Scholar Publishing Group
International Journal of Neural Network
https://doi.org/10.38007/NN.2022.030401
ISSN 2790-2692 Vol. 3, Issue 4: 1-8

Neural Network Classifier Improvements Based on Ant Colony Algorithm

Amar Sadia*

Jimma University, Ethiopia

**corresponding author*

Keywords: Ant Colony Algorithm, Neural Network Classifier, Classification Effect, BP Neural Network

Abstract: Compared with other classifiers, neural network (NN) classifiers are faster and have more advantages. However, the NN classifier is unique, because the number of features often changes when selecting features and the classifier is classified by the number of hidden layer nodes, and a NN with a fixed number of hidden layer nodes is used to evaluate different types of feature subsets obviously inappropriate. Therefore, the NN classifier is optimized, and the ant colony optimization (ACO) algorithm is introduced to obtain a diverse solution, which will be faster than the NN classifier to avoid it falling into the problem of local solutions. As a classic feedforward NN, BP neural network (BPNN) is currently the most widely used NN for classification problems. Therefore, this paper constructs a BPNN classifier. By comparing the traditional BPNN classifier and the improved ACO-BPNN classifier on malicious web page detection and classification effects of different sizes of data sets, the results show that, compared with the traditional BPNN classification model, ACO-BPNN The classification accuracy of the BPNN classification model is higher and the classification effect is better.

1. Introduction

With the increasing development of information technology, people have more and more extensive application requirements for artificial intelligence in life. For these requirements, BPNN has attracted great attention as a tool to realize artificial intelligence. It can establish a simplified model and use it to simulate the decision-making process of biological NN analysis [1]. Among them, the classification problem, as the basic means in the human learning process, is the most basic way of processing information in people's daily life, and it provides a judgment basis for various decisions.

Using NNs to solve classification problems has always been the research topic of many researchers, and the research on NN classifiers has also achieved good results. For example, some researches realize the adjustment of the network structure by adding an interference pattern layer to the BPNN. In the learning of network parameters, the original input samples are mapped through the interference pattern layer, and then the BP algorithm is used to realize the next level of BP. Weight adjustment for the network part. In this way, the parameter selection dependence of the original BPNN is reduced, and the classification performance is improved. However, the kernel type selected by the hidden nodes of the interference pattern layer and the consistency of the kernel parameters cannot guarantee the validity of the kernel mapping, and the final classification performance will also be affected [2-3]. Some studies have pointed out that the RLC NN model can cluster clusters of any shape without knowing the number of clusters in advance. This is because the model has high efficiency, fast learning speed, and strong induction ability. The number of training times required for sample data can be very small, so the RLC NN model can complete unsupervised learning efficiently and quickly, and can classify high-risk data. However, due to the lack of prior knowledge, the optimal kernel parameters cannot generally be found, and the kernel space classifier cannot be completely guaranteed to be linearly separable, resulting in the limitations of the linear classification algorithm for optimizing the output of the RBF network [4-5]. On the whole, the classifiers established by the NN have their own advantages and disadvantages, and it is necessary to further improve the performance of the classifier to improve the classification effect.

This paper first introduces the concepts of ACO and BPNN algorithm, then uses ant colony algorithm to improve the weight of BPNN, then designs BPNN classifier, and finally compares the BPNN classifier optimized by ant colony algorithm through experiments and the classification effect of traditional BPNN classifier verifies the effectiveness of the classifier proposed in this paper in malicious web page detection and data classification of different scales.

2. Related Algorithms

2.1. Ant Colony Optimization

ACO is an optimization algorithm inspired by the behavior of ant colonies in nature. In nature, ants can find the optimal value through cooperation. When they look for food, they will leave pheromone, and the ant colony will find the pheromone intensity in the direction to find the path, and the ant colony will strengthen the original pheromone in the process of moving, that is, increase the concentration [6]. So far, this phenomenon is called an iteration of the ant colony algorithm.

