# Machine Learning and Multiple Regression-Based Study of the Medicinal Properties of Food

Jiasheng Zhou [1,*], Zhihong Fan [1], Wenxin Zhang [2]

[1] Fuzhou University, College of Computer and Data Science/College of Software, Fujian, Fuzhou, 350108

[2] Fuzhou University, College of Physics and Information Engineering, Fujian, Fuzhou, 350108

* Corresponding Author Email: 1246511834@qq.com

**Abstract.** We study the relationship between food ingredients and hot and cold properties based on the idea that "medicine and food have the same origin". Firstly, we classify foods with known hot and cold properties as typical representatives of flat, warm and cold properties, and use various machine learning algorithms to classify the typical food representatives. In order to further improve the reliability and accuracy of the classification, we applied Bayesian optimization to the SVM and the SVM was able to classify the typical food representatives again with an accuracy of 96.53%. Based on the above findings, we further analysed which chemical components played a key role in the hot and cold properties. We then applied multivariate logistic regression for quantitative analysis, using stepwise forward regression to minimise complete multicollinearity and OLS + robust standard errors to eliminate the effect of heteroscedasticity. Based on the analysis of the coefficients of the regression equations and the significance test results, the conclusions reached were consistent with the qualitative analysis, with the four main components of energy, water, minerals and fat playing a major role in the chilling and heating properties of food. In conjunction with the analysis in the full text, we make recommendations for the development of functional foods where heat and cold are the principles.

**Keywords:** Factor analysis, Clustering algorithms, Food classification.

## 1. Introduction

The phrase medicine and food have the same origin' refers to the fact that there is no absolute line of demarcation between many foods and medicines, which demonstrates the correct understanding of ancient Chinese medicine scholars of the relationship between medicine and food and between pharmacotherapy and food therapy. Combining the clinical experience of Chinese medicine with modern medical theory, the hot and cold properties of food can be roughly classified into three categories: flat, warm and cold. Tang Xueyang [1] outlined the development and application of medicinal foods, Zhang Huafeng [2] explained the terms related to medicinal foods, Gao Yue [3] presented his views on marine medicinal foods, Zhao Degang [4], Xiong Chaoyong [5] and others analysed the properties of plant-based foods, and many other scholars studied medicinal herbal medicines [6-8], but the current studies all lack systematization and sufficient theoretical support.

## 2. Building and solving food classification models

We apply several models of machine learning for classification. In order to evaluate how good the classification is, we first introduce several machine learning related evaluation metrics.

### 2.1. Machine learning related evaluation metrics

(i) F1 score

First, let's take a binary classification problem as an example. In a binary classification problem, we have to determine the positive class first for the convenience of classification, and in machine learning we usually define the events that we are more concerned with as positive class events. The number of samples with high potassium is small, so we are more interested in whether the artefacts

of type high potassium are correctly classified, so we define high potassium as its positive class. The relevant indicators in Table 1 were then calculated separately.

**Table 1.** Indicators related to secondary classification

| Indicators | Meaning |
|---|---|
| True Positive (TP) | Number of predictions of positive classes into positive classes |
| False Negative (FN) | Number of positive classes predicted as negative classes |
| False Positive (FP) | Number of negative classes predicted as positive classes |
| True Negative (TN) | Number of negative classes predicted as negative classes |

In order to evaluate the effectiveness of machine learning classification models, the following four metrics are defined.

(1) Accuracy (ACC): The ratio of the number of correctly classified samples to the total number of samples.

$$ACC = \frac{TP + TN}{TP + FN + FP + TN} \tag{1}$$

(2) Recall (R): The proportion of samples that are actually positive classes that are predicted correctly.

$$R = \frac{TP}{TP + FN} \tag{2}$$

(3) Precision (P): The proportion of samples where the prediction was a positive class that were predicted correctly.

$$P = \frac{TP}{TP + FN} \tag{3}$$

In machine learning, the accuracy rate is not an accurate reflection of the effectiveness of a machine learning model, so we introduce a fourth indicator, the score.

