Received: January 26, 2024. Revised: February 20, 2024.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Reducing the Electronic Nose Sensor Array for Asthma Detection Using Firefly Algorithm

Muhammad Rivai[1]* Dava Aulia[2] Sheva Aulia[3]

[1]*Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia*
[2]*Department of Informatics, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia*
[3]*Department of Information Systems, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia*
* Corresponding author's Email: muhammad_rivai@ee.its.ac.id

Abstract: Exhaled breath analysis comprises chemical compounds that can be utilized for diagnostic purposes, including asthma detection. An electronic nose can be offered as a means of monitoring patient circumstances. A significant problem often occurs when determining the appropriate number of gas sensors while maintaining high accuracy. The firefly algorithm (FA) is very effective because of its exploratory capabilities, presents theories that are easy to understand and has relatively fewer parameters. This study aims to reduce and determine the appropriate number of gas sensors for an electronic nose in differentiating healthy and asthmatic subjects using the FA and exhaled breath analysis. The experimental results indicate that the FA provides only four gas sensors that still maintain high performance. The convolutional neural network model was favored for its ability to classify the entire asthma dataset, making it the best machine learning model for the electronic nose, with an accuracy of 97.8%.

Keywords: Asthma, Diseases, Electronic nose, Firefly algorithm.

1. Introduction

Asthma is a chronic respiratory disorder initiated by various factors, especially chemical exposure [1]. This stimulus yields narrowing and inflammation of the respiratory tract and airway obstruction due to excessive mucus production. As a result, asthma sufferers often have difficulty breathing [2].

Asthma is a non-communicable disease. This illness occurs at all ages, and environmental factors can cause most asthma [2, 3]. Common symptoms are shortness of breath, coughing, wheezing, and chest tightness. According to WHO, the World Health Organization, in 2019, asthma concerned more than 262 million people worldwide in countries with limited financial resources [4]. Therefore, this required immediate attention.

One of the biomarkers for asthma can be seen in exhaled breath analysis [5]. This approach does not involve body contact, makes a person comfortable, and promises more convenient and safe measurements [6]. Therefore, this analysis offers an easily accessible way to detect biomarkers.

Exhaled breath contains chemical compounds used for diagnostic purposes, most commonly hydrogen sulfide, nitric oxide, and volatile organic compounds [7]. The latter two substances are often found in asthma sufferers [8]. Gas chromatography, a technique for separating and analyzing gas mixtures, requires longer analysis time and is expensive. An electronic nose is a device that employs an array of chemical gas sensors to recognize the characteristics of various types of gas. This approach has been applied to detect diseases in the human body, such as lung disease [6, 9, 10], and diabetes [11, 12].

Using the appropriate number of gas sensors can determine the effectiveness of measuring the gas content in exhaled breath. Therefore, the gas sensor array on the electronic nose needs to be optimized to provide better outcomes, including cost-effectiveness, increased efficiency, and accurate results [10]. The principle of the sensor selection method is to

eliminate sensors with similar capabilities that do not significantly affect the classification phase.

Some optimization methods can be applied to this purpose. The correlation coefficient, cluster analysis, and distinguishing performance value techniques can give the most effective and eliminate redundant sensors [13]. Cluster analysis achieves sensor independence, and the associated sensor arrays are validated. However, this optimizer method is complex, requires human expertise, and is often applied to specific cases. The random forest (RF) can be offered as a sensor selection method, which is determined based on feature importance [14]. The optimizer assigns an importance score to each sensor, representing each sensor's contribution to the method, and sensors are selected based on the highest ranking. Nevertheless, this approach can introduce risks of bias and instability, require expensive computations, and are complex to implement. Genetic algorithm (GA) is part of an optimization technique based on the principles of gene evolution, which has advantages in overcoming feature selection problems [10]. This optimizer can solve noise and collinearity problems by utilizing three genetic operators: selection, crossover, and mutation. However, this method has a complicated theory and more parameters, which can affect the complexity of the technique and require a longer calculation time. Its performance can be sensitive to parameters that require careful experimentation and adjustment.

The firefly algorithm (FA) is an optimization technique that mimics the behavior of fireflies to solve complex optimization problems by simulating firefly interactions based on the degree of brightness intensity. Fireflies represent a possible solution, where their brightness can be illustrated as an indicator of their quality or fitness value [15]. They communicate with each other through their attraction ability, which is determined by the level of light emitted by other fireflies, and approach fireflies of higher brightness [16]. Therefore, fireflies with brighter brightness offer a better solution [17]. FA is considered very effective and efficient because of its brief and independent exploration capabilities, showing that this approach can work reliably in various problems [18]. In addition, this algorithm presents a simple theory and fewer parameters compared to GA. For this reason, these advantages can provide benefits, including reducing model complexity, speeding up the time required for calculations, and making it easier to adjust to acquire the considerable informative variables [19]. As a result, several studies show that this algorithm is favored as an optimization method [16, 20, 21]. Moreover, no previous study investigated the FA as a sensor array optimization in the electronic nose system.

The novelty of this study includes reducing and determining the appropriate number of gas sensors for the electronic nose using the FA method. Then, several machine learning models are configured with various architectures, which can still provide better evaluation performance for electronic nose systems in predicting healthy and asthmatic subjects. This paper is regulated as follows. Section 1 concerns the research background. Section 2 proposes the proposed method and experimental arrangement. Section 3 demonstrates experimental results and discussion. Section 4 summarizes meaningful results and forthcoming research.

2. Materials and methods

2.1 Subject and proposed research design

Exhaled breath data is accumulated at Dr. Soetomo General Hospital, East Java, Surabaya, Indonesia. Subjects were 30 healthy people and 30 asthmatic suspects, approved by a pulmonary specialist, where the inclusions were men and women aged between 30 and 60 years who had no acute or chronic conditions and were non-smokers [22]. The severity of asthma was assessed using the global initiative for asthma (GINA) and the asthma control test (ACT), classified into controlled, partially controlled, and uncontrolled asthma.

Fig. 1 shows the proposed method design. All participants were prepared to provide breath samples into 1-liter Tedlar bags. An electronic nose is utilized as a measurement instrument. Gas sensor response signal data forms different curves, which can be organized into binary and multiclass datasets. Three pattern recognition models, including support vector machine (SVM), RF, and extreme gradient boosting (XGBoost), were evaluated, and the best one was selected based on the highest accuracy. Data analysis includes data preprocessing and feature scaling to confirm that the model performs effectively. The firefly algorithm, with SVM as its fitness function, evaluates a binary dataset to identify the number of gas sensor arrays. The selected model assesses the number of candidate gas sensors, and the best candidate has the lowest number that maintains high accuracy. In the final stage, the optimal number of gas sensors is tested using the entire dataset with other classification models, including artificial neural network (ANN), 1-dimensional convolutional neural network (1D-CNN), long short-term memory (LSTM), bidirectional long short-term memory (BiLSTM), and gated recurrent unit (GRU). The

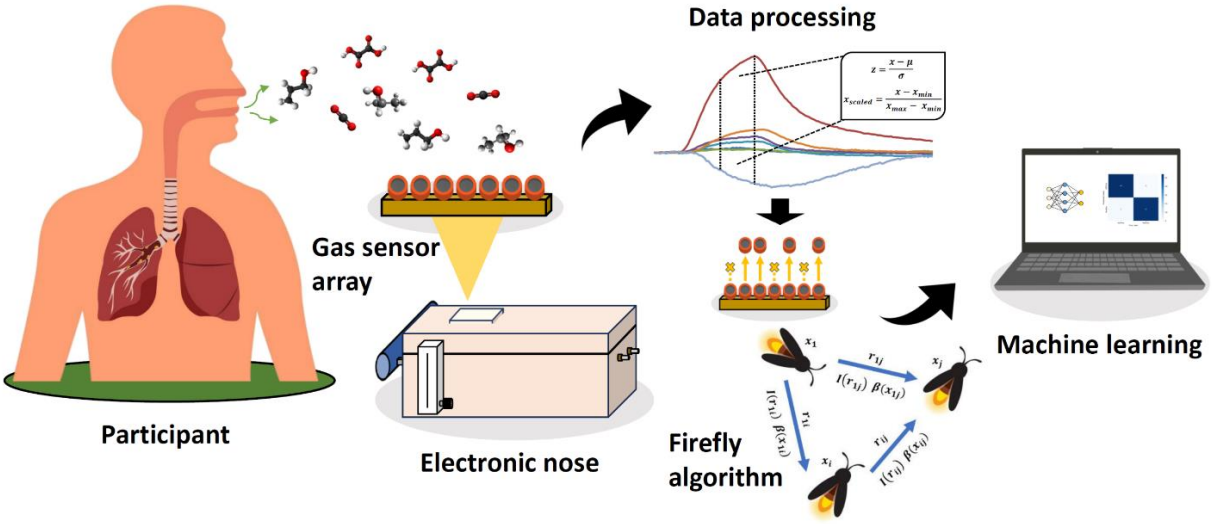
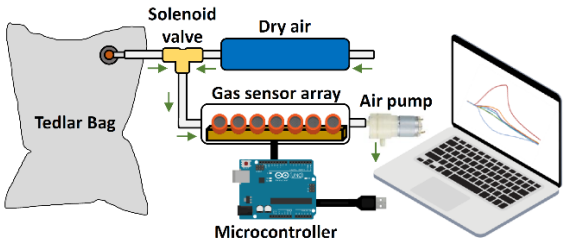


Figure. 1 The proposed research diagram



(a)

(b)

Figure. 2 The electronic nose setup in the experiment: (a) systematic diagram and (b) implementation

principal component analysis (PCA) technique is applied as a final comparison. Model evaluation includes accuracy, precision, recall, and F1 score.

