Heliyon 10 (2024) e32189

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Review article

Application of machine learning approach on halal meat authentication principle, challenges, and prospects: A review

Abdul Mustapha [a], Iskandar Ishak [a,b], Nor Nadiha Mohd Zaki [a,c], Mohammad Rashedi Ismail-Fitry [a,d], Syariena Arshad [a], Awis Qurni Sazili [a,c,*]

[a] *Halal Products Research Institute, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor, Malaysia*
[b] *Department of Computer Science, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang, 43400, Malaysia*
[c] *Department of Animal Science, Faculty of Agriculture, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor, Malaysia*
[d] *Department of Food Technology, Faculty of Food Science and Technology, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor, Malaysia*

A R T I C L E I N F O

Keywords:
Adulteration
Authentication
Halal meat
Machine learning
Supervised
Unsupervised

A B S T R A C T

Meat is a source of essential amino acids that are necessary for human growth and development, meat can come from dead, alive, Halal, or non-Halal animal species which are intentionally or economically (adulteration) sold to consumers. Sharia has prohibited the consumption of pork by Muslims. Because of the activities of adulterators in recent times, consumers are aware of what they eat. In the past, several methods were employed for the authentication of Halal meat, but numerous drawbacks are attached to this method such as lack of flexibility, limited application, time,consumption and low level of accuracy and sensitivity. Machine Learning (ML) is the concept of learning through the development and application of algorithms from given data and making predictions or decisions without being explicitly programmed. The techniques compared with traditional methods in Halal meat authentication are fast, flexible, scaled, automated, less expensive, high accuracy and sensitivity. Some of the ML approaches used in Halal meat authentication have proven a high percentage of accuracy in meat authenticity while other approaches show no evidence of Halal meat authentication for now. The paper critically highlighted some of the principles, challenges, successes, and prospects of ML approaches in the authentication of Halal meat.

1. Introduction

Meat is an animal product, a major source of essential nutrients such as amino acids, protein, minerals, and fat-soluble vitamins [1], and essential fatty acids which are necessary for growth and development [2,3]. This definition of meat is broad without any ambit, this means meat should be from any species of animal irrespective of whether dead or alive, while in Sharia law (Islamic law) meat to be consumed should be Halal meaning the meat should come from live and specified species of animals as prescribed by Sharia, including way and manner the animal is being killed (Halal slaughter). That is, the meat must be from Halal animals and slaughtered using Halal method. Halal meat is defined as meat obtained from Halal animals and killed in accordance with Quran and Sunnah [4]. In recent times, there have been cases of adulteration of Halal meat with meat from unlawful sources such as the meat of dogs and pork [5]. Which are usually intentional and economically motivated [6]. Intentionally, pig derivatives are used in pharmaceutical industries

* Corresponding author. Halal Products Research Institute, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor, Malaysia.
E-mail address: awis@upm.edu.my (A.Q. Sazili).

https://doi.org/10.1016/j.heliyon.2024.e32189
Received 29 January 2024; Received in revised form 20 May 2024; Accepted 29 May 2024
Available online 30 May 2024
2405-8440/

for various medical purposes due to their physicochemical properties, such as in gelatin and insulin production, as heparin is used as an anticoagulant and lard for specific drug production [7,5]. There are several markers used for the analysis of lard and heparin, these include; Platelet activation, fatty acid, molecular weight determination, microscopic analysis, specific biochemical assays, and so on (see Table 3, Figs. 1 and 2).

Most of these adulterators have no religious consciousness with unethical profit-driven business models. Due to issues arising from their activities, consumers are now very much aware and pay more attention to the meat and meat products in the market [8]. Consumption of some species of animals by Muslims is prohibited (The Holy Qur'an, 1:173; 5:3; 6:145; 16:115). Medically, the consumption of pork is discouraged due to potential health risks associated with it, parasitic infection, high-fat content, and toxic accumulation of the meat [9]. Several methods were deployed for the authentication of Halal meat such as DNA based method, Spectroscopic methods (Fourier-Transform Infrared Spectroscopy (FTIR) and Near-infrared Spectroscopy (NIRS), Mass spectrometry, Microscopy, chemical analysis, stable isotope analysis, high-performance liquid chromatography (HPLC) and gas chromatography (GC) [10,11,12,13,14] but the drawbacks to these measures are burdensome, time-consuming, destructive, require expertise in its operation, and are not appropriate for immediate/onsite authentication of meat within the supply chain [15]. Recently, Machine Learning (ML) has been reported to successfully assist researchers in the area of detection and authentication of Halal meat and products when combined with some of the methods mentioned [16,2]. ML is a panacea adulteration problem in meat industries, one of the advantages of ML is that it can be used for on-site meat authentication [17]. ML is the concept of learning through the development and application of algorithms from given data [18]. Additionally [19], defined Machine Learning as the process of creating a computer-based model that automatically learns and gets better over time.

Previously, review works were conducted on different Machine Learning in the field of food and agriculture [20]. ML is automated in its authentication process, thus reducing manual labour and human intervention, the model can be easily deployed and scaled across different locations, the use of Machine Learning in meat adulteration detection and authentication, its fast, cost-effective, and high degree of accuracy ([21,22]. ML is superior in many ramifications compared with other traditional methods [23].

Studies have indicated the different analytical methods used in the authentication of Halal meat as presented in Table 1. These traditional methods ~~are~~ faced with various forms of limitations and constraints such as lack of flexibility, limited application, time consuming, expensive and low level of accuracy and sensitivity. ML in Halal meat authentication is fast, flexible, cheaper, highly sensitive and high level of accuracy in its operations. For this reason, the paper intends to review articles and review papers published on Halal meat authentication and Halal meat adulteration and detection using Machine Learning (Supervised and Unsupervised) techniques from 2018 to 2024. The current review is aimed at highlighting the principles, challenges, successes, and prospects of Machine Learning approaches in the authentication of Halal meat within the timeframe (see Table 2).

Search for related research articles and review papers was conducted on numerous scientific databases such as Google Scholar, Scopus, and PubMed. Using the words; "ML", "Supervised" "Unsupervised" "Halal meat" "Authentication" "Adulteration". The study focused on articles that were published in the English language, related to the research topic, methodologies, and key findings within the publication period of 2018–2024.

2. Machine learning (ML)

ML is the concept of learning through the development and application of algorithms from given data [44]. [45] also defined Machine Learning is a process of developing a model in a computer that automatically learns and improves with experience, there are many approaches to Machine Learning. In recent years ML has been used in the authentication and detection of fraud in agricultural products and yielding very positive results. ML techniques are broadly classified into three categories based on Learning Paradigms; supervised Learning, Unsupervised Learning and reinforcement Learning [46].

2.1. *Supervised learning*

Supervised Learning: Supervised Learning uses a labeled dataset for grouping into categories, in other words, supervised learning

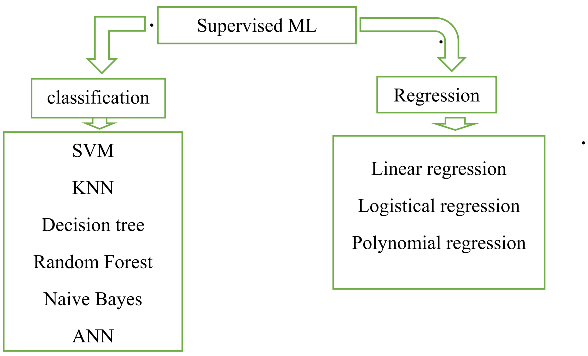

Fig. 1. Classification of supervised machine learning.

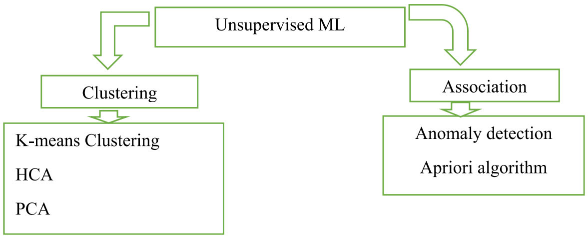

Fig. 2. Classification of unsupervised machine learning.

Table 1
Various techniques used in Halal meat authentication.

Method	Marker	Advantage	Limitation	References
Mass Spectrometry	enzymes	✓Early detection adulteration in raw and processed meat. ✓Increased sensitivity ✓Robust	✓Tedious and time-consuming. ✓Required trained personnel. ✓Very expensive	[24,25]
Stable isotope analysis	Ratio protein fat or organic compound	✓Small samples are required ✓Very high sensitivity	✓Expensive ✓Time consuming	[26]
Physio chemical methods	dielectric	✓Easy to use ✓Less expensive ✓simple	✓Time consuming ✓difficult to interpret	[27]
Near-infrared spectroscopy	Molecular composition of meat	✓non-destructive ✓non-invasive	✓Difficult in calibration ✓Time consuming ✓Low specification	[28]
Polymerase chain reaction (PCR)	DNA	✓Fast ✓High accuracy ✓Low detection limit	✓Expensive ✓Low accuracy for heat treat sample ✓Contamination may affect the result	[29,30,31]
Heat stable peptide	protein	✓Fast ✓Small samples are required ✓High sensitivity	✓Expensive ✓Laborious ✓Required skill	[32]
Hyperspectral imaging/ multispectral imaging	Meat colour	✓Non-destructive ✓Suitable for both raw and processed meat	✓Very complicated ✓Required large volume of data	[33]

Table 2
Supervised Machine Learning Approaches in authentication of meat.

Reference	Purpose	Model	Accuracy
[34]	classification of adulterated pork in beef and chicken meat	KNN	98.33 %
[35]	Minced chicken meat adulterated with pork	SVM	97.78 %
[36]	Detection of beef adulterated with pork meat	ANN	91.27 %
[37]	Determination of minced beef adulteration	Random Forest	87.78 %
[38]	Adulteration detection of minced meat	Decision Trees	95. 33 %
[39]	Classification of beef and pork meat	Naive Bayes	75 %

Table 3
Unsupervised Machine Learning Approaches in authentication of meat.

