Systematic Combat Effectiveness Evaluation Model Based on Xgboost

Bei Xiao, Peng-Cheng Luo, Zhi-Jun Cheng, Xiao-Nan Zhang, Xin-Wu Hu
Department of management science and engineering
National University of Defense Technology
Changsha,China
975312616@qq.com

***Abstract*—Systematic combat effectiveness evaluation is an important part of system combat research, which is of special significance for our army to adapt to modern war. The scale of systematic combat is large and the structure is complex, so it is difficult to complete the evaluation work with the traditional empirical method or mathematical method. In recent years, the simulation method which is widely used has achieved good results, but there are also problems such as huge computational cost. On the basis of system combat simulation, this paper uses xgboost to build an intelligent evaluation model of system combat effectiveness, which effectively solves the problem of traditional methods in calculating costs. Compared with the SVM method, the good performance of this method is proved.**

Keywords—Systematic combat effectiveness,Xgboost, Intellige nt modeling technology Introduction

I. INTRODUCTION

In recent years, systematic combat effectiveness evaluation has also begun to explore the field of artificial intelligence. Through the methods of machine learning and deep learning, the researchers used a large amount of data to train a good model, which can effectively reduce the effectiveness evaluation of time overhead. At the same time it also can ensure the good accuracy. This becomes the new research direction of the system combat effectiveness evaluation.

With the continuous development of artificial intelligence, many scholars at home and abroad have begun to study the evaluation of combat effectiveness by means of machine learning or deep learning. Cheng Kai[4] set up the force effectiveness evaluation model based on SVM in 2011, completed the nonlinear mapping from index to the efficiency value. But the simulation case is simple, which does not reach the complexity of the system operation. Shi Yanbin[5] used BP neural network to evaluate the effectiveness of a certain type of combat equipment and achieved good evaluation results. However, there are some problems. For example, this method cannot ignore the dependence on the initial weights and thresholds of the neural network, which is still limited in practical application. In 2009, liu ling[6] applied LS-SVM method to establish the air combat effectiveness evaluation model of modern fighters. Good results have been achieved. However, its research still needs to establish the index evaluation system, which has certain limitations.

Based on the large amount of data generated by current simulation, this paper constructs an intelligent evaluation model of systematic combat effectiveness through xgboost algorithm. This model can be effectively applied to complex simulation data, which can effectively reduce the amount of calculation on the basis of high precision, and is of great significance to the development of system warfare.

II. SYSTEMATIC COMBAT EFFECTIVENESS EVALUATION MODEL

A. *Description and hypothesis of the problem*

At present, the widely used system combat effectiveness evaluation method is the simulation driving method. In system operation, due to the large number of combat units and the complex operation environment, the data and calculation involved in the simulation are often huge. Even with a large increase in computing power, it still costs a lot of computing. Every time the parameters are modified, the simulation needs to be repeated. The cost of the assessment is relatively high. Especially for some work that needs to constantly change parameters for experimental analysis, the simulation simulation drive method often takes a long time and a high cost.

This paper uses intelligent model to solve the above problems. This model mainly uses machine learning algorithm to train and work from the data generated in the simulation process. The quality of the model is closely related to the simulation work. Therefore, this model is based on the following assumptions:

(1) simulation has been completed for the system combat problem to be studied, and the effect of simulation is good. Moreover, the simulation results should be of high reference value.
(2) the above simulation has carried out multiple experiments and produced a large number of effective data.

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

(3) there is a correlation between the output data and the input data, and the output data changes regularly with the change of the input data, rather than randomly.

B. Data acquisition and preprocessing

First of all, sufficient abundant simulation data should be obtained. The amount of data should be as rich as possible, and the data coverage should be extensive to ensure the effectiveness of training. At the same time, the format of the data obtained is often different. It could be text, graphs, tables, etc. According to the actual situation, data should be unified into tabular data, and text description and graphic description should be transformed into Numbers.

