ISBN: 978-989-8533-40-1 © 2015

BEHAVIORAL FEATURE EXTRACTION TO DETERMINE LEARNING STYLES IN E-LEARNING ENVIRONMENTS

Somayeh Fatahi[1, 2], Hadi Moradi[1, 3] and Elaheh Farmad[1]
[1]*School of Electrical & Computer Engineering, University of Tehran, Tehran, Iran*
[2]*Department of Computer Science Department, Halifax, Canada*
[3]*Intelligent Systems Research Institute, SKKU, South Korea*

ABSTRACT

Learning Style (LS) is an important parameter in the learning process. Therefore, learning styles should be considered in the design, development, and implementation of e-learning environments. Consequently, an important capability of an e-learning system could be the automatic determination of a student's learning style. In this paper, a set of features which are important in extracting the learning style automatically from students' behavior has been determined. These features, which are recognized based on Myers-Briggs Type Indicator's (MBTI), play a key role in predicting learning styles in an online course. The features are determined and ranked using pattern recognition techniques, such as K-means clustering algorithm, to show which features can be better to separate learning style dimensions. The results show several features can be used to predict learning styles with high precision.

KEYWORDS

Learning style, e-learning, MBTI, learner's behavior

1. INTRODUCTION

Today, access to the web and the general use of computers have created several opportunities for e-learning systems, such as fully online and blended learning systems. E-learning environments, like all other tools, offer advantages such as access to different online resources, self-directed learning, and self-paced learning. Despite all the advantages, this kind of learning environment lacks the necessary attractiveness and the dynamic characteristic of the face-to-face learning setups. In other words, it is crucial to pay attention to the human characteristics and use them in the design and development of e-learning environments, aiming to make them more realistic and attractive [1]. This characteristic can give students the sensation that there is a human tutor involved in the process who follows their learning progress and cares about them [2]. Consequently, matching an e-learning environment with individuals' needs and personality characteristics helps them to learn more efficiently, than a system without adaptation to the learners' characteristics. One of the most important characteristics in the learning process is the learning style which can be different between each two students. For example, while a group of students may like to listen and speak, the other prefer to analyze a text, or just listen to a visual media. When students cannot grasp a course's materials, because of a mismatch with their learning styles, they lose their motivation for further learning [3]. This indicates that the better learning takes place when e-learning systems consider the learning style of the students. Thus, an automatic learning style recognition system significantly improves interactive training and learning systems.

Several studies have been carried out in order to consider human characteristics such as learning styles in human computer interaction such as e-learning environments. In 2011, Gong and Wang [4] used support vector machine (SVM) in order to determine learning style in an e-learning environment. Haron suggested a learning system including a learning module which can be adapted to each learner and utilizes fuzzy logic and MBTI personality test [5]. Abrahamian and his colleagues designed a user interface for learners based on their personality type using MBTI test [6]. They considered two dimensions of MBTI in their research. In [7], Bayesian networks is used in detecting the learning style of a student in a tutoring system. In [8], the students' behavior in a web-based course were studied. Results indicated that students with different learning styles use different ways to learn and interact with the course. The authors also presented a framework for automatic detection of a learner's learning style based on the Felder-Silverman model. Fatahi and her

colleagues [9] designed and implemented a virtual tutor and virtual classmate agent that had personality and emotion features similar to a human being. They used the OCC model for emotion modeling and MBTI for personality modeling. The system was tested and the results showed that the learning quality and the learners' satisfaction are improved by incorporating the human features in the interaction with the learners.

As mentioned above, several studies have been carried out in order to consider human characteristics in e-learning environments. Although there have been few studies trying to extract a student's learning style and attempt to adapt the courses to it, none of these studies has taken into account the relationship between learning contents and learning styles. In this paper, we extracted a set of features which are important in determining learning styles from students' behavior automatically based on the MBTI model. The proposed approach uses students' behavioral features during the interaction with a learning content management system. We use clustering methods such as K-means to extract the best features that can distinguish learning style dimensions. In addition, we use MBTI personality type for detecting learning styles of learners whereas most of previous studies focused on Felder-Silverman learning styles.

