

Review

Machine Learning Enhances the Performance of Bioreceptor-Free Biosensors

Kenneth E. Schackart III ¹  and Jeong-Yeol Yoon ^{1,2,*} 

¹ Department of Biosystems Engineering, The University of Arizona, Tucson, AZ 85721, USA; schackartk@email.arizona.edu

² Department of Biomedical Engineering, The University of Arizona, Tucson, AZ 85721, USA

* Correspondence: jyyoon@arizona.edu

Abstract: Since their inception, biosensors have frequently employed simple regression models to calculate analyte composition based on the biosensor's signal magnitude. Traditionally, bioreceptors provide excellent sensitivity and specificity to the biosensor. Increasingly, however, bioreceptor-free biosensors have been developed for a wide range of applications. Without a bioreceptor, maintaining strong specificity and a low limit of detection have become the major challenge. Machine learning (ML) has been introduced to improve the performance of these biosensors, effectively replacing the bioreceptor with modeling to gain specificity. Here, we present how ML has been used to enhance the performance of these bioreceptor-free biosensors. Particularly, we discuss how ML has been used for imaging, Enose and Etongue, and surface-enhanced Raman spectroscopy (SERS) biosensors. Notably, principal component analysis (PCA) combined with support vector machine (SVM) and various artificial neural network (ANN) algorithms have shown outstanding performance in a variety of tasks. We anticipate that ML will continue to improve the performance of bioreceptor-free biosensors, especially with the prospects of sharing trained models and cloud computing for mobile computation. To facilitate this, the biosensing community would benefit from increased contributions to open-access data repositories for biosensor data.

Keywords: label-free biosensor; machine learning; support vector machine; artificial neural network; principal component analysis



Citation: Schackart, K.E., III; Yoon, J.-Y. Machine Learning Enhances the Performance of Bioreceptor-Free Biosensors. *Sensors* **2021**, *21*, 5519. <https://doi.org/10.3390/s21165519>

Academic Editors: Zihuai Lin and Wei Xiang

Received: 15 July 2021

Accepted: 13 August 2021

Published: 17 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The field of biosensing has exploded into nearly all areas of research, from medical applications [1] to environmental monitoring [2]. Some of the greatest appeals of biosensors are their specificity and sensitivity. These properties are primarily due to bioreceptors, which are selected for their inherent specificities such as enzymes [3], antibodies [4], and aptamers [5]. However, the very aspect that makes biosensors so specific and sensitive can also limit the sensor stability due to the degradation of the bioreceptor [6]. Additionally, as the bioreceptor is specific to an individual analyte, the particular sensor's scope is limited to the specific analyte to which the bioreceptor can bind.

To obviate these issues, many nature-inspired sensors have emerged that are bioreceptor-free. Some of the most notable examples that have made great progress include the electronic nose (Enose) [7–11] and electronic tongue (Etongue) [12–16]. Additionally, surface enhanced Raman spectroscopy (SERS)-based sensors have demonstrated incredible chemosensing ability [17–21]. Without a bioreceptor, however, there is the risk of significantly compromised biosensor performance including the limit of detection (LOD) and specificity. Researchers have introduced machine learning (ML) to bioreceptor-free biosensors to bridge this trade-off gap, improving the LOD and specificity [22]. In a sense, ML can be used to take the place of a bioreceptor by reintroducing specificity during data analysis. This is made possible by powerful ML techniques capable of detecting subtle patterns in sensor responses.

While this approach has demonstrated success, there are still several challenges that these systems must overcome. A major challenge being faced is model generalizability. Since many models rely on subtle patterns in the data, they can be quite sensitive to underlying data changes. This can make the models susceptible to error when faced with sensor drift or replacing parts of the system [14].

Since the scope of this review is quite large and covers all bioreceptor-free biosensors that utilize ML, there are a few points to clarify. Many subsets of our scope have received thorough attention and review. For instance, the use of ML for Enose and Etongue [23–27] and SERS-based biosensors [28] have previously been described. Since the literature is rich in these areas, we realize that all recent original research cannot be adequately covered here. Rather, our intent is to provide a unified discussion of the relevant methods and challenges to give a bigger picture. We also would like to acknowledge that there is a complementary review in the literature discussing the use of ML in biosensing in general [29], but not for biosensors that are bioreceptor-free.

In this review, we will give the current state of using ML to enhance the performance of bioreceptor-free biosensors. Section 2 briefly introduces the types of biosensors that have most benefited from ML. Section 3 provides some background on machine learning algorithms and how their performance can be assessed. Section 4 covers electrochemical biosensors, with particular emphasis on Enose and Etongue. Successful methods are discussed as well as some of the challenges and how they are being addressed with ML. Section 5 discusses optical biosensors, notable for imaging- and SERS-based biosensors. Additional considerations and future perspectives are discussed in Section 6 including what currently prevents many of these systems from being commercialized and what directions may be taken. We also present some considerations on best practices for ML in biosensing, especially regarding communication of methods and reproducibility.

2. How Biosensors Can Benefit from Machine Learning

Biosensors in the classic definition are sensors that utilize a bioreceptor such as antibody, enzyme, peptide, nucleic acid, etc. A bioreceptor binds to a target biological molecule and generates a signal when coupled with a transducer. Biosensors have evolved to a wide range of transducer types including electrochemical, optical, and spectroscopic biosensors. Traditionally, it is the bioreceptor that provides specificity and sensitivity to the biosensor. Increasingly, however, researchers are developing biosensors that lack a specific bioreceptor. A typical example is a semi-specific chemical sensor array, termed Enose (from gas), or Etongue (in solution). Since such a sensor's specificity is not provided by the bioreceptor, a fingerprinting technique is used to recognize signal patterns indicative of a particular analyte. Frequently, machine learning techniques are employed to detect these patterns and provide specificity.

The use of machine learning to enhance the performance (e.g., specificity, sensitivity, and LOD) of bioreceptor-free biosensors is not limited to chemical sensor arrays. It has been employed in various biosensor mechanisms. Some of the most famous examples aside from Enose and Etongue are imaging-based biosensors and SERS-based biosensors. Additionally, the use of machine learning for biosensors is not limited to those that lack bioreceptors. Cui et al. [29] cover several examples of traditional biosensors employing machine learning to enhance performance.

Table 1 provides an overview of the tasks for which machine learning has been applied, the specific algorithms used, and the relevant papers. More information on the algorithms themselves can be found in Section 3. Additionally, Table 2 gives a comparison of each of the major biosensing mechanisms including data type and appropriate feature engineering and ML methods. All information in Table 2 comes from Table 1 and serves as a higher-level summary.

Table 1. Machine learning tasks and algorithms used in biosensing.

Biosensing Mechanism	Task	Target	Algorithm	Ref.
ELECTROCHEMICAL				
CV	Regression	Maleic hydrazide	ANN	[30]
CV	Classification	Industrial chemicals	LSTM, CNN	[31]
Enose	Feature extraction	Harmful gases	PCA	[32]
	Classification		DT, RF, SVM	
Enose	Regression	Formaldehyde	SVR	[33]
	Classification		BPNN	
Enose	Classification	Chinese wines	BPNN	[34]
	Target task change	Chinese liquors	Transfer learning	
Enose	Sensor drift compensation for classification	Gases	JDA	[35]
			DTBLS	[36]
			TrLightGBM	[37]
			ELM	[38]
Enose	Sensor drift compensation & noise reduction	Bacteria	ELM	[39]
EIS	Classification	Breast tissue	ELM + SVM	[40]
EIS	Classification	Milk adulteration	<i>k</i> -NN	[41]
EIS	Classification	Breast tissue	RBFN	[42]
EIS	Feature extraction	Avocado ripeness	PCA	[43]
	Classification		SVM	
EIS & EIT	Classification	Prostatic tissue	SVM	[44]
Etongue	Taste classification	Tea storage time	CNN	[45]
	Increase generalizability		Transfer learning	
Etongue	Feature Extraction	Beverages	<i>t</i> -SNE	[46]
	Classification		<i>k</i> -NN	
Etongue	Classification	Cava wine age	LDA	[47]
Etongue	Regression	Black tea theaflavin	Si-CARS-PLS	[48]
OPTICAL				
Colorimetric	Classification	Plant disease VOCs (blight)	PCA	[49]
Diff. contrast microscopy	Digital staining & domain adaptation	Leukocytes	GAN	[50]
Fluorescence imaging	Classification	Microglia	ANN	[51]
FTIR imaging	Digital staining	H&E stain	Deep CNN	[52]
Lens-free imaging	Image reconstruction	Blood & tissue	CNN	[53,54]
		Herpes		[55]
Lens-free imaging	Image reconstruction & classification	Bioaerosol	CNN	[56]
Multi-modal multi-photon microscopy	Digital staining & modal mapping	Liver tissue	DNN	[57]
Multispectral imaging	Classification	Pollen species	CNN	[58]
Quantitative phase imaging	Digital staining	Skin, kidney & liver tissue	GAN	[59]
Raman spectroscopy	Feature extraction	Thyroid dysfunction biomarker	PCA	[60]
	Classification		SVM	

Table 1. Cont.

Biosensing Mechanism	Task	Target	Algorithm	Ref.
TLC-SERS	Feature extraction	Histamine	PCA	[61]
	Quantification		SVR	
SERS	Exploratory analysis	Malachite green & crystal violet	PCA	[37,62]
	Quantification		PLSR	
SERS	Quantification	Methotrexate	PLSR	[63]
SERS	Classification	Oil vs lysate spectra Leukemia cell lysate	<i>k</i> -means clustering	[64]
	Dimension reduction		PCA	
	Classification		SVM	
SERS	Dimension reduction	Levofloxacin	PCA	[38,65]
	Regression		PLSR	
SERS	Quantification	Potassium sorbate & sodium benzoate	PLSR	[66]
SERS	Dimension reduction & regression	Congo red	PCR	[39,67]
SERS	Dimension reduction	Mycobacteria	PCA	[40,68]
	Classification		LDA	
SERS	Quantification	Biofilm formation	PLSR	[41,69]
SERS	Feature extraction	Non-structural protein 1	PCA	[70,71]
	Classification		BPNN, ELM	
SERS	Exploratory analysis	Pollen species	PCA, HCA	[72]
	Classification		ANN	
SERS	Feature extraction	Human serum	KPCA	[73]
	Classification		SVM	

Note. CV = cyclic voltammetry; ANN = artificial neural network; LSTM = Long short-term memory; PCA = principal component analysis; DT = decision tree; RF = random forest; SVM = support vector machine; SVR = support vector regression; BPNN = back-propagation neural network; JDA = joint distribution adaptation; DTBLS = domain transfer broad learning system; GBM = gradient boost machine; ELM = extreme learning machine; EIS = electrical impedance spectroscopy; EIT = electrical impedance tomography; *k*-NN = *k*-nearest neighbor; RBFN = radial basis function network; CNN = convolutional neural network; *t*-SNE = *t*-distributed stochastic neighbor embedding; Si-CARS-PLS = synergy interval partial least square with competitive adaptive reweighted sampling; FTIR = Fourier transform infrared; VOC = volatile organic compound; GAN = generative adversarial network; DNN = deep neural network; TLC = thin layer chromatography; SERS = surface enhance Raman spectroscopy; PLSR = partial least squares regression; PCR = principal component regression; LDA = linear discriminant analysis; HCA = hierarchical cluster analysis; KPCA = kernel principal component analysis.

Table 2. Summary of data types and useful ML methods for biosensing mechanisms.

Biosensing Mechanism	Description of Data	Feature Extraction	ML Model
CV	Cyclic voltammogram		ANN, LSTM, CNN
EIS	Nyquist plot	PCA	<i>k</i> -NN, ELM, SVM, RBFN
Enose	Multivariate	PCA	DT, RF, ELM, SVM, BPNN
Etongue	Multivariate	PCA, <i>t</i> -SNE	LDA, <i>k</i> -NN, CNN, PLS
Lens-free imaging	Image		CNN
Digital staining	Image		Deep learning, GAN
SERS	Spectrum	PCA, KPCA	PLSR, LDA, SVM, SVR, BPNN, ELM

3. A Brief Tour of Machine Learning

In simple terms, machine learning aims to learn patterns in data to make predictions on new data. Generally, this prediction is either categorical classification (into one of a set of classes) or regression (continuous numerical output). In machine learning terms, the data used for prediction (i.e., biosensor data) are termed features or predictors. The set of features associated with one “observation” (e.g., biosensor data from one sample) is termed the feature vector.

3.1. Feature Engineering

Frequently, the predictor variables (feature vector) are not the raw biosensor data. One of the most challenging parts of using machine learning is the construction of the feature vector from the raw data. This process is termed feature engineering and mostly entails finding the relevant information from the data to aid the machine learning algorithm's performance. Common feature engineering steps include denoising, normalization, and rescaling.

One of the most powerful feature engineering processes is dimension reduction. This reduces a large number of features to a smaller number of features while minimizing information loss. Perhaps the most common method of dimension reduction is principal component analysis (PCA) [74], which reduces the original set of variables to a smaller set of independent variables termed principal components (PCs). The effectiveness of PCA to represent the data can be assessed by the amount of variance in the data explained by the PCs. Since PCA determines the PCs based on the eigenvectors' directions in the feature space, data must first be centered and rescaled to avoid bias toward those variables with a larger magnitude. Another common dimension reduction algorithm is linear discriminant analysis (LDA), which also produces a smaller number of variables but is supervised and optimally maximizes class separation [75]. Other more complex dimension reduction methods exist including artificial neural networks (ANN), as discussed in Section 3.3. ANN is typically used as a supervised machine learning method, while it has occasionally also been used for dimension reduction.

3.2. Unsupervised vs. Supervised

The two broad categories of machine learning algorithms are unsupervised and supervised [76]. In unsupervised methods, data labels are not provided during model training, while in supervised methods, they are. An example of an unsupervised algorithm is cluster analysis, used to group similar data. Unsupervised methods are less common in biosensing since we generally know what kind of prediction(s) we would like the model to make. A notable exception is PCA, as mentioned in Section 3.1. While PCA may be considered an unsupervised machine learning method, its use has recently been limited to dimension reduction (one of feature engineering processes) prior to supervised machine learning analyses.

3.3. Classification Algorithms

Among the supervised methods, classification algorithms are some of the most well-known. Classification gives prediction in the form of a class label (e.g., which bacteria species is present), thus, the output is inherently categorical. Briefly, some of the most common classification algorithms are presented in the following.

k-nearest neighbors (*k*-NN): One of the simplest classification algorithms, *k*-NN is a distance-based classifier. Class is predicted as the most common class of the *k*-nearest neighbors in the feature space [77]. In the example shown in Figure 1, the feature space is two dimensional (with variables x_1 and x_2) and the value of *k* is 4. In *k*-NN, the number of neighbors used for assignment, *k*, is a hyperparameter (i.e., a model parameter that is not optimized during the training process itself). As with most ML models, hyperparameter selection may strongly influence performance [78].

