An Artificial Neural Network Based Early Prediction of Failure-Prone Students in Blended Learning Course

Otgontsetseg Sukhbaatar (✉), Tsuyoshi Usagawa
Kumamoto University, Kumamoto, Japan
otgon@st.cs.kumamoto-u.ac.jp

Lodoiravsal Choimaa
National University of Mongolia, Ulaanbaatar, Mongolia

Abstract—One of the objectives of the performance measurement of grade-based higher education is to reduce the failure rate of students. To identify and reduce the number of failing students, the learning activities and behaviors of students in the classroom must be continuously monitored; however, monitoring a large number of students is an extremely difficult task. A penetration of web-based learning systems in academic institutions revealed the possibility of evaluating student activities via these systems. In this paper, we propose an early prediction scheme to identify students at risk of failing in a blended learning course. We employ a neural network on the set of prediction variables extracted from the online learning activities of students in a learning management system. The experiments were based on data from 1110 student who attended a compulsory, sophomore-level course. The results indicate that a neural-network-based approach can achieve early identification of students that are likely to fail; 25% of the failing students were correctly identified after the first quiz submission. After the mid-term examination, 65% of the failing students were correctly predicted.

Keywords—Blended learning course, educational data mining, failure prediction, higher education, neural network

1 Introduction

One of the main objectives of higher-education institutions is to provide high-quality education to their students. Educational quality can be measured by the academic performance and success of students. The success rate of every individual subject can impact the overall completion rate of the educational program. To increase academic success, it is essential to identify students at risk of failing as early as possible. This can be achieved by monitoring the learning activities and achievements of students during the course. However, it is almost impossible to track student activities in conventional teaching environments, particularly if the number of students is relatively large. An introduction of web-based teaching and learning systems in higher

education enabled the possibility of processing and evaluating activities in web-based educational settings [1, 2].

Web-based educational systems, namely learning management systems (LMSs), generate a large amount of fine-grain data on student learning activities. In popular LMSs, instructors are able to check basic learning activity data. However, no functions are available that can help instructors predict the possible outcome and identify students in need of assistance. Data-mining and machine-learning techniques have enabled the modeling of several time-variant and time-invariant features of students in online learning environments [3]. With the increased interest in using data-mining methods for educational purposes, several practices have been presented during the past two decades. Different data-mining techniques can be applied to LMS data, depending on the desired application [4, 5].

A primary area of application has been the usage of prediction methods for future outcomes. Beck and Woolf presented the successful implementation of a student model using previous user data to predict responses for problem solving [6]. Since then, prediction became one of the most dominant research domains in this field. Researchers have proposed different schemes in which various data-mining methods have been employed to predict student performance. Among the techniques used in student-performance prediction, the most popular ones are the decision tree, k-nearest neighbors, support vector machine, and neural networks [7]. These predictive models can be used in performance prediction as a warning system to inform students and instructors during the semester. Pistilli and Arnold have presented a leading example of an internally developed early warning system to accommodate the needs of at-risk students in academic institutions [8]. However, a significant amount of work should be conducted to achieve revolutionary prediction results using academic and web-based learning environment data.

In this paper, we present a prediction model for failure-prone students that uses neural networks in a blended learning course. Five years of data were used to train the models; the models were validated by using student data from the beginning of the semester for different academic years. Semester data-based cross-validation was conducted to ensure generalization. The results were based on undergraduate student data, particularly LMS log data, online quiz scores, mid-term scores, and the final grade information of a compulsory, sophomore-level course of the Kumamoto University, Japan. In general, blended learning is considered as an efficient method in academic institutions to deliver distance education in terms of student experience, as well as instructor experience and preference. However, blended-learning courses have a limited degree of activities in LMS compared with pure online courses. This can raise challenges for educators in terms of analyzing LMS data and achieving results for action in the same manner that they would in e-learning courses. The accuracy of the results presented in this paper proves that there is great potential for early warning systems using blended learning course data in academic institutions.

The remainder of this paper is organized as follows. In Section 2, the findings of related study literature are reported. The methodology of the study, namely the description of the course and dataset, and the details of machine learning techniques are described in Section 3. In Section 4, the experimental results are presented and their

limitations are discussed. Finally, Section 5 concludes the present work and presents the main findings, and the future research directions of this study are outlined.

2   Related Work

In the past decade, several institutions have started deploying analytical tools to achieve various goals [9]. A number of researchers have focused on performance prediction schemes in higher-education institutions. Several studies have been published and are available for literature review. These studies can be categorized in terms of data used and methods applied for the prediction task.

In studies in which a prediction of student performance was presented, a combination of student attributes was typically used, such as high-school background, demographics, and academic data, e.g., the cumulative grade point average (CGPA) [10, 11, 12, 13, 14]. Apart from using past data, in certain studies, attributes collected during the course were used. The engagement of a student in LMS and the assessments during the progression of the course have revealed a dropout prediction accuracy of 75–85% in the first sections of the e-learning courses [15]. In a different study, similar course assessment attributes were used for the dropout prediction task with high accuracy in the early weeks of the course timeline [16]. Shahiri et al. reported that the internal course activity and assessment data can yield more accurate results regardless of the applied methods [7].