2.2 Electronic nose

Fig. 2 depicts the electronic nose setup, which consists of an array of gas sensors, a microcontroller, and a computer. The electronic circuit has been designed using a printed circuit board (PCB) to reduce noise. Metal oxide semiconductor gas sensors were chosen for their benefits: robustness, fast response, and high sensitivity to specific gases [23]. Seven gas sensors, namely MQ-7, MQ-8, MQ-131, MQ-136, MQ-137, MQ-138, and TGS4161, as listed in Table 1, were employed because they were able to detect gas in exhaled breath effectively [6, 9, 10]. The gas sensor array is allocated to the 240mL chamber to prevent interference from external gas. A flow meter helps monitor gas flow rates.

Table 1. Types of gas sensors on the electronic nose

Sensor	Gas	Manufacturer
MQ-7	Carbon monoxide	Hanwei Electronics Co., Ltd
MQ-8	Hydrogen	
MQ-131	Ozone	
MQ-136	Hydrogen sulfide	
MQ-137	Ammonia	
MQ-138	Volatile organic compounds	
TGS4161	Carbon dioxide	Figaro Engineering Inc.

The sample measurement procedure is explained as follows: First, the sensor is cleaned for 15 seconds using dry air. Then, the gas sample in the Tedlar bag was measured for 40 seconds. Finally, a cleaning step was performed for 95 seconds in preparation for the subsequent sample measurement. The output voltage signal from the gas sensor is converted into digital data and sent to the computer to generate an Excel file. The dataset consists of multiple response curves for data preprocessing and feature scaling analysis.

2.3 Data preprocessing and feature scaling

Data preprocessing consists of forming three datasets, namely binary and multiclass. Each class includes healthy and asthmatic with degrees of severity and is labeled using the one-hot encoding

Table 2. Kernel variations in SVM

Kernel	Formula	Parameter
Linear	$K(x, x_i) = x \cdot x_i$	C
Radial basis function (RBF)	$K(x, x_i) = exp\ (-\gamma\|\|x - x_i\|\|^2)$	C, γ
Sigmoid	$K(x, x_i) = tanh\ (\gamma(x \cdot x_i) + coef0)$	$C, \gamma, coef0$
Polynomial	$K(x, x_i) = (\gamma(x \cdot x_i) + coef0)^d$	$C, \gamma, d, coef0$

Table 3. Arrangement of SVM parameters

Parameter	Explanation
Kernel	Linear, RBF, sigmoid, polynomial
C	1 – 100
γ	0.001 – 1
d	2 – 6
$coef0$	0.0 – 0.3

Table 4. Configurable RF parameters

Parameter	Explanation
Criterion	Entropy, Gini index
Max depth	None
Max features	$sqrt, log_2$
Min samples leaf	1 – 3
Min samples split	2 – 6
Trees	1 – 200

Table 5. Interpretation of XGBoost parameter values

Parameter	Explanation
Max depth	None
Trees	1 – 200
Gamma	0.1 – 1
Learning rate	0.01, 0.05, 0.1, 0.15, 0.2
Min child weight	1 – 4

technique. Feature scaling contains standardization, robust standardization, min-max scaling, and normalizer, which are illustrated in Eq. (1), Eq. (2), Eq. (3), and Eq. (4), respectively.

$$z = \frac{x-\mu}{\sigma} \quad (1)$$

$$z_{robust} = \frac{x-x_{med}}{x_{Q3}-x_{Q1}} \quad (2)$$

$$x_{scaled} = \frac{x-x_{min}}{x_{max}-x_{min}} \quad (3)$$

$$x_{normalized} = \frac{x}{||x||} \quad (4)$$

where x, x_{med}, x_{Q1}, x_{Q3}, x_{min}, x_{max}, and $||x||$ are the data value, median value, first quartile, third quartile, minimum value, maximum value, and $L2$ norm of the x value, respectively. μ and σ are the mean and standard deviation of the features, respectively.

Table 6. ANN layer configuration

Model	Hidden layer			
	1	2	3	4
ANN 1	5	-	-	-
ANN 2	20			
ANN 3	35			
ANN 4	50			
ANN 5	5	5		
ANN 6	20	5		
ANN 7	20	20		
ANN 8	35	20		
ANN 9	35	35		
ANN 10	50	35		
ANN 11	50	50		
ANN 12	5	5	5	
ANN 13	20	5	5	
ANN 14	20	20	5	
ANN 15	20	20	20	
ANN 16	35	20	20	
ANN 17	35	35	20	
ANN 18	35	35	35	
ANN 19	50	35	35	
ANN 20	50	50	35	
ANN 21	50	50	50	
ANN 22	5	5	5	5
ANN 23	20	5	5	5
ANN 24	20	20	5	5
ANN 25	20	20	20	5
ANN 26	20	20	20	20
ANN 27	35	20	20	20
ANN 28	35	35	20	20
ANN 29	35	35	35	20
ANN 30	35	35	35	35
ANN 31	50	35	35	35
ANN 32	50	50	35	35
ANN 33	50	50	50	35
ANN 34	50	50	50	50

2.4 Machine learning algorithms

SVM separates data into individual classes using decision boundaries called hyperplanes [24]. The SVM kernel function influences and constructs decision boundaries [25]. Each kernel has its parameters, including cost function (C), gamma (γ), degree (d), and coefficient (*coef0*), as shown in Table 2, and it can be modified as described in Table 3.

RF utilizes bagging methods, which combine multiple decision trees to construct individual outputs and are assessed using majority voting techniques to produce a conclusion [26]. Gini index and entropy determine the optimal splitting nodes for building the best decision tree [27]. Table 4 shows the RF parameters that can be configurable.

XGBoost is established on the boosting method. This approach improves the performance of a weak classifier by adding new trees to rectify errors made by previous trees [28]. The parameters modified are tree architecture depth, number of trees, gamma, learning rate, and minimum weight for each child, as explained in Table 5.

ANN is uplifted by the working principles of the human brain to create a decision-making system. There are three primary layers, namely input, hidden, and output. Neurons in the hidden layer conduct network computing, starting from the weighted sum calculation step. ANN presents two methods, including feed-forward and back-propagation [29]. Table 6 describes the composition, including the number of hidden layers and neurons. All neurons are fully interconnected and attended by a rectified linear unit (ReLU) activation function. The output layer employs a softmax activation function.

1D-CNN is organized to process one-dimensional sequential data, such as gas sensor response signals [12]. This model uses convolution layers, called conv, which provide filters or kernels to extract and analyze features from the sensor response curve more accurately. This approach permits more profound differences in information regarding healthy and asthmatic subjects [10]. Table 7 describes the CNN architecture composition. The convolution layer is connected to an NN layer with five neurons. To provide an adequate model, the candidate was tested again with different NN neurons, i.e., 20, 35, and 50.

LSTM is conceived to predict information on sequential data utilizing memory cell features that can store long-term collections of information or eliminate irrelevant data [30]. The memory cell scheme, including input, forget, and output gates, accompanied by activation functions, manages the flow of information [31]. Table 8 displays the configuration of LSTM, BiLSTM, and GRU. These algorithms do not involve hidden neural network layers. All computational layers in each model employ ReLU and softmax in the output layer.

