Journal of The Electrochemical Society

Review–Modern Data Analysis in Gas Sensors

To cite this article: Md. Samiul Islam Sagar *et al* 2022 *J. Electrochem. Soc.* **169** 127512

Journal of The Electrochemical Society, 2022 169 127512

Review–Modern Data Analysis in Gas Sensors

Md. Samiul Islam Sagar,[1] Noah Riley Allison,[1] Harikrishnan Muraleedharan Jalajamony,[2] Renny Edwin Fernandez,[2] and Praveen Kumar Sekhar[1,z]

[1]*School of Engineering and Computer Science, Washington State University, Vancouver, Washington 98686, United States of America*
[2]*Department of Engineering, Norfolk State University, Robinson Technology Center 410k, Norfolk, Virginia-23504, United States of America*

Development in the field of gas sensors has witnessed exponential growth with multitude of applications. The diverse applications have led to unexpected challenges. Recent advances in data science have addressed the challenges such as selectivity, drift, aging, limit of detection, and response time. The incorporation of modern data analysis including machine learning techniques have enabled a self-sustaining gas sensing infrastructure without human intervention. This article provides a birds-eye view on data enabled technologies in the realm of gas sensors. While elaborating the prior developments in gas sensing related data analysis, this article is poised to be an entrant for enthusiast in the domain of data science and gas sensors.
 [DOI: 10.1149/1945-7111/aca839]

Manuscript submitted September 27, 2022; revised manuscript received November 29, 2022. Published December 14, 2022.

Gas sensors are ubiquitous with myriad of applications. There is increasing demand for accurate and fast gas sensing system for applications ranging from environmental screening to space exploration.[1–13] Rational materials choice, simpler fabrication techniques and modern data analysis have enabled electronic gas sensors to be deployed in sophisticated and hazardous environments.[14–18] There exist different types of gas sensors, which uses different transduction principles to detect specific gases. Such modalities include electrical, optical, chemical, and electrochemical mechanisms.[19–22] An arrayed version of the gas sensors for detecting multiple gases simultaneously is called the Electronic nose (E-Nose). E-nose works by mimicking human odor-sensing or olfaction mechanisms.[23,24] In general, E-nose consists of a separator unit, a sensor or array of sensors, and a processing unit for signal processing and data analysis.[25] Recent advancements in materials and semiconductor technology have ensured reliable odor sensing. Moreover, compact and powerful micro- ICs have enabled miniaturization and fast processing of data without compromising on performance.[26]

It is essential to select, pre-process, and extract significant features for sensor post-data analysis[27–31] for monitoring its performance. External and internal factors degrade E-nose performance. Such factors include vicious change in temperature, pressure, excessive exposure to gases, vibrations, flexing of the substrates and/or, changing of the materialistic properties of the active material themselves.[32] These factors often cause drift in sensors. The drift often leads to a decrement in the life span of the sensors.

Moreover, the detection or, frequent calibration of sensors is expensive or, sometimes, impossible. Asymmetric parameter changes in the E-nose array leads to performance shift. The sensor shift, unlike the drift problem, can cause rapid sensor degradation. The detection of faulty sensor is also crucial when it comes to screening for a particular gas, where different data driven techniques can be handy. There exist different review articles to provide details on data driven topologies and analysis in the gas sensors. However, majority of them are either concise or cater to specific application of gas sensors, without considering the holistic significance of data analysis in gas sensor research.[33–47]

The goal of this manuscript is to summarize a collection of the different data aided techniques that are incorporated in modern gas sensing systems that include sampling and feature selection, data augmentation, classification, and data visualization. Next, the article illustrates important sensor features for data analysis. The evolution of gas sensor signal management tools has been described in detail. Then, different approaches for data analysis to address gas sensor challenges are presented. Finally, the article provides a summary and guidelines for the reader to adopt suitable tools for gas sensor analysis.

Salient Attributes of A Gas Sensor Signal

For any type of signal analysis, it is critical to identify the significant information that can be extracted from that signal. The performance of any data analysis algorithm depends on the quality of derived attributes. For gas sensors, the signal collected from the sensors can be classified into three parts: increasing transient response, the steady-state response, and the decreasing transient response,[48] shown in Fig. 1.

Different time domain and frequency domain features of the signal allows to identify performance metrics. The gas of interest often needs to be mixed up with oxygen and sometimes heated up prior exposing to the sensor. The ratio of oxygen,[49] temperature variation,[50] gas concentration,[51] and flow rate play important roles in the gas sensor response. Specifically, the response time, recovery time, selectivity, and sensitivity are some of the factors which are influenced by different excitation conditions.[52] Some derivative features of sensors are employed in research investigations to describe the transient behavior of the sensors.[53] For model development and data analysis, an easily accessible dataset is critical. Such a dataset can be found from Wijaya et al.[54] In this dataset, ten gas sensors have been used to form a time series data having four distinct classes of beef and continuous microbial population index as target labels with two intrinsic and 4 extrinsic sensor features. In another dataset from a prior article,[55] an E-nose of six sensors was implemented to record 235 wine quality tester data. To classify Chronic Obstructive Pulmonary Disease (COPD) affected people, Acevedo et al.[56] gathered 78 samples from different class of people using E-nose. A large dataset for drift analysis of gas sensors was collected and open-sourced in UCI repository.[57] In this dataset, there are sixteen sensors with eight features each combining 128 features for six gases to be classified.

Gas Sensor Data Analysis

Modern data driven techniques such as regression models, classifiers, deep learning techniques, and machine learning have paved the way of converting raw features into actual and meaningful information. Successful feature extraction and selection techniques improve the performance of such data driven algorithms. This Section describes the feature and model-based approaches that were developed in the past few years in the gas sensors area.

[z]E-mail: praveen.sekhar@wsu.edu

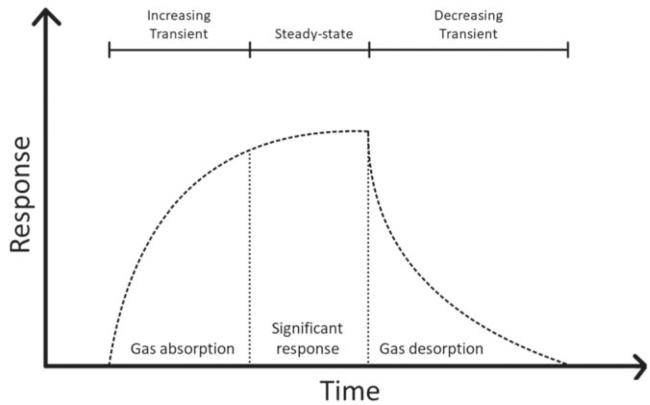

Figure 1. A typical gas sensor output in time domain.

Feature oriented systems.—Feature is a meaningful data that holds dense information for making certain decisions with respect to field deployability of sensors. Such features include correlation and convoluted information among different sensor responses, distribution based information of a dataset, maximum variance and neighborhood information of sensor data, and so on. It is important to choose the right features to ensure accurate decision making. Feature extraction involves algorithms.

The performance of these algorithm can be increased in many ways resulting in decreased time of operation, real time analysis, online monitoring ability, decreased computation complexity, and computational cost. Further, many mathematical models have come up with different solutions to turn the non-linear features of a gas sensor into linear trends, which enables better results with simpler learning models. Also, different augmentation techniques ameliorate the usage of data to extract unknown and interconnected valuable features from the original dataset.

Attribute and instance selection.—It is necessary for feature selection algorithms to have proper selection of samples to achieve maximum efficiency. Yu et al.[58] suggested active sampling approach for query by committee (QBC) algorithm which showed significantly increased performance with low dataset size. Further, Liu et al.[59] proposed a novel min-max confidence strategy for selecting samples toward an online classifier which works better than the conventional QBC approach when addressing gas sensor drift.

Later, Liu et al.[60] proposed an adaptive sampling method for drift compensation (via online learning) inspired by QBC approach where posterior probability needs to be calculated for optimal instance selection.

Unlike the conventional approach, the study offered a pool-based criteria selection method for better adaptation for classifier models. An active and transfer sub sampling-based approach provided significant variance among classes of a dataset reported in a study.[61] An active instance selection strategy on mixed kernel

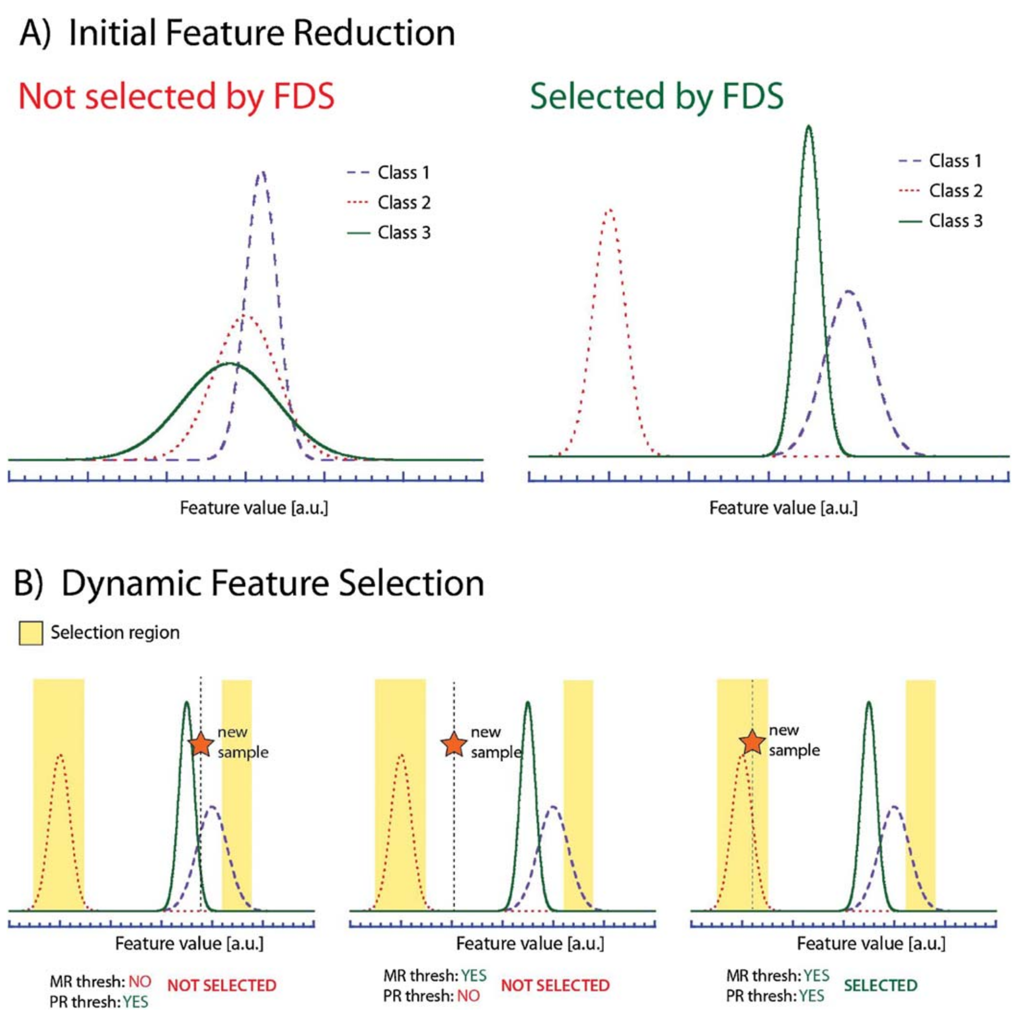

Figure 2. Fisher Discriminant Score (FDS) based feature selection scheme in a prior article. (A) Comparison between FDS and non-FDS approaches for initial feature reduction. The feature will be selected if at least one class distribution is different from the others. (B) three different cases of dynamic feature selection. A feature is selected only if the sample can be attributed to only one class distribution.[66]

mechanism was proposed by Liu et al.[62] which can compensate drift. The study showed that the algorithm employed with active learning provides maximum recognition accuracy of 83.59%. For sensors exhibiting drift, a calibration sample selection method was employed by Liu et al.[63,64] to tackle class imbalance problem. This showed an average accuracy up to 97.64% over other passive sampling-based approaches. Feature selection has become an emerging research topic in recent years due to its importance in data aided performance improvement of gas sensors. In this context, Deng et al.[65] proposed a unique feature fusion framework using separability and dissimilarity index of features which provided improved classification accuracy with simple models like K-nearest neighbor (KNN). A dynamic feature selection based on the closeness to expected class distribution was proposed in another article[66] where Fisher Discriminant Score (FDS) was used for data sorting based on the application (Fig. 2). According to the Fig. 2, the samples are dynamically selected to achieve expected variance to the selected dataset. Multiple feature selection algorithms such as chi-square, reliefF, and gini index was employed in another study[67] which seemingly improved the quality of the learning models. Further, Czarnowski et al.[37] described a weighted ensemble technique of instance selection and oversampling in the case of data imbalance.