In terms of the applied methods, researchers have used various data-mining classification techniques. Decision trees have been extensively used in several studies owing to their simple interpretation and assurance, even for a small amount of data [17, 18, 19]. In several studies in the literature, the Naïve Bayes and/or the support vector machine (SVM) techniques have been employed to predict student performance [20, 17, 21, 22]. Researchers tend to employ these methods together with other data-mining techniques to compare the accuracy of the prediction.

Neural networks are a popular machine learning technique. They are vastly used in educational data mining field owing to their high prediction accuracy for data with nonlinear variable dependencies. In the study of Lykourentzou et al., the results of the dropout prediction using a feed-forward neural network reached an overall accuracy of up to 96% [15]. This study was one of the successful examples of early prediction of at-risk students in e-learning courses. Arsad et al. used neural networks to predict the CGPA at the 8th semester of undergraduate students based on their grade points in fundamental courses [23]. Most recently, the study presented high accuracy results of performance prediction using neural networks in massive online course learner's data [24]. The study, which reviewed research works on performance prediction reported that neural networks presented the highest prediction accuracy compared with other data-mining techniques [7].

Furthermore, certain studies exist on performance prediction, particularly on the prediction of whether students will pass or fail the course. In the study of Tanner and Toivonen, the results revealed an early prediction of students with a high risk of fail-

ing using a k-nearest neighbor algorithm [25]. Romero et al. proposed a failure-prediction scheme based on attributes from online discussion forums [26].

In a very recent study, an early prediction of the failure risk of the students applying four different methods was presented [27]. In this study, neural networks, SVMs, decision-trees, and Naïve Bayes methods were compared for failure prediction in terms of prediction effectiveness. The novelty of the study relied on experimental prediction results based on academic data from both e-learning and on-campus courses.

From the literature review, it is clear that different techniques deliver different prediction accuracies depending on data characteristics. In this work, we will apply the method with the best-reported accuracies, namely a neural network, to identify students with a high risk of failing in a blended learning course.

3 Methodology

3.1 Course setting

In this study, we used data obtained from the blended learning style course "Digital Signal Processing". It is a sophomore-level, compulsory-credit course taught in the Engineering Faculty of Kumamoto University. The course is offered once a year, i.e., in the Fall semester; data from six semesters were used in this study.

The course is organized with face-to-face lectures, on-campus final examination, and online activities on LMS, including regular weekly quizzes, reading material, and monitored mid-term examination [28].

The online activities are delivered through Moodle LMS, which is integrated to the university portal. The course contents are scheduled for a 15-week semester; however, LMS activities are prolonged until the final examination date. The weekly quiz section consists of two to five multiple-choice questions, which are given as homework with a specific deadline. Every quiz allowed a maximum of five attempts for submission. The mid-term examination is a monitored online test after the winter break, which students can take in on-campus classrooms. The final examination is a conventional paper-based on-campus examination. All online and on-campus activities contribute to the final score, which ranges between 0–100; the grading system may grant students AA, A, B, C, and F. The final examination accounts for 50% of the overall grade and the LMS activities (including the mid-term examination) constitute the remaining 50%. To pass the course, the minimum requirement for the final grade (i.e., earn the credit) is to earn 60 points, which will be calculated for each student after the final examination by summing up the final-examination and the online-activity scores. Because the present study is focused on the failing students, we assume that the student has failed if he/she has earned less than 60 points in total and has received an "F" grade.

The course is a fundamental engineering course. Therefore, the contents of the course do not change significantly over the years. Moreover, the course structure is the same each year, i.e., the same number of quiz sections is provided and the exami-

nations have similar characteristics. We conducted an exploratory data analysis to examine whether similarities exists among the LMS activities over the years. Figure 1 presents the accumulated activity of students for six consecutive semester data. Because the course is offered in the Fall semester of the academic year, each set of semester data comprises the activity between October 10th and January 21st. In Figure 1(a), we may observe that each semester presents peak points around the second week of January, when the online mid-term examination took place. As expected, similarity exists among the semesters; the semester pattern is illustrated in Figure 1(b). The random positive and negative peaks in Figure 1(c) can be explained by the difference in the course schedule over the years (mid-term examination dates do not occur on the same date each year). The figure clearly shows that there is a similar activity pattern among the different annual offsets of LMS data, which implies similar activity characteristics during each semester. This data behavior offers the possibility to obtain reasonable results, which can be applied in future course semesters.
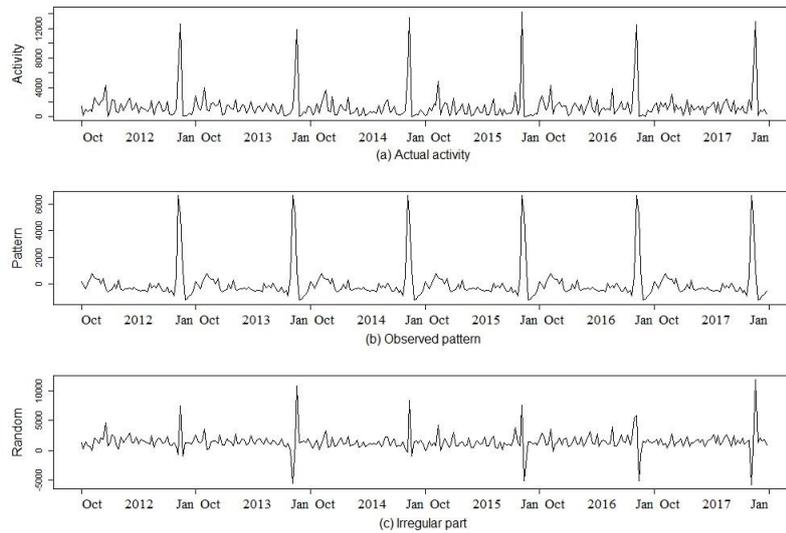

