

Research Article

Effect Evaluation of WTO Dispute Settlement Mechanism Based on Artificial Neural Network

Mengfan Xu 

King's College London, London, UK

Correspondence should be addressed to Mengfan Xu; 5302117076@email.ncu.edu.cn

Received 21 April 2022; Revised 2 June 2022; Accepted 8 June 2022; Published 5 July 2022

Academic Editor: Kuruva Lakshmana

Copyright © 2022 Mengfan Xu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In today's rapid economic development, trade exchanges between countries are increasingly close, and there are inevitable trade frictions that follow as a consequence. The World Trade Organization (WTO) dispute settlement mechanism (DSM) has become a predominant choice to deal with trade frictions for majority of the WTO member states. As part of the process of applying the Sino-US trade friction to the WTO DSM, the rise of US trade protectionism has caused serious damage to the WTO DSM. Also, the DSM similarly exposes its own inherent defects. Hence, the research on improvement of the WTO DSM under the background of trade friction has profound practical value. Based on this background, this paper proposes a framework for the effective evaluation of WTO DSM based on the artificial neural network. The contributions of the paper include the following: (1) Introduction of domestic and foreign scholars' research on different issues in the WTO DSM, wherein the retaliation mechanism is introduced in detail. Also, it is conducive for further deepening of the purpose and significance of the retaliation system in the WTO DSM. (2) The characteristics and structure of the CNN are introduced, and a suitable evaluation system for analyzing the impact of WTO DSM is constructed. Further, an evaluation model based on the CNN is proposed. (3) As per the dimension principle of CNN interlayer calculation, the network structure parameters are selected considering an optimal comprehensive evaluation index of the network. The results verify the effectiveness of the proposed method. Finally, when compared with the other state-of-the-art network models, it is found that CNN generates the highest evaluation accuracy.

1. Introduction

International dispute settlement is an important part of modern international law. The study of international dispute settlement is the study of the laws and procedures aimed at resolving disputes in the international community, especially those designed to avoid resorting to violence to resolve disputes [1]. In the field of international trade, in order to settle disputes between countries in international trade, some special arrangements have gradually emerged in the international community since the middle of the last century. The international DSM that occupies an important seat in the mechanism is the World Trade Organization (WTO) dispute mechanism [2]. The arbitration mechanism included in the WTO DSM does not seem to be well known. The news media more often confuses the WTO "panel" with "arbitral tribunal," treating the panel procedure as an arbitration procedure [3]. Arbitral tribunal or arbitral commis-

sion is a panel which consists of adjudicators who are unbiased in their opinion. The arbitral tribunal is convened, and the members sit to resolve dispute cases through arbitration. The tribunal generally includes a single, two, or multiple arbitrators headed by a chairperson or an umpire. The members in the panel are professionals having expertise in resolving disputes through arbitration as well as friendly mediation. During the process of arbitration, the dispute is submitted to the arbitrator in the arbitral tribunal based on the agreement of both parties. The tribunal panel, which consist of arbitrators, finally gives its decision at the end of the arbitration process pertaining to the dispute case binding both parties. The WTO DSM is settled by the "General Agreement on Tariffs and Trade" (GATT). It is a new type of international DSM based on "Understanding on Rules and Procedures Governing the Settlement of Disputes" (DSU), which integrates various political and legal dispute settlement methods. Much of its success stems from the

successful operation of its unique WTO panel process and the appellate body appellate review process, but this panel process is not the only legal solution provided by the WTO dispute settlement rules. Article 25 of the DSU states that arbitration, as an important part of the WTO DSM, is another alternative legal way to resolve international trade disputes within the WTO system [4]. Different from the international arbitration mechanism in other international treaties, the arbitration system in the WTO DSM is a new type of international arbitration system consisting of different arbitration modes and procedures scattered in the DSU and other WTO agreements. It includes DSU article 25 arbitration, DSB recommendations and arbitration for determining a reasonable time limit in the execution stage of the award and retaliation arbitration as well as special matter arbitration under the “General Agreement on Trade in Services,” the “Agreement on Subsidies and Countervailing Measures,” and arbitration outside the DSU system; these WTO arbitrations have their own characteristics and are unified. Since the establishment of the WTO in 1995, there have been dozens of trade dispute cases resolved by WTO members through the above-mentioned WTO arbitration procedures [5]. Since its formal accession to the WTO on December 11, 2001, China has participated as the complainant, the respondent, and a third party [6]. The number of WTO trade dispute cases is increasing year by year, and the domestic academic circles are increasingly studying the WTO DSM. In the future, with the growth of economy and international trade, China will use various WTO dispute settlement procedures to deal with trade disputes with other WTO members. On the other hand, with the development of the new round of trade negotiations, how to grasp the new development of the WTO DSM, how to actively participate in the new multilateral trade negotiation process, and how to effectively participate in the formulation of multilateral trade rules, all these require a comprehensive understanding and in-depth study of the current WTO DSM [7]. In this context, this paper uses ANN to evaluate the effect of the WTO DSM. In recent years, the successful application of ANNs in digital recognition, speech recognition, machine translation, and other fields has attracted great attention from academia and industry. ANN has the characteristics of self-learning and strong data mining ability. The nonlinear relationship between the evaluation index and the evaluation model realizes the effect evaluation of the WTO DSM based on the ANN.