(4) $F_1$ score: Reconciled average of completeness and accuracy rates.

$$F_1 = \frac{2TP}{2TP + FP + FN} \tag{4}$$

As the food classification problem is a multiclassification problem rather than a biclassification problem, $F_1$ scores cannot be calculated directly by summing the average of the full and accurate rates. There are two types of multiclassified $F_1$ scores, $micro\ F_1$ scores and $macro\ F_1$ scores. Unlike the dichotomous $F_1$ score, the $micro\ F_1$ and $macro\ F_1$ scores are calculated as a composite score.

The two types of $F_1$ scores have different definitions and calculations. The $micro\ F_1$ fraction is calculated by summing the values of TP, FP and FN calculated for each category as a positive class to calculate TP, FP and FN for all samples, and then applying the formula for the $F_1$ fraction in the dichotomous problem; the $macro\ F_1$ fraction is calculated for each category as a positive class and then averaged over all $F_1$ fractions. To ensure the robustness of the results, the $micro\ F_1$ and $macro\ F_1$ scores are used separately.

(ii) ROC curve and AUC

When performing classification, machine learning calculates the predicted probability of a sample. Depending on the set threshold, a positive class is classified when the predicted probability is greater than the threshold, and a negative class is classified when the opposite is true. The threshold setting directly affects the generalisation ability of the classifier, and the ROC curve is a curve comparing the generalisation ability of the same classifier based on different thresholds, with the true class rate (TPR) as the vertical coordinate and the false positive class rate (FPR) as the horizontal coordinate.

The AUC is defined as the area enclosed by the ROC curve and the axis below it. The range of the AUC is between [0, 1] and the larger the AUC the better the model will classify. In general, if the AUC is less than or equal to 0.5, the model cannot be used, and models with an AUC greater than 0.85 perform better.

(iii) Cross-validation

Cross-validation is used to prevent over-fitting problems caused by overly complex models. The data samples are first cut into K smaller subsets, each subset is made into a test set once and the rest is used as the validation set. The cross-validation is repeated K times. In cross-validation, ten-fold cross-validation is more commonly used, and Figure 1 shows the process of cross-validation graphically.

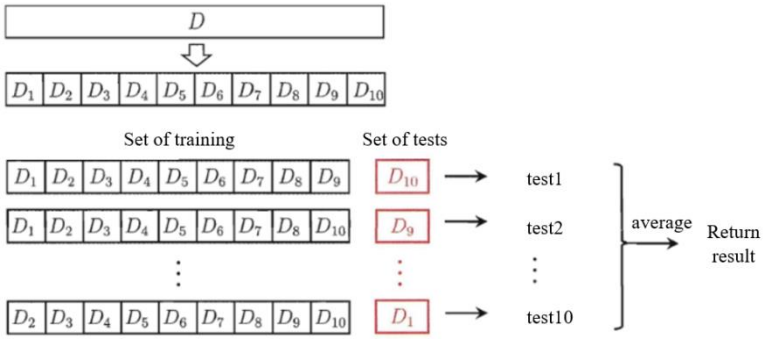


**Figure 1.** Ten fold cross validation diagram [3]

## 2.2. Theories related to machine learning

(i) Support Vector Machine (SVM)

Support Vector Machine (SVM) are a data mining method based on statistical learning theory that can handle regression problems and pattern recognition problems very successfully and can be extended to domain disciplines such as prediction and comprehensive evaluation. The mechanism of an SVM is to find an optimal classification hyperplane that satisfies the classification requirements such that the hyperplane maximises the blank area on either side of the hyperplane while ensuring classification accuracy [9]. Support vector machines can be broadly classified as linear or non-linear.