Fig. 1. Accumulated activity in LMS (6 years data)

3.2 Dataset description

With due respect to privacy issues, all personal data were eliminated and every individual is anonymously presented in the dataset.

The course data consist of:

- The online activity data of the students from LMS
- The final grades as the performance data of the students.

In this study, six semester data within 2012–2017 were used; the total number of enrolled students was 1167. In the data-preprocessing step, students without any online

activity were elimi- nated; hence, the failure prediction was solely based on the LMS activity of the students. As a result, the total number of students was 1110. Figure 2 shows the number of participating and failing students.
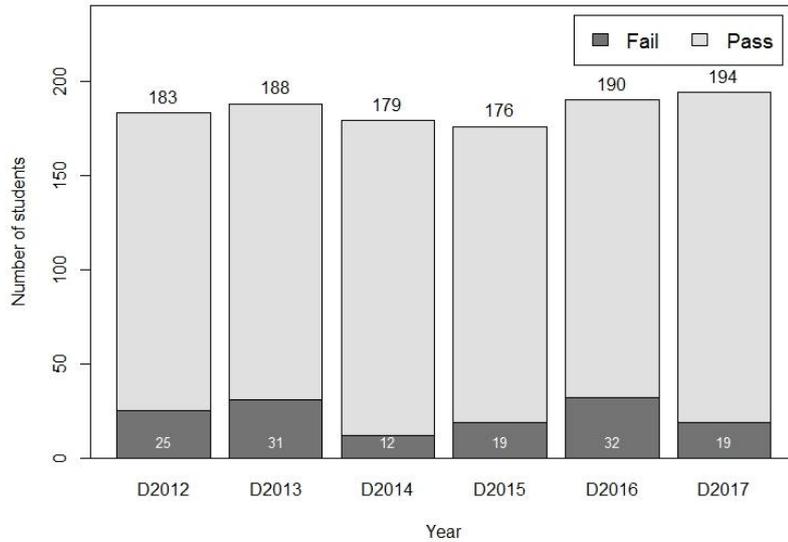

Fig. 2. The number of participating students and failing students

From the LMS log files, time variant variables were extracted for different time periods during the course timeline. The quiz activities were found to be the most dominant and influencing activities in the LMS. Owing to the ease of use and uniformity, all quiz scores were scaled to 100 points. The extracted variables used for training and validation are listed in Table 1.

Table 1. Variables used for the training and testing of neural networks

Name	Description	Range
Section activity	Student activity in quiz section	Greater than 0
Quiz Attempts	Number of attempts in quiz sections, accumulated	Greater than 0
Active days	Number of days student was active in LMS	Greater than 0
Attempted quiz	Number of attempted quizzes	0–13
Completed quiz	Number of quizzes with a completed mark	0–13
Quiz score	Average quiz score	0–100
Quiz solving time	Accumulated time spent on quiz submission	Greater than 0
Exam score*	Mid-term examination score	

*Only used in after 12th quiz deadline

3.3 Training and validation

The dataset was separated into two sets, i.e., a training set and a test set. It is common practice to split the entire dataset into training and test sets in a random manner using certain ratio; further cross-validation can be used to generalize the results [29]. According to the conventional separation of the training and the test set, we should have randomly formed the training and the test set from a total population of 1110 students. Instead, we used one set of semester data as the test set and the data of the remaining years as the training set. The data were divided in this manner because we wished to examine the possibility of failure prediction using entire sets of separate semester data, where data from previous years would be used for the training of the model. This approach would give the opportunity for instructors to train the models using course data from previous years and to apply the resulting models to the upcoming semester throughout the course. For generalization purposes, a six-fold cross-validation was implemented. For example, D2012, D2013, D2014, D2015, and D2016 (data obtained during 2012–2016) were used for the training of the model and D2017 was used for the validation. Next, D2012 through D2015 and D2017 were used for the training and D2016 was used for the validation; the same process was followed for the remaining year ranges. If the prediction method and the aforementioned data splitting could produce valid results, it would prove that this approach would be a good generalization and would enable further application using the data of the upcoming year.