Support vector machine (SVM) is a non-probabilistic, binary, linear classifier [79]. SVM relies on the construction of hyper-plane boundaries in the feature space to separate data of different classes. Although SVM itself only accounts for linear separation of classes (i.e., hyper-plane boundaries must be "flat"), the data may be mapped to a higher-dimensional feature-space using the "kernel trick" [80]. Some of the most common kernels are radial basis function and Gaussian. When the hyperplane boundaries are projected back into the original feature space, they allow for non-linear boundaries, as shown in Figure 1. Additionally, there are methods allowing SVM to be used for multi-class prediction [81]. The placement of hyperplanes is determined by minimizing the distance between the

hyperplane and several of the points closest to the boundary between classes. SVM's robustness against outliers is improved by a soft margin. This allows for a certain quantity of misclassifications, which are presumably outliers, to improve the separation of the other observations [82]. While SVM shows resilience against outliers and performs well in high-dimension feature spaces, it is prone to over-fitting, especially when using non-linear kernels [83]. Overfitting is when the model performs well on training data but performs poorly when generalized to unseen data.

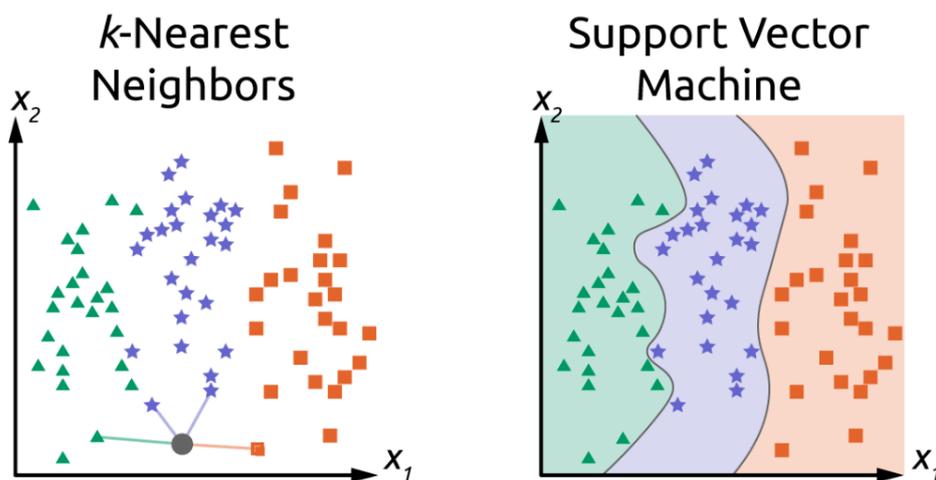


Figure 1. Comparison of classification technique using k -NN and SVM. In k -NN, four nearest neighbors are shown contributing to the gray point's assignment. Classification of the gray point is the blue star class. In hypothetical SVM with nonlinear kernel, new data are classified in which region the point lies. In both examples, the feature space consists of two dimensions. Classification could be, for example, bacterial species like *E. coli*, *Salmonella* spp., *Pseudomonas* spp., *Staphylococcus* spp., *Enterococcus* spp., etc. In practical applications, the feature space has many more dimensions, where decision boundaries for SVM are hyperplanes in the $(n-1)$ dimension for an n -dimensional feature vector.

Linear discriminant analysis (LDA): In addition to dimension reduction, LDA can be used for classification. Other related algorithms allow for non-linear classification such as quadratic discriminant analysis (QDA) [84]. One of the limitations of LDA and its relatives is that they assume the data are normally distributed.

Decision tree (DT) and random forest (RF): In tree-based models such as decision tree (DT), the feature vector starts at the tree's "trunk," and at each branching point a decision is made based on the learned decision rules. The end classification would then be at the terminal or "leaf" node that the instance results. DTs can be used for classification and regression [85]. When the target variable is categorical, it is referred to as a classification tree; when the target variable is numerical and continuous, it is referred to as a regression tree [86]. Random forest (RF) is so called because it can be considered a forest of decision trees (Figure 2) [87]. There are many RF architectures, but in all instances, the classification from each decision tree contributes to the overall classification for an observation.

Artificial neural network (ANN) draws inspiration from biological neural networks (i.e., neurons in the brain) and is composed of a collection of connected nodes called artificial neurons (see Figure 3). ANNs can be used for classification and regression. As mentioned earlier, ANN can be used for dimension reduction prior to supervised machine learning. There are a large variety of ANN structures such as (1) recurrent neural network (RNN) [88], (2) extreme learning machine (ELM) [89], and (3) deep learning algorithms such as the convolutional neural network (CNN) [90], deep belief network [91], and back-propagation neural network (BPNN) [92]. "Deep" indicates several hidden layers. ANN architectures have many hyperparameters such as the number of hidden layers, connectedness, and activation functions [93].

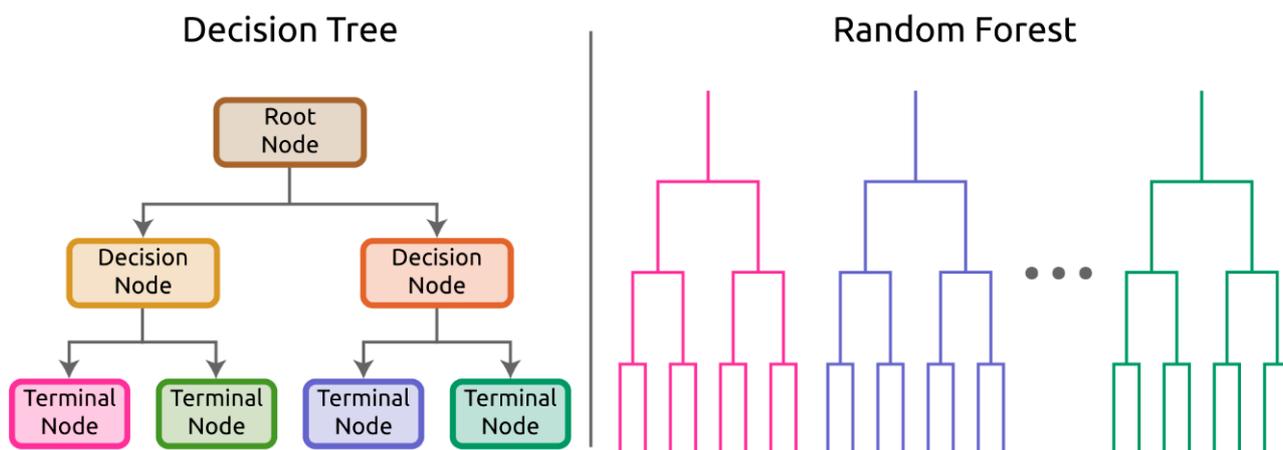


Figure 2. Decision tree (DT) showing nodes at which binary decisions are made on features. Terminal node dictates model prediction. Actual DTs have many more nodes than shown here. Random forest (RF) shown as a series of distinct decision trees.

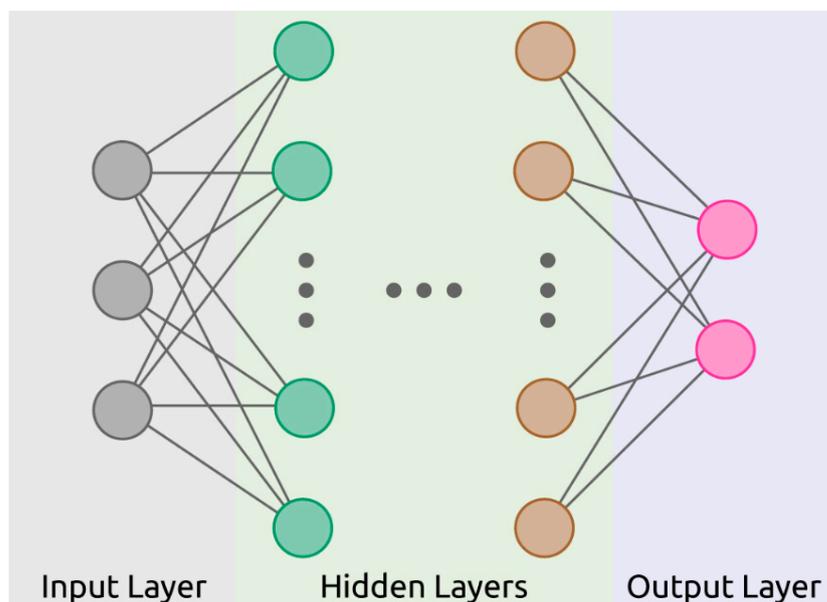


Figure 3. Artificial neural network (ANN) showing nodes of the input, hidden, and output layers.

One of the aspects that makes ANN so powerful is that features do not need to be well-defined real numbers. This allows them to excel at working with data such as images for which extracting numerical features would be difficult and inefficient. One limitation of ANNs is that they require a large amount of data for effective training. In some settings, training data sparsity can be mitigated through a generative adversarial network (GAN) using back propagation [94].

Common classification model performance metrics are accuracy, precision, sensitivity (also known as recall), specificity, and $F1$. For binary classification with labels “positive” and “negative”, they are defined as follows:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$precision = \frac{TP}{TP + FP} \quad (2)$$

$$sensitivity = \frac{TP}{TP + FN} \quad (3)$$

$$specificity = \frac{TN}{TN + FP} \quad (4)$$

$$F1 = \frac{2 \times precision \times sensitivity}{precision + sensitivity} \quad (5)$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

3.4. Regression Algorithms

In contrast to classification, the prediction made by a regression algorithm is a numeric value from a continuous scale (e.g., glucose concentration in blood). A simple regression example fits a linear model of the form $y = mx + b$, where a model is built for the prediction of the output variable y based on the input variable x , and the coefficients m and b are “learned” from the data. The learning is typically done by the least-squares regression approach, minimizing the sum of the squared residuals. The following are some of the most common regression algorithms.

Multilinear regression (MLR) is a simple regression model, which expands the above linear model example, accounting for multiple input variables. This model shows how it can be difficult to determine when an algorithm becomes sophisticated enough to be considered “machine learning”.

Support vector regression (SVR) is an adaptation of SVM used for regression problems. Like SVM, SVR can utilize kernels to allow for non-linear regression. An advantage of SVR over traditional regression is that one need not assume a model that might not be accurate. For instance, with linear regression, there is an assumption that the data distribution is linear. SVR does not require such pre-determined assumptions [95].

Regression tree is an adaptation of DT for regression. Regression tree has the advantage that it is non-parametric, implying that no assumptions are made about the underlying distribution of values of the predictors [86].

Artificial neural network (ANN) is also widely used for regression problems, and many varieties exist, some of which were mentioned previously.

A large variety of metrics exist for regression model performance. Since there are too many to define here, for further reading, we suggest the study by Hoffman et al. [96] to learn more. Some of the most common metrics are briefly presented here. Root mean squared error (RMSE) and mean absolute error (MAE) have the benefit that their units are the same as the output (predicted) variables, but this makes the metrics less universally understandable. Normalized root mean squared error (NRMSE) partially resolves that. Coefficient of determination, R^2 , on the other hand, is unitless and $R^2 \leq 1$, where a value near 1 is generally considered good performance (although this is a bit oversimplified).

3.5. Model Performance Assessment

Frequently, researchers will try various models and compare their performance. The value of the performance metrics listed above can be treated as random variables and statistical analyses can be used to test hypotheses regarding which model is better [96]. While this sounds simple, it can be nuanced: for instance, when working with a classification model, which metric is most important for your application? In some cases, specificity may be more important than accuracy, for instance. Additionally, when using statistical tests to compare model performances, certain assumptions are made, and their validity should be assessed such as when using NRMSE, as it is assumed that noise affecting the output is random and normally distributed.

The best practice for model selection, tuning, and performance assessment is to split the data into 3 sets: training, testing, and validation. For example, if the database consists of 1000 observations, 100 (10%) are assigned to the validation set and the remaining 900 (90%) are split between the training and test sets as 810 (90%) for training, 90 (10%) for testing. The model is then trained on the labeled training set. Model selection and hyperparameter tuning is conducted based on model performance when challenged using the test set. In addition to train–test splitting, cross-fold validation can be used on the training set

when tuning hyperparameters or comparing models [97]. Train–test splitting and cross-validation are most important when you intend to generalize the model to predict new, unseen data [96]. Final model performance validation is conducted on the validation set, which should not be used until all model selection and hyperparameter tuning have been completed.

4. Electrochemical Bioreceptor-Free Biosensors

Since their inception, electrochemical biosensors have become extremely popular. In traditional electrochemical biosensors, the bioreceptor interacts with the target to generate a signal at the electrical interface. A widespread scheme is an enzyme (e.g., glucose dehydrogenase or glucose oxidase) interacting directly with the target analyte (e.g., glucose), catalyzing a redox reaction that generates a signal at the electrical interface [98]. Electrical interfaces include metal electrodes, nanoparticles, nanowires, and field-effect transistors (FET) [99].

It is also possible to eliminate the biorecognition element (=bioreceptor, e.g., an enzyme) in electrochemical biosensors. Voltammetric sensors described in Section 4.1 can detect biomolecules based on direct interaction with the electrical interface [30]. Electrical impedance spectroscopic biosensors can also detect subtle differences in a solution or material's electrical impedance, as discussed in Section 4.2. Alternatively, we can use an array of chemical or physical sensors varying the electrical interface to create multi-dimensional data. Machine learning-based pattern recognition is used to identify the target analyte. Two of the most common sensor arrays are termed Enose and Etongue, which are covered in Section 4.3.

4.1. Cyclic Voltammetry (CV)

Voltammetry sensors apply electric potential to a “working” electrode and measure the current response, which is affected by analyte oxidation or reduction [100]. Cyclic voltammetry (CV) is a specific voltammetry technique in which the potential is swept across a range of values, and current response is recorded. These CV curves (cyclic voltammograms) can serve as a fingerprint of the sensor response. A typical CV curve is shown in Figure 4A.

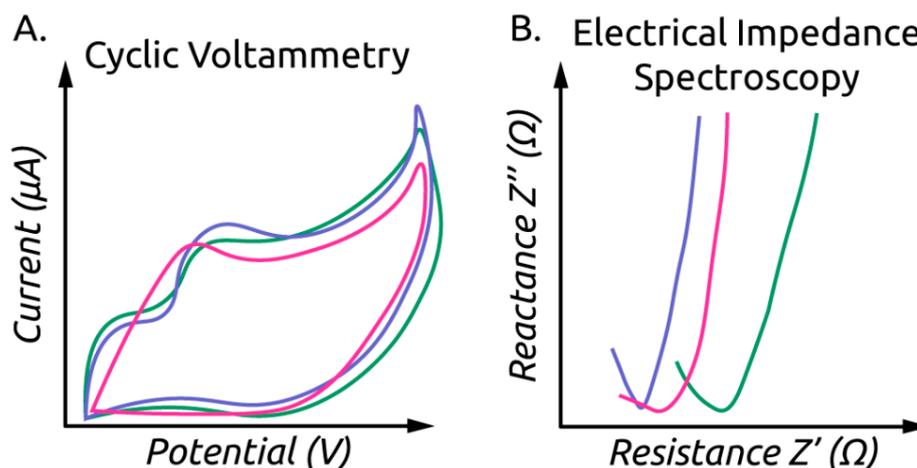


Figure 4. (A) Hypothetical cyclic voltammograms for three samples. (B) Hypothetical Nyquist plot obtained through EIS showing curves for three samples.