2. Related Work

Based on a preliminary review of domestic and foreign documents, there are many relevant research documents on retaliation in the WTO DSM, but most of the research content comes from foreign documents, and there are relatively few domestic special research innovations. There is a lack of further research and analysis on the issue of retaliation that has not been resolved for a long time, and the suggestions and countermeasures against China lag behind the times and lack realistic guidance. In terms of research on the retaliation system, Reference [8] introduces the rele-

vant content of the retaliation system, conducts research and analysis on the specific content in each chapter, and analyzes the improvement and perfection of the retaliation system and the establishment of the retaliation authorization implementation mechanism. Reference [9] compares and analyzes trade retaliation in GATT1947 and the WTO DSM and puts forward suggestions for improving foreign trade legislation in retaliation. Reference [10] pointed out that retaliation in lieu of enforcement will face issues of legitimacy and the stability and predictability of the WTO. The General Agreement on Tariffs and Trade (GATT) was signed on October 30, 1947, by 23 countries. It was a legal agreement that helped to minimize the barriers pertaining to international trade by completely eliminating quotas, tariffs, and subsidies ensuring significant protocols and regulations. The main objective was to revive the sinking economy post-World War II by liberalizing global trade. Reference [11] analyzes the flaws in the WTO system, strengthens the review of options for retaliation, and establishes recommendations for collective retaliation. In terms of research on specific issues, Reference [12] conducts a systematic legal and economic analysis based on the arbitration report of the US-Highland cotton subsidy case and points out the current difficulties and problems of WTO retaliation. Reference [13] discusses the nature and principles of cross-retaliation in combination with the EC-Bananas III case, the US-Gambling case, and the US-Upland Cotton case; analyzes the dilemma in practice; and puts forward suggestions to strengthen the supervision of the enforcement of the ruling and the combination of cross-retaliation. Reference [14] explains the provisions and principles of Article 22.3 of the DSU on choosing the form of retaliation and the analysis of the determination of infeasibility or invalidity, and the situation is serious enough in arbitration with reference to cases. Article 22.3 states the fact that in general, any retaliation should include suspensions of obligations and concessions that have an effect on the same sector or others falling under the same agreement. However, in case it is found to be unsatisfactory, concessions and obligations falling under another agreement would be suspended as part of cross-retaliation. Reference [15] affirms cross-retaliation while paying attention to potential negative factors. Reference [16] explains in detail the provisions on the implementation of the compliance review procedure and the retaliation procedure, analyzes the problems of the application sequence and the practices of the members in practice, and gives suggestions on the revision of the provisions. In terms of the resolution of retaliation, Reference [17] analyzes the value-oriented theory of WTO dispute settlement in domestic academic circles from the perspective of implementing dispute settlement in the fastener case. Reference [18] analyzed the enforcement dispute in the kretek cigarette case in the United States. Indonesia proposed to authorize the suspension of concessions. The United States refuted it on the grounds of public health. For the first time, social and public interests were taken into consideration, which led to reflection on the retaliation system. Reference [19] proposed the significance of cross-retaliation to developing countries and the importance of doing research on cross-retaliation.

Reference [20] studies and analyzes the dilemma of retaliation against developed countries based on the enforcement of arbitral award retaliation and cross-retaliation in the United States-Gaming case. Reference [21] studies the advantages of the monetary compensation system and points out that issues such as the determination of the monetary amount, retroactivity, and procedural principles should be fully considered in the design of the system. Some scholars believe that in the WTO DSM, developing member countries are never rule-makers but only participants in the system. References [22, 23] point out that GATT proceedings tend to focus more on principle than specifics; by contrast, the WTO focuses on the nuances of factual disputes. Foreign countries mostly discuss and study the alternation of Sino-US trade frictions in the form of news and seldom discuss it in the context of the WTO Department of Administration and Management (DAM). Reference [24] believes that the trade imbalance between China and the United States is one of the main reasons for the trade friction between the two countries, and the trade imbalance is caused by the lack of transparency of China's policies and the failure to fully implement the agreements signed by the two countries. References [25, 26] discuss institutional factors in the strategic trade theory, emphasizing that one factor in international trade friction may be the difference in economic institutions between countries and, at the same time, expressing that structural barriers and institutional differences between countries are an important factor in friction. The study in [27] presented a framework for conflict resolution that employed machine learning techniques for conflict resolution using negotiation and mediation. The study incorporated artificial intelligence and multiattribute-based decision theoretic techniques that enabled conflict resolution with an optimum level of efficiency. The study in [28] used multiple machine learning techniques to resolve public-private partnership project dispute issues. The machine learning techniques used included multilayer perceptron (MLP) neural networks, decision trees, support vector machines (SVM), naïve Bayes, and k -nearest neighbor algorithms wherein the MLP and decision tree models yielded optimum accuracy. The study in [29, 30] used Blockchain as part of a smart contract for contract management between the data owner and data purchaser. To sum up, by studying the problems in the WTO DSM, and then analyzing the solutions to the problems, it can provide reference value for China to formulate legal and effective strategies in participating in international disputes. Combined with the research on developing countries' response to the retaliation system in foreign literature, as well as the opinions and suggestions of domestic experts and scholars, this paper analyzes how countries can strengthen the capacity building of participating in the WTO DSM, so as to better safeguard national rights and interests through the WTO DSM.

