# A Machine Learning-Based Anomaly Detection Method and Blockchain-Based Secure Protection Technology in Collaborative Food Supply Chain

Yuh-Min Chen, National Cheng Kung University, Taiwan
Tsung-Yi Chen, Nanhua University, Taiwan*
Jyun-Sian Li, National Cheng Kung University, Taiwan

## ABSTRACT

The complexity of the collaborative food supply chain has resulted in the frequent occurrence of food safety incidents, which harm people's health and life. Therefore, the maintenance of food safety has become a key value. This study expects to solve the food safety problem and bring more benefits to people using intelligent systems. To meet the safety needs of the collaborative food supply chain, this study designed a food safety protection system architecture which collects the supply and sales data of various suppliers, as well as the data of equipment used in production. The architecture can carry out anomaly detections with machine learning to make a preliminary judgement on whether a problem has occurred in this batch of food during the transaction, and then implement in-depth anomaly detections with the supplier's equipment to determine the stage at which this problem occurred. The proposed system can help food operators achieve effective food monitoring, prediction, prevention, and improvement, thereby improving food safety.

## KEYWORDS

Anomaly detection, Blockchain, Food safety, Food supply chain, Machine learning

## 1. INTRODUCTION

The delicacy and complexity of food has resulted in the ever-increasing issue of food safety (George et al., 2019). Therefore, food safety has become a policy of concerned in many countries around the world. With globalization and the current complications of collaborative food supply chains, the complexity and uncertainty of food safety issues have increased, and the maintenance of food safety has become more difficult. Any part of the supply chain may be a cause of food hazard due to human, machinery, material, regulatory/procedural, or environmental factors.

DOI: 10.4018/IJeC.315789
*Corresponding Author

Thus, food safety management should include the entire food supply chain in the monitoring and management scope, and consider all stakeholders and influencing factors (ISO, 2018). The stakeholders of food safety include all the operators in the collaborative food supply chain, governmental food supervision departments, international food safety organizations, and consumers. The influencing factors of food safety include the attributes of human, machinery, material, regulatory/procedural, and environmental factors in the food lifecycle. In terms of raw material supply, the complicated sources and increased pollution sources have resulted in more and more contaminated foods, or foods with illegal chemical additives (Resende-Filho & Hurley, 2012). At the same time, some suppliers in the food supply chain lack food safety management capacity, are unaware of their responsibilities, and lack integrity, thus resulting in deliberate falsification or cover up, and leading to information asymmetry between consumers and manufacturers (Sun & Wang, 2019). The above-mentioned problems may cause food safety incidents. Therefore, the question of how to effectively and comprehensively manage the food safety issues of collaborative supply chains is difficult in both method and technology.

The advancements in computer and network technologies have enable the rapid accumulation of data and information, as well as the application of big data in various industries. The significant increase in data and the upgrading of the computing speed gives rise to the effective utilization of artificial intelligence and machine learning. In recent years, some scholars have applied machine learning in the food industry for the prediction of crop yields (Chlingaryan et al., 2018; Mehra et al., 2018), life cycle evaluation (Nabavi-Pelesaraei et al., 2017), and logistics and distribution management (Krisztin, 2018). However, the abovementioned research only focused on the quality prediction and management of the various stages of the supply chain, but lacked quality management and anomaly protection within the entire supply chain.

Food safety supervision and management of the entire supply chain is necessary, and must be able to: (1) track the supply chain of food raw materials; (2) ensure the autonomous management of all supply points in a supply chain; and (3) detect anomalies in the supply chain and analyze the causes of such problems, while simultaneously track the other affected foods (Wang & Yue, 2017).

Food safety is realized upon a comprehensive food protection architecture, as well as the collection and analysis of the data related to food safety in supply chains (Hazen et al., 2014). However, the food supply chain is complicated and diversified with various types of data from each stage of the lifecycle. How to define and collect useful data related to food safety in the supply chain, in order to facilitate food safety supervision and management in the food lifecycle, is an important topic for food safety management (Kamble et al., 2020). Information sharing and coordination among organizations have become complicated in today's globalized food chain (Morgan et al., 2018). In addition, technologies for the understanding and prediction of food safety incidents need to be further developed based on the food safety data in the supply chain, so as to prevent anomalies.

This study designed a food safety protection architecture to meet the requirements of food safety in collaborative supply chains. It also performs anomaly detections by machine learning for the preliminary determination of a problem occurrence in a food batch during the transaction. Then, this architecture conducts an in-depth investigation into the supplier's equipment for anomaly detection, in order to determine at which stage the problem occurred. This can help food operators to achieve effective food monitoring, prediction, prevention, and improvement so as to enhance the effectiveness of food safety management.

## 2. LITERATURE REVIEW

A food supply chain consists of five steps, including raw material supply, production and processing, warehousing logistics, sales, and consumption. Through continuous data collection in different steps, all members of the supply chain can monitor the real-time location and the status of raw materials (tracking data) to make decisions regarding production, logistics planning, and resource allocation (Gaukler et al. 2008; Galvez et al., 2018).

With the development of information technology, supply chains have changed from the traditional linear model to the network model, thus, increasing the complexity of supply chains, and potential supply chain risks are derived in turn (Sharma et al., 2020), including supply and sales risks (Guan et al., 2011; Assefa et al., 2017), equipment and quality risks (Shirani & Demichela, 2015), and environmental risks (Freise & Seuring, 2015). Thus, new methods or technologies, such as blockchain, IoT, and machine learning, have been applied to eliminate such risks.

At present, the most common IoT technologies for obtaining supply chain information include quick response code (QR-code), radio frequency identification (RFID), and wireless sensor node (WSN). Many fresh food supply chains use RFID for temperature measurement and WSN for data collection, as they show advantages in quality, safety, and analysis, and can optimize the supply chains (Óskarsdóttir & Oddsson, 2019).

Ndraha et al. (2018) mentioned the specific problem of maintaining the temperature of transport containers in the food supply chain. Even if the temperature of the container is slightly higher or lower than the appropriate temperature, this slight temperature change may cause the food to spoil or greatly reduce the quality. Moreover, it has been observed that the problems of cold-sensitive fruits and vegetables being transported at too low temperatures, as well as heat-sensitive food being transported at too high temperatures, frequently occur, which leads to at least 50% of wasted food. However, food suppliers are neither aware of it nor able to respond appropriately (Nunes et al., 2009). As can be seen from the above examples, the IoT technologies can improve food safety and integrity, such as identifying and connecting specific resources in series to reduce food anomalies.

The blockchain, as invented by Nakamoto (2008), is a distributed ledger technology, and its main feature is that it cannot be tampered with. In other words, it is impossible to eliminate any deployed information without prior consensus (Dinh & Thai, 2018).

Currently, centralized cloud servers are commonly seen. However, malicious servers may sign duplicate transactions or problematic transactions (Fraga & Fernández-Caramés, 2019), or centralized cloud servers may stop working due to denial of service attacks, software issues, or single point failure (Kshetri 2018; Makhdoom et al., 2019). In contrast, the blockchain, which blocks malicious nodes by linking blocks and using a consensus mechanism, has the advantage of not being easily changed by malicious parties, transparency, and tamper-proof (Wang et al., 2018). These characteristics can ensure the decision-making in machine learning to be more transparent and reliable (Dinh & Thai, 2018; Mohanta et al., 2020), and also enable each supply point in the supply chain to access its upstream and downstream food raw material records (Creydt & Fischer, 2019). Thus, it is possible to conduct a current situation control, prediction, and prevention analysis without infringing product secrets, thereby enhancing the trust, transparency, and safety of the entire supply chain.

Machine learning can be divided into two types, supervised learning and unsupervised learning (Bishop, 2016). According to a training data set with target variables, supervised learning establishes a model that connects input variables with target variables. The algorithms include logistic regression, support vector machine, naive bayes, decision tree, and random forest, thus, it can solve classification and regression problems. Unsupervised learning is based on similarities, differences, and rules of the training data without target variables. The algorithms of K-means and Principal Components Analysis (PCA) can solve various problems, such as clustering and dimensionality reduction (Ayoub, 2020).

Many studies have suggested that the random forest algorithm performs better than other popular machine learning models in solving classification problems (Bhattacharyya et al., 2011; Dal Pozzolo et al., 2017). This study thus used the random forest algorithm as the supply and marketing anomaly detection model, and compared the the results with the abovementioned popular machine learning algorithms for implementation verification.

Deep learning is a subset of machine learning to enable anomaly detection (Lindemann et al. 2021; Pang et al., 2021), most of which are neural network-like models (Xu et al., 2020). The commonly known neural networks for time series include the recurrent neural network (RNN) (Mikolov et al. 2010) and the long short term memory (LSTM) (Hochreiter & Schmidhuber, 1997; Nguyen et al.,

2021). As the long-term recursion of RNN tends to have the vanishing gradient problem (Hochreiter 1998), errors will increase when processing long-term sequence data. In order to achieve long-term and short-term memory, LSTM uses gates and memory units to control the data flow; hence, it can effectively improve the problem of RNN gradient disappearance. As the equipment anomaly detection in this study needs to consider the relevance of long-term time, the LSTM prediction model, which could improve the vanishing gradient problem of RNN, was selected as the prediction model. Nguyen et al. (2021) proposed forecast and anomaly detection approaches using LSTM for making better decisions in supply chain management.

There are many data imbalances in reality due to the domination of majority samples, and it may be impossible to learn all the patterns and laws of minority samples well, resulting in a failure to detect minority samples. Sampling optimization enables minority samples to achieve the effect of expansion, and then, improves the situation where the algorithm cannot learn well due to the imbalanced data volume (Chen et al., 2018).