BiLSTM can acquire a more comprehensive understanding of the context of sequential data due to its ability to apply bidirectional computing. This model has two separate LSTM layers, including one forward LSTM layer and one backward LSTM layer [32]. This approach captures the context of past and future time steps, making the information apprehended more valuable.

GRU offers update and reset gates to control the flow of information within the unit, allowing it to keep and discard information adaptively based on sequence requirements over various periods [33]. This model has a simpler architecture and fewer parameters, making its computation faster and more efficient [34]. GRU does not annihilate essential features but maintains relevant information through its sequential gate mechanism.

Table 7. CNN's convolution layer setup

Model	Filter		Kernel		Neuron
	Conv 1	Conv 2	Conv 1	Conv 2	
CNN 1	5	-	2	-	5, 20, 35, 50
CNN 2	20				
CNN 3	35				
CNN 4	50				
CNN 5	5		3	-	
CNN 6	20				
CNN 7	35				
CNN 8	50				
CNN 9	5	5	2	2	
CNN 10	20	5			
CNN 11	20	20			
CNN 12	35	20			
CNN 13	35	35			
CNN 14	50	35			
CNN 15	50	50			
CNN 16	5	5	3	2	
CNN 17	20	5			
CNN 18	20	20			
CNN 19	35	20			
CNN 20	35	35			
CNN 21	50	35			
CNN 22	50	50			

Table 8. LSTM, BiLSTM, and GRU formatting

Model	Memory cell
LSTM/BiLSTM/GRU 1	5
LSTM/BiLSTM/GRU 2	20
LSTM/BiLSTM/GRU 3	35
LSTM/BiLSTM/GRU 4	50
LSTM/BiLSTM/GRU 5	65
LSTM/BiLSTM/GRU 6	80
LSTM/BiLSTM/GRU 7	95
LSTM/BiLSTM/GRU 8	110
LSTM/BiLSTM/GRU 9	125
LSTM/BiLSTM/GRU 10	140
LSTM/BiLSTM/GRU 11	155
LSTM/BiLSTM/GRU 12	170
LSTM/BiLSTM/GRU 13	185
LSTM/BiLSTM/GRU 14	200

2.5 Hyperparameter tuning using Bayesian optimization

Bayesian optimization plays a role in discovering proper pattern recognition algorithm parameters iteratively and sequentially through black-box

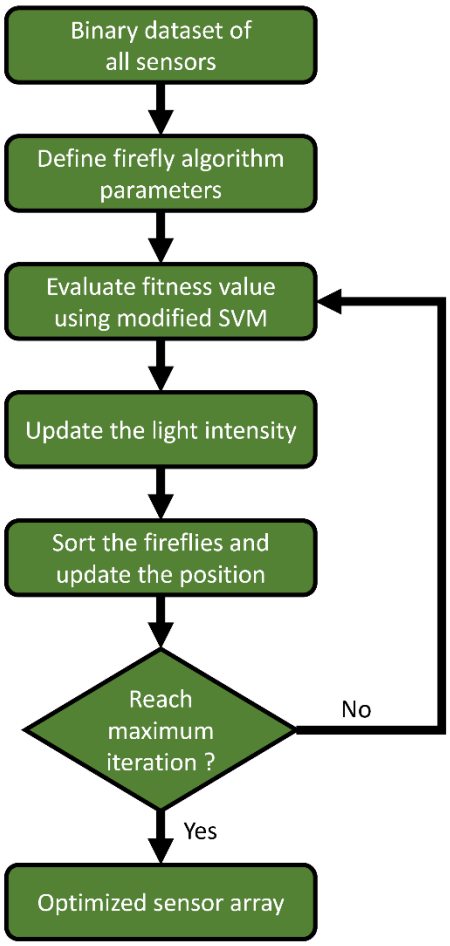


Figure. 3 FA-based sensor array optimization

optimization procedures [35]. This optimization is based on Bayesian and probabilistic inference principles, exploring applicable parameters through the probability distribution of function values at different points in the input space [36]. The surrogate model determines a probabilistic estimate of the next point for the objective function [37]. The acquisition function decides the hyperparameter configuration to evaluate. This optimization is applied to the SVM, RF and XGBoost algorithms.

2.6 Principal component analysis (PCA)

PCA is usually used as a dimensionality reduction technique, converting the original dataset into PCA coordinates that characterize the location of the data in PCA space using an orthogonal linear transformation [12]. The converted data is ranked based on the amount of variance in the data. This transformation is conducted by specifying principal components that are linear combinations of the original characteristics. In this study, PCA converts dataset values and attributes into principal component score value format or PCA dataset.

2.7 Firefly algorithm (FA)

Based on the manners of fireflies, this algorithm utilizes the concepts of attraction, brightness, and movement to efficiently examine and explore optimal solutions in the problem space [18]. There are three principles to explain these fireflies [38]:

1) The gender of the fireflies is unisex, which does not admire gender when close to each other.
2) Attraction specifies the distance significance between fireflies based on the intensity of their respective lights.
3) The degree of brightness of the fireflies is illustrated as a fitness weight that influences the quality of potential solutions.

The calculations of this algorithm are explained as follows:

The intensity of one firefly depends on the distance r with a fixed absorption coefficient value γ, written in Eq. (5).

$$I = I_0 e^{-\gamma r} \tag{5}$$

The function of the attractiveness can be generalized in the form Eq. (6).

$$\beta_{ij}(r) = \beta_0 e^{-\gamma r_{ij}^u}, u \geq 1 \tag{6}$$

Eq. (7) depicts the Euclidean distance between two fireflies i and j at positions x_i and x_j, respectively. x_{ik} is the k^{th} variable from the x_j spatial location points of the i^{th} firefly. The symbol D represents the number of dimensions.

$$r_{ij} = || x_i - x_j|| = \sqrt{\sum_{k=1}^{D}(x_{ik} - x_{jk})^2} \tag{7}$$

The movement of the i^{th} firefly towards the j^{th} brightest firefly is conveyed in Eq. (8).

$$x_i = x_i + \beta_0 e^{-\gamma r_{ij}^u}\left(x_j - x_i\right) + \alpha \in \tag{8}$$

where α is the step size parameter using a random number with the value of $[0, 1]$, and $\in$ is a random number between $[-0.5, 0.5]$.

This study uses FA to eliminate gas sensors based on their fitness. Fig. 3 portrays the working principle of the firefly optimization algorithm. SVM acts as a fitness function, then modified using Bayesian for better fitness quality. The configurations of SVM parameters are explained in Table 3. This step employs a binary dataset of all sensors that has not been standardized and features scaled. The sample

Table 9. Firefly algorithm parameters

Parameter	Explanation
Population (P)	5, 10, 15
Randomness (α)	0.01 – 1
Attractiveness (β)	0.1 – 1
Absorption (γ)	0.1 – 10

Table 10. The number of subjects from the healthy and asthmatic groups

Category	ACT	Total Subjects	Total Data
Healthy	-	30	180
Controlled asthma	25	10	60
Partly-controlled asthma	20 - 24	10	60
Uncontrolled asthma	< 20	10	60
All		60	360

dataset is divided using stratified k-fold cross-validation by dividing each class according to its portion. In this function, there are three parameters, including a number of folds, random state, and shuffle, and they are set to five, None, and True, respectively.

Table 9 is the arrangement of FA optimization parameters. The population size reflects the number of fireflies, which impacts the computational cost of the algorithm. The absorption coefficient is related to the decay of light intensity in the environment to decide how quickly the light intensity diminishes with the distance between the fireflies. The attraction coefficient is the ability of a firefly to attract neighboring fireflies based on the light intensity or fitness value. Randomness presents a measure of the random motion of fireflies to investigate.

3. Experimental results

3.1 Dataset preparation

The number of subjects from the healthy and asthmatic groups is shown in Table 10. The asthma group consists of three degrees of severity assessed by ACT measurements. All these groups can be divided into binary and multiclass datasets. The binary dataset includes both healthy and asthmatic groups. Meanwhile, multiclass assesses all degrees of asthma severity, followed by the healthy group.

The gas sensor response curves for each group are shown in Fig. 4. Each group has a different response pattern, which makes it easy to analyze. Fig. 5 demonstrates the data extraction strategy. This approach begins by selecting the area of the non-overlapping curves before entering a steady state. This region is located in the 45th to 55th-second range. Therefore, six samples were taken, including 45, 47, 49, 51, 53, and 55 data. As a result, the healthy group of 30 subjects was multiplied by six samples to produce 180 data, as explained in Table 10. This phase detours errors in the model training process, including over and underfitting [39].