3. Method

3.1. Network Basic Structure. ANN is an information processing system developed by simulating neuron signal transmission in the biological nervous system. The ANN is

composed of several perceptrons according to a certain connection mode. The perceptron is the basic unit of the network, also known as the neuron. The basic structure of the perceptron is generally t inputs and 1 output, and the specific structure is shown in Figure 1. The calculation method of the output y of the perceptron is usually as follows: multiple inputs of a specific perceptron and their corresponding weights are multiplied and then summed, and the value obtained by the summation is used as the input of the activation function f . However, the output y of the perceptron is also related to the perceptron itself, so an influence bias p is usually taken into account after calculating the summation, which is called the bias of the perceptron.

Based on the description of the above principle, the output of this neuron can be expressed as

$$y = f\left(\sum Q_i W_i + p\right), \quad (1)$$

where f is the activation function and the calculated output value y is compared with the threshold value of the perceptron.

If the output value is greater than the threshold value, the perceptron state is an active state in the network. On the contrary, the perceptron is in an inhibited state in the network and does not participate in the training of the network. Therefore, the state of the neuron can be determined by the training of the network, and the network also uses this rule to filter out the important unit nodes related to the output.

The degree to which the perceptron is activated or inhibited in the network is related to the type of activation function. There are three commonly used activation functions: sigmoid function, tanh function, and ReLU function. The first two activation functions are generally used in network structures without deep layers, and the third activation function is generally used in DNNs to avoid gradient disappearance during the training process and affect the convergence of the network. Connecting multiple perceptrons according to a certain architecture can form a simple multilayer perceptron. Each layer consists of several perceptrons. The perceptrons in each layer are independent of each other, and all the perceptrons in the upper layer are one way with all the perceptrons in the lower layer. Connected, multilayer perceptrons of this structure are also known as feedforward neural networks. The feedforward neural network usually consists of the input layer, hidden layer, and output layer; the specific structure is shown in Figure 2. The feedforward neural network learns the relationship between input and output and iteratively updates the training process to continuously approximate the nonlinear relationship between the two. Its good fitting effect is suitable for solving nonlinear separable problems.

The multilayer perceptron receives the external signal as the input quantity X and transmits the information to the neurons of the next layer through the connection between neurons. The hidden layer transforms the information through the learning algorithm and then transmits the transformed information to the output layer. The calculation method of the output quantity can be expressed as

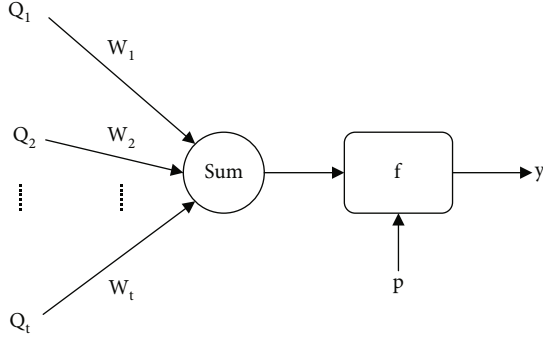


FIGURE 1: Perceptron structure.

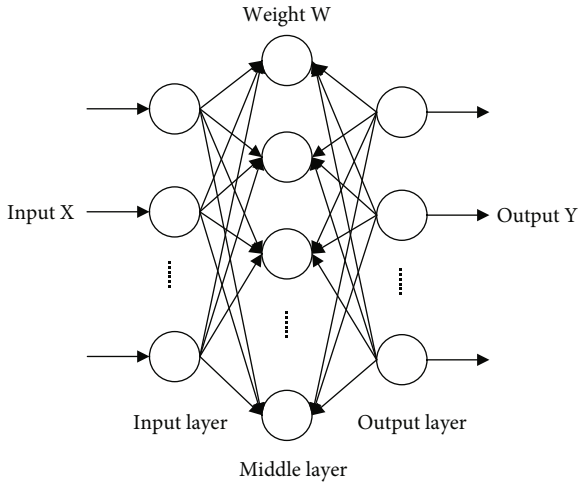


FIGURE 2: Multilayer perceptron architecture.

$$Y = f(WX + p), \quad (2)$$

where X is the input matrix, W is the weight matrix of the interconnected neurons, and p is the bias term.

3.2. Regularization Techniques. A vital step in training a deep learning model with more parameters than the training dataset is to use regularization technology to assure the algorithm's capacity to generalize. The introduction of regularization technology can avoid the problem of algorithm overfitting. Regularization techniques mainly include L_1 regularization, L_2 regularization, and dropout techniques.

3.2.1. L_1 Regularization Technique. The idea of regularization is to add an indicator that describes the complexity of the model to the loss function of the machine learning model, and the regularized model loss function is shown in the following formula.

$$\varphi = \arg \min_{\varphi} \frac{1}{N} \sum_{i=1}^N (L(\hat{y}_i, y) + \mu R(\omega)), \quad (3)$$

where φ is the objective function that minimizes the loss function, \hat{y}_i is the predicted value of the sample, y is the actual value of the sample, L is the square of the difference

between the predicted value and the actual value of the i th sample, $R(\omega)$ is the added regular term, and μ is the penalty coefficient of the regular term.

L_1 regularization adds a regularization term $R_{L_1}(\omega)$ to the loss function to reduce the sum of the absolute values of the parameters, and the regularization terms are shown in the following formula.

$$R_{L_1}(\omega) \triangleq \sum_{m=1}^M \|W\|_1, \quad (4)$$

where $\|W\|_1$ is the L_1 norm of the weights in the network.

The L_1 norm regular term can make the weight corresponding to some nodes equal to 0 through the penalty factor of the network. Therefore, L_1 regularization can perform feature selection on the data, making the model sparsity. That is, when applying L_1 regularization to train the network, the original m feature variables only have less than m features that contribute to the training of the network. Therefore, the L_1 regularization technique realizes feature selection to a certain extent and reduces the complexity of the model, and the formed sparse model is helpful for the convergence of the network.

