RESEARCH ARTICLE

Spanish Journal of Agricultural Research 23 (2)
April-June, 2025, 21120
ISSN-L: 1695-971X, eISSN: 2171-9292
https://doi.org/10.5424/sjar/2025232-21120

Optimizing gas sensor array to classify three edible ripening stages of date fruit using machine learning and feature selection (case study: cv. Shahani)

Ehsan Savand-Roumi[1] and Mahdi Keramat-Jahromi[2*]

[1]*Department of Agricultural Machinery Engineering, Faculty of Agricultural Engineering and Technology, University College of Agriculture and Natural Resources, University of Tehran, 31587-77871 Karaj, Iran.*
[2]*Mechanical Engineering of Biosystems Department, Faculty of Agriculture, Shiraz University, 71441-13131 Shiraz, Iran.*

***Correspondence** should be addressed to Mahdi Keramat-Jahromi: mahdikeramat@shirazu.ac.ir

Abstract

Aim of study: This study aimed to analyze aroma changes in date fruits (*Phoenix dactylifera* L.) across three ripening stages (*Khalal*, *Rutab*, and *Tamr*) using an e-nose as a non-destructive monitoring method for fruit quality management. The study also evaluated feature selection methods to optimize the sensor array for a repeatable model based on external data.

Area of study: This study was carried out at a date palm garden cultivated by cv. Shahani in Jahrom, west-south of Iran.

Material and methods: An e-nose profiled the aroma of date fruits from a date palm garden across three ripening stages over two years. Classification models, developed using one year's data as internal data, were externally validated using data from the other year as external data. Feature selection optimized the sensor array, improving the prediction accuracy on the external dataset.

Main results: The classification models had similar test accuracy, but their predictive performance varied on external data. Using F-test and principal component analysis (PCA) for feature selection to optimize the sensor array, with support vector machine (SVM) as the classification algorithm, resulted in a highly repeatable model. This study suggests that e-noses are a promising tool for monitoring aroma changes during date fruit ripening in artificial ripening or storage processes.

Research highlights: Aroma profiles from edible ripening stages in data fruits were classified with high accuracy using e-nose; Feature selection methods can affect linear discriminant analysis (LDA) and SVM algorithms differently on external data; In SVM, some feature selection methods decreased prediction accuracy on external data compared to using the full sensor array; PCA effectively determined the optimal number of features for optimizing the sensor array.

Keywords: cv. Shahani; date fruit; E-nose; feature selection; ripening.

Optimización de una matriz de sensores de gas para clasificar tres etapas de maduración comestible de dátiles mediante aprendizaje automático y selección de características (estudio de caso: cv. Shahani)

Resumen

Objetivo del estudio: Este estudio tuvo como objetivo analizar los cambios en el aroma de los dátiles (Phoenix dactylifera L.) en tres etapas de maduración (*Khalal*, *Rutab* y *Tamr*) utilizando una "nariz electrónica" (*e-nose*) como método de monitoreo no destructivo para la gestión de la calidad de la fruta. El estudio también evaluó métodos de selección de características para optimizar la matriz de sensores en un modelo repetible basado en datos externos.

Área de estudio: Este estudio se llevó a cabo en un huerto de palmeras datileras cultivadas con el cultivar Shahani, en Jahrom, al suroeste de Irán.

Materiales y métodos: Un *e-nose* perfiló el aroma de los dátiles de un huerto de palmeras datileras en tres etapas de maduración durante dos años. Los modelos de clasificación, desarrollados utilizando los datos de un año como datos internos, fueron validados externamente utilizando los datos del otro año como datos externos. La selección de características optimizó la matriz de sensores, mejorando la precisión de la predicción en el conjunto de datos externos.

Resultados principales: Los modelos de clasificación tuvieron una precisión de prueba similar, pero su rendimiento predictivo varió en datos externos. El uso de la prueba F y el análisis de componentes principales (PCA) para la selección de características con el fin de optimizar la matriz de sensores, utilizando la máquina de soporte vectorial (SVM) como algoritmo de clasificación, resultó en un modelo altamente repetible. Este estudio sugiere que la *e-nose* es una herramienta prometedora para monitorear los cambios de aroma durante la maduración de los dátiles en procesos de maduración artificial o almacenamiento.

Aspectos destacados de la investigación: los perfiles de aroma de las etapas de maduración comestibles en los dátiles, fueron clasificados con alta precisión utilizando *e-nose*; los métodos de selección de características pueden afectar de manera diferente al análisis discriminante lineal (LDA) y a los algoritmos SVM en datos externos; en SVM, algunos métodos de selección de características disminuyeron la precisión de la predicción en datos externos en comparación con el uso de toda la matriz de sensores; PCA determinó de manera efectiva el número óptimo de características para optimizar la matriz de sensores.

Palabras clave: cv. Shahani; dátiles; maduración; nariz electrónica; selección de características.

Citation: Savand-Roumi, E; Keramat-Jahromi, M (2025). Optimizing gas sensor array to classify three edible ripening stages of date fruit using machine learning and feature selection (case study: cv. Shahani). Spanish Journal of Agricultural Research, Volume 23, Issue 2, 21120. https://doi.org/10.5424/sjar/2025232-21120

Received: 02/04/2024. **Accepted:** 18/01/2025. **Published:** 24/11/2025.

Introduction

Fruit, a key crop in worldwide agriculture, requires quality checks to ensure freshness, taste, and safety. Its stages of ripening are an important factor in determining fruit quality, which is usually manually classified by field experts. Thus, automation in fruit ripeness classification is necessary in this field (Rizzo et al., 2023). Therefore, fruit ripeness as a research field has been investigated for years through various methods across different fruits. For example, many research projects in 2024 focused on identifying fruit ripeness, for example, identifying apple ripeness from digital images (Xiao et al., 2024) and infrared spectral models (Pan et al., 2024), classifying makasar fruit extract (*Brucea javanica* L. Merr) based on ripeness level using spectroscopic methods with chemometrics (Banon et al., 2024), and developing a flexible dual-mechanism piezoelectric sensor on a robotic arm to assess fruit ripeness (Huang et al., 2024).

Production of date fruit (*Phoenix dactylifera* L.) as a staple food has been increasingly produced in recent decades. Moreover, date palm cultivation has recently spread out of traditional regions (Echegaray et al., 2021), and its cultivars vary widely in quality and are suitable for different applications that have gained widespread recognition and acceptance by consumers globally (Alam et al., 2023). On the other side, in-situ date palm fruit ripening is significantly delayed due to climatic variation, and some fruits cannot ripen naturally on the tree and will go to waste. Therefore, the artificial ripening of unripe date fruit can be performed to solve this issue (Mohammed et al., 2021). Three edible stages in the ripening of date fruit are known as *Khalal* (hard ripe), *Rutab* (partially ripe), and *Tamr* (fully ripe) based on their Arabic names in order of ripeness (Al-Kaabi, 2020). These ripening stages, which are essential in the ripening process of date fruit, were investigated to characterize based on different parameters and ripeness indices such as physicochemical composition and volatile components (Amira et al., 2011). The volatile profile is a critical examination in food processing because it can be an introduction to applying tools sensitive to the aroma in food process control. The fruit's developmental stages can affect the fruit's aroma (Moon et al., 2018). There is valuable research on date fruit that 71 aroma volatiles were identified by the gas chromatography/mass spectrometry (GC/MS) method in the Kingdom of Saudi Arabia, and also differences in maturation stages in terms of aroma compounds were reported (Ali et al., 2018). Recent research demonstrates that the sensor accurately monitors concentration changes indicating fruit ripeness and emphasizes that developing gas-sensing technology for fruit ripening is highly significant (Zhang et al., 2023).

