Accurate detection and discrimination of pollutant gases using a temperature modulated MOX sensor combined with feature extraction and support vector classification

Oussama Djedidi, Mohand Djeziri, Nicolas Morati, Jean-Luc Seguin, Marc Bendahan, Thierry Contaret

Accurate Detection and Discrimination of Pollutant Gases Using a Temperature Modulated MOX Sensor Combined with Feature Extraction and Support Vector Classification

Oussama Djedidi[a,*], Mohand A. Djeziri[a], Nicolas Morati[b], Jean-Luc Seguin[b], Marc Bendahan[b], Thierry Contaret[b]

[a] *Aix-Marseille University, Université de Toulon, CNRS, LIS, Marseille, France*
[b] *Aix-Marseille University, Université de Toulon, CNRS, IM2NP, Marseille, France*

Abstract

Gas detection and discrimination have been, until recently, sensors-specific, with different sensors and techniques used for each of the gases. In this work, we describe a novel approach relying on a single physical sensor in conjunction with data-driven algorithms for detecting the presence of one of the three dangerous gases: CO, NO_2, and O_3 individually or in mixtures. The approach uses a single Metal Oxide (MOX) sensor coupled with two heaters in its hardware part. Then, its software part uses a supervised machine learning model. The sensor is exposed to the different gases and their mixtures and would react accordingly with a change in its electric signals. These raw signals, along with the readings from the heaters, constitute the primary dataset for the discrimination.

To further enhance the classification results, the raw dataset is augmented by calculating several time-domain features of each of the measurements. Then, the features are ranked, and the ones with the best results to solve the classification problem are selected. Once the pretreatment of the data is finished, the selected features are used to train and validate a multi-support vector machine model. Finally, the results showcased in this paper highlight the effectiveness of the proposed approach.

Keywords: Air pollution, Air quality monitoring, Air quality sensor, Feature extraction, Feature selection, Gas detection, Gas sensors, Metal Oxide sensor, Multi-class classification, Support Vector Machines, ReliefF, Smart sensor, Temperature modulation, Time-domain feature extraction

*Corresponding author
Email addresses: oussama.djedidi@lis-lab.fr (Oussama Djedidi), mohand.djeziri@lis-lab.fr (Mohand A. Djeziri), nicolas.morati@im2np.fr (Nicolas Morati), jean-luc.seguin@im2np.fr (Jean-Luc Seguin), marc.bendahan@im2np.fr (Marc Bendahan), thierry.contaret@im2np.fr (Thierry Contaret)

Contents

1. Introduction

The recent and quickening technological and industrial advancements made by man are a double-edged sword. While they helped elevate the quality of life is to its best level, the negatives effects of human activities are also observed environmentally and, more crucially, health-wise [1]. The concerns raised by these side effects have led, in part, to the development of monitoring tools [2, 3, 4] and regulation of air quality [5, 6]. In this work, we undertook the issue of detecting and identifying three of the major pollutant particles in the air, those are carbon monoxide (CO), ozone (O_3) and nitrogen dioxide (NO_2).

The study of air quality and the development of sensors for monitoring pollutants is not limited to its environmental aspect such as the detection of CO/CO_2 [7], NO_2 [8], and methane [9, 10] but extends into the scientific [11] and industrial applications such as the detection of volatile organic compounds (VOC) [12, 13, 14]. Additionally, the research in this area is approached out from several angles. Traditionally,

researchers have focused on the materials used in the sensors. Most of these studies are listed in several reviews [1, 15, 16, 17, 18, 19]. Of the more generalized studies, Alrammouz et al. [1] summarized the advancements made in the field of flexible gas sensors, listing several sensing materials in the process such as carbon nanotubes, conducting polymers, and the type of sensors used in this work Metal Oxides (MOX). The latter—alongside Field-effect transistor (FET) devices—were also the focus of a work by Sharma et al. [16] studying the advancement in sensing H_2, whereas Dey [17] and [18] narrowed their focus into sensors based on Semiconductor Metal Oxide (SOM) and p-type Nickel Oxide, respectively. Alternatively, Yao et al. [20] and Wang [15] analyzed sensors based on metal-organic frameworks (MOFs), with the first work highlighting material design and performance of the sensors [20], and the second showcasing their application and the application of MOF-Based materials.

With the rise of machine learning (ML), data-driven models were combined with the sensors to obtain higher accuracy and achieve multi-gas selectivity [21]. For instance, multicomponent analysis techniques—notably Principal Component Analysis (PCA)—were one of the first techniques used in combination with sensors to monitor air quality [11, 22]. More recently, PCA was also used for the analysis of gas mixtures [23], and the detection of up to three pollutants simultaneously [24]. Besides PCA, Montoliu et al. [25] used multivariate curve resolution to enhance the sensitivity of MOX sensors to different gases.

The complexity of the task, especially in the presence of several gases or mixtures, led the researchers to use more advanced techniques such as neural networks [26] and Bayesian regularization [27]. In cases similar to this study (detection of air pollutants), Chu et al. [28] combined a sensor array with neural networks to detect CO, O_3, and NO_2 amongst other gases. Similarly, Esposito et al. [29] demonstrated a low-cost sensor and used dynamic neural networks to estimate the concentration of these pollutants, while Topalović et al. [30] compared the use of several ML approaches to achieve the same tasks.

The richness of the literature with research exploring different materials and methodologies is proof of the challenges faced in accurately identifying different gases present in gas mixtures, which remains an open issue. Most of the established approaches either focus on a specific pollutant and filter out any interference [9, 31] or try to overcome the challenge of detecting pollutants in gas mixtures by calibrating multiple sensors (sensor arrays) [28, 29, 30, 32, 33].

In this paper, we tackle the problem of identifying multiple gases in mixtures using a single sensor, by combining a temperature modulated MOX microsensor with a data-driven model. The method uses data from the MOX sensor and processes it in 3 steps. The first consists of changing the operating point of the sensors through temperature modulation, in order to increase the available information in the raw dataset. Then, the dataset is further augmented by extracting key temporal features. These features are automatically ranked by the ReliefF algorithm, with the highest-ranked selected by a custom algorithm for the classification step. Finally, Multi-class SVMs are used for the automatic classification of gases as they offer a wide choice of kernel functions allowing for the separation of classes (gases) by hyperplanes of varying

complexity.

Several established works have tackled the problem of discriminating multiple gases using a single sensor. However, they have either exhibited limited accuracy [12] or cannot be confidently generalized due to the weak number of the sample used in testing [34]. The results presented in this work, on the other hand, avoid these shortcomings and showcase the effectiveness of the proposed approach.

This case study focuses on three gases: carbon monoxide (CO), ozone (O_3), and nitrogen dioxide (NO_2). Nevertheless, the discrimination method is intended to be scalable by an automatic selection of new subsequent features depending on the gases, thus increasing the features space that can be used to train new SVM classifiers. Moreover, the wide choice of kernel functions will allow the algorithm to be adapted to the detection and identification of gases in mixtures with a higher number of gases.

This paper is divided into 7 sections. Section 2 is devoted to the description of the test bench, the characteristics of the MOX sensor, and the data acquisition procedure. Then, an overview of the gas discrimination methodology is given in Section 3. In Section 4, the method of augmenting the database by extracting time-domain features is presented. The ranking and selection of relevant features is explained in Section 5, and the gas classification algorithm is presented in Section 6. The results and discussions are detailed and analyzed in Section 7.

2. Sensor description and setup

2.1. The MOX sensor

The detection and identification of the toxic gases in this work are achieved by two components; a microsensor and an ML-Based model. The process starts with the sensor reacting to the presence of a gas with a change in its reading. Then, the model uses the readings to correctly classify the gas.

To generate the measurement, in this study, we rely on a MOX microsensor, that is the result of a newly developed and patented technology by the *Institut Matériaux Microélectronique Nanosciences de Provence* (IM2NP) and manufactured by Nanoz-SAS[1] [35]. The sensor is composed of four detection zones and two heaters. The detection zones can be configured to allow for the use of these sensors either in single or multi-sensor modes.

In this device, the heaters are located with the sensor on a SiO_2 membrane. At first, the heaters and the electrodes are built on a sputtered platinum thin film using photolithography. After which the sensitive Tungsten trioxide (WO_3) layer is deposited through reactive Radio Frequency Magnetron Sputtering. The WO_3 layer is then annealed for two hours at temperatures higher than the operating range for improved nano-crystallization and stability.

[1] An operating exclusive worldwide license has been granted to NANOZ SAS, which develops manufactures and sells the sensors.

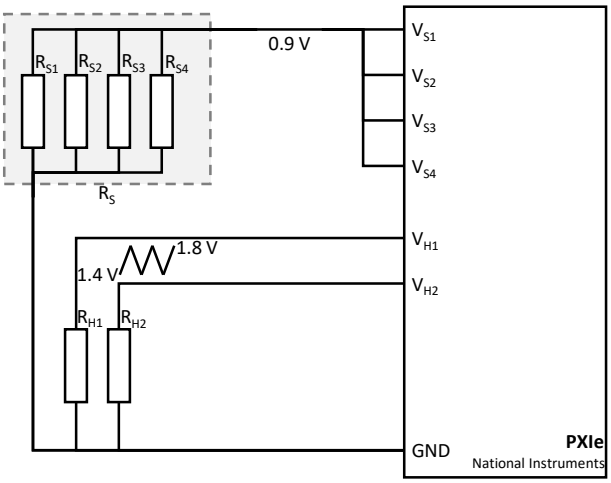


Figure 1: The electric diagram of the MOX sensor [35]. V_{S1}–V_{S4} are the voltages applied to the sensor resistances R_{S1}–R_{S4}, respectively. Similarly, V_{H1} and V_{H2} are the voltage values across heating resistances R_{H1} and R_{H2}, respectively.

The electrical configuration for the MOX sensor is shown in Figure 1. The figure also shows how the device is polarized by a National Instrument PXIe 4140 source meter. In this study, we have chosen to operate the device in a single sensor mode. The sensor is characterized by its resistance R_S and is polarized by 0.9 V [24]. The heaters, on the other hand, are powered by varying voltage—ranging between 1.4–1.8 V, as will explained in the next paragraph.

2.2. Experimental setup and data collection

The characterization and evaluation of the sensors are achieved by exposing it to the studied gases in a controlled environment, to enhances the gas flow hydrodynamics [36]. Therefore, the device is encased in a 3D-printed prototype test chamber that would improve the gas concentration homogeneity and reduce the flow recirculation and dead volumes [37]. The test chamber is then placed in a Faraday cage, in which the sensor is exposed to the different gases to measure its response. For a comprehensive test of the sensor thresholds, the concentration of the tested gases is varied through a dilution system [24]. The latter allows for the concentration of pollutant gas to be adjusted between 20 ppb and 16 ppm, in diluted in dry air under atmospheric pressure and ambient temperature. It is also capable of creating mixtures for a wider range of tests. The concentrations used in this work are displayed in Table 1. They also constitute the base concentrations values for all the possible gas mixtures.

Figure 2 gives an explanatory representation of the experimental setup, where the pollutants are mixed and have their concentration adjusted in the dilution system. The gases are then admitted to the test chamber where the sensor responds with a change in resistance. The response of the sensor as well as the states of heaters are recorded continuously in a dataset composed 7 variables :

- The sensors voltage (V_S, a constant in this study)

Table 1: The different concentrations used for each of the pollutant gases in this study.

Gases	Concentrations		
	C1	C2	C3
CO	2 ppm	8 ppm	16 ppm
NO2	20 ppb	100 ppb	200 ppb
O3	80 ppb	110 ppb	160 ppb

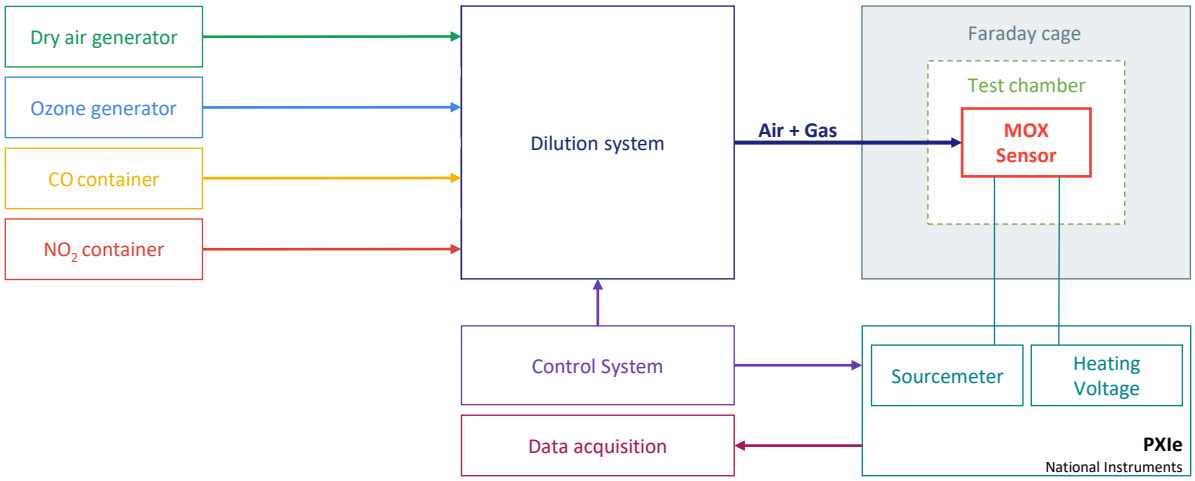


Figure 2: Experimental setup for the microsensor characterization, and the gas classifier training and testing.