3.2.2. L_2 Regularization Technique. In L_2 regularization, the purpose of adding the regularization term $R_{L_2}(\omega)$ is to reduce the sum of the parameter squares, and the regularization terms are shown in the following formula.

$$R_{L_2}(\omega) \triangleq \|W\|_2^2, \quad (5)$$

where $\|W\|_2^2$ is the L_2 norm of the weights in the network.

The L_2 regularization technique makes the solution of the model biased towards the weight W with a smaller norm through continuous iteration and limits the space of the model by limiting the size of W , so that the overfitting phenomenon that occurs during network training can be solved in most cases. However, this regularization method does not have the ability to generate sparse solutions, and the obtained coefficients still require all the features in the data to calculate the predicted results, which has not improved in terms of computational complexity. Therefore, L_1 regularization techniques are more widely used in feature selection and avoiding overfitting problems.

3.2.3. Dropout Technology. Dropout technology is a regularization method to prevent overfitting. Its main idea is to randomly select some neurons to not work when training the network, the weights of the selected neurons are still the weights saved in the previous iteration, and the output of the neuron is set to 0. The neuron selected this time still retains the weights of the previous iteration in the next iteration. Randomly selected neuron weights in repeated iterations follow this process. In the process of each iteration, the network structure will change to a certain extent, and the network structure change is shown in Figure 3.

After applying dropout technology to a neural network with N nodes, the original network becomes a set of $2n$

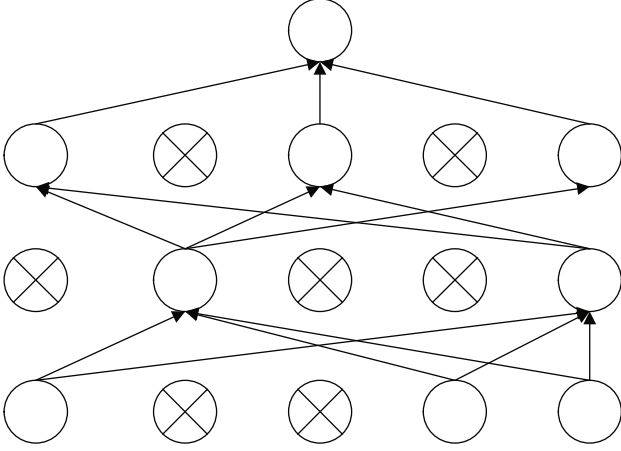


FIGURE 3: Dropout network structure.

models. At this time, although the number of parameters to be trained is the same as that in the network before application, the time cost of network training is greatly reduced. Dropout technology forces one of the neurons to participate in training with randomly selected neurons, which weakens the dependence between neurons in this way and also weakens the influence of some neurons on the network output, thus preventing overfitting, thus enhancing the generalization ability of the network.

3.3. Convolutional Neural Networks

3.3.1. Network Structure and Principle. Through the study of cat visual cortex cells, researchers found that the complexity of the neural network feedback mechanism can be effectively reduced in the deep network architecture. Based on this discovery, the concept of CNN was proposed. One of the first models of CNN was a multilayer feedforward neural network or neural perceptron. Weight-sharing and local connection are two of the network's properties. Because of CNN's unique structure and nature, which resembles biological neural networks more closely, the network model may be simplified and the number of parameters required for training the network can be reduced. Convolutional and pooling layers are used for feature extraction in the network's fundamental structure; these layers are located in the feature extraction layer. Usually, the sample takes the form of a two-dimensional matrix as the input of the network, and then through the local learning of the convolution kernel in the convolution layer, the feature matrix of the input sample is obtained, and the pooling matrix and the feature matrix in the pooling layer are pooled. That is, the next layer of the convolutional layer is the pooling layer, the next layer of the pooling layer is the convolutional layer, and so on. Local feature extraction and abstract learning of raw data is conducted through the deep architecture of the network. The special network structure of CNN has a high fault tolerance for input data, so that it can more accurately express and abstract the data. The convolutional layers of CNN are composed of different feature matrices. Each feature matrix is learned by a convolution kernel, and multiple

convolution kernels can be set to learn in parallel to learn different features of the data. The calculation method of the feature matrix is shown in the following formula.

$$h_{1,\alpha} = f(x^* W^{1,\alpha} + p^{1,\alpha}), \quad (6)$$

where $h_{1,\alpha}$ and $W^{1,\alpha}$ represent the weight matrix of the α th convolution surface and the α th convolution kernel of H_1 , respectively; $p^{1,\alpha}$ is the bias; f is the activation function; and the activation function in this paper adopts the sigmoid function.

The pooling layer of a CNN, also known as a sampling layer, is usually the next layer to a convolutional layer. Similar to the convolutional layer, the pooling matrix of the pooling layer is also composed of multiple feature matrices and corresponds to the feature matrix calculated by the feature matrix of the convolutional layer, so it is unique. The pooling layer consists of all pooling matrices. Commonly used pooling methods include max pooling, random pooling, and mean pooling. This paper adopts the mean pooling method. The pooling layer of CNN downsamples the output feature matrix of the previous layer to screen out important features, thereby playing the role of secondary feature extraction. The pooling layer can further reduce the number of parameters that need to be trained in the network, is less prone to overfitting, and maintains the scale invariance of features to a certain extent. The calculation method is shown in the following formula.

$$h_{2,\alpha} = g(\beta_2 \text{down}_{\sigma_2 \varepsilon_2}(h_{1,\alpha}) + \chi_2), \quad (7)$$

where $h_{2,\alpha}$ represents the α th pooling matrix of H_2 , the weight β_2 generally takes the value of 1, the bias χ_2 generally takes the 0 matrix, and $g(x) = x$.