The training process consisted of 12 steps, which covered data from different periods of the semester. These periods were defined using the first 12 quiz sections as a deadline. Typically, the submission period for every quiz section is one week (or more for certain cases). In the first training step, the variables that had been acquired until the deadline of the first quiz section were considered. For the second step, the variables that had been collected from the beginning to the end of the second-quiz deadline were used (activities of the first and second quiz sections). For the twelve-step training, neural-network inputs were acquired from the beginning of the semester until the quiz-section deadline. In each training phase, 82.5–84.1% of the total student participants was considered.

The testing process was conducted using the same scheme as that used in the training phase. The prediction variables were accumulatively extracted and tested for different semester time stamps. For the validation of the prediction models, the test set covered 15.9–17.5% of the total number of students.

3.4 Neural networks

A neural network is an information processing paradigm of an artificial intelligence field. A neural network consists of numerous processing nodes, referred to as neurons, and the connectivity between these neurons. In a neural network, the processing is performed as weights of connection among neurons and through its ability to learn from the training set. A neural network is organized in three layers of neurons, namely an input layer, a hidden layer, and an output layer. The transfer function
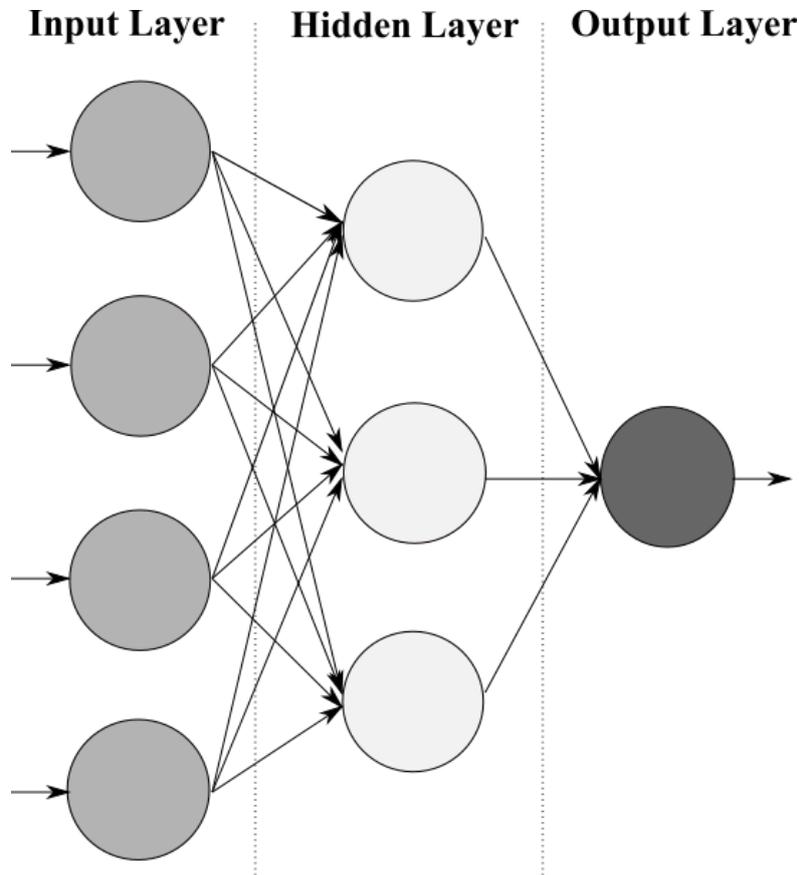
Input Layer   Hidden Layer   Output Layer

Fig. 3. Basic structure of a neural network

of every neuron in each layer individually processes data in the input to the output. The hidden layer can consist of more than one sub-layers. More layers added in the hidden layer will result in a large network with an increased complexity in training the model. Figure 3 illustrates the simple architecture of the neural network. At present, neural networks are extensively used for various types of tasks, including recognition, prediction, signal processing, control, and anomaly detection. In this study, eleven 7-3-1 networks and one 8-3-1 network were employed. Here, the each of the three network configuration numbers indicates the number of neurons in the input layer, the number of neurons in the hidden layer, and the number of neurons in the output layer in the order of their appearance. The reason why different input neurons exist for the input layer is that we added mid-term examination score as an input variable at the 12th step of training. The output of the network is binary, where 1 indicates a student whol failed and 0 for a student that is a completer. RStudio v.1.1463 was used for the processing and the visualization of the data used in this study.

3.5   Accuracy metrics

To evaluate the accuracy of the prediction, the following metrics were examined:

$$accuracy = \frac{TP+TN}{N} \tag{1}$$

$$sensitivity = \frac{TP}{TP+FN} \tag{2}$$

$$precision = \frac{TP}{TP+FP} \tag{3}$$

$$f1 = 2 \times \frac{precision \times sensitivity}{precision + sensitivity} \tag{4}$$

where

TP – True Positive: the number of students that were predicted to be at risk of failure and failed to pass the course

TN – True Negative: the number of students that were predicted to be completers and successfully completed the course

FP – False Positive: the number of students that were predicted to be at risk of failing but completed the course

FN – False Negative: the number of students that were predicted to be completers but failed to pass the course

4   Experimental Results

In this section, the experimental results will be presented and discussed.