- The sensors current (I_S)
- The sensors Resistance (R_S)
- Heater I voltage (V_{H1})
- Heater I current (I_{H1})
- Heater I power (P_{H1})
- Heater II voltage (V_{H2})
- Heater II current (I_{H2})
- Heater II power (P_{H2})

To obtain a larger set of sensor data and generate more accurate results with the classification model, the device is operated with variable heater temperatures, i.e. temperature modulation [19, 38]. In this operating mode, the heating voltage is varied triangularly with a staircase waveform of an amplitude of 0.4 V centered

on 1.6 V. Each step of the staircase signal has a value of ±0.01 V and a duration of 1 s[24]. The range of the heater voltages (1.4–1.8 V) results in temperatures ranging between 199–252 °C which correspond to the edge of detection and the optimal detection temperatures for the studied gases, respectively [24]. This technique has been proven valuable in the literature. It has been used to increase the experimental data, thus extending the available information for processing and feature extraction by the use of wavelet transform [39] or an energy vector approach [32, 40], for instance. On the other hand, it was also used to enhance the sensitivity and selectivity of the sensors to different gas species [41, 42], with very encouraging results in the case of MOX sensors and sensors arrays [23, 25, 31, 43].

3. Gas discrimination methodology

Raw data provided by the sensor reflect the variation the latter undergoes when the different gases are present in the closed chamber. The idea is that each reading of the sensor's resistance and current combined with the power information from the two heaters would correspond to a specific gas or a gas mixture, which leads to the classification problem.

Ideally, the readings from the sensors would lead to a direct—and visual—separation between the different classes. However, as seen in Figure 3, initial tests based solely on the readings from the system establish that the gases strongly overlap with each other, particularly in the cases of gas mixtures ($CO + NO_2$, $CO + O_3$, $NO_2 + O_3$, ...). The overlapping of the classes is a complex issue being studied in the literature [44]. It is even more complex in this work since gas mixtures are themselves overlapping classes of the pure gases. Hence, to simplify the classification problem, each of the gas mixtures will be considered as a distinct class, resulting in 7 classes overall: CO, NO_2, O_3, $CO + NO_2$, $CO + O_3$, $NO_2 + O_3$, and All. In these mixtures, the three levels of concentration indicated in Table 1 are used, to cover all the possible concentrations. For instance, in the case of the CO + NO_2 gas, nine configurations are tested in total, ranging from CO [2ppm] + NO_2 [20ppb] to CO [16ppm] + NO_2 [200ppb].

While considering gas mixtures as distinct classes eliminates one level of the overlapping classes problem, the classification problem is further accentuated by the use of a single sensor to generate readings, thus outlining the minimalism of characteristic data that could help distinguish the classes. To overcome this limitation, one viable solution is to extend the available dataset. This is possible through the extraction of the temporal, frequency, and energy features of the signals from the sensor and heaters.

The newly extracted features would, however, greatly increase the number of available features and it would be computationally inefficient to use all of them in the classification model. Thus, the next steps would be to rank the features according to their influence on the output, and select the feature that would generate the best results. These selected features would then be used to train a supervised ML model. The steps of the proposed approach are illustrated in Figure 4.

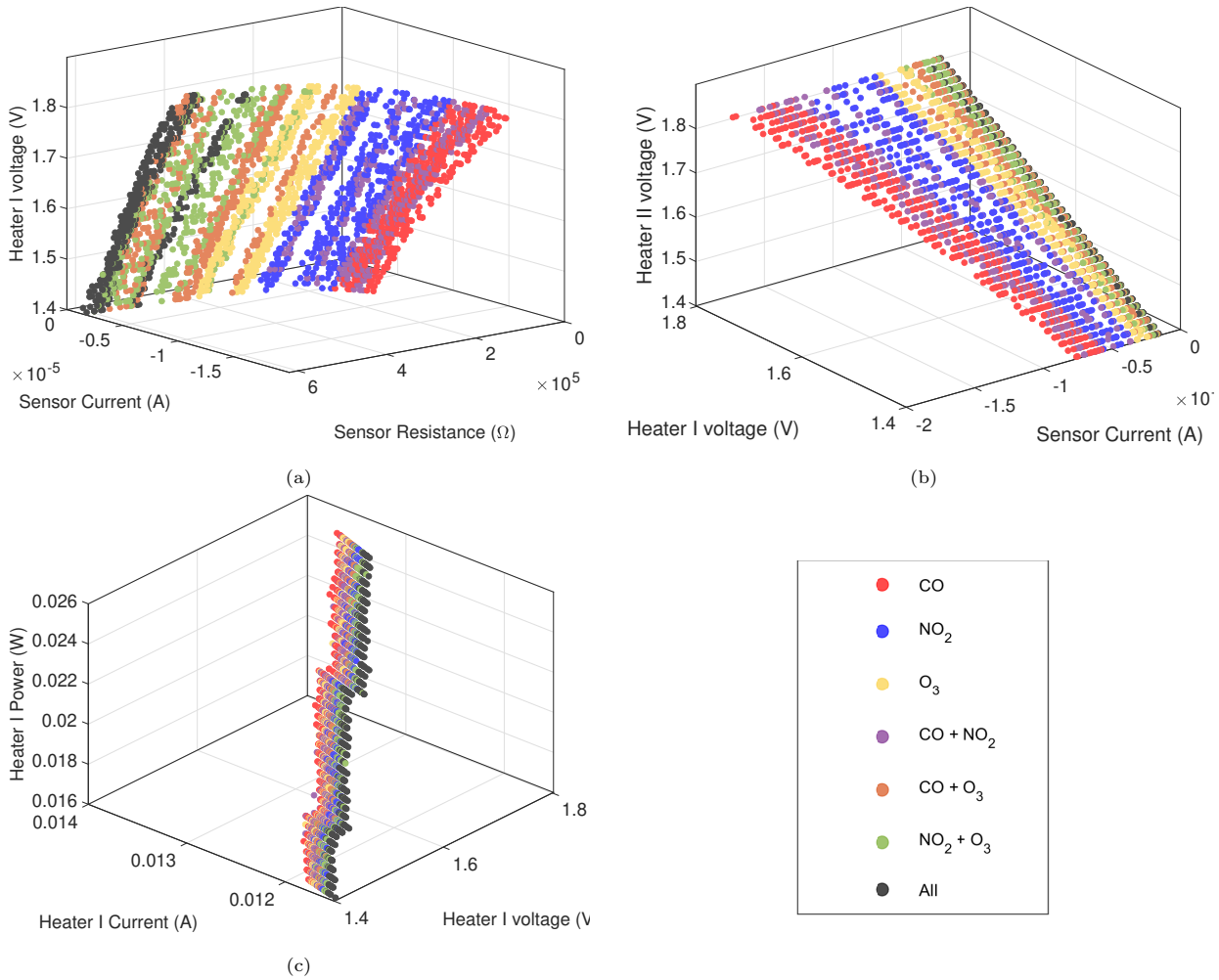


Figure 3: Visual representation of the distribution of the classes according to the initial features. It is impossible to separate the classes linearly.

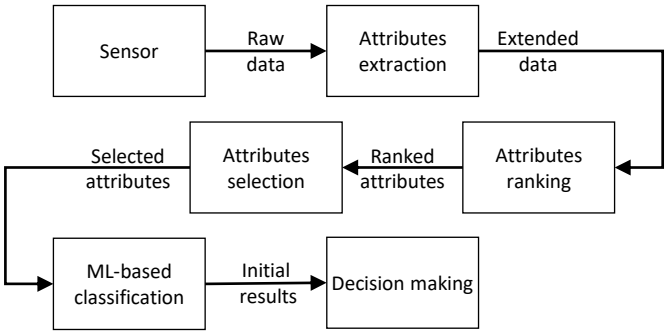


Figure 4: Overview of the gas detection and identification process.

4. Dataset augmentation

Attributes or feature extraction aims to manipulate the data to obtain further information that is not carried by the raw values of the signal but rather by its tendencies, distributions, and frequency-domain characteristics, amongst others [45].

The type and number of features to use depend heavily on the nature of the raw signals. For instance, periodic signals would benefit more from frequency domain analysis rather than time-domain feature extraction. Whereas noisy signal or those used for fault detection—like vibrations—would necessitate both time and frequency feature extraction [46]. The signals in this work, on the other hand, are non-periodic and present no noise, hence making time-domain features the most suitable for the present application. Furthermore, Amir Sattari et al. [34] have recently explored the use of feature extraction in the detection of the type of flow and the concentration of the gas in a two-phase flow meter. Their study yielded great results despite the weak number of testing samples preventing better generalization.

4.1. Chosen time-domain features

The most common time-domain features are statistical features that help differentiate between the different sequences of the signal attached to different classes (gases). These are the means ($\bar{x}$ and the absolute mean $|\bar{x}|$), the standard deviation (σ), and the Root Mean Square (RMS) [47]. On the distribution side, Kurtosis and the Skewness are features used to examine the probability density function (PDF) of each signal sequence. Table 2 highlights all the features used to augment the dataset in this study, starting with the aforementioned features and their formulas, and continuing to other dimensionless time-domain features that have been proven useful in the literature to solve problems of detection and identification [45, 47].

4.2. Features calculation

The features presented in Table 2 are all computed over n samples, and each computed feature over these samples would constitute one entry n the extended dataset. Hence, to build the extended dataset, every vector (x) of the raw dataset would be split into k sequences composed of n samples. The number n can either be defined theoretically (using the signal period, for instance) or empirically by tuning it to generate the best results. In this work, the initial number n was set to be approximately one second of data acquisition. However, knowing a sampling period that is equal to $T_s \approx 0.2\,\mathrm{s}$ would result in $n = 5$ which might be insufficient. We doubled the the value of n inspired by inverse of the Nyquist–Shannon sampling theorem. To be able to empirically compare the results of different cases, we doubled n a second time. Thus:

$$\begin{cases} n_1 = \dfrac{1}{T_s} = 5 \\ n_2 = \dfrac{2}{T_s} = 10 \\ n_3 = \dfrac{4}{T_s} = 20 \end{cases} \tag{1}$$

Table 2: Time-domain features of to be computed for each of the data vectors.

Feature	Formula
Mean value ($\bar{x}$)	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
Absolutes mean value ($\lvert\bar{x}\rvert$)	$\lvert\bar{x}\rvert = \frac{1}{n}\sum_{i=1}^{n} \lvert x_i\rvert$
Standard deviation (σ)	$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$
Root mean square (RMS)	$RMS = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$
Peak value (PV)	$PV = max\left(\lvert x\rvert\right)$
Kurtosis value (KV)	$KV = \frac{1}{n}\frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{\sigma^4}$
Skewness ($Skew$)	$Skew = \frac{1}{n}\sum_{i=1}^{n}\frac{(x_i - \bar{x})^3}{\sigma^3}$
Crest Factor (CF)	$CF = \frac{PV}{RMS}$
Margin Factor (MF)	$MF = \frac{PV}{\left(\frac{1}{n}(\sum_{i=1}^{n}\sqrt{\lvert x_i\rvert})\right)^2}$
Impulse Factor (IF)	$IF = \frac{PV}{\lvert\bar{x}\rvert}$
Form Factor (FF)	$FF = \frac{RMS}{\lvert\bar{x}\rvert}$

Such configuration strikes balance between the speed of detection and accuracy. As, theoretically, the three value of n would allow the system to gather enough data to accurately detect and identify the gas, while still being quick to alarm users to the presence of danger. Once the raw vector (x) is split into k sequences, the features of each sequence are computed and organized into a matrix (x_{ext}) of k rows and 11 columns (the number of the features), as shown in Figure 5. Finally, the extended dataset of each class is then composed of the 8 extended matrices, hence raising the number of features from 8 in the raw dataset, into 88 in the extended one.