3.2 Determining an initial classification model

Three machine learning models, including SVM, RF, and XGBoost, were considered to determine the best model based on the highest accuracy. The

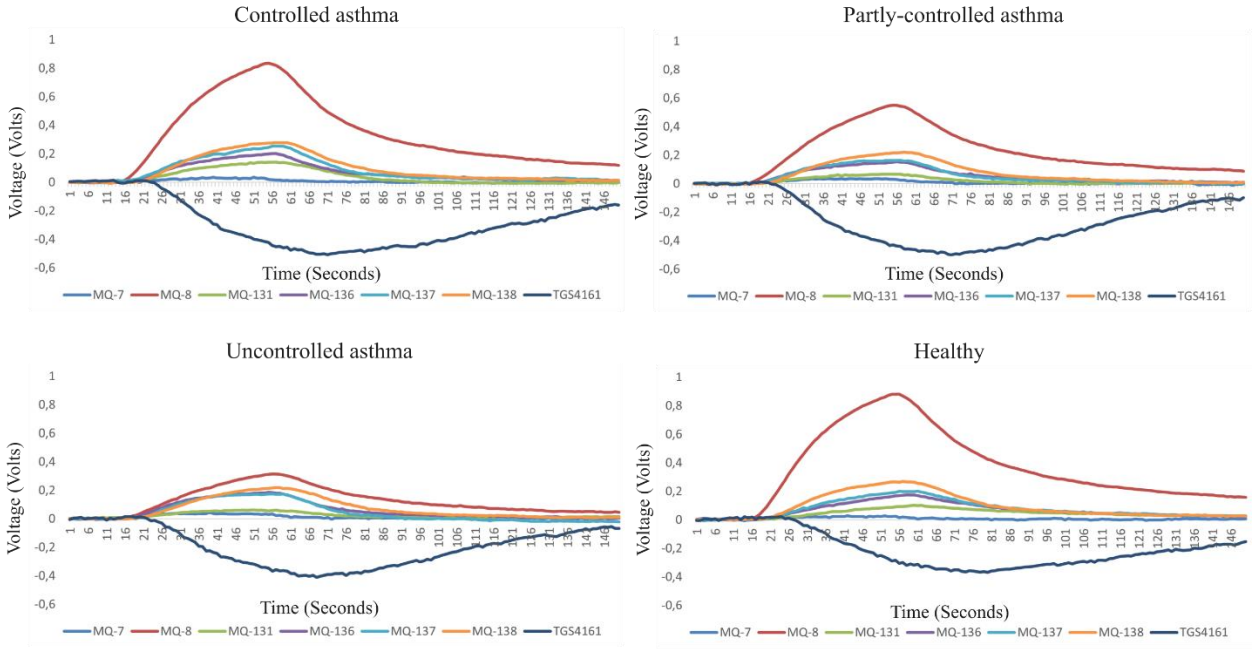


Figure. 4 Gas sensor response characteristics of healthy and asthmatic with degrees of severity

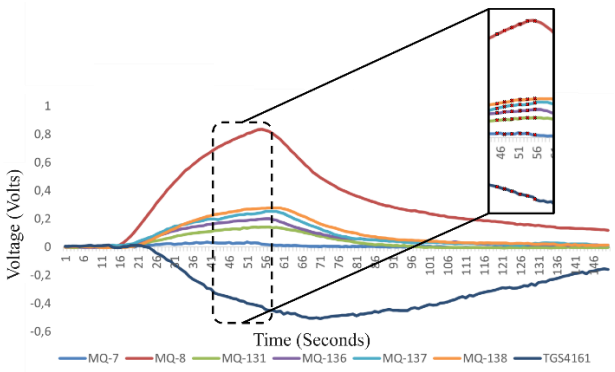


Figure. 5 A sampling strategy that represents a dataset

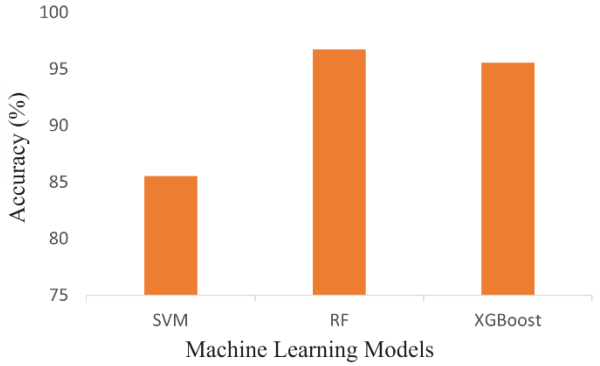


Figure. 6 Average accuracy results from the three models

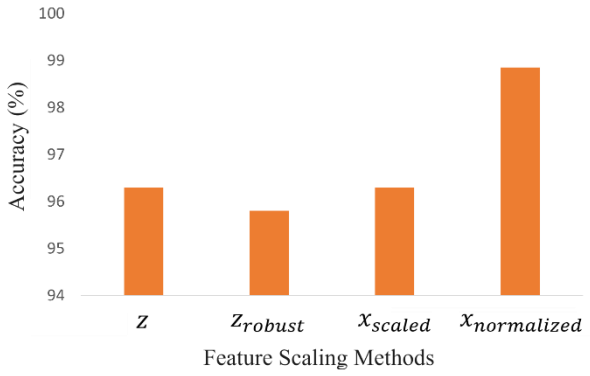


Figure. 7 Comparison of results from all feature scaling techniques using RF

selected model was used for further analysis. Each model is given standard parameters mentioned in the respective libraries. Fig. 6 depicts the average accuracy value of each model. SVM was the lowest ranked, with an average accuracy of 85.55%. At the same time, RF has taken the lead and achieved the highest average accuracy of 96.75%. With its ability to utilize bagging techniques, this model can provide better results compared to other models.

3.3 Determination a feature scaling technique

RF was tasked with defining statistical methods for feature scaling. Fig. 7 is the final result of all statistical methods evaluated by RF. These four techniques provide excellent performance, resulting in accuracy above 95%. The standardization and min-max scaling has a similar accuracy of 96.3%, which shows that they can both scale well. However, the normalization approach broke the highest record, namely 98.85%. Therefore, this technique was chosen because it involves computing the $L2$ norm, which can normalize the dataset's features well.

Table 11. Evaluation results of four SVM kernels

Linear		**RBF**		**Sigmoid**		**Polynomial**	
C	Acc (%)	C, γ	Acc (%)	$C, \gamma, coef0$	Acc (%)	$C, \gamma, coef0, d$	Acc (%)
69	95.4	95, 0.876	94.4	24, 0.136, 0.0	88.9	95, 0.904, 0.2, 2	93.5
79	95.4	81, 0.864	95.4	93, 0.036, 0.1	87.9	85, 0.278, 0.1, 2	91.7
77	95.4	95, 0.966	94,4	95, 0.064, 0.1	79.6	98, 0.85, 0.1, 3	96.3
14	94.4	88, 0.878	93.5	16, 0.196, 0.1	77.8	83, 0.95, 0.3, 3	95,4
72	92.6	85, 0.834	94.4	88, 0.088, 0.0	92.6	97, 0.956, 0.3, 3	94.4
74	94.4	87, 0.27	97.2	3, 0.392, 0.0	66.7	88, 0.826, 0.3, 2	97.2
42	96.3	78, 0.878	96.3	42, 0.108, 0.0	81.5	83, 0.788, 0.2, 2	95.4
88	97.2	72, 0.694	93.5	23, 0.132, 0.0	78.7	80, 0.886, 0.3, 3	89.8
99	93.5	91, 0.842	93.5	91, 0.056, 0.1	93.5	74, 0.962, 0.3, 2	95.4
69	92.6	84, 0.964	97.2	15, 0.186, 0.0	82.4	94, 0.796, 0.3, 4	94.4
Avg	94.7		95.0		82.9		94.4

3.4 Optimizing the firefly algorithm

3.4.1. Building an SVM as an estimator

Based on Fig. 3, the performance ability of the fitness function can impact the type of sensor selected by the FA. Therefore, the SVM parameters need to be modified using Bayesian optimization to provide more accurate fitness results. Data ratio distribution employs the stratified k-fold cross-validation method

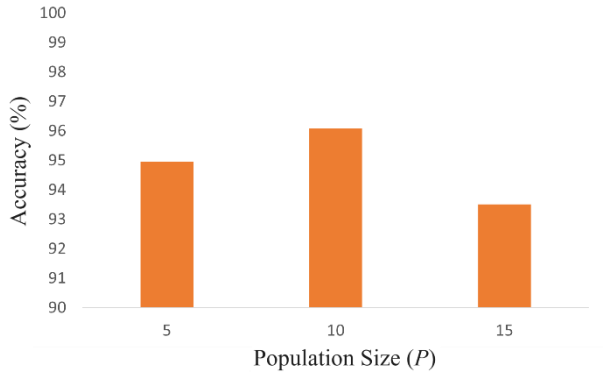


Figure. 8 Average accuracy results from various populations

with k=5, making 80% for training and 20% for testing. The random sampling or shuffle is set to "true". Table 11 shows the highest accuracy results from various kernels. Three kernels, including linear, RBF, and polynomial, occupy positions above 94%. However, RBF was ranked first with the best interpretation value of 95%. As a result, SVM with fewer parameters in the RBF kernel can contribute well. For this reason, the other kernels are omitted. Two sets of RBF parameters from ten experiments presented a similar accuracy of 97.2%. Therefore, the simplest set of parameters is selected, including $C = 84$, $\gamma = 0.964$, to be utilized as the fitness function.

