tea quality is still carried out manually through organoleptic testing by expert tea testers by relying on their human senses. Unfortunately, the physiological and psychological factors during the test could impact the assessment's results [12], as well as the environmental factor. This leads to inconsistent, subjective, and expensive assessment outcomes, and complicated and challenging processes for ordinary people to understand. Factors like sensitivity and fatigue also have an impact on the outcomes of the trained tea tester's sensory analysis [13]. Therefore, testing in this way cannot be used to evaluate the quality of a product with a large number of samples [7]. Although sensory evaluation based on human smell is subjective, careful planning and comprehensive training of assessors enable it to become more objective but remain a more challenging alternative.

Additionally, some analytical techniques, including gas chromatography/mass spectrometry (GC/MS), a sophisticated tool that detects aroma compounds, can be used but requires specific analytical skills by an expert. Sample preparation costs are relatively expensive and inappropriate for real-time monitoring. In addition, since the tool is generally placed in laboratories, it is not lightweight and easy to use/carry outside. Therefore, there is a vital necessity to develop reliable, nonexpensive, precise, and rapid systems to serve as monitoring devices on the process of tea quality assessment. Due to their ability for odor analysis, electronic noses with multisensor systems have been demonstrated as an alternative way to detect and evaluate food quality [14].

The electronic nose (e-nose) is an electronic sensing device that mimics the human sense of smell and is used to detect, recognize, or quantify volatile substances [15]. Gardner stated that "an electronic nose is a system that uses an array of electronic chemical sensors and pattern recognition algorithm, capable of identifying simple or complicated odors," [16]. An electronic nose generally consists of hardware and software. The hardware on the electronic nose is a combination of a series of gas sensors as an aroma extraction system that functions to produce a sensor array as the sensor response signal outputs. Furthermore, the software on the electronic nose is a control and measurement system with a pattern recognition method to evaluate the sensor array patterns for analyzing the aroma [15, 18, 19]. The arrangement of sensor nodes with various functions can be used to develop smart sensing systems for different applications [20]. Artificial intelligence (AI) can be adopted throughout the value chain of agricultural products, making work efficient and reducing costs [21, 22].

Machine learning techniques enable accurate odor identification with qualitative and quantitative analysis. The e-nose signal pattern analysis in the tea quality assessment has been done using general machine learning techniques consisting of preprocessing and feature extraction stages, machine learning modeling, and drift compensation. The identification performance depends on the e-nose system setup and target gas [23]. Xu et al. obtained perfect results for the qualitative identification of tea quality grades based on a feature-level fusion strategy with support vector machine (SVM) and random forest (RF) [24]. Zhi et al. and Dai et al. also applied procedures to tea quality identification including feature-level fusion (fuse the time-domain-based feature and frequency-domain-based feature) and decision-level fusion (D-S evidence to combine the classification results from multiple classifiers) [25, 26]. Experimental results showed that the fused features could better represent signal characteristics compared with the single features and indicated that nonlinear algorithms were more effective in feature selection than linear algorithms, and the highest recognition rate was reached by feature fusion and KLDA-K-Nearest Neighbor (KNN) model. Rouabeh et al. designed fast and low-cost multipurpose e-nose for rapid gas identification [27]. This research compared Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) to project the extracted dataset to demonstrate the discriminative skills of the chosen features and also compared the KNN and SVM radial basis function-based (SVM-RBF) classifiers. These obtained results show that the LDA performed better than the PCA in terms of discrimination and the SVM-RBF has the best, stable, and consistent classification accuracy. Benabdellah et al. identified bad odor diffuses from two rotten types of meat that have almost the same rottenness odor using an electronic nose system [28]. Classification and identification of these odors are realized by PCA and Discriminative Factorial Analysis (DFA) methods and perform a good separation between classes. Xu et al. also compared three data reduction methods including PCA, LDA, and multidimensional scaling (MDS), and two

classifiers including SVM and logistic regression (LR) for qualitative evaluation of tea [29]. The results indicated that LDA outperformed original data, PCA, and MDS in both LR and SVM models, and SVM had an advantage over LR in developing classification models. Tudu et al. assessed eight normalization techniques and multilayer perceptron (MLP) classifier for black tea aroma classification using an electronic nose [30]. It is observed that the normalization techniques influence the results considerably and careful choice of the appropriate method is significant. However, none of the normalization techniques and machine learning approaches hold good results for each application, thus the performance differs based on the application [29, 30].

The results of manually performed preprocessing and feature extraction cannot guarantee the optimization of the machine learning model performance. The classification method without feature extraction is possible using deep learning algorithms. Unlike machine learning, the deep learning approach does not require us to manually extract or select features [31]. This is because deep learning is composed of several processing layers based on artificial neural networks to learn data representations with different levels of abstraction so that no additional processing is needed to represent the data [32]. MLP is a learning algorithm with one or many layers called a hidden layer that contains one or more neurons to carry out learning that can represent the characteristics of the input data. In addition, MLP can study nonlinear and complex models and adapt to data, so there is no requirement for an explicit functional form specification and model distribution [33].

On this basis, this research aims to develop a portable electronic nose system to assess tea quality based on the aroma of tea dregs by relying on the aromatic capture process through sensors. This is done by using a learning algorithm to classify tea aroma quality into three classes, i.e., fresh, burnt, and strange smell. Thus, the first part of this paper elaborates on the following steps: (1) developing a portable electronic nose system based on the fusion features from odor sensor arrangement; (2) data acquisition from various aromas of tea dregs with different quality; and (3) analysis the classification algorithms to determine tea quality by comparing the traditional machine learning and MLP algorithm. Furthermore, section II presents the portable electronic nose, data acquisition, classifiers, and experimental setup, while section III presents results and discussion, and section IV concludes this paper.

