

Machine Learning Applied to the Detection of Mycotoxin in Food: A Systematic Review

Alan Inglis ^{1,*} , Andrew C. Parnell ¹ , Natarajan Subramani ² and Fiona M. Doohan ² 

¹ Hamilton Institute, Eolas Building, Maynooth University, W23 F2H6 Maynooth, Kildare, Ireland; andrew.parnell@mu.ie

² School of Biology and Environmental Science, University College Dublin, D04 C1P1 Dublin, Ireland; natarajan.subramani@ucd.ie (N.S.); fiona.doohan@ucd.ie (F.M.D.)

* Correspondence: alan.inglis@mu.ie

Abstract: Mycotoxins, toxic secondary metabolites produced by certain fungi, pose significant threats to global food safety and public health. These compounds can contaminate a variety of crops, leading to economic losses and health risks to both humans and animals. Traditional lab analysis methods for mycotoxin detection can be time-consuming and may not always be suitable for large-scale screenings. However, in recent years, machine learning (ML) methods have gained popularity for use in the detection of mycotoxins and in the food safety industry in general due to their accurate and timely predictions. We provide a systematic review on some of the recent ML applications for detecting/predicting the presence of mycotoxin on a variety of food ingredients, highlighting their advantages, challenges, and potential for future advancements. We address the need for reproducibility and transparency in ML research through open access to data and code. An observation from our findings is the frequent lack of detailed reporting on hyperparameters in many studies and a lack of open source code, which raises concerns about the reproducibility and optimisation of the ML models used. The findings reveal that while the majority of studies predominantly utilised neural networks for mycotoxin detection, there was a notable diversity in the types of neural network architectures employed, with convolutional neural networks being the most popular.

Keywords: machine learning; predictive model; mycotoxin; food safety; systematic review



Citation: Inglis, A.; Parnell, A.C.; Subramani, N.; Doohan, F.M. Machine Learning Applied to the Detection of Mycotoxin in Food: A Systematic Review. *Toxins* **2024**, *16*, 268. <https://doi.org/10.3390/toxins16060268>

Received: 29 April 2024

Revised: 31 May 2024

Accepted: 6 June 2024

Published: 12 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Key Contribution: Recent developments in machine learning present promising approaches to improve the precision and efficiency of detecting mycotoxins. This review comprehensively gathers and examines the latest research at the juncture of machine learning and mycotoxin detection in food items. It offers a detailed assessment of the methods used, accomplishments, and potential future developments.

1. Introduction

Mycotoxins are a group of naturally occurring toxic chemical compounds produced by certain species of moulds (fungi) during growth on various crops and foodstuffs, including cereals, nuts, spices, and dairy products [1]. The ingestion of certain mycotoxins has been linked to a range of harmful health impacts on both humans and animals, from short-term poisoning to long-term consequences such as liver cancer and, in some cases, death [2–4]. Mycotoxins are secondary metabolites (that is, compounds produced by an organism that are not essential for its primary life processes) and are often produced during the pre-harvest, harvest, and storage phases under favourable conditions of humidity and temperature [3,5]. The most prevalent mycotoxins include aflatoxins, tricothecenes, fumonisins, zearalenones, ochratoxins, and patulin, and are produced by certain plant-pathogenic species of *Aspergillus*, *Fusarium*, and *Penicillium* [6]. Mycotoxin contamination in crop

products has been found to vary significantly across different geographical locations and is influenced by annual weather conditions [7,8]. However, since 2012, there has been a noted increase in the occurrence of mycotoxins in Europe, with the impacts of climate change being most likely a contributing factor [9,10]. An estimated 60–80% of the world's crop supply is contaminated by mycotoxins, and an estimated 20% of those crops surpass the legally mandated food safety thresholds set by the European Union (EU) [11].

With the world's food supply chain being highly interconnected, the presence of mycotoxins not only endangers human health but also has an impact on the stability of agricultural markets and trade [3,12]. The economic impact of mycotoxin contamination is substantial, with a global estimate in the billions of euros for detection, regulation enforcement, and mitigation efforts to manage mycotoxin presence in food and feeds annually [13]. It is estimated that, between 2010 and 2019, approximately 75 million tonnes of wheat in Europe, which constitutes 5% of the wheat intended for human consumption, surpassed the maximum threshold for DON contamination. This excess led to the reclassification of this contaminated wheat grain as 'animal feed', resulting in an economic loss of around EUR 3 billion [14]. Additionally, [15] shows that, between 2010 and 2020, aflatoxins were responsible for the demotion of 4.2% of wheat intended for food, which potentially represented an additional economic loss of EUR 2.5 billion. As a result, the detection and management of mycotoxins in crops and food products is crucial for ensuring food safety and safeguarding consumer health worldwide as well as contributing to economic stability.

According to [16], the standard methodology for mycotoxin detection comprises three main steps: sampling, sample preparation, and analytical determination. Chromatographic techniques, such as liquid chromatography mass spectrometry (LC–MS), high-performance liquid chromatography (HPLC), and gas chromatography mass spectrometry (GC–MS), along with immunoassay-based methods like enzyme-linked immunosorbent assays (ELISAs), are widely recognised as the most prevalent analytical approaches for the detection of mycotoxins [17,18]. The mycotoxin level in a bulk load is determined by measuring a sample taken from the food source. From this, the concentration of mycotoxins in the entire load is assumed to be the same as the concentration of the sample. However, these techniques often require extensive sample preparation, sophisticated equipment, and highly trained personnel, leading to significant costs and time delays in the analytical process. Furthermore, the varied and intricate nature of different foods requires customised detection methods, which can add complexity to the screening process [19,20].

While traditional detection methods such as LC–MS, HPLC, GC–MS, and ELISA generate reliable data, they often result in large, complex datasets that require extensive interpretation and analysis. Machine learning (ML) approaches for both the detection and prediction of the presence of mycotoxins have seen a rise in recent years as an alternative to traditional detection methods (see Figure 1). At its core, ML employs statistical methods to create algorithms that allow computers to learn from data and make decisions based on identified patterns and inferences, without being explicitly programmed for each specific task. ML methods offer a sophisticated approach to deciphering the complex patterns hidden within the data and are adept at processing and analysing large datasets and extracting meaningful patterns that are not immediately apparent. By leveraging ML algorithms, researchers can gain deeper insights into the data and offer a significant advantage, when compared with traditional lab analysis, in terms of efficiency, cost, and scalability, as well as maintaining or improving the accuracy of mycotoxin detection [21].

ML methods can be, broadly, broken into three categories, that is, supervised learning (SL), unsupervised learning (UL), and reinforcement learning (RL). In SL, an algorithm is trained using a dataset that includes both inputs and the corresponding outputs. The model learns to associate the inputs with the outputs. After training, the model can apply this learned relationship to predict the outputs for new, unseen inputs [22]. In UL, an algorithm is presented with only the input data and identifies patterns and structures in the data based only on the inputs. After training, it can classify new inputs based on the patterns it has found. In RL, an algorithm learns to make decisions by performing actions to achieve a goal.

It processes feedback through rewards or penalties associated with its actions, using this information to develop a decision-making framework that aims to maximise rewards [23].

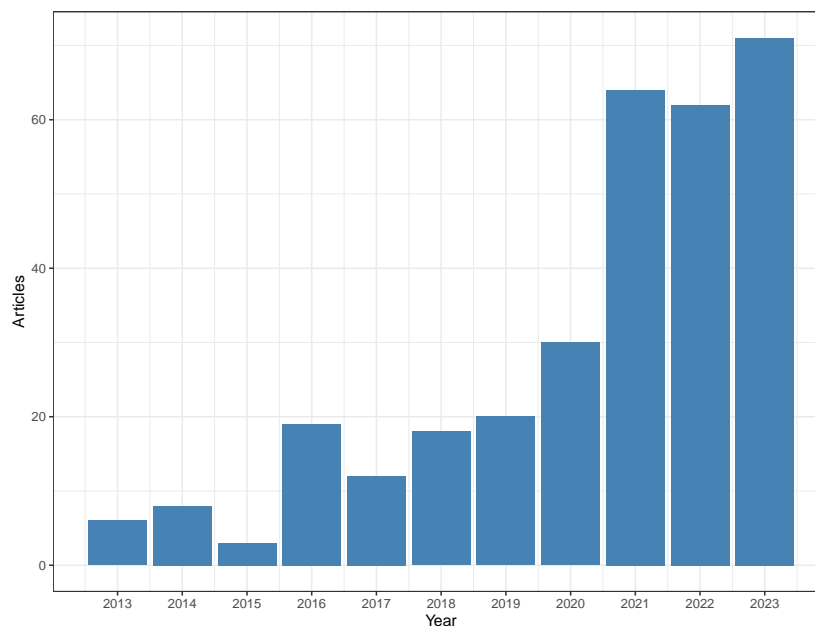


Figure 1. Number of publications between 2013 and 2023 found by our systematic search criteria in Scopus.

Within these categories, many different types of ML models exist and are used based on the specificity of the problem. The most popular of these models, as found by this research, are discussed in detail below. Although ML applications in food safety and mycotoxin detection are widespread, there appears to be a lack of comprehensive reviews that cover the broad spectrum of ML methodologies specifically tailored to mycotoxin analysis, as most studies tend to concentrate on individual techniques. For example, Ref. [24] uses neural networks (NNs) for the prediction of contamination from the mycotoxin fumonisin in corn. Additionally, NNs have been used to forecast the accumulation of the trichothecene mycotoxin deoxynivalenol (DON) in barley seeds [25] and to predict fungal growth [26]. For a comprehensive review of the use of NNs in food science, see Ref. [27]; for a review of ML methods in general in the field of food safety, see Ref. [28]; and in agriculture, see Ref. [21].

ML techniques can alleviate some of the current burdens of mycotoxin detection by providing an efficient and low-cost solution [29]. Additionally, with the impact of climate change, the need for these models to provide reliable predictions at the farm level is increasingly crucial, especially in terms of food safety and health. In this work, we present a comprehensive systematic review of some of the more popular ML techniques used in the detection and prediction of mycotoxin on a range of foods and crops. Our review also identifies critical areas in the current body of work that warrant attention. A notable concern is the often insufficient discussion on the selection and tuning of hyperparameters in ML models, which is crucial for understanding and replicating study results. This lack of details creates issues with the reproducibility of the reviewed methods and also hinders the advancement and application of these techniques.

The organisation of our article is as follows: In Section 2, we provide details regarding our literature search methodology. This includes a description of the search criteria and keywords and discussing the prevalence of each ML method. In Section 3, we provide a short introduction to the ML process and describe some of the common terms. In Section 4, we give a brief introduction to the main ML algorithms used (and their hyperparameters) and discuss the outcomes of the articles reviewed based on the type of machine learning model used. Finally, in Section 5, we provide some concluding remarks.

2. Literature Search Methodology

The literature search for this review was primarily conducted using Scopus (<https://www.scopus.com>, URL accessed on 10 November 2023), a widely recognised academic search engine that indexes scholarly articles across various disciplines. To ensure the relevance of the research, the search was restricted to articles published within the last 10 years (since November 2023). This time frame was chosen to capture the most recent advances and trends in the application of machine learning to mycotoxin detection in crops. The search engine was used to identify key studies, reviews, and seminal works pertinent to the topic at hand.

Search Criteria and Overview

A comprehensive search was conducted on the Scopus database and focused on publications between the years 2013 and 2023. The search was conducted using the primary keyword “mycotoxin” in combination with these machine learning-related terms: “artificial intelligence”, “bagging”, “Bayesian network”, “boosting”, “decision tree”, “deep learning”, “ensemble”, “gradient boost”, “k-means”, “k-nearest neighbour”, “knn”, “machine learning”, “neural network”, “principal component analysis”, “random forest”, “supervised learning”, “support vector machine”, “SVM”, and “unsupervised learning”. The search terms were motivated by a similar search used in a review of machine learning for the monitoring and prediction of food safety by [28]. This strategy was employed to ensure a wide coverage of potential articles at the intersection of mycotoxin detection and machine learning methodologies.

This search yielded 313 documents on Scopus. Figure 1 shows the results obtained from Scopus over the years 2013 to 2023. There is a general increasing trend across the years, with a marked rise after the year 2021.

To limit the search further, only peer-reviewed articles in English in the fields of agricultural and biological sciences, environmental science, computer science, and mathematics were chosen. This reduced the search size to 91. After examining the abstracts of all the 91 articles, 30 were selected for their relevance and included in this study. A flow diagram demonstrating our selection process can be found in Appendix A. From these articles, the predominant ML technique used was neural networks (NNs), followed by random forests (RFs) and gradient boosting (GB), and then support vector machines (SVMs), decision trees (DTs), and Bayesian networks (BNs). Figure 2 shows the frequency of each ML algorithm used in the literature.

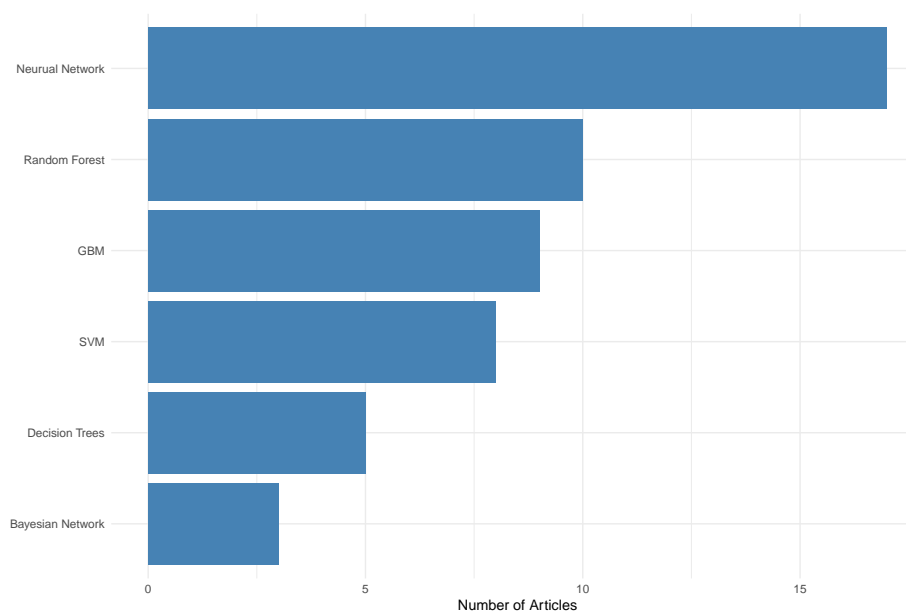


Figure 2. Most popular machine learning methods reviewed in this work.