3.4.2. Determining FA parameters

This study determined FA parameters by evaluating population size, randomness, attractiveness, and absorption. The first step assesses the population with numbers 5, 10, and 15. Other parameters are set at 0.1 as a base state. Fig. 8 shows a comparison of the average accuracy results from various populations. Fifteen populations showed lower results, with a value of 93.5%. Therefore, a higher firefly population does not contribute much to enhancing the grade of the solution. As a result, the five and ten were kept and re-evaluated.

The following analysis is to test these populations with differences in β and γ values, as shown in Table 12. It can be seen that the highest performance was 96.08%, which was achieved by seven sets of parameters, including ($P = 5$, $\beta = 0.25$, $\gamma = 1$), ($P = 5$, $\beta = 0.5$, $\gamma = 0.5$), ($P = 10$, $\beta = 0.1$, $\gamma = 0.1$), ($P = 10$, $\beta = 0.25$, $\gamma = 0.75$), ($P = 10$, $\beta = 0.5$, $\gamma = 0.25$), ($P = 10$, $\beta = 0.5$, $\gamma = 0.5$), and ($P = 10$, $\beta = 1$, $\gamma = 0.25$). For this reason, ($P = 5$, $\beta = 0.25$, $\gamma = 1$), was chosen for simplicity, where a low β encourages more exploratory manners in the solution space.

Table 12. Accuracy results of various parameters β and γ

P	α	β	γ	Accuracy (%)
5	0.1	0.1	0.1	94.96
			0.25	94.4
			0.5	94.4
			0.75	94.4
			1	94.96
		0.25	0.1	94.4
			0.25	94.96
			0.5	94.96
			0.75	95.52
			1	96.08
		0.5	0.1	94.4
			0.25	94.96
			0.5	96.08
			0.75	94,4
			1	95.52
		0.75	0.1	94.96
			0.25	95.52
			0.5	95.52
			0.75	94.4
			1	94.96
		1	0.1	94.96
			0.25	91.06
			0.5	95.9
			0.75	95.52
			1	94.96
10		0.1	0.1	96.08
			0.25	95.52
			0.5	94.96
			0.75	95.52
			1	94.4
		0.25	0.1	94.96
			0.25	95.52
			0.5	94.96
			0.75	96.08
			1	94.4
		0.5	0.1	95.52
			0.25	96.08
			0.5	96.08
			0.75	95.52
			1	94.96
		0.75	0.1	94.96
			0.25	94.96
			0.5	94.96
			0.75	95.52
			1	95.52
		1	0.1	95.52
			0.25	96.08
			0.5	95.52
			0.75	92.18
			1	94.96

The next stage is configuring the γ parameter from 1 to 10 to ensure differences in accuracy values. Fig. 9 shows the FA performance with various values

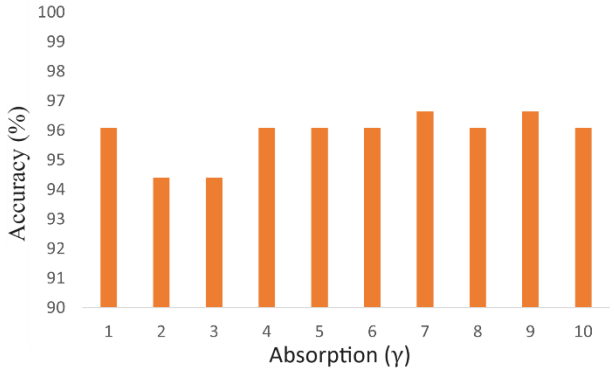


Figure. 9 Influence of differences in parameters γ

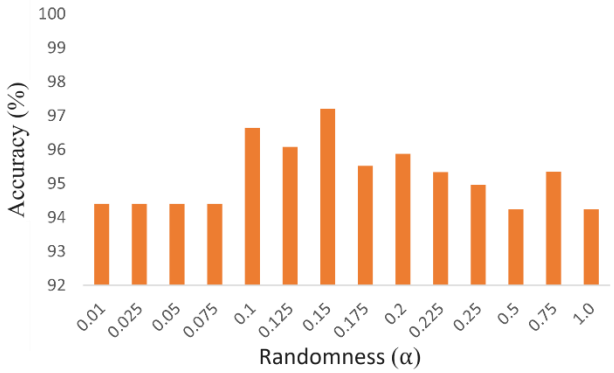


Figure. 10 Variation in performance from different α values

of γ. It can be seen that $\gamma = 7$ and $\gamma = 9$ have the same ability with an accuracy of 96.64%. Therefore, $\gamma = 7$ is chosen because it is smaller. Compared with the previous value of γ, a high γ results in faster convergence to a promising solution. As a result, the current parameters are ($P = 5$, $\alpha = 0.1$, $\beta = 0.25$ and $\gamma = 7$).

The subsequent phase was to modify the α value, beginning from 0.01, 0.025, 0.05, 0.075, 0.1, 0.125, 0.15, 0.175, 0.2, 0.225, 0.25, 0.5, 0.75, and 1. Fig. 10 shows the effect of changing the α value. Low α, i.e., 0.01, 0.025, 0.05, and 0.075, do not give a substantial difference in accurateness. These risks getting stuck in a local optimum, compelling the algorithm to miss out on potentially better solutions. Meanwhile, high α produces inconsistent performance and tends to decrease, inhibiting convergence. However, α in 0.15 significantly improves, bringing the accuracy to 97.2% from 96.64%. As a result, α equal to 0.15 was selected. For this reason, the optimal parameters for FA are ($P = 5$, $\alpha = 0.15$, $\beta = 0.25$ and $\gamma = 7$).

3.5 Gas sensor arrays determination

The optimized FA assigned to decide the appropriate group of gas sensor arrays. Table 13 shows the given group of gas sensor arrays with their

Table 13. The given group of gas sensor arrays with their respective accuracies based on FA

6 Sensors		
No	**Sensors**	**Acc (%)**
1	MQ-8, MQ-131, MQ-136, MQ-137, MQ-138, TGS4161	97.2
2	MQ-8, MQ-131, MQ-136, MQ-137, MQ-138, TGS4161	97.2
3	MQ-8, MQ-131, MQ-136, MQ-137, MQ-138, TGS4161	97.2
4	MQ-8, MQ-131, MQ-136, MQ-137, MQ-138, TGS4161	97.2
5	MQ-7, MQ-8, MQ-131, MQ-136, MQ-138, TGS4161	94.4
5 Sensors		
No	**Sensors**	**Acc (%)**
1	MQ-8, MQ-131, MQ-136, MQ-138, TGS4161	94.4
2	MQ-131, MQ-136, MQ-137, MQ-138, TGS4161	95.3
3	MQ-8, MQ-131, MQ-136, MQ-138, TGS4161	94.4
4	MQ-131, MQ-136, MQ-137, MQ-138, TGS4161	95.3
5	MQ-7, MQ-8, MQ-131, MQ-138, TGS4161	94.4
4 Sensors		
No	**Sensors**	**Acc (%)**
1	MQ-131, MQ-136, MQ-138, TGS4161	94.4
2	MQ-131, MQ-137, MQ-138, TGS4161	94.4
3	MQ-8, MQ-131, MQ-138, TGS4161	94.4
4	MQ-131, MQ-137, MQ-138, TGS4161	94.4
5	MQ-7, MQ-131, MQ-138, TGS4161	94.4
3 Sensors		
No	**Sensors**	**Acc (%)**
1	MQ-131, MQ-138, TGS4161	93.5
2	MQ-131, MQ-137, TGS4161	95.3
3	MQ-8, MQ-131, TGS4161	91.6
4	MQ-131, MQ-138, TGS4161	95.3
5	MQ-131, MQ-138, TGS4161	93.5
2 Sensors		
No	**Sensors**	**Acc (%)**
1	MQ-131, MQ-138	74.0
2	MQ-131, TGS4161	94.4
3	MQ-131, TGS4161	94.4
4	MQ-131, TGS4161	94.4
5	MQ-131, MQ-138	74.0

