

Application of Machine Learning Techniques to Web-Based Intelligent Learning Diagnosis System

Chenn-Jung Huang, Ming-Chou Liu, San-Shine Chu, Chin-Lun Cheng

Institute of Learning Technology, National Hualien Teachers College, Hualien, Taiwan

E-mail: cjhuang@mail.nhltc.edu.tw

Abstract. This work proposes an intelligent learning diagnosis system that supports a Web-based thematic learning model, which aims to cultivate learners' ability of knowledge integration by giving the learners the opportunities to select the learning topics that they are interested, and gain knowledge on the specific topics by surfing on the Internet to search related learning courseware and discussing what they have learned with their colleagues. Based on the log files that record the learners' past online learning behavior, an intelligent diagnosis system is used to give appropriate learning guidance to assist the learners in improving their study behaviors and grade online class participation for the instructor. The achievement of the learners' final reports can also be predicted by the diagnosis system accurately. Our experimental results reveal that the proposed learning diagnosis system can efficiently help learners to expand their knowledge while surfing in cyberspace Web-based "theme-based learning" model.

Keywords: Web-based learning, Theme-based learning, Fuzzy expert system, K-nearest neighbor, Naïve Bayesian classifier, Support vector machines, Learning diagnosis

1 Introduction

The surprising development of information technology has created a new vision for network learning that its influence has already spread over the world to facilitate educational innovation. Therefore, many countries have been paying attention to computer technology and expect it can facilitate the education reform in an effective and efficient ways. It is well known that the application of computer and Internet teachings to traditional teaching requires some kind of transformation. Consequently, the research and development of proper learning model has to seriously consider the mutual interaction between the users and the computers, the instructor and the learners, and the interaction among the learners. Embed the related research issues to the above process, the splendid research results then can be expected.

The theme-based learning is to learn an integrated knowledge by defining a central "theme" at the very start and compose related knowledge surrounds the cen-

tral theme from various aspects. Such a learning model emphasizes the training of the learners with the competency of knowledge integration. Compared with traditional teaching, which teaches fragmentary information within the limitation of subjects, units, chapters, and sections, the intention of theme-based learning is to take a theme as a starting point and stretch out of it based on the learners' interests. Accordingly, the learners can voluntarily construct their own knowledge since the theme is strongly connected with our daily life and developed from learners' willingness.

A theme-based learning process can be divided into exterior circulation and interior circulation as illustrated in Figure 1 [7]. Exterior circulation activities are 1) Identify a central theme, 2) Identify related subject domains based on learner's interest, 3) Collect information for the specific topics, 4) Integrate collected information to build shared knowledge, and 5) Exhibit learning outcomes and share with others. The activities of the exterior circulation are explicit learning behaviors. On the other hand, the interior circulation consists of implicit mental activities, which are Plan, Action, and Introspection, respectively. When learners engage in the theme-based learning processes on Web, they are experiencing the activities of exterior and interior circulation synchronously. Since the explicit feature of the learning processes can be controlled or guided effectively by the careful design and implementation of the Web-based learning environment, it is expected that the interior circulation, which represents the invisible mental behavior of the learners, can make great progress simultaneously.

The exterior circulation of the theme-based learning model, as Figure 1 illustrates, can be implemented as a Web based system that helps to manage the learning processes. The learning activities with regard to student learning can be divided into five stages as follows:

(1) Identify a central theme

The learners engaged into the theme-based learning can propose their own interested topics to ask for feedback from other team members. Meanwhile, every learner can also join other member's proposed topic. After interaction and brain storming, the ones who are interested in the same topic are formed as a learning team, and this topic is the central theme that this team would investigate. The motivation of such an arrangement is that "a student can learn better if he/she was interested

in the learning topic". The theme should be closely connected with the learners' daily life and an extensive range of survey which is not limited in a specific field is encouraged.

(2) Identify related subject domains based on learner's interest

At this stage, the theme is defined and the learning team for each theme is formed. Based on the learner's own specific interest, each team member tries to find the issues in the related subject domain derived from the theme. Notably, the interaction of learners on the learning platform can influence the relatively inactive learners to trigger their interests effectively on some specific topics through the events and activities originated by their team members.

(3) Collect information for the specific topics

Team members will cooperate with each other to collect data and information related to the interested topic at this stage. With the help of search engine and other assistant tools, wealthy knowledge related to the interested topic can be built up. If the data collected from the cyberspace is not enough, real-world resources such as libraries, face-to-face interviews *et al.* also can be utilized. The collected data or information is then processed to form the knowledge stored in the learner's long-term memory.

