Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 1097-1107
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

Classification of Hyperspectral Milk Varieties Based on SSA-BP Neural Network

Meichen Liu#, Jiangping Liu#, Heru Xue*
College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot Inner Mongolia, China
*Corresponding Author.
#These authors contributed equally to this work

Abstract:

Based on hyperspectral technology, a BP neural network classification model based on sparrow algorithm was proposed to solve the shortage of fast classification technology of milk. The experimental samples were five kinds of milk with different nutrient contents, and the wavelength data within 400-1000nm were collected. The competitive adaptive weighted sampling algorithm and continuous projection algorithm were used to extract the characteristic wavelength from the pre-processed spectral data, and the sparrow algorithm was used to optimize the BP neural network to find the optimal weight threshold training model. The classification results were compared with the support vector machine model and BP neural network model. The results showed that the training set accuracy of CARS-SSA-BP model was 100%, and the test set accuracy was 98.14%. The classification effect of CARS-SSA-BP model was better than that of SVM model and BP neural network model. Sparrow algorithm can effectively optimize the weight and threshold value of BP neural network, and it is feasible for the identification of milk varieties.

Keywords: *Hyperspectral imaging, Competitive adaptive weighted sampling algorithm, Sparrow search algorithm, Milk classification method, Back propagation neural network.*

I. INTRODUCTION

The high-spectral image combines image information and spectral information, which can display the internal spectral data according to the external feature. It is very small to detect milk using high spectroscopy techniques [1]. At present, only Golden Huan and others have proposed to use near infrared spectroscopy and main component analysis. [2]; Zhang Lifu is based on FISS imaging 12 Milk mixed, accurate rate up to 95.5% [3]; Muhaio et al. Based on Fourier infrared spectroscopy, the soft independent mode classification method for milk is classified, except for low fatty milk, the ex-tent of 80% [4]. Foreign scholars ShimaBehkami

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 1097-1107
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

classified milk by establishing feed-forward multi-layer perception [5], Asmakhan identified three kinds of milk powder according to different granularities [6]. Due to the use of high spectroscopy, it is less research to classify milk, and the accuracy of the identification method of the milk species currently needs to be improved. The sparrow algorithm has a small parameter, high stability and can effectively solve the characteristics of the BP neural network easily causing the local extract. Therefore, this paper proposes that the sparrow algorithm optimizes the BP neural network to classify five kinds of milk, and the classification accuracy is up to 98.14%.

II. MATERIALS AND METHODS

2.1 Experimental sample

Experimental milk is milk from the five fat content purchased on the market. Each of the milk is stored in normal temperature, and the refrigeration is stored for one week and purchases three types of storage on the same day for sample production. Including Deluxe Milk (Fat content 4.4g / 100ml), Ely thick milk (fat content 4.6g / 100ml), Mengniu high calcium (fat content is 3.7g / 100ml), Ely skimmed milk (fat content is 0g / 100ml), QQ star (3.5g / 100ml). When the sample was configured, a part was poured into a beaker using a glass rod to stir after one minute, placed in a culture dish having a high diameter of 90 mm height of 8 mm. Each type of milk selected 50 samples, with randomly selected training sets and test sets accounted for 35 and 15 respectively.

2.2 Method

2.2.1 High spectrum data collection

The collection of milk image data in this experiment adopts high spectroscopy techniques. High spectroscopy techniques are not only obtained from objects from objects, but also obtain influence information on a spectral band of a certain image simultaneously. When the image acquisition, the external or less interference factors, such as the difference in the sensor in the sensor, which causes the spectral image due to noise when the spectral image is processed. Satisfactory. Therefore, in order to solve this problem, it is necessary to calibrate the high-spectral image before collecting, the parameter condition of the calibration is fixed setting, the exposure time is selected, the number of pixels is 6 times, the resolution is 4.8 nm, 30cm distance collection halogen lamp Image A of a uniformly illuminated whiteboard; then cover the upper lens cover, capture the dark current image B; then according to the formula (1), the high-spectral image data is calibrated, and the acquired absolute image I is converted into a relative image C [7].

