

Article

A Fast and Cost-Effective Electronic Nose Model for Methanol Detection Using Ensemble Learning

Bilge Han Tozlu

Department of Electrical Electronics Engineering, Engineering Faculty, Hitit University, 19000 Çorum, Türkiye; bilgehantozlu@hitit.edu.tr

Abstract: Methanol, commonly used to cut costs in the production of counterfeit alcohol, is extremely harmful to human health, potentially leading to severe outcomes, including death. In this study, an electronic nose system was designed using 11 inexpensive gas sensors to detect the proportion of methanol in an alcohol mixture. A total of 168 odor samples were taken and analyzed from eight types of ethanol–methanol mixtures prepared at different concentrations. Only 4 features out of 264 were selected using the feature selection method based on feature importance. These four features were extracted from the data of MQ-3, MQ-4, and MQ-137 sensors, and the classification process was carried out using the data of these sensors. A Voting Classifier, an ensemble model, was used with Linear Discriminant Analysis, Support Vector Machines, and Extra Trees algorithms. The Voting Classifier achieved 85.88% classification accuracy before and 81.85% after feature selection. With its cost effectiveness, fast processing time, and practicality, the recommended system shows great potential for detecting methanol, which threatens human health in counterfeit drink production.

Keywords: feature selection optimization; electronic nose; methanol detection; voting classifier; extra trees classifier



Citation: Tozlu, B.H. A Fast and Cost-Effective Electronic Nose Model for Methanol Detection Using Ensemble Learning. *Chemosensors* **2024**, *12*, 225. <https://doi.org/10.3390/chemosensors12110225>

Academic Editor: Tao Liu

Received: 18 September 2024

Revised: 18 October 2024

Accepted: 21 October 2024

Published: 29 October 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Ethanol is a compound commonly used in alcoholic beverages and widely used as an industrial solvent, fuel, and antiseptic [1]. However, undesirable contaminants such as methanol can lead to serious health risks [2–4]. Methanol is sometimes used by mixing it with ethanol due to its low cost, but this is extremely dangerous [5]. Methanol can be metabolized into formaldehyde and formic acid in the human body, causing toxic effects that can result in blindness, organ failure, and even death [6]. It is stated that the harm of methanol to the human body starts from 10 mL (milliliter) levels and that this dose can cause blindness. However, when the dose increases to 30–240 mL, it can be fatal, depending on the person's body weight, metabolism, and other factors [7].

Consumption of counterfeit alcohol around the world poses a severe problem for public health, and many people die every year for this reason. It is known that the production and, therefore, consumption of fake alcohol is common, especially in low- and middle-income countries. According to the World Health Organization (WHO), approximately 2.6 million people worldwide die every year from causes related to alcohol use. Some of these deaths are caused by the consumption of fake alcohol [8].

Determination of the presence of methanol in ethanol is of critical importance for public health and safety. For this reason, reliable and sensitive analysis methods have been developed to ensure the purity of products containing ethanol. Different determination methods include gas chromatography (GC) [9], high-performance liquid chromatography (HPLC) [10], and techniques such as Fourier transform infrared spectroscopy (FTIR) and RAMAN Spectroscopy [11,12]. These methods are very effective in detecting the presence of low concentrations of methanol in ethanol. In recent years, electronic nose (e-nose) technology has emerged as an innovative tool in chemical analysis. The e-nose consists of an array of sensors that mimic the odor-detecting ability of the human nose. These

sensors are sensitive to different chemical compounds and can identify the characteristic odor profiles of these compounds. E-nose systems can detect the composition of gases by detecting volatile organic compounds (VOCs) and thus can also detect the presence of methanol in ethanol.

The determination of methanol using e-nose technology has many advantages compared to traditional methods. These advantages include its rapid analysis time, low cost, portability, and non-invasive nature. Additionally, e-noses are highly resilient to changes in environmental conditions. These features make e-nose technology ideal for field applications and routine quality control analyses.

In a study on methanol detection using an e-nose, researchers reported detecting methanol in whiskey at concentrations of 1%, 5%, 10%, and 20% using three Metal Oxide Semiconductor (MOS) technology gas sensors they developed. These sensors were made from pure tin oxide (SnO_2), 0.5% carbon nanotube (CNT)- SnO_2 , and 1% CNT- SnO_2 mixtures. In this study, it was stated that the gas phase of the whiskey samples with different concentrations of methanol placed in the olfactory section of the e-nose was transported to the sensor room with a clean air pump and was classified with high classification performance. However, no information was provided on how many samples or how much data were taken, how many classes were classified, or what percentage accuracy was provided. It was stated that it was classified with high classification performance with the (PCA) algorithm [13]. Another study reported that carbon nanotube–titanium dioxide hybrid nanostructures can contribute to classification algorithms in detecting volatile organic compounds (VOCs) [14]. In a similar study on this subject, researchers conducted gas detection experiments of water (H_2O), methanol (MeOH), and ethanol (EtOH) with two different graphene field effect transistor (GFET) sensors they produced (Pristine GFET and ALD- RuO_2 GFET). Their study classified MeOH gases at 10–20–30% concentrations and EtOH gases at 10–20–30% concentrations under 0–20–40–60% humidity conditions using these sensors and a Multilayer Perceptron (MLP) classifier from an artificial neural network. In the study, which did not provide details about the data set, it was stated that three different gas types (water, ethanol, and methanol) were separated from each other with 96.2% accuracy with a Pristine GFET sensor and with 100% accuracy with the ALD- RuO_2 -GFET sensor. However, no performance information was provided regarding the classification of gases at different concentrations. Although it was stated in the study that each sensor conductivity recording period lasted 10 min and 10 data points were taken by recording 1 data point per minute, it was not stated how much data was used in the classification study [15]. However, in another study on this subject, researchers investigated the recognition of excess methanol in alcoholic beverages by gas chromatography and MOS sensors. The research team used four MOS gas sensors based on SnO_2 , Pd-loaded SnO_2 , WO_3 , and Ru-loaded WO_3 , which they specially designed in the laboratory environment and produced with a screen-printing technique. They classified the liquor samples containing 0–10% methanol into two classes and the liquor samples containing 0–10–20–30–40–50% methanol into six classes. In the study, the odor of methanolic beverage samples was not directly detected by the sensors. The smell recording experiment was conducted as follows: 50 μL of the drink sample taken with a pipette tip was placed in a 20 mL headspace sample bottle, where the sample changed from a liquid to gaseous state by being transferred to the sensor chamber by a pump with a flow rate of approximately 300 mL per minute. Data were recorded every second while the sensors were heated and cooled in 30-s cycles between 200 °C and 400 °C for 20 min. In the first classification, 576 samples from two species were taken (480 training, 96 validation) and classified with 94% accuracy. In the second classification, 216 samples from six species were taken (180 training, 36 validation) and classified with 92% accuracy [16].