Reference	Purpose	Model	Accuracy
[40]	Detection of adulteration in beef meatballs containing pork meat	D-ELM and PCA	99.97 %
[41]	Adulteration of dog meat in beef meatballs.	PCA	Acceptable statistical results
[42]	Classification between Halal gelatin and non-Halal gelatin	HCA, PCA, and PLS-DA	100 %
[43]	Classification of lard from other animal fats	HCA, PCA and CA	98 %

uses the input variable (x-data) and the corresponding output (y-data) to train a learning algorithm to predict the relationship between the input and output (P(Y|X) [47]. This type of technique is commonly used to solve classification and regression problems, as it takes into consideration the actual knowledge of the output target [48].

The supervised learning (SL) approach which is also known as supervised Machine Learning uses a labeled training dataset to teach a model to produce a desirable output [49]. This process would allow the machine to learn over time and make an accurate output. Based on their functions, the SL can be grouped into (1) Classification and (2) Regression. Classification tasks are the most recurring tasks performed by the supervised learning approaches [50].

A process that automatically classifies several features into one of several predefined categories [51]. It deals with issues such as fraud detection, digit identification, disease diagnosis, face recognition, object recognition, and image classification [18,52] and these are a few examples of tasks where the model learns from the labeled data and predict the category or class of new output. It identifies specific features within the test data and makes a meaningful conclusion.

The most commonly used supervised learning approaches for classification and regression purposes are Artificial Neural Networks (ANN), Decision Trees, Forests, K-Nearest Neighbor (KNN), Random Naïve Bayes, and Support Vector Machine (SVM). These methods perform excellently in the authentication of Halal meat and record various degrees of accuracy.

Regression is used in showing a relationship between two variables i.e., dependent and independent variables. Linear regression, logistical regression, and polynomial regression are the most common regression algorithms.

2.2. Unsupervised learning

In unsupervised Learning the inputs do not have a complete and clean labeled dataset, the structure and grouping do not have any prior knowledge, and it works on cluster analysis [53]. Unlike supervised learning, unsupervised learning uses unlabeled input X value only and learns to make predictions (P(X)), find underlying patterns and predict output accordingly [47]. The unlabeled input in this technique is asked to look out for hidden features and cluster the data based on their similarity [54].

2.3. Reinforcement learning

Reinforcement learning is the type of learning that combines the input X with an acting phase (critic (C)) to simultaneously learn and achieve a self-optimizing feature [48]. The learning algorithm interacts with the perspective of the feedback mechanism and improves its performance-based goal explicitly [55]. The learning algorithm is not synchronized but would find an activity that yields the best result, by attempting several activities in a steady progression [56].

2.3.1. Decision tree

Decision Tree is one of the most commonly supervised learning approaches used for classification [57]. The classification procedures in this method represent a form of a tree-like from roots to leaves, with three nodes; root nodes, leaf nodes, and internal nodes [58], the internal nodes signify a test on a feature, and each branch represents the result of a test, and each leaf node provides final classification accuracy and denote class label followed by the performance of decision after reflecting all the features [59].

The root node also denotes the beginning of the decision-support process [60,61]. Decision Trees consist of sets of data that consist of trait vectors, which in turn contain a set of classified traits describing the vector and a class quality assigning the data entry to a specified class. Because of the tree-like structure and quick overview Decision tree can easily be visualized [62].

A Decision Tree is built iteratively which breaks the data set on its characteristics and separates the data into different existing classes until a certain specification is reached [63]. The common Decision tree algorithms development is Iterative Dichotomizer 3 ID3, C4.5, C5 [64] and Breiman's Classification and Regression Tree [65]. The decision tree algorithm is increasingly gaining popularity in the field of animal science including meat science [66,67]. The feature space is recurrently divided into rectangular subregions using Decision trees for classification, where the predicted class is the most prevalent. A tree algorithm runs through each iteration's potential split points for each feature to identify a division [68].

The merit attached to the use of decision trees in the classification process is the non-requirement of the creation of dummy variables by the algorithms [20]. [69] used a decision trees classifier for coffee classification based on civet and non-civet and recorded a 97 % accuracy level.

Decision Tree Algorithms can handle a variety of data types including numeric, categorical and ratings data. They can also manage missing data in responses and independent variables which are mainly used in classification problems [70].

The fact that the working process of the decision tree algorithm is similar to human thinking, made the model more acceptable in classification issues of agricultural products. Decision Tree is the most common model in classification and prediction. Several studies involved the use of a decision tree as a classifier indicating a >90 % classification accuracy rate which shows the robustness of the model [20].

[71] combined Fourier Transform Infrared Spectroscopy (FTIR) with multivariate classification methods in the classification of minced meat from different species of animals (Beef, Lamb, Chicken, and Pork), a Decision Tree was employed to optimize two different situations, hundred per cent correct identification of pork meat and false-positive and false-negative rate balance tuned. According to Ref. [72], in authentication of meat with an increased number of Decision Tree models would result in better generalization and also prevent overfitting. The drawback of this model is the problem of the large growth of the trees resulting in one leaf per observation [73], and it is impossible to review a decision once the training data set has been divided for solving a problem [74].

2.3.2. Random forest

A random forest can be referred to as an advanced version of a decision tree, which consists of multiple trees that are used for classification and regression tasks, it makes use of decision trees with a randomly chosen portion of training data and replacement. In each set, a random selection of features is made from the collection of features. The tree-growing procedure is then repeated until the set of the classifier with the highest classification accuracy is produced [75]. Each grown tree makes a forecast about its target class at the time of prediction, just like the decision tree does.

The classifier proposed the class that the decision trees predict with the highest accuracy. In other words, random forest uses the mean of decision trees with significant individual variance that can be combined to create a more reliable model and is less prone to overfitting [76]. Reducing the quantity of the bootstrap samples, which increases the randomness of the random forest and, in turn, lessens the issues of overfitting, will however, have unfavourable effects on performance [77]. For a satisfactory exchange between bias and variance, the bootstrap sample size should ideally match the number of samples in the original training data set [78]. The overfitting of the training data that would be on the single decision tree will be corrected while a random forest will be used to make the selection of the decision class [79].

The approach involved the collection of uncorrelated decision trees and then merging them to reduce variation and develop more precise data prediction. The creation of trees will be used to form a forest, a decision class will be selected on the basis that has been developed by the larger number of trees in the forest [80].

Random forest is now gaining much more popularity in different fields of animal science for the prediction of animal weight, animal feed intake, animal waste management, and other aspects of animal production ([81,82,83,84,85] due to its flexibility and accuracy in classification and regression compared to other algorithms [86]. Yet the application of the model in the area of food authentication is scarce [37]. assessed the effectiveness of hyperspectral reflectance spectroscopy with multivariate methods for the detection of adulteration of minced beef, samples were prepared at different adulteration levels of thirty samples as pure minced beef and ninety samples of adulterated beef. A Random Forest model was used to detect the adulteration, to increase resilience, the best wavelengths were chosen using the successive projection method (SPA). The authors concluded that RF performed better in prediction with an accuracy rate of 96.87 %.

In another study [87], used complementing spectrum data to enhance the classification accuracy of beef, mutton, and pork meat tissues based on combined LIBS (laser-induced breakdown spectroscopy) and Raman spectroscopy, including three choices of LIBS, Raman, and LIBS-Raman. The three different types of beef tissues were classified using a BPNN (back propagation neural network) with input variables optimized using RF (random forest). The three classification models were assessed using the 10-fold cross-validation method. The outcomes of the three approaches were then compared. The combined LIBS-Raman model offers the highest classification accuracy, up to 99.42 %.

Some of the challenges of the model include difficulty in interpretation as compared to other algorithms [88], and the accuracy of the model is affected when the number of trees is not enough [89].

2.3.3. K-nearest neighbor (K-NN)

K-nearest neighbor (K-NN) is a non-parametric technique that is used to classify a data point according to how closely it resembles the data that is already available [68]. The classifier algorithm is built on the closest training instances in the feature space since the operation is under the assumption that comparable data points are relatively close to one another. The distance between the data points is often measured using the Euclidean distance, and the most common category is then assigned [68].

K-NN is involved in gathering all available examples. Then classifying new cases based on similarities. The algorithm looks for the k samples in the training dataset that are closest to the point to be categorized based on the distance measure that is selected [90]. For achieving a fair balance between underfitting and overfitting, the optimal value of k is crucial. If k is too large or too small, the neighborhood may contain points from other classes. If k is too tiny, it will be more prone to noise points [91].

The main advantages of K-NN its high accuracy with low calculation time required, it requires zero cost in the learning process, no optimization is required and it's very easy to use, however, when the dataset is large the calculation time will be lengthened, making it less appealing for classification issues [92,93].

The K-NN algorithm is commonly used in engineering and pattern recognition. Because of its simplicity and high accuracy, the algorithm is now applied in different areas of Agriculture [94,95]. [34], used KNN in the classification of beef adulterated with pork meat based on E-nose data and the result obtained was satisfactory. The research work compared five different classification methods (K-nearest neighbor (KNN), Logistic regression (LR), SVM, LDA, and Naive Bayes) to distinguish between beef, pork, and pork adulteration in beef, the result indicated that KNN produces the best classification result with 98.33 % accuracy.

An experiment was conducted using 4 Metal Oxide Sensors (MOS) and gas sensors to differentiate meat samples (beef and pork) based on their odour, the odour profile was produced by data extracted using a mean feature, K-Nearest Neighbors (KNN) was used for the classification and recorded the highest classification performance of 99.24 % [96]. In another study. k-nearest neighbor and support vector machine as classifiers together with hyperspectral imaging technique with multivariate analyses were used in chicken breed fraud detection. Images of breast meat of four different breeds were taken with a near-infrared range between 900 and 1700 nm, different pre-treatments were individually done for spectra extraction, and the result showed that the models recorded a classification accuracy rate of 98 % [97].

[98] combined near-infrared spectroscopy and electronic nose for the detection of different adulteration lamb meat with duck meat, used KNN, RF, SVM, and BPNN as classifiers, for effective analyses of the data the F1-score-MRE (F1-score-based Model Reliability Estimation) was used, the result of the study showed that KNN (85.21 %), RF (81.94 %), SVM (84.51 %), BPNN 88.73 % and SSA-BPNN (94.36) were recorded. The method is one of the lazy learning methods, which requires no computation to be performed on the data before a query is given to the system. This method is different from eager learning methods such as Decision Trees, which try to structure the data before receiving queries. Tuning parameter k is estimated using cross-validation of the training set.

