Gamification for blended learning in higher education

Mewati Ayub, Hapnes Toba, Maresha C. Wijanto, Steven Yong & Bryan Wijaya

Maranatha Christian University
Bandung, West Java, Indonesia

ABSTRACT: The main objective of this study was to evaluate the use of gamification in an extended learning management system (LMS). Tournament and leaderboards are the gamification features that were implemented in the LMS. As case studies, a number of experiments were performed in an informatics Bachelor programme subject, i.e. Introduction to Programming (IP). Education data mining (EDM) techniques were used to extract from the experiments compelling patterns in the form of association rules and decision tree. Analysis of the technical evaluation results and the questionnaire shows that students need learning activity outside the classroom to sharpen and deepen the learning materials given during college hours.

INTRODUCTION

Some benefits of blended learning are an increase in learning effectiveness, convenience and access [1][2]. The aim of this research was to explore educational data mining (EDM) [3-6], by using an extended learning management system (LMS) at the Faculty of Information Technology, Maranatha Christian University in Indonesia. The Faculty has a blended learning system with full face-to-face instruction. In this case, the LMS was extended to accommodate gamification techniques in blended learning.

In this study, the authors discuss the results of gamification as an extension of previous work [7][8]. The work aims to increase students' motivation in their learning through blended learning in the classroom. Gamification methods implemented in the extended LMS are tournament and leaderboard features. The extended LMS was evaluated in the Introduction to Programming (IP) course taught during the first semester of the 2018 academic year.

SYSTEM DEVELOPMENT AND RESEARCH METHOD

To enhance the learning system, gamification was implemented as extended features to support blended learning [7][8]. In this study, the focus was on tournament and leaderboard features. The experiment was performed over 11 weeks of the IP course. During the experiment, each student had to take on-line quizzes in and outside the classroom using the extended LMS.

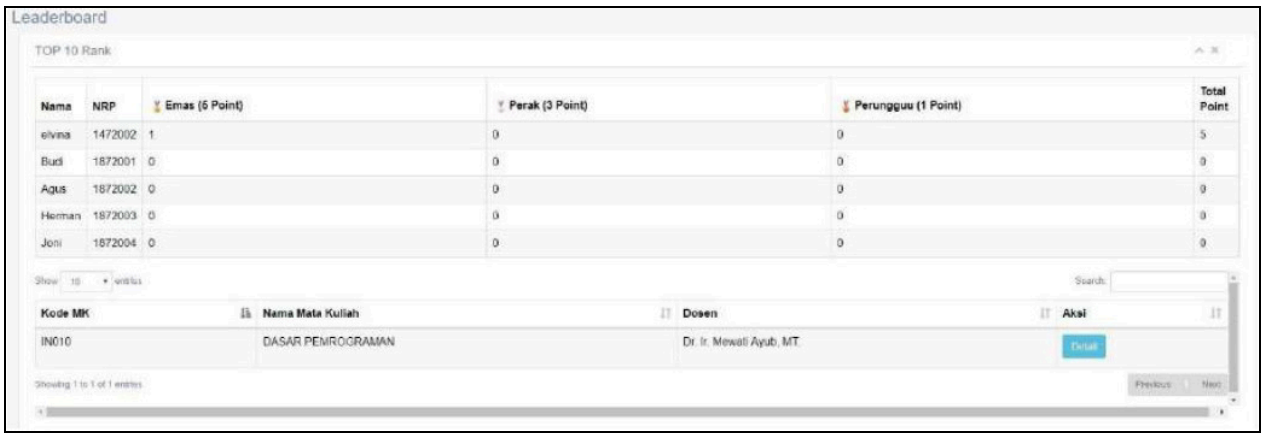

Figure 1: An example of the leaderboard.

On-line quizzes were run as a tournament and the result of each tournament updated the course leaderboard. There were four quizzes during class and three outside the classroom. The on-line quizzes started in the third week and ended in the sixth. The quizzes outside the classroom were in the weeks following the classroom quizzes. Besides on-line quizzes, there were two written quizzes in the fourth and seventh week.