Sampling optimization includes undersampling and oversampling. The most well-known oversampling method is the Synthesized Minority Oversampling Technique (SMOTE), which reduces an imbalanced number by generating new minority samples (Chawla et al., 2002). Regarding the undersampling method, Wilson (1972) proposed the edited nearest neighbor (ENN) rule, which states that if a sample's neighboring samples are different from its own category, it will be deleted. This study suggested that combining SMOTE with ENN can expand the minority samples while separating the majority samples from the minority samples.

Determining a set of hyperparameters can show the best performance of the model, and this process is called hyperparameter optimization (Wu et al., 2019). Bergstra and Bengio (2012) proposed the random search, which reduces the insignificant hyperparameters to improve the overall model efficiency and obtains the approximate solution. As the hyperparameters of deep learning are more complicated, traditional optimization techniques are unsuitable. Bayesian optimization is an algorithm that can solve this type of optimization problem, as it combines the prior information of the unknown function with the sample information, uses the Bayesian formula to obtain the posterior information of function distribution, and then, infers where the function obtains the best value, based on the posterior information (Wu et al., 2019). This study used Bayesian optimization to optimize the LSTM model, and random search to optimize the random forest model.

## 3. METHODOLOGY

### 3.1 Blockchain-based Food Safety Protection Architecture

In the food supply chain, each decision requires communication and cooperation with all suppliers to achieve the desired results; however, due to the complexity of the food supply chain, the information and processes are not transparent, and there are high risks for both buyers and sellers. On the other hand, as there are more and more elements that affect food quality, and food is more sensitive to the impact of quality than other sectors, it is necessary to conduct analyses and detections at all stages of the food lifecycle.

In order to solve the above problems, this study designed a food safety protection system architecture (Figure 1), where each supplier manages independently. The blockchain was used as the system environment for supply chain communication to provide the footprint as collected by autonomous management in the blockchain, in order to grasp the current situation of the supply chain, and realize prediction and prevention within the supply chain.

The bottom layer of the architecture is composed of the food supply chain, where all members can conduct real-time and accurate data collection through various collection media, such as QR-code, RFID, sensor device, and GPS. Based on the food safety data model, the sales and production data in the food lifecycle are mainly collected. With the easily deployed and flexible communication

protocols in the communication transmission layer, the data read by the sensor are transmitted to the monitoring and management layer, which are used as a relay station for data deployment to the blockchain for data storage, maintenance, monitoring, and management.

The next layer is the blockchain layer which aims to promote data transparency and improve the security of the food supply chain. The data in the monitoring and management layer are processed according to the blockchain structure to form a series of blocks which are connected in chronological order to form a blockchain. During the process of formation, a consensus mechanism is required to ensure consistency and fault tolerance between the distributed systems. Moreover, a reward mechanism is also needed to reward the stakeholders for their accurate recording and maintenance of the data. In order to protect the supply chain, all members of the supply chain and supervisory units should jointly act as the consensus nodes. While the supply chain members can execute transactions through smart contracts, they cannot tamper with the data without destroying the HASH link.

Finally, when a complete food blockchain is formed, users can read or write data into the blockchain through various devices (e.g., smartphones or tablets) and analyze and apply the data recorded by the blockchain. For example, food suppliers can predict environmental changes. This study used the data of the blockchain for application and machine learning to detect whether food is abnormal. Detection was carried out on two levels.

(1) Supply and sales anomaly detection: Abnormal behaviors are detected at the supply network level. Most of the interactions between suppliers consist of transactions (supply and sales), and such transaction processes leave a lot of data. The analysis of these data can obtain a model that can judge the transaction risks, while detection with this model can initially assess whether a batch of products has any anomalies.
(2) Equipment anomaly detection: The supply network is composed of many factories. Taking a factory as a unit can collect numerous dynamic data, including the data regarding people, machines, materials, and environment. These data all have specific patterns or laws corresponding to food quality. Therefore, this study suggested that it is necessary to preliminarily evaluate the status of each batch of products with supply and sales anomaly detection, and then, perform detailed detection according to the correctness of the data of each factory.

## 3.2 Food Anomaly Detection Method

In the anomaly detection method, detection is mainly conducted on the supply network layer and inside of each factory, as these two layers have different anomalies (Table 1). All the anomalies described in Table 1 affect the quality of food, thus, the anomaly degree of a product in the production stage can be determined through the preliminary evaluation of the product status via supply and sales anomaly detection. Afterwards, an in-depth investigation of the supplier regarding the concerned batch of products is conducted in order to perform equipment anomaly detection.

### *3.2.1 Supply and Sales Anomaly Detection*

The supply and sales anomaly detection technology is shown in Figure 2. First, the supply and sales information is collected from the suppliers and deployed in the food blockchain, then the supervisory unit marks any anomaly data found in the chain. Next, pre-processing is performed for the marked supply and sales data, and the best feature combination is selected through feature selection. Random forest model training is performed, and the best supply and sales anomaly detection model is obtained through model optimization and model evaluation. After the model is obtained, it is necessary to conduct a real-time supply and sales detection on the data, which is in an unlabeled state. After the data meet the detection data pattern through data pre-processing, the optimized model is applied to detect issues, and the final supply and sales anomaly detection result is obtained to inform the user when the current supply and sales status is abnormal.

**Figure 1. The blockchain-based secure food protection architecturre**

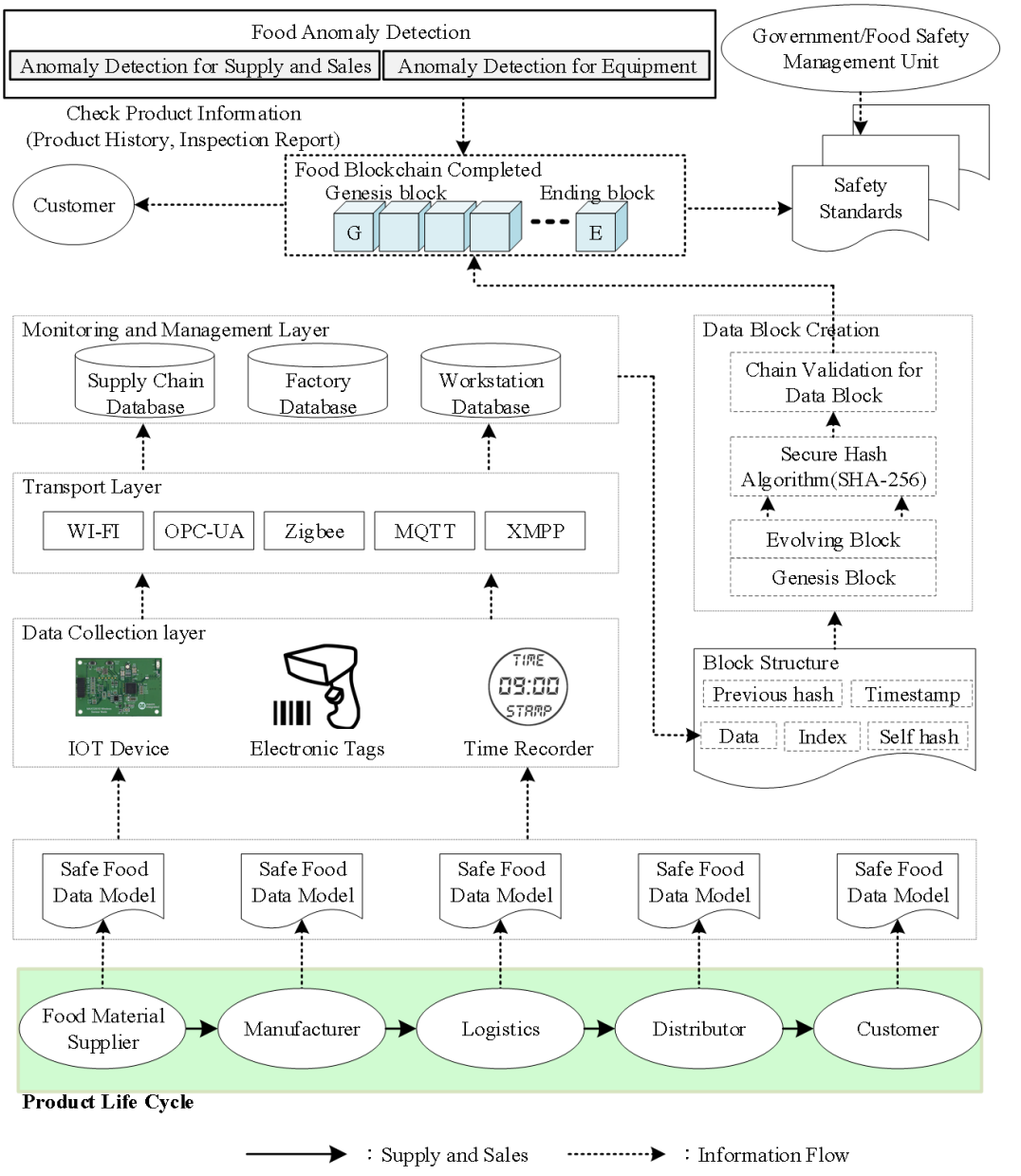
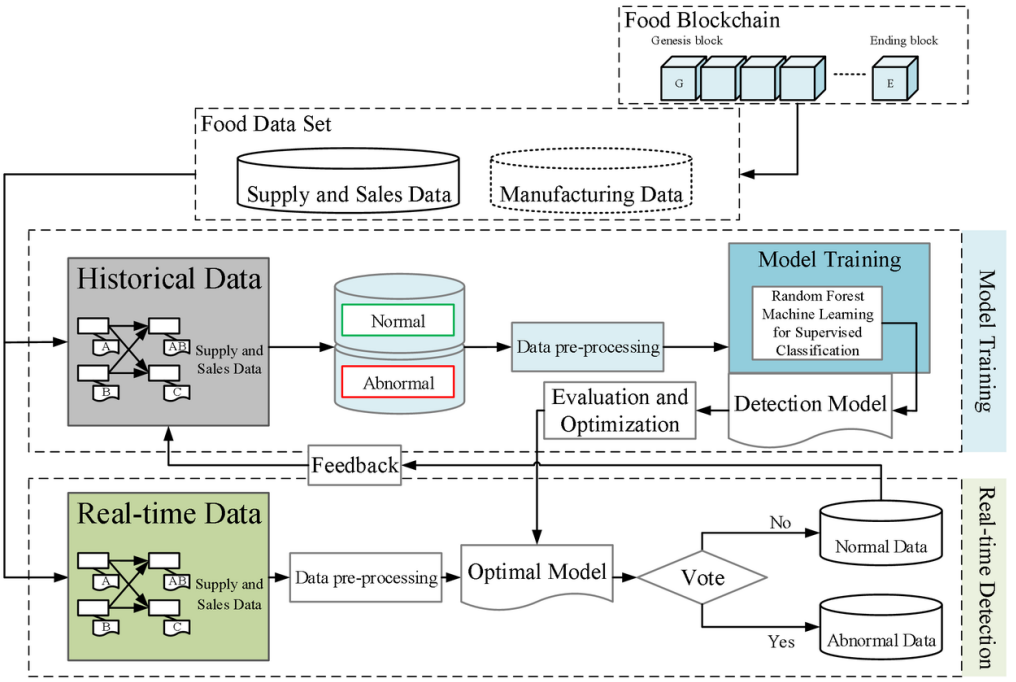