3. A Brief Introduction to Machine Learning

In this section, we provide a general overview of the ML process. This foreknowledge is useful when discussing the ML approaches reviewed later in this document, though those already with experience in this topic may skip this section. To begin, we describe the typical process of creating an ML model.

3.1. Typical Machine Learning Process

Figure 3 shows a typical ML process for unsupervised and supervised learning methods.

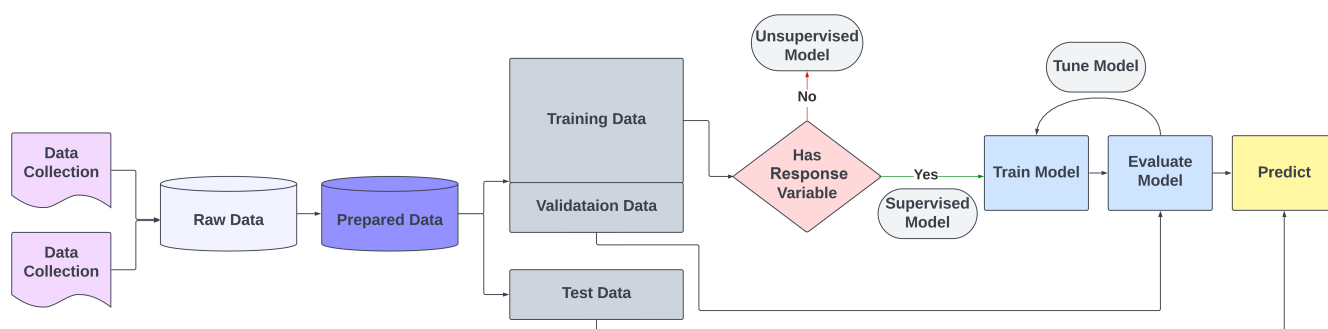


Figure 3. Typical machine learning process.

We can break up the process outlined in Figure 3 into five distinct steps. These are as follows:

1. **Data Collection:** The process starts with the collection of raw data, which can be from many sources or sites.
2. **Data Preparation:** These raw data are then prepared for analysis. This process typically involves cleaning and formatting the data.
3. **Data Splitting:** After preparation, the data can be split into three parts. These are training data, validation data, and test data (discussed more below).
4. **Model Selection:** Depending on the type of data, either an unsupervised or supervised learning model (or models) is chosen.
5. **Model Training, Evaluation, and Prediction:** This process involves training the model with training data, optimising the hyperparameters of the model using the validation data, and then evaluating its overall performance using test data.

3.2. Training, Validation, and Test Data

In ML, validation and test data are crucial for developing and evaluating models. Validation data are a separate subset of the original data, not used in training the model (see Figure 3). It helps in fine-tuning the model's parameters (known as hyperparameters), which are pre-set configurations of the model. This fine-tuning of hyperparameters during the validation process is essential to optimise the model's performance. One common technique used during this process is regularisation. Regularisation involves adding a penalty to the model's complexity, which helps prevent overfitting by ensuring that the model generalises well to new, unseen data rather than just memorising the training data. Validation also assists in selecting the best version of the model by providing feedback on its performance. This step is essential to prevent overfitting, ensuring that the model learns to generalise from the data and makes accurate predictions on new, unseen data. Test data are used after the model has been trained and validated (see Figure 3). It is another distinct subset of the dataset, not used in either training or validation. The test data are used to evaluate the final model's performance, providing an unbiased assessment of how well the model is likely to perform in real-world scenarios.

In all the referenced studies we cover below, model performance is quantified by evaluating the model performance on the test dataset, unless otherwise stated. Sometimes authors also report the training or validation dataset performance, but for the reasons

outlined above, these should be discarded as a measure of model performance. The common performance metrics used in these studies include the following:

- R^2 : This statistic measures the proportion of the variance in the dependent variable that can be explained by the independent variables in the model. An R^2 value closer to 1 indicates that the model accounts for a significant amount of the variance in the dependent variable.
- MSE and RMSE: Mean Square Error (MSE) is the average of the squares of the errors, which are the differences between predicted and actual values. Lower MSE values indicate a better fit of the model to the data. Root Mean Square Error (RMSE) is the square root of MSE. It has the same units as the quantity being estimated (for regression problems) and provides a measure of the differences between a model's predicted values and the actual observed values. Like MSE, a lower RMSE is better.
- Accuracy: This metric is commonly used for classification tasks and represents the ratio of correctly predicted observations to the total observations. High accuracy indicates that the model can correctly classify instances with high reliability.
- AUC: Area Under the Receiver Operating Characteristic Curve (AUC) is used in binary classification to measure a model's ability to distinguish between classes. An AUC of 1 represents perfect classifier performance, while an AUC of 0.5 denotes a model with no discriminative power.

4. Application of Machine Learning to Mycotoxin Data

In this section, we first include a brief discussion on common data types in mycotoxin detection. We then discuss the most common ML algorithms (from Figure 2) and review their application to mycotoxin data. Each subsection is dedicated to a single ML method in which we describe the basic algorithm, how it makes predictions/detections, some advantages and disadvantages of the algorithm, and finally a review of the literature using these methods. In cases where the reviewed studies employ multiple machine learning models, we categorise each paper based on the highest-performing model used in that particular work.

4.1. Types of Data Used in Mycotoxin Detection

In the context of ML applications for mycotoxin detection, the literature highlights the use of various data types, including weather parameters (temperature, rainfall, and relative humidity), crop phenology, agronomic data, and spectral imaging. Additionally, spatiotemporal data, which include information collected over time and across different spatial locations, play a vital role in understanding and predicting mycotoxin contamination by incorporating key environmental variables and temporal dynamics. Each type of data offers unique characteristics and applications. Understanding the context and conditions under which these data are collected is essential for interpreting the results and evaluating the effectiveness of different ML models.

4.1.1. Weather Data

Weather variables, including temperature, relative humidity, precipitation, and carbon dioxide levels, play a significant role in mycotoxigenic fungal growth and subsequent mycotoxin formation on agricultural commodities [30–32]. ML models can leverage historical and real-time weather data to predict the likelihood of mycotoxin contamination. For example, continuous monitoring of these variables in the field can help create more dynamic and responsive models. Incorporating these factors allows for a more comprehensive understanding of the conditions that favour mycotoxin contamination and can improve the predictive power of ML models.

Ref. [33] proposed a Convolutional Neural Network model based on CO₂ respiration rate and the visual appearance of mold formation for classifying mycotoxin contamination in wheat grains stored in sealed containers, which achieved an accuracy of 83.3%. Ref. [34] constructed a predictive model that incorporated multiple data sources, such as historical

records of aflatoxin and fumonisin in corn, daily weather conditions, satellite imagery, dynamic geospatial soil characteristics, and land usage information. Using both a gradient boosting machine and a neural network, the study demonstrated that the NN models exhibited high class-specific accuracy for predicting mycotoxin levels over a 1-year period, with accuracies of 73% for aflatoxin and 85% for fumonisin, demonstrating their efficacy in forecasting annual mycotoxin levels.

4.1.2. Agronomic Data

The impact of agronomic factors on mycotoxin occurrence has been extensively studied in various research. These factors include previous crop details, the use of fungicides, cropping patterns, and cultivar selection, all of which have been found to significantly affect mycotoxin levels [35–37]. In a study by Ref. [38], data on cropping system factors were used as input variables to predict aflatoxins and fumonisins in corn. Additionally, soil properties, when combined with meteorological data and historical aflatoxin content, have been used in gradient boosting machine models to distinguish aflatoxin-contaminated corn [39].

4.1.3. Crop Phenology and Cultivar-Specific Data

Another important aspect of spatiotemporal data is the inclusion of specific cultivars. Different crop varieties can exhibit varying levels of susceptibility to fungal colonisation and mycotoxin contamination [37,40]. Including data on specific cultivars in ML models can help tailor predictions and interventions to the particular characteristics of each crop variety. Certain wheat varieties may be more resistant to *Fusarium* head blight, while others might be more prone to infection. By incorporating cultivar-specific data, ML models can provide more accurate risk assessments and suggest more effective mitigation strategies [41]. This approach enhances the precision of mycotoxin contamination forecasts and supports targeted agricultural practices, such as selecting the most resistant varieties for planting in high-risk areas. Additionally, integrating crop phenology data, such as growth stages and development timelines, can improve the temporal accuracy of predictions [42].

4.1.4. Spectral Data

Spectral data are one of the most common types used in mycotoxin detection, valued for their non-invasive nature. This data type involves capturing the reflectance or absorbance of light at various wavelengths from the material being analysed. Spectral data can be further categorised into multispectral and hyperspectral data, each offering different levels of detail and information.

Multispectral Imaging: This imaging technique captures data at a few specific wavelength bands, making it effective for distinguishing between different materials based on their spectral signatures. Unlike hyperspectral imaging, which captures continuous spectral information across a wide range of wavelengths, multispectral imaging focuses on discrete bands, making data collection and processing less complex while still providing valuable information for specific applications. For instance, multispectral images can be captured in controlled greenhouse environments, where conditions such as temperature, humidity, and lighting are regulated to optimise data quality. This controlled setting allows for consistent and repeatable measurements, crucial for precise analysis. An example of this application is a study [43] that used hyperspectral data to detect *Fusarium* head blight in wheat under greenhouse conditions, demonstrating the potential of spectral imaging in plant pathology. Moreover, multispectral imaging can be integrated with advanced computational techniques for enhanced analysis. In another study, Ref. [44] used ML combined with multispectral imaging and image processing techniques to detect aflatoxin contamination in figs.

Hyperspectral Imaging: Hyperspectral imaging is a technique that captures data across a continuous spectrum of wavelengths, providing significantly more detailed information compared with multispectral imaging. This method is particularly valuable for

the precise identification of toxigenic fungal contaminants and mycotoxins [45]. Hyperspectral images can be acquired using various platforms, including ground-based systems and unmanned aerial vehicles (UAVs). UAV hyperspectral imagery showed to effectively monitor *Fusarium* head blight in wheat fields, highlighting its potential for large-scale agricultural monitoring [46]. In another study, Ref. [47] used a visible and near-infrared hyperspectral imaging system operating in the range of 400–900 nm under ultraviolet excitation. They successfully differentiated spectral characteristics between corn kernels inoculated with aflatoxigenic *A. flavus* strains and naturally infected kernels from the same field. Furthermore, Ref. [48] explored the combination of fluorescence and reflectance visible and near-infrared hyperspectral images for detecting aflatoxin contamination in inoculated corn kernels in the field.

Ground vs. Intact Material: The context in which spectral data are collected can also vary. In some cases, imaging occurs on ground material, where samples are collected and analysed in a laboratory setting. This approach allows for controlled conditions and high-resolution data. In other instances, imaging is performed on intact material, such as whole peanut grains [49], to assess contamination directly in the field or during processing.

4.1.5. Limitations in Image Analysis

While image analysis using spectral data is a powerful tool for detecting mycotoxins, there are notable limitations and challenges. One significant factor is that visual features of an image, such as plant damage or fungal presence, may not always directly correlate with the presence of specific mycotoxins [50]. This is particularly relevant when different species of fungi, capable of producing various mycotoxins, are involved [51]. For example, certain fungi can cause visible damage or contamination on crops, which may be detected by ML models. However, these visual features might not indicate the presence of the specific mycotoxin of interest [50]. As a result, models focusing on plant damage or fungal contamination might not accurately reflect the levels of regulated mycotoxins. This discrepancy underscores the importance of integrating spectral imaging features that are more closely associated with the specific mycotoxins being regulated. Addressing this challenge requires combining image analysis with other data types, such as chemical analysis or molecular techniques, to improve the specificity and accuracy of mycotoxin detection. By doing so, ML models can better distinguish between general fungal contamination and the presence of specific harmful mycotoxins.

4.2. Neural Networks

Neural networks (NNs), first introduced by [52], are a class of machine learning algorithms modelled loosely after the human brain [53]. They are designed to identify patterns and make predictions by learning from data and can be used for supervised or unsupervised problems. NNs are made up of interconnected nodes and edges, where the nodes represent the *neurons* and the edges are the links between the neurons. The nodes are organised into layers, where the first layer is called the input layer, the last layer is the output layer, and all intermediate layers are called hidden layers. Typically, in an NN, the data are fed to the input layer; then one or more hidden layers perform computations and learn from the data, and finally, predictions (or classifications) are provided by the output layer. A simple diagram of an NN can be seen in Figure 4.

Every neuron in a hidden layer applies a weighted sum of the inputs to transform the data. This is followed by a function, referred to as an *activation* function [53]. The network fine-tunes the weights associated with each neuron by employing optimisation algorithms throughout the training phase. There are numerous hyperparameters associated with NNs. Some of the main hyperparameters include (i) the learning rate, which determines how much the weights are changed at each iteration; (ii) the number of epochs, which refers to how many times the entire training dataset is passed forward and backward through the neural network; (iii) the batch size, which controls the number of training examples used in one iteration; and (iv) activation functions like ReLU (Rectified Linear Unit), sigmoid, and

tanh, which determine the output value of a node given an input or a set of inputs. After training, the NN is capable of generating predictions for new, unseen data by passing the input across the layers to produce an output.

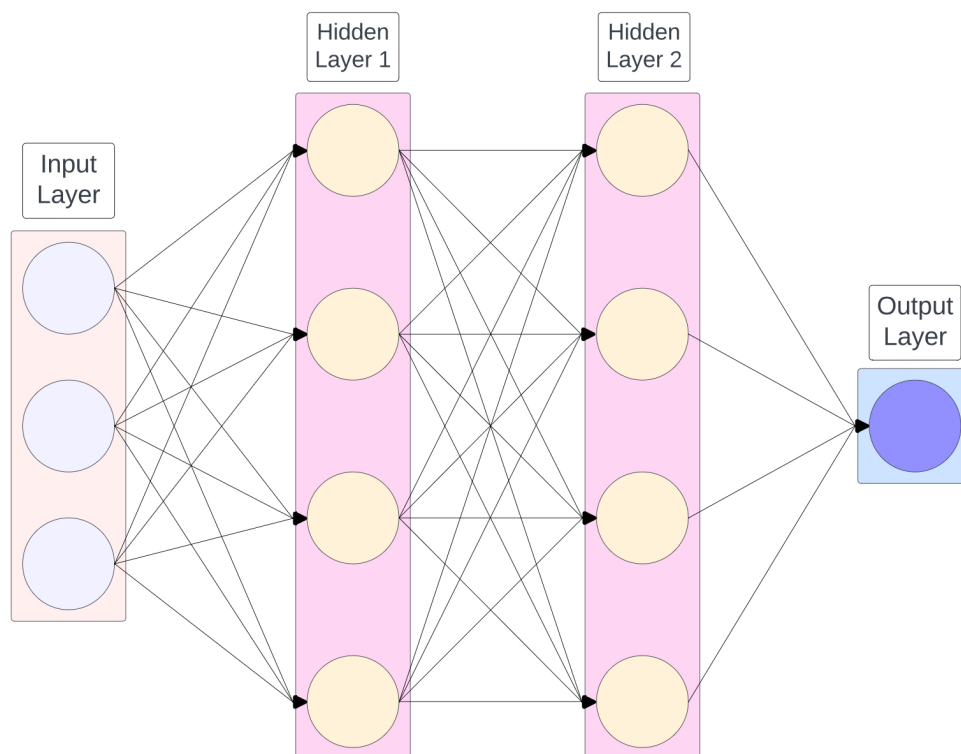


Figure 4. Basic neural network structure, showing an input layer, two hidden layers, and an output layer, where each circle represents a neuron, and these neurons are interconnected by lines symbolising neural connections. The input layer receives the initial data, which are then processed through successive hidden layers using weights and activation functions, refining the information before it reaches the output layer.