Many unsupervised optimization algorithms are employed to extract intrinsic features from actual data. A modified binary ant colony optimization (ACO) was proposed by Shu et al.[68] to select features with minimum redundancy to reduce computational time. A feature selection approach based on the orthogonal correlation among features was proposed in prior articles[69,70] where the actual features are converted to a orthogonal feature space using orthogonalized scores and weights as described in Eq. 1, where $t_\perp$ is the orthogonal score, and $p'_\perp$ is the loadings. Substituting their product from the actual data space X provides orthogonal components free subspace X_{OSC}.

$$X_{OSC} = X - t_\perp p'_\perp \tag{1}$$

Data augmentation.—Researchers are looking into inherent correlation between different raw sensor attributes. Such correlations can be generated using many mathematical models[71] to address sensor performance challenges under field conditions. In most cases, the gas sensor responses are non-linear in nature. This non-linearity often leads to incompatibility in optimization of any learning algorithm or result in an expensive and complex system design.

Vergara et al.[48] provided an approach where the non-linear transient features have been converted into linear features using the Eq. 2. The $y[k-1]$ is the exponential moving average value of the sample $k-1$ in the discrete time series, α is the weighting factor, and $r[k]$ and $r[k-1]$ are the two consecutive responses from the discrete time series response.

$$y[k] = (1-\alpha)y[k-1] + \alpha(r[k] - r[k-1]) \tag{2}$$

Burgues et al.[72] proposed a logarithmic transformation of the non-linear data for low concentrations of the analyte. Utilizing cosine similarities between different combinations of features defined in Eq. 3, Rehman et al.[73] extracted metaheuristic drift-insensitive features from raw data. In the Eq. 3, t represents the test sample, m is the median of the selected features in training data from class k, and s is the total number of selected features.

$$CS = \frac{\sum_{i=1}^{s} t_i m_i^k}{\sqrt{\sum_{i=1}^{s} t_i^2}\sqrt{\sum_{i=1}^{s} (m_i^k)^2}} \tag{3}$$

Another study employing the cosine similarity-based feature extraction method showed improved performance in long term drift affected dataset.[74] The Tree Structured Cosine Similarity based Shuffled Frog Leaping Optimization (TSCS- SFLO) algorithm proposed in a prior study showed an overall detection accuracy of 91.34% over three years period of drifted gas sensors data. However, the study from an earlier article[65] showed correlation-based distance measurement exhibiting better performance over cosine based distance calculation algorithms.

A canonical correlation analysis of data was proposed in an earlier study[41] to utilize the maximization of correlation between two sets of data for improved sensor performance. Yan et al.[40] proposed a method of subspace alignment among features by optimizing the distribution discrepancy between source and target domain. To minimize the maximum mean discrepancy among the features, many researchers employed joint distribution adaptation-based transfer learning where Principal Component Analysis (PCA) was used to measure the maximum variance of the dataset. Liu et al.[75] extracted features using Fisher linear discriminant approach for drift compensation which decreases the inter-concentration discrepancy of the feature distribution, resulting in concentration independent features. Balanced distribution adaptation algorithm showed better performance over joint distribution adaptation according to Jiang et al.[76] where the marginal and conditional distributions are considered. An earlier article[77] combined simple classifiers with these features, resulting in an improved accuracy.

Orthogonal representation of a signal often provides significant information on domain distribution of a dataset. A kernel PCA based approach with XGBoost classifier resulted in enhanced the accuracy, the sensitivity, and the specificity of gas sensor when recognizing different lung cancer stages.[78] Apart from PCA and orthogonal projection, projection on convex set showed improved performance while combining with extreme machine learning approaches to achieve global calibration instead of local calibration.[79,80]

A recursive feature elimination-based approach was proposed in a prior article[81] to extract critical features that compensates small scale drifting. PCA can also be used for feature reduction purposes.[82]

Wozniak et al.[83] utilized Fast Fourier Transform (FFT) to extract features from the frequency domain of the sensor signal via a simple regression model. In another study, Discrete Fourier Transform (DFT) analysis was utilized to observe the H_2S gas absorption with respect to the doping of osmium.[84] A wavelet packet decomposition algorithm was adopted in another study[85] for no load data decomposition to correct the drifting effect over time. Similar approach of sample test time window (SMTW)[86] successfully rectified long term drift component from the gas sensor dataset. A Wasserstein Distance Learned Feature Representations (WDLFR) was proposed in another study[87] to optimize the Wasserstein distance instead of Euclidean distance for domain invariant feature extraction in noise and drift affected gas sensor dataset.

Features sometimes need to be reshaped for data analysis. Instead of using 1D sensor signal data, an augmentation was incorporated[88] to make the sensor signal compatible with the 3D input feature space of a conventional convolutional neural network (CNN) as illustrated in Fig. 3. A deep recurrent neural network illustrated by Wang et al.[89] derived extreme non-linear features of drift contaminated gas sensors without incorporating any further formulations. Most of these complex models can use the original sensor data directly and learn latent information about the sensor data.

Signal denosing applications.—Denoising is an integral part of signal pre-processing for many sensing applications. Noise in the sensor data often cause false predictions. There are two types of noises: random and systematic. Random noise in sensor signals is inevitable and need transformational algorithms to change the initial domain of dataset into another (such as time domain to frequency domain) for better estimation. The systematic errors can be caused by drift or, shift and require recalibration of the sensor, which is both time consuming and expensive. Researchers are working on different techniques to reduce in gas sensor dataset,[35,90] which will be demonstrated later in this section.

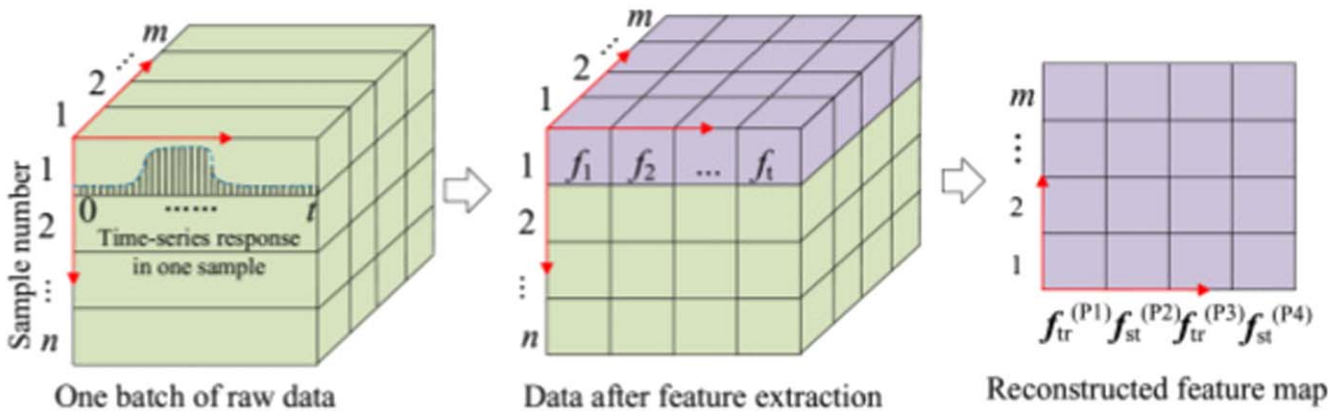

Figure 3. E-nose data transformation for augmented convolutional neural network.[88]

Zhifang et al. categorized interference in E-nose sensors into two types.[35] Interference caused by the operating condition of the sensor, such as sudden environmental change, carrier gas, and dynamic interference. Another type includes system effect which refers to the change in sensor characteristics due to drift, hardware failure etc. These interferences are non-linear in nature, which makes it difficult to suppress. Mojtaba et al. implemented a MANOVA based statistical approach to track the significance among independent features, to learn if a dataset needs baseline correction due noise.[91] Discrete wavelet based tranformations (DWT) are popular in removing baseline noise. DWT was illustrated in a article by Shen et al. in a gas detection system for baseline correction[92] and by Wijaya et al. in beef quality monitoring application.[93] The same author, in a later article, proposed a variant of discrete wavelet transform combining with long short term memory (LSTM), which utilizes the inherent features of the sequence of reading from an E-nose sensor.[94] LSTM was employed by Wang et al. with CNN to reduce noise.[95] Yang et al. showed that by optimizing the K values while using K-means clustering could achieve significant noise reduction.[96]

Another popular algorithm to eliminate noise from gas sensor dataset is the moving averge method. The algorithm was implemented by Ima et al. to suppress noise.[97] Another article, by Albert et al. also used moving average technique for noise cancellation.[98] PCA was utilized to reduce noise while random tree was used for mixed gas classification by Yonghui et al.[99] Morati et al. implemented PCA for noise reduction for single MOX gas sensor.[100] Alongside a filtering algorithm, Liu et al. proposed a PCA based squared prediction error technique to eliminate random noise classified as outliers.[101] Jianyou et al. proposed a sparse optimization algorithm for removing the coupling noise in a distributed acoustic sensing (DAS) system.[102] Maximum independence of concentration features was calculated to reduce the inter-concentration discrepancy of distribution thereby improving the noise influence on the dataset.[75] Chao et al. employed systematic and measurement noise inside a Kalman filter equation, which was successful in reducing overall noise, hence increasing the signal to noise ratio.[103] A CNN based denoising autoencoder technique showed significant efficiency in noise removal.[104] Zhifang et al.[41] showed a sparse reconstruction using domain transfer technique to eliminate noise resulting in increased gas sensor performance.

Model based algorithms.—Model based algorithms have now become an eminent tool for decision making in many unforeseen and unprecedented circumstances observed in real time gas sensor applications. There exist different types of modeling approaches according to the data availability, performance, level of complexity, and usability. Based on the learning pattern, the model-based algorithms have been primarily divided into supervised and unsupervised learning.

Supervised approaches.—Supervised learning refers to the technique of learning internal structure of the features with the help of labels. They are getting popular due to the increased availability of labeled data.