CV biosensors often employ bioreceptors to provide specificity in the interaction between target analyte and electrode surface. However, there has also been research on utilizing more complex electrode surface structures and modifications to allow for semi-specific interaction with the target analyte without the need of a bioreceptor. Sheng et al. [30] describe a compound electrode utilizing Cu/PEDOT-4-COOH particles for CV detection of the phytoinhibitor maleic hydrazide. They found that several regression models had poor

performance for modeling the sensor current response with respect to target concentration. However, they employed an ANN with great success for the same regression task. The result is that their detection range is broader than comparable methods by an order of magnitude at each extreme (detection range = 0.06–1000 μM and LOD = 0.01 μM).

4.2. Electrical Impedance Spectroscopy (EIS)

Electrical or electrochemical impedance spectroscopy (EIS) is an analytical technique that provides a fingerprint of the electrical properties of a material. EIS is performed by applying a sinusoidal electric potential to a test sample and recording the impedance (both resistance and reactance expressed in a complex number) over a range of frequencies [101]. Frequently, an equivalent circuit model is fitted to EIS data to provide a fingerprint of the material properties [101]. Figure 5 shows an equivalent circuit diagram for EIS being performed on a single cell suspension. An example EIS spectrum is shown in Figure 4B. It is the classification and regression on such fingerprints that machine learning tends to be well suited.

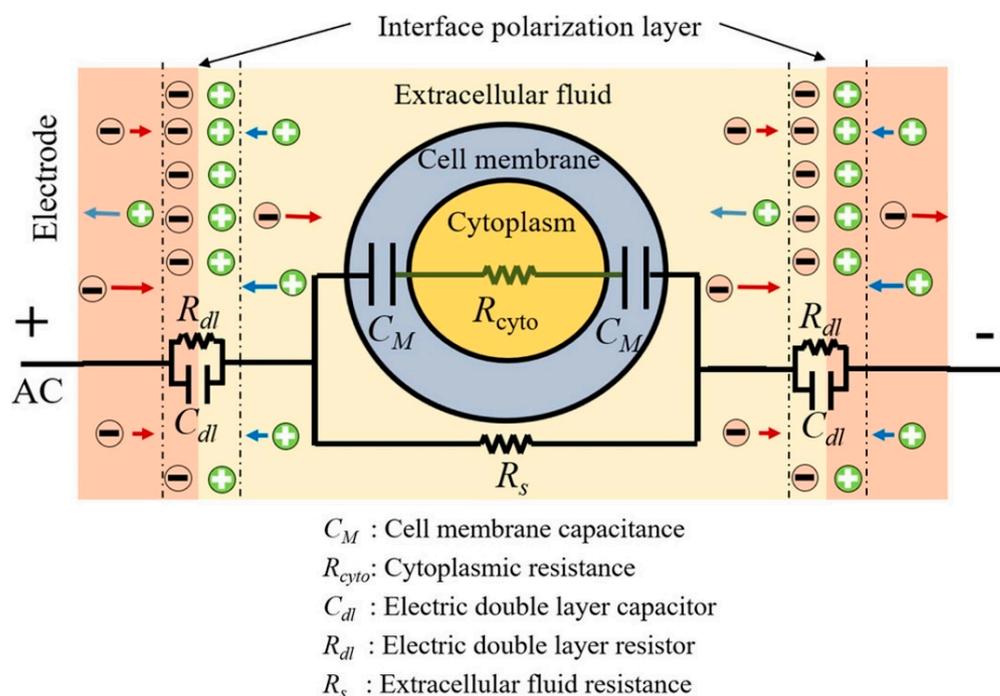


Figure 5. Equivalent circuit diagram of single cell suspension. Reproduced with permission from [102] without modification. Copyright 2020 John Wiley and Sons.

A simple example of this is the use of k -NN on EIS data for the detection of adulteration in milk [41]. In this work, the feature space was composed of resistance at a certain temperature and pH. They demonstrated good accuracy of 94.9%. However, the data were highly imbalanced, and in the example classification plot [41], one of the three unadulterated samples were misclassified, a 66% specificity.

More robust classification has been performed using SVM. One example is for the assessment of avocado ripeness [43]. This work describes using PCA for feature extraction, resulting in two PCs that explain >99.3% of the variance. SVM is then used for classification based on the first two PCs. SVM for EIS was also described by Murphy et al. [44] for classification of malignant and benign prostatic tissue. However, instead of using PCA for feature extraction, equivalent electrical circuit model parameters were used as predictors. The feature vector size was 2160, consisting of four electrical features for each of eleven frequencies across multiple electrode configurations. Classification was also performed on electrical impedance tomography (EIT) data from the same samples using SVM. Both showed good classification performance, though the authors mention that EIT may be

preferable since the measurements are not dependent on probe electrical properties, and thus can be compared more easily to other studies.

While SVM is renowned for its tolerance of outliers, this is a trade-off in that data points not near the boundary between classes do not contribute to defining class attributes. However, ANNs preserve more of this information for prediction. When the number of observations or predictors are small, this can lead to overfitting. However, with sufficient data size, ANNs can preserve predictive information and be robust against outliers and overfitting. These attributes have been utilized for EIS based classification of breast tissue [40,42]. Both works use the same publicly available dataset of EIS measurements from freshly excised breast tissue [103], made available on the University of California, Irvine (UCI) Machine Learning Repository [104]. The dataset contains nine spectral features from EIS. Daliri [40] describes using three ELMs, each with different numbers of nodes, and feeding the output of the three ELMs (extreme learning machines) into SVM for classification. This method showed improved performance over previous methods for the same dataset such as LDA [105]. Helwan et al. [42] compared both BPNN and radial basis function network (RBFN) for the same task. Both methods showed an improvement over ELM-SVM as described by Daliri [40], with RBFN performing better than the BPNN including improved generalizability (i.e., classification performance on new data).

It is seen that in the case of EIS classification, node-based models have shown improved performance over other models. This can be seen most clearly when comparing classification accuracy for those methods that utilized the same dataset. The RBFN and BPNN had the highest classification accuracy, with 93.39% and 94.33%, respectively [42]. The next best performance was achieved by the ELM-SVM, achieving 88.95% accuracy [40]. These results show marked performance increase over LDA [105]. Model performance is greatest in those models that do not utilize distance for classification (i.e., SVM and LDA). While distance-based classifiers are robust to outliers, in these EIS datasets, performance benefitted by node-based classification.

4.3. Enose and Etongue

Enose and Etongue are named in analogy to their respective animal organs. Both sensor types rely on an array of semi-specific sensors, each of which interacts to a different degree with a wide range of analytes. Figure 6 shows a comparison between Enose and Etongue alongside the analogy to their respective biological systems [27,106]. The sensor arrays can be composed of any variety of sensors. The following chemical gas sensors have been used in Enose systems: metal oxide (MOX) gas sensor, surface or bulk acoustic wave (SAW and BAW) sensors, piezoelectric sensor, metal oxide semiconductor field-effect transistor (MOSFET) sensor, and conducting polymer (CP) based sensor [107]. Similarly, a variety of sensors can be employed in Etongue systems such as ion-selective field-effect transistor (ISFET) and light-addressable potentiometric sensor (LAPS) [108].

Analyte presence, or a more general attribute such as odor or taste, is detected through pattern recognition of the sensor array response. For pattern recognition on this naturally high-dimensional data, machine learning techniques are an obvious choice. Scott et al. provided a relevant and succinct paper on data analysis for Enose systems [23]. As discussed in Section 3 of this review, feature engineering is critical in any machine learning pipeline. Yan et al. [24] provide a review article on the feature extraction methods for Enose data. For non-linear feature extraction of Etongue data, Leon-Medina et al. [46] give a great comparison of seven manifold learning methods.

A vast number of papers exist detailing such systems and their use of machine learning. As such, it would be infeasible to cover all of them adequately. For this review, a higher-level analysis is presented by looking at the conclusions reached in the review papers covering this topic as well as a few notable examples of specific papers. Of particular interest is which algorithms had the most success with Enose and Etongue sensors or applications.

A common task of Enose is the prediction of “scent”, which is a classification problem. Before the application of the classification algorithm, it is common to perform dimension

reduction. PCA is the most common choice for this task, although independent component analysis (ICA, a generalization of PCA) has shown success [25]. PCA has been shown to improve the performance over classification algorithms alone for the piezoelectric Enose [25]. The two classifiers most commonly in use are SVM [109,110] and various ANN methods [25,111]. In addition to classification problems, Enose may be used for analyte concentration prediction. One example is the use of MOS (metal oxide semiconductor) gas sensors for formaldehyde concentration assessment. In this case, the back-propagated neural network (BPNN) outperformed radial basis function network (RBFN) and SVR [33]. In another instance, with the single nickel oxide (NiO) gas sensor, PCA with SVR was utilized for harmful gas classification and quantification [32]. In cases where the amount of data are not large, SVM may be advantageous over node-based models (ANNs) for its resilience against outliers and overfitting.

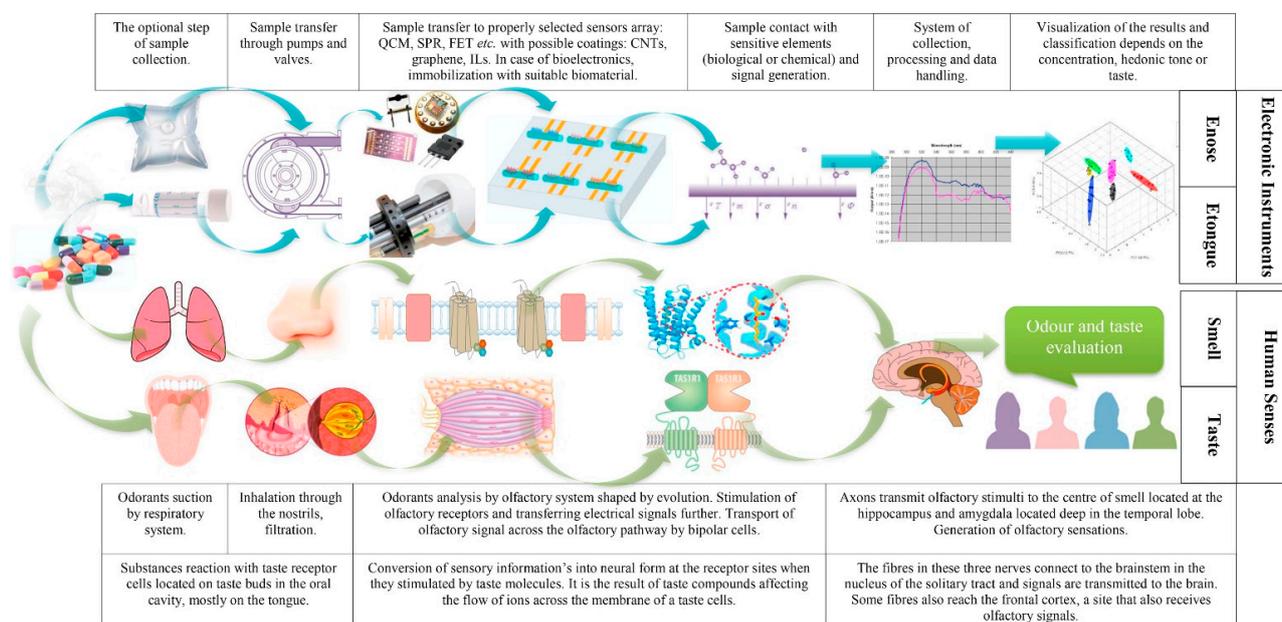


Figure 6. Comparison of operation principle of Enose and Etongue, and the analogy to the biological systems. Reproduced with permission from [27] without modification. Copyright 2019 Elsevier.

While Enose and Etongue systems have shown great promise for non-destructive analytical devices, there are challenges that have limited their use in commercial settings. Several challenges involve changes in the sensor data, which affect the performance of the trained model. A common phenomenon is when the sensor array response changes over time or upon prolonged expose under identical conditions. Such change in sensor response is referred to as sensor drift and can greatly affect the trained models' performance [14]. Another way in which the sensor response may change is if a sensor in the array becomes defective and must be replaced, as it is difficult to replace it with one that responds identically, largely due to variability in manufacturing [112,113]. For both challenges, time consuming and computationally expensive recalibration may be necessary.

The issue of needing retraining due to underlying data distribution changes is commonly addressed through transfer learning in many machine learning settings. Transfer learning is a computational method for minimizing the need for retraining when either the data distributions change (e.g., sensor array response to an analyte) or the task changes (e.g., new classes of analytes are being detected).

Transfer learning has been extensively employed to counter Enose sensor drift and reduce the need for complete retraining [35–38]. It has also been used to reduce the deleterious effect of background interference [39,114]. Although several of the above papers [35,36,38,39] demonstrate the efficacy of their approach on a shared sensor drift dataset shown in Figure 7 [115], ranking of the methods is difficult due to inconsistent

benchmarking metrics. As mentioned previously, the data distribution may also change due to replacing a sensor with a new sensor, or when attempting to apply a trained model to a theoretically identical array with differences due to manufacturing variability. Transfer learning, specifically using ANN, has demonstrated decent recalibration [116].

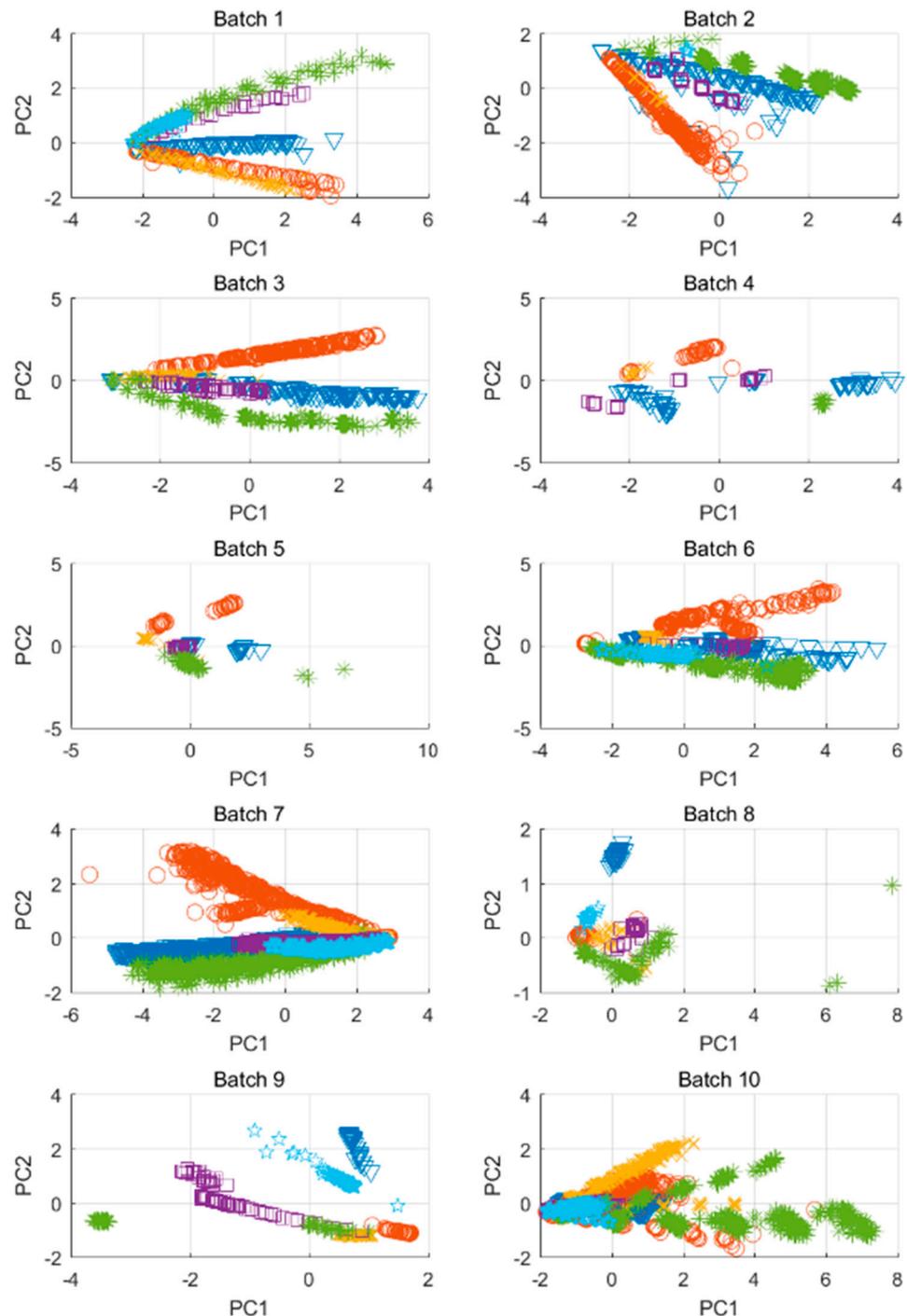


Figure 7. Gas sensor drift dataset from [36]. Each color represents a different gas. Each panel represents a measurement “batch” at various times spanning 36 months. Reproduced from [36] without modification, under Creative Commons Attribution 4.0 License.

