Review Article
Food Science and Technology

ISSN 0101-2061 (Print)
ISSN 1678-457X (Online)

DOI: https://doi.org/10.1590/fst.42021

Study on food safety risk based on LightGBM model: a review

Wang JING[1*], Bi QIAN[2], Li YANNIAN[3]

Abstract

Accurately detecting risk points is crucial to food safety risk assessment and prewarning in food safety risk management because it helps solve food safety problems at their source. With the advancement of informationization in the food industry, a vast quantity of food safety data generated throughout sample inspection, transportation, storage, food processing, and raw material production has become urgently necessary to develop and use. Nevertheless, the existing food safety risk warning system has several flaws, including a high personnel cost, a low data utilization rate, and a crude risk measurement system. As a result, we described the data attributes for further analysis and sorted the food safety data in this study. In the meantime, to fully exploit the high dimension and the data's large amount, a mixture of fuzzy hierarchy partition and prior risk probability could be used to calculate fuzzy comprehensive risk values depending on multiple traits as the predicted outcome of a predictive model which can forecast and confirm risk levels, created with the use of a light gradient boosting machine (LightGBM) and skilled adjustment procedures. Finally, the outcomes of the multiple methods are compared using the same training and test data in order to verify the efficiency of LightGBM. The risk analysis results presented in this study, including attribute importance distribution and the risk values, can be useful to decision-makers.

Keywords: risk points; predictive model; risk management; fuzzy comprehensive risk values.

Practical Application: Food safety risk management.

1 Introduction

Food safety is a global issue that affects public health, economic development, and human social stability (Molajou et al., 2021a; Molajou et al., 2021b). In recent years, there have been many public health incidents caused by food safety issues at home and abroad, such as "lead crabs" and "parasitic kimchi" in South Korea, "horse beef" in Europe, and "plasticizer health products" and "sulphur ginger" in China (Aiyar & Pingali, 2020; Bouzembrak et al., 2019; Deng et al., 2021; Fung et al., 2018). In order to improve the level of safety supervision and control in the food industry, a series of legal documents have been enacted around the world to support risk control in the food industry. The EU has completed the enactment of the EU Food Law in two phases: market-oriented and food safety-oriented, and further established the European Food Safety Authority (EFSA) to cover all aspects of risk assessment in the food supply chain (Authority, 2011; Klintman & Kronsell, 2010; Merten et al., 2011; Portier et al., 2016). In China, the Food Safety Law was enacted in 2009 to replace the Food Sanitation Law, placing greater emphasis on the need for food risk assessment (Chen et al., 2015a; Gale & Buzby, 2009; Jia & Jukes, 2013).

In this study, in order to take full advantages of the data features of large quantity and high dimension, the combination of prior risk probability and fuzzy hierarchy partition was employed to calculate fuzzy comprehensive risk values based on various attributes for use as the expected output of a predictive model that can predict and validate risk values, generated using light gradient boosting machine (LightGBM) combined with experts' modification operations. Finally, data on meat products and aquatic products were used to illustrate how to use this method, and its superiority and reasonability were validated.

2 Food safety data sources and characteristics

The sources of food safety data can be broadly summarized into three areas: 1) static data, 2) dynamic data, and 3) expert experience data (McMeekin et al., 2006; Yusianto & Hardjomidjojo, 2019). Dynamic data or transactional data is information that is periodically updated, meaning it changes asynchronously over time as new information becomes available. Data that is not dynamic is considered either static (unchanging) or persistent, which is data that is infrequently accessed and not likely to be modified (Chen et al., 2014). Static data refers to data that will not change over a period of time once defined, such as standard data in-laws and regulations (passing standard line, minimum, maximum detection limit, test basis, judgment basis, etc.), information data of sales and production enterprises (enterprise-scale, establishment years, major food categories, production and sales areas, procurement locations, etc.), etc. The data are mostly available in local industry and commerce (Salmon et al., 2012). Most of these data exist in the enterprise registration information records of local industrial and commercial

Received 10 Jun., 2021
Accepted 16 Nov., 2021
[1] *School of Humanities, Southwest Jiaotong University, Chengdu, China*
[2] *School of Electrical Engineering, Sichuan University, Chengdu, China*
[3] *Qinghai Provincial Party School of the CPC, Xining, China*
*Corresponding author: wangjing12021@outlook.com

bureaus. The annual inspection data and sampling data (food category, test items, detected content, production time, sampling time, etc.) generated by the enterprise food safety system testing (generation, processing, circulation, consumption and other aspects of routine inspection records and violation records, etc.) and routine food sampling are affected by time and geographical factors, and the inspection results and sampling records are dynamically accumulated and updated (Arpanutud et al., 2009).

Data are collected and entered through local food safety monitoring and management systems and stored in the corresponding databases for retrieval and query (Lam et al., 2013; Wu & Chen, 2018). In contrast to the first two types of data, the data given by experts based on experience or literature research, such as the probability of occurrence of a contaminant in a particular type of food, the metric value of the contamination indicator, the risk level of the item, etc., can be both stable over time and dynamically adjusted to different application contexts (Hammouri et al., 2015; Sierra-Soler et al., 2015). Food safety-related data involves food information, information of production enterprises, information of sales enterprises, information of inspection agencies, national laws and regulations, expert experience indicators, inspection standards of each contaminant, etc. Their attribute types can be broadly summarized as discrete character-based attributes (enterprise size, production province, inspection items, etc.), discrete numerical attributes (production date, sampling date, sample status, etc.), and continuous numerical attributes (enterprise turnover, detected content, etc.) (Jia & Jukes, 2013; Zhong et al., 2020). Many attribute categories, attribute value format is chaotic so that the data has the characteristics of high dimensionality and high complexity. Therefore, the data mostly show an indistinguishable linear state. The distribution pattern is hidden in incomplete data, missing data, incorrectly entered data and other noise interference, which increases the difficulty of risk analysis (Deng et al., 2021; Donaghy et al., 2021; Savelli & Mateus, 2020).

