A Comparative Data Mining Technique for David Kolb's Experiential Learning Style Classification

Phanthipha Petchboonmee, Duangkamol Phonak, and Monchai Tiantong

Abstract—The objective of this research is to study the results of learning style classification and compare the efficiency of David Kolb's learning style classification of students in the Department of Computer Information System, Rajamangala University of Technology Lanna (Tak Campus). Thereby, the algorithms used in this research include J48, NBTree and NaiveBayes. The 10-fold Cross Validation was used to create and test the model, and the data was analyzed by the WAKA program. The data was collected by means of questionnaire from 502 students in the 1st semester of academic year 2013. The results show that the efficiency of classification by means of J48 technique had the highest value of Correct at 85.65% and it could be applied to develop David Kolb's learning style, which was correct and precise to classify the learning style.

Index Terms—Comparative, data mining technique, David Kolb's experiential learning style, classification.

I. INTRODUCTION

There has been a change of learning style in educational management in Thailand, which, at first, was the teacher-centered style where the teacher was the only teaching source the students could learn knowledge from, to the learner-centered where the teaching style emphasized the students' interest. According to the National Education Act of 1999, mentioned in Act 4, section 22, the educational management acknowledges that every learner is capable of learning and improving themselves. They are the center of learning and it is essential that the learning management must facilitate learners' full educational improvement. The Act 24, additionally, states that the educational management of any lessons or activities must be appropriate to the competence and interest of the learners, related to individual differences, which is effective enough to reflect learners' learning efficiency. Traditionally, to classify individual differences, an implemented learning style to analyze a learner's individual competence was applied. The learning styles are the perception and the learning each learner individually possesses, which are integrated with a learner's physical appearance, emotion, and perception and these show how effective each learner learns. Moreover, the learning styles are considered changeable habits. Nowadays, there exists a learner-centered education system in Thailand .The instructors have to base the contents and activities on interests and skills of individual learners. Accordingly, the learning style of each learner has an important effect on learning performance of the learners.

In this research, the theory of David Kolb's learning style was applied.

A. David Kolb's Learning Style

Kolb's learning style is based on his "experiential learning theory". David A. Kolb believes "learning is the process whereby knowledge is created through the transformation of experience". The theory presents a cyclical model of learning, consisted of four stages shown in the below diagram (Concrete Experience, Reflective Observation, Abstract Conceptualization, Active Experimentation). One may begin at any stage, but must follow each other in the sequence.

This cyclical model consists of two dimensions, perception dimension and processing dimension. Concrete experience and abstract conceptualization are two polars of perception dimension, which represent the perception context or experimentation means of individual preferences, while reflective observation and active experimentation are two polars of processing dimension, which represents the processing or means of transforming future information of individual preferences. Kolb believes that the learning style is the way of individual's information perception and processing. In the combination of these two dimensions of experiential learning, different preferences lead to different learning styles, i.e. Divergent, Assimilator, Converger, and Accommodator.

A divergent takes experience and deep thought, thus diverging from a single experience to multiple possibilities in terms of what this might mean. They are generally influenced by other people and like to receive constructive feedback. They like to learn via logical instruction or hands one exploration with conversations that lead to discovery.

An assimilator has the most cognitive approach, preferring to think than to act. The question 'What is there to know?' takes organized and structured understanding. They prefer lectures for learning, with demonstrations where possible, and will respect the knowledge of experts. They will also learn through conversation that takes a logical and thoughtful approach. They often prefer the clean and simple predictability of internal models to external messiness. The best way to teach an assimilator is with lectures that start from high-level concepts and work down to the detail. Give them reading material, especially academic papers and they'll gobble it down. Do not teach through play as they like to stay serious.

The converger thinks about things and then tries them out to see if they work in practice. They like to ask 'How' about a situation. Understanding how things work in practice, they like facts and will seek to make things efficient by making small and careful changes. They prefer to work by themselves, thinking carefully and acting independently.

They learn through interaction and computer-based learning which is more effective with them than other methods.

Accommodators have the most hands-on approach, a strong preference for doing rather than thinking. They like to ask 'What if?' and 'Why not?' to support their action-first approach. They do not like routine and will take creative risks to see what happens. They like to explore complexity by direct interaction and learn better by themselves than with other people. As might be expected, they like hands-on and practival learning rather than lectures [1]-[3].

B. Data Mining

Regarding the analysis of learning style, the data mining technique has been widely used in data classification today. Data mining (DM) or Knowledge Discovery in Database (KDD), is an approach to discover useful information from large amount of data. DM techniques apply various methods in order to discover and extract patterns from stored data. The pattern found will be used to solve a number of problems occurred in many fields such as education, economic, business, statistics, medicine, and sport. The large volume of data stored in those areas demands for the DM approach because the resulting analysis is much more precise and accurate [4].