The ant colony path optimization formula is as follows:

$$\tau_{ij}(n+1) = \rho \times \tau_{ij}(n) + \sum_{d=1}^{m} \Delta \tau_{ij}^{d} \tag{1}$$

$$\Delta \tau_{ij}^{d} = Q / L \tag{2}$$

Among them, d is the individual ant colony, i is the current position of the ant colony, j is the position that the individual ant colony can choose, τ_{ij} is the pheromone concentration between positions i and j, Q is the amount of pheromone released by the ant colony, and ρ is the information Evaporation coefficient, L is the distance ants crawl.

2.2. BPNN Algorithm

The BPNN algorithm iteratively adjusts the weights between the layers based on the error value to construct a NN model suitable for the current sample data, which requires that the learning sample not only contains the input signal, but also needs to provide corresponding to the current input signal. The target output of [7].

The standard BPNN algorithm iteratively adjusts the weights between layers by following the Delta learning rule during the learning and training process, so as to obtain a NN model suitable for the current sample data [8]. On the whole, the delta learning rule adjusts the connection weight matrix between neurons according to the actual output of the NN and the error function value of the target output. When training the BPNN, the minimum value solved by the gradient descent method is to minimize the error function value between the output value of the NN and the target output [9].

3. Improvement of NN Classifier Based on ACO

3.1. Application of ACO to Improve BPNN Weights

NN, as its name suggests, is inspired by human neurology, and ACO is based on the study of group life and ant movement laws in nature. The BPNN algorithm and ACO have different advantages respectively. However, these two algorithms have their own unavoidable shortcomings when applied separately. These shortcomings are constantly discovered through the research of BPNN. In the research, it was recognized that if ACO and BPNN algorithms are combined in some way to form a new training method, namely ACO-BPNN, many defects can be effectively solved. The ACO-BPNN algorithm is an algorithm with strong optimization ability. Its implementation is simpler than other algorithms, and it needs to adjust relatively few parameters. This is an important advantage of this algorithm, so it can be used in the training of NNs. It is widely used by various scholars [10-11]. For example, the connection weights of the BP neural system are trained, the structure of the BPNN is modified, and the optimal solution of the number of network nodes is determined. Based on this theoretical basis, the BPNN classifier is optimized.

Using ACO to optimize the BPNN classifier, the main optimization method is to use ACO to optimize the classification performance of the traditional BPNN algorithm, that is, to improve the training results of the BPNN algorithm on the connection weights [12]. Since ACO has great advantages in finding the optimal value, it can quickly and accurately find the global optimal solution, and only need to modify a few parameters to achieve good optimization results. Therefore, ACO is selected to train the connection weights to optimize the BPNN classifier [13].

3.2. Optimal Design of NN Classifier

(1) Design of the number of network layers

The BPNN with a three-layer network structure can represent all nonlinear functions. Only when training some data features with high dimensionality, multiple hidden layers are required, which may be relatively complicated when training the BPNN. , if increasing the number of hidden layer nodes still cannot get better training results, you can consider increasing the number of hidden layers to improve the optimization performance of the network model, but it will also increase the difficulty of training [14-15]. Therefore, a single hidden layer is selected as the training model of the BPNN, which can speed up the training speed of the BPNN without affecting the performance of the network model.

(2) Design of hidden layer nodes

Generally, increasing the number of network nodes in the function solving process is more effective and easier to implement than increasing the number of network layers. However, it will be limited by many basic conditions, such as many training parameters and slow training speed, because these problems limit the final result and optimization effect of BP network training [16].

Therefore, this paper uses the ant colony algorithm to optimize the structure of the network model, and selects the number of nodes in different hidden layers for comparative analysis, and finally determines the appropriate number of nodes in the hidden layer, which better solves the problem of complex BP network structure [17]]. Using ACO to optimize the structure of the BPNN produces a better optimization effect, and in the network model with a single hidden layer, selecting the appropriate number of hidden layer nodes, the accuracy of the BPNN classification will be higher, and the training effect will be more effective. .