Like all machine learning models, NNs come with their own set of advantages and disadvantages. For example, NNs excel at identifying and modelling non-linear interactions present in data, which are common in biological processes. They are also flexible and can handle a wide range of data types, such as numerical and categorical, text, and image data. Despite their advantages, neural networks also have limitations. One of the major limitations is interpretability. NNs are considered *black-box* algorithms, meaning that it is difficult to understand why specific predictions are being made [54]. Second, like many of the other ML approaches we cover, they are not probabilistic models, making it hard to accurately quantify the uncertainty in the predictions. Overfitting can also be an issue for NNs. Without appropriate regularisation, NNs can become too complex, capturing the noise in the training data instead of generalising to the underlying pattern [55]. Finally, training large NNs requires a significant amount of computing power. The computational cost of NNs will increase with the complexity of the model [56]. In the following subsections, we review the use of NNs on different types of mycotoxin data.

4.2.1. NNs Applied to Spatiotemporal Data

NNs have been widely applied to spatiotemporal data, despite them not forming part of the traditional suite of spatiotemporal analytics techniques. In the field of mycotoxin study, NNs have been used for a variety of tasks and data types. For example, Ref. [38] used data from several sites in Northern Italy over the years 2005 to 2018. Their goal was to predict the presence of mycotoxins (specifically, aflatoxin and fumonisins) using NNs in

corn. In their work, they trained two NNs to predict if the contamination levels were above legal thresholds at the time of harvest. Both models performed well, achieving an accuracy of greater than 75% on the test data. However, they recommend, for future research, that improvements can be made to the modelling by taking into account the co-occurrence of aflatoxin and fumonisins in corn and their complex interaction, which may be due to the effects of climate change.

Ref. [57] applied NNs to analyse the concentration of mycotoxins in winter wheat grain. They examined 23 winter wheat genotypes with different *Fusarium* resistances from three different sites in Poland during the years 2011 to 2013. They developed three NN models; however, only two of these are concerned with the detection of mycotoxins, that is, the DONANN model, which is used to detect DON, and the NIVANN model, which examines the nivalenol content. The DONANN and NIVANN models were designed using an automatic network designer using Statistica v7.1 software [58], and were evaluated among a set of 10,000 generated networks. The performance of these models was assessed on several statistical metrics, but the primary focus was on the correlation coefficient (which, in this case, would be the correlation between the predicted values from the model and the actual observed values) and the mean absolute error (MAE), which is the absolute differences between the predicted values and the actual values. For the best-performing DONANN model, a low MAE of 0.37 was reported; however, the correlation coefficient was exceptionally high at 0.99, indicating an almost perfect linear relationship between the predicted and actual values. The best-performing NIVANN model, while exhibiting a slightly lower correlation coefficient of 0.81 and an MAE of 0.02, still performed within acceptable ranges. The architecture of the created models was designed as a multi-layer perceptron (MLP) type of NN, with two hidden layers. Despite reporting training, validation, and test errors, the authors did not specify the dataset on which the correlation and MAE metrics were based.

In a novel application of NNs, Ref. [59] used a transformer-based deep learning method, called *GPTransformer*. A transformer-based deep learning algorithm refers to a type of NN architecture that relies on a mechanism called *attention* to boost the performance of the model [60]. In their work, the authors proposed a transformer-based genomic prediction model for predicting *Fusarium* head blight disease levels and associated DON concentration in barley data collected in three locations in Canada over the years 2014 to 2015. One of their goals was to compare the accuracy of the GPTransformer model to existing genomic prediction methods such as decision tree algorithms (DT), linear regression (LReg), and traditional statistical algorithms like best linear unbiased prediction (BLUP). The authors used the Pearson correlation coefficient (PCC) as a measure of performance, which calculates the linear relation between the true output and the predicted output. They showed that the GPTransformer model (and all of the used ML models) did not significantly outperform the statistical method of BLUP in terms of predictive accuracy. However, GPTransformer did perform better than both the DT and LReg methods. The authors note that the ML methods used are able to capture non-additive genetic elements, and as such, the predictions provided might include some of these interactions in their estimations.

4.2.2. NNs Applied to Spectral Data

Hyperspectral (or just spectral) data refer to the capture and processing of information from across the electromagnetic spectrum [61]. Refs. [43,62,63] applied NN classification algorithms to pixels of hyperspectral image data to examine wheat for *Fusarium* head blight infection. Each author used a convolutional NN (CNN), which captures spatial patterns or motifs by identifying and calculating weights from the images according to how often the motif appears.

In Ref. [43], the authors investigated four distinct methods for converting hyperspectral imaging data. They then evaluated the performance of eight different CNN models in classifying pixels as either healthy or infected with *Fusarium* head blight. The effectivenesses of these models were compared based on their classification accuracy. They

found that a particular type of CNN called *DarkNet 19* [64] performed the best, with an accuracy of close to 100% across all data conversion methods, on both the validation and test data. For Ref. [63], tests showed that the CNN model is effective in detecting images that contain the blight and achieved an R^2 value of 0.80, and the mean average accuracy for the testing dataset was 92%. In Ref. [62], the authors compared the accuracies of the different NNs to determine which is the best at identifying diseased regions of the wheat kernel. They showed that a two-dimensional convolutional bidirectional gated recurrent unit NN performed the best, with an accuracy of 84.6% on the validation dataset and an F1 score and accuracy of 0.75 and 74.3%, respectively, on the test data.

Ref. [49] used a combination of hyperspectral data and NNs to detect aflatoxin in peanuts. They showed the CNN's efficacy in classifying infected peanuts and achieved a test set accuracy of 95%. They later expanded their work and used a one-dimensional CNN (1D-CNN) to classify aflatoxin infection in corn and peanuts. This time, they achieved accuracies of 96.4% for peanuts and 92.1% for corn [65].

In a research conducted by [66], infrared (IR) spectroscopy and ML algorithms were used to detect fungal contamination in corn. In their study, 183 naturally infected samples (contaminated with different *Fusarium* DON species and at different concentrations) were obtained from the seed production Linz of Austria (SBL) and from the Cereal Research Centre of Hungary (CRC). The authors assessed several classification ML models, including multi-layer perceptron (MLP) neural networks, random forests, support vector machines, and adaptive boosting, for their accuracy in correctly classifying contaminated from non-contaminated samples. Their results showed that the MLP approach correctly classified 94% of the non-contaminated samples and 91% of the contaminated samples. The authors note that while this approach yields promising results, these findings are specific to a contamination threshold of 1250 mg/kg, which is the EU regulatory limit, and that subsequent research will aim to evaluate the performance of the classification methods across various contamination levels.

4.2.3. NNs with an Electronic Nose

An electronic nose (e-nose) is a device intended to detect chemical compounds in gasses. E-noses have been extensively used in the detection of aflatoxins [67,68], fumonisins [69], and DON [70] in corn. However, Ref. [71] used an e-nose supported by NNs for the detection of aflatoxin and fumonisins in corn. In their work, they compared three different approaches, that is, NN, logistic regression (LR), and discriminant analysis (DA), to examine the e-nose's ability to discriminate between samples contaminated with concentrations either exceeding or falling below legal thresholds on data spanning 5 years. They showed that all methodologies achieve an accuracy of above 70%, with the NN performing the best with an accuracy of 78% for aflatoxin detection and 77% for fumonisin detection. They went on to suggest that the e-nose, when supported by an NN, can provide a fast screening tool for classifying samples.

4.2.4. NN Summary

Neural Networks have been widely adopted as the ML algorithm of choice for analysing mycotoxin data, especially in the field of hyperspectral imaging. However, as of yet, there seems to be a gap between research applications and the wider use in industry. The application of NNs in hyperspectral data for mycotoxin detection (and food safety in general) is a relatively new process, and the implementation of an NN approach to hyperspectral data in industrial quality control faces various challenges, mainly due to hardware limitations, such as the cost of operating imaging equipment [72]. However, in research, NNs for use in hyperspectral imaging have seen an increase in popularity with many of the reviewed works being widely cited, for example, Refs. [62,63].

4.3. Random Forests

A random forest (RF) [73] is an ensemble learning method used for classification and regression. The RF algorithm creates a *forest* of decision trees, where each tree in the forest is built from a sample drawn with replacement (that is, a bootstrap sample) from the training set and selects splits from a random subset of features.

While Section 4.6.1 provides a comprehensive examination of decision trees, this section offers a concise introduction to familiarise readers with the basic concepts and terminologies associated with decision trees. Figure 5 shows an example of a single decision tree. In constructing each decision tree, the root node is the starting point, and it represents the entire dataset, which gets split based on a feature that provides the best separation according to a certain criterion [like Gini impurity [74]]. The decision nodes are the points where the data are split further. Each decision node represents a decision rule on a specific feature. The process continues recursively until a stopping criterion is met, such as reaching the tree's maximum depth, attaining a minimum sample count in a leaf, or achieving adequate purity within the leaf nodes. The leaf/terminal nodes represent the final output of the decision process. Each branch/sub-tree represents a possible outcome of the decision made at the decision node, leading to further sub-trees or leaf nodes.

For RF classification tasks, each tree in the forest votes for a class, and the class receiving the majority of votes becomes the model's prediction. For regression tasks, the forest takes the average of the outputs by individual trees. Figure 6 shows a summary of the RF algorithm.

One of the main advantages of using RFs is their versatility. They are capable of performing both regression and classification tasks, as well as handling large datasets. Additionally, they require very little tuning and can perform well without much hyperparameter optimisation. Some of the main hyperparameters associated with RF include the following: (i) Number of trees: this is the number of trees in the forest. Generally, more trees increase performance but also increase the computational cost. (ii) Maximum depth of trees: the maximum depth of each tree. Deeper trees can model more complex patterns but might lead to overfitting. (iii) Minimum samples split: the smallest number of samples needed to split an internal node. Setting higher values helps prevent the model from learning overly specific patterns, which can lead to overfitting. As with NNs, RFs are a black-box algorithm, and so interpretability can be an issue. Each decision tree upon which the RF is built can be easy to interpret, but since RFs consist of a large number of decision trees averaged together, the decision process by which a prediction is made can be somewhat opaque.

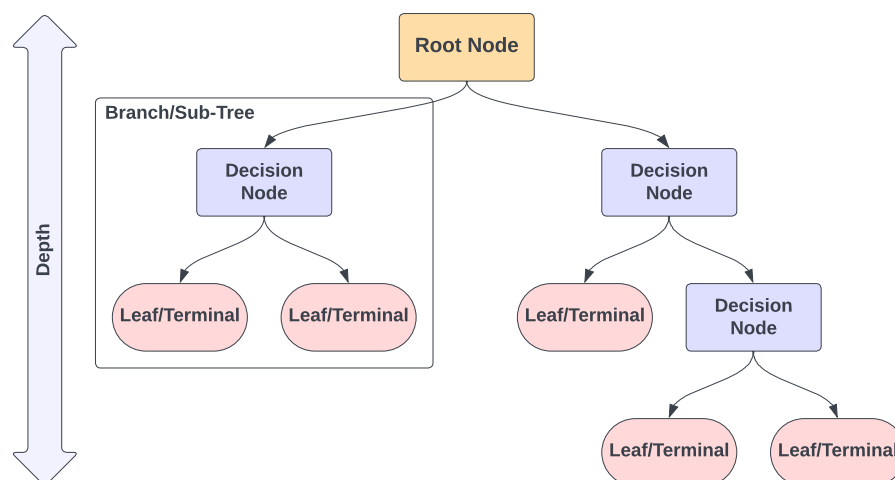


Figure 5. Decision tree process demonstrating the structure of a decision tree, including the root node, branching to decision nodes, and culminating in leaf/terminal nodes. The depth of the tree is indicated, showing the levels of decision making from the root to the leaves.

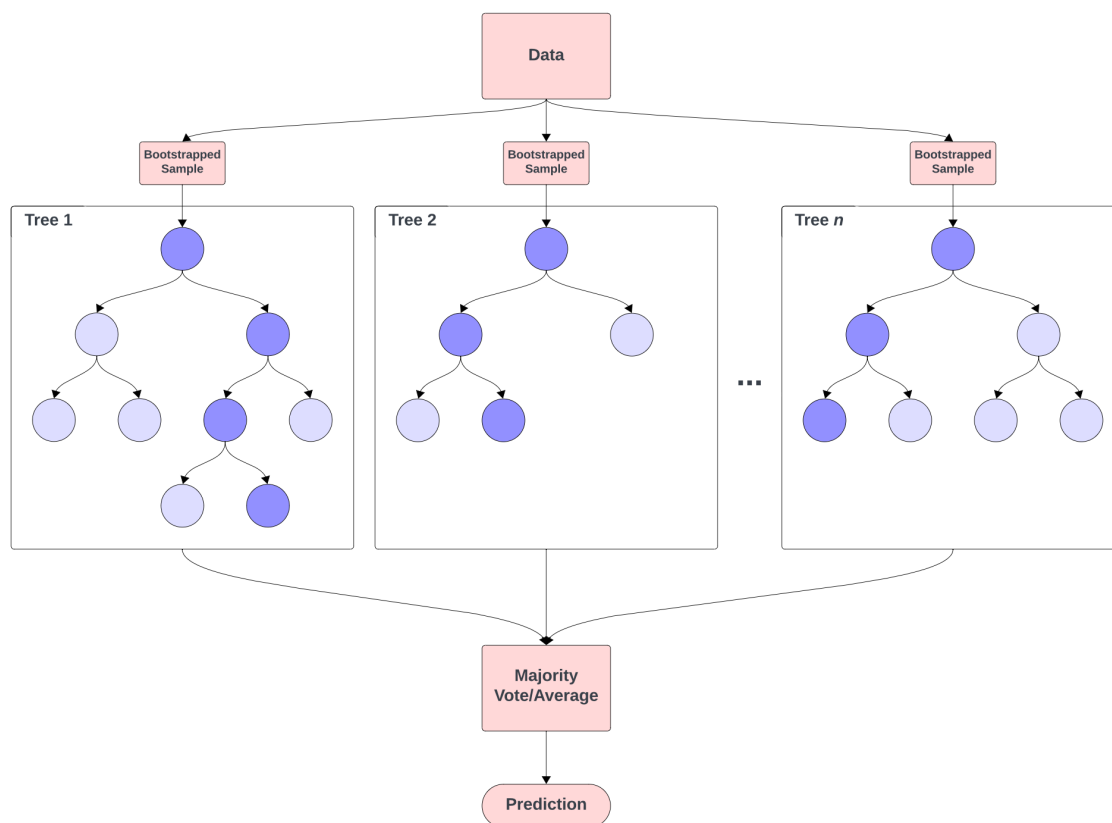


Figure 6. The random forest algorithm constructs an ensemble of decision trees, with each tree built from a unique bootstrapped sample of the original dataset. Nodes are colored light blue to represent the regular decision nodes of the trees. Distinct paths through each tree are shown, highlighted by the darker blue nodes, and represent a sequence of decisions made from the root to a leaf node based on the input features. The final prediction of the random forest is determined by aggregating the predictions of all trees, using majority voting for classification tasks or mean prediction for regression tasks.

