Research on Dynamic Reputation Evaluation Model Based on Machine Learning

Xiang Chen

Jiangsu Normal University Kewen College,221000,China

Keywords: machine learning; dynamic reputation evaluation model; support vector machine

Abstract: Considering the security risks of Internet online service acquisition, it is necessary to strengthen the service dynamic reputation assessment. Based on this, this paper combines the dynamic reputation evaluation model to establish the requirements, proposes the use of machine learning to complete the support vector machine model, and enhances the dynamic evaluation of service reputation level through self-learning. Based on the introduction of model training and updating methods, it is found that the performance of the model can be obtained with high accuracy by verifying the performance of the model.

Introduction

The Internet has a certain openness, which provides users with an opportunity to generate online attacks when they obtain online services. Strengthen the service dynamic reputation assessment, establish a good trust relationship between users and resource providers, and ensure the security of network resources. Therefore, advanced methods such as machine learning should be introduced to complete the establishment of a dynamic reputation assessment model to promote the sound development of the network environment.

1. Dynamic Reputation Assessment Model to Establish Requirements

In the Internet environment, users often need to obtain online services. However, the network environment is dynamic and cannot guarantee that all services are trusted. Therefore, between users and service providers, dynamic reputation assessment needs to be implemented so that users can select according to the evaluation results. Reliable service. Complete the establishment of dynamic reputation assessment model, which can realize the definition and quantitative expression of each factor of credit, implement reputation evaluation from online service acquisition and calculation agent, and complete the dynamic evaluation of service resource reliability[1]. Direct online The service needs to be evaluated according to the service access history. To realize the indirect evaluation of the service, it is necessary to evaluate the past reputation of the service resource. Combined with the direct evaluation and the indirect evaluation result, the service reputation dynamic evaluation can be completed to ensure that the evaluation result is dynamically adapted. Sex.

2. Dynamic Reputation Evaluation Model based on Machine Learning

2.1 Model establishment ideas

In the actual establishment of the dynamic reputation evaluation model, the fuzzy rules can be used to complete the direct and indirect evaluation results. Only the single-dimensional evaluation of the service subject's reputation can be carried out. It is difficult to evaluate the quality of the service's various dimensions, which leads to the objectiveness of the evaluation results being affected. The machine learning method can be used to utilize the training samples, and the optimal classifier model can be constructed in a multi-dimensional space to realize the service classification of different reputation categories. Multi-dimensional evaluation of the service can complete the

  DOI: 10.25236/icme.2019.106

dynamic reputation evaluation and promote the model in In the service environment, it is better to adapt. In the process of dynamic evaluation, due to the uncertainty of the network environment, it is necessary to use the fuzzy reasoning method to generate dynamic evaluation rules to determine the dynamic reputation level of the service. Actually, in the network environment, the number of training samples is limited, relying on a small amount. It is difficult to accurately describe the regularity of the data in the sample data. In order to establish an effective dynamic reputation evaluation model, it is necessary to use the support vector machine to complete the establishment of the classifier model, and use a small number of manually labeled samples and a large number of unlabeled samples for model training to make the model generalization ability. Get enhanced[2]. as a structure. The commonly used algorithm of classifier, support vector machine can realize the introduction of kernel function by using the principle of risk minimization, realize sample mapping in linear multidimensional space, and determine global optimal solution by solving convex optimization problem. In dynamic reputation evaluation, it can combine more discriminant Features get more support vectors to ensure that the evaluation results get more decision information and ensure the objectivity of the results.

2.2 Model establishment method

2.2.1 Modeling process

Before modeling, you should also specify the service sample xi to have the reputation class label yj, get the sample instance s={xi, yj}, and add it as a training sample to the sample set to get the sample. Collection S. Usually there are at least two types of online service reputation tags, which are multi-classification problems. It is also necessary to implement the conversion of the two-category problem to simplify the model analysis process. Assuming that xi={a1,...,ad,...,an} is an n-dimensional feature vector, a plurality of two classifier configurations should also be completed. In order to ensure a balanced distribution of different types of tags, a one-to-one classification method is also used to complete the construction of the training sample set D for the service q classification, and q(q-1)/2 sample sets[3] are obtained. In the sample set D[1,2] with the reputation labels 1 and 2, the sample rejection of the reputation labels 3, 4, ..., q needs to be completed. In the same way, a sample set of multiple reputation tag groups can be obtained. The resulting two classifier can be recorded as C[k,l], k, l = 1, 2, ..., q, and returns 1 or 0 depending on the equality of the equation. Each classifier has a corresponding prediction class label, combines the prediction results of each classifier, and regards the label with the most cumulative number as the final label, and can obtain the dynamic reputation evaluation result.