After we have collected and sorted out the data, we need to deal with the outliers. The maximum and minimum analysis method is usually used for testing. This method is simple and feasible, and it is not restricted by data distribution and has good robustness. The specific method is as follows: Q_1 is set as the lower quartile, Q_2 as the median, and Q_3 as the upper quartile. The calculation formula of the maximum estimated value T_{max} and the minimum estimated value T_{min} can be given as follows:

$$T_{max} = Q_3 + k(Q_3 - Q_1) \quad (1)$$
$$T_{min} = Q_1 + k(Q_3 - Q_1)$$

The value beyond Tmax or less than Tmin is the exception value. When k is 1.5, it is the moderate abnormal value; when k is 3, it is the special abnormal value.

The outliers detected can be handled according to the actual situation. If the simulation model is well interpreted and the abnormal value is small, the detected abnormal value can be reintroduced into the simulation model for verification to see whether it is an error or a special case. When the data is abundant enough, the method of discarding abnormal data can also be adopted. In addition, according to the actual situation, the median, average and mode can be selected to replace the outliers.

Xgboost cannot handle discrete features directly. Therefore, after the exception value is processed, we deal with the discrete type features in the data by one-hot method. The method is to use the n-bit status register to encode N states, each of which has its own register bits, and at any time, only one of them is valid.

For example:

The natural status code is: 1,2,3,4,5,6

The one-hot code is:

000001,000010,000100,001000,010000,100000

It can be understood that for every feature, if it has m possible values, then it becomes m binary features after unique heat coding. Also, these features are mutually exclusive, with only one activation at a time. Therefore, the data will become sparse.

After all processing is completed, data should be divided into training sets and test sets in a certain proportion for training and testing of models.

C. Model training and testing

Send training data to the model and conduct supervised learning training. To be specific, the xgboost algorithm cannot perform multi-metric output, so if the evaluation indicator is set with n dimensions, n xgboost learners need to be trained accordingly.

When the training data is input, the training process of each xgboost is as follows: initialize the loss function and calculate the loss function reduction according to the loss function. The optimal split attributes and split points are determined based on the loss function reduction, and the splitting process is iterated until the maximum depth of the specified tree is reached. After completing the training of the first tree, according to the training results of the first tree, update the loss function and repeat the above process to complete the training of the next tree. Iterate the process of updating the loss function and training tree until the specified number of trees is reached. The corresponding calculation formula is given below.

Settings: X represents the training sample set. N is the number of training samples. Y is the true value of the training sample. F represents the weak learner set. T is the number of learners. F_m represents the m-th generation of weak learners. L_m represents the number of leaf nodes in f_axial m, while m represents the output value of f_m. Based on the above Settings, the loss function can be defined:

$$\Psi_m = \sum_{i=1}^{N}\left[g_i f_m(x_i) + \frac{1}{2}h_i f_m^2(x_i)\right] + \Omega(f_m) \quad (2)$$

The notation in the formula is as follows:

$$g_i = \frac{d\Psi\big(y_i, F_{m-1}(x_i)\big)}{dF_{m-1}(x_i)}$$
$$h_i = \frac{d^2\Psi\big(y_i, F_{m-1}(x_i)\big)}{dF_{m-1}(x_i)^2} \quad (3)$$
$$\Omega(f_m) = \sum_{m=0}^{T}\left(\gamma L_m + \frac{1}{2}\lambda\|\omega_m\|^2\right)$$

γ, λ is the regularization coefficient, which can play a strong control on the complexity and output value of the algorithm.

P represents the index set of all samples at that node, and now the node is split. So P_L and P_R are set to represent the index set of samples of the left and right sub-nodes after the split. The loss function reduction is as follows:

$$\Delta\Psi = \frac{1}{2}\left[\frac{\left(\sum_{i\epsilon P_L} g_i\right)^2}{\lambda + \sum_{i\epsilon P_L} h_i} + \frac{\left(\sum_{i\epsilon P_R} g_i\right)^2}{\lambda + \sum_{i\epsilon P_R} h_i} - \frac{\left(\sum_{i\epsilon P} g_i\right)^2}{\lambda + \sum_{i\epsilon \mathrm{P}} h_i}\right] - \gamma \quad (4)$$

The specific training steps are as follows:

Step1: Initialize the CART tree number k, the maximum depth of z, loss function Ψ, trees set Q.

