applied sciences

MDPI

Review

Enhancing Food Integrity through Artificial Intelligence and Machine Learning: A Comprehensive Review

Sefater Gbashi * and **Patrick Berka Njobeh**

Department of Biotechnology and Food Technology, Faculty of Science, University of Johannesburg, Doornfontein Campus, P.O. Box 17011, Gauteng 2028, South Africa; pnjobeh@uj.ac.za
* Correspondence: sefatergbashi@gmail.com; Tel.: +27-1-919-316-8640

Abstract: Herein, we examined the transformative potential of artificial intelligence (AI) and machine learning (ML) as new fronts in addressing some of the pertinent challenges posed by food integrity to human and animal health. In recent times, AI and ML, along with other Industry 4.0 technologies such as big data, blockchain, virtual reality, and the internet of things (IoT), have found profound applications within nearly all dimensions of the food industry with a key focus on enhancing food safety and quality and improving the resilience of the food supply chain. This paper provides an accessible scrutiny of these technologies (in particular, AI and ML) in relation to food integrity and gives a summary of their current advancements and applications within the field. Key areas of emphasis include the application of AI and ML in quality control and inspection, food fraud detection, process control, risk assessments, prediction, and management, and supply chain traceability, amongst other critical issues addressed. Based on the literature reviewed herein, the utilization of AI and ML in the food industry has unequivocally led to improved standards of food integrity and consequently enhanced public health and consumer trust, as well as boosting the resilience of the food supply chain. While these applications demonstrate significant promise, the paper also acknowledges some of the challenges associated with the domain-specific implementation of AI in the field of food integrity. The paper further examines the prospects and orientations, underscoring the significance of overcoming the obstacles in order to fully harness the capabilities of AI and ML in safeguarding the integrity of the food system.

Keywords: artificial intelligence; machine learning; food integrity; food safety; food quality control; food hazards; nutritional health

Citation: Gbashi, S.; Njobeh, P.B. Enhancing Food Integrity through Artificial Intelligence and Machine Learning: A Comprehensive Review. *Appl. Sci.* **2024**, *14*, 3421. https://doi.org/10.3390/app14083421

Academic Editor: Marco Iammarino

Received: 7 January 2024
Revised: 17 March 2024
Accepted: 8 April 2024
Published: 18 April 2024

1. Introduction

Food integrity is a vital aspect of the food industry and encompasses the preservation of the safety, quality, and authenticity of food throughout the entire food supply chain, from its production and handling to its delivery and consumption. Important elements of food integrity encompass the prevention of food fraud, the assurance of food safety, the guarantee of traceability across the supply chain, and the promotion of sustainability and transparency in food production and distribution processes, with the ultimate goal of protecting the health and well-being of consumers. With regard to food safety, it continues to present serious hazards to public health, given the persistent threats posed by contamination incidents and food-associated disease outbreaks [1–5], being driven by the ever-increasing demand for cheaper products, low-cost manufacturing, and the increasing globalization of the food supply chain, amongst other factors. According to the World Health Organization (WHO) estimates, consumption of unsafe food results in the loss of 33 million healthy years of life annually [6,7]. Diarrheal diseases alone are implicated in about 50% of the global burden of foodborne diseases, causing illness in at least 550 million people and resulting in 230,000 annual deaths annually [8]. In fact, it is acknowledged that these figures are conservative [9] as the net societal impact of food safety and microbiological food hazards including morbidity and consequent immobility, disability, mortality, costs of diagnosis/treatment, etc., are difficult to accurately estimate but unequivocally prodigious. Moreover, the disproportionate impact of food safety on designated societal groups (e.g., children, the elderly, the poor, and other vulnerable members of society who bear the brunt of this effect) further exacerbates this dire situation.

In response, government agencies and relevant health actors along the food supply chain often review food safety policies and tighten regulations in order to mitigate public health concerns. For example, recent high levels of ochratoxin A (OTA), a potent mycotoxin, in certain food products prompted the European Commission to take action by issuing Regulation (EU) 2022/1370 on 5 August 2022 [10], amending Regulation (EC) No. 1881/2006 and adding maximum levels (MLs) for OTA in certain foodstuffs. While previous regulations already established MLs for OTA in selected foods, the absence of limits for other food items contributing to human exposure necessitated this amendment. Additionally, the updated regulation introduced new MLs for various foodstuffs and tightened existing limits for specific food items. These changes aimed to enhance public health by ensuring stricter adherence to food processing and handling standards [10–12].

Indeed, food integrity poses a complex and multidimensional challenge and traditional methods of monitoring food quality and safety rely on manual inspection and testing, which can be time-consuming and prone to human error. To effectively address these concerns, food businesses must develop robust monitoring systems that can detect and prevent food quality issues early in the production process. This is where AI and ML technologies are emerging as a beacon of promise and a transformative frontier in addressing these critical challenges within the food system. AI and ML have the potential to tackle complex problems and unearth innovative solutions with real-life applications, including improving the accuracy, efficiency, and effectiveness of food safety and quality monitoring systems. The US Food and Drug Administration (FDA) also acknowledges the role of AI and ML in food safety. In an April 2019 publication, the FDA outlined measures to lead the US into a more advanced phase of food safety, advocating the utilization of AI and ML applications in the field [13,14]. These technologies can analyze vast amounts of data in real-time, enabling rapid and precise identification of potential hazards or deviations from quality standards, and determine which products should be given priority for monitoring, as well as forecast future food safety events or outcomes. These models also provide insights into the timing and location of monitoring activities throughout the food supply chain, as well as identifying the origins of ingredients, enabling quicker and more precise recalls in the event of safety violations. AI can assist food firms in complying with rigorous regulatory standards and frameworks and other essential quality criteria.

In recognition of the potential of AI and ML to revolutionize food integrity practices, this paper delves into the diverse applications of AI in the field of food integrity, offering a comprehensive review of recent developments and their implications. From leveraging computer vision for food inspection, sorting, and grading to the prediction of pathogenic microbial contamination patterns, each application represents a step toward a more secure and reliable food supply chain. In looking forward, we deliberate the prospects and directions that will be pivotal in harnessing the full potential of AI and ML for the benefit of food safety and public health on a global scale.

2. Methodology

A comprehensive semi-formal literature search methodology was adopted in this study following a modified workflow described by Gbashi et al. [15] and Vinci et al. [16].

2.1. Data Sources

Various internet search engines were utilized for digital data collection. Google Scholar and Google search engines were queried using numerous English-language keywords related to the study. In addition, Scopus, Web of Science, PubMed, and other databases were used to retrieve relevant publications and information on the study title. Some of the query terms included in the search are artificial intelligence, machine learning, food integrity, food quality, food integrity, food contamination, food pathogens, applications of AI and ML in food safety and quality, etc.

2.2. Study Inclusion and Exclusion Criteria

Following an internet search, an inclusive data selection and collection method was implemented and the final consulted literature included relevant systematic and meta-analytic reviews, non-systematic critical reviews, technical papers, theses, communications, reports, blog posts, and web pages, among others, published in the English language. The returned publications were critically scrutinized and evidently irrelevant articles were excluded. After the inclusion criteria had been applied, all duplicate papers were removed while all pertinent articles were retrieved, collated, reviewed a second time, and relevant information abstracted.

3. Artificial Intelligence (AI) and Machine Learning (ML)

AI and ML are two related but separate topics in computer science. AI seeks to replicate human intelligence in machines, while ML focuses on creating algorithms that allow computers to learn from data (Figure 1 and Table 1). In the succeeding sections of this article, the differences and characteristics of these two complementary fields will be explored more in-depth.

Table 1. Differences between AI and ML.

Aspect	AI	ML
Description	AI is the emulation of human intelligence processes by machines, particularly computer systems. It entails creating systems or algorithms capable of carrying out activities that usually demand human intelligence, like comprehending natural language, identifying patterns, acquiring knowledge from experience, and making judgements	ML is a branch of AI that concentrates on creating algorithms and statistical models to empower computers to carry out tasks without absolute instructions.
Goal/objective	To create systems with the ability to carry out activities that usually demand human intelligence	To allow machines to enhance their performance on a particular activity by learning from data without the need for explicit programming
Focus/scope	Has a broader scope, encompassing a wide range of approaches, including ML	More focused on extracting patterns and insights from data to enhance the performance of a certain activity or application.
Techniques and methodologies	Incorporates a variety of methods and techniques, including computer vision, natural language processing, expert systems, etc.	Primarily focuses on mathematical and statistical techniques algorithms for analyzing data and forecasting outcomes
Data dependency	May or may not depend strongly on acquired data	Strongly data-dependent for model training and predictions
Human intervention	Capable of functioning autonomously or with human involvement	Dependent on human involvement for training, validation, and adjusting parameters
Decision making	Can make decisions based on predetermined rules or acquired patterns	Utilizes data-driven patterns to make decisions.
Adaptability	Capable of adjusting to novel circumstances by following predetermined guidelines or assimilating information from fresh data	Enhances and refines its performance through increased data and iterations
Examples	Chatbots, expert systems, virtual assistants, autonomous cars, robotics, natural language processing, computer vision, game-playing AI, etc.	Regression, classification, clustering algorithms, neural networks, decision trees, SVM, and reinforcement learning algorithms are examples of ML techniques

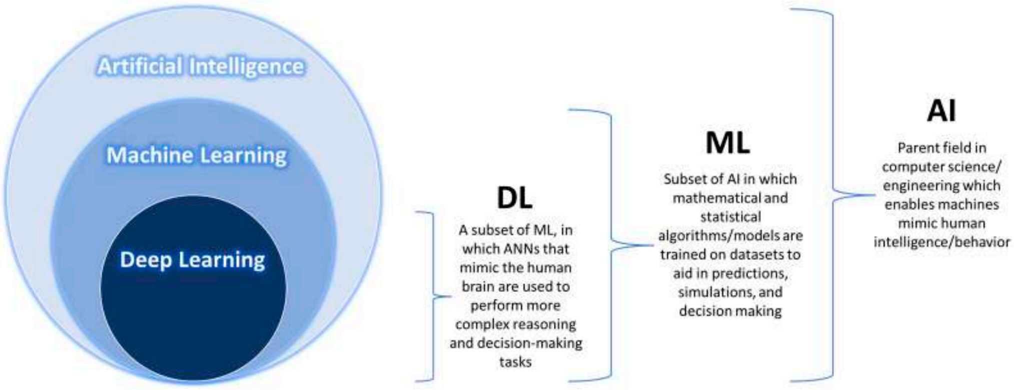

Figure 1. Artificial intelligence vs. machine learning.

3.1. Artificial Intelligence (AI)

AI is a wide-ranging domain within the discipline of computer science, focusing on the creation and advancement of computer systems and related technologies that possess the ability to execute tasks that conventionally necessitate human intelligence [17–19]. The aforementioned tasks encompass a range of cognitive abilities, such as logical thinking, experiential learning, problem solving, comprehension of natural language, perception, and the capacity to interact with and modify the surrounding environment [18,20].

Historically, AI emerged as an academic discipline in the 1950s when researchers aimed to develop robots with the ability to think, solve problems, and learn independently [21,22]. The field remained largely unpopular until the 2000s [21], with repeated cycles of optimism, followed by periods of disappointment and loss of funding, referred to as AI winter [23]. The introduction of data-driven algorithms and ML revolutionized the field by offering the unique dimension of learning from historical information to enhance problem-solving. Funding and interest significantly rose after 2012 when deep learning methodologies exhibited significant learning and predictive capabilities, outperforming perhaps all previous AI techniques [24,25], and after 2017 with the introduction of the transformer architecture [25,26]. In the early 2020s, a surge in AI development occurred, primarily led by entities in the US such as big technology firms, universities, and laboratories, resulting in notable progress in the field.

AI can be categorized into two distinct types: Narrow or Weak AI and General or Strong AI [27]. Weak AI is specifically engineered to carry out a particular activity within set limits (i.e., they are specialized and have a restricted range of use). They do not possess the broad cognitive capabilities seen in human intelligence. Weak AI's primary advantage is its practical utility in several fields, allowing for automation, optimization, and improvement in particular jobs. These systems are commonly utilized in several industries like healthcare, finance, manufacturing, and customer service to optimize processes, increase decision making, and improve user experiences. Weak AI also has significant limitations. These systems have limited comprehension outside of their planned scope, which makes it difficult for them to deal with unforeseen inputs or situations. Concerns about data privacy, prejudice, and ethical implications emerge when these systems handle sensitive information and impact decision-making processes. On the other hand, strong AI seeks to mimic human intelligence in several areas and possesses the capacity to comprehend, acquire knowledge, and apply it across a wide spectrum of tasks, akin to human intelligence [27–30].