2. Experiment Setup

2.1. Tea samples

This study used black tea with varying aroma collected from PT Pagilaran in Batang Regency, Central Java Province, Indonesia as the research samples. Unit of Production (UP) Pagilaran is the main estate owned by PT Pagilaran with an area of 998 hectares and a capacity of 27 tons of tea shoots per day. In accordance with the rule of the Indonesian National Standard 1902:2016 regarding black tea [34], the conditions based on the aromas of the tea dregs are divided into three classes, i.e., good, burnt, and strange smell. Good condition is obtained from good quality tea so as to produce the aroma of tea dregs when brewed with a distinctive aroma of fresh tea. Meanwhile, the burnt condition is gathered from charred tea samples, usually caused by an error in the manufacturing process that causes a burnt aroma to appear on the tea dregs when brewed. Lastly, the strange smell is obtained from tea, in which its leaf processing is mixed with other components, such as grass, flowers, fruits, etc., causing the tea brewed to emit a strange smell.

2.2. Electronic nose unit

In this study, a portable electronic nose system was designed and developed to evaluate the aromas of the tea dregs into three classes, i.e., good, burnt, and strange smell. An electronic nose generally consists of four parts: an aroma extraction system, a sensor array, a control and measurement system, and a pattern recognition method.

The sampling method extracts the aroma and transports the volatile compounds from the tea aroma samples to the sensor chamber. It significantly contributes to an odor-sensing system's capability and reliability. There are two primary types of aroma extraction systems, i.e., the sample flow system and the static system. In the sample flow system, the sensors are positioned in the vapor flow, allowing for a rapid vapor exchange and quick measurement of many samples. Meanwhile, in the static system, measurements are usually made on the steady-state responses of the sensors exposed to vapor at a constant concentration since there is no vapor flow around the sensor [15]. Our portable electronic nose unit uses a static system for extracting aroma.

The most significant part of an electronic nose is the detection system, which consists of chemical sensors to convert a chemical change in the environment

into an electric signal in the gas sensors and respond to the concentration of particular compounds from gases [35]. Chemical semiconductor sensors can detect gases in samples through a chemical reaction that happens when the gases come in direct contact with the sensor surface. Since the electrical resistance in the sensor changes when exposed to the monitored gas, it is possible to detect both the chemical reaction and the presence of the gases. The measurement of the change in resistance makes it possible to identify the presence of gas and estimate the gas concentration [15]. The portable electronic nose used in this research was equipped with a metal oxide semiconductor (MOS) gas sensor array. It is composed of 6 different MOS gas sensors with two sensors supplied from MQ (Hanwei, China) and four sensors from TGS (Glenview, USA). The MOS gas sensors were placed at the top of the sensory chamber of the portable electronic nose system. The electronic diagram of the portable electronic nose is shown in Figure 1, and the name and specifications of the sensor array for some specific volatile compounds are presented in Table 1.

The control and measurement system is required to measure signals produced by the sensors. Before the signal is processed by a computer, it must be converted into a digital format. An analog carries it out to a digital converter, followed by a multiplexer to produce a digital signal connecting to a serial microprocessor's serial port. The single-board computer and the microcontroller are programmed to perform several tasks, including preprocessing the sensor signals to compute the input vectors and get the output. Finally, the output of the sensor array is gathered to be used for odor classification [15]. In this study, a teensy microcontroller and Arduino Nano directly connected to an LCD are used as the control and measurement system to store and process the input vectors. The LCD can display the output of the obtained sensor array and the aroma classification result.

The primary objective of an electronic nose is to identify odorant samples, estimate their concentration, and possibly classify them. The multivariate information obtained by the sensor array, an electronic identification of the volatile compound measured, can be sent to a computer to perform automated analysis

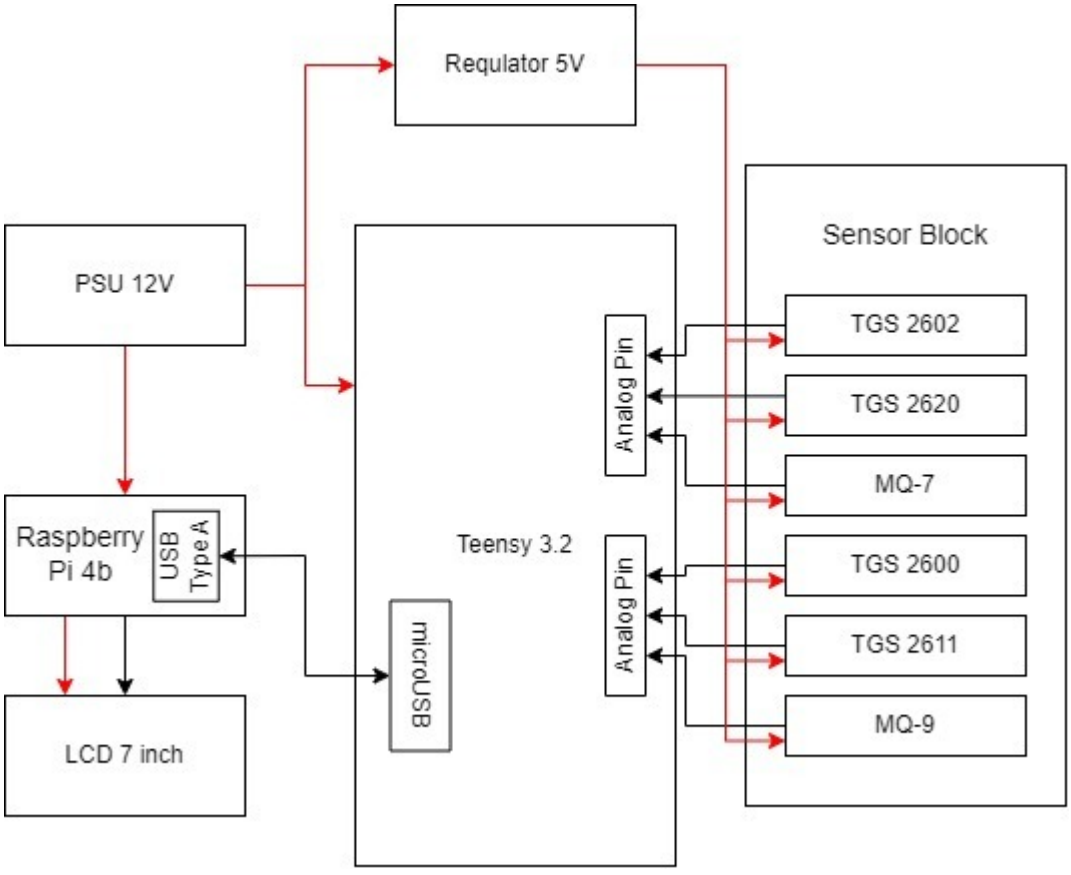


Figure 1: Electronic diagram of portable electronic nose.