(3) Design of the classifier structure

The BP network model shows good predictive ability because of its unique structure, so it can accurately predict according to the characteristics of the data in the classification process [18]. It is divided into two parts in the prediction process: first, the similarity between the predicted results of the classified samples and the real results is calculated, and then the category with the largest similarity is selected as the final output result according to the comparison results. The classification process is shown in Figure 1 below.

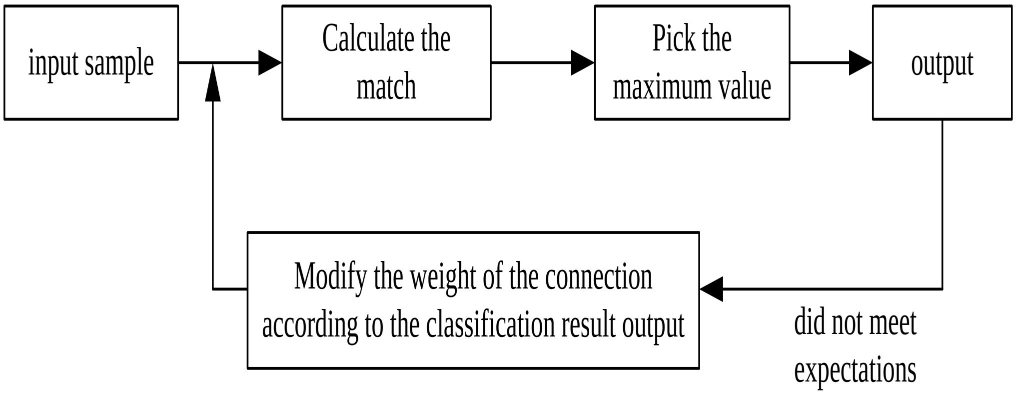

Figure 1. Classifier classification process

The BPNN classifier needs to firstly train the data features before classification. The training mainly uses the process of continuous updating of weights, so that the network model has the function of classification and prediction. The trained BPNN can be applied to sample classification. The samples to be identified are sent to the BPNN classifier, which consists of two steps. First calculate the weight combination of each layer in the network model, output according to the results of the connection weights, and compare the classification results of the classifier with the actual classification results. If the expected value is not reached, the output of the classifier is fed back to the classifier. And correct the connection weights according to the algorithm, and recalculate until the training is completed [19].

4. Improved BPNN Classifier Application Experiment

4.1. Malicious Web Page Classification and Detection

This paper evaluates and analyzes the detection results of malicious web pages from the three indicators of accuracy, recall and precision. The network classifier designed in this paper is to classify and detect web pages by building a network model to verify the classification effect. Before detecting the performance of the malicious web page classifier, it is necessary to classify the data to be detected. The following is the detection of malicious web pages of three sets of data, each set of data contains 1000 samples, and the results are shown in Figure 2.

