Learning Style Classification Based on Student's Behavior in Moodle Learning Management System

Manal Abdullah, Asmaa Alqahtani, Jawhara Aljabri, Reem Altowirgi and Ruqiah Fallatah
Department of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia
maaabdullah@kau.edu.sa; alraizz@stu.kau.edu.sa; jaljabri0004@stu.kau.edu.sa;
ral-towirgi@stu.kau.edu.sa; rfallatah0008@stu.kau.edu.sa

ABSTRACT

In learning field, each student has his own learning style that affects his way of get, process, understand and percept information. Determining the learning style of students enhances the performance of learning process. Two methods are commonly used to acquire student's learning style: static by questionnaire and dynamic by tracing student's navigation on e-learning environment.

In this paper, a new approach to classify students dynamically depending on their learning style was proposed. This approach was experimented on 35 students for Data Structure online course created using Moodle. By extracting students' behavior, data from Moodle log, the learning style for each student was identified according to Felder and Silverman model. Also, learning style based on the behavior have been compared with a quiz results conducted at the end of the course.

Receiver operating characteristic (ROC) curve have been used to evaluate the quality of classification results comparing with quiz results. Good results with average accuracy of 76% are achieved. Students' data have been divided into four training and testing sets with different splitting ratio. Different testing accuracy values are obtained for the different ratios using each dimension of Felder-Silverman learning style model (FSLSM).

Keywords: E-learning, Moodle, learning styles, Felder-Silverman learning style, classification.

1 Introduction

Many studies in cognitive psychology indicate that each person differs in the ways in which he/she uses to analyze the information. With the development of psychological studies, there is an increasing desire to study the differences in getting and processing of information.

In the environment of education, there are many ways of thinking, analyzing information, and solving problems. Every student has a personal way to get, handle and understand the information that called student's learning style [1]. For example, some students prefer to learn by seeing graphs, pictures and presentation; some others like to learn by listening; some others like to learn by doing experiments.

Knowing the learning style of students will help educators to provide the learning content in an appropriate format that matching the student's learning style to get greater learning process [2]. There are many ways to know student learning style. The most popular method is the questionnaire. Acquiring student's learning style from this static method is time consuming and may be not accurate [3]. After the
emergence of e-learning and using Learning Management System (LMS) in education, it becomes more appropriate to acquire the learning style of students dynamically and indirectly while they navigate through e-learning environment.

In this experiment, Moodle LMS was used to develop Data structure course which is a foundation course for computing study. Data Structure course has rich course topics that can be represented using different component forms provided by Moodle LMS such as texts, graphs, animation, exercises. Also, interaction between students are possible by using forums, so student can post topics and reply to others. At the end of the course, quiz should be conducted. 35 students in computer science department were involved in the study. All student activities through the course are stored in Moodle LMS log and the wanted features can be extracted any time.

In this research, we want to identify learning style of students dynamically based on their behavior within the course and compare these styles with the quiz results to evaluate the proposed approach.

This research is organized as follows: section 2 provides background material. Section 3 presents related work. Section 4 describes the system model. Section 5 is methodology and experiment. Sections 6 discuss the results of the proposed system. Section 7 gives the conclusion of the study and the future work.

2 Background

Learning is known as procedures that the teacher used to achieve some relatively permanent change in how students think in obtaining useful information in a short and clever way [4]. Learning is divided into two major categories: traditional learning and e-Learning. Traditional learning is a simple oral spelling where students sat quietly at their places, listened carefully to the teacher whose primary task is assigning, and listening to his students, find out their mistakes and inform them about the correct ideas [5]. E-Learning is a set of learning services and technology tactics aim to provide high value full learning; anytime, anyplace [6].

2.1 Learning Styles

There are many learning styles applicable to both traditional learning and e-learning, which can be described as a set of characteristics and behaviors that define the way of learning [7]. Different styles affect the form that learners learn, how they can interact in interactive methods with learners. Each style has its features and does best in certain circumstances. Learning styles contribute to improve the performance of both teaching and learning process, so that the teacher, using learning styles, can affect not only the way used to deliver the information easily, but also can effect emotionally on his students. As a result, the course will be likable for the students even in case of lack in physical and financial possibilities.