A mid-semester examination was run at the eleventh week. An example of the leaderboard user interface is shown in Figure 1. The students' data attributes are shown in Table 2. In this study, the data were extracted from 55 students, who had a minimum 75 percent attendance for the IP course. The statistics of the students' grades are presented in Table 1.

Table 1: The mean and standard deviation of students' score.

	Written quiz	Mid-semester examination	On-line quiz in class	On-line quiz outside class
Mean	60.89	62.35	61.36	70.13
SD	28.02	24.71	17.05	13.77

Table 2: Students' attributes data set.

Attribute name	Description	Possible values
GradeWQ	Grade of written quiz	(Excellent, Good, BelowAvg)
GradeOQ_in	Grade of on-line quiz in class	(Excellent, Good, BelowAvg)
GradeOQ_out	Grade of on-line quiz outside the class	(Excellent, Good, BelowAvg)
ActivityQ	Activity level in on-line quiz	(High, Medium, Low)
GradeMid	Mid-semester grade	(Excellent, Good, BelowAvg)

The data were analysed by using association rules and classification techniques [6][9]. Association rule mining was used to obtain general rules which indicate the contribution of on-line quizzes (as tournaments) during the learning process. Classification techniques, in the form of a decision tree, were used to analyse the robustness of the generic rules [4][8][9].

RESULTS AND DISCUSSIONS

In the association rule mining, the minimum support was set as 0.1, the minimum confidence was 0.8 and the lift had to be greater than 1.0. Shown in Table 3 is a set of extracted rules using the data attributes from Table 2. The rules indicate a strong relationship between the results of written and on-line quizzes. Rule numbers 1 to 3 show that BelowAvg grades of the mid-semester were determined by the results of the students' written and on-line quizzes.

Table 3: Extracted rules for mid-semester grades in the IP course.

No.	Association rules	Support	Confidence	Lift
1	GradeWQ=BelowAvg, GradeOQ_in=BelowAvg ==> GradeMid=BelowAvg	0.25	0.93	2.27
2	GradeOQ_in=BelowAvg, GradeOQ_out=Good ==> GradeWQ=BelowAvg	0.16	0.9	2.15
3	GradeOQ_in=BelowAvg ==> GradeWQ=BelowAvg	0.27	0.83	1.99
4	GradeWQ=Good, GradeOQ_in=Good, GradeOQ_out=Good ==> GradeMid=Good	0.15	0.89	2.44
5	GradeWQ=Good, GradeOQ_out=Good ==> GradeOQ_in=Good	0.16	0.82	1.55
6	GradeWQ=Good, GradeOQ_out=Good ==> GradeMid=Good	0.16	0.82	2.25
7	GradeMid = Excellent ==> GradeWQ=Excellent	0.24	0.81	2.35

Similarly rules number 4 to 6 reveal that Good grades of the mid-semester are also determined by written quiz and on-line quizzes. Those rules are also confirmed by rule number 7, which says that if the grade of the mid-semester is Excellent, then the grade of the written quiz is Excellent, but in this case the on-line quizzes have no significant contribution. This suggests that students, who have excellent academic records, have a stronger intention to display their abilities in lectures.

Explored further in this study is the relationship between mid-semester grades and other attributes using J48 classification with tenfold cross validation. The classification was used to derive general rules from the data set to indicate whether on-line quizzes affect the students' mid-semester grades. The first classification involved all of the attributes of the data set, and the second classification involved only three attributes, i.e. grade of the on-line quiz in class, grade of the on-line quiz outside the class and grade of the mid-semester examination.

The tree for the first classification shown in Figure 2 indicates that the most effective attribute in predicting the grade of the mid-semester was the grade of the written quiz. The accuracy of this classification was 69.09%; there were 38 correctly classified instances and 17 incorrectly classified instances. Table 4 summarises the rules generated from the tree in Figure 2.