3.2. Machine Learning (ML)

ML, on the other hand, is a subset of AI that is focused on the creation and refinement of algorithms and statistical models that can effectively learn from historical data and thus accomplish tasks (such as making decisions or predictions) without explicit instructions/programming tailored to the specific task at hand [19,22]. The fundamental principles and methodologies of ML include supervised learning, unsupervised learning, and reinforcement learning [31,32].

Supervised learning: In supervised learning, models are trained using labeled data, where each input datum point is associated with a specific target label. During the training, the algorithm establishes the relationship between the input and output by modifying its internal parameters to reduce the discrepancy between the predicted output and the real output, known as the error or loss. The trained algorithm can thus be utilized for predicting or categorizing new unseen data [31–33]. Supervised learning can be further classified into two primary types: classification and regression [34]. Classification involves predicting the category or class label of future observations using past observations with known labels. The output variable is categorical, indicating that it assumes a finite set of values. Whereas regression problems involve predicting a continuous numerical value using input features. The output/response variable is continuous, allowing it to assume any value within a specific range. Some of the underlying algorithms of supervised learning include linear regression, support vector machine (SVM), logistic regression, ANN, gradient-boosted regression/classification, random forest, etc. [34–37].

Unsupervised learning: Unsupervised learning involves models identifying patterns in data without the need for labeled examples or explicit guidance [38]. The goal is to investigate the inherent organization of data and decipher concealed distributional patterns, clusters, or connections in the data without predetermined labels or intended results [38]. Unsupervised learning is especially beneficial in situations where there is a lack of labeled data, when obtaining such data is costly, or when the data's underlying structure is not fully comprehended. Tasks, methods, and approaches associated with unsupervised learning include clustering methods, dimensionality reduction strategies, anomaly detection, association rule learning algorithms, etc. [38–40]. Clustering methods seek to categorize related data points into groups or clusters using a similarity measure. The goal is to divide the data into separate groups where data points in the same group are more alike to each other than to those in other groups. Clustering algorithms such as hierarchical clustering and K-means clustering are commonly used [38,39]. Dimensionality reduction strategies focus on decreasing the number of features, characteristics, or variables in the data while retaining critical data information. This can aid in visualizing data with many dimensions, eliminating interference, and expediting further processing. Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) are popular unsupervised ML techniques used for dimensionality reduction and data exploration [41,42]. Anomaly identification, sometimes referred to as outlier detection, is the process of discovering uncommon or unusual patterns or data points that differ significantly from the typical data distribution [43]. Anomaly detection techniques encompass distance-based methods, density-based techniques, isolation forests, etc.

Reinforcement learning: This category of ML entails models learning through interactions with an environment and obtaining feedback, in the form of incentives or punishments [44–46]. The fundamental elements of a reinforcement learning system include the agent, the one who engages with the environment to learn or make decisions, and the environment, the external system the agent interacts with. Actions refer to the range of potential choices or maneuvers available to the agent. The state refers to the present condition or arrangement of the surroundings. Rewards are feedback from the environment that indicates whether the agent's activities were successful or not. Policy refers to the systematic approach or plan that guides an agent in selecting actions based on different states to make judgments [44,47,48]. In general, reinforcement learning aims to identify the most effective strategy that maximizes the total reward received over a period of time through interacting with the environment, obtaining rewards or penalties as feedback, and adjusting its policy accordingly [45–47]. Table 1 provides a systematic overview of the distinctions between AI and ML.

4. Application of AI in Food Integrity

The application of AI and ML in food integrity has many dimensions (Figure 2). Until recently, a greater part of food safety and quality evaluation procedures depended on manual human labor and inspection, which involved performing arduous, repetitive, and time-consuming assays with increased vulnerabilities to human bias and errors. These factors underscore the importance of the increasing capabilities of AI, ML, and analog technologies in playing a viable role in food integrity activities throughout the food supply chain. Below, some of the applications of AI and ML in ensuring food integrity are discussed.

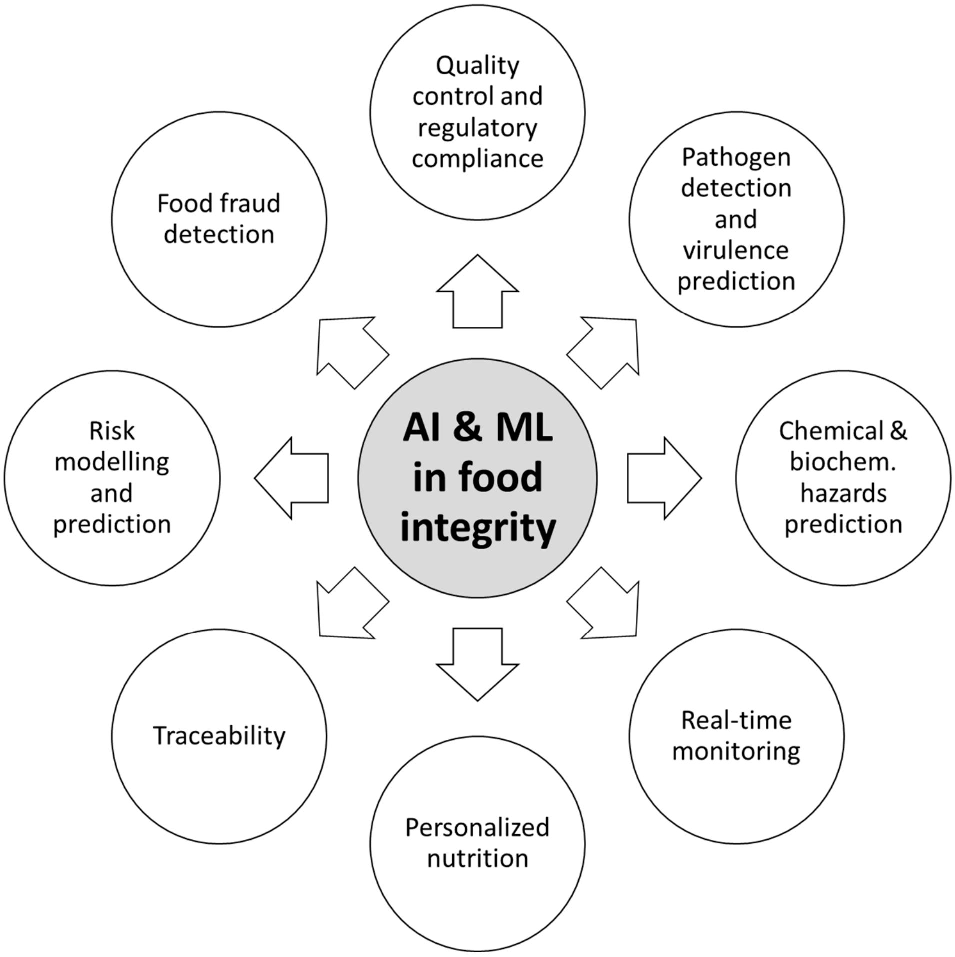

Figure 2. Dimensions of AI and ML in food integrity.

4.1. AI and ML in Food Quality Control and Regulatory Compliance

Relevant players in the food value chain are progressively embracing AI and ML technologies to enhance safety inspection and quality control operations, as well as optimize their operations in order to minimize errors and deviations from established quality standards. A case in point is that of AgShift [49], a US-based food company that implements an AI-powered automated food inspection system. The company's primary product, the "Hydra" platform, utilizes AI, IoT, and computer vision to automate and standardize evaluations of food quality. The system employs deep learning techniques in an automated workflow to examine photos of food products, detecting flaws and evaluating them to guarantee adherence to industry standards and regulations. This approach improves the efficiency of inspections, decreases human errors, upholds consistent quality control, and empowers food firms to limit waste and provide safer superior products to consumers [49,50].

Computer vision, one of the critical technologies implemented by AgShift, facilitates the extraction of intricate insights from digital images, videos, and other visual inputs that are consumed and comprehended by AI algorithms that use such to make informed judgments and take suitable actions. In food processing, computer vision is employed for quality control by identifying defects, ensuring proper labeling, and assessing overall product quality. Manual food quality inspections and quality control in the process stream are suboptimal in many ways. On the one hand, the task is frequently perceived as monotonous, demanding sustained attention from individuals for extended durations, a cognitive aspect in which human brains are notoriously deficient. Nonetheless, with the aid of AI and ML technologies, automated systems are able to establish accurate and consistent quality characteristics of different food products ranging from differences in shapes, sizes, colors, textures, and nutraceutical compositions of different food commodities and maintain quality during processing.

Chen et al. [51] implemented an ML-driven computer vision system to discriminate defective rice grains with flaws such as fracture, chalkiness, or spotting. The method employed near-infrared pictures and a support vector machine (SVM) classifier for computer intelligence and decision making. The results of the study indicated that the system was very effective, with an accuracy of 99.3% in identifying fractured kernels, a 96.3% accuracy in identifying chalky grain regions, and a 93.6% accuracy in identifying impaired/spotted grain areas, all within a record average execution time of 0.15 s. Alavi [52] proposed a Mamdani fuzzy inference system (MFIS) as a decision-making tool for categorizing and grading Mozafati dates according to three quality parameters: juice quantity, size, and freshness. Their ML-based grading method demonstrated an 86% conformity with the human-based grading method. Moreso, the model was more accurate than that of human specialists in date classification. Ireri et al. [53] developed an ML-powered computer vision system that accurately identified and graded tomatoes based on visual appearance (from RGB images). For defect detection, the system deployed a radial basis function (RBF) kernel-SVM classifier, which effectively detected calyx and stalk scars, achieving an accuracy of nearly 95% for both defect-free and damaged tomatoes. Overall, the system demonstrated great promise as an inline tomato sorting tool, capable of ensuring adherence to quality standards while streamlining the process.

Another computer vision approach being implemented for food quality inspection is hyperspectral imaging for compositional analysis of food products, which does not require any form of sample preparation [54]. The utilization of spectral signatures of different chemical constituents facilitates the process of composition mapping in food materials and products. Kamruzzaman et al. [55] utilized near-infrared hyperspectral imaging (900–1700 nm) in combination with partial least squares (PLS) regression to predict the nutritional/chemical composition of mutton (lamb meat). The constructed models demonstrated strong predictive ability with coefficients of determination (R^2) ranging from 0.63 to 0.88. In a different study, Rivera et al. [56] demonstrated the capability of KNN modeling of near-infrared hyperspectral images for the detection and classification of mechanically-induced damage in ‘Manila’ mangos at different ripeness stages. A classification accuracy of up to 97.9% was achievable on the third-day post-damage using the KNN model. Liu et al. [57] detected fecal contamination in apple skins using hyperspectral reflectance imaging techniques and various image processing algorithms, followed by PCA dimensional reduction. Gómez-Sanchis and colleagues [58] utilized a hyperspectral imaging system combined with various algorithms such as classification and regression trees (CART) and linear discriminant analysis (LDA) for early detection of rottenness caused by *Penicillium digitatum* in mandarins. Lee et al. [59] described a novel computer vision system utilizing an AI model based on the Mask region-based convolutional neural network (Mask-RCNN) architecture to detect tomato fruits and estimate their size and weight. Their trained model showed strong efficiency at identifying and distinguishing multiple occurrences of tomato fruit in intricate environmental conditions and the estimated dimensions by the model exhibited a promising correlation with the real dimensions.

4.2. AI and ML in Food Supply Chain Traceability

The US FDA Food Safety Modernization Act (FSMA) addresses the need for rapid and effective food tracking and tracing [60]. Food traceability refers to the ability to systematically track (both backward and forward) the origin, location, and transformations of a food product and its ingredients at any stage of the supply chain, including manufacturing, processing, transport, distribution, and storage. Globalization and the increased complexity of the food supply chain, characterized by numerous actors and logistical steps, coupled with recent high-profile fraud and safety incidents, have fueled a heightened demand for transparency in food provenance and traceability [61]. Traceability helps consumers, food companies, inspectors, and government agencies to locate the source of a food product, assess the magnitude of potential food safety issues, and determine when and where such issues may have occurred, aiding in faster elimination of the affected product from the supply chain, thus minimizing public health situations. Traceability also ensures compliance with guidelines, reduces risks, promotes trust and responsible sourcing, improves supply chain efficiency and transparency, and enhances fair competition.