In a linear support vector machine, the hyperplane for the sample space can be described using the following linear equation.

$$w^T x + b = 0 \tag{5}$$

The hyperplane of the maximum margin needs to be solved for, with the constraint that the distance from the sample point to the decision boundary is greater than or equal to 1.

$$\begin{aligned} &\max_{w,b} \frac{2}{\| w\|} \Longleftrightarrow \min_{w,b} \frac{1}{2} \| w\|^2 \\ &s.t.\ y_i(w^T X_i + b) \geqslant 1 \end{aligned} \tag{6}$$

A new optimisation problem is constructed by introducing a stochastic loss function, taking into account the classification error. The segmented values of the hinge loss function are handled using slack variables.

$$\begin{aligned} &\min_{w,b} \frac{1}{2} \| w\|^2 + C \sum_{i=1}^{N} \xi_i \\ &s.t.\ y_i(w^T X_i + b) \geqslant 1 - \xi_i,\ \xi_i \geqslant 0 \end{aligned} \tag{7}$$

Solving the above equation usually makes use of the duality of the optimisation problem to obtain the dual of the above equation and solve it.

$$\begin{aligned} &\max_{\alpha} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} [\alpha_i y_i (X_i)^T (X_j) y_j \alpha_j] \\ &s.t. \ \sum_{i=1}^{N} \alpha_i y_i = 0, \ 0 \leqslant \alpha_i \leqslant C \end{aligned} \tag{8}$$

For non-linear support vector machines, a kernel function $K(x,z)$ can be used to map a low-dimensional data set to a higher-dimensional data set, where the data becomes linearly separable. In other words, given a kernel function $K(x,z)$, a non-linear classification problem can be solved in the same way as a linear classification problem. The commonly used kernel functions are polynomial kernel function and Gaussian kernel function. Based on the characteristics of the model, we use Gaussian kernel function to construct the non-linear support vector machine.

$$K(x,z) = \exp\left(-\frac{|x-z|^2}{2\sigma^2}\right) \tag{9}$$

The corresponding support vector machine is a Gaussian radial basis function classifier. In this case, the classification decision function is

$$f(x) = sign\left(\sum_{i=1}^{N_s} a_i^* y_i \exp\left(-\frac{\| x - x_i \|^2}{2\sigma^2}\right) + b^*\right) \tag{10}$$

(ii) k-Nearest Neighbor (KNN)

The core idea of the K nearest neighbour algorithm is that a sample is most similar to K samples in a dataset, and if most of these K samples belong to a particular class, then that sample also belongs to that class and has the same characteristics as these K samples. In this paper, we choose the Euclidean distance to determine the similarity of two instances.

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{11}$$

The flow of the algorithm for K nearest neighbours is shown in Figure 2.

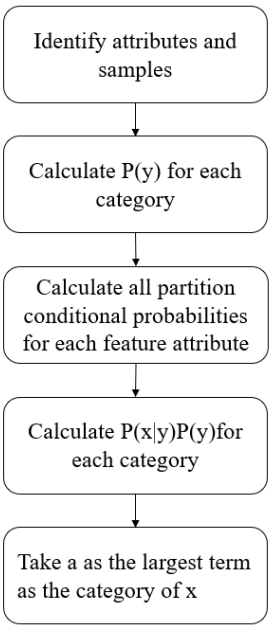


**Figure 2.** Algorithm flow for K-nearest neighbours

(iii) Random forests (RF)

Random Forest is an integrated algorithm. As individual classifiers often have problems such as poor classification accuracy and tendency to overfit, integrated learning is the aggregation of individual classifiers to integrate all classifiers. A random forest is a combined classification model consisting of many decision tree classification models. When a sample to be classified is input, the output of the random forest is determined by a simple vote on the classification result of each decision tree in the following steps.

1) randomly select N samples from the original dataset of N samples and build a decision tree model based on these N samples, as shown in Figure 3.

2) repeat the above process K times to construct K different decision trees.

3) input random variables and generate predictions for each decision tree, yielding K classification results.