4.3.1. RFs for Spectral Data

As with NNs, RFs have been applied to hyperspectral data. For example, Ref. [75] used a RF classification model to classify corn silage for high or low mycotoxin contamination using near-infrared spectroscopy (NIR). In their study, 155 samples were collected from several sites in the Po Valley (Italy) and from Sardinia over the years 2017 to 2019. Their aim was to develop qualitative models capable of distinguishing corn silage based on either the total concentrations or the total counts of various groups of mycotoxins (in this case, *Fusarium* and *Penicillium* toxins). To evaluate various classification strategies, different distinct threshold levels were established for each mycotoxin contamination. These thresholds were used to categorise each sample as having either a high or low contamination level in relation to these specified values. To predict the contamination level, an RF classification model was fitted, using the wavelength of light as the predictors, and achieved an out-of-sample accuracy of above 90% for the classification of both *Fusarium* and *penicillium* toxins.

In a 2023 study, Ref. [76] utilised NIR spectroscopy for detecting DON in oat samples from Spain and Sweden collected over the years 2021–2022. The authors applied two different transformation techniques to the spectral data and examined which allowed for greater classification of the data using four different ML algorithms (k-nearest neighbours, naïve Bayes, NN, and RF). Both preprocessing transformation methods achieved similar results for all ML methods, with RFs performing the best with an accuracy of 77.8% and an

area under the curve (AUC) of around 0.77. However, they noted that other similar studies have been conducted that achieved a higher classification accuracy, such as [77].

In a similar study, Ref. [78] constructed a biosensor array for identifying mycotoxins in peanuts and corn, produced by *Aspergillus flavus*, using six ML models, including partial least square determination analysis (sPLS-DA), linear support vector machine (svmLinear), radial support vector machine (svmRadial), RF, NN, and high-dimensional discriminant analysis (HDDA). The authors used the classification models for three separate purposes: to distinguish healthy from infected samples, to distinguish the pre-mould status in infected samples, and to distinguish between infected peanuts or corn samples. To distinguish the pre-mould status, the aim was to create a three-class model to predict either the control or 1 or 2 days after inoculation. Their approach achieved a reported 100% accuracy in distinguishing healthy from infected samples and RF accuracies of 95% and 98% in identifying pre-mould status in peanuts and corn, respectively. However, such high levels of accuracy warrant further investigation, as such high accuracy rates can often be indicative of issues in the experimental design, such as the creation of non-representative test sets or overfitting, especially if the test sets are not properly randomised.

4.3.2. RFs for Mycotoxin Treatment

ML models in mycotoxin treatment can be used to predict mycotoxin contamination risk and optimise mitigation strategies. This application can boost accuracy in prediction and effectiveness in deploying targeted anti-fungal treatments. In a study conducted by [79], the authors employed machine learning techniques to predict the growth of *Fusarium culmorum* and *Fusarium proliferatum*, as well as their production of mycotoxins, in environments where ethylene vinyl alcohol copolymer films are used. These films contain pure components of essential oils, which are used to inhibit the growth of the fungi and their mycotoxin production. In their work, they studied fungal growth on corn in vitro and modelled the fungal growth and toxin production under different environmental scenarios and with different treatments applied. The ML models used were NNs, RF, extreme gradient boosted trees (XGB), and multiple linear regression (MLR). The performance of the ML methods was assessed using the root mean square error (RMSE). It was found that RF performed the best in predicting the growth rates of *Fusarium culmorum* and *Fusarium proliferatum* and mycotoxin production, having consistently the lowest RMSE value.

Ref. [80] evaluated the anti-fungal properties of specific lactic acid bacteria strains against *Fusarium* species found in cereals. To achieve this, various machine learning algorithms, including NN, RF, XGB, and MLR, were employed to predict the extent of fungal growth inhibition resulting from the application of the tested lactic acid bacteria strains. As with the previous study, the RMSE was the metric used to assess the performance of the model, in conjunction with the R^2 value. In this work, both RF and XGB showed comparable performances, reporting similar RMSE (0.0604 and 0.0581, respectively) and R^2 values (0.992 and 0.992, respectively) on the test data, in predicting the percentage of growth inhibition.

Several other studies exist on the topic of using ML models (and specifically RF) to predict mycotoxin growth in the presence of treatments. In the interest of brevity and space, we name them here but do not provide additional details of the studies. In each of these studies, the authors used multiple ML models, with a general consensus that RF models performed the best at their given tasks. See Refs. [81–83] for more details.

4.3.3. Random Forest Summary

RFs have emerged as a robust and versatile tool in the field of mycotoxin detection and treatment and have gained popularity due to their ease of use, computational speed, and predictive performance. These studies collectively underline the significant potential of RF in enhancing food safety measures, although it is crucial to acknowledge the necessity for rigorous validation and testing to ensure the reliability of these models.

4.4. Gradient Boosting

Gradient boosting (GB) [84] builds on the concept of boosting, where weak learners are converted into strong ones through an iterative process. The GB framework builds boosted regression models by sequentially training a weak classifier (such as a linear regression or simple decision tree) successively on the data using the residuals from previous model fits (as shown in Figure 7). This process ensures that each new weak classifier addresses the inaccuracies of its predecessors, thereby enhancing the prediction accuracy. The final model aggregates the outputs from all these weak classifiers to form a robust, ‘strong’ classifier through an ensemble approach. The term *gradient* in gradient boosting refers to the method’s use of gradient descent, a numerical optimisation algorithm, to minimise the loss or the difference between the actual and predicted values.

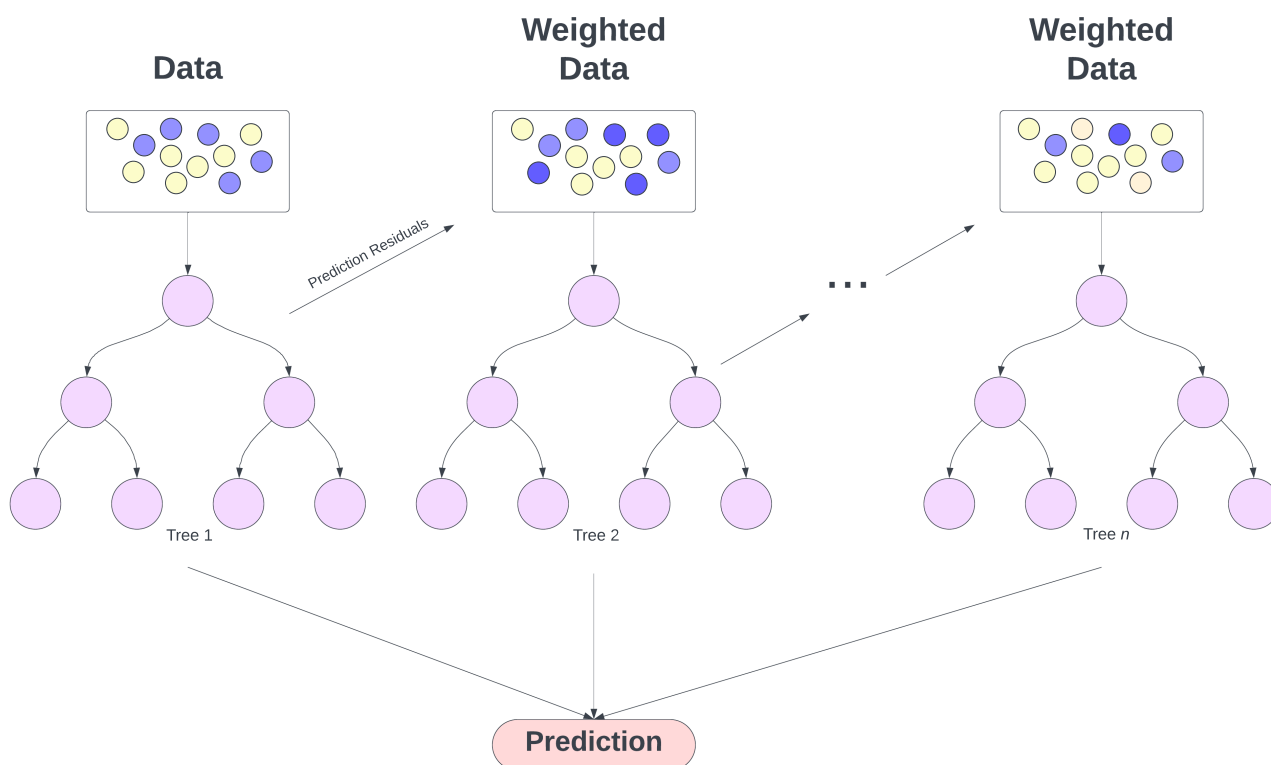


Figure 7. Gradient boosting process. Here, the weak learners are trees that are trained sequentially on weighted data with iteratively adjusted weights based on previous prediction errors. The light yellow circles represent data points with lower residuals (errors), the light blue circles represent data points with moderate residuals, and the dark blue circles represent data points with higher residuals from previous model predictions. The pink circles within the trees indicate the decision nodes of each weak learner. The final prediction is made by aggregating the outputs from all weak learners.

In gradient boosting, when the weak learners are decision trees, each tree is grown in a greedy manner, but unlike random forests, trees are grown sequentially. After the first tree is built and predictions are made, the errors (residuals) from those predictions are used to build the next tree. The subsequent tree aims to predict the residuals from the previous tree. This process is continued, with each new tree correcting the residuals of the ensemble of all previous trees. The final prediction is made by summing the predictions from all trees, which can be thought of as a weighted vote where trees that reduce the error the most have more influence.

An advantage of GB models is their strong predictive capability and adaptability, especially in dealing with complex non-linear relationships between independent variables and the dependent variable. They adapt to various prediction problems by supporting different loss functions, making them suitable for both regression and classification tasks.

However, these models have their challenges. Without careful tuning and regularisation, there is a risk of overfitting, a problem exacerbated by noisy data [85]. Additionally, their sequential boosting process is computationally intensive and time-consuming compared with methods like random forests that build trees in parallel. This complexity can be a significant drawback in scenarios where computational resources or time are limited. Some of the main hyperparameters associated with GB are as follows: (i) Number of weak learners: this defines the number of boosting stages or learners to be created. More learners can lead to a more powerful model, but also increase the risk of overfitting and raise computational cost. (ii) Learning rate: this parameter scales the contribution of each learner. A smaller learning rate requires more weak learners but can yield a more generalised model. In the case of the weak learner being trees, (iii) the maximum depth of trees determines the maximum depth of each individual tree. Deeper trees can model more complex patterns but can also lead to overfitting. An extension of a GBM model is called eXtreme Gradient Boosting (XGB) [86], with the key difference between the two being performance. In general, XGB models are faster and have better optimisation. Additionally, XGB models have the ability to deal with missing values.

4.4.1. GB for Spatiotemporal Data

In a study by [87], the authors designed a program for aflatoxin monitoring in feed products (peanuts and soy beans), while considering both the performance of the model and the cost of monitoring. In the study, they applied four different ML algorithms (namely, GB, LR, SVM, and DT) to historical data concerning monitoring for the presence of aflatoxins in feed products. The data were collected from several sites around the world, including China, Brazil, and Argentina, over the years 2005 to 2018. The ML algorithms were compared to predict which feed batches are high risk and which should be considered for further aflatoxin analysis. In their work, they found that all the ML models performed well and used several error metrics to assess their models. They obtained an accuracy of over 90% for all models and an AUC and recall of over 0.8 and 0.6, respectively. However, the XGB model performed better than all other models, and the authors proposed a reduction to the monitoring cost of up to 96% for the years 2016 to 2018.

In Ref. [88], the authors proposed to use un-targeted metabolomics and ML techniques to mine biomarkers of the species *Aspergillus* on peanut data collected from several sites in China over the years 2013 to 2018. They initially used an RF model to determine *Aspergillus* species with 97.8% accuracy. They then went on to use XGB to create a decision rule to help regulators in evaluating risk prioritisation with a claimed accuracy of 87.2%. However, the authors noted that they built the XGB model using only a single tree and used this tree to create an operable decision workflow for risk assessment. Although using a single tree can reduce complexity, it also increases the likelihood of less robust predictions. Part of the strength of XGB (and GBM) models is that they iteratively correct the mistakes of previous trees, a process that is lost if only a single tree is used.

Ref. [39] conducted a study with the objective of evaluating the performance of GBM models to predict the presence of aflatoxins in corn at two risk thresholds, that is, 20 ppb and 5 ppb. These cut-off values were chosen based on the U.S. Food and Drug Administration's (FDA) action level for corn (20 ppb) [89], whereas the lower cut off is based on the European standard of 5 ppb [90]. Additionally, the authors performed feature engineering, which is the process of transforming raw data into meaningful and informative features with the intention of enhancing the performance of ML algorithms [91]. The data used were historical climate, soil, and aflatoxin data, collected in several sites in Iowa in the years 2010, 2011, 2012, and 2021. As the data had many missing values, the authors used an imputation method; however, they noted that data from the months of January, February, and December had to be excluded from the model as there were too many missing values to accurately impute the data. The authors reported that the GBM model performed well, achieving high accuracy rates of 96.8% for the 20 ppb threshold and 90.3% for the 5 ppb threshold. The study highlighted the significant influence of the vegetative index (which

is a quantitative measure that uses satellite imagery to assess the amount and health of plant life in a specific area) in August on aflatoxin risk for both thresholds, indicating the critical environmental and ecological impact of drought conditions during this month. Additionally, predictors related to soil properties (such as hydraulic conductivity, pH, and bulk density) were found to potentially affect aflatoxin contamination levels before harvest.