Table 14. Groups of optimal sensor arrays

Groups	**Sensors**
6	MQ-8, MQ-131, MQ-136, MQ-137, MQ-138, TGS4161
5	MQ-131, MQ-136, MQ-137, MQ-138, TGS4161
4	MQ-131, MQ-137, MQ-138, TGS4161
3a	MQ-131, MQ-137, TGS4161
3b	MQ-131, MQ-138, TGS4161
2	MQ-131, TGS4161

Table 15. RF configuration to determine the optimal number of sensor array

Gini index			
$sqrt$	**Acc (%)**	log_2	**Acc (%)**
Min sample leaf = 1, Min sample split = 2, Trees =199	97.2	Min sample leaf = 1, Min sample split = 4, Trees =141	100
Min sample leaf = 2, Min sample split = 5, Trees =100	98.1	Min sample leaf = 1, Min sample split = 2, Trees =83	96.2
Min sample leaf = 1, Min sample split = 4, Trees =51	100	Min sample leaf = 1, Min sample split = 2, Trees =90	98.1
Average	98.4	Average	98.1
Entropy			
Min sample leaf = 1, Min sample split = 3, Trees =199	97.2	Min sample leaf = 1, Min sample split = 2, Trees =191	100
Min sample leaf = 1, Min sample split = 4, Trees = 96	99.0	Min sample leaf = 1, Min sample split = 5, Trees =177	99.0
Min sample leaf = 1, Min sample split = 5, Trees =105	96.2	Min sample leaf = 1, Min sample split = 4, Trees =91	97.2
Average	97.5	Average	98.7

respective accuracies based on FA. Five pairs of gas sensor arrays in six categorical groups have varying types. On six sensors, the fifth pair, including MQ-7, MQ-8, MQ-131, MQ-136, MQ-138, and TGS4161, was removed due to insufficient accuracy values. On five sensors, the second and fourth pairs have mutually acceptable interpretations. Therefore, other groups were annihilated. The second and fourth pairs provide good results on four, three, and two sensors. However, these two pairs have different sensor types for three sensors, including MQ-131, MQ-137, TGS4161, and MQ-131, MQ-138, TGS4161. Therefore, these arrays were retained for further evaluation. Table 14 is the result of determining the optimal gas sensor array groups.

3.6 Selection of RF configuration to determine the fewest number of gas sensor

The selected RF model needs to be improved with Bayesian optimization. This purpose defines the fewest number of gas sensor while maintaining high accuracy. Standardized binary datasets are involved at this stage. Table 15 illustrates the RF configuration results. Two criteria accompanied by the $sqrt$ and log_2 functions provide an average accuracy above 97%. However, RF with entropy criteria accompanied by log_2 was chosen because it was better, which was then followed by other parameters, including min sample leaf = 1, min sample split = 2, and trees =191, which produced perfect accuracy.

Table 16. Evaluation results of sensor array groups with RF

Groups	**Sensors**	**Acc (%)**
7	MQ-7, MQ-8, MQ-131, MQ-136, MQ-137, MQ-138, TGS4161	98.1
6	MQ-8, MQ-131, MQ-136, MQ-137, MQ-138, TGS4161	99.0
5	MQ-131, MQ-136, MQ-137, MQ-138, TGS4161	96.6
4	MQ-131, MQ-137, MQ-138, TGS4161	95.5
3a	MQ-131, MQ-137, TGS4161	91.5
3b	MQ-131, MQ-138, TGS4161	91.8
2	MQ-131, TGS4161	85.5

Table 16 shows the RF classification results for the entire array of gas sensors. It can be seen that the four sensors have maintained their accuracy above 95%. As a result, the array was reduced as optimal gas sensors, including MQ-131, MQ-137, MQ-138, and TGS4161.

3.7 Modifying of All classification models

After determining the appropriate number of gas sensors, the machine learning must be improved to obtain a robust system. This study modified RF, ANN, 1D-CNN, LSTM, BiLSTM, and GRU with various configurations. This approach is tested on a binary group with an optimal gas sensor array.

For RF, two criteria accompanied by the log_2 provide average accuracy above 95%, with Gini index and entropy of 95.41% and 95.96%, respectively. Therefore, a set of RF parameters, including criterion = entropy, max features = log_2, min samples leaf = 1, min samples split = 2, tress = 37, was selected as the optimal RF that individually provides an accuracy of 99%.

The performance of ANN, 1D-CNN, LSTM, BiLSTM, and GRU is decided based on the architectural modifications described in Tables 6, 7, and 8, respectively. Fig. 11 (a) shows the overall ANN results. It can be seen that simple ANN structures, including ANN 1 to ANN 6, ANN 12, ANN 13, ANN 22 and ANN 23, provide performance below 90%. However, the accuracy can be improved as the number of neurons increases, as in other ANNs. As a result, ANN 32, which involves a more complex network, presented an accuracy of 99.2%. This model was chosen because of its excellent capabilities compared to ANN 33 and 34, which tend to experience overfitting.

The overall CNN architecture assessment results are shown in Fig. 11 (b). For CNN, no model exceeds

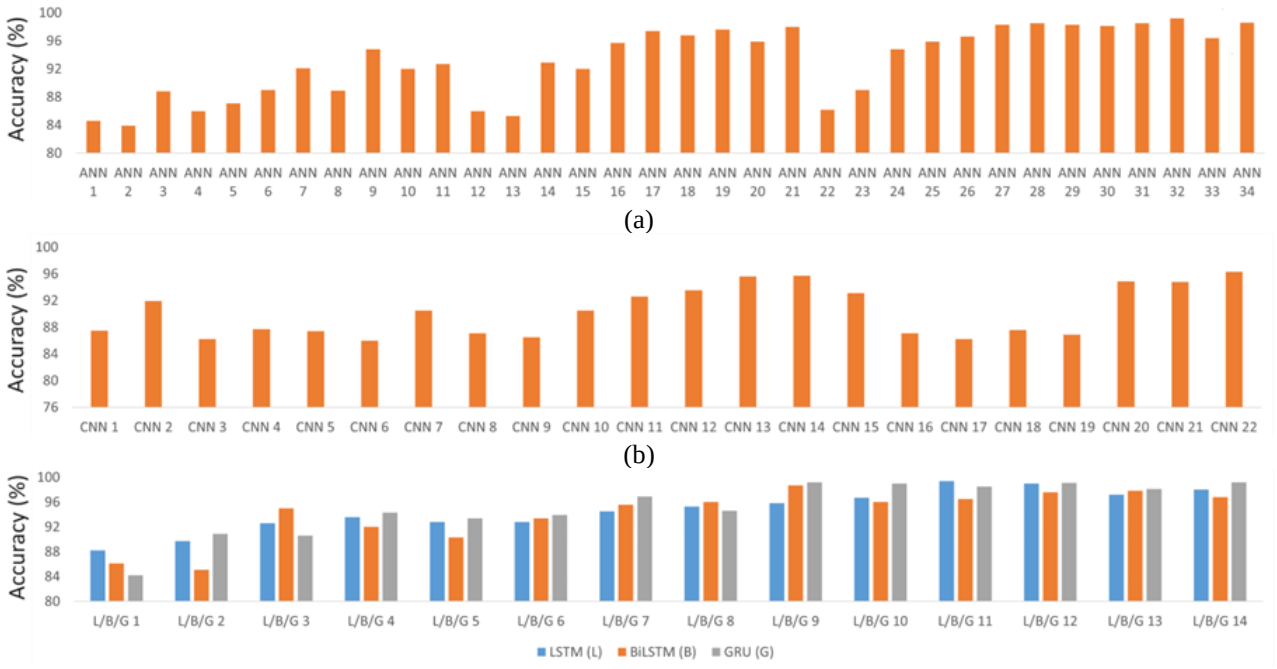


Figure. 11 Performance of: (a) ANN, (b) CNN, and (c) Memory-cell-based architectures