4) Vote on each record to determine its final classification based on the results of the K classifications.

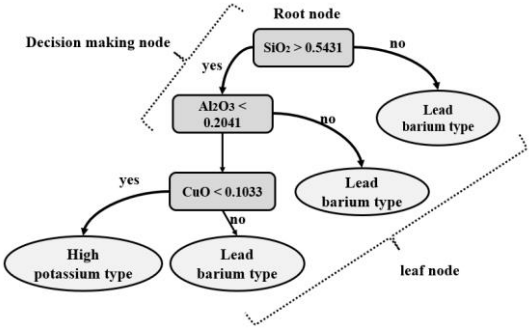


**Figure 3.** Decision tree

### 2.3. Model comparison

In order to achieve effective classification of the food items, we used support vector machine, K-nearest neighbour, and random forest to classify the food items in Annex II by supervised learning and cross-validation, respectively, plotting the ROC curves and confusion matrices, and calculating each parameter such as AUC, $micro\ F_1$, $macro\ F_1$ etc. The results are shown in Table 2 and Figure 4.

**Table 2.** Comparison of evaluation metrics of food classification models based on machine learning methods

| Model | ACC | $micro\ F_1$ | $macro\ F_1$ | AUC |
|---|---|---|---|---|
| SVM | 0.9583 | 0.9583 | 0.9588 | 0.9968 |
| CNN | 0.9514 | 0.9514 | 0.9519 | 0.9925 |
| RF | 0.9444 | 0.9444 | 0.9450 | 0.9484 |

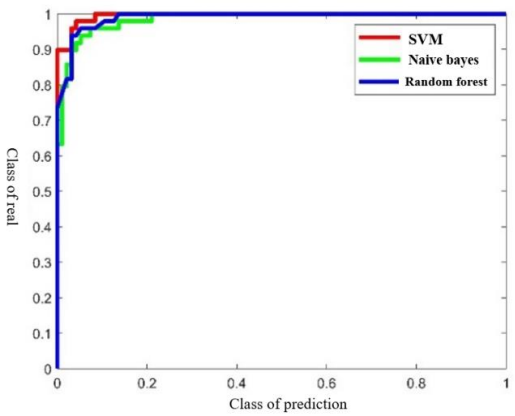


**Figure 4.** Comparison of ROC curves of food classification models based on machine learning methods

Analysis of the results in Table 2 shows that the classification accuracy of SVM is as high as 95.83%, the highest than the other two, $micro\ F_1$, $macro\ F_1$, and the AUC is also both better than CNN and random forest, and the AUC is greater than 0.85, the model performs better; in addition, according to the ROC curve in Figure 4, we can see that the ROC curve of SVM is always on the leftmost side. From this, we conclude the result that SVM has a better classification effect than CNN and random forest. To obtain a better model, we performed Bayesian optimization on the SVM.

### 2.4. SVM with Bayesian optimization

Bayesian optimization is a heuristic tuning strategy. The basic idea is that given an optimised objective function, the posterior distribution of the objective function is updated by continuously adding sample points to better adjust the current parameters. Bayesian conditioning also has the advantages of fewer iterations, faster and more robust to various problems.

The Bayesian optimization framework consists of two main core components [10]: a probabilistic agent model and a collection function, which is implemented through an iterative process consisting of three main steps.

Step 1: select the next most "promising" assessment point based on the maximisation function $x_t$.

Step 2: evaluate the objective function value $y_t = f(x_t) + \varepsilon_t$ based on the selected evaluation point $x_t$.

Step 3: the newly obtained input-observation pairs $\{x_t, y_t\}$ are added to the historical observation set $D_{1:t-1}$ and the probabilistic proxy model is updated in preparation for the next iteration.

Algorithm 1 is a Bayesian optimization framework pseudo-code, as shown in table 3.