One instance of utilizing transfer learning for target task change was demonstrated by Yang et al. by training an Enose classifier on wines (source task) and applying it to classify Chinese liquors (target task) while only retraining the output layer [34]. Interestingly,

transfer learning has been used much less commonly for Etongue systems, although they also face sensor drift. However, Yang et al. utilized transfer learning to improve the generalizability of their Etongue [45]. In this work, they demonstrate the superiority of their transfer learning trained CNN over other methods such as BPNN, ELM, and SVM for tea age classification.

A trend that has been gaining traction is data fusion to combine Enose and Etongue systems. The value of this can again be appreciated in how closely the senses of smell and taste are linked in animals [117], complementing each other to provide the most accurate assessment. Similarly, by using information from both Enose and Etongue, better analysis can be conducted. As illustrated in Figure 8, data fusion can be performed at three levels: low, mid, and high [118]. Recently, mid-level fusion schemes have shown promising results for fusion of Enose and Etongue data [119,120], especially when performing PCA on the two systems and using those features for fusion before model training [121–123]. Such systems have also benefitted from the inclusion of a computer vision system in data fusion [121,124].

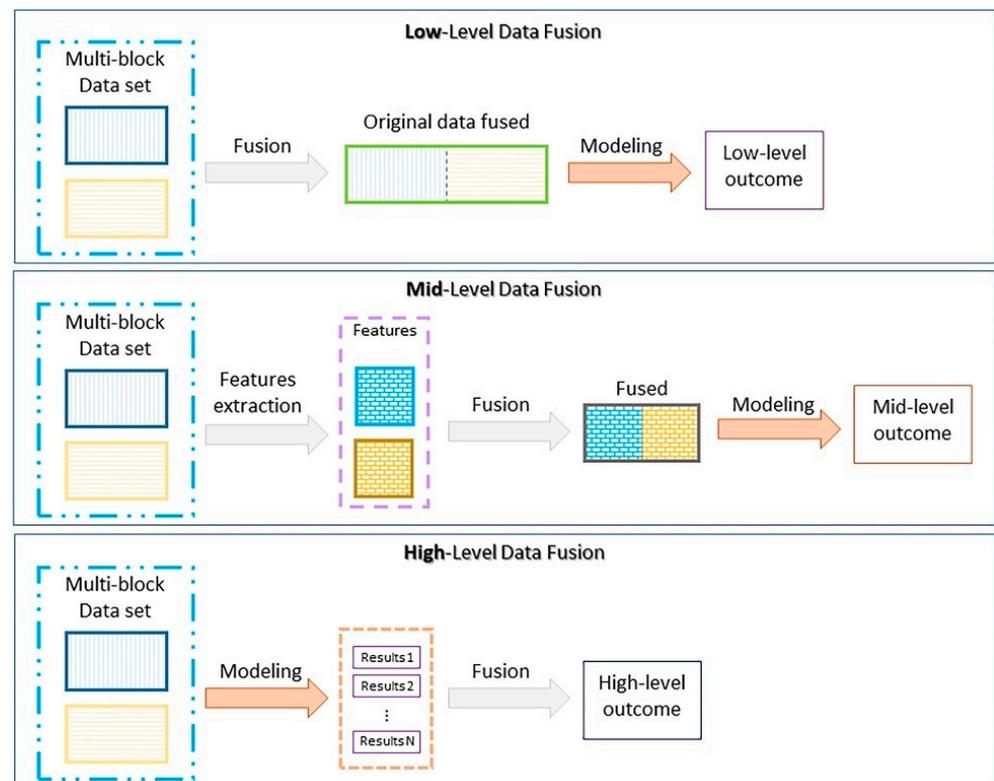


Figure 8. General scheme depicting the main differences among low-, mid-, and high-level data fusion. Reproduced with permission from [118] without modification. Copyright 2019 Elsevier.

Currently, another class of systems exist with the same goals as Enose and Etongue that utilize biochemical recognition elements, termed bioelectronic nose (bEnose) and tongue (bEtongue). These devices utilize biological elements such as taste receptors, cells, or even tissues for sensing [106,125]. These systems show impressive selectivity and sensitivity, especially when coupled with nanomaterials to aid in signal transduction from the biochemical recognition element [106,126]. Their major challenges, as with most biosensors, is stability and reproducibility of the biological element [106]. For these reasons, Enose and Etongue remain popular for their sensor stability. Continued efforts are necessary to improve sensitivity closer to their bioelectronic counterparts, especially regarding sensor design and feature extraction methods.

With such a large variety of sensors in use for Enose and Etongue systems, data processing can vary significantly. Of particular interest is finding appropriate feature

extraction methods [23,24]. A huge variety of machine learning classification and regression methods have been employed, both on unsupervised dimensionally reduced feature vectors and classically extracted features. Transfer learning methods have been successful in allowing target task change with minimal retraining, especially when using node-based models. However, the challenges posed by sensor drift and manufacturing variability are still significant and will likely remain a focus for researchers over the next several years.

4.4. Summary of Electrochemical Bioreceptor-Free Biosensing

Many electrochemical bioreceptor-free biosensors employ chemical or physical sensor arrays coupled with machine learning. These are most obvious in Enose and Etongue systems, inspired by nature (humans and animals). Other systems generate multivariate spectral data also coupled with machine learning. In both cases, machine learning models can aid in analyte classification or quantification. Especially when using distance-based models, choice of feature extraction method is important to optimally capture the features relevant to the task (i.e., classification or regression). Node-based models, primarily ANNs often require less feature extraction pre-processing as this step is built into the model learning. Additionally, node-based models offer a great solution to target task change and noise elimination through transfer learning, often aided by integration through the back-propagation step so that only the final layer needs to be refined [34].

5. Optical Bioreceptor-Free Biosensors

The mechanisms of optical detection in biosensing are diverse. A classic example is the colorimetric lateral flow assay [127–129]. Mechanisms beyond colorimetry include fluorescence [130–132], luminescence [133], surface plasmon resonance [134], and light scattering [135,136].

Machine learning has been widely employed in optical biosensors. An example with similarities to Enose and Etongue is the bacterial bioreporter panel. Each bacterial bioreporter responds to target analytes in a semi-specific manner. Machine learning is used to discover patterns in the bioreporter panel response and relate them to analyte presence or concentration [137,138]. However, this review's focus is to discuss cases in which the bioreceptor is absent, so such sensors are not covered in detail.

Another prevalent use of machine learning for analyzing images as biosensor data is for image processing, especially segmentation [139–142]. The literature is rich in reviews on machine learning for image segmentation, and this technology is in no way specific to biosensors, so this review will not discuss those examples. However, the topic is essential to many biosensors, so it must be mentioned.

5.1. Imaging

Imaging sensors utilize an array of optical sensors such as a CMOS array (complementary metal-oxide-semiconductor array; the most used image sensor for digital cameras). Images of the specimen can be used to identify the target presence and concentration as the molecules exhibit different coloration, fluorescence, or light scattering, with varying morphology and spatial distribution. In this manner, several imaging biosensors have been developed to eliminate the need for labels and bioreceptors.

A growing field of imaging-based biosensors utilizes lens-free imaging techniques [143,144]. Since the images from lens-free imaging are not in focus, computational techniques are needed for image reconstruction, the most common of which is deep learning (mostly based on ANN with “deeper” layers) [53,54,145]. Lens-free imaging may be used to detect the aggregation of particles caused by bioreceptor–analyte interaction [55] (Figure 9). However, an exciting application is the direct, label-free classification of particles by lensless holography. Wu et al. [56] presented a lensless holography biosensor for classifying pollen and spores. As with many of these systems, a CNN was used for image reconstruction. In this work, another CNN was used to classify the particles, yielding > 94% accuracy.

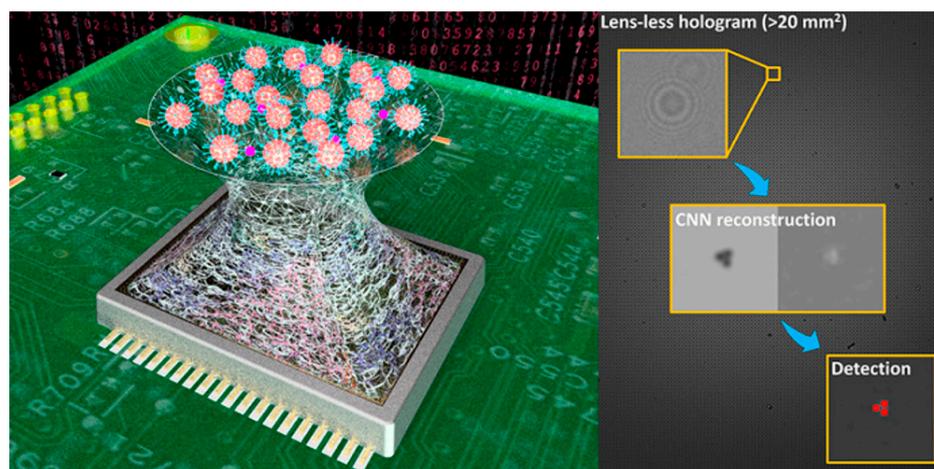


Figure 9. Lensless hologram reconstruction via CNN for particle aggregation detection. Reprinted with permission from [55] without modification. Copyright 2019 American Chemical Society.

Another work on the imaging classification of pollen utilizes multispectral imaging [58]. Again, a CNN was trained for classification, and a species-averaged accuracy of 96% was achieved for 35 plant species.

Artificial neural networks (ANNs) have also found great success in the developing field of digital staining. Hematoxylin and eosin (H&E) stain is the most common stain for histology [146]. However, the quality of tissue staining is subject to many factors that can affect the diagnosis. Digital staining is an alternative in which tissue sections are imaged unstained, and a trained model generates an image simulating stained tissue (Figure 10). Deep learning has been applied for digital staining on images acquired from a variety of methods including quantitative phase imaging [59], Fourier transformed infrared spectroscopy (FTIR) [52], and multi-modal multi-photon microscopy [57]. To overcome the issue of data scarcity and overfitting, researchers have frequently employed generative adversarial neural network (GAN) for medical imaging [147], which has shown promising results for digital staining model training [148]. Additionally, transfer learning has improved the model's generalizability to multiple domains [50].

Fluorescence-based imaging biosensors are also worthy of mention. Sagar et al. [51] presented a microglia classification based on fluorescence lifetime utilizing ANN.

The applications of imaging biosensors are extensive. Indeed, the scope is too large to analyze all papers in this review. However, of particular importance to imaging biosensors is the ANN, especially the CNN. This preference is expected since CNN has shown exceedingly good performance in a variety of image classification contexts [149,150].

5.2. Colorimetry

One class of optical biosensors is the colorimetric biosensor. Currently, the applications of machine learning to enhance the performance of bioreceptor-free colorimetric biosensors are limited. This limitation is because the colorimetric biosensors (most notably lateral flow assays) mostly utilize bioreceptors (e.g., antibodies, enzymes, and aptamers) [98]. One example of such a bioreceptor-free biosensor is non-invasive plant disease diagnosis by Li et al. [49]. They utilized an array of plasmonic nanocolorants and chemo-responsive organic dyes that interact with volatile compounds from the plant. Their technique is similar to Enose and Etongue since it is a fingerprinting approach to the array response for classification. They used PCA, but do not cite an actual classifier, although they give performance metrics such as accuracy. At this time, it is unclear how the classification was performed on the PCA-transformed data.

Most colorimetric biosensors do not require machine learning due to their simplicity for readout. However, the arrays of bioreceptor-free (semi-specific) colorimetric sensors require machine learning-based classification in a way similar to Enose and Etongue.

In these instances, they will likely benefit from the same treatment, namely dimension reduction by PCA and SVM classification.

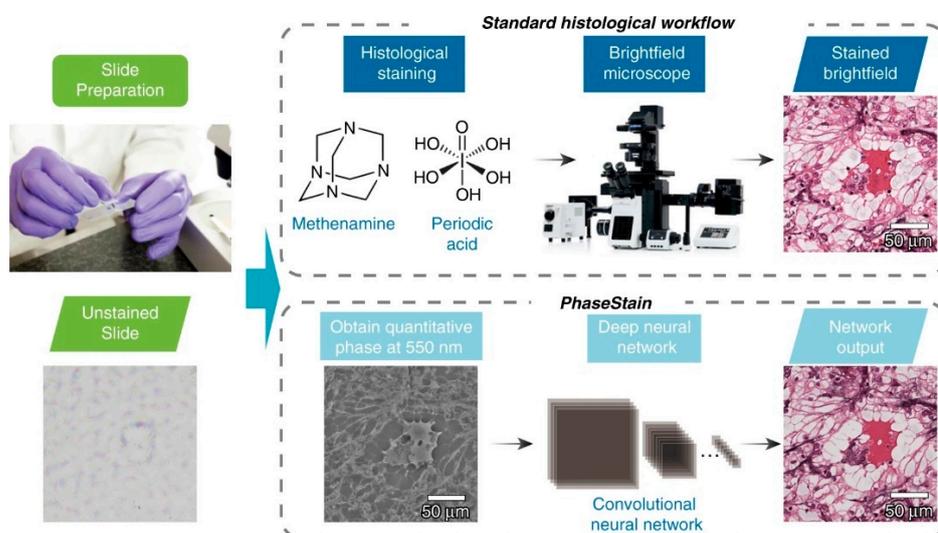


Figure 10. A quantitative phase image of a label-free specimen is virtually stained by a deep neural network, bypassing the standard histological staining procedure that is used as part of clinical pathology. Reproduced from [59] without modification under Creative Commons Attribution 4.0 License.