3 Progress of domestic and international research on food safety risk assessment

The adverse effects of food safety events and the nature of food safety data have led to the enactment of national laws and regulations and increasing demand for food safety risk warnings (Li et al., 2020).

Effective risk warning models can be used to extract a priori knowledge to establish patterns and analyse risk factors, risk levels, or predict risk values, and are important for governments to rationally allocate limited resources, correctly identify risk points, and address food safety issues at the source (Viscusi, 1988). There has been a great deal of research and application by domestic and foreign scholars. The Delphi method is a method for incorporating the opinions of a wide range of experts from different regions and fields and involves repeating multiple rounds of feedback and revising subsequent questionnaires based on intermediate feedback (Pérez-Castellanos, 2004). With three rounds of expert feedback based on the importance of safety issues, the Delphi method was used in the Brazilian Food Trade Risk Assessment Tool (Auad et al., 2018; Ribeiro & Quintanilla, 2015; Sossa et al., 2019). The final assessment tool consisted of 39 risk items, including additions and refinements to the original list of factors by experts.

However, all the evaluation indicators in this method, from input attribute indicators to output indicators, are artificially determined by experts, which is less efficient and has high labour cost, and cannot meet the demand of timeliness of risk warning system. The analytic hierarchy process (AHP) deals with complex multi-objective decision problems by establishing a three-level structure of objective, criterion, and solution levels, so it can be used to calculate the weights of attribute indicators and classify the evaluation levels, and then has the functions of attribute reduction and indicator denoising (Chaiyaphan & Ransikarbum, 2020; Geng et al., 2019; Sossa et al., 2019). The research used a single AHP method to develop a quantitative risk assessment model for the Indian food supply chain by initializing the comparison matrix with the indicator preferences of supply chain experts and outputting the indicator weights to identify the weak links (Ilbahar et al., 2018). Delphi and AHP methods are relatively mature risk assessment methods (Chen, 2015).

In addition, risk matrix and grey theory methods have been integrated and applied to food safety risk assessment and have made some progress (Julong, 1989). However, the above methods suffer from the shortcomings of small attribute coverage, low index accuracy, high human cost, long assessment process, inability to update dynamically, and weak adaptability, resulting in low accuracy of assessment results and lack of ability to pinpoint risk points (Kamble & Raut, 2019).

The rapid development of computer hardware and software is driving the process of informatization in the food industry. The accumulation of food safety-related data has created conditions for applying intelligent computing methods in food safety risk assessment (Cui et al., 2006; Leng et al., 2019).

According to the predicted value, the BP neural network with two implicit layers was constructed. The main sources of contamination for the following week were predicted to be "excessive pathogenic bacteria" and "veterinary drug residues," according to the predicted value.

Using the same BP neural network model, a study selected 13 attributes such as "province of the production company" and "sampling location" of the heavy metal "lead" sampling records as inputs to predict the "test results" of the records (Gao, 2021). Here, the value of the "test result" attribute is pass or fail, which is an existing attribute of the sampling record and does not require human-made labelling, which improves the accuracy of the prediction result. In addition, another study took the products of a dairy company as the object of analysis, and extracted seven influencing factors such as transportation time, temperature, season, and packaging method as rule mining attribute items, and used the association rule and the Apriori algorithm to generate a rule base, and used the support and confidence filtering rules to retain the most frequent rule combinations (Bu et al., 2020). The results are used as the process combinations that should be avoided in supply chain linkage development.

A study first used principal component analysis to filter the evaluation indicators of yak milk dregs quality and then used clustering to group the data into clusters based on the two main indicators of appearance and colour and nutritional quality, and finally used p-values to evaluate the cluster variability and derive intra-cluster patterns (Chi et al., 2021). Another research used Grey relational analysis to determine the index weights of the sampled data, developed the label's risk value, and then used Hidden Markov Method to train the model and predict the risk value (Jin et al., 2013). The AHP method, as an important method for calculating index weights and risk classification with low complexity, is usually combined with other theoretical methods to improve the accuracy and interpretability of risk assessment.

Another study first used AHP to calculate the weights of risk indicators predefined by experts and then combined with Dempster-Shafer's theory to synthesize the final risk values (Lyu et al., 2020). Similarly, after calculating the weights of each risk indicator by AHP to develop a risk value label, research applied an extreme learning machine and a deep radial basis neural network, respectively, to predict the risk value by using each indicator value as input (Mojrian et al., 2020). In addition, support vector machines have also been used for risk assessment of a small amount of sample data. Figure 1 summarizes the above-mentioned risk warning methods and their general algorithmic flow.

Under the guidance of the State Administration of Market Supervision and Administration, local food and drug supervision and management agencies have been actively engaged in regulatory, technological innovation, and technical difficulties to overcome, and have continuously strengthened the construction of food supervision information technology (Fu, 2014; Liu et al., 2019). However, compared with other countries, the long-standing food culture and unique geographical location have created the existing complex food system in China, and the diversity in food types, processing methods, packaging and storage methods, additives, and additional methods is far greater than that of foreign countries. In this regard, after a comprehensive analysis of data characteristics, in this review a combination of boundary value division hierarchy and Bayesian prior probability are used to calculate the expected output of the model and LightGBM is used with minimal training cost and good accuracy for risk value prediction (Lin & Sun, 2020; Wu, 2020). In practical applications, experts can correct the prediction results to help improve the accuracy of the output rules until the prediction model gradually approximates the optimal decision solution.