The overall accuracy of the results is defined via Eq. (1). This metric evaluates the number of successful prediction results, including the prediction of the number of failures and completers.

The overall accuracy of the test results is illustrated in Figure 4. As shown in the figure, the models present stable and significantly high prediction accuracy results (> 84%) from the beginning of the semester. These results prove that the prediction method generalizes well.

Moreover, it shows that student performance can be accurately predicted early in the semester, whether students eventually fail or complete the course, in a blended learning course. However, the overall accuracy does not fully represent the evaluation of the prediction capabilities.

As defined in Eq. (1), the overall accuracy metric includes both the correctly predicted completers and the students who failed. Typically, the final grade distribution of the course is negatively skewed, which means that the number of completers is significantly higher than the number of students who have failed. The percentage of failing students in the datasets that were used in this study ranges between 7–17%.
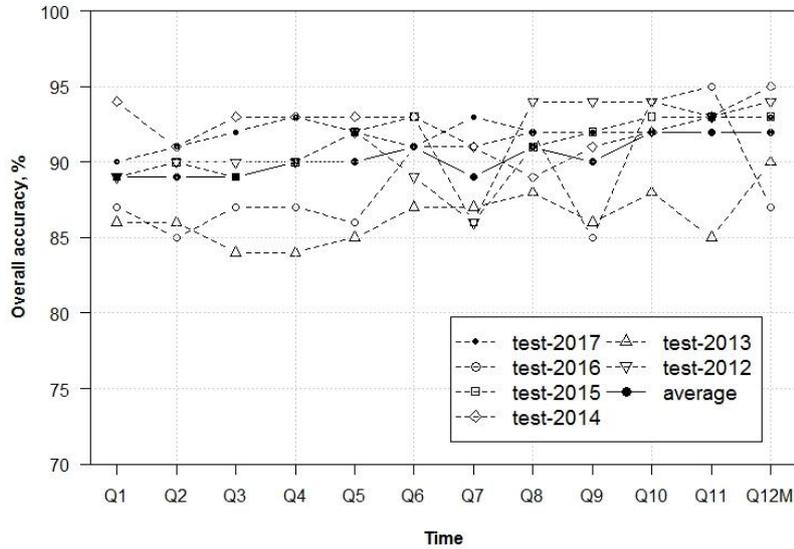
Fig. 4. Overall accuracy of prediction results (176-194 students of test sets)
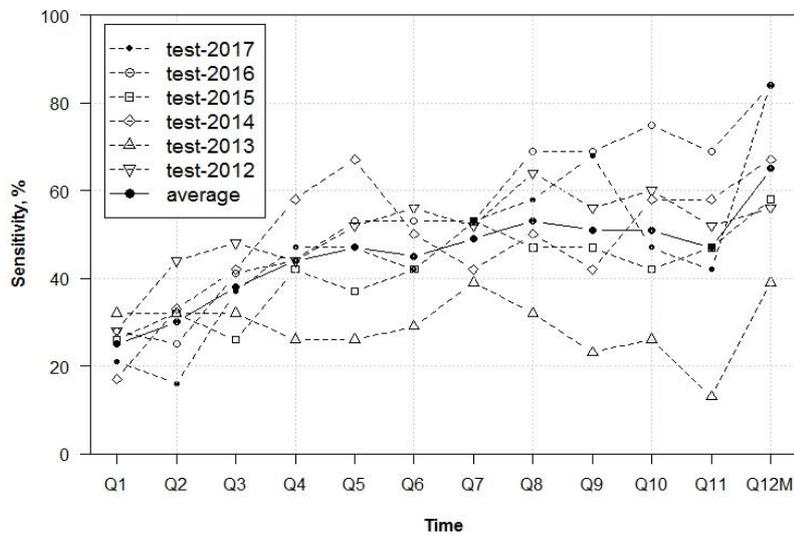

Fig. 5. Sensitivity of prediction results (out of 12–32 failing students from the test sets)

In this work, we focused on the failing students; our main task was the prediction of students at risk of failing. Therefore, sensitivity measures were considered to illustrate the prediction accuracy of the failing students. The sensitivity metric, which is defined in Eq. (2), measures the proportion of the correctly predicted at-risk students

from the total number of failing students. Figure 5 shows the sensitivity results of the prediction on the test sets. The results show the correct prediction over the total failure ratios of 4/19, 9/32, 5/19, 2/12, 10/31, and 7/25 for test-2017, test-2016, test-2015, test-2014, test-2013, and test-2012, respectively. On average, 25% of the failure-prone students can already be identified at the beginning of the semester. In the middle of the semester (Q8), the ratios increase to 11/19, 22/32, 9/19, 6/12, 10/31, and 16/25 with an average sensitivity of 53%. After the mid-term examinations (and 12th quiz section), the ratios increase to 16/19, 27/32, 11/19, 8/12, 12/31, and 14/25; the average of the correctly predicted failing students was 65%. The relatively poor prediction results in a test set from 2013 may be attributed to the difference in the student body characteristics, i.e., the student efforts and achievements can change year by year. These results promise the possibility of early identification of failure-prone students by the instructors in order for instructors to take appropriate actions to notify, encourage, and support these students. We acknowledge that the accuracy obtained in this study is not as high as those reported in the literature, which are typically based on pure online-course data. However, we believe that this study can provide compelling evidence on the prediction of failure-prone students in a blended-learning course.