5.3. Spectroscopy

Of the spectroscopic biosensing techniques, surface-enhanced Raman spectroscopy (SERS) has shown great success [151,152]. SERS is a vibrational surface sensing technique that enhances Raman scattering based on surface characteristics. Briefly, SERS utilizes incident laser light to induce inelastic scattering (Raman scattering) from the target analyte. The intensity of the Raman scattering is enhanced by interaction with the conduction electrons of metal nanostructures (SERS substrate). The enhancement of the Raman scattering is what makes SERS so sensitive. Researchers have reported enhancement factors of up to ten or eleven orders of magnitude [153]. Figure 11 illustrates a SERS sensor for the analysis of breath volatile organic compound (VOC) biomarkers [154]. Due to the complex nature of the obtained spectral signal, various machine learning algorithms have been used to process SERS data in multiple contexts [28].

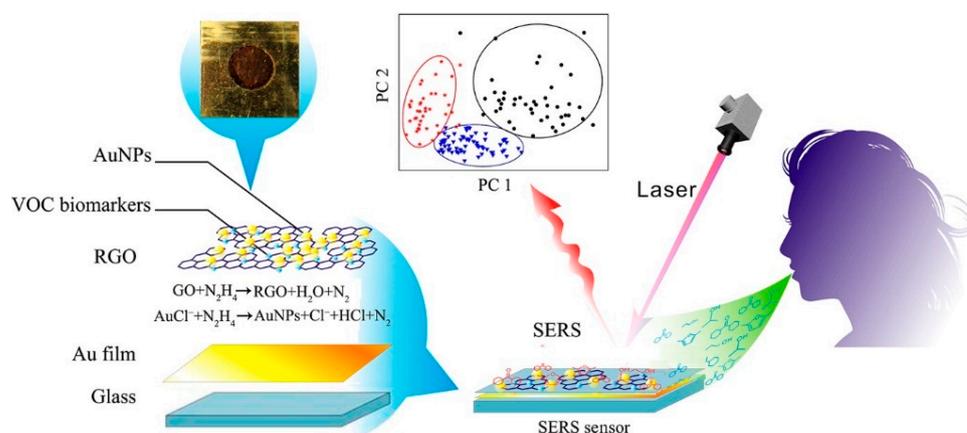


Figure 11. SERS sensor for analysis of breath VOC biomarkers utilizing AuNPs. Reprinted with permission from [154] without modification. Copyright 2016 American Chemical Society.

Although bioreceptors may be used to allow for specific binding of the target analyte to the SERS sensing surface [155,156], direct detection is also possible. Robust classification

and regression algorithms can bring specificity and sensitivity to these biosensors. A simple yet effective method for SERS based quantification is partial least squares regression (PLSR). PLSR has been used for a variety of quantification applications such as biofilm formation monitoring [69], blood serum methotrexate concentration [63], aquaculture toxins [62], and food antiseptics [66]. PLSR has the advantage of model simplicity with well-defined parameters, but it may be insufficient in modeling data with significant sources of noise.

Since the spectra have high dimensionality, dimension reduction is a frequent preprocessing step (Figure 12). PCA is again popularly used as a dimension reduction or feature extraction step [60,61,64,65,68,70,71,73], or for exploratory analysis [62,72,157]. Once the spectra are remapped using PCA, a classifier or regression model is employed such as an extreme learning machine (ELM) [71], LDA [68], SVM [60,64,73], PLSR [65], or ANN [70]. An alternative to dimension reduction is utilizing the high dimensionality spectral data directly with a node-based algorithm such as ANN [72,158,159] and CNN [160,161].

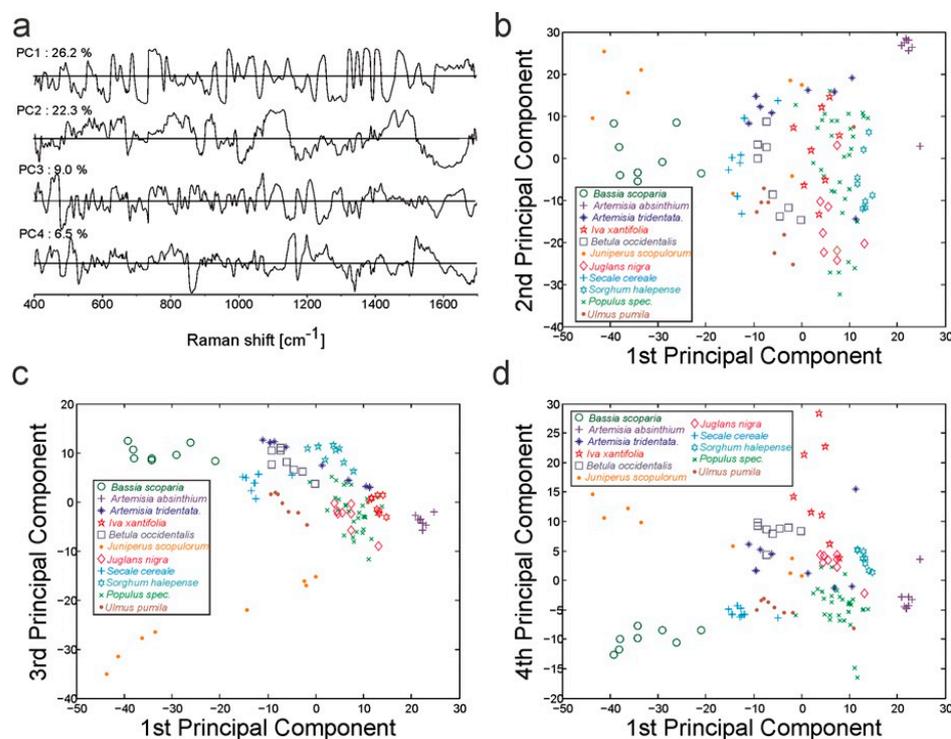


Figure 12. PCA results using the spectral range of 400–1700 cm^{-1} of 112 average SERS spectra from 14 different commercially available pollen species. Loadings of the first four PCs (a) as well as the scores of the first and second (b), first and third (c), and first and fourth PC (d) are shown. PCA was done with standardized first derivatives of the mean spectra of 500 vector-normalized spectra. Reprinted with permission from [72] without modification. Copyright 2016 John Wiley and Sons.

The reusability and generalizability of the trained models are often limited. Spectral response is affected not just by analyte presence but surface structure. Therefore, for the model to be reused on a new SERS biosensing dataset, the surface characteristics must be very similar. In terms of transfer learning, this is an issue of changes in the underlying data distributions. However, if the surface structure methods are well documented and reproducible, transfer learning could be employed on a spectral library [28]. Ideally, researchers could contribute to this library in an open-access manner and use these spectra for model training. In this case, the quality of the attached metadata would be a crucial factor.

Clearly, machine learning has been used extensively in the context of SERS sensors. The most common pipeline is to perform unsupervised dimensionality reduction/feature extraction for which PCA is generally the preferred method. Less consistency is seen in the algorithms used for classification and regression. Alternatively, ANNs can be used directly on the data, and the advantage of one approach over the other is not clearly illustrated

in the literature. We anticipate, however, that like in the case of electrochemical sensors, node-based models would allow for more efficient transfer learning to accommodate target task change.

5.4. Summary of Optical Bioreceptor-Free Biosensing

A variety of optical sensing methods have benefited from machine learning techniques, with the preferred method being dependent on the data type. For image type data, CNN is the most obvious choice for its ability to detect features as well as reconstruct images obtained by lensless systems. For spectral data, the approach is similar to spectral data obtained with electrochemical sensors. In those instances, dimensionality reduction coupled with a classification/regression algorithm may perform nearly as well as node-based methods. Indeed, they may be preferable in instances where the quantity of training data is small.

6. Considerations and Future Perspectives

Biosensor research has shown great success and promise. For both systems with and without bioreceptor, ML has demonstrated huge success in going from large, complex sensor datasets to getting meaningful measurements and classification of analytes. However, in many of these systems, a key challenge is consistency in device manufacturing. This manifests itself regarding sensor reproducibility for Enose and Etongue, or as substrate reproducibility for SERS. Since the models used to process these data often rely on subtle signals in the data, even small changes in sensor response characteristics can lead to poor performance. These issues have effectively limited widespread commercial adoption of these technologies. There has been some success in accommodating these inconsistencies through computational methods, notably with transfer learning for Enose. More work, both from a manufacturing and computational standpoint, needs to be done before many of these systems are robust enough for widespread adoption.

One area in which these systems have pushed to increase commercial potential is through miniaturization and modularity. There have been efforts with several of the methods presented here to develop compact standalone devices that rival their bulkier counterparts in terms of performance [16,47,162–166]. We believe that cloud computing may be a key element to the success of these endeavors. Some of the models in use, especially for image-based sensors, are computationally expensive. By offloading the computational work to cloud computing, the device footprint imposed by processing and memory needs is greatly reduced.

A central question is what the relative advantages and disadvantages are between systems that utilize a bioreceptor and those that do not. A key advantage of those that eliminate the bioreceptor addresses one of the barriers to commercialization—manufacture variability. By eliminating the bioreceptor, device manufacture is simplified, and may decrease manufacture variability. Additionally, sensor longevity is generally improved because the long-term stability of the bioreceptor is often limited [6]. However, to match LOD and specificity of bioreceptors, improvements must be made. Nanomaterials show promise for improving device performance [167].

There have been studies that attempt to gain the advantages of both systems by creating artificial bioreceptors, notably nanomaterials with enzymatic properties referred to as nanozymes [168,169]. While exciting progress has been made in this field, current nanozyme-based biosensors have inferior catalytic activity and specificity to their biological alternatives [170,171]. Nanozyme catalytic activity is also currently limited to oxidase-like activity [171]. If researchers can broaden nanozyme activity and improve selectivity, these biosensors may become a competitive alternative for biological bioreceptors.

In addition to device considerations, there are computational challenges to consider. Although some ML algorithms have been in use for decades such as PCA and SVM, the field of ML is advancing rapidly with new algorithms being described frequently. While many areas are quick to adopt the new methods, improper usage is common and certainly

not limited to biosensing. Some common mistakes are inappropriate data splitting, hidden variables serving as bad predictors, and mistaking the objective of the model [172]. Great emphasis must be placed on the importance of reporting appropriate performance metrics. A great example of a misleading metric is reporting accuracy on highly imbalanced data such as in Durante et al. [41]. It can often be difficult to determine if the proper pre-processing and model assumption checks are being performed. This may be centering and re-scaling prior to PCA, or normality checks for LDA.

Some of these issues can be solved with better methods reporting, especially regarding computational methods. Certain key details are frequently left out, making critical evaluation difficult and reproducibility impossible [173]. One of the most striking examples from the literature described herein is reporting classification metrics, without reporting what classifier was used on PCA processed data [49]. Perhaps the best way to make methods clear and reproducible is to release all associated code, preferably publicly.

Increased availability in general can greatly improve this field. More open access repositories of training sets may allow researchers to improve model robustness by exposing them to more diverse datasets [16]. Some examples currently exist such as the gas sensor drift dataset [115] and the EIS breast tissue dataset [103], both available in the UCI Machine Learning repository [104]. One vision would be to have large repositories of gas sensor responses to many analytes under various experimental conditions. Models could be trained on such repositories to improve generalizability. Ideally, with such repositories and improved manufacturing consistency, trained models could be shared directly and need only minimal recalibration.

7. Conclusions

In this review we have explored the ways in which bioreceptor-free biosensors can benefit from ML methods. Robust ML models bring specificity and accuracy to array-based biosensors such as Enose and Etongue by learning the patterns in the sensor responses. Notably, PCA has shown great performance as a feature extraction technique for these systems. Similar power of PCA has been demonstrated for optical biosensors that generate spectra such as Raman spectra or SERS. ANNs using deep learning generate impressive results for imaging-based sensors including lensless holography and digital staining. ML has also been used in creative ways such as for data fusion of multiple biosensors, and transfer learning for noise correction, sensor drift compensation, and domain adaptation.

However, many practical challenges still exist. Many of the methods presented here are not widely used in commercial settings. This is due to many reasons including variability in manufacturing and the ability to make compact versions of the biosensors while maintaining performance. ML models that can adapt to differences in sensor response are at an advantage, and transfer learning shows promise to be part of the solution.

In recent years, ML has garnered strong research interest in many fields including biosensing, as evidenced in this review. If this review has inspired interest to learn more about how machine learning is being used for one of the methods presented here, we encourage you to seek more specific reviews for the subject. There are great reviews in the literature, many of which were referenced, that take a closer look at the methods presented in this review.

Author Contributions: Conceptualization, K.E.S.III and J.-Y.Y.; Methodology, K.E.S.III; Formal analysis, K.E.S.III; Investigation, K.E.S.III; Data curation, K.E.S.III; Writing—original draft preparation, K.E.S.III; Writing—review and editing, J.-Y.Y.; Supervision, J.-Y.Y.; Project administration, J.-Y.Y.; Funding acquisition, K.E.S.III and J.-Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the U.S. National Institutes of Health under the grant T32GM132008.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study did not report any data.