Several learning style models have been used to identify students' learning styles such as Felder Silverman Learning Style Model (FSLSM), Dunn and Dunn model, Honey and Mumford, Kolb's theory of experimental learning and Howard Gardner Multiple Intelligence. FSLSM [7] is used in this study to detect students' learning style. The reason of choosing FSLSM is that the development of the hypermedia learning system with incorporated learning components including navigation tool, different forms of presenting course materials (graphics, video, sound) is suitable to the dimensions of this model.
This model categorized student's learning styles into four dimensions. Table 1 summarize the characteristics of learners based on these dimensions.

<table>
<thead>
<tr>
<th>Table 1: Felder Silverman learning style model dimensions [7]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Processing dimension: Describes the processing of information</strong></td>
</tr>
<tr>
<td><strong>Active</strong></td>
</tr>
<tr>
<td>Learners impart by making something with the information. They need to talk about information to process it.</td>
</tr>
<tr>
<td><strong>Perception dimension: Describes the perceiving of information</strong></td>
</tr>
<tr>
<td><strong>Sensing</strong></td>
</tr>
<tr>
<td>Learners impart by focusing on details, facts, figures and dealing with proven functions that have real applications.</td>
</tr>
<tr>
<td><strong>Input dimension: Describes the presenting of information</strong></td>
</tr>
<tr>
<td><strong>Visual</strong></td>
</tr>
<tr>
<td>Learners impart by focusing on diagrams, charts or anything can present information.</td>
</tr>
<tr>
<td><strong>Understanding dimension: Describes the understanding of information</strong></td>
</tr>
<tr>
<td><strong>Sequential</strong></td>
</tr>
<tr>
<td>Learners impart by focusing on managing information in a linear way and dealing with logic and followed steps.</td>
</tr>
</tbody>
</table>

Felder and Silverman have developed Index of Learning Style (ILS) questionnaire that consists of 44 questions with two possible answers, A or B. This questionnaire used to determine the learning style of students explicitly [8].

2.2 Learning Management System (Moodle)

LMS is a system aims to support learning content development depending on the web. LMS has many features including course creation, information delivery, students’ enrolment and navigation tracking [9]. In LMS, educators can deliver the information to learners; create course materials, conducting exercises and quizzes. In addition, they can manage learning and engage students’ discussion.

Many examples of LMSs are Moodle, Blackboard and webCT. Moodle [10] is used in this study for building the e-learning system. It is an acronym for Modular Object-oriented Dynamic Learning Environment. Moodle offers an important feature for the e-learning systems developed using it. It provides a huge amount of students’ information that is very helpful for analyzing their behavior on the course.

2.3 Classification

Classification is one type of the Data Mining techniques. It is a process of identifying understandable patterns and providing the meaningful information from the given data set. It is mainly used in computer science fields such as pattern recognition, statistics, and data base management in order to
analyze a given dataset and to take each instance of it. In addition, it is used to find models that define significant data classes within the given dataset, and extract the relation between large sets of data [11].

### 2.3.1 Classification Procedures

Two classification types can be defined: supervised and unsupervised classifications. Supervised is defined as processes that enable the user to select a training dataset and perform the classification algorithm on it, and then creates a model that can measure the performance and accuracy of test dataset. On the other hand, unsupervised classification outcomes are based on the software analysis of the elements, without the user defining sample classes. The computer process techniques to determine which items are related, and it must belong to actual features in order to group them into classes [11]. In this research, we rely on supervised classification which consists of the following steps [12]:

**2.3.1.1 Data Collection and Feature Extraction**

Pre-processing operation is the first step performed on raw data collections. In feature extraction, it is better to discard the samples rows that have no values, and the attribute columns for which data is not found.

**2.3.1.2 Sampling**

After extracting features from the raw data, the dataset must be defined randomly into training and test dataset. The training dataset will be used to practice the model. The test dataset then will be used to evaluate the performance of the final model.

**2.3.1.3 Validation**

Validation is one of the most useful techniques to test different combinations of feature selection, dimensionality reduction, and learning algorithms.

**2.3.1.4 Normalization**

Normalization feature technique is applying in order to make comparisons between various attributes, especially, if the attributes are measured on different dimensions, and it is an essential requirement for machine learning algorithms.