Against the backdrop of novel technological inventions such as blockchain, 5G, big data, and IoT, food traceability in the Industry 4.0 age seems to greatly prosper, addressing critical vulnerabilities to fraudulent activities in the food supply chain that have heretofore remained elusive. The incorporation of these technologies presents novel opportunities for attaining more intelligent food traceability. AI, which is based on the principles of data science, is compatible with blockchain's data storage approach, such as "Hyperledger", as it embodies data intelligence. The effectiveness of the convergence of these technologies is attributed to their collective emphasis on data. The application of blockchain's data decentralization, immutability, consensus algorithms, and smart contracts can enhance the development of reliable AI systems [62]. An example in point is the collaboration between Walmart and IBM to implement a food traceability system based on the Hyperledger Fabric, allowing for swift source tracking of agricultural commodities within seconds. In China, this technology was utilized to track pork in the supply chain [63]. Furthermore, IBM, Walmart, and JD.com, in partnership with Tsinghua University National Engineering Laboratory for E-Commerce Technologies, declared the establishment of the Blockchain Food Safety Alliance [64]. The solution, created by IBM's blockchain platform, has a global scope and involves major participants such as Dole, Nestlé, and Tyson Foods. Its objective is to enhance efficiency, transparency, and authenticity in food supply chains on a global scale [64].

Shahbazi et al. [65] presented a food traceability system for perishable foods based on blockchain, ML, and fuzzy logic. This novel system is called the Blockchain Machine Learning-based Food Traceability System (BMLFTS) and it incorporates a shelf-life management component. According to the authors, the proposed food traceability system enhances the efficiency of the current supply chain environment and produces reliable tracking, monitoring, and food quality outcomes. Li et al. [66] investigated the geographical origin of refined sugar using high-resolution mass spectrometry and the SVM algorithm, with a demonstrated classification accuracy of 83.3%. In a different study, the geographical origin of Saanen goat milk from Guangdong, Shaanxi, and Inner Mongolia Provinces in China was traced using mass spectrometric data and ML (OPLS-DA) modeling [67]. Liu et al. [68] successfully employed a lipidomics technique in combination with PCA and partial least squares-discriminate analysis (PLS-DA) to determine the geographical origin of *Crassostrea gigas* oyster species from three different maritime locations in China. In another study, Balamurugan et al. [69] described the development of a refined ML classifier to detect anomalies such as contaminated food items along the supply chain that may need to be recalled. The monitoring system was enhanced by integrating internet-of-things (IoT) and the Bayes classifier algorithm with the Food Supply Chain Management (FSCM) system, which resulted in a secure framework for effective monitoring, tracing, and administration of the entire food supply chain, encompassing producers, exporters, and buyers. Wang et al. [70] described an enhanced traceability system for food quality assurance and evaluation based on AI and ML technologies. Their proposed system integrates fuzzy classification and ANN models that allow for both forward tracking and varied tracing along the supply chain to offer real-time evaluations of food quality along the entire supply chain. The researchers conducted a case study with a pork producer and the findings revealed the system's efficacy in guaranteeing food quality and delivering precise assessments. Using an open-source IoT-enabled spectrometer, Violino et al. [71] demonstrated an AI-powered approach for tracking the geographical origin of Italian Extra Virgin Olive Oil (EVOO). Their trained ANN model had a 94.6% accuracy in the classification of the oil samples.

4.3. AI and ML in Predictive Analytics for Risk Assessment

Through rigorous analysis of extensive food safety datasets, predictive AI and ML models can accurately estimate the probability of future occurrences and identify specific areas that require attention, as well as assist in developing effective food safety management systems [72]. This information is used to build interventions that allocate resources more effectively while helping food safety specialists make well-informed decisions to guarantee a more secure food supply chain [50,72]. In the study by Wu and Weng [73], numerous ensemble learning models were utilized to forecast food safety hazards, with a specific emphasis on enhancing border inspection techniques for imported food in Taiwan. Online implementation of their approach models resulted in a notable improvement in the non-conformity hit rate, demonstrating the efficacy of ensemble learning in predicting food safety risks. In a different approach, Zhang [74] constructed a food safety risk intelligent early warning system leveraging the predictive prowess of SVM that can assess and signal food safety risks.

Liu et al. [75] created an ML-based automated food safety early warning system for the dairy supply chain with the goal of detecting indicators that precede the emergence of food safety hazards from adjacent domains of the food supply chain, referred to as "drivers of change". The European dairy supply chain was used as a case study during the implementation of the system. The authors adopted a Bayesian network to detect chemical food safety risks in milk. Throughout the study period (2008–2019), anomalies were regularly detected in the data across all nations examined, offering the opportunity for prompt preventive actions by food manufacturers or inspectors to tackle emerging food safety issues. In a different study, Rortais et al. [76] utilized an unsupervised ML model, specifically the Latent Dirichlet Allocation topic model, to analyze the Europe Media Monitor Medical Information System (EMM/MEDISYS) corpus in order to promptly identify instances of food fraud (beeswax adulteration) reported in the media. A total of 2276 news articles were collected and categorized into 10 subjects. Each topic showed a different level of connection to beeswax adulteration. The themes also indicated the potential for new hazards in the cosmetic and food packaging industries. Additionally, some of the themes emphasized the rise of fresh market prospects for beeswax. Overall, the adopted Latent Dirichlet Allocation topic model demonstrated abilities in efficiently analyzing vast media communications and creating targeted food fraud filters on EMM/MEDISYS, which could directly aid stakeholders responsible for monitoring, evaluating, and addressing food fraud instances.

Marvin et al. [77] developed an AI model to accurately estimate the likelihood of food fraud events. The Bayesian network model developed in the study was derived from an analysis of 1393 instances of food fraud data from 15 distinct sources. The model successfully identified product categories with the greatest likelihood of being involved in fraudulent activities with a prediction accuracy of 91.5%. Elsewhere, a similar model was developed to forecast significant food safety risks and their corresponding food products, with an emphasis on herbs and spices [78]. The model's predictive accuracy exceeded 85%. Zhang et al. [79] investigated the use of an extreme learning machine (ELM) model to predict food safety risks in dairy products. This early-warning model implemented the kernel-based extreme learning machine (K-ELM), which achieved a prediction accuracy of 86%, showing its potential to guarantee the quality and safety of food products. Nogales et al. [80]

evaluated different ML models, both neural and non-neural, for the prediction of food safety risks using data from the European Union Rapid Alert System for Food and Feed (RASFF). The models that comprised logistic regression, decision trees, random forest, boosted tree, SVM, Support Vector Regression (SVR), and multilayer perceptron (MLP) were applied at three specific stages of a simplified RASFF system, each contributing to an intermediate output. Results from the experiments revealed that utilization of deep learning with entity embedding yielded the most accurate results, with accuracies ranging from 82.31 to 88.94% in the three distinct stages of the simplified RASFF procedure.

4.4. AI and ML in Real-Time Monitoring of Food in the Supply Chain

With the decentralized nature of the food supply chain, which presents difficulties for governmental oversight, the integration of AI and IoT in these supply enterprises offers peculiar benefits for real-time monitoring and assessment of food product quality and safety. Real-time monitoring constitutes the continuous and instantaneous collection, analysis, presentation, and reporting of data as events unfold. It entails the utilization of technologies and systems that offer prompt or nearly instantaneous feedback on a process, system, or environment. The main objective of real-time monitoring is to facilitate timely decision making, intervention, or response using the most current information accessible. With the integration of AI and ML technologies, real-time monitoring often involves the use of smart sensors and other data collection devices at critical stages of the supply chain. The gathered data are thereafter transmitted to AI systems for immediate processing and predictive modeling with the goal of identifying any deviations from the specified standards, enabling prompt corrective measures, and improving overall process effectiveness and safety.

Alfian et al. [81] developed and demonstrated an ML-enhanced RFID-based system in combination with sensors to track and monitor the temperature and humidity of perishable foods along the supply chain (i.e., during storage and transportation). In the study by Liu et al. [82], an IoT-based solution was proposed for food safety and quality control. This system was evaluated via a pilot project in China known as the Internet of Agricultural Things (AIoT), which incorporated cutting-edge technologies to create a streamlined approach for monitoring and tracing the food supply chain, specifically targeting food safety issues. The data and information synthesized from the so-called AIoT system were presented in an easily accessible and user-friendly manner, allowing both consumers, suppliers, and supervisors to make well-informed decisions when purchasing and supervising food supplies, presenting a hopeful resolution to augment the transparency, traceability, and ultimately, the security of food supply chains. Khan et al. [83] proposed a combination of advanced deep learning (ADL) and IoT-blockchain technologies to optimize the provenance of the food supply chain within the context of Food Industry 4.0. Their proposed system allowed food consumers to verify the origin and method of food distribution before consuming it, promoting transparency. The technology also ensures that goods are maintained at the proper temperature throughout the whole supply chain.

An AI-powered early warning system based on an analytic hierarchy process integrated extreme learning machine (AHP-ELM) was proposed by Geng et al. [84] to address the complexities and intricacies of food safety inspection data. The algorithm extracts effective process characteristic information (PCIs) from the inspection data by employing the analytic hierarchy process (AHP) model. The study concluded that the proposed AHP-ELM model was effective and feasible in processing complex food inspection data, contributing to improving the quality of food products, ensuring food safety, and reducing the overall risk of food safety incidents. Tian [85] conducted a study showcasing a decentralized real-time food safety supply chain traceability system that integrated Hazard Analysis and Critical Control Points (HACCP), blockchain, and IoT in order to make information available to all participants and players along the food supply chain, effectively minimizing the hazards associated with centralized information systems and aiding in strengthening food safety protocols and restoring consumer trust in the food sector.

Sadilek et al. [86] investigated an ML-powered system for the real-time detection of foodborne illness in a real-world setting. They created a Foodborne Illness Detector in Real-time (FINDER), an ML algorithm that uses location and anonymized aggregated web search data to predict foodborne disease in real-time. FINDER determined the percentage of patrons who went to a specific restaurant and then looked up terms related to food poisoning online. Using this data, eateries/restaurants that may be hazardous can be identified, enabling inspection agencies and health workers to focus their restaurant inspection efforts. According to the study, restaurants identified by FINDER are 3.1 times more likely to be found hazardous during inspections than restaurants identified by other existing methods. Tutul et al. [87] presented an advanced and intelligent system that enables continuous monitoring of food products in real-time. The system integrated multiple IoT devices (such as temperature, humidity, and gas concentration sensors) and ML technologies to constantly monitor food products' freshness in real-time. The system also includes a web-based dashboard that presents real-time data and generates alarms when food products exceed predetermined threshold levels. Users have the ability to remotely access the dashboard via a mobile application, allowing for convenient monitoring from any location. The system is designed to be cost-effective and intuitive, making it accessible for both commercial and personal use. Evaluation of the performance of the system showed it was effective in predicting the freshness of food with high accuracy.

4.5. AI and ML in Food Pathogen Detection, Classification, and Virulence/Resistance Prediction

The utilization of AI and ML methodologies in foodborne pathogen source fingerprinting, prediction of antibiotic resistance, and rapid detection and assessment of foodborne outbreaks and associated risks has been facilitated by the emergence of pathogen genomes and novel data streams, including textual data [14]. Wang et al. [88] developed a classification system based on ML for predicting microorganisms that cause foodborne illnesses. The study attempted to examine connections between several factors (such as area, time, and exposed food) and foodborne disease pathogens using case data from the National Foodborne Disease Surveillance Reporting System. Multiple ML models were used to categorize these infections and the best model (i.e., the gradient boost decision tree) predicted *Salmonella*, *Norovirus*, *E. coli*, and *Vibrio parahaemolyticus* with a 69% accuracy rate.

Yan et al. [89] used Raman spectroscopy in conjunction with ML to rapidly identify food-borne pathogens at the single-cell level. They collected 15,890 single-cell Raman spectra from 23 common strains from 7 different genera. Individual bacterial cells were analyzed and distinguished at the serotype level using a decision tree technique. The adopted methodology demonstrated an average correct prediction rate of 86.23% on an independent test set. The approach showed great promise for the quick diagnosis of pathogenic bacteria, demonstrating the ability to swiftly identify food-borne pathogens at the single-cell level in contaminated food. In a different study, Pesesky et al. [90] compared ML and rule-based techniques for predicting antimicrobial resistance profiles in Gram-negative *bacilli*, specifically *Enterobacteriaceae* isolates, including *E. coli*. Three curated antibiotic resistance sequencing databases were compared to determine resistance genes of 78 previously known whole genome-sequenced (*Enterobacteriaceae*) isolates. The rule-based system used hard-coded resistance profiles based on curated information for identified genes, whereas the ML approach treated all resistance genes equally and predicted based on algorithmically deduced patterns in the data. Both classifiers performed similarly, agreeing with phenotypic antibiotic susceptibility testing (AST) for over 90% of the isolates tested.