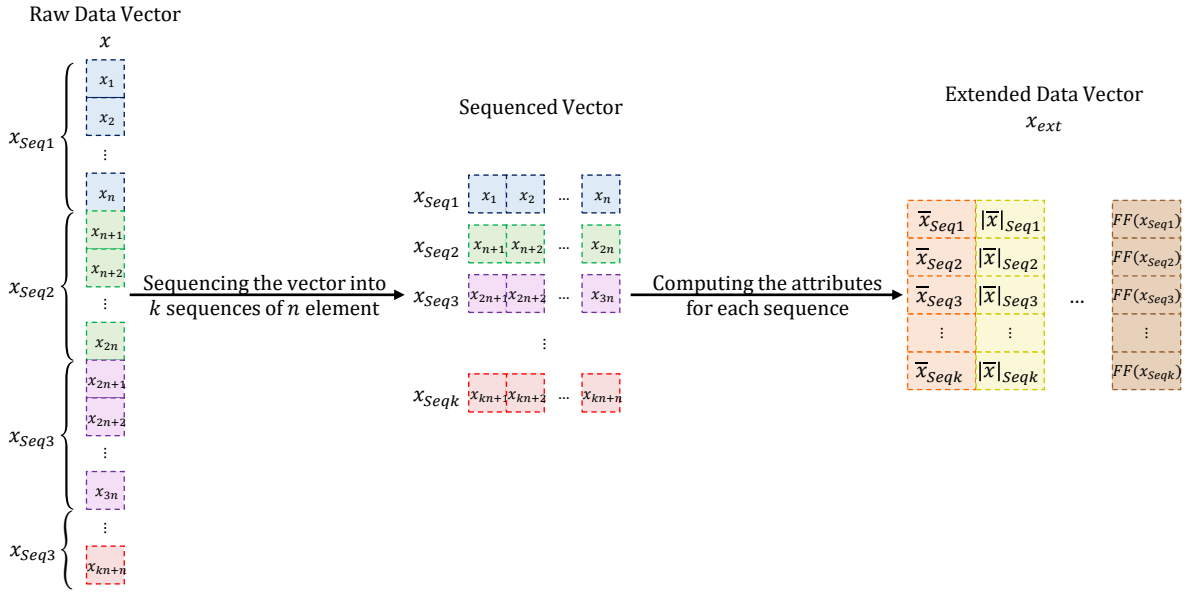


Figure 5: Sequencing and calculation of the features of a raw data vector x.

5. Feature ranking and selection

The extended dataset resulting from the feature extraction has 88 possible predictors for the classification model. While a classification model with an input this size is feasible, it is not computationally efficient both in training and online testing. Thus, its size needs to be reduced with a minimum loss of information.

Reducing the size of a database can be done through several approaches. The most known of which is the PCA [48] and its variants (such as independent component analysis [49]). With these techniques, reducing the size of a database is achieved by projecting the data into a reduced space by a linear or a non-linear mathematical transformation. The key advantage of these methods is preserving most of the variability even with a much-reduced size [11, 22, 50, 33]. However, the dimensions of the reduced space lose their physical meaning.

The literature is also rich with other approaches for the the selection of the features such as the heuristic

and iterative Sequential Forward Selection [51] and Sequential Backward Selection [52], or the graphical Branch and Bound [53] and its subsequent improvements [54, 55], or the algorithmic FOCUS [56, 57].

All of the previously mentioned methods have been used to achieve great results in the literature. However for this study, we aimed to keep the physical meaning of the signals—hence ruling out PCA—and make it computationaly possible for an implementation in an embedded system with the sensor—ruling out the heuristic methods. Thus, we have opted to use the ReliefF methods for the ranking and its scores for the selection.

5.1. The ReliefF Features ranking algorithm

The Relief method is a widely used approach for feature ranking and selection, and is based on the nearest neighbor principle [58]. It has had a lot of variance since its release [59, 60, 61]. One of which is its adaptation for multi-class problems; the ReliefF [62].

In a summary, the ReliefF computes a weight for each of the features. The value of the weight depends on the performance of the feature in detecting classes. In the initialization phase, all the weights W are set to 0. Then, for n iteration, the algorithm randomly selects an instance x_r from the feature currently weighted A. With k being a predefined constant, the algorithm proceeds to find the nearest k neighbors to the instance x_r of the same class called Hits (H_{x_r}), and the nearest k neighbors to this instance of different classes (called Misses, M_{x_r}). The k nearest neighbors are computed using the Manhattan distance [62]. The weight of the feature A, during the iteration i, is computed by Equation 2 [59].

$$W[A] = W[A] - \frac{\Delta\left(A, x_r, H_{x_r}\right)}{n \cdot k} + \frac{\Delta\left(A, x_r, M_{x_r}\right)}{n \cdot k} \tag{2}$$

In this work, the features are numerical and continuous. Thus, the difference function Δ is equal to [59]:

$$\Delta\left(A, x_r, I_{x_r}\right) = \frac{\|A(x_r) - A(I_{x_r})\|}{\max(A) - \min(A)} \tag{3}$$

This operation is repeated for all the available features.

5.2. Features Selection

The ReliefF method would result in the ranking of all the features from best to worst, according to their weights. Nevertheless, the number of features remains unchanged. Some of the features would have negative weights associated with them and can be directly discarded, but the number of the remaining features would remain relatively high. In such cases, the number of features can be selected empirically from the highly ranked ones through trial and error. In this study, on the other hand, the features are selected by applying a median filter to the numerical differentiation of the ranked weights. This method allows us to keep highly ranked features and stop when the values of the weights drop suddenly.

The selection algorithm starts by sorting only the positive features and defining a minimum number of features to be selected. Then, it computes the numerical differentiation (Δ_{SA}) and selects features with

weights greater than the median value of Δ_{SA}. Algorithm 1 describes a pseudo-code of the features selection methodology. In this algorithm, the selection is applied three times generating three levels of features. Each subsequent level contains more information and will be used to build a different model for better accuracy and comparison. The results of each model will be given in Section 7.

Algorithm 1 Pseudo-code of the 3 levels features selection algorithm.

Begin
Define $minimumFeatures, k$
$[ranking, weights] = ReliefF(x_{ext}, k)$ ▷ Run the ReliefF and compute the ranks and weights
$rankedFeatures = sort(x_{ext}, rankings)$ ▷ Sort the extended features according to their ranks from best to worst
$rankedWeights = sort(weights)$ ▷ Sort the weights from best to worst
$rankedWeights(rankedWeights \leq 0) = [\,]$ ▷ Delete the weights less or equal than 0
$rankedFeatures = rankedFeatures[0 : rankedWeights.length - 1]$ ▷ Delete the features with negative or zero weights
$selectedFeatures = [0, 0, 0]$
for $j = 1 : selectedFeatures.length$ **do** ▷ Computing the numerical differentiation
    $\Delta_{SA}(j) = rankedFeatures(j) - rankedFeatures(j-1)$
$currentFeatures = rankedFeatures$
for $i = 0 : (selectedFeatures.length - 1)$ **do**
    $j = 0$
    while $currentFeatures(j) > median(\Delta_{SA})$ **do**
        $j = j + 1$
    if $j < minimumFeatures$ **then**
        $selectedFeatures(i) = minimumFeatures$
    else
        $selectedFeatures(i) = j$
    $currentFeatures = rankedFeatures[j : rankedFeatures.length]$
End

6. Gas discrimination using Multi-Class Support Vector Machines

Support Vector Machines (SVM) are binary classification methods used for both supervised and unsupervised learning. In this method, the algorithm searches for a seperation hyperplane with an optimal margin between the two classes. This approach was proposed by Cortes and Vapnik [63], and is formally described as follows. Let us consider that a training data matrix (x) is composed of m features and a corresponding assigned label vector y of values $C1$ and $C2$, for the two classes. The classifier builds a model which predicts

the target class y from the input data (the features), by searching for an optimal hyperplane optimizing a quadratic problem (Equation 4).

$$\begin{aligned} &\min J(a) = \tfrac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j g_i(x)\, g_j(x)\, k(x,x) - \sum_{i=1}^{N} a_i \\ &s.t: \sum_{i=1}^{N} a_i g_i(x) = 0,\ 0 \le a_i \le D\ for\ i = 1, ..., N \end{aligned} \tag{4}$$

where $g(x) = 1$ if $x \in C1$ and $g(x) = -1$ if $x \in C2$, $a = [a_1, a_2, ..., a_N]^T$ are the Lagrange multipliers, D is the penalty parameter, and $k(x, x)$ is the Kernel function. Then, the output is estimated by the decision function :

$$y = \begin{cases} 1 & \text{, if } sign\left(\sum_{i=1}^{S} a_i^S g_i^S k(x_i^S, x) + b\right) = 1 \\ -1 & \text{, elsewhere} \end{cases} \tag{5}$$

whereas:

$$b = \frac{1}{S}\sum_{j=1}^{S}\left(g_j^S - \sum_{i=1}^{S} a_i^S g_i^S k(x_i^S, x_j^S)\right) \tag{6}$$

In this work, since there 7 gases, we use a Multi-Class Support Vector Machine (MC-SVM) classifier. The version used in this work is the error-correcting output codes (ECOC) classifier, which consists of several binary SVM used in tandem to solve the multi class problems [64]. The ECOC is used in both in its "One against One" and "One against All" configurations [65]. In the first configuration, the classifier trains $n * (n-1)/2$ binary SVM, whereas in the second one n binary SVM classifiers are built, with n being the number of classes. The classification results are obtained by a voting strategy: a pattern is classified to the class where the maximum number of votes is obtained.

7. Experimentation and results

This section is divided into three parts. The first part is dedicated to the division of the data and the extraction of the features. The second one focuses on the ranking and the selection of the features. Finally, the section ends with the results from the SVM classifiers, and comparative analysis of the results between the different approaches.

7.1. Data preprocessing

As described in Section 2, the sensor delivers 9 measurements per sample. One of these measurements—the sensors voltage—is constant and thus carries no information. Consequently, only the other 8 measurements are to be used for feature extraction and classification.

Furthermore, the data gathered from the experimentation are not uniformly distributed. Table 3 shows the number of samples per class and highlights the inequality between them. As it is, this dataset would

cause an issue of imbalanced learning for SVMs. Since, the latter, by design, require the same number of samples for each class [66]. This problem has been and is still being studied in the literature [67, 68], resulting in several creative solutions [69, 70].

Table 3: Raw and Extended datasets distribution amongst the 7 classes. FE: Feature Extraction.*$n = \frac{1}{T_s}$

Class	Label	Raw data samples	Samples after FE		
			$n = 5^*$	$n = 10$	$n = 20$
CO	1	24003	4800	2400	1200
NO_2	2	24003	4800	2400	1200
O_3	3	16002	3200	1600	800
CO + NO_2	4	72009	14401	7200	3600
CO + O_3	5	72009	14401	7200	3600
NO_2 + O_3	6	72009	14401	7200	3600
All	7	40005	8001	4000	2000

The study of imbalanced learning largely comes into play when there is a great difference in the representation of the classes. However, implementing such solutions would add additional processing layers, and is not the purpose of this study, especially with the number of instances for the least represented class is sufficient for training and testing the model. Therefore, to avoid the problems related to imbalanced learning, the classes datasets are balanced by determining the set with the fewest samples N—in this case, O_3—and then dividing it into a 70% training set and 30% test set. For the rest of the classes, the test sets are built by randomly selecting $0.3 \times N$ samples for each class. These sets are set apart and only used to obtain final test results. The training sets, on the other hand, are built by randomly sampling $0.7 \times N$ instances at each new step of the cross-validation. As can be seen in Table 3, from the number of samples after the extraction of the features with different values of n, the class O_3 has the fewest samples. Accordingly, the number of samples used for the training and the test phases will be based on its size.

7.2. Features ranking and selection

Once the features are extracted, the ReliefF algorithm is applied to the resultant training extended sets. During our testings, the best results were obtained with the number of nearest neighbors $k = 12$. The 30 highest ranked are displayed in Table 4. The latter shows that the highest-ranked features are mostly independent of the number of samples per instance (n). It also indicates—counter-intuitively—that the highest-ranked features are not the ones extracted from the sensor's measurements, but rather from the measurements of the first heater V_{H1}, I_{H1}, with the features of I_S coming after them.