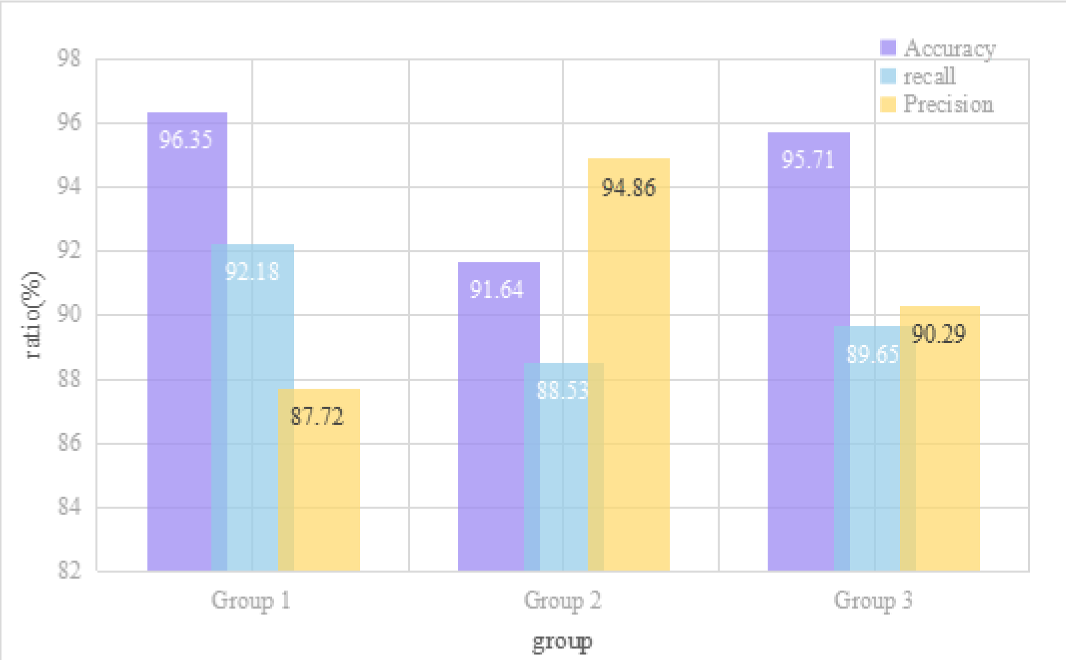

Figure 2. Precision, recall, and precision (%) of three sets of data

According to the data shown in Figure 2, the first group of data has the highest classification accuracy and recall rate, 96.35% and 92.18%, respectively, and the second group of data has the highest classification precision, reaching 94.86%. In order to better describe the performance measured by the classifier, the improved BPNN classifier and the traditional BPNN classifier are used to comprehensively predict the three sets of data to obtain the accuracy, recall and precision of the three datasets. See Figure 3 for details.

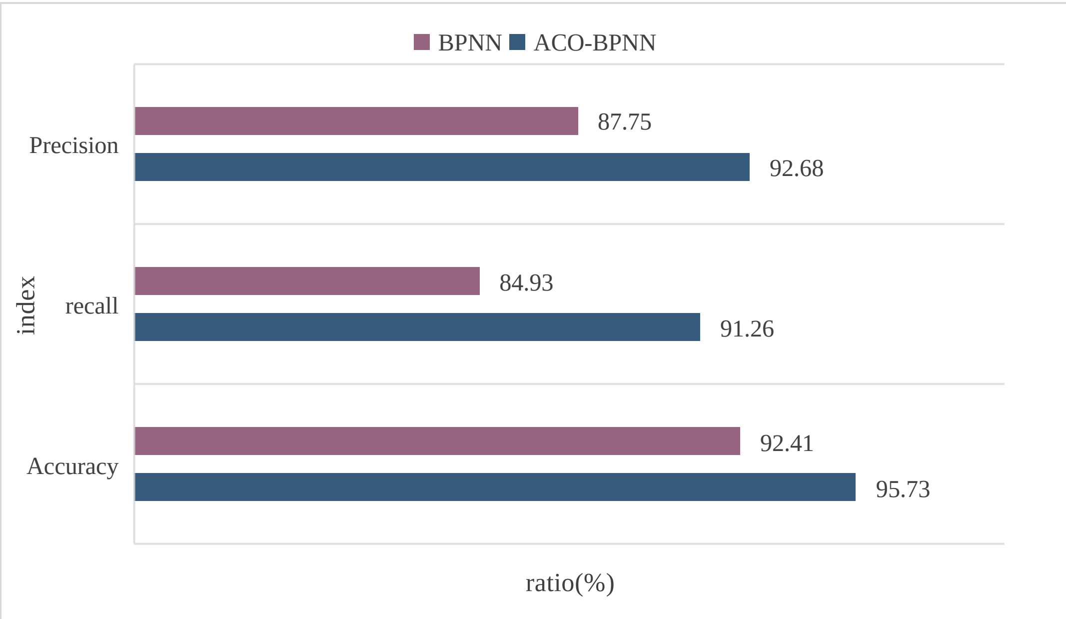

Figure 3. Malicious webpage classification effect of different network models (%)

As can be seen from Figure 3, compared with the performance of the traditional BPNN classification model, the ACO-based BPNN classification model proposed in this paper has higher evaluation indicators for the comprehensive test of the three sets of data, indicating that the performance of the classifier proposed in this paper is equally Outperforms the traditional BPNN classification model. The ACO-BPNN classifier not only ensures the accuracy of the detection model, but also greatly improves the efficiency of the classifier to detect malicious web pages.