Instruments like electronic noses (e-noses) could produce high accuracies for monitoring the physical, physicochemical, and sensory parameters of different food materials due to the formation, transformation, and decomposition of some of the volatile compounds contained by them (Aouadi et al., 2020). Non-destructive instruments like e-noses have been used in many practical applications to analyze and evaluate the quality of fruits and vegetables in the product chain during ripening, production, storage, and distribution (Beghi et al., 2017). Recently, numerous studies have been done using e-noses. In a case study, a prototype e-nose was developed to identify the maturation of the Eragil peach in the pre-harvest growth cycle to determine the harvesting point in the peach trees (Voss et al., 2020). Quality assessment of roasted date pits (Rahman et al., 2018), discriminating kiwifruit at different ripening stages (Du et al., 2019), the ripening status of bananas (Kim et al., 2022), and ripeness detection of cashew apples (Sudha & Mohan, 2023) are the other researches that have been reported for e-nose application.

Selecting an optimal sensor array in an e-nose is fundamental and affects the success or failure of a certain application (Ulivieri et al., 2006).

In the present research, date fruit cv. Shahani was sampled during edible ripening stages, i.e., *Khalal, Rutab*, and *Tamr*, using a laboratory-made e-nose including eight gas sensors and a specified sampling system. To our knowledge, the e-nose as a non-destructive and feasible method was first applied to classify date fruit's aroma in the ripening process, which is a promising method for monitoring the post-harvest process in date fruit. The classification models were examined for repeatability by external datasets. Some popular variable selection methods were used to gain an optimized sensor array to best classification accuracy on an internal dataset and prediction accuracy on an external dataset.

Material and methods

The research and data analysis process are shown in Fig. 1.

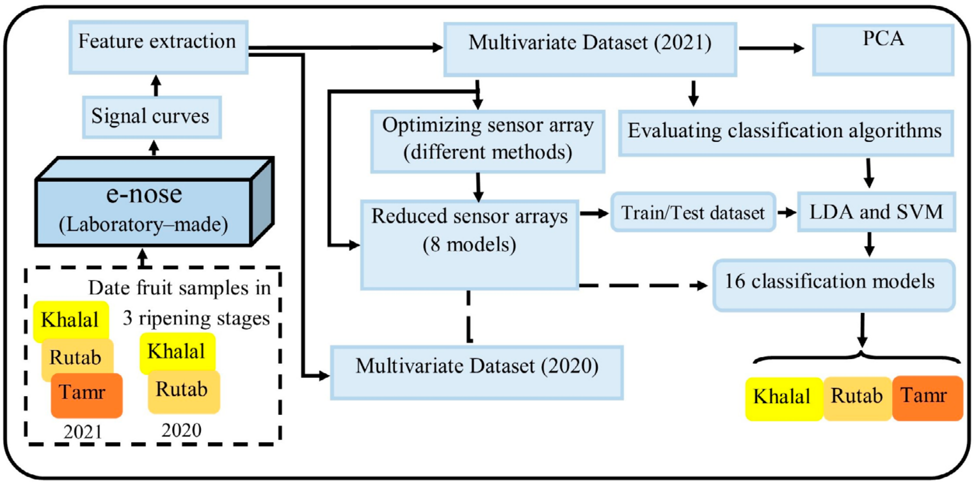

Figure 1. Research flowchart to optimize a gas sensor array to classify three edible ripening stages of date fruit.

Sample collection procedure

The samples were randomly selected from a date palm garden cultivated by the cv. Shahani in Jahrom, located west-south of Iran [1]. The samples were picked up from the branch based on their color in three classes, i.e., *Khalal, Rutab*, and *Tamr*. The date fruit color is yellow in the *Khalal* stage, bright brown in *Rutab*, and brown in the *Tamr* stage. In this study, the date fruit samples were divided into two groups. The first group, as the internal dataset, sampled in late 2021, was applied to establish the classification models. The second group, as an external dataset sampled in late 2020 by an independent study, was applied for external validation to evaluate the repeatability performance of the classification models. The external validation dataset was restrained by two classes, i.e., *Khalal* and *Rutab*, based on the aim of the independent study in 2020.

Gas sensors

Metal oxide semiconductor (MOS) gas sensors are one of the most common sensors due to their high sensitivity, low cost, simplicity (Gao & Zhang, 2018), and ease of integration. Multiple MOS sensor arrays can be applied to the e-nose system for improving the problem of low gas selectivity (Wang et al., 2024). A laboratory-made e-nose (Mohtasebi et al., 2019) based on MOS sensors listed in Table 1 was used in this study.

Table 1. The gas sensors used as the gas sensor array of the laboratory e-nose system.

Sensor name	The sensor sensitivity
TGS 2611	Methane
TGS 2602	Volatile organic compounds (VOCs) and odorous gases
MQ136	Hydrogen sulfide (H_2S)
MQ131	Ozone (O_3)
TGS 2620	Alcohol and VOCs
MQ3	Alcohol
MQ138	VOCs including n-hexane, benzene, ammonia, alcohol, …
TGS 822	VOCs such as ethanol

The e-nose was programmed to 500 seconds for each sampling test, i.e., 200 seconds for baseline, 200 seconds for aroma sampling, and 100 seconds for purging the sensor chamber. The times were set based on the pre-test of date fruit samples, i.e., how long the sensors needed to reach baseline and steady-state response. The intake air into the e-nose was treated with an

1. Google map location: 28.497506, 53.523486 (28°29’51.0”N, 53°31’24.6”E)

active carbon filter as an adsorbent of smells and moisture of environmental air in laboratory space. Five date fruits in each sample were put in a 200-ml sealed container for 10 min before running an e-nose test until the container space was saturated from the sample aroma.

In this innovative e-nose shown in Fig. 2, a gas circulation system creates a closed loop between the sample chamber and the sensor array chamber in the aroma sampling stage to prevent loss of aroma concentration. In other words, there is no exhaust from the e-nose in the aroma sampling stage, and the aroma air circulates along the path of the sensor array and the sample chamber. All experiments were carried out in the laboratory at a temperature range of 27–29 °C and relative humidity (RH) of 22-25%.

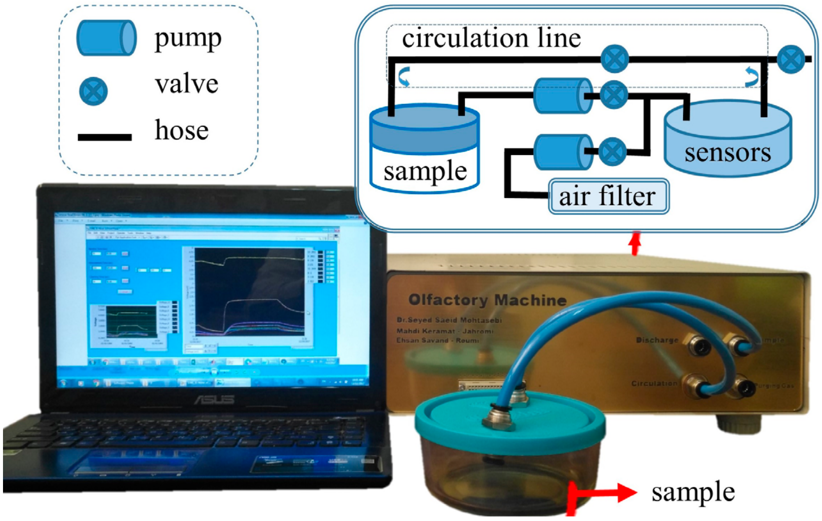

Figure 2. The laboratory e-nose technology used to optimize a gas sensor array to classify three edible ripening stages of date fruit.

