Predicting Students’ Performance based on Instructors’ Pattern of Metacognitive Scaffolding

Nurul Farhana Jumaat\(^\text{1}\) and Zaidatun Tasir\(^\text{1}\)

\(^\text{1}\)Faculty of Education, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia.

\(^\text{2}\)Corresponding author: nfarhana@utm.my, Tel: 607-5538843, Fax: 607-5534884

**Abstract:** Metacognitive scaffolding is a strategy provided by an online instructor to assist students to develop their own understanding and improve their achievements in learning. To date, studies have focused on the effect of metacognitive scaffolding on students’ performance. However, researches that focus on the understanding of the pattern of metacognitive scaffolding and its effect on students’ learning success are still lacking. The purpose of this study is to predict students’ performance based on the instructors’ pattern of metacognitive scaffolding through data mining: a decision tree analysis. This study used a pre-experimental, post-test only research design. One online instructor and 37 postgraduate’s students from the Educational Technology program participated in the study. The data were mined based on the frequency of metacognitive scaffolding posted by the instructor in Facebook group discussion and also students’ scores in the performance test. The decision tree analysis predicts that students who achieved grade A in their study were prone to receive a combination of guidance that focused on the process of learning (MS4), the assistance from the instructor presenting the rationale for tasks and activities (MS1), the guidance from the instructor who encouraged the relationship among participants (MS5) and supervised text comprehension (MS7) by the instructor such as giving response and feedbacks towards students’ comments. Thus, the results serve as a guideline for online instructors to prompt these dominant mechanisms of metacognitive scaffolding in order to increase students’ performance in learning.

**Keywords:** educational data mining, decision tree analysis, metacognitive scaffolding, pattern discovery

1. **INTRODUCTION**

Scaffolding in education is best reflected by the works carried out by Lev Vygotsky in 1978 and his renowned Social Development Theory. The Social Development Theory suggests that social interaction plays an important role in cognitive development. In particular, Vygotsky emphasized that community plays an important role in the process of cognition development in which learning occurs through social interaction with skillful others [1]. The two main principles highlighted by Vygotsky in this theory are: More Knowledgeable Others (MKO) and Zone of Proximal Development (ZPD).

Vygotsky referred MKO as teachers, parents or peers who have a higher level of ability than the learners. In other words, MKO are the persons whom the learners seek guidance in understanding a particular task, process or concept. The Zone of Proximal Development (ZPD) is related to the concept of MKO. It is defined as the area between what is known and what is not known by a child during learning. The ZPD is the area in which a child receives guidance or instruction from MKO. As a result, the child is able to develop skills on his own and develop his understanding. With the existence of MKO, individuals are able to coordinate their cognition throughout the process of learning. Understanding this process of cognition is known as metacognition.

Meanwhile, John Flavell published a paper in 1976 which describes metacognition as a concept of defining “one’s knowledge concerning one’s own cognitive processes” (p. 232), or simply thinking about one’s own thinking. It consists of both knowledge of cognition and the ability to control and regulate the cognitive processes [2]. Other pioneer researchers who studied metacognition include Reese and Brown [3] who considered metacognition as a term that is generally referred as “individuals’ ability to understand and manipulate their own cognitive processes” (p.3). However, failing to acquire metacognitive abilities will result in poor academic performance especially in problem-solving tasks [4][5][6]. An individual’s metacognition can be guided through the use of specific strategies known as metacognitive scaffolding.

To date, studies have focused on the effect of metacognitive scaffolding on students’ learning process and their learning performance. However, researches that focus on understanding the pattern of metacognitive scaffolding and its effect on students’ learning success are absent. Hence, the purpose of this study is to predict students’ performance based on the pattern of instructors’ metacognitive scaffolding by using data mining, a decision tree analysis. Data mining is known as a powerful tool that is able to uncover hidden information from large datasets. Recently, researchers have used this application to analyze educational data to understand and predict students’ learning style, behavior, performance and examination especially in online learning environment.

2. **THEORETICAL BACKGROUND**

2.1 Metacognitive Scaffolding

Fouché and Lamport [7] stated that metacognition is not easily taught, nor can it be easily transferred. However, many researchers proved that students’ metacognition can be trained [8]. Accordingly [9], metacognitive scaffolding can support and teach students’ metacognition in learning. Luckin and Hemmerton [9] studied how the technique can become a strategy to support learners’ ability to challenge their own skill level and look for appropriate assistance.