**Table 3.** Bayesian optimization framework pseudo-code

| **Algorithm 1: Bayesian optimization framework [10]** |
|---|
| 1: **for** t=1, 2, ... **do** |
| 2: Maximize the collection function to get the next evaluation point: $x_t = arg\max_{x\in\chi}\alpha(x\|D_{1:t-1})$; |
| 3: Evaluate the objective function value $y_t = f(x_t) + \varepsilon_t$; |
| 4: Integrate the data: $D_t = D_{t-1}\cup\{x_t, y_t\}$ and update the probabilistic proxy model; |

The Bayesian tuning process is shown in Figure 5. According to the minimum classification error map of the SVM, the optimal solution for each parameter can be obtained after 30 iterations.

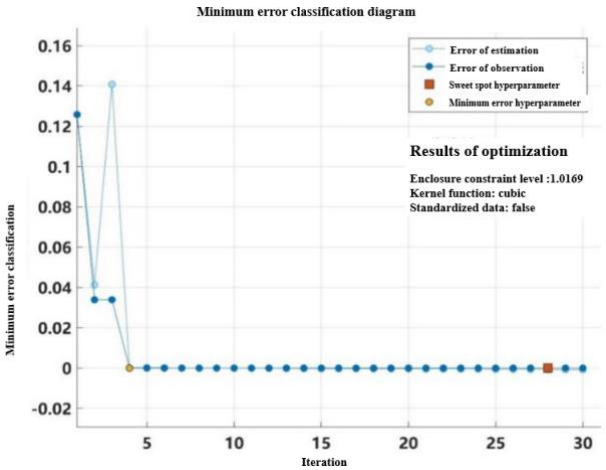


**Figure 5.** SVM minimum classification error plot

The tuned SVM was used to reclassify typical food samples with a classification accuracy of 96.53%. The ROC curves are shown in Figure 6, which shows that the SVM with the tuned parameters has the best results compared with the original three models. In addition, considering the large randomness of the samples selected for cross-validation, we conducted several more validations on

the samples, and the results also proved that the performance of the SVM after cross-validation was significantly improved.

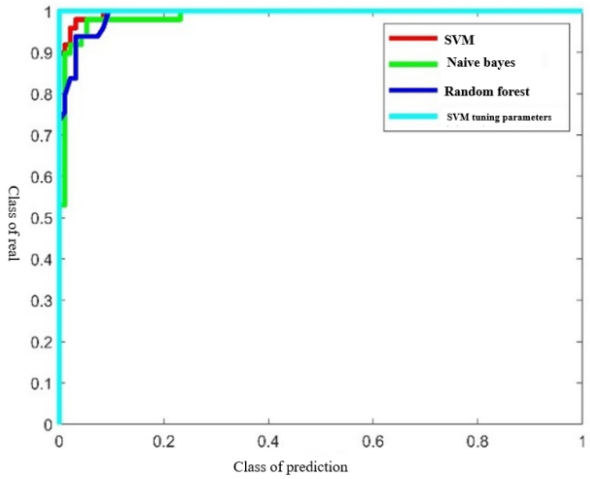


**Figure 6.** Comparison of ROC curves of food classification models based on machine learning methods

Finally, we applied the tuned SVM to all foods to classify all foods. From the results, we can see that meat is mostly warm food, beans and fruits are mostly cool food, and rice and pasta are mostly flat food, which proves that our model is accurate and reliable.

In addition, we also found that the classification of the same food is not invariable, and different cooking methods may result in the same food belonging to different categories. This is because different cooking methods for different foods can lead to significant changes in their original components, so we need to analyse specific foods.

## 3. Analysis of the relationship between food composition and medicinal properties

### 3.1. Qualitative analysis based on SVM

In order to determine which types of ingredients play a major role in food chilling and heatiness, we first used a support vector machine to classify the samples according to each of these nine food ingredients individually, and the food ingredients with better accuracy and related evaluation indicators indicate that they play a major role in chilling and heatiness. The results of the runs are shown in Table 4, where energy, water, minerals and fat were classified better compared to the other indicators, indicating that the four food groups play a more important role in the hot and cold properties.