Furthermore, while the highest-ranked features remain largely unchanged for all values of n, their weights vary significantly enough with every value of n to be a major factor in changing the outcome of the selection algorithm. The other major factor is the minimum number of features per iteration, which was set to 4 in

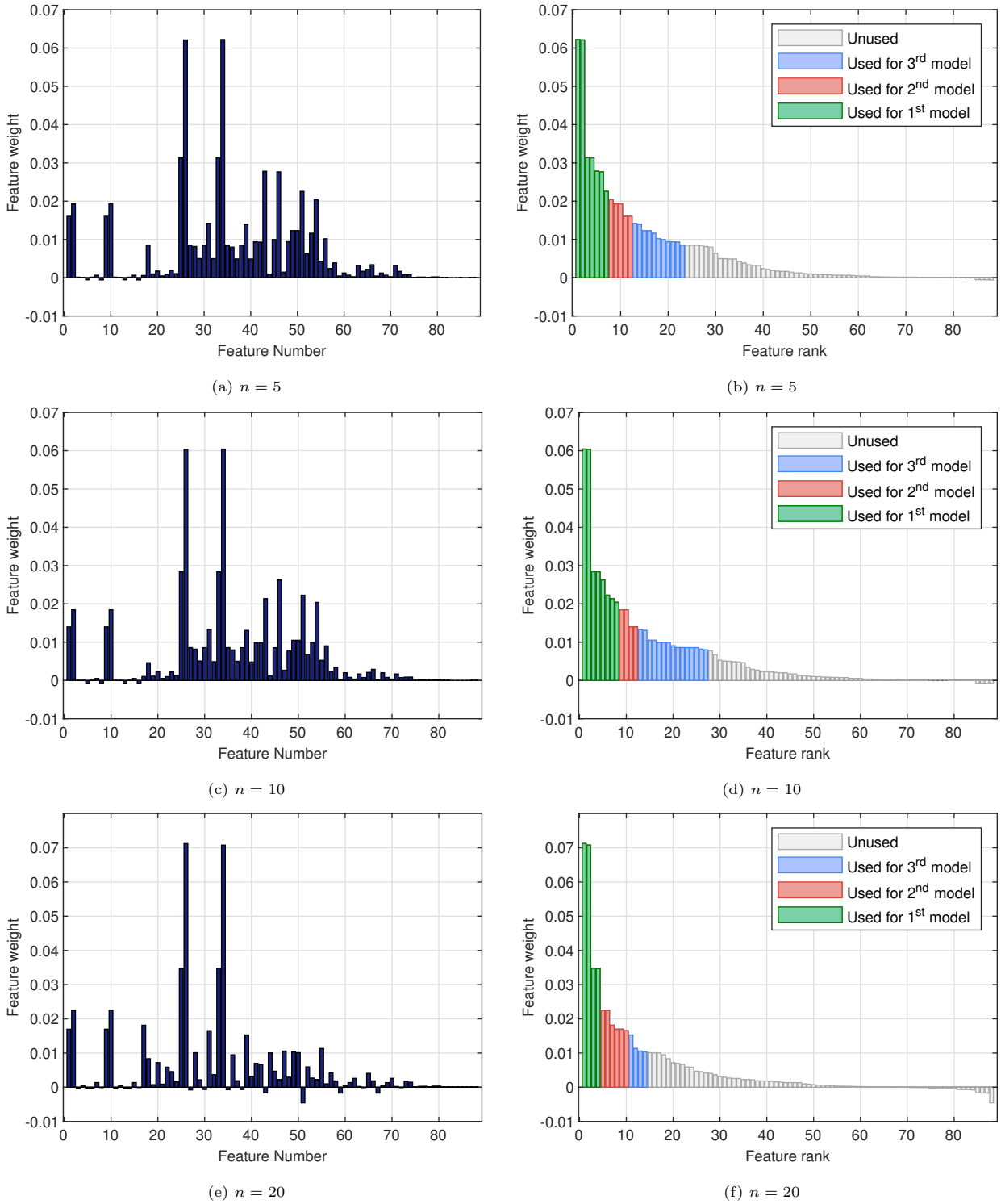
this study (c.f. Algorithm 1). Thus, the number of predictors used to train the Support Vector Classifiers varies with n. In Table 4, row colors indicate the features used to train each of the three models per n value. These features are also shown ranked from highest to lowest in Figure 6, showing the difference between the three levels. The figure also shows the weights of all the 88 features for the three values of n.

Finally, Figure 7 highlights a better visual separability between the classes according to the highest-ranked features, compared to using raw readings (Figure 3). The separability is especially visible for the classes O_3 and All in Figures 7a and 7c. The classes CO and NO_2 also show better potential for separability in Figures 7c, indicating the potential of achieving high accuracy with the MC-SVM.

s

Table 4: Results of the ReliefF : the 30 highest ranked features for different values of n. **FN**: Feature Number. **VN**: Variable Name. **FN**: Feature Name. **W**: Weight. **Green**: Features used for the first model. **Green** + **Red**: Features used for the second model. **Green** + **Red** + **Blue**: Features used for the third model. **Grey** : Unused features.

n	**5**				**10**				**20**			
Rank	**FN**	**VN**	**AN**	**W**	**FN**	**VN**	**AN**	**W**	**FN**	**VN**	**AN**	**W**
1	26	V_{H1}(V)	PV	0.07658	26	V_{H1}(V)	PV	0.06924	26	V_{H1}(V)	PV	0.07122
2	34	I_{H1}(A)	$\lvert\bar{x}\rvert$	0.07616	34	I_{H1}(A)	$\lvert\bar{x}\rvert$	0.06919	34	I_{H1}(A)	$\lvert\bar{x}\rvert$	0.07077
3	25	V_{H1}(V)	RMS	0.03942	33	I_{H1}(A)	$\bar{x}$	0.03324	33	I_{H1}(A)	$\bar{x}$	0.03476
4	33	I_{H1}(A)	$\bar{x}$	0.03940	25	V_{H1}(V)	RMS	0.03310	25	V_{H1}(V)	RMS	0.03468
5	2	I_S(A)	σ	0.02417	2	I_S(A)	σ	0.02265	2	I_S(A)	σ	0.02247
6	10	I_S(A)	IF	0.02417	10	I_S(A)	IF	0.02265	10	I_S(A)	IF	0.02247
7	1	I_S(A)	$\lvert\bar{x}\rvert$	0.01922	31	V_{H1}(V)	MF	0.01751	17	$R_S(\Omega)$	$Skew$	0.01812
8	9	I_S(A)	MF	0.01922	1	I_S(A)	$\lvert\bar{x}\rvert$	0.01715	1	I_S(A)	$\lvert\bar{x}\rvert$	0.01696
9	17	$R_S(\Omega)$	$Skew$	0.01899	9	I_S(A)	MF	0.01715	9	I_S(A)	MF	0.01696
10	31	V_{H1}(V)	MF	0.01619	17	$R_S(\Omega)$	$Skew$	0.01713	31	V_{H1}(V)	MF	0.01652
11	39	I_{H1}(A)	$Skew$	0.01508	39	I_{H1}(A)	$Skew$	0.01623	39	I_{H1}(A)	$Skew$	0.01528
12	49	P_{H1}(W)	KV	0.01284	49	P_{H1}(W)	KV	0.01305	55	V_{H2}(V)	$\bar{x}$	0.01131
13	50	P_{H1}(W)	$Skew$	0.01267	50	P_{H1}(W)	$Skew$	0.01285	47	P_{H1}(W)	RMS	0.01056
14	28	V_{H1}(V)	$Skew$	0.01110	28	V_{H1}(V)	$Skew$	0.01157	49	P_{H1}(W)	KV	0.01032
15	36	I_{H1}(A)	RMS	0.01013	36	I_{H1}(A)	RMS	0.01062	50	P_{H1}(W)	$Skew$	0.01011
16	18	$R_S(\Omega)$	CF	0.00862	47	P_{H1}(W)	RMS	0.00845	28	V_{H1}(V)	$Skew$	0.01008
17	20	$R_S(\Omega)$	MF	0.00755	18	$R_S(\Omega)$	CF	0.00773	44	P_{H1}(W)	$\bar{x}$	0.01003
18	41	I_{H1}(A)	MF	0.00639	41	I_{H1}(A)	MF	0.00700	36	I_{H1}(A)	RMS	0.00947
19	44	P_{H1}(W)	$\bar{x}$	0.00612	44	P_{H1}(W)	$\bar{x}$	0.00672	18	$R_S(\Omega)$	CF	0.00834
20	42	I_{H1}(A)	MF	0.00606	42	I_{H1}(A)	MF	0.00669	20	$R_S(\Omega)$	MF	0.00719
21	47	P_{H1}(W)	RMS	0.00598	20	$R_S(\Omega)$	MF	0.00660	41	I_{H1}(A)	MF	0.00699
22	55	V_{H2}(V)	$\bar{x}$	0.00557	55	V_{H2}(V)	$\bar{x}$	0.00638	42	I_{H1}(A)	MF	0.00669
23	32	V_{H1}(V)	IF	0.00485	22	V_{H1}(V)	$\bar{x}$	0.00521	52	P_{H1}(W)	MF	0.00593
24	52	P_{H1}(W)	MF	0.00484	32	V_{H1}(V)	IF	0.00501	22	V_{H1}(V)	$\bar{x}$	0.00586
25	22	V_{H1}(V)	$\bar{x}$	0.00442	40	I_{H1}(A)	CF	0.00445	45	P_{H1}(W)	$\lvert\bar{x}\rvert$	0.00468
26	40	I_{H1}(A)	CF	0.00434	29	V_{H1}(V)	CF	0.00397	23	V_{H1}(V)	$\lvert\bar{x}\rvert$	0.00457
27	29	V_{H1}(V)	CF	0.00404	45	P_{H1}(W)	$\lvert\bar{x}\rvert$	0.00396	57	V_{H2}(V)	σ	0.00418
28	21	$R_S(\Omega)$	IF	0.00379	57	V_{H2}(V)	σ	0.00376	65	V_{H2}(V)	IF	0.00404
29	37	I_{H1}(A)	PV	0.00376	37	I_{H1}(A)	PV	0.00361	32	V_{H1}(V)	IF	0.00366
30	57	V_{H2}(V)	σ	0.00355	65	V_{H2}(V)	IF	0.00360	40	I_{H1}(A)	CF	0.00312



(a) $n = 5$ (b) $n = 5$

(c) $n = 10$ (d) $n = 10$

(e) $n = 20$ (f) $n = 20$

Figure 6: The rank and the weights of the all features fro all the values of n.

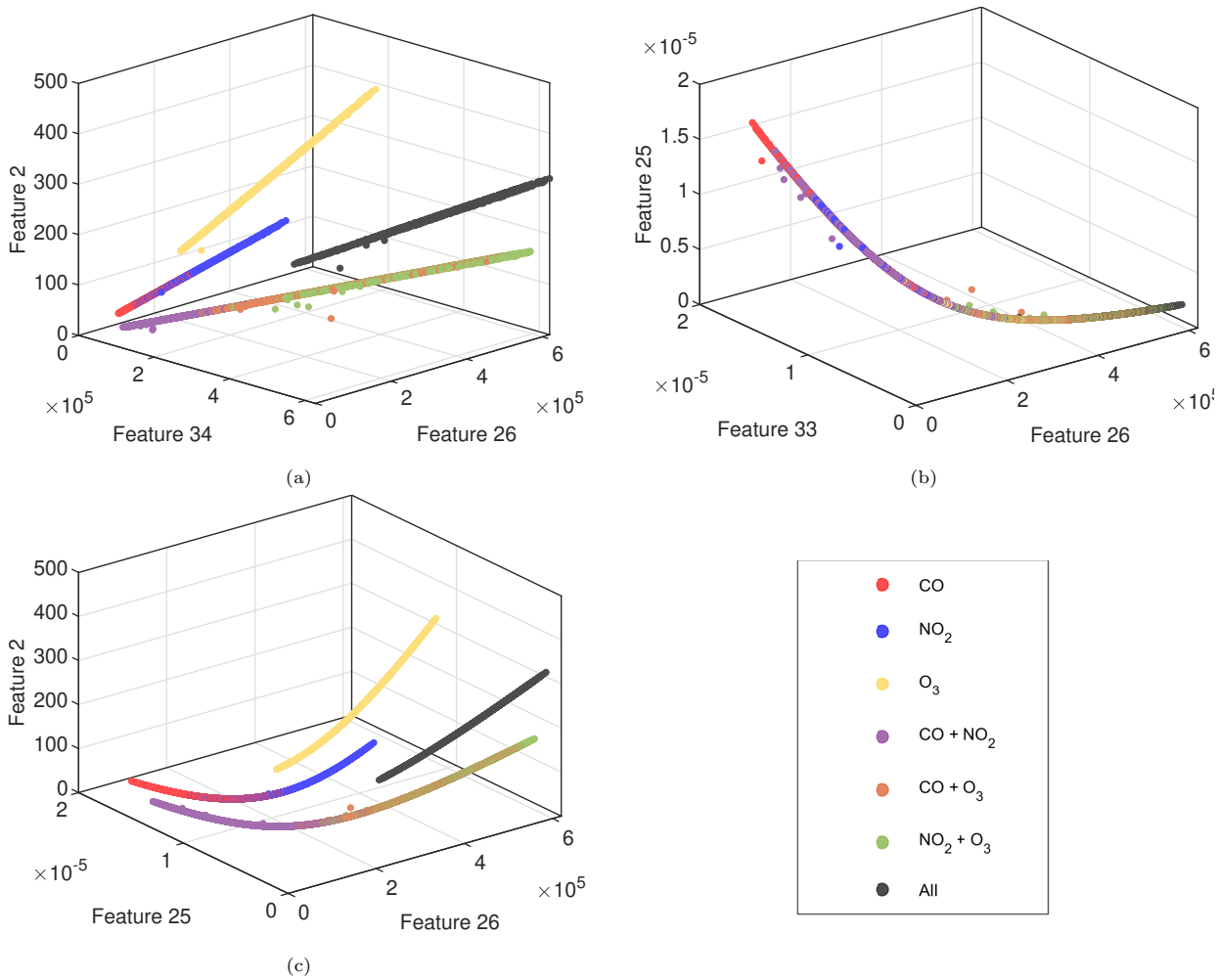


Figure 7: Visual representation of the distribution of the classes according to the highest-racked features for $n = 20$.