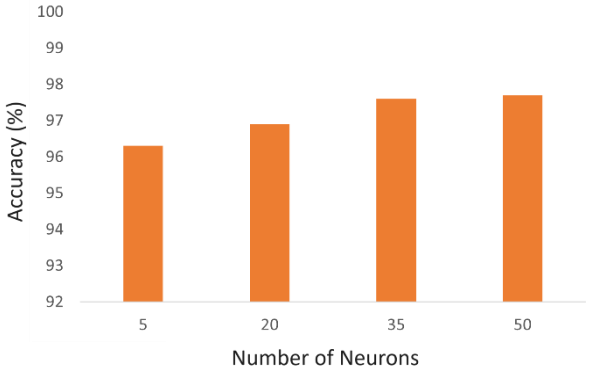


Figure. 12 Enhanced performance on CNN 22

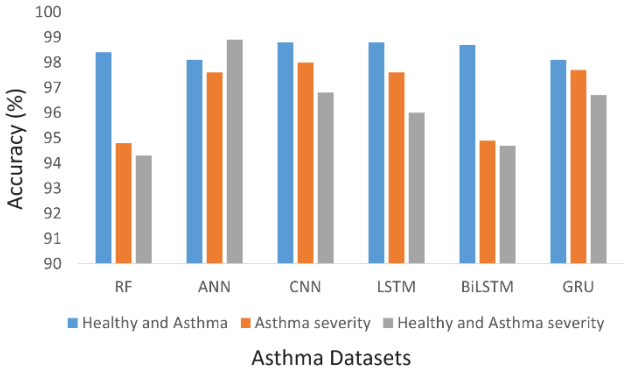


Figure. 13 Performance of the six models on the entire asthma datasets

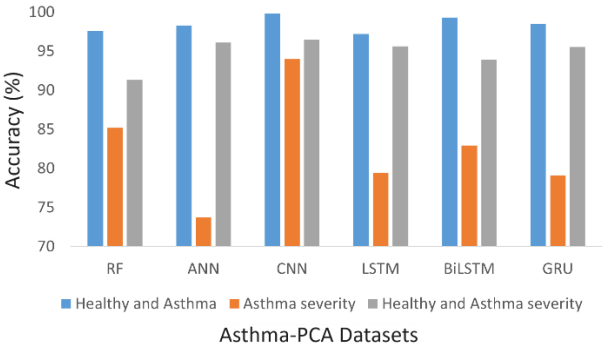


Figure. 14 Evaluation of models on the asthma-PCA datasets

the ANN performances. Meanwhile, CNN 22, the most complex candidate in its class, provides an accuracy of 96.3%. Therefore, CNN 22 was selected and will be necessarily modified to the NN hidden layers described in Table 7 with 20, 35, and 50 neurons. Fig. 12 is the final result of the enhanced performance on the CNN 22 model. By increasing the number of neurons, each neuron provides improved performance to the CNN, where the model with 50 neurons achieved the highest accuracy of 97.7%, an increase from 96.3%. Therefore, the enhanced CNN 22 was confirmed as the selected model.

The results of the modified model based on memory cells are shown in Fig. 11 (c). Most of the architectures of these three models produce accuracy above 90%. LSTM 11 is ranked first with almost perfect results, namely 99.4%. Meanwhile, BiLSTM 9 and GRU 9 obtained accuracies of 98.7% and 99.2%, respectively. These three models were chosen as optimal models. As a result, six selected models were tested in the following analysis.

Table 17. Comparative results from previous related studies

Reference	Dataset	Optimizer	The Number of Sensors	Classifier	Accuracy (%)
[22]	Healthy = 30; Controlled asthma = 10; Partly-controlled asthma = 10; Uncontrolled asthma = 10	-	5	SVM	89.5
[10]		GA	5	CNN	96.6
This work		FA	4	CNN	97.8

3.8 Evaluate the entire asthma datasets

This stage uses an optimal sensor array to evaluate the selected models with the entire asthma dataset. This goal is to decide the best model to be employed as an electronic nose classification algorithm. Fig. 13 shows the performance of the six models on all datasets. In the binary group, all models provide the same accuracy results above 98%. However, CNN showed its capabilities and achieved the highest record of 98.8%. In the multiclass group of three classes, RF and BiLSTM could not give good accuracy, providing accuracy below 95%. However, ANN, CNN, LSTM, and GRU are still the best, with an accuracy of 97.6%, 98%, 97.6%, and 97.7%, respectively. As a result, CNN was selected. Finally, in the four classes, CNN performance tends to decrease compared to ANN performance, which is 97.2% from 98.9%. In these conditions, it is challenging to determine the best model for the electronic nose that still produces high accuracy. Therefore, all models will be re-evaluated with the PCA with three principal components to reduce the dimensionality of the original feature space. The evaluation results are shown in Fig. 14. CNN still performs the best, with an average accuracy of 96.8%. Therefore, CNN was chosen because of its ability to apply convolution layers, which makes the information retrieved more adequate. As a final result, CNN is the preferred model for application in electronic nose systems. The detailed performance of CNN on the asthma dataset includes accuracy of 97.8%, precision of 97.9%, recall of 97.4%, and F1-Score of 97.6%. Table 17 compares previous related studies with the same number and conditions of input datasets, including 30 healthy people and 30 asthma suspects with different severity. This shows that the FA optimization method can provide the smallest array comprising only four gas sensors with the highest accuracy of 97.8%.

4. Conclusion

This study optimized the electronic nose system for classifying asthmatic subjects, including reducing the number of gas sensors and adjusting several machine learning models. The electronic nose was designed with seven metal oxide semiconductor gas sensors. The FA evaluated the optimal number of groups of gas sensor arrays with SVM as a fitness function. The RF determined appropriate sensor arrays by evaluating each group based on the most elevated accuracy. The experimental results demonstrated that the FA with modified RF could generate only four gas sensors, i.e., MQ-131, MQ-137, MQ-138, and TGS4161, which provide accuracy above 95%. Additionally, all classification models presented the best implementation on the asthma dataset, making it more challenging to define a reasonable model for the electronic nose system. Therefore, the PCA technique as dimension reduction was involved in creating the asthma-PCA dataset for the final evaluation step. As a result, CNN still performs the best. Therefore, CNN was preferred as an electronic nose model with the best average accuracy of 97.8%, precision of 97.9%, recall of 97.4%, and F1-Score of 97.6%.

This electronic nose is still developed with a large size, including a sensor chamber and control device systems. For further study, the size of the sensor chamber will be reduced to accelerate the response of the gas sensor. The machine learning model will also be implemented on a single-board computer to deliver a more compact and accessible electronic nose system.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

In this research article author contribution are as follows: Conceptualization and Methodology, Muhammad Rivai; Software, Dava Aulia and Sheva Aulia; Data curation, Dava Aulia; Writing original draft preparation, Sheva Aulia; Writing-review and editing by Muhammad Rivai and Sheva Aulia; Supervision, Muhammad Rivai.

Acknowledgments

This work was supported in part by the Indonesian Ministry of Education, Culture, Research,

and Technology, and Institut Teknologi Sepuluh Nopember.

References

[1] M. A. Awal, M. S. Hossain, K. Debjit, N. Ahmed, R. D. Nath, G. M. M. Habib, M. S. Khan, M. A. Islam, and M. A. P. Mahmud, "An early detection of asthma using BOMLA detector", *IEEE Access*, Vol. 9, pp. 58403–58420, 2021.

[2] D. Kim, S. Cho, L. Tamil, D. J. Song, and S. Seo, "Predicting asthma attacks: Effects of indoor PM concentrations on peak expiratory flow rates of asthmatic children", *IEEE Access*, Vol. 8, pp. 8791–8797, 2020.

[3] T. Mattila, T. Santonen, H. R. Andersen, A. Katsonouri, T. Szigeti, M. Uhl, W. Wąsowicz, R. Lange, B. Bocca, F. Ruggieri, M. K. Gehring, D. A. Sarigiannis, and H. Tolonen, "Scoping review - The association between asthma and environmental chemicals", *Int. J. Environ. Res. Public Health*, Vol. 18, No. 1323, pp. 1–13, 2021.

[4] World Health Organization, *Chronic respiratory diseases : asthma*, 2020.

[5] A. Tiotiu, "Biomarkers in asthma: state of the art", *Asthma Res. Pract.*, Vol. 4, No. 10, pp. 1–10, 2018.

[6] H. Hendrick, R. Hidayat, G. J. Horng, and Z. H. Wang, "Non-invasive method for tuberculosis exhaled breath classification using electronic nose", *IEEE Sens. J.*, Vol. 21, No. 9, pp. 11184–11191, 2021.