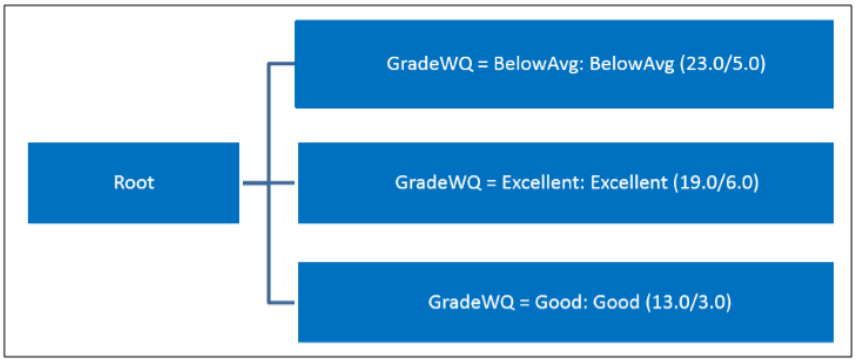

Figure 2: J48 tree for the first classification.

Table 4: Rules for mid-semester grades from the first classification.

Rule no.	Rule's premise	GradeMid Percentages of instances		
		Excellent	Good	BelowAvg
1	IF GradeWQ = BelowAvg	-	-	78.26%
2	IF GradeWQ = Excellent	68.42%	-	-
3	IF GradeWQ = Good	-	76.92%	-

In Figure 2, the tree indicates that there are five instances, which have BelowAvg grade of written quiz, but the grades of mid semester are not BelowAvg.

Shown in Figure 3 is the decision tree for the second classification. The J48 tree indicates that the grades of on-line quizzes in the classroom also affected the mid-semester grades. For those who achieve a good grade for on-line quizzes in the classroom, the results of their on-line quizzes outside the classroom will also contribute to their mid-semester grades. The accuracy of this classification was 67.27%; there were 37 correctly classified instances and 18 incorrectly classified instances.

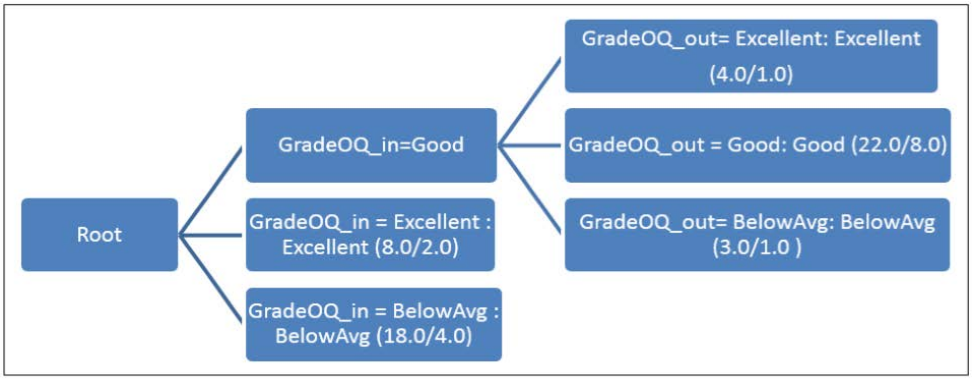

Figure 3: J48 Tree of the second classification.

Summarised in Table 5 are the rules generated from the tree in Figure 3.

Table 5: Rules for the mid-semester grades from the second classification.

Rule no.	Rule's premise	GradeMid Percentages of instances		
		Excellent	Good	BelowAvg
1	IF GradeOQ_in = Good and GradeOQ_out = Excellent	75%	-	-
2	IF GradeOQ_in = Good and GradeOQ_out = Good	-	63.63%	-
3	IF GradeOQ_in = Good and GradeOQ_out = BelowAvg	-	-	66.67%
4	IF GradeOQ_in = Excellent	75%	-	-
5	IF GradeOQ_in = BelowAvg	-	-	77.78%

Revealed in Table 6 are five instances in Figure 2; although they have BelowAvg grades on the written quiz, the grades on the mid-semester examination are Good or Excellent. Also indicated in Table 6 is that they have Good grades in the on-line quiz and had High or Medium activity. These outliers suggest that the students used the on-line system to achieve better results in their studies.

Table 6: Outliers of BelowAvg grades on the written quiz.