7.3. Support vector classification results

The results presented in this paragraph showcase all the approaches and the tuning done to the classifiers to generate the best possible results. Thus, these results are divided into 4 Tables (Tables 5—8). Each table contains the results of the classification per class, then averaged results of the 7 classes. The training results represent the results after a k-fold cross-validation on the training set ($k = 5$), whereas the test results are gathered from the predictions of the model on a separate test dataset (c.f. Paragraph 7.1). The results are reported in terms of metrics that are well-established in the literature. These metrics are all computed by comparing the model's predictions and the targets. This comparison results in 4 types of predictions per class: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). Thus, models' performance metrics are listed below.

- **The Accuracy (ACC):** The most basic metric. It is the ratio of the correctly classified instances

over the total number of the instances.

$$\text{ACC}(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{7}$$

- **The Mis-classification Rate (MR):**

$$\text{MR}(\%) = 100 - \text{ACC} \tag{8}$$

- **The Precision (PRCN):** The ratio of the correctly identified positive class.

$$\text{PRCN}(\%) = \frac{TP}{TP + FP} \times 100 \tag{9}$$

- **The Recall:** The true positive rate, or the ratio of correctly identified positives of all the positive class instances.

$$\text{Recall}(\%) = \frac{TP}{TP + FN} \times 100 \tag{10}$$

- **The F1_Score (F1):** Most commonly used for the comparison between the models. It is takes into account both the Recall and the Precision.

$$\text{F1} = 2 \cdot \frac{\text{PRCN} \cdot \text{Recall}}{\text{PRCN} + \text{Recall}} \tag{11}$$

The results in Tables 5– 8 are the best results with these predictors after all the tuning was done. They are color-coded to facilitate their investigation and the comparison between the different models and configurations. Perfect scores are colored in purple, while the rest of the scores are colored from green (best) to red (worst).

Starting with Table 5, it showcases the results from two preliminary models built to serve as references of base performance. The first model—which we will call Power Model, hereafter—is a classifier that uses the minimum amount of data, with only three inputs: the power readings from the sensor and two heaters. The second model is a support vector classifier that relies on the non-transformed raw data from the sensors. So, it has 8 inputs. This model will be referred to as the Raw Data Model. While both models deliver good performance overall, notably in the case of the *One against all* classifiers, they both take a hit in performance with the binary gas mixtures. However, they still deliver better performance than the first level models that rely on feature extraction and feature selection using the ReliefF. For instance, in Table 6 ($n = 5$), the results of the first models with 6 predictors are far worse than those of the Power the Raw Data Models. The same conclusion can be drawn for the first models in Table 7 (8 predictors) and those in Table 6 (4 predictors).

The low performance of these models cannot be explained by the number of predictors, as the Power Model contains even fewer predictors. It can be explained, however, by the predictors used. In these first

models, most of the predictors come from readings of the first heater (notably V_{H1}). This is especially relevant for $n = 20$, where only 4 features extracted from V_{H1} were used as predictors and no information from the sensor whatsoever.

The first tier of models highlights that—while the ranking of the feature and the difference between their weights are important factors for the selection—the distribution of the information is also a primordial criterion, during the selection of the features. In Algorithm 1, the minimum number of features was included to mitigate such shortcomings. While its value of 4 was not high enough for the first tier of models, nevertheless, it forced the addition of 4 features to the second model for the case of $n = 10$, instead of two as computed by the median filter (Figure 6d), resulting in the use of 12 predictors. Table 7 shows how the performance for this tier of models increased significantly, achieving even perfect test results. As for $n = 20$, results from Table 8 show that, for the second tier of models, the inclusion of features from the sensor's measurements lead to achieving perfect results, even with fewer predictors compared to the case of $n = 10$, thanks to the greater number of samples per instance.

The duality of the number of predictors and the number of samples per instance is also visible in the case of $n = 5$, where a greater number of predictors (23) is needed to achieve 100% test results because of the fewer samples used to compute every feature instance.

The results presented in this paragraph demonstrate that achieving perfect discrimination is possible with multiple configurations. However, a compromise needs to be struck between the speed of detection and the available computational power, as more samples per instance require more time to be collected, and more features require a greater computational capacity to be computed.

Finally, Figure 8 and Figure 9 show examples for separating hyperplanes for the cases of $n = 10$ and $n = 20$, respectively.

Table 5: Results of the classification of the gases using the minimum amount of data (power measurements as predictors) in a first model test and the raw data set in the second one.

Model		**Power Measuremets**				**Raw Data**			
Number of predictors		**3**				**8**			
Kernel function		**RBF**				**RBF**			
Muticlass method		**One vs one**		**One vs all**		**One vs one**		**One vs all**	
Results		**Training set**	**Test Set**	**Training set**	**Test Set**	**Training set**	**Test Set**	**Training set**	**Test Set**
CO	ACC (%)	93.18	92.96	97.33	97.33	93.15	92.92	97.33	97.33
	PRCN (%)	70.35	69.68	90.05	90.05	67.92	67.15	90.05	90.05
	MR (%)	6.82	7.04	2.67	2.67	6.85	7.08	2.67	2.67
	Recall (%)	90.29	89.84	91.42	91.42	98.61	98.69	91.42	91.42
	F1 (%)	79.08	78.48	90.73	90.73	80.44	79.92	90.73	90.73
NO2	ACC (%)	93.37	93.45	99.37	99.37	96.28	96.59	99.37	99.37
	PRCN (%)	69.70	69.99	97.28	97.28	80.01	81.52	97.28	97.28
	MR (%)	6.63	6.55	0.63	0.63	3.72	3.41	0.63	0.63
	Recall (%)	94.80	94.77	98.31	98.31	98.61	98.44	98.31	98.31
	F1 (%)	80.34	80.52	97.79	97.79	88.34	89.19	97.79	97.79
O3	ACC (%)	97.73	97.70	99.99	99.99	97.73	97.77	99.99	99.99
	PRCN (%)	87.40	87.17	99.96	99.96	86.63	86.75	99.96	99.96
	MR (%)	2.27	2.30	0.01	0.01	2.27	2.23	0.01	0.01
	Recall (%)	98.31	98.38	99.98	99.98	99.49	99.56	99.98	99.98
	F1 (%)	92.53	92.44	99.97	99.97	92.62	92.72	99.97	99.97
CO + NO2	ACC (%)	86.93	86.79	96.70	96.70	89.39	89.47	96.70	96.70
	PRCN (%)	61.88	60.55	89.48	89.48	92.96	93.04	89.48	89.48
	MR (%)	13.07	13.21	3.30	3.30	10.61	10.53	3.30	3.30
	Recall (%)	22.10	21.64	87.13	87.13	27.83	28.39	87.13	87.13
	F1 (%)	32.56	31.89	88.29	88.29	42.83	43.50	88.29	88.29
CO + O3	ACC (%)	94.00	93.79	99.01	99.01	93.52	93.49	99.01	99.01
	PRCN (%)	89.70	89.13	97.03	97.03	79.95	79.58	97.03	97.03
	MR (%)	6.00	6.21	0.99	0.99	6.48	6.51	0.99	0.99
	Recall (%)	65.56	64.40	95.98	95.98	72.95	73.23	95.98	95.98
	F1 (%)	0.76	0.75	0.97	0.97	0.76	0.76	0.97	0.97
NO2 + O3	ACC (%)	90.62	90.36	98.11	98.11	89.54	89.59	98.11	98.11
	PRCN (%)	69.16	67.79	93.50	93.50	65.27	65.58	93.50	93.50
	MR (%)	9.38	9.64	1.89	1.89	10.46	10.41	1.89	1.89
	Recall (%)	61.94	61.99	93.25	93.25	57.28	57.11	93.25	93.25
	F1 (%)	65.35	64.76	93.38	93.38	61.01	61.06	93.38	93.38
All	ACC (%)	94.67	94.74	99.09	99.09	94.23	94.32	99.09	99.09
	PRCN (%)	75.12	75.60	96.18	96.18	77.65	77.92	96.18	96.18
	MR (%)	5.33	5.26	0.91	0.91	5.77	5.68	0.91	0.91
	Recall (%)	93.76	93.27	97.50	97.50	83.71	84.07	97.50	97.50
	F1 (%)	83.41	83.51	96.83	96.83	80.57	80.87	96.83	96.83
Averaged	ACC (%)	92.93	92.83	98.51	98.51	93.41	93.45	98.51	98.51
	PRCN (%)	74.99	74.73	94.81	94.81	76.20	76.26	94.81	94.81
	Recall (%)	75.25	74.90	94.80	94.80	76.92	77.07	94.80	94.80
	microF1 (%)	75.12	74.81	94.80	94.80	76.56	76.66	94.80	94.80
	macroF1 (%)	72.72	72.34	94.78	94.78	74.59	74.79	94.78	94.78

Table 6: The results from support vector classifiers using Attributes Extraction (AE) from data augmentation and ReliefF for predictor selection. Case of $n = 5$.

