Chen, W. et al. (Eds.) (2017). Proceedings of the 25th International Conference on Computers in Education.
New Zealand: Asia-Pacific Society for Computers in Education

Students' Performance Prediction Using Data of Multiple Courses by Recurrent Neural Network

Fumiya OKUBO[a*], Takayoshi YAMASHITA[b], Atsushi SHIMADA[c] & Shin'ichi KONOMI[a]

[a] *Faculty of Arts and Science, Kyushu University, Japan*
[b] *Department of Computer Science, Chubu University, Japan*
[c] *Faculty of Information Science and Electrical Engineering, Kyushu University, Japan*
* fokubo@artsci.kyushu-u.ac.jp

Abstract: In this paper, we show a method to predict students' final grades using a recurrent neural network (RNN). An RNN is a variant of a neural network that handles time series data. For this purpose, the learning logs from 937 students who attended one of six courses by two teachers were collected. Nine kinds of learning logs are selected as the input of the RNN. We examine the prediction of final grades, where the training data and test data are the logs of courses conducted in 2015 and in 2016, respectively. We also show a way to identify the important learning activities for obtaining a specific final grade by observing the values of weight of the trained RNN.

Keywords: learning analytics, recurrent neural network, learning log, prediction of student's performance

1. Introduction

At Kyushu University, a learning support system, called M2B system, was introduced in 2014, and a mechanism for recording various logs of learning activities was developed. All students in Kyushu University have their own PC and they can use M2B system both inside and outside classroom. Hence, a large amount of their learning logs are accumulated in M2B system. By March 2017, about 45 million logs have been collected in M2B system. Using these learning logs, a number of investigations on learning analytics have been conducted at Kyushu University (e.g., Ogata et al. (2015), Ogata et al. (2017), Okubo et al. (2016), Okubo et al. (2015)).

In the field of educational data mining, predicting students' performance is important topic, which is useful for teachers and students. Particularly, an early detection of students who are likely to fail or drop out of class, i.e., students refered to as "at-risk" students has been intensively investigated such as in Hlosta, Zdrahal & Zendulka. (2017). In Marbouti, Diefes-Dux & Madhavan. (2016), the prediction method for identifying "at-risk" students in educational data mining is summarized, where Logistic Regression, Support Vector Machine, Decision Tree, Multi-Layer Perceptron, Naïve Bayes Classifier, and K-Nearest Neighbor are guided.

We consider the prediction of students' performance using an artificial neural network which has has recently become a hot topic. With a general neural network, it is impossible to consider the relationship between past data and current data in the time series. This situation often appears in the case of learning logs. A recurrent neural network (RNN) is a variant of a neural network, that can address this problem (Bodén, M. (2002)). An RNN can obtain the output value based on past and current information using the internal loops of the network. In Okubo et al. (2017), we proposed a method for predicting students' performance with an RNN from the log data collected in M2B system. It was confirmed that an RNN is effective in the early prediction of final grade of students in cases where training data and test data were collected from the same course. In order to apply the results of prediction at actual education sites, the model must be trained using data on courses that were held in the past. In addition, the teacher of the course for the training data may be different from the one who will consult the course to be predicted.

In this paper, for the sake of validation of the method of the prediction by an RNN in a more realistic situation, we collect learning logs from six courses. These courses were conducted by two different teachers in 2015 and 2016, who followed the same syllabus for eight week. We examined the method of prediction by an RNN by using the data on courses held in 2015 and 2016 as the training data and the test data. We also compare the results with predictions by multiple regression analysis.

2. Data Collection

2.1 Active Learner Point

Many kinds of logs of learning activities are stored in M2B system. For analyzing and visualizing these data easily, we select nine major learning activities, and evaluate them of each student from 0 to 5 points for each week of a course. A vector of these nine evaluations is called an Active Learner Point (ALP). The nine selected learning activities are summarized in Table 1. We note that

- The logs of a total time of slide views are only calculated for viewing outside of class.
- The logs of the number of course views, total time of slide views, number of markers, number of actions, and word count in a journal are evaluated relatively in a course. For example, the top 10% of students who are doing a lot about the learning activity obtain a score of 5.

Table 1: Criteria for Active Learner Point.