Teyhouee et al. [91] studied the prospective detection of foodborne illness outbreaks using ML algorithms. The research looked at the performance of Hidden Markov Models (HMMs) for syndromic surveillance monitoring and disease outbreak detection under two different data collection regimes. A sentinel population was used in one regime, with smartphone-based software tracking the location of food consumption and subclinical reporting. To compare the results with the HMM, the researchers used an SVM technique. According to the findings, depending purely on clinical data has a low potential for automatic epidemic detection. However, when guided by subclinical reporting, the use of HMMs indicated remarkable potential for detecting foodborne illness outbreaks, even from a relatively small sentinel group representing 4% of the population. Vangay et al. [92] investigated the persistence of *Listeria monocytogenes* in retail delicatessen environments using expert elicitation and ML. Using retail delicatessen operations as a template, the study extracted parameters utilized by food safety specialists in estimating bacterial persistence in the environment. Likewise, with the Delphi approach, the researchers undertook an expert elicitation with 10 food safety experts from academia, business, and government to classify *L. monocytogenes* persistence based on environmental sampling results from 30 retail delicatessen stores gathered over six months. Variations in random forest, SVM, logistic regression, and linear regression were used to model the results. With average validation errors of 3.1% and 2.2% (n = 144), the SVM and random forest models properly classified the data.

The burden of atypical mutations in protein-coding genes across independently evolved *Salmonella enterica* lineages was measured and the data was used to train an ML classifier (i.e., random forest) in order to identify strains associated with extraintestinal disease [93]. The classifier was able to accurately distinguish between invasive and long-established gastrointestinal *Salmonella serovars*. Furthermore, in immunocompromised individuals in sub-Saharan Africa, the model demonstrated the ability to distinguish between the recently identified *Salmonella enteritidis* and *Typhimurium* lineages linked to invasive illness, as well as within-host adaptation to invasive infection. In their study, Yi et al. [94] presented an AI-biosensing framework intended for automated and accelerated pathogen detection in a variety of water samples, including liquid food and agricultural water. Target bacteria were identified and quantified by the framework using a deep learning model that was trained on the bacteria's microscopic patterns generated by unique interactions with bacteriophages. The model was then applied to real-world water samples with environmental noises that were not observed during training. The AI model demonstrated its capacity to generalize to unseen data by making quick predictions in less than 5.5 h on real-world water samples, even though it was only trained on lab-cultured bacteria. The model's accuracy ranged from 80% to 100% and it can be effectively utilized to evaluate the microbiological quality of water during culinary and agricultural processes. Elsewhere, a pan-genome-based ML technique was used in the Her and Wu [95] study to forecast the antimicrobial resistance behaviors of *E. coli* strains. Findings from the study revealed that the group of AMR genes located in the accessory region of the pan-genome had the highest prediction accuracy, indicating the significant contribution of these gene clusters to AMR activities in *E. coli*, a well-known persistent food pathogen. Additionally, the researchers were able to select subsets of AMR genes for various antibiotic medications using a genetic algorithm (GA). Compared to the gene sets reported in the literature, the subsets chosen by the GA performed better in predictions.

4.6. AI and ML for Chemical and Biochemical Food Hazard Prediction and Analysis

AI and ML systems have the capability to thoroughly examine chemical and biochemical data pertaining to the composition of food and the presence of contaminants and other unwanted substances in order to forecast future risks. In the study by Chakraborty et al. [96], various ML models were used to classify and predict aflatoxin B1 (AFB1) levels in maize kernels. The hyperspectral images of 240 maize kernels that were exposed to six varying levels (25, 40, 70, 200, 300, and 500 ppb) of AFB1 were preprocessed and used to generalize a number of ML models, including partial least square discriminant analysis (PLS-DA) and KNN. The choice of the wavelength (508 nm) was determined by analyzing the loadings of PCA in order to differentiate between noncontaminated and contaminated maize kernels. A classification accuracy of 94.7% was attained by employing PLS-DA with standard normal variate (SNV) pre-processed data. Moreover, the KNN model using the raw data showed the highest efficiency, with an accuracy of 98.2%.

In their study on the detection of pesticide residue levels in crops, Ye et al. [97] employed visible/near-infrared (Vis-NIR) (376–1044 nm) and NIR (915–1699 nm) hyperspectral imaging systems (HSIs) together with ML to detect the levels of pesticide residues in grapes. Three grape cultivars were exposed to four degrees of pesticide application. Logistic regression, SVM, random forest, convolutional neural network (CNN), and residual neural network (ResNet) were utilized to build classification models for the pesticide residue levels in the crop. ResNet proved to be the most effective model for Vis-NIR spectra, achieving an accuracy rate above 93%. With regards to the NIR spectra data, logistic regression had the highest performance, surpassing 97% accuracy, while SVM, CNN, and ResNet also demonstrated comparable and satisfactory prediction outcomes. Baghel et al. [98] employed ML to optimize pesticide spray on crops. The algorithm pinpoints the areas on the crop that do not necessitate additional pesticide spraying, enabling farmers to exclusively administer pesticides to the required regions. This system reduced pesticide repetition by at least 20% compared to the standard method (the exit concept). Inferences from the study indicated that implementation of the concept has the potential to decrease the usage of pesticides in India by 72.5%. In a different study, Shen et al. [99] adopted ML models to predict pesticide dissipation half-life intervals. Four machine learning models, namely gradient boosting regression tree (GBRT), random forest, support vector classifier (SVC), and logistic regression, were developed to predict pesticide dissipation half-life intervals using temperature, plant component class, extended connectivity fingerprints (ECFP), and plant type as model inputs. The GBRT-ECFP model exhibited superior performance compared to the other models. In their study, Bhatia and Albarrak [100] presented a novel approach using Explainable Artificial Intelligence (XAI) and Faster RCNN to digitize food product information, retrieve it, evaluate the composition of the foods, and identify the hazardous ingredients that may compromise food integrity.

Linear and non-linear ML models were employed by Petrea et al. [101] to ascertain the levels of heavy metals present in turbot muscle and liver tissues. The models were constructed using data that were obtained from the scientific literature and included eleven heavy metals (As, Ca, Cd, Cu, Fe, K, Mg, Mn, Na, Ni, and Zn) that were found in the liver and muscle tissues of specimens of turbot. Over 70% prediction accuracy was attained by the non-linear tree-based random forest prediction models for As, Cd, Cu, K, Mg, and Zn in muscle tissue and As, Ca, Cd, Mg, and Fe in turbot liver tissue. The adopted multiple linear regression (MLR) and random forest models were found to be appropriate for predicting the levels of heavy metals in the liver and muscle of turbot. Using a different approach, Yu et al. [102] estimated heavy metal concentrations in winter wheat leaves from a typical sewage irrigation area based on canopy reflectance spectra and ML modeling. Spectral data from winter wheat canopies were collected at 61 sampling stations in Longkou City, Shandong province, China, and the Pb, Zn, Cd, Cr, Ni, and Hg contents were measured. Back-propagation neural network (BPNN), partial least squares regression, and stepwise multiple linear regression methods were used to build eight estimation models, which were trained with reflectance spectra, a first-order derivative of spectral reflectance (FDR), a second-order derivative of spectral reflectance (FDR), and spectral parameters (SPs). The study revealed that the BPNN model with SPs performed best for Pb, Zn, and Cd, while the BPNN model with FDR performed best for Cr, Ni, and Hg.

4.7. AI and ML in the Formulation of New Recipe and Personalized Nutrition for Improved Quality and Safety

Beyond the prediction of food safety hazards, AI and ML are driving product development innovation, for improved food quality and nutritional health. AI-based algorithms benefit from the fundamental understanding of the intricacies and interrelationships of the various food components and ingredients and vast datasets containing consumer preferences, market trends, and nutritional profiles and are able to suggest varied ingredient combinations and design tailored food products for consumer-specific needs and tastes. AI-driven product development improves not only customer satisfaction and food safety/quality but also market competitiveness by knowing consumer preferences and market dynamics and responding to changing trends. For example, Park et al. [103] assembled various culinary recipes and food chemical information to create FlavorGraph, a large-scale food-compound network graph. The food-specific metapath graph embedding method was combined with a chemical structure learning layer to create an intricate representation of food vectors for FlavorGraph. The food representation vectors were then used to make meal pairing recommendations and predict new food-compound relationships. FlavorGraph, according to the authors, can be used to better comprehend the cooking and medicinal uses of food. Furthermore, the deep learning algorithms described in the study can be used as the foundation for meal pairing and food-relationship prediction problems.

Iwendi and colleagues [104] proposed a deep learning solution for health-based medical datasets, which autonomously determines the appropriate food and makes recommendations for patients based on their specific diseases and other features such as age, gender, weight, calories, protein, fat, sodium, fiber, and cholesterol. Their study implemented several ML and deep learning algorithms such as logistic regression, naive Bayes, recurrent neural network (RNN), MLP, gated recurrent units (GRU), and long short-term memory (LSTM) in accomplishing the research goal. The medical dataset, which was gathered from the internet and hospitals, included information from 30 patients with 13 attributes related to various diseases and 1000 products, each with an 8-feature set. The performance of the different ML and deep learning models was evaluated and it was observed that the LSTM model surpassed other methods in terms of prediction accuracy (97.74%), recall, precision, and F1 measures. Vairale and Shukla [105] developed a diet recommendation system by employing hybrid collaborative filtering learning methods. The nutrition support system was created by utilizing KNN and collaborative filtering models to recommend well-balanced diets for thyroid patients that fulfill their specific nutrient needs. The system manages the patient's dietary intake by providing the necessary nutrients required for those with thyroid issues.

Sowah and colleagues [106] created a diabetes management system that uses several ML and AI algorithms to improve diabetes control based on patient calorie intake through food consumption. The system essentially implements a meal suggestion system with food identification capabilities, aimed at providing users with daily individualized meal plans based on their nutritional needs and previous meal choices. The system allows users to input food images to decide whether a meal is safe to consume or not. For specific calorie intakes, the food recognition and categorization model reached an accuracy level of more than 95%. In recognizing personalized eating habits and the value of individualized diets in maintaining food consumer health, Naik [107] proposed an intelligent food recommendation system using deep learning wherein food consumers are recommended products based on the experiences of other customers who have used the same product. A web crawler with a review collection technique is used in the implementation to collect reviews about food products and save them in the application. User-specific nutrition questions contribute to the creation of a health profile for each user. In the recommendation process, the genetic algorithm builds the association between the product and the user and makes appropriate food recommendations.

4.8. AI and ML in the Detection of Food Fraud

Food fraud refers to any deliberate misrepresentation or deception related to food products, which includes the substitution, adulteration, or mislabeling of ingredients in order to achieve economic gain. AI and ML models can scrutinize ingredient information, supply chain records, and consumer feedback, to identify patterns and anomalies and detect risky features that indicate potential fraud in the food industry. Not only can AI and ML technology aid in the detection of food fraud but it can also contribute to prevention efforts. For instance, by analyzing historical data and detecting fraudulent patterns, ML algorithms can help food organizations identify vulnerable points in their supply chains and implement measures to prevent fraudulent activities.

Buyuktepe et al. [108] developed a Deep Neural Network (DNN) model and employed XAI technologies to predict food fraud categories, analyze the outcomes, and explain the forecasts generated by the AI model. Bouzembrak and Marvin [109] employed Bayesian network (BN) modeling to predict several types of food fraud using notification data from the Rapid Alert System for Food and Feed (RASFF) spanning from 2000 to 2013. The study sought to forecast anticipated forms of food fraud for imported items based on their known product categories and countries of origin, thereby enabling focused enforcement efforts. The model was trained on 749 RASFF fraud warning data classified into six distinct types and further validated with an additional 88 RASFF fraud notifications from the year 2014. The model demonstrated an accuracy of 80% when the fraud type, country, and food category constituted pre-existing knowledge and 52% accuracy when this information was not previously documented in the RASFF database. Such a system could be utilized by risk managers and controllers at border inspection checkpoints to effectively prioritize fraud-type checks while dealing with goods imports. Elsewhere, Marvin et al. [77] constructed a BN model to scrutinize all relevant driving factors influencing food fraud cases, utilizing data from RASFF and the European Medicines Agency (EMA). Their BN model attained a 91.5% accuracy in forecasting food fraud categories, providing significant information for risk managers.

In their study, Mithun et al. [110] utilized deep learning techniques to discriminate between bananas that have ripened naturally and those that have been artificially ripened. According to the authors, a risk exists that bananas that have been artificially ripened may have been treated with compounds that can cause cancer, such as calcium carbide. The deep learning model implemented in the study achieved a classification accuracy of 90%, while accuracies of up to 98.74% and 89.49% were also achievable using the random forest and MLP feed-forward neural network classifiers, respectively. Elsewhere, the work undertaken by Pulluri and Kumar [111] involved the development of a smart electronic nose (SE-Nose) designed to rapidly detect and measure food adulteration, particularly in recognizing the presence of pork in beef. The methodology employed classification models to do qualitative analysis of adulteration and regression models to conduct quantitative analysis. SVM classification and regression models yielded exceptional outcomes, with an accuracy rate of up to 99.996%. Moreover, the SVM models significantly lowered the detection time for identifying pork adulteration in beef to 40 s, which is a 33%-time decrease compared to the other approaches.

de Santana et al. [112] combined infrared spectroscopy and ML to identify instances of food adulteration. Their proposed system was utilized in two investigations on food adulteration: one using evening primrose oils analyzed with ATR-FTIR spectroscopy and the other including ground nutmeg analyzed with NIR diffuse reflectance spectroscopy. The proposed methodology, which employs the random forest algorithm with artificial outlier generation as a one-class classifier, showed improved performance in comparison to the PLS-DA and SIMCA methods. Specificity values of 0.9988 and 0.9286 were observed for evening primrose oil and ground nutmeg, respectively. Lim et al. [113] employed deep learning techniques to differentiate between 10 different types of plant oils by examining the fatty acid profiles in the oils. Their ML approach consisted of supervised end-to-end learning methodologies to ascertain the oil composition of various oil mixes. The model was trained using a substantial dataset of simulated oil mixes and showed effectiveness during validation with an independent test dataset.

