

# Mining Students' Interaction Data from a System that Support Learning by Reflection

Rajibussalim<sup>1,2</sup>

raji0050@uni.sydney.edu.au

<sup>1</sup>Physics Department, Syiah Kuala University

<sup>2</sup>School of Information Technology, the University of Sydney

**Abstract.** In this paper we utilising some popular educational data mining (EDM) methods to explore and mine educational data resulted from a system that supports reflection for learning called Reflect. Our objective is to study whether using the Reflect helps students learn better in the subject. First, we explore data by means of statistical analyses. Then, we used clustering and classification techniques to identify learning behaviours associated with positive or negative outcomes. The results suggest that there is a positive correlation between reflection and student performance characterised by the level of activities carried out in the Reflect and the exam mark. We claim that our approach has resulted in the identification of some behaviour that leads to the increase in the students' performances.

## 1 Introduction

Literature suggested that reflection defined as “*individuals engage to explore their experiences in order to lead to a new understanding and appreciation*” [1] improves learning. Learning by reflection is considered important since one's learning effectiveness can increase when learners adapt to their learning experiences by reflecting on their learning process and the state of their knowledge [1, 2, 3]. The same researchers also suggested that self-assessment helps learners to develop the ability to identify their strengths and weaknesses and focus their study to the area that is need for improvement. However, most of the existing work that suggested this theory was not utilising education data mining (EDM) methods to draw their conclusion. In this paper, we explore the use of EDM methods on data gathered from an online learning system that support learning by reflection called Reflect (more discussion on Reflect presented in Section 2). In particular our goal was to evaluate the effectiveness of EDM methods for: (1) extracting knowledge about the impact of reflection on learning, (2) gaining knowledge about students' learning behaviour, and (3) identifying which behavioural patterns lead to positive or negative outcomes. This knowledge can be useful for teachers to better understand their students' learning behaviour and to inform students if their current behaviour was associated with negative outcomes in the past. The information can also be benefited if presented to students in which they can reflect on how effective their learning habits are against their current performances.

In section 2, we describe Reflect and how it supports student's self-assessment, followed by section 3 that outlines the subjects and data. Section 4 presents data pre-processing, the main data exploration, application of two educational data mining methods: clustering and classification techniques, we present the results and discussion. Section 5 summarises our conclusion.

## **2 Reflect**

The Reflect system is a web-based educational learning system used for learning, practise and testing one's programming knowledge. This system is not intended as a sole mean for learning a topic but it serves as an additional learning medium to face to face teaching. The Reflect system was used by students as an alternative source of learning and practices of the topic Software Construction I (SOFT2130/2830) at School of IT, the University of Sydney. Reflect can offers a great flexibility in learning experience. It offers extra attributes that differ and not commonly found in other LMSs or ITSs such as when students assess an example solution, they can view the tutor's assessment that are comparable against the students' assessment. It will show a discrepancy table if the student's assessment differ from the tutor's ones. Students may also view their progress in a graph showing how close they are getting to the tutor's assessment.

The Reflect system is unique because it aims to promote learner reflection and student self-assessment through scrutable learner model that shows the learner a model of their progress [4]. Kay et al. [4] suggested that there are two key elements of self-assessment: (1) identification of criteria and standards to be applied to their own work and (2) evaluation of their work compared to these criteria and standards. In Reflect, these two key elements of self-assessment were carried out by students when they performed the following steps in order to solve the problems:

1. Students read a task and assess example solutions made available in the system.
2. Students provide their own solutions to solve the problem and
3. They self-assess those solutions compared to the criteria teachers defined for the task.

By performing these steps, the students had practised to self-assess themselves and their state of knowledge of a particular task. Thus, they carried out learning by reflecting to their learning experiences.

## **3 Subject and Data**

### ***3.1 Subjects***

The subjects used in this research are the students who undertook the Software Construction I (SOFT2130/2830) course at School of Information Technology, the University of Sydney in the second semester of 2007. There were 175 students enrolled for the course, however only 156 (89.1%) students managed to completed it. Out of 156 students who completed the course, 109 (69.9%) students passed and 47 (30.1%) of them failed the course.

### ***3.2 Data***

The data used in this research are coming from two sources: (1) students' activities data gathered from Reflect system and (2) students assessment data gained from the lecture. In experiments both of data were used. Student activity data are gathered from Reflect database in the form of XML format. The data are organised in a hierarchical structure i.e. each task consists of a task name and a number of learning objectives. The numbers

of learning objectives for each task are varying from one to five learning objectives. Moreover, a learning objective may appear in one or more task. The student assessment data are spread over several worksheets of Microsoft Excel format. These include lecture quiz marks, homework and labs marks (including practical exam marks on week 4, 8 and 12), quizzes marks, and the final exam mark of students who enrolled for Software Construction I topic for the year of 2007. These data are obtained from the lecturer who taught and organised the course.