Table 1: Sensor array used in the portable electronic nose and its specifications.

Name	Specification
TGS 2602	High sensitivity to gaseous air contaminants and VOCs such as ammonia, H_2S, ethanol, and toluene.
TGS 2620	High sensitivity to alcohol and organic solvent vapors and other volatile vapors, such as hydrogen, carbon monoxide, methane, and iso-butane.
MQ-7	High sensitivity to carbon monoxide.
MQ-9	High sensitivity to carbon monoxide, methane, and LPG.
TGS 2600	High sensitivity to low concentrations of gaseous air contaminants such as hydrogen, carbon monoxide, methane, iso-butane, and ethanol.
TGS 2611	High sensitivity and selectivity to methane gas, response to ethanol, hydrogen, and iso-butane.

VOCs, volatile organic compounds.

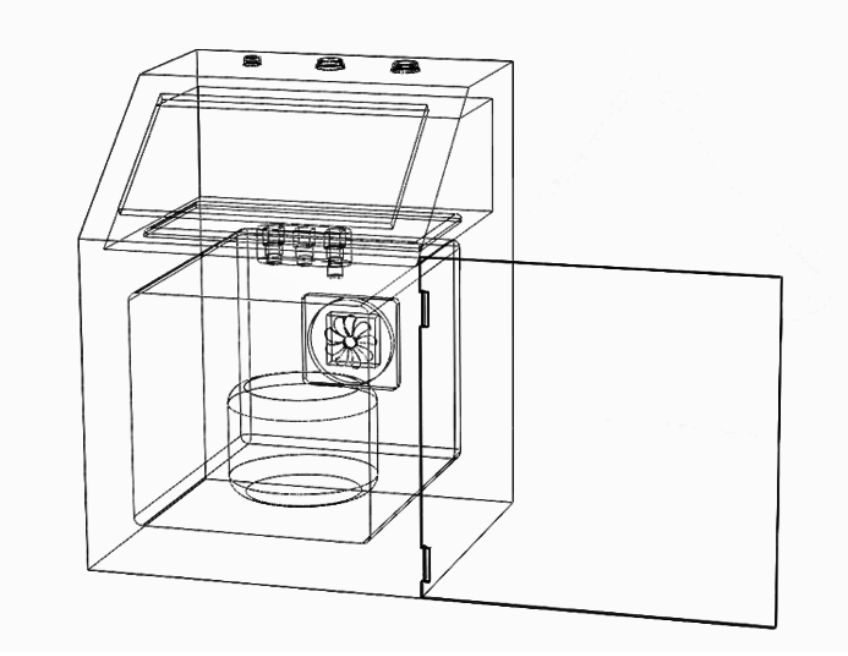


Figure 2: The design of the portable electronic nose.

and simulate the human sense of smell. This automated analysis from pattern recognition methods is vital to developing an electronic nose to detect and identify different volatile compounds responsible for sensing [15]. This study conducted several experiments to obtain the best pattern recognition method to detect the quality of tea aroma based on the smell of tea dregs using a MLP and compare it with several traditional machine learning models.

The design of a portable electronic nose developed in this study is shown in Figure 2, which generally consists of a sensory chamber for extracting the aroma of samples, sensor arrays, and control and measurement systems with a pattern recognition method to evaluate the quality of tea aroma. The sensory chamber for aroma extraction is box-shaped at the bottom equipped with a door. In the middle is a coaster for placing a teacup containing a sample of brewed tea dregs to make it closer to the sensor array placed at the top of this chamber. The back of the sensory chamber is equipped with a fan to suck air. This airflow aims to neutralize the aroma in the box. The top of this system has a button to turn on and turn off the fan. The sensor array consists of TGS 2602, TGS 2620, MQ-7, MQ-9, TGS 2600, and TGS 2611 placed at the top of the sensory chamber. This



Figure 3: Photograph of a portable electronic nose.

sensor array is connected to a teensy microcontroller and Arduino Nano as a control and measurement system. It has been programmed with a pattern recognition method to assess tea aroma quality. Arduino Nano is also directly connected to an LCD embedded in this tool to display the output obtained and the quality aroma classification result.

The portable electronic nose developed in this study is illustrated in Figure 3 which depicts the appearance of the portable electronic nose, particularly by highlighting its outer, inner, and overall look when used to analyze the brewed tea dregs sample in the cup.

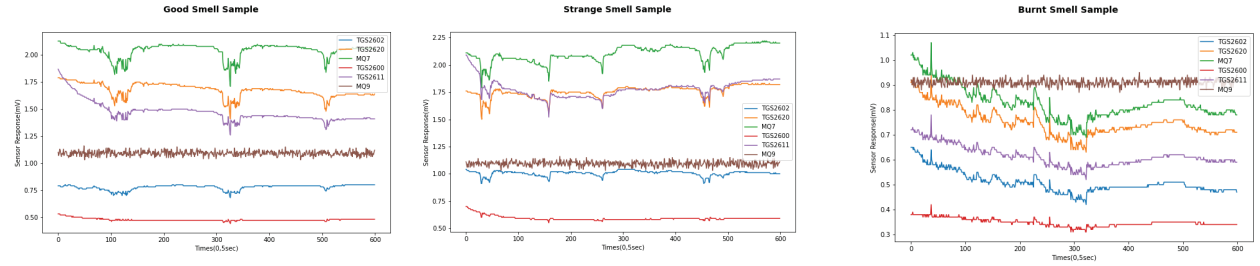
2.3. Data acquisition

The data acquisition process began with making a sample of tea dregs aroma by putting 5.6 g of tea into a cup before pouring 250 mL of water at a temperature of 90–95°C into it, and waiting for 5 min with the cup tightly closed and airproofed. Subsequently, the tea-steeping water was poured into the other container and the tea-brewing dregs were kept. The remaining tea dregs were left in the cup and covered. Before the sample was put inside the electronic nose, the aroma in the electronic nose was neutralized by turning on the fan for ±3 min. Afterward, the tea dregs sample in the cup was placed in the chamber sensory and the door was closed.