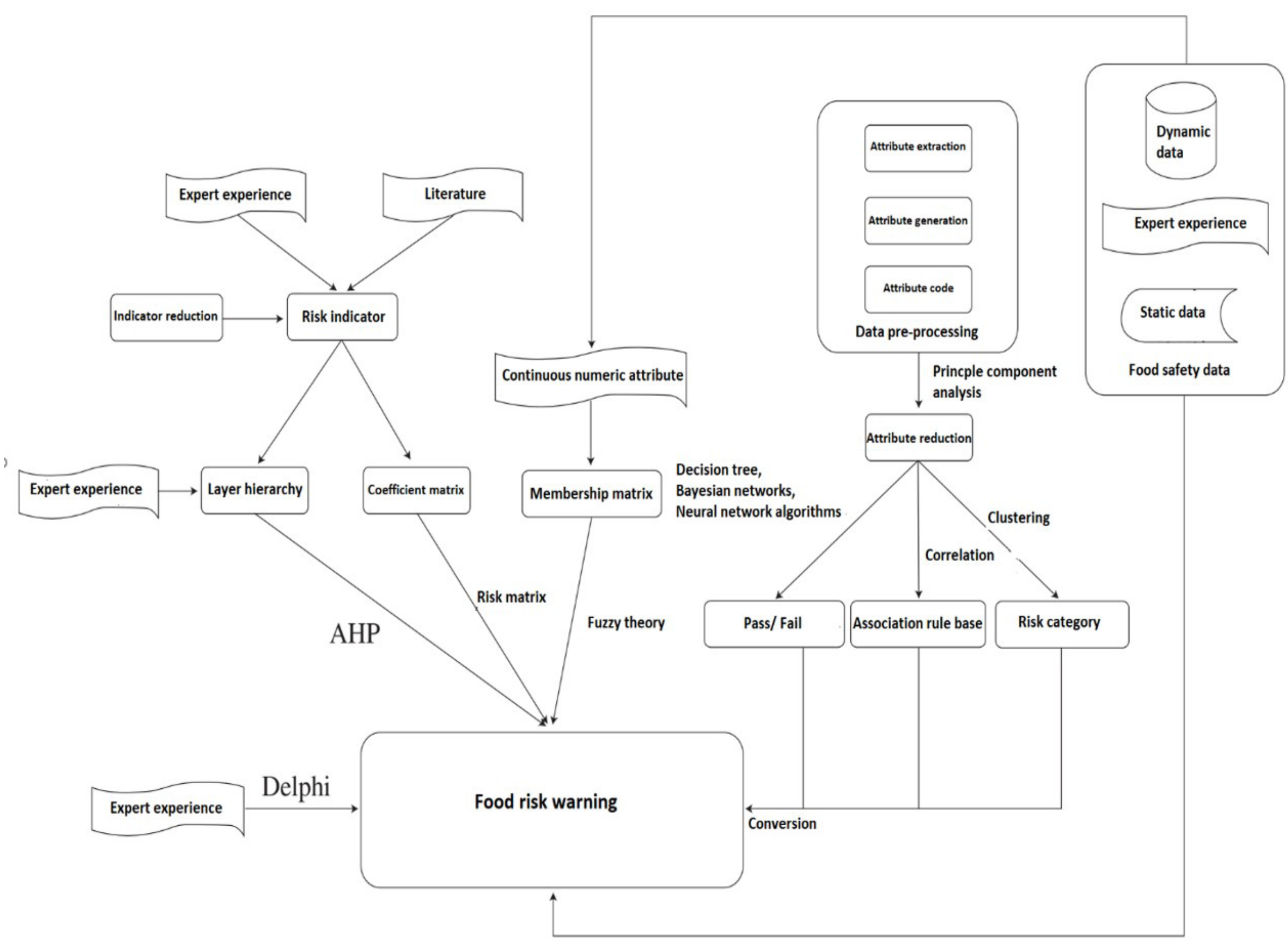

Figure 1. Classification of the previous algorithm studies.

4 Using the LightGBM model to predict the outcome

Previous studies have mostly used numerical features from sampling data as input for risk value prediction, such as the content of each test item, production sampling quarter, temperature, yield, etc. (Geng et al., 2019; Ma et al., 2020; Marvin et al., 2016; Williams et al., 2011; Zhang et al., 2018). The small number of features increases the error between the risk prediction and the actual risk value, resulting in a "false risk" contrary to the fact and even misleading decisions. The inability to consider the full range of risk point attributes causes valuable relationships and patterns in the data to be ignored. In addition, the labeling of training data only takes into account the influence of expert factors on index weights. It ignores the significant statistical information that already exists in a large amount of historical data. The labeling is done directly without a method to check the error of both, which inevitably leads to mislabeling. In order to ensure the reliability of the experimental data and reduce the risk of false warnings, the a priori risk probability of the historical data combined with the fuzzy hierarchical division method is used to calculate the specific eigenvalue weights of each attribute, and the result of the weighted sum of the discrete and continuous attribute eigenvalues and then normalization is used as the risk value. In order to learn the above-established rules, discrete attributes are processed using unique thermal coding and used for outcome prediction using the LightGBM model, with the consequent incorporation of expert intervention strategies for accuracy verification and outcome correction. As shown in Figure 2, the raw data is divided into two processing steps: on the one hand, it is used for model expected output value calculation; on the other hand, it is used for model output feature processing. The processing result features of the two steps are divided into the training set and test set, which are used for model training and testing to verify the results, respectively.