StudentID	GradeWQ	GradeMid	GradeOQ_in	GradeOQ_out	Activity
1872001	BelowAvg	Good	Good	Good	High
1872018	BelowAvg	Excellent	Good	Good	Medium
1872022	BelowAvg	Good	Good	Good	High
1872027	BelowAvg	Good	BelowAvg	Good	High
1872048	BelowAvg	Good	Good	Good	High

Listed in Table 7 are the top ten students in the leaderboard. All of the students had Good or Excellent in mid-semester, written quizzes and on-line quizzes.

Table 7: Top tenth of leaderboard.

StudentID	GradeWQ	GradeMid	GradeOQ_in	GradeOQ_out	Activity
1872002	Excellent	Excellent	Excellent	Good	High
1872004	Excellent	Excellent	Excellent	Excellent	Medium
1872006	Good	Good	Good	Good	High
1872015	Good	Good	Excellent	Excellent	High
1872020	Excellent	Excellent	Excellent	Excellent	High
1872025	Excellent	Excellent	Excellent	Good	High
1872035	Excellent	Excellent	Good	Excellent	High
1872049	Excellent	Excellent	Good	Good	High
1872057	Excellent	Good	Excellent	Good	Medium
1872061	Excellent	Excellent	Good	Good	High

IMPACT OF THE LMS EXTENDED FEATURES

At the end of the technical evaluation, the students were given a questionnaire to evaluate the impact of the new system in a blended learning situation. The questions are shown in Table 8. Questions 1 to 15 used a 1-5 Likert scale, where 1 means strongly disagree, 2 disagree, 3 neutral, 4 agree and 5 strongly agree.

Table 8: Questionnaire to evaluate the impact of blended learning and the extended LMS.

ID	Statement
Q1	On-line quiz in class helps a student to understand the theory given in class
Q2	On-line quiz in class helps a student to understand a simple program
Q3	On-line quiz in class helps a student to understand the conditional statement
Q4	On-line quiz in class helps a student to understand the loop (while or for)
Q5	On-line quiz in class helps a student to write a computer program
Q6	On-line quiz outside the class helps a student to understand the theory given in class
Q7	On-line quiz outside the class helps a student to understand a simple program
Q8	On-line quiz outside the class helps a student to understand the conditional statement
Q9	On-line quiz outside the class helps a student to understand the loop (while or for)
Q10	On-line quiz outside the class helps a student to write a computer program
Q11	Features in the system helps a student to read the quiz
Q12	Features in the system helps a student in answering the quiz
Q13	Features in the system helps a student in checking the answers to the quiz
Q14	Features in the system helps a student in viewing the results of the quiz
Q15	The system supports the learning of Introduction to Programming in and outside the classroom

The questions can be divided into three sections, viz. Q1 - Q5 are intended to evaluate the quizzes in the classroom; Q6 - Q10 to evaluate the quizzes outside the classroom; and Q11 - Q15 to evaluate the extended LMS. In Table 9 are the means and standard deviations for the three sections. The detailed response of each question can be seen in Figure 4.

Responses to all questions had a greater value than 4 (agree), which suggests that students mostly are in favour of the blended learning process and appreciate activities outside the classroom. The students also gave specific comments for

Q15 that indicate most of them are strongly motivated to compete with each other. The competition in the leaderboard gives extra motivation to completing the tasks and quizzes, either in or out of the classroom.

Table 9: The means and standard deviations for the questionnaire.

Questions	Evaluation section	Mean	SD
Q1 - Q5	Quizzes in the classroom	4.12	0.76
Q6 - Q10	Quizzes outside the classroom	4.04	0.90
Q10 - Q15	Blended learning system in LMS	4.06	0.86