Activities	**5**	**4**	**3**	**2**	**1**	**0**
Attendance	Attendance		Being late			absence
Quiz	Above 80%	Above 60%	Above 40%	Above 20%	Above 10%	Otherwise
Report	Submission		Late			None
Course views	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	Otherwise
Slide views	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	Otherwise
Markers	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	Otherwise
Memos	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	Otherwise
Actions	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	Otherwise
Word count	Upper 10%	Upper 20%	Upper 30%	Upper 40%	Upper 50%	Otherwise

2.2 Courses

We collected learning logs regarding ALPs from six “Information Science” courses for first grade students, as shown in Table 2. These courses were conducted, basically, by following the same syllabus for eight weeks, i.e., they are quarter courses. These courses were conducted in diffrent years and terms with different teachers, and participating students. The prepared materials were needed to implement a standard plan for “Information Science” courses. The teachers were permitted to change this plan so long as the objectives of the course were not changed, e.g., by changing the order of subjects to be taught, adding the extra teaching materials, reports, quizzes, and so on. Hence, the structure and progress varied in each course. The students in the course are evaluated by each teacher with the final grades A, B, C, D, F, based on final exam and several learning activities including quiz and reports. The number of students who obtained each final grade is shown in Table 3.

Table 2: Course information.

Course No.	Year	Term	Teacher	The number of students
1	2015	Spring	A	130
2	2015	Spring	A	209
3	2015	Autumn	B	81
4	2015	Autumn	B	120
5	2016	Spring	A	161
6	2016	Spring	A	236

Table 3: Frequency of final grades obtained in the six courses.

Grade	A	B	C	D	F
The number of students	673	157	69	19	19

3. Method

3.1 Recurrent Neural Network

A recurrent neural network (RNN) is a variant of neural networks that handles time series data. An RNN has been applied in the various fields, such as speech recognition and machine translation. In Figure 1 (a) shows a graphical illustration of a structure of an RNN. By inputting data to an RNN, an output value corresponding to the input value is obtained through a hidden layer. At this time, the internal information of the hidden layer based on the past data is input into an RNN, together with the information of input of the present time. Thus, it is possible to output in consideration of the past state. Figure 1 (b) shows the unfolding in the time of the computation of an RNN. Since the information of the hidden layer at time t-1 is propagated to the same network at time t, an RNN theoretically can output with consideration of all the past information. We can select a method to construct hidden layers, such as Long Short Term Memory and Gated Recurrent Unit (GRU), depending on the way of consideration of the past information. In this paper, we deploy GRU. For the details of RNN, see Bodén, M. (2002).

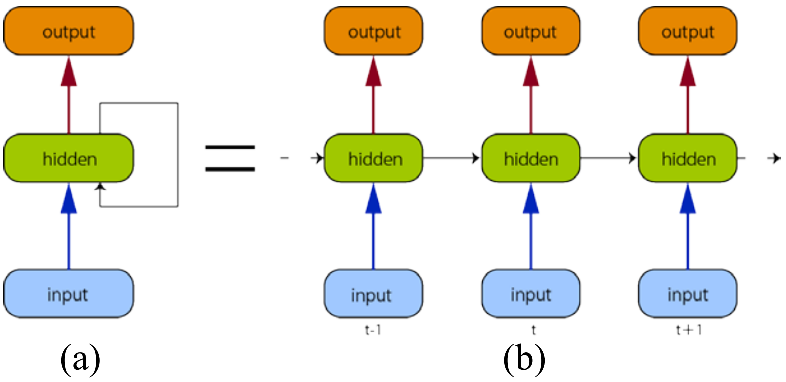

Figure 1. A structure of recurrent neural network.

3.2 Prediction of Students' Final Grades

A vector of nine kinds of points for each week, that is, an ALP (introduced in Section 2.1) of a student is input into the RNN for each time. The final grade A, B, C, D, or F of the student is considered as the output. Let a number of GRUs included in a hidden layer be 32. The time series data of vectors of nine kinds of points is fed into the RNN, and in each time, the final grade is predicted by the trained RNN. For the training of the RNN, we apply the Back Propagation Through Time (BPTT) to repeatedly update parameters of network and learn optimal parameters.

We also examined the prediction of final grades using multiple regression analysis. For this aim, the final grades A, B, C, D, and F are replaced with 95, 85, 75, 65, and 55, respectively. For each week, the multiple regression analysis using the training data is performed. Then, the final grade is predicted as the final grade whose corresponding value is the closest to the value obtained by applying the multiple regression equation to the test data.

4. Experimental Results

4.1 Prediction by Recurrent Neural Network

We evaluate the prediction performance of the proposed method where, the data on courses conducted in 2015 (Course 1, 2, 3, and 4) is used as the training data, while the data on courses in 2016 (Course 5 and 6) is used as the test data. The numbers of the training data and test data are 540 and 397, respectively. Table 4 shows the accuracy of RNN predictions for each week of the experiment. In the eighth week, the prediction has an accuracy of 84.6%. we also show the accuracy of prediction by the RNN for each final grade in Table 4. We can see that the accuracy of prediction for the final grade A is very high, e.g., the 91.9% in the eighth week. On the other hand, the accuracy of prediction for other final grades is approximately 50%. This may be attributed to a bias in final grades because many students obtained s final grade of A, while the other final grades have few samples. Among them, the prediction for the final grade F is over 50% in the eighth week. As the number of weeks increases, the accuracy of predictions tends to increase.

