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

Alan Inglis

Hamilton Institute, Maynooth University

Andrew Parnell

Hamilton Institute, Maynooth University

Natarajan Subramani

School of Biology and Environmental Science, University College Dublin

Fiona Doohan

School of Biology and Environmental Science, University College Dublin

April 25, 2024

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 as well as 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

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 (The World Health Organization (WHO), 2023). 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 (Mavrommatis et al., 2021; Marroquín-Cardona et al., 2014; Liu and Wu, 2010). 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 (Marroquín-Cardona et al., 2014; Van der Fels-Klerx et al., 2022). 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* (Tola and Kebede, 2016). Mycotoxin contamination in crop products has been found to vary significantly across different geographical locations and is influenced by annual weather conditions (Logrieco et al., 2021; Leggieri et al., 2020). 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 (Zingales et al., 2022; Medina et al., 2017). 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) (Eskola et al., 2020).

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 (Alshannaq and Yu, 2017; Marroquín-Cardona et al., 2014). The economic impact of mycotoxin contami-

nation is substantial, with a global estimate in the billions of euro for detection, regulation enforcement, and mitigation efforts to manage mycotoxin presence in food and feeds annually (Wu, 2015). 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 €3 billion (Johns et al., 2022). Additionally, Latham et al. (2023) show 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 €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 Whitaker (2003), the standard methodology for mycotoxin detection comprises three main steps: sampling, sample preparation, and analytical determination. Chromatographic techniques, such as high-performance liquid chromatography (HPLC) and gas chromatography mass spectrometry (GC-MS), along with immunoassay-based methods like enzyme-linked immunosorbent assays (ELISA), are widely recognised as the most prevalent analytical approaches for the detection of mycotoxins (Anfossi et al., 2016; Maragos, 2004). 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 (Soares et al., 2018; Renaud et al., 2019).

While traditional detection methods such as 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 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 to traditional lab analysis, in terms of efficiency, cost, and scalability, as well as maintaining or improving the accuracy of mycotoxin detection (Liakos et al., 2018).

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 (Baştanlar and Özuysal, 2014). 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 (Alpaydin, 2020).

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, Torelli et al. (2012) use neural networks (NN) for the prediction of contamination from the mycotoxin fumonisin in corn. Additionally, NN have been used to forecast accumulation of the trichothecene mycotoxin deoxynivalenol (DON) in barley seeds (Mateo et al., 2011) and to predict fungal growth (Panagou and Kodogiannis, 2009). For a comprehensive review of the use of NN in food science see Zhou et al. (2019), and for a review of ML methods in general in the field of food safety, see Wang et al. (2022a), and in agriculture, see Liakos et al. (2018).

ML techniques can alleviate some of the current burdens of mycotoxin detection by providing an efficient and low-cost solution (Bernardes et al., 2022). 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 detail 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, as well as 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[1], 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.

2.1 Search Criteria and Overview

A comprehensive search was conducted on the Scopus database and focused on publications between the years 2013 to 2023. The search was conducted using the primary keyword *"mycotoxin"* in combination with 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 to the monitoring and prediction of food safety by Wang et al. (2022a). This strategy was em-

---

ployed 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.
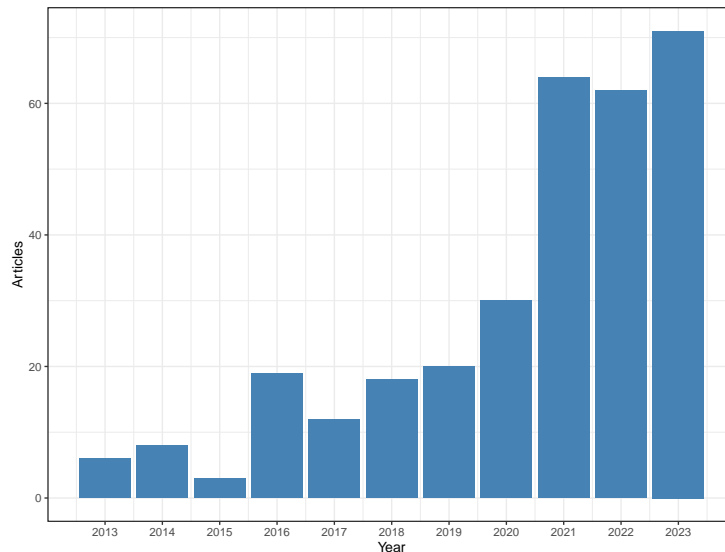

Figure 1: Number of publications between 2013 and 2023 found by our systematic search criteria in Scopus.

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 to search size to 91. After examining the abstracts of all 91 articles, 45 were selected for their relevance and included in this study. From these articles, the predominant ML technique used was neural networks (NN), followed by random forests (RF) and gradient boosting (GB), then support vector machines (SVM), decision trees (DT), and Bayesian networks (BN). Figure 2 shows the frequency of each ML algorithm used in the literature.
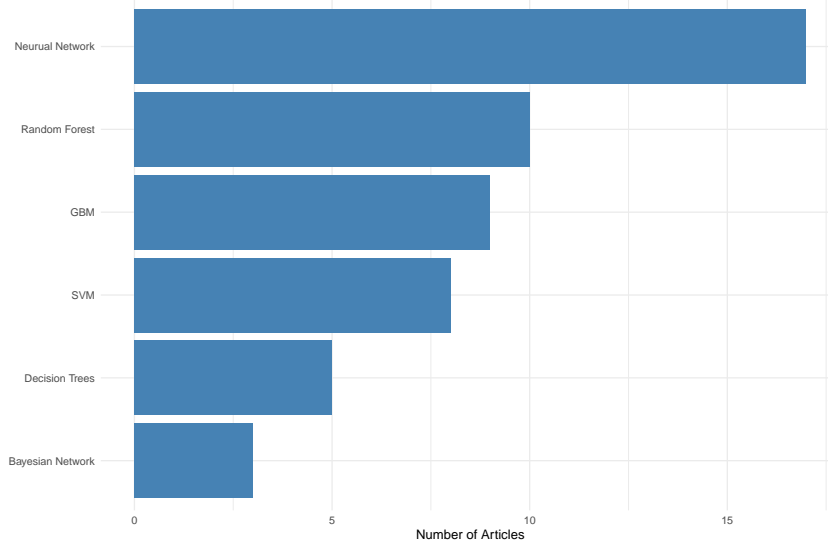
Figure 2: Most popular machine learning methods reviewed in this work.