### 2.3.2 Classification Models

Classification includes a particular outcome depend on a given input. The algorithm processes a training dataset consist of a set of attributes to providing outcomes. The algorithm tries to find out relationships between the characteristics that make it possible to predict the outcome. There are numbers of different most used learning algorithms, but the decision tree classification model is the most popular model used [11].

**2.3.2.1 Decision Tree Classification Model**

A decision tree is defined as a classifier in the form of data structure to analyze, recognize, and decide a particular pattern. Decision tree starts the test from the root of the tree. Then, test moves through the tree until the leaf node. Using the pruning process to stop tree splitting and decide leaf nodes with a small number of points of error or some fixed percentage of the total training set [12].
2.3.2.2 Types of Decision Tree

Decision tree has three main types that are: Classification tree analysis, Regression tree, and CART analysis “Classification and Regression Trees”.

- Classification tree analysis (CTA): is an algorithm of machine learning used when the outcome of the decision for tree is a class to which the data belongs. The CTA is based on the C4.5 decision tree algorithm [12].
- Regression tree: is used when the predicted outcome is a real ordered number.
- CART analysis: CART is the abbreviation of “Classification and Regression Trees”. This type of the decision tree combines the previous two types of the decision tree [14].

3 Related work

Several studies have been conducted to detect learning style of students based on their behaviors in LMS by using different techniques and approaches.

Graf et al. [15] proposed an automatic student modeling approach for identifying learning styles in LMS. This approach is developed in a generic way, based on commonly used features in LMSs including: content objects, outlines, examples, self-assessment test, exercises, and discussion forums. Data about students' behavior can be used as hints for learning style preferences depending on FSLSM. Then, the respective learning style for each student will be calculated by applying a simple rule-based mechanism on these hints.

Khribi et al. [16] intended to adopt a learner model with three components: learner's profile, learner's knowledge and learner's educational preferences. Based on a web usage mining mechanism and literature-based approach, learning styles have been identified using indication from learners' behaviors based on FSLSM. After learner model has built, they apply a hierarchical multilevel model to assign learners with common preferences and interests to the same groups, so that feedback from one learner can serve as a guideline for information delivery to the other learners within the same group.

Dung et al. [17] promoted a new method to estimate the learning style depending on the number of visits and time spent on a specific learning object. Authors developed architecture for multi-agent adaptive learning system and implemented their own web based LMS called POLCA based on this architecture. After identifying student's learning style using literature-based method, the system automatically adapts the contents to match the detected learning style.

Fasihuddin et al. [18] focused on personalization of Massive Open Online Courses (MOOCs) learning environments based on learning styles theory using FSLSM. The main goal of this study was determining the patterns that provide hints to identify students' learning styles in MOOC learning environment. Many patterns of student's interaction have monitored regarding for each learning dimension of FSLSM. After identifying learning style for students, adaptation to the learning content occurs through navigational support.

Reddy et al. [19] has collected data from engineering students using Moodle tool with ILS Questionnaire. In this research, how the student should select the different courses based on their learning styles in different levels is derived. Different classifiers are applied to frame set of rules to select suitable courses based on their learning style.
All these researches were identifying students' learning style. However, these studies depend only on students' behavior in e-learning environment and they lacked to validate their classification process. Our study focuses on obtaining behavior data of students when they are interacting with Moodle LMS and normalizing these data with ILS questionnaire results they have filled to get high precision classification of students based on their learning styles according to FSLSM. Also, students' quiz data have been used to validate the classification process.

4 Moodle E-learning System Model

In this section, the model of learning style classification depending on student's behavior in Moodle LMS will be described. Figure 1 shows the architecture of the system model.

The main components of the system model are as in the following:

- **Moodle**: A LMS used for developing Data Structure course to enable students to interact with the course material.
- **LOG**: All students’ activities in Moodle including students' behaviors and students' quiz results will be stored directly in LOG File.
- **Students' Data Extraction**: it is the process of choosing the appropriate features of students that matches FSLSM dimensions from data stored in LOG file. These features includes: **Behavior Features** and **Quiz Features**.
- **ILS Questionnaire**: Results of ILS questionnaire were collected and analyzed for using in normalization process.
- **Normalization**: Is the process of organizing the students’ behavior data collected with ILS questionnaire results, then finding relationship in order to represent them in appropriate format to create **Behavior Classification Rules**.

![Figure 1: System Model](image-url)
- **Rule Estimation**: Is the process of designing rules depending on the Quiz results to create Quiz Classification Rules.