5 Gradient boosting tree

In classification problems, the decision tree uses the Gini coefficient or information gain as an indicator of how well an attribute feature distinguishes between categories, which is used to determine the splitting nodes. Cart regression trees use mean square error or exponential error as an indicator to select the best splitting point when processing regressions (Ahmad et al., 2018; Mahjoobi & Etemad-Shahidi, 2008; Su et al., 2004). In supervised learning, the expectation is to obtain a stable model with high accuracy and generalization power, but due to limitations in the amount of data and the method itself, only multiple biased models, or weakly supervised models, can be obtained. Integrated learning combines these weakly supervised models by voting to smooth out noise correction errors and obtain a strongly supervised model with better performance, i.e., the idea of "bagging." With the decision tree as a sub-model, the random forest completes the model generation by two random selections: random selection of the training set and random selection of the sub-model splitting features (Rad & Ayubirad, 2017). The random sampling with put-back ensures the independent homogeneous distribution of the sub-model training set, and the splitting by selecting some features helps prevent overfitting. It can be seen that in constructing the random forest, each weakly supervised model generation process has the same sampling priority for the data. The gradient boosting decision tree (GBDT) also uses Cart regression tree as a weak learner, but unlike the parallel construction of sub-models of RF, GBDT adopts the idea of "boosting" for the progressive construction of sub-model association and solves the problem that the loss function is a general function by fitting a negative gradient between sub-models, in order to achieve the purpose of the fastest loss reduction. Therefore, there are errors in such function estimation. Based on this, research proposed an eXtreme Gradient Boosting (XGBoost) model, which was extended into a second-order Taylor expansion with a regular term correction, which greatly improved the accuracy of the model (Chen et al., 2015b).

6 LightGBM principle and advantage analysis

In the GBDT-like model mentioned above, the best model is obtained by tuning in the function space using the gradient descent method. However, in constructing sub-models, each feature has to traverse all the sample points to select the optimal segmentation point, which is a very time-consuming operation. In this regard, the LightGBM model was proposed by Sun et al. (2020). He uses gradient-based one-sided sampling on the training data and mutually exclusive feature bundling on the features to improve the learner's training speed and generalization ability, respectively. Each time the weak learner is updated, one-sided sampling compresses the training data set and reduces the computational effort without changing the distribution of feature values and losing accuracy. In addition, sampling increases the diversity of the weak learner, which in turn improves the generalization ability. However, the features of high-dimensional data are often sparse. This results in a large number of mutually exclusive features in the data, i.e., there are usually no non-zero values for a record at the same time, as in the case of One-Hot encoded features. Therefore, LightGBM bundles mutually exclusive features to improve operational efficiency. There are two issues involved: which features should be tied together and how they should be tied. By constructing a weighted undirected graph, LightGBM models the construction of a feature set as a graph colouring problem, using a greedy-like algorithm to obtain the result with a complexity of $O(\#feature^2)$. Although the complexity is higher when dealing with high-dimensional features, it only needs to be processed once. The algorithm uses a partitioned histogram to bundle the mutually exclusive features in the feature set regarding the bundling method. The mutually exclusive features are combined by recording the number of blocks per feature and shifting the defined range (Al Daoud, 2019; Ke et al., 2017).

In this way, the weak learner finds the segmentation nodes by simply iterating through all the block values to determine the segmentation point, with an iterative complexity of $O(\#blocks * feature^2)$.

Although the segmentation points found by discretizing the feature values are not as accurate as the original feature values, the decision tree itself is a weak learner. The error caused by bin segmentation serves as a regular term.

