Classification of Learning Styles in Virtual Learning Environment using Data Mining: A Basis for Adaptive Course Design

Renato Racelis Maaliw III

Faculty, College of Industrial Technology, Southern Luzon State University, Lucban, Quezon, Philippines
rmaaliw@slsu.edu.ph or renatomaaliw3@yahoo.com

---***---

Abstract - *The objective of this research is to study the results and compare several classifiers such as Bayes and Decision trees in classifying student's learning styles in a Virtual Learning Environment. This approach was experimented initially on 108 students of Computer Programming 1 online course created using Moodle. Student's behaviors have been extracted from Moodle log data and the learning style for each student was mapped according to Felder-Silverman Learning Style Model. A 10-fold cross validation was used to evaluate the selected classifiers. Classification accuracy and Kappa statistics have been observed to measure the performance of each classifier. The results show that the efficiency of classification by means of J48 technique had the highest average value of correctly classified instances at 89.91% accuracy and it could be used to infer the learning styles of students in a Virtual Learning Environment.*

Key Words: classification, educational data mining, Felder-Silverman learning style model, learning styles, virtual learning environment

1. INTRODUCTION

There are increasing research interest in utilizing data mining in the field of education. This new emerging discipline is known as Educational Data Mining (EDM). Its primary concern is developing methods for exploring the diverse and unique types of data that come from educational settings. At present, Virtual Learning Environments (VLEs) increasingly serve as a vital infrastructure of most universities that enable teachers to provide students with different representations of knowledge and to enhance interaction between teachers and students, and even amongst students themselves.

Virtual learning Environments usually provide online tools for assessment, communication, uploading of content and various features. Whilst traditional teaching methods, such as face-to-face lectures, tutorials, lab assignments, and mentoring remain dominant in the educational setting, universities are heavily investing in learning technologies to facilitate improvements with respect to the quality of learning [1]. Despite the ever-increasing practice of using e-learning in educational institutions, most of these applications perform poorly in motivating students to learn. There are many issues that are not addressed due to the very complex and varying ideas in the development. It fails to meet the needs of students and fail to serve the ultimate goal of having on-line learning.

But what is almost completely overlooked is a vast collection of data that resides inside these specific environments. All of this data represents a potentially valuable source which is not adequately considered. The data stored in these VLEs can be used to improve the learning and pedagogical process to make it more efficient for both teachers and learners. Specifically, it can be used in the identification or classification of student's learning styles (LS). Notable educational theorist and researchers consider learning style as an important factor that affects the learning process. Understanding how different individual learn is the key to a successful teaching and learning.

The study is based on a widely accepted theory that each learner has an individual or specific learning style. A learner with specific learning style can face difficulties while learning, when their learning style is not supported by the teaching environment thus as a precursor to an adaptive Virtual Learning Environment the research initially focuses on the automatic identification of student's learning styles using data mining techniques based on their behaviors on a Virtual Learning Environment. In terms of learning style model, Felder-Silverman learning style model (FSLSM) was used for the reason that is often used in technology-enhanced learning [3]. Moreover, FSLSM describes the learning style of a learner in more detail, distinguishing between preferences on four dimensions as compared to other learning style models that classify learners in only a few groups.

2. RELATED LITERATURE

2.1 Learning Styles (LS)

A learning style is a student's consistent way of responding to and using stimuli in the context of learning. Reference [4] defines learning styles as the composite of characteristic cognitive, affective, and physiological factors that serve as relatively stable indicators of how a learner perceives, interacts with, and responds to the learning environment. Reference [5] defines learning style as those educational conditions under which a student is most likely to learn. They are not concerned with what learners learn, but rather how they prefer to learn. Learning styles are points along a scale that help discovers the different forms of mental representations. When individual tries to learn something new they prefer to learn it by listening to someone, talk to someone, or perhaps they prefer to read about a concept to learn it, or perhaps would like to see a demonstration.

Learning styles can be defined, classified, and identified in many different ways. It can also be describe as a set of factors, behaviors, and attitudes that enhance learning in any situation. How the students learn and how the teachers teach, and how the two interact with each other are influenced by different learning styles. Each person is born with and has certain innate tendencies towards a particular style, and these biological characteristics are influenced by external factors such as cultures, personal experiences, and developments. Each learner has a different and consistent preferred ways of perception, organization and retention. These learning styles are the indicators of how learners perceive, interact with, and respond to the learning environments. Students have different styles of learning, and they learn differently from one another. There are sufficient evidences for the diversity in individual's thinking and ways of processing various types of information, and shown that students will learn best if taught in a method deemed appropriate for their learning styles [6].