2.2.2 Model training

In the model training, the support vector machine is used to realize the manual labeling initial training sample data input, and then the self-learning analysis of the unlabeled samples can realize the expansion of the training sample set, so that the service data distribution can be accurately reflected to ensure the obtained decision. Functions can fit in multidimensional space. The iterative loop of the self-learning process is completed, which can optimize the performance of the model and minimize the impact of the wrong training samples on the model. For the transplant training sample set S and the unlabeled sample set U, after randomly selecting z service samples to complete manual labeling, the sample set L can be obtained, and the sample sets S and U are updated to obtain U:S=L, and U= UL, the initial sample set S = L, using the model to achieve the prediction of the sample class label in U.With self-learning technology, it is necessary to complete the establishment of the classifier model from a small amount of sample training, and then judge the reputation class label.By screening the classification results, the most unidentified samples and their corresponding class labels can be determined, the training set is expanded, the model is retrained, and new evaluation results are obtained.Using the model to predict the label of the sample class can easily lead to inconsistency between the predicted result and the real situation, resulting in the

accumulation of errors in the iterative calculation, and ultimately affecting the class label.Therefore, in order to improve the method, the distance constraint condition is also used to judge, so that the accuracy of class label screening is improved.For the unlabeled sample xu, the model can be used to determine the label to obtain yu.According to the Euclidean distance pair xu and the sample distance D(xu, xi) in the sample set S, the label of the nearest neighbor point N can be completed, and the class label yn is obtained.Compare two sample reputation class tags, add unlabeled samples to the initial sample set when the class labels are equal, or add them to the unlabeled sample set.Taking this self-learning method, it is possible to determine the unknown sample reputation class label by model prediction, and to ensure that the sample class label is accurate.

2.2.3 Model Optimization

In the process of model optimization, it is necessary to use the new screening sample to complete the sample set update, re-train the model, and implement the model loop iterative update according to the distance correlation. Continuously adding new samples to the sample training set can enrich the amount of information and improve the performance of the model. To improve the generalization ability of the model, it is necessary to achieve dimension reduction of features. Principal component analysis can be used. When the set of xi attributes is A={a1,...,ad,...,an}, the principal component of the dynamic evaluation of reputation is evaluated from the subset of principal components. Screening is performed to complete irrelevant principal component culling. By adopting such an algorithm, the main information of the original data can be carried by the independent data, and the complexity of the model sample data is reduced, thereby improving the efficiency of the model analysis. When the model training and test data sources are the same, it is easy to cause training errors. To optimize the model from this perspective, cross-validation needs to be completed. The sample set is divided into training set and test set, and the model training and testing are carried out respectively to evaluate the performance of the model. In the verification process, the optimal kernel function parameters are selected from the original data set. In fact, the radial basis kernel function K(xi, x)=e-γ|x-xi|2 can be used to complete the 10-fold cross-validation and the kernel function. The parameters in the optimization are used to provide useful information for model validation. After the effective sign extraction of the model is completed, the cross-validation method is used to improve the model, and a dynamic reputation evaluation model with strong generalization ability can be obtained.

2.3 Model Validation Analysis

2.3.1 Verification method

In order to determine the validity of the model for online service dynamic reputation assessment, it is necessary to analyze the accuracy of each credit category assessment in the test set and the overall accuracy of the dynamic reputation assessment, and carry out corresponding verification experiments. Comparing the evaluation results with the results of the reputation evaluation based on fuzzy rules can reflect the application performance of the model. The selected sample data is the transaction evaluation statistic of 1500 online merchants in Jingdong.com, including 10 commodity types and 15 dimension attributes, which are independent samples.80% of them were taken as training samples, and the rest were used as test samples. Considering that the test sample reputation class label should have the same subjective preference as the human, the manual labeling of the test sample set class label can be completed. During the experiment, the verification operation needs to be performed on the 3.3GHz i3 processor. The system is Window10, and the kernel function is provided by the LibSVM toolkit to complete the experimental scheme design and model cross-validation. With the Matlab toolkit, the elimination of incomplete samples in the experimental data can be completed, and the number of samples can be equalized, so that the data noise problem can be solved. When measuring the model, it is necessary to analyze the method from the perspective of average accuracy, so that the accuracy of the model is reflected and provide a basis

for the model.

2.3.2 Verification results

From the verification results, the support vector machine is used for dynamic reputation evaluation, the accuracy rate can reach 0.897, and the accuracy based on fuzzy rules is 0.886.The reason for the analysis can be found that compared with the fuzzy rule algorithm, the support vector machine is used for dynamic evaluation and classification of reputation, which can complete the evaluation of multi-service attributes of sample data, fuse direct evaluation results and indirect evaluation results, and determine different attributes for sample data. The dynamic images brought by the reputation finally get objective evaluation results. The actual analysis of the service samples, using the support vector machine can solve the problem of multi-heterogeneous data of reputation data, and complete the optimal allocation of services. Because in the model training process, the identification of the class label information in the service attribute can be completed according to the user's Fenqueque score, and the reputation discriminant feature is obtained, and then the accuracy of the evaluation result is improved.

Conclusion

In summary, the acquisition of online services in the Internet environment requires a dynamic reputation assessment of the service providers to ensure that users can obtain security services. Considering that there are a large number of service data with unlabeled attributes in the network environment, the support vector machine is used to complete the establishment of the machine learning model, and the performance optimization of the dynamic reputation evaluation model is completed. From the perspective of model performance verification, the accuracy of the evaluation results can be improved, so that the service dynamic reputation evaluation requirements can be met.

references:

[1] Liu Houqin. Credit risk prediction model of machine learning algorithm [j]. Microcomputer Application, 2019, 35 (02): 70-73.

[2] Yang Tidong, Fu Xiaodong, Liu Wei, et al. Online service reputation measurement based on multi-dimensional evaluation information [j]. Small microcomputer system, 2018, 39(12): 2625-2631.

[3] Zhang Huibing, Li Chao, Hu Xiaoli, et al. Dynamic reputation assessment of subjective and objective integration in Internet of Things search [j]. Journal of Communications, 2015, 36(12): 106-113.