4.4.2. GB for Spectral Data

Ref. [92] conducted a study on aflatoxin and fumonisin contamination in a single kernel corn. They argued that bulk sampling of the corn may not produce accurate results, and thus focus solely on single kernels. In their study, they performed measurements to show the skewness of the data and calculated weighted sums of toxin contamination. Additionally, they aimed to improve single kernel classification performance through the use of different ML applications. Their methodology was to take corn kernels that were already contaminated and scan them using the NIR technique. The samples were then ground and measured for both toxins using the ELISA method (discussed in Section 1). In their work, they used five different ML models to classify both mycotoxins. They are GBM, RF, least absolute shrinkage and selection operator (LASSO), elastic-net regularised generalised linear models (GLMNETs), and support vector machines (SVMs). They additionally applied ML algorithms for classifying each individual mycotoxin. For aflatoxin, they used bagged AdaBoost, linear discriminant analysis (LDA), and penalised logistic regression (PLR). For fumonisin classification, GBM and penalised discriminant analysis (PDA) were used. For aflatoxin, they found that GBM was the best-performing model, with an accuracy of 83%, on both the training and the test data. For fumonisin, the PDA model performed the best with an accuracy of 86% on the test data. However, the authors noted that, for future studies, opportunities for better classification exist, including increasing the proportion of samples so the algorithm can learn the characteristics of contaminated corn kernels better.

4.4.3. Gradient Boosting Summary

The application of GBM models across various datasets, from spatiotemporal to spectral data, demonstrate their versatility and potential in predicting mycotoxin contamination in agricultural products. While GBM models generally exhibit high accuracy, there are criticisms concerning the robustness of these models when applied with limited trees, as in the case of [88], or when handling datasets with substantial missing values, as noted by [39]. The high accuracy rates reported should be examined for potential overfitting or lack of generalisation to broader datasets. The approach of ref. [92] to single kernel analysis opens avenues for improved precision in toxin detection, but also indicates the need for larger sample sizes to enhance model learning.

4.5. Support Vector Machines

Support vector machines (SVMs) [93] are a set of supervised learning methods used for classification, regression, and outlier detection. To make predictions, SVMs identify the optimal hyperplane that maximises the margin between the two classes (where the margin is defined as the distance between the nearest data points of each class and the dividing hyperplane). The data points that are closest to the hyperplane and that influence its position and orientation are known as support vectors, as they *support* or define the hyperplane. Figure 8 illustrates an SVM in action. One of the key advantages of SVMs is their versatility as they can be used on a variety of data types, and are particularly useful for image recognition [94]. Additionally, they are memory efficient since they only use a subset of training points, called support vectors, in the decision function. However, SVMs require careful tuning of the hyperparameters and an appropriate kernel choice. A kernel is a function used to transform data into a higher-dimensional space. By projecting the data into a higher dimension, a kernel makes it possible to find a hyperplane that can effectively separate the classes. Some of the common kernels include [95]:

1. Linear: No non-linear transformation, suitable for linearly separable data.
2. Polynomial: Suitable for non-linearly separable data, involves higher degree terms of the features.
3. Radial basis function: Good for non-linear data, uses a Gaussian distribution.
4. Sigmoid: Similar to the sigmoid function in logistic regression.

Additional hyperparameters include the following: (i) Gamma: This is needed for all kernels except linear. It determines the extent of the influence that a single training example has. Low values indicate a wide reach, and high values indicate a close reach. A high gamma value can cause the model to overfit. (ii) Degree: This is only relevant for a polynomial kernel. It defines the degree of the polynomial used in the kernel. A higher degree can model more complex relationships but increases the risk of overfitting. (iii) Coef0: This is a parameter for polynomial and sigmoid kernels that adjusts the independent term in the kernel function. It is often called the kernel bias.

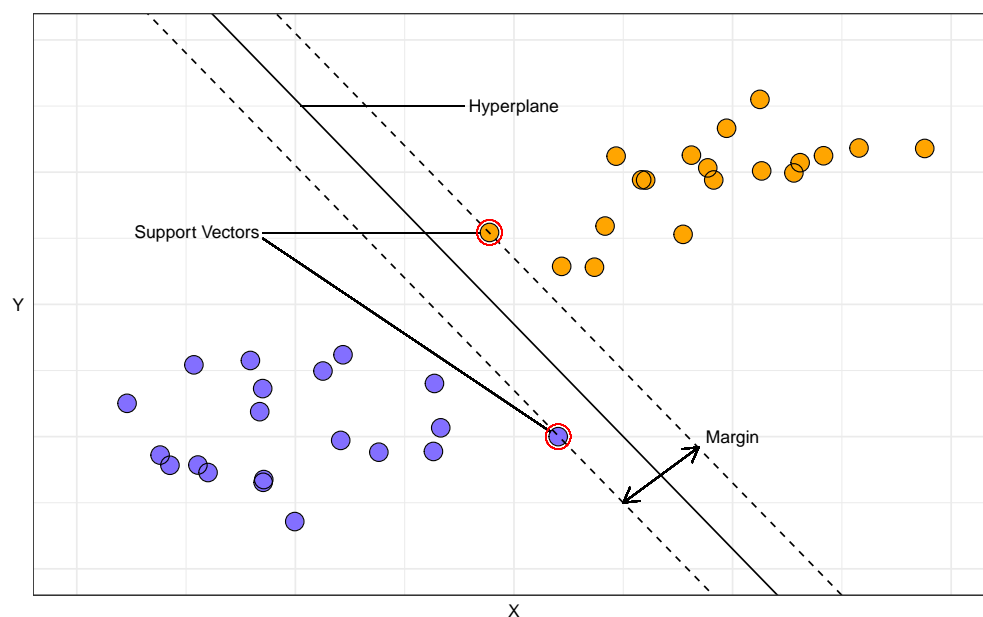


Figure 8. Support vector machine process. The diagram illustrates the SVM's method of finding the optimal hyperplane that maximises the margin between two classes, depicted by the blue and orange points. The support vectors, which are the data points closest to the decision boundary, define the margin.

4.5.1. SVMs for Spectral Data

In the review of the literature concerning the use of SVMs in mycotoxin detection, it was found that they were overwhelmingly used for image recognition and, as such, primarily used spectral data. For example, ref. [45] used several ML models (SVM, NN, and LR) for the classification of *Fusarium* head blight in wheat, using spectral data. The data were collected in the years 2020 to 2021 at a single site in Belgium, with the experiment using eight varieties of wheat. They found that the SVM model outperformed both the NN and LR method in classifying contaminated wheat in every variety, with a classification accuracy of 96.5% on the test data (with NN and LR achieving accuracies of 82.9% and 82.5%, respectively).

In a similar study, Ref. [96] used three different imaging methods alongside ML classification models to test ground corn samples for the presence of aflatoxin and fumonisin, both as individual contaminants and in combination. Two classification models were used, partial least squares-discriminant analysis (PLS-DA) and SVM, using specific threshold values for each mycotoxin. The naturally contaminated corn samples were obtained from the Office of Texas State Chemist, which in turn collected the samples from different feed companies located around Texas. They found that the SVM performed better than

the PLS-DA with classification accuracies of 89.1%, 71.7%, and 95.7% for each imaging technique. The imaging method with the highest accuracy was the short-wave infrared (SWIR) method.

In a study concerning the detection of *Aspergillus parasiticus* in corn kernels using NIR hyperspectral imaging, conducted by [97], the authors used SVMs to compare the performances of multiple different preprocessing and imaging techniques. For their study, corn kernels were harvested from Hefei City, Anhui Province, China, in 2015. Each day (for a period 7 days), 36 sterilised corn kernels were inoculated with *Aspergillus parasiticus* and were grouped into four groups depending on the day of inoculation. From this, an SVM was used to determine which groups were infected using different preprocessing techniques. Additionally, this study examined the orientation of the kernel in the image to determine if this property had an effect on predictive performance. They found that the best preprocessing method was a combination of the standard normal variate (SNV) and moving average smoothing (MAS) methods, with an accuracy of 91.67% for detecting contaminated kernels using the validation data. They also found that the performance of the classified models was influenced by orientation; however, the models built using data from a mix of kernels with their germs facing both up and down still achieved an accuracy of 84.38% on the validation data.

4.5.2. Support Vector Machine Summary

In the reviewed work, SVMs demonstrated considerable accuracy in mycotoxin detection through spectral data analysis. However, as with other ML methods reviewed, the consistently high classification accuracy reported raises questions about potential overfitting and the representativeness of the datasets used. Moreover, factors such as kernel orientation (which refers to the way in which the kernel function transforms the input data into a higher-dimensional space to find an optimal boundary between classes) significantly influenced SVM performance, indicating that model robustness may be context dependent. The choice of kernel and its parameters, like orientation, scale, and type, is critical in shaping the decision surface and, thus, the SVM's ability to generalise from training to unseen data.

4.6. Other ML Methods

In this section, we cover the remaining ML methods. These include decision trees and Bayesian networks and have been grouped together as they make up a minority of the reviewed work. As such, they are not separated by the type of data used, and all data types are discussed together.

4.6.1. Decision Trees

Decision tree (DT) learning is a type of non-parametric supervised learning algorithm used for both classification and regression tasks [74,98]. A DT is a flowchart-like structure, resembling a tree structure with branches representing decision paths and leaves (or terminal nodes) representing predicted outcomes (see Figure 5 in Section 4.3). A DT splits the data into subsets based on the value of input features. Splits are chosen to maximise the separation of the classes based on measures like Gini impurity or information gain [74]. This process continues recursively until a stopping criterion is met, resulting in a tree where each path represents a decision pathway that leads to a predicted outcome. The advantages of decision trees include their simplicity, interpretability, and ability to handle both numerical and categorical data. However, DTs have a tendency to overfit, especially when a tree is particularly deep [74]. This can be mitigated by pruning the tree or setting a maximum depth of the tree via the use of hyperparameters. As this method is a tree-based approach, there is an overlap with RF and GB in terms of hyperparameters. Some of these include maximum depth, minimum samples split, and minimum samples leaf (i.e., the minimum number of samples needed to be at a leaf node). Setting this parameter can ensure

that each leaf node represents a reasonable number of samples, which can smooth the model, particularly for regression tasks, and prevent overfitting).

The use of DTs in the field of mycotoxin detection is quite varied. For example, in a study conducted by [99], in which they assessed the use of an electronic nose to identify DON contamination of wheat samples, an extension of decision trees called Classification and Regression Trees (CART) [74] was used to classify samples based on four thresholds of DON contamination (1750, 1250, 750, and 500 $\mu\text{g}/\text{kg}$). For this study, 214 wheat samples were collected from Northern Italy during the years 2014–2015 and 2017–2018. For the threshold values of ≥ 1250 $\mu\text{g}/\text{kg}$, the accuracy of sample classification was the highest, ranging between 88% and 92%. The lower thresholds of ≤ 750 $\mu\text{g}/\text{kg}$ were found to be the least accurate, with an accuracy of $< 83\%$. The authors proposed that the reduced sensitivity of the instrument at lower DON concentrations might explain this drop in accuracy.

Ref. [99] examined the classification of DON mycotoxin-contaminated corn and peanuts at regulatory limits using spectral data. The spectral data were analysed using a bootstrap-aggregated (bagged) DT approach, focusing on the protein and carbohydrate absorption bands of the spectrum. The corn samples were obtained by Saatbau Linz (Linz, Austria) and the Cereal Research Centre (Szeged, Hungary). For the peanuts, 92 different infected samples were purchased from public markets in Tanzania, Mozambique, and Burkina Faso. The authors demonstrated that the DT method could classify corn samples at the 1750 and 500 $\mu\text{g}/\text{kg}$ thresholds for DON with accuracies of 79% and 85%, respectively. Additionally, it was able to classify peanut samples for aflatoxin at 8 $\mu\text{g}/\text{kg}$ with a 77% accuracy.

In a study related to identifying and predicting risks related to the presence of fumonisins in breakfast cereal products, Ref. [100] developed a model specifically designed to predict the risk of fumonisin contamination, with a particular emphasis on a mixture of ingredients. In their research, fifty-eight distinct breakfast products were purchased from local grocery stores in Florence, Italy, during 2019. The selection criteria for purchasing breakfast products included (i) products with packaging sizes ranging from 200 to 500 g, including both plastic and non-plastic materials; (ii) items sourced from retail shops; and (iii) products primarily made of wheat, corn, dry fruits, rice, and oats. Principal component analysis (PCA) and k-means clustering were employed to explore the connection between cereal ingredients, their composition and packaging, and the concentration of fumonisins. The findings suggested that the fumonisin concentration might be linked to complex non-linear interactions among various factor variables. To explore this potential and identify the factors most closely linked with high concentrations, DTs were employed. Two decision trees (DTs) were developed, with the first indicating a relationship between high concentrations of fumonisins and cereal products rich in corn, particularly when combined with high levels of sodium or rice. The second tree highlighted a link between corn and either high sodium or high-fat concentrations. In both models, the presence of plastic packaging appeared to mitigate the concentration of fumonisins to a certain degree.

4.6.2. Bayesian Network

Bayesian networks (BN) are a type of probabilistic graphical model that uses Bayesian statistics to represent and infer the conditional dependencies between different variables in a dataset [101]. The networks are structured as a directed acyclic graph (DAG), with feature nodes representing variables and edges indicating probabilistic relationships between them. Predictions in BNs are made through a process called probabilistic inference, which involves calculating the likelihood of certain outcomes based on known information and the network's structure. In contrast with linear regression models, BN models excel at analysing variable dependencies, handling non-linear interactions, and incorporating diverse types of data [102]. The strengths of BN include the handling of uncertainty, the integration of prior knowledge with observed data (thereby enhancing the model's predictive capabilities), and interpretability. However, some disadvantages of using BNs exist. As the number

of variables increases, the complexity of the network and the computational resources required for inference can grow exponentially.