Experimental data

The e-nose signals were evaluated as experimental data to classify three ripening stages of date fruit. The maximum value feature from the original response curves of sensors was chosen in this study. The maximum value, which reflects the maximum reaction degree change of sensors responding to odors, is the most common and straightforward e-nose feature (Yan et al., 2015). In the recent research on sensor array optimization in an e-nose using different features, including maximum value, average value, integral value, and wavelet transform, the maximum was shown as a better feature extraction method than the other three (Tong et al., 2022). Then, standardization was applied by ‘StandardScaler’ from Scikit-learn as preprocessing data, a common requirement for many machine learning estimators (Pedregosa et al., 2011).

Analyzing experimental data

The methods of principal component analysis (PCA) and linear discriminant analysis (LDA) are clearly stated as outperformer classification methods in robustness and simplicity (Naik & Wunnava, 2020). Also, they are among the most common methods in analyzing e-nose data as unsupervised and supervised methods, respectively (Karakaya et al., 2020). Actually, PCA as the exploratory data analysis is proper to visualize the distribution of e-nose samples using a bi-plot (the scores and correlation loadings) to show how the sensors affected the positions of the samples in the scores plot (Savand-Roumi et al., 2024). This study applied PCA to visualize the distribution dataset in two-dimensions along with sensor vectors in a bi-plot and explore important sensors to optimize the initial sensor array.

Different algorithms, including LDA, support vector machine (SVM), decision tree (DT), and random forest (RF), were evaluated to identify the proper algorithms for classifying the dataset in this study. Comparing the algorithms was performed on an internal dataset based on responses of eight sensors as initial sensor arrays and validated by Leave One Out Cross-Validation (LOOCV). After selecting two algorithms, an internal dataset was split into a train set (70%) and a test set (30%) to develop and test classification models.

Optimizing sensor array

The sensor array of the e-nose is for general purposes; it can be optimized to reduce the sensor array and is low-cost for special purposes. Moreover, optimizing sensor arrays can improve classification models’ training time, accuracy, and overfitting problems (Tong et al., 2022). A feature selection algorithm is used to obtain features that are more efficient and lead to higher accuracy performances and reduced feature size (Cıbuk et al., 2019). Studies on feature selection methods showed that they are influenced by dataset type (Bhalaji et al., 2018; Dhanya et al., 2019), and the choice of feature selection methods can differ among application areas (Jović et al., 2015). Therefore, this study applied different feature selection methods to select eight reduced sensor arrays. The models of reduced sensor arrays were compared by classification accuracy in training, testing, and repeatability on an external dataset to identify optimized sensor arrays.

The feature selection methods applied in this study are briefly introduced as follows:

Principal component analysis

Principal component analysis as a feature selection method can be applied according to the magnitude of their loadings on the most significant principal components (PCs). For example, in research on water demand, PCA was applied along with six variable selection methods in the context of linear regression. The results showed that PCA has the potential to identify the influential variables better than the other statistical methods, and in the research, variable loadings in PCs were considered as a criterion to select variables (Haque et al., 2018). Other research to select optimal variables using the PCA-based method was done based on the highest loading values corresponding to the most significant PC to identify important variables, and the results showed that the model with reduced variables has a very high prediction accuracy and minimizes overfitting of the model (Wang et al., 2018).

F-test

In this study, the F-test calculates the variance ratio between two classes and within the class for a sensor as a feature or variable. The analyst judges the calculated F-ratio of features based on an F-critical value (F table), and the features cause significant differences between classes that have an F-ratio greater than F-critical (Watson et al., 2016).

Random forest (RF)

As a popular feature selection method, RF is applied to optimize an e-nose's sensor array and estimate the importance of features (Tong et al., 2022).

Feature importance from coefficients

The importance of the features is calculated using the 'RidgeCV' estimator, and the features with the highest absolute coefficient value are considered the most important (Pedregosa et al., 2011).

Sequential feature selection

Sequential feature selection (SFS) is another way of selecting features using 'SequentialFeatureSelector', and the best feature is chosen based on a cross-validation score. Forward SFS starts with empty features and chooses the best single feature with the highest score, and the procedure is repeated until it reaches the desired number of selected features. Backward SFS starts with all the features and greedily chooses features to remove one by one (Pedregosa et al., 2011).

Analyzing data and optimizing models were performed using Spyder 5 (Python 3.11) and the sklearn library (Pedregosa et al., 2011), and the main modules are presented in Supplementary Table S1. The F-test was also done by Unscrambler X 10.4.

Models repeatability

In addition to verifying the results or classification models, repeatability by external validation in this study was done to decrease the gap between research and industrial works. Taking into account different data as replicable work (Pineau et al., 2021) to ensure consistent reproducibility under similar conditions. The external dataset is the same as the internal dataset in e-nose, laboratory conditions, and garden but differs in harvest year. All classification models, after developing by dividing the internal dataset into training and testing data, were evaluated by external data. Actually, this process assesses the challenge of sensor drift in e-nose and its effect on the repeatability of classification models.

Results and discussion

Aroma patterns of date fruit by e-nose sensors

The patterns of date fruit aroma were extracted by the mean of normalized data for each sensor in different edible ripening stages presented by the radar chart in Fig. 3.

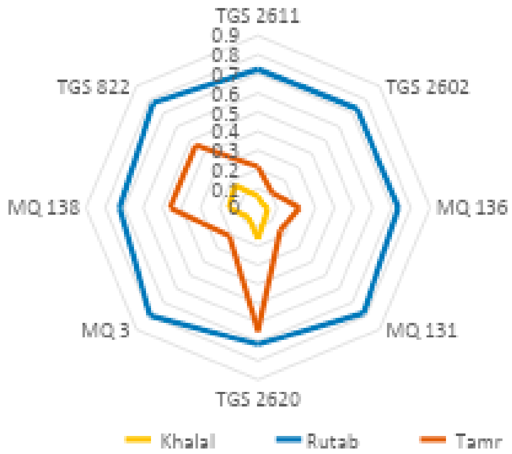

Figure 3. Aroma patterns for different ripening stages of date fruit samples based on mean responses.

There are clearly disferences between the patterns of the three ripening stages, as the highest responses are related to the *Rutab* stage and the least related to the *Khalal* stage.

Visualizing data

This section's results help make data easily understandable and visible using PCA.

Visualization of internal dataset

The e-nose data as observations or samples were introduced to PCA to visualize the spread of the samples along two PCs (PC-1 and PC-2) to interpret the relations in the dataset. According to Fig. 4, the first two main components (PC-1 and PC-2) are explained by more than 95% of variations of the original data. A very high proportion of variability explained by the first two PCs provides a solid ground for the conclusions by PCA (Jollife & Cadima, 2016), and on another side, the samples of each ripening stage are close to each other and separately from other stages. Therefore, it can be concluded that sensory fingerprints were correlated with aroma in different ripening stages. It is expected that the date fruit samples can be classified by the classification models into three classes of edible ripening stages.

