Rough Set Approach for Classifying Student's Learning Style : A Comparative Analysis

Nor Bahiah Ahmad, Siti Mariyam Shamsuddin and Norsham Idris
Faculty of Computing, Universiti Teknologi Malaysia, 81310, Johor Bahru, Johor, Malaysia
[e-mail: bahiah@utm.my]
*Corresponding author: Nor Bahiah Ahmad

Abstract

The student's interaction in e-learning which were captured in the log file can be intelligently examined to diagnose students' learning style. This is important since a student's behaviour while learning online is among the significant parameters for adaptation in e-learning system. Currently, Felder Silverman (FS) is a common learning style model that is frequently used by many researchers for personalizing learning materials based on learning style. There are four learning style dimensions in FS model and most researches need to develop four classifiers to map the characteristics into the dimensions. Such approach is quite tedious in terms of data pre-processing and it also time consuming when it comes to classification. Therefore, this study propose mapping the students' characteristics into Integrated Felder Silverman (IFS) learning styles, by combining the four learning dimensions in FS model into sixteen learning styles. However, the most crucial problem for IFS model is the difficulties in identifying the significant pattern for the classifier that has high dimension and large number of classes. In this study, fifteen features have been identified as the granule learning features for IFS. Comparative analysis of the Rough Set performance between IFS classifier and the conventional four classifiers shows that the proposed IFS gives higher classification accuracy and rule coverage in identifying student's learning style. However, Rough Sets generate very large rules for IFS compared to the conventional FS four classifiers.

Keywords**:** Felder Silverman, learning style, Rough Set, rule generation, rule support, rule length

1. Introduction

Learning style has become a significant factor contributing in learner progress and many researchers agree that incorporating learning styles in education has potential to make learning easier for students and increases learning efficiency [1,2,3]. The importance of learning style that can increase student's performance has led to the efforts in developing an adaptive learning system that adapt the course content based on the user features such as the student's learning style, background and preferences [2,4].

Felder Silverman (FS) model is among the most adopted learning style model. This model was initially developed by Felder and Silverman for engineering students [5]. This model categorized a student's dominant learning style along a scale of four dimensions: active-reflective (how information is processed), sensing-intuitive (how information is perceived), visual-verbal (how information is presented) and global–sequential (how information is understood).

Early research has focused on student's

This research is supported by Fundamental Research Grant Scheme (FRGS), vote number 4F496 and Flagship Grant, vote number 02G69.

learning style by using questionnaire to assess the student's learning characteristics [6,7,8]. However, the exploitation of questionnaires is time consuming and unreliable approach for acquiring learning style characteristics and may not be accurate [9]. Most questionnaires are too long, hence, causing students to choose answers arbitrarily instead of thinking seriously about them. Due to these problems, several studies have been conducted in detecting student's learning style that are based on the student's browsing behavior [10, 2, 4]. This approach can be implemented successfully since the style of student's interaction with the system can be inferred accurately and can be used as attributes for adaptation purposes.

Student's learning characteristics and the choice of learning materials in e-learning environment have been used in previous research to classify student's learning style. Previous research need to conduct four classifiers to predict student's learning style into four FS learning dimensions. Thus it is desirable to integrate the FS four dimensions into 16 learning styles with only one classifier. This paper intend to compare the performance of Rough Set classifier for identifying student's learning style using conventional four classifiers with Rough Sets performance for IFS with only one classifier.

2. Previous Work

Several studies focusing on learning style detection based on user interaction in the system have been conducted. Currently, there is no benchmark data that can be used for classification of learning styles. All researchers in this area developed their own e-learning system, either in LMS environment or developed a prototype version. Various e-learning systems for different subjects have been deployed and used by students [2,10]. The student's interactions are recorded in the log data. The data is preprocessed and transformed into numerical format before it can be fed to the classifier.

Various techniques, approaches and purposes of detecting learning styles have been implemented in the researches in order to find the appropriate classifier for each case study. Among the popular techniques being used are rule classifier [2,4], Neural Network [9,11], Decision Tree [12], Bayesian Networks [13,14] and Genetic Algorithm [15]. Each technique has its own strength in detecting the student's characteristics. All researches develop four classifiers in order to detect the four learning dimensions in FS, except the study conducted by [10] and [16] that combined the FS four dimension into 16 learning styles. The four classifiers approach is quite time consuming in identifying and pre-processes the data for the four classifiers. Therefore, integrated FS is needeed for faster learning style classification.

3. Research Methodology

Fig. 1 shows the phases conducted in this study. In phase 1, an e-learning system has been developed that contain the learning resources structured into components that are suitable for FS learning dimensions. Among the resource materials provided in the learning systems are forum, animation, source codes demonstration, hypertext, power point slides, on-line exercises and on-line assessment.