**Table 4.** Comparison of evaluation indicators using SVM alone for the nine component categories

| Indicators | ACC | $micro\ F_1$ | $macro\ F_1$ | AUC |
|---|---|---|---|---|
| Energy | 0.7431 | 0.7431 | 0.7431 | 0.7724 |
| Moisture | 0.7292 | 0.7292 | 0.7308 | 0.7715 |
| Minerals | 0.7153 | 0.7153 | 0.7162 | 0.7708 |
| Fats | 0.7083 | 0.7083 | 0.7083 | 0.7597 |
| Vitamins | 0.6254 | 0.6138 | 0.6002 | 0.6485 |
| Dietary fibre | 0.6248 | 0.6254 | 0.6125 | 0.6389 |
| Cholesterol | 0.6024 | 0.6027 | 0.6327 | 0.6281 |
| Protein | 0.6186 | 0.6253 | 0.6024 | 0.6166 |
| Carbohydrates | 0.4682 | 0.4592 | 0.4468 | 0.3841 |

In addition, statistical analyses were carried out for each of the known nutritional indicators for warm, cold and flat foods to analyse the sum of the nutrient contents of each category.

### 3.2. Quantitative judgements based on multivariate logistic regression

In order to quantify the role of these components on the hot and cold properties of food, we used multiple logistic regression. Multiple logistic regression is an extension of binary logistic regression and addresses the situation where there are a large number of dependent variables. The dependent variables 0, 1 and 2 were recorded as warm, cold and flat, respectively, and the indicators were used as independent variables in the regression of the dependent variables, with the following expressions.

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_9 x_9 + \mu_i,\ i = 0, 1, 2 \tag{12}$$

For the convenience of the subsequent analysis, it was assumed that the individual nutrient composition of the food was determined by these indicators only and that there was no endogeneity problem due to omission of indicators. Considering that the indicators are not completely independent of each other, there may be a complete covariance between fat and energy, carbohydrate and water, i.e. there exists the relationship Eq.

$$C_0 + C_1 x_{1i} + C_2 x_{2i} + \cdots + C_m x_{mi} = 0, i = 1, 2, \cdots, n \tag{13}$$

To confirm our conjecture, we detected multicollinearity by calculating the variance inflation factor (VIF), which is the coefficient of determination when the ith variable is regressed on all other variables, and it is generally accepted that a VIF greater than 10 is a problem of multicollinearity.

$$VIF_i = \frac{1}{1 - R_i^2} \tag{14}$$

**Table 5.** VIF calculation results

| Indicators | VIF | Indicators | VIF |
|---|---|---|---|
| Moisture | 217.72 | Protein | 81.69 |
| Energy | 182.08 | Vitamins | 66.48 |
| Carbohydrate | 152.16 | Minerals | 58.34 |
| Dietary fibre | 111.02 | Cholesterol | 35.1 |
| Fat | 101.57 | | |

The results of the VIF calculation are shown in Table 5. It can be seen that there is more serious multicollinearity in the regression equation, and for this reason, the forward stepwise regression method was used to reduce complete multicollinearity as much as possible. The basic idea of the stepwise regression method is to introduce new variables one by one. If the partial regression squared is tested to be significant, the new variables can be considered as independent explanatory variables and cannot be linearly represented by other explanatory variables, otherwise it means that the new variables are not independent [11], as shown in the specific steps in Algorithm 2, as shown in table 6.