Mu et al. [114] presented a partial least square model for predicting oil adulteration with errors below 2%. The researchers further employed ANN and SVM algorithms to categorize pure and mixed oils, specifically focusing on differentiating between olive, rapeseed, peanut, and blend oils. The study determined that by utilizing their approach, a classification accuracy of 100% was attainable. Laga and Sarno [115] utilized different ML algorithms to distinguish between pure beef and mixed beef. This distinction was made by analyzing various beef parameters, including temperature, strain, and humidity. By employing an electronic nose equipped with electrochemical and air sensors (capable of detecting gaseous emissions), data were collected and used to generalize the ML models to determine whether the beef is blended or unadulterated. The model performance, ranked from highest to lowest accuracy at room temperature, was as follows: KNN with an average ROC value of 0.886, Naïve Bayes with an average ROC value of 0.856, Random Forest with an average ROC value of 0.839, and SVM with an average ROC value of 0.821. Table 2 summarizes some of the studies reviewed herein on the applications of AI and ML in the domain of food safety and integrity.

Table 2. Applications of AI and ML in food integrity.

S/No.	Intelligent Approach	Model(s)	Application	Objective	Food Commodity	Reference
1.	AI		Food quality control and regulatory compliance	Food inspection system	Raspberries, almonds, strawberries, cashew, carrots, and blueberries	[49]
2.	ML	SVM	Food quality control and regulatory compliance	Discriminate defective grains	Rice	[51]
3.	ML	Fuzzy system	Food quality control and regulatory compliance	Categorizing and grading	Mozafati dates	[52]
4.	ML	Radial basis function (RBF) kernel-SVM classifier	Food quality control and regulatory compliance	Identify and grade food products	Tomatoes	[53]
5.	ML	Partial least squares (PLS) regression	Food quality control and regulatory compliance	Predict nutritional/chemical composition	Mutton	[55]
6.	ML	KNN	Food quality control and regulatory compliance	Detection and classification of mechanically-damaged fruits	'Manila' mangos	[56]
7.	ML	PCA	Food quality control and regulatory compliance	Detection of fecal contamination	Apple skins	[57]
8.	AI	Mask-RCNN	Food quality control and regulatory compliance	Fruit detection, identification, and dimensions estimation	Tomatoes	[59]
9.	ML	classification and regression trees (CART) and LDA	Food quality control and regulatory compliance	Early detection of rottenness caused by *Penicillium digitatum*	Mandarins	[58]
10	AI and ML		Food supply chain traceability	Swift source tracking of agricultural commodities	Various food products	[63]
11.	ML	Bayesian regression and random forest	Food supply chain traceability	Food supply chain tracing	Perishable foods	[65]
12.	ML	SVM	Food supply chain traceability	Trace geographical origin of food	Refined sugar	[66]
13.	ML	OPLS-DA	Food supply chain traceability	Trace geographical origin of food	Saanen goat milk	[67]
14.	ML	PCA and PLS-DA	Food supply chain traceability	Trace geographical origin of food	*Crassostrea gigas* oyster species	[68]

Table 2. *Cont.*

S/No.	Intelligent Approach	Model(s)	Application	Objective	Food Commodity	Reference
15.	ML	Bayes classifier algorithm	Food supply chain traceability	Detect anomalies such as contaminated food items along the supply chain that may need to be recalled	Various food commodities	[69]
16.	AI and ML	Fuzzy classification and ANN models	Food supply chain traceability	Both forward tracking and varied tracing along the supply chain with real-time evaluations of food quality along the entire supply chain	Various food commodities. Case study conducted on a pork producer	[70]
17.	AI	ANN	Food supply chain traceability	Trace geographical origin of food	Italian Extra Virgin Olive Oil (EVOO)	[71]
18.	ML	Numerous ensemble learning models including Bagging-Logistic, Bagging-CART, Bagging-C5.0, Bagging-NB (Bayesian classification), and Bagging-RF.	Predictive analytics for risk assessment	Forecast food safety hazards, with a specific emphasis on enhancing border inspection techniques for imported food	Various food commodities	[73]
19.	ML	SVM	Predictive analytics for risk assessment	Food safety risk intelligent early warning system	Various food commodities	[74]
20.	ML	Bayesian network	Predictive analytics for risk assessment	Food safety early warning system with the goal of detecting indicators that precede the emergence of food safety hazards from adjacent domains of the food supply chain	Dairy supply chain	[75]
21.	ML	Latent Dirichlet Allocation topic model	Predictive analytics for risk assessment	Promptly identify instances of food fraud reported in the media	Beeswax	[76]
22.	ML	Bayesian network	Predictive analytics for risk assessment	Estimate the likelihood of food fraud events	Various food commodities including fish and seafood", meat, and fruits and vegetables	[77]
23.	ML	Bayesian network	Predictive analytics for risk assessment	Forecast significant food safety risks and their corresponding food products	Herbs and spices	[78]
24.	ML	Kernel-based extreme learning machine (K-ELM)	Predictive analytics for risk assessment	Predict food safety risks	Dairy products	[79]

Table 2. *Cont.*

S/No.	Intelligent Approach	Model(s)	Application	Objective	Food Commodity	Reference
25.	AI and ML	Logistic regression, decision trees, random forest, boosted tree, SVM, Support Vector Regression (SVR), and multilayer perceptron (MLP)	Predictive analytics for risk assessment	Predict food safety risks	Various food commodities	[80]
26.	ML	XGBoost	Real-time monitoring of food in the supply chain	Monitor the temperature and humidity of perishable foods. Also establish the direction of movement of the products along the supply chain	Perishable foods	[81]
27.	ML		Real-time monitoring of food in the supply chain	Real-time monitoring and tracing the food supply chain, specifically targeting food safety issues	Various food commodities	[82]
28.	AI and ML	A combination of advanced deep learning (ADL) and IoT-blockchain technologies	Real-time monitoring of food in the supply chain	Optimize provenance of the food supply chain	Various food commodities	[83]
29.	AI	Analytic hierarchy process integrated extreme learning machine (AHP-ELM)	Real-time monitoring of food in the supply chain	Real-time food monitoring and early warning	Various food commodities	[84]
30.		Blockchain and IoT	Real-time monitoring of food in the supply chain	Real-time food safety supply chain traceability	Various food commodities	[85]
31.	ML		Real-time monitoring of food in the supply chain	Real-time detection of foodborne illness	Various food commodities	[86]
32.	ML	PCA and t-distributed Stochastic Neighbor Embedding (t-SNE)	Real-time monitoring of food in the supply chain	Continuous monitoring of food products in real-time	Various food commodities	[87]

Table 2. *Cont.*

S/No.	Intelligent Approach	Model(s)	Application	Objective	Food Commodity	Reference
33.	ML	Decision tree, random forest, gradient boost decision tree (GBDT), and adaptive boosting models	Food pathogen detection, classification, and virulence/resistance prediction	Prediction of microorganisms that cause foodborne illnesses		[88]
34.	ML	PCA, Kernel principal component analysis (KPCA), decision tree (DT), CART, and PCA-SVM classifier	Food pathogen detection, classification, and virulence/resistance prediction	Rapidly identify food-borne pathogens at the single-cell level		[89]
35.	ML	Rules-based (RB) algorithm and logistic regression algorithm	Food pathogen detection, classification, and virulence/resistance prediction	Prediction of antimicrobial resistance profiles		[90]
36.	ML	Hidden Markov Models (HMMs) and SVM	Food pathogen detection, classification, and virulence/resistance prediction	Syndromic surveillance monitoring and disease outbreak detection	Various food commodities	[91]
37.	ML	Variations of random forest, SVM, logistic regression, and linear regression	Food pathogen detection, classification, and virulence/resistance prediction	Classification and estimation of pathogenic bacterial persistence	Various food commodities	[92]
38.	ML	Random forest	Food pathogen detection, classification, and virulence/resistance prediction	Distinguish between invasive and long-established gastrointestinal *Salmonella serovars*		[93]
39.	AI	R-CNN deep learning model	Food pathogen detection, classification, and virulence/resistance prediction	Rapid pathogen detection	From liquid food to agricultural water	[94]
40.	ML	SVM, Naïve Bayes, Adaboost, and random forest	Food pathogen detection, classification, and virulence/resistance prediction	Forecast antimicrobial resistance behaviors of *E. coli* strains		[95]
41.	ML	Different ML models including PCA, PLS-DA, and KNN	Chemical and biochemical food hazards prediction and analysis	Classification and prediction of aflatoxin B1 (AFB1) levels	Maize kernels	[96]

Table 2. *Cont.*

S/No.	Intelligent Approach	Model(s)	Application	Objective	Food Commodity	Reference
42.	AI and ML	Logistic regression, SVM, random forest, convolutional neural network (CNN), and residual neural network (ResNet)	Chemical and biochemical food hazards prediction and analysis	Detection of pesticide levels	Grapes	[97]
43.	ML	Logistic regression classification, polynomial regression, and KNN	Chemical and biochemical food hazards prediction and analysis	Optimization of pesticides spray on crops	Various crops	[98]
44.	ML	Gradient boosting regression tree (GBRT), random forest, support vector classifier (SVC), and logistic regression	Chemical and biochemical food hazards prediction and analysis	Prediction of pesticide dissipation half-life intervals in plants		[99]
45.	AI	Deep reinforcement learning-based supply chain management (DR-SCM) and explainable artificial intelligence-based faster regions with convolutional neural networks (XAI-based Faster RCNN)	Chemical and biochemical food hazards prediction and analysis	Identification of hazardous chemical components in foods	Various food commodities	[100]
46.	ML	Stepwise multiple linear regression and random forest	Chemical and biochemical food hazards prediction and analysis	Estimation of the concentrations of heavy metals food commodities	Turbot muscle and liver tissues	[101]
47.	AI and ML	Back-propagation neural network (BPNN), PLS, and stepwise multiple linear regression	Chemical and biochemical food hazards prediction and analysis	Estimation of the concentrations of heavy metals food commodities	Wheat leaves	[102]

Table 2. *Cont.*

S/No.	Intelligent Approach	Model(s)	Application	Objective	Food Commodity	Reference
48.	AI and ML	Logistic regression, naive Bayes, recurrent neural network (RNN), MLP, gated recurrent units (GRU), and long short-term memory (LSTM)	Formulation of new recipe and personalized nutrition	Diet recommendation. Automatically determine the appropriate food and make recommendations for patients based on their specific diseases and other features such as age, gender, weight, calories, protein, fat, sodium, fiber, and cholesterol	Different food recipes	[104]
49.	ML	KNN with alternating least squares (KNN-ALS) and KNN with stochastic gradient decent (KNN-SGD)	Formulation of new recipe and personalized nutrition	Diet and exercise recommendation system for a balanced diet for thyroid patients	Various diets	[105]
50.	ML	KNN	Formulation of new recipe and personalized nutrition	Meal recommendation system for diabetic patients	Various diets	[106]
51.	AI	Word Embedding and deep learning	Formulation of new recipe and personalized nutrition	Intelligent food recommendation	Various diets	[107]
52.	AI	DNN and explainable artificial intelligence (XAI) techniques such as Local Interpretable Model-Agnostic Explanations (LIME), Shapley Additive exPlanations (SHAP), and What-If Tool (WIT)	Detection of food fraud	Predict food fraud categories and interpret the predictions of the AI model	Various food commodities	[108]
53.	ML	BN	Detection of food fraud	Prediction of several types of food fraud	Various food commodities	[109]
54.	ML	BN	Detection of food fraud	Prediction of food fraud categories	Various food commodities	[77]
55.	AI and ML	Deep learning, random forest, and MLP	Detection of food fraud	Discrimination between bananas that ripened naturally and those that were artificially ripened	Bananas	[110]

Table 2. *Cont.*

S/No.	Intelligent Approach	Model(s)	Application	Objective	Food Commodity	Reference
56.	AI and ML	SVM and MLP	Detection of food fraud	Detection of meat adulteration	Beef and pork	[111]
57.	ML	Random forest, PLS-DA, and SIMCA	Detection of food fraud	Detection of food adulteration	Evening primrose oils and ground nutmeg	[112]
58.	AI and ML	t-stochastic neighborhood embedding, PCA, PLS2, Gaussian mixture model	Detection of food fraud	Discrimination of different types of plant oils	Vegetable oils including groundnut oil, high-erucic acid rapeseed oil, high-oleic acid sunflower oil, low-erucic acid rapeseed oil, linseed oil, low-oleic acid sunflower oil, maize oil, rice bran oil, soybean oil, and sesame oil	[113]
59.	AI and ML	PLS, ANN, and SVM	Detection of food fraud	Prediction of oil adulteration	Olive oil, rapeseed oil, peanut oil, and blend oils	[114]
60.	ML	KNN, SVM, Naïve Bayes, and random forest	Detection of food fraud	Distinction between pure beef and mixed beef	Beef	[115]

5. Study Limitations

While this study aims to offer a thorough and extensive analysis of the present state of research in enhancing food integrity via AI and ML, it is critical to recognize various limitations. Firstly, it is important to note that the understanding of the results and the conclusions made in this review article depends on the quality and accuracy of the research included. The scope of this review is constrained by the existing literature and may not include every pertinent study or advancement in the subject. The authors acknowledge that the choice of search engines and databases, as well as the search terms/phrases and syntax employed to obtain data from the internet, may limit the accessibility of information. In this regard, the authors made deliberate efforts to significantly broaden the range of search phrases and utilized many databases in order to consolidate as much relevant information as possible.

