Predicting Student Success according to Online Activities in a Blended Course using Artificial Neural Networks

Murat Karakaya[1], Meltem Eryilmaz[2], Atila Bostan[3]

[1, 2,3]Department of Computer Engineering, Atilim University,

Ankara, TURKEY

[1]murat.karakaya@atilim.edu.tr, [2] meltem.eryilmaz@atilim.edu.tr , [3]atila.bostan@atilim.edu.tr

ABSTRACT

In a blended course, some portion of the classes is held as the traditional face-to-face approach whereas the rest is conducted as a web-based online learning approach. In this paper, we focus on a selected blended course in order to observe the effects of online activities on the final success of students in that course. We opt to calculate the success of a student in a course with two metrics. The first success metric is the letter grade and the second is if the student passes or fails. We would like to predict the student success by applying an artificial neural network (ANN) model. In the model, we provided different set of features from the collected features

The experiment results indicate that predicting grade letters is much more prone to errors than predicting the pass/fall result. The tests show that we can predict if a student pass or fail with about 81% accuracy considering only the number of online activities. These results indicate that, in a blended course, student success is not only determined by the quantity of online activities but also it might be related with the quality of the face-to-face interaction with the instructor.

Keywords – artificial neural network (ANN), Prediction, student success, blended course

1. INTRODUCTION

As the mobile platforms and applications are getting popular in our daily lives, the education approaches aim to utilize these opportunities as well to improve learning experiences. One of the novel approaches in education is providing classes as a blended course. In a blended course, some portion of the classes is held as the traditional face-to-face approach whereas the rest is conducted as a web-based online learning approach. Blended courses are also named as hybrid or mixed-mode courses.

In this paper, we focus on a selected blended course in order to observe the effects of online activities on the final success of students in that course. Thus, we collected the online activity logs of 152 students from the used Learning Management System (LMS). Then, we determine the number of all these activities under three important topics: Course module viewed, Discussion module viewed, Quiz module viewed. Moreover, we added the demographic information about the students such as sex and department.

We opt to calculate the success of a student in a course with two metrics. The first success metric is the letter grade and the second is if the student passes or fails. We would like to predict the student success by applying an artificial neural network (ANN) model. In the model, we provided different set of features from the collected features, that is, number of Course module viewed, Discussion module viewed, Quiz module viewed, and demographic information. The target values to be predicted are the success metrics.

The paper is organized as follows. In the following section, we provided a summary of the related work. Section 3 explains how the data are collected and processed. Conclusion section presents the implications and conclusions which can be drawn from the study.

2. RELATED WORK

Blended Learning can be described as "*a learning program where more than one delivery mode is being used with the objective of optimizing the learning outcome and cost of program delivery*" [1]. These delivery modes can be face-to-face and computer-based learning methods. Thus, teachers can use a variety of instructional techniques.

As in any learning methodology, Blended Learning (BL) has its own benefits and challenges for students and teachers [2, 4]. The benefits for the students can be time flexibility, rich and multimedia contents among the others. Teachers can also benefit from Blended Learning having enhanced students-teacher interaction, increased student involvement, and ease to monitor student improvement. The challenges for students can be time management, self-discipline, responsibility, engagement, etc. Similarly, for the instructors, there are some difficulties such as keeping up with the technology, time commitment, losing control over class, etc. In order to manage Blended Learning successfully, instructors should apply best practices by combining current technology and the teaching methodology [6].

As there are risks in Blended Learning (BL), one major question is whether it can increase the success of learning/teaching considerably [3, 5]. To evaluate the success of the blended courses, one might need to use a series of techniques to study the usefulness of the BL environment and its key components. In [3], the authors applied a series of three online questionnaires and face-to-face interviews which were given to the students, at the beginning, middle and end of the semester.

In our previous work, we applied a satisfaction metric to evaluate if integrating a social networking platform into a blended course can increase the overall satisfaction of the students [8]. In this work, we do not aim to apply questionnaires and face-to-face interviews for observing the students satisfaction or evaluating the success of the applied BL. On contrary, we opted to examine direct measurements of student's interactions with the Learning Management System (LMS) used in BL with respect to student success metrics. We believe that one of the acceptance and the key elements determining success of a blended learning implementation is the e-learning component, i.e. LMS. Therefore, we observe the interactions of the students with the LMS.