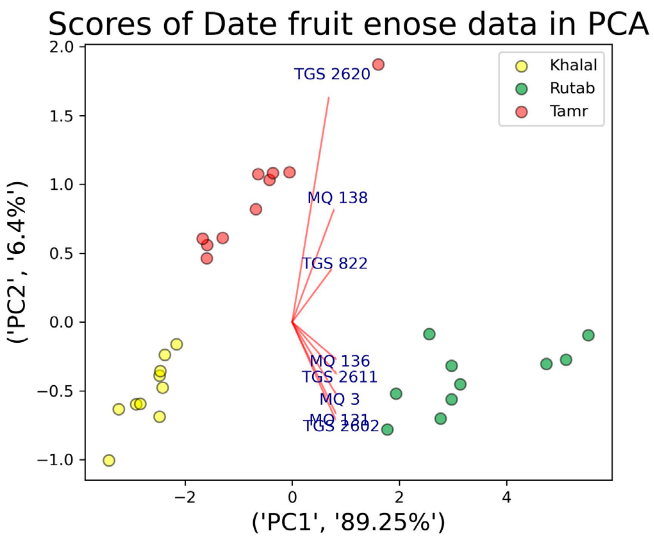

Figure 4. Scores in a bi-plot of PCA with an initial sensor array for different ripening stages samples.

The position of sensors in Fig. 4 shows that all sensors positively correlate to PC-1. Therefore, according to the position of the projection of samples on PC-1, the sensors' responses are higher in the *Rutab* stage than in the other stages due to more aroma concentration. Research on identifying the volatile constituents of date palm fruits by the GC/MS reported that qualitative and quantitative differences in compounds responsible for aroma were noted across ripening stages, i.e., at the Besser (*Khalal*), *Rutab*, and *Tamr*, so that phenolic compounds were at the highest level in the *Rutab* stage than the other ripening stages in varieties of the Khenazy, Khalas, and Sokary (Ali et al., 2018).

Integrating internal and external dataset

Integrating external data with internal data on a PCA plot was performed in two ways based on Fig. 5 (a). Actually, Fig. 5 (b) shows that sensor drift or other unwanted changes in experimental conditions cause a change in the distribution of input data with the passing of a year, and that is why the similar classes in the internal dataset (i.e., *Khalal*, *Rutab*, and *Tamr*) and the external dataset (i.e., Ex *Khalal* and Ex *Rutab*) are laid far away from each other. This problem has been described in the applications of e-nose; for example, the factors including the aging and poisoning of gas sensors and the change in operating conditions make the distribution of test samples different from that of training samples (Yan & Zhang, 2016); the sensor drift is a major disadvantage in the gas sensors that the data distribution may change over time (Se et al., 2024); and "data drift" or "data shift" causes degrade the performance of predictive machine learning models with time (Rahmani et al., 2023). Then, after scaling the datasets by standardizing, the similar classes in the internal and external datasets are close to each other on the PCA plot in Fig 5 (c). Therefore, the standardization can compensate for the sensor drift in this research.

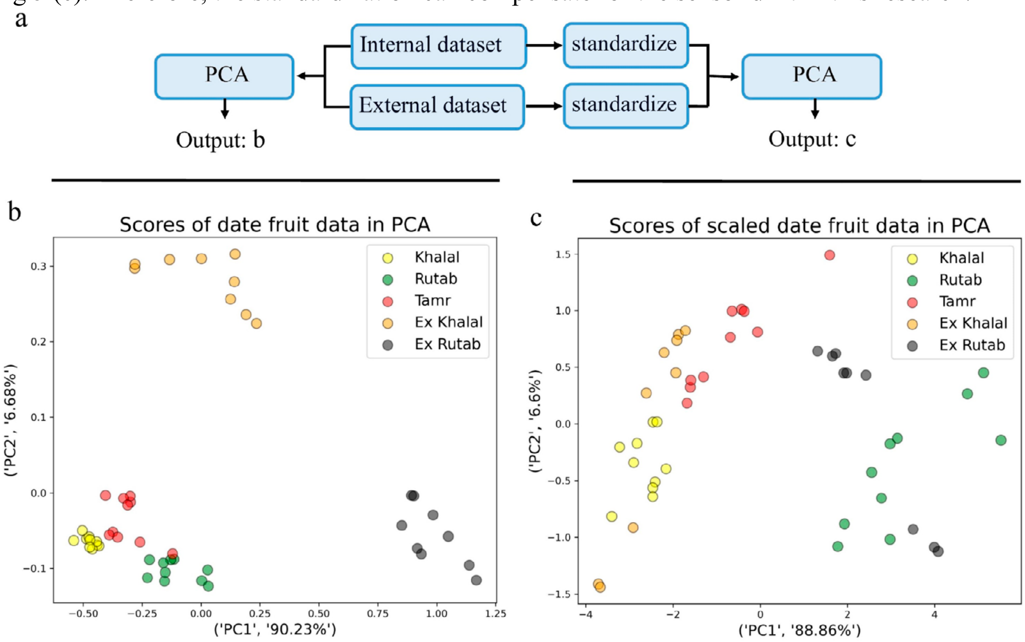

Figure 5. Methods introducing the datasets to PCA (a), data situation on PCA plot before standardization (b), and after standardization (c), Ex: External data.

Evaluating classification algorithms

Classification of three ripening stages of date fruit (internal dataset) based on responses of eight sensors was performed by different algorithms, including LDA, SVM, DT, and RF. The mean accuracy estimate was obtained 100% using LOOCV with zero standard deviation in all algorithms. Hence, robust work leads to the same result on the same data but different analysis methods (Pineau et al., 2021). Therefore, the first two algorithms as different analysis methods, i.e., LDA and SVM, were chosen to develop the models in this study.

Reduced sensor arrays

Eight models of reduced sensor arrays were extracted and evaluated to achieve an optimized sensor array for recognizing three ripening stages of date fruit. The results of sensor selection methods are explained in detail for each model in the following.

Prinicpal Component Analysis

PCA in a bi-plot way aids the identification of important variables, i.e., sensors in this study. The main question in selecting sensors is how many can be selected initially. Based on this point that enough information is provided by the first q PCs, it is often only necessary to keep (q+1) or (q+2) variables (Jolliffe, 2002); therefore, in the first step of sensor selection based on the result of PCA in Fig. 4, the first two PCs explained more than 95% of the total variance, the number of sensors was chosen as three in the beginning.

The important sensors were selected based on two methods by the PCA plot in Fig. 4, and values of the sensor loadings on PC-1 and PC-2 are reported in Table 2.

Table 2. F ratio of each sensor response between different ripening stages and sensor loadings in three first PCs.

