2017 9th International Conference on Knowledge and Systems Engineering(KSE)

Predicting Students' performance based on Learning Style by using Artificial Neural Networks

Hoang Tieu Binh* and Bui The Duy**

*Hanoi National University of Education
**University of Engineering and Technology - Hanoi National University

ABSTRACT

In recent years, many education researchers concentrated on learning styles and its implications. They relied that people who have different personality types tend to have different learning styles, which in turn influence to students' performance in each type of subjects. To measure the relationship between learning styles and student performance in a subject or entire course, we conducted an online survey with a participation of students in various courses to analysis and show the effects of different learning styles on students' performance. This test based on the Felder-Soloman questionnaire with 44 test questions divided into four dimensions. We have filtered data to select 316 undergraduate students for the analysis. According questionnaire results, we have applied statistical methods to analyze data and issued some important notes. Besides these, we have built an artificial neutral network to predict academic performance based on students' learning style. We have shown some significant relationships between learning style and performance of the testers, which have implications for the studying of a course.

Keywords: Learning Styles, Educational Data Mining, Artificial Neural Networks.

I. INTRODUCTION

Learning styles nowadays play an important role in the whole educated progress of a person. Learning styles refers to the preferential way in which the student absorbs, processes, comprehends and retains information. Many researchers showed that the matching of teaching and learning styles will lead very different results between students, especially in online education environment. Richard M.Felder in [9] showed that mismatches exist between learning styles of students and teaching styles of professors can cause students become bored and inattentive in class, do poorly on tests, get discouraged about the courses...In some cases, students tend to drop out class or make a poor attendance.

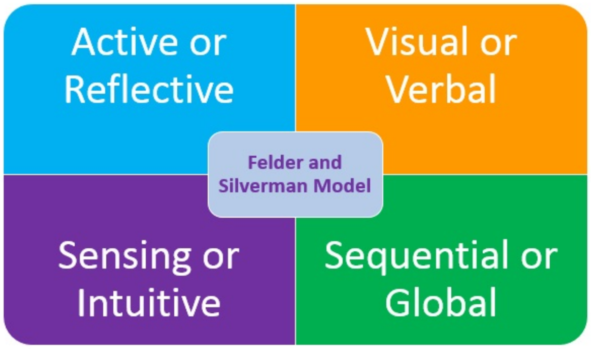
In online learning courses or e-learning systems, discovering the attitude of student is helpful for systems to adjust the content or method of teaching. Adaptive learning system nowadays can be an alternative to the tradition "one-size-fits-all". Based on learning style survey, we can find out the favorite learning and teaching models for both students and teachers. According to the results, we can suggest the best way for them in organizing and pursuing a online/offline course.

This research has conducted a survey to collect data from more than three hundred students and analyses data based on statistical methods, the result was post-processed by a machine learning method to categorize the accumulated study results combining with learning styles. We also proposed a method of combination data mining methods and learning style data for predicting students performance. These might be a good suggestion for many intelligent tutoring systems or testing systems for supporting adaptation ability.

II. LEARNING STYLES

Among many learning style models in intelligent tutoring systems, Felder Silverman Learning Styles Model (FSLSM) [9] is the most popular model for adaptive learning. This model divides learning styles into four dimensions and each dimension follows two opposite directions: active/reflective, sensing/intuitive, visual/verbal, and sequential/global.

- **Active/Reflective**: Active learners like to interact with object, do the experiments, and study by try to do somethings, by discussing or explaining it to others. They want to join into a group to solve problems/ Reflective learners prefer to think first and



Felder and Silverman model

do nothing physical, they study by analyzing. They want to solve everything by themselves.

- **Sensing/Intuitive**: Sensing learners interesting in detailed information, they learn by facts and good at memorizing facts and doing hands-on (laboratory) work/Intuitive learners like concepts, creations, and theories. They want to find the meaning and are more innovative than sensors.
- **Visual/Verbal**: Visual learners like diagrams, images and graphs. They remember best what they have seen/Verbal learner want to read or listen to information. They find solution by verbal.
- **Sequential/Global**: Sequential learner like information is expressed in linear steps. They sort all the detailed events to show the whole picture./Global learners want a method with a systematic and globalized approaching. They firstly tend to see the global picture and then fill with details.
- **Balanced learning styles**: When you know what is your leaning styles, you can extend and develop your method to make it more balanced, it could help you improve your performance in studying and make it more efficient. The focus of this model is the balance, whereby you should not push your learning preferences to any poles, because it can narrow your ability to acquire knowledge and understand information quickly, accurately and efficiently.