Table 4: The accuracy of prediction for each final grade by RNN.

Week	1	2	3	4	5	6	7	8
All students	71.3%	79.3%	79.8%	84.6%	84.1%	84.9%	85.1%	84.6%
A	84.0%	93.4%	90.4%	95.5%	94.9%	94.3%	93.4%	91.9%
B	12.1%	6.1%	27.3%	30.3%	33.3%	42.4%	42.4%	54.5%
C	0.0%	11.8%	35.3%	41.2%	23.5%	29.4%	47.1%	41.2%
D	0.0%	14.3%	14.3%	28.6%	28.6%	42.9%	28.6%	28.6%
F	0.0%	0.0%	12.5%	0.0%	25.0%	25.0%	50.0%	50.0%

4.2 *Prediction by Multiple Regression Analysis*

We also conduct an experiment using multiple regression analysis. The data on courses held in 2015 and 2016 are used as the training data and test data, respectively.

Figure 2 shows the accuracy of prediction when applying the method using multiple regression analysis to the training data and test data for each week, along with the results by RNN. Here, we also show the accuracy of predictions for each final grade A, B, C, D, and F. In the fourth week, the accuracy for all grades is 64.7%. From the fifth week, the accuracy exceeds 80% and very close to the accuracy of predictions by RNN. As the case of RNN, the accuracy of prediction for the final grade A is high, but the accuracy for other grades is low. This tendency seems to be stronger in multiple regression analysis than with the RNN. In fact, the accuracy of predictions for grades B, C, D, and F is less than 33.3% in the eighth week. Especially, the accuracy of prediction for a grade F is 50% in the case of RNN in the eighth week, whereas it is 0% for each week by the multiple regression analysis.

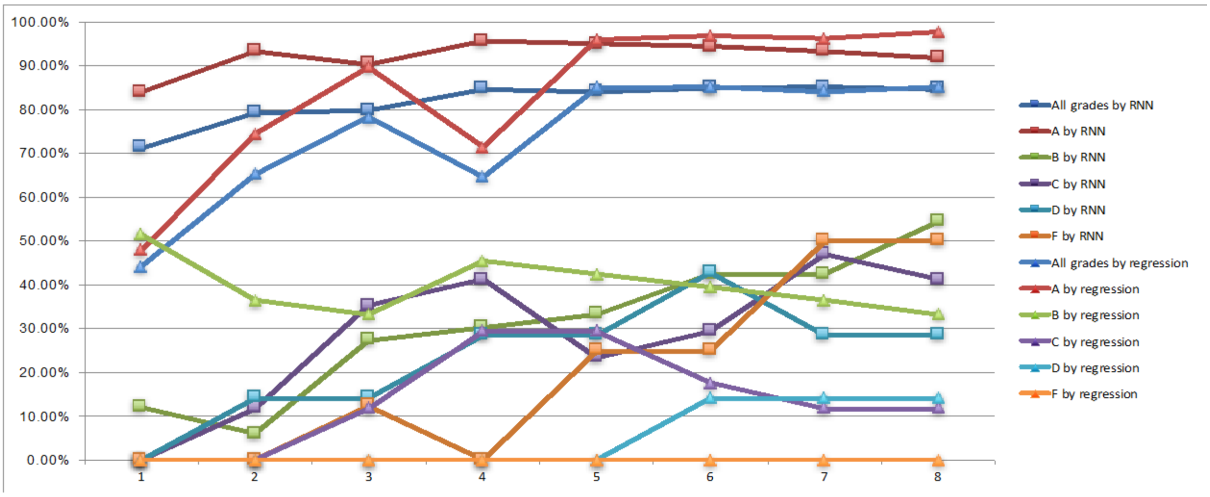

Figure 2. The accuracy of prediction by multiple regression analysis and by RNN.