3 A Brief Introduction to Machine Learning

In this section we provide a general overview of the ML process. This fore-knowledge 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.
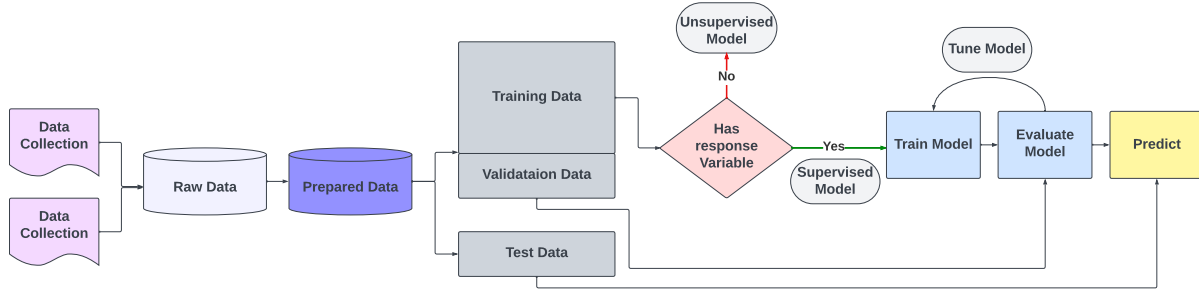
Figure 3: Typical machine learning process.

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

1. **Data Collection:** The process starts with the collection of raw data, which can be from many sources or sites.

2. **Data Preparation:** This raw data is 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, 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 is 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. 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 doesn't just memorise the training data but learns to generalise from it, making accurate predictions on new, unseen data. Test data is used after the model has been trained and validated (see Figure 3). It is another distinct subset of the data set, not used in either training or validation. The test data is 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 the model performance is quantified by evaluating the model performance on the test data set, unless otherwise stated. Sometimes authors also report the training or validation data set 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:

- R^2: This statistic measures the proportion of 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 regresion problems) and provides a measure of the differences between 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 the instances

10

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 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 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   Neural Networks

Neural networks (NNs), first introduced by McCulloch and Pitts (1943), are a class of machine learning algorithms modelled loosely after the human brain (Gurney, 2018). 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 is 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 (Gurney, 2018). 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 that determine the output value of a node given an input or 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.
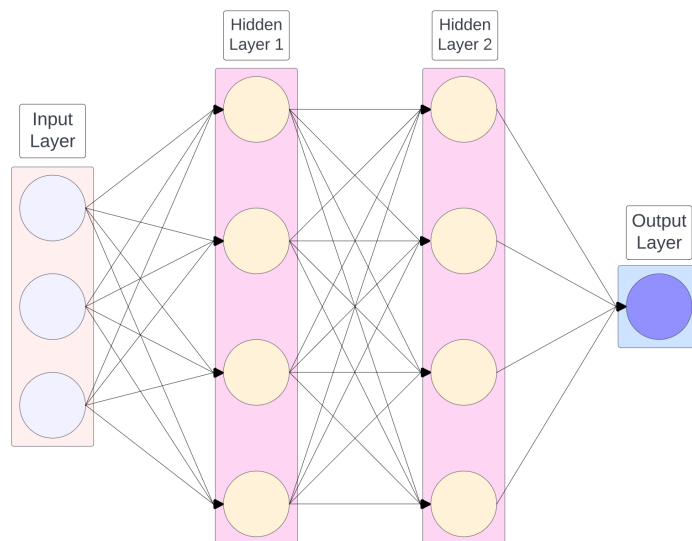
Figure 4: Basic neural network structure, showing an input layer, two hidden layers, and an output layer, where each circle represents a neuron and are interconnected by lines symbolising neural connections. The input layer receives the initial data, which is then processed through successive hidden 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 (Montavon et al., 2018). Secondly, 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 (Srivastava et al., 2014).

Finally, training large NNs requires a significant amount of computing power. The computational cost of NNs will increase with the complexity of the model (Canziani et al., 2016). In the following subsections, we review at the use of NN on different types of mycotoxin data.

4.1.1   NNs applied to Spatiotemporal Data

NNs have been widely applied to spatio-temporal data, despite them not forming part of the traditional suite of spatio-temporal analytics techniques. In the field of mycotoxin study, NNs have been used for a variety of tasks and data types. For example, Camardo Leggieri et al. (2021) use 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.

Niedbała et al. (2020) applied NNs to analyse the concentration of mycotoxins in winter wheat grain. They examined 23 winter wheat genotypes with different Fusarium resistance, 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 (Stat-Soft, Inc., 2020), 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 of training, validation, and test errors, the authors did not specify the data set on which the correlation and MAE metrics were based.

In a novel application of NNs, Jubair et al. (2021) use 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 (Vaswani et al., 2017). In their work, the authors propose a transformer-based genomic prediction model for predicting Fusarium head blight disease levels and associated Fusarium 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 use the Pearson correlation coefficient (PCC) as a measure of performance which calculates the linear relation between the true output and predicted output. They show 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.

4.1.2 NNs Applied to Spectral Data

Hyperspectral (or just spectral) data refers to capture and processing of information from across the electromagnetic spectrum (Grahn and Geladi, 2007). Jin et al. (2018), Qiu et al. (2019), and Rangarajan et al. (2022), all apply NN classification algorithms to pixels of hyperspectral image data to examine wheat for Fusarium head blight infection. Each author uses 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 Rangarajan et al. (2022) 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 effectiveness of these models was compared based on their classification accuracy. They found that a particular type of CNN called *DarkNet 19* (Redmon and Farhadi, 2017) performed the best, with an accuracy of close to 100% across all data conversion methods, on both the validation and test data. For Qiu et al. (2019), 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 data set was 92%. In Jin et al. (2018), the authors compared the accuracy 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 data set and an F1 score and accuracy of 0.75 and 74.3%, respectively, on the test data.

Han and Gao (2019) use a combination of hyperspectral data and NNs to

detect aflatoxin in peanuts. They showed the CNNs efficacy in classifying infected peanuts and achieve a test set accuracy of 95%. They later expanded their work and use a one dimensional CNN (1D-CNN) to classify aflatoxin infection in corn and peanuts. This time they achieved an accuracy of 96.4% for peanuts and 92.1% for corn (Gao et al., 2021).