Many researches related to learning style models such as Felder-Silverman, Myers-Briggs, Honey and Mumford, Kolb's model... are now developing parallely. Myers-Briggs model with the Myers-Briggs Type Indicator (MBTI) is based on Blooms Taxonomy, The MBTI measures preferences on four dimensions as Energizing, Attending, Deciding and Living [1]. Honey and Mumford [13] which identified four distinct learning styles or preferences: Activist, Theorist, Pragmatist and Reflector. David Kolb published his learning styles model in 1984 from which he developed his learning style inventory included 4 cycle stages: Concrete Experience, Reflective Observation, Abstract Conceptualization and Active Experimentation [15]. Several studies showed that the ILS questionnaire provides a very precise quantitative estimation of a learners preference for each dimension of FSLSM [19].

In our studies, the Felder-Silverman learning style was chosen for the survey thus this theory occupies about 70.6% in adaptive learning systems, much greater than Honey & Mumford and Kolb model which were used totaly 7.8% in intelligent tutoring systems [23]. Beside these, Felder-Silverman model was tested widely and was illustrated that it is important for relating the learning style model with the features of online environment [12], even it is considered as one of the best model to use in adaptive systems [16].

III. Educational Data Mining

In recent years, the term Educational Data Mining is widely used, following the website http://www.educationaldatamining.org/ [18], Educational Data Mining (EDM) is an emerging discipline, concerned with developing methods for exploring the unique and increasingly large-scale data that come from educational settings, and using those methods to better understand students, and the settings which they learn in. EDM is growing reflected in the increasing number of contributions published every year, and the number of specific tools specially developed for applying data mining algorithms in educational data/environments. The full integration of data mining in the educational environment will become a reality, and fully operative implementations could be made available not only for researchers and developers but also for external users [21].

IV. Related work

Learning styles and related issues are widely researched in recent years. In some adaptive systems, learning styles are detected automatically based on tracking behavior of learners. Crockett et al. in 2013 [4] proposed a model predicting learning styles using fuzzy rule base which used tutorial data with applying genetic algorithm. Sabine Graf have [11] detected learner styles based on the behavior of learners during an online course and calculated and filled student model of a Learning Management System (LMS). Silvia et al [22] have used data driven method for detecting characteristics of learners. Crockett in [4] and [5] used a fuzzy model to predict learning styles in Conventional Conversational Intelligent Tutoring

Systems. Other researchers concentrated on discovering the relationship between learning styles and academic performance, such as Bruce Brunton [2] and Daniel Alan Seiver [6] (both of them used Kolb Learning Style Inventory [15] and Myers-Briggs Type Indicator [1] respectively instead). The first author used data from nine introductory microeconomics classes to test the effect of student learning style on academic performance while the second measured the impact of learning styles on student performance in the introductory finance course in the university . In fact, not many researchers predict learning performance base on learning styles. This why we try to find out the relationship between students' performance and learning style by using some advanced techniques of data mining.