In previous e-nose studies, the performance of sensors has generally been evaluated by critical metrics such as detection limit, response and recovery times, and stability. For example, one study reported a response time of approximately 20 s for ethanol. However, the recovery time was longer, lasting more than a minute [17]. That study highlights

that sensors exhibit short response times, especially for ethanol. Furthermore, the study evaluated sensor stability with long-term tests, indicating that ethanol detection sensitivity remained at 95% for 21 days. The performance ranges of the sensors in this study also align with the findings in such literature, with both response times and stability consistent with the effective operation of the e-nose system.

Studies have shown that electronic nose systems exhibit high sensitivity to volatile organic compounds (VOCs) such as methanol and that sensors can distinguish between methanol and some other compounds [18]. While these sensors are generally reliable, potential cross-sensitivities between methanol and other substances, like ethanol, may occur. To address this, advanced classification algorithms were applied to ensure accurate differentiation between the compounds. In this study, the classification of eight different ratios of ethanol–methanol mixtures in the home environment with an e-nose obtained with sensors sold at very affordable prices on the market was investigated. The alcohol mixtures prepared in this study were classified by placing them directly in liquid form in a glass into the odor chamber of the e-nose without any processing.

2. Materials and Methods

2.1. Electronic Nose System and Data Recording Procedures

The e-nose used in the study was designed for this study at a very low cost and is also presented as a particularly cost-effective solution to increase access to this technology and encourage wider audiences. The sensor block of the e-nose is designed with MQ brand sensors that can be purchased at very low prices. The sensors used in the sensor block and the types of gas they detect are given in Table 1. To electrically read the changes that the sensors show in the density of the relevant gas, it is necessary to establish a separate measurement circuit for each sensor. All sensors were purchased and used with their kits to alleviate this burden.

Table 1. Used gas sensors in the sensor block.

Sensor No.	Sensor Name	Target Gas
1	MQ-2	Methane, butane, LPG, and smoke
2	MQ-3	Alcohol, ethanol, and smoke
3	MQ-4	Methane and CNG gas
4	MQ-5	Natural gas and LPG
5	MQ-6	LPG and butane gas
6	MQ-7	Carbon monoxide
7	MQ-8	Hydrogen gas
8	MQ-9	Carbon monoxide and flammable gases
9	MQ-131	Ozone
10	MQ-135	Carbon monoxide, ammonia, benzene, alcohol, and smoke (for air quality)
11	MQ-137	Ammonia

For ease of use, the sensor kits are arranged on a perforated circuit board, and after this block was placed in the odor chamber, the relevant connections were taken out of the box through a narrow air-tight hole with a PCI-E Riser Cable. The voltages of the sensors were supplied from a GW Instek brand power supply, and the analog data outputs of the sensors were connected to the analog inputs of two Arduino Uno microcontroller cards. Analog/digital converted sensor data were transferred to a computer via USB cables and saved there. The recording program was prepared in Lab-View (2016). In this software, 10 data points per second are taken and recorded from each sensor. The odor recording

time is 40 s, and a data matrix of size 11×401 is generated after each odor recording. An image of the sensor block and the prepared e-nose system is shown in Figure 1.

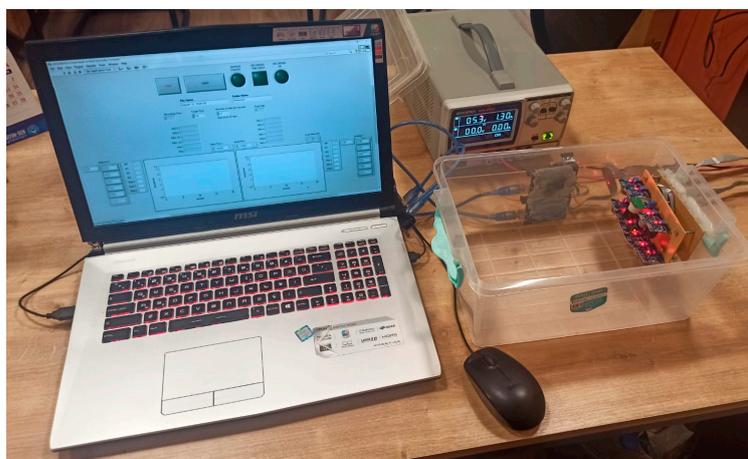


Figure 1. The sensor block, recording software, and the electronic nose system.

In preliminary experiments, it was observed that MOS sensors provide a fast and stable response within 40 s of exposure time. Also, in the literature, it is seen that this type of sensor is used with optimal response times in the range of 30–60 s [16]. To more clearly represent the acquisition of data, a graph of the values obtained from the MQ-137 sensor in the odor recording of a sample is given in Figure 2 below. The graph shows 400 voltage values taken in 40 s. Since the MQ-137 gas sensor detected ammonia gas, the graph shows the change in ammonia gas over 40 s.

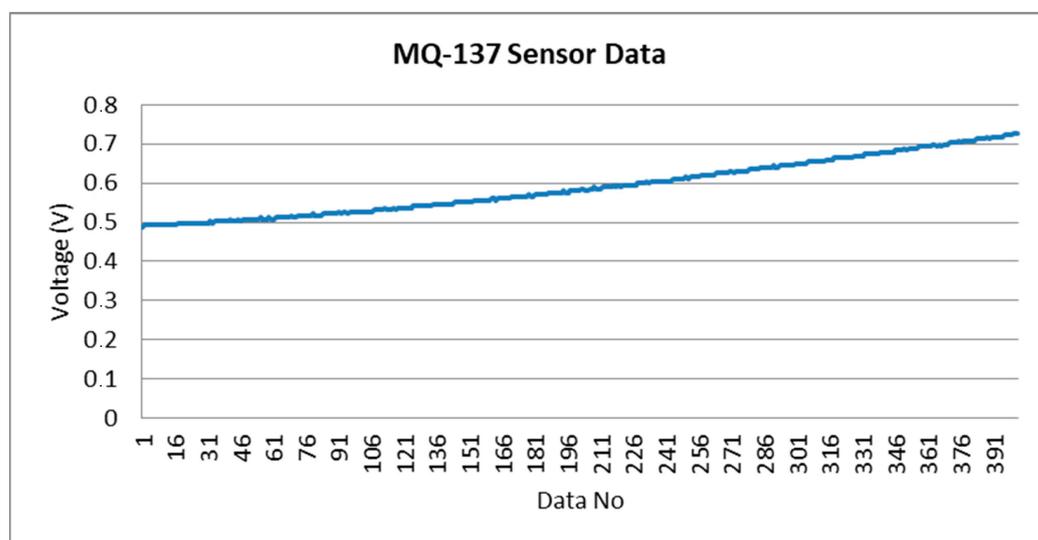


Figure 2. Output values of the MQ-137 sensor during a 40 s recording.