3. DATA COLLECTION

In this paper, we focus on a selected blended course given in Atılım University Turkey. The selected course is Introduction to Computers and Information Systems given to engineering students in 4 different sections by the same instructor during 2017 spring semester [9]. The Learning Management System (LMS) used in this course is Course Networking (CN) system [7].

In order to observe the effects of online activities on the final success of students in that course, we collected the online activity logs of 152 students from the CN system. Then, we determine the number of all these activities under three important topics: Course module viewed, Discussion module viewed, Quiz module viewed. Moreover, we added the demographic information about the students: sex and department.

The number of "Course module viewed" activity is the number of a student visiting the online materials shared on the CN system. The "Discussion module viewed" activity is counted when a student either starts a discussion or visits a discussion forum. The last activity "Quiz module viewed" indicates how many times a student attempts to take quizzes. We also sum up all the numbers of these activities as "Total logged-on" which can indicate the total number of student interactions with LMS.

4. PREDICTION METHOD AND RESULTS

4.1 Succes Metrics

As discussed above, our goal is to observe if we can predict the student success by using the number of their activities on the LMS. We first define the student success by using two metrics. The first one is the grade letter that a student received at the end of the semester. The grade letters used at Atılım University are AA, BA, BB, CB, CC, DC, DD, FD, and FF.

The second success metric is if the student passes or fails from a course. According to the established rules at Atılım University, when a student receives FD or FF grade letter, he or she is accepted as "failed", otherwise as "passed" the course.

4.2 Artificial Neural Network Model

For prediction method, we opt to use Artificial Neural Network (ANN). ANN is a well-known pattern recognition and classification algorithm. We specifically use the MATLAB tool as the development environment.

During experiments, 70% of the data is selected as the training data; the rest of data is equally distributed as test and validation data. The ANN model consists of 1 hidden layer with 10 neurons and necessary number of inputs and outputs. In Figure 1, the ANN model using 3 input features (Course module viewed, Discussion module viewed, and Quiz module viewed) to classify student's success to 9 grade letters is given.

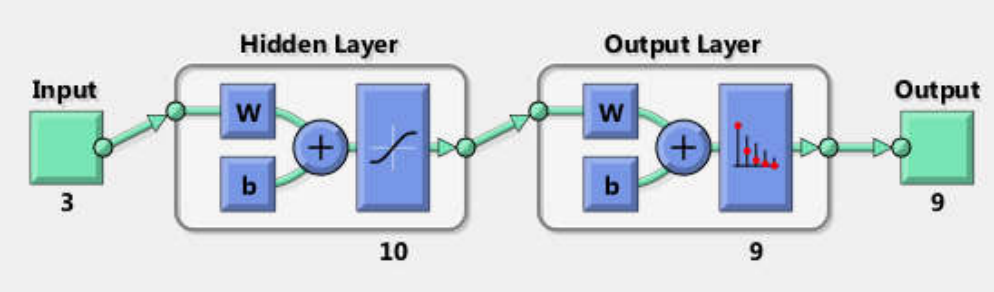

Figure 1. *The ANN model used in experiments.*

4.3 Results

We have conducted several tests to observe the effect of selected features on the prediction accuracy. The tests and their results are summarized at Table 1.

In the first experiment we use all the types of interaction numbers as inputs to ANN model. Using these inputs, we would like to predict a student's grade letter. As seen in the table, the prediction accuracy is very low, about 20%.

However, for the second example, by using the same features of the experiment 1, we try to predict the student's pass/fail status. In that case, the accuracy is considerably increased to a level of 80%. Even though, it is not at a desirable level, these results may support the expectation which suggests "a relationship between levels of student's involvement to the LMS with the level of their results". In order to examine this expectation in details, we also conducted the following tests.

Test 4 and 5 are designed to observe the effect of total number of logged-on activity. For both of success metrics, the prediction accuracy is increased with respect to previous results. Using a single feature (total number of logged-on) can attain 81.6% accuracy for the Pass or Fail metric. However, the prediction of grade letters still indicates very low accuracy.

In tests 5 and 6, we added Sex and Department features into the feature set as well. As seen at Table 1, adding new features has increased the prediction accuracy of the grade letters from 20% to 50%. On the other hand, the added features negatively affect the Pass/Fail prediction accuracy by dropping it about 3%.