V. Experiment

We have conducted a survey with The Index of Learning Styles (ILS) [17] in Vietnamese language located in http://daotao.hnue.edu.vn/questionnaire/ to access students' learning style. The objects of the survey were students in the second and third year at Hanoi National University of Education. The questionnaire consists of 44 two-choice questions distributed along the four learning style dimensions. Each dimension includes 11 quesions which measure separately one preference. The questions of Active/Reflective dimension are [1,5,9,13,17,21,25,29,33,37,41] which collect information about the processing of learners, the Sensing/Intuitive preference includes the questions [2,6,10,14,18,22,26,30,34,38,42] which measure students' perception, the Visual/Verbal is a scale for input information which are series of questions containing [3,7,11,15,19,23,27,31,35,39,43] and the last dimension Sequential/Global with [4,8,12,16,20,24,28,32,36,40,44] questions get understanding ways of learners. The result of each question returned an integer value -1 or 1. With 11 questions for one group, we have got a value ranging [-11,11] corresponding to each dimension. The number columns which are calculated as shown in Table VII by subtractions for each pair of dimensions, for example AR column is gained by subtracting Reflective from Active column and so on. The test was implemented with the joining of more than seven hundred students, however the result was chosen from these students to elect 316 from two main majors: Natural Science and Social Science, including 256 female students and 60 male students. These students are selected from three recent courses of six main specialities: Mathematics, Physics, Chemistry, Philology, History and Geography with represent almost students of the university. Marks of all students are also extracted from Credit-Based Learning Management System to engage to learning style. All of this data then was trained by an artificial neural network to find the relation between learning style and learning performance. In our study, we propose a method of combining an effective machine learning algorithm to mine data for detecting the relationship between learning style and learning performance.

The data was used to train includes all of 316 records with known as training examples. The extract of data is shown in Table VII. The column [Avg4] is the results of accumulate mark of students till the survey time. These values were discretized to the nominal attributes for the classify process. We divided it into 4 ranges: [2-2.4], [2.5-3.1], [3.2-3.5], [over 3.5] correspond to ”Average”, ”Good”, ”Very Good”, ”Excellent” respectively. Some columns were eliminated before running algorithm for the dependency.

We used an artificial neural networks (ANN) with a 3-layer perceptron for predicting academic performance based on learning styles. The parameters of our ANN were optimized through experimentation using suggested values/ranges from many trials. As the results the 10 folds cross validation was chosen when performed training process to ensure generalization to the independent datasets. The groups of subject and gender are included to find out any relationship to students' achievement.

VI. Statistics analysis of the Index of Learning Styles

A. Reliability

We consider the reliability of test result by calculating Cronbach's alpha which is a measure of internal consistency. It is used to measure scale reliability of the items which are grouped together within an instrument to form scales. Coefficient alpha is the most efficient measure of reliability in terms of the classical model of error measurement, and in ideal circumstances [14]. However, it is difficult to do statistical analizing if we do not convert the student responses to [0,1], because of the bipolar nature of the scales in the ILS. [24]. To calculate this coefficient, we had to convert the result of the dichotomous questions to 1 if the answer is 'a' and value 0 if the answer is 'b' and vice versa.

Cronbach's alpha can be expressed as follows [3]:

$$\alpha = \frac{N}{N-1}\left(1 - \frac{\sum_{i=1}^{N}\sigma^2(Y_i)}{\sigma_X^2}\right)$$

where σ_X^2 is the variance of the observed total test scores, and $\sigma^2(Y_i)$ is the variance of item i for the current sample of students. $X = Y_1 + Y_2 + \cdots + Y_N, N$ is equal to the number of items.

The code was implemented by Python version 3.5:

```
def CronbachAlpha(data):
   itemvars = np.var(data,axis=0, ddof=1)
   tscores = np.sum(data,axis=1)
   N = len(data[0])
   return N / (N-1)
         * (1 - np.sum(itemvars)
         / np.var(tscores,ddof=1))

a = CronbachAlpha(ILSdata)
```

The Cronbach's alpha results are shown in Table I:

	α coefficients
Active-Reflective	0.343
Sensing-Intuitive	0.521
Visual-Verbal	0.507
Sequential-Global	0.276

TABLE I: Internal reliability of ILS scales - Cronbach Alpha

These coefficients α of Sensing-Intuitive and Visual-Verbal dimensions are the same with others reliability scales tested before, however the results of Active-Reflective and Sequential-Global scales are lower than the others tests of Felder in [8] or Felkel et al. [10] or Zywno in [25].

B. Correlation

Correlation is expressed for the strength of the linear relationship between two variables [7]

Variables	**AR**	**SI**	**VV**	**SG**
AR	1	0.110	0.236	-0.070
SI	0.110	1	0.052	0.139
VV	0.236	0.052	1	-0.068
SG	-0.070	0.139	-0.068	1

TABLE II: Correlation matrix (Pearson)

The inter-correlations between the scales on the ILS are shown in Table II. Only the Active-Reflective and Visual-Verbal (AR-VV) scales show any significant correlations with each other. The two others appear not to be related either to each other or to these scales. This supports the suggestion that at least three of the four ILS scales generally seem to be orthogonal.