4.2. The Influence of Data Sample Size on Classification Effect

Table 1. Dataset sample size

	Dataset name	Sample size
Large sample	Dataset A	522
	Dataset B	781
Small sample	Dataset C	113
	Dataset D	96

This experiment tests the effect of data sample size on the classification effect of NN classifier. As shown in Table 1, data sets A and B are divided into large sample data, and data sets C and D are divided into small sample data. The sample sizes of datasets A, B, C, and D are 522, 781, 113, and 96, respectively.

The classification experiment results are shown in Table 2. In order to more intuitively express the classification effect of the network, Table 4 describes the classification accuracy ACC and standard deviation SDev of the classification results of each dataset, where the value of ACC is the number of correct classification results obtained by the network classification output 10 times and the target The ratio of classification results. The classification accuracy is used to describe the classification accuracy of the classifier, the larger the value, the higher the accuracy; the standard deviation is used to describe the stability of the classifier, the smaller the value, the stronger the stability.

Table 2. Dataset classification experimental results

		Dataset A	Dataset B	Dataset C	Dataset D
ACO-BPNN	ACC	85.16%	83.08%	94.87%	96.32%
	SDEV	1.54	1.25	1.33	1.47
BPNN	ACC	77.35%	74.24%	81.61%	82.56%
	SDEV	1.60	1.37	1.39	1.52

From the analysis of the classification experimental results in Table 2, it can be concluded that the average absolute error of the improved ACO-BPNN classification for each data set is smaller than that of the traditional BPNN classification, and it also obtains a lower average absolute error. Classification standard deviation; whether it is the traditional BPNN classification model or the ACO-BPNN classification model, the classification effect of small-scale data sets is better than that of large-scale data sets, which shows that the classifier system can classify small-scale data. To a certain extent, the classification accuracy is improved. On the whole, the classification effect of the BPNN classifier based on the ant colony algorithm is better.

5. Conclusion

As a traditional NN model, BPNN has good prediction ability in practical applications. However, with the gradual in-depth study of the BP network model, the BPNN also highlights its shortcomings. The algorithm has great stability and prediction accuracy. Limit the application of NN. Therefore, this paper uses the ant colony algorithm to optimize the NN classifier and applies it to the detection of malicious web pages. In the comparison test of the classification effect with the traditional BPNN optimization model, the ACO-BPNN model selected in this paper shows good results in classification accuracy, recall rate, standard deviation, etc., which greatly improves the performance of the classifier.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Alotaibi A, Hamdaoui R. Improvement of Semi-Supervised Document Classification based on Fine Tuning Naive Bayesian Classifier. European Journal of Scientific Research, 2021, 158(3):181-190.

[2] Adams J W, Zhang Z, Noetscher G M, et al. Application of a NN Classifier to Radiofrequency-Based Osteopenia/Osteoporosis Screening. IEEE Journal of Translational Engineering in Health and Medicine, 2021, PP(99):1-1. https://doi.org/10.1109/JTEHM.2021.3108575

[3] Ghazikhani H, Rouhani M. A Deep NN Classifier for P300 BCI Speller Based on Cohen's Class Time-Frequency Distribution. Turkish Journal of Electrical Engineering and Computer Sciences, 2021, 29(2):1226-1240. https://doi.org/10.3906/elk-2005-201

[4] Azzam R, Kong F H, Taha T, et al. Pose-Graph NN Classifier for Global Optimality Prediction in 2D SLAM. IEEE Access, 2021, PP(99):1-1. https://doi.org/10.1109/ACCESS.2021.3084599

[5] Pamela V. An Efficient Intrusion Detection System Using Improved Bias Based Convolutional NN Classifier. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021, 12(6):2468-2482. https://doi.org/10.17762/turcomat.v12i6.5689

