

Research on Intelligent Analysis of Illegal Food Safety Behavior Based on Deep Learning Algorithm

Bing Yang^{1,2}, Kai Chen^{1,2}, Yajie Wang^{1,2,*}, Hong Tan^{1,2}, Fugui Wang² and Minghui Wang²

¹ Guizhou Food Safety Inspection Engineering Technology Research Center Co., Ltd., Guiyang, Guizhou, 550000, China

² Guizhou Academy of Sciences Big Data Co.LTD, Guiyang, Guizhou, 550000, China

Abstract. Food safety has been a major concern in recent years as a result of numerous food safety events in many nations. This could increase the health risks associated with eating low-quality food, lowering customer confidence in food safety. It is critical to overcome this challenge and gain consumer trust in order to improve food quality and safety. To address this issue, we suggested an intelligent deep learning method for identifying which foods are potentially harmful to human health based on chemical and additive qualities, which could have a significant impact on consumer health. The findings of our survey show that deep learning surpasses other methods such as manual feature extractors, as well as the promising findings of categorization of hazardous food, further research efforts to apply deep learning to the field of food will be made in the future.

1 Introduction

Health of human depend upon how much healthy food he/she is eating on a regular diet. Naturally grown food have been used for long as food, and they can be treated according to the demand of customer [6]. Types, balance, vitamin, and processing strategies of food should be natural which are important considerations for a balanced diet [10]. It is true that people from different parts of the world eat in different ways. It is critical to understand food features types, balance, vitamin, and various other factors in order to inspect food quality and safety for consumers all over the world.

The economic effect and disturbed health of humans from recent food problems demonstrate the need of food safety [3]. If food quality and standards are not met, any product will be lost on a huge scale, causing economic problems as well as having an environmental impact. As a result, any food product must be made while adhering to strict safety regulations [11]. Many of these occurrences, however, may be traced back to food manufacturers' failure to follow food hygiene and/or food safety procedures. As consumer awareness of food safety improves, more people are ready to pay higher costs for safe food, which benefits both the country's economy and company profits. As a result, initiatives for food traceable, food certification marks, and product ingredient labelling are required to ensure food safety.

In this paper we proposed a deep learning strategy for verifying food safety based on food additive and chemical properties in this research. For example, we may take the Vitamin A, Vitamin B, and Fat percent per 100 g and estimate if the food is safe to sell in the market or not based on these features. The encoder was used to embed

the data features, and a neural network was utilised to forecast the outcome. Hyperopt was utilised to optimise the settings of our network [1].

2 Literature Survey

Exposure to various institution standard of quality control enhances the food supply chain's susceptibility in the component of international purchasing. Proactive and reactive methods, including as rules, laws, and control mechanisms, have been implemented to mitigate risks in this circumstance. Many research is being done to better understand entire supply chain policies and their impact on food system performance [16]. Despite these efforts and the importance of safety in quality management, the frequency of contaminated incidents demonstrates that the phenomena is still poorly understood. There are very less studies which have attempted to solve that problem using deep learning, thus we employed deep learning as a technique to classify food based on its environmental and substantial features.

3 Methods

In this section we will go through we will go through strategies we utilize in our Deep Learning Model and what additional features we used for better understanding of data and making our model more accurate for prediction.

* Corresponding author: wangyajie@gzbd.com

3.1 Label Smoothing for Overfitted model

When we were training our neural network model, one issue that kept cropping up was that our network prediction was not getting any closer to the goal value. One of the reasons for this was the loss measure we were employing, Binary Cross Entropy with Logits loss [7], which penalises very confident inaccurate predictions and so harmed our neural network. To solve this problem, the weight for each target is set to $\frac{\log(F_{min}+100)}{\log(F_i+100)}$ where F_i is the number of samples with 1 value in target column. The constant term 100 is included to prevent a target with a low frequency from having a significant impact on the model, and log conversion is used to smooth it out [4].

3.2 Additional Features

For getting more understanding of the food quality standards we used additional handcrafted features which can help in making better classification model some of that features are: -

(1) Ash % per 100g in food: - The inorganic residue (minerals) left following full oxidation of organic matter is referred to as ash content. The assessment of ash content is frequently utilised in the flour milling industry as a quality control metric for flour extraction. When ash levels in flour are excessively high, it suggests that the wheat has been contaminated with bran [14]. The formula for calculating ash percent is as follows:

$$\%ash = \frac{(wt\ of\ ashing - wt\ of\ crucible)}{original\ wt} \times 100 \quad (1)$$

(2) Solubility: - The expression "like dissolves like" is frequently used to describe the solubility of organic compounds. This indicates that molecules having a lot of polar groups (polar molecules) are more soluble in polar solvents, while molecules with few or no polar groups (non-polar molecules) are more soluble in non-polar solvents. As a result, depending on their chemical structures, vitamins are either water-soluble or fat-soluble [12].