2. MATERIALS AND METHODS

2.1 Learning Styles and MBTI

According to the Keefe's definition [10], "Learning style consists of cognitive, emotional, and physiological features which are used to recognize how the learner understands the concepts, interacts with the learning environment, and reacts to the environment". There are many questionnaires categorizing each person in line with his/her learning styles [11]; such as Kolb's, Honey and Mumford's [12], GRSLSS's learning styles [12] and Felder-Silverman [13]. In comparison to the other questionnaires, the MBTI has been widely used and validated in the educational sciences [14]. It should be noted that MBTI can be considered as one of the most powerful tools to determine a learner's learning styles.

MBTI is an assessment tool based on the Jung Personality Theory [15]. Based on Jung's personality theory, every person has instinctive priorities that specify their behavior in different conditions [16]. Jung's type theory specifies three dimensions: Extraversion/Introversion (E/I); Sensing/Intuition (S/N); and Thinking/Feeling (T/F). Myers and Briggs added another dimension to Jung's typological model, i.e. Judging/Perceiving (J/P) [17]. These dimensions determined the learner's learning style [17] [18]. Sixteen personality types result from mixing four two-dimensional functions. Individuals are categorized into these types after filling out the questionnaire [19]. For example, people in ESTJ group are all extrovert, sensing, thinking, and judging.

2.2 Research Design

In this paper, we focus on the learning style features that can be obtained automatically through a learner's behavior. In the following, a procedure for extracting the human behavior traits, based on a selected learning style model in an e-learning environment, is proposed here.

//The algorithm for selecting the best online behavioral features, in an e-learning environment, to determine
//learning Style (LS)
//Output: The vector of best online behavioral features for predicting LS
Step 1: Select a learning style (LS) model
Step 2: Determine the behavioral features related to the selected LS
Step 3: Determine the mapping between LS behavioral features and users' behavior in the e-learning environment
Step 4: Data collection
    For all students
        a. Determine their LS
        b. Determine features' values based on logged behavior
Step 5: Feature selection
    For each user's behavioral features
        Find the relationship between each set of all dimensions of LS
    For each cluster of all LS dimensions
        If the relationship between LS dimensions and students' behavioral feature is positive
            Add this feature to the vector of best features

In this paper, in the first step of the procedure, we chose MBTI which is an assessment tool related to Jung's personality theory. According to the Center for Applications of Psychological Type [20], MBTI is the most commonly used personality inventory in history, with approximately 2,000,000 people have used MBTI for their personality detection every year. Moreover, the validity of MBTI theory has been widely shown [14] [21] [22]. The MBTI model is chosen and the general human behavior traits of this model related to the selected e-learning environment have been selected. In step 2, for each dimension, a few human behavior traits which are general and independent of the e-learning domain have been selected. These behaviors expressed in the most of studies about MBTI [18] . In step 3, we defined a mapping between LSs and a user's behavior presented through the interaction logs based on psychological studies [18]. Our goal is to define features that can be extracted from e-learning logs which correspond to the LS behavioral features. For example, in the MBTI learning style model, psychologists suggest that extroverts are interested in group working [18]. Consequently, we expect them to have a large number of interactions in group-based and collaborative tasks such as forums, group projects, and team-based homework. In contrast, introverts prefer listening more than talking [18]; hence, it is anticipated that these learners prefer viewing forums and read posts more than writing in forums. The list of the features are presented in table 1.