## **4 Process of EDM: Results and Discussion**

### ***4.1 Data pre-processing***

Data pre-processing tasks consist of cleaning the data from incomplete or inconsistency data and errors and also integrating and transforming the data to an appropriate format for mining.

#### *Cleaning the Data*

After data extracted from the Reflect database in XML format, the next step is to clean the data from inconsistencies and errors before they are ready to be mined. We did the process of data cleansing manually by searching for incomplete data or errors and removed them from database. The errors may come from unintended users who did not enrol for the topic. These users include Reflect system administrator who used the system for testing and other students who did not enrol for the course but used the system for learning and practising their C programming skills. The error may also come from a user who had a double user logins. This could happen when a student changed their user login at some points during the semester.

#### *Integrating and Transforming the Data*

Data integration is a process of integrating or merging the data from multiple tables or databases into a coherent data store [5]. Educational data mining processes often involve retrieving and analysing multiple data attributes that spread across several tables or databases. The data, therefore, needs to be summarised into a new summarisation table consisted all necessary attributes for mining. As mentioned earlier in Section 3.2, our data sources are coming from: the students' activity data gathered from the Reflect database and students assessment spreadsheet obtained from the course coordinator. To perform the data mining out of these data, we are required to merge necessary data attributes into a summary table. Table 1 shows the data attributes of this summary table.

Data transformation may involve a number of techniques including smoothing i.e. to remove noise from the data; aggregation i.e. summary operation applied to the data; generalisation where the low level data are replaced by higher-level concepts and normalisation where the attributes data are scaled thus they fall within a small range i.e. 0.0 to 1.0. [5]. In regards to the Reflect data, we used a smoothing technique to remove a noise from a student with an excessive self assessment input. We also used aggregation and normalisation techniques to smooth our data. We used aggregation to summarise the

weekly Homework/Lab marks and quiz marks into Total Homework/Lab and Total quizzes marks, while we used normalisation to normalise the attribute values of Total Lecture quizzes and Lecture quiz attendances. Finally, we transformed our data into ARFF and csv for mining with Weka.

**Table 1. List of data attributes in summary table**

Attribute Name	Description
L2_mem_diagram	Week 2 lecture quiz mark (numeric)
L5_pointers	Week 5 lecture quiz mark (numeric)
L8_scope	Week 8 lecture quiz mark (numeric)
Tot_Lec_qz	Total lecture quiz marks (%)
Lec_qz_attd	Total lecture quiz attendances (%)
Tot_HWL	Total homework/lab marks (%)
HWL_attd	Total homework/lab attendances (%)
Tot_qz	Total quiz marks (%)
qz_attd	Total quiz attendances (%)
Exam	Final exam mark (numeric)
exam_cate	Categorical final exam mark (ordinal)
n_task	Number of tasks done in the Reflect (numeric)
n_lo	Number of learning objectives done in the Reflect (numeric)

### *Constructing Dataset for Experiments*

From the list of data attributes showed in Table 1, we constructed two sets of data: (1) a data set that consist only numerical data and percentages, and (2) a data set that consist of numerical data, percentages and categorical final exam mark. The first data set is used for statistical and clustering analyses while the second data set is mainly used for classification analyses.

## **4.2 Data Exploration**

The statistical analyses are often providing a starting point for data analyses. Therefore we carried out a number of correlation analyses to study the relationships and to measure if one data attribute is significantly correlated to another. In addition, a statistical analysis can be used to determine which variable is best explain the differences between two or more groups, that is a variable that can distinguish one group from another. Our objective is to utilise statistical analyses to find the relationships and correlation analyses between: (1) lecture quizzes score and final exam mark, and (2) lecture quizzes attendances and final exam mark.

### *Discussion*

Statistical analyses suggested that although there were positive correlations for both attending (Lec\_qz\_attd) and performing quizzes (Tot\_Lec\_qz) set by the teachers in the lectures and the final exam mark, these correlation are not considered high enough. The

correlation score ( $r$ ) between Lec\_qz\_attd and exam is only 0.462 ( $p < 0.01$ ) and correlation ( $r$ ) between Tot\_Lec\_qz and exam mark is 0.413 ( $p < 0.01$ ).

Meanwhile, another analyses suggested that correlation score between Total Homework/Lab (Tot\_HWL) and exam is reasonably high ( $r = 0.618$ ;  $p < 0.01$ ). However, the correlation between Homework/Lab attendances (HWL\_attd) and exam is not high enough ( $r = 0.354$ ;  $p < 0.01$ ).