(4) Integrate collected information to build shared knowledge

Each team member tries to organize the data or information collected at the previous stage and to generate a thematic report. The reports can be shared with other teammates through peer review and online discussion.

(5) Exhibit learning outcomes and share with others

The thematic report for the each learner is expected to be refined to certain degree through consecutive discussion with the teammates and constructive suggestion offered by the teacher. The elaborated report is finally displayed on the public area to make it accessible to the teachers and all the learners.

Besides putting the learning activities that correspond to the exterior circulation of the theme-based learning model into practice, an intelligent diagnosis system is also incorporated in the proposed Web-based thematic learning platform. Notably, a fuzzy expert system and a composite classifier are used to give the learning guidance to the learners, and assist the instructor in grading each learner's online class participation and predicting the performance of each learner's final written report. The remainder of the paper is organized as follows. Section 2 gives a brief description of architecture of the proposed Web-based thematic learning platform. In section 3, we will show the details of the intelligent diagnosis system. Section 4 reviews and discusses the experimental results. Conclusions and the future work are made in Section 5.

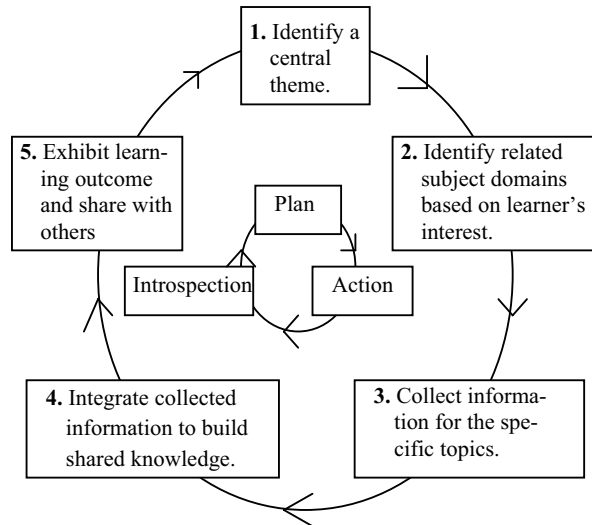


Figure 1. A theme-based learning model.

2 Architecture of the Theme-Based Learning Website

A theme-based learning system is composed of three functional modules as shown in Figure 2. They are User Interface Agent, Learner Profile Management Agent, and Learning Diagnosis System, respectively.

2.1 User Interface Agent

The learners can login the theme-based learning system through User Interface Agent to participate in the learning activities such as data searching, data managing, discussing with the colleagues and the teacher online, posting and replying the articles, etc.

2.2 Learning Profile Management Agent

The connections and the interactions for the learners with the system and other learners are built up in this module. The system can generate the learners' learning profiles, including the total time that the learners stayed in the platform, the frequency of login sessions, the learning materials collected by the learners, the articles posted or replied by the learners, and the online group discussion time spent by the learners, etc., which provide the teachers and Learner Diagnosis System to follow the learners' learning status such that effective scaffolding and constructive suggestion or analysis for the learners can be given timely.

2.3 Learning Diagnosis System

The learners are expected to make progress based on certain proper learning advices given by the Learning Diagnosis System. The system also generates the online

participation assessment at the end of each learning activity stage as mentioned in Section 1, according to the learner profile logged in the system. The teacher can either take this class participation assessment as a proportion of the learners' final grade, or use this assessment to uncover the learners that fall behind at the end of each learning activity stage. Meanwhile, the system can also predict the performance of the learners' final report so that the teacher can use this predicted achievement for further analysis on the learner's study behavior when there is a gap between expected result and the learner's actual performance.

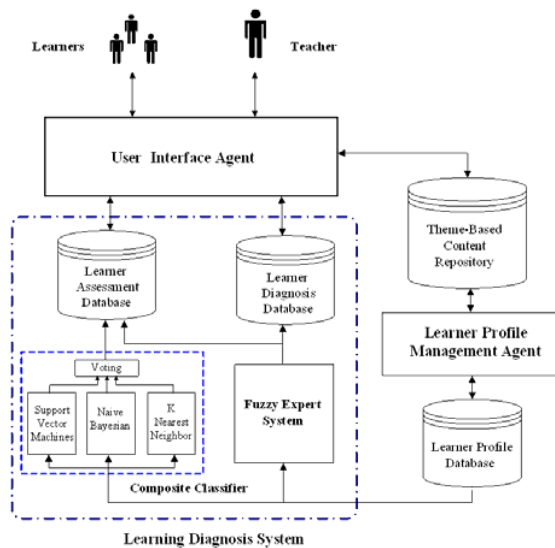


Figure 2. Architecture of the theme-based learning website.