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 1097-1107
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

$$C=(I-B)/(A-B) \tag{1}$$

The original spectrum acquired is shown in Fig 1.

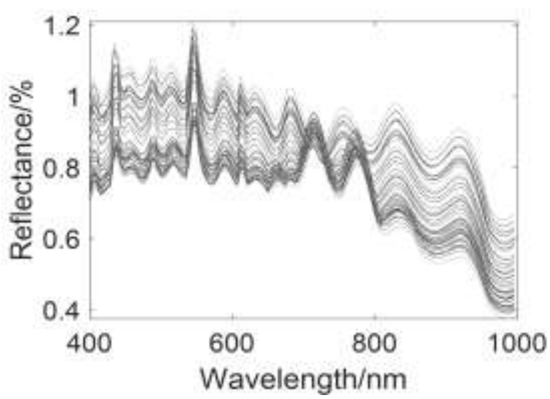


Fig 1: Raw spectrum of the five kinds of milk

2.2.2 SSA-BP Neural network model

Since the BP neural network is due to the slow convergence speed and randomly select the characteristic of the initial weight, it has a certain limitation that is not easy to obtain the optimal value, so the sparrow algorithm is introduced to the BP neural network [**Error! Reference source not found.**]. Design ideas are shown in Fig 2.

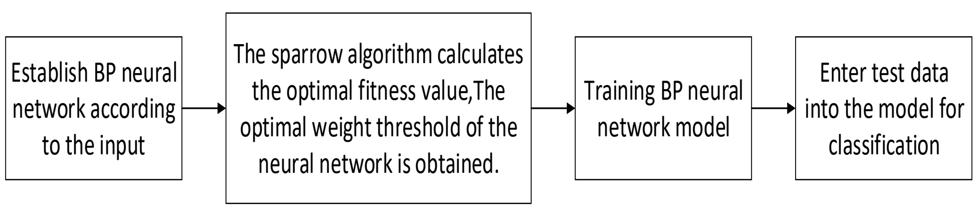


Fig 2: Classification model flow chart

Sparrow search algorithm(SSA) is an intelligent optimization algorithm based on predation behavior of discovers and followers in sparrow populations. The sparrow population contains three behaviors: prey, followed, investigates, and sparrows represent their properties[9]. The founder enjoys the best predation location and a larger food exploration range, and each iterates the location update is as follows:

$$X_{i,j}^{t+1}=\begin{cases} X_{i,j}^{t}\cdot exp\left(\frac{-i}{\alpha\cdot iter_{max}}\right), R_2<ST \\ X_{i,j}^{t}+Q\cdot L \end{cases} \tag{2}$$

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 1097-1107
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

Formula 2 shows the position of the i-th sparrow in the J dimension when the iterations. is the maximum number of iterations, R2 is the security value, and ST is alert value. α and Q are (0, 1] random number and random number of obey normal distribution. When the security value is less than the alert value, the sparrow population is in a safe location, and there is no detection to the predator, and the dis-covers search for food. When safe When the value is greater than or equal to the warning value, it means that the sparrow population detects the predator and quickly escapes to the safe area. Follower position update is as follows:

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot exp\left(\frac{x_{worst}^{t} - x_{i,j}^{t}}{i^2}\right), i > n/2 \\ X_{p}^{t+1} + |x_{i,j}^{t} - X_{p}^{t+1}| \cdot A^{+} \cdot L \end{cases} \tag{3}$$

The X_{Worst} and XP are the worst positions in the overall location and the best position of the current discovery. A is a matrix of 1 * D, When item i is only in the second half, the adaptation is low, so the obtained food is less, requiring non-belt other areas. During the foraging process, the sparrow will observe if the for-aging area has predator, if the danger is updated, the formula is as follows:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^{t} + K \cdot \left(\frac{\left|x_{i,j}^{t} - x_{worst}^{t}\right|}{f_i - f_w + \varepsilon}\right), f_i = f_g \\ X_{best}^{t} + \beta \cdot \left|x_{i,j}^{t} - X_{best}^{t}\right|, f_i > f_g \end{cases} \tag{4}$$

In the formula (4), the X_{best} represents the optimal position in the global , indicates the best adaptivity value of the sparrow at the global optimal location, is the adaptivity value of sparrow at the worst position, K is a random value in the interval [-1,1].It can be seen that when , when the sparrow is in the edge of the group, it will be easily discovered by predator. When the sparrow is in a good location at this time, and the small range moves close to other sparrows to reduce danger.