2.2 Felder-Silverman Learning Style Model (FSLSM)

One of the most widely used models of learning styles is the Index of Learning Styles (ILS) [7] developed by Richard Felder and Linda Silverman. The learning style model unlike other model is based on tendencies, indicating that learners with a high preference for certain behavior can also act sometimes differently. FSLSM [8] is used very often in advanced learning technologies and technology-enhanced education. According to reference [9], the FSLSM model is most appropriate for multimedia courseware and online-teaching. Reference [10] confirmed this by conducting a comparison of learning models with respect to the application in Web-based learning systems. The result of their research confirmed that the use of FSLSM is the most appropriate model for technology-enhanced education environments. There are four dimensions in FSLSM such as Perception, Input, Information Processing and Understanding. Each learner is characterized by a specific preference for each of these dimensions.

These dimensions are based on major dimensions in the field of learning styles and can be viewed independently from each other. They show how learners prefer to process (active/reflective), perceive (sensing/intuitive), receive (verbal/visual), and understand (sequential/global) information. While these dimensions are not new in the field of learning styles, the way in which they describe a learning style of a student can be seen as new and innovative. While most learning style models, which include two or more dimensions, derived statistically prevalent learner types from these dimensions such as models by Myers-Briggs [11], Gregorc [12], Kolb [13], and Honey and Mumford [14].

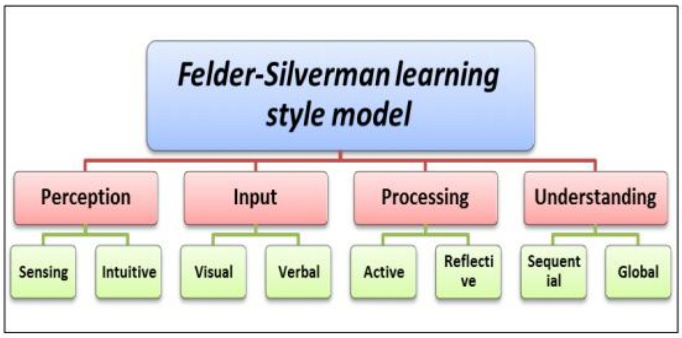

Fig -1: Felder-Silverman Learning Style Model

The *active/reflective* dimension is analogous to the respective dimension in Kolb's model [15]. Active learners learn best by working actively with the learning material, by applying material, and by trying things out. Furthermore, they tend to be more interested in communicating with others and preferred to learn by working in groups where they can discuss about the learned material. In contrast, reflective learners prefer to think about and reflect on the material. In contrast, reflective learners prefer to think about and reflect on the material. Regarding communication, they prefer to work alone.

The *sensing/intuitive* dimension is taken from the Myers-Briggs Type Indicator [11] and has also similarities to the sensing/intuitive dimension in Kolb's model [13]. Learners with sensing learning styles prefer to learn facts and concrete materials, using their sensory experiences of particular instances as a primary source. They like to solve problems with standard approaches and also tend to be more patient with details. They tend to be more practical than intuitive learners and like to relate the learned material to the real world. In contrast, intuitive learners prefer to learn abstract learning material, such as theories and their underlying meanings, with general principles rather than concrete instances being a preferred source of information.

The third, *visual/verbal* dimension deals with the preferred input mode. The dimension differentiates learners who remember best what they have seen (e.g. pictures, diagrams), from learners who get more out of textual or text-based representation, regardless of the fact whether they are written or spoken.

In the fourth dimension, learners are distinguished between *sequential and global* way of understanding. This dimension is based on the learning style model by Pask [16], where sequential learners refer to serial learners and global learners refer to holistic learners. Sequential learners learn in small incremental step and therefore have a linear learning progress. They tend to follow logical stepwise paths in finding solutions. In contrast, global learners use a holistic thinking process and learn in large leaps. They tend to absorb learning material almost randomly without seeing connections but after they have learned enough material they suddenly get the whole picture. Because the whole picture is important for global learners, they tend to be more interested in overviews and in a broad knowledge, whereas sequential learners are more interested in details.

3. METHODOLOGY

3.1 Context, Participants and Data Source

The study is based on data from Computer Programming 1 course which is taught during the first semester for Computer Technology course in Southern Luzon State University. Aside from traditional classroom setup, it is accompanied by a supplementary Moodle [17] course that is composed of eight chapters that includes learning objects ranging from textual, visual, concrete and abstract learning materials. There are also different exercises that allow students to practice programming skills. Self-Assessment tests were also provided for each chapter overall. Students also were encouraged to use the forums in order to interact and solve problems with other students during the course. This particular course was selected for the investigation of individual learning styles for it is found to have large number of enrolled students in the Moodle course and the structure of the course is most appropriate for the selected learning style model.