Table 1. *The prediction experiments and their results.*

No	Features as inputs	Predicted Target Metric	Accuracy
1	Course module viewed, Discussion module viewed, Quiz module viewed	Grade Letters	20.4%
2	Course module viewed, Discussion module viewed, Quiz module viewed	Pass or Fail	80.9%
3	Total logged-on	Grade Letters	23.7%
4	Total logged-on	Pass or Fail	**81.6%**
5	Course module viewed, Discussion module viewed, Quiz module viewed, Total logged-on, Sex, Department	Grade Letters	**50.0%**
6	Course module viewed, Discussion module viewed, Quiz module viewed, Total logged-on, Sex, Department	Pass or Fail	78.3%
7	Sex	Grade Letters	18.4%
8	Sex	Pass or Fail	77.6%
9	Quiz module viewed	Grade Letters	17.8%
10	Quiz module viewed	Pass or Fail	78.9%
11	Course module viewed	Grade Letters	25.6%
12	Course module viewed	Pass or Fail	**81.6%**
13	Discussion module viewed	Grade Letters	21.7%
14	Discussion module viewed	Pass or Fail	78.3%
15	Department	Grade Letters	15.1%
16	Department	Pass or Fail	77.6%

Beginning from experiment 7, we investigated the effect of single features over the prediction accuracy. In these experiments, we observed that the feature "Course module viewed" attained the highest accuracy for "Pass or Fail" metric with 81.6% and for "Grade Letters" metric with 25.6%. When all the results considered, these results are important. For "Pass or Fail" metric, "total logged-on" feature lead the same accuracy. However, using all 6 features generates higher accuracy (50.0%) for "Grade Letters" metric.

Department and Sex features generates the lowest accuracy (77.6%) for "Pass or Fail" metric. For "Grade Letters" metric, the lowest accuracy (15.1%) is generated by the Department feature.

4.3 Discussion

The above results indicate that prediction of Grade Letters with high level of accuracy has not been accomplished in the experiments (maximum 50%). This result could be due to the fact that the collected data is composed of only 152 students and number of classes (letter grades) is relatively very high (9). Thus, ANN algorithm couldn't create a proper model for the given inputs with respect to classes.

On the other hand, prediction of "Pass or Fail" has been done with considerable better level of accuracy 81.6%. Two separate features, Total logged-on and Course module viewed, can attain this level. Using all features can also generate very close accuracy (80.9%).

We also observe that sex and department features do not have much impact on prediction of success metrics alone.

5. CONCLUSIONS

The experiment results indicate that predicting grade letters is much more prone to errors than predicting the pass/fall result. The tests show that we can predict if a student pass or fail with about 81% accuracy considering only the number of online activities.

As expected, using only demographic information (sex and department) does not lead a good prediction at all. Using both the number of online activities and demographic information do not increase the prediction accuracy considerably either.

These results indicate that, in a blended course, student success is not only determined by the quantity of online activities but also it might be related with the quality of the face-to-face interaction with the instructor.

As a future work, we would like to extend this work by collecting more data and online activity features. We hope that with more data and features, we can attain higher accuracy in the predictions of the success metrics.

6. REFERENCES

[1] Singh, H. and Reed, C., 2001. "A white paper: Achieving success with blended learning." Centra software, 1, pp.1-11, Available: http://www.leerbeleving.nl/wbts/wbt2014/blend-ce.pdf

[2] Graham, C.R., Allen, S. and Ure, D., 2005. Benefits and challenges of blended learning environments. In Encyclopedia of Information Science and Technology, First Edition (pp. 253-259). IGI Global.

[3] Boyle, T., Bradley, C., Chalk, P., Jones, R. and Pickard, P., 2003. Using blended learning to improve student success rates in learning to program. Journal of educational Media, 28(2-3), pp.165-178.

[4] Vaughan, N., 2007. Perspectives on blended learning in higher education. International Journal on ELearning, 6(1), p.81.

[5] Holsapple, C.W. and Lee-Post, A., 2006. Defining, assessing, and promoting e-learning success: An information systems perspective. Decision sciences journal of innovative education, 4(1), pp.67-85.

[6] Graham, C.R., 2013. Emerging practice and research in blended learning. Handbook of distance education, 3, pp.333-350.

[7] The Course Networking system, Available: https://www.thecn.com

[8] Meltem Eryılmaz, Murat Karakaya, 2017, The Impact of Using Social Media in Blended Learning Environments on Student Satisfaction, submitted to Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi.

[9] Course web page, Available: http://compe.atilim.edu.tr/academicprogramcourses/view/id/1155?lang=en