Data acquisition was performed using a teensy microcontroller device and a program running in Arduino Nano. The acquisition process was shown on the LCD, which was considered stable and final when the weight concentration reached a steady value. The output of each MOS sensor sampled every 0.5 s was considered an aroma feature. Six aroma features were acquired and recorded for each measurement for 5 min. A rectangular data matrix of 6 × 600 was organized as raw data. The measurements using a portable electronic nose were performed 30 times for each type of tea aroma sample. Hence, 90 samples of three different tea qualities were used in this experiment. For the “fresh” class of aroma, the experiment used black tea from many grades including BOP, BOPF, F I, F II, PF, PF II, PF III, Fann II, Dust II, and Bohea. The aroma of the tea was converted into electrical signals (voltage) by gas sensors that represent the gas sensor response. The sensor response is a description of each gas sensor that responds to the aroma samples with its characteristics. Typical response signals from gas sensors are shown in Figure 4.

2.4. Classification methods

This study conducted several experiments to obtain the best pattern recognition method. Pattern recognition can be used to classify the quality of tea based on the aroma of tea dregs into three classes, i.e., fresh, burnt, and strange smell, using MLP and compare it with several traditional machine learning models. The experiment used a total of 90 aroma samples from three different classes. Each sample represents gas sensor response from six sensor measurements in a data matrix of 6 × 600. The whole dataset is split into training and testing datasets using a stratified



(A) (B) (C)

Figure 4: Typical response of sensor array to tea aroma (A) good, (B) strange, and (C) burnt.

shuffle split strategy, with 72 (80%) of the data used for training and the remaining 18 (20%) used for testing. Following this, the training dataset was split again using a stratified shuffle split strategy with 4 numbers of splitting, so 18 (25%) data from training data were used for the validation dataset, and the remaining 54 (80%) were used for training.

2.4.1. Machine learning classifier

Five machine learning classifiers, namely Naïve Bayes (NB), KNN, SVM, Decision Tree (DT), and RF, were trained for the classification of aroma in this study. Raw data consisted of a rectangular data matrix with a size of 6 × 600. First, the data were preprocessed by changing the shape of the dimension into one-dimensional so that each data has 3600 elements. Then, the data were normalized using min-max normalization that scales the data in the range (0, 1). The target labels also encode with values between 0 and 2. Grid search cross-validation was used with a k-fold value of 4 to get the best hyperparameters over a parameter grid of different classifiers and it was examined to reach the validation dataset. The optimized hyperparameters using grid search for training classifiers are depicted in Table 2. Once the best model for each split was revealed, it was used to evaluate the testing dataset.

2.4.2. Multilayer perceptron

The training process using a MLP started with transforming the dimensions of raw data, which was originally a matrix of 6 × 600 into a one-dimensional vector, and thus each data had 3600 features. Furthermore, preprocessing was done by standardizing the data to have the same scale using standard feature scaling by transforming each feature into a value in the range (–1, 1). Since labels on data were categorical data, they must be and are converted into the numeric form using one-hot encoding. The training process was performed using a MLP algorithm for each split and the validation data was used to assess the model to get the best model. Once the best model for each split was obtained, it was used to evaluate the testing dataset.

MLP is a simple and efficient deep-learning framework that works on the principle of backpropagation algorithm. In backpropagation, the data are processed in forward direction to estimate the target and calculate the loss, then processed in backward direction by adjusting the weights to minimize the loss. This forward and backward process continues until optimal values are obtained [36]. MLP consists of three layers: the input layer, hidden layer(s), and output layer. The input layer consists of several nodes representing the data's features. The input used in classification with MLP is raw data that has been previously transformed. The output layer consists of three nodes representing the output target class. The Softmax activation function was used as the activation function of the output layer since this task is a multiclassification problem. The illustrative details of MLP architecture are displayed in Figure 5.

The algorithm run by MLP to get the optimal weight is as follows:

1. Initialize all weights with small random numbers.
2. If the stopping conditions have not been met, proceed to steps 2–8.
3. For each pair of training data, perform steps 3–8.
4. Each input unit receives a signal and forwards it to the hidden unit.
5. Calculate all outputs in the hidden unit $z_j\,(j = 1,2,\ldots,p)$ using Eq. (1) and an activation function.

$$z_j = f\left(z_{net_j}\right) = f\left(v_{j0} + \sum_{i=1}^{n} x_i v_{ji}\right) \quad (1)$$

Table 2: Hyperparameter settings for training using machine learning classifiers

Classifier	Hyperparameter	Value
NB (type = multinomial NB)	alpha	[0, 0.1, 0.01, 0.001, 1.0]
	fit_prior	[True, False]
KNN	n_neighbors	1–10
	P	[1, 2]
	weights	['uniform', 'distance']
SVM (type = SVC)	C	[1, 10, 100, 1000]
	gamma	[0.1, 0.01, 0.001, 0.0001]
	kernel	['linear', 'poly', 'rbf', 'sigmoid']
DT	criterion	['gini', 'entropy']
	max_depth	[5, 10, 15, 20, 25, 30]
	max_features	['sqrt', 'log2', None]
	min_samples_leaf	[2, 3, 4, 5]
	min_samples_split	[2, 3, 4, 5]
	splitter	['best', 'random']
RF	n_estimators	[10, 25, 50, 100]
	criterion	['gini', 'entropy', 'log_loss']
	max_depth	[5, 10, 15, 20, 25, 30]
	max_features	['sqrt', 'log2', None]
	min_samples_leaf	[2, 3, 4, 5]
	min_samples_split	[2, 3, 4, 5]

DT, Decision Tree; KNN, K-Nearest Neighbor; NB, Naïve Bayes; RF, Random Forest; SVM, Support Vector Machine.