Sensor	F ratio* between classes			Loadings in PCA		
	Khalal - Rutab	*Rutab - Tamr*	*Khalal - Tamr*	PC-1	PC-2	PC-3
TGS** 2611	**18.36999**	**5.51077**	**3.333471**	**0.369934**	-0.16856	0.081822
TGS 2602	57.68604	2.136036	27.00612	0.361208	**-0.31638**	0.058341
MQ** 136	**30.12675**	**3.520655**	**8.557142**	**0.368272**	-0.12114	0.226743
MQ 131	26.52692	2.847566	9.315649	0.364295	-0.29723	0.091698
TGS 2620	6.916918	1.229028	5.627957	0.310179	**0.739704**	**0.295814**
MQ 3	**12.91047**	**3.343843**	**3.860967**	**0.368219**	-0.23537	0.027759
MQ 138	5.290115	1.257427	6.651933	0.352405	0.369175	0.086405
TGS 822	1.075948	1.688174	1.569012	0.329176	0.172108	**-0.91342**

*F table: 3.178893

** The MQ series gas sensor, and TGS is a commercial brand for gas sensors produced by Figaro Engineering Inc.

- The first method is based on the position of samples, the direction of sensors, and which sensors are most strongly loadings with each component, i.e., PC-1 and PC-2 in the scores plot in Fig. 4. All sensors are in the positive direction on PC-1. Therefore, TGS 2611, as the highest loading on PC-1, was kept, and other sensors were removed. From the point of view of PC-2 in Fig. 4 and Table. 2, TGS 2620 and TGS 2602 were chosen as the highest loadings in opposite directions. Therefore, three sensors, i.e., TGS 2611, TGS 2602, and TGS 2620, were selected as a reduced sensor array named Model 2.
- The second method is based on only the PC-1, which explains 89.25% of the total variation, and the projections of the samples on the PC-1 axis can be separated into three classes. Hence, the three first sensors with the highest loading on the PC-1 were selected, as shown in Table. 2, i.e., TGS 2611, MQ 136, and MQ 3, and this array is introduced as Model 4 (it will be shown this method can lead to quite the same sensors in Model 4 through the F-test method).

F-test

The F-test is applied to analyze the variance based on the F statistic. A F-test on the sensor data was done between three ripening stages to essay the effect of aroma in ripening stages on each sensor. The results of the F-test are presented in Table. 2. If the F ratio is larger than the value in the F table, the effect is significant. Models 3, 4, and 5 were extracted based on Table. 2 as follows:

- Model 3: The sensors were selected based on the highest F ratio between each class. TGS 2602 (highest F ratio between *Khalal* and *Rutab*), TGS 2611 (highest F ratio between *Rutab* and *Tamr*), and TGS 2602 (highest F ratio between *Khalal* and *Tamr*). Therefore, TGS 2602 and TGS 2611 were selected as models 3.
- Model 4: The sensors were selected based on having a significant F ratio between all three classes. TGS 2611, MQ 136, and MQ 3 have a significant F ratio between all three classes, bolded in Table 2. Therefore, TGS 2611, MQ 136, and MQ 3 were selected as model 4, and this method can lead to quite the same sensors in the second method in PCA.
- Model 5: Models 3 and 4 have the best accuracy in predicting the external data, but model 3 has one sensor less than model 4. MQ 3 had a lower F ratio than TGS 2611 and MQ 136, then it was removed, and Model 5 was developed by TGS 2611 and MQ 136 as a reduced sensor array of Model 4.

Random forest (RF)

The relative importance of RF sensors is shown in Fig. 6, which are an average of 10 runs on the an internal dataset. The three first sensors, i.e., TGS 2611, MQ 3, and TGS 2620, were selected as model 6.

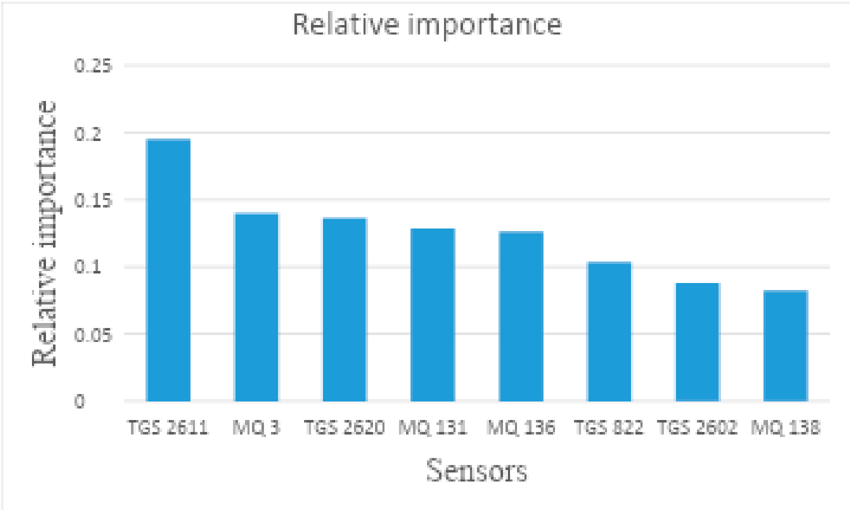

Figure 6. Relative importance of an initial sensor array using an RF algorithm to classify three edible stages of date fruit ripening.

Sequential feature selection (SFS)

Sequential feature selection was applied in forward and backward methods to identify the three most important sensors to develop models 7 and 8. Selected sensors by forward SFS are MQ 136, MQ 131, and TGS 2620 in model 7, and selected sensors by backward SFS are TGS 2620, MQ 3, and TGS 822 in model 8. TGS 2620 is a common sensor between forward SFS and backward SFS.

Sensor importance from coefficients

The importance of sensors based on the 'feature importance from coefficients' method was applied to develop Model 9, and its result is shown in Fig. 7. Therefore, TGS 822, TGS 2620, and MQ 136 were selected as the three most important sensors in Model 9.

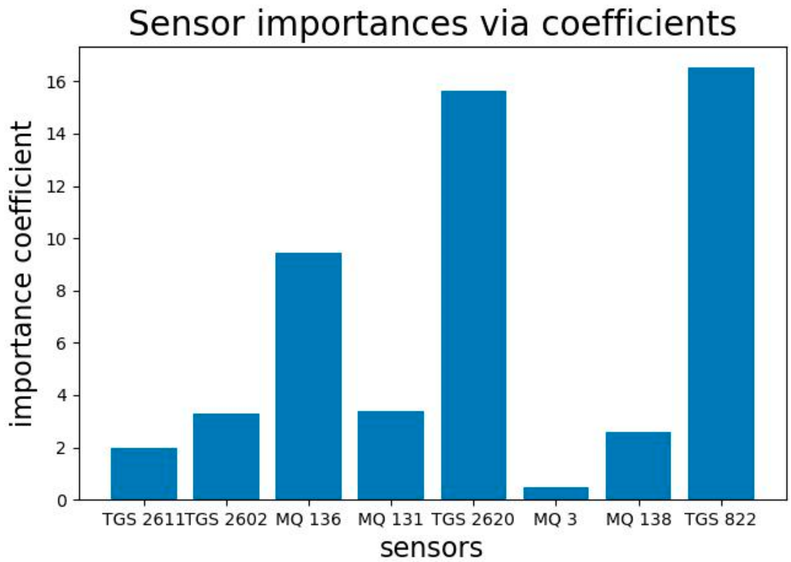

Figure 7. The importance of sensors in the initial sensor array to classify three edible stages of date fruit using the 'feature importance from coefficients' method.

Finally, the results of evaluating the reduced sensor arrays by different methods are presented to compare with the initial sensor array in Table 3 to choose the optimized sensor array. The results show that the different variable selection methods can lead to different reduced sensor arrays but the same result in classification accuracy of training and testing on the internal datasets. Therefore, in the next section, these models will be compared by classifying performance on external datasets.

Table 3. Categorizing three edible date fruit stages by comparing a new reduced sensor array to the original sensor array.