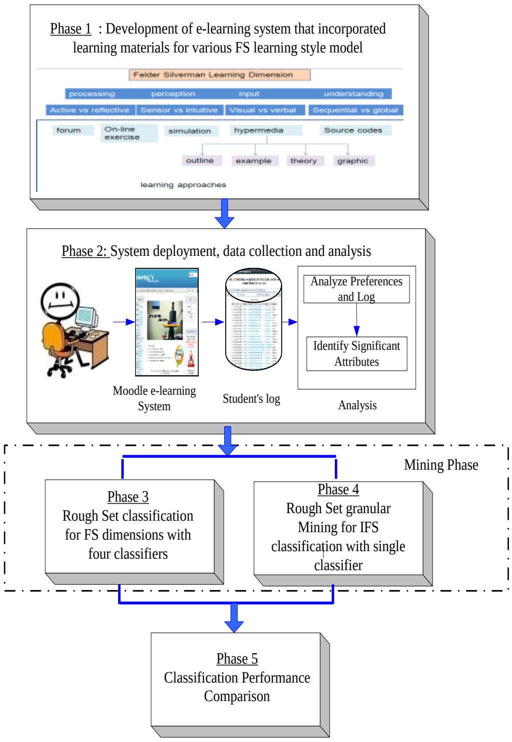

Fig. 1. Operational framework of the study.

In order to determine which characteristics of the students can be used to identify their learning style, two experiments have been conducted for two semesters among Computer Engineering students and Computer Science students. There were 136 students participated in this study. The experiments took place at Faculty of Computing, Universiti Teknologi Malaysia (UTM). The subject chosen for this experiment is Data Structure and Algorithms.

Table 1 shows the Integrated Felder Silverman (IFS) that incorporates FS dimensions such as processing, perception, input and understanding into 16 (2^4) learning styles.

Table 1. Sixteen learning styles in IFS.

IFS Learning Styles	**Label**
Active/Sensor/Visual/Sequential	ASViSq
Reflective/Sensor/Visual/Sequential	RSViSq
Active/Intuitive/Visual/Sequential	AIViSq
Reflective/Intuitive/Visual/Sequential	RIViSq
Active/Sensor/Verbal/Sequential	ASVbSq
Reflective/Sensor/Verbal/Sequential	RSVbSq
Active/Intuitive/Verbal/Sequential	AIVbSq
Reflective/Intuitive/Verbal/Sequential	RIVbSq
Active/Sensor/Visual/Global	ASViG
Reflective/Sensor/Visual/Global	RSViG
Active/Intuitive/Visual/Global	AIViG
Reflective/Intuitive/Visual/Global	RIViG
Active/Sensor/Verbal/Global	ASVbG
Reflective/Sensor/ Verbal/Global	RSVbG
Active/Intuitive/Verbal/Global	AIVbG
Reflective/Intuitive/Verbal/Global	RIVbG

Rough Set technique introduced by [17] is used to compare the performance of the two classification models, which are FS and IFS model. Rough Sets steps involve seven processes and listed as follows:

i) Mapping of information into the decision system format
ii) Data Completion and pre-processing.
iii) Data Discretization
iv) Splitting Data into training and testing
v) Reduct Computation
vi) Rules Generation and Classification

4. Rough Set Experimental Result

4.1 Rough Set Classification for Four FS Dimensions

Rough set offers some important techniques in managing an information system and consists of several steps leading towards the final goal of generating rules from information systems. The main steps of Rough set approach are data preparation, discretization, reducts computation, rules generation and classification. This section describes how Rough Set is used to generate four classifiers to classify learner into active/reflective, sensor/intuitive, visual/ verbal or sequential/global. Table 2 summarizes the characteristics for every FS classifier. The table reveals that large number of attributes tends to generate larger number of reducts and rules such as shown for classifying Active/Reflective. Meanwhile, classifying sequential/global learners with only 3 attributes generates least reduct and rules.

Table 2. Characteristics for FS classifiers

FS Dimension	**No. of Attri-butes**	**No. of Redu cts**	**No. of Rules**	**Rule Leng th**	**Rule Supp ort**	**Accu racy (%)**
Active/ Reflective	7	96	682	2-6	1-172	93.44
Sensor/ Intuitive	6	40	159	1-5	1-457	98.8
Visual/ Verbal	5	22	176	2-5	1-280	93.24
Sequential/ Global	3	3	10	1-2	1-102 7	100

Overall, Rough Set performance is good in classifying the behavior data set into four FS dimensions with 93.44% accuracy for processing dimension, 98.8% for perception dimension, 93.24% for input dimension and 100% accuracy for understanding dimension.