Acknowledgments: The authors would like to thank Lane E. Breshears for her contribution to the collection of papers and for discussions on how to organize this manuscript. The authors would also like to thank Kattika Kaarj (now at Mahidol University) for helpful discussions and proofreading this manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Metkar, S.K.; Girigoswami, K. Diagnostic biosensors in medicine—A review. *Biocatal. Agric. Biotechnol.* **2019**, *17*, 271–283. [[CrossRef](#)]
2. Justino, C.I.L.; Duarte, A.C.; Rocha-Santos, T.A.P. Recent progress in biosensors for environmental monitoring: A review. *Sensors* **2017**, *17*, 2918. [[CrossRef](#)] [[PubMed](#)]
3. Nguyen, H.H.; Lee, S.H.; Lee, U.J.; Fermin, C.D.; Kim, M. Immobilized enzymes in biosensor applications. *Materials* **2019**, *12*, 121. [[CrossRef](#)] [[PubMed](#)]
4. Hock, B. Antibodies for immunosensors a review. *Anal. Chim. Acta* **1997**, *347*, 177–186. [[CrossRef](#)]
5. Lim, Y.C.; Kouzani, A.Z.; Duan, W. Aptasensors: A review. *J. Biomed. Nanotechnol.* **2010**, *6*, 93–105. [[CrossRef](#)] [[PubMed](#)]
6. Massah, J.; Asefpour Vakilian, K. An intelligent portable biosensor for fast and accurate nitrate determination using cyclic voltammetry. *Biosyst. Eng.* **2019**, *177*, 49–58. [[CrossRef](#)]
7. Esfahani, S.; Shanta, M.; Specht, J.P.; Xing, Y.; Cole, M.; Gardner, J.W. Wearable IoT electronic nose for urinary incontinence detection. In Proceedings of the 2020 IEEE Sensors, Virtual Conference, Rotterdam, The Netherlands, 25–28 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–4. [[CrossRef](#)]
8. Pelosi, P.; Zhu, J.; Knoll, W. From gas sensors to biomimetic artificial noses. *Chemosensors* **2018**, *6*, 32. [[CrossRef](#)]
9. Wilson, A.D. Noninvasive early disease diagnosis by electronic-nose and related VOC-detection devices. *Biosensors* **2020**, *10*, 73. [[CrossRef](#)]
10. Szulczyński, B.; Armiński, K.; Namieśnik, J.; Gębicki, J. Determination of odour interactions in gaseous mixtures using electronic nose methods with artificial neural networks. *Sensors* **2018**, *18*, 519. [[CrossRef](#)]
11. Zambotti, G.; Soprani, M.; Gobbi, E.; Capuano, R.; Pasqualetti, V.; Di Natale, C.; Ponzoni, A. Early detection of fish degradation by electronic nose. In Proceedings of the 2019 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), Fukuoka, Japan, 26–29 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–3. [[CrossRef](#)]
12. Podrażka, M.; Bączynska, E.; Kundys, M.; Jeleń, P.S.; Witkowska Nery, E. Electronic tongue—A tool for all tastes? *Biosensors* **2018**, *8*, 3. [[CrossRef](#)]
13. Chen, X.; Xu, Y.; Meng, L.; Chen, X.; Yuan, L.; Cai, Q.; Shi, W.; Huang, G. Non-parametric partial least squares—Discriminant analysis model based on sum of ranking difference algorithm for tea grade identification using electronic tongue data. *Sens. Actuat. B Chem.* **2020**, *311*, 127924. [[CrossRef](#)]
14. Kovacs, Z.; Szöllösi, D.; Zaukuu, J.-L.Z.; Bodor, Z.; Vitális, F.; Aouadi, B.; Zsom-Muha, V.; Gillay, Z. Factors influencing the long-term stability of electronic tongue and application of improved drift correction methods. *Biosensors* **2020**, *10*, 74. [[CrossRef](#)] [[PubMed](#)]
15. Guedes, M.D.V.; Marques, M.S.; Guedes, P.C.; Contri, R.V.; Kulkamp Guerreiro, I.C. The use of electronic tongue and sensory panel on taste evaluation of pediatric medicines: A systematic review. *Pharm. Dev. Technol.* **2020**, *26*, 119–137. [[CrossRef](#)] [[PubMed](#)]
16. Ross, C.F. Considerations of the use of the electronic tongue in sensory science. *Curr. Opin. Food Sci.* **2021**, *40*, 87–93. [[CrossRef](#)]
17. Guerrini, L.; Alvarez-Puebla, R.A. Chapter 19—Surface-enhanced Raman scattering chemosensing of proteins. In *Vibrational Spectroscopy in Protein Research*; Ozaki, Y., Baranska, M., Lednev, I.K., Wood, B.R., Eds.; Academic Press: London, UK, 2020; pp. 553–567. [[CrossRef](#)]
18. Feng, J.; Hu, Y.; Grant, E.; Lu, X. Determination of thiabendazole in orange juice using an MISPE-SERS chemosensor. *Food Chem.* **2018**, *239*, 816–822. [[CrossRef](#)]
19. Langer, J.; Jimenez de Aberasturi, D.; Aizpurua, J.; Alvarez-Puebla, R.A.; Auguie, B.; Baumberg, J.J.; Bazan, G.C.; Bell, S.E.J.; Boisen, A.; Brolo, A.G.; et al. Present and future of surface-enhanced Raman scattering. *ACS Nano* **2020**, *14*, 28–117. [[CrossRef](#)]
20. Zheng, X.S.; Jahn, I.J.; Weber, K.; Cialla-May, D.; Popp, J. Label-free SERS in biological and biomedical applications: Recent progress, current challenges and opportunities. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2018**, *197*, 56–77. [[CrossRef](#)] [[PubMed](#)]
21. Krafft, C.; Osei, E.B.; Popp, J.; Nazarenko, I. Raman and SERS spectroscopy for characterization of extracellular vesicles from control and prostate carcinoma patients. *Proc. SPIE* **2020**, *11236*, 1123602. [[CrossRef](#)]
22. Sang, S.; Wang, Y.; Feng, Q.; Wei, Y.; Ji, J.; Zhang, W. Progress of new label-free techniques for biosensors: A review. *Crit. Rev. Biotechnol.* **2016**, *36*, 465–481. [[CrossRef](#)]
23. Scott, S.M.; James, D.; Ali, Z. Data analysis for electronic nose systems. *Microchim. Acta* **2006**, *156*, 183–207. [[CrossRef](#)]

24. Yan, J.; Guo, X.; Duan, S.; Jia, P.; Wang, L.; Peng, C.; Zhang, S. Electronic nose feature extraction methods: A review. *Sensors* **2015**, *15*, 27804–27831. [[CrossRef](#)] [[PubMed](#)]
25. Hotel, O.; Poli, J.-P.; Mer-Calfati, C.; Scorsone, E.; Saada, S. A review of algorithms for SAW sensors Enose based volatile compound identification. *Sens. Actuat. B Chem.* **2018**, *255*, 2472–2482. [[CrossRef](#)]
26. Da Costa, N.L.; da Costa, M.S.; Barbosa, R. A review on the application of chemometrics and machine learning algorithms to evaluate beer authentication. *Food Anal. Meth.* **2021**, *14*, 136–155. [[CrossRef](#)]
27. Wasilewski, T.; Migoń, D.; Gebicki, J.; Kamysz, W. Critical review of electronic nose and tongue instruments prospects in pharmaceutical analysis. *Anal. Chim. Acta* **2019**, *1077*, 14–29. [[CrossRef](#)]
28. Lussier, F.; Thibault, V.; Charron, B.; Wallace, G.Q.; Masson, J.-F. Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering. *TrAC-Trends Anal. Chem.* **2020**, *124*, 115796. [[CrossRef](#)]
29. Cui, F.; Yue, Y.; Zhang, Y.; Zhang, Z.; Zhou, H.S. Advancing biosensors with machine learning. *ACS Sens.* **2020**, *5*, 3346–3364. [[CrossRef](#)] [[PubMed](#)]
30. Sheng, Y.; Qian, W.; Huang, J.; Wu, B.; Yang, J.; Xue, T.; Ge, Y.; Wen, Y. Electrochemical detection combined with machine learning for intelligent sensing of maleic hydrazide by using carboxylated PEDOT modified with copper nanoparticles. *Microchim. Acta* **2019**, *186*, 543. [[CrossRef](#)] [[PubMed](#)]
31. Dean, S.N.; Shriver-Lake, L.C.; Stenger, D.A.; Erickson, J.S.; Golden, J.P.; Trammell, S.A. Machine learning techniques for chemical identification using cyclic square wave voltammetry. *Sensors* **2019**, *19*, 2392. [[CrossRef](#)] [[PubMed](#)]
32. Tonezzer, M.; Le, D.T.T.; Iannotta, S.; Van Hieu, N. Selective discrimination of hazardous gases using one single metal oxide resistive sensor. *Sens. Actuat. B Chem.* **2018**, *277*, 121–128. [[CrossRef](#)]
33. Xu, L.; He, J.; Duan, S.; Wu, X.; Wang, Q. Comparison of machine learning algorithms for concentration detection and prediction of formaldehyde based on electronic nose. *Sens. Rev.* **2016**, *36*, 207–216. [[CrossRef](#)]
34. Yang, Y.; Liu, H.; Gu, Y. A model transfer learning framework with back-propagation neural network for wine and Chinese liquor detection by electronic nose. *IEEE Access* **2020**, *8*, 105278–105285. [[CrossRef](#)]
35. Leon-Medina, J.X.; Pineda-Muñoz, W.A.; Burgos, D.A.T. Joint distribution adaptation for drift correction in electronic nose type sensor arrays. *IEEE Access* **2020**, *8*, 134413–134421. [[CrossRef](#)]
36. Liu, B.; Zeng, X.; Tian, F.; Zhang, S.; Zhao, L. Domain transfer broad learning system for long-term drift compensation in electronic nose systems. *IEEE Access* **2019**, *7*, 143947–143959. [[CrossRef](#)]
37. Wang, X.; Gu, Y.; Liu, H. A transfer learning method for the protection of geographical indication in China using an electronic nose for the identification of Xihu Longjing tea. *IEEE Sens. J.* **2021**, *21*, 8065–8077. [[CrossRef](#)]
38. Yi, R.; Yan, J.; Shi, D.; Tian, Y.; Chen, F.; Wang, Z.; Duan, S. Improving the performance of drifted/shifted electronic nose systems by cross-domain transfer using common transfer samples. *Sens. Actuat. B Chem.* **2021**, *329*, 129162. [[CrossRef](#)]
39. Liang, Z.; Tian, F.; Zhang, C.; Sun, H.; Song, A.; Liu, T. Improving the robustness of prediction model by transfer learning for interference suppression of electronic nose. *IEEE Sens. J.* **2018**, *18*, 1111–1121. [[CrossRef](#)]
40. Daliri, M.R. Combining extreme learning machines using support vector machines for breast tissue classification. *Comput. Meth. Biomech. Biomed. Eng.* **2015**, *18*, 185–191. [[CrossRef](#)]
41. Durante, G.; Becari, W.; Lima, F.A.S.; Peres, H.E.M. Electrical impedance sensor for real-time detection of bovine milk adulteration. *IEEE Sens. J.* **2016**, *16*, 861–865. [[CrossRef](#)]
42. Helwan, A.; Idoko, J.B.; Abiyev, R.H. Machine learning techniques for classification of breast tissue. *Proc. Comput. Sci.* **2017**, *120*, 402–410. [[CrossRef](#)]
43. Islam, M.; Wahid, K.; Dinh, A. Assessment of ripening degree of avocado by electrical impedance spectroscopy and support vector machine. *J. Food Qual.* **2018**, *2018*, 4706147. [[CrossRef](#)]
44. Murphy, E.K.; Mahara, A.; Khan, S.; Hyams, E.S.; Schned, A.R.; Pettus, J.; Halter, R.J. Comparative study of separation between ex vivo prostatic malignant and benign tissue using electrical impedance spectroscopy and electrical impedance tomography. *Physiol. Meas.* **2017**, *38*, 1242–1261. [[CrossRef](#)] [[PubMed](#)]
45. Yang, Z.; Miao, N.; Zhang, X.; Li, Q.; Wang, Z.; Li, C.; Sun, X.; Lan, Y. Employment of an electronic tongue combined with deep learning and transfer learning for discriminating the storage time of Pu-erh tea. *Food Control* **2021**, *121*, 107608. [[CrossRef](#)]
46. Leon-Medina, J.X.; Anaya, M.; Pozo, F.; Tibaduiza, D. Nonlinear feature extraction through manifold learning in an electronic tongue classification task. *Sensors* **2020**, *20*, 4834. [[CrossRef](#)]
47. Giménez-Gómez, P.; Escudé-Pujol, R.; Capdevila, F.; Puig-Pujol, A.; Jiménez-Jorquera, C.; Gutiérrez-Capitán, M. Portable electronic tongue based on microsensors for the analysis of Cava wines. *Sensors* **2016**, *16*, 1796. [[CrossRef](#)]
48. Ouyang, Q.; Yang, Y.; Wu, J.; Liu, Z.; Chen, X.; Dong, C.; Chen, Q.; Zhang, Z.; Guo, Z. Rapid sensing of total theaflavins content in black tea using a portable electronic tongue system coupled to efficient variables selection algorithms. *J. Food Compos. Anal.* **2019**, *75*, 43–48. [[CrossRef](#)]
49. Li, Z.; Paul, R.; Ba Tis, T.; Saville, A.C.; Hansel, J.C.; Yu, T.; Ristaino, J.B.; Wei, Q. Non-invasive plant disease diagnostics enabled by smartphone-based fingerprinting of leaf volatiles. *Nat. Plants* **2019**, *5*, 856–866. [[CrossRef](#)]
50. Tomczak, A.; Ilic, S.; Marquardt, G.; Engel, T.; Forster, F.; Navab, N.; Albarqouni, S. Multi-task multi-domain learning for digital staining and classification of leukocytes. *IEEE Trans. Med. Imaging* **2020**. [[CrossRef](#)]
51. Sagar, M.A.K.; Cheng, K.P.; Ouellette, J.N.; Williams, J.C.; Watters, J.J.; Eliceiri, K.W. Machine learning methods for fluorescence lifetime imaging (FLIM) based label-free detection of microglia. *Front. Neurosci.* **2020**, *14*, 931. [[CrossRef](#)]