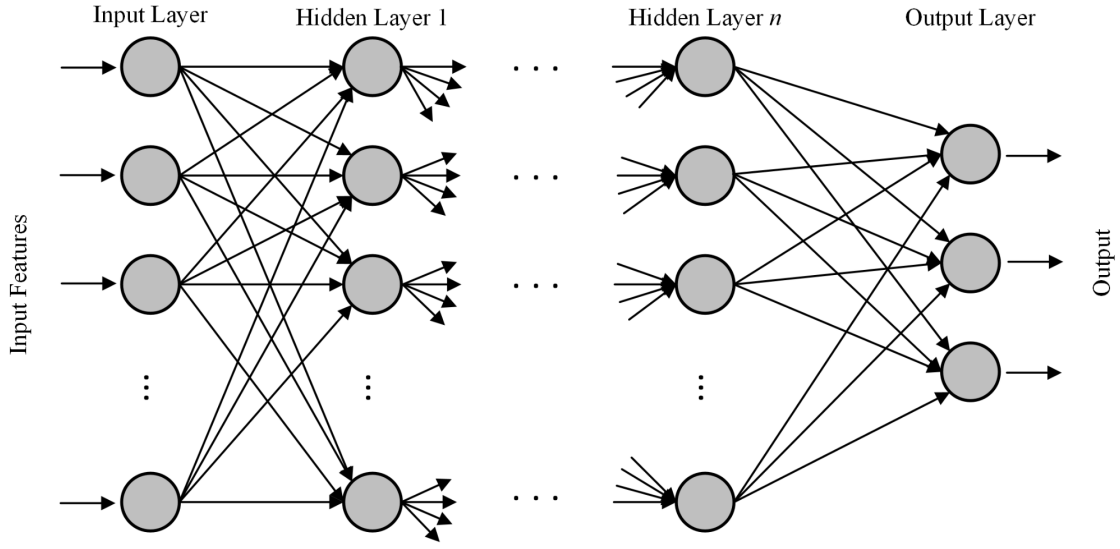


Figure 5: The architecture of MLP. MLP, multilayer perceptron.

where f is an activation function, v_{j0} is bias vector, x_i is the input vector or output from the previous layer, and v_{ji} is weights.

6. Calculate all outputs in the output unit $y_k\left(k=1,2,\ldots,m\right)$ using Eq. (2) and an activation function.

$$y_k = f\left(y_{net_k}\right) = f\left(w_{k0} + \sum_{j=1}^{p} z_j w_{kj}\right) \quad (2)$$

7. Calculate δ factor of output unit based on error in each output unit $y_k\left(k=1,2,\ldots,m\right)$ using Eq. (3).

$$\delta_k = \left(t_k - y_k\right) f'\left(y_{net_k}\right) = \left(t_k - y_k\right) y_k\left(1-y_k\right) \quad (3)$$

t_k is target, δ_k is the error unit used for updating the previous weight. Compute weight increment $\Delta w_{kj}(k=1,2,\ldots,m; j=0,1,2,\ldots,p)$ with learning rate α using Eq. (4).

$$\Delta w_{kj} = \alpha \delta_k z_j \quad (4)$$

Update all new weights $w_{kj}(k=1,2,\ldots,m; j=0,1,2,\ldots,p)$ in output unit using Eq. (5)

$$w_{kj}\left(new\right) = w_{kj}\left(old\right) + \Delta w_{kj} \quad (5)$$

8. Calculate δ factor of hidden unit based on the error in each hidden unit $z_j\left(j=1,2,\ldots,p\right)$ using Eq. (6).

$$\delta_j = \delta_{net_j} f'\left(z_{net_j}\right) = \sum_{k=1}^{m} \delta_k w_{kj} z_j\left(1-z_j\right) \quad (6)$$

Compute weight increment $\Delta v_{ji}(j=1,2,\ldots,p; i=1,2,\ldots,n)$ with learning rate α using Eq. (7).

$$\Delta v_{ji} = \alpha \delta_j x_i \quad (7)$$

Update all new weights $v_{ji}(j=1,2,\ldots,p; i=1,2,\ldots,n)$ in hidden unit using Eq. (8).

$$v_{ji}\left(new\right) = v_{ji}\left(old\right) + \Delta v_{ji} \quad (8)$$

The number of hidden layers and the number of nodes in the hidden layer play an essential role in MLP classification. The architecture of an artificial neural network is trial and error, where there is no definite architecture to solve a problem, making it necessary to conduct experiments by varying the number of hidden layers and the number of nodes in the hidden layer. Therefore, in this study, the number of hidden layers and nodes will be varied to obtain the optimum model for classification. In addition, this study also varied the activation function by comparing the activation function of the Rectified Linear Unit (ReLU) and Sigmoid. The Softmax function was used as the activation function of the output unit as this task is a multiclassification problem. The variations in the activation functions, the number of hidden layers, and the number of nodes in the hidden layer of MLP are presented in Table 3. This study used several other parameters with fixed values, including the maximum number of an epoch of 100, the learning rate of 0.00001, and the batch size of 1, and used Adam as the optimization function. "EarlyStoppingCallback" with early stopping patience 10 to choose the best model was also applied so as to stop the training

Table 3: Variation of the number of hidden layers, the number of hidden nodes, and the activation function of MLP architecture

Activation Function	Hidden Layer	Hidden Node
ReLU	2	50, 50
		100, 100
	3	50, 50
		100, 100
	4	50, 50
		100, 100
Sigmoid	2	50, 50
		100, 100
	3	50, 50
		100, 100
	4	50, 50
		100, 100

MLP, multilayer perceptron; ReLU, Rectified Linear Unit.

process when the validation accuracy score worsened for ten evaluation calls.

2.4.3. Model evaluation

The machine learning and MLP models with the best hyperparameters gathered for each split were evaluated with the testing dataset. In this study, four evaluation metrics were applied to evaluate the model performance for each split: accuracy, precision, recall, and F1-score. Accuracy is the proportion of correctly classified samples of the total number of samples in the testing dataset. A recall is the proportion of true positive predictions of the actual positive samples. Precision is the proportion of true positive predictions of the total number of optimistic predictions. The F1-score is the harmonic mean of precision and recall. In the subsequent step, the average value of each evaluation metric was calculated.