During each test, 100 mL of an alcohol mixture in a large-diameter glass beaker was placed in the same spot of the odor chamber where the sensor block was located. The lid of the odor chamber was closed, and each test was performed in the same windless and constant temperature environment. The gases resulting from the evaporation of the mixtures for 40 s reached the sensors passively without the use of any pump or fan. All tests were performed in this way under homogeneous and consistent conditions. In this way, it was ensured that each odor record was taken with the same standards.

2.2. Sample Preparation Process

In this study, eight different ratios of ethanol–methanol mixtures were prepared. These mixtures were prepared at the Hitit University Faculty of Engineering Chemical Engineering Laboratories. The ratios of ethanol–methanol mixtures are listed in Table 2.

Table 2. Concentrations of ethanol–methanol mixtures.

No.	Sample Name	Concentration Ratios (%)
1	Ethanol	0 methanol–100 ethanol
2	MeOH 10–EtOH 90	10 methanol–90 ethanol
3	MeOH 20–EtOH 80	20 methanol–80 ethanol
4	MeOH 30–EtOH 70	30 methanol–70 ethanol
5	MeOH 40–EtOH 60	40 methanol–60 ethanol
6	MeOH 50–EtOH 50	50 methanol–50 ethanol
7	MeOH 60–EtOH 40	60 methanol–40 ethanol
8	Methanol	100 methanol–0 ethanol

The values given in Table 2 represent the rates. As stated in Section 1, since the amount of methanol that threatens human health starts from 10 mL, the mixtures were placed in the scent chamber of the e-nose in 100 mL glasses. Thus, the amount of methanol in mixture number 2, which contained methanol with the lowest ratio, was set at the limit level of methanol that begins to threaten human health.

Twenty-one odors were recorded from each of the samples. As a result, $21 \times 11 \times 401$ data points were taken from each sample, and a 3-dimensional data matrix with $168 \times 11 \times 401$ data points in total was obtained.

2.3. Feature Extraction and Feature Selection Method

Feature extraction is one of the most critical steps in a machine learning process. Features are extracted from raw data, which are generally complex, high-dimensional, and contain much information. A single data set is obtained and then expressed with as many values as the extracted feature. Here, obtaining meaningful and informative representations (features) of the data directly affects the model's performance. The data are often high-dimensional, but not all of the data contribute to the learning capacity of the model. By using more compelling features in the decision mechanism of the model, higher success, less computational cost, and shorter processing time can be achieved. Additionally, obtaining meaningful information from high-dimensional data increases the model's generalizability and prevents overlearning.

Since a tree-based algorithm was used in this study, as explained below, a wealthy feature set was extracted from the raw sensor data in the hope that it would be helpful. By extracting the data's statistical, time, and frequency domain properties, attributes that reflect all the dynamics of the data set were obtained. The extracted features are mean, standard deviation, total, median, minimum, maximum, first quartile (Q1), third quartile (Q3), variance, RMS, skewness, kurtosis, fifth moment, energy, range, harmonic mean, geometric mean, mean absolute deviation, coefficient of variation, zero crossing rate, interquartile range (IQR), signal-to-noise ratio (SNR), log variance, Holder mean values and their formulas are given below, respectively:

Here, n is the total number of elements in the data set, μ is the arithmetic mean of the data set, σ is the standard deviation, ϵ is a tiny constant added to prevent 0 in logarithm operations, and p is the force used in the Holder mean formula.

Mean	$\mu = \frac{1}{n} \sum_{i=1}^n x_i$	(1)
Standard Deviation	$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$	(2)
Sum	$Sum = \sum_{i=1}^n x_i$	(3)
Median	$\tilde{x} = x_{(\frac{n+1}{2})}$ $\tilde{x} = \frac{x_{(\frac{n}{2})} + x_{(\frac{n+1}{2})}}{2}$	If n is odd If n is even
Min	$Min = \min(x_1, x_2, \dots, x_n)$	(5)
Max	$Max = \max(x_1, x_2, \dots, x_n)$	(6)
Q1—First Quartile	$Q1 = x_{(\frac{n+1}{4})}$	when the data set is sorted x_1, x_2, \dots, x_n (7)
Q3—Third Quartile	$Q3 = x_{(\frac{3(n+1)}{4})}$	when the data set is sorted x_1, x_2, \dots, x_n (8)
Variance [19]	$Var(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$	(9)
RMS (Root Mean Square) [20]	$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$	(10)
Skewness [19]	$Skewness = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^3$	(11)
Kurtosis [21]	$Kurtosis = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^4 - 3$	(12)
5. Moment (Fifth Moment) [22]	$M_5 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^5$	(13)
Energy [23]	$Energy = \sum_{i=1}^n x_i^2$	(14)
Range	$Range = \max(x) - \min(x)$	(15)
Harmonic Mean [24]	$HarmonicMean = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$	(16)
Geometric Mean [25]	$GeometricMean = \left(\prod_{i=1}^n x_i \right)^{1/n}$	(17)
Mean Absolute Deviation [26]	$MAD = \frac{1}{n} \sum_{i=1}^n x_i - \mu $	(18)
Coefficient of Variation [27]	$CV = \frac{\sigma}{\mu}$	(19)
Zero Crossing Rate [28]	$ZCR = \frac{1}{n-1} \sum_{i=1}^n 1_{[x_i \cdot x_{i+1} < 0]}$	(20)
Interquartile Range (IQR) [28]	$IQR = Q3 - Q1$	(21)
Signal-to-Noise Ratio [29]	$SNR = \frac{\mu}{\sigma}$	(22)
Log Variance [30]	$LogVariance = \log(Var(X) + \epsilon)$	(23)
Holder Mean [31]	$HolderMean = \left(\frac{1}{n} \sum_{i=1}^n x_i^p \right)^{1/p}$	(24)

When more features are provided to classification algorithms, the ability of the tree-type classification algorithm to make accurate predictions generally increases because every new piece of information makes the differences between data points more apparent. However, using too many features may not always be the most efficient approach because each additional feature increases the processing load and cost, negatively affecting the model’s performance. Therefore, optimizing the use of the least possible amount of input information and features is necessary without reducing the classification performance too much. This approach creates more efficient and faster models free from unnecessary data, thus reducing processing time and cost [32].