C. Statistics Analysis

First, we analyzed the distribution of learning style preferences for each dimension. According to Table III, 64.6% of student in our survey were found to have an Active preference, 87.7% a Sensing preference, 75.6% a Visual preference, and 51.9% a Global preference. At this level of information analysis, it is simple to realize that most of students (87.7%) trend to study by facts and do experiments.

Act	Ref	Sen	Int	Vis	Ver	Seq	Glo
112	204	277	39	239	77	152	164
35.4%	64.6%	87.7%	12.3%	75.6%	24.4%	48.1%	51.9%

TABLE III: Learning style preferences

Descriptive statistics for the ILS questionnaire are shown in Table IV. As the data displayed, the standard deviation between variables is similar and approximately equal to 2. These mean that the learning styles of every students are not absolutely different.

Statistic of students' response		
	Mean	Standard deviation
Active/Reflective	-1.203	3.843
Sensing/Intuitive	4.595	3.8
Visual/Verbal	2.873	4.166
Sequential/Global	-0.196	3.7
Active	4.899	1.921
Reflective	6.101	1.921
Sensing	7.797	1.9
Intuitive	3.203	1.9
Visual	6.937	2.083
Verbal	4.063	2.083
Sequential	5.402	1.85
Global	5.598	1.85

TABLE IV: Statistic results

Looking into Table V we can see the balance between dimensions except Sensing/Intuitive. This can be explained that students trend to learn by facts and do their exercises or work in a laboratory. These can help university's administrators in developing experiment centers or practical equipment.

With more detail, we want to know how strong of each preference of students. We built a table with 3 levels for each dimension based on the strength of preferences. Based on original survey data, we defined how strong preference of each student for one side of dimension. If he/she got a score between 5-11 (either plus or minus), he/she had a strong preference, and if the score was between -3 $\rightarrow$ 3 (data is shown in Table VII), he/she had a moderate or balanced preference. This statistic is shown in Table V. Concretely, we concentrate on Table VI, in which we separate data into two groups: Nature Science and Social Science. From this point of view, we realize the balance in almost dimensions, especially for the Social Science group and there is no significant difference between these groups.

Active/Reflective			Sensing/Intuitive			Visual/Verbal			Sequential/Global		
strong Active	mod	strong Reflective	strong Sensing	mod	strong Intuitive	strong Visual	mod	strong Verbal	strong Sequential	mode	strong Global
33 10.4%	206 65.2%	77 24.4%	202 63.9%	103 32.6%	11 3.5%	137 43.4%	156 49.4%	23 7.3%	42 13.3%	217 68.7%	57 18.0%

TABLE V: Strength of preferences

	Active/Reflective			Sensing/Intuitive			Visual/Verbal			Sequential/Global		
	strong Active	mod	strong Reflective	strong Sensing	mod	strong Intuitive	strong Visual	mod	strong Verbal	strong Sequential	mode	strong Global
Nature Science	22 11.4%	119 61.7%	52 26.9%	115 59.6%	69 35.8%	9 4.7%	85 44.0%	96 49.7%	12 6.2%	29 15.0%	132 68.4%	32 16.6%
Social Science	11 8.9%	87 70.7%	25 20.3%	87 70.7%	34 27.6%	2 1.6%	52 42.3%	60 48.8%	11 8.9%	13 10.6%	85 69.1%	25 20.3%