Step2: Initialize the tree q.

Step3: Calculate the loss function reduction for all the candidate features.

Step4: Find the feature which have the minimum reduction as the splitting point and split q. The feature was removed from the candidate feature.

Step5: Check whether the depth of q z. if so, enter step6, otherwise return to step3.

Step6: Add the q to Q. And according to the Q, update Ψ loss function by the above formula.

Step7: check whether the number of elements in Q reaches the specified amount k. If it reaches, the training is finished, otherwise return to Step2.

In the process of using the model, the preprocessed data were input into n xgboost learners respectively for prediction, and the n-dimensional evaluation indicators were obtained, as shown in figure 1.

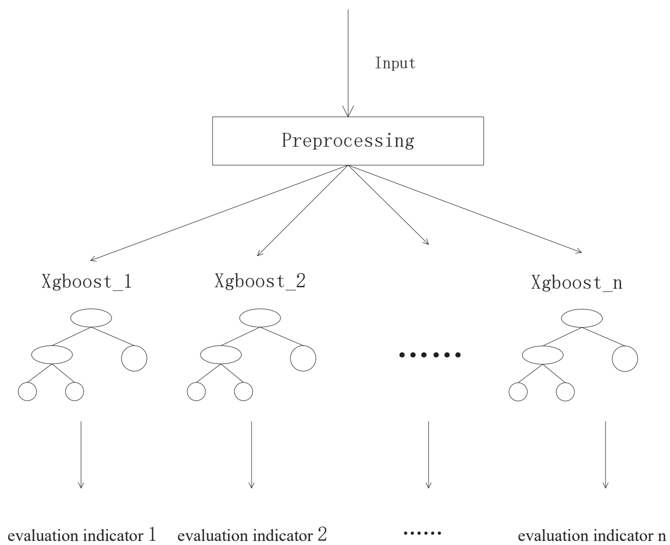

Fig. 1. Model diagram

As can be seen from the figure, the model established in this paper is simple to use and the results are clear. Moreover, when the training is completed, the calculation amount of the indicator prediction using this model is very small, far lower than the workload required by the simulation. It can save a lot of time and accelerate the research.

III. Case Study

A. Introduction of instance data

In order to verify the validity of the model, simulation data are used to carry out experiments. It is proved that the model has good practical application effect and lays a foundation for further study. The experimental data selected in this paper are briefly introduced. The experimental data selected in this paper are from a case study of a systematic conduct. In this case, it was decided that the blue side first launched an attack on the red side at sea. The red side sent a fleet of aircraft carriers to fight back against the blue side. There was a huge maritime clash between the red and blue sides.

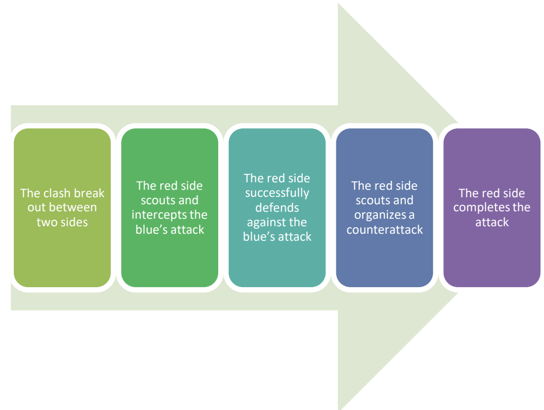

Fig. 2. Combat process

The case analysis research spectrum density of the jammer, anti-ship missile warhead weight, carrier-based fighter multi-target attack ability, anti-ship missile terminal speed, maximum effective distance of awacs radar eight parameters, such as. By setting different parameter values as input for these eight parameters, the relevant indicators are calculated to complete the analysis. The experimental data input dimension selected in this paper has 8 dimensions and the output dimension has 49 dimensions. The input dimensions are all discrete values, and each dimension has three possible input values. A total of 6,561 samples were obtained through the simulation experiment through the arrangement and combination of different values in 8 dimensions. This paper divided them into training set and test set according to the ratio of 8:2 for experiments. A prominent feature of the selected data in this paper is that the indicator difference between different samples is very small, so a high requirement is put forward for the accuracy of the indicator prediction.