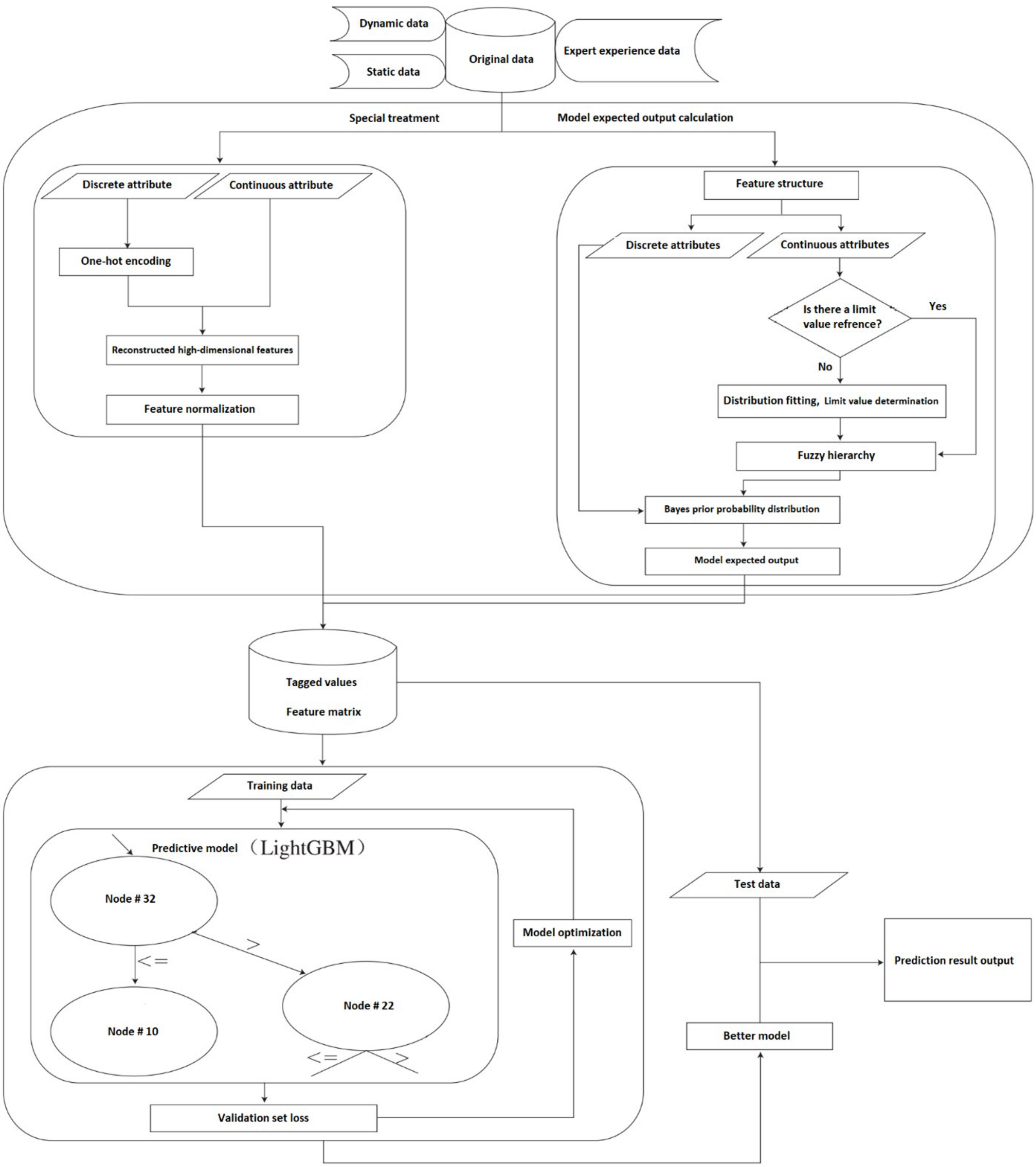

Figure 2. Flow chart of the algorithm using LightGBM.

7 Comparison of prediction results by models

To verify the effectiveness of LightGBM, the results are compared between the different methods using the same training data and test data. Similarly, the training data is still partitioned by 20% to obtain a validation set adjusted to the super-reference of the other methods. The comparison methods include BP neural networks, RBF networks, random forests, the usual GBDT, XGBoost, and LightGBM models, and Table 1 shows the results of the tuned parameters, the minimum validation set loss during tuning, and the prediction error on the same test set for the current dataset size. For all other parameters, the default values are used.

As a result, a combination of fuzzy hierarchy partition and prior risk probability could be used to calculate fuzzy comprehensive

Table 1. Parameters of comparative methods.

Method	Parameter Setting	Minimum validation set	Test set
		Loss	loss
Backpropagation (BP)	Single hidden layer; 10 hidden layer nodes; optimization algorithm Adam; activation function "Relu"; number of iterations 2000	0.1051	0.1397
Radial Basis Function (RBF)	170 hidden layer nodes; randomly selected centroids; range radius β = 2.0; optimization algorithm Adam; number of iterations 2000	0.2050	0.2087
Random Forest (RF)	max_feature=None;n_estimators=1000;min_sample_leaf=100	0.0658	0.0752
General Gradient Boosted Decision Trees (GBDT)	learning_rate=0.01;n_estimators=1000; max_depth=-1;max_feature=None	0.0612	0.0723
eXtreme Gradient Boosting (XGBoost)	learning_rate=0.01;n_estimators=1200;max_depth=9	0.0528	0.0698
Light Gradient Boosting Machine (LightGBM)	{learning_rate=0.01;max_bin=200; max_depth=12;num_leaves=21;n_estimators=900}	0.0524	0.0688

risk values based on multiple traits as the predicted outcome of a predictive model that can forecast and confirm risk levels, created with the use of a LightGBM and skilled adjustment procedures, to fully exploit the high dimension and large amount of data. Finally, the results of the various techniques are compared using the same training and test data to ensure that LightGBM is effective. The results of this study's risk analysis, including the attribute significance distribution and risk levels, might be valuable to decision-makers.

8 Conclusion

In summary, the output of this research method consists of three parts: risk value prediction, risk analysis conclusion, and attribute value importance distribution. The risk value prediction enables the rapid calculation of more accurate risk values for newly entered data and incorporates dynamic intervention strategies for experts, breaking the limitations of fixed rules in the original method.

Based on the derived risk values, the statistically significant risk analysis results provide the experts with a reference for intervention. Finally, the importance distribution of attribute values comprehensively illustrates the contribution of continuous and discrete attributes to risk, and decision-makers can make any combination of attribute values according to the distribution law to develop more accurate risk prevention and control strategies, such as targeted sampling, high-frequency sampling, and contaminant tracing. Food safety is an important issue for people's health.

In this paper, we first summarize and analyse the food safety data and the intelligent methods used in the past.

According to the characteristics of the data and the shortcomings of the existing methods, we propose a risk value calculation rule combining a priori risk probability and fuzzy hierarchy and apply the LightGBM model combined with expert empirical intervention strategies for risk value correction and prediction. However, there are still many shortcomings in the method. In terms of data application, it is crucial to take into account the time series to obtain a more refined time-series correlation of risk patterns. In terms of model application, it is also important to address the issue of "data silos" in food safety by integrating data from various sources and coordinating training to obtain a more general model that meets actual needs.

References

Ahmad, M. W., Reynolds, J., & Rezgui, Y. (2018). Predictive modelling for solar thermal energy systems: a comparison of support vector regression, random forest, extra trees and regression trees. *Journal of Cleaner Production*, 203, 810-821. http://dx.doi.org/10.1016/j.jclepro.2018.08.207.

Aiyar, A., & Pingali, P. (2020). Pandemics and food systems-towards a proactive food safety approach to disease prevention & management. *Food Security*, 12(4), 749-756. http://dx.doi.org/10.1007/s12571-020-01074-3. PMid:32837645.

Al Daoud, E. (2019). Comparison between XGBoost, LightGBM and CatBoost using a home credit dataset. *International Journal of Computer and Information Engineering*, 13(1), 6-10.

Arpanutud, P., Keeratipibul, S., Charoensupaya, A., & Taylor, E. (2009). Factors influencing food safety management system adoption in Thai food-manufacturing firms: model development and testing. *British Food Journal*, 111(4), 364-375. http://dx.doi.org/10.1108/00070700910951506.

Auad, L. I., Ginani, V. C., Leandro, E. S., Farage, P., Nunes, A. C. S., & Zandonadi, R. P. (2018). Development of a Brazilian food truck risk assessment instrument. *International Journal of Environmental Research and Public Health*, 15(12), 2624. http://dx.doi.org/10.3390/ijerph15122624. PMid:30477105.