TABLE VI: Strength preferences group by Social/Natural Science

SID	Group	sex	AR	SI	VV	SG	A	R	Sen	I	Vis	Ver	Seq	G	Avg4
635101004	Nat	F	-3	5	1	-3	4	7	8	3	6	5	4	7	3.4
635101101	Nat	F	-1	5	1	3	5	6	8	3	6	5	7	4	2.9
635101109	Nat	M	1	-1	-3	-5	6	5	5	6	4	7	3	8	3.16
635101161	Nat	M	-3	7	11	-5	4	7	9	2	11	0	3	8	2.71
635103047	Nat	F	-5	9	-1	3	3	8	10	1	5	6	7	4	3.58
635602066	Soc	M	-7	-3	-5	-1	2	9	4	7	3	8	5	6	3.673
635603110	Soc	F	-1	5	11	1	5	6	8	3	11	0	6	5	3.27
645602068	Soc	F	-3	3	-7	-5	4	7	7	4	2	9	3	8	3.23
655101045	Nat	M	1	9	-1	1	6	5	10	1	5	6	6	5	3.43
655101055	Nat	M	-3	3	3	1	4	7	7	4	7	4	6	5	3.73
655101061	Nat	F	5	1	5	5	8	3	6	5	8	3	8	3	3.75
655101063	Nat	F	-3	5	5	1	4	7	8	3	8	3	6	5	3.27
655101072	Nat	F	-5	3	3	-1	3	8	7	4	7	4	5	6	2.88
655101083	Nat	M	-3	7	-1	3	4	7	9	2	5	6	7	4	3.62
655101103	Nat	F	-5	9	9	1	3	8	10	1	10	1	6	5	3.61
655101105	Nat	M	-7	-11	5	-5	2	9	0	11	8	3	3	8	3.65
655101128	Nat	F	-3	5	-1	-3	4	7	8	3	5	6	4	7	3.8
655601056	Soc	F	3	7	7	-3	7	4	9	2	9	2	4	7	3
655601140	Soc	F	1	9	9	-5	6	5	10	1	10	1	3	8	3.03
655602010	Soc	F	-3	-1	5	1	4	7	5	6	8	3	6	5	3.64

TABLE VII: The first 20 students of survey data

VII. TRAINING RESULTS

We use a Multilayer Perception to train current data. A network with a = attributes + classes hidden layers of sigmoid nodes, a learning rate of 0.3, momentum of 0.2 and 500 learning cycles were used. The training results are shown in Table VIII. We did not use any normalizing methods for estimating because the data was verified before. We trained the network with varies of input parameters for optimization. We also divided data into 10 folds with 90% for training set and 10% for test set.

Classify	Number	Accuracy
Correctly Classified Instances	254	80.63%
Incorrectly Classified Instances	62	19.37%

TABLE VIII: Training Results

	Very Good	Excellent	Good	Average
Very Good	61	20	6	0
Excellent	12	48	0	0
Good	2	0	136	10
Average	0	0	12	9

TABLE IX: Confusion Matrix

We can observe in the Table VIII that the accuracy of estimation is approximate 80.63% with 254 students are correctly classified. This is a significant improvement comparing to 74% of the research of V.O. Oladokun et all in [20]. Conversely, incorrect classified instances are 62, occupied 19.37%.

The confusion matrix as shown in Table IX shows the performance of the classification model. As it shown, the best percentage rate of accuracy is [Good] level with

136/148, the following is [Excellent] with 48/60 and [Very Good], [Average] with 61/87, 9/21 respectively.

VIII. Conclusion and future work

In this paper, we have concretely described an approach on multilayer perceptron to predict academic result based on learning styles. As future work, we make a plan to expand our research by increasing the number of students in survey (increasing training examples may increase accuracy rate in ANN) and by deeply analyzing with variety of data mining algorithms. Beside this, we are able to conduct the test with other popular learning style such as MBTI, Kolb... for comparison. The results from these researches are also applied in e-learning environment to build adaptive model for supporting learners.

Acknowledgments

This work has received support from the project "Multimedia application tools for intangible cultural heritage conservation and promotion - DTDL.CN-34/16" funded by Ministry of Science and Technology of the Socialist Republic of Vietnam

References

[1] Catherine Bishop-Clark and Daniel D Wheeler. The Myers-Briggs personality type and its relationship to computer programming. *Journal of Research on Computing in Education*, 26(3):358, 1994.

[2] Bruce Brunton. Learning Styles and Student Performance in Introductory Economics. *Journal of Education for Business*, 90(2):89–95, 2014.

[3] Edward G Carmines and Richard A Zeller. *Reliability and validity assessment*, volume 17. Sage Publications, Inc, 1979.