The term metacognitive scaffolding was first introduced by Hannafin [10]. Metacognitive scaffolding is referred as one of the types of scaffolding in an online learning environment that assists learners in establishing what is known and how to think during a learning process. Thus, metacognitive scaffolding is a strategy provided by the instructor to scaffold or assist students to develop their own understanding and improve their achievements in learning.
In this study, the strategy was adapted from the seven mechanisms of metacognitive scaffolding developed by Reingold et al. [12]. The mechanisms include (MS1) Presenting rationale for task and activities, (MS2) Presenting the relationship between reading items, course objectives and tasks, (MS3) Supporting reflective writing, (MS4) Focusing on the process of learning, (MS5) Encourage relationships among participants, (MS6) Discriminating between conclusion/fact/opinion/hypothesis and, (MS7) Supervising text comprehension. Table 1 shows the details on the descriptions of the mechanisms above. In order to discover the pattern of metacognitive scaffolding, this study used the data mining approach; a decision tree algorithm.

2.2 Data Mining
Data mining is the process of extracting information from large datasets [13]. It is also known as a tool to uncover hidden information from a large volume of data [13]. The process is done through the use of techniques and algorithms drawn from the data mining software. The algorithms used include clustering, association rule mining and decision tree analysis. Data mining has been widely used across different fields including medicine, business, and marketing & sales. Recently, data mining analysis is also applied on educational data.

The mining of educational data has yielded better outcome in terms of students' performance, course development, students’ retention and may contribute to the development of institutional standard [14]. Usually, the applications are used to predict students’ behavior, performance and examination. Such prediction will help the instructors to pinpoint student’s weaknesses in learning and assist them to achieve better performance. It also serves as a guideline for the instructors to provide students with sufficient support for a meaningful learning.

Other than that, additional useful information mined from the data can help the instructors to detect useful patterns of students’ preferences in learning especially in the online learning environment. Valuable data from the database such as student’s academic background, financial status and online learning participation level (log data from online learning management system) are examples of the useful data that can be used to understand their preferences in learning. Thus, this will help institutions to predict future behaviors of students from a certain area of concern [14][15]. For instance, Kovacic [16] conducted a case study on educational data mining that utilized student’s enrollment data to predict student’s performance in learning.

Data mining has attracted a considerable amount of attention in turning digital data into useful information and knowledge [15]. A number of data mining techniques such as decision tree analysis, association rule mining and clustering analysis have been conducted to perform different knowledge tasks [15]. Due to its proficiency and high accuracy, the decision tree algorithm is ideal for classifying datasets [17]. In this study, we have conducted a decision tree analysis to predict student’s performance based on the pattern of metacognitive scaffolding provided by the instructor in an online learning environment.

3. METHOD
3.1 Participant
A total of 37 postgraduate students from the Educational Technology program participated in the study, out of which 27 (73%) were female and 10 (27%) were male. More than half of the participants (22 participants) aged between 25 and 30 years old, while 13 participants aged less than 40 years old and 2 participants aged 41 years old or more. 27 of the participants hold a Degree in Education, 8 of the participants hold a Degree in Computer Science and 2 of participants hold a Degree in Social Science (Psychology & Counseling).

3.2 Instrument
The instruments selected for the research were online discussion transcripts and students’ scores in a performance test. The online discussion transcripts were extracted from the Facebook group page. The instructor’s postings were coded based on the mechanisms of metacognitive scaffolding by Reingold et al. [12]. Table 1 denotes the mechanisms of metacognitive scaffolding and the examples of the instructor’s posting in the Facebook group page. Students’ scores were taken from the performance test which includes 12 structured questions of topics in learning Adobe Flash software. The rubric used for students’ scores are reported in Table 2.

The content validity of the performance test was performed by the content expert. Besides, a test-retest reliability for the performance test was conducted by running the same test at different points of time. The purpose was to determine the consistency of the test. The test was given to a group of subjects in at least two separate occasions. The correlation between the scores across these two tests would determine its reliability. The test is considered reliable if each student receives similar scores between the first test and the second test. If the scores are high, the tests are considered highly reliable. The correlation coefficient value for test-retest in this study is 0.738 which is considered good.