4 Data Preprocessing and Analysis

This section contains details on the data preprocessing and analysis procedures used in the studies. All of the procedures utilised are theoretically described, with references provided.

4.1 Data Cleansing

After gathering all of the data from various sources, we perform general data cleaning and formatting before running our data through our prediction model. For handling the missing values in the categorical columns which contain information about food in that column there are NaN values for handling that string replacement has been done because these values can cause an error when using the encoding techniques that each model needs. For the columns that include information about the Date-Time

when the entry for that food item was made, the format of that column has been altered for our encoder model [5].

4.2 Data Analysis

As we know that fast development of many economies, quality standards have focused on consumers' demand for safe food and beverage. General ecological factors and influences, such as chemicals, pesticides, food hygiene, ethical trade and production, are among the components of a quality standard set. When we did analysis for finding the distribution of chemical and physical properties of food where target value is 1 [15] there was certain range of values where food is considered good for selling. For example, as we can see that Egg product, Fat (%), Solids (%), Vitamin A, C and D is calculated for per100g of item for finding properties of it.

Table1. Physical and Chemical Properties

Properties	Product		
	Yogurt	Egg	Fish
Fat (%)	3-15	2.7-4.2	7.9-9.26
Solids (%)	7.3-8.5	24-41	36-40
Protein (%)	2.1-2.9	12.3-48	31-62.82
Moisture (%)	9.35-21	45.67-48.9	1.1-1.5
Ash (%)	0.70-0.82	0.08-0.15	1.1-2.0
Solubility	10.2-12	10.2-12	5.1-12.3
Vitamin A	0-130	100-126	350-700
Vitamin D	0-1.33	2.9-3.2	25-45
Vitamin C	4.1-16.2	4.1-16.3	4.1-16.4

5 Model

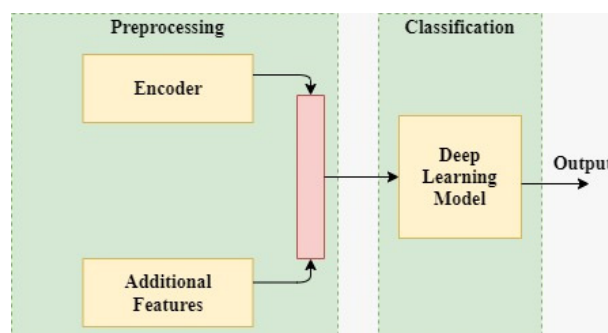


Figure 1. System Architecture

Figure 1 depicts our network's system model, with the preprocessing module as the initial module. Our encoder model takes category, chemical, and physical information as input. We then combined the new handcrafted features with the encoder model's output. The goal of this concatenation is to improve data quality. After the preprocessing module, the concatenated data is passed to the classification module, which is performed using our proposed deep learning model. The classification result is our network's output. The input (x_1, x_2, \dots, x_n) is passed through squeeze and excitation block in our deep

learning model because it brings a building structure for our model that improves model assurance at almost negligible cost for machines [8] and with the help of the

weighting of each feature map can be modified adaptively by the model.