2.3.4. Support vector machine (SVM)

Support Vector Machine is another supervised learning model developed and used for both data classification and regression, it is known for its excellent classification power with the hyperplane which is also known as the decision boundary that separates the two classes of data points [99].

The algorithm is used to determine the best decision boundary between vectors that belong to different categories. Vectors are lists of numbers that represent a set of coordinates in some space [100]. When SVM determines the hyperplane, it indicates where to draw the lines with optimum deviation, into two subspaces, which are also referred to as categories [101].

SVM works with the principle of classifying data, by creating a task that splits the data points into two categories with the corresponding labels (a) with the least possible number of errors or (b) with the largest possible margin. The larger space next to the splitting task results in fewer errors, due to the labels that are better distinguished from one another [62]. Due to the larger margin created and linear separation by the algorithm, the classification is more precise. SVM is one of the robust techniques used in the detection of meat adulteration in its products in recent times [102].

The SVM model has been used for the authentication of meat in several studies and has recorded remarkable successes [103]. reported that SVM was used in the authentication of Halal meat and the accuracy was magnificent because the model was not only designed for classification. [104,105], but can also perform regression through the Support Vector Regression (SVR) channel. The model has been widely used in classification and is compatible with many techniques. [106], in their study that used SVM in the classification of pure horse, pure beef meat and adulterated levels at 0. 60 80 and 100 in freshly ground and stored meat samples, the result indicated the model performance based on OCC (One Class Classification) clearly distinguished with 95 % accuracy rate, only one sample was misclassified.

[35]reported that multispectral images and SVM were used for the classification of chicken and pork meat under different cases and stimulation of adulteration, the result indicated a 95.31 % accuracy level [107]. reported Fourier transform infrared spectroscopy (FTIR) combined with multivariate methods for the classification of minced beef, lamb, and chicken adulterated with pork meat, SVM and partial least squares-discriminant analysis with radial basis function (FBF) were used for the evaluation of meat speciation, the findings indicated that SVM showed highest accuracy rate of 98 % in classification performance.

[108], developed an optimized electronic nose system (OENS) to detect different adulteration levels of 10 %, 25 %, 50 %, 75 %, and 90 % of pork in beef, the optimization of algorithms indicated that the SVM classifier recorded the best classification result with an accuracy level of 98.10 %. Furthermore [109], mixed lard with beef, lamb, and chicken in a ratio of 10–50 % v/v to obtain an adulterated sample, the samples are labeled as Pure and Adulterated. For pure pork, functional groups were found using absorbance values from the FTIR spectrum, two regions of difference (RoD) at wavenumbers of 1700–1800 cm^{-1} and 2800–3000 cm^{-1} were branded for the samples, multi-support vector machine (M-SVM) was used for the classification of the samples into pure and adulterated cluster, it was concluded that M-SVM model can be used for rapid and accurate classification adulteration of lard.

2.3.5. Artificial neural network (ANN)

Artificial Neural Networks (ANNs) model was developed to mimic how the human brain works based on the functioning principle of biological neurons [110]. The algorithm contains several interconnected neurons with weights, thresholds, and an activation function [111]. Perceptron which is a mathematical model of a biological neuron is composed of three layers; an input layer, a hidden layer, and an output layer [112], in each layer certain number of neutrons are present which increases the performance accuracy [113], and also increases the computation load [114]. These multiple neurons are connected to one another by weighted links in a complex and non-linear manner [115]. The classification model has been used in different fields of study and recorded numerous successes due to its high classification accuracy, ability to deal with complex relationships, robustness, automation, and simplicity [116,117,118]. A smart electronic nose (SE-nose) was developed and used for qualitative and quantitative adulteration detection of pork in beef meat using ANN-C and SVM-C as classification models. The authors used sample slicing window protocol, pattern recognition, normalization, and output block as datasets, these datasets were used for validation of the SE-nose, and the first and second datasets were used for classification and regression purposes, the result of the study showed that ANN-C recorded high performance with a classification accuracy of 99.996 % [119]. [71] assessed the authentication of (pork in beef, pork in lamb, and pork in chicken) meat with the use of a combined shortwave infrared hyperspectral imaging (SWIR-HSI) (1116–1670 nm) and visible-near infrared hyperspectral imaging (Vis-NIR-HSI) (400–1000 nm) which are HIS camera employed for animal speciation and adulteration detection, multivariate methods were used with SVM and ANN-BP (backward propagation) the results reveal that the used of these models performed better in meat adulteration detection [120]. used also ANN as a classifier to distinguish lard adulteration from the fats of other animals based on their dielectric spectra which were measured in 100 Hz–100 kHz at 45 °C–55 °C, gas chromatography-mass spectrometry (GCMS) was used to study the fatty acid composition of the fats, the finding of the study showed that ANN model recorded an accuracy level of 85 % in the classification of the fats. ANN has been widely used in the detection and authentication of meat in many studies [101,121,22,122,42,123,124].

One of the challenges of Artificial Neural Networks is that the data must be transformed into binary as the algorithm does not allow the use of discrete variables, which is also characterized by a high risk of overfitting [125,126,127]. To achieve optimum accuracy, ANNs require a lot of computation [114], it is often referred to as the black box by researchers due to a lack of understanding of some of the decisions made [46].

2.3.6. Naive Bayes

Naïve Bayes is another simple and powerful supervised Machine Learning classification model that uses the concept of conditional probability (Bayes' theorem) to identify the result of related tasks [128]. The model operates under the presumption that each trait is independent of others and is equally distributed among them, showing that the value of one feature is unrelated to the value of other features [20]. The algorithm is based on Bayes' Theorem which describes the probability of an event happening based on previous knowledge that an event has occurred and makes the probability of each feature belonging to a class to make a prediction [129].

Naïve Bayes has three models; The multinomial model, the Bernoulli model, and the Poisson model most studies indicate the use of the Multinomial in classification [130]. The model is mostly used in text classification such as spam detection and is also used in the detection of adulteration in agricultural produce with high accuracy [131]. An electronic nose system was used to distinguish between fresh pork and beef meat, eight Metal Oxide Semi-conductor gas sensors and an Arduino micro-controller array were used to create the electronic system, which identified fresh beef and pork, a Naïve Bayes classifier was employed, followed by min-max magnitude scaling. Based on k-fold cross-validation, the results show that the system could classify beef and pork with 75 % classification accuracy [39]. [132] employed a cross-sensitive sensor array in an electronic taste system to deliver a worldwide liquid and taste perception for meat's soluble flavour components. Sensors' taste reactions to tainted mutton were recorded, and multivariate data processing techniques were used to examine the results.

Bayes discriminant analysis and canonical discriminant analysis (CDA) are used to discriminate between different meat species and content. The result showed that CDA and BDA can classify and predict chicken with pork adulteration with optimal precision. The model, apart from its simplicity, has high speed in classification [133]. However, the challenge of Naïve Bayes is its inability to relate between two predictor features because of the assumption of conditional independence [134].

2.4. Unsupervised Machine Learning approaches

The unsupervised ML technique uses unlabeled input and asks it to look out for hidden features and cluster the data based on their similarity [54]. Unsupervised learning leads to further grouping as clustering and association. Clustering and association are not the only grouping techniques in unsupervised Machine Learning, but they are the most commonly applied techniques with various applications. In clustering, the objects are divided into clusters based on similar and dissimilarity features, the clustering approach aims to cluster or group the data without any prior knowledge (label). One of the major challenges in this approach is that it requires expert domain knowledge to determine the optimal number of clusters or interpretation of the obtained clusters. Whereas association identifies relationships, patterns, or associations among the dataset or attributes, it aims to discover the probability of co-occurrence of items in a collection.

The most widely used unsupervised Machine Learning approaches are K-means Clustering, Hierarchical Cluster Analysis (HCA), Anomaly detection, Principal Component Analysis (PCA), Independent Component Analysis and Apriori algorithm [135,136,137]. Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) are the most widely used in food detection and authentication.

2.4.1. K-means clustering

The k-means clustering algorithm is a clustering approach in unsupervised Machine Learning that partitions the data into k clusters based on the high and low similarity between intra and inter-clusters [138]. k-means algorithm involves organizing data into a pre-defined number of clusters based on their similarity of features. An item can belong to only one cluster since it produces a definite number of non-hierarchical and disjoint clusters.

The process in the K-means model starts with the random selection of picked centroids which is followed by a designated k symbol. Centroids is the average arithmetic mean of all the points. The distances between the points and the centroid are calculated using Euclidean metric [139]. The algorithm then assigns each data point to the nearest cluster of centroids and then apprises the centroids based on the new assignments. This continues until convergence of the cluster is achieved. K-mean is simple and efficient in computation, easy in interpretation and scales with large datasets. These made the model ahead of other algorithms for classification purposes [140]. used k-means clustering for the classification of pork samples using hyperspectral data and recorded a significant success. The major challenge of the K-mean approach is its sensitivity to outliers, it requires a specified number of clusters (k) in advance.

2.4.2. Hierarchical Cluster Analysis (HCA)

HCA is an unsupervised Machine Learning approach that creates a hierarchy of clusters either by agglomerative (bottom to top) or divisive (top to down) approach. HCA generates a graph structure known as a dendrogram from the iterative coupling of clusters according to similarity and grouping criteria. The HCA algorithm assesses and determines the highest similarity and dissimilarity within and between classes, then followed by a clustering process which could be either distance, scale, sample, linkage method, or variable [141].

According to Ref. [142], HCA is one of the widely used classifications in the field of sciences. HCA and PCA are chemometric techniques applied in the detection of non-Halal components such as lard and gelatin in food products [42]. have used FTIR spectroscopy combined with chemometrics of principal component analysis (PCA) and cluster analysis (CA) for the confirmation of lard and other edible fats and oils. Lard, obtained from rendering adipose tissues of pigs [143]. also applied PCA and HCA with combined techniques of liquid Chromatography and tandem mass spectrometer to detect bile acids, sterols, and acylcarnitine from humans, mice, and pigs [144]. reported the use of a portable Raman spectrometer combined with multiple chemometrics of HCA and others to detect lard fat adulteration with other fats, the concentration range of 0%–100 % (w/w) and different adulteration levels of lard fat content were used, HCA dendrogram showed a detailed categorization pattern, while 3-D score plots of PCA analysis showed the similarities and contrasts of multivariate data. The findings of the study suggested the techniques could be employed by industry in the detection of unwanted materials in foods.