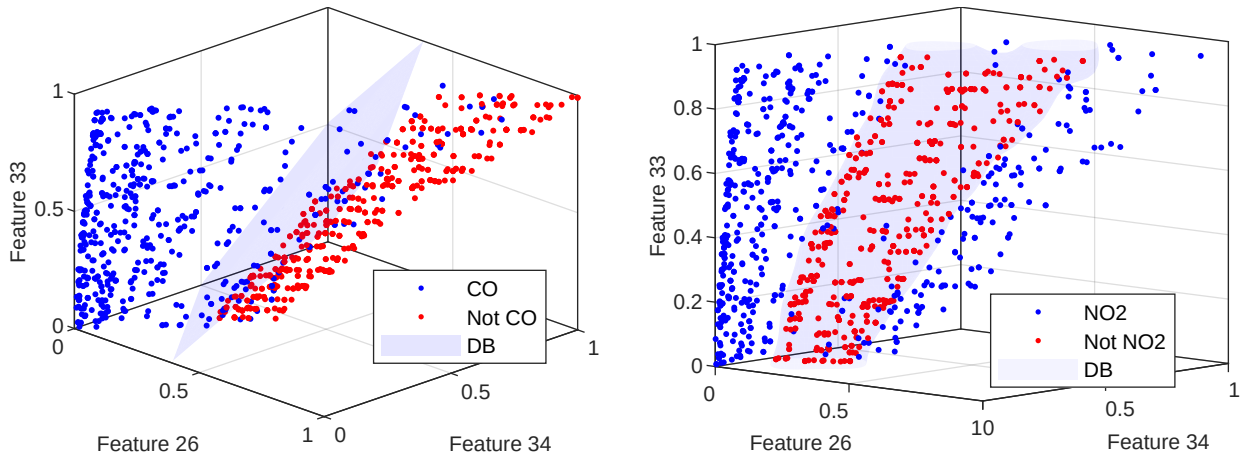
Model		**AE + ReliefF, $n = 5$**											
Number of predictors		**7**				**12**				**23**			
Kernel function		**Linear**		**RBF**		**Linear**		**RBF**		**Linear**		**RBF**	
Muticlass method		**One vs one**											
Results		**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**
CO	ACC (%)	92.67	92.34	98.61	92.22	98.60	98.97	99.50	98.82	99.96	**100.00**	99.96	**100.00**
	PRCN (%)	71.04	70.73	93.71	72.90	95.82	96.87	98.38	97.42	99.69	**100.00**	99.69	**100.00**
	MR (%)	7.33	7.66	1.39	7.78	1.40	1.03	0.50	1.18	0.04	**0.00**	0.04	**0.00**
	Recall (%)	81.97	82.57	96.73	75.75	94.32	96.19	98.12	94.59	100.00	**100.00**	100.00	**100.00**
	F1 (%)	76.11	76.19	95.20	74.30	95.06	96.53	98.25	95.98	99.84	**100.00**	99.84	**100.00**
NO2	ACC (%)	92.02	92.31	98.05	92.46	98.65	98.97	99.55	98.78	100.00	**100.00**	100.00	**100.00**
	PRCN (%)	69.20	69.18	90.34	72.18	94.48	96.07	98.15	94.43	100.00	**100.00**	100.00	**100.00**
	MR (%)	7.98	7.69	1.95	7.54	1.35	1.03	0.45	1.22	0.00	**0.00**	0.00	**0.00**
	Recall (%)	80.44	83.23	96.81	76.77	96.24	96.77	98.72	97.19	100.00	**100.00**	100.00	**100.00**
	F1 (%)	74.40	75.56	93.46	74.41	95.35	96.42	98.43	95.79	100.00	**100.00**	100.00	**100.00**
O3	ACC (%)	94.41	94.94	98.40	95.27	100.00	100.00	100.00	99.96	100.00	**100.00**	100.00	**100.00**
	PRCN (%)	77.92	80.11	91.52	79.63	100.00	100.00	100.00	99.78	100.00	**100.00**	100.00	**100.00**
	MR (%)	5.59	5.06	1.60	4.73	0.00	0.00	0.00	0.04	0.00	**0.00**	0.00	**0.00**
	Recall (%)	85.34	82.78	97.96	86.89	100.00	100.00	100.00	99.89	100.00	**100.00**	100.00	**100.00**
	F1 (%)	81.46	81.42	94.63	83.10	100.00	100.00	100.00	99.83	100.00	**100.00**	100.00	**100.00**
CO + NO2	ACC (%)	86.08	85.85	97.18	86.80	100.00	100.00	100.00	99.99	100.00	**100.00**	100.00	**100.00**
	PRCN (%)	51.93	52.32	94.86	55.25	100.00	100.00	100.00	100.00	100.00	**100.00**	100.00	**100.00**
	MR (%)	13.92	14.15	2.82	13.20	0.00	0.00	0.00	0.01	0.00	**0.00**	0.00	**0.00**
	Recall (%)	34.90	32.28	84.87	49.44	100.00	100.00	100.00	99.90	100.00	**100.00**	100.00	**100.00**
	F1 (%)	41.74	39.92	89.59	52.18	100.00	100.00	100.00	99.95	100.00	**100.00**	100.00	**100.00**
CO + O3	ACC (%)	87.38	87.49	95.78	89.52	89.02	89.14	97.50	91.74	99.99	**100.00**	99.99	**99.91**
	PRCN (%)	58.02	56.82	91.54	64.11	62.81	63.01	91.86	70.08	100.00	**100.00**	100.00	**99.57**
	MR (%)	12.62	12.51	4.22	10.48	10.98	10.86	2.50	8.26	0.01	**0.00**	0.01	**0.09**
	Recall (%)	40.99	43.45	77.55	56.87	56.14	53.89	90.46	71.35	99.96	**100.00**	99.96	**99.79**
	F1 (%)	48.04	49.25	83.97	60.27	59.29	58.09	91.15	70.71	99.98	**100.00**	99.98	**99.68**
NO2 + O3	ACC (%)	86.21	85.71	96.55	87.08	89.03	89.14	97.51	91.76	100.00	**100.00**	100.00	**99.91**
	PRCN (%)	51.90	49.80	89.82	55.57	60.34	60.38	90.63	71.41	100.00	**100.00**	100.00	**99.79**
	MR (%)	13.79	14.29	3.45	12.92	10.97	10.86	2.49	8.24	0.00	**0.00**	0.00	**0.09**
	Recall (%)	41.66	38.56	85.43	46.39	66.73	68.97	91.97	70.22	100.00	**100.00**	100.00	**99.58**
	F1 (%)	46.22	43.46	87.57	50.57	63.37	64.39	91.30	70.81	100.00	**100.00**	100.00	**99.69**
All	ACC (%)	88.41	88.24	96.55	90.43	99.96	100.00	99.96	100.00	99.96	**100.00**	99.96	**100.00**
	PRCN (%)	56.49	57.00	83.43	64.68	100.00	100.00	100.00	100.00	100.00	**100.00**	100.00	**100.00**
	MR (%)	11.59	11.76	3.45	9.57	0.04	0.00	0.04	0.00	0.04	**0.00**	0.04	**0.00**
	Recall (%)	79.45	81.26	94.42	76.80	99.73	100.00	99.73	100.00	99.73	**100.00**	99.73	**100.00**
	F1 (%)	66.03	67.00	88.59	70.22	99.86	100.00	99.86	100.00	99.86	**100.00**	99.86	**100.00**
Averaged	ACC (%)	89.60	89.55	97.30	90.54	96.47	96.60	99.15	97.29	99.99	**100.00**	99.99	**99.97**
	PRCN (%)	64.36	64.05	90.37	67.37	87.55	88.01	97.01	90.60	99.96	**100.00**	99.96	**99.91**
	Recall (%)	63.63	63.44	90.58	66.81	87.68	88.25	97.02	90.60	99.96	**100.00**	99.96	**99.91**
	microF1 (%)	63.99	63.75	90.48	67.09	87.61	88.13	97.02	90.60	99.96	**100.00**	99.96	**99.91**
	macroF1 (%)	62.00	61.83	90.43	66.44	87.56	87.92	97.00	90.44	99.96	**100.00**	99.96	**99.91**

Table 7: The results from support vector classifiers using Attributes Extraction (AE) from data augmentation and ReliefF for predictor selection. Case of $n = 10$

Model		**AE + ReliefF, $n = 10$**											
Number of predictors		**8**				**12**				**27**			
Kernel function		**Linear**		**RBF**		**Linear**		**RBF**		**Linear**		**RBF**	
Muticlass method		**One vs one**											
Results		**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**
CO	ACC (%)	99.02	98.93	99.40	98.78	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.88**
	PRCN (%)	98.91	98.03	98.76	98.01	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	MR (%)	0.98	1.07	0.60	1.22	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.12**
	Recall (%)	94.36	94.33	97.14	93.28	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.16**
	F1 (%)	96.58	96.15	97.94	95.59	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.58**
NO2	ACC (%)	99.02	98.93	99.40	98.72	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.82**
	PRCN (%)	94.50	94.71	97.12	93.42	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**98.80**
	MR (%)	0.98	1.07	0.60	1.28	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.18**
	Recall (%)	98.94	98.17	98.76	98.17	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	F1 (%)	96.67	96.41	97.93	95.74	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.39**
O3	ACC (%)	100.00	100.00	100.00	99.94	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.94**
	PRCN (%)	100.00	100.00	100.00	100.00	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	MR (%)	0.00	0.00	0.00	0.06	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.06**
	Recall (%)	100.00	100.00	100.00	99.60	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.60**
	F1 (%)	100.00	100.00	100.00	99.80	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.80**
CO + NO2	ACC (%)	100.00	100.00	100.00	99.97	**100.00**	**100.00**	**100.00**	**99.97**	**100.00**	**100.00**	**100.00**	**99.97**
	PRCN (%)	100.00	100.00	100.00	99.78	**100.00**	**100.00**	**100.00**	**99.78**	**100.00**	**100.00**	**100.00**	**99.78**
	MR (%)	0.00	0.00	0.00	0.03	**0.00**	**0.00**	**0.00**	**0.03**	**0.00**	**0.00**	**0.00**	**0.03**
	Recall (%)	100.00	100.00	100.00	100.00	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	F1 (%)	100.00	100.00	100.00	99.89	**100.00**	**100.00**	**100.00**	**99.89**	**100.00**	**100.00**	**100.00**	**99.89**
CO + O3	ACC (%)	89.71	89.97	97.93	94.91	**100.00**	**100.00**	**100.00**	**99.97**	**100.00**	**100.00**	**100.00**	**99.52**
	PRCN (%)	67.88	69.25	96.66	84.91	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.79**
	MR (%)	10.29	10.03	2.07	5.09	**0.00**	**0.00**	**0.00**	**0.03**	**0.00**	**0.00**	**0.00**	**0.48**
	Recall (%)	52.07	53.85	88.49	78.38	**100.00**	**100.00**	**100.00**	**99.79**	**100.00**	**100.00**	**100.00**	**96.88**
	F1 (%)	58.93	60.58	92.39	81.51	**100.00**	**100.00**	**100.00**	**99.90**	**100.00**	**100.00**	**100.00**	**98.31**
NO2 + O3	ACC (%)	89.71	89.97	97.93	94.88	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.43**
	PRCN (%)	62.20	61.92	89.72	79.57	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**96.54**
	MR (%)	10.29	10.03	2.07	5.12	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.57**
	Recall (%)	76.19	75.84	97.05	85.92	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.58**
	F1 (%)	68.49	68.18	93.24	82.63	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**98.04**
All	ACC (%)	100.00	100.00	100.00	99.94	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.94**
	PRCN (%)	100.00	100.00	100.00	100.00	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	MR (%)	0.00	0.00	0.00	0.06	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.06**
	Recall (%)	100.00	100.00	100.00	99.59	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.59**
	F1 (%)	100.00	100.00	100.00	99.79	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**99.79**
Averaged	ACC (%)	96.78	96.83	99.24	98.16	**100.00**	**100.00**	**100.00**	**99.99**	**100.00**	**100.00**	**100.00**	**99.79**
	PRCN (%)	87.98	88.37	97.19	93.46	**100.00**	**100.00**	**100.00**	**99.97**	**100.00**	**100.00**	**100.00**	**99.24**
	Recall (%)	88.66	88.92	97.32	93.58	**100.00**	**100.00**	**100.00**	**99.97**	**100.00**	**100.00**	**100.00**	**99.25**
	microF1 (%)	88.32	88.65	97.26	93.52	**100.00**	**100.00**	**100.00**	**99.97**	**100.00**	**100.00**	**100.00**	**99.25**
	macroF1 (%)	88.67	88.76	97.36	93.56	**100.00**	**100.00**	**100.00**	**99.97**	**100.00**	**100.00**	**100.00**	**99.26**

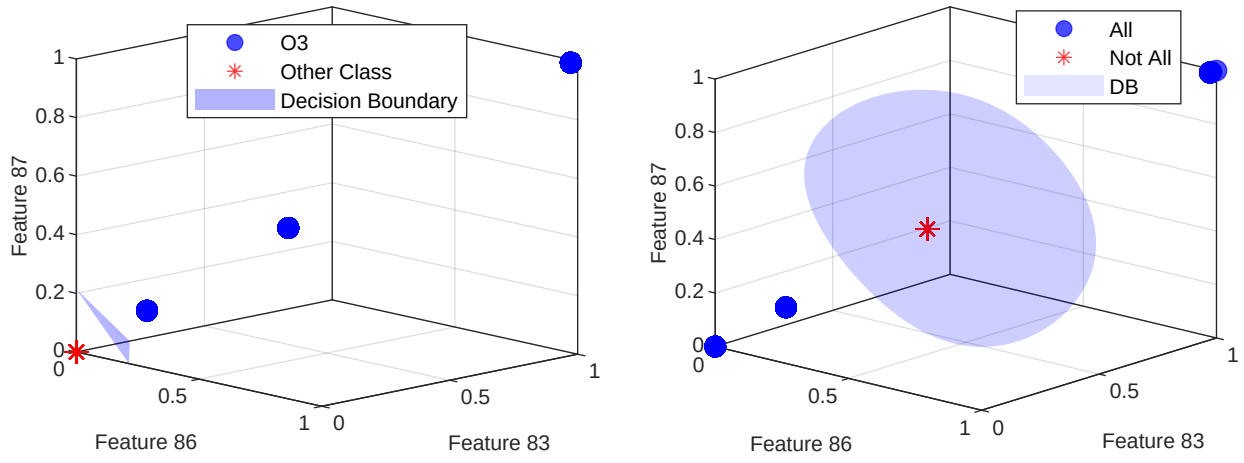
Table 8: The results from support vector classifiers using Attributes Extraction (AE) from data augmentation and ReliefF for predictor selection. Case of $n = 20$