B. Analysis of evaluation results

First of all, in the analysis of the data, it is found that the data are all continuous values, and the regression forecasting method should be adopted. Therefore, in this paper, 49 xgboost regressions are built for 49-dimensional output, and the above training data are used for training and testing with test data.

In order to make the effect more clear and intuitive, a representative indicator mission1 missile consumption number in the output indicator is selected. Different experiments were set up to observe and analyze the prediction results of the learner on this indicator. The xgboost prediction results are compared with the actual results on the test data, as shown in figure 3.

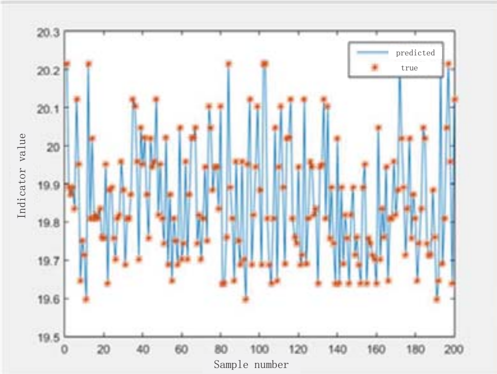

Fig. 3. Based on the xgboost mission1 missile consumption prediction test

It is not hard to see from figure 3 that the indicator prediction results obtained through the xgboost learner are very close to the real value. The error is almost zero and the prediction effect is good, which proves that the model has excellent performance and can be applied to specific research.

In order to better verify the accuracy of the model, SVR, a widely used SVM regression tool, is introduced in this paper. The same data were used for training and testing, and the test samples were tested. The above indicators were predicted, and the results were as follows.

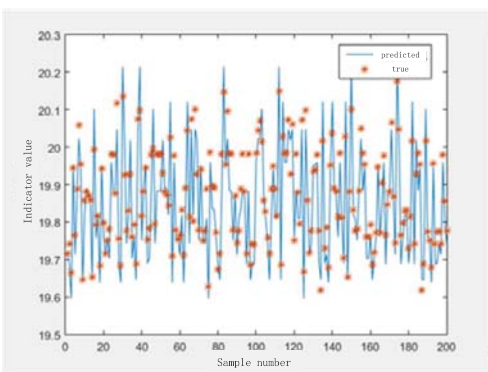

Fig. 4. Based on the svm mission1 missile consumption prediction test

It can be seen from figure 4 that the prediction results of SVM algorithm reflect the variation trend of this indicaotr in the sample to some extent, but the error is large. The real value of many samples is far from the prediction curve, which indicates that this algorithm cannot truly reflect the characteristics of samples in this indicator and cannot meet the high precision requirements mentioned above. Therefore, the results produced by this method cannot be applied to further research.

In order to further illustrate the problem, the 5-dimensional output indicators are randomly selected from the 49-dimensional output indicators in this paper.

TABLE I XGBOOST VS. SVM

Index number / method	8	11	13	15	36
svm	2.15E-3	9.32E-05	1.13E-3	1.74E-06	5.35E-3
xgboost	3.46E-11	7.60E-12	6.10E-11	9.23E-12	2.96E-13

It can be found from table 1 that when using the SVM method, the maximum Mean square error is 2.15E-3 and the minimum is 1.74E-06. Such precision is acceptable in conventional data. However, in the data used in this paper, the changes in indicators between different samples are very small, so such precision is not enough. It can be seen from the table that the Mean square error of the above five indicators is at most 3.46E-11 in the predicted results using xgboost algorithm, which is much smaller than SVM algorithm, which fully shows that the performance of xgboost algorithm is better than SVM. Moreover, according to the characteristics of this data, the Mean square error can be controlled within a small range, which can meet the precision requirements, and the calculation of evaluation can be completed well.