The merit of this approach is that it does not need several clusters in advance, it can produce interpretable dendrograms. The major drawback of HCA is its computation is expensive for large datasets and it is sensitive to noise and outliers.

2.4.3. Independent Component Analysis (ICA)

ICA is an advanced Principal Component Analysis (PCA), which is a powerful dimensional reduction technique useful for mining information or source signals from the original data [20]. ICA is a statistical signal processing technique that aims to find a statistical representation of the data, where the components are maximally independent. It separates the observed data into statistically independent subcomponents [145]. ICA computes linear transformation of the observed data and creates a set of statistically independent sources.

The advantages of this algorithm are its usefulness in finding latent variables or blind source separation, it can capture hidden factors leading to the observed data, rather than just the correlated factor. The algorithm works with the assumption that the data is linearly a combination of those independent signals. By computing the real independent components, the model can distinguish different sources mixed, the same principle is applied when detecting meat adulteration or non-Halal components. ICA has been used by some researchers to identify and distinguish between the various constituents present in a meat sample [146] applied ICA in the authentication of Halal meat using an electronic nose through the detection of compounds emitted by the samples. The application of ICA for Halal meat authentication is limited. The drawback to this model is that it requires a sufficient amount of data for an accurate estimation, limitation application in non-linear mixing, and sensitivity to outliers.

2.4.4. Anomaly detection

The anomaly detection technique sometimes referred to as the Gaussian Mixture Model (GMM) is a model that is usually used to identify observations in a dataset that deviate significantly from the expected normal behaviour ([147]. The aim of the model is to detect a sample that deviates significantly from the expected pattern, the algorithm studies various characteristics or features of a sample (meat) to form a normal distribution, any deviation from the normal is considered as an anomaly. It estimates indices such as weight, mean and covariance. GMM is flexible when dealing with complex data distribution, it can detect abnormalities in numerical and categorical data. Anomaly detection techniques can be used to detect fraud in the meat industry such as adulteration of non-Halal meat.

The model was used for the authentication of Halal meat using near-infrared spectroscopy, in the study Mahalanobis distance to detect non-Halal meat samples. The major challenge of using this model for Halal authentication is the lack of comprehensive and standardized datasets for Halal products and for effective detection, sufficient and diverse datasets are required for training, validation, and testing ([17,148].

2.4.5. Principal component analysis (PCA)

PCA is an unsupervised, non-linear statistical approach, it significantly reduces dimensionality and discovers a set of orthogonal components called Principal Components (PC). The model linearly reduces the number of variables in the original data into PC, that contain most of the variable data for simplification and identification of significant variation in the data. The most widely used multivariate statistical method is probably PCA, which has been used in almost all scientific fields [149].

The algorithm is widely used for the identification and confirmation of pigs and other non-Halal derivatives in meats and other consumable products [150]. PCA is an exploratory analysis tool for data interpretation, although some studies used it as a classification model [149,151], which is a scientific blunder that still exists [152]. PCA has greatly increased the ML classifier accuracy in classification, for that, it is considered the robust dimensionality reduction technique [153]. [42] combined PCA with chemometrics in the detection of adulteration in beef meatballs by differentiating pure beef meatballs and beef meatballs containing pork.

[40] combined Deep limit Learning Machines (D-ELM) and PCA for the detection of pork adulteration in beef and recorded an accuracy rate of 99.97 %. Several studies were conducted on the use of PCA for Halal meat authentication [154,155,156,157,158,159,146,108,25,160,143].

3. Overview

ML has the potential of revolutionizing Halal meat authentication compared to the traditional methods, by holistic detection of fraudulent or adulterated meat that does not meet the Halal standard. ML approaches have been applied to address issues related to meat fraud in the meat industry, the techniques have performed tremendously in the last decade. ML is broadly categorized into supervised and unsupervised ML.

The supervised ML approach required labeled data (Halal or non-Halal) samples for the classification of the dataset. The various types of SL used for classification purposes include ANN, DT, RF, SVM, and KNN. The principle behind these algorithms is they learn and use patterns and relationships from the labeled datasets to produce new unlabeled samples. Supervised ML has yielded positive results with a high accuracy rate.

The unsupervised ML on the other hand requires unlabeled datasets for its classification, the algorithm is designed to discover hidden patterns and structures in the sample (Halal or non-Halal). The algorithm also used similarities and dissimilarities features among the Halal and non-Halal meat samples for classification. The types of unsupervised ML approaches are K-mean, clustering analysis, and anomaly detection.

The research gap that exists in this study is the need for standardized and validated datasets for Halal meat authentication.

4. Challenges of ML application in halal meat authentication

The availability of representative and reliable standard datasets is a major challenge to the Halal meat authentication process, which is a very important aspect in model training and the accuracy of the machine, researchers are using different datasets in their various studies leading to variability. There is a need for an accurate, reliable, and robust that can handle different scenarios and data variations.

Another challenge is the quality of data required for authentication which depends on the ML approach used, example SVM requires a sufficient amount of data for accuracy authentication but if the data is large, it would affect the accuracy level of the model, while ANN requires very large quality of data to perform accurately.

5. Conclusion and future prospective

It is apparent that consumers are very much aware of what they eat, and this has continued to increase the demand for Halal meat globally, it becomes paramount to ensure that approaches for Halal meat authentication are fast, scale with large datasets, efficiently compute data with high accuracy, ML is the approach that fit to do that.

Machine Learning approaches such as Artificial Neural Networks (ANN), Decision trees, Forests, K-nearest neighbor (KNN), Random Naïve Bayes, Support Vector Machine (SVM), K-means Clustering, Hierarchical Cluster Analysis (HCA), Principal Component Analysis (PCA) and Independent Component Analysis have proven to be more efficient and effective in Halal meat authentication. The advancement in technology has pushed meat scientists to produce tissue meat or laboratory meat. In such a scenario, Machine Learning Machine Learning can be able to authenticate whether the tissue is from a Halal or non-Halal source by analysing various parameters such as composition, amino acid profile, and microbial content of the meat to determine its authenticity.

In the future, it's expected that supervised machine-learning approaches will continue to play a crucial role in Halal meat authentication. Technological advancement coupled with the integration of the availability of datasets will enable more robust, high accuracy and precision in Halal meat authentication. Furthermore, the use of Machine Learning techniques can potentially reduce energy, time, and cost in Halal meat authentication, making it more accessible to consumers.

Data availability

The data that support the findings of this study are available on request from the corresponding author.

CRediT authorship contribution statement

Abdul Mustapha: Writing – original draft. **Iskandar Ishak:** Writing – review & editing. **Nor Nadiha Mohd Zaki:** Writing – review & editing. **Mohammad Rashedi Ismail-Fitry:** Writing – review & editing. **Syariena Arshad:** Writing – review & editing. **Awis Qurni Sazili:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S.A. Rabia, I. Ali, B.H. Mohammed, Meat Science and Nutrition, 2018, https://doi.org/10.5772/intechopen.77045.
[2] T.A. Shaikh, W.A. Mir, T. Rasool, S. Sofi, Machine learning for smart agriculture and precision farming: towards making the fields talk, Arch. Comput. Methods Eng. 29 (7) (2022) 4557–4597, https://doi.org/10.1007/s11831-022-09761-4.
[3] M. Zheng, P. Mao, X. Tian, Q. Guo, L. Meng, Effects of Dietary Supplementation of alfalfa meal on growth performance, carcass characteristics, meat and egg quality, and intestinal microbiota in Beijing chicken, J. Poultry Sci. 98 (5) (2019) 2250–2259, https://doi.org/10.3382/ps/pey550.
[4] A. Benzertiha, B.A.R.T.O.S.Z. Kierończyk, M. Rawski, A. Jozefiak, J. Mazurkiewicz, D. Jozefiak, S.Y.L.W.E.S.T.E.R. Świątkiewicz, Cultural and practical aspects of Halal slaughtering in food production, Med. Weter. 74 (6) (2018) 371–376, https://doi.org/10.21521/mw.6023.
[5] B.O. Alao, A.B. Falowo, A. Chulayo, V. Muchenje, The potential of animal by-products in food systems: production, prospects, and challenges, Sustainability 9 (7) (2017) 1089, https://doi.org/10.3390/su9071089.
[6] Z. Dachuan, O. Shuyu, C. Minqing, Z. Haoyang, D. Shaozhen, L. Dongliang, C. Pengli, L. Yingying, Hu Qian-Nan, FADB-China: a molecular-level food adulteration database in China based on molecular fingerprints and similarity algorithms prediction expansion, Food Chem. 327 (2020) 127010, https://doi.org/10.1016/j.foodchem.2020.127010. ISSN 0308-8146.
[7] A.N.M. Alamgir, A.N.M. Alamgir, Bioactive compounds and pharmaceutical excipients derived from animals, marine organisms, microorganisms, minerals, synthesized compounds, and pharmaceutical drugs. Therapeutic Use of Medicinal Plants and their Extracts, Phytochemistry and Bioactive Compounds 2 (2018) 311–406, https://doi.org/10.1007/978-3-319-92387-1_4.
[8] N. Tarulevicz, Discursively globalized: Singapore and food safety, Food Cult. Soc. 23 (2) (2020) 193–208, https://doi.org/10.1080/15528014.2019.1682890.
[9] A.K. Das, P.K. Nanda, A. Das, S. Biswas, Hazards and safety issues of meat and meat products, in: Food Safety and Human Health, Academic Press, 2019, pp. 145–168, https://doi.org/10.1016/B978-0-12-816333-7.00006-0.
[10] C. Brooks, L. Parr, J.M. Smith, D. Buchanan, D. Snioch, E. Hebishy, A review of food fraud and food authenticity across the food supply chain, with an examination of the impact of the COVID-19 pandemic and Brexit on food industry, Food Control 130 (2021) 108171, https://doi.org/10.1016/j.foodcont.2021.108171.