**Table 1. Abnormal behavior**

| Layer | Focus | Abnormal behavior | Impact |
|---|---|---|---|
| Supply network | Abnormal supply and sales | • Delayed delivery<br>• Abnormal quality of raw materials<br>• Incorrect quantity of raw material<br>• Smuggling of raw materials | • the shelf life of food<br>• purchaser's senses<br>• subsequent production<br>• Make the purchaser without protection |
| Inside the factory | Equipment abnormal | • Equipment noise<br>• Damaged equipment parts<br>• Consumables have not been replaced (cutting tools) | • Unstable production quality<br>• Production line stopped<br>• Affect food quality |

**Figure 2. Supply and sale of anomaly detection technology**




This study used the random forest algorithm as the supply and sales anomaly detection model. In fact, as there are various problems, such as missing data, uniform distribution, and feature interference, the data are processed before model training to ensure they are standardized and have a consistent format. This way, model training would not result in a poor model due to data flaws. The data pre-processing method applied in this study is as follows.

(1) Data cleaning: Meaningless features are deleted after filling in missing values, as they are not helpful in model training and have no association with the features.
(2) Data conversion: All features are classified values and must be converted into easy-to-understand digital formats. For example, the supplier name "M348934600" is converted into code 34 among all suppliers, and the features with smaller classification numbers are used as one-hot encoding for discretization.

According to the standard process of machine learning, the model must be optimized after model training, and the optimization method is shown in Figure 3. This study used SMOTE combined with ENN for sampling optimization and random search for hyperparameter optimization.

Hyperparameter optimization is to select the optimal hyperparameter value based on the evaluation indicators. In this study, random search is selected for hyperparameter optimization, where the principle of operation (Figure 4) is to first separate the training and test data from the original data set, and use the training data for model training. The range of hyperparameters and the number of search interactions must be defined before model training. According to its range and set number of interactions, uniform sampling and random search are performed, and finally, the best model is selected by cross validation.

This study used SMOTE combined with ENN for data sampling and designed the algorithm shown in Figure 5 which is explained as follows:

**Figure 3. Model optimization method**

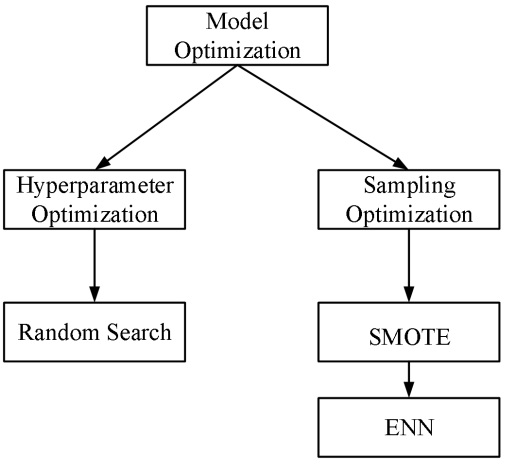
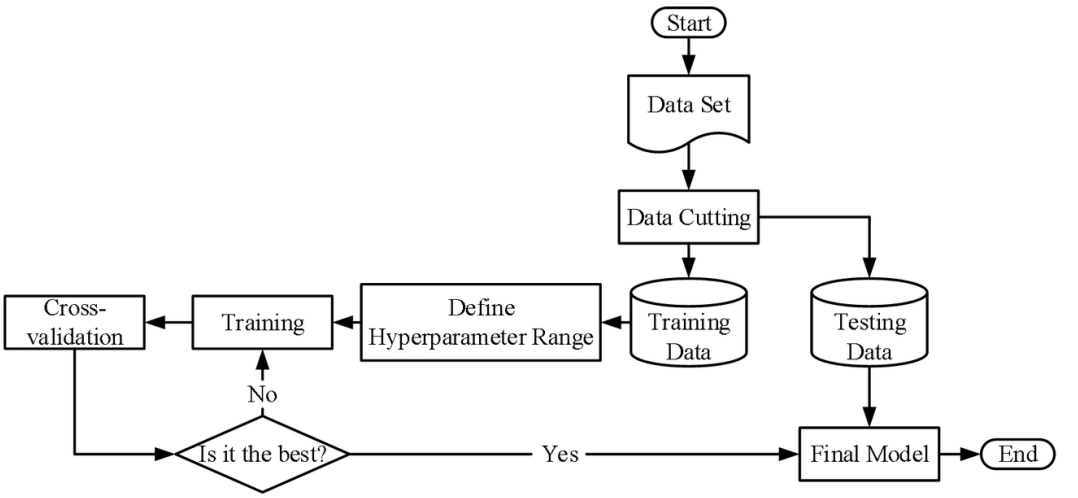



**Figure 4. Random search principle**




(1) The calculated distance $d\left(X,Y\right)=\sqrt{\sum_{i=1}^{n}\left(x_i-y_i\right)^2}$, $X=\left(x_1,x_2,\ldots,x_n\right)$ represents one anomaly data, and $Y=\left(y_1,y_2,\ldots,y_n\right)$ denotes that the rest of the anomaly data all have n features. Each anomaly data uses distance to find the k-nearest neighbors closest to itself, which is called $Z_j$.

(2) Calculate difference $\left(Z_j-X\right)$ between $X$ and $Z_j$.

(3) Generate a random number between 0 and 1.

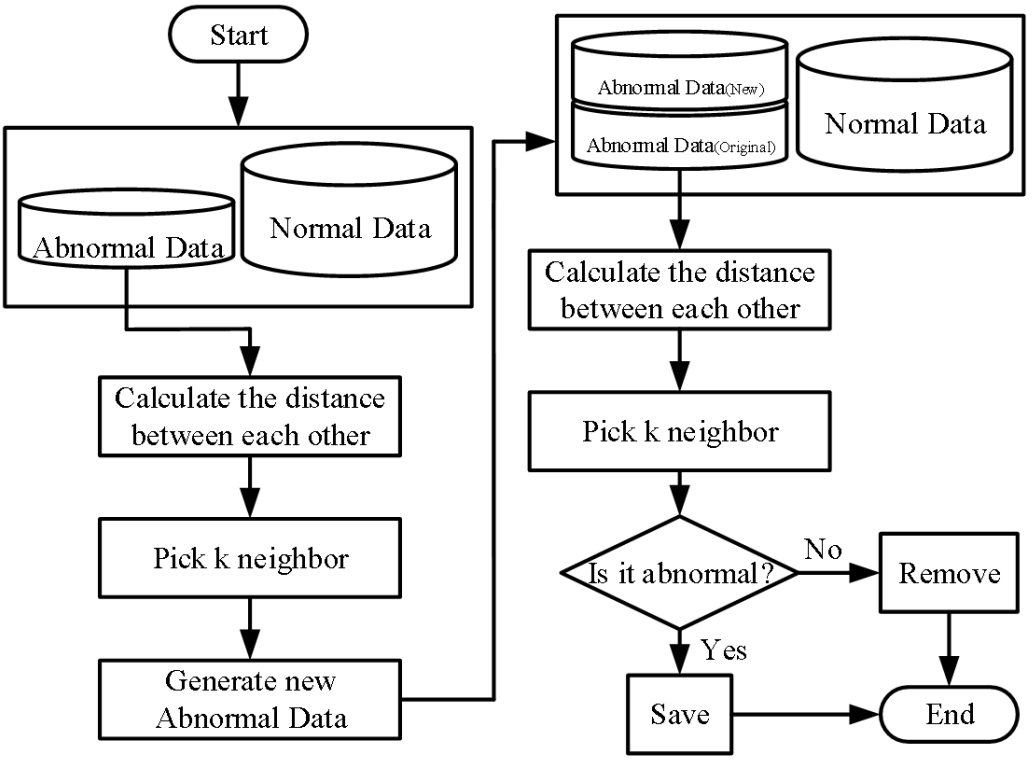
(4) Generate new anomaly data $P_j=X+random\left(0,1\right)\times\left(Z_j-X\right), j=1,2,\ldots k$.

(5) Obtain a new data set, and the anomaly data of this data set has been increased.

(6) Repeat Step (1), and select k-nearest neighbors for each anomaly data.

(7) Determine whether the neighbors are also anomaly data; if not, delete the neighbor, otherwise, continue.