- **Rule Validation**: Is the process of validating Behavior Classification Rules with Quiz Classification Rules.

- **Behavior Classification Model**: This model is used to classify students’ learning styles based on their behavior using Behaviors Classification Rules.

- **Quiz Classification Model**: This model is used to classify students’ learning styles based on their quiz results using Quiz Classification Rules.

- **Classification**: It is a process of classifying students to their learning styles.

- **Accuracy**: It will be calculated depending on the error rate between learning styles given by Behavior Classification Model and by Quiz Classification Model.

5 System Methodology

5.1 Experimental Setup

Moodle LMS is used to build Data Structure course for undergraduate computing students at Majmaah University in Saudi Arabia. Course contents had been chosen to reflect FSLMS dimensions. An outline of the course topics was presented. Each course topic is represented in many forms; such as: text, video, simulation animation, PowerPoint slideshow. For each topic, some examples and exercises are provided. Sample code is presented to students for running and modifying. The system also contains forums for discussion. Figure 2 shows the used e-learning system for the selected course for this research, data Structure course provides codes, structural figures, and explanation text which serve each dimension of FSLMS learning style.

![Figure 2: Screenshot of Moodle system developed for Data Structure course](http://dx.doi.org/10.14738/tmlai.31.868)
At the end of the course, quiz is conducted and submitted. Quiz questions were taken from each course topic presented in the system and from all components introduced in the course that reflect FSLSM dimension. Quiz questions are in true/false form. Screenshot of the quiz is shown in Figure 3.

![Figure 3: Screenshot of quiz presented in Moodle for Data Structure course](image)

54 students were enrolled to the course. Each student should login to the course, browsing its contents, understanding its topics in a way she prefers. Then, the student should solve and submit the quiz. Only 35 students have logged in the course and navigate through its contents. ILS Questionnaire was provided to those 54 students to fill and submit. Only 40 students have responded. Questionnaire results were collected and analyzed for later use in student classification.

5.2 Training phase

Students’ data can be acquired from LOG provided by Moodle LMS. These data includes student name, IP address, activity done, component accessed, and date and time for each activity. Students' data have been collected by tracking the students' navigation in the course, and by analyzing their answers of the quiz. Screenshot of LOG is shown in Figure 4.

![Figure 4: Screenshot of LOG provided by Moodle system](image)
5.2.1 Students' Behavior Data

For each student, navigation activity was collected and categorized according to 10 attributes which represent the students' characteristics based on FSLSM [3]. A zero-one matrix shown in Figure 5 is used to reflect and analyze the student's behavior given from LOG.

For each student, attributes of behaviors are normalized with her questionnaire attributes to estimate the classification rules depending on the student behavior. Rules used to identifying student learning style depending on the behavior in Data Structure course are shown in Table 2.

5.2.2 Students' Quiz Data

For each student, quiz data is collected from LOG. Students have been classified according to FSLSM dimensions based on their solved question.

Behavior classification rules are validated using quiz data for each student. Rules results of students' classification depending on behavior data were extremely identical with results of students' quiz classification rules.

![Figure 5: Students' behavior attributes](image)

### Table 2: Learning style classification rules based on student's behavior

<table>
<thead>
<tr>
<th>Learning Content</th>
<th>Learning Content Capabilities</th>
<th>Student Expected Behavior on Learning Content</th>
<th>Student Learning Style</th>
<th>Affected Learning Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forum</td>
<td>Conducting discussion between students based on certain topics</td>
<td>post and reply</td>
<td>active</td>
<td></td>
</tr>
<tr>
<td>Sample code</td>
<td>Providing source codes for algorithms contained in the subject</td>
<td>view and no-action</td>
<td>reflective</td>
<td></td>
</tr>
<tr>
<td>Text</td>
<td>Describing course topics in text form</td>
<td>visiting text</td>
<td>verbal</td>
<td></td>
</tr>
<tr>
<td>PowerPoint Video Animation</td>
<td>Presenting learning materials in the form of pictures, videos and animations</td>
<td>viewing media</td>
<td>visual</td>
<td></td>
</tr>
<tr>
<td>Examples</td>
<td>Providing examples of the topics discussed in the class</td>
<td>visiting exercises</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>On-line exercises</td>
<td>Providing exercises in multiple choice questions and returning the feedback immediately to students</td>
<td>visiting exercises</td>
<td>sensor</td>
<td></td>
</tr>
<tr>
<td>Course outlines</td>
<td>Providing the content outlines of the course</td>
<td>no-visit</td>
<td>intuitive</td>
<td></td>
</tr>
<tr>
<td>Navigation</td>
<td>Navigation between course materials</td>
<td>viewing course outline</td>
<td>global</td>
<td></td>
</tr>
</tbody>
</table>