**Table 6.** Stepwise forward regression pseudo-code

| **Algorithm 2: Stepwise forward regression algorithm** |
|---|
| 1: **while** variables are still introduced |
| 2: Build a k-regression model with y for n regressors or subsets of regressors (k=1, 2, ...) |
| 3: Calculate the value $x_i$ of the corresponding regression coefficient F test statistic, denoted as $F_1^{(k)}, \cdots, F_p^{(k)}$ denote $F_{i1}^{(k)} = \max\{F_1^{(k)}, \cdots, F_p^{(k)}\}$ |
| 4: if $F_{i1}^{(k)} \geqslant F^{(k)}$ |
| Introduce $X_{i_1}$ into the regression model |

After stepwise forward regression, a relatively optimal multiple regression equation was developed. Considering that the possible non-spherical perturbation terms in the equation could lead to heteroskedasticity and thus invalidate the hypothesis test. Therefore, we performed BP test and White's test for heteroskedasticity and the results are shown in Table 7. The results of both tests are significant and there is indeed some heteroskedasticity.

**Table 7.** BP test and White's test results

| White test | |
|---|---|
| $\chi^2(151)$ | Fprob > $\chi^2(151)$ |
| 198.95 | 0.0054 |
| BP test | |
| $\chi^2(24)$ | Fprob > $\chi^2(24)$ |
| 59.54 | 0.0001 |

We use the heteroskedasticity robust standard error method for correction. The ordinary least squares estimator is still used first, and then the variance of the estimator is corrected to eliminate the effect of heteroskedasticity. The final coefficients of the regression equation were obtained as shown in Table 8.

**Table 8.** Table of regression equation coefficients

| | Mild | Warm | Cold | | Mild | Warm | Cold |
|---|---|---|---|---|---|---|---|
| $\beta_0$ | 1.131 | 1.302 | 1.297 | $\beta_5$ (Fats) | 0.523 | -1.682 | 1.236 |
| $\beta_1$ (Moisture) | -1.102 | 2.298 | -1.471 | $\beta_6$ (Protein) | -0.195 | 0.372 | -0.229 |
| $\beta_2$ (Energy) | -1.114 | 3.798 | -2.827 | $\beta_7$ (Vitamins) | 0.128 | 0.012 | -0.082 |
| $\beta_3$ (Carbohydrates) | 0.000 | 0.000 | 0.000 | $\beta_8$ (Minerals) | -0.258 | 0.470 | -0.282 |
| $\beta_4$ (Dietary fibre) | -0.197 | 0.304 | -0.166 | | | | |

The results were tested for significance and the results of the significance test are shown in Table 8, where it can be seen that the coefficients for energy, water, protein, minerals and vitamins are large and significant, indicating that these components play a major role in the coldness and hotness of the food.

**Table 9.** Results of significance tests

| Indicators | Moisture | Energy | Carbohydrates | Dietary fibre | Fats | Protein | Vitamins | Minerals | Cholesterol |
|---|---|---|---|---|---|---|---|---|---|
| P | 0.000 | 0.673 | 0.174 | 0.856 | 0.562 | 0.809 | 0.378 | 0.896 | 0.490 |

As shown in table 9, it is hypothesised that the more energy there is, the more energy the food contains and the more likely it is to be warm; as fat and other energy substances contain more hydrocarbons and less oxygen, water provides the raw material for the release of these energy substances and therefore plays a major role. In addition, minerals and vitamins are also associated with the hot and cold properties of food, but their role is not as significant as that of energy and other components.

The above analysis and the review of the relevant literature have led to the following conclusions on the classification of hot and cold foods, which can be used as a reference for the development of functional foods based on the principle of hot and cold properties:

Making full use of the traditional classification of the cold and hot properties of foods to make a preliminary identification and classification of cold and hot foods, distinguishing the cold and hot properties of foods in terms of colour, flavour, growing environment and growing season.

The material basis of the cold and hot attributes of meat contains three major energy substances: sugar, fat and protein. The possibility of a gradual increase in the ranking of food properties from cool-cold to flat to warm as the content of energy substances increases.s

Minerals play a more dominant role in the cold and hot properties of foods, but the relationship between the type and content of specific minerals and the cold and hot properties of foods needs to be further explored and explored. Current research [12] shows that foods with higher magnesium and copper content are colder, while foods with higher iron, selenium and zinc content are hotter.