Table 1. List of features extracted from Moodle

Forums	Homework	Chat	Folder	Other
• The number of viewing discussions • The number of deleting discussions • The number of adding discussions • The number of viewing posts • The number of deleting posts • The number of adding posts • The number of updating posts • The frequency of viewing forums that includes students' discussion • The frequency of viewing forums that includes TA news	• The time of the first view of HWs files since uploading • The upload time of a HW • The number of viewing feedback from a HW • The number of viewing HW file from the second update to last update • The number of viewing homework file since the first to second update • The number of viewing homework file to the first update • The number of times viewing homework solutions	• The number of messages sent • The chat view times **Project** • The frequency of viewing discussion about projects • The time of the first view of project files since uploading • The number of times viewing project files **Quiz** • The number of viewing quiz page • The time of the first view of a quiz since uploading • The time spent on answering quiz • The number of reviews in each question • The number of viewing quiz grade	• The number of times viewing slides of the book • The number of times viewing the chapters of the book folder • The number of times viewing the summery slides folder • The first time viewing the slides of the book • The time of the first book files' view since uploading • The time of the first book chapter folder viewing • The first time viewing the summery slides folder	• The notes view times • The blogs view times • The number of times using the search option • The changing the account password times • The number of times updating the profile • The number of times viewing the units information folder • The number of viewing the final exam sample files • The number of viewing the midterm exam sample files • The first time viewing the exam sample files

In Step 4, in order to find the most important behavioral features to automatically determine the LS dimensions of a student, adequate data should be collected that can show a significant correlation between the behavioral features and the learning styles. Our goal is to select the best behavioral features that can determine the learning styles dimensions. Thus, in step 5, each feature is used to cluster the students into different groups. If the clustered groups highly correlate to a set of learning styles dimensions, then the feature is added to a feature vector suitable for determining the learning styles dimensions of a learner.

To extract the important features based on the features explained in Table 1, we define features based on the context (Context-based Learning Activity Feature or CLAF for short). Since the context of learning activities could impact the interaction of a student with it, the CLAFs are determined considering the context of the activities. For instance, in our study, there were 9 homework in our course with different context resulting into nine CLAFs for a Learning Activity Feature. In our course, we came up with 112 CLAFs.

3. EXPERIMENTS

This study is based on a blended learning environment for the “Introduction to computing systems and programming” (ICSP) course. The course is taught to the first-year students at the school of electrical and computer engineering at the University of Tehran in Iran. The course runs for 18 weeks, and is composed of two lectures and a laboratory part each week. The online part of the course includes general discussion forums, a place to receive and submit homework, download slides and notes, and check grades which are managed through Moodle Learning Content Management System.

To evaluate our proposed algorithm, we have collected data from two hundred and twenty six students who participated in this study. At first, they registered for the course and filled the MBTI questionnaire to determine their learning styles, as our ground truth. However, only two hundred and three of these students filled out the questionnaire. The time student spent to complete the MBTI questionnaire was recorded. The data of students who spent less than three minutes or more than eighteen minutes on the MBTI questionnaire were discarded to avoid contaminating the data with unreliable data. Consequently, data from five students were discarded. Finally, data from one hundred ninety eight students were used in this study. Table 2 shows the learning styles' distribution for all dimensions. In addition, Figure 1 shows the distribution of personality types in each dimension of MBTI.

Table 2. The distribution of the learning styles for the students

Learning Styles	Extroversion	Introversion	Sensing	Intuition	Thinking	Feeling	Judging	Perceiving
Number of Students	69	129	108	90	149	54	106	92

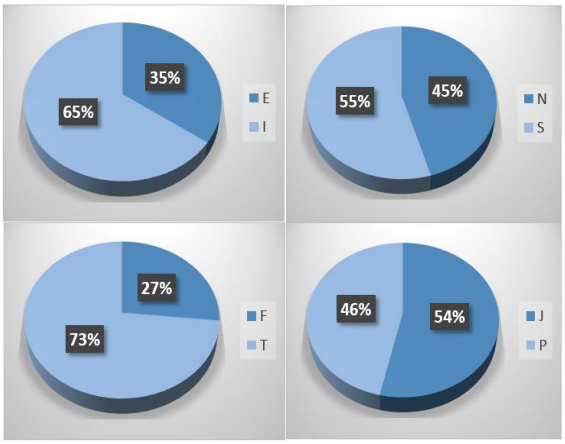

Figure. 1. Distribution of personality types at the school of ECE, University of Tehran, in each dimension of MBTI

Although there are not equal number of people in each dimension, we selected equal number of samples to select best features in each dimensions. if we do not consider equal number of samples in each dimension, the clustering`s results are biased toward the class with more samples.