[6] Sedova N, Sedov V, Bazhenov R, et al. NN Classifier for Automatic Course-Keeping Based on Fuzzy Logic. Journal of Intelligent and Fuzzy Systems, 2021, 40(2):1-12. https://doi.org/10.3233/JIFS-201495

[7] Gautam S, Singhai J. Cosine-similarity watershed algorithm for water-body segmentation applying deep NN classifier. Environmental Earth Sciences, 2021, 81(9):1-16. https://doi.org/10.1007/s12665-022-10376-y

[8] Oluwole A H , Ositola O M , Olufisayo O , et al. Women's Behavioral Patterns in Domestic Tasks in Western Nigeria: Hazards Forecasting with NN Classifier . Journal of International Women's Studies, 2020, 21(1):241-254.

[9] Masuda H , Nakai T , Yoshida K , et al. Black-Box Adversarial Attack against Deep NN Classifier Utilizing Quantized Probability Output. Journal of Signal Processing, 2020, 24(4):145-148. https://doi.org/10.2299/jsp.24.145

[10] Zadkarami M , Safavi A A , Taheri M , et al. Data driven leakage diagnosis for oil pipelines: An integrated approach of factor analysis and deep NN classifier:. Transactions of the Institute of Measurement and Control, 2020, 42(14):2708-2718. https://doi.org/10.1177/0142331220928145

[11] Shahin I, Nassif A B. Novel Cascaded Gaussian Mixture Model-Deep NN Classifier For Speaker Identification in Emotional Talking Environments. Neural Computing and Applications, 2020, 32(7):2575-2587. https://doi.org/10.1007/s00521-018-3760-2

[12] Giveki D, Rastegar H, Karami M. Erratum to: A New NN Classifier Based on Atanassov's Intuitionistic Fuzzy Set Theory. Optical Memory and NNs, 2019, 28(3):237-237. https://doi.org/10.3103/S1060992X19030111

[13] Abdullah S M. A Deep NN Classifier for Android Malwares Detection using Feature Combination. Journal of Engineering and Applied Sciences, 2019, 14(23):8761-8768. https://doi.org/10.36478/jeasci.2019.8761.8768

[14] Daud K, Abidin A F, Ismail A P, et al. Evaluating Windowing-Based Continuous S-Transform with NN Classifier for Detecting and Classifying Power Quality Disturbances. Indonesian Journal of Electrical Engineering and Computer Science, 2019, 13(3):1136-1142. https://doi.org/10.11591/ijeecs.v13.i3.pp1136-1142

[15] Jayanthi S, Vennila C. Advanced Satellite Image Classification of Various Resolution Image Using a Novel Approach of Deep NN Classifier. Wireless Personal Communications, 2019, 104(1):357-372. https://doi.org/10.1007/s11277-018-6024-7

[16] Hidayat A A, Purwandari K, Cenggoro T W, et al. A Convolutional NN-based Ancient Sundanese Character Classifier with Data Augmentation. Procedia Computer Science, 2021, 179(2):195-201. https://doi.org/10.1016/j.procs.2020.12.025

[17] Anjani S, Ch S. An efficient facial emotion recognition system using novel deep learning nn-regression activation classifier. Multimedia Tools and Applications, 2021, 80(6):1-26. https://doi.org/10.1007/s11042-021-10547-2

[18] Satti S K, Devi K S, Dhar P, et al. Detecting Potholes on Indian Roads Using Haar Feature-Based Cascade Classifier, Convolutional NN, And Instance Segmentation. Soft Computing, 2021, 26(18):9141-9153. https://doi.org/10.1007/s00500-022-07265-8

[19] Madhusudhana S K, Symes L B, Klinck H. A Deep Convolutional NN Based Classifier for Passive Acoustic Monitoring Of Neotropical Katydids. The Journal of the Acoustical Society of America, 2019, 146(4):2982-2982. https://doi.org/10.1121/1.5137323