In research conducted by (Oener et al., 2019), 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 concentraions) were obtained from the seed production Linz of Austria (SBL), and from the Cereal Research Centre of Hungary (CRC). The authors assess 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 show 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 1,250 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.

4.1.3 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 (Ottoboni et al., 2018; Campagnoli et al., 2009), fumonisins (Gobbi et al., 2011), and DON (Lippolis et al., 2014) in corn. However, Leggieri et al. (2021) use 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 abil-

ity to discriminate between samples contaminated with concentrations either exceeding or falling below legal thresholds on data spanning five 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 go on to suggest that the e-nose, when supported by a NN, can provide a fast screening tool for classifying samples.

4.1.4 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 (Saha and Manickavasagan, 2021). 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, Jin et al., 2018; Qiu et al., 2019).

4.2 Random Forests

A random forest (RF; Breiman, 2001) 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 4.5.1 provides a comprehensive examination of decision trees, this Section offers a concise introduction to familiarise readers with the basic concepts and terminology 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 Breiman et al., 1984). The decision nodes are the points where the data is 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.
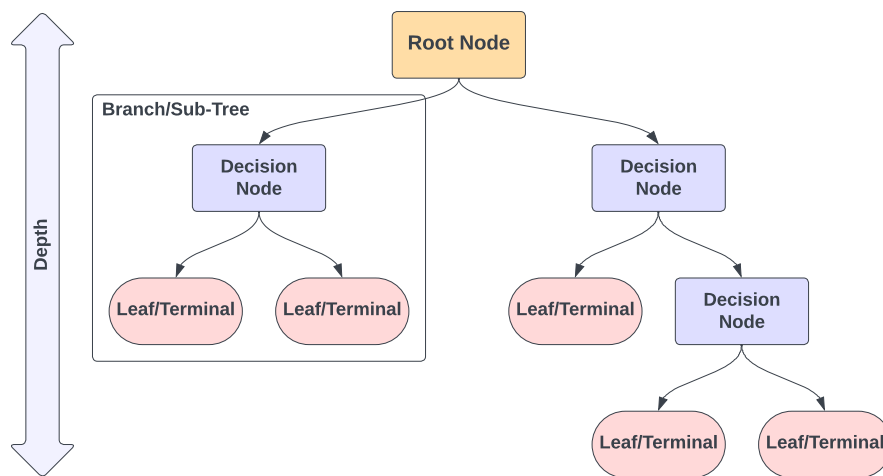

Figure 5: Decision tree process demonstrating the structure of a decision tree, including the root node, branching to decision nodes, and culminating in leaf/terminal nodes. The depth of the tree is indicated, showing the levels of decision-making from the root to the leaves.

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 (i) The number of trees: this is the number of trees in the forest. Generally, more trees increase performance but also increase 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.
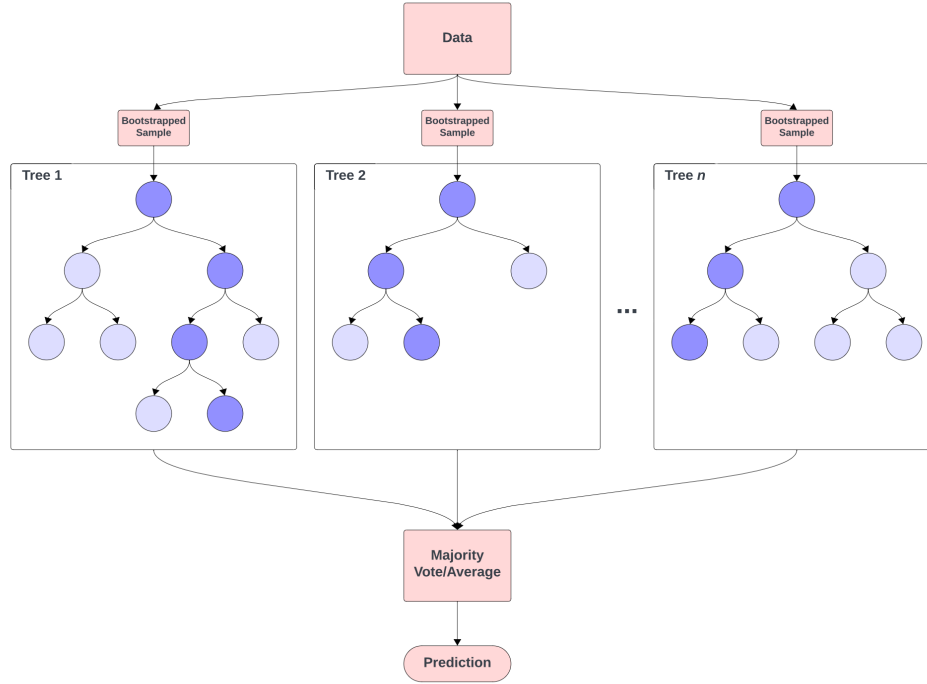
Figure 6: The Random Forest algorithm constructs an ensemble of decision trees, with each tree built from a unique bootstrapped sample of the original dataset. Distinct paths through each tree are shown, highlighted by the darker blue nodes, and represent a sequence of decisions made from the root to 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.

4.2.1   RFs for Spectral Data

As with NNs, RFs have been applied to hyperspectral data. For example, Ghilardelli et al. (2022) use 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, Teixido-Orries et al. (2023) utilised NIR spectroscopy for detecting DON in oat samples from Spain and Sweden collected over the years 2021-2022. The authors apply two different transformation techniques to the spectral data and examine which allows 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 note that other similar studies have been conducted that achieve a higher classification accuracy, such as Femenias et al. (2021).

In a similar study, Ma et al. (2023) 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 use 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 one-or-two days post-inoculation. Their approach achieved a reported 100% accuracy in distinguishing healthy from infected samples, and an RF accuracy 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.

4.2.2 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 Tarazona et al. (2021a), the authors employ 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 rate of *Fusarium culmorum* and *Fusarium proliferatum* and mycotoxin production, having consistently the lowest RMSE value.