3. Results and Discussion

This section provides and discusses the evaluation results of the classification methods using machine learning and MLP models. Table 4 displays the evaluation performances for machine learning classifiers, including the average accuracy, precision, recall, and F1-score.

According to the evaluation results using machine learning classifiers as illustrated in Table 4, the NB classifier has the lowest average accuracy of 0.5972 with a standard deviation of 0.0241 and an average F1-score of 0.5432 with a standard deviation of 0.0181. Moreover, the DT classifier produced slightly better results than NB with an accuracy of around 0.5972, but with a higher standard deviation of 0.0992 and an F1-score of around 0.6002 with a standard deviation of 0.0932. The high standard deviation scores indicate that metric scores are still not constant and change significantly. The RF classifier performs slightly better than the DT, with an accuracy of 0.6667 ± 0.0556 and an F1-score of 0.6572 ± 0.0444. The best machine learning classifiers are SVM and KNN. The average accuracy of these two models is identical with around 0.7778. However, the standard deviation for KNN is 0 and the SVM is 0.0393. Thus, the KNN result is more consistent than the SVM. Furthermore, according to the F1-score, there is a slightly different result between KNN and SVM. KNN obtained an F1-score of 0.7575 ± 0.0112, while a higher score was obtained with the SVM which gathered an F1-score of 0.7702 ± 0.0433.

Table 5 demonstrates the evaluation performances for MLP with variations of the number of hidden layers, the number of hidden nodes, and the activation function. The evaluation metrics include the average accuracy, precision, recall, and F1-score.

Evaluation performances in Table 5 indicate that the MLP model using the ReLU activation function is better than the Sigmoid activation function. It also revealed that increasing the number of hidden layers and nodes does not necessarily enhance its performance. The best architecture obtained the highest average accuracy of 0.8750 with a standard deviation of 0.0241 and F1-score of 0.8703 with a standard deviation of 0.0246 when using the MLP model with three hidden layers and 100 hidden nodes in each layer.

Overall, the performance evaluation results using the MLP model are higher than the performance results of traditional machine learning classifiers. The highest performance results in experiments using machine learning classifiers were achieved with the SVM method with an F1-score of 0.7702 ± 0.0433. Meanwhile, the performance results when using the MLP model with three hidden layers and 100 hidden nodes in each layer produced an F1 score of 0.8703 ± 0.0246.

Table 4: The evaluation performance of machine learning classifiers

Classifier	Accuracy	Precision	Recall	F1-score
NB	0.5972 ± 0.0241	0.6513 ± 0.0877	0.5972 ± 0.0241	0.5432 ± 0.0181
KNN	0.7778 ± 0.0000	0.7897 ± 0.0253	0.7778 ± 0.0000	0.7575 ± 0.0112
SVM	0.7778 ± 0.0393	0.7792 ± 0.0473	0.7778 ± 0.0393	0.7702 ± 0.0433
DT	0.5972 ± 0.0992	0.6196 ± 0.0894	0.5972 ± 0.0992	0.6002 ± 0.0932
RF	0.6667 ± 0.0556	0.6649 ± 0.0443	0.6667 ± 0.0556	0.6572 ± 0.0444

DT, Decision Tree; KNN, K-Nearest Neighbor; NB, Naïve Bayes; RF, Random Forest; SVM, Support Vector Machine.

Table 5: The evaluation performance of MLP

Activation function	Hidden layer	Hidden node	Accuracy	Precision	Recall	F1-score
ReLU	2	50	0.8472 ± 0.0461	0.8714 ± 0.0545	0.8472 ± 0.0461	0.8394 ± 0.0489
		100	0.8750 ± 0.0461	0.8921 ± 0.0464	0.8750 ± 0.0461	0.8683 ± 0.0511
	3	50	0.8611 ± 0.0278	0.8831 ± 0.0362	0.8611 ± 0.0278	0.8537 ± 0.0323
		100	0.8750 ± 0.0241	0.8921 ± 0.0361	0.8750 ± 0.0241	0.8703 ± 0.0246
	4	50	0.8472 ± 0.0461	0.8564 ± 0.0581	0.8472 ± 0.0461	0.8399 ± 0.0494
		100	0.8472 ± 0.0241	0.8712 ± 0.0307	0.8472 ± 0.0241	0.8362 ± 0.0290
Sigmoid	2	50	0.7361 ± 0.0241	0.6907 ± 0.0486	0.7361 ± 0.0241	0.6816 ± 0.0385
		100	0.7500 ± 0.0278	0.7738 ± 0.0636	0.7500 ± 0.0278	0.7149 ± 0.0247
	3	50	0.7083 ± 0.0241	0.7401 ± 0.0736	0.7083 ± 0.0241	0.6486 ± 0.0305
		100	0.7222 ± 0.0000	0.7817 ± 0.0413	0.7222 ± 0.0000	0.6583 ± 0.0185
	4	50	0.6806 ± 0.0461	0.5978 ± 0.1663	0.6806 ± 0.0461	0.5978 ± 0.0888
		100	0.7083 ± 0.0241	0.7153 ± 0.1564	0.7083 ± 0.0241	0.6191 ± 0.0495

MLP, multilayer perceptron; ReLU, Rectified Linear Unit.

4. Conclusion

This study developed a portable electronic nose that can assess tea quality based on tea dregs' aroma by relying on the aromatic capture process through sensors and then using MLP to classify the quality of tea aroma. Tea quality was divided into three classes, i.e., fresh, burnt, and strange smell. The sensors used in the TeaNose were TGS 2602, TGD 2620, MQ-7, MQ-9, TGS 2600, and TGS 2611. The dataset in this study consisted of three aroma quality classes, each consisting of 30 samples. A MLP with variations of the activation function, the number of hidden layers, and the number of nodes were used in this study. The MLP was then compared with five machine learning classifiers. According to the machine learning classifier, the SVM obtained the highest average accuracy of 0.7778 ± 0.0393 and an F1-score of 0.7702 ± 0.0433. The results revealed that the classification using the MLP model with ReLU activation function and three hidden layers with 100 hidden nodes obtained the highest average accuracy of 0.8750 ± 0.0241 and F1-score of 0.8703 ± 0.0246. Based on these experiments, it is noteworthy that the MLP model using the ReLU activation function is better than the Sigmoid activation function while increasing the number of hidden layers and hidden nodes does not always enhance its performance. Overall, the MLP model produces better performance evaluation results than traditional machine learning classifiers.