C. *Sensitivity analysis of learning rate*

In the course of the experiment, it is found that the setting of learning rate has a very important influence on the final prediction results. At present, there is no strict theoretical support for the setting of learning rate. Researchers usually compare the prediction accuracy through repeated experiments, and finally determine the learning rate of learning devices. The relevant experiments were also carried out in this paper. The results of the learning rate of 0.1, 0.5 and 1 in the experimental process were shown in figure 5.

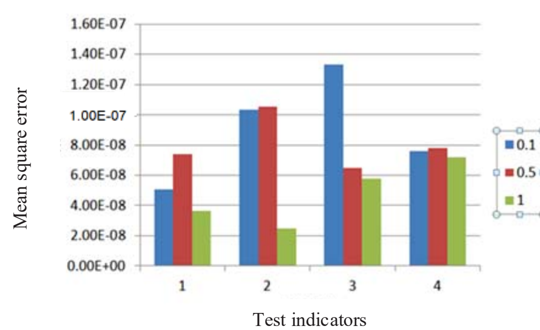

Fig. 5. Performance of each learning rate

From figure 5, it can be seen that the same learning rate has a great difference in the performance of different indicators. In test indicator 1, the mean variance of 0.1 learning rate is obviously lower than that of 0.5 learning rate. The two are similar in test indicator 2. In test index 3, the mean variance of 0.1 learning rate is significantly higher than 0.5 learning rate. Secondly, it can be seen that in test indicator 1, the performance of each learning rate is quite different, but in test indicator 4, the performance of each learning rate is quite similar. The first reason for this phenomenon is that different indicators have different internal relations with output. When using the same algorithm to predict, the performance of learning rate will also be different. Secondly, xgboost is based on the gradient descent algorithm. In the training process, if the step size is too small, the training may end without reaching convergence. At this point, the model is not optimal and therefore unstable. The detailed analysis is given below.

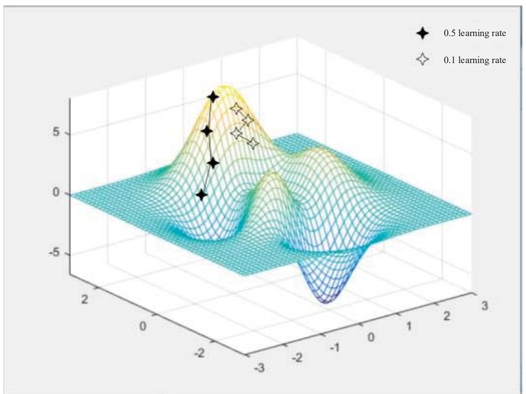

Fig. 6. Training processing 1

It is assumed that the number of training rounds is very small, only three. In figure 6, a case where the learning rate is 0.5 and the learning rate is 0.1 is given. It can be seen that when the starting point of the two is similar, after three rounds, the result of 0.5 learning rate is better than that of 0.1 learning rate, but neither of them is optimal, which cannot mean that the model has been trained.

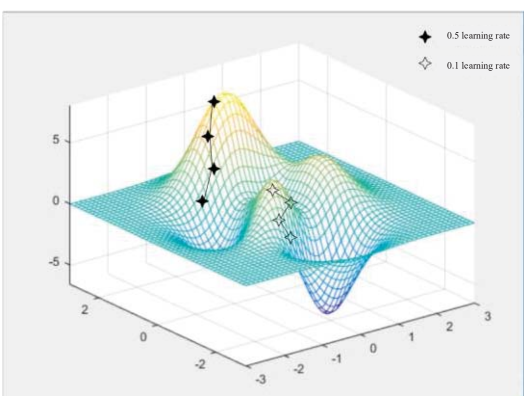

Fig. 7. Training processing 2

Another situation is given, as shown in figure 7. Since the initial position of 0.1 learning rate is good, after three steps, the result is significantly better than that of 0.5 learning rate. But again, neither of them is optimal. Therefore, we can give an explanation for the phenomenon that the learning rate performance mentioned above is quite different: the setting of learning rate is not reasonable enough. In machine learning, the consequences of setting various learning rates are shown in figure 8.