**Figure 5. SMOTE-ENN**




### *3.2.2 Equipment Anomaly Detection*

Equipment anomaly detection is mainly for the detection of food processing equipment. All suppliers continuously collect equipment data through sensors and store them in the food blockchain; the historical data marked as normal is pre-processed, and then, neural network model training is performed to obtain the best equipment anomaly detection model through model optimization and model evaluation. Finally, real-time equipment status detection is performed, and the anomaly degree is estimated according to the error of model detection. The users are informed whether the current equipment status is correct.

This study used LSTM as the neural network for equipment anomaly detection as it is effective in handling time series problems. Since data features have many different types and minority features may affect model learning, data pre-processing is required, as shown in Figure 6. The processing actions are described as follows.

(1) Data cleaning: Meaningless features are removed from the file after filling in missing values.
(2) Normalization: In order to eliminate the unevenness of data distribution in the features, the variance of different features or the same feature in different data is reduced, in order to improve the convergence speed and accuracy of the model. The MinMaxScaler method is applied for normalization, where normalization is in the range between 1 and -1, and the formula is (1). $X_{min}$ is the minimum value in the feature data; $X_{max}$ is the maximum value in the feature data; $max$ is the maximum value of the expected zoom; and $min$ is the minimum value of the expected zoom.

$$X_{scaled} = \frac{\left(X - X_{min}\right)}{\left(X_{max} - X_{min}\right)} \times \left(max - min\right) + min \tag{1}$$

(3) Feature selection: Pearson correlation coefficient is used in univariate feature selection as the feature selection algorithm, as shown in formula (2), where $x_i$ is the first data of the feature $x$; $y_i$ is $i$ th data of feature $y$; and there are a total of n data, where $\mu_x$ and $\mu_y$ denotet the mean of the features $x$ and $y$, respectively. Pearson's correlation coefficient mainly measures the degree of "linear" correlation between two variables.

$$\rho = \frac{\sum_{i=1}^{n}\left(x_i - \mu_x\right)\left(y_i - \mu_y\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \mu_x\right)^2 \sum_{i=1}^{n}\left(y_i - \mu_y\right)^2}} \tag{2}$$

(4) Dimensionality reduction: Combining features that are highly similar can increase the efficiency of the model. PCA uses the same features obtained after feature selection to reduce dimensionality, and its algorithm is as follows:

(a) For the input data set, take two features as an example, $x = \begin{Bmatrix} x_1, x_2, x_3, \ldots, x_n \\ y_1, y_2, y_3, \ldots, y_n \end{Bmatrix}$ includes $\left\{x_1, x_2, x_3, \ldots, x_n\right\}$ of feature $x$ and $\left\{y_1, y_2, y_3, \ldots, y_n\right\}$ of feature $y$, and each feature has *n* data. Multiple features should follow the same rule.

(b) Centralize each feature, in other words, each sample of each feature subtracts its respective average value, and the average value is calculated as $\hat{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

(c) First, calculate the covariance, as shown in formula (3), where $x_i$ denotes the $i$ th data of feature $x$; $y_i$ is $i$ th data of feature $y$; and there are a total of *n* data, thus, $\mu_x$ and $\mu_y$ represent the mean of features $x$ and $y$, respectively. The covariance matrix $A = \frac{1}{n}XX^T$ is formed by the covariance, where the calculation of $X$ is shown in formula (4).

$$cov\left(x, y\right) = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \mu_x\right)\left(y_i - \mu_y\right) \tag{3}$$

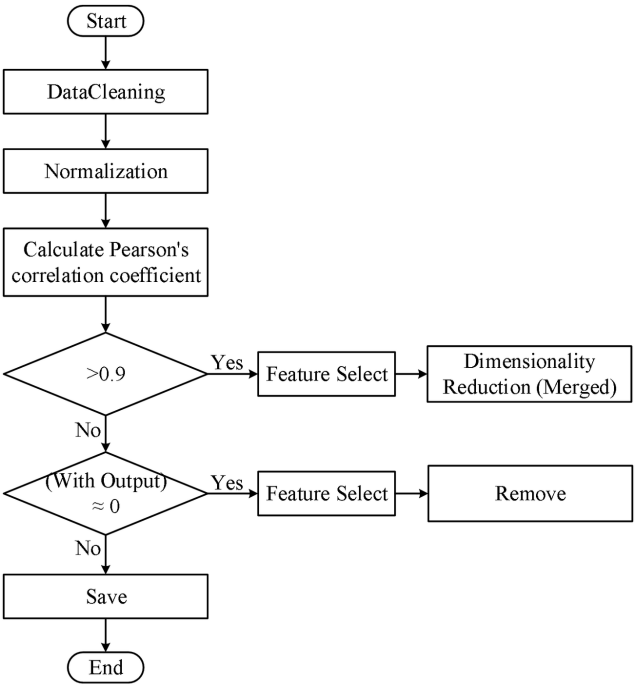
$$X = \begin{bmatrix} cov\left(x, x\right) & cov\left(x, y\right) \\ cov\left(y, x\right) & cov\left(y, y\right) \end{bmatrix} \tag{4}$$

(d) Use the relationship formula $\left(A - \lambda\right)v = 0$ between the eigenvalue and eigenvector to find eigenvalue $\lambda$ and eigenvector $v$ of the covariance matrix.

(e) Sort the eigenvalues from large to small, and select the largest *k* values among them. Then, use the corresponding *k* eigenvectors as row vectors to form eigenvector matrix *P*.

(f) Convert the data to a new space, as constructed by the *k* feature vectors, to produce dimensionality reduction data $z$, i.e., $z = Px$.

**Figure 6. Data pre-processing**



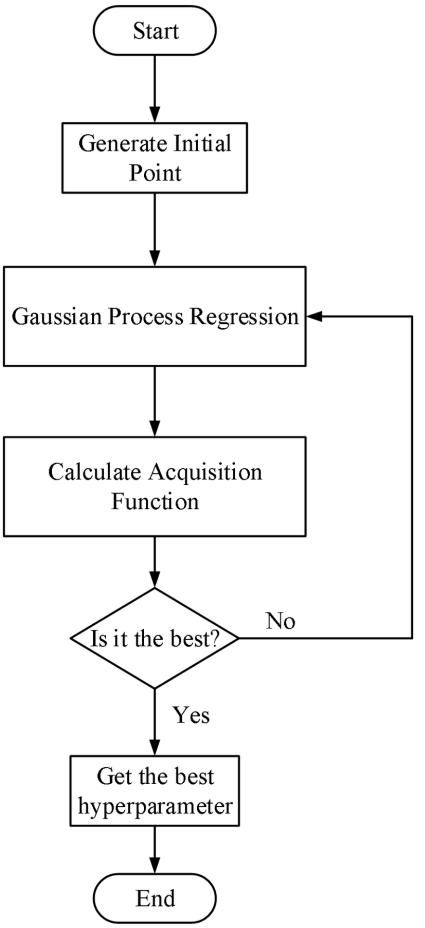

This study used Bayesian optimization to optimize the LSTM model, as shown in Figure 7. The probability model is established to determine the global optimal value of the model. Its mathematical expression is $\mathrm{X}^{*} = \arg_{\mathrm{x}\in\mathrm{S}}\mathrm{minf}\left(\mathrm{x}\right)$, where $S$ is the range to be adjusted by the model hyperparameters; $x$ is a set of hyperparameters in ; and function $f\left(x\right)$ represents the LSTM model. Different hyperparameter combination model functions yield different results.

(1) First, as $f\left(x\right)$ is an unknown situation, it is necessary to assume a prior function. Gaussian process regression is used, which assumes that the Gaussian distribution is satisfied by $f\left(x\right)$, and the prior model is corrected with historical data to make it closer to actual distribution.
(2) After the revised function is obtained in the first step, the acquisition function should be used to select the next $x$ from the revised prior function. The Expected Improvement (EI) method, as proposed by Mockus et al. (1978), is employed as the acquisition function.
(3) After a new $x$ is selected, judge whether $f\left(x\right)$ is the optimal solution. By selecting the number of iterations, we can control how many times we need to find a new $x$, and repeat the first and second steps to find them. The more iterations, the more accurate the result.

### *3.3.3 Estimate of Abnormality*

In some cases, equipment anomalies cannot be directly judged; for example, the abnormal range suddenly returning to the normal range for an instant cannot be judged as normal. The reason for this issue is that the equipment is of the sequential operation mode, and the values before and after must be considered to determine whether the current case is an instantaneous return or a stable return. Therefore, the degree of anomaly is calculated by multiplying the past error value with the exponentially

**Figure 7. Bayesian optimization**




decreasing threshold, and adjusting the time period to be considered according to the threshold. As expressed in formula (5), $value_m$ is the current abnormal value, $RMSE_1$ represents the current root mean square error value, $RMSE_2$ is the root mean square error value at the previous time point, $RMSE_n$ is the root mean square error value at the previous *n* time points, $\alpha$ is the threshold control parameter, and *N* denotes how many previous values need to be taken. The $\alpha$ value can be controlled by adjusting *N*. The higher the power and further back of $\left(1-\alpha\right)$, the longer the distance is ignored. Finally, the number of anomalies in all production cycles is calculated. Herein, the production cycle is set as 1000, as shown in formula (6), where $\varepsilon$ is the anomaly threshold. After experiments, the anomaly threshold is set to 2.5 in this study.

$$value_j = \alpha \times \sum_{i=1}^{n}\left[\left(1-\alpha\right)^{i-1} \times RMSE_i\right] = \alpha \times \left[RMSE_1 + \left(1-\alpha\right) \times RMSE_2 + \cdots + \left(1-\alpha\right)^{n-1} \times RMSE_n\right], \alpha = \frac{2}{N+1}, N \in \mathbb{N} \quad (5)$$

$$\text{Anomaly Score} = \sum\nolimits_{m=1}^{n=1000} f\left(value_j, \varepsilon\right), f\left(value_j, \varepsilon\right) = \begin{cases} 1; value_i \geq \varepsilon \\ 0; value_i < \varepsilon \end{cases} \quad (6)$$

## 4. RESULTS AND DISCUSSIONS

### 4.1 Supply and Sales Anomaly Detection Results

As complete food safety data are difficult to acquire, as this requires long-term cooperation of food manufacturers, this study used the same data set used by Lopez-Rojas et al. (2014) as the simulation data. Vaughan (2018) used the same data set for his research, as shown in Table 2.