Mateo et al. (2023) 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 performance of the model, in conjunction with the R^2 value. In this work, both RF and XGB showed comparable performance, 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 use multiple ML models, with a general consensus that RF models perform the best at their given tasks. See Mateo et al. (2021); Tarazona et al. (2021b); Srinivasan et al. (2022) for more details.

4.2.3   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.

4.3   Gradient Boosting

Gradient Boosting (GB; Friedman, 2001) 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 optimisa-

tion algorithm, to minimise the loss, or the difference between the actual and predicted values.
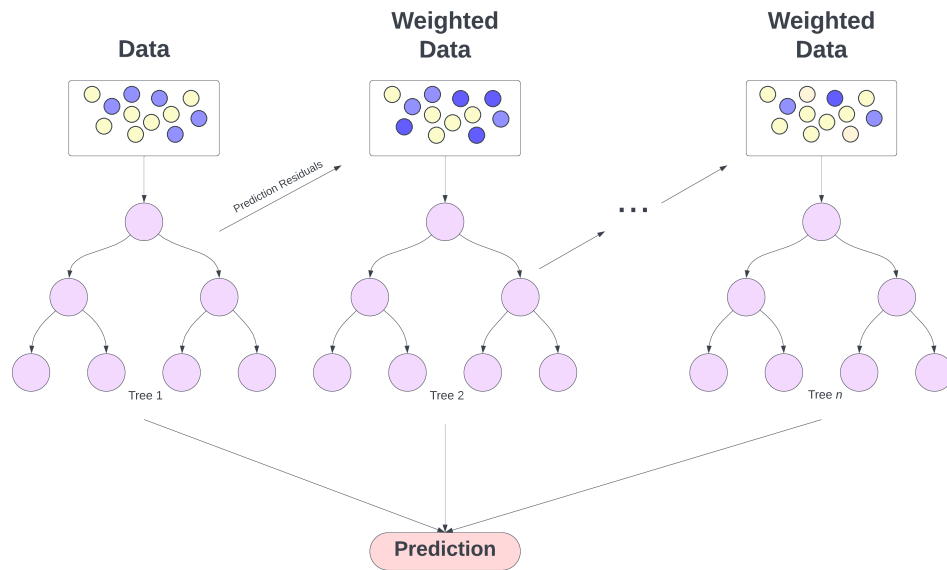

Figure 7: Gradient Boosting Process. Here the weak learners are trees that are trained sequentially on weighted data with iteratively adjusted weights based on previous prediction errors.

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

25

for both regression and classification tasks. However, these models have their challenges. Without careful tuning and regularisation, there is a risk of over-fitting, a problem exacerbated by noisy data (Natekin and Knoll, 2013). Additionally, their sequential boosting process is computationally intensive and time-consuming compared to 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; (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) 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) (Chen and Guestrin, 2016), 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.

4.3.1 GB for Spatio-Temporal Data

In a study by Wang et al. (2022b), the authors design a program for aflatoxin monitoring in feed products (peanuts and soy beans), whilst considering both the performance of the model and the cost of monitoring. In the study, they apply 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 was 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 propose a reduction to the monitoring cost of up to 96% for the years 2016 to 2018.

In Xie et al. (2022), the authors propose 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 use an RF model to determine *Aspergillus* species with 97.8% accuracy. They then go 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 note that they build the XGB model using only a single tree and use 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 which is lost if only a single tree is used.

Branstad-Spates et al. (2023) 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's) action level for corn (20*ppb*) (FDA, 2020), whereas the lower cut off is based on the European standard of 5*ppb* (EFSA, 2013). 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 (Zheng and Casari, 2018). The data used was 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 use an

imputation method, however they note 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 report that the GBM model performed well, achieving high accuracy rates of 96.8% for the $20ppb$ threshold and 90.3% for the $5ppb$ 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 aflatoxins 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.

4.3.2   GB for Spectral Data

Chavez et al. (2022) conducted a study on aflatoxin and fumonisin contamination in single kernel corn. They argue 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, as well as calculating weighted sums of toxin contamination. Additionally, they aim to improve single kernel classification performance through the use of different ML applications. Their methodology was to take corn kernels that are 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 1 Section). 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 (GLMNET); and support vector machines (SVM). They additionally applied ML algorithms for classifying each individual mycotoxin. For aflatoxin they used: bagged AdaBoost; linear discriminant analysis (LDA); and penalised logistic regression

28

(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 note 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.

4.3.3   Gradient Boosting Summary

The application of GBM across various datasets, from spatio-temporal to spec-
tral data, demonstrates 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 Xie et al. (2022), or when
handling datasets with substantial missing values, as noted by Branstad-Spates
et al. (2023). The high accuracy rates reported should be examined for potential
overfitting or lack of generalisation to broader datasets. Chavez et al. (2022)'s
approach 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.

4.4   Support Vector Machines

Support Vector Machines (SVMs; Vapnik, 1999) are a set of supervised learning
methods used for classification, regression, and outlier detection. To make pre-
dictions, SVMs identify the optimal hyperplane which 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 which influence its position and
orientation are known as support vectors, as they *support* or define the hyper-

plane. 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 (Burges, 1998). 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 (Cristianini and Shawe-Taylor, 2000):

1. Linear: No nonlinear 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: (i) Gamma: 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 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.
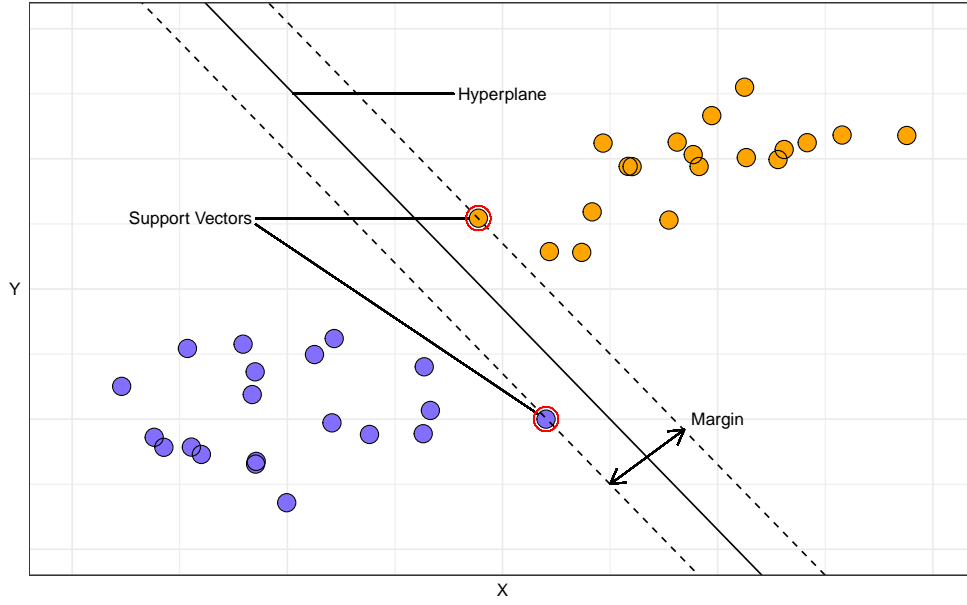
Figure 8: Support Vector Machine Process. The diagram illustrates the SVM's method of finding the optimal hyperplane that maximises the margin between two classes, depicted by the blue and orange points. The support vectors, which are the data points closest to the decision boundary, define the margin.