The algorithm step can be summarized as follows:

1. Set the sparrow population parameters: sparrow, maximum iteration number , discovery and followers;
2. Build a BP neural network, initialize the weight threshold of the network;
3. Adaptive function setting, Where TrainER is the error rate of the training set, TestER is the error rate of the test set. The adaptation function of the BP net-work is the target function. When the two error rates are very suitable, the network is best, and the network is also best;
4. Update the sparrow position according to follow 2. - 4.
5. Calculate the adaptivity value and perform global optimal value update;
6. Get the network optimal weight threshold;
7. calculate the error value of the BP neural network;
8. Determine whether to meet iterative conditions, if the maximum number of iterations is

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 1097-1107
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

reached;

9. The best sparrow brought into the network, sort the iterative result, select the best sparrow of the adaptivity value, bring its weight threshold into BP neural network training, otherwise return to step 4[10].

III. RESULT

3.1 Data pretreatment

Due to the difference in the environment, the spectrum will contain background noise and other physical factors interference, resulting in a high precision of the model, and take appropriate spectral pretreatment method to prepare milk data. Processing to achieve interference and better improvement of model accuracy. This experiment uses Multiple Scattering Correction(MSC) Pretreatment. The spectral image after the MSC treatment method is shown in Fig 3.

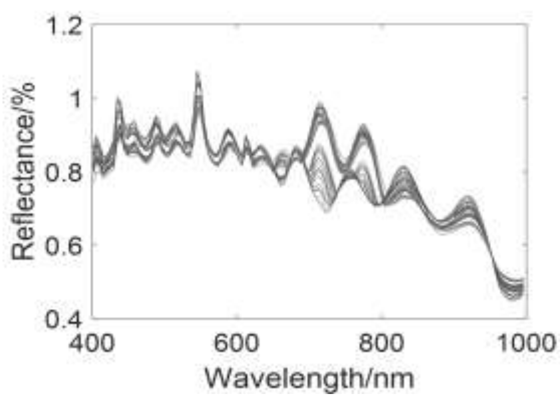


Fig 3: Spectra of milk after MSC pretreatment

3.2. Character selection

3.2.1. Successive Projection Algorithm

Since the original spectrum is 125 bands, the selection of milk data pre-processed after the multi-scattering correction algorithm is used in the Successive Projection Algorithm(SPA) algorithm[11], as shown in Fig 4, the experiment is selected according to the SPA algorithm 17 feature wavelengths [12].

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 1097-1107
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

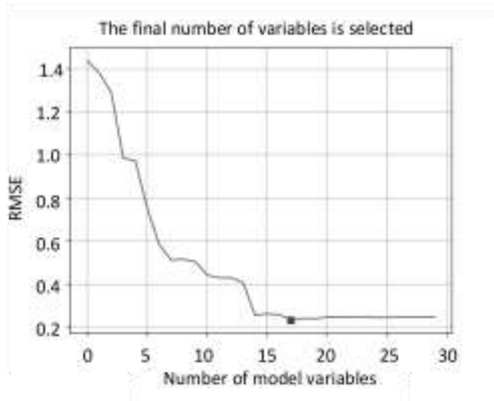


Fig 4. SPA feature band selection

3.2.2. Competitive Adaptive Reweighted Sampling

Competitive adaptive reweighted sampling(CARS) is a feature extraction method based on Monte Carlo Sampling ARS and the least squares PLS[13]. Calculate the wavelength of the wavelength by the PLS algorithm, the wavelength of the weight value is reserved, and the wavelength of the weight value is set to 0 and sieved. Ac-cording to the attenuation index method (EDF), the number of wavelength variables is used to establish the number of modes. Reserved variables are adaptive to the adaptive reweighted sampling(ARS) select variable subset and establish a PLS model. Introducing cross-validation, continuously optimizing RMSECV, selecting the minimum subset of RMSECV[14], that is, the characteristic wavelength combination of the model accuracy. As can be seen from Figure 5, when the number of iterations is 7 times, the RMSE value is the smallest, and the feature wavelength is 23.

