Journal of Computer Science 7 (9): 1423-1425, 2011
ISSN 1549-3636

Statistical Analysis and Learning Method on Users' Feedbacks

Doris Hooi-Ten Wong and Sureswaran Ramadass
National Advanced IPv6 Centre (NAv6),
University Sains Malaysia, 11800 Penang, Malaysia

Abstract: Problem statement: The purpose of this study was constructing an effective algorithm in order to learn the users' feedbacks from their displayed visualization. This is due to existing visualization tools typically involve presenting network data regardless of considering level of network data knowledge among different levels of computer users. The machine learning algorithm has been applied in order to find the most effective statistical analysis and learning algorithm in learning users' feedbacks. **Approach:** The objectives of this study were to conduct statistical analysis and learning algorithm model for different levels of computer users' feedbacks and procedure to test the classifier. **Results:** WEKA the machine learning workbench that supports many activities of machine learning practitioners will be used to implement the proposed algorithm. The implemented program will work as training testing model. **Conclusion:** We can produce an adaptive visualization to the different levels of computer users as we have learnt their feedbacks (behavior) and update the classifier model.

Key words: Statistical analysis, learning method, user feedback, machine learning, effective algorithm, displayed visualization, existing visualization, network data, different levels of computer, learning algorithm

INTRODUCTION

Each computer users have different levels of system skills, different behavior as well as different perspective view on the same displayed visualization (Ghazarian and Noorhosseini, 2010). The purpose of this study was constructing an effective algorithm in order to learn the users' feedbacks on their displayed visualization. This is due to existing visualization tools typically involve presenting network data regardless of considering level of network data knowledge or feedback among different levels of computer users. We believed that by having the learning algorithm, the system managed to generate adaptive visualization which is comprehensive and can be understood by the different levels of computer users.

Statistical analysis and learning users' feedback aiming to study on the applied machine learning algorithm in order to find the most effective statistical analysis and learning algorithm in learning users' feedbacks (Wong and Chai, 2010) . Generally, there are number of visualization tools that have applied on the network data visualization. However, those existing visualization tools working without the expert-aware approach and statistical analysis learning approach, which we believed it will become a unique contribution for the visualization tools.

In this study, we focus on the users' feedback statistical analysis and learning algorithm. Additionally, users' knowledge levels change dynamically as users gain more experience in a user interface. In order to adapt comprehensive user interfaces to the different needs of user groups with different feedbacks, automatic methods of feedback learning are required.

MATERIALS AND METHODS

In order to achieve all the mentioned objectives of this study, the proposed users' feedbacks statistical analysis and learning system focuses more on the following four tasks:

- Collection of the users' feedbacks
- Construction of the training model
- Construction of the testing model
- Evaluate proposed algorithm effectiveness

Feedback collection: Basically, Intuitive Visualization Component will display and show the visualization to the computer users (Wong and Manickam, 2010). Users are allowed to send their feedback on the displayed visualization or what types of visualization that they requested. Feedbacks on the visualization are sending back to the Intelligent Analysis and Learning Component. Feedbacks will be collected and being the input for the construction of the training model which will be discussed more in the training model

Corresponding Author: Doris Hooi-Ten Wong, National Advanced IPv6 Centre (NAv6), University Sains Malaysia, 11800 Penang, Malaysia

construction. 50 users from beginner level will be asked to give their feedback on the displayed visualization. We concentrate on the beginner users as they are our scope target in the research.

Training model construction: The training model (classifier) is built using the supervised machine learning algorithm, namely Support Vector Machine (SVM) algorithm. Visualization provided some of the features to be chosen by the computer users. These features (attributes) are either explicit feedbacks from users' relevance judgments or implicit feedbacks by observing the users' search and browsing activities. The attributes become the inputs to construct the training model, Attribute values, such as age (years), data selection (counts), menu interactions (counts) and help buttons clicking (counts) were used as inputs to our models.

Testing model construction: The testing model is being built after the completion of the training model. Same features will be taken as the inputs for the testing model. We believed that users' knowledge levels will change dynamically as computer users gain more experience from the displayed visualization. The training model will be improved based on the consecutive inputs from the testing procedure. This can ensure the best adaptive comprehensive visualization being displayed to the computer user.