In many gas sensing applications, supervised learning methods showed tremendous success in improving the performance, robustness, and device reliability.[105] Different supervised algorithms such as support vector machine (SVM), random forest, XGBoost, K-nearest neighbor (KNN), different neural networks are widely being implemented to address challenges like drifting, fault detection, calibration, and classification etc.[106–110] Models like SVM and KNN provide expected performance in online active learning applications even when encountering sensor drifting challenges.[59,111]

Matthews et al.[61] described a non-linear classification approach implementing an active and transfer sampling method for low cost SnO_2 based semiconducting gas sensors. While using 75% of the total dataset which is 1,35,000, this approach achieved accuracy above 90% with Radial Basis Function (RBF) based SVM. Shu et al.[68] described a random forest regression algorithm to compensate drift of gas sensors. An online inertial learning enabled SVM classifier was proposed in an article[112] where for each sample, the training model is poised to change in a short amount of time, making it conducive for real world applications. In another study, XGBoost facilitated better results compared to SVM when combining kernel PCA based feature extraction.[78] In lung cancer diagnosis application,[113] random forest with 5-fold cross- validation provided good accuracy over logistic regression with an accuracy of 85.38% and 0.87 area under the curve (AUC). KNN provided promising results (when integrated with enhance feature selection algorithms) with an increased accuracy up to 97.5% in a drifted E-nose dataset.[76] It was used alongside balanced distribution adaptation (BDA) optimization of features. In another study, KNN showed promising result in determining the perishable quality of shrimp when used with softmax regressor.[114]

A domain transfer-based algorithm was described in a prior study[115] where the feature level and the decision level information of the training data was utilized using joint distribution adaptation mechanism. Another domain transfer-based approach proposed by Liang et al.[41] utilized external factor interference suppression in classifying gases. A feature transfer-based algorithm was proposed in another article[116] to increase accuracy for drifted E-nose sensors. A joint distribution adaptation-based algorithm was demonstrated in another study[77] where KNN was implemented after minimizing the maximum mean discrepancy (MMD) among features. The study showed an accuracy of 97.29% for some of the batches. Another study that employed MMD[117] where the combination of manifold learning and domain adaptation was used to increase performance of the classifiers resulting in the reduction of the long- and short-term drifting of gas sensors. To eliminate domain discrepancy, Zhengkun et al.[82] proposed an approach of aligning source and target domain distribution, which showed an average accuracy of 66.05% in gas classification.

The domain adaptation-based algorithm described in prior articles[33,118,119] was based on optimal subspace selection for each dataset to achieve better accuracy. Yi et al.[120] proposed another approach of neighborhood preservation using two novel terms: intraclass and interclass distance. After optimizing the distance among different classes, the resulting features showed enhanced performance when using a regular classifier to classify drifted gas sensor dataset.

The idea of the intra class and inter class distance analysis between features is gaining popularity due to their ability to model performance enhancement. The authors from the previous article[121] proposed a similar idea where the mean distance between samples was optimized prior classification. This approach showed promising results in drift affected sensors with an average accuracy of 76.31%. A Time-Wasserstein dynamic distribution alignment algorithm was proposed in an article[122] for learning the feature distribution dynamically in time domain thus minimizing the distribution discrepancy in drifting sensors.

Ensembling techniques provide better performance when compared to standalone models with a single classifier.[123,124] An ensemble technique was proposed in a previous study[67] and evaluated considering the F-measure and mean squared error (MSE), where MSE was found to be less than 0.001 for different classifiers. An ensemble tree mechanism was adopted prior[125] for overcoming delayed response of models incorporating only the transient feature of the sensor. This method achieved an overall accuracy of 87.34%. A similar type of ensemble technique was found in an earlier article[126] where a small number of labeled samples was required for training to achieve the desired accuracy. A 2D Convolutional Neural Network (CNN) based ensemble technique showed significant accuracy with drifted dataset as demonstrated in a prior article.[127] The study claims of an accuracy of 91%. Ensemble techniques can perform better with generalized neural network.[128] A multi-dimensional CNN ensembling technique was adopted by Chaudhri et al.[129] which outperformed simpler classifiers such as SVM providing a robust classification for drifted gas sensors. A decision level drift compensation scheme was proposed by Tian et al.[130] where a unified classification model was implemented incorporating a Gaussian deep belief classification network.

Feng et al.[88] proposed an augmented CNN modeling approach where a feedback network updates the base model with necessary adjustment in case of drifted dataset. Another CNN model with 10 layers was proposed in a prior study[131] for volatile organic compound (VOC) detection which achieved promising results with minimized root mean squared error (RMSE). For specific elimination of the effect of temperature and humidity in sensor signal, a deep back propagation neural network was proposed in a prior study[132] where 14 hidden layers have been used and optimized for analyte classification. A novel memory mechanism was proposed to address sensor drift in gas sensing applications[133] where a simple feedback algorithm has been utilized to let the model learn about the deviation and thus calibrate the model parameters. A multilayered perceptron SimResNet-9 was implemented in another article[134] which improved accuracy of the model from 88.0% to 93.7%. A hybrid KNN-ANN network proposed in a prior study[135] that achieved 96.51% of accuracy for drifted gas sensor data classification.

For detecting fault in the sensor systems, Tan et al.[136] employed Naive Bayes and probabilistic neural network (PNN) to monitor and classify faulty sensors and systems. Chaudhuri et al.[137] proposed an attention based gated recurrent unit (GRU) modeling for drift compensation in gas sensors. The use of the model increased the average accuracy to 93% for drift affected dataset. Using specific concentrations from the target, Cheng et al.[138] demonstrated an algorithm to compensate drift from a time varying dataset.

Liu et al.[75] proposed a concentration independent drift compensation approach utilizing transfer learning technique with an average accuracy of 76.17%. A short-term memory and SVM incorporated ensemble technique was mentioned in an earlier article.[139] The results showed an average accuracy upto 89.3% for a time series drift affected dataset of an E-nose system.

Apart from classification, some applications need regression models where the concentration of a particular gas needs to be measured. Partial Least Squared (PLS) algorithm has been used in many applications for fast and accurate prediction.[140] Wozniak et al.[83] proposed a simple PLS regression model which provided as low RMSE (7.34) toward estimating gas concentrations. Another study adopting PLS with domain adaptation showed reduced relative standard deviation of the sensor signals ranging between 91.5% and 99.7% depending on long or, short sequences.[141] PLS was also used in another study[142] for recognizing volatile organic gases. Zhang et al.[143] proposed a target domain free approach which provided promising results without incorporating the data from the unknown target domain (Fig. 4).

According to the Fig. 4, the technique converts raw input data to a stream of sensor delimited data which enters into multiple convoluting blocks to extract inherent features, then ensembles the features for gas prediction. Wang et al.[144] proposed a domain reconstruction-based approach to suppress drifting effect of gas sensors which achieved 86.88% of average accuracy without any target domain dataset for the training.

Unsupervised approaches.—The main drawback of supervised learning is the need for a lot of labeled data. It is practically hard to have enough labeled data for supervised learning as it becomes expensive to collect and label gas sensor data manually. Unsupervised and semi-supervised learning approaches are gaining popularity in recent years.

Different optimization algorithms based on natural phenomena such as particle swarm, ant colony, genetic algorithm is showing significant performance with scarcity or absence of a labeled dataset.

A fuzzy clustering approach was proposed earlier[145] to eradicate the overlapping problem of clusters while differentiating aromatic and non- aromatic rice (Fig. 5). A semi-supervised classifier was proposed earlier[66] where a novel feature selection approach was utilized to better sensor performance. Another study[60] showed that the average accuracy of active learning on adaptive confidence rule was greater than the conventional supervised and semi-supervised approaches. Further, semi-supervised approach for drift rectification of gases was proposed in an earlier study[146] which could classify unlabeled samples collected real time.

Rehman et al.[73] described such an approach where a discrete particle swarm optimization (PSO) technique was utilized. PSO can be found alongside with many feature augmenting and optimization algorithms for increasing the performance of different classifiers and regressors.[76] Another study[147] provided evidence of efficient parameter correction using PSO for practical unsupervised drift compensation of gas sensors.

Another PSO based recursive metaheuristic optimization technique was proposed prior[148] for anti-drift feature selection thus increasing the performance of simple classifiers such as random forest.

It is often difficult to obtain enough data to calibrate gas sensors over time. To minimize the cost of the system, a correlation-based approach was proposed by Maho et al.[149] where the sensors were calibrated blindly without the aid of any data. Yan et al.[40] proposed a PSO based subspace alignment method of classifying gas sensor data with an average accuracy of 90.07% using an existing gas sensor drift dataset. Ouqamra et al.[150] described a blind source feature separation framework that can eliminate baseline drifts as well as mixed noises without actual labeled data.

A hybrid pattern recognition model based on PCA and K-means clustering was proposed in another study[151] for explosive detection. An example of deep reinforcement learning was reported in a prior study[152] where a deep Q-network was used to self-calibrate the sensors of a remote system. Genetic algorithm is also getting popular

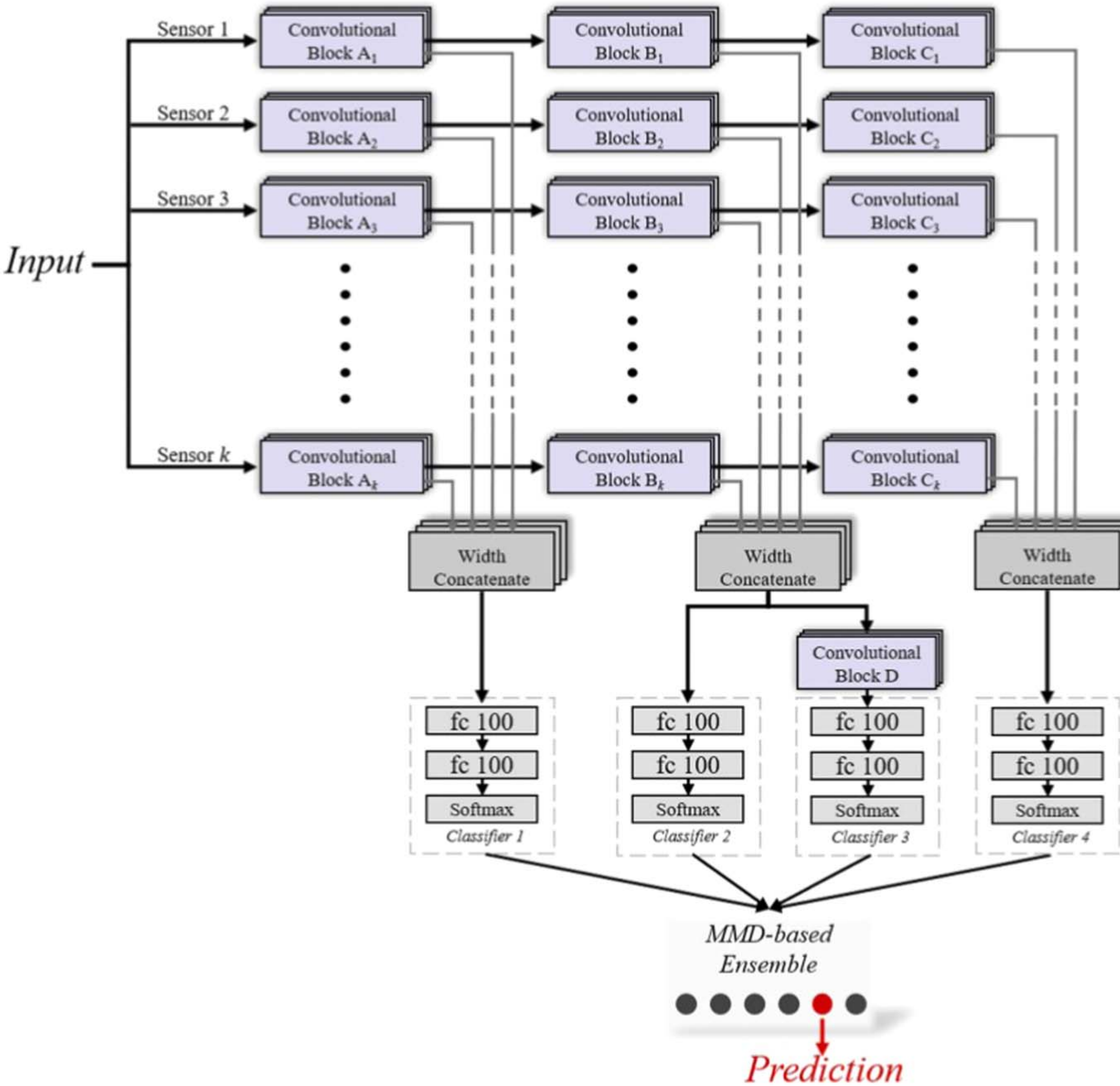

Figure 4. Example of an augmented ensemble classifier from Ref. 143.

toward feature selection for fast and reliable decision-making.[124] In Table I, a collection of different data-driven studies to address gas sensor challenges and their performance metrics is illustrated.