The hot and cold properties of foods are closely related to the content of vitamins [12], and the content of vitamins C and E is positively correlated with the hotness of foods, while the flatness of foods is mainly determined by vitamin E and vitamin F.

## 4. Conclusion

We first classified typical food representatives with known properties as flat, warm and cold, and then used three machine learning algorithms, namely Support Vector Machine (SVM), Random Forest (RF) and K-Nearest Neighbour (KNN), to classify the typical food representatives respectively. In order to further improve the reliability and accuracy of the classification, we applied Bayesian optimization to the SVM and the accuracy of the Bayesian optimized SVM was 96.53%.

Based on the above findings, we first used qualitative analysis to classify each of the nine food categories individually using support vector machines (SVM), and ranked the accuracy of each classification, with high accuracy indicating that such ingredients play a major role in the cold and heat properties of the food, with higher accuracy for energy, water, minerals and fat. We then applied multiple logistic regression for quantitative analysis, using stepwise forward regression to minimise complete multicollinearity and OLS + robust standard errors to eliminate the effect of heteroskedasticity. Based on the analysis of the coefficients and significance test results of the regression equations, the conclusions reached were consistent with the qualitative analysis, with the four components of energy, water, minerals and fat playing a major role in food chilling and caloric properties. In conjunction with the analysis in the full text, we make recommendations for the development of functional foods where cold and heat properties are the principles.

## References

[1] Tang Xueyang, Xie Guozhen, Zhou Rongrong, et al. An overview of the development and application of medicinal food ingredients [J]. China Modern Traditional Chinese Medicine 2020, vol. 22, no. 9, pp. 1428-1433, ISTIC CA, 2020.

[2] Zhang Huafeng. Problems and countermeasures of terminology related to medicinal and food ingredients [J]. Chinese scientific and technical terminology, 2019, 21(4): 65.

[3] Gao Yue, Xu Qiannan, Cai Minggang, et al. Research progress on the development and utilization of medicinal and food products of marine origin [J]. Chinese herbal medicine, 2021, 52(17): 10.

[4] Zhao Degang. Research on plants of medicinal and edible origin [J]. Journal of Plant Physiology, 2021.

[5] Xiong Zhaoyong, Chen Xia. Progress of research on the medicinal and food-related wild vegetable rootlet garlic [J]. Collection, 2019, 20.

[6] Tian Yu, Ding Yanping, Shao Baoping, et al. Interaction of Astragalus membranaceus and other herbal medicines as functional foods with intestinal flora [J]. Chinese Journal of Traditional Chinese Medicine, 2020, 45(11): 2486.

[7] Jiang Yutong, Zhao Yinxu, Jiang Haoxuan, et al. Research progress on the prevention and treatment of Alzheimer's disease with herbs from the same source of medicine and food [J]. Henan TCM, 2022.

[8] Changqing, Wang Jie, Liu Weihai, et al. The application of herbal medicines based on literature analysis in the treatment of osteoporosis [J]. Journal of Chinese medicine, 2020, 26(15): 173-176.

[9] Ding S.F., Qi B.J., Tan H.Y.. A review of support vector machine theory and algorithm research [J]. Journal of the University of Electronic Science and Technology, 2011, 40(01): 2-10.

[10] Cui Jiaxu, Yang Bo. A review of Bayesian optimization methods and applications [J]. Journal of Software, 2018, 29(10): 3068-3090. DOI:10.13328/j.cnki.jos.005607.

[11] You Soldier, Yan Yan. Stepwise regression analysis and its application [J]. Statistics and decision making, 2017(14): 31-35. DOI:10.13546/j.cnki.tjyjc.2017.14.007.

[12] Zhang D, An YX. Advances in the study of cold and heat properties of foods [J]. Asia-Pacific Traditional Medicine, 2017, 13(07): 52-54.