Evaluation: To evaluate the accuracy of statistical analysis and learning algorithm, we have conducted a testing on 50 computer users based on the training model in order to achieve comprehensive visualization to the beginner computer users. The evaluation has been run for fifth times, total of 250 users have given their feedbacks.

RESULTS

In this study, WEKA the machine learning workbench which is famous in supporting developments and activities of machine teaching practitioners (Bouckaert *et al*., 2010). WEKA will be used to implement the proposed algorithm. The implemented program will work as training as well as testing model.

Training model: 50 users who have been categories as beginner group have been requested to give their feedback on their displayed visualization. They will response to the following features which will be used as the input for the training model construction. The inputs have been run and the result as shown below. Table 1 has shown the result by applying SVM using WEKA.

Table 1: Result from training model

Training size	True positive rate	False positive rate	Precision	Recall
50	0.600	0.216	0.646	0.600
100	0.770	0.122	0.826	0.770
150	0.820	0.094	0.831	0.820
200	0.785	0.109	0.790	0.787
250	0.800	0.101	0.804	0.800

Table 2: Testing result based on 50, 100,150, 200 and 250 training model

Testing size	True positive rate	False positive rate	Precision	Recall
50	0.72	0.146	0.769	0.72
50	0.60	0.206	0.640	0.60
50	0.64	0.169	0.651	0.64
50	0.86	0.072	0.861	0.86
50	0.92	0.041	0.920	0.92

Testing model: After construct the training model, another 50 users with beginner knowledge of computer will test the model. Each update from each computer user will be update to the training model. The testing dataset involved 250 users. Table 2 has shown the result of the five times of testing.

DISCUSSION

As mentioned earlier, this study focuses more on the beginner computer users and do not discuss about the intermediate as well as advanced computer users. We consider beginner users because we believed that beginner users have more requirements on what is displayed to them.

The result has shown the increasing on the precision and recall. The more feedbacks we get, the more adaptive visualization will be achieved in order to present to the beginner user.

CONCLUSION

In this study, we have proposed and designed statistical analysis and learning algorithm and implemented evaluation on 250 computers users in order to prove that the users' feedback learning able to generate adaptive visualization to different levels of computer users based on the training model. Our initial testing proved that our adaptive system can intelligently and automatically analysis the same feedback and produce comprehensive visualization to different levels of users. The comprehensive visualization screens will be able to enhance the visualization and assist computer users easily to interpret huge volumes of network data into understandable knowledge.

ACKNOWLEDGEMENTS

The researchers would like to acknowledge Institute of Postgraduate Studies (IPS), Universiti Sains Malaysia (USM) for funding of USM Fellowship Scheme 2011 and Postgraduate Research Grant Scheme (USM-RU-PRGS). Thanks to Dr Kok-Soon Chai for his valuable guidance and motivation.

REFERENCES

Bouckaert, R.R., E. Frank, M.A. Hall, G. Holmes and B. Pfahringer *et al*., 2010. WEKA-experiences with a java open-source project. J. Machine Learn. Res., 11: pp. 2533-2541.

Ghazarian, A. and S.M. Noorhosseini, 2010. Automatic detection of users' skill levels using high-frequency user interface events. User Model. User-Adapted Interaction, 20: 109-146. DOI: 10.1007/S11257-010-9073-5

Wong, D.H.T. and K.S. Chai, 2010. Statistical analysis learning approach: The use of artificial intelligence in network data visualization system. Proceeding of the International Conference on Computer Applications and Industrial Electronics, Kuala Lumpur, Malaysia, IEEE, pp: 206-210. DOI: 10.1109/ICCAIE.2010.5735076

Wong, D.H.T. and S. Manickam, 2010. Intelligent Expertise Classification Approach: An innovative artificial intelligence approach to accelerate network data visualization. Proceeding of the 3rd International Conference on Advanced Computer Theory and Engineering, Chengdu, China, IEEE, pp: 6437-6440. DOI: 10.1109/ICACTE.2010.5579790