4.2 Rough Set Classification for IFS

This section describes the experiment for identifying the students IFS learning style. The combination of 4 FS dimensions into 16 IFS learning styles has led to 15 significant attributes being identified to classify students into 16 integrated FS learning styles. The attributes have been chosen based on the analysis of the student's behavior while using e-learning system. In this study, the high number of attributes and large number of classes in the decision table are generated; hence, classification and the analysis protocol become difficult due to the complexity of the student's behavior data set.

The experiment conducted reveal that Boolean Reasoning discretization always give the highest accuracy compare to others, while Genetic Algorithm with object reduct always give the highest accuracy compare to other reducers. Therefore, these two techniques are chosen for Rough Sets rule generation for IFS. The next step is to generate the rules from reducts based on BR discretization and GA object reduct. Rule generation plays an important role in classifying the output for the classifier. Rough set classifier was run using 10 fold data sets and it is observed that the sample from fold 2 and fold 5 has the highest classification accuracy (97.61%) compared to the other folds. The characteristics of the generated rules for fold 2 and 5 are shown in Table 3. The characteristics consist of the number of reducts, number of rules, rule support, rule length and classification accuracy for each

Table 3. The rule characteristics for fold 2 and fold 5 dataset.

Rule Characteristics	**Fold 2**	**Fold 5**
No. of Reducts	***16280***	**16316**
No. of Rules	***82287***	**82914**
Rule Support	***1-67***	**1-67**
Rule Length	***2-11***	**2-10**
Classification Accuracy	**97.61**	**97.61**

From the experiment, there is no obvious relationship between the number of reducts and the number of rules, and also the relationship between the number of rules and the classification accuracy. However, it can be observed that Rough Set classifier generates a very large number of rules in all samples. The next step is to filter unimportant rules based on the rule length and rule support [18]. Table 4 reveal the result after filtering. It can be seen that after filtering the rules, the classification result are still significant, even though the classification accuracy decrease. The most interesting result can be seen that even after only 12% rules left the classification accuracy are still high for fold 2 (94.25) and fold 5 (93.38%).

Table 4. Result after Rules are filtered

	All Rule (100%)	**Filtered Rule (98.6%)**	**Filtered Rule (67.3%)**	**Filtered Rule (12%)**
Length	**2-11**	**4-8**	**4-8**	**4-8**
Support	**1-67**	**1-67**	**2-67**	**8-67**
No Rules	**82287**	**81757**	**55828**	**9953**
Fold 2 accuracy	98.53	98.53	98.53	***94.12***
Fold 5 accuracy	97.79	97.79	97.79	***93.38***
Average	98.16	98.16	98.16	***93.75***

5. Comparative Analysis of Rough Set Classifiers

The experimental results reveal that classifying student behavior into IFS is able to give better performance compared to classifying FS dimensions. Table 5 gives the comparative analysis of Rough Sets performance on classifying both models. It can be seen that classifying the student behavior using IFS gives higher classification accuracy in all dimension as compared to the classification using conventional approach with four classifiers. However, both approaches give 100% accuracy result for understanding dimension.

Table 5. Comparative analysis of IFS and FS performance.

	Processing Active/ Reflective	**Perception** Sensor/ Intuitive	**Input** Visual/ Verbal	**Understanding** Sequential/ Global
IFS Testing (Fold 2)	99.6	99.8	98.21	100
IFS Testing (Fold 5)	99.6	99.6	98.4	100
FS Testing	93.44	98.8	93.24	100

Filtering the rules based on rule length and rule support implies that the rules with lower support are not important since eliminating these rules doesn't have effect on the accuracy, unlike rules with higher support, which are really vital for higher accuracy. By filtering the rules up to 12% the accuracy rate is decreasing by 4%. In this domain, since the number of rules generated is very large, it is recommended to use fewer rules with significant result [19].

6. Conclusions

This paper discusses the approach of classification of four FS dimensions and IFS using Rough Sets. Rough Set is able to classify all 4 dimensions with significant result. Four classifiers have been generated and the rules for every classifier have been extracted. Mapping four dimensions of FS into IFS that only need one classifier to do the classification show that IFS performance is better than FS performance. However, the high number of attributes and classes in IFS (15 attributes and 16 classes) generates very excessive amount of reducts and rules compare to FS classifiers with small number of attributes and only two classes. Filtering the rules by eliminating the rules with less support and highest length show that the rules with less support and longer length are not important since by eliminating the rules with these criteria, the classifier can still maintain a good classification accuracy result. After filtering the rules, the classification result is still significant with the average of 93.75% accuracy even though with only 12% rules left for classification.