The feature of the fully connected layer of CNN is that the neurons of the fully connected layer are all connected to the neurons of the previous layer in one direction, which is used to integrate the local features extracted layer by layer through the feature extraction layer, and the function of this layer is to add a bias term to each neuron. By training a large number of samples, the important local features of the data are learned, and the output value of the final fully connected layer is passed to the output layer after the activation function. The output layer of CNN is also called the Softmax layer, which is a commonly used classifier for CNN. Usually, after the last pooling layer, the learned local features need to be mapped to implement the classification problem of the input data. The function expression of Softmax is as follows:

$$F(i) = \frac{e^{-\delta_i^T x}}{\sum_{i=1}^k e^{-\delta_i^T x}}. \quad (8)$$

In the classification problem, the Softmax function maps $F(i)$ in the range of $[0, 1]$ and outputs the category of the predicted sample according to the set threshold. In formula (8), δ is the parameter to be determined, and the optimal parameter when $F(i)$ is maximized is searched through the training of the network.

3.3.2. Training Algorithm of CNN. As a supervised deep learning model, the training process of CNN is similar to that of the feedforward neural network. The training network is divided into two stages: forward calculation and back propagation. First, the forward calculation process is to use the current parameters to represent the input samples in the dataset and then perform transfer calculations based on the current parameters to learn the error function between the predicted output and the actual sample label in the last layer of the network; secondly, in the reverse direction, the propagation process uses the chain rule to calculate the gradient of each layer parameter and then updates the parameter gradient in the opposite direction.

3.3.3. Network Parameter Selection. The convolution kernel in the convolution layer of the CNN and the window dimension of the pooling matrix in the pooling layer should be determined according to the dimension principle: the input matrix is subjected to the convolution operation of the convolution layer to obtain the output matrix of this layer, and the dimension of the input matrix is $m \times s$, the convolution kernel window is $k \times k$, and the dimension of the output matrix is $n \times v$, and the dimensional relationship of the three should satisfy $n = m - k + 1$, $v = s - k + 1$. The dimension $j \times j$ of the pooling matrix in the pooling layer is determined by the output matrix dimension $n \times v$ of the previous layer. It is required that the dimension of the pooling matrix can be divisible by the number of rows and columns of the output matrix; that is, the dimension relationship $r = n/j$, $u = v/j$, where $r \times u$ is the dimension of the output matrix obtained by the pooling operation. Network parameter selection sets the number of layers of the network at first; on this basis, all the window combination parameters that satisfy the dimension principle are found; that is, we search for the combination of the convolution kernel and the window dimension of the pooling matrix, and the obtained window set can be expressed as

$$T = \{(k_1, j_1), (k_2, j_2), \dots, (k_t, j_t)\}, \quad (9)$$

where T is the total number of window combinations that satisfy the window dimension principle.

A predetermined ratio is used to split the data sample set into training and testing sets. The model is trained using the data in the training set. The test set is used to evaluate the model's evaluation performance under a specific set of window combination settings that are utilized in each training session. The optimal evaluation index is the goal to determine the optimal network structure parameters.

3.3.4. WTO Dispute Settlement Mechanism Effectiveness Evaluation Index. The WTO DSM is both a judicial and a political DSM, and it is a complete one. Rather than merely defining the parties' success or failure in relevant instances or sanctioning a party, its primary goal is to effectively settle trade disputes and restore and preserve the balance of important rights and duties between the parties to the dispute. The following are some of its most distinguishing features: the "reverse consensus principle" should be estab-

lished. To put it another way, the "reverse consensus" rule of the DSB is entirely different from the GATT1947 rule of "forward consensus." DSB adheres to the "reverse consensus principle," which means that a decision on a procedural item may be approved by just one member's positive vote, such as the admission of cases. In a roundabout way, this shows how the WTO's dispute resolution process has been expanding its authority over conflicts. Secondly, a strict time limit for resolving disputes is set. Although DSB is not a judicial institution, it has rigorous and clear procedural limitations on the time limit for hearing cases in order to maintain the effectiveness of conflict settlement recognized by DSB. Every disagreement case must be finished within 15 months, and urgent matters must be resolved within three months. As a result, the WTO dispute resolution mechanism's decisions may be put into action even more rapidly and efficiently. As a result, as indicated in Table 1, this research develops an appropriate assessment method based on the features of the aforementioned WTO dispute resolution procedure. The output set in this work has three grades of A, B, and C, with grade A being the best effect and grade C representing the worst impact according to the evaluation index system created.

4. Experiment and Analysis

4.1. Neural Network Model and Parameters. According to the evaluation indicators of the effect of the WTO DSM, this paper designs a dataset, which contains 500 sets of data, of which 400 sets are used as training sets and 100 sets are used as test sets. At present, the parameter determination of CNN still lacks a clear guiding theory and still relies on manual experience. It is necessary to continuously adjust the parameters and determine an appropriate value through comparison. The model in this chapter consists of five convolutional layers, five pooling layers, and three fully connected layers. The output dimensions of each layer of the fully connected layer are 80, 20, and 4 in turn. For the first two layers of the fully connected layer, dropout with a parameter of 0.15 is performed. The convolution kernel used in the first layer of the convolution kernel in this paper is 35. At present, the parameter adjustment of the neural network still relies on experience and constant comparison and adjustment for setting. The following is the comparison and selection process of some main parameters of the model.