In recent years, there has been an increasing interest in the use of DM to investigate the educational field. Educational Data Mining (EDM) is concerned with developing methods and analyzing educational content to enable a better understanding of students' performance. It is also important to enhance teaching and learning process [5], [6].

This research applied three data classification techniques: J48, NBTree and NaïveBayes to classify David Kolb's experiential learning style.

C. Data Classifier

Classification analysis is the organization of data in given classes. Also known as supervised classification, the classification uses five class labels to order the objects in the data collection. Classification approaches normally use a training set where all objects are already associated with known class labels. The classification algorithm learns from the training set and builds a model. The model is use to classify new objects [6].

J48 is an open source Java implementation of the C4.5 algorithm in the WEKA data mining tool. C4.5 is an algorithm used to generate a decision tree. C4.5 is a software extension and thus improvement of the basic ID3 algorithm. The decision trees generated by C4.5 can be used for classification, and for this reason, C4.5 is often referred to as a statistical classifier. C4.5 is an evolution and refinement of ID3 that accounts for unavailable values, continuous attribute value ranges, pruning of decision trees, rule derivation, and so on. A set of records are given [6]-[8].

NBTree is a hybrid algorithm with Decision Tree and Naïve-Bayes. The algorithm is similar to the classical recursive partitioning schemes except that the leaf nodes create Naïve-Bayes categorizers instead of node predicting a single class [9].

NaïveBayes algorithm is a simple probabilistic classifier that calculates a set of probabilities by counting the frequency and combinations of values in a given data set. The algorithm uses Bayes theorem and assumes all attributes to be independent given the value of the class variable. This conditional independence assumption rarely holds true in real world applications, hence the characterization as Naïve yet the algorithm tends to perform well and learn rapidly in various supervised classification problems [6], [10].

D. WEKA

WEKA is a data mining system developed by the Universit of Wailato in New Zealand that implements data mining algorithms using the JAVA language. WEKA is a state of the art facility for developing machine learning (ML) techniques and their application to real-world data mining problems. It is a collection of machine learning algorithms for data mining tasks. The algorithms are applied directly to a dataset. WEKA implements algorithms for data preprocessing, classification, regression, clustering and association rules; It also includes visualization tools. The new machine learning schemes can also be developed with this package. WEKA is open source software issued under General Public License.

The data file normally used by Weka is in ARFF file format, which consists of special tags to indicate different things in the data file (foremost: attribute names, attribute types, attribute values and the data). The main interface in Weka is the Explorer. It has a set of panels, each of which can be used to perform a certain task. Once a data set has been loaded, one of the other panels in the Explorer can be used to perform further analysis [11], [12].

II. LITERATURE REVIEW

H. Bahiah Hj Ahmad and S. Mariyam Shamsuddin compared performance of several classifiers provided in WEKA such as bayes, decision tree and classification rules in classifying student's learning style. The experiment showed that the tree classifiers have high accuracy with more than 91% accuracy. The sizes of the tree and the number of leaves among the tree classifies techniques have also been observed [13].

S. Pal used a data mining application to generate predictive models for engineering student's dropout management. The results show that the machine learning algorithm is able to establish effective predictive model from the existing student dropout data [14].

S. Kumar Yadav, B. Bharadwaj and S. Pal obtained the university students data such as attendance, class test, seminar and assignment marks from the students' database, to predict the performance at the end of the semester using three algorithms ID3, C4.5 and CART and shoes that CART is the best algorithm for classification of data [15].

M. Wook, Y. Hani Yamaya, N. Wahab, M. Rizal Mohd Isa, N. Fatimah Awang and H. Yann Seong compared two data mining techniques which are: Artificial Neural Network and the combination of clustering and decision tree classification techniques for predicting and classifying student's academic performance. As a result, the technique that provides accurate prediction and classification was chosen as the best model. Using the proposed model, the pattern that influences the student's academic performance was identified [16].

N. Thai Nghe, P. Janecek and P. Haddawy compared the accuracy of decision tree and Bayesian network algorithms for predicting the academic performance of undergraduate

and postgraduate students at two very different academic institutes. These predictions are most useful for identifying and assisting failing students, and better determine scholarships. As a result, the decision tree classifier provides better accuracy in comparison with the Baysian network classifier [17].

III. METHODOLOGY

This research applied the data mining technique of David Kolb's experiential learning style classification, including the following steps:

1) Data Selection: The data herein was collected by means of rating-scale questionnaire, which was divided into 2 parts. Part 1 was about general information with several variables for data analysis, i.e. gender, education level, former education background, preferred learning styles, learning styles that the subjects were skilled at, and learning styles (Table I). Part 2 contained 32 items of David Kolb's experiential learning style classification questionnaire. Then, the data was collected from 502 students in Department of Computer Information System, Rajamangala University of Technology Lanna (Tak Campus) in the 1st semester of academic year 2013.