3 Learning Diagnosis System

The Learning Diagnosis System employed in this work consists of two major parts. One is a fuzzy expert system which not only gives appropriate diagnosis messages to the learners but also delivers each learner's online participation assessment to the teacher at the end of each learning activity stage based on the learners' profile. The other part is a so called composite classifier which is used to predict the learners' accomplishment on the final report. The motivation of using a fuzzy expert system to give diagnosis and class participation assessment whereas using a composite classifier to predict the learners' final accomplishment is that the fuzzy expert system can function more like human experts who explain the reasoning processes behind their recommendation. On the other hand, it is not too difficult to find some advanced machine learning techniques combined with wrapper attribute selection method possess a better prediction capability than a fuzzy expert system.

3.1 Architecture of the fuzzy expert system

An expert system is a program that behaves like an expert in some problem domain. The principle use of expert systems is to seek information from a variety of sources including databases and the users to solve finite, well-defined problems [3]. To deal with uncertain and incomplete information, the fuzzy expert system incorporates fuzzy logic into the reasoning process and knowledge representation scheme [1].

The input to a fuzzy expert system is a crisp value that is given by the learner profile database. The fuzzy rule base is composed of a set of fuzzy if/then rules and the intersection or minimum operation is employed to generate a corresponding fuzzy subset for each fuzzy rule. The aggregator then combines all of the fuzzy subsets assigned to the output variable together to form a single fuzzy subset for the output. Lastly the aggregated linguistic values from the inferred fuzzy control action are fed into the defuzzifier to generate a non-fuzzy control output. Notably, the generalized bell-shaped membership function is chosen for three antecedents and the consequent in the fuzzy expert system. The three antecedents are the number of articles posted/replied by the learners, the number of the learning materials that each learner collected, and the frequencies of visiting the platform by the learners, respectively. Meanwhile, the Mamdani defuzzification method is employed to compute the centroid of membership function for the aggregated output, where the area under the graph of membership function for the aggregated output is divided into two equal subareas.

The inferential rules of the fuzzy expert system are treated differently when they are used to generate some appropriate suggestion or diagnosis messages for the teacher and the learners. The system may give a feedback message to the learners when the membership grades of linguistic variables such as "low" or "high" is the largest among the three for each input. The learning diagnosis system will not only give the teacher a summary report of suggestion messages that the learners received, but also offer the teacher each learner's online participation assessment based on the non-fuzzy output of the defuzzifier. The teacher can locate the learners that fall behind and give them individual guidance by examining the class performance record given by the system. Notably, although the exact calculation of class participation assessment involves the 27 inferential rules, only 11 inferential rules are used in actual computation since the rest are considered unreasonable.

3.2 Architecture of the composite classifier

The motive of using the composite classifier in our

work is that the composite classifier has the advantage of making decisions more reliable and accurate than a single classifier although the combined model is typically hard to analyze in intuitive terms what factors are contributing to the improved decisions [3]. The composite classifier in this work is mainly composed of three independent classifiers, i.e., a K nearest neighbor classifier, a naïve Bayesian classifier, and a support vector machines classifier, respectively. Each of three independent classifiers uses wrapper approach to select the desirable input parameters during training. The training data is taken from the learners' profiles database maintained for the learning groups in past learning programs. A vote is taken if there are conflicts among the prediction results of the three classifiers. The output of the composite classifier is the predicted grade for the learners' final report. The teacher can either cite this predicted grade as portion of the learners' final achievements or perform a further investigation if there is a discrepancy between the learners' actual achievement and study behavior.

● **Wrapper attribute selection method**

It is well known that the performance of most machine learning algorithms can be deteriorated by some irrelevant or unhelpful attributes. Thus it is common to precede classification work with an attribute selection stage which strives to eliminate all but the most relevant attributes. This is also one of the major reasons that the prediction capability of a fuzzy expert system is worse than some advanced machine learning algorithms since the inputs to the fuzzy expert system are always chosen by the human experts and these selected attributes might not be the most promising ones for the fuzzy expert systems.