An additional measure of the accuracy of the prediction is the precision. The precision of the prediction indicates the percentage of correct predictions among all positive outcomes. The precision results presented in Figure 6 indicate the extent to which the prediction model can accurately predict the failed students among students that
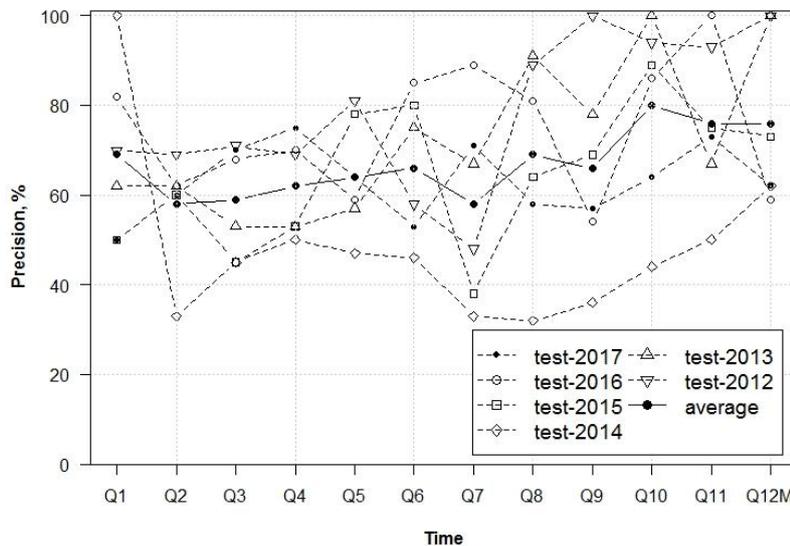

Fig. 6. Precision of prediction results

were predicted to be prone to failure. At the beginning of the semester, the FPs ranged between 0–6 students. At the end of the examined period (Q12M), the FP predictions ranged between 0–19 students. A total of 38 students (for 6 test sets) were incorrectly

predicted to be at risk of failure after the mid-term examination. The results are reasonable; the performance of all these students was below average, except for one student with grade "A". The anomalous result in the 2014 test set may be the result of irregular quiz submission deadlines that did not align with data from other years.

To evaluate the overall effectiveness of the method, we examined the f-measure (i.e., the f1 score), which is the most common metric used in classification problems [30]. Equation 4 represents the f-measure estimation, in which both the precision and the sensitivity are considered. Figure 7 presents the estimated f-measure results of the test sets. The average f-measure value of the test sets ranges between 36–66%, from the beginning of the semester (Q1) to the end of 12th quiz section (Q12M). This result may be interpreted as follows: after the first quiz submission, the prediction model could identify students at risk of failure with an effectiveness of 36%.

Limitations of the results: Although we admit that certain limitations exist, we think that the experimental results of the study reveal the possibility of being able to predict and identify the students who are at risk of failure in a blended learning course.

The experimental results cannot be generalized for all blended courses because the data used in this study involve only the student learning activity in one particular course. The study can be extended using other course data and/or, more preferably, data from different institutions. Moreover, owing to the blended learning style, in this study, the failure prediction of the students was solely based on quiz-related activities.

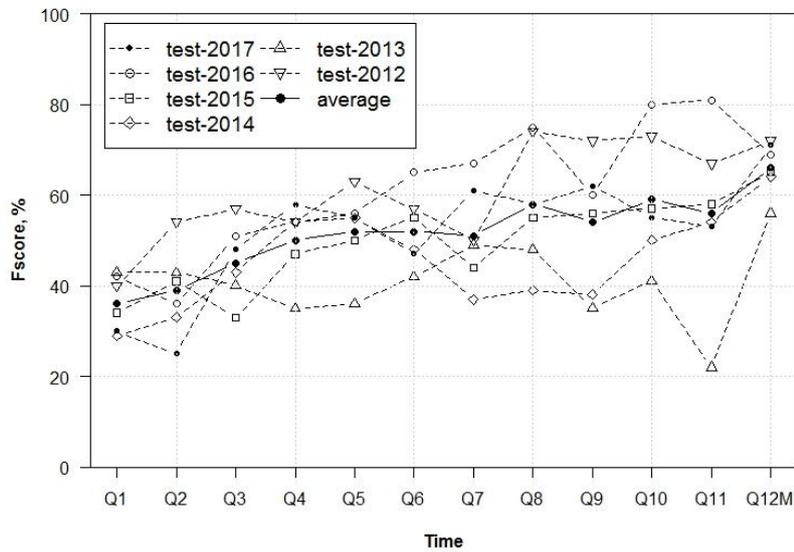Our primary intention was to examine the possibility of failure prediction using

Fig. 7. F1-score measures of the results

One-semester enrollment data in the future. Owing to this reason, the training and test sets were not randomly split for cross-validation. Hence, the experimental results

presented irregular accuracy results for different test sets. The overall student characteristics differed in each course enrollment, although the academic background was fairly the same.