Total of 35,155 interaction records were collected of the students from the Moodle’s log file. Each record includes "time", "IP address", "action", "URL", "info", "username", "first name", "last name" and "email of the corresponding student". The "time" shows the duration of time students did an activity and the "info" includes an id uniquely assigned to the page accessed/used by the students.

4. RESULTS

As mentioned earlier, the data were collected, i.e. the learning styles of the students and their e-learning behavioral interactions with the system, for 198 records of students. As mentioned earlier, in our study, there were 9 homework in our course with different context resulting into nine CLAFs for a Learning Activity Feature. In our course, we came up with 112 CLAFs.

To determine the best features for determining the learning styles, we run many important clustering methods through Weka tools. K-means (k=2) was the suitable method for separating two MBTI dimensions. It seems clear that the k-means is one of the simplest algorithm which uses unsupervised learning method to solve known clustering issues. It is fast, robust and easier to understand. The k-means produce tighter clusters than other methods for example hierarchical clustering, especially if the clusters are globular. It gives best result when data set are distinct or well separated from each other. The disadvantage of the k-means method is predict k-value. In this research, this problem solved because there are two clusters in each dimension of personality type.

Finally, k-means reports the better results rather than others. We run k-means method on all features. Tables 3, 4, 5, and 6 report which features have the high percentage of correctly clustered instances in each dimension in MBTI.

Table 3. The list of context-based learning activity features to determine LSs in E/I dimension

Feature	Personality Dimension	Extroversion	Introversion
The response time of the Quizzes	Precision	0.62	0.63
	Recall	0.66	0.57
	Accuracy	0.62	0.62
	F-measure	0.64	0.60

Table 4. The list of context-based learning activity features to determine LSs in S/N dimension

Feature	Personality Dimension	Sensing	Intuition
The frequency of the folder view (the folders which includes the summery slide about each chapter of the course)	Precision	0.63	0.94
	Recall	0.63	0.42
	Accuracy	0.7	0.70
	F-measure	0.63	0.58
The frequency of the homework solution view (the homework which is about the assembly language)	Precision	0.70	0.59
	Recall	0.44	0.81
	Accuracy	0.62	0.62
	F-measure	0.54	0.68
The frequency of the homework solution view (the homework which is about the assembly language)	Precision	0.70	0.59
	Recall	0.44	0.81
	Accuracy	0.62	0.62
	F-measure	0.54	0.68
The number of times viewing HW_1 file (which is about flowchart)	Precision	0.70	0.59
	Recall	0.44	0.81
	Accuracy	0.62	0.62
	F-measure	0.54	0.68

Table 5. The list of context-based learning activity features to determine LSs in T/F dimension

Feature	Personality Dimension	Thinking	Feeling
The frequency of viewing the forum (that includes the TAs news)	Precision	0.67	0.73
	Recall	0.78	0.61
	Accuracy	0.70	0.70
	F-measure	0.72	0.67
The number of times viewing the units information folder	Precision	1	0.64
	Recall	0.44	1
	Accuracy	0.72	0.72
	F-measure	0.61	0.78
The number of messages sent in the chat rooms	Precision	0.61	0.87
	Recall	0.94	0.39
	Accuracy	0.67	0.67
	F-measure	0.74	0.54

Table 6. The list of context-based learning activity features to determine LSs in J/P dimension

Feature	Personality Dimension	Perceiving	Judging
The time of the first view of the phase 3 file of the project since uploading	Precision	0.65	0.63
	Recall	0.62	0.66
	Accuracy	0.64	0.64
	F-measure	0.63	0.64

This paper aimed at determining the relationship between human behavioral features and learning styles dimensions in e-learning systems. Results in table 3, 4, 5 and 6 show there is a significant relationship between the human behaviors and the learning styles that can help to determine the learning style dimensions of a learner.