Table 1. Metacognitive scaffolding and examples of instructor’s posting

<table>
<thead>
<tr>
<th>Metacognitive Scaffolding (MS)</th>
<th>Example of instructor’s posting</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS 1: Presenting rationale for task and activities</td>
<td>“All developed a text entry question in flash and he was using static text as a place where users will type the answer of the question. When he executed the file, the user can see the blinking cursor in a place where they can type the answer.”</td>
</tr>
<tr>
<td>MS 2: Presenting the relationship between reading items, course, objectives and tasks</td>
<td>“How can you relate between the concept of what we have learned before (in Topic 1, 2 and 3) and creating different kinds of applications in flash (in Topic 4)?”</td>
</tr>
<tr>
<td>MS 3: Supporting reflective writing</td>
<td>“Thank you for your opinion, keep up the good work.”</td>
</tr>
<tr>
<td>MS 4: Focusing on the process of learning</td>
<td>“If you increase the number of fps, what will happen to your animation and your file size in byte?”</td>
</tr>
<tr>
<td>MS 5: Encourage relationships among participants</td>
<td>“What are your comments on your friend’s answer?”</td>
</tr>
<tr>
<td>MS 6: Discriminating between conclusion/fact/opinion/hypothesis</td>
<td>“Describe the differences between motion tweening and shape tweening? Your explanation should come with an appropriate example.”</td>
</tr>
</tbody>
</table>
3.3 Procedure and Analysis

The study was conducted for seven weeks. We developed the decision tree using J48 Random Tree classifier using WEKA software. The WEKA software is an open source. It can work efficiently with limited data and also provides convenient data preprocessing, cleaning and handles missing values [14].

The decision tree is able to predict the students’ performance based on the mechanisms of metacognitive scaffolding posted by the instructor. The types of data involved in developing the decision tree include the frequencies of metacognitive scaffolding received by each student, as well as the performance scores categorized into the grading scheme as shown in Table 2.

Table 2. Grading scheme

<table>
<thead>
<tr>
<th>Grade</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>80-100</td>
</tr>
<tr>
<td>B</td>
<td>70-79</td>
</tr>
<tr>
<td>C</td>
<td>60-69</td>
</tr>
<tr>
<td>D</td>
<td>50-59</td>
</tr>
<tr>
<td>E</td>
<td>0-49</td>
</tr>
</tbody>
</table>

4. RESULTS & DISCUSSION

A decision tree which consists of two nodes which are the root node and leaf node was generated. Each branch node represents a choice between a number of alternatives, and each leaf node represents a decision [18]. It starts with a root node which allows the researchers to take actions. From this node, users can split each node recursively based on the decision tree learning algorithm. The final result is a decision tree in which each branch represents a possible scenario of decision and its outcome. Figure 1 denotes the decision tree structure.

![Figure 1. Decision tree structure](image)

The decision tree structure was developed using the data mining software which is WEKA version 3.6.6. The tree model was generated using a full dataset with training set mode. It was constructed using Random Tree classifier. The output shows how the classifier uses the attributes of the instructors’ metacognitive scaffolding (MS) to make decision. Overall, there are a total of 21 nodes derived from this model. The leaf nodes (square box) indicate student’s score in the post-test assessment. The number of brackets indicates the class label. The number of students that gained specific scores followed by the number of students is incorrectly classified as a result. From the output, none of the students are incorrectly classified.

The dataset was split into several attributes with MS4 (mechanisms that allow students to focus on the process of learning) at the top of the tree structure. This shows that MS4 is the highest metacognitive scaffolding prompted by the instructor throughout the study.

The tree indicates that 21 students obtained grade A. 11 students obtained grade B and 5 students obtained grade C in this study. Moreover, from the results mentioned in this study, one can predict the pattern or flow of metacognitive scaffolding utilized by the students to achieve certain grades in their tests. Besides, the certain number of metacognitive scaffolding posts from the instructor is clearly demonstrated.

5. CONCLUSION

Data mining is a powerful tool that enables researchers especially educators and educational institutions to better allocate resources and efficiently manage student outcomes in learning. It is able to uncover hidden information from large datasets and allow educational institutions to achieve better educational standard. As such, the goal of education to serve the nation and inspire the society to acquire knowledge and skills that will benefit them in the future can be reached. This study is intended to enhance the quality of learning.

The decision tree predicts that students who achieved grade A in their study were prone in receiving the combination of guidance that aimed for them to focus on the process of learning (MS4), the assistance of instructor presenting rationale for task and activities (MS1), the guidance from instructor who encourage the relationship among participants (MS5) and supervised text comprehension by the instructor (MS7) such as giving response and feedbacks towards students’ comments. This proved that the combination of these mechanisms affects the students’ performance in learning. It is interesting to note that these mechanisms were the dominant mechanisms of metacognitive scaffolding used by the instructor as reported earlier in the study.

Thus, it is not surprising that most of the students (21 out of 37 students) obtained grade A in the present research. The prediction model is beneficial as it informs the instructor regarding which mechanism of metacognitive scaffolding can be used to assist students in achieving certain grades.

ACKNOWLEDGMENT

This project was funded by Research University Grant and Universiti Teknologi Malaysia (R.J130000.7831.4F755).

REFERENCES