Model		**AE + ReliefF,** $n = 10$											
Number of predictors		**4**				**10**				**14**			
Kernel function		**Linear**		**RBF**		**Linear**		**RBF**		**Linear**		**RBF**	
Muticlass method		**One vs one**											
Results		**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**	**Training set**	**Test set**
CO	ACC (%)	91.66	90.18	97.73	90.95	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	PRCN (%)	66.15	57.26	92.59	61.03	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	MR (%)	8.34	9.82	2.27	9.05	**0.00**	**0.00**	**0.00**	0.00	**0.00**	**0.00**	**0.00**	**0.00**
	Recall (%)	74.51	67.34	89.88	65.33	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	F1 (%)	70.08	61.89	91.21	63.11	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
NO2	ACC (%)	91.94	92.44	97.91	92.44	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	PRCN (%)	70.18	73.61	90.69	75.76	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	MR (%)	8.06	7.56	2.09	7.56	**0.00**	**0.00**	**0.00**	0.00	**0.00**	**0.00**	**0.00**	**0.00**
	Recall (%)	79.14	80.61	95.69	76.05	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	F1 (%)	74.39	76.95	93.12	75.90	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
O3	ACC (%)	93.44	93.93	99.21	97.02	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	PRCN (%)	76.12	80.65	95.25	85.92	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	MR (%)	6.56	6.07	0.79	2.98	**0.00**	**0.00**	**0.00**	0.00	**0.00**	**0.00**	**0.00**	**0.00**
	Recall (%)	74.52	78.74	99.05	96.06	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	F1 (%)	75.31	79.68	97.11	90.71	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
CO + NO2	ACC (%)	84.49	83.10	96.51	85.89	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	PRCN (%)	47.75	39.57	89.49	51.80	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	MR (%)	15.51	16.90	3.49	14.11	**0.00**	**0.00**	**0.00**	0.00	**0.00**	**0.00**	**0.00**	**0.00**
	Recall (%)	36.05	30.20	86.90	46.94	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	F1 (%)	41.09	34.26	88.18	49.25	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
CO + O3	ACC (%)	86.53	85.06	95.94	91.43	**100.00**	**100.00**	**100.00**	99.70	**100.00**	**100.00**	**100.00**	**99.46**
	PRCN (%)	55.37	45.54	90.72	73.80	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	MR (%)	13.47	14.94	4.06	8.57	**0.00**	**0.00**	**0.00**	0.30	**0.00**	**0.00**	**0.00**	**0.54**
	Recall (%)	46.21	39.48	80.86	59.23	**100.00**	**100.00**	**100.00**	97.85	**100.00**	**100.00**	**100.00**	**96.14**
	F1 (%)	50.38	42.30	85.51	65.71	**100.00**	**100.00**	**100.00**	98.92	**100.00**	**100.00**	**100.00**	**98.03**
NO2 + O3	ACC (%)	86.56	84.46	95.71	88.57	**100.00**	**100.00**	**100.00**	99.70	**100.00**	**100.00**	**100.00**	**99.35**
	PRCN (%)	55.53	44.44	86.36	59.52	**100.00**	**100.00**	**100.00**	97.96	**100.00**	**100.00**	**100.00**	**95.62**
	MR (%)	13.44	15.54	4.29	11.43	**0.00**	**0.00**	**0.00**	0.30	**0.00**	**0.00**	**0.00**	**0.65**
	Recall (%)	40.31	35.00	83.94	62.50	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	F1 (%)	46.71	39.16	85.13	60.98	**100.00**	**100.00**	**100.00**	98.97	**100.00**	**100.00**	**100.00**	**97.76**
All	ACC (%)	88.44	87.98	97.09	92.50	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**99.88**
	PRCN (%)	56.40	56.63	86.06	73.26	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**100.00**
	MR (%)	11.56	12.02	2.91	7.50	**0.00**	**0.00**	**0.00**	0.00	**0.00**	**0.00**	**0.00**	**0.12**
	Recall (%)	83.54	76.42	94.99	76.83	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**99.19**
	F1 (%)	67.34	65.05	90.31	75.00	**100.00**	**100.00**	**100.00**	100.00	**100.00**	**100.00**	**100.00**	**99.59**
Averaged	ACC (%)	89.01	88.16	97.16	91.26	**100.00**	**100.00**	**100.00**	99.91	**100.00**	**100.00**	**100.00**	**99.81**
	PRCN (%)	61.70	60.04	89.94	70.16	**100.00**	**100.00**	**100.00**	99.70	**100.00**	**100.00**	**100.00**	**99.32**
	Recall (%)	61.03	58.96	89.92	69.82	**100.00**	**100.00**	**100.00**	99.71	**100.00**	**100.00**	**100.00**	**99.36**
	microF1 (%)	61.36	59.49	89.93	69.99	**100.00**	**100.00**	**100.00**	99.71	**100.00**	**100.00**	**100.00**	**99.34**
	macroF1 (%)	60.76	57.04	90.08	68.66	**100.00**	**100.00**	**100.00**	99.70	**100.00**	**100.00**	**100.00**	**99.34**



(a) (b)

Figure 8: Graphical representation of of the separation hydroplanes obtained by the support vector classifiers. Case $n = 10$. (a): Linear kernel function. (b): RBF kernel function.



(a) (b)

Figure 9: Graphical representation of of the separation hydroplanes obtained by the support vector classifiers. Case $n = 20$. (a): Linear kernel function. (b): RBF kernel function.

8. Conclusion

In this paper, we have investigated the use of a single MOX sensor, temperature modulated, to detect and identify multiple gas species and mixtures of these species. The temperature modulation allows the permanent changing of the sensor operating point, and thus its sensitivity to different gases. Then, a data-driven method that exploits and augments the signals delivered by the microsensor is proposed. The data augmentation is achieved through time-domain feature extraction. Increasing the raw database by

extracting temporal features enrich the prior knowledge available while maintaining a clear physical sense of all features. The significant increase in the feature space gives the potential to extend the algorithm to detect and identify a larger number of gases.

The newly extracted features were then ranked by the ReliefF algorithm. The highest-ranking features are then automatically selected as predictors for the MC-SVMs. The automatic selection of the relevant features (which carry the information necessary for the detection and identification of gases) reduces the space of the features used in the learning step and, therefore, the complexity of the MC-SVM classifiers. The Radial Basis Function is used as a kernel and gives promising results.

The experimental results, obtained using data from a test bench of the IM2NP laboratory at Aix-Marseille University show the relevance of temperature modulation, as the most ranked features, selected by the ReliefF algorithm are the heaters current and voltage. The performance analysis results highlights the effectiveness of this method by recording a 100% of accuracy with different metrics (Accuracy, Racall and F1 metric). Additionally, thanks to the availability of a wide choice of kernel functions, using the MC-SVM classifier allowed us to build models able to achieve perfect test scores using several configurations. This flexibility is the result of the well-chosen features that improved the separability of the classes. Furthermore, the ability to use multiple configurations allows us to strike balance between the speed of detection and the number of features necessary for accurate discrimination.

Acknowledgment

The authors would like to thank *Région Sud* of France and Nanoz- SAS for their financial support. We also thank Tomas Fiorido for his technical support throughout this work.

References

[1] R. Alrammouz, J. Podlecki, P. Abboud, B. Sorli, R. Habchi, A review on flexible gas sensors: From materials to devices, Sensors and Actuators, A: Physical 284 (2018) 209–231. doi:10.1016/j.sna.2018.10.036.

[2] H. Chojer, P. T. Branco, F. G. Martins, M. C. Alvim-Ferraz, S. I. Sousa, Development of low-cost indoor air quality monitoring devices: Recent advancements (jul 2020). doi:10.1016/j.scitotenv.2020.138385.

[3] J. van den Broek, D. Klein Cerrejon, S. E. Pratsinis, A. T. Güntner, Selective formaldehyde detection at ppb in indoor air with a portable sensor, Journal of Hazardous Materials 399 (2020) 123052. doi:10.1016/j.jhazmat.2020.123052.

[4] A. Schieweck, E. Uhde, T. Salthammer, L. C. Salthammer, L. Morawska, M. Mazaheri, P. Kumar, Smart homes and the control of indoor air quality, Renewable and Sustainable Energy Reviews 94 (2018) 705–718. doi:10.1016/j.rser.2018.05.057.

[5] P. Kumar, A. N. Skouloudis, M. Bell, M. Viana, M. C. Carotta, G. Biskos, L. Morawska, Real-time sensors for indoor air monitoring and challenges ahead in deploying them to urban buildings (aug 2016). doi:10.1016/j.scitotenv.2016.04.032.

[6] M. Penza, Chapter 12 - Low-cost sensors for outdoor air quality monitoring, in: E. Llobet (Ed.), Advanced Nanomaterials for Inexpensive Gas Microsensors, Micro and Nano Technologies, Elsevier, 2020, pp. 235–288. doi:https://doi.org/10.1016/B978-0-12-814827-3.00012-8.

[7] M. Mahbub, M. M. Hossain, M. S. A. Gazi, IoT-Cognizant cloud-assisted energy efficient embedded system for indoor intelligent lighting, air quality monitoring, and ventilation, Internet of Things 11 (2020) 100266. doi:10.1016/j.iot.2020.100266.

[8] I. J. Choi, B. J. Kim, S. H. Lee, B. J. Jeong, T. Nasir, Y. S. Cho, N. Kim, J.-H. Lee, H. K. Yu, J.-Y. Choi, Fabrication of a room-temperature NO2 gas sensor using morphology controlled CVD-grown tellurium nanostructures, Sensors and Actuators B: Chemical (2020) 128891doi:10.1016/j.snb.2020.128891.

[9] Y. Zhao, S. Wang, W. Yuan, S. Fan, Z. Hua, Y. Wu, X. Tian, Selective detection of methane by Pd-In2O3 sensors with a catalyst filter film, Sensors and Actuators B: Chemical 328 (2021) 129030. doi:10.1016/j.snb.2020.129030.
URL https://linkinghub.elsevier.com/retrieve/pii/S0925400520313770

[10] B. Yang, Z. Zhang, C. Tian, W. Yuan, Z. Hua, S. Fan, Y. Wu, X. Tian, Selective detection of methane by HZSM-5 zeolite/Pd-SnO2 gas sensors, Sensors and Actuators, B: Chemical 321 (2020) 128567. doi:10.1016/j.snb.2020.128567.

[11] K. C. Persaud, A. M. Pisanelli, S. Szyszko, M. Reichl, G. Horner, W. Rakow, H. J. Keding, H. Wessels, Smart gas sensor for monitoring environmental changes in closed systems: Results from the MIR space station, Sensors and Actuators, B: Chemical 55 (2) (1999) 118–126. doi:10.1016/S0925-4005(99)00168-9.

[12] S. Acharyya, B. Jana, S. Nag, G. Saha, P. K. Guha, Single resistive sensor for selective detection of multiple VOCs employing SnO2 hollowspheres and machine learning algorithm: A proof of concept, Sensors and Actuators, B: Chemical 321 (2020) 128484. doi:10.1016/j.snb.2020.128484.

[13] F. James, T. Fiorido, M. Bendahan, K. Aguir, Development of MOX Sensors for Low VOCs Concentrations Detection: Responses Comparison for WO3, SnO2, and ZnO Sensitive Layers with Interfering Gases as CO and CO2, International Journal on Advances in Systems and Measurements 10 (3 \& 4) (2017) 158–162.

[14] R. S. Lakshmi, A. Sivakumar, G. Rajaram, V. Swaminathan, K. Kannan, A novel hypergraph-based feature extraction technique for boiler flue gas components classification using PNN – A computational model for boiler flue gas analysis, Journal of Industrial Information Integration 9 (2018) 35–44. doi:10.1016/j.jii.2017.11.002.

[15] L. Wang, Metal-organic frameworks for QCM-based gas sensors: A review, Sensors and Actuators A: Physical 307 (2020) 111984. doi:https://doi.org/10.1016/j.sna.2020.111984.

[16] B. Sharma, A. Sharma, J. S. Kim, Recent advances on H2 sensor technologies based on MOX and FET devices: A review (jun 2018). doi:10.1016/j.snb.2018.01.212.

[17] A. Dey, Semiconductor metal oxide gas sensors: A review, Materials Science and Engineering: B 229 (2018) 206–217. doi:https://doi.org/10.1016/j.mseb.2017.12.036.

[18] T. P. Mokoena, H. C. Swart, D. E. Motaung, A review on recent progress of p-type nickel oxide based gas sensors: Future perspectives, Journal of Alloys and Compounds 805 (2019) 267–294. doi:10.1016/j.jallcom.2019.06.329.

[19] N. Barsan, M. Schweizer-Berberich, W. Göpel, Fundamental and practical aspects in the design of nanoscaled SnO2 gas sensors: A status report (1999). doi:10.1007/s002160051490.
URL https://link.springer.com/article/10.1007/s002160051490

[20] M. S. Yao, W. H. Li, G. Xu, Metal–organic frameworks and their derivatives for electrically-transduced gas sensors, Coordination Chemistry Reviews 426 (2021) 213479. doi:10.1016/j.ccr.2020.213479.

[21] B. Firtat, C. Moldovan, C. Brasoveanu, G. Muscalu, M. Gartner, M. Zaharescu, P. Chesler, C. Hornoiu, S. Mihaiu, C. Vladut, I. Dascalu, V. Georgescu, I. Stan, Miniaturised MOX based sensors for pollutant and explosive gases detection, Sensors and Actuators, B: Chemical 249 (2017) 647–655. doi:10.1016/j.snb.2017.04.032.

[22] Y. Yin, H. Yu, H. Zhang, A feature extraction method based on wavelet packet analysis for discrimination of Chinese vinegars using a gas sensors array, Sensors and Actuators, B: Chemical 134 (2) (2008) 1005–1009. doi:10.1016/j.snb.2008.07.018.

[23] R. Gosangi, R. Gutierrez-Osuna, Active temperature modulation of metal-oxide sensors for quantitative analysis of gas mixtures, Sensors and Actuators, B: Chemical 185 (2013) 201–210. doi:10.1016/j.snb.2013.04.056.