Acknowledgements

This paper is supported by University Flagship Grant, vote number 02G69 and Fundamental Research Grant Scheme (FRGS), vote number 4F496. The authors would like to express their deepest gratitude to Ministry of Education Malaysia and Research Management Centre, Universiti Teknologi Malaysia for the support in R & D, UTM Big Data Centre and Soft Computing Research Group for the inspiration in making this study a success

References

[1] Mampadi, F., Chen, S.Y.H., Ghinea, G. and Chen, M.P, "Design of Adaptive Hypermedia Learning Systems: A Cognitive Style Approach". *Computers & Education*, Vol. 56(4), pp.1003-1011, 2011.

[2] Popescu E., "Adaptation Provisioning with Respect to Learning Styles in a Web-based Educational System: An Experimental Study". *Jounal of Computer Assisted Learning*. Vol. 26. Pp. 243-257, 2010.

[3] Kinshuk, Liu T-C and Graf S., "Coping with Mismatched Courses: Student's Behaviour and Performance in Course Mismatched to Their Learning Styles". *Education Tech Research Dev*. Vol. 57, pp. 739-752, 2009.

[4] Graf, S. "Adaptivity in Learning Management System Focusing on Learning Styles". *Phd Thesis*, Vienna University of Technology, 2007.

[5] Felder R. and Silverman L., "Learning And Teaching Styles In Engineering Education", *Engineering Education*, Vol. 78 (7), pp. 674–681, 1988.

[6] Wolf C., "iWeaver: Towards an Interactive Web-Based Adaptive Learning Environment to Address Individual Learning Styles", *Proceedings Fifth Australasian Computing Education Conference (ACE2003)*, Adelaide, Australia. pp. 273-279, 2003.

[7] Papanikolau K., Grigoriadou M., Knornilakis H., and Magoulas G., "Personalizing the Interaction in a Web-based Educational Hypermedia System: the case of INSPIRE", *User Modeling and User-Adapted Interaction* Vol. (13), pp. 213 – 267, 2003.

[8] Triantafillou E., Pomportsis, A., A., and Georgiadou, E., "AES-CS: Adaptive Educational System based on Cognitive Styles", *Second International Conference on Adaptive Hypermedia and Adaptive Web-based Systems*, Malaga, Spain, May 29-3, 2002.

[9] Villaverde J., Godoy D. and Amanda A. "Learning Styles' Recognition in E-Learning Environments with Feed-Forward Neural Networks", *Journal of Computer Assisted Learning*, Vol. 22(3), pp. 197—206, 2006.

[10] Klašnja-Milićević A., Vesin B., Ivanović M. and Budimac Z., "E-learning

Personalization Based On Hybrid Recommendation Strategy and Leaning Style Identification". *Computers & Education*, Vol. 56, pp. 885-899, 2011.

[11] Lo J. and Shu P., "Identification Of Learning Styles Online by Observing Learners' Browsing Behaviour Through A Neural Network", *British Journal of Educational Technology*. Vol 36 (1). pp 43–55, 2005.

[12] Cha, H.J., Kim, Y.S., Lee, J.H. and Yoon, T.B., "An Adaptive Learning System with Learning Style Diagnosis Based On Interface Behaviors". In: *Workshop Proceedings of International Conference on E-Learning and Games*, 2006. pp. 513–524.

[13] Garcia E., Romero C., Ventura S. and Calders T., "Drawbacks and Solutions of Applying Asoociation Rule Mining in Learning Management Systems". *Proceedings of the International Workshop on Applying Data Mining in e-Learning (ADML 07)* 2007, pp. 13-22.

[14] Kelly D. and Tangney B., "Predicting Learning Characteristics In A Multiple Intelligence Based Tutoring System", *LNCS* Vol. 220/2004. Springer Berlin/Heidelberg. pp.678-688, 2004.

[15] Yaannibelli V., Godoy D. and Amanda A., "A Genetic Algorithm Approach to Recognize Students' Learning Styles", *Interactive Learning Environments*. Vol. 14(1), pp. 55-78. 2006.

[16] Fabiano A. D., Luciano V. L., M´arcia A. F. and Carlos R. L., "A New Approach to Discover Students Learning Styles in Adaptive Educational Systems". (2011)

[17] Pawlak Z. "Rough Set Theory and its Applications to Data Analysis". *Cybernetics and System.* Vol. 29(7): pp. 661-688, 1998.

[18] Bose I. (2006). "Deciding the Financial Health of Dot-coms Using Rough Sets". *Information and Management*, pp. 835-846, 2006.

[19] Pai P-F, Lyu Y-J and Wang Y-M. "Analyzing Academic Achievement of Junior High School Students by in Improved Rough set Model". *Computers & Education,* Vol. 54, pp.889-900, 2010.