This study first selected six machine learning models to compare five cases, respectively. Due to the imbalance between the ratio of positive and negative samples of the test data representing the actual situation in anomaly detection, effective evaluation with the accuracy rate as an evaluation indicator cannot be carried out. Hence, recall was taken as the evaluation indication for comparison, as shown in Table 3. The results indicate that random forest was better. Finally, this study conducted a comparison with the F1-Score as the comprehensive evaluation indicator, as shown in Table 4. The results show that SVM (Zhu et al. 2015; Zhu et al. 2016) and Naive Bayes performed poorly, and the

**Table 2. Supply and sales data**

| | Symbol | Feature | Description |
|---|---|---|---|
| Input | $S_1$ | step | The number of days of transaction history, the range of this data set is 180 days, and the categories include 0~179. |
| | $S_2$ | customer_id | This represents the customer ID, which does not contain any sensitive information about the customer, and the categories include “C1093826151”, “C2054744914”, etc. |
| | $S_3$ | annualrevenue | The customer’s annual operating revenue, this category includes “0” to “6” revenue from low to high, and “U” denotes unknown. |
| | $S_4$ | phase | The stage in the supply chain where the customer is, the categories include “M” manufacturing, “S” supply, “L” logistics, and “U” unknown. |
| | $S_5$ | merchant_id | Represents the merchant ID and does not contain any sensitive data. The categories include “M348934600”, “M1823072687”, etc. |
| | $S_6$ | zip_code_ori | Represents the commodity source code, but all commodity source codes are the same. Therefore, as it no impact on model training, this study filtered out this feature. |
| | $S_7$ | zip_merchant | Represents the merchant code, which is the same as the product source code, thus, this feature was filtered out. |
| | $S_8$ | category | Represents the category of the purchased product. The categories include “es_Meat” for meat, “es_Cereal” for cereals, and other categories. |
| | $S_9$ | amount | Represents the amount of each transaction in US dollars. |
| Output | $Y_1$ | fraud | Represents whether the transaction is abnormal. If the category contains “1”, it means the transaction is abnormal, and “0” means the transaction is normal. |

**Table 3. Comparison of machine learning models (Recall)**

| | KNN | Logistic Regression | Naive Bayes | Decision Tree | SVM | Random Forest |
|---|---|---|---|---|---|---|
| Case 1 | 0.60556 | 0.52292 | 0.74653 | 0.71319 | 0.71667 | 0.76250 |
| Case 2 | 0.58333 | 0.54097 | 0.72500 | 0.69444 | 0.41042 | 0.75208 |
| Case 3 | 0.60486 | 0.54167 | 0.74653 | 0.70139 | 0.25486 | 0.74931 |
| Case 4 | 0.60486 | 0.54236 | 0.72431 | 0.70694 | 0.46667 | 0.74861 |
| Case 5 | 0.58889 | 0.54375 | 0.73264 | 0.68611 | 0.22153 | 0.73889 |
| Average Recall | 0.59750 | 0.53833 | 0.73500 | 0.70041 | 0.41403 | 0.75027 |

**Table 4. Comparison of machine learning models (F1-Score)**

| | KNN | Logistic Regression | Naive Bayes | Decision Tree | SVM | Random Forest |
|---|---|---|---|---|---|---|
| Case 1 | 0.69289 | 0.63145 | 0.20287 | 0.78069 | 0.49568 | 0.82837 |
| Case 2 | 0.68655 | 0.65517 | 0.19960 | 0.78064 | 0.57351 | 0.82014 |
| Case 3 | 0.70044 | 0.66130 | 0.20368 | 0.77872 | 0.40308 | 0.81898 |
| Case 4 | 0.69792 | 0.65685 | 0.19613 | 0.77268 | 0.62222 | 0.81205 |
| Case 5 | 0.68831 | 0.64524 | 0.20416 | 0.77097 | 0.35944 | 0.80667 |
| Average F1-Score | 0.69322 | 0.65000 | 0.20128 | 0.77674 | 0.49078 | 0.81724 |

random forest method was the best. The reason may be that the Naive Bayes classification method was based on the independence of features, thus, the correlation between features led to poor results. Moreover, as the negative samples of the data set may be mixed with positive samples, SVM failed to have better segment results.

Finally, all models were evaluated in five cases, as shown in Table 5 and Figure 8. As seen, the optimization degree of SMOTE-ENN was completely better than that of single SMOTE. The reason is that SMOTE sampling is grown by the neighbors of the positive sample, thus, the sample is too close to the negative sample. As a result, some positive samples and negative samples being mixed during model training, and the model fails to detect them. In addition, SMOTE-ENN can clearly train the rules regarding positive and negative samples with the help of ENN. Compared with the initial model, the recall of the final model of this study (RF+Random search+SMOTE-ENN) is highly improved.

The comparison of the results of this study and Vaughan (2018) are shown in Table 6. As the evaluation method of Vaughan (2018) is different from that of this study, the evaluation method of this study was converted to the same as that used by Vaughan (2018), including the misclassification rate, false negative rate, and false positive rate, and these evaluation methods are all the lower the better. To ensure that abnormal transactions are not missed, the normal data omission rate was used as the key indicator of the evaluation in this study; the lower the value, the higher the rate of correctly identifying anomaly data. As the final model of this research (RF+Random search+SMOTE-ENN) has a higher rate of identifying anomaly data, it is considered to be better than the research model.

## 4.2 Equipment Anomaly Detection Result

The data collected by Kirchgässner et al. (2019), namely the status data of the model collected by sensors, are shown in Table 7. The data were used as the simulation data for equipment anomaly detection in this study.

**Table 5. Overall model evaluation**

| Model Case | RF | RF+SMOTE | RF+SMOTE-ENN | RF+Random search+SMOTE | RF+Random search+SMOTE-ENN |
|---|---|---|---|---|---|
| Case 1 | 0.77083 | 0.78542 | 0.83611 | 0.83403 | 0.87083 |
| Case 2 | 0.75764 | 0.76875 | 0.82778 | 0.84310 | 0.86181 |
| Case 3 | 0.7625 | 0.78056 | 0.83542 | 0.83681 | 0.87708 |
| Case 4 | 0.76094 | 0.76181 | 0.82015 | 0.84677 | 0.86250 |
| Case 5 | 0.75715 | 0.75972 | 0.82292 | 0.83478 | 0.85972 |
| Average Recall | 0.76181 | 0.77125 | 0.82847 | 0.83909 | 0.86638 |

**Figure 8. Overall model comparison (Recall)**

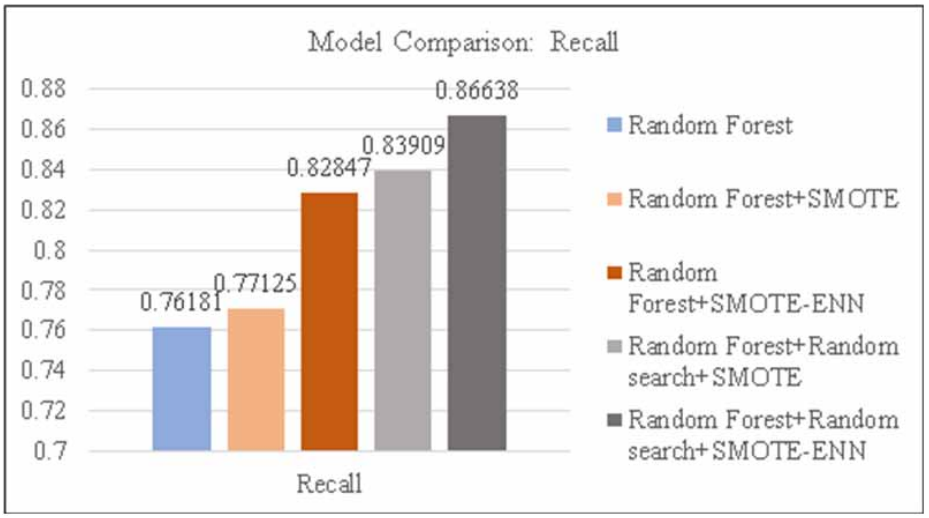



**Table 6. Compare with Vaughan (2018)**

| | misclassification rate | false negative rate | false positive rate |
|---|---|---|---|
| The final model of this research | 0.68% | 13.47%(193/1440) | 0.52%(610/117489) |
| Model 1 by Vaughan (2018) | 0.52% | 29.20%(420/1440) | 0.16%(187/117489) |
| Model 2 by Vaughan (2018) | 0.57% | 24.01%(345/1440) | 0.27%(317/117489) |

The two models were trained and tested with all cases, respectively. The results were evaluated according to the evaluation indicators, as shown in Table 8. The averages of the cases are presented in Figure 9, indicating that the final result of LSTM is better than the traditional RNN. Thus, LSTM was used as the equipment anomaly detection model.

Bayesian optimization was set to search 20 prior times, and the best hyperparameter combination was automatically searched to perform model training, as based on this combination, and compared with the unoptimized model. All cases were used to train and evaluate the optimized LSTM model, and the evaluation results are shown in Table 9. The averages of the cases are shown in Figure 10. The results confirm that the error rate of the optimized LSTM model was indeed reduced.