The prediction efficiency was not as high as the one reported in the literature. For early prediction purposes, later semester periods were not considered. Certain compelling accuracy results in the literature were based on full-course data; high-efficiency results were mostly achieved in the later in the course. However, as per our intention, considering the application of the prediction results in upcoming semesters, the later semester periods would be too late for the instructors to take action and intervene or provide support to students at risk of failure.

We did not fine-tune the method applied to the experimental results. There is a great possibility that the tuning structure and the parameters of the neural network can increase the overall effectiveness of the results.

Moreover, other student characteristics may affect student performance. As far as educators are concerned, the relationship between the learning activity and the student performance can be quite complex. Different student characteristics and other extracted variables may greatly impact the failure-prediction results. Therefore, examining a wide variety of variables is necessary for the improvement and validation of the results.

5 Conclusion

Presently, the online learning activities of students are a crucial part of their learning process; therefore, there is a definite requirement for developing an efficient method to monitor and report LMS activities. Norris et al. emphasized the power of academic data analysis in higher education and the possibility of data being used to take appropriate action [31]. The experimental results presented in this study showed how a machine-learning technique can be used to improve student performance in universities. More specifically, the results suggested the possibility of realizing an early warning system using the online activity data of a blended course in degree programs.

It was fascinating that 25% of the failing students could be correctly predicted immediately after the first quiz section. The prediction accuracy gradually increased week by week, reaching 53% after the 8th quiz and 65% after the mid-term examination.

Future work should be conducted to overcome the limitations of the present study. First, in future works, other possible attributes should be investigated and datasets from other courses should be included to increase the accuracy and to generalize the results. In this study, the variable extraction and the data preparation steps are realized manually. An instructor-friendly plug-in tool could be developed to automate the entire procedure. These types of tools could assist in the acquisition and the processing of data in a timely manner, and they could offer access to periodical (e.g., weekly) results to instructors and students.

6   References

[1] R. Agrawal and R. Srikant (1995). Mining sequential patterns. Proceedings of the 11th International Conference on Data Engineering, Taipei.

[2] A. L. Ingram (1999). Using web server logs in evaluating instructional web sites. Journal of Educational Technology Systems, 28(2), pp. 137-157. https://doi.org/10.2190/r3ae-ucry-njvr-ly6f

[3] O. R. Zaïane (2001). Web usage mining for a better web-based Learning envronment. In Proceedings of conference on advanced technology for education, Banff, Canada.

[4] C. Romero and S. Ventura (2007). Educational data mining: a survey from 1995 to 2005. Expert Systems with Applications 33 , pp. 135-146. https://doi.org/10.1016/j.eswa.2006.04.005

[5] Suhirman, J. M. Zain, H. Chiroma and T. Herawan (2014). Data mining for education decision support: a review. International Journal of Emerging Technologies in Learning, 9(6), pp. 4-19. https://doi.org/10.3991/ijet.v9i6.3950

[6] J. E. Beck and B. P. Woolf (2000). High-level student modeling with machine learning. In Proceedings of the International Conference on Intelligent tutoring systems, Montreal, Canada.

[7] A. M. Shahiri, W. Husain and N. A. Rashid (2015). A review on predicting student's performance using data mining techniques. Procedia Computer Science, 72. https://doi.org/10.1016/j.procs.2015.12.157

[8] M. D. Pistilli and K. E. Arnold (2010). Purdue Signals: Mining Real-Time Academic Data to Enhance Student Success. About Campus, 15(3), p. 22–24. https://doi.org/10.1002/abc.20025

[9] U. b. Mat, N. Buniyamin, P. M. Arsad and R. Kassim (2013). An overview of using academic analytics to predict and improve students' achievement: a proposed proactive intelligent intervention. IEEE 5th Conference on Engineering Education (ICEED), Kuala Lumpur, Malaysia. https://doi.org/10.1109/iceed.2013.6908316

[10] Z. Ibrahim and D. Rusli (2007). Predicting students academic performance: comparing artificial neural network, decision tree and linear regression. 21st Annual SAS Malaysia Forum, Kuala Lumpur, Malaysia.

[11] E. Osmanbegovich and M. Suljic (2012). Data mining approach for predicting student performance. Economic Review, 10(1).

[12] V. Oladokun, A. Adebanjo and O. Charles-Owaba (2008). Predicting students academic performance using artificial neural network: A case study of an engineering course. The Pacific Journal of Science and Technology, 9(1), pp. 72-79.