The results indicated that Homework/Lab performance is considered significantly important toward achieving a good mark in the final exam. However, this is not necessarily the case for lecture quiz and its attendances as their correlation with the exam mark are not significantly high. Nevertheless, the results showed positive correlations that indicated the importance of attending both and performing well in both lecture quizzes and Homework/Lab activities.

### ***4.3 Clustering Students***

Clustering is an unsupervised classification used for grouping objects into classes of similar objects [6]. A number of researchers have implemented clustering techniques for mining e-learning data with various purposes such as for finding groups of students who have similar learning characteristics, to encourage group-based collaborative learning and to provide incremental learner diagnosis [7].

Our work utilised a K-means clustering algorithms to cluster group of students based on their similar behaviour in using the Reflect system. We choose K-means clustering algorithm because it is considered as one of the most popular and mostly used clustering algorithm in broad data mining research community [8]. Another reason for choosing K-means is because it is available and ready to use in Weka system.

#### *Lecture quizzes versus final exam score*

We utilised both numerical data and a combination of numerical and nominal data set. Before running the K-means algorithm, the data set are transformed into comma separated value (csv) format. The reason is because, a csv file format is easy to use and it is one of file format acceptable in the Weka system. Here, our objectives are first, to distinguish stronger group of students from the weaker ones and then to study learning characteristics that differentiate each group. To achieve this objective, we utilised numerical data set with categorical exam mark.

#### *Discussion*

Clustering analyses are able to provide more detailed information about students beyond what simple statistical analyses may offer. Based on the clustering results, we are able to categorise students into several clusters based on their performances and attendances in the lecture quizzes. Each group is characterised by how often their attended the lecture quizzes and their performance on those quizzes.

**Table 2. Student cluster using K-means (k=4)**

Cluster size and (percentages)	Lecture quizzes Mark	Lecture quizzes attendance	Exam Mark*	Cluster label
52 students (33%)	Average High mark of all lecture quizzes	Very high % of lecture quizzes attendances (99%)	High Moderate (highest among other clusters) exam	Strong group
14 students (9%)	Average Low mark of the lecture quizzes	High % of lecture quizzes attendances	Moderate exam mark	Moderate group
47 students (30%)	Average Low mark of the lecture quizzes	Medium % of lecture quizzes attendances	Low exam mark	Lower Moderate group
43 students (28%)	Lecture quiz marks not available	Very low % of lecture quizzes attendances (15%).	Very Low (almost failed) exam mark	Weak group

\*Exam mark: fail < 40, Low >=40 and <55; Moderate >=55 and <70; Good >=70 and < 80; Very Good >=80.

The results summarised in the Table 2, suggested that 33% of students who achieved higher exam mark had at least average high mark on lecture quizzes and its attendances. In the other hand, 28% students who achieved very unsatisfactory (very low) exam results had not attended enough lecture quizzes (very low attendances).

These results highlighted the important of every lecture quiz toward the increasing of students' understanding to the topic. In other words, attending and performing tasks and exercises in the lecture quizzes may have a direct impact on student's knowledge of the topic and may affect their performance in the exam. This is maybe because attending and completing tasks and exercises in the "exam-like" environment, such as in a lecture quiz, during the lectures may make the student familiar and well prepared with the type of question being asked in the final exam. Meanwhile, the students with poor class attendances would miss the opportunity to familiar themselves to the type of question that may appear in the exam. This result supports our hypothesis that "*students who attended and have good marks on all lecture quizzes (week 2, 5 and 8) performed well in the final exam*".

#### **4.4 Classifications**

Classification is a supervised classification in which the labels of pre-classified patterns are identified. This pre-classified pattern is known as training data set. Within the training data set there is a class attribute that will be used to label a newly encountered, still unlabelled pattern [HanK06\_Data\_mining]. Our approach is to use a well known C4.5 (J48) algorithm from Weka data mining tools. The J48 algorithms would generate decision trees that might be used to extract classification rules. In our case, the objective is to classify students into different groups or branches with almost equals final exam mark. The model resulted from the experiments can be used to predict the final exam