In order to perform anomaly detection, the anomaly data generated in this study was divided into two types of anomaly data in the normal data. The first type is transient anomaly (TA), as shown in Figure 11, meaning when the actual output torque of equipment may suddenly increase due to the current instability of the input voltage, and this sudden increase in torque affects the quality of food production. The second type is persistent anomaly (PA), as shown in Figure 14, meaning when the actual situation may be that the output torque is not in the normal range due to equipment failure. This is continuous instability, and such output causes the entire batch of food to be ruined or the production line is stopped.

The above TA and PA were directly calculated with the root mean square error, and the results are shown in Figure 12 and Figure 15. However, this high fluctuation made it impossible to judge the abnormal area according to the threshold; thus, conversion calculation of the anomaly degree estimation was required. The calculations of anomaly degree estimation in this study all considered the previous state; hence, any sudden return to an abnormality in the anomaly range can be effectively

**Table 7. Equipment data**

| | Symbol | Feature | Description |
|---|---|---|---|
| Input | $X_1$ | ambient_tem | Ambient temperature. |
| | $X_2$ | coolant | Equipment coolant temperature. |
| | $X_3$ | u_d | voltage d component. |
| | $X_4$ | u_q | voltage q component. |
| | $X_5$ | motor_speed | Motor speed. |
| | $X_6$ | i_d | Current d component. |
| | $X_7$ | i_q | Current q component. |
| | $X_8$ | pm | Equipment rotor temperature. |
| | $X_9$ | stator_yoke | Stator yoke temperature. |
| | $X_{10}$ | stator_tooth | Stator tooth temperature. |
| | $X_{11}$ | stator_winding | Stator winding temperature. |
| | $X_{12}$ | profile_id | Equipment ID. |
| Output | $Y_1$ | torque | Equipment torque. |

**Table 8. Algorithm comparison (RMSE of all cases)**

| | RNN | LSTM |
|---|---|---|
| Case 1 | 0.006927 | 0.006066 |
| Case 2 | 0.006530 | 0.006691 |
| Case 3 | 0.006895 | 0.005837 |
| Case 4 | 0.007106 | 0.005849 |
| Case 5 | 0.006423 | 0.005944 |
| Average RMSE | 0.006776 | 0.006077 |

judged. The actual anomaly range of the TA detection score was 0~100, as shown in Figure 13, and the actual anomaly range of the PA detection score was 400~800, as shown in Figure 16. It is clear that the anomaly area has a higher anomaly value; therefore, the anomaly can be judged by setting the threshold of the anomaly score. This study was based on the anomaly threshold of 2.5, and was judged as abnormal if it exceeded 2.5.

**Figure 9. Model comparison (Case average)**

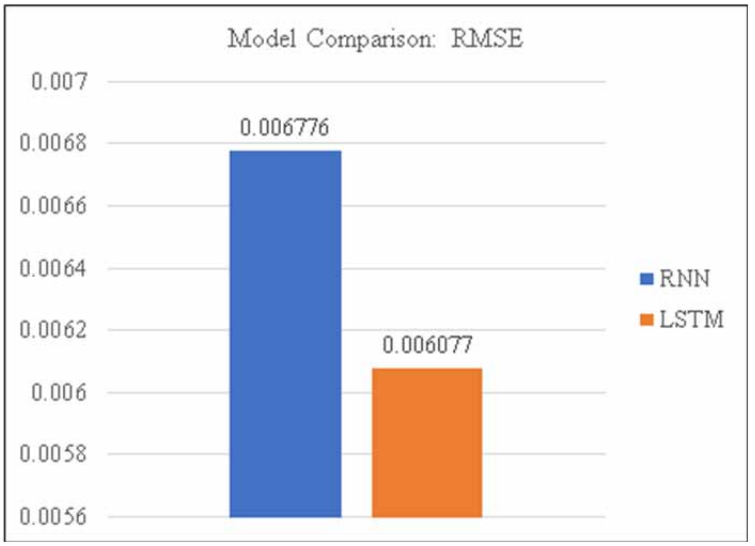



**Table 9. Hyperparameter optimization comparison**

| | LSTM | LSTM Hyperparameter Optimization |
|---|---|---|
| Case 1 | 0.005919 | 0.004001 |
| Case 2 | 0.005556 | 0.003997 |
| Case 3 | 0.004696 | 0.004115 |
| Case 4 | 0.006538 | 0.004097 |
| Case 5 | 0.005372 | 0.004354 |
| Average RMSE | 0.005616 | 0.004112 |