The study used the acquisition of data coming from the virtual learning environment database. Specifically, the data from student's logs and activities on the virtual learning environment were carefully examined. The student's learning styles are obtained by using the Index of Learning Styles (ILS) questionnaires that are answered by the students who are enrolled and completed the selected course. An initial data for 108 students out of the possible 547 students who completed the Computer Programming Course were collected. These sets of students are also enrolled with the corresponding Moodle course during those periods.

3.2 Feature Extraction

Table 1 provides the list of learning styles mapping of relevant student's behavior on virtual learning environments. The features were mapped according from the Felder-Silverman Learning Style Model [8].

Table-1: Learning style mapping of relevant student's behavior

Learning Object	Relevant Behavior	Learning Style	Dimension
Forum	post/reply	active	Processing
	view/read	reflective	
Self-Assessments	attempt	active	
Text-based	view/visit	reflective	
Concrete materials	view/visit	sensor	Perception
Abstract materials	view/visit	intuitive	
Examples	view/visit	sensor	
Exercises	revisions	intuitive	
Visual materials	view/visit	visual	Input
Text-based	view/visit	verbal	
Video materials	view/visit	visual	
Forum	post/reply	verbal	
Course overviews	view/visit	global	Understanding
Detailed Activities	view/visit	sequential	
Navigation	navigate linearly	sequential	
	navigate globally	global	

3.3 Data Preparation and Analysis Procedure

Figure 2 illustrates the data preparation and analysis procedure of the study. Every student logs and activities are recorded in the virtual learning environment databases. Primarily, a search query (SQL) was conducted to retrieve a variety of data from the VLE, and then log records are saved. The data processing exploration and analysis process included the following: the data preprocessing phase, the data mining phase, and the pattern analysis phase. The data preprocessing phase was performed by reducing the log file, which was cleaned by removing all useless, irregular, and missing data from the original VLE common log files. Feature extractions filtered out was based on the related characteristics of the Felder-Silverman Learning Style Model. Derived variables was extracted through calculating or accumulating variable data such as number of views, number of visits, number of posts, and number of exercises attempts just to name a few. In addition, another field was created in order to accommodate the results of each dimension for the Index of Learning Style questionnaire answered by participating students in the study. These questionnaires identified each preferences of each student when it comes to their learning styles. These variables together with the feature extracted will be transformed into fields, assigned with proper data attributes, and stored in a file.

Classification algorithm techniques were applied to build a classification model of the student's behavior in Virtual Learning Environment. The behavior pattern analysis phase includes data interpretation and evaluation of the results. This phase was needed in order to identify meaningful results from outcomes of the data mining phase. An open-source data mining software package such as the Waikato Environment for Knowledge Analysis (WEKA) [18] was used to perform data analysis on the derived datasets to uncover the most accurate classification model that will be used for the future development of a software prototype. The learning algorithms implemented in WEKA are Bayes network classifier [19] (Naïve Bayes, BayesNet) and classification tree with pruning algorithms [20] (J48 and NBTree). A 10-fold cross validation was used to evaluate the classifiers.

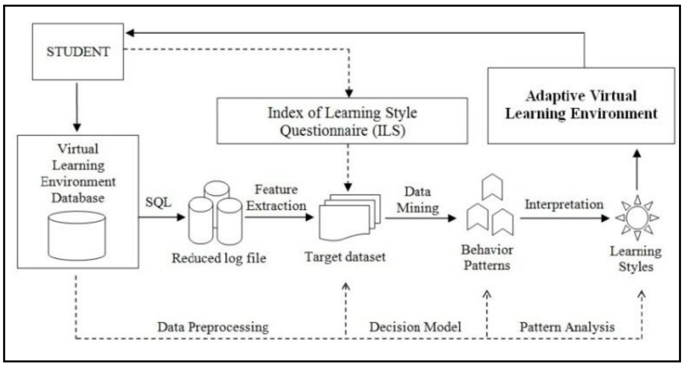

Fig-2: Data preparation and analysis procedures

4. RESULTS AND DISCUSSIONS

To empirically investigate the performance of the classifiers on the extracted data sets, classification algorithms such as Bayes classifier and Decision Tree with pruning classifier are selected. A 10-fold cross validation for every classifier was used. Classifications are tested on processing dimension, perception dimension and input of Felder-Silverman Learning Style Model. The understanding dimension is left out for the reason that there are no available data sets yet to identify user's behavior in terms of their navigational patterns in the Virtual Learning Environment. The results of the tests are shown in Table 2 and are summarized based on correctly and incorrectly classified instances and Kappa statistics.