[13] T. M. Christian and M. Ayub (2014). Exploration of classification using nbtree for predicting students' performance. International Conference on Data and Software Engineering, Bandung, Indonesia. https://doi.org/10.1109/icodse.2014.7062654

[14] M. Tan and P. Shao (2015). Prediction of student dropout in e-learning program through the use of machine learning method. International Journal of Emerging Technologies in Learning, 10(1), pp. 11-17. https://doi.org/10.3991/ijet.v10i1.4189

[15] I. Lykourentzou, I. Giannoukos, V. Nikolopoulos, G. Mpardis and V. Loumos (2009). Dropout prediction in e-learning courses through the combination of machine learning techniques. Computers & Education, 53, pp. 950-965. https://doi.org/10.1016/j.compedu.2009.05.010

[16] C. Burgos, M. L. Campanario, D. d. l. Pena, J. A. Lara, D. Lizcano and M. A. Martinez (2018). Data mining for modeling students' performance: a tutoring action plan to prevent

academic dropout. Computers and Electrical Engineering, 66, pp. 541-556. https://doi.org/10.1016/j.compeleceng.2017.03.005

[17] V. Ramesh, P. Parkavi and K. Ramar (2013). Predicting student performance: a statistical and data mining approach. International Journal of Computer Applications, 63(8), pp. 35-39. https://doi.org/10.5120/10489-5242

[18] S. Natek and M. Zwilling (2014). Student data mining solution-knowledge management system related to higher education institutions. Expert Systems with Applications, 41, pp. 6400-6407. https://doi.org/10.1016/j.eswa.2014.04.024

[19] C. Romero, S. Ventura, P. G. Espejo and C. Hervas (2008). Data mining algorithms to classify students. Educational Data Mining, pp. 8-17.

[20] S. T. Jishan, R. I. Rashu, N. Haque and R. M. Rahman (2015). Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique. Decision Analytics, 2(1). https://doi.org/10.1186/s40165-014-0010-2

[21] W. Hämäläinen and M. Vinni (2006). Comparison of machine learning methods for intelligent tutoring systems. International Conference on Intelligent Tutoring Systems . https://doi.org/10.1007/11774303_52

[22] G. Gray, C. McGuinness and P. Owende (2014). An application of classification models to predict learner progression in tertiary education. International Advance Computing Conference, Gurgaon, India. https://doi.org/10.1109/iadcc.2014.6779384

[23] P. M. Arsad, N. Buniyamin and J.-l. A. Manan (2013). A Neural Network Students' Performance Prediction Model (NNSPPM). International Conference on Smart Instrumentation, Measurement and Applications, Kuala Lumpur, Malaysia. https://doi.org/10.1109/icsima.2013.6717966

[24] Y. Zhang and W. Jiang (2018). Score prediction model of MOOCs learners based on neural network. International Journal of Emerging Technologies in Learning, 13(10), pp. 171-182. https://doi.org/10.3991/ijet.v13i10.9461

[25] T. Tanner and H. Toivonen (2010). Predicting and preventing student failure - using the k-nearest neighbour method to predict student performance in an online course environment. International Journal of Learning Technology, 5(4), pp. 356-377. https://doi.org/10.1504/ijlt.2010.038772

[26] C. Romero, M.-I. López, J.-M. Luna and S. Ventura (2013). Predicting students' final performance from participation in on-line discussion forums. Computers & Education, pp. 458-472. https://doi.org/10.1016/j.compedu.2013.06.009

[27] E. B. Costa, B. Fonseca, M. A. Santana, F. F. d. Araújo and J. Rego (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. Computers in Human Behavior, 73, pp. 247-256. https://doi.org/10.1016/j.chb.2017.01.047

[28] K. Ogata and T. Usagawa (2017). Lecture management of parallel classes in a blended learning style: the case of Digital Signal Processing I as a compusory course. Acoustical Science and Technology, 38(4), pp. 203-212. https://doi.org/10.1250/ast.38.203

[29] C. Schaffer (1993). Selecting a classification method by cross-validation. Machine Learning, 13, pp. 135-143. https://doi.org/10.1007/bf00993106

[30] D. L. Olson and D. Delen (2008). Advanced data mining techniques. Berlin: Springer.

[31] D. Norris, L. Baer, J. Leonard, L. Pugliese and P. Lefrere (2008). Action analytics: measuring and improving performance that matters in higher education. Educause Review, pp. 42-66.

7   Authors

Otgontsetseg Sukhbaatar earned her master degree in computer science and electrical engineering from Seoul National University in 2009. She has worked as a lecturer in School of Engineering and Applied Sciences, National University of Mongolia until 2016. She is currently a doctoral student at Graduate School of Science and Technology, Kumamoto University. Her areas of research interest include learning analytics and educational data analysis.

Lodoiravsal Choimaa received his Ph.D. in Electronics Engineering from the National University of Mongolia (with a sandwich program of Juelich Research Center, Germany). He is an Associate Professor at the School of Engineering and Applied Sciences, National University of Mongolia. He has worked as a Dean of the school between 2009 and 2013. Since 2015, he has been working as a head of Machine Intelligence Laboratory, National University of Mongolia. His current research fields are industrial automation, machine vision, and online learning.

Tsuyoshi Usagawa joined Kumamoto University in 1983 right after he received an M.E. degree from Tohoku University. In 1988, he received a Dr. Eng. from Tohoku University. Since 2004 he has been a professor, and he is a member of the IEEE, ASA, ASJ, INCE/J, JSET, and JSAI. He is interested in e-learning contents and systems, and acoustic signal processing.