52. Lotfollahi, M.; Berisha, S.; Daeinejad, D.; Mayerich, D. Digital staining of high-definition Fourier transform infrared (FT-IR) images using deep learning. *Appl. Spectrosc.* **2019**, *73*, 556–564. [[CrossRef](#)]
53. Rivenson, Y.; Zhang, Y.; Günaydin, H.; Teng, D.; Ozcan, A. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light Sci. Appl.* **2018**, *7*, 17141. [[CrossRef](#)]
54. Rivenson, Y.; Wu, Y.; Ozcan, A. Deep learning in holography and coherent imaging. *Light Sci. Appl.* **2019**, *8*, 85. [[CrossRef](#)]
55. Wu, Y.; Ray, A.; Wei, Q.; Feizi, A.; Tong, X.; Chen, E.; Luo, Y.; Ozcan, A. Deep learning enables high-throughput analysis of particle-aggregation-based biosensors imaged using holography. *ACS Photon.* **2019**, *6*, 294–301. [[CrossRef](#)]
56. Wu, Y.; Calis, A.; Luo, Y.; Chen, C.; Lutton, M.; Rivenson, Y.; Lin, X.; Koydemir, H.C.; Zhang, Y.; Wang, H.; et al. Label-free bioaerosol sensing using mobile microscopy and deep learning. *ACS Photon.* **2018**, *5*, 4617–4627. [[CrossRef](#)]
57. Borhani, N.; Bower, A.J.; Boppart, S.A.; Psaltis, D. Digital staining through the application of deep neural networks to multi-modal multi-photon microscopy. *Biomed. Opt. Express* **2019**, *10*, 1339–1350. [[CrossRef](#)]
58. Dunker, S.; Motivans, E.; Rakosy, D.; Boho, D.; Mäder, P.; Hornick, T.; Knight, T.M. Pollen analysis using multispectral imaging flow cytometry and deep learning. *New Phytol.* **2021**, *229*, 593–606. [[CrossRef](#)]
59. Rivenson, Y.; Liu, T.; Wei, Z.; Zhang, Y.; de Haan, K.; Ozcan, A. PhaseStain: The digital staining of label-free quantitative phase microscopy images using deep learning. *Light Sci. Appl.* **2019**, *8*, 23. [[CrossRef](#)] [[PubMed](#)]
60. Zheng, X.; Lv, G.; Du, G.; Zhai, Z.; Mo, J.; Lv, X. Rapid and low-cost detection of thyroid dysfunction using Raman spectroscopy and an improved support vector machine. *IEEE Photon. J.* **2018**, *10*, 1–12. [[CrossRef](#)]
61. Tan, A.; Zhao, Y.; Sivashanmugan, K.; Squire, K.; Wang, A.X. Quantitative TLC-SERS detection of histamine in seafood with support vector machine analysis. *Food Control.* **2019**, *103*, 111–118. [[CrossRef](#)]
62. Chen, X.; Nguyen, T.H.D.; Gu, L.; Lin, M. Use of standing gold nanorods for detection of malachite green and crystal violet in fish by SERS. *J. Food Sci.* **2017**, *82*, 1640–1646. [[CrossRef](#)] [[PubMed](#)]
63. Fornasaro, S.; Marta, S.D.; Rabusin, M.; Bonifacio, A.; Sergo, V. Toward SERS-based point-of-care approaches for therapeutic drug monitoring: The case of methotrexate. *Faraday Discuss.* **2016**, *187*, 485–499. [[CrossRef](#)] [[PubMed](#)]
64. Hassoun, M.; Rüger, J.; Kirchberger-Tolstik, T.; Schie, I.W.; Henkel, T.; Weber, K.; Cialla-May, D.; Krafft, C.; Popp, J. A droplet-based microfluidic chip as a platform for leukemia cell lysate identification using surface-enhanced Raman scattering. *Anal. Bioanal. Chem.* **2018**, *410*, 999–1006. [[CrossRef](#)]
65. Hidi, I.J.; Jahn, M.; Pletz, M.W.; Weber, K.; Cialla-May, D.; Popp, J. Toward levofloxacin monitoring in human urine samples by employing the LoC-SERS technique. *J. Phys. Chem. C* **2016**, *120*, 20613–20623. [[CrossRef](#)]
66. Hou, M.; Huang, Y.; Ma, L.; Zhang, Z. Quantitative analysis of single and mix food antiseptics basing on SERS spectra with PLSR method. *Nanoscale Res. Lett.* **2016**, *11*, 296. [[CrossRef](#)]
67. Kämmer, E.; Olschewski, K.; Stöckel, S.; Rösch, P.; Weber, K.; Cialla-May, D.; Bocklitz, T.; Popp, J. Quantitative SERS studies by combining LOC-SERS with the standard addition method. *Anal. Bioanal. Chem.* **2015**, *407*, 8925–8929. [[CrossRef](#)]
68. Mühlig, A.; Bocklitz, T.; Labugger, I.; Dees, S.; Henk, S.; Richter, E.; Andres, S.; Merker, M.; Stöckel, S.; Weber, K.; et al. LOC-SERS: A promising closed system for the identification of mycobacteria. *Anal. Chem.* **2016**, *88*, 7998–8004. [[CrossRef](#)]
69. Nguyen, C.Q.; Thrift, W.J.; Bhattacharjee, A.; Ranjbar, S.; Gallagher, T.; Darvishzadeh-Varcheie, M.; Sanderson, R.N.; Capolino, F.; Whiteson, K.; Baldi, P.; et al. Longitudinal monitoring of biofilm formation via robust surface-enhanced Raman scattering quantification of *Pseudomonas aeruginosa*-produced metabolites. *ACS Appl. Mater. Interfaces* **2018**, *10*, 12364–12373. [[CrossRef](#)] [[PubMed](#)]
70. Othman, N.H.; Lee, K.Y.; Radzol, A.R.M.; Mansor, W. PCA-SCG-ANN for detection of non-structural protein 1 from SERS salivary spectra. In *Intelligent Information and Database Systems*; Nguyen, N.T., Tojo, S., Nguyen, L.M., Trawiński, B., Eds.; Springer: Cham, Switzerland, 2017; pp. 424–433. [[CrossRef](#)]
71. Othman, N.H.; Yoot Lee, K.; Mohd Radzol, A.R.; Mansor, W.; Amanina Yusoff, N. PCA-polynomial-ELM model optimal for detection of NS1 adulterated salivary SERS spectra. *J. Phys. Conf. Ser.* **2019**, *1372*, 012064. [[CrossRef](#)]
72. Seifert, S.; Merk, V.; Kneipp, J. Identification of aqueous pollen extracts using surface enhanced Raman scattering (SERS) and pattern recognition methods. *J. Biophotonics* **2016**, *9*, 181–189. [[CrossRef](#)]
73. Sun, H.; Lv, G.; Mo, J.; Lv, X.; Du, G.; Liu, Y. Application of KPCA combined with SVM in Raman spectral discrimination. *Optik* **2019**, *184*, 214–219. [[CrossRef](#)]
74. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52. [[CrossRef](#)]
75. Cunningham, P. Dimension reduction. In *Machine Learning Techniques for Multimedia: Case Studies on Organization and Retrieval, Cognitive Technologies*; Cord, M., Cunningham, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 91–112. [[CrossRef](#)]
76. Alpaydin, E. *Introduction to Machine Learning*; MIT Press: Cambridge, MA, USA, 2020.
77. Peterson, L.E. K-nearest neighbor. *Scholarpedia* **2009**, *4*, 1883. [[CrossRef](#)]
78. Balaprakash, P.; Salim, M.; Uram, T.D.; Vishwanath, V.; Wild, S.M. DeepHyper: Asynchronous hyperparameter search for deep neural networks. In Proceedings of the 2018 IEEE 25th International Conference on High Performance Computing (HiPC), Bengaluru, India, 17–20 December 2018; IEEE: Piscataway, NJ, USA, 2019; pp. 42–51. [[CrossRef](#)]
79. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, COLT'92, Pittsburgh, PA, USA, 27–29 July 1992; Association for Computing Machinery: New York, NY, USA, 1992; pp. 144–152. [[CrossRef](#)]
80. Zhang, J. A complete list of kernels used in support vector machines. *Biochem. Pharmacol.* **2015**, *4*, 2. [[CrossRef](#)]

81. Mathur, A.; Foody, G.M. Multiclass and binary SVM classification: Implications for training and classification users. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 241–245. [CrossRef]
82. Shawe-Taylor, J.; Cristianini, N. On the generalization of soft margin algorithms. *IEEE Trans. Inf. Theory* **2002**, *48*, 2721–2735. [CrossRef]
83. Han, H.; Jiang, X. Overcome support vector machine diagnosis overfitting. *Cancer Inform.* **2014**, *13*, 145–158. [CrossRef]
84. Ghojogh, B.; Crowley, M. Linear and quadratic discriminant analysis: Tutorial. *arXiv* **2019**, arXiv:1906.02590. Available online: <https://arxiv.org/abs/1906.02590v1> (accessed on 2 August 2021).
85. Myles, A.J.; Feudale, R.N.; Liu, Y.; Woody, N.A.; Brown, S.D. An introduction to decision tree modeling. *J. Chemom.* **2004**, *18*, 275–285. [CrossRef]
86. Lewis, R.J. An introduction to classification and regression tree (CART) analysis. In *The 2000 Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, California*; Society for Academic Emergency Medicine: Des Plaines, IL, USA, 2000; Available online: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.95.4103&rep=rep1&type=pdf> (accessed on 4 August 2021).
87. Criminisi, A.; Shotton, J.; Konukoglu, E. Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Found. Trends Comput. Graph. Vis.* **2012**, *7*, 81–227. [CrossRef]
88. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Phys. D Nonlinear Phenom.* **2020**, *404*, 132306. [CrossRef]
89. Huang, G.B.; Zhu, Q.Y.; Siew, C.-K. Extreme learning machine: Theory and applications. *Neurocomput. Neural Netw.* **2006**, *70*, 489–501. [CrossRef]
90. Fukushima, K. Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Netw.* **1998**, *1*, 119–130. [CrossRef]
91. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [CrossRef]
92. Hecht-Nielsen, R. Theory of the backpropagation neural network. In *Neural Networks for Perception, Volume 2: Computation, Learning, and Architectures*; Academic Press: Boston, MA, USA, 1992; pp. 65–93. [CrossRef]
93. Alto, V. Neural Networks: Parameters, Hyperparameters and Optimization Strategies. Available online: <https://towardsdatascience.com/neural-networks-parameters-hyperparameters-and-optimization-strategies-3f0842fac0a5> (accessed on 2 August 2021).
94. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *arXiv* **2014**, arXiv:1406.2661. Available online: <https://arxiv.org/abs/1406.2661> (accessed on 13 April 2021). [CrossRef]
95. Zhang, F.; O'Donnell, L.J. Chapter 7—Support vector regression. In *Machine Learning*; Mechelli, A., Vieira, S., Eds.; Academic Press: London, UK, 2020; pp. 123–140. [CrossRef]
96. Hoffmann, F.; Bertram, T.; Mikut, R.; Reischl, M.; Nelles, O. Benchmarking in classification and regression. *WIREs* **2019**, *9*, e1318. [CrossRef]
97. Shao, J. Linear model selection by cross-validation. *J. Am. Stat. Assoc.* **1993**, *88*, 486–494. [CrossRef]
98. Yoon, J.-Y. *Introduction to Biosensors: From Electric Circuits to Immunosensors*, 2nd ed.; Springer: New York, NY, USA, 2016. [CrossRef]
99. Grieshaber, D.; MacKenzie, R.; Vörös, J.; Reimhult, E. Electrochemical biosensors—Sensor principles and architectures. *Sensors* **2008**, *8*, 1400–1458. [CrossRef] [PubMed]
100. Dhanjai; Sinha, A.; Lu, X.; Wu, L.; Tan, D.; Li, Y.; Chen, J.; Jain, R. Voltammetric sensing of biomolecules at carbon based electrode interfaces: A review. *TrAC—Trends Anal. Chem.* **2018**, *98*, 174–189. [CrossRef]
101. Grossi, M.; Riccò, B. Electrical impedance spectroscopy (EIS) for biological analysis and food characterization: A review. *J. Sens. Sens. Syst.* **2017**, *6*, 303–325. [CrossRef]
102. Yao, J.; Wang, L.; Liu, K.; Wu, H.; Wang, H.; Huang, J.; Li, J. Evaluation of electrical characteristics of biological tissue with electrical impedance spectroscopy. *Electrophoresis* **2020**, *41*, 1425–1432. [CrossRef]
103. Jossinet, J. Variability of impedivity in normal and pathological breast tissue. *Med. Biol. Eng. Comput.* **1996**, *34*, 346–350. [CrossRef]
104. Dua, D.; Graff, C. *UCI Machine Learning Repository* (<https://archive.ics.uci.edu/>); University of California, Irvine, School of Information and Computer Sciences: Irvine, CA, USA, 2017.
105. Estrela da Silva, J.; Marques de Sá, J.P.; Jossinet, J. Classification of breast tissue by electrical impedance spectroscopy. *Med. Biol. Eng. Comput.* **2000**, *38*, 26–30. [CrossRef]
106. Wasilewski, T.; Kamysz, W.; Gębicki, J. Bioelectronic tongue: Current status and perspectives. *Biosens. Bioelectron.* **2020**, *150*, 111923. [CrossRef] [PubMed]
107. Röck, F.; Barsan, N.; Weimar, U. Electronic nose: Current status and future trends. *Chem. Rev.* **2008**, *108*, 705–725. [CrossRef] [PubMed]
108. Ha, D.; Sun, Q.; Su, K.; Wan, H.; Li, H.; Xu, N.; Sun, F.; Zhuang, L.; Hu, N.; Wang, P. Recent achievements in electronic tongue and bioelectronic tongue as taste sensors. *Sens. Actuat. B: Chem.* **2015**, *207*, 1136–1146. [CrossRef]
109. Chen, C.Y.; Lin, W.C.; Yang, H.Y. Diagnosis of ventilator-associated pneumonia using electronic nose sensor array signals: Solutions to improve the application of machine learning in respiratory research. *Respir. Res.* **2020**, *21*, 45. [CrossRef]