In a study aimed for predicting DON contamination in wheat, ref. [103] compared three different modelling approaches. These are a mixed effect LR method, a mechanistic model (which simulates the mechanisms of plant and fungus development stages and their interactions) adapted to the current data, and a BN. These were all used to predict DON contamination. The data used were collected in the Netherlands over the years 2001 to 2013. The results of the experiments showed that all three models performed well, with the LR method performing the best, achieving an accuracy of 88% for detecting DON contamination. However, the authors noted that this model is greatly reliant on both the specific location and the available data, and it requires that all input data be present. The mechanistic model achieved an accuracy of 80%, while the BN achieved an 86% accuracy. However, the authors noted that the BN is easier to implement when the data are incomplete, when compared with the other methods.

Ref. [104] constructed transcriptional regulatory networks (TRNs) using a BN algorithm called the *module network* algorithm. TRNs are complex systems in biology that describe the relationships and interactions between various proteins and genes involved in the process of *transcription* [105], where transcription is the process by which the information encoded in a section of the DNA is transcribed to produce a complementary RNA strand. The goal of their work was to understand how specific gene groups (modules) in the fungus *Fusarium graminearum* regulate biological processes. The authors reported that their network inference is of high credibility, with 81.8% of the evaluable modules classified as high or moderate confidence based on their validation against a variety of evidence sources. This suggests a robust alignment of the inferred network with the existing understanding of the biological processes within *Fusarium graminearum*.

4.6.3. Summary of Other ML Methods

Decision trees have shown varying degrees of effectiveness in detecting mycotoxins, as evidenced by diverse research outcomes. The use of CART to classify contaminated wheat samples achieved higher accuracy at certain thresholds but showed diminished performance at lower contamination levels. A bagged DT approach showed moderate success, suggesting that while DTs are capable classifiers, their accuracy can vary significantly based on the mycotoxin levels and sample types. The application of these methods includes potential issues with model sensitivity, particularly at lower toxin concentrations, and a reliance on the quality of the data. These factors underscore the need for a careful calibration and validation of DTs in diverse settings for reliable mycotoxin detection.

BNs have shown effectiveness in mycotoxin detection, as demonstrated in various studies, but with some limitations. Ref. [103] compared BNs with other models for predicting DON contamination in wheat, achieving a respectable 86% accuracy. However, they highlighted BNs' advantage in handling incomplete data, a significant benefit over other methods like logistic regression. The reviewed applications show BNs' flexibility and efficiency, though their performance can be contingent on data completeness and specific biological contexts, which may limit their broader applicability.

4.7. Summary and Comparison of Case Studies

To provide a comprehensive overview of the specific case studies discussed, here, we include a summary table in Table 1 that highlights the key findings by describing the data types, ML models used, application contexts, and reported accuracies. In cases where more than one ML model is used, the highest-performing model is reported in the accuracy column.

Examining Table 1, we can see that the most frequently used ML model in the reviewed studies is the neural network, with convolutional neural networks also being highly prevalent. The most common data type used is spatiotemporal data, followed by hyperspectral data. The research covers a range of crops, including corn, wheat, barley, peanuts, and

oats, with a primary focus on detecting contaminants such as aflatoxin, fumonisins, and Fusarium head blight. However, the most commonly studied crop is corn. Many studies achieved high accuracy rates, often above 90%, showcasing the potential of ML models to enhance mycotoxin detection in agriculture. However, it is important to consider that these high accuracies may be influenced by the controlled environments of individual laboratories, which can lead to overfitting and potentially less reliable performance in real-world applications (see Section 5 for a discussion on this).

Table 1. Summary of reviewed mycotoxin detection studies. In cases where more than one ML model is used, the highest-performing model is reported.

Study	Data Type	ML Model	Application Context	Accuracy
Camardo et al., 2021 [38]	Spatiotemporal	NN	Corn in Northern Italy (2005–2018)	>75%
Niedbala et al., 2020 [57]	Spatiotemporal	NN	Winter wheat in Poland (2011–2013)	DONANN: 99%, NIVANN: 81%
Jubair et al., 2021 [59]	Spatiotemporal	GPTransformer	Barley in Canada (2014–2015)	Not significantly better than BLUP
Rangarajan et al., 2022 [43]	Hyperspectral	CNN (DarkNet 19)	Wheat for Fusarium head blight	100%
Qiu et al., 2019 [63]	Hyperspectral	CNN	Wheat for Fusarium head blight	92%
Jin et al., 2018 [62]	Hyperspectral	CNN (2D conv. bidirectional GRU)	Wheat for Fusarium head blight	84.6%
Han et al., 2019 [49]	Hyperspectral	CNN	Peanuts for aflatoxin	95%
Han et al., 2019 [49]	Hyperspectral	CNN	Corn and peanuts for aflatoxin	96% corn, 92% peanuts
Gao et al., 2021 [65]	Hyperspectral	1D-CNN	Peanuts and corn for aflatoxin	Peanuts: 96.4%, corn: 92.1%
Öner et al., 2019 [66]	Infrared spectroscopy	MLP NN, RF, SVM, adaptive boosting	Corn for fungal contamination	MLP: 91%
Leggieri et al., 2021 [71]	E-nose	NN, LR, DA	Corn for aflatoxin and fumonisins	NN: 78% (aflatoxin), 77% (fumonisins)
Ghilardelli et al., 2022 [75]	NIR	RF	Corn silage	>90%
Teixidó et al., 2023 [76]	NIR	RF	Oat for DON	77.8%
Ma et al., 2023 [78]	Biosensor	sPLS-DA, svmLinear, svmRadial, RF, NN, HDDA	Peanuts and corn	RF: 95–98%
Tarazona et al., 2021 [79]	Spatiotemporal	NN, RF, XGB, MLR	Corn	RF: Best performance
Mateo et al., 2023 [80]	Spatiotemporal	NN, RF, XGB, MLR	Cereals for lactic acid bacteria	RF, XGB: Similar performance
Chávez et al., 2022 [92]	NIR	GBM, RF, LASSO, GLMNET, SVM	Single kernel corn	GBM: 83%
Liu et al., 2018 [103]	Spatiotemporal	BN, LR, mechanistic	Wheat for DON	LR: 88%, BN: 86%
Guo et al., 2020 [104]	TRNs	BN	Fusarium graminearum	High confidence modules: 81.8%
Kim et al., 2024 [33]	Weather	CNN	Wheat grains stored in sealed containers	83.3%
Castano et al., 2023 [34]	Weather	NN	Corn in the US	Aflatoxin: 73%, fumonisin: 85%
Branstad et al., 2023 [39]	Spatiotemporal	GBM	Corn in Iowa	20 ppb: 96.8%, 5 ppb: 90.3%
Wang et al., 2022 [87]	Spatiotemporal	XGB	Peanuts and soybeans in China, Brazil, Argentina	>90%
Xie et al., 2022 [88]	Spatiotemporal	XGB	Peanuts in China	98%
Xie et al., 2022 [88]	Spatiotemporal	XGB with decision rule	Peanuts in China	87.2%
Zhao et al., 2017 [97]	Hyperspectral	SVM	Corn in China	91.67%
Kim et al., 2023 [96]	Hyperspectral	SVM	Ground corn samples from Texas	95.7%
Almoujahed et al., 2022 [45]	Spectral	SVM	Fusarium head blight in wheat in Belgium, 2020–2021	97%
Kos et al., 2016 [99]	Spatiotemporal	CART	Wheat in Italy	88–92%
Purchase et al., 2023 [100]	Spatiotemporal	DT	Breakfast cereals in Italy	High fumonisin risk in high sodium or high-fat cereals

5. Conclusions

Our research focuses on highlighting and evaluating different ML models for monitoring and predicting the presence of mycotoxins in common crops. We conducted an extensive literature review of over 30 studies performed within the years 2013 to 2023. The number of publications in each field has grown significantly over the 10 years reviewed; however, the application of ML in the area of monitoring and predicting mycotoxins is still in its infancy, and despite the promise of ML methods in mycotoxin detection, their adoption in industry has been cautious. This is likely due to the high operational costs associated with advanced techniques like hyperspectral imaging, as opposed to the use of ML methods themselves. The prevalence of such data-intensive methods raises questions about the feasibility of widespread implementation, particularly in resource-constrained settings.

We found that the most common data type was spectral or image data, and as such, the most common ML method used was NNs, as they can be readily applied to image data. RFs were the second most popular ML method and have gained traction due to their robustness and ease of implementation. Additionally, most of the studies reviewed used classification ML techniques to distinguish contaminated from healthy crops. The high predictive accuracy reported in the reviewed studies suggests that these methods represent a promising approach for mycotoxin detection and enhancing food safety in general. However, a point to note is that the reported high accuracy of the ML model's predictions, often exceeding 90%, may not fully account for the homogeneity of training and test sets within individual laboratories. This homogeneity can result in overfitting, where models appear highly accurate in a controlled setting but may not perform as well under the variable conditions of real-world applications.

Although this work focused on the application of the most popular ML methods, numerous other ML and statistical techniques have been applied to mycotoxin detection data. For example, in a study by [106], classification models such as partial least squares-discriminant analysis (PLS-DA) and principal component-linear discriminant analysis (PC-LDA) were employed to distinguish between wheat samples with high and low contamination. Additionally, statistical techniques like PCA are often used as a dimension reduction method. Refs. [107–109] used PCA when dealing with high-dimensional image data.

A critical bottleneck in the development of ML applications for food safety is the lack of detailed hyperparameter descriptions, which further complicates the landscape, as these parameters are crucial for the replication and validation of ML models. Without clear reporting on hyperparameter tuning, the ability to reproduce results and validate findings becomes challenging, hindering the progression towards robust and reliable ML applications in food safety. The majority of the reviewed studies do not provide open access to code, and many have limited access to data, further impeding the reproducibility of the described methods.

Despite these challenges, the future prospects of ML in food safety are promising. As the field matures, there is a need for standardisation in reporting practices and for developing models that can reliably perform across diverse laboratory conditions and datasets. Extensive research could be conducted that directly compares different ML models under a standardised set of hyperparameters, providing clearer insights into the most effective techniques in specific contexts related to mycotoxin detection.

As the field is growing, there are numerous avenues for future work. One such avenue is model interpretability. Given the critical nature of food safety, future research could also focus on improving the interpretability of ML models. Techniques like SHAP (SHapley Additive exPlanations) [110] and LIME (Local Interpretable Model-Agnostic Explanations) [111] can be used to make the models' decisions more transparent and trustworthy. Furthermore, addressing the current bottlenecks, such as the high operational costs and the need for data standardisation, will be crucial. Future research should explore cost-effective techniques and advocate for open-access datasets and standardised reporting practices to enhance reproducibility and application in diverse settings.

Author Contributions: Conceptualization, A.I. and A.C.P.; investigation, A.I.; data curation, A.I.; writing—original draft preparation, A.I.; writing—review and editing, A.C.P., F.M.D. and N.S.; visualization, A.I.; supervision, A.C.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was conducted as part of the Mycotox-I project, which is kindly supported by the Department of Agriculture, Food, and the Marine (DAFM) and the Department of Agriculture, Environment, and Rural Affairs (DAERA), grant number 2021R460. Andrew Parnell’s work was supported by the SFI Centre for Research Training in Foundations of Data Science 18/CRT/6049 and the SFI Research Centre award 12/RC/2289_P2. For the purpose of open access, the author has applied a CC BY public copyright licence to any author-accepted manuscript version arising from this submission.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analysed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NN	Neural Network
CNN	Convolutional Neural Network
RF	Random Forest
GBM	Gradient Boosted Machine
XGB	eXtreme Gradient Boosted Machine
DT	Decision Trees
CART	Classification and Regression Trees
SVM	Support Vector Machine
BM	Bayesian Models
BN	Bayesian Network
LDA	Linear Discriminant Analysis
PDA	Penalised Discriminant Analysis
LReg	Linear Regression
LR	Logistic Regression
MLR	Multiple Linear Regression
LASSO	Least Absolute Shrinkage and Selection Operator
GLMNET	Elastic-Net Regularized Generalised Linear Models
PLS-DA	Partial Least Squares-Discriminant Analysis
sPLS-DA	Sparse Partial Least Squares-Discriminant Analysis
PCA	Principal Component Analysis
MLP	Multi-Layer Perceptron
BLUP	Best Linear Unbiased Prediction
PCC	Pearson Correlation Coefficient
RMSE	Root Mean Square Error
R^2	Coefficient of Determination
AUC	Area Under the Curve
NIR	Near-Infrared Spectroscopy
DON	Deoxynivalenol

Appendix A

The quality of review has been assessed according to PRISMA guidelines [112]. This review has not been registered in a public registry. Figure A1 shows a flow chart demonstrating the selection process used in this work.

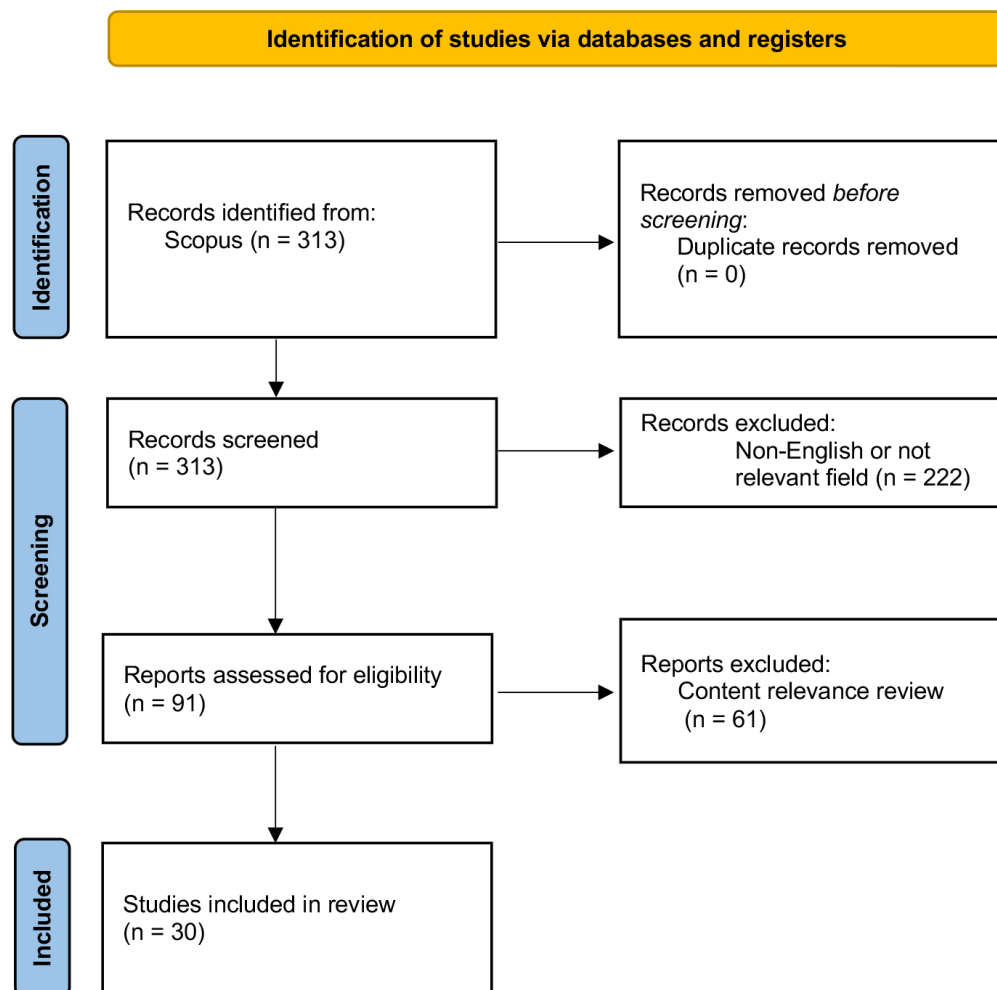


Figure A1. PRISMA flowchart of literature search strategy.