Evolution of Managerial Tools

Critical tasks for any data-oriented system are to properly control, calibrate, segment, visualize, and direct the flow of the raw data gathered from the gas sensors online and offline.

Interface design.—With the rapid advancement of graphical user interface, it has become necessary to improve the visualization and ease of control in gas sensing systems. Many studies adopted feature reduction processes prior to visualizing signal data as it is hard to demonstrate the distribution of features. Researchers explore different correlation and component selection-based approaches such as principal component analysis, exponential transformations, and variance calculation for better visualization of features.[40,41,48,118,121] In many real-time applications, E-noses are used with proper visualization tools.[153–155]

Debabhuti et al.[156] developed a graphical user interface (GUI) for Quartz crystal Microbalance sensors and implemented in a Raspberry Pi module. A GUI with interactive control and visualization system was developed by Djelouat et al.[30] which can also analyze the responses and display the extracted features of the gas sensor system real time. To learn different features from the data distribution, Liu et al.[29] developed a data visualization-based approach, where CNN was utilized to extract small scale visual features from different graphs automatically instead of manual calculations.

A user interface for fast and reliable response-recovery analysis was developed earlier[157] which analyzed 8 samples of data from gas sensors simultaneously and provide result instantaneously. PyQt library alongside with NumPy, SciPy, Pandas, Matplotlib libraries were used to develop this graphical user interface (Fig. 6).

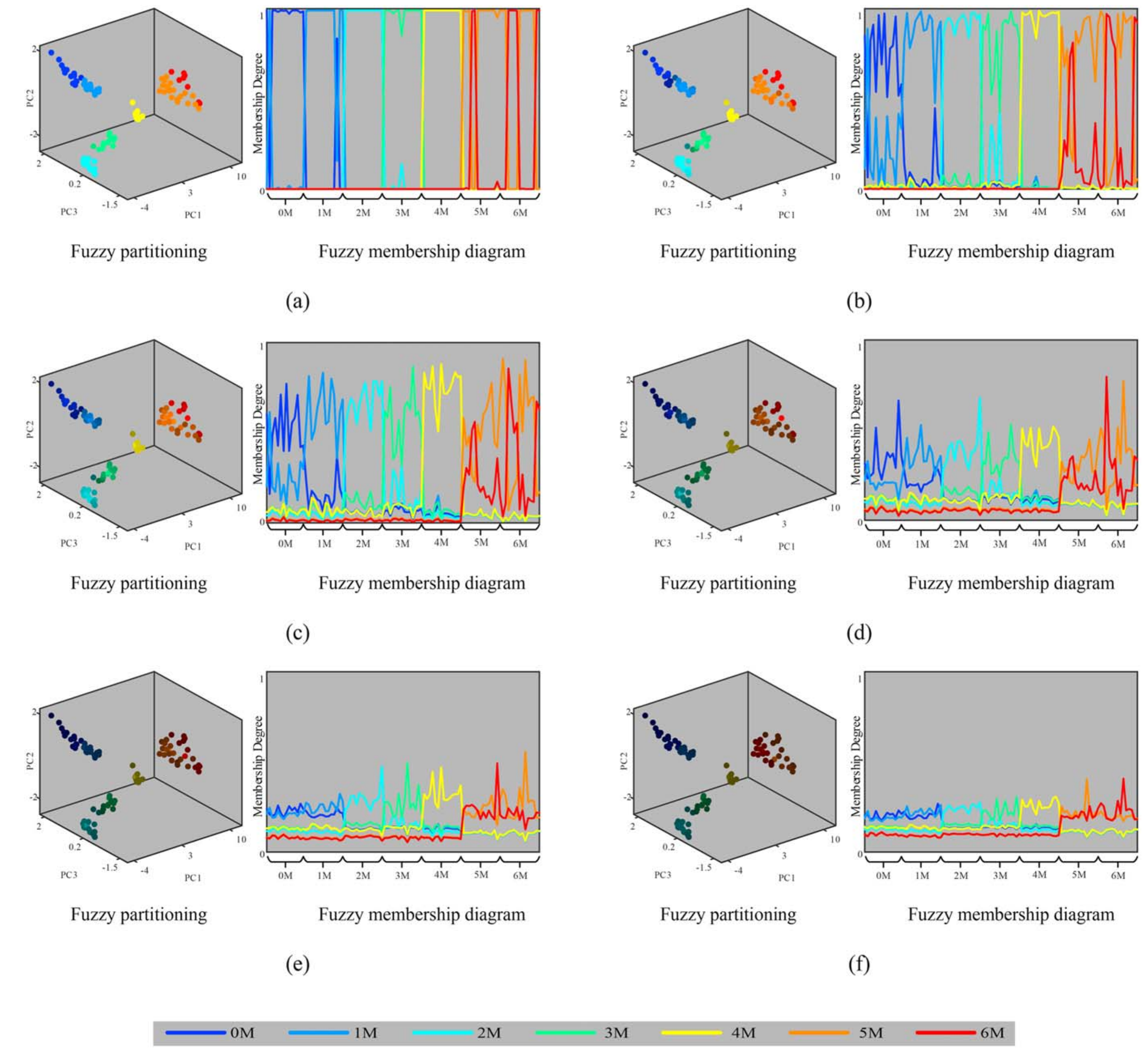

Figure 5. Fuzzy partitionings of non-aromatic rice samples and related membership diagrams for different fuzzifier values: (a) f = 1.1, (b) f = 1.5, (c) f = 2.0, (d) f = 3.0, (e) f = 4.0, and (f) f = 5.0.[145]

A simulator was developed in a prior investigation[158] to identify faulty sensors using sensor validity index. Again, Bastuck et al.[28] designed a Matlab toolbox to extract different features from cyclic sensor data with a GUI. The toolbox can also be used for different data driven operations such as exhaustive parameter search, automatic data fusion and model hierarchies (Fig. 7). A mobile application of e-nose was found in another study[18] where a simple threshold-based algorithm is employed with a GUI interface to detect victims in hazardous area in real time.

Dataflow system.—Dataflow is one of the most important areas of a gas sensor system design. Real-time data transformation, drift compensation, calibration and analysis often lead to trade off in performance over time required for a selected algorithm for a definite application. Liu et al.[59] implemented SVM and KNN based models for online active learning applications. The performance of such systems can be increased by extracting as much information as possible from the raw dataset. Rehman et al.[73] provided such an approach where cosine similarity was utilized to minimize the learning time and memory usage of an online system. Active sampling showed promising outcome in online instance selection for increased sensor system performance.[61]

An unsupervised approach was described in an earlier article[66] for online applications where fault or high-level noise can risk the performance of the sensor system. Rehman et al.[125] provided a faster approach by considering less features than usual to maintain high accuracy. The multiclassifier tree model showed significant robustness in handling drifted data. Active sampling provides promising results than passive sampling for online drift calibration according to Liu et al.[63] Universal calibration of the system was achieved in an earlier article[79] by using projection on to convex set (POCS) and extreme learning machine (ELM). To properly track the air quality, a stochastic model was proposed earlier[27] that illustrates the behavioral pattern of Graphene-based electrochemical gas sensors by simulating specially mapped sensor data related to the microscopic features of a material. For self-calibration of Internet of things (IoT)

Table I. Performance study of different data-driven techniques in gas sensing technology.

Algorithms	Applications	Performance matrics	References
SVM	Drift compensation, Classification	Accuracy: 89.98%–96.62%	59, 61
XGBoost	Classification	Accuracy: 96.62%, Sensitivity: 95.60%, Specificity: 91.09%	61, 78
Random forest	Classification	Accuracy: 85.38%, AUC: 0.87	113
KNN	Drift compensation	Accuracy: 80.74%–97.5%	59, 65, 76, 114
CDC	Drift compensation	Avg accuracy: 91.85%	138
KNN-ANN	Drift suppressed classification	Accuracy: 96.51%	135
GDBCN	Drift compensation	Avg accuracy: 73.51%	130
PLS regressor	Gas concentration prediction	RMSE: 7.34	83, 176
AL-ISSMK	Drift suppressed classification	Avg accuracy: 74%–83.59%	62
OPLS	Drift suppressed classification	Accuracy: 91%	69
PLS-DA	Drift suppressed feature augmentation	Accuracy: 91%	141
AL-ACR	Classification	Avg accuracy: 66.60%	60
Domain transfer	Drift and interference suppression	Avg accuracy: 79.4%–95.92%	41, 77, 116
DANP and DANP+	Drift suppressed classification	Avg accuracy: 77.83%–79.76%	120
TWDDA	Drift suppressed classification	Avg accuracy: 80.99%–92.44%	122
Multiclassifier ensemble	Fast response classifier	Avg accuracy: 87.34%	125, 126
LDSP	Drift suppressed classification	Avg accuracy: 76.31%	121
SAELM	Drift suppressed classification	Avg accuracy: 90.07%	40
ANN	Classification	Accuracy: 91.26%	61
CSFT-AELM	Drift suppressed classification	Avg accuracy: 92.44%	115
SSCA-DA	Drift suppressed classification	Avg accuracy: 86.54%	118
LME-CDSL	Drift suppressed classification	Avg accuracy: 70.95%–73.96%	117
PNN	Fault classification	Accuracy: 83.05%–99.89%	136
ACNN	Drift compensation	Accuracy increased over 30% worst case	88
MLPNN	Gas concentration estimation	Error decreased 7%–19% worst case	167
1d CNN	Hazard monitoring and classification	Accuracy: 99.6%	165, 166
2d CNN	Online drift suppressed classification	Accuracy: 91%	127
Deep CNN	Real-time classification	Accuracy: 98.1%	162
CNN ensemble	Classification	Accuracy: 99.72%	129
TDACNN	Drift suppressed classification	Avg accuracy: 81.48%	143
CGDA	Drift suppressed classification	Avg accuracy: 93%	137
PSO	Drift compensation	Accuracy: 86.01%	73
WWH-SSO	Online drift suppressed classification	Avg accuracy: 78.03%	146
LSTM-SVM ensemble	Drift suppressed classification	Avg accuracy: 89.3%	139
WDLRF	Feature extraction	Avg accuracy: 82.55%	87
SFLO-TSCS	Feature extraction and classification	Avg accuracy: 91.34%	74

based sensor systems, the hidden Markov model was used for characterization of a single sensor.[159] You et al.[152] proposed a complete scheme of calibrating IoT gas sensors which eradicates the Markov decision process problem by utilizing a deep Q-network. To suppress faulty sensor data from inserting to the classifier, Magna et al.[160] proposed a self-repairing scheme, where the faulty sensor will be replaced by a replica to increase the consistency of model performance. An ultrafast algorithm based on opto-E-noses was proposed in a prior study[161] to detect infectious bacteria in metallic nanoparticles. A deep CNN model was employed for real time gas identification with E-nose sensors in a prior study.[162] Another CNN based approach for real-time gas classification can be found in an earlier article[163] with 100% accuracy.