Moreover, data quality and availability are crucial factors for the successful implementation of AI and ML systems, as amongst other things, the accuracy of model decisions is directly influenced by data quality. AI and ML algorithms can be influenced by biases present in the data used for training, potentially resulting in unintended effects or prejudiced outputs, particularly in crucial domains like food integrity. Researchers frequently encounter difficulties as a result of the limited availability of data, emphasizing the importance of standardizing and exchanging data across many fields. Our study may have overlooked the significant obstacles associated with data gathering, standardization, and privacy concerns, as well as details about the quality of data used to generalize the models described in the respective studies reviewed herein. These challenges have the potential to affect the dependability and scalability of AI/ML solutions in many food integrity scenarios. Finally, the field of AI and ML is vast and constantly evolving. Given the ever-changing nature of research in this domain, it is possible that new studies and advancements have been published since the completion of this review. Therefore, it is important to seek updated information in addition to that presented in this study to ensure its relevance.

6. Future Prospects and Directions

Going forward, the use of AI and ML in food integrity will continue to play a crucial role in transforming the food and beverage industry landscape. These technologies have demonstrated their efficacy in many other scientific fields and likewise provide unique benefits in the food and beverage industry by predicting and identifying food safety hazards, enhancing production processes, minimizing food waste, improving product monitoring, etc. Indeed, the incorporation of AI and ML technologies in food processes is rather essential in the present competitive and quality-focused food industry. Organizations that do not adopt these technologies early enough face the risk of lagging behind. Moreover, the continuous advancement of AI and ML, for example via artificial general intelligence (AGI) and its adaptive learning capabilities, will continue to benefit the creation of customized food safety solutions tailored to the specific needs of diverse food industry segments, fostering a more resilient and responsive approach to safety management and staying ahead of evolving food integrity threats.

Although AI and ML offer several prospects for the food and beverage industry, it is crucial to acknowledge and tackle some of the pertinent challenges to the adoption of these technologies within the domain. For example, it is vital to meticulously tackle data privacy and security concerns in order to guarantee the safeguarding of delicate information throughout the manufacturing process. Another challenge is that many food safety records are highly dispersed across various domains (such as food, health, and agriculture) and frequently not digitized [116]. Consolidation of these records in a digital and computer-friendly format, in addition to the integration of numerous data sources to produce a larger dataset, has the potential to improve modeling performance. By merging data from diverse sources, one can capture a greater range of characteristics and factors that influence food safety outcomes. This comprehensive approach enables a more comprehensive knowledge of the complex relationships and dynamics within the food safety domain. For example, mycotoxins, which are harmful substances produced by certain molds, are affected by environmental factors such as temperature and humidity. By including climatic data in the modeling process, one can acquire insight into the environmental conditions that influence mycotoxin levels in food products. This bigger dataset not only enriches the information available for modeling but it also allows for the identification of detailed linkages and patterns that may not be obvious when using a restricted scope of data. It enables a more nuanced and accurate portrayal of the many elements influencing food safety, ultimately leading to more robust and effective models. However, it is critical to approach the integration of disparate data sources with caution, taking into account data quality, relevance, and potential biases.

Continued innovation and strategic collaborative efforts among food companies, governments, regulatory agencies, academic specialists, consumers, and various other stakeholders along the food supply chain are critical to drive AI advancement and integration in the domain. As the food industry continues to adopt these new technologies, it is critical to prioritize ethical considerations, regulatory compliance, and security measures to guarantee responsible and secure implementations. Lastly, it is important to acknowledge that as the integrated domain of AI and food integrity evolves, there will be significant disruptions and changes in the way food safety specialists and industry players manage risks and make decisions, resulting in the emergence of potentially new operational risks. Therefore, it is imperative to foresee these risks and create plans for contingency mitigation.

7. Conclusions

The food industry, which is characterized by perpetual evolving consumer trends, competitiveness, emerging food safety risks, and the obligation for industry players to adhere to stringent safety standards, faces the difficult issue of ensuring food integrity and ensuring quality in manufacturing processes. In this review paper, we explored the general research trends and transformative potential of AI and ML in food integrity, ranging from the detection of food fraud, quality control, supply chain traceability and transparency, analytics for risk assessment and hazard prediction, real-time monitoring, food pathogen detection and prediction, and personalized nutrition for improved quality and safety. From the literature examined herein, it is deduced that through the utilization of AI and ML and related technologies such as big data, predictive analytics, IoT, blockchain, real-time sensor technologies, virtual reality, etc., the food and beverage industry can actively reduce risks, improve product safety, and strengthen consumer trust. The paper further underlines the need for data quality, privacy, and data protection, as well as careful attention to ethical considerations for realizing the full benefits of AI in the industry. The paper concludes with a call to action for continuing research, innovation, and collaboration, highlighting the immense promise of AI and ML in conjunction with related technologies in building a secure and resilient global food supply chain.

Author Contributions: S.G. conceived and designed the review, conducted the literature review, and drafted the manuscript. P.B.N. provided expertise, supervised the study, and performed a critical review of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by the South African National Research Foundation (NRF) via the Innovation Postdoctoral Fellowship, grant number 138455 and the Maize Trust, grant number MTM23-01.

Data Availability Statement: Not applicable.

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

References

1. Cao, Z.; Yemets, M.; Muneem, S.; Shannon, K.; Gowher, F.; Malloy, J.; Kafka, E.; Mitchell, C.; King, J.; Urban, S. Food Fraud: Insights from Investigating a near-Fatal Poisoning with Global Implications. *medRxiv* **2023**, 2023–2029. [CrossRef]
2. Kemsawasd, V.; Jayasena, V.; Karnpanit, W. Incidents and Potential Adverse Health Effects of Serious Food Fraud Cases Originated in Asia. *Foods* **2023**, *12*, 3522. [CrossRef]
3. Wang, T.; Zhang, M.; Sun, N.; Chen, H.; Zhang, C.; Wang, Q.; Zhang, W. Analysis of the Economically Motivated Food Adulteration in China Based on 6477 Events from 2000 to 2020. *Qual. Assur. Saf. Crops Foods* **2023**, *15*, 222–231. [CrossRef]
4. Lawrence, S.; Elliott, C.; Huisman, W.; Dean, M.; van Ruth, S. The 11 Sins of Seafood: Assessing a Decade of Food Fraud Reports in the Global Supply Chain. *Compr. Rev. Food Sci. Food Saf.* **2022**, *21*, 3746–3769. [CrossRef]
5. Gwenzi, W.; Makuvara, Z.; Marumure, J.; Simbanegavi, T.T.; Mukonza, S.S.; Chaukura, N. Chicanery in the Food Supply Chain! Food Fraud, Mitigation, and Research Needs in Low-Income Countries. *Trends Food Sci. Technol.* **2023**, *136*, 194–223. [CrossRef]
6. WHO. Estimating the Burden of Foodborne Diseases. Available online: https://www.who.int/activities/estimating-the-burden-of-foodborne-diseases#:~:text=Eachyearworldwide,unsafefood,numberislikelyanunderestimation (accessed on 21 October 2022).
7. Lee, H.; Yoon, Y. Etiological Agents Implicated in Foodborne Illness World Wide. *Food Sci. Anim. Resour.* **2021**, *41*, 1. [CrossRef]
8. WHO. *WHO Estimates of the Global Burden of Foodborne Diseases*; World Health Organization: Geneva, Switzerland, 2015.
9. WHO. *The Burden of Foodborne Diseases in the WHO European Region*; World Health Organization: Geneva, Switzerland, 2017.
10. European Commission. Commission Regulation (EU) 2022/1370 of 5 August 2022 Amending Regulation (EC) No 1881/2006 as Regards Maximum Levels of Ochratoxin A in Certain Foodstuffs. *Off. J. Eur. Union* **2022**, *L 206*, 11–14.
11. Phytocontrol New Limits for Ochratoxin A. Available online: https://www.phytocontrol.com/en/regulatory-watch/new-limits-for-ochratoxin-a/ (accessed on 12 February 2024).
12. SGS EU Updates Maximum Levels for Ochratoxin A in Certain Foodstuffs. Available online: https://www.digicomply.com/blog/eu-updates-maximum-levels-for-ochratoxin-a-in-certain-foodstuffs (accessed on 12 February 2024).
13. Sharpless, N.E. *Statement from Acting FDA Commissioner Ned Sharpless, MD, and Deputy Commissioner Frank Yiannas on Steps to Usher the US into a New Era of Smarter Food Safety*; US Food and Drug Administration (FDA): Silver Spring, MD, USA, 2019.
14. Deng, X.; Cao, S.; Horn, A.L. Emerging Applications of Machine Learning in Food Safety. *Annu. Rev. Food Sci. Technol.* **2021**, *12*, 513–538. [CrossRef]
15. Gbashi, S.; Adebo, O.A.; Doorsamy, W.; Njobeh, P.B. Systematic Delineation of Media Polarity on COVID-19 Vaccines in Africa: Computational Linguistic Modeling Study. *JMIR Med. Inf.* **2021**, *9*, e22916. [CrossRef]
16. Vinci, G.; Ruggieri, R.; Ruggeri, M.; Prencipe, S.A. Rice Production Chain: Environmental and Social Impact Assessment—A Review. *Agriculture* **2023**, *13*, 340. [CrossRef]
17. Ehret, D.L.; Hill, B.D.; Helmer, T.; Edwards, D.R. Neural Network Modeling of Greenhouse Tomato Yield, Growth and Water Use from Automated Crop Monitoring Data. *Comput. Electron. Agric.* **2011**, *79*, 82–89. [CrossRef]