Authority, E. F. S. (2011). Evaluation of the FoodEx, the food classification system applied to the development of the EFSA Comprehensive European Food Consumption Database. *EFSA Journal*, 9(3), 1970.

Bouzembrak, Y., Klüche, M., Gavai, A., & Marvin, H. J. (2019). Internet of Things in food safety: literature review and a bibliometric analysis. *Trends in Food Science & Technology*, 94, 54-64. http://dx.doi.org/10.1016/j.tifs.2019.11.002.

Bu, K., Li, X., Wang, K., & Li, Y. (2020). Data analysis of public food safety cases based on Apriori. In A. Xue, & D. Xue (Eds.), *2020 Chinese Control and Decision Conference (CCDC)* (pp. 343-348). Piscataway: IEEE. http://dx.doi.org/10.1109/CCDC49329.2020.9163958.

Chaiyaphan, C., & Ransikarbum, K. (2020). Criteria analysis of food safety using the Analytic Hierarchy Process (AHP)-a case study of Thailand's fresh markets. In M. Sriariyanun, Y. S. Cheng, K. Rattanaporn, P. Yasurin, & W. Rodiahwati (Eds.), *E3S Web of Conferences: Vol. 141* (pp. 02001). Piscataway: IEEE.

Chen, F. (2015). *An investigation and evaluation of risk assessment methods in Information systems* (Master thesis). Göteborg: Chalmers University of Technology.

Chen, K., Wang, X., & Song, H. (2015a). Food safety regulatory systems in Europe and China: a study of how co-regulation can improve regulatory effectiveness. *Journal of Integrative Agriculture*, 14(11), 2203-2217. http://dx.doi.org/10.1016/S2095-3119(15)61113-3.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: a survey. *Mobile Networks and Applications*, 19(2), 171-209. http://dx.doi.org/10.1007/s11036-013-0489-0.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., & Cho, H. (2015b). *Xgboost: extreme gradient boosting. R Package Version 0.4-2* (pp. 1–4). Vienna: R Foundation for Statistical Computing.

Chi, F., Tan, Z., Gu, X., Yang, L., & Luo, Z. (2021). Bacterial community diversity of yak milk dreg collected from Nyingchi region of Tibet, China. *Lebensmittel-Wissenschaft & Technologie*, 145, 111308. http://dx.doi.org/10.1016/j.lwt.2021.111308.

Cui, L., Zhang, C., Zhang, C., & Huang, L. (2006). Exploring e-government impact on Shanghai firms' informatization process. *Electronic Markets*, 16(4), 312-328. http://dx.doi.org/10.1080/10196780600999734.

Deng, X., Cao, S., & Horn, A. L. (2021). Emerging applications of machine learning in food safety. *Annual Review of Food Science and Technology*, 12(1), 513-538. http://dx.doi.org/10.1146/annurev-food-071720-024112. PMid:33472015.

Donaghy, J. A., Danyluk, M. D., Ross, T., Krishna, B., & Farber, J. (2021). Big Data impacting dynamic food safety risk management in the food chain. *Frontiers in Microbiology*, 12, 668196. http://dx.doi.org/10.3389/fmicb.2021.668196. PMid:34093486.

Fu, W. (2014). Research of food safety supervision system in China. *Applied Mechanics and Materials*, 644(650), 5995-5998. http://dx.doi.org/10.4028/www.scientific.net/AMM.644-650.5995.

Fung, F., Wang, H.-S., & Menon, S. (2018). Food safety in the 21st century. *Biomedical Journal*, 41(2), 88-95. http://dx.doi.org/10.1016/j.bj.2018.03.003. PMid:29866604.

Gale, F., & Buzby, J. C. (2009). *Imports from China and food safety issues*. Washington: Dept. of Agriculture, Economic Research Service.

Gao, J. (2021). Performance evaluation of manufacturing collaborative logistics based on BP neural network and rough set. *Neural Computing & Applications*, 33(2), 739-754. http://dx.doi.org/10.1007/s00521-020-05099-9.

Geng, Z., Shang, D., Han, Y., & Zhong, Y. (2019). Early warning modeling and analysis based on a deep radial basis function neural network integrating an analytic hierarchy process: a case study for food safety. *Food Control*, 96, 329-342. http://dx.doi.org/10.1016/j.foodcont.2018.09.027.

Hammouri, N., Al-Qinna, M., Salahat, M., Adamowski, J., & Prasher, S. O. (2015). Community based adaptation options for climate change impacts on water resources: the case of Jordan. *Journal of Water and Land Development*, 26(1), 3-17. http://dx.doi.org/10.1515/jwld-2015-0013.

Ilbahar, E., Karaşan, A., Cebi, S., & Kahraman, C. (2018). A novel approach to risk assessment for occupational health and safety using Pythagorean fuzzy AHP & fuzzy inference system. *Safety Science*, 103, 124-136. http://dx.doi.org/10.1016/j.ssci.2017.10.025.

Jia, C., & Jukes, D. (2013). The national food safety control system of China–a systematic review. *Food Control*, 32(1), 236-245. http://dx.doi.org/10.1016/j.foodcont.2012.11.042.

Jin, X., Xu, X., Song, X., Li, Z., Wang, J., & Guo, W. (2013). Estimation of leaf water content in winter wheat using grey relational analysis–partial least squares modeling with hyperspectral data. *Agronomy Journal*, 105(5), 1385-1392. http://dx.doi.org/10.2134/agronj2013.0088.

Julong, D. (1989). Introduction to grey system theory. *Journal of Grey System*, 1, 1-24.

Kamble, S. S., & Raut, R. D. (2019). Evaluating the factors considered for procurement of raw material in food supply chain using Delphi-AHP methodology-a case study of potato chips processing company in India. *International Journal of Productivity and Quality Management*, 26(2), 176-189.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). Lightgbm: a highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30, 3146-3154.