Model number	Sensor selection method	Selected sensors
Model 1	Initial sensor array	TGS 2611, TGS 2602, MQ 136, MQ 131, TGS 2620, MQ 3, MQ 138, TGS 822
Model 2	PCA	TGS 2611, TGS 2602, TGS 2620
Model 3	F-test	TGS 2611, TGS 2602
Model 4	F-test, PCA	TGS 2611, MQ 136, MQ 3
Model 5	F-test, PCA	TGS 2611, MQ 136
Model 6	Random forest	TGS 2611, TGS 2620, MQ 3
Model 7	Forward SFC	MQ 136, MQ 131, TGS 2620
Model 8	Backward SFC	TGS 2620, MQ 3, TGS 822
Model 9	Feature importance from coefficient	MQ 136, TGS 2620, TGS 822

Evaluating repeatability of models

Since the test performance on the internal dataset was equal in reduced sensor arrays and the initial sensor array, an external dataset, besides evaluating the repeatability in models, also compares them to the initial sensor array.

The results of the repeatability in the models on the external dataset are reported in Table 4. The results show that model No. 1, the initial sensor array of 8 sensors with 100% accuracy, has poor repeatability on an external dataset. The effect of feature selection methods is different on repeatability based on classification algorithms, i.e., LDA and SVM. All reduced sensor arrays improve the repeatability in LDA models better than the initial sensor array (model 1), especially in models No. 3, 4, and 5, where the sensors were selected by the F-test and PCA methods. However, the feature selection methods have positive and negative effects on repeatability than the initial sensor array in the SVM models, and like LDA, models No. 3, 4, and 5 have positive effects and improve the repeatability to 100% in prediction accuracy.

Table 4. The models' performance as smaller sensor arrays (described in Section 3.4)

Model No.	Training accuracy		Testing accuracy		Prediction accuracy of external data	
	LDA	SVM	LDA	SVM	LDA	SVM
Model 1	1	1	1	1	0.33	0.72
Model 2	1	1	1	1	0.61	0.67
Model 3	0.9	0.86	1	1	0.94	1
Model 4	0.9	0.86	1	1	0.94	1
Model 5	0.9	0.86	1	1	0.94	1
Model 6	1	1	1	1	0.61	0.67
Model 7	1	1	1	1	0.39	0.67
Model 8	1	1	1	1	0.5	0.67
Model 9	1	1	1	1	0.61	0.67

LDA: Linear discriminant analysis; SVM: Support vector machine

There is an interesting point about TGS 2611 and TGS 2620 that are selected in five models, both of which are in model 2. Therefore, a comparison between the performance of models 2 and 3 (model 3 is like model 2 but without TGS 2620) will explain which of them is more efficient in classifying the classes in this study. Based on Table 4, the classifying performance on the external dataset improves credibly significantly by LDA and SVM in model 3 when TGS 2620 is removed from model 2. On the other side, TGS 2620 was selected in models 6, 7, 8, and 9 developed by popular feature selection methods in machine learning, and these models suffer from overfitting in the external dataset in this study.

Comparison of initial and optimized sensor arrays

The performance of the initial sensor array and the optimized sensor arrays is compared by PCA and confusion matrix in this section.

According to Table 4, models 3, 4, and 5 show the best repeatability results among the eight reduced sensor array models. The internal and external datasets are introduced to PCA through the models to assess the effect of these models on the position of the samples in the PCA plot. In these models, shown in Fig. 8 (b, c, and d), the samples are closer together within each class than in model 1 in Fig. 8(a). Also, these models capture more information by PC-1 and PC-2 than model 1. Based on Fig 8 (C), a few samples in classes *Khalal* and *Tamr* overlap partially on an imaginary separation border using model 4. However, the samples in models 3 and 5, shown in Figs. 8(b) and (d), are clearly separated. Therefore, models 3 and 5 are introduced as optimized sensor arrays based on equal properties like accuracy in classification, clear separation between samples of three classes in the scores plot of PCA, and finally, the number of sensors.

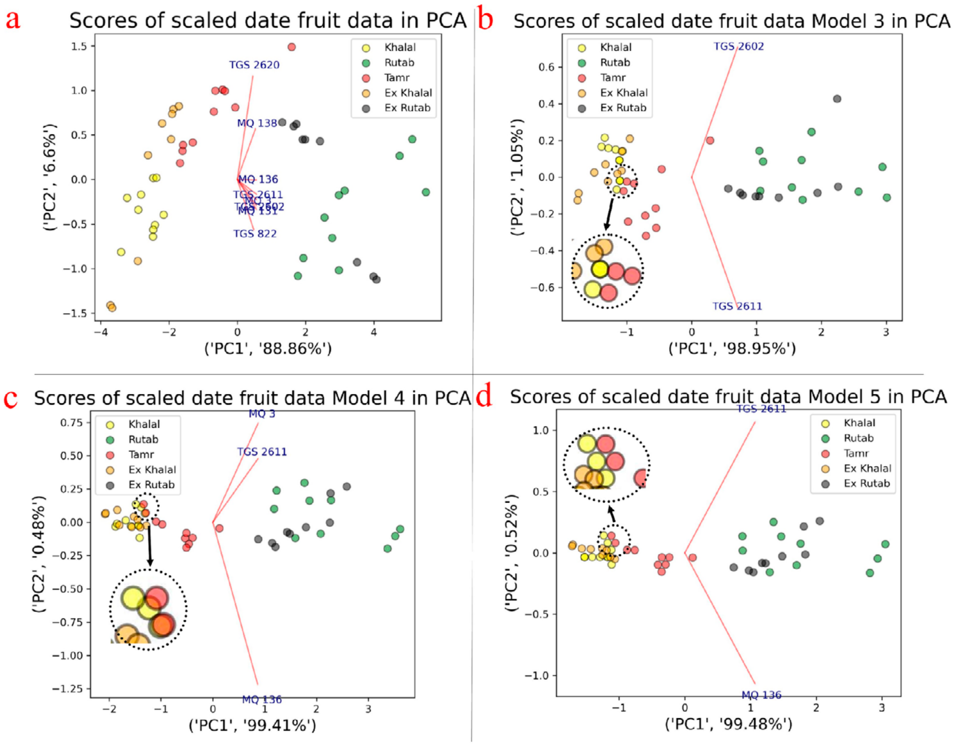

Figure 8. Comparison of initial sensor array and optimized sensor array using PCA.

The performance of prediction models 1, 3, and 5 using SVM on the external dataset is presented by the confusion matrix in Fig. 9 (the confusion matrices in models 3 and 5 are equal). The effect of optimized sensor arrays and the initial sensor array is investigated on prediction performance by comparing the confusion matrices. There are some misclassifications between classes *Khalal* (1) and *Tamr* (3) based on Fig. 9 (a) using the model of the initial sensor array, i.e., model 1. Then, the misclassifications are thoroughly corrected based on Fig. 9 (b) by models 3 and 5 using optimized sensor arrays.