The features used are carefully selected from both statistical and signal-processing metrics to reflect different aspects of the signal. For example, features such as zero crossing rate, the signal-to-noise ratio, and kurtosis are quite successful in capturing the dynamic structure and distribution of the signal. In particular, these features, which represent the signal’s amplitude, spread, and shape properties, provide a comprehensive treatment of the data from different perspectives.

Feature selection is a critical step in optimizing the performance of machine learning models and reducing processing costs. Methods such as Recursive Feature Elimination (RFE) generally identify the most compelling features in data sets and purify the model from unnecessary information [33]. This study used a feature selection method based on feature importance. By calculating the usage rates of the features used by the classifiers in the classification algorithm, the features used at a rate above the reference value were selected, the classification model was re-trained using these features, and the classes of the test data were determined [34]. This way, unnecessary or low-contributing features were eliminated, and the model was rebuilt with less input information. Thus, while the model's performance was maintained, its efficiency was increased, the classification process was completed in a shorter time and with less processing load, and both the cost and processing time were significantly reduced.

2.4. Classification Process

Features extracted from the data were classified using Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Extra Trees classification algorithms frequently used in classification problems, and the Voting Classifier algorithm created by combining these classifiers.

The LDA classification algorithm is a supervised learning algorithm used in classification problems. It is a classifier that minimizes the intra-class variance among the data and maximizes the separation between classes. LDA tries to separate classes by drawing a linear line and defining the boundaries for each class using the mean and covariance of the data set in classification. Therefore, it is effective on low-dimensional data [35].

Support Vector Machines (SVM) is a robust machine learning algorithm frequently used in classification problems. SVM finds the most appropriate separation plane that allows the best separation of data into classes by trying to maximize the most comprehensive distance (margin) between two classes. Although it is effectively used in linearly separable data, it can also be used in non-linear data by transforming it into linear form in a high-dimensional space using appropriate kernel functions. SVM classifier, which performs better in small data sets, minimizes overfitting by trying to find the widest margin and separating the data with a general model [36,37].

The Extra Trees "Extremely Randomized Trees" classification algorithm is a derivative of the Random Forest algorithm and is an ensemble learning algorithm based on decision trees. This classifier creates a decision tree consisting of many random samples in the data set, and the node splits of each tree are randomly selected cut points. This way, the diversity between trees increases, and the risk of overfitting decreases. It performs exceptionally well on high-dimensional datasets because it can use the information it receives from every tree. However, since it uses many trees, it may require higher computational costs and is less accurate but faster than Random Forest [38].

Ensemble learning methods improve overall performance by taking advantage of the predictive ability of more than one model. Voting Classifier, one of the ensemble learning methods, also makes its classification by combining the prediction decisions of different machine learning algorithms. This model collects each algorithm's predictions and makes its final classification using the majority or weighted vote method [39]. Voting Classifier is an advantageous method, especially for balancing the weaknesses of different model types, and is often preferred to increase the overall accuracy rate. This method aims to achieve better prediction performance by combining the advantages of different classification algorithms [40]. This study used a Voting Classifier that makes decisions with the majority of votes from LDA, SVM, and Extra Trees Classifiers.

Classifier classification accuracy (CA), sensitivity (SE), and specificity (SF) metrics were used to evaluate the performance of the classifiers. Here, CA is calculated as the ratio of correctly classified samples to all. SE and SF are calculated separately for each class, followed by their average. These metrics' mathematical expressions are given in Equations (25)–(29) [41].

$$CA = \frac{CCT}{TT} \quad (25)$$

$$SE_j = \frac{TP_j}{TP_j + FN_j} \quad (26)$$

$$SE = SE_{avg} = \frac{1}{8} \sum_{j=1}^8 SE_j \quad (27)$$

$$SF_j = \frac{TN_j}{TN_j + FP_j} \quad (28)$$

$$SF = SF_{avg} = \frac{1}{8} \sum_{j=1}^8 SF_j \quad (29)$$

Here, SE_j stands for the sensitivity of the j^{th} class. TP_j refers to the number of correctly classified samples of the j^{th} class, and FN_j refers to the number of false-negative classified samples of the j^{th} class. SF_j refers to the specificity of the j^{th} class, TN_j refers to the number of correctly classified negative samples of the j^{th} class, and FP_j refers to the number of false-positive classified samples belonging to the j^{th} class.

In the classification study, 60% of the data set is divided into training data, 20% is validation data, and 20% is test data. The classification process consists of training the classifier with the training data, optimizing the model settings with the validation data, and then classifying the test data. The classification performance is evaluated according to the performance metrics mentioned above. Since a result obtained as an outcome of good or bad matching of the training–validation–test data sets will not fully reflect reality, the classification process was carried out 100 times with randomly selected data sets. The performances of these classifications were averaged and presented with standard deviation values. Figure 3 shows the classification flow chart with selected features and all features. The data analysis for this study was conducted using the Python programming language (version 3.8).

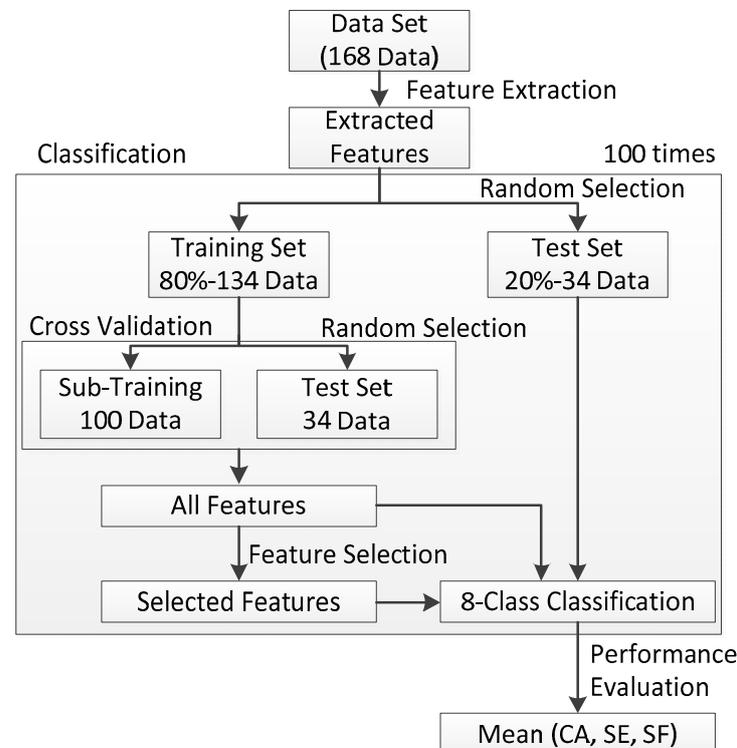


Figure 3. Flow chart of the classification.