4.4.1 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 use spectral data. For example, Almoujahed et al. (2022) use several ML models (SVM, NN, and LR) for the classification of Fusarium head blight in wheat, using spectral data. The data was collected in the years 2020 to 2021 at a single site in Belgium, with the experiment using eight varieties of wheat. The 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 an accuracy of 82.9% and 82.5%, respectively).

In a similar study, Kim et al. (2023) use three different imaging methods alongside ML classification models to test ground corn samples for the presence

31

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

In a study concerning the detection of *Aspergillus parasiticus* in corn kernels using NIR hyperspectral imaging, conducted by Zhao et al. (2017), the authors use SVMs to compare the performance 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.

4.4.2 Support Vector Machine Summary

In the reviewed work, SVMs have 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 influence SVM performance, indicating that model robustness may be context-dependent. The choice of kernel and its parameters, like orientation, scale, and type, are critical in shaping the decision surface and thus the SVM's ability to generalise from training to unseen data

4.5 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.

4.5.1 Decision Trees

Decision tree (DT) learning is a type of non-parametric supervised learning algorithm used for both classification and regression tasks (Quinlan, 1986; Breiman et al., 1984). 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 4.2). 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 (Breiman et al., 1984). This process continues recursively until a stopping criterion is met, resulting in a tree where each path represents a decision path-

way 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 (Breiman et al., 1984). 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 Kos et al. (2016), in which they assess the use of an electronic nose to identify DON contamination of wheat samples, an extension of decision trees called Classification And Regression Trees (CART) (Breiman et al., 1984) was used to classify samples based on four thresholds of DON contamination (1750, 1250, 750, and 500 $\mu g/kg$). For this study, 214 wheat samples were collected from northern Italy during the years 2014–2015 and 2017–2018. For the threshold values of $\geq$ 1250 $\mu g/kg$, the accuracy of sample classification was the highest, ranging between 88% and 92%. The lower thresholds of $\leq$ 750 $\mu g/kg$ 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.

Kos et al. (2016) 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 $\mu g/kg$ thresholds for DON with an accuracy of 79% and 85%, respectively. Additionally, it was able to classify peanut samples for aflatoxin at 8 $\mu g/kg$ with a 77% accuracy.

In a study related to identifying and predicting risks related to the presence of fumonisins in breakfast cereal products, Purchase et al. (2023) developed a model specifically designed to predict the risk of fumonisins 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 grams, 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. Findings suggested that the fumonisins 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.

4.5.2 Bayesian Network

Bayesian networks (BN) are a type of probabilistic graphical model that use Bayesian statistics to represent and infer the conditional dependencies between different variables in a dataset (Jensen and Nielsen, 2007). 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 to linear regression models, BN models excel at analysing variable dependencies, handling non-linear interactions, and incorporating diverse types of data (Buriticá and Tesfamariam, 2015). 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, Liu et al. (2018) compare 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 was 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 note 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 note that the BN is easier to implement when the data is incomplete, when compared to the other

methods.

Guo et al. (2020) construct 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* (Babu et al., 2004), where transcription is the process by which the information encoded in a section of 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 report 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*.

4.5.3   Other ML Methods Summary

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 showing 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 include 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 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. Liu et al. (2018) compared BNs with other models for predicting DON contamination in wheat, achieving a re-

spectable 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.

5   Conclusion

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 45 studies performed within the years 2013 to 2023. The number of publications in each field has grown significantly over the ten 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.

A critical observation from this review is that the lack of detailed hyperparameter descriptions 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. 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.

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 De Girolamo et al. (2019), 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. Shen et al. (2019); Jha et al. (2021), and Milićević et al. (2019) all use PCA when dealing with high-dimensional image data.

As the field is growing, there are numerous avenues for future work. More extensive research could be conducted that directly compares different ML models under a standardised set of hyperparameters. This will provide clearer insights into the most effective techniques in specific contexts related to mycotoxin detection. Another avenue in model interpretability. Given the critical nature of food safety, future research could also focus on improving the interpretability

39

of ML models. Techniques like SHAP (SHapley Additive exPlanations) (Shapley et al., 1953) and LIME (Local Interpretable Model-agnostic Explanations) (Ribeiro et al., 2016) can be used to make the models' decisions more transparent and trustworthy.

6 Appendix

In Table 1 we provide a list of the used abbreviations and nomenclature used in this work.

Abbreviation	Meaning
NN	Neural Network
CNN	Convoluted 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 regularised generalised linear models
PLS-DA	Partial Least Squares–Discriminant Analysis
sPLS-DA	sparse Partial Least Square Determination Analysis
PCA	Principal Components 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

Table 1: Nomenclature used in our paper.

References

Almoujahed, M. B., Rangarajan, A. K., Whetton, R. L., Vincke, D., Eylenbosch, D., Vermeulen, P., and Mouazen, A. M. (2022), "Detection of fusarium head blight in wheat under field conditions using a hyperspectral camera and machine learning," *Computers and Electronics in Agriculture*, 203, 107456.

Alpaydin, E. (2020), *Introduction to machine learning*, MIT press.

Alshannaq, A. and Yu, J.-H. (2017), "Occurrence, toxicity, and analysis of major mycotoxins in food," *International journal of environmental research and public health*, 14, 632.