**Figure 10. Model comparison (Hyperparameter optimization)**

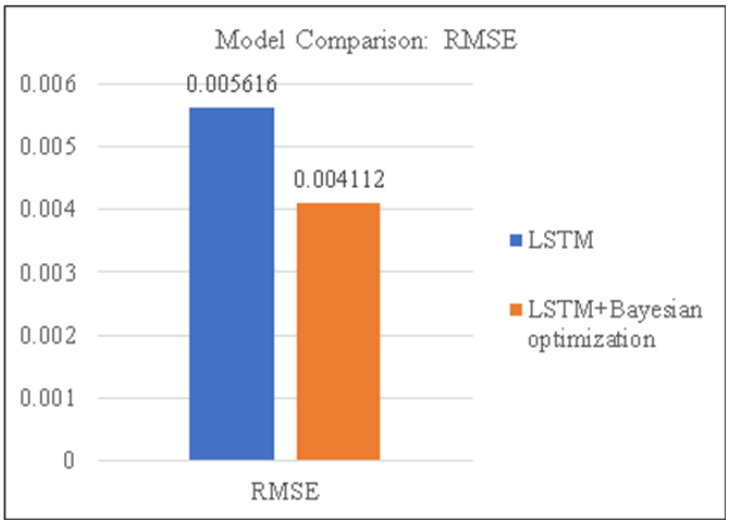


**Figure 11. Transient anomaly**

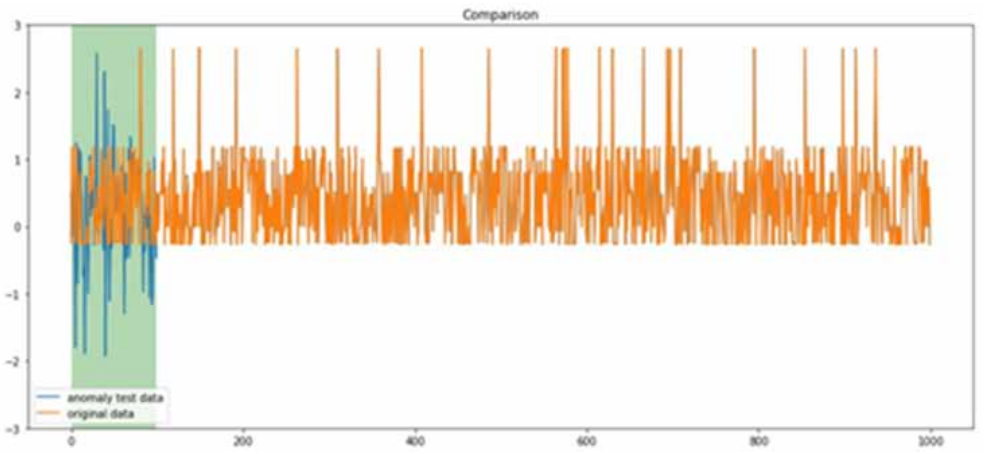



**Figure 12. Transient anomaly (RMSE)**

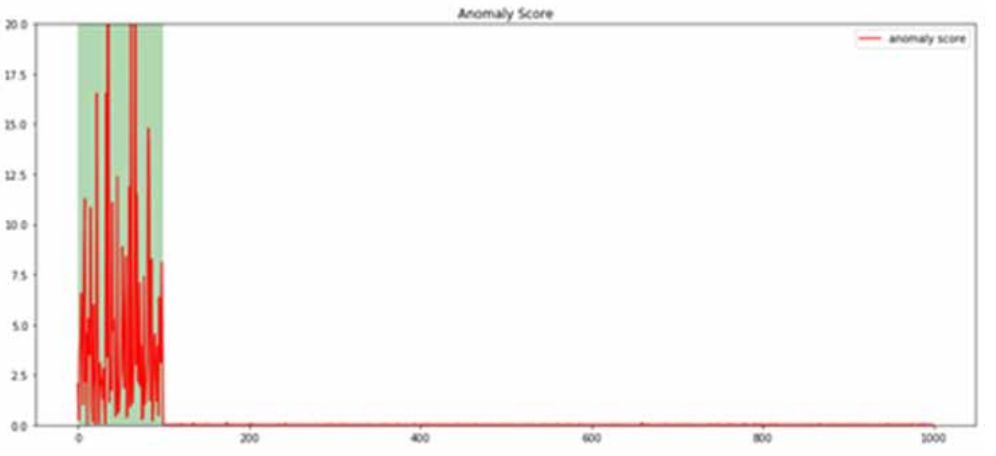



**Figure 13. Transient anomaly (Estimate of abnormality)**

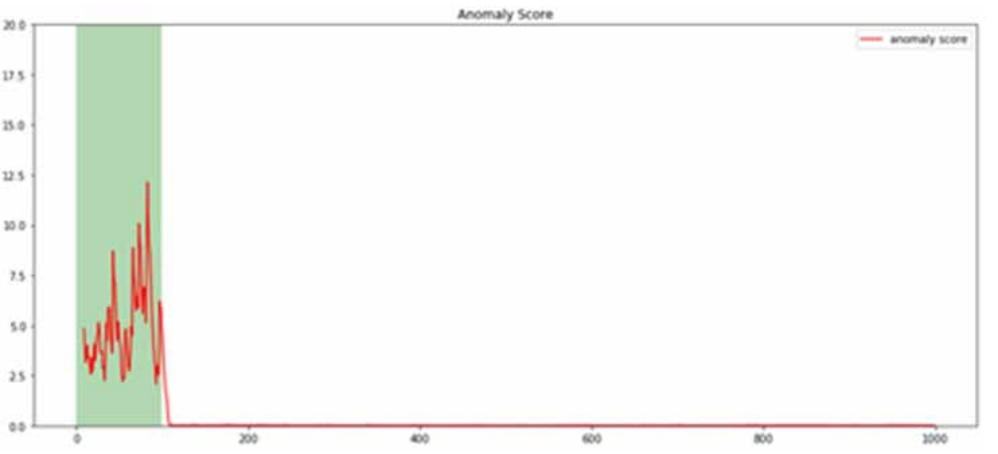


**Figure 14. Persistent anomaly**

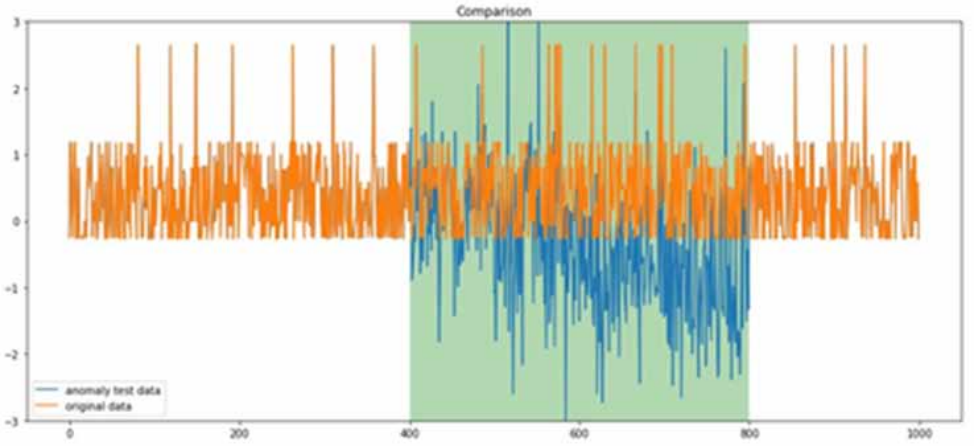



**Figure 15. Persistent anomaly (RMSE)**

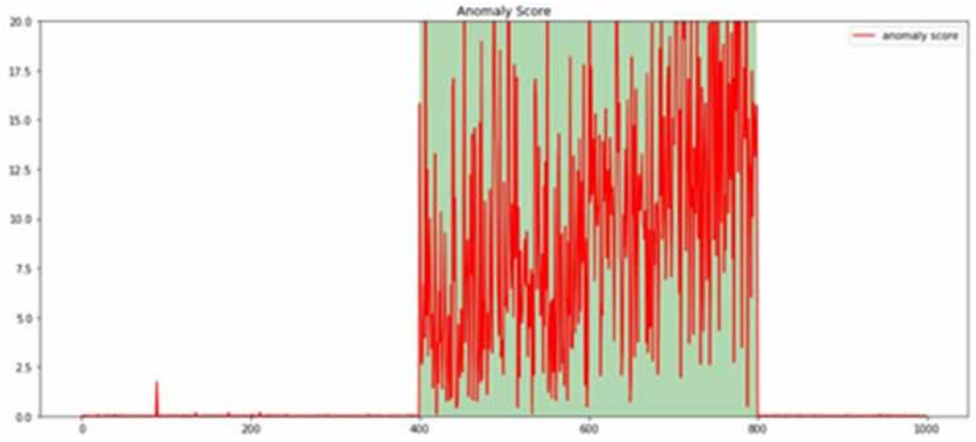
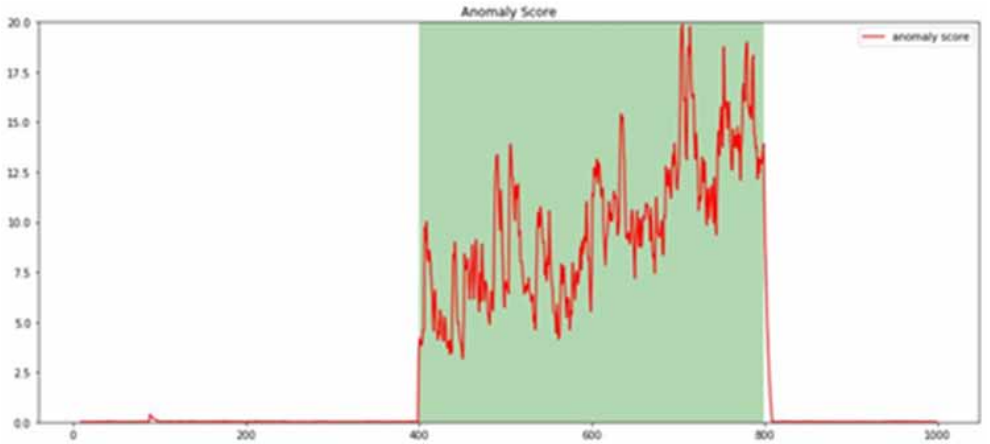



**Figure 16. Persistent anomaly (Estimate of abnormality)**



Finally, this study compared the RMSE and the estimate of abnormality, and the threshold of the two was set to 2.5 to determine whether the estimate of abnormality in this study is better than the direct RMSE calculation, as shown in Tables 10 and 11. The results indicate that the estimates of abnormality were better in both cases.

## 5. CONCLUSION

Food safety has become a concern due to the frequent occurrence of food safety incidents, thus, a set of methods is required to maintain food quality. This study developed the food safety protection system architecture to control and maintain food quality in the entire collaborative supply chain. The proposed architecture provides a complete history of food supply, sales, and production from suppliers to sellers. Consumers can view the detailed product history to confirm whether or not the food is abnormal, and at which stage the anomaly occurred.

With this food anomaly detection method, the supply, sales, and production data stored in the blockchain can be used to determine whether a batch of food is abnormal. The use of the machine learning method can be fine-tuned according to the trends, and be compared with standard specifications. Then, the supplier of this batch of food can conduct equipment anomaly detection to determine when an anomaly occurred, and at which stage of the production cycle. Moreover, the machine learning method can be applied to determine the normal mode of the equipment. By estimating the degree of an anomaly, the time period of an anomaly can be estimated, and then, action can be taken according to the result to prevent distribution or destroy food in the abnormal stage, thereby avoiding selling the foods to consumers. The proposed method features automated detection, which

**Table 10. Transient anomaly detection**

| | **Transient anomaly (RMSE)** | **Transient anomaly (Estimate of abnormality)** |
|---|---|---|
| Case 1 | 0.57 | 0.89 |
| Case 2 | 0.56 | 0.87 |
| Case 3 | 0.57 | 0.90 |
| Case 4 | 0.55 | 0.87 |
| Case 5 | 0.56 | 0.88 |
| Average recall | 0.562 | 0.882 |

**Table 11. Persistent anomaly detection**

| | **Persistent anomaly (RMSE)** | **Persistent anomaly (Estimate of abnormality)** |
|---|---|---|
| Case 1 | 0.865 | 0.98 |
| Case 2 | 0.860 | 0.979 |
| Case 3 | 0.861 | 0.975 |
| Case 4 | 0.851 | 0.973 |
| Case 5 | 0.866 | 0.977 |
| Average recall | 0.8606 | 0.9768 |

means that potential food safety problems can be identified, and abnormal batches of food can be automatically detected with high accuracy .

Since there are many factors that affect food quality, this study only analyzed supply, sales, and equipment, quality analysis including the chemical analysis of raw materials. Future studies can define food quality with fuzzy theories, in order to provide a clear explanation on the production process concerning food quality issues. In addition, this study only conducted anormaly detection; in the future, in-depth research on prevention, cause analysis, and prescription can be conducted. Due to human negligence and cost considerations, it is impossible to conduct complete data marking in the actual production process, thus, by determining potential rules, an algorithm with unsupervised learning can be used to perform anomaly detection, prediction, and prevention.

## REFERENCES

Assefa, T. T., Meuwissen, M. P., & Lansink, A. G. O. (2017). Price risk perceptions and management strategies in selected European food supply chains: An exploratory approach. *NJAS Wageningen Journal of Life Sciences*, *80*(1), 15–26. doi:10.1016/j.njas.2016.11.002

Ayoub, M. (2020). A review on Machine learning algorithms to predict daylighting inside buildings. *Solar Energy*, *202*, 249–275. doi:10.1016/j.solener.2020.03.104

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*(1), 281–305.

Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. *Decision Support Systems*, *50*(3), 602–613. doi:10.1016/j.dss.2010.08.008

Bishop, C. M. (2016). *Pattern Recognition and Machine learning*. Springer Publishing.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. doi:10.1613/jair.953

Chen, Z., Yan, Q., Han, H., Wang, S., Peng, L., Wang, L., & Yang, B. (2018). Machine learning based mobile malware detection using highly imbalanced network traffic. *Information Sciences*, *433*, 346–364. doi:10.1016/j.ins.2017.04.044

Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture*, *151*, 61–69. doi:10.1016/j.compag.2018.05.012

Creydt, M., & Fischer, M. (2019). Blockchain and more-Algorithm driven Food Traceability. *Food Control*, *105*, 45–51. doi:10.1016/j.foodcont.2019.05.019

Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2017). Credit card fraud detection: A realistic modeling and a novel learning strategy. *IEEE Transactions on Neural Networks and Learning Systems*, *29*(8), 3784–3797. PMID:28920909

Dinh, T. N., & Thai, M. T. (2018). AI and Blockchain: A disruptive integration. *Computer*, *51*(9), 48–53. doi:10.1109/MC.2018.3620971

Fraga-Lamas, P., & Fernández-Caramés, T. M. (2019). A review on Blockchain technologies for an advanced and cyber-resilient automotive industry. *IEEE Access: Practical Innovations, Open Solutions*, *7*, 17578–17598. doi:10.1109/ACCESS.2019.2895302

Freise, M., & Seuring, S. (2015). Social and environmental risk management in supply chains: A survey in the clothing industry. *Logistics Research*, *8*(1), 2. doi:10.1007/s12159-015-0121-8

Galvez, J. F., Mejuto, J. C., & Simal-Gandara, J. (2018). Future challenges on the use of Blockchain for food traceability analysis. *Trends in Analytical Chemistry*, *107*, 222–232. doi:10.1016/j.trac.2018.08.011

Gaukler, G. M., Özer, Ö., & Hausman, W. H. (2008). Order progress information: Improved dynamic emergency ordering policies. *Production and Operations Management*, *17*(6), 599–613. doi:10.3401/poms.1080.0066

George, R. V., Harsh, H. O., Ray, P., & Babu, A. K. (2019). Food quality traceability prototype for restaurants using Blockchain and food quality data index. *Journal of Cleaner Production*, *240*, 118021. doi:10.1016/j.jclepro.2019.118021

Guan, G. F., Dong, Q. L., & Li, C. H. (2011) Risk identification and evaluation research on F-AHP evaluation based supply chain. In *18th International Conference on Industrial Engineering and Engineering Management* (pp. 1513-1517). IEEE. doi:10.1109/ICIEEM.2011.6035447

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, *154*, 72–80. doi:10.1016/j.ijpe.2014.04.018

Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, *6*(2), 107–116. doi:10.1142/S0218488598000094

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. doi:10.1162/neco.1997.9.8.1735 PMID:9377276

ISO, I. (2018) 22000: Food safety management systems—Requirements for any organization in the food chain. *International Standard*, 1-48.

Kamble, S. S., Gunasekaran, A., & Gawankar, S. A. (2020). Achieving sustainable performance in a data-driven agriculture supply chain: A review for research and applications. *International Journal of Production Economics*, *219*, 179–194. doi:10.1016/j.ijpe.2019.05.022

Krisztin, T. (2018). Semi-parametric spatial autoregressive models in freight generation modeling. *Transportation Research Part E, Logistics and Transportation Review*, *114*, 121–143. doi:10.1016/j.tre.2018.03.003

Kshetri, N. (2018). 1 Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management*, *39*, 80–89. doi:10.1016/j.ijinfomgt.2017.12.005

Lindemann, B., Maschler, B., Sahlab, N., & Weyrich, M. (2021). A survey on anomaly detection for technical systems using LSTM networks. *Computers in Industry*, *131*, 103498. doi:10.1016/j.compind.2021.103498

Makhdoom, I., Abolhasan, M., Abbas, H., & Ni, W. (2019). Blockchain's adoption in IoT: The challenges, and a way forward. *Journal of Network and Computer Applications*, *125*, 251–279. doi:10.1016/j.jnca.2018.10.019

Mehra, M., Saxena, S., Sankaranarayanan, S., Tom, R. J., & Veeramanikandan, M. (2018). IoT based hydroponics system using Deep Neural Networks. *Computers and Electronics in Agriculture*, *155*, 473–486. doi:10.1016/j.compag.2018.10.015

Mikolov, T., Karafiát, M., Burget, L., Černocký, J., & Khudanpur, S. (2010) Recurrent neural network based language model. In *Eleventh annual conference of the international speech communication association*, 1045-1048. ISCA.

Mohanta, B. K., Jena, D., Satapathy, U., & Patnaik, S. (2020) Survey on IoT Security: Challenges and Solution using Machine learning. *Artificial Intelligence and Blockchain Technology. Internet of Things,* 100227.

Morgan, T. R., Richey, R. G. Jr, & Ellinger, A. E. (2018). Supplier transparency: Scale development and validation. *International Journal of Logistics Management*, *29*(3), 959–984. doi:10.1108/IJLM-01-2017-0018

Nabavi-Pelesaraei, A., Rafiee, S., Mohtasebi, S. S., Hosseinzadeh-Bandbafha, H., & Chau, K. W. (2017). Energy consumption enhancement and environmental life cycle assessment in paddy production using optimization techniques. *Journal of Cleaner Production*, *162*, 571–586. doi:10.1016/j.jclepro.2017.06.071

Nakamoto, S., & Bitcoin, A. (2008) *A peer-to-peer electronic cash system*. Bitcoin. https://bitcoin.org/bitcoin.pdf

Ndraha, N., Hsiao, H. I., Vlajic, J., Yang, M. F., & Lin, H. T. V. (2018). Time-temperature abuse in the food cold chain: Review of issues, challenges, and recommendations. *Food Control*, *89*, 12–21. doi:10.1016/j.foodcont.2018.01.027

Nguyen, H. D., Tran, K. P., Thomassey, S., & Hamad, M. (2021). Forecasting and Anomaly Detection approaches using LSTM and LSTM Autoencoder techniques with the applications in supply chain management. *International Journal of Information Management*, *57*, 102282. doi:10.1016/j.ijinfomgt.2020.102282

Nunes, M. C. N., Emond, J. P., Rauth, M., Dea, S., & Chau, K. V. (2009). Environmental conditions encountered during typical consumer retail display affect fruit and vegetable quality and waste. *Postharvest Biology and Technology*, *51*(2), 232–241. doi:10.1016/j.postharvbio.2008.07.016

Óskarsdóttir, K., & Oddsson, G. V. (2019). Towards a decision support framework for technologies used in cold supply chain traceability. *Journal of Food Engineering*, *240*, 153–159. doi:10.1016/j.jfoodeng.2018.07.013

Pang, G., Shen, C., Cao, L., & Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. [CSUR]. *ACM Computing Surveys*, *54*(2), 1–38. doi:10.1145/3439950

Resende-Filho, M. A., & Hurley, T. M. (2012). Information asymmetry and traceability incentives for food safety. *International Journal of Production Economics*, *139*(2), 596–603. doi:10.1016/j.ijpe.2012.05.034

Sharma, R., Kamble, S. S., Gunasekaran, A., Kumar, V., & Kumar, A. (2020). A systematic literature review on Machine learning applications for sustainable agriculture supply chain performance. *Computers & Operations Research*, *119*, 104926. doi:10.1016/j.cor.2020.104926

Shirani, M., & Demichela, M. (2015). Integration of FMEA and human factor in the food chain risk assessment. International Journal of Social, Behavioural, Educational, Economic. *Business and Industrial Engineering*, *12*, 4103–4106.

Sun, S., & Wang, X. (2019). Promoting traceability for food supply chain with certification. *Journal of Cleaner Production*, *217*, 658–665. doi:10.1016/j.jclepro.2019.01.296

Wang, J., Li, M., He, Y., Li, H., Xiao, K., & Wang, C. (2018). A Blockchain based privacy-preserving incentive mechanism in crowdsensing applications. *IEEE Access: Practical Innovations, Open Solutions*, *6*, 17545–17556. doi:10.1109/ACCESS.2018.2805837

Wang, J., & Yue, H. (2017). Food safety pre-warning system based on data mining for a sustainable food supply chain. *Food Control*, *73*, 223–229. doi:10.1016/j.foodcont.2016.09.048

Wilson, D. L. (1972). Asymptotic properties of nearest neighbor rules using edited data. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-2*(3), 408–421. doi:10.1109/TSMC.1972.4309137

Wu, J., Chen, X. Y., Zhang, H., Xiong, L. D., Lei, H., & Deng, S. H. (2019). Hyperparameter optimization for Machine learning models based on bayesian optimization. *Journal of Electronic Science and Technology*, *17*(1), 26–40.

Xu, X., Cao, D., Zhou, Y., & Gao, J. (2020). Application of neural network algorithm in fault diagnosis of mechanical intelligence. *Mechanical Systems and Signal Processing*, *141*, 106625. doi:10.1016/j.ymssp.2020.106625

Zhu, F., Yang, J., Gao, C., Xu, S., Ye, N., & Yin, T. (2016). A weighted one-class support vector machine. *Neurocomputing*, *189*, 1–10. doi:10.1016/j.neucom.2015.10.097

Zhu, F., Yang, J., Xu, S., Gao, C., Ye, N., & Yin, T. (2016). Relative density degree induced boundary detection for one-class SVM. *Soft Computing*, *20*(11), 4473–4485. doi:10.1007/s00500-015-1757-7

*Yuh-Min Chen is currently the Distinguished Professor of Department of Computer Science and Information Engineering, and Institute of Manufacturing Information and Systems, College of Electric Engineering and Computer Science, National Cheng Kung University, Taiwan, ROC. His research has been focused on the applications of knowledge engineering and management in virtual enterprising, including areas of Collaborative Engineering Knowledge Management, Semantic Content Management, Semantic-based Image Management, Web-based Business Intelligence, and KM (Knowledge Management) -based e-Learning.*

*Tsung-Yi Chen is currently a Professor of Department of Information Management, and Institute of Information Management, College of Science and Technology, Nanhua University, Taiwan, ROC. He gained his Ph.D. and M.S. degrees from Institute of Manufacturing Engineering, National Cheng Kung University, Taiwan, ROC in 2006 and 2001 respectively. His research interests include Enterprise System and Information Integration, Knowledge Access Control and Sharing for Virtual Enterprises, Security in Social Network and Virtual Community, Electronic Commerce Business Model Innovation, Advanced Network Marketing Technology, and Cloud-based Enterprise Information System. He has published his research results in various journals.*

*Jyun-Sian Li is currently an Engineer of Delta Electronics, Taiwan, ROC. He gained his M.S. degree from Institute of Manufacturing Information and Systems Manufacturing Engineering, National Cheng Kung University, Taiwan, ROC in 2020.*