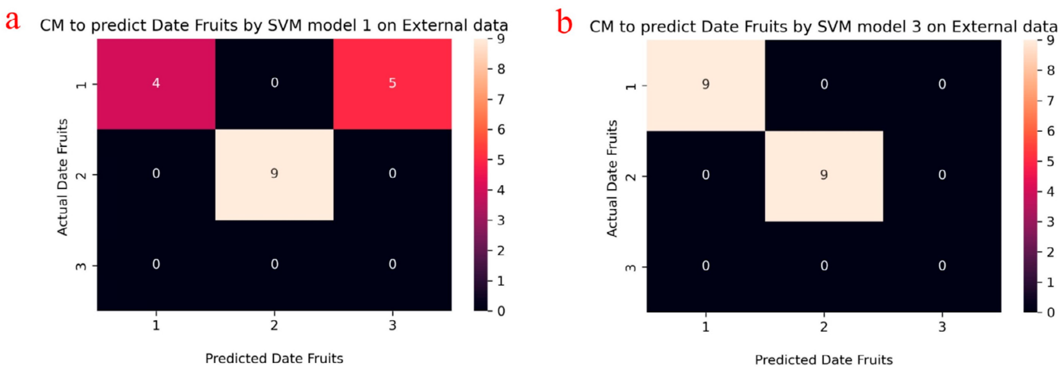

Figure 9. Comparison confusion matrix between model 1 (a) as initial sensor array and model 3 (b) as optimized sensor array (the confusion matrices in models 3 and 5 are equal) to predict external dataset using SVM into three classes: 1: *Khalal*, 2: *Rutab*, 3: *Tamr*.

Finally, some points are summarized from the results in terms of sensor selection in this study as follows:

- Test accuracy in all models of the reduced sensor array was the same and equal to 100%.
- The effect of feature selection on the model repeatability (prediction performance on an external dataset) is related to the classification algorithm.
- In the models based on LDA, all reduced sensor arrays had a positive effect on repeatability, but the models based on SVM had positive and negative effects on repeatability than the initial sensor array.
- The repeatability performance of SVM was better than LDA in each model of a reduced sensor array.
- PCA was an effective and easy method to address the efficient number of sensors (variables) at the beginning of the sensor selection process.
- The reduced sensor arrays that were selected by F-test and PCA lead to the optimized sensor arrays in this study.

Conclusions

The reduced models based on F-test and PCA showed better repeatability results than the popular feature selection methods in ML that were applied in this study.

An e-nose was applied in laboratory condition to extract aroma patterns from date fruits to classify and predict three edible ripening stages _*Khalal, Rutab*, and *Tamr*_ from the cv. Shahani in the Jahrom region of Iran. Using an internal dataset, the LDA and SVM models achieved 100% accuracy in classifying these stages. However, prediction accuracy significantly dropped on an external dataset, highlighting issues with repeatability. To improve repeatability, some optimization methods were applied, and ultimately, an optimized reduced sensor array improved model repeatability to 100% in the SVM model. In general, the results indicate that e-nose successfully detected aroma changes in ripening of date fruits, highlighting its potential for future post-harvest research and the development of monitoring devices for date fruit processing. Future research will expand with new samples from various regions and varieties to assess a comprehensive model for date fruits. Additionally, the results in optimization methods indicate that the different methods showed the same results in test accuracy on the internal dataset and different results in the external dataset or repeatability, which shows the importance of assessing both internal and external validation metrics to evaluate model performance. The model repeatability was affected by feature selection method and classification algorithm.

Supplementary material (Table S1): accompanies the paper on SJAR's website.

Competing interests: The authors have declared that no competing interests exist.

Statement about use of generative AI: The following instructions refer solely to the composition process; they do not address the application of AI instruments for data analysis. Authors are required to include a disclosure that details the usage of generative AI and AI-assisted technologies in the writing process, prior to the References list:

The authors employed QuillBot (https://quillbot.com/) in order to translate from English to Spanish while preparing this work. Following their use, the author(s) assumed full responsibility for the publication's content and reviewed and edited it as necessary.

This statement does not cover the use of simple grammar, spelling, reference, and other checkers. There's no need to include this statement if there's nothing to reveal.

Authors' contributions: Ehsan Savand-Roumi: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Mahdi Keramat-Jahromi:** Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding: The authors received no specific funding for this work.

References

Al-Kaabi HJ, 2020. Effect of modified atmosphere packaging on quality of barhi dates at Khalal stage. International Journal of Food Studies, 9.https://doi.org/10.7455/ijfs/9.SI.2020.a10

Alam MZ, Al-Hamimi S, Ayyash M, Rosa CT, Yahia EM, Haris S, Al-Marzouqi AH, Kamal-Eldin A, 2023. Contributing factors to quality of date (*Phoenix dactylifera* L.) fruit. Scientia Horticulturae 321. https://doi.org/10.1016/j.scienta.2023.112256

Ali H, Al-Khalifa AR, Farouk A, Shaheen M, 2018. Effect of maturation stages on flavor profile and antioxidant activity of date palm fruits (*Phoenix dactylifera*) grown in Saudi Arabia. International Journal of Pharmacology 14(3). https://doi.

org/10.3923/ijp.2018.407.414

Amira EA, Flamini G, Behija SE, Manel I, Nesrine Z, Ali F, Mohamed H, Noureddine HA, Lotfi A, 2011. Chemical and aroma volatile compositions of date palm (*Phoenix dactylifera* L.) fruits at three maturation stages. Food Chemistry 127(4). https://doi.org/10.1016/j.foodchem.2011.02.051

Aouadi B, Zaukuu JLZ, Vitális F, Bodor Z, Fehér O, Gillay Z, Bazar G, Kovacs Z, 2020. Historical evolution and food control achievements of near infrared spectroscopy, electronic nose, and electronic tongue—critical overview. Sensors 20(19). https://doi.org/10.3390/s20195479

Banon C, Triawan DA, Adfa M, Sihite MSA, Nurwidiyani R, Rahmi W, 2024. Classification of Makasar Fruit Extract (*Brucea javanica* L. Merr.) Based on The Level of Ripeness Using a Combination of FTIR and UV-Vis Spectroscopy with Chemometrics. Jurnal Kimia Sains Dan Aplikasi 26(12). https://doi.org/10.14710/jksa.26.12.483-488

Beghi R, Buratti S, Giovenzana V, Benedetti S, Guidetti R, 2017. Electronic nose and visible-near infrared spectroscopy in fruit and vegetable monitoring. Reviews in Analytical Chemistry 36(4). https://doi.org/10.1515/revac-2016-0016

Bhalaji N, Sundhara Kumar KB, Selvaraj C, 2018. Empirical study of feature selection methods over classification algorithms. International Journal of Intelligent Systems Technologies and Applications 17(1–2). https://doi.org/10.1504/IJISTA.2018.091590

Cıbuk M, Budak U, Guo Y, Cevdet Ince M, Sengur A, 2019. Efficient deep features selections and classification for flower species recognition. Measurement. Journal of the International Measurement Confederation 137. https://doi.org/10.1016/j.measurement.2019.01.041

Dhanya R, Paul IR, Sindhu Akula S, Sivakumar M, Nair JJ, 2019. A comparative study for breast cancer prediction using machine learning and feature selection. International Conference on Intelligent Computing and Control Systems, ICCS 2019. https://doi.org/10.1109/ICCS45141.2019.9065563

Du D, Wang J, Wang B, Zhu L, Hong X, 2019. Ripeness prediction of postharvest Kiwifruit using a MOS E-nose combined with chemometrics. Sensors 19(2). https://doi.org/10.3390/s19020419

Echegaray N, Gullón B, Pateiro M, Amarowicz R, Misihairabgwi JM, Lorenzo JM, 2021. Date Fruit and Its By-products as Promising Source of Bioactive Components: A Review. Food Reviews International. https://doi.org/10.1080/87559129.2021.1934003

Gao X, Zhang T, 2018. An overview: Facet-dependent metal oxide semiconductor gas sensors. Sensors and Actuators, B: Chemical 277. https://doi.org/10.1016/j.snb.2018.08.129