- (1) Hyperparameters of the fully connected layer after the convolutional layer and the pooling layer extract features from the data, the fully connected layer is usually used to combine these features, and the output value is finally given. If there are too few neurons in the fully connected layer or the depth is too shallow, it is easy to cause the model to not learn enough features and it is difficult to fit the data; if there are too many neurons or the number of layers is too deep, it is easy to cause the model to overfit, which reduces the generalization ability of the model. Although the number of fully connected layers and the number of neurons in each layer cannot be tested one by one, it is still possible to select suitable

TABLE 1: WTO dispute settlement mechanism effectiveness evaluation index system.

Index	Label
Accept consultation or mediation procedures	X1
Quick and efficient dispute resolution	X2
Both parties to the dispute accept the ruling	X3
The losing party fulfills its adjudication obligations on time	X4
No retaliation occurred	X5
No appellate body review	X6

hyperparameters through various attempts. In this paper, several hyperparameter combinations are selected for comparison, as shown in Figure 4. Through the comparison of the test set accuracy of the models under different hyperparameters, it is reasonable to set the fully connected layer to three layers. Among them, the test set has the highest diagnostic accuracy under the hyperparameter combination of 80-20-4. The 80-20-4 refers to the output dimensions of each layer of the fully connected layer as 80, 20, and 4 in turn

- (2) *Minibatch parameter*: the application of minibatch technology can improve the convergence speed of the model. The parameter batch size is the number of samples used in training. When the batch size is too large, it is not much different from not using the minibatch technology; when this value is too small, it will make it difficult for the model to converge, resulting in insufficient fitting accuracy. In this section, we select 16, 32, 64, 128, 192, 256, and 7 cases without minibatch technology, set a fixed number of 40 epochs, and compare the convergence process (see it in Table 2)

The comparison results are shown in Table 2. According to the convergence speed of the model, the final batch-size value is 64.

- (3) *Model parameter verification results*: this paper uses MATLAB to train the neural network, and the training set and test set of the data are divided according to 4:1. During training, the minibatch size is 64, the dropout coefficient is taken as 0.15, and the L_2 regularization coefficient is taken as 0.02. The number of training rounds is 20. During the training process, the accuracy of the training set and the test set is shown in Figure 5

Among them, the accuracy rate of the training set finally reached 0.995, and the accuracy rate of the test set also converged above 0.95. It can be seen that the model has extremely high accuracy in predicting the effect of the WTO DSM.

4.2. Comparison of Effect Prediction Performance of Different Algorithms. In order to verify the accuracy of CNN in the evaluation of the effect of the WTO DSM, different model

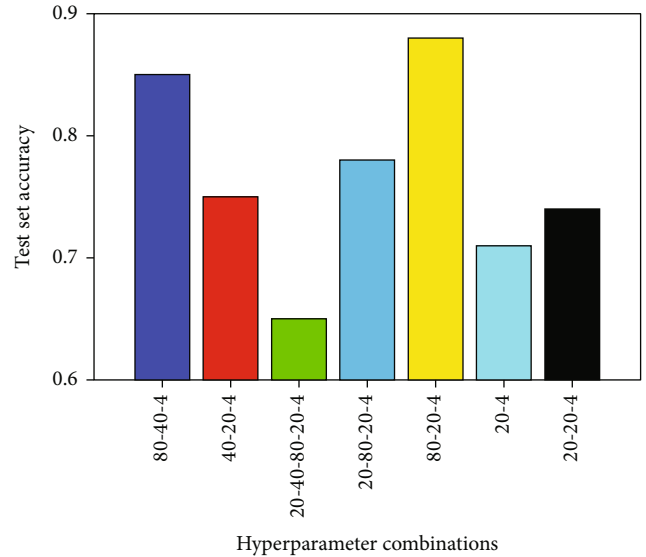


FIGURE 4: Fully connected layer hyperparameter combination and test set accuracy.

TABLE 2: Minibatch parameters and their comparison.

Batch size	Time required for 40 epochs (s)	Time required for the accuracy to converge to 95% (s)
16	46	15.8
32	42	8.5
64	32	8.1
128	35	9.2
192	37	10.8
256	40	11.2
None	42	12.6

algorithms are tested on the same dataset, and the evaluation performance of CNN is compared with SVM, DT, and ANN. Among them, the default kernel function of SVM adopts RBF, and the optimal structural parameters are obtained by 5-fold cross-validation and the grid search method. The value range of both is $[2^{-8}, 2^8]$, the penalty function $C = 12$ is obtained, and the kernel parameter $\gamma = 12$. DT adopts the C4.5 algorithm, and the confidence factor adopts the default value of 0.25, which has a good effect in most scenarios. The number of neurons in the input layer of ANN is equal to the number of values in the input matrix, the number of neurons in the output layer and the number of output categories are equal to 2, the number of layers in the hidden layer is set to 1, and the number of neurons is determined by the traversal method as 18. The training algorithm is the same as that of CNN, both of which are gradient descent. The evaluation results of different evaluation models are shown in Figure 6.