3. Results

The classification results are given in two different parts, without feature selection and with feature selection.

3.1. Classification Without Feature Selection

In this study, the odors of eight different ratios of the ethanol–methanol mixtures shown in Table 2 were recorded for 40 s. A data set of $[168 \times 11 \times 401]$ size was obtained after recording 21 odors taken from each of the eight samples with the 11-sensor e-nose. A total of 264 features were obtained from this data set by removing the 24 features described in the feature extraction section. These features are classified using the classification algorithms given in Section 2.4. The mean and standard deviation results of CA, SE, and SF values obtained from the classification results performed on 100 different training–validation–test data sets with four different classification algorithms are given in Table 3.

Table 3. Classification results without feature selection.

Classifier	CA	SE	SF
Ensemble Voting Classifier	0.8588 ± 0.0506	0.8588 ± 0.0509	0.9798 ± 0.0073
Extra Trees Classifier	0.8206 ± 0.0555	0.8213 ± 0.0571	0.9744 ± 0.0080
SVM Classifier	0.8059 ± 0.0629	0.8060 ± 0.0610	0.9722 ± 0.0090
LDA Classifier	0.3539 ± 0.0798	0.3579 ± 0.0828	0.9080 ± 0.0114

Table 3 shows that the Voting Classifier gives the most successful classification performance. The average confusion matrix of 100 classifications of this model as a percentage is presented in Table 4.

Table 4. Confusion matrix of voting classifier without feature selection.

Accuracy: 85.88% MeOH-EtOH%		Predicted							
		0–100%	10–90%	20–80%	30–70%	40–60%	50–50%	60–40%	100–0%
Real	0–100%	97.66	0.00	0.00	0.78	0.00	0.00	0.00	1.56
	10–90%	0.00	78.62	10.69	8.39	0.76	0.00	0.76	0.00
	20–80%	0.00	12.90	74.18	5.65	1.60	0.00	5.65	0.00
	30–70%	0.00	7.87	5.51	86.61	0.00	0.00	0.00	0.00
	40–60%	0.00	0.00	0.78	0.00	79.86	6.98	12.40	0.00
	50–50%	0.00	0.00	0.00	0.00	13.07	84.62	2.31	0.00
	60–40%	0.00	0.00	2.78	4.17	3.47	2.08	87.50	0.00
	100–0%	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

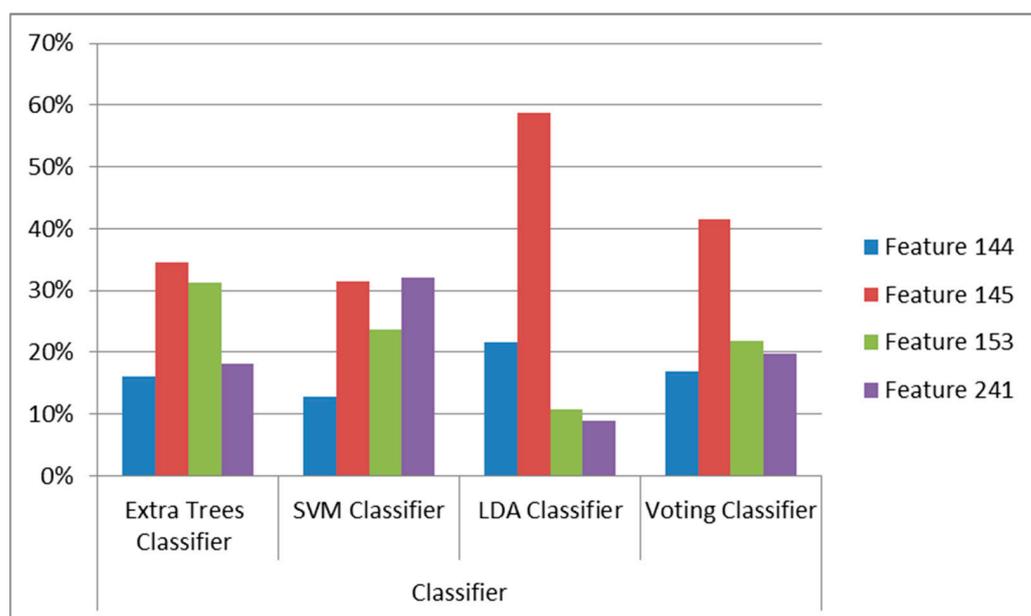
3.2. Optimizing Classification by Feature Selection

Here, a feature selection method based on feature importance was applied. Features used above a specific percentage value were selected, and classification was carried out using only the selected features. According to LDA, SVM, and Extra Trees Classifiers, the importance levels of each feature were calculated separately, threshold values were determined starting from 4%, and classification procedures were carried out. The classification process continued by increasing the threshold value. Table 5 gives the features selected according to the threshold value and the classifier performances for these features. This table clearly shows the effects of different threshold values on the classification performances.

Table 5. Selected features according to threshold values and classifier performances.

Threshold Value	Used Feature Numbers	Voting Classifier CA	Extra Trees Classifier CA	SVM Classifier CA	LDA Classifier CA
4%	22, 24, 26, 27, 29, 30, 32, 127, 130, 143, 144, 145, 146, 148, 149, 150, 151, 153, 231, 233, 241	0.8356 ± 0.0576	0.7524 ± 0.0991	0.7932 ± 0.0676	0.4609 ± 0.2401
5%	22, 26, 32, 143, 144, 145, 146, 148, 149, 150, 151, 153, 233, 241	0.8297 ± 0.0616	0.7576 ± 0.0973	0.7924 ± 0.0676	0.4435 ± 0.2459
6%	143, 144, 145, 146, 148, 151, 153, 241	0.8221 ± 0.0610	0.7518 ± 0.0913	0.7959 ± 0.0635	0.4421 ± 0.2337
7%	144, 145, 153, 241	0.8185 ± 0.0725	0.7582 ± 0.0865	0.7718 ± 0.0679	0.4376 ± 0.2271
8%	144, 145	0.7565 ± 0.1642	0.7138 ± 0.1643	0.7174 ± 0.1358	0.3768 ± 0.1443

The classifiers were re-trained with features that were used above the threshold values, and the test data were classified using these features. Figure 4 shows the usage percentages of the features of the classifiers for the 7% threshold value. Since a Voting Classifier decided by majority vote was used in this study, the usage rate of the Voting Classifier was calculated as the arithmetic average of the usage rates of LDA, SVM, and Extra Trees Classifiers.