[24] N. Morati, T. Contaret, J.-L. Seguin, M. Bendahan, O. Djedidi, M. DJEZIRI, Data Analysis-Based Gas Identification with a Single Metal Oxide Sensor Operating in Dynamic Temperature Regime, in: ALLSENSORS 2020, The Fifth International Conference on Advances in Sensors, Actuators, Metering and Sensing, Vol. 2020, Valencia, Spain, 2019, pp. 20–23.
URL https://hal-amu.archives-ouvertes.fr/hal-02575436

[25] I. Montoliu, R. Tauler, M. Padilla, A. Pardo, S. Marco, Multivariate curve resolution applied to temperature-modulated metal oxide gas sensors, Sensors and Actuators, B: Chemical 145 (1) (2010) 464–473. doi:10.1016/j.snb.2009.12.051.
[26] Y.-C. Bo, P. Wang, X. Zhang, B. Liu, Modeling data-driven sensor with a novel deep echo state network, Chemometrics and Intelligent Laboratory Systems 206 (2020) 104062. doi:https://doi.org/10.1016/j.chemolab.2020.104062.
[27] S. De Vito, M. Piga, L. Martinotto, G. Di Francia, CO, NO2 and NOx urban pollution monitoring with on-field calibrated electronic nose by automatic bayesian regularization, Sensors and Actuators B: Chemical 143 (1) (2009) 182–191. doi:https://doi.org/10.1016/j.snb.2009.08.041.
URL http://www.sciencedirect.com/science/article/pii/S092540050900673X
[28] J. Chu, W. Li, X. Yang, Y. Wu, D. Wang, A. Yang, H. Yuan, X. Wang, Y. Li, M. Rong, Identification of Gas Mixtures via Sensor Array Combining with Neural Networks, Sensors and Actuators B: Chemical (2020) 129090doi:10.1016/j.snb.2020.129090.
URL https://linkinghub.elsevier.com/retrieve/pii/S0925400520314325
[29] E. Esposito, S. De Vito, M. Salvato, V. Bright, R. L. Jones, O. Popoola, Dynamic neural network architectures for on field stochastic calibration of indicative low cost air quality sensing systems, Sensors and Actuators B: Chemical 231 (2016) 701–713. doi:https://doi.org/10.1016/j.snb.2016.03.038.
URL http://www.sciencedirect.com/science/article/pii/S092540051630332X
[30] D. B. Topalović, M. D. Davidović, M. Jovanović, A. Bartonova, Z. Ristovski, M. Jovašević-Stojanović, In search of an optimal in-field calibration method of low-cost gas sensors for ambient air pollutants: Comparison of linear, multilinear and artificial neural network approaches, Atmospheric Environment 213 (2019) 640–658. doi:https://doi.org/10.1016/j.atmosenv.2019.06.028.
URL http://www.sciencedirect.com/science/article/pii/S1352231019304194
[31] N. Masson, R. Piedrahita, M. Hannigan, Approach for quantification of metal oxide type semiconductor gas sensors used for ambient air quality monitoring, Sensors and Actuators, B: Chemical 208 (2015) 339–345. doi:10.1016/j.snb.2014.11.032.
[32] A. Vergara, E. Martinelli, E. Llobet, F. Giannini, A. D'Amico, C. Di Natale, An alternative global feature extraction of temperature modulated micro-hotplate gas sensors array using an energy vector approach, Sensors and Actuators, B: Chemical 124 (2) (2007) 352–359. doi:10.1016/j.snb.2006.12.050.
[33] S. Zhang, C. Xie, M. Hu, H. Li, Z. Bai, D. Zeng, An entire feature extraction method of metal oxide gas sensors, Sensors and Actuators, B: Chemical 132 (1) (2008) 81–89. doi:10.1016/j.snb.2008.01.015.
[34] M. Amir Sattari, G. Hossein Roshani, R. Hanus, E. Nazemi, Applicability of time-domain feature extraction methods and artificial intelligence in two-phase flow meters based on gamma-ray absorption technique, Measurement: Journal of the International Measurement Confederation 168 (2021) 108474. doi:10.1016/j.measurement.2020.108474.
[35] K. Aguir, M. Bendahan, V. M. Laithier Martini, Heated sensitive layer gas sensor (aug 2020).
[36] F. E. Annanouch, G. Bouchet, P. Perrier, N. Morati, C. Reynard-Carette, K. Aguir, V. Martini-Laithier, M. Bendahan, Hydrodynamic evaluation of gas testing chamber: Simulation, experiment, Sensors and Actuators, B: Chemical 290 (2019) 598–606. doi:10.1016/j.snb.2019.04.023.
[37] C. Wang, L. Yin, L. Zhang, D. Xiang, R. Gao, Metal oxide gas sensors: Sensitivity and influencing factors (mar 2010). doi:10.3390/s100302088.
URL /pmc/articles/PMC3264469/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3264469/
[38] P. R. Van Der Meer, G. C. Meijer, M. J. Vellekoop, H. M. Kerkvliet, T. J. Van Den Boom, A temperature-controlled smart surface-acoustic-wave gas sensor, Sensors and Actuators, A: Physical 71 (1-2) (1998) 27–34. doi:10.1016/S0924-4247(98)00166-6.
[39] R. Ionescu, E. Llobet, Wavelet transform-based fast feature extraction from temperature modulated semiconductor gas sensors, Sensors and Actuators, B: Chemical 81 (2-3) (2002) 289–295. doi:10.1016/S0925-4005(01)00968-6.
[40] A. Vergara, E. Llobet, E. Martinelli, C. Di Natale, A. D'Amico, X. Correig, Feature extraction of metal oxide gas sensors using dynamic moments, Sensors and Actuators, B: Chemical 122 (1) (2007) 219–226. doi:10.1016/j.snb.2006.05.028.
[41] C. Arul, K. Moulaee, N. Donato, D. Iannazzo, N. Lavanya, G. Neri, C. Sekar, Temperature modulated Cu-MOF based gas sensor with dual selectivity to acetone and NO2 at low operating temperatures, Sensors and Actuators B: Chemical (2020) 129053doi:10.1016/j.snb.2020.129053.
URL https://linkinghub.elsevier.com/retrieve/pii/S0925400520314003
[42] T. H. Kim, S. Y. Jeong, Y. K. Moon, J. H. Lee, Dual-mode gas sensor for ultrasensitive and highly selective detection of xylene and toluene using Nb-doped NiO hollow spheres, Sensors and Actuators, B: Chemical 301 (2019) 127140. doi:10.1016/j.snb.2019.127140.
[43] J. Burgués, J. M. Jiménez-Soto, S. Marco, Estimation of the limit of detection in semiconductor gas sensors through linearized calibration models, Analytica Chimica Acta 1013 (2018) 13–25. doi:10.1016/j.aca.2018.01.062.
[44] S. Sridhar, A. Kalaivani, A survey on methodologies for handling imbalance problem in multiclass classification, in: Advances in Intelligent Systems and Computing, Vol. 1163, Springer Science and Business Media Deutschland GmbH, 2021, pp. 775–790. doi:10.1007/978-981-15-5029-4_67.
URL https://link.springer.com/chapter/10.1007/978-981-15-5029-4{_}67
[45] W. Caesarendra, T. Tjahjowidodo, A Review of Feature Extraction Methods in Vibration-Based Condition Monitoring and Its Application for Degradation Trend Estimation of Low-Speed Slew Bearing, Machines 5 (4) (2017) 21. doi:10.3390/machines5040021.
[46] X. Wang, Y. Zheng, Z. Zhao, J. Wang, Bearing Fault Diagnosis Based on Statistical Locally Linear Embedding, Sensors 15 (7) (2015) 16225–16247. doi:10.3390/s150716225.
[47] M. Djeziri, S. Benmoussa, E. Zio, Review of Health Indices extraction and Trend Modeling methods for Remaining Useful Life Estimation, in: M. Sayed-Mouchaweh (Ed.), Artificial Intelligence Techniques for a Scalable Energy Transition, 1st

Edition, Springer, 2020, pp. VIII, 292.
[48] K. Pearson, On lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2 (11) (1901) 559–572. doi:10.1080/14786440109462720.
URL https://www.tandfonline.com/doi/full/10.1080/14786440109462720
[49] A. Hyvärinen, Independent component analysis: Recent advances (feb 2013). doi:10.1098/rsta.2011.0534.
[50] I. T. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 374 (2065) (2016) 20150202. doi:10.1098/rsta.2015.0202.
URL https://royalsocietypublishing.org/doi/10.1098/rsta.2015.0202
[51] T. Marill, D. M. Green, On the Effectiveness of Receptors in Recognition Systems, IEEE Transactions on Information Theory 9 (1) (1963) 11–17. doi:10.1109/TIT.1963.1057810.
[52] A. W. Whitney, A Direct Method of Nonparametric Measurement Selection, IEEE Transactions on Computers C-20 (9) (1971) 1100–1103. doi:10.1109/T-C.1971.223410.
[53] P. M. Narendra, K. Fukunaga, A Branch and Bound Algorithm for Feature Subset Selection, IEEE Transactions on Computers C-26 (9) (1977) 917–922. doi:10.1109/TC.1977.1674939.
[54] X. W. Chen, An improved branch and bound algorithm for feature selection, Pattern Recognition Letters 24 (12) (2003) 1925–1933. doi:10.1016/S0167-8655(03)00020-5.
[55] P. Somol, P. Pudil, J. Kittler, Fast branch & bound algorithms for optimal feature selection, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (7) (2004) 900–912. doi:10.1109/TPAMI.2004.28.
[56] H. Almuallim, T. G. Dietterich, Learning with many irrelevant features, in: Proceedings of the Ninth National Conference on Artificial Intelligence, 1990, pp. 547–552.
URL https://dl.acm.org/doi/10.5555/1865756.1865761
[57] H. Almuallim, T. G. Dietterich, Efficient Algorithms for Identifying Relevant Features, Tech. rep., USA (1992).
[58] K. Kira, L. A. Rendell, A Practical Approach to Feature Selection, in: Machine Learning Proceedings 1992, Elsevier, 1992, pp. 249–256. doi:10.1016/b978-1-55860-247-2.50037-1.
[59] R. J. Urbanowicz, R. S. Olson, P. Schmitt, M. Meeker, J. H. Moore, Benchmarking relief-based feature selection methods for bioinformatics data mining, Journal of Biomedical Informatics 85 (2018) 168–188. arXiv:1711.08477, doi:10.1016/j.jbi.2018.07.015.
[60] D. Koller, M. Sahami, Toward Optimal Feature Selection, in: Proceedings of the Thirteenth International Conference on International Conference on Machine Learning, ICML'96, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1996, pp. 284–292.
[61] H. Liu, H. Motoda, L. Yu, Feature Selection with Selective Sampling, in: Proceedings of the Nineteenth International Conference on Machine Learning, ICML '02, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002, pp. 395–402.
[62] I. Kononenko, Estimating attributes: Analysis and extensions of RELIEF, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 784 LNCS, Springer Verlag, 1994, pp. 171–182. doi:10.1007/3-540-57868-4_57.
URL https://link.springer.com/chapter/10.1007/3-540-57868-4{_}57
[63] C. Cortes, V. Vapnik, Support-vector networks, Machine Learning 20 (3) (1995) 273–297. doi:10.1007/bf00994018.
URL https://link.springer.com/article/10.1007/BF00994018
[64] T. G. Dietterich, G. Bakiri, Solving multiclass learning problems via error-correcting output codes, J. Artif. Int. Res. 2 (1) (1995) 263–286.
[65] C.-W. Hsu, C.-J. Lin, A comparison of methods for multiclass support vector machines, IEEE Transactions on Neural Networks 13 (2) (2002) 415–425. doi:10.1109/72.991427.
[66] R. Batuwita, V. Palade, Class Imbalance Learning Methods for Support Vector Machines, in: Imbalanced Learning, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2013, pp. 83–99. doi:10.1002/9781118646106.ch5.
URL http://doi.wiley.com/10.1002/9781118646106.ch5
[67] H. He, Y. Ma, Imbalanced learning: Foundations, algorithms, and applications, Wiley, 2013. doi:10.1002/9781118646106.
URL https://onlinelibrary.wiley.com/doi/book/10.1002/9781118646106
[68] R. Akbani, S. Kwek, N. Japkowicz, Applying support vector machines to imbalanced datasets, in: Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science), Vol. 3201, Springer Verlag, 2004, pp. 39–50. doi:10.1007/978-3-540-30115-8_7.
URL https://link.springer.com/chapter/10.1007/978-3-540-30115-8{_}7
[69] Y. Liu, X. Yu, J. X. Huang, A. An, Combining integrated sampling with SVM ensembles for learning from imbalanced datasets, Information Processing and Management 47 (4) (2011) 617–631. doi:10.1016/j.ipm.2010.11.007.
[70] H. He, Y. Bai, E. A. Garcia, S. Li, ADASYN: Adaptive synthetic sampling approach for imbalanced learning, in: Proceedings of the International Joint Conference on Neural Networks, 2008, pp. 1322–1328. doi:10.1109/IJCNN.2008.4633969.