Anfossi, L., Giovannoli, C., and Baggiani, C. (2016), "Mycotoxin detection," *Current opinion in biotechnology*, 37, 120–126.

Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M., and Teichmann, S. A. (2004), "Structure and evolution of transcriptional regulatory networks," *Current opinion in structural biology*, 14, 283–291.

Baştanlar, Y. and Özuysal, M. (2014), "Introduction to machine learning," *miRNomics: MicroRNA biology and computational analysis*, 105–128.

Bernardes, R. C., De Medeiros, A., da Silva, L., Cantoni, L., Martins, G. F., Mastrangelo, T., Novikov, A., and Mastrangelo, C. B. (2022), "Deep-learning approach for fusarium head blight detection in wheat seeds using low-cost imaging technology," *Agriculture*, 12, 1801.

Branstad-Spates, E. H., Castano-Duque, L., Mosher, G. A., Hurburgh Jr, C. R., Owens, P., Winzeler, E., Rajasekaran, K., and Bowers, E. L. (2023), "Gradient boosting machine learning model to predict aflatoxins in Iowa corn," *Frontiers in Microbiology*, 14.

Breiman, L. (2001), "Random forests," *Machine learning*, 45, 5–32.

Breiman, L., Friedman, J., Olshen, R., and Stone, C. (1984), *Classification and Regression Trees*, Wadsworth.

Burges, C. J. (1998), "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, 2, 121–167.

Buriticá, J. A. and Tesfamariam, S. (2015), "Consequence-based framework for electric power providers using Bayesian belief network," *International Journal of Electrical Power & Energy Systems*, 64, 233–241.

Camardo Leggieri, M., Mazzoni, M., and Battilani, P. (2021), "Machine learning for predicting mycotoxin occurrence in maize," *Frontiers in Microbiology*, 12, 661132.

Campagnoli, A., Cheli, F., Savoini, G., Crotti, A., Pastori, A., and Dell'Orto, V. (2009), "Application of an electronic nose to detection of aflatoxins in corn," *Veterinary research communications*, 33, 273–275.

Canziani, A., Paszke, A., and Culurciello, E. (2016), "An analysis of deep neural network models for practical applications," *arXiv preprint arXiv:1605.07678*.

Chavez, R. A., Cheng, X., Herrman, T. J., and Stasiewicz, M. J. (2022), "Single kernel aflatoxin and fumonisin contamination distribution and spectral classification in commercial corn," *Food Control*, 131, 108393.

Chen, T. and Guestrin, C. (2016), "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794.

Cristianini, N. and Shawe-Taylor, J. (2000), *An introduction to support vector machines and other kernel-based learning methods*, Cambridge university press.

De Girolamo, A., von Holst, C., Cortese, M., Cervellieri, S., Pascale, M., Longobardi, F., Catucci, L., Porricelli, A. C. R., and Lippolis, V. (2019), "Rapid screening of ochratoxin A in wheat by infrared spectroscopy," *Food Chemistry*, 282, 95–100.

EFSA (2013), "Aflatoxins (sum of B1, B2, G1, G2) in cereals and cereal-derived food products," .

Eskola, M., Kos, G., Elliott, C. T., Hajšlová, J., Mayar, S., and Krska, R. (2020), "Worldwide contamination of food-crops with mycotoxins: Validity of the widely cited 'FAO estimate'of 25%," *Critical reviews in food science and nutrition*, 60, 2773–2789.

FDA (2020), "Guidance for Industry: Action Levels for Poisonous or Deleterious Substances in Human Food and Animal Feed," https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-industry-action-levels-poisonous-or-deleterious-substances-human-food-and-animal-feed#afla, Date Accessed: 05-11-2023.

Femenias, A., Gatius, F., Ramos, A. J., Sanchis, V., and Marín, S. (2021), "Near-infrared hyperspectral imaging for deoxynivalenol and ergosterol estimation in wheat samples," *Food Chemistry*, 341, 128206.

Friedman, J. H. (2001), "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, 1189–1232.

Gao, J., Zhao, L., Li, J., Deng, L., Ni, J., and Han, Z. (2021), "Aflatoxin rapid detection based on hyperspectral with 1D-convolution neural network in the pixel level," *Food Chemistry*, 360, 129968.

Ghilardelli, F., Barbato, M., and Gallo, A. (2022), "A preliminary study to classify corn silage for high or low mycotoxin contamination by using near infrared spectroscopy," *Toxins*, 14, 323.

Gobbi, E., Falasconi, M., Torelli, E., and Sberveglieri, G. (2011), "Electronic nose predicts high and low fumonisin contamination in maize cultures," *Food Research International*, 44, 992–999.

Grahn, H. and Geladi, P. (2007), *Techniques and applications of hyperspectral image analysis*, John Wiley & Sons.

Guo, L., Ji, M., and Ye, K. (2020), "Dynamic network inference and association computation discover gene modules regulating virulence, mycotoxin and sexual reproduction in Fusarium graminearum," *BMC genomics*, 21, 1–14.

Gurney, K. (2018), *An introduction to neural networks*, CRC press.

Han, Z. and Gao, J. (2019), "Pixel-level aflatoxin detecting based on deep learning and hyperspectral imaging," *Computers and Electronics in Agriculture*, 164, 104888.

Jensen, F. V. and Nielsen, T. D. (2007), *Bayesian networks and decision graphs*, vol. 2, Springer.

Jha, S. N., Jaiswal, P., Kaur, J., and Ramya, H. (2021), "Rapid detection and quantification of aflatoxin B1 in milk using fourier transform infrared spectroscopy," *Journal of The Institution of Engineers (India): Series A*, 102, 259–265.

Jin, X., Jie, L., Wang, S., Qi, H. J., and Li, S. W. (2018), "Classifying wheat hyperspectral pixels of healthy heads and Fusarium head blight disease using a deep neural network in the wild field," *Remote Sensing*, 10, 395.

Johns, L. E., Bebber, D. P., Gurr, S. J., and Brown, N. A. (2022), "Emerging health threat and cost of Fusarium mycotoxins in European wheat," *Nature Food*, 3, 1014–1019.

Jubair, S., Tucker, J. R., Henderson, N., Hiebert, C. W., Badea, A., Domaratzki, M., and Fernando, W. (2021), "GPTransformer: A transformer-based deep learning method for predicting Fusarium related traits in barley," *Frontiers in plant science*, 12, 761402.