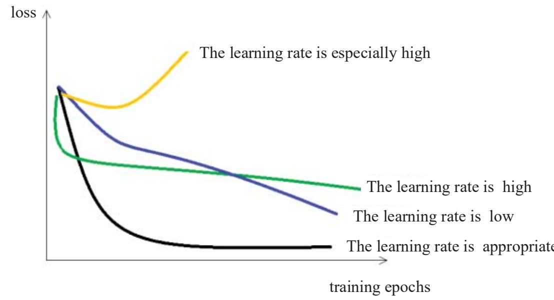

Fig. 8. The training process of different learning rates

As can be seen from the figure 8, when the learning rate is good, the training can reach the convergence quickly and the training effect is also very good. When the learning rate is high, the model converges rapidly, but because of the high learning rate, the learner "misses" the global optimum. When the learning rate is low, the learning machine training is slow and takes much time. After a considerable amount of training, it still does not reach convergence. There is also a special case, when the learning rate is very high, the training completely deviates from the set goal, and the training effect becomes worse and worse.

It can be seen from figure 5 that the learning rate of 1 in each dimension has achieved a good effect. The mean square error of learning rate is less than that of the other two. It shows that the learning rate of 1 is better than the other two. Therefore, the learning rate of the learning device is set as 1.

IV. CONLUSION

In this paper, an intelligent evaluation model of system combat effectiveness is established based on xgboost. The main work of this paper is to construct the training process and solving process of the whole model. The model is successfully applied to systematic combat effectiveness evaluation, avoiding the problem of large amount of calculation based on simulation evaluation method.

At the same time, the simulation data are used to carry out experimental analysis on the established model and verify the validity of the model. In the experiment, by training and testing the model and comparing it with the traditional SVM method, it is proved that the model has a good prediction accuracy. The experimental results show that the research results of this paper have certain significance to the study of systematic warfare.

REFERENCES

[1] zhao chao, wenchuanyuan. Exploration on the comprehensive effectiveness evaluation method of combat system [J]. Electro-optical and control,2001(01):63-65.

[2] li hui. Research on the operational effectiveness evaluation method of photoelectric defense system [D]. Graduate school of Chinese academy of sciences (changchun institute of optical precision machinery and physics),2012.

[3] Yang mi, Chen jianzhong, niu yingtao. Evaluation method of cloud-bp neural network for combat effectiveness of communication electronic defense [J]. Communication technology.2017,(4) : 746-752.

[4] Cheng kai, che xianming, zhang hongjun, zhang rui, Dan lili. Military operational effectiveness evaluation based on support vector machines [J]. Systems engineering and electronics technology.2011,(5) : 1055-1058.

[5] Shi Yanbin，Zhang An， Guo Jian． Research on the sample training of BP neural network in effectiveness evaluation [C]//2007 International Conference on Wireless Communications，Networking and Mobile Computing． Shanghai，China: ICWC，2007: 6649-6652

[6] Liu ling, xu haojun, guo hui. Research on the ls-svm based modern fighter aircraft air combat effectiveness evaluation model [J]. Practice and understanding of mathematics.2009,(3) : 36-42.

[7] Yun Bai,Yong Li,Xiaoxue Wang,Jingjing Xie,Chuan Li. Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions[J]. Atmospheric Pollution Research,2016.

[8] H. Jiang,J. Tang,F. Ouyang. A New Method for the Prediction of the Gasoline Yield of the MIP Process[J]. Petroleum Science and Technology,2015.

[9] Valiant L G. A Theory of the Learnable. Communications of the ACM, 1984.

[10] Wen-sheng zhang, Yu Ting. Boosting algorithm theory and application research [J]. Journal of China university of science and technology. 2016, (3) : 222-230.

[11] Han xiulong. Modeling of user credit score based on XGBOOST [J]. Computer knowledge and technology.2018,(5) : 7-8.

[12] Zhang Hongxia Guo He, jin-xia wang, Xu Yanyan Lv Bin, Chang Jia, Hu Guangrui, Wang Xue, hong-jun li, Liu Tianji, yan-lin li, zhi-qiang zhao, NiuXiaoJiang. Type 2 diabetes accurate prediction model based on XGBoost algorithm study [J]. China laboratory diagnostics. 2018, (3) : 408-412.