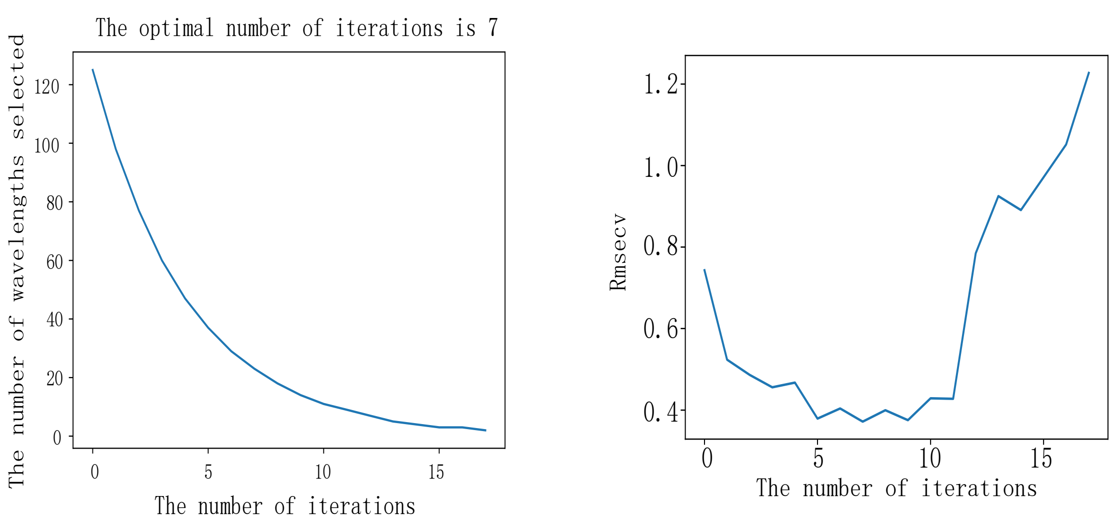


Fig 5. Cars feature band selection

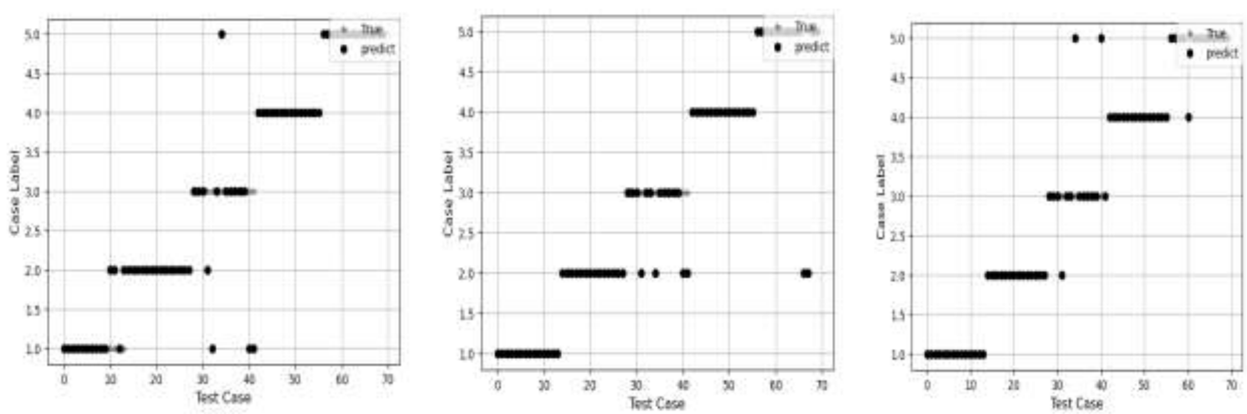
Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 1097-1107
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

3.3. Classification results

This paper proposes to use the sparrow search algorithm to classify the milk variety for milk varieties, of which BP neural network parameters set: iterative parameters 200; the learning rate is 0.01; the training target error is 0.01; Sparrow algorithm parameter setting: population number 50; iterative number 20; discovery and access ratio is 7: 3, the detection risk of sparrow accounts for 0.2, and the warning value is set to 0.6.