The attribute selection methods can be divided into two broad categories in the literature [4], i.e. filter methods and wrapper methods. Filter methods select predictive subset of the attributes using heuristics based on characteristics of the data, whereas wrapper methods make use of the classifier actually used to evaluate the accuracy of attribute subsets. Wrapper methods generally result in better performance than filter methods because the latter suffers from the potential drawback that the attribute selection principle and the classification step do not necessarily optimize the same objective function.

In the wrapper approach, the learner is applied to subsets of attributes and tested on a hold-out set. From the results of these tests, a good subset of attributes is selected. For example, for *forward selection*, a classifier is built for each attribute individually; and the most accurate attribute is "accepted" into the subset of good attributes. That attribute is removed, and the process is repeated, adding each of the remaining attributes and evaluating its performance. The "best" set of two attributes is thus created. This proceeds incrementally until

an attribute set with maximal accuracy is achieved. Similarly, *backward selection* proceeds by eliminating one attribute at a time, finding the least beneficial attribute and eliminating it, and repeating the process, eliminating the least accurate attributes until eliminating further attributes decreases accuracy.

● **K nearest neighbor classifier**

To classify an unknown data sample X , the k nearest neighbor classifier simply examines the k closest training samples to X and assigns it to the most common class among these k closest samples. In other words, we are seeking those training samples that are most similar to X , and then classify X into the most heavily represented class among these most similar objects [2]. Notably, "closeness" is defined in terms of Euclidean distance.

A k nearest neighbor classifier has several attractive properties. For example, it is easy to program and no optimization or training is required; Extension to multiple classes is straightforward. Although a potential drawback of the k nearest neighbor classifier is that it does not build a model, relying instead on retaining all of the training data set points. Thus, searching through a large training data set to find the k nearest can be a time-consuming process. However, this problem can be evaded here since small data sets in low dimension are used in this work.

● **Naïve Bayesian classifier**

The naïve Bayesian classifier predicts an unknown data sample, X , belonging to the class with highest posterior probability, conditioned on X [6]. Bayesian classifiers have minimum error rate in comparison to all other classifiers theoretically. Even though this is not the case in practice due to inaccuracies in the simplified assumptions made for its use, such as class conditional independence and the lack of available probability data, empirical studies given in the literature shows the performance of naïve Bayesian classifier is still comparable to other complex machine learning techniques such as neural networks. We thus adopt this simple but effective approach as an element of the composite classifier.

● **Support vector machines**

Support vector machines (SVM) have recently gaining popularity due to its numerous attractive features and eminent empirical performance [4]. The main difference between the SVM and conventional regression techniques is that it adopts the structural risk minimization (SRM) approach, as opposed to the empirical risk minimization (ERM) approach commonly used in statistical learning. The SRM tries to minimize an upper bound on the generalization rather than minimize the training error, and is expected to perform better than the traditional ERM approach. Moreover, the SVM is a convex optimization, which ensures that the local minimization is the unique minimization.

To solve a nonlinear regression or functional ap-

proximation problem, the SVM nonlinearly map the input space into a high-dimensional feature space via a suitable kernel representation, such as polynomials and radial basis functions with Gaussian kernels. This approach is expected to construct a linear regression hyperplane in the feature space, which is nonlinear in the original input space. Then the parameters can be found by solving a quadratic programming problem with linear equality and inequality constraints.

As the SVM outperforms other conventional regression methods in the application of time series and internet traffic predictions in the literature [4], we thus try to replace the fuzzy logic module in the bandwidth reservation scheme with the SVM as illustrated in the previous section to estimate the reserved bandwidth in the neighboring cells.

4 Experimental Results and Analyses

To examine the educational effect of the theme-based learning system, two fifth grade classes at an elementary school have been chosen to practice theme-based learning in classroom teaching. One of the two classes was experimented with the proposed Web-based thematic learning platform in a Natural Science course, wherein a fuzzy expert system is incorporated into the intelligent diagnosis system to grade students' class participation and learning guidance to the learners based on learning profiles so that the students can receive just-in-time support or suggestion to help them gain better learning achievement. The second experiment was conducted in another Natural Science course for the other fifth grade class whereas the diagnosis system was removed from the Web-based thematic learning platform in order to demonstrate the performance of the proposed diagnosis system.

Table 1 compares the pupils' achievement in two classes. The statistical results were obtained by running t test with the SPSS software package. The average score received by the 25 pupils whose study behavior was rectified by the diagnosis system is apparently better than the average score of 18 pupils in the other experiment, although the gap between these two mean scores are not quite significant due to the different grading standard for the two teachers. However, we observe that the proportion of the pupils that failed in the course substantially lowered down with the aid of the diagnosis system. Meanwhile, the diagnosis system greatly reduces the teaching load of the teacher so that the teacher can have more free time to give individual guidance to the specific pupils that fall behind or behave inactively.