TABLE I: VARIABLES FOR DATA CLASSIFICATION

Variable	Posible values
Gender	
Male	1
Female	2
Education level	
Bachelor degree (4 years)	1
Bachelor degree (2 years)	2
Former education background	
High school (grade 12)	1
Vocational Certificate	2
HighVocational Certificate	3
Perferred learning styles	
Listen to the lectures of instructors	1
Self-learning	2
Practice by oneself	3
Mixed styles	4
Learning style that the subject were skilled at	
Read	1
Write	2
Listen	3
Doing	4
Learning style	
Divergent	1
Assimilator	2
Accommodator	3
Converger	4

2) Data Transformation: The data derived from the questionnaire was transformed into the proper format in order to be analyzed based on selected algorithms. The data was represented by numbers and stored in the form of a CSV file.

3) Creation and test of the data classification model were conducted by WEKA program with the algorithms J48, NBTree and NaïveBayes. The model was tested by means of 10 - fold cross-validation to find out the values of Correctly Classified, Precision, Recall and F-Measure. Then, the results of the tests were compared in terms of efficiency of each data classification technique.

IV. RESULT OF THE RESEARCH

The Table II contains the results of efficiency analysis of each data classification technique, showing Correctly Classified Instances and Incorrectly Classified Instances. In addition, the table presents the values of Precision, Recall and F-Measure, classified as to the following learning styles.

TABLE II: CLASSIFICATION ACCURACY TEST RESULES

Algorithm	Precision	Recall
J48	0.983	0.783
NBTree	0.962	0.796
NaiveBayes	0.760	0.887

According to Table II, the results of efficiency analysis on David Kolb's experiential learning style classification for students in the Department of Computer Information System, Rajamangala University of Technology Lanna (Tak Campus) show that the efficiency of classification by means of J48 technique had the highest value of Correct at 85.65%, followed by NBTree 85.45%, and NaiveBayes 74.49%, respectively.

TABLE III: PERCISION, RECALL AND F-MEASURE FOR DIVERGENT LEARNING STYLE

Algorithm	Precision	Recall	F-measure
J48	0.983	0.783	0.872
NBTree	0.962	0.796	0.917
NaiveBayes	0.760	0.887	0.818

According to Table III, the results of efficiency analysis on David Kolb's experiential learning style classification for students in the Department of Computer Information System, Rajamangala University of Technology Lanna (Tak Campus) show that Divergent, analyzed by means of J48, had the highest value of Precision at 0.983, NaiveBayes had the highest value of Recall at 0.887, and NBTree had the highest value of F-Measure at 0.917.

TABLE IV: PERCISION, RECALL AND F-MEASURE FOR ASSIMILATOR LEARNING STYLE

Algorithm	Precision	Recall	F-measure
J48	0.71	1	0.831
NBTree	0.698	0.973	0.824
NaiveBayes	0.652	0.626	0.637

According to Table IV, the results of efficiency analysis on David Kolb's experiential learning style classification for students in the Department of Computer Information System, Rajamangala University of Technology Lanna (Tak Campus) show that Assimilator, analysed by means of J48, had the value of Precision at 0.71, with the value of Recall 1, and F-Measure 0.831, all of which were higher than NBTree and NaiveBayes.

TABLE V: PERCISION, RECALL AND F-MEASURE FOR ACCOMMODATOR LEARNING STYLE

Algorithm	Precision	Recall	F-measure
J48	0.959	0.922	0.940
NBTree	0.959	0.922	0.940
NaiveBayes	0.935	0.569	0.707

According to Table V, the results of efficiency analysis on David Kolb's experiential learning style classification for students in the Department of Computer Information System,

Rajamangala University of Technology Lanna (Tak Campus) show that Accommodator, analyzed by means of J48 and NBTree, had the value of Precision at 0.959, with the value of Recall 0.922, and F-Measure 0.707, all of which were higher than NaiveBayes.

TABLE VI: PERCISION, RECALL AND F-MEASURE FOR CONVERGER LEARNING STYLE

Algorithm	Precision	Recall	F-measure
J48	0.90	0.759	0.824
NBTree	0.90	0.759	0.824
NaiveBayes	0.861	0.747	0.80

According to Table VI, the results of efficiency analysis on David Kolb's experiential learning style classification for students in the Department of Computer Information System, Rajamangala University of Technology Lanna (Tak Campus) shown that Converger, analyzed by means of J48 and NBTree, had the value of Precision at 0.90, with the value of Recall 0.759, and F-Measure 0.824, all of which were higher than NaiveBayes.