110. Tan, J.; Xu, J. Applications of electronic nose (Enose) and electronic tongue (Etongue) in food quality-related properties determination: A review. *Artif. Intell. Agric.* **2020**, *4*, 104–115. [[CrossRef](#)]
111. Sanaeifar, A.; ZakiDizaji, H.; Jafari, A.; de la Guardia, M. Early detection of contamination and defect in foodstuffs by electronic nose: A review. *TrAC—Trends Anal. Chem.* **2017**, *97*, 257–271. [[CrossRef](#)]
112. Zhang, L.; Tian, F.; Kadri, C.; Xiao, B.; Li, H.; Pan, L.; Zhou, H. On-line sensor calibration transfer among electronic nose instruments for monitoring volatile organic chemicals in indoor air quality. *Sens. Actuat. B Chem.* **2011**, *160*, 899–909. [[CrossRef](#)]
113. Ciosek, P.; Wróblewski, W. Sensor arrays for liquid sensing—Electronic tongue systems. *Analyst* **2007**, *132*, 963–978. [[CrossRef](#)]
114. Liu, H.; Li, Q.; Li, Z.; Gu, Y. A suppression method of concentration background noise by transductive transfer learning for a metal oxide semiconductor-based electronic nose. *Sensors* **2020**, *20*, 1913. [[CrossRef](#)]
115. Vergara, A.; Vembu, S.; Ayhan, T.; Ryan, M.A.; Homer, M.L.; Huerta, R. Chemical gas sensor drift compensation using classifier ensembles. *Sens. Actuat. B Chem.* **2012**, *166–167*, 320–329. [[CrossRef](#)]
116. Zhang, L.; Tian, F.; Peng, X.; Dang, L.; Li, G.; Liu, S.; Kadri, C. Standardization of metal oxide sensor array using artificial neural networks through experimental design. *Sens. Actuat. B Chem.* **2013**, *177*, 947–955. [[CrossRef](#)]
117. De Vere, R. Disorders of taste and smell. *Continuum* **2017**, *23*, 421. [[CrossRef](#)]
118. Biancolillo, A.; Boqué, R.; Cocchi, M.; Marini, F. Chapter 10—Data fusion strategies in food analysis. In *Data Handling in Science and Technology*; Cocchi, M., Ed.; Elsevier: Amsterdam, The Netherlands, 2019; Volume 31, pp. 271–310. [[CrossRef](#)]
119. Banerjee, M.B.; Roy, R.B.; Tudu, B.; Bandyopadhyay, R.; Bhattacharyya, N. Black tea classification employing feature fusion of E-nose and E-tongue responses. *J. Food Eng.* **2019**, *244*, 55–63. [[CrossRef](#)]
120. Men, H.; Shi, Y.; Fu, S.; Jiao, Y.; Qiao, Y.; Liu, J. Mining feature of data fusion in the classification of beer flavor information using E-tongue and E-nose. *Sensors* **2017**, *17*, 1656. [[CrossRef](#)] [[PubMed](#)]
121. Buratti, S.; Malegori, C.; Benedetti, S.; Oliveri, P.; Giovanelli, G. E-nose, E-tongue and E-eye for edible olive oil characterization and shelf life assessment: A powerful data fusion approach. *Talanta* **2018**, *182*, 131–141. [[CrossRef](#)]
122. Dai, C.; Huang, X.; Huang, D.; Lv, R.; Sun, J.; Zhang, Z.; Ma, M.; Aheto, J.H. Detection of submerged fermentation of *Tremella aurantialba* using data fusion of electronic nose and tongue. *J. Food Proc. Eng.* **2019**, *42*, e13002. [[CrossRef](#)]
123. Tian, X.; Wang, J.; Ma, Z.; Li, M.; Wei, Z. Combination of an E-nose and an E-tongue for adulteration detection of minced mutton mixed with pork. *J. Food Qual.* **2019**, *2019*, e4342509. [[CrossRef](#)]
124. Di Rosa, A.R.; Leone, F.; Cheli, F.; Chiofalo, V. Fusion of electronic nose, electronic tongue and computer vision for animal source food authentication and quality assessment—A review. *J. Food Eng.* **2017**, *210*, 62–75. [[CrossRef](#)]
125. Cave, J.W.; Wickiser, J.K.; Mitropoulos, A.N. Progress in the development of olfactory-based bioelectronic chemosensors. *Biosens. Bioelectron.* **2019**, *123*, 211–222. [[CrossRef](#)]
126. Ahn, S.R.; An, J.H.; Song, H.S.; Park, J.W.; Lee, S.H.; Kim, J.H.; Jang, J.; Park, T.H. Duplex bioelectronic tongue for sensing umami and sweet tastes based on human taste receptor nanovesicles. *ACS Nano* **2016**, *10*, 7287–7296. [[CrossRef](#)] [[PubMed](#)]
127. Edachana, R.P.; Kumaresan, A.; Balasubramanian, V.; Thiagarajan, R.; Nair, B.G.; Thekkedath Gopalakrishnan, S.B. Paper-based device for the colorimetric assay of bilirubin based on in-vivo formation of gold nanoparticles. *Microchim. Acta* **2019**, *187*, 60. [[CrossRef](#)]
128. Mutlu, A.Y.; Kılıç, V.; Kocakuşak Özdemir, G.; Bayram, A.; Horzum, N.; Solmaz, M.E. Smartphone-based colorimetric detection via machine learning. *Analyst* **2017**, *142*, 2434–2441. [[CrossRef](#)]
129. Solmaz, M.E.; Mutlu, A.Y.; Alankus, G.; Kılıç, V.; Bayram, A.; Horzum, N. Quantifying colorimetric tests using a smartphone app based on machine learning classifiers. *Sens. Actuat. B Chem.* **2018**, *255*, 1967–1973. [[CrossRef](#)]
130. Lin, B.; Yu, Y.; Cao, Y.; Guo, M.; Zhu, D.; Dai, J.; Zheng, M. Point-of-care testing for streptomycin based on aptamer recognizing and digital image colorimetry by smartphone. *Biosens. Bioelectron.* **2018**, *100*, 482–489. [[CrossRef](#)] [[PubMed](#)]
131. Song, E.; Yu, M.; Wang, Y.; Hu, W.; Cheng, D.; Swihart, M.T.; Song, Y. Multi-color quantum dot-based fluorescence immunoassay array for simultaneous visual detection of multiple antibiotic residues in milk. *Biosens. Bioelectron.* **2015**, *72*, 320–325. [[CrossRef](#)] [[PubMed](#)]
132. Wang, L.; Chen, W.; Ma, W.; Liu, L.; Ma, W.; Zhao, Y.; Zhu, Y.; Xu, L.; Kuang, H.; Xu, C. Fluorescent strip sensor for rapid determination of toxins. *Chem. Commun.* **2011**, *47*, 1574–1576. [[CrossRef](#)] [[PubMed](#)]
133. Chen, X.; Lan, J.; Liu, Y.; Li, L.; Yan, L.; Xia, Y.; Wu, F.; Li, C.; Li, S.; Chen, J. A paper-supported aptasensor based on upconversion luminescence resonance energy transfer for the accessible determination of exosomes. *Biosens. Bioelectron.* **2018**, *102*, 582–588. [[CrossRef](#)] [[PubMed](#)]
134. Guo, X. Surface plasmon resonance based biosensor technique: A review. *J. Biophoton.* **2012**, *5*, 483–501. [[CrossRef](#)] [[PubMed](#)]
135. Heinze, B.C.; Yoon, J.-Y. Nanoparticle immunoagglutination Rayleigh scatter assay to complement microparticle immunoagglutination Mie scatter assay in a microfluidic device. *Colloids Surf. B Biointerfaces* **2011**, *85*, 168–173. [[CrossRef](#)]
136. Park, T.S.; Li, W.; McCracken, K.E.; Yoon, J.-Y. Smartphone quantifies Salmonella from paper microfluidics. *Lab Chip* **2013**, *13*, 4832–4840. [[CrossRef](#)] [[PubMed](#)]
137. Elad, T.; Benovich, E.; Magrisso, S.; Belkin, S. Toxicant identification by a luminescent bacterial bioreporter panel: Application of pattern classification algorithms. *Environ. Sci. Technol.* **2008**, *42*, 8486–8491. [[CrossRef](#)]
138. Jouanneau, S.; Durand, M.-J.; Courcoux, P.; Blusseau, T.; Thouand, G. Improvement of the identification of four heavy metals in environmental samples by using predictive decision tree models coupled with a set of five bioluminescent bacteria. *Environ. Sci. Technol.* **2011**, *45*, 2925–2931. [[CrossRef](#)] [[PubMed](#)]

139. Gou, T.; Hu, J.; Zhou, S.; Wu, W.; Fang, W.; Sun, J.; Hu, Z.; Shen, H.; Mu, Y. A new method using machine learning for automated image analysis applied to chip-based digital assays. *Analyst* **2019**, *144*, 3274–3281. [[CrossRef](#)] [[PubMed](#)]
140. Ker, J.; Wang, L.; Rao, J.; Lim, T. Deep learning applications in medical image analysis. *IEEE Access* **2018**, *6*, 9375–9389. [[CrossRef](#)]
141. Uslu, F.; Icoz, K.; Tasdemir, K.; Yilmaz, B. Automated quantification of immunomagnetic beads and leukemia cells from optical microscope images. *Biomed. Signal Proc. Control.* **2019**, *49*, 473–482. [[CrossRef](#)]
142. Zeng, N.; Wang, Z.; Zhang, H.; Liu, W.; Alsaadi, F.E. Deep belief networks for quantitative analysis of a gold immunochromatographic strip. *Cogn. Comput.* **2016**, *8*, 684–692. [[CrossRef](#)]
143. Roy, M.; Seo, D.; Oh, S.; Yang, J.-W.; Seo, S. A review of recent progress in lens-free imaging and sensing. *Biosens. Bioelectron.* **2017**, *88*, 130–143. [[CrossRef](#)]
144. Wu, Y.; Ozcan, A. Lensless digital holographic microscopy and its applications in biomedicine and environmental monitoring. *Methods* **2018**, *136*, 4–16. [[CrossRef](#)]
145. Greenbaum, A.; Luo, W.; Su, T.W.; Göröcs, Z.; Xue, L.; Isikman, S.O.; Coskun, A.F.; Mudanyali, O.; Ozcan, A. Imaging without lenses: Achievements and remaining challenges of wide-field on-chip microscopy. *Nat. Methods* **2012**, *9*, 889–895. [[CrossRef](#)]
146. Slaoui, M.; Fiette, L. Histopathology procedures: From tissue sampling to histopathological evaluation. In *Drug Safety Evaluation: Methods and Protocols*; Gautier, J.-C., Ed.; Humana Press: Totowa, NJ, USA, 2011; pp. 69–82. [[CrossRef](#)]
147. Yi, X.; Walia, E.; Babyn, P. Generative adversarial network in medical imaging: A review. *Med. Image Anal.* **2019**, *58*, 101552. [[CrossRef](#)]
148. Rana, A.; Yauney, G.; Lowe, A.; Shah, P. Computational histological staining and destaining of prostate core biopsy RGB images with generative adversarial neural networks. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*; IEEE: Piscataway, NJ, USA, 2018; pp. 828–834. [[CrossRef](#)]
149. Affonso, C.; Rossi, A.L.D.; Vieira, F.H.A.; de Leon Ferreira de Carvalho, A.C.P. Deep learning for biological image classification. *Expert Syst. Appl.* **2017**, *85*, 114–122. [[CrossRef](#)]
150. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [[CrossRef](#)]
151. Yang, T.; Guo, X.; Wang, H.; Fu, S.; Wen, Y.; Yang, H. Magnetically optimized SERS assay for rapid detection of trace drug-related biomarkers in saliva and fingerprints. *Biosens. Bioelectron.* **2015**, *68*, 350–357. [[CrossRef](#)]
152. Zhang, D.; Huang, L.; Liu, B.; Ni, H.; Sun, L.; Su, E.; Chen, H.; Gu, Z.; Zhao, X. Quantitative and ultrasensitive detection of multiplex cardiac biomarkers in lateral flow assay with core-shell SERS nanotags. *Biosens. Bioelectron.* **2018**, *106*, 204–211. [[CrossRef](#)] [[PubMed](#)]
153. Schlücker, S. Surface-enhanced Raman spectroscopy: Concepts and chemical applications. *Angew. Chem. Int. Ed.* **2014**, *53*, 4756–4795. [[CrossRef](#)]
154. Chen, Y.; Zhang, Y.; Pan, F.; Liu, J.; Wang, K.; Zhang, C.; Cheng, S.; Lu, L.; Zhang, W.; Zhang, Z.; et al. Breath analysis based on surface-enhanced Raman scattering sensors distinguishes early and advanced gastric cancer patients from healthy persons. *ACS Nano* **2016**, *10*, 8169–8179. [[CrossRef](#)]
155. Banaei, N.; Moshfegh, J.; Mohseni-Kabir, A.; Houghton, J.M.; Sun, Y.; Kim, B. Machine learning algorithms enhance the specificity of cancer biomarker detection using SERS-based immunoassays in microfluidic chips. *RSC Adv.* **2019**, *9*, 1859–1868. [[CrossRef](#)]
156. Guselnikova, O.; Postnikov, P.; Pershina, A.; Svorcik, V.; Lyutakov, O. Express and portable label-free DNA detection and recognition with SERS platform based on functional Au grating. *Appl. Surf. Sci.* **2019**, *470*, 219–227. [[CrossRef](#)]
157. Goodacre, R.; Graham, D.; Faulds, K. Recent developments in quantitative SERS: Moving towards absolute quantification. *TrAC—Trends Anal. Chem.* **2018**, *102*, 359–368. [[CrossRef](#)]
158. Guselnikova, O.; Trelin, A.; Skvortsova, A.; Ulbrich, P.; Postnikov, P.; Pershina, A.; Sykora, D.; Svorcik, V.; Lyutakov, O. Label-free surface-enhanced Raman spectroscopy with artificial neural network technique for recognition photoinduced DNA damage. *Biosens. Bioelectron.* **2019**, *145*, 111718. [[CrossRef](#)] [[PubMed](#)]
159. Thrift, W.J.; Cabuslay, A.; Laird, A.B.; Ranjbar, S.; Hochbaum, A.I.; Ragan, R. Surface-enhanced Raman scattering-based odor compass: Locating multiple chemical sources and pathogens. *ACS Sens.* **2019**, *4*, 2311–2319. [[CrossRef](#)] [[PubMed](#)]
160. Thrift, W.J.; Nguyen, C.Q.; Wang, J.; Kahn, J.E.; Dong, R.; Laird, A.B.; Ragan, R. Improved regressions with convolutional neural networks for surface enhanced Raman scattering sensing of metabolite biomarkers. *Proc. SPIE* **2019**, *11089*, 1108907. [[CrossRef](#)]
161. Thrift, W.J.; Ragan, R. Quantification of analyte concentration in the single molecule regime using convolutional neural networks. *Anal. Chem.* **2019**, *91*, 13337–13342. [[CrossRef](#)]
162. Cheng, L.; Meng, Q.H.; Lilienthal, A.J.; Qi, P.F. Development of compact electronic noses: A review. *Meas. Sci. Technol.* **2021**, *32*, 062002. [[CrossRef](#)]
163. Jiang, L.; Hassan, M.M.; Ali, S.; Li, H.; Sheng, R.; Chen, Q. Evolving trends in SERS-based techniques for food quality and safety: A review. *Trends Food Sci. Technol.* **2021**, *112*, 225–240. [[CrossRef](#)]
164. Pang, S.; Yang, T.; He, L. Review of surface enhanced Raman spectroscopic (SERS) detection of synthetic chemical pesticides. *TrAC—Trends Anal. Chem.* **2016**, *85*, 73–82. [[CrossRef](#)]
165. Barreiros dos Santos, M.; Queirós, R.B.; Geraldés, Á.; Marques, C.; Vilas-Boas, V.; Dieguez, L.; Paz, E.; Ferreira, R.; Morais, J.; Vasconcelos, V.; et al. Portable sensing system based on electrochemical impedance spectroscopy for the simultaneous quantification of free and total microcystin-LR in freshwaters. *Biosens. Bioelectron.* **2019**, *142*, 111550. [[CrossRef](#)]

166. Huang, X.; Li, Y.; Xu, X.; Wang, R.; Yao, J.; Han, W.; Wei, M.; Chen, J.; Xuan, W.; Sun, L. High-precision lensless microscope on a chip based on in-line holographic imaging. *Sensors* **2021**, *21*, 720. [[CrossRef](#)]
167. Zhang, R.; Belwal, T.; Li, L.; Lin, X.; Xu, Y.; Luo, Z. Nanomaterial-based biosensors for sensing key foodborne pathogens: Advances from recent decades. *Compr. Rev. Food Sci. Food Saf.* **2020**, *19*, 1465–1487. [[CrossRef](#)]
168. Wang, Q.; Wei, H.; Zhang, Z.; Wang, E.; Dong, S. Nanozyme: An emerging alternative to natural enzyme for biosensing and immunoassay. *TrAC—Trends Anal. Chem.* **2018**, *105*, 218–224. [[CrossRef](#)]
169. Zhu, X.; Liu, P.; Ge, Y.; Wu, R.; Xue, T.; Sheng, Y.; Ai, S.; Tang, K.; Wen, Y. MoS₂/MWCNTs porous nanohybrid network with oxidase-like characteristic as electrochemical nanozyme sensor coupled with machine learning for intelligent analysis of carbendazim. *J. Electroanal. Chem.* **2020**, *862*, 113940. [[CrossRef](#)]
170. Mahmudunnabi, G.R.; Farhana, F.Z.; Kashaninejad, N.; Firoz, S.H.; Shim, Y.-B.; Shiddiky, M.J.A. Nanozyme-based electrochemical biosensors for disease biomarker detection. *Analyst* **2020**, *145*, 4398–4420. [[CrossRef](#)]
171. Zhang, X.; Wu, D.; Zhou, X.; Yu, Y.; Liu, J.; Hu, N.; Wang, H.; Li, G.; Wu, Y. Recent progress in the construction of nanozyme-based biosensors and their applications to food safety assay. *TrAC—Trends Anal. Chem.* **2019**, *121*, 115668. [[CrossRef](#)]
172. Riley, P. Three pitfalls to avoid in machine learning. *Nature* **2019**, *572*, 27–29. [[CrossRef](#)] [[PubMed](#)]
173. Stodden, V.; McNutt, M.; Bailey, D.H.; Deelman, E.; Gil, Y.; Hanson, B.; Heroux, M.A.; Ioannidis, J.P.A.; Tauber, M. Enhancing reproducibility for computational methods. *Science* **2016**, *354*, 1240–1241. [[CrossRef](#)] [[PubMed](#)]