URL: [http://dx.doi.org/10.14738/tmlai.31.868](http://dx.doi.org/10.14738/tmlai.31.868)
5.3 Testing phase

When a student login to the system and navigate through it, the behavior data extracted from the LOG. The student's behavior will be classified based on Behavior Classification Model to obtain the learning style based on the student's navigation. Also, Student's quiz data were extracted from LOG. Quiz data will classified based on Quiz Classification Model to get the student's learning style based on the solved questions.

Learning style based on the student's behavior will compared with learning style based on the student's quiz for calculating the accuracy of matching between the two learning styles. Accuracy of matching will be discussed in details in the results section.

6 Results and Discussion

The proposed approach has been evaluated using Receiver Operating Characteristic (ROC) to measure the quality of classification results. ROC is a graphical curve plotted using the True Positive Rate (TPR) and the False Positive Rate (FPR) of the classification results [20]. For each FSLSM dimension, four TPR and FPR values have been calculated and presented in Table 3. Four ROC curves representing classification accuracy for each dimension are shown in Figure 6.

<table>
<thead>
<tr>
<th>Table 3: True Positive Rate and False Positive Rate values for each learning style dimension.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True Positive and False Positive values For Each Learning Style Dimension</strong></td>
</tr>
<tr>
<td>Processing Dimension</td>
</tr>
<tr>
<td>FPR TPR Accuracy</td>
</tr>
<tr>
<td>0.05 0.23 57%</td>
</tr>
<tr>
<td>0.12 0.47 63%</td>
</tr>
<tr>
<td>0.32 0.68 69%</td>
</tr>
<tr>
<td>0.45 0.92 82%</td>
</tr>
</tbody>
</table>

Figure 6: Four ROC curves representing classification accuracy for each dimension

Students’ data have been divided into four different training and testing sets with 20%, 40%, 60%, 80% sample splitting ratio. Table 4 illustrates testing accuracy of the different sets' partitions for each
dimension of FSLSM, as shown in the table 4. The classification of 40% training data and 60% testing data gives the highest accuracy for almost the four dimensions. Using this splitting ratio, it is obvious from ROC curve that input and understanding dimensions are nearly the same and not differentiate this course type. Also, computing the area under ROC curve that represents classification quality, input, processing and understanding are very close areas. This means that based on these three dimensions, the proposed model is a good classifier for data structure students and can help instructors of this type of courses to adapt their course materials with student's type to achieve high outcomes.

Table 4: Testing accuracy of classification with various partition of training and testing data

<table>
<thead>
<tr>
<th>Training Data</th>
<th>Testing Data</th>
<th>Testing Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>80%</td>
<td>Processing Dim.</td>
</tr>
<tr>
<td>40%</td>
<td>60%</td>
<td>Input Dim.</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>Perception Dim.</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>Understanding Dim.</td>
</tr>
</tbody>
</table>

7 Conclusion and Future work

In this study, we have proposed a model that classifies students depending on their behavior on Moodle course according to FSLSM. The model is experimented on 35 students enrolled in Data structure online course in Majmaah University at KSA.

The flexibility of Moodle allow instructor to track the students behavior during the course and within the quiz. Student behavior data extracted from Moodle log are normalized with ILS questionnaire to reach the final learning style. Comparing the final learning styles of each student with the results of quiz was used for validation. We have found that the accuracy of classification results is satisfactory with minimum rate of errors.

This study gives hints to the educators to format the course contents in an appropriate form matching the students' learning style to get best performance of learning process.

For the future work, we will propose a dynamic system that adapts its contents to match the learning style of student and responds immediately to her needs. Also, it is recommended to test other learning styles dimensions that may more suits computing study.

REFERENCES


URL: [http://dx.doi.org/10.14738/tmlai.31.868](http://dx.doi.org/10.14738/tmlai.31.868)