[4] Keeley Crockett, Annabel Latham, David Mclean, and James O'Shea. A fuzzy model for predicting learning styles using behavioral cues in an conversational intelligent tutoring system. *IEEE International Conference on Fuzzy Systems*, 2013.

[5] Keeley Crockett, Annabel Latham, and Nicola Whitton. On Predicting Learning Styles in Conversational Intelligent Tutoring Systems using Fuzzy Decision Trees. *International Journal of Human-Computer Studies*, 97:98–115, 2016.

[6] Andrew Do Daniel Alan Seiver, Kamal Haddad. Student learning styles and performance in an introductory finance class. *American Journal of Business Education*, 7(3):183–191, 2014.

[7] S.M. Dowdy, S. Wearden, and D.M. Chilko. *Statistics for Research*. Wiley series in probability and statistics. Wiley-Interscience, 2004.

[8] Richard M Felder and Joni Spurlin. Applications, Reliability and Validity of the Index of Learning Styles. *International Journal of Engineering Education*, 21:103 – 112, 2005.

[9] RM Felder and LK Silverman. Learning and teaching styles in engineering education. *Engineering education*, 78(June):674–681, 1988.

[10] Brian H. Felkel and Ross M. Gosky. A Study of Reliability and Validity of the Felder-soloman Index of Learning Styles for Business Students. In *Proceedings of the Annual International Conference on Technology in Collegiate Mathematics*, 2013.

[11] Sabine Graf, P. Kinshuk, and Kinshuk. An Approach for Detecting Learning Styles in Learning Management Systems. *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*, pages 161–163, 2006.

[12] Sabine Graf, Silvia Rita Viola, Tommaso Leo, and Kinshuk. In-Depth Analysis of the Felder-Silverman Learning Style Dimensions. *Journal of Research on Technology in Education*, 40:79–93, 2007.

[13] Peter Honey and Alan Mumford. *Using your learning styles*. Maidenhead: Honey, 1986.

[14] Paul Kline. *A Handbook of Test Construction*. Taylor & Francis Ltd, 2015.

[15] D.A. Kolb. *Experiential Learning: Experience as the Source of Learning and Development*. Prentice-Hall, 1984.

[16] Jasna Kuljis and Fang Liu. A Comparison of Learning Style Theories on the Suitability for elearning. In *Web Technologies, Applications, and Services*, 2005.

[17] Richard M.Felder. Index of Learning Styles Questionnaire. http://www.engr.ncsu.edu/learningstyles/ilsweb.html. Accessed: 2016-June-10.

[18] Educational Data Mining. International Educational Data Mining Society. http://www.educationaldatamining.org/. Accessed: 2017-Apr-19.

[19] Jelena Nakić, Sabine Graf, and Andrina Granić. Exploring the adaptation to learning styles: The case of AdaptiveLesson module for moodle. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lec ture Notes in Bioinformatics)*, 7946 LNCS:534–550, 2013.

[20] V O Oladokun, D Ph, A T Adebanjo, B Sc, and D Ph. Predicting Students' Academic Performance using Artificial Neural Network: A Case Study of an Engineering Course . *The Pacific Journal of Science and Technology*, 9(1):72–79, 2008.

[21] Cristóbal Romero. Educational Data Mining: A Review of the State-of-the-Art. *Transactions on Systems, Man, and Cybernetics*, XX(X):1–19, 2016.

[22] Kinshuk Silvia Rita Viola, Sabine Graf and Tommaso Leo. Analysis of Felder - Silverman Index of Learning Styles by a data-driven Statistical approach. In *Proceedings of the Eighth IEEE International Symposium on Multimedia (ISM'06)*, pages 1–24, 2006.

[23] Huong May Truong. Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities. *Computers in Human Behavior*, 55:1185–1193, 2015.

[24] N. Van Zwanenberg, L. J. Wilkinson, and A. Anderson. Felder and Silverman's Index of Learning Styles and Honey and Mumford's Learning Styles Questionnaire How do they compare and do they predict academic performance? *Educational Psychology*, 20(3):365–380, 2000.

[25] Malgorzata S Zywno. A Contribution to Validation of Score Meaning for Felder- Soloman ' s Index of Learning Styles. *Engineering Education*, pages 1–16, 2003.