**Figure 4.** Usage rates of selected features (%).

4. Discussion

In this study, the sensor data of alcohol mixtures prepared at different concentrations mixed into ethanol were classified to detect methanol, the alcohol used in producing illicit alcohol, by an electronic nose just using smell. In the classification study, 24 different feature extraction methods were applied to the data obtained from 11 gas sensors, and 264 features were obtained. As a result of the classification performed using LDA, SVM, Extra Trees Classifiers, and a Voting Classifier created together, the Voting Classifier offered the highest classification performance with a CA of 0.8588 with all of these features.

Tree-based classifiers increase classification performance by using almost all of the features available to them to a greater or lesser extent. However, using 264 features here is quite troublesome in terms of both cost and processing load and time because it is necessary

to use 11 gas sensors for such a detection. As a result of the feature selection optimization applied, Table 6 was prepared to show how much of a burden the system has been relieved from without reducing the performance too much. Table 6 also gives the usage rates of features and sensors in the classification process carried out with the selected features. When the feature types used in the classification of odor signals obtained from gas sensors were evaluated, the energy feature was calculated by taking the sum of the squares of the data of the relevant sensors, and the SNR feature was calculated by dividing the average value of the sensor data by the standard deviation.

Table 6. Usage percentages of features and sensors according to voting classifier.

	MQ-2	MQ-3	MQ-4	MQ-5	MQ-6	MQ-7	MQ-8	MQ-9	MQ-131	MQ-135	MQ-137
Mean	F0	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Standard Deviation	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21
Sum	F22	F23	F24	F25	F26	F27	F28	F29	F30	F31	F32
Median	F33	F34	F35	F36	F37	F38	F39	F40	F41	F42	F43
Min	F44	F45	F46	F47	F48	F49	F50	F51	F52	F53	F54
Max	F55	F56	F57	F58	F59	F60	F61	F62	F63	F64	F65
Q1—First Quartile	F66	F67	F68	F69	F70	F71	F72	F73	F74	F75	F76
Q3—Third Quartile	F77	F78	F79	F80	F81	F82	F83	F84	F85	F86	F87
Variance	F88	F89	F90	F91	F92	F93	F94	F95	F96	F97	F98
RMS (Root Mean Square)	F99	F100	F101	F102	F103	F104	F105	F106	F107	F108	F109
Skewness	F110	F111	F112	F113	F114	F115	F116	F117	F118	F119	F120
Kurtosis	F121	F122	F123	F124	F125	F126	F127	F128	F129	F130	F131
5.Moment (Fifth Moment)	F132	F133	F134	F135	F136	F137	F138	F139	F140	F141	F142
Energy	F143	16.86	41.55	F146	F147	F148	F149	F150	F151	F152	21.87
Range	F154	F155	F156	F157	F158	F159	F160	F161	F162	F163	F164
Harmonic Mean	F165	F166	F167	F168	F169	F170	F171	F172	F173	F174	F175
Geometric Mean	F176	F177	F178	F179	F180	F181	F182	F183	F184	F185	F186
Mean Absolute Deviation	F187	F188	F189	F190	F191	F192	F193	F194	F195	F196	F197
Coefficient of Variation	F198	F199	F200	F201	F202	F203	F204	F205	F206	F207	F208
Zero Crossing Rate	F209	F210	F211	F212	F213	F214	F215	F216	F217	F218	F219
Interquartile Range (IQR)	F220	F221	F222	F223	F224	F225	F226	F227	F228	F229	F230
Signal-to-Noise Ratio	F231	F232	F233	F234	F235	F236	F237	F238	F239	F240	19.72
Log Variance	F242	F243	F244	F245	F246	F247	F248	F249	F250	F251	F252
Holder Mean	F253	F254	F255	F256	F257	F258	F259	F260	F261	F262	F263
TOTAL IMPACT	-	16.86	41.55	-	-	-	-	-	-	-	41.59

As seen in Table 6, the ethanol–methanol mixtures prepared at eight different concentrations were separated with an accuracy of 0.8185 CA using MQ-3, MQ-4, and MQ-137 gas sensors.

The main difference between this study and other similar studies in the literature is the practicality and low cost of the proposed method. By keeping the alcohol mixture placed in a glass in the odor chamber of the e-nose system for only 40 s without any other process, the methanol content in the mixture was successfully detected. While previous studies on the subject in the existing literature required complex sensor arrays and high-cost hardware, the proposed method obtained results using only MQ-3, MQ-4, and MQ-137 gas sensors and a simple Arduino Uno card. This approach offers a significant difference and advantage over other methods in the literature by providing the advantage of low-cost, easy, and fast processing.

One of the most critical limitations of this study is the small number of samples used. However, only a limited number of experiments have been conducted with specific alcohol mixtures, which may raise some questions about the generalizability of the results. Additionally, only MQ-brand gas sensors were used in the e-nose's sensor block; the potential levels of sensitivity and accuracy that could be achieved using a more comprehensive sensor array have not been investigated.

In future studies, the generalizability of the results can be increased by conducting experiments based on different alcohol mixtures and a larger sample group. Furthermore, the accuracy and sensitivity of the system can be improved by using different sensor technologies.

5. Conclusions

In this study, an e-nose system consisting of 11 sensors was established, and 168 odor data points obtained from eight ethanol–methanol mixtures at different concentrations were classified. By applying a feature selection method based on feature importance, only 4 out of 264 features were selected, and the classification process was carried out using only MQ-3, MQ-4, and MQ-137 sensors. In this way, the number of sensors was reduced, and the complexity of the model was diminished, resulting in a faster and more effective classification. LDA, SVM, Extra Trees, and their combination Voting Classifier were used for classification. The most successful classification result was obtained with the Voting Classifier. Before feature selection, 0.8588 CA, 0.8588 SE, and 0.9798 SF performance values were achieved, and after feature selection, 0.8185 CA, 0.8136 SE, and 0.9734 SF performance values were achieved. These results reveal that the Voting Classifier performs better than other classifiers and that feature selection increases the system's overall effectiveness, although it slightly reduces classification performance. The proposed system demonstrates the success of the feature selection and classification methods applied here with its accuracy, cost-effectiveness, practicality, and fast processing time. The proposed system offers a feasible and effective solution with significant potential for detecting methanol, which is threatening human health.

Funding: This research received no external funding.

Institutional Review Board Statement: This study did not involve humans or animals.

Informed Consent Statement: Not applicable. This study did not involve humans.

Data Availability Statement: The data supporting the reported results are part of ongoing research and are not publicly available at this time. The data may be made available upon reasonable request after the completion of further studies.