[11] A.Y. Khaled, C.A. Parrish, A. Adedeji, Emerging nondestructive approaches for meat quality and safety evaluation—a review, Compr. Rev. Food Sci. Food Saf. 20 (4) (2021) 3438–3463, https://doi.org/10.1111/1541-4337.12781.
[12] A.M. Rady, A. Adedeji, N.J. Watson, Feasibility of utilizing color imaging and machine learning for adulteration detection in minced meat, Journal of Agriculture and Food Research 6 (2021) 100251, https://doi.org/10.1016/j.jafr.2021.100251.
[13] S.F. Sim, M.X.L.Jeffrey Chai, A.L. Kimura, Prediction of lard in palm olein oil using simple linear regression (SLR), multiple linear regression (MLR), and partial least squares regression (PLSR) based on Fourier-transform infrared (FTIR), J. Chem. (2018) 1–8, https://doi.org/10.1155/2018/7182801.
[14] H.T. Zhao, Y.Z. Feng, W. Chen, G.F. Jia, Application of invasive weed optimization and least square support vector machine for prediction of beef adulteration with spoiled beef based on visible near-infrared (Vis-NIR) hyperspectral imaging, Meat Sci. 151 (2019) 75–81, https://doi.org/10.1016/j.meatsci.2019.01.010.
[15] F.W. Harun, Fourier transform infrared spectroscopy as a technique for multivariate analysis of lard adulteration in food products: a review, J Fatwa Manag. Res 17 (July 2019) (2019) 1–13, https://doi.org/10.33102/jfatwa.vol17no1.
[16] V. Bischoff, K. Farias, J.P. Menzen, G. Pessin, Technological support for detection and prediction of plant diseases: a systematic mapping study, Comput. Electron. Agric. 181 (2021) 105922, https://doi.org/10.1016/j.compag.2020.105922.
[17] P.C. Ng, N.A.S. Ahmad Ruslan, L.X. Chin, M. Ahmad, S. Abu Hanifah, Z. Abdullah, S.M. Khor, Recent advances in halal food authentication: challenges and strategies, J. Food Sci. 87 (1) (2022) 8–35, https://doi.org/10.1111/1750-3841.15998.
[18] J. Alzubi, A. Nayyar, A. Kumar, Machine learning from theory to algorithms: an overview, in: Journal of Physics: Conference Series, vol. 1142, IOP Publishing, 2018, November 012012, https://doi.org/10.1088/1742-6596/1142/1/012012.
[19] I. El Naqa, M.J. Murphy, What Is Machine Learning? Springer International Publishing, 2015, pp. 3–11, https://doi.org/10.1007/978-3-319-18305-3_1.
[20] D. Saha, A. Manickavasagan, Machine learning techniques for analysis of hyperspectral images to determine the quality of food products: a review, Curr. Res. Food Sci. 4 (2021) 28–44, https://doi.org/10.1016/j.crfs.2021.01.002.
[21] M. Imdoukh, I. Ahmad, M.G. Alfailakawi, Machine learning-based auto-scaling for containerized applications, Neural Comput. Appl. 32 (13) (2020) 9745–9760.
[22] S. Othman, N.R. Mavani, M.A. Hussain, N. Abd Rahman, J.M. Ali, Artificial intelligence-based techniques for adulteration and defect detections in food and agricultural industry: a review, Journal of Agriculture and Food Research (2023) 100590.
[23] J.P. Bharadiya, N.T. Tzenios, M. Reddy, Forecasting of crop yield using remote sensing data, agrarian factors and machine learning approaches, Journal of Engineering Research and Reports 24 (12) (2023) 29–44.
[24] M. Mortas, N. Awad, H. Ayvaz, Adulteration detection technologies used for halal/kosher food products: an overview, Discover Food 2 (1) (2022) 15.
[25] Suratno, A. Windarsih, H.D. Warmiko, Y. Khasanah, A.W. Indrianingsih, A. Rohman, Metabolomics and proteomics approach using LC-Orbitrap HRMS for the detection of pork in tuna meat for halal authentication, Food Anal. Methods 16 (5) (2023) 867–877.
[26] C. Li, X. Kang, J. Nie, A. Li, M.A. Farag, C. Liu, Y. Yuan, Recent advances in Chinese food authentication and origin verification using isotope ratio mass spectrometry, Food Chem. 398 (2023) 133896.
[27] I. Usman, S. Sana, M. Afzaal, A. Imran, F. Saeed, A. Ahmed, M.R. Khan, Advances and challenges in conventional and modern techniques for halal food authentication: a review, Food Sci. Nutr. 12 (3) (2024) 1430–1443.
[28] J.T.L. Müller-Maatsch, Y.J.A. Weesepoel, E.A.M. Roetgerink, A.M. Wijtten, M. Alewijn, Are low-cost, hand-held NIR sensors suitable to detect adulterations of halal meat? Proceedings of the OCM 2021—Optical Characterization of Materials (2021) 1–10.
[29] P. Qin, W. Qu, J. Xu, D. Qiao, L. Yao, F. Xue, W. Chen, A sensitive multiplex PCR protocol for simultaneous detection of chicken, duck, and pork in beef samples, J. Food Sci. Technol. 56 (2019) 1266–1274.
[30] N. Shahidan, A. Amid, Challenges faced by halal meat industry: a review, Halalpshere 3 (1) (2023) 55–63.
[31] S.M.K. Uddin, M.M. Hossain, S. Sagadevan, M. Al Amin, M.R. Johan, Halal and Kosher gelatin: applications as well as detection approaches with challenges and prospects, Food Biosci. 44 (2021) 101422.
[32] A. Stachniuk, A. Sumara, M. Montowska, E. Fornal, Liquid chromatography–mass spectrometry bottom-up proteomic methods in animal species analysis of processed meat for food authentication and the detection of adulterations, Mass Spectrom. Rev. 40 (1) (2021) 3–30.
[33] C.H. Feng, Y. Makino, S. Oshita, J.F.G. Martín, Hyperspectral imaging and multispectral imaging as the novel techniques for detecting defects in raw and processed meat products: current state-of-the-art research advances, Food Control 84 (2018) 165–176.
[34] M. Malikhah, R. Sarno, S.I. Sabilla, Ensemble learning for optimizing classification of pork adulteration in beef based on electronic nose dataset, International Journal of Intelligent Enginerring and Systems 14 (4) (2021), https://doi.org/10.22266/ijies2021.0831.05.
[35] L.C. Fengou, P. Tsakanikas, G.J.E. Nychas, Rapid detection of minced pork and chicken adulteration in fresh, stored, and cooked ground meat, Food Control 125 (2021), https://doi.org/10.1016/j.foodcont.2021.108002.
[36] F. Han, X. Huang, J. H. Aheto, D. Zhang, F. Feng, Detection of beef adulterated with pork using a low-cost electronic nose based on colorimetric sensors, Foods 9 (2) (2020) 193, https://doi.org/10.3390/foods9020193.
[37] B. Guo, J. Zhao, S. Weng, X. Yin, P. Tang, Rapid determination of minced beef adulteration using hyperspectral reflectance spectroscopy and multivariate methods, IOP Conf. Ser. Earth Environ. Sci. 428 (1) (2020) 012049, https://doi.org/10.1088/1755-1315/428/1/012049. IOP Publishing.
[38] I.C. Setiadi, A.M. Hatta, S. Koentjoro, S. Stendafity, N.N. Azizah, W.Y. Wijaya, Adulteration detection in minced beef using a low-cost color imaging system coupled with a deep neural network, Front. Sustain. Food Syst. 6 (2022) 1073969, https://doi.org/10.3389/fsufs.2022.1073969.
[39] D.R. Wijaya, R. Sarno, A.F. Daiva, Electronic nose for classifying beef and pork using Naïve Bayes, in: 2017 International Seminar on Sensors, Instrumentation, Measurement and Metrology (ISSIMM), IEEE, 2017, August, pp. 104–108, https://doi.org/10.1109/ISSIMM.2017.8124272.
[40] J. Du, M. Gan, Z. Xie, C. Zhou, M. Li, M. Wang, L. Zhu, Current progress on meat food authenticity detection methods, Food Control 109842 (2023), https://doi.org/10.1016/j.foodcont.2023.109842.
[41] W.S. Rahayu, A. Rohman, S. Martono, S. Sudjadi, Application of FTIR spectroscopy and chemometrics for halal authentication of beef meatball adulterated with dog meat, Indonesian Journal of Chemistry 18 (2018) 376, https://doi.org/10.22146/ijc.27159.
[42] A. Rohman, A. Windarsih, The application of molecular spectroscopy in combination with chemometrics for halal authentication analysis: a review, Int. J. Mol. Sci. 21 (14) (2020) 5155, https://doi.org/10.3390/ijms21145155.
[43] N.A.S. Sapian, M.A.M. Roslan, A.M. Hashim, M.N.M. Desa, M. Halim, Y.N.A. Manaf, H. Wasoh, Differentiation of lard from other animal fats based on n-Alkane profiles using chemometric analysis, Food Res. Int. 164 (2023) 112332, https://doi.org/10.1016/j.foodres.2022.112332.
[44] M. Meenakshi, Machine learning algorithms and their real-life applications: a survey, in: Proceedings of the International Conference on Innovative Computing & Communications (ICICC), 2020, May.
[45] X. Wu, L. Xiao, Y. Sun, J. Zhang, T. Ma, L. He, A survey of human-in-the-loop for machine learning, Future Generat. Comput. Syst. 135 (2022) 364–381.
[46] I.H. Sarker, Machine learning: algorithms, real-world applications, and research directions, SN Computer Science 2 (3) (2021) 160, https://doi.org/10.1007/s42979-021-00815-1.
[47] K.M.M. Uddin, N. Biswas, S.T. Rikta, S.K. Dey, Machine learning-based diagnosis of breast cancer utilizing feature optimization technique, Computer Methods and Programs in Biomedicine Update 140 (2023) 145–156, https://doi.org/10.1016/j.cmpbup.2023.100098, 3, 100098.
[48] J.H. Lee, J. Shin, M.J. Realff, Machine learning: overview of the recent signs of progress and implications for the process systems engineering field, Comput. Chem. Eng. 114 (2018) 111–121, https://doi.org/10.1016/j.compchemeng.2017.10.008.
[49] B. Jin, M. Milling, M.P. Plaza, J.O. Brunner, C. Traidl-Hoffmann, B.W. Schuller, A. Damialis, Airborne pollen grain detection from partially labeled data utilising semi-supervised learning, Sci. Total Environ. 891 (2023) 164295, https://doi.org/10.1016/j.scitotenv.2023.164295.
[50] M. Amini, A. Rahmani, Machine learning process evaluating damage classification of composites, International Journal of Science and Advanced Technology 9 (2023) (2023) 240–250.
[51] F.Y. Osisanwo, J.E.T. Akinsola, O. Awodele, J.O. Hinmikaiye, O. Olakanmi, J. Akinjobi, Supervised machine learning algorithms: classification and comparison, Int. J. Comput. Trends Technol. 48 (3) (2017) 128–138.