5. Discussion

5.1 Consideration of the Results

From the experimental results, we can say that the accuracy of the prediction of final grades by RNN improves as the courses progresses and the data for training and predicting increases. Comparing the results by the RNN and the case of multiple regression analysis shows that the accuracy of predictions after the fifth week is almost same. However, the accuracy by the RNN before the fourth week by RNN is higher than that of multiple regression analysis. Hence, the RNN method is effective in the early prediction of final grades, which is important for both students and teachers. Moreover, the accuracy of prediction for other grades than the final grade A by RNN is generally higher. Table 3 implies that the final grade with a relatively small amount of data can be predicted by RNN. Regarding the final grade F, in particular, the accuracy of prediction is always 0% in regression analysis, whereas RNN achieves 50% accuracy in the seventh and eighth weeks.

In the experiment of this paper, we used the data on the four courses as the training data, where the two courses were taught by Teacher A, and the other two courses by Teacher B. On the other hand, the two courses treated as the test data are conducted by only Teacher A. The difference between the sets of teachers in the training data and test data can be the reason for the decrease in accuracy of predictions.

5.2 Finding the Important Learning Activities which Distinguish the Final Grades

The learning activities including in Active Learner Point that contribute to obtaining a certain final grade can be inferred from the weight of the obtained RNN. First, by observing the values of weight between each input and each GRU of a hidden layer, we can determine which learning activities are important for a unit due to firing. Similarly, by observing the values of weight between each GRU of a hidden layer and each output, we can also determine which units' firing is important for predicting the final grades of students. By combining these two observations, the items of input that are important to obtain a certain final grade can be inferred.

In our experiment, the following are implied from the obtained RNN:

- To obtain the grade A, a submission of report is important.
- To obtain the grade B, the number of actions in BookLooper and a quiz score are important.
- To obtain the grade C, the logs of Moodle and BookLooper are emphasized evenly.
- To obtain the grade D, the number of markers in BookLooper is important.
- To obtain the grade F, It is important that the number of memos in BookLooper is small.

6. Conclusion

In this paper, we presented a method to predict students' final grades using a recurrent neural network (RNN). For this purpose, we collected the learning logs from 937 students who attended one of the six courses shown in Section 2.2 for eight weeks. Two different teachers conducted these courses basically following the same syllabus. The nine selected learning logs stored in M2B system are evaluated from 0 to 5 points for each student in each week of the course, and the obtained vector of these nine evaluations is called Active Learner Point (ALP). The ALPs of students and the final grades are treated as the input and output, respectively, of the RNN.

There remain many issues to be investigated along the research direction presented in this paper. Points of particular importance include the following:

- Confirming that the prediction method shown in this paper is effective even when the set of teachers for the training data is completely different from that for test data. This enables us to help teachers who have not yet conducted a course.

- It may be valuable to determine whether giving students feedback about the results of prediction, i.e., their potential final grades, in the early stage of a course, would influence their learning activities.

Acknowledgements

The research results have been achieved by "Research and Development on Fundamental and Utilization Technologies for Social Big Data" (178A03), the Commissioned Research of National Institute of Information and Communications Technology (NICT), Japan.

References

Bodén, M. (2002). A Guide to Recurrent Neural Networks and Backpropagation, *The Dallas Project, SICS Technical Report*, pp.1-10.

Hlosta, M., Zdrahal, Z., Zendulka, J. (2017). Ouroboros: Early identification of at-risk students without models based on legacy data, *proc. LAK2017*, pp.6-15.

Marbouti, F., H. A. Diefes-Dux & Madhavan, K. (2016) Models for early prediction of at-risk students in a course using standards-based grading, *Computer & Education*, pp. 1-15.

Ogata, H., Taniguchi, Y., Suehiro, T., Shimada, A., Oi, M., Okubo, F., Yamada, M., Kojima, K. (2017). M2B System: A Digital Learning Platform for Traditional Classrooms in University, *Practitioner Track Proc. of LAK2017*, pp.154-161.

Ogata, H., Yin, C., Oi, M., Okubo, F., Shimada, A., Kojima, K. & Yamada, M. (2015). E-Book-based Learning Analytics in University Education, *proc. ICCE2015*, pp.401-406.

Okubo, F., Hirokawa, S., Oi, M., Shimada, A., Kojima, K. & Yamada, M. & Ogata, H. (2016). Learning Activity Features of High Performance Students, *Proc. of the 1st International Workshop on Learning Analytics Across Physical and Digital Spaces (Cross-LAK 2016)*, pp.28-33.

Okubo, F., Shimada, A., Yin, C. & Ogata, H. (2015). Visualization and Prediction of Learning Activities by Using Discrete Graphs, *proc. ICCE2015*, pp.739-744.

Okubo, F., Yamashita, T., Shimada, A. & Ogata, H. (2017). A Neural Network Approach for Students' Performance Prediction, *proc. LAK2017*, pp.598-599.

You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning, *Internet and Higher Education*, vol.29, pp.23-30.