Acknowledgments: I would like to express my gratitude to İbrahim Bilici from the Department of Chemical Engineering, Faculty of Engineering, Hitit University, for his assistance in the procurement of the alcohols used in this study.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Simon Araya, S.; Liso, V.; Cui, X.; Li, N.; Zhu, J.; Sahlin, S.L.; Jensen, S.H.; Nielsen, M.P.; Kær, S.K. A Review of The Methanol Economy: The Fuel Cell Route. *Energies* **2020**, *13*, 596. [[CrossRef](#)]
2. Doreen, B.; Eyu, P.; Okethwangu, D.; Biribawa, C.; Kizito, S.; Nakanwagi, M.; Nguna, J.; Nkonwa, I.H.; Opio, D.N.; Aceng, F.L.; et al. Fatal Methanol Poisoning Caused by Drinking Adulterated Locally Distilled Alcohol: Wakiso District, Uganda, June 2017. *J. Environ. Public Health* **2020**, *2020*, 5816162. [[CrossRef](#)] [[PubMed](#)]
3. Tephly, T.R.; McMartin, K.E. Methanol Metabolism and Toxicity. In *Aspartame*; CRC Press: Boca Raton, FL, USA, 1984; ISBN 978-1-00-306528-9.
4. Nekoukar, Z.; Zakariaei, Z.; Taghizadeh, F.; Musavi, F.; Banimostafavi, E.S.; Sharifpour, A.; Ebrahim Ghuchi, N.; Fakhar, M.; Tabaripour, R.; Safanavaei, S. Methanol Poisoning as a New World Challenge: A Review. *Ann. Med. Surg.* **2021**, *66*, 102445. [[CrossRef](#)] [[PubMed](#)]
5. Charapitsa, S.; Sytova, S.; Kavalenka, A.; Sabalenka, L.; Zayats, M.; Egorov, V.; Leschev, S.; Melsitova, I.; Vetokhin, S.; Zayats, N. Intelligent Use of Ethanol for the Direct Quantitative Determination of Methanol in Alcoholic Beverages. *J. Food Compos. Anal.* **2022**, *114*, 104772. [[CrossRef](#)]
6. Tonezzer, M.; Bazzanella, N.; Gasperi, F.; Biasioli, F. Nanosensor Based on Thermal Gradient and Machine Learning for the Detection of Methanol Adulteration in Alcoholic Beverages and Methanol Poisoning. *Sensors* **2022**, *22*, 5554. [[CrossRef](#)]
7. Hovda, K.E.; McMartin, K.; Jacobsen, D. Methanol and Formaldehyde. In *Critical Care Toxicology: Diagnosis and Management of the Critically Poisoned Patient*; Brent, J., Burkhart, K., Dargan, P., Hatten, B., Megarbane, B., Palmer, R., White, J., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 1769–1786. ISBN 978-3-319-17900-1.
8. Manning, L.; Kowalska, A. Illicit Alcohol: Public Health Risk of Methanol Poisoning and Policy Mitigation Strategies. *Foods* **2021**, *10*, 1625. [[CrossRef](#)]
9. Dural, E.; Karakus, A.; Aliyev, V.; Kayaalti, Z.; Yalçın, S.; Kaya, S.; Mergen, G.; Söylemezoğlu, T. *Simultaneous Headspace-GC–FID Analysis for Methanol and Ethanol in Blood, Saliva, and Urine: Validation of Method and Comparison of Specimens*; LCGC North Am.: Iselin, NJ, USA, 2010; Volume 28, pp. 540–543.
10. Quintans-Júnior, L.J.; Gandhi, S.R.; Passos, F.R.S.; Heimfarth, L.; Pereira, E.W.M.; Monteiro, B.S.; dos Santos, K.S.; Duarte, M.C.; Abreu, L.S.; Nascimento, Y.M.; et al. Dereplication and Quantification of the Ethanol Extract of *Miconia Albicans* (Melastomaceae) by HPLC-DAD-ESI-/MS/MS, and Assessment of Its Anti-Hyperalgesic and Anti-Inflammatory Profiles in a Mice Arthritis-like Model: Evidence for Involvement of TNF- α , IL-1 β and IL-6. *J. Ethnopharmacol.* **2020**, *258*, 112938. [[CrossRef](#)]
11. Thanasi, V.; Caldeira, I.; Santos, L.; Ricardo-da-Silva, J.M.; Catarino, S. Simultaneous Determination of Ethanol and Methanol in Wines Using FTIR and PLS Regression. *Foods* **2024**, *13*, 2975. [[CrossRef](#)]
12. Boyaci, I.H.; Genis, H.E.; Guven, B.; Tamer, U.; Alper, N. A Novel Method for Quantification of Ethanol and Methanol in Distilled Alcoholic Beverages Using Raman Spectroscopy. *J. Raman Spectrosc.* **2012**, *43*, 1171–1176. [[CrossRef](#)]
13. Wongchoosuk, C.; Wisitsoraat, A.; Tuantranont, A.; Kerdcharoen, T. Portable Electronic Nose Based on Carbon Nanotube-SnO₂ Gas Sensors and Its Application for Detection of Methanol Contamination in Whiskeys. *Sens. Actuators B Chem.* **2010**, *147*, 392–399. [[CrossRef](#)]
14. Shooshtari, M.; Salehi, A. An Electronic Nose Based on Carbon Nanotube -Titanium Dioxide Hybrid Nanostructures for Detection and Discrimination of Volatile Organic Compounds. *Sens. Actuators B Chem.* **2022**, *357*, 131418. [[CrossRef](#)]
15. Hayasaka, T.; Lin, A.; Copa, V.C.; Lopez, L.P.; Loberternos, R.A.; Ballesteros, L.I.M.; Kubota, Y.; Liu, Y.; Salvador, A.A.; Lin, L. An Electronic Nose Using a Single Graphene FET and Machine Learning for Water, Methanol, and Ethanol. *Microsyst. Nanoeng.* **2020**, *6*, 50. [[CrossRef](#)] [[PubMed](#)]
16. Liu, H.; Wu, R.; Guo, Q.; Hua, Z.; Wu, Y. Electronic Nose Based on Temperature Modulation of MOS Sensors for Recognition of Excessive Methanol in Liquors. *ACS Omega* **2021**, *6*, 30598–30606. [[CrossRef](#)]
17. Hu, J.; Ma, H.; Zhou, Y.; Ma, L.; Zhao, S.; Shi, S.; Li, J.; Chang, Y. Gas-Sensing Properties and Mechanisms of 3D Networks Composed of ZnO Tetrapod Micro-Nano Structures at Room Temperature. *Materials* **2024**, *17*, 203. [[CrossRef](#)] [[PubMed](#)]
18. Fong, C.-F.; Dai, C.-L.; Wu, C.-C. Fabrication and Characterization of a Micro Methanol Sensor Using the CMOS-MEMS Technique. *Sensors* **2015**, *15*, 27047–27059. [[CrossRef](#)]
19. Usta, I.; Kantar, Y.M. Mean-Variance-Skewness-Entropy Measures: A Multi-Objective Approach for Portfolio Selection. *Entropy* **2011**, *13*, 117–133. [[CrossRef](#)]
20. Ning, F.; Cheng, Z.; Meng, D.; Wei, J. A Framework Combining Acoustic Features Extraction Method and Random Forest Algorithm for Gas Pipeline Leak Detection and Classification. *Appl. Acoust.* **2021**, *182*, 108255. [[CrossRef](#)]
21. Borman, R.I.; Rossi, F.; Jusman, Y.; Rahni, A.A.A.; Putra, S.D.; Herdiansah, A. Identification of Herbal Leaf Types Based on Their Image Using First Order Feature Extraction and Multiclass SVM Algorithm. In Proceedings of the 2021 1st International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS), Yogyakarta, Indonesia, 15–16 October 2021; pp. 12–17.
22. Mohammed, I.J.; George, D.L.E. Lie Detection and Truth Identification Form EEG Signals by Using Frequency and Time Features. *J. Algebr. Stat.* **2022**, *13*, 4102–4121.
23. Nsaif, Y.M.; Hossain Lipu, M.S.; Hussain, A.; Ayob, A.; Yusof, Y.; Zainuri, M.A.A.M. A Novel Fault Detection and Classification Strategy for Photovoltaic Distribution Network Using Improved Hilbert–Huang Transform and Ensemble Learning Technique. *Sustainability* **2022**, *14*, 11749. [[CrossRef](#)]