References

1. The World Health Organization (WHO). Food Safety. The World Health Organization. 2023. Available online: <https://www.who.int/news-room/fact-sheets/detail/mycotoxins> (accessed on 5 November 2023).
2. Mavrommatis, A.; Giamouri, E.; Tavrizelou, S.; Zacharioudaki, M.; Danezis, G.; Simitzis, P.E.; Zoidis, E.; Tsiplakou, E.; Pappas, A.C.; Georgiou, C.A.; et al. Impact of mycotoxins on animals' oxidative status. *Antioxidants* **2021**, *10*, 214. [[CrossRef](#)] [[PubMed](#)]
3. Marroquín-Cardona, A.; Johnson, N.; Phillips, T.; Hayes, A. Mycotoxins in a changing global environment—A review. *Food Chem. Toxicol.* **2014**, *69*, 220–230. [[CrossRef](#)]
4. Liu, Y.; Wu, F. Global burden of aflatoxin-induced hepatocellular carcinoma: A risk assessment. *Environ. Health Perspect.* **2010**, *118*, 818–824. [[CrossRef](#)]
5. Van der Fels-Klerx, H.; Liu, C.; Focker, M.; Montero-Castro, I.; Rossi, V.; Manstretta, V.; Magan, N.; Krska, R. Decision support system for integrated management of mycotoxins in feed and food supply chains. *World Mycotoxin J.* **2022**, *15*, 119–133. [[CrossRef](#)]
6. Tola, M.; Kebede, B. Occurrence, importance and control of mycotoxins: A review. *Cogent Food Agric.* **2016**, *2*, 1191103. [[CrossRef](#)]
7. Logrieco, A.; Battilani, P.; Leggieri, M.C.; Jiang, Y.; Haesaert, G.; Lanubile, A.; Mahuku, G.; Mesterházy, A.; Ortega-Beltran, A.; Pasti, M.; et al. Perspectives on global mycotoxin issues and management from the MycoKey Maize Working Group. *Plant Dis.* **2021**, *105*, 525–537. [[CrossRef](#)]
8. Leggieri, M.C.; Lanubile, A.; Dall'Asta, C.; Pietri, A.; Battilani, P. The impact of seasonal weather variation on mycotoxins: Maize crop in 2014 in northern Italy as a case study. *World Mycotoxin J.* **2020**, *13*, 25–36. [[CrossRef](#)]
9. Zingales, V.; Taroncher, M.; Martino, P.A.; Ruiz, M.J.; Caloni, F. Climate change and effects on molds and mycotoxins. *Toxins* **2022**, *14*, 445. [[CrossRef](#)] [[PubMed](#)]
10. Medina, A.; Akbar, A.; Baazeem, A.; Rodriguez, A.; Magan, N. Climate change, food security and mycotoxins: Do we know enough? *Fungal Biol. Rev.* **2017**, *31*, 143–154. [[CrossRef](#)]
11. Eskola, M.; Kos, G.; Elliott, C.T.; Hajšlová, J.; Mayar, S.; Krska, R. Worldwide contamination of food-crops with mycotoxins: Validity of the widely cited 'FAO estimate' of 25%. *Crit. Rev. Food Sci. Nutr.* **2020**, *60*, 2773–2789. [[CrossRef](#)]

12. Alshannaq, A.; Yu, J.H. Occurrence, toxicity, and analysis of major mycotoxins in food. *Int. J. Environ. Res. Public Health* **2017**, *14*, 632. [[CrossRef](#)] [[PubMed](#)]
13. Wu, F. Global impacts of aflatoxin in maize: Trade and human health. *World Mycotoxin J.* **2015**, *8*, 137–142. [[CrossRef](#)]
14. Johns, L.E.; Bebbler, D.P.; Gurr, S.J.; Brown, N.A. Emerging health threat and cost of Fusarium mycotoxins in European wheat. *Nat. Food* **2022**, *3*, 1014–1019. [[CrossRef](#)]
15. Latham, R.L.; Boyle, J.T.; Barbano, A.; Loveman, W.G.; Brown, N.A. Diverse mycotoxin threats to safe food and feed cereals. *Essays Biochem.* **2023**, *67*, 797–809. [[PubMed](#)]
16. Whitaker, T.B. Standardisation of mycotoxin sampling procedures: An urgent necessity. *Food Control* **2003**, *14*, 233–237. [[CrossRef](#)]
17. Anfossi, L.; Giovannoli, C.; Baggiani, C. Mycotoxin detection. *Curr. Opin. Biotechnol.* **2016**, *37*, 120–126. [[CrossRef](#)] [[PubMed](#)]
18. Maragos, C.M. Emerging technologies for mycotoxin detection. *J. Toxicol. Toxin Rev.* **2004**, *23*, 317–344. [[CrossRef](#)]
19. Soares, R.R.; Ricelli, A.; Fanelli, C.; Caputo, D.; de Cesare, G.; Chu, V.; Aires-Barros, M.R.; Conde, J.P. Advances, challenges and opportunities for point-of-need screening of mycotoxins in foods and feeds. *Analyst* **2018**, *143*, 1015–1035. [[CrossRef](#)] [[PubMed](#)]
20. Renaud, J.B.; Miller, J.D.; Sumarah, M.W. Mycotoxin testing paradigm: Challenges and opportunities for the future. *J. AOAC Int.* **2019**, *102*, 1681–1688. [[CrossRef](#)] [[PubMed](#)]
21. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine learning in agriculture: A review. *Sensors* **2018**, *18*, 2674. [[CrossRef](#)]
22. Baştanlar, Y.; Özuysal, M. Introduction to machine learning. In *miRNomics: MicroRNA Biology and Computational Analysis*; Humana Press: Totowa, NJ, USA, 2014; pp. 105–128.
23. Alpaydin, E. *Introduction to Machine Learning*; MIT Press: Cambridge, MA, USA, 2020.
24. Torelli, E.; Firrao, G.; Bianchi, G.; Saccardo, F.; Locci, R. The influence of local factors on the prediction of fumonisin contamination in maize. *J. Sci. Food Agric.* **2012**, *92*, 1808–1814. [[CrossRef](#)] [[PubMed](#)]
25. Mateo, F.; Gadea, R.; Mateo, E.M.; Jiménez, M. Multilayer perceptron neural networks and radial-basis function networks as tools to forecast accumulation of deoxynivalenol in barley seeds contaminated with *Fusarium culmorum*. *Food Control* **2011**, *22*, 88–95. [[CrossRef](#)]
26. Panagou, E.Z.; Kodogiannis, V.S. Application of neural networks as a non-linear modelling technique in food mycology. *Expert Syst. Appl.* **2009**, *36*, 121–131. [[CrossRef](#)]
27. Zhou, L.; Zhang, C.; Liu, F.; Qiu, Z.; He, Y. Application of deep learning in food: A review. *Compr. Rev. Food Sci. Food Saf.* **2019**, *18*, 1793–1811. [[CrossRef](#)] [[PubMed](#)]
28. Wang, X.; Bouzembrak, Y.; Lansink, A.O.; van der Fels-Klerx, H. Application of machine learning to the monitoring and prediction of food safety: A review. *Compr. Rev. Food Sci. Food Saf.* **2022**, *21*, 416–434. [[CrossRef](#)] [[PubMed](#)]
29. Bernardes, R.C.; De Medeiros, A.; da Silva, L.; Cantoni, L.; Martins, G.F.; Mastrangelo, T.; Novikov, A.; Mastrangelo, C.B. Deep-learning approach for fusarium head blight detection in wheat seeds using low-cost imaging technology. *Agriculture* **2022**, *12*, 1801. [[CrossRef](#)]
30. Magan, N.; Medina, A. Integrating gene expression, ecology and mycotoxin production by *Fusarium* and *Aspergillus* species in relation to interacting environmental factors. *World Mycotoxin J.* **2016**, *9*, 673–684. [[CrossRef](#)]
31. Verheecke-Vaessen, C.; Diez-Gutierrez, L.; Renaud, J.; Sumarah, M.; Medina, A.; Magan, N. Interacting climate change environmental factors effects on *Fusarium langsethiae* growth, expression of Tri genes and T-2/HT-2 mycotoxin production on oat-based media and in stored oats. *Fungal Biol.* **2019**, *123*, 618–624. [[CrossRef](#)]
32. Natarajan, S.; Balachandar, D.; Paranidharan, V. Inhibitory effects of epiphytic *Kluyveromyces marxianus* from Indian senna (*Cassia angustifolia* Vahl.) on growth and aflatoxin production of *Aspergillus flavus*. *Int. J. Food Microbiol.* **2023**, *406*, 110368. [[CrossRef](#)]
33. Kim, Y.; Kang, S.; Ajani, O.S.; Mallipeddi, R.; Ha, Y. Predicting early mycotoxin contamination in stored wheat using machine learning. *J. Stored Prod. Res.* **2024**, *106*, 102294. [[CrossRef](#)]
34. Castano-Duque, L.; Winzeler, E.; Blackstock, J.M.; Liu, C.; Vergopolan, N.; Focker, M.; Barnett, K.; Owens, P.R.; van der Fels-Klerx, H.; Vaughan, M.M.; et al. Dynamic geospatial modeling of mycotoxin contamination of corn in Illinois: unveiling critical factors and predictive insights with machine learning. *Front. Microbiol.* **2023**, *14*, 1283127. [[CrossRef](#)] [[PubMed](#)]
35. Orlando, B.; Barrier-Guillot, B.; Gourdain, E.; Maumene, C. Identification of agronomic factors that influence the levels of T-2 and HT-2 toxins in barley grown in France. *World Mycotoxin J.* **2010**, *3*, 169–174. [[CrossRef](#)]
36. Edwards, S.G. Influence of agricultural practices on *Fusarium* infection of cereals and subsequent contamination of grain by trichothecene mycotoxins. *Toxicol. Lett.* **2004**, *153*, 29–35. [[CrossRef](#)] [[PubMed](#)]
37. Edwards, S.G.; Jennings, P. Impact of agronomic factors on *Fusarium* mycotoxins in harvested wheat. *Food Addit. Contam. Part A* **2018**, *35*, 2443–2454. [[CrossRef](#)] [[PubMed](#)]
38. Camardo Leggieri, M.; Mazzoni, M.; Battilani, P. Machine learning for predicting mycotoxin occurrence in maize. *Front. Microbiol.* **2021**, *12*, 661132. [[CrossRef](#)] [[PubMed](#)]
39. Branstad-Spates, E.H.; Castano-Duque, L.; Mosher, G.A.; Hurburgh, C.R., Jr.; Owens, P.; Winzeler, E.; Rajasekaran, K.; Bowers, E.L. Gradient boosting machine learning model to predict aflatoxins in Iowa corn. *Front. Microbiol.* **2023**, *14*, 1248772. [[CrossRef](#)] [[PubMed](#)]
40. Wegulo, S. Factors influencing deoxynivalenol accumulation in small grain cereals. *Toxins* **2012**, *4*, 1157–1180. [[CrossRef](#)] [[PubMed](#)]