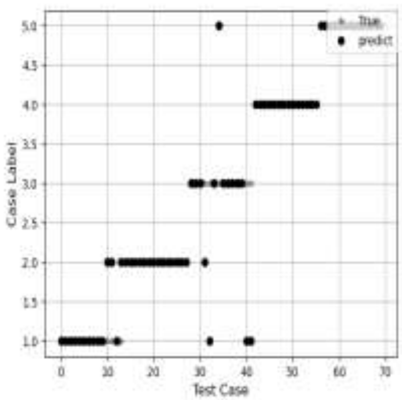
In order to visually see the classification effect of the BP neural network model after the SSA improvement, the experiment is compared with the BP neural network model and the SVM classification model[15].

The 13 and 21 feature bands selected for CARS and SPA input variables, establish a classification model, in order to obtain a more obvious study result, experimentally established three models for 125-dimensional full spectral data, and the classification results are shown in Fig 6.

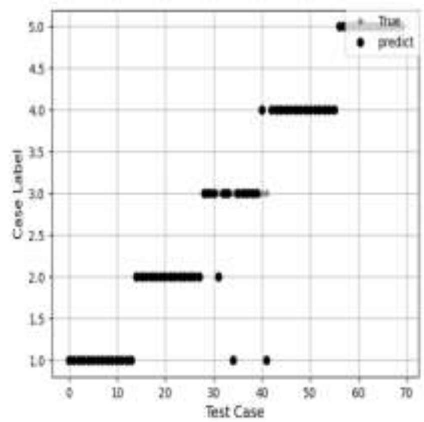


(a) (b) (c)

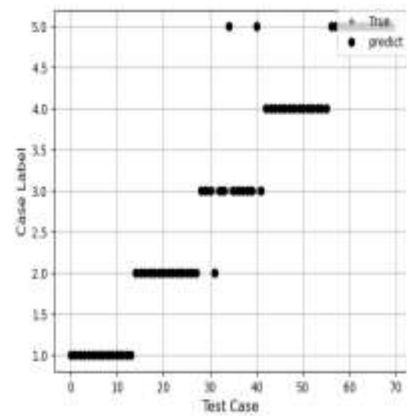
Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 1097-1107
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021



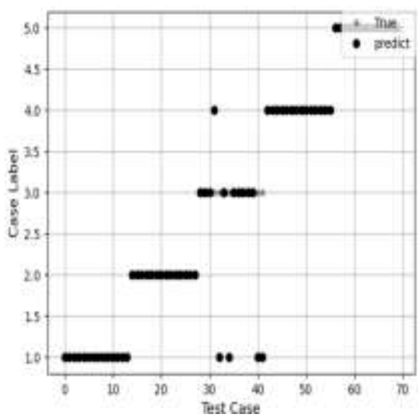
(**d**)



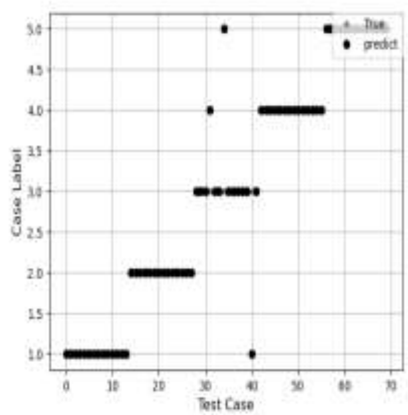
(**e**)