24. Jalal, A.; Batool, M.; Kim, K. Stochastic Recognition of Physical Activity and Healthcare Using Tri-Axial Inertial Wearable Sensors. *Appl. Sci.* **2020**, *10*, 7122. [[CrossRef](#)]
25. Takacs, P.; Bourrat, P. The Arithmetic Mean of What? A Cautionary Tale about the Use of the Geometric Mean as a Measure of Fitness. *Biol. Philos.* **2022**, *37*, 12. [[CrossRef](#)]
26. Dessouky, A.M.; Abd El-Samie, F.E.; Hamed, H.F.A.; Salama, G.M. Optimum Model Selection and Statistical Analysis for DNA Sequences. *Nucleosides Nucleotides Nucleic Acids* **2021**, *40*, 808–820. [[CrossRef](#)] [[PubMed](#)]
27. Jiang, M.; Hooper, A.; Tian, X.; Xu, J.; Chen, S.-N.; Ma, Z.-F.; Cheng, X. Delineation of Built-up Land Change from SAR Stack by Analysing the Coefficient of Variation. *ISPRS J. Photogramm. Remote Sens.* **2020**, *169*, 93–108. [[CrossRef](#)]
28. Khan, I.; Choi, S.; Kwon, Y.-W. Earthquake Detection in a Static and Dynamic Environment Using Supervised Machine Learning and a Novel Feature Extraction Method. *Sensors* **2020**, *20*, 800. [[CrossRef](#)] [[PubMed](#)]
29. Sharma, C.; Ojha, C.S.P. Statistical Parameters of Hydrometeorological Variables: Standard Deviation, SNR, Skewness and Kurtosis. In Proceedings of the Advances in Water Resources Engineering and Management; AlKhaddar, R., Singh, R.K., Dutta, S., Kumari, M., Eds.; Springer: Singapore, 2020; pp. 59–70.
30. Khan, R.A.; Rashid, N.; Shahzaib, M.; Malik, U.F.; Arif, A.; Iqbal, J.; Saleem, M.; Khan, U.S.; Tiwana, M. A Novel Framework for Classification of Two-Class Motor Imagery EEG Signals Using Logistic Regression Classification Algorithm. *PLoS ONE* **2023**, *18*, e0276133. [[CrossRef](#)]
31. Altulea, A.H.; Jalab, H.A.; Ibrahim, R.W. Fractional Hölder Mean-Based Image Segmentation for Mouse Behavior Analysis in Conditional Place Preference Test. *Signal Image Video Process.* **2020**, *14*, 135–142. [[CrossRef](#)]
32. Ramos-Pérez, I.; Barbero-Aparicio, J.A.; Canepa-Oneto, A.; Arnaiz-González, Á.; Maudes-Raedo, J. An Extensive Performance Comparison between Feature Reduction and Feature Selection Preprocessing Algorithms on Imbalanced Wide Data. *Information* **2024**, *15*, 223. [[CrossRef](#)]
33. Guyon, I.; Elisseeff, A. An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
34. Alhamidi, M.R.; Jatmiko, W. Optimal Feature Aggregation and Combination for Two-Dimensional Ensemble Feature Selection. *Information* **2020**, *11*, 38. [[CrossRef](#)]
35. Qu, L.; Pei, Y. A Comprehensive Review on Discriminant Analysis for Addressing Challenges of Class-Level Limitations, Small Sample Size, and Robustness. *Processes* **2024**, *12*, 1382. [[CrossRef](#)]
36. Yang, X.; Song, Q.; Cao, A. Weighted Support Vector Machine for Data Classification. *Int. J. Pattern Recognit. Artif. Intell.* **2007**, *21*, 859–864. [[CrossRef](#)]
37. Naser, A.; Aydemir, O. Classification of Pleasant and Unpleasant Odor Imagery EEG Signals. *Neural Comput. Appl.* **2023**, *35*, 9105–9114. [[CrossRef](#)]
38. Krupski, J.; Iwanowski, M.; Graniszewski, W. Extraction of Minimal Set of Traffic Features Using Ensemble of Classifiers and Rank Aggregation for Network Intrusion Detection Systems. *Appl. Sci.* **2024**, *14*, 6995. [[CrossRef](#)]
39. Cui, S.; Han, Y.; Duan, Y.; Li, Y.; Zhu, S.; Song, C. A Two-Stage Voting-Boosting Technique for Ensemble Learning in Social Network Sentiment Classification. *Entropy* **2023**, *25*, 555. [[CrossRef](#)] [[PubMed](#)]
40. Rokach, L. Ensemble-Based Classifiers. *Artif. Intell. Rev.* **2010**, *33*, 1–39. [[CrossRef](#)]
41. Grandini, M.; Bagli, E.; Visani, G. Metrics for Multi-Class Classification: An Overview. Available online: <https://arxiv.org/abs/2008.05756v1> (accessed on 7 September 2024).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.