41. Dhakal, K.; Sivaramakrishnan, U.; Zhang, X.; Belay, K.; Oakes, J.; Wei, X.; Li, S. Machine learning analysis of hyperspectral images of damaged wheat kernels. *Sensors* **2023**, *23*, 3523. [[CrossRef](#)] [[PubMed](#)]
42. Wang, X.; Liu, C.; van der Fels-Klerx, H. Regional prediction of multi-mycotoxin contamination of wheat in Europe using machine learning. *Food Res. Int.* **2022**, *159*, 111588. [[CrossRef](#)]
43. Rangarajan, A.K.; Whetton, R.L.; Mouazen, A.M. Detection of fusarium head blight in wheat using hyperspectral data and deep learning. *Expert Syst. Appl.* **2022**, *208*, 118240. [[CrossRef](#)]
44. Kalkan, H.; Güneş, A.; Durmuş, E.; Kuşçu, A. Non-invasive detection of aflatoxin-contaminated figs using fluorescence and multispectral imaging. *Food Addit. Contam. Part A* **2014**, *31*, 1414–1421. [[CrossRef](#)] [[PubMed](#)]
45. Almoujahed, M.B.; Rangarajan, A.K.; Whetton, R.L.; Vincke, D.; Eylenbosch, D.; Vermeulen, P.; Mouazen, A.M. Detection of fusarium head blight in wheat under field conditions using a hyperspectral camera and machine learning. *Comput. Electron. Agric.* **2022**, *203*, 107456. [[CrossRef](#)]
46. Liu, L.; Dong, Y.; Huang, W.; Du, X.; Ma, H. Monitoring wheat fusarium head blight using unmanned aerial vehicle hyperspectral imagery. *Remote Sens.* **2020**, *12*, 3811. [[CrossRef](#)]
47. Hruska, Z.; Yao, H.; Kincaid, R.; Brown, R.; Cleveland, T.; Bhatnagar, D. Fluorescence excitation–emission features of aflatoxin and related secondary metabolites and their application for rapid detection of mycotoxins. *Food Bioprocess Technol.* **2014**, *7*, 1195–1201. [[CrossRef](#)]
48. Zhu, F.; Yao, H.; Hruska, Z.; Kincaid, R.; Brown, R.L.; Bhatnagar, D.; Cleveland, T.E. Integration of fluorescence and reflectance visible near-infrared (VNIR) hyperspectral images for detection of aflatoxins in corn kernels. *Trans. ASABE* **2016**, *59*, 785–794.
49. Han, Z.; Gao, J. Pixel-level aflatoxin detecting based on deep learning and hyperspectral imaging. *Comput. Electron. Agric.* **2019**, *164*, 104888. [[CrossRef](#)]
50. Del Fiore, A.; Reverberi, M.; Ricelli, A.; Pinzari, F.; Serranti, S.; Fabbri, A.A.; Bonifazi, G.; Fanelli, C. Early detection of toxigenic fungi on maize by hyperspectral imaging analysis. *Int. J. Food Microbiol.* **2010**, *144*, 64–71. [[CrossRef](#)] [[PubMed](#)]
51. Serranti, S.; Cesare, D.; Bonifazi, G. The development of a hyperspectral imaging method for the detection of Fusarium-damaged, yellow berry and vitreous Italian durum wheat kernels. *Biosyst. Eng.* **2013**, *115*, 20–30. [[CrossRef](#)]
52. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [[CrossRef](#)]
53. Gurney, K. *An Introduction to Neural Networks*; CRC Press: Boca Raton, FL, USA, 2018.
54. Montavon, G.; Samek, W.; Müller, K.R. Methods for interpreting and understanding deep neural networks. *Digit. Signal Process.* **2018**, *73*, 1–15. [[CrossRef](#)]
55. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
56. Canziani, A.; Paszke, A.; Culurciello, E. An analysis of deep neural network models for practical applications. *arXiv* **2016**, arXiv:1605.07678.
57. Niedbała, G.; Kurasiak-Popowska, D.; Stuper-Szablewska, K.; Nawracała, J. Application of artificial neural networks to analyze the concentration of ferulic acid, deoxynivalenol, and nivalenol in winter wheat grain. *Agriculture* **2020**, *10*, 127. [[CrossRef](#)]
58. StatSoft, Inc. STATISTICA (Data Analysis Software System), Version 7.1. 2020. Available online: <http://www.statsoft.com> (accessed on 28 April 2024).
59. Jubair, S.; Tucker, J.R.; Henderson, N.; Hiebert, C.W.; Badea, A.; Domaratzki, M.; Fernando, W. GPTransformer: A transformer-based deep learning method for predicting Fusarium related traits in barley. *Front. Plant Sci.* **2021**, *12*, 761402. [[CrossRef](#)] [[PubMed](#)]
60. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems*; Curran Associates Inc.: Red Hook, NY, USA, 2017; Volume 30.
61. Grahn, H.; Geladi, P. *Techniques and Applications of Hyperspectral Image Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2007.
62. Jin, X.; Jie, L.; Wang, S.; Qi, H.J.; Li, S.W. Classifying wheat hyperspectral pixels of healthy heads and Fusarium head blight disease using a deep neural network in the wild field. *Remote Sens.* **2018**, *10*, 395. [[CrossRef](#)]
63. Qiu, R.; Yang, C.; Moghimi, A.; Zhang, M.; Steffenson, B.J.; Hirsch, C.D. Detection of fusarium head blight in wheat using a deep neural network and color imaging. *Remote Sens.* **2019**, *11*, 2658. [[CrossRef](#)]
64. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271.
65. Gao, J.; Zhao, L.; Li, J.; Deng, L.; Ni, J.; Han, Z. Aflatoxin rapid detection based on hyperspectral with 1D-convolution neural network in the pixel level. *Food Chem.* **2021**, *360*, 129968. [[CrossRef](#)]
66. Oener, T.; Thiam, P.; Kos, G.; Krska, R.; Schwenker, F.; Mizaikoff, B. Machine learning algorithms for the automated classification of contaminated maize at regulatory limits via infrared attenuated total reflection spectroscopy. *World Mycotoxin J.* **2019**, *12*, 113–122. [[CrossRef](#)]
67. Ottoboni, M.; Pinotti, L.; Tretola, M.; Giromini, C.; Fusi, E.; Rebutti, R.; Grillo, M.; Tassoni, L.; Foresta, S.; Gastaldello, S.; et al. Combining E-nose and lateral flow immunoassays (LFIA) for rapid occurrence/co-occurrence aflatoxin and fumonisin detection in maize. *Toxins* **2018**, *10*, 416. [[CrossRef](#)]
68. Campagnoli, A.; Cheli, F.; Savoini, G.; Crotti, A.; Pastori, A.; Dell’Orto, V. Application of an electronic nose to detection of aflatoxins in corn. *Vet. Res. Commun.* **2009**, *33*, 273–275. [[CrossRef](#)]

69. Gobbi, E.; Falasconi, M.; Torelli, E.; Sberveglieri, G. Electronic nose predicts high and low fumonisin contamination in maize cultures. *Food Res. Int.* **2011**, *44*, 992–999. [CrossRef]
70. Lippolis, V.; Pascale, M.; Cervellieri, S.; Damascelli, A.; Visconti, A. Screening of deoxynivalenol contamination in durum wheat by MOS-based electronic nose and identification of the relevant pattern of volatile compounds. *Food Control* **2014**, *37*, 263–271. [CrossRef]
71. Leggieri, M.C.; Mazzoni, M.; Fodil, S.; Moschini, M.; Bertuzzi, T.; Prandini, A.; Battilani, P. An electronic nose supported by an artificial neural network for the rapid detection of aflatoxin B1 and fumonisins in maize. *Food Control* **2021**, *123*, 107722. [CrossRef]
72. Saha, D.; Manickavasagan, A. Machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review. *Curr. Res. Food Sci.* **2021**, *4*, 28–44. [CrossRef] [PubMed]
73. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
74. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. *Classification and Regression Trees*; Wadsworth: Hartford, CT, USA, 1984.
75. Ghilardelli, F.; Barbato, M.; Gallo, A. A preliminary study to classify corn silage for high or low mycotoxin contamination by using near infrared spectroscopy. *Toxins* **2022**, *14*, 323. [CrossRef] [PubMed]
76. Teixido-Orries, I.; Molino, F.; Femenias, A.; Ramos, A.J.; Marín, S. Quantification and classification of deoxynivalenol-contaminated oat samples by near-infrared hyperspectral imaging. *Food Chem.* **2023**, *417*, 135924. [CrossRef]
77. Femenias, A.; Gatiús, F.; Ramos, A.J.; Sanchis, V.; Marín, S. Near-infrared hyperspectral imaging for deoxynivalenol and ergosterol estimation in wheat samples. *Food Chem.* **2021**, *341*, 128206. [CrossRef]
78. Ma, J.; Guan, Y.; Xing, F.; Eltzov, E.; Wang, Y.; Li, X.; Tai, B. Accurate and non-destructive monitoring of mold contamination in foodstuffs based on whole-cell biosensor array coupling with machine-learning prediction models. *J. Hazard. Mater.* **2023**, *449*, 131030. [CrossRef]
79. Tarazona, A.; Mateo, E.M.; Gómez, J.V.; Gavara, R.; Jiménez, M.; Mateo, F. Machine learning approach for predicting *Fusarium culmorum* and *F. proliferatum* growth and mycotoxin production in treatments with ethylene-vinyl alcohol copolymer films containing pure components of essential oils. *Int. J. Food Microbiol.* **2021**, *338*, 109012. [CrossRef]
80. Mateo, E.M.; Tarazona, A.; Aznar, R.; Mateo, F. Exploring the impact of lactic acid bacteria on the biocontrol of toxigenic *Fusarium* spp. and their main mycotoxins. *Int. J. Food Microbiol.* **2023**, *387*, 110054. [CrossRef] [PubMed]
81. Mateo, E.M.; Gómez, J.V.; Tarazona, A.; García-Esparza, M.Á.; Mateo, F. Comparative analysis of machine learning methods to predict growth of *F. sporotrichioides* and production of T-2 and HT-2 toxins in treatments with ethylene-vinyl alcohol films containing pure components of essential oils. *Toxins* **2021**, *13*, 545. [CrossRef] [PubMed]
82. Tarazona, A.; Mateo, E.M.; Gómez, J.V.; Romera, D.; Mateo, F. Potential use of machine learning methods in assessment of *Fusarium culmorum* and *Fusarium proliferatum* growth and mycotoxin production in treatments with antifungal agents. *Fungal Biol.* **2021**, *125*, 123–133. [CrossRef] [PubMed]
83. Srinivasan, R.; Lalitha, T.; Brintha, N.; Sterlin Minish, T.; Al Obaid, S.; Alharbi, S.A.; Sundaram, S.; Mahilraj, J. Predicting the Growth of *F. proliferatum* and *F. culmorum* and the Growth of Mycotoxin Using Machine Learning Approach. *BioMed Res. Int.* **2022**, *2022*, 9592365. [CrossRef] [PubMed]
84. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
85. Natekin, A.; Knoll, A. Gradient boosting machines, a tutorial. *Front. Neurobot.* **2013**, *7*, 21. [CrossRef] [PubMed]
86. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
87. Wang, X.; Bouzembrak, Y.; Oude Lansink, A.; Van der Fels-Klerx, H. Designing a monitoring program for aflatoxin B1 in feed products using machine learning. *NPJ Sci. Food* **2022**, *6*, 40. [CrossRef] [PubMed]
88. Xie, H.; Wang, X.; van der Hoof, J.J.; Medema, M.H.; Chen, Z.Y.; Yue, X.; Zhang, Q.; Li, P. Fungi population metabolomics and molecular network study reveal novel biomarkers for early detection of aflatoxigenic *Aspergillus* species. *J. Hazard. Mater.* **2022**, *424*, 127173. [CrossRef] [PubMed]
89. FDA. Guidance for Industry: Action Levels for Poisonous or Deleterious Substances in Human Food and Animal Feed. 2020. Available online: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-industry-action-levels-poisonous-or-deleterious-substances-human-food-and-animal-feed#afla> (accessed on 5 November 2023).
90. EFSA. *Aflatoxins (Sum of B1, B2, G1, G2) in Cereals and Cereal-Derived Food Products*; EFSA: Parma, Italy, 2013.
91. Zheng, A.; Casari, A. *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*; O'Reilly Media, Inc.: Newton, MA, USA, 2018.
92. Chavez, R.A.; Cheng, X.; Herrman, T.J.; Stasiewicz, M.J. Single kernel aflatoxin and fumonisin contamination distribution and spectral classification in commercial corn. *Food Control* **2022**, *131*, 108393. [CrossRef]
93. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1999.
94. Burges, C.J. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [CrossRef]
95. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000.
96. Kim, Y.K.; Baek, I.; Lee, K.M.; Kim, G.; Kim, S.; Kim, S.Y.; Chan, D.; Herrman, T.J.; Kim, N.; Kim, M.S. Rapid Detection of Single-and Co-Contaminant Aflatoxins and Fumonisins in Ground Maize Using Hyperspectral Imaging Techniques. *Toxins* **2023**, *15*, 472. [CrossRef] [PubMed]

97. Zhao, X.; Wang, W.; Chu, X.; Li, C.; Kimuli, D. Early detection of *Aspergillus parasiticus* infection in maize kernels using near-infrared hyperspectral imaging and multivariate data analysis. *Appl. Sci.* **2017**, *7*, 90. [[CrossRef](#)]
98. Quinlan, J.R. Induction of decision trees. *Mach. Learn.* **1986**, *1*, 81–106. [[CrossRef](#)]
99. Kos, G.; Sieger, M.; McMullin, D.; Zahradnik, C.; Sulyok, M.; Öner, T.; Mizaikoff, B.; Krska, R. A novel chemometric classification for FTIR spectra of mycotoxin-contaminated maize and peanuts at regulatory limits. *Food Addit. Contam. Part A* **2016**, *33*, 1596–1607. [[CrossRef](#)] [[PubMed](#)]
100. Purchase, J.; Donato, R.; Sacco, C.; Pettini, L.; Rookmin, A.D.; Melani, S.; Artese, A.; Purchase, D.; Marvasi, M. The association of food ingredients in breakfast cereal products and fumonisins production: Risks identification and predictions. *Mycotoxin Res.* **2023**, *39*, 165–175. [[CrossRef](#)] [[PubMed](#)]
101. Jensen, F.V.; Nielsen, T.D. *Bayesian Networks and Decision Graphs*; Springer: Berlin/Heidelberg, Germany, 2007; Volume 2.
102. Buritica, J.A.; Tesfamariam, S. Consequence-based framework for electric power providers using Bayesian belief network. *Int. J. Electr. Power Energy Syst.* **2015**, *64*, 233–241. [[CrossRef](#)]
103. Liu, C.; Manstretta, V.; Rossi, V.; Van der Fels-Klerx, H. Comparison of three modelling approaches for predicting deoxynivalenol contamination in winter wheat. *Toxins* **2018**, *10*, 267. [[CrossRef](#)] [[PubMed](#)]
104. Guo, L.; Ji, M.; Ye, K. Dynamic network inference and association computation discover gene modules regulating virulence, mycotoxin and sexual reproduction in *Fusarium graminearum*. *BMC Genom.* **2020**, *21*, 179. [[CrossRef](#)] [[PubMed](#)]
105. Babu, M.M.; Luscombe, N.M.; Aravind, L.; Gerstein, M.; Teichmann, S.A. Structure and evolution of transcriptional regulatory networks. *Curr. Opin. Struct. Biol.* **2004**, *14*, 283–291. [[CrossRef](#)]
106. De Girolamo, A.; von Holst, C.; Cortese, M.; Cervellieri, S.; Pascale, M.; Longobardi, F.; Catucci, L.; Porricelli, A.C.R.; Lippolis, V. Rapid screening of ochratoxin A in wheat by infrared spectroscopy. *Food Chem.* **2019**, *282*, 95–100. [[CrossRef](#)]
107. Shen, F.; Zhao, T.; Jiang, X.; Liu, X.; Fang, Y.; Liu, Q.; Hu, Q.; Liu, X. On-line detection of toxigenic fungal infection in wheat by visible/near infrared spectroscopy. *LWT* **2019**, *109*, 216–224. [[CrossRef](#)]
108. Jha, S.N.; Jaiswal, P.; Kaur, J.; Ramya, H. Rapid detection and quantification of aflatoxin B1 in milk using fourier transform infrared spectroscopy. *J. Inst. Eng. Ser. A* **2021**, *102*, 259–265. [[CrossRef](#)]
109. Miličević, D.; Petronijević, R.; Petrović, Z.; Djinović-Stojanović, J.; Jovanović, J.; Baltić, T.; Janković, S. Impact of climate change on aflatoxin M1 contamination of raw milk with special focus on climate conditions in Serbia. *J. Sci. Food Agric.* **2019**, *99*, 5202–5210. [[CrossRef](#)]
110. Shapley, L.S. *A Value for n-Person Games*; Princeton University Press: Princeton, NJ, USA, 1953.
111. Ribeiro, M.T.; Singh, S.; Guestrin, C. “Why should I trust you?” Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144.
112. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews *Br. Med. J. Publ. Group* **2021**, *372*, n71. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.