Remote management and automation.—The importance of gas sensors cannot be emphasized more where direct human interaction in hazardous conditions is not possible. For such applications, it is important to manage and control the systems remotely.

Tan et al.[136] provided a data driven automatic fault detection and classification scheme utilizing naive bayes and probabilistic neural network. To prevent hazardous gas leakage, a monitoring system employed with an adjoint probabilistic algorithm was proposed by Zhou et al.[164] where the response time of the gas monitoring sensors was used to estimate the gas leakage point of a system. In the underground coal mining system, it is necessary to monitor harmful, combustible, and noxious gases for the safety of the workers. Instead of using fuzzy, rule-based statistical approaches which are inefficient in complex scenario, Sharma et al. and Pareek et al.[165,166] approached with 1-D CNN models powered by Dempster Shafer evidence theory that seem to perform in accuracy and the number of training parameters. A multilayer perceptual neural network model was used for natural gas monitoring application by an array of infrared sensors earlier.[167] To monitor NO_2 in air, Laref et al.[168] proposed an approach of calibration transfer to reduce the long-term drift of the sensors. Targeting continuous monitoring of ambient air quality, Bax et al.[69] proposed an approach of removing all the uncorrelated features from the feature vector using orthogonal signal correction based method. Unsupervised approaches are more suitable for online monitoring and classification as they do not require labeled dataset. A smartphone based remote management and basic data manipulation from a wireless sensor has been proposed by Alexander et al.[169]

Prior articles[80,146] provided such algorithms which can counteract drifting of sensors employing a semi-supervised approach. A 2-D CNN was implemented for fast and accurate online classification of drifted sensors ensuring 91% accuracy.[127] A genetic algorithm-based feature optimization enabled fast ensemble classifier was proposed by Manna et al.[124] A novel energy efficient additive neural network-based leakage detection system assisted by generative adversarial network was proposed earlier[170,171] which performs better with consistency than regular and convolutional neural networks. Another study to determine and eliminate sensor

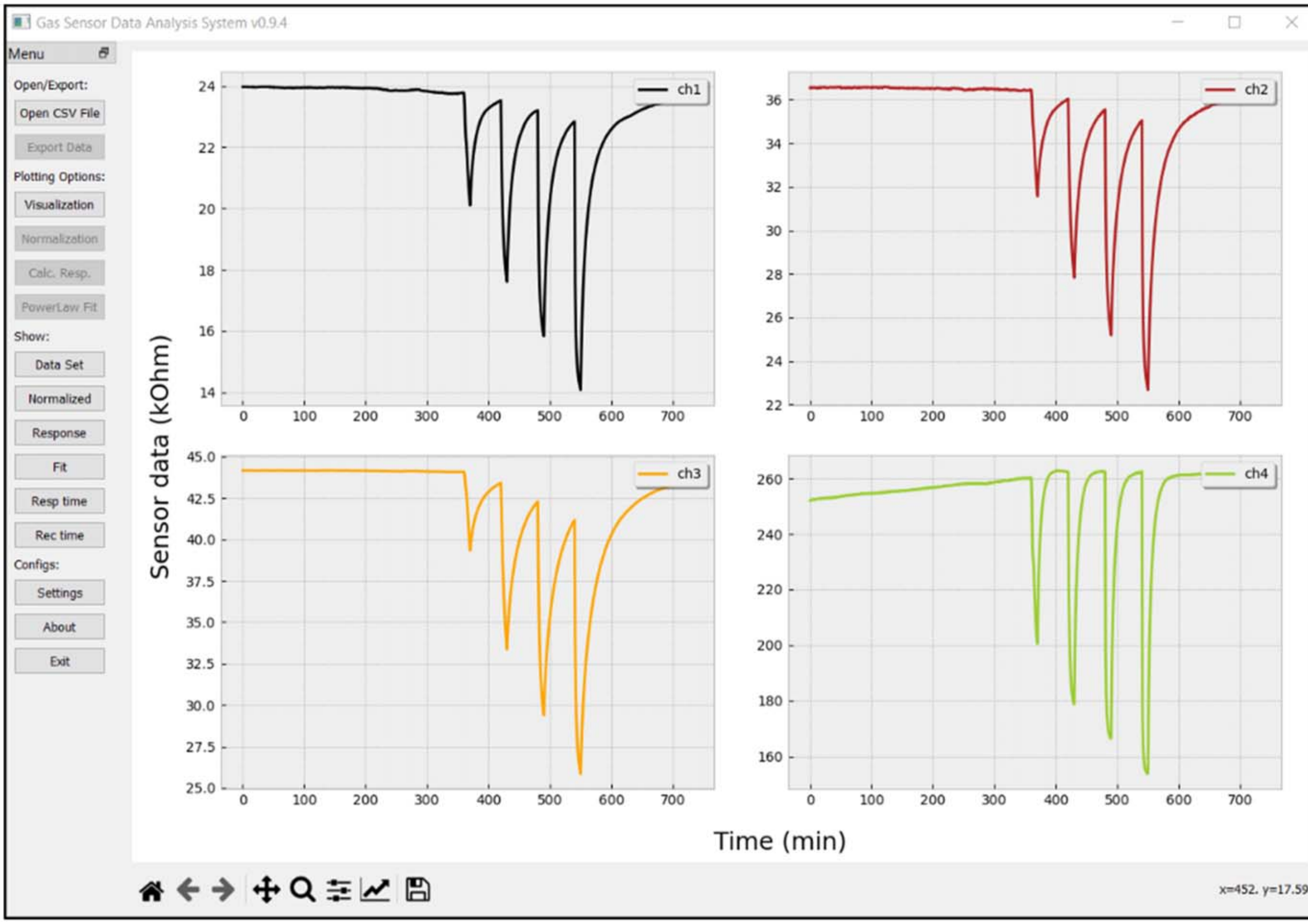

Figure 6. Graphical User Interface showing the main menu and plot area.[157]

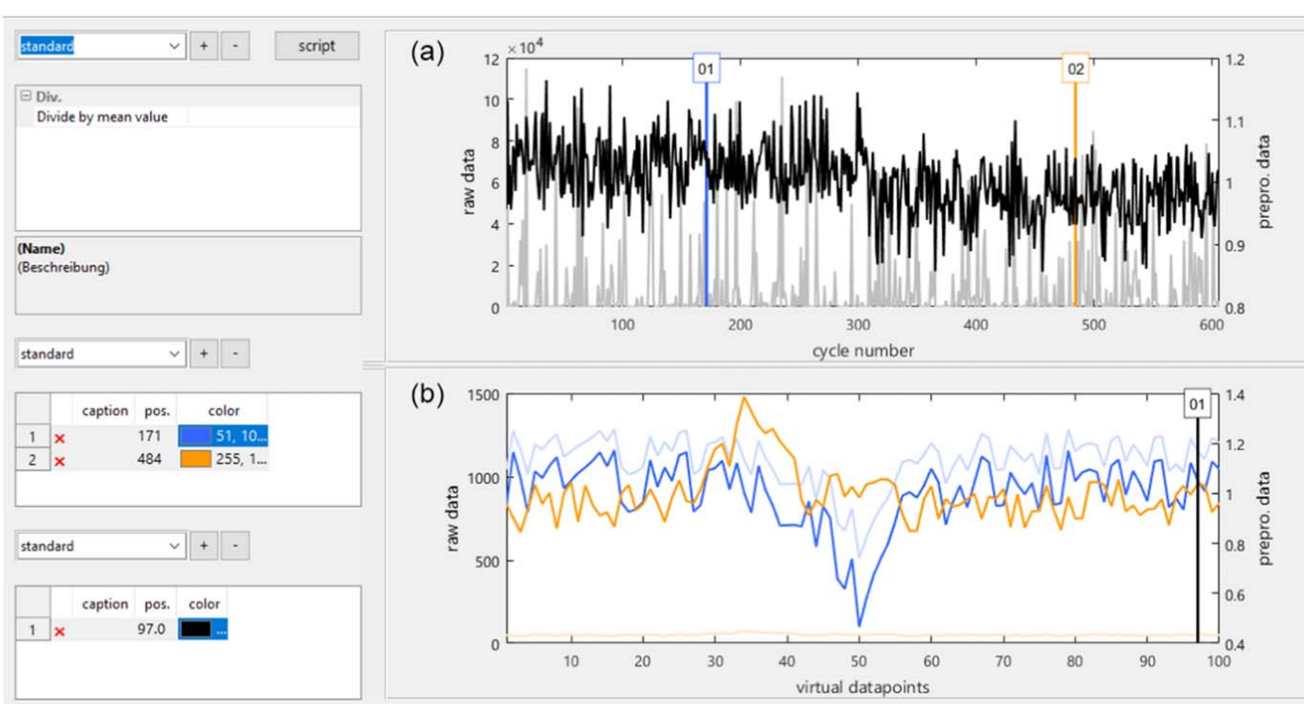

Figure 7. Preprocessing module showing the quasistatic (a) and cyclic (b) view of the preprocessed data, with pale raw data in the background for comparison.[28]

Table II. **Data enabled techniques of different sectors in gas sensing technology.**

Sector of implementation	Applications	Algorithms	References
Feature oriented systems			
Attributes and instance selection	Fast response applications and data imbalance reduction	KNN, QBC, ACO, PCA	37, 66, 58.–65., 67.–70.
Data manipulation	Model compatibility and latent feature extraction of data	Cosine similarity, correlation analysis, PCA, FFT, Wavelet transform, augmented CNN, and RNN	40, 41, 48, 65, 71–89
Signal denoising applications	Signal to noise ratio improvement, outlier reduction, noise cancellation	MANOVA, DWT, CNN-LSTM, PCA, Sparse optimization	41, 75, 91–104
Model based algorithms			
Supervised learning	Classification, regression, drift compensation, and fault detection	SVM, KNN, RF, XGBoost, PLS, BDA, MMD optimization, domain transfer, classifier ensembling, ANN, CNN, GRU	33, 41, 59, 61, 67, 68, 75–78, 82, 83, 88, 105–139, 143, 140–142, 144
Unsupervised learning	Classification, regression, calibration, and drift compensation	PSO, GA, ACO, semi-supervised learning, K-means clustering, Q-network	40, 60, 66, 73, 76, 124, 145–152
Managerial systems			
Interface design	Control and visualization	PyQt, Python, Matlab	18, 29, 30, 40, 41, 48, 118, 121, 157, 153–156, 158
Dataflow systems	Fast classification, self-calibration and proper sampling	SVM, KNN, POCS, ELM, CNN, Reinforcement learning	27, 66, 59, 61, 63, 73, 79, 125, 162, 152, 159–161, 163
Remote management and automation	Mobile system development and fault management	SVM, KNN, 1d CNN, PLSR	69, 80, 103, 124, 127, 136, 176, 167, 165, 166, 146, 164, 168–175