[52] S. Dargan, M. Kumar, M.R. Ayyagari, G. Kumar, A survey of deep learning and its applications: a new paradigm to machine learning, Arch. Comput. Methods Eng. 27 (2020) 1071–1092, https://doi.org/10.1007/s11831-019-09344-w.
[53] H. Liu, B. Lang, Machine learning and deep learning methods for intrusion detection systems: a survey, Appl. Sci. 9 (20) (2019) 4396, https://doi.org/10.3390/app9204396.
[54] H. Ma, Z. Zhang, W. Li, S. Lu, Unsupervised human activity representation learning with multi-task deep clustering, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 5 (1) (2021) 1–25, https://doi.org/10.1145/3448074.
[55] Z.E. Imel, B.T. Pace, C.S. Soma, M. Tanana, T. Hirsch, J. Gibson, D.C. Atkins, Design feasibility of an automated, machine-learning-based feedback system for motivational interviewing, Psychotherapy 56 (2) (2019) 318, https://doi.org/10.1037/pst0000221.
[56] J. Peksa, D. Mamchur, State-of-the-Art on brain-computer interface technology, Sensors 23 (13) (2023) 6001, https://doi.org/10.3390/s23136001.
[57] B. Charbuty, A. Abdulazeez, Classification based on decision tree algorithm for machine learning, Journal of Applied Science and Technology Trends 2 (1) (2021) 20–28, https://doi.org/10.38094/JASTT20165.
[58] T. Ge, X. Luo, Y. Wang, M. Sedlmair, Z. Cheng, Y. Zhao, B. Chen, Optimally ordered orthogonal neighbor joining trees for hierarchical cluster analysis, IEEE Trans. Visual. Comput. Graph. (2023), https://doi.org/10.1109/TVCG.2023.3284499.
[59] V.G. Costa, C.E. Pedreira, Recent advances in decision trees: an updated survey, Artif. Intell. Rev. 56 (5) (2023) 4765–4800, https://doi.org/10.1007/s10462-022-10275-5.
[60] A. Ahmim, L. Maglaras, M.A. Ferrag, M. Derdour, H. Janicke, A novel hierarchical intrusion detection system based on decision tree and rules-based models, in: 2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS), IEEE, 2019, pp. 228–233, https://doi.org/10.1109/DCOSS.2019.00059.
[61] G. Gonzalez Rodriguez, J.M. Gonzalez-Cava, J.A. Méndez Pérez, An intelligent decision support system for production planning based on machine learning, J. Intell. Manuf. 31 (5) (2020) 1257–1273, https://doi.org/10.1007/s10845-019-01510-y.
[62] P.C. Sen, M. Hajra, M. Ghosh, Supervised classification algorithms in machine learning: a survey and review, in: Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph 2018, Springer, Singapore, 2020, pp. 99–111, https://doi.org/10.1007/978-981-13-7403-6_11.
[63] S. Neelakandan, D. Paulraj, A gradient-boosted decision tree-based sentiment classification of Twitter data, Int. J. Wavelets, Multiresolut. Inf. Process. 18 (4) (2020) 2050027, https://doi.org/10.1142/S0219691320500277.
[64] A.A. Dehghani, N. Movahedi, K. Ghorbani, S. Eslamian, Decision tree algorithms, in: Handbook of Hydroinformatics, Elsevier, 2023, pp. 171–187, https://doi.org/10.1016/B978-0-12-821285-1.00004-X.
[65] H. Kocyigit, Process of machine learning methods, in: Advancement in Business Analytics Tools for Higher Financial Performance, IGI Global, 2023, pp. 1–38, https://doi.org/10.4018/978-1-6684-8386-2.ch001.
[66] S. Barbon, A.P.A.D. Costa Barbon, R.G. Mantovani, D.F. Barbin, Machine learning applied to near-infrared spectra for chicken meat classification, Journal of Spectroscopy 2018 (2018), https://doi.org/10.1155/2018/8949741.
[67] L. Velásquez, J.P. Cruz-Tirado, R. Siche, R. Quevedo, An application based on the decision tree to classify the marbling of beef by hyperspectral imaging, Meat Sci. 133 (2017) 43–50, https://doi.org/10.1016/j.meatsci.2017.06.002.
[68] A.X. Wang, S.S. Chukova, B.P. Nguyen, Ensemble k-nearest neighbors based on centroid displacement, Inf. Sci. 629 (2023) 313–323, https://doi.org/10.1016/j.ins.2023.02.004.
[69] S. Wakhid, R. Sarno, S.I. Sabilla, D.B. Maghfira, Detection and classification of Indonesian civet and non-civet coffee based on statistical analysis comparison using E-nose, International Journal of Intelligent Engineering & Systems 13 (4) (2020).
[70] K. Kirasich, T. Smith, B. Sadler, Random forest vs logistic regression: binary classification for heterogeneous datasets, SMU Data Science Review 1 (3) (2018) 9.
[71] A. Dashti, J. Müller-Maatsch, E. Roetgerink, M. Wijtten, Y. Weesepoel, H. Parastar, H. Yazdanpanah, Comparison of a portable Vis-NIR hyperspectral imaging and a snapscan SWIR hyperspectral imaging for evaluation of meat authenticity, Food Chem. X 18 (2023) 100667.
[72] S. Weng, B. Guo, P. Tang, X. Yin, F. Pan, J. Zhao, D. Zhang, Rapid detection of adulteration of minced beef using Vis/NIR reflectance spectroscopy with multivariate methods, Spectrochim. Acta Mol. Biomol. Spectrosc. 230 (2020) 118005.
[73] A. Kulikov, A. Loskutov, D. Bezdushniy, I. Petrov, Decision tree models and machine learning algorithms in the fault recognition on power lines with branches, Energies 16 (14) (2023) 5563, https://doi.org/10.3390/en16145563.
[74] S.K. Parhi, S.K. Patro, Compressive strength prediction of PET fiber-reinforced concrete using Dolphin echolocation optimized decision tree-based machine learning algorithms, Asian Journal of Civil Engineering (2023) 1–20, https://doi.org/10.1007/s42107-023-00826-8.
[75] G.R. Faqe Ibrahim, A. Rasul, H. Abdullah, Improving crop classification accuracy with integrated Sentinel-1 and Sentinel-2 data: a case study of barley and wheat, Journal of Geovisualization and Spatial Analysis 7 (2) (2023) 22, https://doi.org/10.1007/s41651-023-00152-2.
[76] A. Abdulhafedh, Comparison between common statistical modeling techniques used in research, including Discriminant analysis vs logistic regression, ridge regression vs LASSO, and decision tree vs random forest, Open Access Library Journal 9 (2) (2022) 1–19, https://doi.org/10.4236/oalib.1108414.
[77] W. Chen, D. Krainak, B. Sahiner, N. Petrick, A regulatory science perspective on performance assessment of machine learning algorithms in imaging, in: Machine Learning for Brain Disorders, Springer US, New York, NY, 2023, pp. 705–752, https://doi.org/10.1016/j.autcon.2021.103606.
[78] A. Moclán, M. Domínguez-Rodrigo, Are highly accurate models of agency in bone breaking the result of misuse of machine learning methods? J. Archaeol. Sci.: Reports 51 (2023) 104150 https://doi.org/10.1016/j.jasrep.2023.104150.
[79] T.T.N. Pragasam, J.V.J. Thomas, M.A. Vensuslaus, S. Radhakrishnan, CEAT: categorising ethereum addresses' transaction behaviour with ensemble machine learning algorithms, Computation 11 (8) (2023) 156, https://doi.org/10.3390/computation11080156.
[80] N. Chaplot, D. Pandey, Y. Kumar, P.S. Sisodia, A comprehensive analysis of artificial intelligence techniques for the prediction and prognosis of genetic disorders using various gene disorders, Arch. Comput. Methods Eng. 30 (5) (2023) 3301–3323, https://doi.org/10.1007/s11831-023-09904-1.
[81] D. Dutta, D. Natta, S. Mandal, N. Ghosh, MOOnitor: an IoT based multi-sensory intelligent device for cattle activity monitoring, Sensor Actuator Phys. 333 (2022) 113271.
[82] A.M. Franco, A.E.M. da Silva, P.J. Hurtado, F.H. de Moura, S. Huber, M.A. Fonseca, Comparison of linear and nonlinear decision boundaries to detect feedlot bloat using intensive data collection systems on Angus x Hereford steers, Animal 100809 (2023).
[83] G. Morota, R.V. Ventura, F.F. Silva, M. Koyama, S.C. Fernando, Big data analytics and precision animal agriculture symposium: machine learning and data mining advance predictive big data analysis in precision animal agriculture, J. Anim. Sci. 96 (4) (2018) 1540–1550.
[84] D.A. Wijeyakulasuriya, E.W. Eisenhauer, B.A. Shaby, E.M. Hanks, Machine learning for modeling animal movement, PLoS One 15 (7) (2020) e0235750.
[85] T. Xiang, T. Li, J. Li, X. Li, J. Wang, Using machine learning to realize genetic site screening and genomic prediction of productive traits in pigs, Faseb. J. 37 (6) (2023) e22961.
[86] N. Haque, A. Jamshed, K. Chatterjee, S. Chatterjee, Accurate sensing of power transformer faults from dissolved gas data using random forest classifier aided by data clustering method, IEEE Sensor. J. 22 (6) (2022) 5902–5910, https://doi.org/10.1109/JSEN.2022.3149409.
[87] H. Sun, C. Song, X. Lin, X. Gao, Identification of meat species by combined laser-induced breakdown and Raman spectroscopies, Spectrochim. Acta B Atom Spectrosc. 194 (2022) 106456.
[88] T. Lan, H. Hu, C. Jiang, G. Yang, Z. Zhao, A comparative study of decision tree, random forest, and convolutional neural network for spread-F identification, Adv. Space Res. 65 (8) (2020) 2052–2061, https://doi.org/10.1016/j.asr.2020.01.036.
[89] M.Á. Carreira-Perpiñán, M. Gabidolla, A. Zharmagambetov, Towards better decision forests: forest alternating optimization, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 7589–7598.
[90] T. Thomas, P. VijayaraghavanA, S. Emmanuel, T. Thomas, P. Vijayaraghavan, A, S. Emmanuel, Nearest neighbor and fingerprint classification, Machine Learning Approaches in Cyber Security Analytics (2020) 107–128, https://doi.org/10.1007/978-981-15-1706-8_6.
[91] O.A. Montesinos López, A. Montesinos López, J. Crossa, Overfitting, model tuning, and evaluation of prediction performance, in: Multivariate Statistical Machine Learning Methods for Genomic Prediction, Springer International Publishing, Cham, 2022, pp. 109–139, https://doi.org/10.1007/978-3-030-89010-0_4.