As mentioned in Section 3, we use three different classifiers to predict the pupils' achievement in the final report based on the pupils' profiles, and the majority of their votes are taken as the final prediction result. To

verify the performance of each individual classifier, we ran a series of tests on each of three classifiers by using a so-called leaving-one-out cross validation technique (LOOCV) [5] with 18 and 25 samples for the two classes, respectively. The inputs to each classifier are identical to those used in the fuzzy expert system. Notably, the LOOCV method removes a single sample in each trial, trains on the rest, and then tests the PNN classifier on the removed single sample.

The performance comparison given in Table 2 reveals that all of the three classifiers can achieve a much higher prediction rate for the class using the theme-based learning platform in which the diagnosis system is embedded. It can be inferred from the results that the learning guidance given by the fuzzy expert system significantly affected the learners' study behavior and boosted the quality of the learners' final reports further.

Table 1 Comparison of the pupils' achievement in two experimental classes

t test	Learning Platform with Diagnosis System	Learning Platform without Diagnosis System
Mean	86.60	82.22
Standard Deviation	9.443	13.225
Standard Error Mean	1.889	3.117
Degree of Freedom	41	
Significance (2-tailed)	0.212	

Table 2 The LOOCV prediction results of three classifiers.

Classifier	Learning Platform with Diagnosis System	Learning Platform without Diagnosis System
Naïve Bayesian	92%	67%
K Nearest Neighbor	88%	50%
Support Vector Machines	84%	67%

5 Conclusions and Future Work

The Web-based thematic learning system implemented in this work mainly practices the exterior circulation of

the Theme-Based Learning model as illustrated in Figure 1. An intelligent diagnosis system, which is composed of a fuzzy expert system and a composite classifier, is proposed to support the Web-based thematic learning platform. Experimental results show that the fuzzy expert system is very effective in offering the learning guidance to the pupils. Besides, a series of leaving-one-out cross validation tests demonstrate the high prediction accuracy of three components of the composite classifier on the final report quality based on the pupils' learning profiles. In case an inconsistency between the learners' predicted result and their actual performance takes place, the teacher can also proceed with a further analysis on the cause of the discrepancy.

In the future work, we plan to build a platform for practicing theme-based learning between cooperative schools so that cross-classes or cross-schools theme-based learning can be spanned. The system will also have new developed mobile learning tools to support outdoor teaching by allowing the learner to access the archived learning resources. Meanwhile, the user interface will be redesigned to be more user friendly, and intelligent detecting mechanism and text mining techniques will be incorporated into the learner profile management agent to locate the distractive learners and questionable plagiaries, respectively.

Acknowledgments

The authors would like to thank the National Science Council of the Republic of China, Taiwan for financially supporting this research under Contract No. NSC 92-2213-E-026-001.

References

- [1] Buckley, J., and Eslami., *An introduction to fuzzy logic and fuzzy sets (advances in soft computing)*. Physica Verlag, 2002.
- [2] Denoeux, T. "A k-nearest neighbor classification rule based on Dempster-Shafer theory," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 25, no. 5, pp. 804-813.
- [3] El-Shal, S.M., and Morris, A.S., "A fuzzy expert system for fault detection in statistical process control of industrial processes," *IEEE Transactions on Systems, Man and Cybernetics, Part C*, vol. 30, no. 2, pp. 281-289, 2000.
- [3] *Applied data mining*. John Wiley & Sons Ltd., 2003.
- [4] Hall, M. and Holmes, G. "Benchmarking attribute selection techniques for discrete class data mining," *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, no. 3, pp. 1437-1447, 2003.
- [5] Hand, D., Mannila, H., and Smyth, P., *Principles of data mining*, MIT press, 2001.
- [6] Liu, J.N.K., Li, B.N.L., and Dillon, T.S., "An improved naive Bayesian classifier technique coupled with a novel input solution method [rainfall prediction]," *IEEE Transactions on Systems, Man and Cybernetics, Part C*, vol. 31, no. 2, pp. 249-256, 2001.
- [7] Liu, M. C., Chen, L. C., and Tung, S. C., "The implementation and study of the Web-based thematic learning model," *Proceeding of the 6th Global Chinese Conference on Computers in Education, Beijing*, 2002.