V. CONCLUSION

This research involves learning style classification and the comparison of efficiency of David Kolb's experiential learning style classification for students in the Department of Computer Information System, Rajamangala University of Technology Lanna (Tak Campus) .The results show that the efficiency of classification by means of J48 technique had the highest value of Correct at 85.65% and it could be applied to develop David Kolb's learning style, which was correct and precise to classify the learning style. Thereby, the research would be beneficial for the instructional management that is suitable for the learning styles of the learners. This is to create achievement motive as well as to enhance the efficiency of the learners.

REFERENCES

[1] Z. Jinewei, "Emprical study on online interaction Based on learning style differences," in *Proc. the 6th International Conference on Computer Science & Education in ICCSE 2011)*, pp. 1023-1026, 2011.

[2] F. Kaya, N. Ozabaci, and O. Tezek, "Inveatigating primary school seconf grade students' learning styles according to the kolb learning style model in terms of demographic variables," *Journal of Turkish Science Education,* April 2009.

[3] Y. Chen. (2011). Learning Styles and Adopting Facebook Technology. *Chinese Culture University, Taiwan.* [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6017676

[4] B. R. Sachin and S. M. Vijay, "A survey and Future Vision of data mining in educational field," in *Proc. International Conference on Advanced Computing & Communication Technologies,* pp. 96-100, 2012.

[5] E. Miranda, "Data mining as a techniqur to analyze the learning styles of students in using the learning management system," presented at Semimar Nasional Aplikasi Teknologi Informasi, June 17-18, 2011.

[6] T. R. Patil and S. S. Sherekar, "Performance analysis of naïve Bayes and J48 classification algoithm for data classification," *International Journal of Computer Science and Applications,* vol. 6, pp. 256-261, 2013.

[7] A. K. Sharma and S. Sahni, "A comparative study of classification algorithms for spam email data analysis," *International Journal on Computer Science and Engineering (IJCES),* vol. 3, pp. 1890-1895, 2011.

[8] T. C. Sharma and M. Jain, "WEKA approach for comparative study of classification algorithm," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 4, pp. 1925-1931, April 2013.

[9] P. Pumpuang, A. Srivihok, and P. Praneetpolgrang, "Comparisons of classifier algorithmss: Bayesian network, C4.5, decision forest and NBTree for course registration planning model of undergraduate students," presented at the IEEE International Conference on Systems, Man and Cubermetics (SMC), 2008.

[10] M. Koklu and Y. Unal, Analysis of a population of diabetic patients databases with classifiers," *World Academy of Science, Engineering and Technology,* vol. 80. pp. 776-778.

[11] M. F. B. Othman and T. M. S. Yau, "Comparison of Different Classification Techniques Using WEKA for Breast Cancer," in *IFMBE Proceedings*, vol. 15, pp. 520-523, 2007.

[12] E. Bhuvaneswari and V. R. S. Dhulipala, "The Study and analysis of classification algorithm for animal kingdom dataset," *Information Engineering,* vol. 2, no. 1, pp. 6-13, March 2013.

[13] N. B. H. Ahmad and S. M. Shamsuddin, "A comparative analysis of mining techniques for automatic detection of student's learning style," presented at the International Conference on Intelligent Systems Design and Applications, 2010.

[14] S. Pal, "Mining Educational data using classification to decrease dropout rate of students," *International Journal of Multidisciplinary Sciences and Engineering,* vol. 3, no. 5, pp. 35-39, May 2012.

[15] S. K. Yadav, B. Bharadwaj, and S. Pal, "Mining education data to predict student's retention: a comparative study," *Intermational Journal of Computer Science and Information Security,* vol. 10, no. 2, 2012.

[16] M. Wook, Y. H. Yahaya, N. Wahab, M. R. M. Isa, N. F. Awang, and H. Y. Seong, "Prediction NDUM student's academic performance using data mining techniques," presented at the International Conference on Computer and Electrical Engimeering, 2009.

[17] N. T. Nghe *et al.*, "A comparative analysis of technique for predicting academix performance," presented at ASEE/IEEE Frontiers in Education Conference, October 10-13, 2007.

Phanthipha Petchboonmee is a Ph.D. candidate in computer education at division of Computer Edication, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok (KMUTNB), Thailand. She is an assistant professor at Division of Computer Information System, Faculty of Business Administation and Liberal Arts, Rajamangala University of Technology Lanna (Tak Campus), Thailand.

Duangkamol Phonak is a Ph.D.. She is working as a lecturer more than 10 years in Computer Education, Faculty of Techmical Education, KMUTNB. She has several specialized areas such as curriculum, teaching methodology, statistical techniques for data mining, programming for software and applications, technologies and applications, the professional qualifications of ICT.

Monchai Tianting is a Ph.D.. He is an associate professor at division of Computer Edication, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok (KMUTNB), Thailand. He is working as an associate professor more than 20 years in Information Technology and Computer Education, Faculty of Technical Education, KMUTNB. He is actually member of AACE, IEEE, ACM, and etc.