[92] N.S. Alotaibi, H.I. Ahmed, S.O.M. Kamel, Dynamic adaptation attack detection model for a distributed multi-access edge computing smart city, Sensors 23 (16) (2023) 7135, https://doi.org/10.3390/s23167135.
[93] K. Takahashi, K. Ichikawa, J. Park, G.M. Pao, Scalable empirical dynamic modeling with parallel computing and approximate k-NN search, IEEE Access (2023), https://doi.org/10.1109/ACCESS.2023.3289836.
[94] M.D. de Lima, R. Barbosa, Methods of authentication of food grown in organic and conventional systems using chemometrics and data mining algorithms: a review, Food Anal. Methods 12 (2019) 887–901, https://doi.org/10.1007/s12161-018-01413-3.
[95] D. Stefas, N. Gyftokostas, E. Nanou, P. Kourelias, S. Couris, Laser-induced breakdown spectroscopy: an efficient tool for food science and technology (from the analysis of Martian rocks to the analysis of olive oil, honey, milk, and other natural earth products), Molecules 26 (16) (2021) 4981, https://doi.org/10.3390/molecules26164981.
[96] N.F. Hamidon Majid, M.S. Najib, M.F. Zahari, S. Zaib, T.S. Tuan Muda, The classification of meat odor-profile using K-nearest neighbors (KNN), in: Proceedings of the 6th International Conference on Electrical, Control and Computer Engineering: InECCE2021, Kuantan, Pahang, Malaysia, Springer Singapore, Singapore, 2022, March, pp. 551–562, 23rd August.
[97] B. Zhang, S. Gao, F. Jia, X. Liu, X. Li, Categorization and authentication of Beijing-you chicken from four breeds of chickens using near-infrared hyperspectral imaging combined with chemometrics, J. Food Process. Eng. 43 (12) (2020) e13553.
[98] W. Jia, Y. Qin, C. Zhao, Rapid detection of adulterated lamb meat using near infrared and electronic nose: a F1-score-MRE data fusion approach, Food Chem. (2023) 138123.
[99] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, A. Lopez, A comprehensive survey on support vector machine classification: applications, challenges, and trends, Neurocomputing 408 (2020) 189–215, https://doi.org/10.1016/j.neucom.2019.10.118.
[100] K. ElHaj, D. Alshamsi, A. Aldahan, GeoZ: a region-based visualization of clustering algorithms, Journal of Geovisualization and Spatial Analysis 7 (1) (2023) 15, https://doi.org/10.1007/s41651-023-00146-0.
[101] N. Jalal, A. Mehmood, G.S. Choi, I. Ashraf, A novel improved random forest for text classification using feature ranking and optimal number of trees, Journal of King Saud University-Computer and Information Sciences 34 (6) (2022) 2733–2742, https://doi.org/10.1016/j.jksuci.2022.03.012.
[102] M. Zareef, Q. Chen, M.M. Hassan, M. Arslan, M.M. Hashim, W. Ahmad, A.A. Agyekum, An overview on the applications of typical non-linear algorithms coupled with NIR spectroscopy in food analysis, Food Eng. Rev. 12 (2020) 173–190, https://doi.org/10.1007/s12393-020-09210-7.
[103] S. Tarannum, Halal Food Identification from Product Ingredients Using Machine Learning (Doctoral Dissertation, United International University, 2023.
[104] A. Kartakoullis, J. Comaposada, A. Cruz-Carrión, X. Serra, P. Gou, Feasibility study of smartphone based Near Infrared Spectroscopy (NIRS) for salted minced meat composition diagnostics at different temperatures, Food Chem. 278 (2019) 314–321, https://doi.org/10.1016/j.foodchem.2018.11.054.
[105] M.M. Oliveira, J.P. Cruz-Tirado, J.V. Roque, R.F. Teófilo, D.F. Barbin, Portable near-infrared spectroscopy for rapid authentication of adulterated paprika powder, J. Food Compos. Anal. 87 (2020) 103403, https://doi.org/10.1016/j.jfca.2019.103403.
[106] A.I. Ropodi, E.Z. Panagou, G.J.E. Nychas, Multispectral imaging (MSI): a promising method for the detection of minced beef adulteration with horsemeat, Food Control 73 (2017) 57–63, https://doi.org/10.1016/j.foodcont.2016.05.048.
[107] A. Dashti, Y. Weesepoel, J. Müller-Maatsch, H. Parastar, F. Kobarfard, B. Daraei, H. Yazdanpanah, Assessment of meat authenticity using portable Fourier transform infrared spectroscopy combined with multivariate classification techniques, Microchem. J. 181 (2022) 107735, https://doi.org/10.1016/j.microc.2022.107735.
[108] R. Sarno, K. Triyana, S.I. Sabilla, D.R. Wijaya, D. Sunaryono, C. Fatichah, Detecting pork adulteration in beef for halal authentication using an optimized electronic nose system, IEEE Access 8 (2020) 221700–221711, https://doi.org/10.1109/ACCESS.2020.3043394.
[109] M.A. Siddiqui, M.H.M. Khir, G. Witjaksono, A.S.M. Ghumman, M. Junaid, S.A. Magsi, A. Saboor, Multivariate analysis coupled with M-SVM classification for lard adulteration detection in meat mixtures of beef, lamb, and chicken using FTIR spectroscopy, Foods 10 (10) (2021) 2405.
[110] E.Ş. Tüzemen, A.G. Yüksek, İ. Demir, S. Horoz, İ. Altuntaş, Modeling of temperature-dependent photoluminescence of GaN epilayer by artificial neural network, Journal of the Australian Ceramic Society (2023) 1–15, https://doi.org/10.1007/s41779-023-00911-w.
[111] O. Taki, K.S. Rhazi, Y. Mejdoub, Stirling engine optimization using artificial neural networks algorithm, in: ITM Web of Conferences, vol. 52, EDP Sciences, 2023 02010, https://doi.org/10.1051/itmconf/20235202010.
[112] X. Feng, G. Ma, S.F. Su, C. Huang, M.K. Boswell, P. Xue, A multi-layer perceptron approach for accelerated wave forecasting in Lake Michigan, Ocean Eng. 211 (2020) 107526, https://doi.org/10.1016/j.oceaneng.2020.107526.
[113] M. Roshani, G.T. Phan, P.J.M. Ali, G.H. Roshani, R. Hanus, T. Duong, E.M. Kalmoun, Evaluation of flow pattern recognition and void fraction measurement in two-phase flow independent of oil pipeline's scale layer thickness, Alex. Eng. J. 60 (1) (2021) 1955–1966, https://doi.org/10.1016/j.aej.2020.11.043.
[114] F. Farooq, W. Ahmed, A. Akbar, F. Aslam, R. Alyousef, Predictive modeling for sustainable high-performance concrete from industrial wastes: a comparison and optimization of models using ensemble learners, J. Clean. Prod. 292 (2021) 126032, https://doi.org/10.1016/j.jclepro.2021.126032.
[115] Vesna Knights, Mirela Kolak, Gordana Markovikj, Jasenka Gajdoš Kljusurić, Modeling and optimization with artificial intelligence in nutrition, Appl. Sci. 13 (13) (2023) 7835, https://doi.org/10.3390/app13137835.
[116] D. Dais, I.E. Bal, E. Smyrou, V. Sarhosis, Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning, Autom. ConStruct. 125 (2021) 103606, https://doi.org/10.1016/j.autcon.2021.103606.
[117] E.H. Houssein, A. Hammad, A.A. Ali, Human emotion recognition from EEG-based brain–computer interface using machine learning: a comprehensive review, Neural Comput. Appl. 34 (15) (2022) 12527–12557, https://doi.org/10.1007/s00521-022-07292-4.
[118] T.M. Tung, Z.M. Yaseen, A survey on river water quality modeling using artificial intelligence models: 2000–2020, J. Hydrol. 585 (2020) 124670, https://doi.org/10.1016/j.jhydrol.2020.124670.
[119] K.K. Pulluri, V.N. Kumar, Qualitative and quantitative detection of food adulteration using a smart e-nose, Sensors 22 (20) (2022) 7789.
[120] M.A. Sairin, S. Abd Aziz, C.P. Tan, S. Mustafa, S.S. Abd Gani, F.Z. Rokhani, Lard classification from other animal fats using Dielectric Spectroscopy technique, Int. Food Res. J. 26 (3) (2019) 773–782.
[121] M. Noshad, B.A. Behbahani, I.K. Karabagias, Volatilomic with chemometrics: a toward authentication approach for food authenticity control, Eur. Food Res. Technol. (2023) 1–12.
[122] A. Rohman, A.R. Putri, The chemometrics techniques in combination with instrumental analytical methods applied in Halal authentication analysis, Indonesian Journal of Chemistry 19 (1) (2019) 262–272.
[123] M.S.A. Sani, N.F.H. Nordin, A.A. Elgharbawy, Halal detection technologies: analytical method approaches, validation and verification, and multivariate data analysis for halal authentication, Innovation of Food Products in Halal Supply Chain Worldwide (2023) 253–271.
[124] H. Xuesong, C. Pu, L. Jingyan, X. Yupeng, L. Dan, C. Xiaoli, Commentary on the review articles of spectroscopy technology combined with chemometrics in the last three years, Appl. Spectrosc. Rev. (2023) 1–60.
[125] C.C. Aggarwal, Neural networks and deep learning, Springer 10 (978) (2018) 3, https://doi.org/10.1007/978-3-319-94463-0.
[126] D. Menegatti, A. Giuseppi, F. Delli Priscoli, A. Pietrabissa, A. Di Giorgio, F. Baldisseri, V. Suraci, CADUCEO: a platform to support federated healthcare facilities through artificial intelligence, Healthcare 11 (15) (2023) 2199, https://doi.org/10.3390/healthcare11152199. MDPI.
[127] S. Uhlig, I. Alkhasli, F. Schubert, C. Tschöpe, M. Wolff, A review of synthetic and augmented training data for machine learning in ultrasonic non-destructive evaluation, Ultrasonics 107041 (2023), https://doi.org/10.1016/j.ultras.2023.107041.
[128] M. Vishwakarma, N. Kesswani, A new two-phase intrusion detection system with Naïve Bayes machine learning for data classification and elliptic envelop method for anomaly detection, Decision Analytics Journal 7 (2023) 100233, https://doi.org/10.1016/j.dajour.2023.100233.
[129] K. Yadav, R. Thareja, Comparing the performance of naive bayes and decision tree classification using R, Int. J. Intell. Syst. Appl. 11 (12) (2019) 11, https://doi.org/10.5815/ijisa.2019.12.02.
[130] Y. Lin, J. Ma, Q. Wang, D.W. Sun, Applications of machine learning techniques for enhancing nondestructive food quality and safety detection, Crit. Rev. Food Sci. Nutr. 63 (12) (2023) 1649–1669, https://doi.org/10.1080/10408398.2022.2131725.