mark of new students. . The J48 decision tree algorithm required that the label in this case is exam mark, must be in the categorical or nominal form. We, therefore, grouped the exam mark into the following category: Fail if the exam mark is  $<40$ ; Low if the exam mark  $\geq 40$  and  $<55$ ; Moderate if the exam mark  $\geq 55$  and  $<70$ ; Good if the exam mark  $\geq 70$  and  $<80$ ; and Very Good if the exam mark  $\geq 80$ . This categorisation is different to the categorisation of the final grade students received at the end of semester. The final grade which is determined by the course coordinator at the end of the semester integrated not only the final exam mark but also other assessment components including quizzes marks, homework/lab marks and assignments scores. The categorisation for the final exam mark is fixed by the course coordinator and agreed by students at the beginning of the course. Some rules were generated from classification analyses. These rules are summarised as follow:

```

IF (Tot_qz <= 0.67) AND (n_task <= 6) THEN exam = Fail
IF (Tot_qz <= 0.67) AND (n_task > 6) AND (HWL_attd <= 0.8) AND
  (Lec_qz_attd <= 0.33 ) THEN exam = Low
IF (Tot_qz <= 0.67) AND (n_task > 6) AND (HWL_attd <= 0.8) AND
  (Lec_qz_attd > 0.33 ) THEN exam = Moderate
IF (Tot_qz between 0.67 and 0.96) AND (n_lo <= 12) THEN exam = Low
IF (Tot_qz between 0.67 and 0.96) AND (n_lo > 12) THEN exam =
  Moderate.
IF (Tot_qz >= 0.96) THEN exam = Good

```

### *Discussion*

The results suggested that the total quizzes mark (Tot\_qz) was the main discriminator of the final marks followed by the number of task students done in the Reflect system. The rules generated by J48 algorithm revealed the characteristics of each group of students. For example, a student should at least complete six tasks or more in the Reflect system and achieved more that 67% of total quizzes mark to reduce the risk of fail in the final exam. Meanwhile a student who can achieved 96% of the total quizzes marks are directly classified as Good, meaning he/she will be among students who are likely to achieve a good mark in their final exam. The other groups of students are classified based on some other activities including homework/lab (HWL\_attd) and lecture quiz attendances (Lec\_qz\_attd).

As we can see from the results, there are set of critical separation point for each class, for example 67% or 96% of total quizzes, 6 or 12 tasks done in Reflect, 33% of lecture quiz attendances (Lec\_qz\_attd) and 80% of homework/lab attendances (HWL\_attd). These separation points are set automatically by J48 classifier.

Using J48 algorithm, our first experiment did not produce a good result. The accuracy level recorded for the first experiment is only 51.28%. This means that out of 156 data, only 80 instances were correctly classified into their classes. Other 76 instances were incorrectly labelled into the other classes.

In the second experiments, the accuracy level increased although not much. The accuracy level for the classifier model became 56.41%. This means that out of 156 instances, 88 of

them are correctly classified into their classes while remaining of 68 instances were still not correctly classify or labelled into their classes.

This information would be very useful for a course coordinator for example the classifier can be used to predict the final exam marks of new students or to promote some types of activities to obtain higher marks.

## 5 Conclusions

We have demonstrated how EDM methods have been utilised to extract some valuable information from the Reflect data. We have discussed how the users of Reflect can perform self-assessment by following certain procedures as described in the section 2. These include to rate their understanding of each learning objectives related to the task, providing answers to the tasks, self assess their answers and sample solution against certain criteria defined by the teacher and comparing the discrepancies between their answers to the teacher's assessment.

A number of issues have emerged during the study either related to the data or the interpretation of the results. First, the present study only used one semester data, hence the analyses is limited. For the further work, it would be more interesting if data mining is conducted for data that have been collected for many years. Secondly, the current data used was not designed in the first place to be suitable for mining. Therefore, the data are complex, contain noise, and heterogeneous.

The results of EDM analyses suggested that the Reflect system helps users to self assess themselves and thus satisfied the aims indentified in the introduction of section that is students learnt more by using system that support learning by reflection such as the Reflect system.

## References

- [1] Boud, D., and Keogh, R. Promoting reflection in learning: A model, Reflection: turning experience into learning, 1985, pp. 18-40.
- [2] Schon, D. A. The reflective practitioner, New York, 1083.
- [3] Schon, D.A. Educating the reflective practitioner, San Francisco, 1987.
- [4] Kay, J., Li, L., and Fekete, A. Learner Reflection in Student Self-assessment, *Proceedings of ACE 2007, 9th Australasian Computing Education Conference*, Australian Computer Society, Ballarat, Victoria, 2007, pp. 89-95.
- [5] Han, J., and Kamber, M. Data Mining Concept and Techniques, Morgan Kaufman, San Francisco, 2006.
- [6] Witten, I.H., and Frank, E. Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufmann, 1999.
- [7] Tang, T., and McCalla, G. Smart Recommendation for an Evolving e-Learning System: Architecture and Experiment. *International Journal on e-Learning* 4, 2005. pp. 105-129.
- [8] Romero, C., Ventura, S., and Garcia, E., Data mining in course management systems: Moodle case study and tutorial. *Computers & Education in Press*, Corrected Proof 2007.