As can be seen from Figure 6, the evaluation accuracy rate of CNN is more than 10% higher than that of ANN, indicating that compared with ANN with only a single hidden layer, CNN can effectively extract the key features hidden in the data by using its own deep structure; the

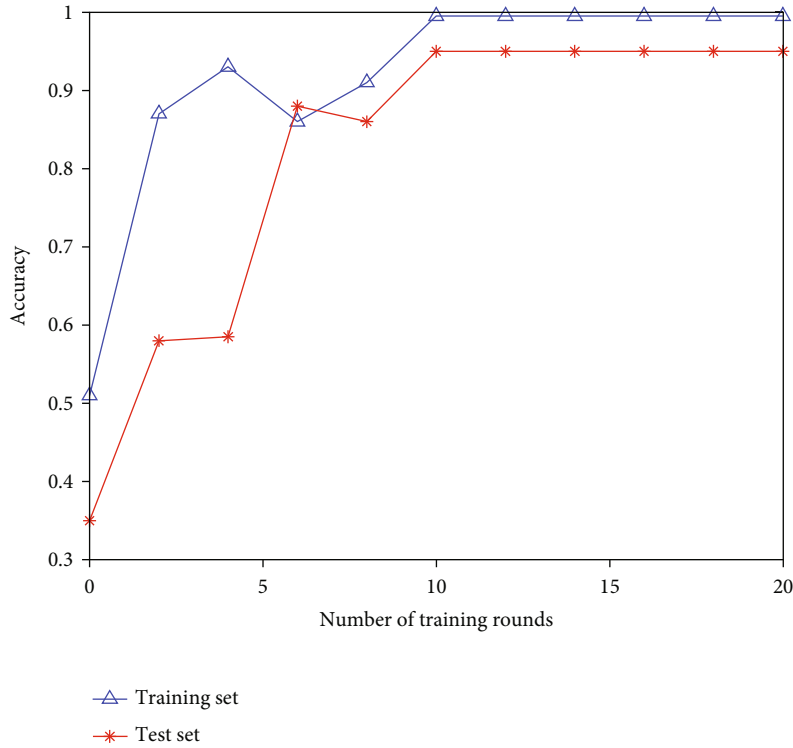


FIGURE 5: Evaluate accuracy of CNN during training.

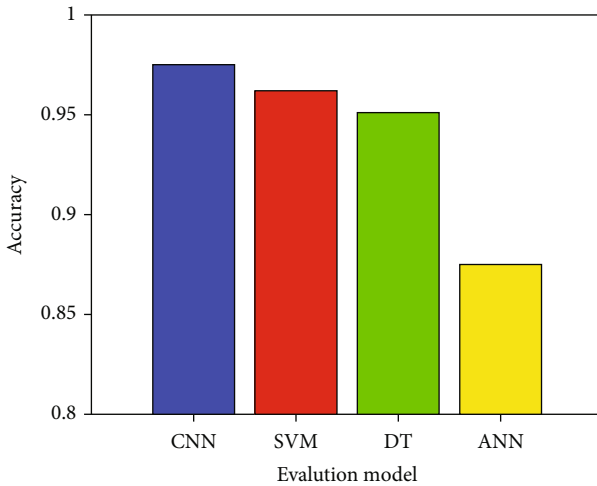


FIGURE 6: Evaluation results of different evaluation models.

ability of abstract representation of data is better than that of shallow learning models. The evaluation accuracy and comprehensive evaluation index of CNN are better than the other three evaluation models, so the performance of CNN is better than the other three shallow evaluation methods.

5. Conclusion

Emerging economies have grown swiftly as a result of economic globalization, and the global economy's center of gravity has changed. The relevance of the WTO DSM is self-evident in the contemporary international economic

and commercial setting. DSM is aimed at protecting and promoting the fundamental principles of global commerce, as well as the integrity of the global economy. China supports the reform of the WTO's DSM, but it must always adhere to the core values and basic principles of the WTO, such as upholding openness, inclusiveness, and nondiscrimination, and adhere to extensive consultations among all parties and proceed in an orderly manner. The rule-oriented global multilateral trading system is the goal of economic globalization and the development interests of most countries. Therefore, the WTO DSM should take the multilateral trading system as the core and the cooperation of all member states as the premise. The evaluation of the effect of the WTO DSM based on ANN caters to the actual needs of the world's exchanges and development and is a great attempt to build a multilateral trading system. Therefore, this paper has completed the following tasks: (1) Introduce the research on different problems in the WTO DSM by scholars at home and abroad and then analyze the solutions to the problems. It introduces the retaliation mechanism in detail, which is conducive to further deepening the purpose and significance of the retaliation system in the WTO DSM and is conducive to clarifying the provisions and principles of the articles and the problems in practice. (2) The characteristics and structure of CNN are introduced, and an evaluation model based on CNN is proposed. Then, based on the characteristics of the WTO DSM, this paper constructs a suitable evaluation system for the effect of the WTO DSM. (3) According to the dimension principle of CNN interlayer calculation, the network structure parameters are selected with the optimal comprehensive evaluation

index of the network. The results verify the effectiveness of the proposed method. The proposed CNN-based model was compared with the state-of-the-art techniques, namely, SVM, ANN, DT, and ANN, considering the network window parameter as 80-20-4. The evaluation results highlighted the superiority of the CNN model yielding higher accuracy. As part of future study, various other metrics could be considered as part of the evaluation process. Also, to further strengthen the explainable inferences, AI could be implemented.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The author declares that he has no conflict of interest.