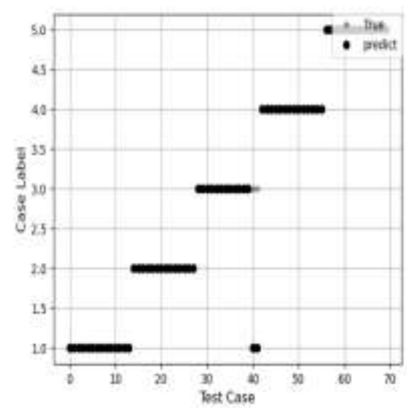
(**f**)



(**g**)



(**e**)



(**f**)

Fig 6. Classification effect diagram obtained by three classification models: (a) All band-SVM; (b) Spa band-SVM;(c) Cars band-SVM;(d) All band-BP; (e) Spa band-BP;(f) Cars band-BP;(g) All band-SSA-BP; (h) Spa band-SSA-BP;(i) Cars band-SSA-BP.

The combination of three models and three different input variables shows the accuracy of the display of TABEL I. It can be seen that the full spectroscopy and feature spectrum on the test data classification.

TABLE I. Classification model based on full spectroscopy and feature spectrum

MODEL	**VARIABLE NUMBER**	**ACCURACY OF TRAIN SET /%**	**ACCURACY OF TEST SET /%**
Full+SVM	125	94.29	88.57

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 1097-1107
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

CARS+SVM	13	98.28	95.71
SPA+SVM	21	96.00	94.29
Full+BP	125	96.57	92.00
CARS+BP	13	98.98	95.75
SPA+BP	21	97.14	94.28
full+SSA-BP	125	99.41	96.33
CARS+SSA-BP	13	100.00	98.14
SPA+SSA-BP	21	98.82	95.71

It can be obtained that the following conclusions: When the variable number changes, the accuracy of the classification will also change, such as the full spectral model relative to the classification accuracy of the feature spectrometer 5,there is a lot of redundancy and physical interference information in the 125-dimensional data. After introducing the SPA algorithm and the CARS algorithm, the classification accuracy of the three classification models has been improved, compared, the CARS algorithm is less than the number of variables selected than the SPA algorithm and higher precision. CARS greatly reduces redundant information and reduces the input variables of the model, It has proven that the classification model introduced by the CARS algorithm has better classification capabilities and stability. The feature band only accounts for only 0.104% of the total band, and the total linear problem between the variables is solved, and the model is more simple and fast and stable.

According to the results of the classification, it can be seen that the sparrow search algorithm proposed in this paper optimizes the classification of the BP Neural Network (SSA-BP) model. It is better than the classic SVM model and BP neural net-work model. As can be seen from Table 1, when the milk is all data as input, the accuracy of the SSA-BP model is 4.33% higher than that of the conventional BP neural net-work model, 7.76% higher than the SVM model. When the feature band extracted by the CARS is used as input data, the accuracy of the SSA-BP model is 2.39% higher than that of the conventional BP neural network model, 2.43% higher than the SVM model. Further explanation of high precision and high stability of the SSA-BP algorithm this article, comparing other milk classification methods experiments easier. Therefore, it is effective to classify milk data using the SSA-BP algorithm.

IV. CONCLUSION

In order to find a method of quickly non-destructive milk variety classification, this paper constructs a model of the sparrow search algorithm to optimize the BP neural net-work, and select the optimal weight threshold to enhance the robustness and classification recognition ability of the BP neural network. The experimental conclusions are as follows:

•The CARS algorithm clearly enhances the correlation between the spectrum, eliminating a large amount of redundant information, and has a good effect on feature extraction of milk high spectroscopy data;

•The two classification algorithms used by SVM, BP neural networks can be clearly seen that the SSA-BP neural network model has the best precision and minimum error. When the full-band data is input, its training set accuracy reaches 99.41, the test set accuracy is 96.33, which is more stable than the above two algorithms, indicating that it meets the classification requirements of milk.

•A more common classification method, the classification accuracy of the SSA-BP neural network model is higher and time consuming, and it provides a new theoretical basis for the classification of milk varieties.