Klintman, M., & Kronsell, A. (2010). Challenges to legitimacy in food safety governance? The case of the European Food Safety Authority (EFSA). *European Integration*, 32(3), 309-327. http://dx.doi.org/10.1080/07036331003646835.

Lam, H.-M., Remais, J., Fung, M.-C., Xu, L., & Sun, S. S.-M. (2013). Food supply and food safety issues in China. *Lancet*, 381(9882), 2044-2053. http://dx.doi.org/10.1016/S0140-6736(13)60776-X. PMid:23746904.

Leng, K., Jin, L., Shi, W., & Van Nieuwenhuyse, I. (2019). Research on agricultural products supply chain inspection system based on internet of things. *Cluster Computing*, 22, 8919-8927. http://dx.doi.org/10.1007/s10586-018-2021-6.

Li, G., Shang, X., & Liu, Q. (2020). Regional food safety testing risk analysis and early warning research. In C. Yang, Y. Pei, & J. Chang (Eds.), *Innovative computing* (pp. 1135-1142). Singapore: Springer. http://dx.doi.org/10.1007/978-981-15-5959-4_139.

Lin, X., & Sun, D.-W. (2020). Recent developments in vibrational spectroscopic techniques for tea quality and safety analyses. *Trends in Food Science & Technology*, 104, 163-176. http://dx.doi.org/10.1016/j.tifs.2020.06.009.

Liu, Z., Mutukumira, A. N., & Chen, H. (2019). Food safety governance in China: from supervision to coregulation. *Food Science & Nutrition*, 7(12), 4127-4139. http://dx.doi.org/10.1002/fsn3.1281. PMid:31890192.

Lyu, H.-M., Zhou, W.-H., Shen, S.-L., & Zhou, A.-N. (2020). Inundation risk assessment of metro system using AHP and TFN-AHP in Shenzhen. *Sustainable Cities and Society*, 56, 102103. http://dx.doi.org/10.1016/j.scs.2020.102103.

Ma, B., Han, Y., Cui, S., Geng, Z., Li, H., & Chu, C. (2020). Risk early warning and control of food safety based on an improved analytic hierarchy process integrating quality control analysis method. *Food Control*, 108, 106824. http://dx.doi.org/10.1016/j.foodcont.2019.106824.

Mahjoobi, J., & Etemad-Shahidi, A. (2008). An alternative approach for the prediction of significant wave heights based on classification and regression trees. *Applied Ocean Research*, 30(3), 172-177. http://dx.doi.org/10.1016/j.apor.2008.11.001.

Marvin, H. J., Bouzembrak, Y., Janssen, E. M., Van Der Fels-Klerx, H., Van Asselt, E. D., & Kleter, G. A. (2016). A holistic approach to food safety risks: food fraud as an example. *Food Research International*, 89(Pt 1), 463-470. http://dx.doi.org/10.1016/j.foodres.2016.08.028. PMid:28460939.

McMeekin, T. A., Baranyi, J., Bowman, J., Dalgaard, P., Kirk, M., Ross, T., Schmid, S., & Zwietering, M. H. (2006). Information systems in food safety management. *International Journal of Food Microbiology*, 112(3), 181-194. http://dx.doi.org/10.1016/j.ijfoodmicro.2006.04.048. PMid:16934895.

Merten, C., Ferrari, P., Bakker, M., Boss, A., Hearty, A., Leclercq, C., Lindtner, O., Tlustos, C., Verger, P., Volatier, J.-L., & Arcella, D. (2011). Methodological characteristics of the national dietary surveys carried out in the European Union as included in the European Food Safety Authority (EFSA) comprehensive european food consumption database. *Food Additives & Contaminants: Part A*, 28(8), 975-995. http://dx.doi.org/10.1080/19440049.2011.576440. PMid:21732710.

Mojrian, S., Pinter, G., Joloudari, J. H., Felde, I., Szabo-Gali, A., Nadai, L., & Mosavi, A. (2020). Hybrid machine learning model of extreme learning machine radial basis function for breast cancer detection and diagnosis; a multilayer fuzzy expert system. In M. Dinh, A. L. Felipe, & D. Dang-Pham (Eds.), *RIVF International Conference on Computing and Communication Technologies (RIVF)* (pp. 1-7). Piscataway: IEEE. http://dx.doi.org/10.1109/RIVF48685.2020.9140744.

Molajou, A., Afshar, A., Khosravi, M., Soleimanian, E., Vahabzadeh, M., & Variani, H. A. (2021a). A new paradigm of water, food, and energy nexus. *Environmental Science and Pollution Research International*. http://dx.doi.org/10.1007/s11356-021-13034-1. PMid:33634401.

Molajou, A., Pouladi, P., & Afshar, A. (2021b). Incorporating social system into water-food-energy nexus. *Water Resources Management*, 35(13), 4561-4580. http://dx.doi.org/10.1007/s11269-021-02967-4.

Pérez-Castellanos, M. S. (2004). Food safety warnings in public health. *Gaceta Sanitaria*, 18(Suppl 1.), 234-238. PMid:15171885.