Based from Table 2, Table 3, Table 4 and Chart 1, we can infer that method of classification tree algorithms with pruning obtained the highest accuracy with an average of 88.59% collectively for all dimension. Comparing both classification tree classifiers, J48 attained the highest average accuracy of 89.81% and NBTree yields 87.37%. The Bayes algorithm, Naïve Bayes and BayesNet yields an average accuracy of around 79.84% and 78.51% respectively. Kappa statistics is used to assess the accuracy of any particular measuring cases, it is usual to distinguish between the reliability of the data collected and their validity. Average Kappa score derived from the algorithms obtained a value of 0.7 to 0.9 which shows that the accuracy of the classification is substantial [21].

Table-2: Classification results for processing dimension

Method	Algorithm	PROCESSING DIMENSION		
		Correctly Classified Instances (%)	Incorrectly Classified Instances (%)	Kappa Statistics
bayes	Naïve Bayes	78.35	21.65	0.753
	BayesNet	77.81	22.19	0.711
tree	J48	**87.47**	12.53	0.815
	NBTree	85.18	14.82	0.704

Table-3: Classification results for perception dimension

Method	Algorithm	PERCEPTION DIMENSION		
		Correctly Classified Instances (%)	Incorrectly Classified Instances (%)	Kappa Statistics
bayes	Naïve Bayes	79.39	20.61	0.768
	BayesNet	78.05	21.95	0.725
tree	J48	**91.25**	8.75	0.908
	NBTree	88.58	11.42	0.891

Table-4: Classification results for input dimension

Method	Algorithm	INPUT DIMENSION		
		Correctly Classified Instances (%)	Incorrectly Classified Instances (%)	Kappa Statistics
bayes	Naïve Bayes	81.78	18.22	0.795
	BayesNet	79.69	20.31	0.713
tree	J48	**90.73**	9.27	0.902
	NBTree	88.36	11.64	0.872

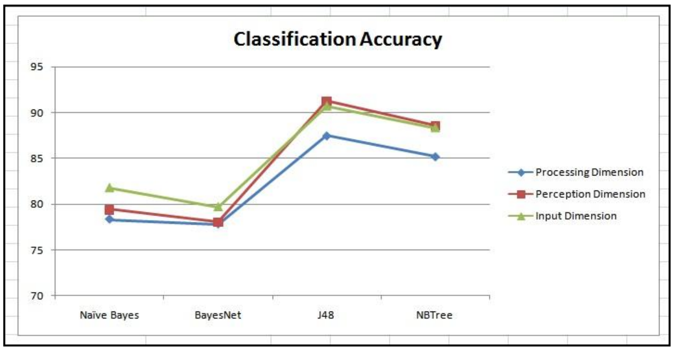

Chart-1: Classification accuracy of selected classifiers

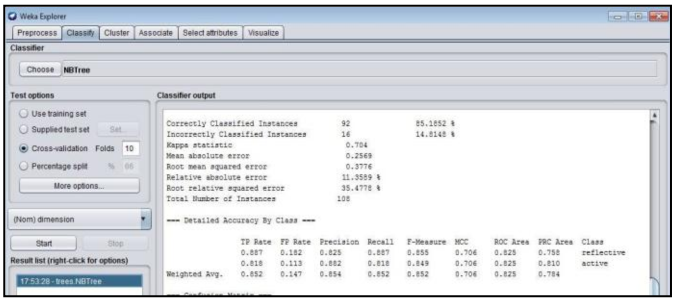

Fig-3: NBTree classifier output for processing dimension

5. CONCLUSIONS AND FUTURE RESEARCH

This paper is part of an initial stage of the study and is still an on-going research that involves detection of learning styles that classifies student based from their behavior on a Moodle course according to Felder-Silverman Learning Style Model. The selected model is implemented on partial data sets of 108 students enrolled in Computer Programming 1 course in Southern Luzon State University at Lucban, Quezon, Philippines. The results show that the efficiency of classification by means of J48 algorithm had the highest average accuracy in terms of correctly classified instances at 89.81%. In current popular Virtual Learning Environments, no functions or features are currently available to automatically identify student's individual learning styles that are based from their relative behaviors. This study can be a basis for educators that students have varied behavior and learning styles. Moreover, this study gives hints to educators to design appropriate course contents that matches the student's learning styles to optimize the learning process.