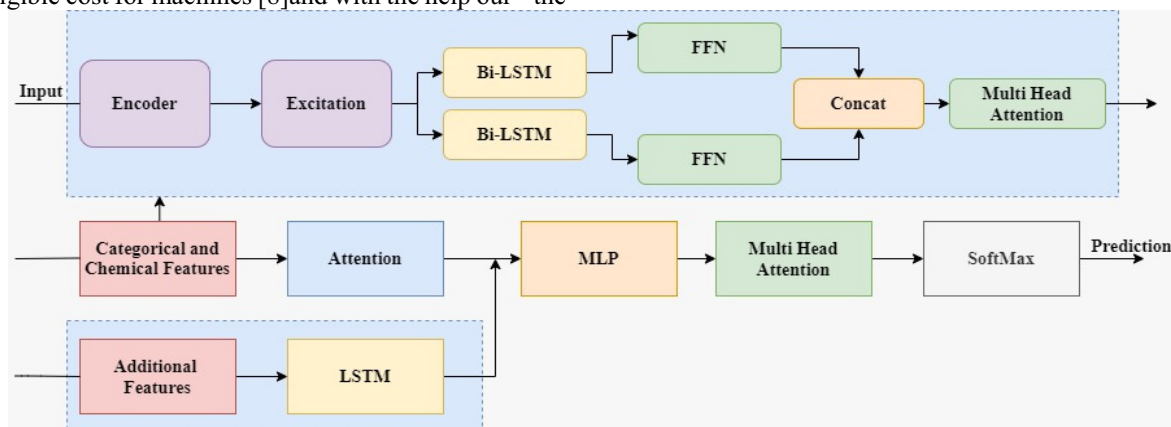


Figure 2. Model Architecture

We inserted the Bi-LSTM block [9] after the squeeze and excitation blocks because Bidirectional LSTMs are an extension of standard LSTMs that can increase model performance on sequence classification challenges. Bidirectional LSTMs train two instead of one LSTM on the input sequence in problems where all time-steps of the input sequence are known. For the embedding's of input features, the output of the two bi-directional LSTMs is transmitted through the FFN block and concatenated for the Multi head attention block [13].

After getting the embedding's for input features we applied LSTM layer to the additional features $y_1, y_2 - y_n$. The input features embeddings and additional features are then fed through MLP, a feedforward artificial neural network that uses backpropagation as a supervised learning strategy for training. MLP is distinguished from a linear perceptron by its numerous layers and non-linear activation. It can tell the difference between data that isn't linearly separable. It is again transferred through Multi head attention from the output MLP. After that we passed the $z_1, z_2, - z_n$ through softmax activation function which also known as softargmax or normalized exponential function, is a generalization of the logistic function to multiple dimensions. The softmax [2] is shown below which takes our output z for final prediction.

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \quad (2)$$

5.1 Bi-LSTM and FFN

Bi-LSTM, often known for two-way directional model, is process in which we have to pass input sequentially functional model that has 2 model, which takes values from single way and different from different direction. This boost the properties which are in model, allowing model to better understand characters rapidly following and word in a sentence. Because they have attempted to learn how and when to forget and when not to employ gates in their architecture, LSTMs and their bidirectional

variations have become popular. Vanishing gradients were a major issue in previous RNN architectures, causing the nets to learn less. The following is the Bi-LSTM formulation: -

$$f_t = \sigma_g(W_f \times x_t + U_f \times h_{t-1} + b_f) \quad (3)$$

$$i_t = \sigma_i(W_i \times x_t + U_i \times h_{t-1} + b_i) \quad (4)$$

$$o_t = \sigma_o(W_o \times x_t + U_o \times h_{t-1} + b_o) \quad (5)$$

It is first and most basic neural network model to be created. The properties in this model flows exclusively in single way that is in from input to output nodes, passing from any nodes. At values for activated and states which are not active can be used to generate as long until the cut-off value is between 0-2. A basic algorithm known for backpropagation which is helpful in making accurate ML/AI networks can be helpful for training. This evaluates differences for computing the reassigning of weights, resulting in a optimizer algorithm. Single-layer networks only learn patterns which has property of separating linearly. Single-layer network is identical to the logistic regression model, which extensively applied at statistical modelling decision. The FFN has logistic prediction is joined with the activation function.

5.2 Loss Function BCE

Cross-entropy loss, or log loss, measures evaluation of a classification model whose output is a probability value between 0 and 1 is measured by cross-entropy loss, also known as log loss. As the projected likelihood differs from the actual label, cross-entropy loss grows. Log loss decreases as the estimated probability approaches 1.

$$loss = \frac{-1}{n \sum_i^N \sum_j^M z_{ij} \log P_{ij}} \quad (6)$$

Then it calculates the score that penalizes the probabilities based on the distance from the expected value.

The log value for the revised probabilities will then be estimated. The reason in adopting the log estimate is that it has a lower penalty for modest variations between predicted and corrected probabilities. The first portion of the formula becomes active and the second portion vanishes when the observation belongs to class 1, and vice versa when the observation's actual class is 0. This is how the Binary cross-entropy is calculated.

6 Future Advancement

Improving cooking items quality best strategies and legislation is critical in future of food item safe. Factories which are involved in keeping food safe market has a high chance of increase at a worth of \$20.1 billion in the future. There are ongoing practices of new rules to assist preserve the industries of food at global level. The following are some examples of future scope: -

(1) Warning judgement: - Identifying potential threats and growing risks for food supply chain, recognising which points in the supply chain are vulnerable to threats, and applying appropriate controllable strategies to limit risk are part in warning judgement process. Cooking items fraud and intentional attacks on the food supply chain are the focus of the warning judgement concept.