[7] M. Maciel, S. Sankari, M. Woollam, and M. Agarwal, "Optimization of metal oxide nanosensors and development of a feature extraction algorithm to analyze VOC profiles in exhaled breath", *IEEE Sens. J.*, Vol. 23, No. 15, pp. 16571–16578, 2023.

[8] L. Savito, S. Scarlata, A. Bikov, P. Carratù, and G. E. Carpagnano, "Exhaled volatile organic compounds for diagnosis and monitoring of asthma", *World J. Clin. Cases*, Vol. 11, No. 21, pp. 4996–5014, 2023.

[9] V. A. Binson, M. Subramoniam, Y. Sunny, and L. Mathew, "Prediction of Pulmonary Diseases with Electronic Nose Using SVM and XGBoost", *IEEE Sens. J.*, Vol. 21, No. 18, pp. 20886–20895, 2021.

[10] D. Aulia, R. Sarno, S. C. Hidayati, and M. Rivai, "Optimization of the electronic nose sensor array for asthma detection based on genetic algorithm", *IEEE Access*, Vol. 11, pp. 74924–74935, 2023.

[11] Misbah, M. Rivai, and F. Kurniawan, "Diabetes detection using carbon nanomaterial coated QCM gas sensors and a convolutional neural network through urine sample", *Int. J. Intell. Eng. Syst.*, Vol. 16, No. 5, pp. 417–427, 2023, doi: 10.22266/ijies2023.1031.36.

[12] Misbah, M. Rivai, F. Kurniawan, Z. Muchidin, and D. Aulia, "Identification of diabetes through urine using gas sensor and convolutional neural network", *Int. J. Intell. Eng. Syst.*, Vol. 15, No. 1, pp. 520–529, 2022, doi: 10.22266/ijies2022.0228.47.

[13] J. Wang, C. Zhang, M. Chang, W. He, X. Lu, S. Fei, and G. Lu, "Optimization of electronic nose sensor array for tea aroma detecting based on correlation coefficient and cluster analysis", *Chemosensors*, Vol. 9, No. 266, pp. 1–20, 2021.

[14] J. Tong, C. Song, T. Tong, X. Zong, Z. Liu, S. Wang, L. Tan, Y. Li, and Z. Chang, "Design and optimization of electronic nose sensor array for real-time and rapid detection of vehicle exhaust pollutants", *Chemosensors*, Vol. 10, No. 496, pp. 1–12, 2022.

[15] T. T. Nguyen, N. V. Quynh, and L. V. Dai, "Improved firefly algorithm: A novel method for optimal operation of thermal generating units", *Complexity*, Vol. 2018, 2018.

[16] J. Wang, M. Zhang, H. Song, Z. Cheng, T. Chang, Y. Bi, and K. Sun, "Improvement and application of hybrid firefly algorithm", *IEEE Access*, Vol. 7, pp. 165458–165477, 2019.

[17] T. Khurshaid, A. Wadood, S. G. Farkoush, C. H. Kim, J. Yu, and S. B. Rhee, "Improved firefly algorithm for the optimal coordination of directional overcurrent relays", *IEEE Access*, Vol. 7, pp. 78503–78514, 2019.

[18] M. Ghasemi, S. K. Mohammadi, M. Zare, S. Mirjalili, M. Gil, and R. Hemmati, "A new firefly algorithm with improved global exploration and convergence with application to engineering optimization", *Decis. Anal. J.*, Vol. 5, No. 100125, pp. 1–18, 2022.

[19] K. A. M. Attia, M. W. I. Nassar, M. B. E. Zeiny, and A. Serag, "Firefly algorithm versus genetic algorithm as powerful variable selection tools and their effect on different multivariate calibration models in spectroscopy: A comparative study", *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.*, Vol. 170, pp. 117–123, 2017.

[20] E. Ergün and O. Aydemir, "Firefly algorithm based feature selection for EEG signal classification", In: *Proc. of 2020 Medical Technologies Congress (TIPTEKNO)*, Antalya, Turkey, pp. 1–4, 2020.

[21] Y. S. Deshmukh, P. Kumar, R. Karan, and S. K.

Singh, "Breast cancer detection-based feature optimization using firefly algorithm and ensemble classifier", In: *Proc. of 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, Coimbatore, India, pp. 1048–1054, 2021.

[22] H. A. Sujono, M. Rivai, and M. Amin, "Asthma identification using gas sensors and support vector machine", *Telkomnika (Telecommunication Comput. Electron. Control.*, Vol. 16, No. 4, pp. 1468–1480, 2018.

[23] D. Xie, D. Chen, S. Peng, Y. Yang, L. Xu, and F. Wu, "A low power cantilever-based metal oxide semiconductor gas sensor", *IEEE Electron Device Lett.*, Vol. 40, No. 7, pp. 1178–1181, 2019.

[24] M. Yalsavar, P. Karimaghaee, A. S. Akbari, M. H. Khooban, J. Dehmeshki, and S. A. Majeed, "Kernel parameter optimization for support vector machine based on sliding mode control", *IEEE Access*, Vol. 10, pp. 17003–17017, 2022.

[25] W. Ding, "SVM-Based feature selection for differential space fusion and its application to diabetic fundus image classification", *IEEE Access*, Vol. 7, pp. 149493–149502, 2019.

[26] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review", *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, Vol. 13, pp. 6308–6325, 2020.

[27] G. Zhang and A. Gionis, "Regularized impurity reduction: accurate decision trees with complexity guarantees", *Data Min. Knowl. Discov.*, Vol. 37, No. 1, pp. 434–475, 2023.

[28] Y. Jiang, G. Tong, H. Yin, and N. Xiong, "A pedestrian detection method based on genetic algorithm for optimize XGBoost training parameters", *IEEE Access*, Vol. 7, pp. 118310–118321, 2019.

[29] M. Rivai, F. Budiman, D. Purwanto, M. S. A. A. Baid, Tukadi, and D. Aulia, "Discrimination of durian ripeness level using gas sensors and neural network", *Procedia Computer Science*, pp. 677–684, 2022.

[30] Y. Liu, Y. X. Huang, X. Zhang, W. Qi, J. Guo, Y. Hu, L. Zhang, and H. Su, "Deep C-LSTM neural network for epileptic seizure and tumor detection using high-dimension EEG signals", *IEEE Access*, Vol. 8, pp. 37495–37504, 2020.

[31] M. U. Abbasi, A. Rashad, A. Basalamah, and M. Tariq, "Detection of epilepsy seizures in neo-natal EEG using LSTM architecture", *IEEE Access*, Vol. 7, pp. 179074–179085, 2019.

[32] L. Xiaoyan and R. C. Raga, "BiLSTM model with attention mechanism for sentiment classification on chinese mixed text comments", *IEEE Access*, Vol. 11, pp. 26199–26210, 2023.

[33] Z. Huang, F. Yang, F. Xu, X. Song, and K. L. Tsui, "Convolutional gated recurrent unit-recurrent neural network for state-of-charge estimation of lithium-ion batteries", *IEEE Access*, Vol. 7, pp. 93139–93149, 2019.

[34] D. M. B. Lent, M. P. Novaes, L. F. Carvalho, J. Lloret, J. J. P. C. Rodrigues, and M. L. Proenca, "A gated recurrent unit deep learning model to detect and mitigate distributed denial of service and portscan attacks", *IEEE Access*, Vol. 10, pp. 73229–73242, 2022.

[35] I. Roman, R. Santana, A. Mendiburu, and J. A. Lozano, "An experimental study in adaptive kernel selection for bayesian optimization", *IEEE Access*, Vol. 7, pp. 184294–184302, 2019.

[36] H. Cho, Y. Kim, E. Lee, D. Choi, Y. Lee, and W. Rhee, "Basic enhancement strategies when using bayesian optimization for hyperparameter tuning of deep neural networks", *IEEE Access*, Vol. 8, pp. 52588–52608, 2020.

[37] G. Sui and Y. Yu, "Bayesian contextual bandits for hyper parameter optimization", *IEEE Access*, Vol. 8, pp. 42971–42979, 2020.

[38] S. Bazi, R. Benzid, Y. Bazi, and M. M. A. Rahhal, "A fast firefly algorithm for function optimization: Application to the control of bldc motor", *Sensors*, Vol. 21, No. 5267, pp. 1–23, 2021.

[39] K. Alomar, H. I. Aysel, and X. Cai, "Data augmentation in classification and segmentation: A survey and new strategies", *J. Imaging*, Vol. 9, No. 46, pp. 1–26, 2023.