Haque MM, Rahman A, Hagare D, Chowdhury RK, 2018. A comparative assessment of variable selection methods in urban water demand forecasting. Water 10(4). https://doi.org/10.3390/w10040419

Huang W, Xia J, Wang Y, Jin X, Zhu HL, Zhang X, 2024. Flexible multimode sensors based on hierarchical microstructures enable non-destructive grading of fruits in cold chain logistics. Materials Today Sustainability 25. https://doi.org/10.1016/j.mtsust.2024.100691

Jolliffe IT, 2002. Principal Component Analysis. Springer-Verlag. https://doi.org/10.1007/b98835

Jollife IT, Cadima J, 2016. Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 374(2065), p. 20150202. https://doi.org/10.1098/rsta.2015.0202

Jović A, Brkić K, Bogunović N, 2015. A review of feature selection methods with applications. 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015. https://doi.org/10.1109/MIPRO.2015.7160458

Karakaya D, Ulucan O, Turkan M, 2020. Electronic Nose and Its Applications: A Survey. International Journal of Automation and Computing. https://doi.org/10.1007/s11633-019-1212-9

Kim C, Kim SJ, Lee Y, Nguyen TM, Lee JM, Moon JS, Han DW, Oh JW, 2022. A phage- and colorimetric sensor-based artificial nose model for banana ripening analysis. Sensors and Actuators B: Chemical, 362. https://doi.org/10.1016/j.snb.2022.131763

Mohammed M, Sallam A, Alqahtani N, Munir M, 2021. The combined effects of precision-controlled temperature and relative humidity on artificial ripening and quality of date fruit. Foods 10(11). https://doi.org/10.3390/foods10112636

Mohtasebi SS, Keramat-Jahromi M, Savand-Roumi E, 2019. Laboratory Olfactory Machine, Iranian Research Organization for Science and Technology (IROST), No. 1397236. DOI: 10.22104/IROST.1397.236

Moon P, Fu Y, Bai J, Plotto A, Crane J, Chambers A, 2018. Assessment of fruit aroma for twenty-seven guava (*Psidium guajava*) accessions through three fruit developmental stages. Scientia Horticulturae 238.https://doi.org/10.1016/j.scienta.2018.04.067

Naik MK, Wunnava A, 2020. Classical 2d face recognition: a survey on methods, face databases, and performance evaluation. Lecture Notes in Networks and Systems109: 375–383. https://doi.org/10.1007/978-981-15-2774-6_45

Pan L, Li H, Hu Z, Zhang M, Zhao J, 2024. Standard-free sample model transfer of visible/near -infrared spectral model of apple ripeness under seasonal variation. Journal of Food Composition and Analysis 128.https://doi.org/10.1016/j.jfca.2024.106028

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É, 2011. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12:2825–2830. http://scikit-learn.sourceforge.net.

Pineau J, Vincent-Lamarre P, Sinha K, Lariviére V, Beygelzimer A, d'Alché-Buc F, Fox E, Larochelle H, 2021. Improving reproducibility in machine learning research (a report from the neurips 2019 reproducibility program). Journal of Machine Learning Research 22.

Rahman MS, Al-Farsi K, Al-Maskari SS, Al-Habsi NA, 2018. Stability of electronic nose (e-nose) as determined by considering date-pits heated at different temperatures. International Journal of Food Properties. https://doi.org/10.1080/10942912.2018.1463540

Rahmani K, Thapa R, Tsou P, Casie Chetty S, Barnes G, Lam C, Foon Tso C, 2023. Assessing the effects of data drift on the performance of machine learning models used in clinical sepsis prediction. International Journal of Medical Informatics 173. https://doi.org/10.1016/j.ijmedinf.2022.104930

Rizzo M, Marcuzzo M, Zangari A, Gasparetto A, Albarelli A, 2023. Fruit ripeness classification: A survey. Artificial Intelligence in Agriculture 7. https://doi.org/10.1016/j.aiia.2023.02.004

Savand-Roumi E, Salehiyoun AR, Mohtasebi SS, 2024. Use of a biogas-specific e-nose with machine learning to identify biogas pattern changes linked to hydraulic retention times in an anaerobic digester: A case study. Fuel 357. 130013. https://doi.org/10.1016/j.fuel.2023.130013

Se H, Song K, Sun C, Jiang J, Liu H, Wang B, Wang X, Zhang W, Liu J, 2024. Online drift compensation framework based on active learning for gas classification and concentration prediction. Sensors and Actuators B: Chemical 398. 134716. https://doi.org/https://doi.org/10.1016/j.snb.2023.134716

Sudha C, Mohan KJ, 2023. E-Nose for Cashew Apple Ripeness Detection for Autonomous Fruit Plucking. SSRG International Journal of Electronics and Communication Engineering 10(10). https://doi.org/10.14445/23488549/IJECE-V10I10P105

Tong J, Song C, Tong T, Zong X, Liu Z, Wang S, Tan L, Li Y, Chang Z, 2022. Design and Optimization of Electronic Nose Sensor Array for Real-Time and Rapid Detection of Vehicle Exhaust Pollutants. Chemosensors 10(12). https://doi.org/10.3390/chemosensors10120496

Ulivieri N, Distante C, Luca T, Rocchi S, Siciliano P, 2006. IEEE1451.4: A way to standardize gas sensor. Sensors and Actuators, B: Chemical 114(1). https://doi.org/10.1016/j.snb.2005.04.044

Voss HGJ, Ayub RA, Stevan SL, 2020. E-nose Prototype to Monitoring the Growth and Maturation of Peaches in the Orchard. IEEE Sensors Journal 20(20). 11741-11750.

Wang X, Zhao W, Ma R, Zhuo J, Zeng Y, Wu P, Chu J, 2024. A novel high accuracy fast gas detection algorithm based on multi-task learning. Measurement 228. 114383. https://doi.org/10.1016/J.MEASUREMENT.2024.114383

Wang Y, Ma X, Qian P, 2018. Wind Turbine Fault Detection and Identification Through PCA-Based Optimal Variable Selection. IEEE Transactions on Sustainable Energy 9(4). https://doi.org/10.1109/TSTE.2018.2801625

Watson NE, Parsons BA, Synovec RE, 2016. Performance evaluation of tile-based Fisher Ratio analysis using a benchmark yeast metabolome dataset. Journal of Chromatography A 1459: 101–111. https://doi.org/10.1016/j.chroma.2016.06.067

Xiao B, Nguyen M, Yan Wq, 2024. Fruit ripeness identification using YOLOv8 model. Multimedia Tools and Applications 83(9). https://doi.org/10.1007/s11042-023-16570-9

Yan J, Guo X, Duan S, Jia P, Wang L, Peng C, Zhang S, 2015. Electronic Nose Feature Extraction Methods: A Review. Sensors 15. 27804–27831. https://doi.org/10.3390/s151127804

Yan K, Zhang D, 2016. Calibration transfer and drift compensation of e-noses via coupled task learning. Sensors and Actuators, B: Chemical 225. https://doi.org/10.1016/j.snb.2015.11.058

Zhang X, Li Y, Hong T, Tegeltija S, Babić M, Wang X, Ostojić G, Stankovski S, Marinković D, 2023. Response Characteristics Study of Ethylene Sensor for Fruit Ripening under Temperature Control. Sensors 23(11). https://doi.org/10.3390/s23115203