Results in tables 3, 4, 5, and 6 indicate that there are features in each dimension of MBTI that are relevant to learners' interaction with the e-learning system. Based on the fact that extroverted people tend to be fast in doing tasks, act quickly and sometimes without thinking, while the opposite is true for introverts. Our result confirms these behaviors because the response time of the quizzes is a feature that can cluster extroverts and introverts. Extrovert students have a high-speed in answering quizzes and do not waste time, but introvert students spend more time for thinking about the quizzes.

Intuitive people attend more to the whole concept than to the details. In contrast, sensing people learn best from an ordered sequence of details, like to know the "right way" to solve problems, and wish to answer homework in details. Our result confirms such behavior. Frequency of folder viewing that includes summary slide about each chapter of the course is a suitable feature for clustering intuitive and sensing people. Results show intuitive people have more tendency to view summary slides than sensing people because they attend more to the whole than the details. Also, frequency of a detailed homework solution viewing, which was Homework 2 in our case, is an appropriate feature to separate sensing and intuitive students. This is because sensing students welcome details, especially difficult subjects such as assembly language which was the subject of the 2nd homework. The frequency of viewing homework solutions, especially related to algorithms and flowcharts, are well-suited features to separate S/N dimension. Sensing people learn best from an orderly sequence of details and love to know the "right way" to solve problems.

Thinking people like specific and logical courses and objective subjects. Feeling people are personal, like warm people relationships, and they are more interested in people than things or ideas. Consequently the frequency of discussion viewing and frequency of posting to forums are the best features to separate feeling students from thinking ones. These features confirm that feeling people tend to interact and relate to other students through forums and discussion rooms. Table 5 shows that feeling students used chat rooms more than thinking people. Since thinking students prefer specific and detailed information, they used and downloaded materials from a special folder that included their program of study charts which explain the courses that the students should take during their four years of study.

While the perceiving people have a flexible lifestyle and do not worry about deadlines, the judging people prefer a systematized life and regulated thoughts and ideas and care about activities which they can do on time. Our results show that the first time viewing files after uploading is a relevant feature that separates judging and perceiving students. Judging students are ordered, organized, and systematic, so they immediately view a file after uploading in the learning content management system.

As the results show, there is one feature which is related to E/I and J/P dimensions while we initially estimated many behavioral features for these dimensions. This is due to the fact that many features, which are detectable in the face-to-face environment, are not detectable in e-learning environment. Consequently, we have a limited number of behavioral features in e-learning environments and the number of features vary based on the each dimension.

5. CONCLUSION

In this paper a set of features suitable for automatically determining the learning style dimensions of a learner is presented. These learners' behavioral features, which have been extracted from the logged interaction of the learners with an e-learning system, have been determined using clustering approaches. We have used the MBTI model for identifying learning styles of students. The automatic recognition of users' learning styles would be an important feature for intelligent tutoring systems' designers to develop adaptive systems capable of adjusting to a user's learning style. If these systems offer courses in harmony with learning styles of students, they have made a substantial progress in the learning process. A significant finding in this paper is

finding a suitable feature vector to determine learning styles of students in automatic way. Furthermore, it has been shown that the context and content play an important role in the learning process based on learning styles.

One limitation of the study is the number and diversity of the participants. Another limitation is the nature of blended learning structure of our e-learning system. Consequently, the students do not show all their behavior through online system and the off-line behaviors are not captured and used for learning style recognition.

The findings in this study can be used as the basis for further research and improvements for providing advanced adaptability, especially in Learning Management Systems (LMS). Future work will deal with automatic student modeling process of learning styles and evaluating its effectiveness in improving the accuracy. Likewise, finding a sequence of human behavior will be of great value in predicting learning style.