The portable electronic nose developed in this research used 6 sensors and limited tea sample data consisting of 3 tea aroma qualities with each sample taken 30 times. Future research can be improved by adding sensors to the portable e-nose and increasing the number of datasets used. Apart from that, the model classifiers used to detect tea quality are still limited to traditional machine learning classifiers. In the future, this research can be explored by using ensemble learning to combine several machine learning models or using deep learning models based on recurrent neural network (RNN) architecture.

Acknowledgments

The authors wish to acknowledge the financial support from the Ministry of Research, Technology and Higher Education of the Republic of Indonesia by 2022 Universitas Gadjah Mada Grants of "Penelitian Terapan Unggulan Perguruan Tinggi" No. 1832/UN1/DITLIT/DitLit/PT.01.03/2022. The authors also thank Mr. M. Ridho Fuadin for helping in the hardware development.

References

[1] V. Sai, P. Chaturvedula, and I. Prakash, "The aroma, taste, color and bioactive constituents of tea," *Journal of Medicinal Plants Research*, vol. 5, no. 11, pp. 2110–2124, 2011, [Online]. Available: http://www.academicjournals.org/JMPR

[2] D. Kalauni, B. Joshi, and A. Joshi, "Production, marketing, and future prospects of Nepali orthodox tea," *Cogent Food Agric*, vol. 6, no. 1, p. 1757227, Jan. 2020, doi: 10.1080/23311932.2020.1757227.

[3] A. Kumar, H. Singh, S. Sharma, and A. Kumar, "Color Analysis of Black Tea Liquor using Image Processing Techniques," *International Journal of Electronics & Communication Technology*, vol. 2, no. 3, pp. 292–296, 2011.

[4] N. Bhattacharyya *et al.*, "Monitoring of black tea fermentation process using electronic nose," *J Food Eng*, vol. 80, no. 4, pp. 1146–1156, Jun. 2007, doi: 10.1016/j.jfoodeng.2006.09.006.

[5] G. W. Sanderson, "The Chemistry of Tea and Tea Manufacturing," in *Structural and functional aspects of phytochemistry: Recent advance in phytochemistry*, V. C. Runeckles and T. C. Tso, Eds., 1st ed.New York: Academic Press, 1972, pp. 247–317.

[6] D. Saikia, P. K. Boruah, and U. Sarma, "Development And Implementation Of A Sensor Network To Monitor Fermentation Process Parameter In Tea Processing," *International Journal on Smart Sensing and Intelligent Systems*, vol. 7, no. 3, pp. 1254–1270, Jan. 2022, doi: 10.21307/ijssis-2017-704.

[7] M. Kawakami, S. N. Ganguly, J. Banerjee, and A. Kobayashi, "Aroma Composition of Oolong Tea and Black Tea by Brewed Extraction Method and Characterizing Compounds of Darjeeling Tea Aroma," *J Agric Food Chem*, vol. 43, no. 1, pp. 200–207, Jan. 1995, doi: 10.1021/jf00049a037.

[8] X. Pang *et al.*, "Development of regression model to differentiate quality of black tea (Dianhong): correlate aroma properties with instrumental data using multiple linear regression analysis," *Int J Food Sci Technol*, vol. 47, no. 11, pp. 2372–2379, Nov. 2012, doi: 10.1111/j.1365-2621.2012.03112.x.

[9] C. Schuh and P. Schieberle, "Characterization of the Key Aroma Compounds in the Beverage Prepared from Darjeeling Black Tea: Quantitative Differences between Tea Leaves and Infusion," *J Agric Food Chem*, vol. 54, no. 3, pp. 916–924, Feb. 2006, doi: 10.1021/jf052495n.

[10] C. Wang, S. Lv, Y. Wu, M. Lian, X. Gao, and Q. Meng, "Study of aroma formation and transformation during the manufacturing process of Biluochun green tea in Yunnan Province by HS-SPME and GC-MS," *J Sci Food Agric*, vol. 96, no. 13, pp. 4492–4498, Oct. 2016, doi: 10.1002/jsfa.7663.

[11] C. Wang *et al.*, "A comparative study of volatile components in Dianhong teas from fresh leaves of four tea cultivars by using chromatography-mass spectrometry, multivariate data analysis, and descriptive sensory analysis," *Food Research International*, vol. 100, pp. 267–275, Oct. 2017, doi: 10.1016/j.foodres.2017.07.013.

[12] A. Akuli *et al.*, "A Machine Vision System For Estimation Of Theaflavins And Thearubigins In Orthodox Black Tea," *International Journal on Smart Sensing and Intelligent Systems*, vol. 9, no. 2, pp. 709–731, Jan. 2016, doi: 10.21307/ijssis-2017-891.

[13] M. F. Njoman *et al.*, "The vulnerability of human sensory evaluation and the promising senses instrumentation," *British Food Journal*, vol. 119, no. 10, pp. 2145–2160, 2017, doi: 10.1108/BFJ-10-2016-0505.

[14] X. Pang *et al.*, "Development of regression model to differentiate quality of black tea (Dianhong): correlate aroma properties with instrumental data using multiple linear regression analysis," *Int J Food Sci Technol*, vol. 47, no. 11, pp. 2372–2379, Nov. 2012, doi: 10.1111/J.1365-2621.2012.03112.X.

[15] M. Ghasemi-Varnamkhasti, C. Apetrei, J. Lozano, and A. Anyogu, "Potential use of electronic noses, electronic tongues and biosensors as multisensor systems for spoilage examination in foods," *Trends Food Sci Technol*, vol. 80, pp. 71–92, Oct. 2018, doi: 10.1016/j.tifs.2018.07.018.

[16] R. Gutierrez-Osuna, H. T. Nagle, B. Kermani, and S. S. Schiffman, "Signal Conditioning and Preprocessing," in *Handbook of Machine Olfaction*, Weinheim, FRG: Wiley-VCH Verlag GmbH & Co. KGaA, 2004, pp. 105–132. doi: 10.1002/3527601597.ch5.