18. Resnick, S. Artificial Intelligence in Eye Care? Naturally! *Clin. Refract. Optom. J.* **2023**, *79*, 82–89. [CrossRef]
19. Cioffi, R.; Travaglioni, M.; Piscitelli, G.; Petrillo, A.; De Felice, F. Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions. *Sustainability* **2020**, *12*, 492. [CrossRef]
20. Burati, M.; Tagliabue, F.; Lomonaco, A.; Chiarelli, M.; Zago, M.; Cioffi, G.; Cioffi, U. Artificial Intelligence as a Future in Cancer Surgery. *Artif. Intell. Cancer* **2022**, *3*, 11–16. [CrossRef]
21. Cannataro, M.; Guzzi, P.H.; Agapito, G.; Zucco, C.; Milano, M. Chapter 3—Artificial Intelligence. In *Artificial Intelligence in Bioinformatics*; Cannataro, M., Guzzi, P.H., Agapito, G., Zucco, C., Milano, M., Eds.; Elsevier: Amsterdam, The Netherlands, 2022; pp. 29–33, ISBN 978-0-12-822952-1.
22. Helm, J.M.; Swiergosz, A.M.; Haeberle, H.S.; Karnuta, J.M.; Schaffer, J.L.; Krebs, V.E.; Spitzer, A.I.; Ramkumar, P.N. Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions. *Curr. Rev. Musculoskelet. Med.* **2020**, *13*, 69–76. [CrossRef]
23. Galanos, V. Expectations and Expertise in Artificial Intelligence: Specialist Views and Historical Perspectives on Conceptualisation, Promise, and Funding. Ph.D. Thesis, University of Edinburgh, Edinburgh, UK, 2023.
24. Le, Q.V. Building High-Level Features Using Large Scale Unsupervised Learning. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 8595–8598.
25. Delipetrev, B.; Tsinaraki, C.; Kostic, U. Historical Evolution of Artificial Intelligence. JRC Publications Repository. Available online: https://publications.jrc.ec.europa.eu/repository/handle/JRC120469 (accessed on 6 January 2024).
26. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
27. Flowers, J.C. Strong and Weak AI: Deweyan Considerations. In Proceedings of the AAAI Spring Symposium: Towards Conscious AI Systems, Palo Alto, CA, USA, 25–27 March 2019; Volume 2287.
28. Fjelland, R. Why General Artificial Intelligence Will Not Be Realized. *Humanit. Soc. Sci. Commun.* **2020**, *7*, 1–9. [CrossRef]
29. Asay, C.D. Artificial Stupidity. *Wm. Mary L. Rev.* **2019**, *61*, 1187.
30. Kuusi, O.; Heinonen, S. Scenarios from Artificial Narrow Intelligence to Artificial General Intelligence—Reviewing the Results of the International Work/Technology 2050 Study. *World Futures Rev.* **2022**, *14*, 65–79. [CrossRef]
31. Varshney, H.; Khan, R.A.; Khan, U.; Verma, R. Approaches of Artificial Intelligence and Machine Learning in Smart Cities: Critical Review. In *Proceedings of the IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2021; Volume 1022, p. 012019.
32. Ullah, Z.; Al-Turjman, F.; Mostarda, L.; Gagliardi, R. Applications of Artificial Intelligence and Machine Learning in Smart Cities. *Comput. Commun.* **2020**, *154*, 313–323. [CrossRef]
33. Saravanan, R.; Sujatha, P. A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification. In Proceedings of the 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 14–15 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 945–949.
34. Dridi, S. Supervised Learning-a Systematic Literature Review. *PERINTIS eJournal* **2020**, *10*, 1–24.
35. Muhammad, I.; Yan, Z. Supervised machine learning approaches: A survey. *ICTACT J. Soft Comput.* **2015**, *5*, 946–952. [CrossRef]
36. Shetty, S.H.; Shetty, S.; Singh, C.; Rao, A. Supervised Machine Learning: Algorithms and Applications. In *Fundamentals and Methods of Machine and Deep Learning: Algorithms, Tools and Applications*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2022; pp. 1–16.
37. Verma, R.; Nagar, V.; Mahapatra, S. Introduction to Supervised Learning. In *Data Analytics in Bioinformatics*; Wiley: Hoboken, NJ, USA, 2021; pp. 1–34.
38. Sindhu Meena, K.; Suriya, S. A Survey on Supervised and Unsupervised Learning Techniques. In Proceedings of the International Conference on Artificial Intelligence, Smart Grid and Smart City Applications: AISGSC, Coimbatore, India, 3–5 January 2019; Springer: Berlin/Heidelberg, Germany, 2020; pp. 627–644.
39. Usama, M.; Qadir, J.; Raza, A.; Arif, H.; Yau, K.-L.A.; Elkhatib, Y.; Hussain, A.; Al-Fuqaha, A. Unsupervised Machine Learning for Networking: Techniques, Applications and Research Challenges. *IEEE Access* **2019**, *7*, 65579–65615. [CrossRef]
40. Rajoub, B. Supervised and Unsupervised Learning. In *Biomedical Signal Processing and Artificial Intelligence in Healthcare*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 51–89.
41. Chauhan, T.; Rawat, S.; Malik, S.; Singh, P. Supervised and Unsupervised Machine Learning Based Review on Diabetes Care. In Proceedings of the 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 19–20 March 2021; IEEE: Piscataway, NJ, USA, 2021; Volume 1, pp. 581–585.
42. Balamurali, M. T-Distributed Stochastic Neighbor Embedding. In *Encyclopedia of Mathematical Geosciences*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 1–9.
43. Naeem, S.; Ali, A.; Anam, S.; Ahmed, M.M. An Unsupervised Machine Learning Algorithms: Comprehensive Review. *Int. J. Comput. Digit. Syst.* **2023**, *13*, 911–921. [CrossRef]
44. Dong, H.; Dong, H.; Ding, Z.; Zhang, S. *Chang Deep Reinforcement Learning*; Springer: Berlin/Heidelberg, Germany, 2020; ISBN 9811540942.
45. Wiering, M.A.; Van Otterlo, M. Reinforcement Learning. *Adapt. Learn. Optim.* **2012**, *12*, 729.

46. Martín-Guerrero, J.D.; Lamata, L. Reinforcement Learning and Physics. *Appl. Sci.* **2021**, *11*, 8589. [CrossRef]
47. Zihan, D.; Huang, Y.; Yuan, H.; Dong, H. Introduction to Reinforcement Learning. In *Deep Reinforcement Learning: Fundamentals, Research and Applications*; Dong, H., Ding, Z., Zhang, S., Eds.; Springer Nature Singapore Pte Ltd.: Gateway East, Singapore, 2020; pp. 47–123.
48. Coronato, A.; Naeem, M.; De Pietro, G.; Paragliola, G. Reinforcement Learning for Intelligent Healthcare Applications: A Survey. *Artif. Intell. Med.* **2020**, *109*, 101964. [CrossRef]
49. AgShift Hydra—Power the Transformation of Food through AI. Available online: https://www.agshift.com/ (accessed on 21 December 2023).
50. Zaynub. How Can AI Technology Be Used in Food Safety Implementation? Available online: https://medium.com/@jenny_sBlogs/how-can-ai-technology-be-used-in-food-safety-implementation-efec411e29f0 (accessed on 29 November 2023).
51. Chen, S.; Xiong, J.; Guo, W.; Bu, R.; Zheng, Z.; Chen, Y.; Yang, Z.; Lin, R. Colored Rice Quality Inspection System Using Machine Vision. *J. Cereal Sci.* **2019**, *88*, 87–95. [CrossRef]
52. Alavi, N. Date Grading Using Rule-Based Fuzzy Inference System. *J. Agric. Technol.* **2012**, *8*, 1243–1254.
53. Ireri, D.; Belal, E.; Okinda, C.; Makange, N.; Ji, C. A Computer Vision System for Defect Discrimination and Grading in Tomatoes Using Machine Learning and Image Processing. *Artif. Intell. Agric.* **2019**, *2*, 28–37. [CrossRef]
54. Huang, H.; Liu, L.; Ngadi, M.O. Recent Developments in Hyperspectral Imaging for Assessment of Food Quality and Safety. *Sensors* **2014**, *14*, 7248–7276. [CrossRef]
55. Kamruzzaman, M.; ElMasry, G.; Sun, D.-W.; Allen, P. Non-Destructive Prediction and Visualization of Chemical Composition in Lamb Meat Using NIR Hyperspectral Imaging and Multivariate Regression. *Innov. Food Sci. Emerg. Technol.* **2012**, *16*, 218–226. [CrossRef]
56. Rivera, N.V.; Gómez-Sanchis, J.; Chanona-Pérez, J.; Carrasco, J.J.; Millán-Giraldo, M.; Lorente, D.; Cubero, S.; Blasco, J. Early Detection of Mechanical Damage in Mango Using NIR Hyperspectral Images and Machine Learning. *Biosyst. Eng.* **2014**, *122*, 91–98. [CrossRef]
57. Liu, Y.; Chen, Y.-R.; Kim, M.S.; Chan, D.E.; Lefcourt, A.M. Development of Simple Algorithms for the Detection of Fecal Contaminants on Apples from Visible/near Infrared Hyperspectral Reflectance Imaging. *J. Food Eng.* **2007**, *81*, 412–418. [CrossRef]
58. Gómez-Sanchís, J.; Gómez-Chova, L.; Aleixos, N.; Camps-Valls, G.; Montesinos-Herrero, C.; Moltó, E.; Blasco, J. Hyperspectral System for Early Detection of Rottenness Caused by Penicillium Digitatum in Mandarins. *J. Food Eng.* **2008**, *89*, 80–86. [CrossRef]
59. Lee, J.; Nazki, H.; Baek, J.; Hong, Y.; Lee, M. Artificial Intelligence Approach for Tomato Detection and Mass Estimation in Precision Agriculture. *Sustainability* **2020**, *12*, 9138. [CrossRef]
60. FDA. Tracking and Tracing of Food. Available online: https://www.fda.gov/food/new-era-smarter-food-safety/tracking-and-tracing-food (accessed on 21 December 2023).
61. Hassoun, A.; Alhaj Abdullah, N.; Aït-Kaddour, A.; Ghellam, M.; Beşir, A.; Zannou, O.; Önal, B.; Aadil, R.M.; Lorenzo, J.M.; Mousavi Khaneghah, A. Food Traceability 4.0 as Part of the Fourth Industrial Revolution: Key Enabling Technologies. *Crit. Rev. Food Sci. Nutr.* **2022**, *64*, 873–889. [CrossRef]
62. Ling, C.; Zeng, T.; Su, Y. Research on Intelligent Supervision and Application System of Food Traceability Based on Blockchain and Artificial Intelligence. In Proceedings of the 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), Chongqing, China, 17–19 December 2021; IEEE: Piscataway, NJ, USA, 2021; Volume 2, pp. 370–375.
63. Archana Sristy Blockchain in the Food Supply Chain—What Does the Future Look Like? Available online: https://tech.walmart.com/content/walmart-global-tech/en_us/news/articles/blockchain-in-the-food-supply-chain.html (accessed on 10 December 2023).
64. Roger Aitken IBM & Walmart Launching Blockchain Food Safety Alliance in China with Fortune 500's JD.Com. Available online: https://www.forbes.com/sites/rogeraitken/2017/12/14/ibm-walmart-launching-blockchain-food-safety-alliance-in-china-with-fortune-500s-jd-com/?sh=22a372277d9c (accessed on 10 December 2023).
65. Shahbazi, Z.; Byun, Y.-C. A Procedure for Tracing Supply Chains for Perishable Food Based on Blockchain, Machine Learning and Fuzzy Logic. *Electronics* **2020**, *10*, 41. [CrossRef]
66. Li, S.; Yu, X.; Zhen, Z.; Huang, M.; Lu, J.; Pang, Y.; Wang, X.; Gao, Y. Geographical Origin Traceability and Identification of Refined Sugar Using UPLC-QTof-MS Analysis. *Food Chem.* **2021**, *348*, 128701. [CrossRef] [PubMed]
67. Shang, J.; Liu, N.; Cheng, J.; Gao, W.; Sun, X.; Guo, M. Analysis and Comparison of Lipids in Saanen Goat Milk from Different Geographic Regions in China Based on UHPLC-QTOF-MS Lipidomics. *Food Res. Int.* **2022**, *157*, 111441. [CrossRef]
68. Liu, Z.; Zhao, M.; Wang, X.; Li, C.; Liu, Z.; Shen, X.; Zhou, D. Investigation of Oyster Crassostrea Gigas Lipid Profile from Three Sea Areas of China Based on Non-Targeted Lipidomics for Their Geographic Region Traceability. *Food Chem.* **2022**, *386*, 132748. [CrossRef]
69. Balamurugan, S.; Ayyasamy, A.; Joseph, K. An Efficient Bayes Classifiers Algorithm for Traceability of Food Supply Chain Management Using Internet of Things. *Int. J. Eng. Adv. Technol.* **2019**, *9*, 2995–3005. [CrossRef]
70. Wang, J.; Yue, H.; Zhou, Z. An Improved Traceability System for Food Quality Assurance and Evaluation Based on Fuzzy Classification and Neural Network. *Food Control* **2017**, *79*, 363–370. [CrossRef]
71. Violino, S.; Ortenzi, L.; Antonucci, F.; Pallottino, F.; Benincasa, C.; Figorilli, S.; Costa, C. An Artificial Intelligence Approach for Italian EVOO Origin Traceability through an Open Source IoT Spectrometer. *Foods* **2020**, *9*, 834. [CrossRef] [PubMed]