Kim, Y.-K., Baek, I., Lee, K.-M., Kim, G., Kim, S., Kim, S.-Y., Chan, D., Herrman, T. J., Kim, N., and Kim, M. S. (2023), "Rapid Detection of Single-and Co-Contaminant Aflatoxins and Fumonisins in Ground Maize Using Hyperspectral Imaging Techniques," *Toxins*, 15, 472.

Kos, G., Sieger, M., McMullin, D., Zahradnik, C., Sulyok, M., Öner, T., Mizaikoff, B., and Krska, R. (2016), "A novel chemometric classification for FTIR spectra of mycotoxin-contaminated maize and peanuts at regulatory limits," *Food Additives & Contaminants: Part A*, 33, 1596–1607.

Latham, R. L., Boyle, J. T., Barbano, A., Loveman, W. G., and Brown, N. A. (2023), "Diverse mycotoxin threats to safe food and feed cereals," *Essays in Biochemistry*, EBC20220221.

Leggieri, M. C., Lanubile, A., Dall'Asta, C., Pietri, A., and Battilani, P. (2020), "The impact of seasonal weather variation on mycotoxins: Maize crop in 2014 in northern Italy as a case study," *World Mycotoxin Journal*, 13, 25–36.

Leggieri, M. C., Mazzoni, M., Fodil, S., Moschini, M., Bertuzzi, T., Prandini, A., and Battilani, P. (2021), "An electronic nose supported by an artificial neural network for the rapid detection of aflatoxin B1 and fumonisins in maize," *Food Control*, 123, 107722.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., and Bochtis, D. (2018), "Machine learning in agriculture: A review," *Sensors*, 18, 2674.

Lippolis, V., Pascale, M., Cervellieri, S., Damascelli, A., and Visconti, A. (2014), "Screening of deoxynivalenol contamination in durum wheat by MOS-based electronic nose and identification of the relevant pattern of volatile compounds," *Food Control*, 37, 263–271.

Liu, C., Manstretta, V., Rossi, V., and Van der Fels-Klerx, H. (2018), "Comparison of three modelling approaches for predicting deoxynivalenol contamination in winter wheat," *Toxins*, 10, 267.

Liu, Y. and Wu, F. (2010), "Global burden of aflatoxin-induced hepatocellular carcinoma: a risk assessment," *Environmental health perspectives*, 118, 818–824.

Logrieco, A., Battilani, P., Leggieri, M. C., Jiang, Y., Haesaert, G., Lanubile, A., Mahuku, G., Mesterházy, A., Ortega-Beltran, A., Pasti, M., et al. (2021), "Perspectives on global mycotoxin issues and management from the MycoKey Maize Working Group," *Plant disease*, 105, 525–537.

Ma, J., Guan, Y., Xing, F., Eltzov, E., Wang, Y., Li, X., and Tai, B. (2023), "Accurate and non-destructive monitoring of mold contamination in foodstuffs based on whole-cell biosensor array coupling with machine-learning prediction models," *Journal of Hazardous Materials*, 449, 131030.

Maragos, C. M. (2004), "Emerging technologies for mycotoxin detection," *Journal of Toxicology: Toxin Reviews*, 23, 317–344.

Marroquín-Cardona, A., Johnson, N., Phillips, T., and Hayes, A. (2014), "Mycotoxins in a changing global environment–a review," *Food and Chemical Toxicology*, 69, 220–230.

Mateo, E. M., Gómez, J. V., Tarazona, A., García-Esparza, M. Á., and Mateo, F. (2021), "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*, 13, 545.

Mateo, E. M., Tarazona, A., Aznar, R., and Mateo, F. (2023), "Exploring the impact of lactic acid bacteria on the biocontrol of toxigenic Fusarium spp. and their main mycotoxins," *International Journal of Food Microbiology*, 387, 110054.

Mateo, F., Gadea, R., Mateo, E. M., and Jiménez, M. (2011), "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*, 22, 88–95.

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. (2021), "Impact of mycotoxins on animals' oxidative status," *Antioxidants*, 10, 214.

McCulloch, W. S. and Pitts, W. (1943), "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, 5, 115–133.

Medina, A., Akbar, A., Baazeem, A., Rodriguez, A., and Magan, N. (2017), "Climate change, food security and mycotoxins: Do we know enough?" *Fungal biology reviews*, 31, 143–154.

Milićević, D., Petronijević, R., Petrović, Z., Djinović-Stojanović, J., Jovanović, J., Baltić, T., and Janković, S. (2019), "Impact of climate change on aflatoxin M1 contamination of raw milk with special focus on climate conditions in Serbia," *Journal of the Science of Food and Agriculture*, 99, 5202–5210.

Montavon, G., Samek, W., and Müller, K.-R. (2018), "Methods for interpreting and understanding deep neural networks," *Digital signal processing*, 73, 1–15.

Natekin, A. and Knoll, A. (2013), "Gradient boosting machines, a tutorial," *Frontiers in neurorobotics*, 7, 21.

Niedbała, G., Kurasiak-Popowska, D., Stuper-Szablewska, K., and Nawracała, J. (2020), "Application of artificial neural networks to analyze the concentration of ferulic acid, deoxynivalenol, and nivalenol in winter wheat grain," *Agriculture*, 10, 127.

Oener, T., Thiam, P., Kos, G., Krska, R., Schwenker, F., and Mizaikoff, B. (2019), "Machine learning algorithms for the automated classification of contaminated maize at regulatory limits via infrared attenuated total reflection spectroscopy," *World mycotoxin journal*, 12, 113–122.

Ottoboni, M., Pinotti, L., Tretola, M., Giromini, C., Fusi, E., Rebucci, R., Grillo, M., Tassoni, L., Foresta, S., Gastaldello, S., et al. (2018), "Combining E-nose and lateral flow immunoassays (LFIAs) for rapid occurrence/co-occurrence aflatoxin and fumonisin detection in maize," *Toxins*, 10, 416.

Panagou, E. Z. and Kodogiannis, V. S. (2009), "Application of neural networks as a non-linear modelling technique in food mycology," *Expert Systems with Applications*, 36, 121–131.