For future work, the researcher will propose to extend the capability of Virtual Learning Environment to adapt its course content and design to match the learning style of each student to respond immediately to their needs based from the model. Furthermore, the researcher would plan on methods on capturing student's navigation behavior patterns in a learning system so that all learning dimensions can be included in the process. Also, experimentally apply the adaptive system to test the relationship between learning styles and academic performance.

REFERENCES

[1] Dumciene, A., Lapeniene, D., Possibilities of Developing Study Motivation in E-Learning Products. Electronics and Engineering. – Kaunas: Technolojia, 2010. – No. (102) pp. 43-46, 2010

[2] Ballera, Melvin A., and Elssaedi, Mosbah Mohamed, New E-learning Strategy Paradigm: A Multi-Disciplinary Approach to Enhance Learning Delivery, Proceedings of the 2nd E-learning Regional Conference, State of Kuwait, pp. 25-27, 2013

[3] Liu, F., Kuljis, J., A Comparison of Learning Style Theories on the Suitability for E-Learning. In M.H. Hamza (Ed.), Proceedings of the IASTED Conference on Web Technologies, Applications and Services, ACTA Press, pp. 191-197, 2005

[4] Keefe, J.W., Learning Style: An Overview. In National Association of Secondary School Principlas (Ed.), Student Learning Styles: Diagnosing and Prescribing Programs, 1979; pp. 1-17

[5] Stewart, K.L., Felicettie, L.A., Learning Styles of Marketing Majors. Educational Research Quarterly, 15(2), pp. 15-23, 1992

[6] Pashler, H., McDaniel, M. Rohrer, D, Bjork, R., Learning Styles: Concepts and Evidence. Psychological Science in the Public Interest, vol. 9, pp. 105-119, 2008

[7] Felder, R.M., and B.A. Soloman, Index of Learning Styles, http://www.ncsu.edu/felder-public/ILSpage.html, 2004, accessed May 2, 2016.

[8] Felder, R.M., and Silverman, L.K., Learning Styles and Teaching Styles in Engineering Education. Presented at the 1987 Annual Meeting of the American Institute of Chemical Engineers, New York, Nov. 1987

[9] Carver, C.A., Howard, R.A., and Lane, W.D., Addressing Different Learning Styles through Course Hypermedia. IEEE Transactions on Education, vol. 42, no.1, pp. 33-38, 1999

[10] Liu, F., Kuljis, J., A Comparison of Learning Style Theories on the Suitability for E-Learning. In M.H. Hamza (Ed.), Proceedings of the IASTED Conference

on Web Technologies, Applications and Services, ACTA Press, pp. 191-197, 2005

[11] Myers, I.B., and McCaulley, M.H., Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator. Consulting Psychologists Press, Palo Alto, CA., 1985

[12] Gregorc, A.F., Style Delineator: A Self-Assessment Instrument for Adults. Gregorc Associates Inc. Columbia, 1985

[13] Kolb, D.A., Experiential Learning: Experiences as the Source of Learning and Development. Prentice-Hall, Englewood Cliffs, New Jersey, 1984

[14] Honey, P., and Mumford, A., The Learning Styles Helper's Guide. Peter Honey Publications Ltd., Maidenhead, 1982

[15] Kolb, D.A., Experiential Learning: Experiences as the Source of Learning and Development. Prentice-Hall, Englewood Cliffs, New Jersey, 1984

[16] Pask, G., A Fresh Look at Cognition and the Individual. International Journal of Man Machine Studies, vol. 4, pp. 211-216, 1972

[17] Moodle Learning Management System. [Online]. Available: http://www.moodle.org.

[18] WEKA at http://www.cs.waikato.ac.nz/~ml/weka. Retrieved April 29, 2016.

[19] Garcia P., Amandi A., Schiaffin S., Campo M., Evaluating Bayesian Networks Precision for Detecting Students' Learning Styles. Computers and Education, 49, pp.794-808, Elsevier, 2007

[20] Cha, H.J., Kim, Y.S., Lee, J.H., and Yoon, T.B. Learning Style Diagnosis Based on Interface Behaviors. Workshop Proceedings of International Conference on E-Learning and Games, Hangzhou, China, April 17-19, pp. 513-524, 2006

[21] Cohen J. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, 20, pp. 37-46, 1960