[17] J. W. Gardner and P. N. Bartlett, "A brief history of electronic noses," *Sens Actuators B Chem*, vol. 18, no. 1–3, pp. 210–211, Mar. 1994, doi: 10.1016/0925-4005(94)87085-3.

[18] S. Kiani, S. Minaei, and M. Ghasemi-Varnamkhasti, "Real-time aroma monitoring of mint (Mentha spicata L.) leaves during the drying process using electronic nose system," *Measurement*, vol. 124, pp. 447–452, Aug. 2018, doi: 10.1016/j.measurement.2018.03.033.

[19] M. Tohidi, M. Ghasemi-Varnamkhasti, V. Ghafarinia, M. Bonyadian, and S. S. Mohtasebi, "Development of a metal oxide semiconductor-based artificial nose as a fast, reliable and non-expensive analytical technique for aroma profiling of milk adulteration," *Int Dairy J*, vol. 77, pp. 38–46, Feb. 2018, doi: 10.1016/j.idairyj.2017.09.003.

[20] A. James, A. Seth, and S. C. Mukhopadhyay, "IoT enabled sensor node: a tutorial paper," *International Journal on Smart Sensing and Intelligent Systems*, vol. 13, no. 1, pp. 1–18, Jan. 2020, doi: 10.21307/ijssis-2020-022.

[21] A. Aryal, Y. Liao, P. Nattuthurai, and B. Li, "The emerging big data analytics and IoT in supply chain

management: a systematic review," *Supply Chain Management*, vol. 25, no. 2. Emerald Group Holdings Ltd., pp. 141–156, Feb. 24, 2020. doi: 10.1108/SCM-03-2018-0149.

[22] C. Ganeshkumar, S. K. Jena, A. Sivakumar, and T. Nambirajan, "Artificial intelligence in agricultural value chain: review and future directions," *J Agribus Dev Emerg Econ*, 2021, doi: 10.1108/JADEE-07-2020-0140.

[23] Z. Ye, Y. Liu, and Q. Li, "Recent Progress in Smart Electronic Nose Technologies Enabled with Machine Learning Methods," *Sensors*, vol. 21, no. 22, p. 7620, Nov. 2021, doi: 10.3390/s21227620.

[24] M. Xu, J. Wang, and L. Zhu, "The qualitative and quantitative assessment of tea quality based on E-nose, E-tongue and E-eye combined with chemometrics," *Food Chem*, vol. 289, pp. 482–489, Aug. 2019, doi: 10.1016/j.foodchem.2019.03.080.

[25] R. Zhi, L. Zhao, and D. Zhang, "A Framework for the Multi-Level Fusion of Electronic Nose and Electronic Tongue for Tea Quality Assessment," *Sensors*, vol. 17, no. 5, p. 1007, May 2017, doi: 10.3390/s17051007.

[26] Y. Dai, R. Zhi, L. Zhao, H. Gao, B. Shi, and H. Wang, "Longjing tea quality classification by fusion of features collected from E-nose," *Chemometrics and Intelligent Laboratory Systems*, vol. 144, pp. 63–70, May 2015, doi: 10.1016/j.chemolab.2015.03.010.

[27] H. Rouabeh, S. Gomri, and M. Masmoudi, "The design and validation of a fast and low-cost multi-purpose electronic nose for rapid gas identification," *Sensor Review*, vol. 42, no. 6, pp. 613–630, Nov. 2022, doi: 10.1108/SR-01-2022-0019.

[28] N. Benabdellah, K. Hachami, M. Bourhaleb, and N. Benazzi, "IDENTIFICATION OF TWO TYPES OF ROTTEN MEAT USING AN ELECTRONIC NOSE FOR FOOD QUALITY CONTROL," *International Journal on Smart Sensing and Intelligent Systems*, vol. 10, no. 3, pp. 1–23, Jan. 2017, doi: 10.21307/ijssis-2017-229.

[29] M. Xu, J. Wang, and L. Zhu, "Tea quality evaluation by applying E-nose combined with chemometrics methods," *J Food Sci Technol*, vol. 58, no. 4, pp. 1549–1561, Apr. 2021, doi: 10.1007/s13197-020-04667-0.

[30] B. Tudu, B. Kow, N. Bhattacharyya, and R. Bandyopadhyay, "Normalization techniques for gas sensor array as applied to classification for black tea," *International Journal on Smart Sensing and Intelligent Systems*, vol. 2, no. 1, pp. 176–189, Jan. 2009, doi: 10.21307/ijssis-2017-344.

[31] N. K. Chauhan and K. Singh, "A Review on Conventional Machine Learning vs Deep Learning," in *2018 International Conference on Computing, Power and Communication Technologies (GUCON)*, IEEE, Sep. 2018, pp. 347–352. doi: 10.1109/GUCON.2018.8675097.

[32] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: review, opportunities and challenges," *Brief Bioinform*, vol. 19, no. 6, pp. 1236–1246, Nov. 2018, doi: 10.1093/bib/bbx044.

[33] U. SIRIPATRAWAN and P. JANTAWAT, "A novel method for shelf life prediction of a packaged moisture sensitive snack using multilayer perceptron neural network," *Expert Syst Appl*, vol. 34, no. 2, pp. 1562–1567, Feb. 2008, doi: 10.1016/j.eswa.2007.01.008.

[34] Badan Standardisasi Nasional, "Standar Nasional Indonesia: Teh Hitam," 2016, [Online]. Available: www.bsn.go.id

[35] H. T. Nagle, R. Gutierrez-Osuna, B. G. Kermani, and S. S. Schiffman, "Environmental Monitoring," in *Handbook of Machine Olfaction*, Weinheim, FRG: Wiley-VCH Verlag GmbH & Co. KGaA, 2004, pp. 419–444. doi: 10.1002/3527601597.ch17.

[36] Hecht-Nielsen, "Theory of the backpropagation neural network," in *International Joint Conference on Neural Networks*, IEEE, 1989, pp. 593–605 vol.1. doi: 10.1109/IJCNN.1989.118638.