REFERENCES

[1] A. Niesler and G. Wydmuch, "User Profiling in Intelligent Tutoring Systems Based on Myers-Briggs Personality Types," in In Proceedings of the International Multi Conference of Engineers and Computer Scientists, 2009.

[2] X. Mao and Z. Li, "Agent based affective tutoring systems: A pilot study," Computers & Education, p. 202–208, 2012.

[3] B. S. Jong, C. H. Lai, Y. T. Hsia and T. W. Lin, "Using Game-Based Cooperative Learning to Improve Learning Motivation: A Study of Online Game Use in an Operating Systems Course," IEEE Transactions on Education, pp. 183-190, 2013.

[4] W. Gong and W. Wang, "Application Research of Support Vector Machine in E-Learning for Personality," in In IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), 2011.

[5] N. Haron and N. Salim, "Empirical Evaluation of Mixed Approach in Adaptive Hypermedia Learning System," in Postgraduate Annual Research Seminar, 2006.

[6] E. Abrahamian, J. Weinberg, M. Grady and . C. M. Stanton, "The Effect of Personality-Aware Computer-Human Interfaces on Learning," Universal Computer Science, pp. 27-37, 2004.

[7] P. García, A. Amandi, S. Schiaffino and M. Campo, "Evaluating Bayesian networks' precision for detecting students' learning styles," Computers & Education, p. 794–808, 2007.

[8] S. Graf and . T. C. Liu, "Analysis of Learners' Navigational Behaviour and their Learning Styles in an Online Course," Journal of Computer Assisted Learning, pp. 116-131, 2010.

[9] S. Fatahi and N. Ghasem-Aghaee, "Design and Implementation of an Intelligent Educational Model Based on Personality and Learner's Emotion," International Journal of Computer Science and Information Security, pp. 1-13, 2010.

[10] L.-f. Zhang, "Thinking styles and personality types revisited," Personality and Individual Differences, pp. 883-894, 2001.

[11] T. Anderson and F. Elloumi, Theory and practice of online learning (1rd ed.)., Athabasca: Athabasca University Press, 2003.

[12] K. Logan and P. Thomas, "Learning Styles in Distance Education Students Learning to Program," in In Proceedings of 14th Workshop of the Psychology of Programming Interest Group, 2002.

[13] R. Felder and L. K. Silverman, "Learning and Teaching Styles," Journal of Engineering, 1988.

[14] J. Kim, A. Lee and H. Ryu, "Personality and its effects on learning performance: Design guidelines for an adaptive e-learning system based on a user model," International Journal of Industrial Ergonomics, pp. 450-461, 2013.

[15] D. Schultz and S. Schultz, Theories of Personality, Wadsworth publishing, 2008.

[16] T. Dewar and D. Whittington, "Online Learners and their Learning Strategies," Journal of Educational Computing Research, pp. 415-433, 2000.

[17] S. Rushton, J. Morgan and M. Richard, "Teacher's Myers- Briggs personality profiles: Identifying effective teacher personality traits," Journal of Teaching and Teacher Education, pp. 432-441, 2007.

[18] "MBTI Web Page," [Online]. Available: http://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/.

[19] S. Jessee, A. Paula , N. O'Neill and R. O. Dosch, "Matching student personality types and learning preferences to teaching methodologies," Journal of Dental Education, pp. 644-651, 2006.

[20] "Center for applications of psychological type.," [Online]. Available: http://www.capt.org/.

[21] M. Malik and . S. Zamir, "The Relationship between Myers Briggs Type Indicator (MBTI) and Emotional Intelligence among University Students," Journal of Education and Practice, 2014.

[22] V. Rodrı´guez Montequı´n and J. M. Mesa Ferna´ndez,, "Using MBTI for the success assessment of engineering teams in project-based learning," in International Journal of Technology Design Education, 2013.