Portier, C. J., Armstrong, B. K., Baguley, B. C., Baur, X., Belyaev, I., Bellé, R., Belpoggi, F., Biggeri, A., Bosland, M. C., Bruzzi, P., Budnik, L. T., Bugge, M. D., Burns, K., Calaf, G. M., Carpenter, D. O., Carpenter, H. M., López-Carrillo, L., Clapp, R., Cocco, P., Consonni, D., Comba, P., Craft, E., Dalvie, M. A., Davis, D., Demers, P. A., Roos, A. J., DeWitt, J., Forastiere, F., Freedman, J. H., Fritschi, L., Gaus, C., Gohlke, J. M., Goldberg, M., Greiser, E., Hansen, J., Hardell, L., Hauptmann, M., Huang, W., Huff, J., James, M. O., Jameson, C. W., Kortenkamp, A., Kopp-Schneider, A., Kromhout, H., Larramendy, M. L., Landrigan, P. J., Lash, L. H., Leszczynski, D., Lynch, C. F., Magnani, C., Mandrioli, D., Martin, F. L., Merler, E., Michelozzi, P., Miligi, L., Miller, A. B., Mirabelli, D., Mirer, F. E., Naidoo, S., Perry, M. J., Petronio, M. G., Pirastu, R., Portier, R. J., Ramos, K. S., Robertson, L. W., Rodriguez, T., Röösli, M., Ross, M. K., Roy, D., Rusyn, I., Saldiva, P., Sass, J., Savolainen, K., Scheepers, P. T., Sergi, C., Silbergeld, E. K., Smith, M. T., Stewart, B. W., Sutton, P., Tateo, F., Terracini, B., Thielmann, H. W., Thomas, D. B., Vainio, H., Vena, J. E., Vineis, P., Weiderpass, E., Weisenburger, D. D., Woodruff, T. J., Yorifuji, T., Yu, I. J., Zambon, P., Zeeb, H., & Zhou, S. F. (2016). Differences in the carcinogenic evaluation of glyphosate between the International Agency for Research on Cancer (IARC) and the European Food Safety Authority (EFSA). *Journal of Epidemiology and Community Health*, 70(8), 741-745. http://dx.doi.org/10.1136/jech-2015-207005. PMid:26941213.

Rad, M. A., & Ayubirad, M. S. (2017). Comparison of artificial neural network and coupled simulated annealing based least square support vector regression models for prediction of compressive strength of high-performance concrete. *Scientia Iranica*, 24(2), 487-496. http://dx.doi.org/10.24200/sci.2017.2412.

Ribeiro, B. E., & Quintanilla, M. A. (2015). Transitions in biofuel technologies: an appraisal of the social impacts of cellulosic ethanol using the Delphi method. *Technological Forecasting and Social Change*, 92, 53-68. http://dx.doi.org/10.1016/j.techfore.2014.11.006.

Salmon, D., Yih, W. K., Lee, G., Rosofsky, R., Brown, J., Vannice, K., Tokars, J., Roddy, J., Ball, R., Gellin, B., Lurie, N., Koh, H., Platt, R., & Lieu, T. (2012). Success of program linking data sources to monitor H1N1 vaccine safety points to potential for even broader safety surveillance. *Health Affairs*, 31(11), 2518-2527. http://dx.doi.org/10.1377/hlthaff.2012.0104. PMid:23129683.

Savelli, C. J., & Mateus, C. (2020). Looking inside the International Food Safety Authorities Network Community website. *Journal of Food Protection*, 83(11), 1889-1899. http://dx.doi.org/10.4315/JFP-20-193. PMid:32556306.

Sierra-Soler, A., Adamowski, J., Qi, Z., Saadat, H., & Pingale, S. (2015). High accuracy Land Use Land Cover (LULC) maps for detecting agricultural drought effects in rainfed agro-ecosystems in central Mexico. *Journal of Water and Land Development*, 26(1), 19-35. http://dx.doi.org/10.1515/jwld-2015-0014.

Sossa, J. W. Z., Halal, W., & Zarta, R. H. (2019). Delphi method: analysis of rounds, stakeholder and statistical indicators. *Foresight*, 21(5), 525-544. http://dx.doi.org/10.1108/FS-11-2018-0095.

Su, X., Wang, M., & Fan, J. (2004). Maximum likelihood regression trees. *Journal of Computational and Graphical Statistics*, 13(3), 586-598. http://dx.doi.org/10.1198/106186004X2165.

Sun, X., Liu, M., & Sima, Z. (2020). A novel cryptocurrency price trend forecasting model based on LightGBM. *Finance Research Letters*, 32, 101084. http://dx.doi.org/10.1016/j.frl.2018.12.032.

Viscusi, W. K. (1988). Predicting the effects of food cancer risk warnings on consumers. *Food, Drug, Cosmetic Law Journal*, 43, 283.

Williams, M. S., Ebel, E. D., & Vose, D. (2011). Framework for microbial food-safety risk assessments amenable to Bayesian modeling. *Risk Analysis: An International Journal*, 31(4), 548-565. http://dx.doi.org/10.1111/j.1539-6924.2010.01532.x. PMid:21105883.

Wu, D. D. (2020). Data intelligence and risk analytics. *Industrial Management & Data Systems*, 120(2), 249-252. http://dx.doi.org/10.1108/IMDS-02-2020-606.

Wu, Y., & Chen, J. (2018). Food safety monitoring and surveillance in China: past, present and future. *Food Control*, 90, 429-439. http://dx.doi.org/10.1016/j.foodcont.2018.03.009.

Yusianto, R., & Hardjomidjojo, H. (2019). Intelligent spatial logistics DSS for tracking and tracing in horticultural food security. In Universitas Dian Nuswantor (Org.), *2019 International Seminar on Application for Technology of Information and Communication (ISemantic)* (pp. 1-5). Piscataway: IEEE.

Zhang, R., Zhou, L., Zuo, M., Zhang, Q., Bi, M., Jin, Q., & Xu, Z. (2018). Prediction of dairy product quality risk based on extreme learning machine. In S. Patnaik (Ed.), *2018 2nd International Conference on Data Science and Business Analytics (ICDSBA)* (pp. 448-456). Los Alamitos: Conference Publishing Services, IEEE Computer Society.

Zhong, Y., He, C., & Tang, J. (2020). Research on the application of Big Data in smart food safety. In Galatasaray Üniversitesi (Org.), *2020 4th Annual International Conference on Data Science and Business Analytics (ICDSBA)* (pp. 36-39). Los Alamitos: Conference Publishing Services, IEEE Computer Society.