References

- [1] M. E. O'Connell, *International Dispute Resolution: Cases and Materials*, Carolina Academic, 2012.
- [2] J. Merrills, *International Dispute Settlement*, Cambridge University Press, New York, 4th edition, 2005.
- [3] Y. Iwasawa, "WTO dispute settlement as judicial supervision," *Journal of International Economic Law*, vol. 5, no. 2, pp. 287–305, 2002.
- [4] C. D. Creamer and Z. Godaimirska, "(De)legitimation at the WTO dispute settlement mechanism," *Vanderbilt Journal of Transnational Law*, vol. 49, p. 275, 2016.
- [5] R. A. Torres, *Use of the WTO trade dispute settlement mechanism by the Latin American Countries Dispelling Myths and Breaking Down Barriers*, World Trade Organization (WTO), 2012.
- [6] J. L. Gertler, *What China's WTO accession is all about*, China and the WTO, 2004.
- [7] H. Gao and C. L. Lim, "Saving the WTO from the risk of irrelevance: the WTO dispute settlement mechanism as a 'common good' for RTA disputes," *Journal of International Economic Law*, vol. 11, no. 4, pp. 899–925, 2008.
- [8] S. Shadikhodjaev, *Retaliation in the WTO Dispute Settlement System*, Kluwer Law International BV, 2009.
- [9] B. Mercurio, "Why compensation cannot replace trade retaliation in the WTO dispute settlement understanding," *World Trade Review*, vol. 8, no. 2, pp. 315–338, 2009.
- [10] K. Anderson, "Peculiarities of retaliation in WTO dispute settlement," *World Trade Review*, vol. 1, no. 2, pp. 123–134, 2002.
- [11] H. Spamann, "The myth of 'rebalancing' retaliation in WTO dispute settlement practice," *Journal of International Economic Law*, vol. 9, no. 1, pp. 31–79, 2006.
- [12] R. Read, "46 dispute settlement, compensation and retaliation under the WTO," in *Handbook on international trade policy*, vol. 497, p. 501, Edward Elgar Publishing, 2007.
- [13] L. E. F. A. Spadano, "Cross-agreement retaliation in the WTO dispute settlement system: an important enforcement mechanism for developing countries?," *World Trade Review*, vol. 7, no. 3, pp. 511–545, 2008.
- [14] M. Moinuddin and V. Sengsavang, "WTO dispute settlement and the problems of compliance: does cross-retaliation under TRIPS provide a remedy?," *Yokohama Journal of Social Sciences*, vol. 15, no. 4, 2010.
- [15] M. Bronckers and N. Van den Broek, "Financial compensation in the WTO," *Journal of International Economic Law*, vol. 8, no. 1, pp. 101–126, 2005.
- [16] G. Sacerdoti, "The nature of WTO arbitration on retaliation," *The Law, Economics and Politics of Retaliation in WTO Dispute Settlement*, vol. 3, p. 23, 2010.
- [17] K. K. Sharma, *Dispute settlement mechanism of World Trade Organization: a critical study in the context of developing countries*, [Ph.D. thesis], Sikkim University, 2018.
- [18] S. Picciotto, *Defending the public interest in TRIPS and the WTO*, Global Intellectual Property Rights. Palgrave Macmillan, London, 2002.
- [19] M. E. Footer, "Developing country practice in the matter of WTO dispute settlement," *Journal of World Trade*, vol. 35, no. 1, pp. 55–98, 2001.
- [20] J. Li, Z. Zhou, J. Wu et al., "Decentralized on-demand energy supply for blockchain in Internet of Things: a microgrids approach," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 6, pp. 1395–1406, 2019.
- [21] W. Duan, J. Gu, M. Wen, G. Zhang, Y. Ji, and S. Mumtaz, "Emerging technologies for 5G-IoV networks: applications, trends and opportunities," *IEEE Network*, vol. 34, no. 5, pp. 283–289, 2020.
- [22] G. N. Horlick, *WTO and NAFTA rules and dispute resolution: selected essays on antidumping, subsidies and other measures*, Cameron May, 2003.
- [23] K. Lakshmana, R. Kaluri, N. Gundluru et al., "A review on deep learning techniques for IoT data," *Electronics*, vol. 11, no. 10, p. 1604, 2022.
- [24] M. Bronckers, "Financial compensation in the WTO: improving the remedies of WTO dispute settlement," *Social Science Electronic Publishing*, vol. 8, no. 1, pp. 101–126, 2008.
- [25] J. Pauwelyn, "Enforcement and countermeasures in the WTO: Rules are rules—toward a more collective approach," *American Journal of International Law*, vol. 94, no. 2, pp. 335–347, 2000.
- [26] N. Gundluru, D. S. Rajput, K. Lakshmana et al., "Enhancement of detection of diabetic retinopathy using Harris hawks optimization with deep learning model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8512469, 13 pages, 2022.
- [27] K. P. Sycara, "Machine learning for intelligent support of conflict resolution," *Decision Support Systems*, vol. 10, no. 2, pp. 121–136, 1993.
- [28] J.-S. Chou, C.-F. Tsai, and L. Yu-Hsin, "Project dispute prediction by hybrid machine learning techniques," *Journal of Civil Engineering and Management*, vol. 19, no. 4, pp. 505–517, 2013.
- [29] W. Xiong and L. Xiong, "Smart contract based data trading mode using blockchain and machine learning," *IEEE Access*, vol. 7, pp. 102331–102344, 2019.
- [30] G. T. Reddy, M. P. K. Reddy, K. Lakshmana et al., "Analysis of dimensionality reduction techniques on big data," *IEEE Access*, vol. 8, pp. 54776–54788, 2020.