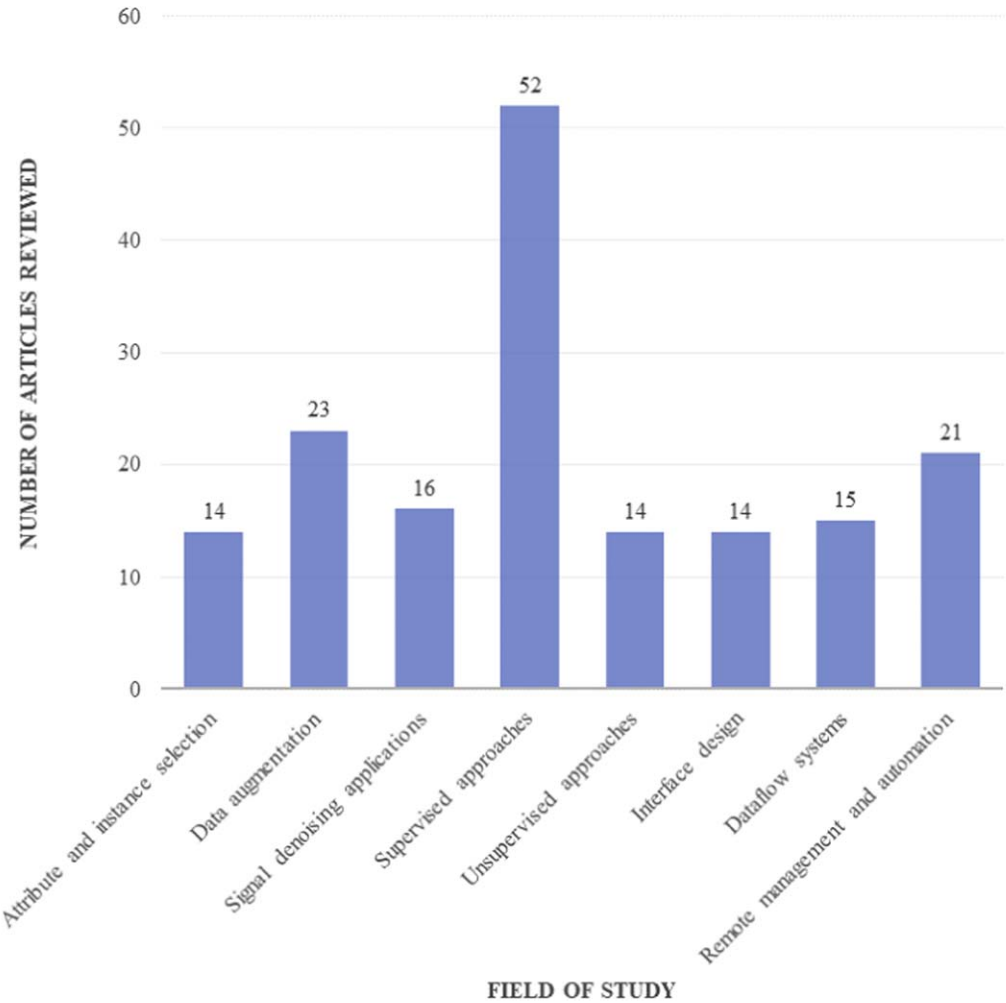

Figure 8. The reviewed articles' distribution according to the field of study.

failure effect from the system was described in a report[172] where KNN, decision tree and linear discriminant analysis was used for food quality assessment. E-nose devices are being used in many quality assessment applications.[173]

A simulation based stochastic model was developed by Schober et al.[174] to properly screen sensor failures. A smartphone-based meat quality inspection system was proposed earlier[175] by tracking CO_2. In another study, a light-weight portable E-nose was developed with Partial Least Square Regression (PLSR) based classifier for waste-water treatment plant. The study implemented a GUI that illustrates different measurements as well as the sensor condition and real time imaging of the sensor location.[176] A fast two-stage method for drift suppressed learning was proposed earlier[103] where one stage extracts the linear information such as temperature change and time series information from the dataset and another stage constantly makes decision based on the linear parameters extracted from the previous stage.

Future Prospects

The evolution of algorithms in recent years have provided numerous opportunities in designing efficient odor sensing systems. E-noses are currently being used to diagnose different chronic diseases such as cancer, diabetes, and halitosis where data analysis helps to distinguish the suitable biomarkers for better pattern recognition.[177] On the other hand, it is important to have better gas recognition ability with as small amount of data as possible to minimize the overall expense of a chemical sensor system.[44] Although recent deep learning based complex machine learning algorithms have provided improved results, they often require extensive amount of data. Moreover, the complexity and depth of the models are inversely proportional to the processing time. Different algorithms like elastic net (Enet), region based CNN (R-CNN), YOLO (you only look once) and their faster variants was developed recently, which can improve the processing time significantly.[178–181] On the other hand, different activation and ensembling approaches such as eLU, ReLU, linear discriminant analysis (LDA), locality preserving projection (LPP) and other neighborhood based methods have increased the convergence performance, even with smaller dataset.

Multi-modality of a gas sensing system often leads to the aging of sensors. Different evolutionary and metaheuristic algorithms can ensure efficient analysis of such systems providing increased life-time of the gas sensors.[182–184]

In many sophisticated applications such as military, nuclear energy research and space aviation, it is critical to have energy efficient and fast response algorithms.[159,185,186] The demand of such advanced algorithms in the field of gas sensing system is becoming a necessity. Table II summarizes the different data enabled techniques to address gas sensor challenges and the corresponding algorithms. Figure 8 illustrates the article distribution in accordance with the research domains of data driven analogies in gas sensing technology.

Conclusions

Malfunctioning of gas sensors can be devastating resulting in accidents, inaccurate decision making, and loss of revenue. The need for a smart gas sensing system cannot be overstated. It is critical to employ modern data analysis for building smart gas sensing systems. Due to specific atomic and chemical structure of the active materials used in any gas sensor, it has been a challenge to detect a range of gases with one single device, making E-noses popular. E-nose sensors are getting complex. It is critical to maintain the sensors as well as analyze the raw data to make meaningful inferences. Moreover, the size of factors that affect the sensor performance have become enormous requiring the use of modern data analysis techniques and management tools. For an automated gas sensor system design, it is critical to select the parameters of the sensor system carefully. From data collection to modeling, every aspect needs to be in place to obtain intended performance out of the system.

Numerous machine learning and statistical analysis based algorithms have been proposed for data enabled sensor systems. However, they cannot be directly implemented in gas sensing applications given the complexity and interdisciplinary nature of gas sensor research. Again, there exist many modeling and feature extraction techniques independently, but there is a lack of holistic sensor system design with data collection, visualization, feature extraction and selection, failure detection and drift compensation with qualitative and quantitative analysis of gas sensors. This article acts as guide to develop an automated gas sensor system with the above-mentioned features. With extensive elaboration and structured way of presenting the cutting-edge research performed during the past years, this mini-review provided an overall view on the utilization of modern data enabled technologies in gas sensor applications.

Acknowledgments

This research is funded by the National Science Foundation, under grant #2104513.