72. Frąckiewicz, M. The Role of Predictive Analytics in Food Safety Risk Assessment. Available online: https://ts2.space/en/the-role-of-predictive-analytics-in-food-safety-risk-assessment/#gsc.tab=0 (accessed on 29 November 2023).
73. Wu, L.-Y.; Weng, S.-S. Ensemble Learning Models for Food Safety Risk Prediction. *Sustainability* **2021**, *13*, 12291. [CrossRef]
74. Zhang, Y. Food Safety Risk Intelligence Early Warning Based on Support Vector Machine. *J. Intell. Fuzzy Syst.* **2020**, *38*, 6957–6969. [CrossRef]
75. Liu, N.; Bouzembrak, Y.; Van den Bulk, L.M.; Gavai, A.; van den Heuvel, L.J.; Marvin, H.J.P. Automated Food Safety Early Warning System in the Dairy Supply Chain Using Machine Learning. *Food Control* **2022**, *136*, 108872. [CrossRef]
76. Rortais, A.; Barrucci, F.; Ercolano, V.; Linge, J.; Christodoulidou, A.; Cravedi, J.-P.; Garcia-Matas, R.; Saegerman, C.; Svečnjak, L. A Topic Model Approach to Identify and Track Emerging Risks from Beeswax Adulteration in the Media. *Food Control* **2021**, *119*, 107435. [CrossRef]
77. Marvin, H.J.P.; Bouzembrak, Y.; Janssen, E.M.; van der Fels-Klerx, H.J.; van Asselt, E.D.; Kleter, G.A. A Holistic Approach to Food Safety Risks: Food Fraud as an Example. *Food Res. Int.* **2016**, *89*, 463–470. [CrossRef] [PubMed]
78. Bouzembrak, Y.; Camenzuli, L.; Janssen, E.; Van der Fels-Klerx, H.J. Application of Bayesian Networks in the Development of Herbs and Spices Sampling Monitoring System. *Food Control* **2018**, *83*, 38–44. [CrossRef]
79. Zhang, R.; Zhou, L.; Zuo, M.; Zhang, Q.; Bi, M.; Jin, Q.; Xu, Z. Prediction of Dairy Product Quality Risk Based on Extreme Learning Machine. In Proceedings of the 2018 2nd International Conference on Data Science and Business Analytics (ICDSBA), Changsha, China, 21–28 September 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 448–456.
80. Nogales, A.; Díaz-Morón, R.; García-Tejedor, Á.J. A Comparison of Neural and Non-Neural Machine Learning Models for Food Safety Risk Prediction with European Union RASFF Data. *Food Control* **2022**, *134*, 108697. [CrossRef]
81. Alfian, G.; Syafrudin, M.; Farooq, U.; Ma'arif, M.R.; Syaekhoni, M.A.; Fitriyani, N.L.; Lee, J.; Rhee, J. Improving Efficiency of RFID-Based Traceability System for Perishable Food by Utilizing IoT Sensors and Machine Learning Model. *Food Control* **2020**, *110*, 107016. [CrossRef]
82. Liu, Y.; Han, W.; Zhang, Y.; Li, L.; Wang, J.; Zheng, L. An Internet-of-Things Solution for Food Safety and Quality Control: A Pilot Project in China. *J. Ind. Inf. Integr.* **2016**, *3*, 1–7. [CrossRef]
83. Khan, P.W.; Byun, Y.-C.; Park, N. IoT-Blockchain Enabled Optimized Provenance System for Food Industry 4.0 Using Advanced Deep Learning. *Sensors* **2020**, *20*, 2990. [CrossRef] [PubMed]
84. Geng, Z.; Zhao, S.; Tao, G.; Han, Y. Early Warning Modeling and Analysis Based on Analytic Hierarchy Process Integrated Extreme Learning Machine (AHP-ELM): Application to Food Safety. *Food Control* **2017**, *78*, 33–42. [CrossRef]
85. Tian, F. A Supply Chain Traceability System for Food Safety Based on HACCP, Blockchain & Internet of Things. In Proceedings of the 2017 International Conference on Service Systems and Service Management, Dalian, China, 16–18 June 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–6.
86. Sadilek, A.; Caty, S.; DiPrete, L.; Mansour, R.; Schenk Jr, T.; Bergtholdt, M.; Jha, A.; Ramaswami, P.; Gabrilovich, E. Machine-Learned Epidemiology: Real-Time Detection of Foodborne Illness at Scale. *NPJ Digit. Med.* **2018**, *1*, 36. [CrossRef] [PubMed]
87. Tutul, M.J.I.; Alam, M.; Wadud, M.A.H. Smart Food Monitoring System Based on Iot and Machine Learning. In Proceedings of the 2023 International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM), Gazipur, Bangladesh, 16–17 June 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.
88. Wang, H.; Cui, W.; Guo, Y.; Du, Y.; Zhou, Y. Machine Learning Prediction of Foodborne Disease Pathogens: Algorithm Development and Validation Study. *JMIR Med. Inf.* **2021**, *9*, e24924. [CrossRef] [PubMed]
89. Yan, S.; Wang, S.; Qiu, J.; Li, M.; Li, D.; Xu, D.; Li, D.; Liu, Q. Raman Spectroscopy Combined with Machine Learning for Rapid Detection of Food-Borne Pathogens at the Single-Cell Level. *Talanta* **2021**, *226*, 122195. [CrossRef] [PubMed]
90. Pesesky, M.W.; Hussain, T.; Wallace, M.; Patel, S.; Andleeb, S.; Burnham, C.-A.D.; Dantas, G. Evaluation of Machine Learning and Rules-Based Approaches for Predicting Antimicrobial Resistance Profiles in Gram-Negative Bacilli from Whole Genome Sequence Data. *Front. Microbiol.* **2016**, *7*, 1887. [CrossRef]
91. Teyhouee, A.; McPhee-Knowles, S.; Waldner, C.; Osgood, N. Prospective Detection of Foodborne Illness Outbreaks Using Machine Learning Approaches. In Proceedings of the Social, Cultural, and Behavioral Modeling: 10th International Conference, SBP-BRiMS 2017, Washington, DC, USA, 5–8 July 2017; Proceedings 10. Springer: Berlin/Heidelberg, Germany, 2017; pp. 302–308.
92. Vangay, P.; Steingrimsson, J.; Wiedmann, M.; Stasiewicz, M.J. Classification of Listeria Monocytogenes Persistence in Retail Delicatessen Environments Using Expert Elicitation and Machine Learning. *Risk Anal.* **2014**, *34*, 1830–1845. [CrossRef]
93. Wheeler, N.E.; Gardner, P.P.; Barquist, L. Machine Learning Identifies Signatures of Host Adaptation in the Bacterial Pathogen Salmonella Enterica. *PLoS Genet.* **2018**, *14*, e1007333. [CrossRef]
94. Yi, J.; Wisuthiphaet, N.; Raja, P.; Nitin, N.; Earles, J.M. AI-Enabled Biosensing for Rapid Pathogen Detection: From Liquid Food to Agricultural Water. *Water Res.* **2023**, *242*, 120258. [CrossRef]
95. Her, H.-L.; Wu, Y.-W. A Pan-Genome-Based Machine Learning Approach for Predicting Antimicrobial Resistance Activities of the Escherichia Coli Strains. *Bioinformatics* **2018**, *34*, i89–i95. [CrossRef] [PubMed]
96. Chakraborty, S.K.; Mahanti, N.K.; Mansuri, S.M.; Tripathi, M.K.; Kotwaliwale, N.; Jayas, D.S. Non-Destructive Classification and Prediction of Aflatoxin-B1 Concentration in Maize Kernels Using Vis–NIR (400–1000 Nm) Hyperspectral Imaging. *J. Food Sci. Technol.* **2021**, *58*, 437–450. [CrossRef] [PubMed]
97. Ye, W.; Yan, T.; Zhang, C.; Duan, L.; Chen, W.; Song, H.; Zhang, Y.; Xu, W.; Gao, P. Detection of Pesticide Residue Level in Grape Using Hyperspectral Imaging with Machine Learning. *Foods* **2022**, *11*, 1609. [CrossRef] [PubMed]

98. Baghel, A.S.; Bhardwaj, A.; Ibrahim, W. Optimization of Pesticides Spray on Crops in Agriculture Using Machine Learning. *Comput. Intell. Neurosci.* **2022**, *2022*, 9408535.
99. Shen, Y.; Zhao, E.; Zhang, W.; Baccarelli, A.A.; Gao, F. Predicting Pesticide Dissipation Half-Life Intervals in Plants with Machine Learning Models. *J. Hazard. Mater.* **2022**, *436*, 129177. [CrossRef] [PubMed]
100. Bhatia, S.; Albarrak, A.S. A Blockchain-Driven Food Supply Chain Management Using QR Code and XAI-Faster RCNN Architecture. *Sustainability* **2023**, *15*, 2579. [CrossRef]
101. Petrea, Ș.-M.; Costache, M.; Cristea, D.; Strungaru, Ș.-A.; Simionov, I.-A.; Mogodan, A.; Oprica, L.; Cristea, V. A Machine Learning Approach in Analyzing Bioaccumulation of Heavy Metals in Turbot Tissues. *Molecules* **2020**, *25*, 4696. [CrossRef]
102. Yu, Q.; Li, J.; Yao, L.; Li, C.; Cao, J.; Qiao, J.; Wu, Q. Estimation of Heavy-Metal Concentrations in Winter Wheat Leaves from Typical Sewage Irrigation Area Based on Canopy Reflectance Spectra. *J. Appl. Remote Sens.* **2018**, *12*, 36019. [CrossRef]
103. Park, D.; Kim, K.; Kim, S.; Spranger, M.; Kang, J. FlavorGraph: A Large-Scale Food-Chemical Graph for Generating Food Representations and Recommending Food Pairings. *Sci. Rep.* **2021**, *11*, 931. [CrossRef]
104. Iwendi, C.; Khan, S.; Anajemba, J.H.; Bashir, A.K.; Noor, F. Realizing an Efficient IoMT-Assisted Patient Diet Recommendation System through Machine Learning Model. *IEEE Access* **2020**, *8*, 28462–28474. [CrossRef]
105. Vairale, V.S.; Shukla, S. Recommendation of Diet Using Hybrid Collaborative Filtering Learning Methods. In *Advances in Computational Intelligence and Informatics, Proceedings of ICACII 2019, Hyderabad, India, 20–21 December 2019*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 309–318.
106. Sowah, R.A.; Bampoe-Addo, A.A.; Armoo, S.K.; Saalia, F.K.; Gatsi, F.; Sarkodie-Mensah, B. Design and Development of Diabetes Management System Using Machine Learning. *Int. J. Telemed. Appl.* **2020**, *2020*, 8870141. [CrossRef] [PubMed]
107. Naik, P.A. Intelligent Food Recommendation System Using Machine Learning. *Nutrition* **2020**, *5*, 616–619.
108. Buyuktepe, O.; Catal, C.; Kar, G.; Bouzembrak, Y.; Marvin, H.; Gavai, A. Food Fraud Detection Using Explainable Artificial Intelligence. *Expert. Syst.* **2023**, e13387. [CrossRef]
109. Bouzembrak, Y.; Marvin, H.J.P. Prediction of Food Fraud Type Using Data from Rapid Alert System for Food and Feed (RASFF) and Bayesian Network Modelling. *Food Control* **2016**, *61*, 180–187. [CrossRef]
110. Mithun, B.S.; Shinde, S.; Bhavsar, K.; Chowdhury, A.; Mukhopadhyay, S.; Gupta, K.; Bhowmick, B.; Kimbahune, S. Non-Destructive Method to Detect Artificially Ripened Banana Using Hyperspectral Sensing and RGB Imaging. In Proceedings of the Sensing for Agriculture and Food Quality and Safety X, Orlando, FL, USA, 17–18 April 2018; Volume 10665, pp. 122–130.
111. Pulluri, K.K.; Kumar, V.N. Qualitative and Quantitative Detection of Food Adulteration Using a Smart E-Nose. *Sensors* **2022**, *22*, 7789. [CrossRef] [PubMed]
112. de Santana, F.B.; Neto, W.B.; Poppi, R.J. Random Forest as One-Class Classifier and Infrared Spectroscopy for Food Adulteration Detection. *Food Chem.* **2019**, *293*, 323–332. [CrossRef] [PubMed]
113. Lim, K.; Pan, K.; Yu, Z.; Xiao, R.H. Pattern Recognition Based on Machine Learning Identifies Oil Adulteration and Edible Oil Mixtures. *Nat. Commun.* **2020**, *11*, 5353. [CrossRef] [PubMed]
114. Mu, T.; Chen, S.; Zhang, Y.; Chen, H.; Guo, P.; Meng, F.D. Portable Detection and Quantification of Olive Oil Adulteration by 473-Nm Laser-Induced Fluorescence. *Food Anal. Methods* **2016**, *9*, 275–279. [CrossRef]
115. Laga, S.A.; Sarno, R. Temperature Effect of Electronic Nose Sampling for Classifying Mixture of Beef and Pork. *Indones. J. Electr. Eng. Comput. Sci.* **2020**, *19*, 1626–1634. [CrossRef]
116. Wang, X.; Bouzembrak, Y.; Lansink, A.O.; van der Fels-Klerx, H.J. Application of Machine Learning to the Monitoring and Prediction of Food Safety: A Review. *Compr. Rev. Food Sci. Food Saf.* **2022**, *21*, 416–434. [CrossRef]