Purchase, J., Donato, R., Sacco, C., Pettini, L., Rookmin, A. D., Melani, S., Artese, A., Purchase, D., and Marvasi, M. (2023), "The association of food ingredients in breakfast cereal products and fumonisins production: risks identification and predictions," *Mycotoxin Research*, 1–11.

Qiu, R., Yang, C., Moghimi, A., Zhang, M., Steffenson, B. J., and Hirsch, C. D. (2019), "Detection of fusarium head blight in wheat using a deep neural network and color imaging," *Remote Sensing*, 11, 2658.

Quinlan, J. R. (1986), "Induction of decision trees," *Machine learning*, 1, 81–106.

Rangarajan, A. K., Whetton, R. L., and Mouazen, A. M. (2022), "Detection of fusarium head blight in wheat using hyperspectral data and deep learning," *Expert Systems with Applications*, 208, 118240.

Redmon, J. and Farhadi, A. (2017), "YOLO9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7263–7271.

Renaud, J. B., Miller, J. D., and Sumarah, M. W. (2019), "Mycotoxin testing paradigm: Challenges and opportunities for the future," *Journal of AOAC International*, 102, 1681–1688.

Ribeiro, M. T., Singh, S., and Guestrin, C. (2016), ""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*, pp. 1135–1144.

Saha, D. and Manickavasagan, A. (2021), "Machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review," *Current Research in Food Science*, 4, 28–44.

Shapley, L. S. et al. (1953), "A value for n-person games," .

Shen, F., Zhao, T., Jiang, X., Liu, X., Fang, Y., Liu, Q., Hu, Q., and Liu, X. (2019), "On-line detection of toxigenic fungal infection in wheat by visible/near infrared spectroscopy," *Lwt*, 109, 216–224.

Soares, R. R., Ricelli, A., Fanelli, C., Caputo, D., de Cesare, G., Chu, V., Aires-Barros, M. R., and Conde, J. P. (2018), "Advances, challenges and opportunities for point-of-need screening of mycotoxins in foods and feeds," *Analyst*, 143, 1015–1035.

Srinivasan, R., Lalitha, T., Brintha, N., Sterlin Minish, T., Al Obaid, S., Alharbi, S. A., Sundaram, S., and Mahilraj, J. (2022), "Predicting the Growth of F. proliferatum and F. culmorum and the Growth of Mycotoxin Using Machine Learning Approach," *BioMed Research International*, 2022.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014), "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, 15, 1929–1958.

StatSoft, Inc. (2020), "STATISTICA (Data Analysis Software System), Version 7.1," http://www.statsoft.com.

Tarazona, A., Mateo, E. M., Gómez, J. V., Gavara, R., Jiménez, M., and Mateo, F. (2021a), "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," *International journal of food microbiology*, 338, 109012.

Tarazona, A., Mateo, E. M., Gómez, J. V., Romera, D., and Mateo, F. (2021b), "Potential use of machine learning methods in assessment of Fusarium culmorum and Fusarium proliferatum growth and mycotoxin production in treatments with antifungal agents," *Fungal biology*, 125, 123–133.

Teixido-Orries, I., Molino, F., Femenias, A., Ramos, A. J., and Marín, S. (2023), "Quantification and classification of deoxynivalenol-contaminated oat samples by near-infrared hyperspectral imaging," *Food Chemistry*, 417, 135924.

The World Health Organization (WHO) (2023), "Food Safety. The World Health Organization," https://www.who.int/news-room/fact-sheets/detail/mycotoxins, Date Accessed: 05-11-2023.

Tola, M. and Kebede, B. (2016), "Occurrence, importance and control of mycotoxins: A review," *Cogent Food & Agriculture*, 2, 1191103.

Torelli, E., Firrao, G., Bianchi, G., Saccardo, F., and Locci, R. (2012), "The influence of local factors on the prediction of fumonisin contamination in maize," *Journal of the Science of Food and Agriculture*, 92, 1808–1814.

Van der Fels-Klerx, H., Liu, C., Focker, M., Montero-Castro, I., Rossi, V., Manstretta, V., Magan, N., and Krska, R. (2022), "Decision support system for integrated management of mycotoxins in feed and food supply chains," *World Mycotoxin Journal*, 15, 119–133.

Vapnik, V. (1999), *The nature of statistical learning theory*, Springer science & business media.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017), "Attention is all you need," *Advances in neural information processing systems*, 30.

Wang, X., Bouzembrak, Y., Lansink, A. O., and van der Fels-Klerx, H. (2022a), "Application of machine learning to the monitoring and prediction of food safety: A review," *Comprehensive Reviews in Food Science and Food Safety*, 21, 416–434.

Wang, X., Bouzembrak, Y., Oude Lansink, A., and Van der Fels-Klerx, H. (2022b), "Designing a monitoring program for aflatoxin B1 in feed products using machine learning," *npj Science of Food*, 6, 40.

Whitaker, T. B. (2003), "Standardisation of mycotoxin sampling procedures: an urgent necessity," *Food control*, 14, 233–237.

Wu, F. (2015), "Global impacts of aflatoxin in maize: trade and human health. World Mycotoxin J. 8, 137–142," .

Xie, H., Wang, X., van der Hooft, J. J., Medema, M. H., Chen, Z.-Y., Yue, X., Zhang, Q., and Li, P. (2022), "Fungi population metabolomics and molecular network study reveal novel biomarkers for early detection of aflatoxigenic Aspergillus species," *Journal of Hazardous Materials*, 424, 127173.

Zhao, X., Wang, W., Chu, X., Li, C., and Kimuli, D. (2017), "Early detection of Aspergillus parasiticus infection in maize kernels using near-infrared hyperspectral imaging and multivariate data analysis," *Applied Sciences*, 7, 90.

Zheng, A. and Casari, A. (2018), *Feature engineering for machine learning: principles and techniques for data scientists*, " O'Reilly Media, Inc.".

Zhou, L., Zhang, C., Liu, F., Qiu, Z., and He, Y. (2019), "Application of deep learning in food: a review," *Comprehensive reviews in food science and food safety*, 18, 1793–1811.

Zingales, V., Taroncher, M., Martino, P. A., Ruiz, M.-J., and Caloni, F. (2022), "Climate change and effects on molds and mycotoxins," *Toxins*, 14, 445.