[131] E.G. Dada, J.S. Bassi, H. Chiroma, A.O. Adetunmbi, O.E. Ajibuwa, Machine learning for email spam filtering: review, approaches, and open research problems, Heliyon 5 (6) (2019) E01802.
[132] X. Tian, J. Wang, R. Shen, Z. Ma, M. Li, Discrimination of pork/chicken adulteration in minced mutton by electronic taste system, Int. J. Food Sci. Technol. 54 (3) (2019) 670–678.
[133] S.H. Alizadeh, A. Hediehloo, N.S. Harzevili, Multi-independent latent component extension of naive Bayes classifier, Knowl. Base Syst. 213 (2021) 106646, https://doi.org/10.1016/j.knosys.2020.106646.
[134] S. Srivastav, K. Guleria, S. Sharma, Predictive machine learning approaches for chronic kidney disease, in: 2023 4th International Conference for Emerging Technology (INCET), IEEE, 2023, May, pp. 1–5, https://doi.org/10.1109/INCET57972.2023.10170425.
[135] Z.S. Alzamil, D. Appelbaum, W. Glasgall, M.A. Vasarhelyi, Applications of data analytics: cluster analysis of not-for-profit data, J. Inf. Syst. 35 (3) (2021) 199–221.
[136] V. Gugueoth, S. Safavat, S. Shetty, Security of Internet of Things (IoT) using federated learning and deep learning-Recent advancements, issues and prospects, ICT Express (2023).
[137] M.I. Razzak, M. Imran, G. Xu, Big data analytics for preventive medicine, Neural Comput. Appl. 32 (9) (2020) 4417–4451.
[138] H.H. Ali, L.E. Kadhum, K-means clustering algorithm applications in data mining and pattern recognition, Int. J. Sci. Res. 6 (8) (2017) 1577–1584.
[139] A.K. Dubey, U. Gupta, S. Jain, Analysis of k-means clustering approach on the breast cancer Wisconsin dataset, International Journal of Computer Assisted Radiology and Rurgery 11 (2016) 2033–2047, https://doi.org/10.1007/s11548-016-1437-9.
[140] A. Kartakoullis, A. Kamilaris, J. Gonzalez, X. Serra, P. Gou, M. Font, Hyperspectral imaging for assessing the quality attributes of cured pork loin, in: 2018 9th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), IEEE, 2018, September, pp. 1–4, https://doi.org/10.1109/WHISPERS.2018.8747235.
[141] W. Leal, E.J. Llanos, G. Restrepo, C.F. Suárez, M.E. Patarroyo, How frequently do clusters occur in hierarchical clustering analysis? A graph theoretical approach to studying ties in proximity, J. Cheminf. 8 (2016) 1–16, https://doi.org/10.1186/s13321-016-0114-x.
[142] M. Moufid, C. Tiebe, N. El Bari, D.A.H. Fakra, M. Bartholmai, B. Bouchikhi, Pollution parameters evaluation of wastewater collected at different treatment stages from wastewater treatment plants based on E-nose and E-tongue systems combined with chemometric techniques, Chemometr. Intell. Lab. Syst. 227 (2022) 104593, https://doi.org/10.1016/j.chemolab.2022.104593.
[143] A. Windarsih, N. K. A.Dachriyanus Bakar, N.D. Yuliana, F.D.O. Riswanto, A. Rohman, Analysis of pork in beef sausages using LC-orbitrap HRMS untargeted metabolomics combined with chemometrics for halal authentication study, Molecules 28 (16) (2023) 5964.
[144] O. Taylan, N. Cebi, M.T. Yilmaz, O. Sagdic, A.A. Bakhsh, Detection of lard in butter using Raman spectroscopy combined with chemometrics, Food Chem. 332 (2020) 127344.
[145] A.M. Judith, S.B. Priya, R.K. Mahendran, Artifact removal from EEG signs using regenerative multi-dimensional singular value decomposition and independent component analysis, Biomed. Signal Process Control 74 (2022) 103452, https://doi.org/10.1016/j.bspc.2021.103452.
[146] L.A. Putri, I. Rahman, M. Puspita, S.N. Hidayat, A.B. Dharmawan, A. Rianjanu, H.S. Wasisto, Rapid analysis of meat floss origin using a supervised machine learning -based electronic nose towards food authentication, npj Science of Food 7 (1) (2023) 31, https://doi.org/10.1038/s41538-023-00205-2.
[147] J. Happa, Insider-threat detection using Gaussian mixture models and sensitivity profiles, Comput. Secur. 77 (2018) 838–859, https://doi.org/10.1016/j.cose.2018.03.006.
[148] S. Thudumu, P. Branch, J. Jin, J. Singh, A comprehensive survey of anomaly detection techniques for high dimensional big data, Journal of Big Data 7 (2020) 1–30.
[149] D. Granato, J.S. Santos, G.B. Escher, B.L. Ferreira, R.M. Maggio, Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: a critical perspective, Trends Food Sci. Technol. 72 (2018) 83–90, https://doi.org/10.1016/j.tifs.2017.12.006.
[150] N. Upadhyay, P. Jaiswal, S. Narayan Jha, Application of attenuated total reflectance Fourier Transform Infrared spectroscopy (ATReFTIR) in MIR range coupled with chemometrics for detection of pig body fat in pure ghee (heat clarified milk fat), J. Mol. Struct. 1153 (2018) 275–281, https://doi.org/10.1016/j.molstruc.2017.09.116.
[151] J. Zhao, Z. Li, Q. Gao, H. Zhao, S. Chen, L. Huang, T. Wang, A review of statistical methods for dietary pattern analysis, Nutr. J. 20 (1) (2021) 1–18, https://doi.org/10.1186/s12937-021-00692-7.
[152] D. Goretzko, T.T.H. Pham, M. Bühner, Exploratory factor analysis: current use, methodological developments, and recommendations for good practice, Curr. Psychol. 40 (2021) 3510–3521, https://doi.org/10.1007/s12144-019-00300-2.
[153] Y. Wan, T. Li, P. Wang, S. Duan, C. Zhang, N. Li, Robust and efficient classification for underground metal target using dimensionality reduction and machine learning, IEEE Access 9 (2021) 7384–7401.
[154] M. Ahda, A. Guntarti, A. Kusbandari, H. Andoyo Nugroho, Identification of adulterated sausage products by pork using FTIR and GC-MS combined with chemometrics, Journal of Chemical Health Risks 13 (2) (2023) 325–332.
[155] D. Cozzolino, D. Bureš, L.C. Hoffman, Evaluating the use of a similarity index (SI) combined with near infrared (NIR) spectroscopy as method in meat species authenticity, Foods 12 (1) (2023) 182.
[156] K.R. Dewi, M. Ismayati, N.N. Solihat, N.D. Yuliana, F. Kusnandar, H. Riantana, S. Kim, Advances and key considerations of liquid chromatography–mass spectrometry for porcine authentication in Halal analysis, Journal of Analytical Science and Technology 14 (1) (2023) 13.
[157] F. Mabood, R. Boqué, A.Y. Alkindi, A. Al-Harrasi, I.S. Al Amri, S. Boukra, I. Kadim, Fast detection and quantification of pork meat in other meats by reflectance FT-NIR spectroscopy and multivariate analysis, Meat Sci. 163 (2020) 108084.
[158] V. Maritha, P.W. Harlina, I. Musfiroh, M. Rafi, F. Geng, M. Muchtaridi, Exploring untargeted metabolomics for halal authentication of triceps brachii, longissimus dorsi, and biceps femoris of meat muscles, Int. J. Food Prop. 26 (2) (2023) 3148–3159.
[159] Mokhtar, N., Hamid, A. A., Ismail, A., Mustafa, S., & Rohman, A. Lipidomic Fingerprinting and Chemometric Application for Halal Authentication of Meat Species. Available at: SSRN 4609182.
[160] M.P. Totaro, G. Squeo, D. De Angelis, A. Pasqualone, F. Caponio, C. Summo, Application of NIR spectroscopy coupled with DD-SIMCA class modelling for the authentication of pork meat, J. Food Compos. Anal. 118 (2023) 105211.