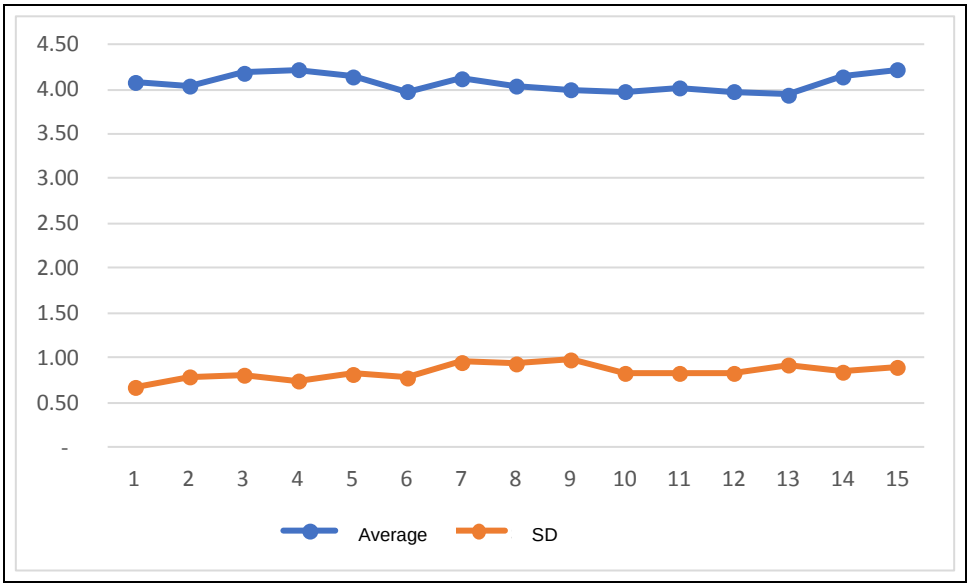

Figure 4: The results of the questionnaire.

In general the information in Figure 4 shows that the students have great enthusiasm for exploring blended learning. Some improvements that the students have suggested are directed at the enhancement of the gamification features, viz. to give more explanation on the use of video for difficult topics, to enhance the scoring system for difficult questions to ensure the fairness of the competition, and deploying an inquiry-based system for communication between the lecturer and the students [10].

CONCLUSIONS

In this research, the evaluation of a blended learning system was explored in the Introduction to Programming course. The evaluation showed that students need more learning activity outside the classroom to sharpen and deepen the understanding of the learning material delivered during college hours. Used in the study were decision trees for classification and questionnaires. It is important to develop an LMS to encourage students' enthusiasm toward undertaking extra efforts outside the classroom.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support provided by the Directorate General of Research and Development Strengthening in the Ministry of Research, Technology and Higher Education of the Republic of Indonesia, under the Research Grant number 0815/K4/KM/2018.

REFERENCES

1. Graham, C.R., *Blended Learning Models*. In: Encyclopedia of Information Science and Technology. Hershey: PA: Idea Group Inc., 375-383 (2009).
2. Pankin, J., Roberts, J. and Savio, M., *Blended Learning at MIT*. Massachusetts Institute of Technology Repository (2015).
3. Baker, R. and Yacef, K., The state of educational data mining in 2009: a review and future visions. *J. of Educational Data Mining*, 1, **1**, 3-16 (2009).
4. Romero, C. and Ventura, S., Educational data mining: a survey from 1995 to 2005. *Expert System with Applications*, 33, **1**, 135-146 (2007).
5. Romero, C., Ventura, S. and Garcia, E., Data mining in course management systems: Moodle case study and tutorial. *Computers and Educ.*, 51, **1**, 368-384 (2008).

6. Romero, C. and Ventura, S., Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3, **1**, 12-27 (2013).
7. Ayub, M., Toba, H., Wijanto, M.C. and Yong, S., Modelling online assessment in management subjects through educational data mining. *Proc. Inter. Conf. on Data and Software Engng*,, Palembang (2017).
8. Ayub, M., Toba, H., Yong, S. and Wijanto, M.C., Modelling students' activities in programming subjects through educational data mining. *Global J. of Engng. Educ.*, 19, **3**, 249-255 (2017).
9. Han, J., Kamber, M. and Pei, J., *Data Mining Concepts and Techniques.* Waltham: Elsevier, Inc., 264-266 (2012).
10. Hrast, Š. and Ferk Savec, V., ICT-supported inquiry-based learning. *World Trans. on Engng. and Technol. Educ.*, 16, **4**, 398-403 (2018).