(2) Fast identifying strategies: - To comply with mandated health strategies, quick identifying approaches are being introduced. New technology and equipment assist confirm items satisfy hygiene criteria and detect spoiling before it becomes a contamination by producing results on-site in real time. This testing approach improves the efficiency of identifying, improves cooking items safety procedures, and contributes to a safer food supply.

7 Conclusion

The tests detailed in this paper show how environmental and substantial attributes can be used to predict quality of product. In terms of performance, we compared deep learning to other prominent approaches and discovered that the deep learning approach outperforms the other approaches in the studies we looked at. We ended our examination of the benefits and drawbacks of deep learning approaches, as well as the obstacles and future prospects of deep learning safety if food. Future studies should rely on the advancement of fully automated information gathering equipment systems with stable signal output for food, as well as global food data exchange platforms, due to the use of automated or even manual information collecting techniques and poor data management and sharing systems. It is still exceedingly difficult to gather big data linked to food. Deep learning technology's data mining potential may be assessed in food-related areas are rarely investigated, like product sensory and intake, supply chain, and so on. From a future perspective, there is a lot more to discover.

Acknowledgement

This work was supported by Science and Technology Projects of Guiyang National High-tech Industrial Development Zone (GXYP-2017-005) and Guizhou Provincial Science and Technology Projects([2015]4005.

References

1. James Bergstra, Dan Yamins, David D Cox, et al. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In Proceedings of the 12th Python in science conference, volume 13, page 20. Citeseer, 2013.
2. Guillaume Bouchard. Efficient bounds for the softmax function, applications to inference in hybrid models. In Presentation at the Workshop for Approximate Bayesian Inference in Continuous/Hybrid Systems at NIPS-07. Citeseer, 2007.
3. Fernando P Carvalho. Agriculture, pesticides, food security and food safety. *Environmental science & policy*, 9(7-8):685–692, 2006.
4. Blair Chen, Liu Ziyin, Zihao Wang, and Paul Pu Liang. An investigation of how label smoothing affects generalization. arXiv preprint arXiv:2010.12648, 2020.
5. Xu Chu, Ihab F Ilyas, Sanjay Krishnan, and Jiannan Wang. Data cleaning: Overview and emerging challenges. In Proceedings of the 2016 international conference on management of data, pages 2201–2206, 2016.
6. Thomas L Friedman. *The world is flat: The globalized world in the twenty-first century*. Penguin London, 2006.
7. Elliott Gordon-Rodriguez, Gabriel Loaiza-Ganem, Geoff Pleiss, and John P Cunningham. Uses and abuses of the cross-entropy loss: case studies in modern deep learning. arXiv preprint arXiv:2011.05231, 2020.
8. Jie Hu, Li Shen, and Gang Sun. Squeeze-andexcitation networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7132–7141, 2018.
9. Zhiheng Huang, Wei Xu, and Kai Yu. Bidirectional lstm-crf models for sequence tagging. arXiv preprint arXiv:1508.01991, 2015.
10. Francisco Jiménez-Colmenero, Jesus Ventanas, and Fidel Toldrá. Nutritional composition of drycured ham and its role in a healthy diet. *Meat Science*, 84(4):585–593, 2010.
11. Marion Koopmans and Erwin Duizer. Foodborne viruses: an emerging problem. *International journal of food microbiology*, 90(1):23–41, 2004.
12. Marion Koopmans and Erwin Duizer. Foodborne viruses: an emerging problem. *International journal of food microbiology*, 90(1):23–41, 2004.

13. Jian Li, Zhaopeng Tu, Baosong Yang, Michael R Lyu, and Tong Zhang. Multi-head attention with disagreement regularization. arXiv preprint arXiv:1810.10183, 2018.
14. Young W Park and Leonard N Bell. Determination of moisture and ash contents of foods. *FOOD SCIENCE AND TECHNOLOGY NEW YORK-MARCEL DEKKER-*, 138(1):55, 2004.
15. John K Taylor and Cheryl Cihon. *Statistical techniques for data analysis*. CRC Press, 2004.
16. Donald Waters. *Supply chain risk management: vulnerability and resilience in logistics*. Kogan Page Publishers, 2011.