ACKNOWLEDGEMENTS

This research was supported by National Natural Science Foundation of China (Grant No. 61461041). By The Key Science-Technology Project of Inner Mongolia (Grant No.2020GG0169), by Inner Mongolia Education Department Project (Grant No. NJZY21491, NJZY18062).

REFERENCES

[1] Premalatha M (2008) Efficient cogeneration scheme for sugar industry. Journal of Scientific & Industrial Research 67:239-242

[1] Munir MT, David IW, Yu W, Young BR (2018) An evaluation of hyperspectral imaging for characterizing milk powders. Journal of Food Engineering 221:1-10

[2] Jin Y, Du B, Zhi XJ (2016) Fast recognition of adulterated milk and the brands of milk aased on the NIR technology. Food Research And Development 37(03):178-181

[3] Huang CP, Zhang LF, Zhang XH, Zheng FL, Tong QX (2011) Study on discrimination of varieties of milk based on FISS imaging spectral data. Spectroscopy and Spectral Analysis 31(01):214-218

[4] Mu HB, Yin X, Ai LZ, Gu XH (2012) Discrimination of Milk by FTIR and Soft Independent Modeling of Class Analogy. Journal of Dairy Science and Technology 35(02): 34-37

[5] Shima B, Sharifuddin MZ, Mehrdad G, Mohd Fared Abdul K (2019) Classification of cow milk using artificial neural network developed from the spectral data of single and three-detector spectrophotometers. Food Chemistry 294:309-315

[6] Khan A, Munir MT, Yu W, Young B (2020) Wavelength Selection FOR Rapid Identification of Different Particle Size Fractions of Milk Powder Using Hyperspectral Imaging. Sensors 20(16)

Forest Chemicals Revew
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2021 Page No. 1097-1107
Article History: Received: 12 May 2021 Revised: 25 June 2021 Accepted: 22 July 2021 Publication: 31 August 2021

[7] Pan L, He CX, Xiang Y, Sun L (2020) Multiscale Adjacent Superpixel-Based Extended Multi-Attribute Profiles Embedded Mul-tiple Kernel Learning Method for Hyperspectral Classification. Remote Sensing 13(1):50-50
[8] Wang YG, Ahmed SH (2021) Determination of Bridge Prestress Loss under Fatigue Load Based on PSO-BP Neural Network. Computational Intelligence and Neuroscience 2021
[9] Song CG, Yao LH, Hua CY, Ni QH (2021) Comprehensive water quality evaluation based on kernel extreme learning machine optimized with the sparrow search algorithm in Luoyang River Basin. Environmental Earth Sciences 80(16)
[10] Liu TZ, Zou GP, Ahmed SH (2021) Evaluation of Mechanical Properties of Materials Based on Genetic Algorithm Optimizing BP Neural Network. Computational Intelligence and Neuroscience
[11] Sanaeifar A, Huang XY, Chen MY, Zhao ZF, Ji YF, Li XL, He Y, Zhu Y, Chen X, Yu X(2020) Nondestructive mon-itoring of polyphenols and caffeine during green tea processing using Vis-NIR spectroscopy. Food science & nutrition 8(11):5860-5874
[12] Nurajijah, Amsury F, Saputra I, Frieyadie, Sulistyowati DN, Rifai B (2020) Approval of Sharia Cooperative Customer Financing Using PSO-Based SVM Classification Algorithm. Journal of Physics: Conference Series 1641(1):012047
[13] Rashima M (2020) Automated emotion state classification using higher order spectra and interval features of EEG. International Journal of Biomedical Engineering and Technology 2020:34(3)
[14] Yuan ZR, Wei LF, Zhang YX, Yu M, Yan XR (2020) Hyperspectral Inversion and Analysis of Heavy Metal Arsenic Conten-tin Farmland Soil Based on Optimizing CARS Combined with PSO-SVM Algorithm. Spectroscopy and spectral analysis 40(02):567-573
[15] Xue JK, Shen B (2020) A novel swarm intelligence optimization approach: sparrow search algorithm. Systems Science & Control Engineering 8(1):22-34