ORCID

Praveen Kumar Sekhar https://orcid.org/0000-0002-4669-535X

References

1. J. Cho and G. Shin, *Polymers*, **14**, 1557 (2022).
2. A. P. Turner, *ECS Sensors Plus*, **1**, 011601 (2022).
3. B.-M. Tuchiu, R.-I. Stefan-van Staden, and J. K. F. van Staden, *ECS Sensors Plus*, **1**, 030601 (2022).
4. R. Lopez, J. Fuentes, A. Gonzalez-Camps, T. Benhaddouch, A. Kaushik, C. L. Metler, S. Bhansali, and D. Dong, *ECS Sensors Plus*, **1**, 035601 (2022).
5. F. D. S. Santos, L. V. da Silva, P. V. S. Campos, C. de Medeiros Strunkis, C. M. G. Ribeiro, and M. O. Salles, *ECS Sensors Plus*, **1**, 013603 (2022).
6. X. Wu, H. Wang, J. Wang, D. Wang, L. Shi, X. Tian, and J. Sun, *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, **632**, 127752 (2022).
7. P. Saxena and P. Shukla, *Computational and Experimental Methods in Mechanical Engineering* (Springer, Berlin) p. 165 (2022).
8. N. Wang, N. Zhang, T. Wang, F. Liu, X. Wang, X. Yan, C. Wang, X. Liu, P. Sun, and G. Lu, *Sensors and Actuators B: Chemical*, **350**, 130854 (2022).
9. S. Agrohiya, S. Dahiya, P. K. Goyal, I. Rawal, A. Ohlan, R. Punia, and A. Maan, *ECS Sensors Plus*, **1**, 043601 (2022).
10. A. Yadav and P. D. Indurkar, *Water Conservation Science and Engineering*, **6**, 175 (2021).
11. R. V. Poonguzhali, E. R. Kumar, T. Pushpagiri, A. Steephen, N. Arunadevi, and S. Baskoutas, *Solid State Commun.*, **325**, 114161 (2021).
12. L. Ge, X. Mu, G. Tian, Q. Huang, J. Ahmed, and Z. Hu, *Frontiers in Chemistry*, **839**, 54 (2019).
13. T. T. T. Toan et al., *ECS Sensors Plus*, **1**, 031603 (2022).
14. C. Gonzalez Viejo and S. Fuentes, *Chemosensors*, **10**, 159 (2022).
15. S. Grassi, S. Benedetti, L. Magnani, A. Pianezzola, and S. Buratti, *Food Control*, **138**, 108994 (2022).
16. S. Fuentes, E. Tongson, R. R. Unnithan, and C. Gonzalez Viejo, *Sensors*, **21**, 5948 (2021).
17. V. Binson and M. Subramoniam, *Acta of Bioengineering & Biomechanics*, **23**, 35 (2021).
18. A. Anyfantis and S. Blionas, *presented at the Panhellenic Conference on Electronics & Telecommunications (PACET)* p. 1 (2019).
19. N. Mohammad Yusof, S. Ibrahim, and S. Rozali, *J. Mater. Res.*, **37**, 1 (2022).
20. S. Lv et al., *Sensors and Actuators B: Chemical*, **354**, 131219 (2022).
21. Z. Gong, H. Li, G. Wu, J. Ma, K. Chen, F. Ma, L. Mei, W. Peng, and Q. Yu, *Microwave and Optical Technology Letters* (2022).
22. J.-y. Lee, K. Ko, and H. Chung, *J. Environ. Management*, **312**, 114893 (2022).
23. D. Hodgins, *Techniques for Analyzing* (Boca Raton, FL)(CRC Press) p. 331 (2020).
24. M. Modesti, I. Taglieri, A. Bianchi, A. Tonacci, F. Sansone, A. Bellincontro, F. Venturi, and C. Sanmartin, *Appl. Sci.*, **11**, 8453 (2021).
25. M. Yakob, D. Mustika, R. N. Ida, and A. P. Rachmad, *J. Phys.: Conf. Ser.*, **1428**, 012062.
26. W. Hu, L. Wan, Y. Jian, C. Ren, K. Jin, X. Su, X. Bai, H. Haick, M. Yao, and W. Wu, *Adv. Mater. Technol.*, **4**, 1800488 (2019).
27. S. A. Schober, C. Carbonelli, A. Roth, A. Zoepfl, and R. Wille, *presented at the 2020 IEEE SENSORS* p. 1 (2020).
28. M. Bastuck, T. Baur, and A. Schütze, *J. Sensors and Sensor Systems*, **7**, 489 (2018).
29. Y.-J. Liu, Q.-H. Meng, and X.-N. Zhang, *IEEE Sens. J.*, **18**, 9360 (2018).
30. H. Djelouat, A. A. S. Ali, A. Amira, and F. Bensaali, *J. Natural Gas Sci. Eng.*, **55**, 612 (2018).
31. Y. A. Alsultanny, *presented at the XIX International Conference onData Analytics andManagementin Data Intensive Domains* p. 350.
32. Z. Kovacs, D. Szöllősi, J.-L. Z. Zaukuu, Z. Bodor, F. Vitális, B. Aouadi, V. Zsom-Muha, and Z. Gillay, *Biosensors*, **10**, 74 (2020).
33. B. Liu, X. Zeng, F. Tian, S. Zhang, and L. Zhao, *IEEE Access*, **7**, 143947 (2019).
34. R. Saeed, H. Feng, X. Wang, Z. Xiaoshuan, and F. Zetian, *Food Control*, **137**, 108902 (2022).
35. Z. Liang, F. Tian, S. X. Yang, C. Zhang, H. Sun, and T. Liu, *Sensors*, **18**, 1179 (2018).
36. S. Al Maskari and X. Li, *Electronic Nose Technologies and Advances in Machine Olfaction* (IGI Global) p. 38 (2018).
37. I. Czarnowski, *J. Comput. Sci.*, **61**, 101614 (2022).
38. A. Gautam, *ECS Sensors Plus*, **1**, 042401 (2022).
39. V. Chaudhary, A. K. Kaushik, H. Furukawa, and A. Khosla, *ECS Sensors Plus*, **1**, 013601 (2022).
40. J. Yan, F. Chen, T. Liu, Y. Zhang, X. Peng, D. Yi, and S. Duan, *Knowledge-Based Systems*, **235**, 107664 (2022).
41. Z. Liang, Q. Xue, F. Tian, C. Xu, C. Wang, L. Yang, and T. Guo, *IEEE Sens. J.*, **22**, 6717 (2022).
42. T. Guo, K. Yu, X. Cheng, and A. K. Bashir, *presented at the IEEE International Conference on Communications Workshops (ICC Workshops)* p. 1 (2021).
43. Z. Ye, Y. Liu, and Q. Li, *Sensors*, **21**, 7620 (2021).
44. U. Yaqoob and M. I. Younis, *Sensors*, **21**, 2877 (2021).
45. M. V. Nikolic, V. Milovanovic, Z. Z. Vasiljevic, and Z. Stamenkovic, *Sensors*, **20**, 6694 (2020).
46. S. Feng, F. Farha, Q. Li, Y. Wan, Y. Xu, T. Zhang, and H. Ning, *Sensors*, **19**, 3760 (2019).
47. L. Zhang, F. Tian, and D. Zhang, *Electronic Nose: Algorithmic Challenges* (Springer, Berlin) (2018).
48. A. Vergara, S. Vembu, T. Ayhan, M. A. Ryan, M. L. Homer, and R. Huerta, *Sensors and Actuators B: Chemical*, **166**, 320 (2012).
49. Z. Wang, Z.-h. Deng, R.-j. Zhu, Y.-h. Zhou, and X. Li, *Sensors and Actuators B: Chemical*, **359**, 131622 (2022).
50. T. Iwata, M. Saeki, Y. Okura, and T. Yoshikawa, *Sensors and Actuators B: Chemical*, **354**, 131225 (2022).
51. S. Wakhid, R. Sarno, and S. I. Sabilla, *Computers and Electronics in Agriculture*, **195**, 106838 (2022).
52. M. A. Franco, P. P. Conti, R. S. Andre, and D. S. Correa, *Sensors and Actuators Reports*, **4**, 100100 (2022).
53. A. Siqueira, M. Melo, D. Giordani, D. Galhardi, B. Santos, P. Batista, and A. Ferreira, *J. Food Eng.*, **221**, 114 (2018).
54. D. R. Wijaya, R. Sarno, and E. Zulaika, *Data in Brief*, **21**, 2414 (2018).
55. J. C. R. Gamboa et al., *Data in Brief*, **25**, 104202 (2019).
56. C. M. D. Acevedo, C. A. C. Vasquez, and J. K. C. Gómez, *Data in Brief*, **35**, 106767 (2021).
57. UCI Machine Learning Repository: Gas Sensor Array Drift Dataset at Different Concentrations Data Set, (Accessed on 05/29/2022).
58. S. Yu, X. Luo, Z. He, J. Yan, K. Lv, and D. Shi, *presented at the Ninth International Conference on Intelligent Control and Information Processing (ICICIP)* p. 224 (2018).
59. T. Liu, D. Li, J. Chen, M. Wu, and Y. Chen, *presented at the V International Conference on Systems and Informatics (ICSAI)* p. 417 (2018).
60. T. Liu, D. Li, J. Chen, Y. Chen, T. Yang, and J. Cao, *Sensors*, **18**, 4028 (2018).
61. T. Matthews, M. Iqbal, and H. Gonzalez-Velez, *presented at the IEEE/ACM V International Conference on Big Data Computing Applications and Technologies (BDCAT)* p. 61 (2018).
62. T. Liu, D. Li, Y. Chen, M. Wu, T. Yang, and J. Cao, *Sensors and Actuators B: Chemical*, **316**, 128065 (2020).
63. T. Liu, D. Li, and J. Chen, *Measurement*, **171**, 108748 (2021).
64. T. Liu, J. Cao, D. Li, Y. Chen, T. Yang, and X. Zhu, *Sensors and Actuators A: Physical*, **312**, 112149 (2020).
65. C. Deng, K. Lv, D. Shi, B. Yang, S. Yu, Z. He, and J. Yan, *Sensors*, **18**, 1909 (2018).
66. G. Magna, F. Mosciano, E. Martinelli, and C. Di Natale, *Sensors and Actuators B: Chemical*, **258**, 1242 (2018).
67. D. R. Wijaya, F. Afianti, A. Arifianto, D. Rahmawati, and V. S. Kodogiannis, *Sensing and Bio-Sensing Research*, **36**, 100495 (2022).
68. H. Shu and R. Y.-N. Wong, *presented at the IEEE XVI Conference on Industrial Electronics and Applications (ICIEA)* p. 1131 (2021).
69. C. Bax et al., *NOSE*, **85**, 13 (2021).
70. Y.-C. Cheng, T.-I. Chou, J.-L. Lee, S.-W. Chiu, and K. T. Tang, *presented at the ECS Meeting Abstracts* (1856).

71. A. Rudnitskaya, *Frontiers in Chemistry*, **433**, 1 (2018).
72. J. Burgués, J. M. Jiménez-Soto, and S. Marco, *Analytica Chimica Acta*, **1013**, 13 (2018).
73. A. U. Rehman and A. Bermak, *IEEE Sens. J.*, **18**, 7173 (2018).
74. A. ur Rehman, A. Bermak, and M. Hamdi, *IEEE Sens. J.*, **19**, 12126 (2019).
75. H. Liu, Q. Li, Z. Li, and Y. Gu, *Sensors*, **20**, 1913 (2020).
76. Z. Jiang, P. Xu, Y. Du, F. Yuan, and K. Song, *Sensors*, **21**, 3403 (2021).
77. J. X. Leon-Medina, W. A. Pineda-Muñoz, and D. A. T. Burgos, *IEEE Access*, **8**, 134413 (2020).
78. K. Chen, L. Liu, B. Nie, B. Lu, L. Fu, Z. He, W. Li, X. Pi, and H. Liu, *Computers in Biology and Medicine*, **131**, 104294 (2021).
79. S. Zhang, F. Tian, J. A. Covington, H. Li, L. Zhao, R. Liu, J. Qian, and B. Liu, *IEEE Trans. Instrum. Meas.*, **70**, 1 (2021).
80. N. J. Pineau, J. F. Kompalla, A. T. Güntner, and S. E. Pratsinis, *Microchimica Acta*, **185**, 1 (2018).
81. M. Ijaz, A. ur Rehman, M. Hamdi, and A. Bermak, *presented at the IEEE International Symposium on Circuits and Systems (ISCAS)* p. 1 (2020).
82. Z. Yi, *Expert Systems with Applications*, **148**, 113238 (2020).
83. L. Wozniak, P. Kalinowski, G. Jasinski, and P. Jasinski, *Microelectron. Reliab.*, **84**, 163 (2018).
84. R. Deji, B. Choudhary, and R. K. Sharma, *Mater. Lett.*, **306**, 130986 (2022).
85. Y. Wang, Y. Yin, F. Ge, and H. Yu, *Sensors and Actuators B: Chemical*, **292**, 217 (2019).
86. Y. Yin, Y. Bai, F. Ge, H. Yu, and Y. Liu, *Measurement*, **139**, 284 (2019).
87. Y. Tao, C. Li, Z. Liang, H. Yang, and J. Xu, *Sensors*, **19**, 3703 (2019).
88. L. Feng, H. Dai, X. Song, J. Liu, and X. Mei, *Sensors and Actuators B: Chemical*, **351**, 130986 (2022).
89. Q. Wang, H. Qi, and F. Liu, *presented at the Chinese Control Conference (CCC)* p. 3479 (2019).
90. M. Zych, R. Hanus, B. Wilk, L. Petryka, and D. Świsulski, *Measurement*, **129**, 288 (2018).
91. M. Tohidi, M. Ghasemi-Varnamkhasti, V. Ghafarinia, S. S. Mohtasebi, and M. Bonyadian, *Measurement*, **124**, 120 (2018).
92. D. Shen and A. Yu, *Journal of Nanoelectronics and Optoelectronics*, **13**, 1533 (2018).
93. D. R. Wijaya, R. Sarno, and E. Zulaika, *Computers and Electronics in Agriculture*, **157**, 305 (2019).
94. D. R. Wijaya, R. Sarno, and E. Zulaika, *Sensors and Actuators B: Chemical*, **326**, 128931 (2021).
95. B. Wang, Y. Guo, D. Wang, Y. Zhang, R. He, and J. Chen, *Mechanical Systems and Signal Processing*, **181**, 109557 (2022).
96. Y. Rui-jun, D. Dan-feng, and C. Yan, *presented at the International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*74 (2019).
97. I. Essiet, Y. Sun, and Z. Wang, *Procedia Manufacturing*, **35**, 629 (2019).
98. A. Miquel-Ibarz, J. Burgués, and S. Marco, *Sensors and Actuators B: Chemical*, **350**, 130769 (2022).
99. Y. Xu, X. Zhao, Y. Chen, and Z. Yang, *Applied Sciences*, **9**, 1728 (2019).
100. N. Morati, T. Contaret, S. Gomri, T. Fiorido, J.-L. Seguin, and M. Bendahan, *Sensors and Actuators B: Chemical*, **334**, 129654 (2021).
101. B. Liu, Y. Zhou, H. Fu, P. Fu, and L. Feng, *Sensors*, **22**, 4315 (2022).
102. J. Chen, J. Ning, W. Chen, X. Wang, W. Wang, and G. Zhang, *Interpretation*, **7**, T373 (2019).
103. C. Zhang, W. Wang, Y. Pan, L. Cheng, S. Zhai, and X. Gao, *Meas. Sci. Technol.*, **33**, 045108 (2022).
104. J. Zhao, Y. Li, and S. Sampath, *J. Eng. for Gas Turbines and Power*, 1 (2022).
105. S. A. Schober, Y. Bahri, C. Carbonelli, and R. Wille, *Chemosensors*, **10**, 152 (2022).
106. M. M. Habib, A. Rodan, and A. Alazzam, *presented at the Sixth HCT Information Technology Trends (ITT)* p. 172 (2019).
107. T. Liu, H. Jiang, and Q. Chen, *Microchemical Journal*, **159**, 105339 (2020).
108. X. Dong, X. Qi, J. Cui, X. Xu, and A. Wan, *presented at the IEEE X International Conference on Electronics Information and Emergency Communication (ICEIEC)* p. 236 (2020).
109. A. Paleczek and A. M. Rydosz, *J. Breath Res.*, **16**, 026003 (2022).
110. S. Upasham, P. Rice, S. Shahub, V. N. Dhamu, and S. Prasad, *ECS Sensors Plus*, **1**, 031602 (2022).
111. T. Liu, D. Li, J. Chen, Y. Chen, T. Yang, and J. Cao, *Sensors*, **19**, 3601 (2019).
112. X. Dong, S. Han, A. Wang, and K. Shang, *Chemosensors*, **9**, 353 (2021).
113. V. Binson and M. Subramoniam, *presented at the Journal of Physics: Conference Series, vol 1950* 012065 (1950).
114. P. Srinivasan, J. Robinson, J. Geevaretnam, and J. B. B. Rayappan, *Sensors and Actuators B: Chemical*, **317**, 128192 (2020).
115. R. Yi, J. Yan, D. Shi, Y. Tian, F. Chen, Z. Wang, and S. Duan, *Sensors and Actuators B: Chemical*, **329**, 129162 (2021).
116. X. Zhu, T. Liu, J. Chen, J. Cao, and H. Wang, *Chemosensors*, **9**, 208 (2021).
117. Y. Tian, J. Yan, D. Yi, Y. Zhang, Z. Wang, T. Yu, X. Peng, and S. Duan, *IEEE Trans. Instrum. Meas.*, **70**, 1 (2021).
118. Z. Liang, L. Yang, T. Guo, and J. Li, *presented at the XIII International Conference on Communication Software and Networks (ICCSN)* 242 (2021).
119. Z. Yi and C. Li, *IEEE Access*, **7**, 170087 (2019).
120. Z. Yi, W. Shang, T. Xu, and X. Wu, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **52**, 3530 (2021).
121. Z. Yi, W. Shang, T. Xu, S. Guo, and X. Wu, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **52**, 247 (2020).
122. Y. Tao, K. Zeng, and Z. Liang, *presented at the IEEE/CIC International Conference on Communications in China (ICCC)* p. 130 (2020).
123. P. Das, A. Manna, and S. Ghoshal, *presented at the International Conference on Renewable Energy Integration into Smart Grids: A Multidisciplinary Approach to Technology Modelling and Simulation (ICREISG)* p. 197 (2020).
124. A. Manna, *presented at the International Conference on Computer Communication and Informatics (ICCCI)* p. 1 (2020).
125. A. U. Rehman, S. B. Belhaouari, M. Ijaz, A. Bermak, and M. Hamdi, *IEEE Sens. J.*, **21**, 6564 (2020).
126. Q. Li, P. Wu, Z. Liang, and Y. Tao, *presented at the XIII International Conference on Communication Software and Networks (ICCSN)* p. 277 (2021).
127. V. Pareek, S. Chaudhury, and S. Singh, *presented at the XI IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Vol 2,* p. 679 (2021).
128. A. ur Rehman and A. Bermak, *presented at the XV International Wireless Communications & Mobile Computing Conference (IWCMC)* p. 848 (2019).
129. S. N. Chaudhri and N. S. Rajput, *IEEE Sens. Lett.*, **6**, 1 (2022).
130. Y. Tian, J. Yan, Y. Zhang, T. Yu, P. Wang, D. Shi, and S. Duan, *IEEE Access*, **8**, 121385 (2020).
131. Y. Robin, J. Amann, T. Baur, P. Goodarzi, C. Schultealbert, T. Schneider, and A. Schütze, *Atmosphere*, **12**, 1487 (2021).
132. Z. Wang, C. Xie, B. Liu, Y. Jiang, Z. Li, H. Tai, and X. Li, *Sensors and Actuators B: Chemical*, **362**, 131812 (2022).
133. X. Dong, H. Duan, X. Xu, and S. Han, *presented at the IEEE XI International Conference on Electronics Information and Emergency Communication (ICEIEC) 2021 IEEE XI International Conference on Electronics Information and Emergency Communication (ICEIEC)* p. 1 (2021).
134. S.-H. Wang, T.-I. Chou, S.-W. Chiu, and K.-T. Tang, *IEEE Sens. J.*, **21**, 6401 (2020).
135. M. Abbatangelo, E. Núñez-Carmona, V. Sberveglieri, E. Comini, and G. Sberveglieri, *Chemosensors*, **8**, 6 (2020).
136. Q. Tan, X. Mu, M. Fu, H. Yuan, J. Sun, G. Liang, and L. Sun, *Measurement*, **194**, 111037 (2022).
137. T. Chaudhuri, M. Wu, Y. Zhang, P. Liu, and X. Li, *IEEE Sens. J.*, **21**, 7908 (2020).
138. Y.-C. Cheng, T.-I. Chou, S.-W. Chiu, and K.-T. Tang, *IEEE Sens. Lett.*, **4**, 1 (2020).
139. X. Zhao, P. Li, K. Xiao, X. Meng, L. Han, and C. Yu, *Sensors*, **19**, 3844 (2019).
140. X. Li, Y. Yang, and Y. Zhu, "A. Ben, J. Qi." *Food Chemistry*, **384**, 132519 (2022).
141. M. Valcárcel, G. Ibáñez, R. Martı, J. Beltrán, J. Cebolla-Cornejo, and S. Roselló, *Analytical and Bioanalytical Chemistry*, **413**, p. 3893 (2021).
142. C. Jaeschke, J. Glöckler, M. Padilla, J. Mitrovics, and B. Mizaikoff, *Analytical Methods*, **12**, 4724 (2020).
143. Y. Zhang, S. Xiang, Z. Wang, X. Peng, Y. Tian, S. Duan, and J. Yan, *Sensors and Actuators B: Chemical*, **361**, 131739 (2022).
144. Z. Wang, J. Yan, F. Chen, X. Peng, Y. Zhang, Z. Wang, and S. Duan, *IEEE Sens. J.*, **21**, 17144 (2021).
145. H. Rahimzadeh, M. Sadeghi, S. A. Mireei, and M. Ghasemi-Varnamkhasti, *Biosystems Engineering*, **216**, 132 (2022).
146. Z. Liang, L. Zhang, F. Tian, C. Wang, L. Yang, T. Guo, and L. Xiong, *Sensors and Actuators B: Chemical*, **349**, 130727 (2021).
147. R. Laref, E. Losson, A. Sava, and M. Siadat, *Sensors*, **21**, 3581 (2021).
148. A. ur Rehman and A. Bermak, *IEEE Sens. J.*, **19**, 1443 (2018).
149. P. Maho, C. Herrier, T. Livache, P. Comon, and S. Barthelmé, *Chemometrics and Intelligent Laboratory Systems*, **225**, 104549 (2022).
150. M. Ouqamra and D. Bouilly, *presented at the IEEE XXIX International Workshop on Machine Learning for Signal Processing (MLSP)* p. 1 (2019).
151. V. S. Palaparthy, S. N. Doddapujar, S. G. Surya, S. A. Chandorkar, S. Mukherji, M. S. Baghini, and V. R. Rao, *IEEE Sens. J.*, **21**, 8011 (2020).
152. Y. You, K. You, H. Chen, and T. J. Oechtering, *IEEE Internet of Things Journal*, **9**, 13848 (2022).
153. Q. Liu, K. Sun, N. Zhao, J. Yang, Y. Zhang, C. Ma, L. Pan, and K. Tu, *Postharvest Biology and Technology*, **153**, 152 (2019).
154. P. Wei, Z. Ning, S. Ye, L. Sun, F. Yang, K. C. Wong, D. Westerdahl, and P. K. Louie, *Sensors*, **18**, 59 (2018).
155. S. Srivastava and S. Sadistap, *Journal of Food Measurement and Characterization*, **12**, 2758 (2018).
156. N. Debabhuti, A. Sengupta, P. Sharma, R. Sen, B. Tudu, and R. Bandyopadhyay, *Proceedings of Industry Interactive Innovations in Science, Engineering & Technology (I3SET2K19)* (2020).
157. B. de Lima, W. A. d. S. Silva, A. L. Ndiaye, V. R. Mastelaro, and J. Brunet, *Chemometrics and Intelligent Laboratory Systems*, **220**, 104460 (2022).
158. M. Dmitrzak, P. Kalinowski, P. Jasinski, and G. Jasinski, *Sensor Review*, **42**, 195 (2021).
159. Y. You and T. J. Oechtering, *presented at the XXVIII European Signal Processing Conference (EUSIPCO)* p. 1717 (2020).
160. G. Magna, C. Di Natale, and E. Martinelli, *Sensors and Actuators B: Chemical*, **297**, 126721 (2019).
161. M. M. Bordbar, J. Tashkhourian, A. Tavassoli, E. Bahramali, and B. Hemmateenejad, *Sensors and Actuators B: Chemical*, **319**, 128262 (2020).
162. M. Kang, I. Cho, J. Park, J. Jeong, K. Lee, B. Lee, D. Del Orbe Henriquez, K. Yoon, and I. Park, *ACS Sens.* (2022).
163. S. N. Chaudhri, N. S. Rajput, and A. Mishra, *Journal of Electrical Engineering*, **73**, 108 (2022).
164. K. Zhou, F. Li, H. Cai, Y. Jing, J. Zhuang, M. Li, and Z. Xing, *Energy and Buildings*, **254**, 111645 (2022).
165. M. Sharma and T. Maity, *IEEE Internet of Things Journal*, (2022).
166. V. Pareek and S. Chaudhury, *Soft Computing*, **25**, 14155 (2021).

167. J. Wang, S. Lian, B. Lei, B. Li, and S. Lei, *Sensors and Actuators A: Physical*, **335**, 113392 (2022).
168. R. Laref, E. Losson, A. Sava, and M. Siadat, *presented at the International Conference on Control, Automation and Diagnosis (ICCAD)* p. 1 (2021).
169. A. Scott, R. Pandey, S. Saxena, E. Osman, Y. Li, and L. Soleymani, *ECS Sensors Plus*, **1**, 014601 (2022).
170. D. Badawi, S. Ozev, J. B. Christen, C. Yang, A. Orailoglu, and A. E. Çetin, *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* p. 8296.
171. D. Badawi, T. Ayhan, S. Ozev, C. Yang, A. Orailoglu, and A. E. Cetin, *IEEE Access*, **7**, 155701 (2019).
172. A. Kaya, A. S. Keçeli, C. Catal, and B. Tekinerdogan, *Sensors*, **20**, 3173 (2020).
173. A. Ruengdech and U. Siripatrawan, *Food Bioscience*, **36**, 100593 (2020).
174. S. A. Schober, C. Carbonelli, and R. Wille, *Engineering Proceedings*, **6**, 88 (2021).
175. I. M. P. de Vargas-Sansalvador, M. M. Erenas, A. Martińez-Olmos, F. Mirza-Montoro, D. Diamond, and L. F. Capitan-Vallvey, *Talanta*, **216**, 120985 (2020).
176. J. Burgués, M. D. Esclapez, S. Doñate, and S. Marco, *Iscience*, **24**, 103371 (2021).
177. H. Tai, S. Wang, Z. Duan, and Y. Jiang, *Sensors and Actuators B: Chemical*, **318**, 128104 (2020).
178. A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, *arXiv:1606.02147*, *ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation,* 2016.
179. B. Cheng, Y. Wei, H. Shi, R. Feris, J. Xiong, and T. Huang, *presented at the Proceedings of the European conference on computer vision (ECCV)* p. 453.
180. Y. Wang and R. Chen, *presented at the IEEE X Data Driven Control and Learning Systems Conference (DDCLS)* p. 236 (2021).
181. J. Redmon and A. Farhadi, *arXiv:1804.02767*, *YOLOv3: An Incremental Improvement*, 2018.
182. D. Loukatos and K. G. Arvanitis, *IoT-based Intelligent Modelling for Environmental and Ecological Engineering* (Springer, Berlin) p. 101 (2021).
183. P. Narkhede, R. Walambe, S. Mandaokar, P. Chandel, K. Kotecha, and G. Ghinea, *Applied System Innovation*, **4**, 3 (2021).
184. A. Polo-Rodriguez, F. Cruciani, C. Nugent, and J. Medina-Quero, *presented at the The International Research & Innovation Forum* p. 89.
185. M. Kaliszewski et al., *Sensors*, **21**, 7608 (2021).
186. R. Ismail and S. Muthukumaraswamy, *Intelligent Manufacturing and Energy Sustainability* (Springer, Berlin) p. 637 (2021).