

On the Prediction of Students' Quiz Score by Recurrent Neural Network

Fumiya Okubo^{1*}, Takayoshi Yamashita², Atsushi Shimada¹, Yuta Taniguchi¹ and Shin'ichi Konomi¹

¹ Learning Analytics Center, Kyushu University, Japan

² Department of Computer Science, Chubu University, Japan

*fokubo@artsci.kyushu-u.ac.jp

ABSTRACT: In this paper, we explore the factor for improving the performance of prediction of students' quiz scores by using a Recurrent Neural Network. The proposed method is applied to the log data of 2693 students in 15 courses that were conducted with following the common syllabus by 10 teachers. The experimental results show that in the case where the same teacher is not included in both training and test data, the accuracy of prediction slightly lower. We also show that at the beginning of a course, it is better to construct a prediction model including various items of learning logs, however, in the latter half, it is better to update the model by using selected information only.

Keywords: Learning log, prediction of quiz score, recurrent neural network

1 INTRODUCTION

At Kyushu University, a learning support system called the M2B system was introduced in October 2014. The M2B system consists of three subsystems, the e-learning system Moodle, the e-portfolio system Mahara, and the e-book system BookRoll, which is enable us to record various types of logs regarding learning activities. Utilizing the collected data, various investigations on learning analytics have been conducted. The details of the M2B system and our investigations are summarized in Ogata et al. (2017) and Ogata et al. (2015).

One of the most important tasks in the field of learning analytics is to find "at-risk" students who are likely to fail or drop out of class. In Baradwaj & Pal (2011), Marbouti et al. (2016), a lot of method for this purpose are intensively investigated such as Regression, Support Vector Machine, Decision Tree. It is also valuable to identify learning activities that have significant effect on obtaining a particular final grade of students (Okubo et al. (2016), You (2016)). Our research group has developed a method for predicting final grades of students by a Recurrent Neural Network (RNN) from the log data stored in the M2B system (Okubo et al. (2017b)) which is collected from courses conducted in classrooms. An RNN is a variant of a deep neural network that handles time series data, hence it is appropriate to deal with weekly learning logs of course. In Okubo et al. (2017a), using the nine types of learning logs in multiple courses following common syllabus, the accuracy of prediction by our method is confirmed.

In this paper, we explore the factor for improving the performance of prediction by RNN. The log data was collected from 15 courses following common syllabus, which were attended by all first grade students in Kyushu University in 2017. In order to observe the prediction performance in

details, the sum of quiz score of the student is regressed as the output, instead of the final grade. Moreover, we confirm the accuracy of prediction from some selected learning activities. The method and result may help teachers to find at-risk students in the course and give appropriate feedback.

2 DATA COLLECTION.

2.1 Active Learner Point

Many kinds of logs of learning activities are stored in the M2B system. To analyze and visualize these data easily, we select nine major learning activities, and evaluate them for each student from 0 to 5 points for each week of a course. The vector of these nine evaluations is called the Active Learner Point (ALP). The nine selected learning activities and the method for evaluating them are summarized in Table 1. The logs of attendance, quizzes, reports, and course views are stored in Moodle. The other logs are stored as shown in Table 1. We note that the aim of this work is to predict the sum of the quiz scores from the other learning logs. Hence, for “quiz”, instead of the criteria of the original ALP, we use “0, 1” to indicate whether a student had taken a quiz.

2.2 Courses

We collected learning logs regarding ALP from 15 courses of “Basics of Cybersecurity” in the spring quarter of 2017. Each first grade student of Kyushu University was assigned to one of these courses on the basis that he/she must attend the course. These courses were conducted, basically, by following the same syllabus for eight weeks. Through attending the courses, students study entire primary cybersecurity matters including basic technologies, laws and morals of cybersecurity. Ten teachers were in charge of these courses; hence, five teachers taught two classes and the other five teachers taught one class. In Table 2, the elementary information of the courses including the number of students who attended each course is summarized. The histogram of the total score of the quizzes in eight lectures for all students in the 15 courses is shown in Figure 1.

Table 1: Active Learner Point
(the criteria for “Quiz” is not used in this paper.)

| Activities | 5 | 4 | 3 | 2 | 1 | 0 |
|-------------------------------|------------|-----------|-----------------|-----------|-----------|---------------|
| Attendance | Attendance | | Being late | | | absence |
| Quiz (rate of collect answer) | Above 80% | Above 60% | Above 40% | Above 20% | Above 10% | o.w. |
| Report | Submission | | Late submission | | | No submission |
| Course views | Upper 10% | Upper 20% | Upper 30% | Upper 40% | Upper 50% | o.w. |
| Slide views in BookRoll | Upper 10% | Upper 20% | Upper 30% | Upper 40% | Upper 50% | o.w. |
| Markers in BookRoll | Upper 10% | Upper 20% | Upper 30% | Upper 40% | Upper 50% | o.w. |
| Memos in BookRoll | Upper 10% | Upper 20% | Upper 30% | Upper 40% | Upper 50% | o.w. |
| Actions in BookRoll | Upper 10% | Upper 20% | Upper 30% | Upper 40% | Upper 50% | o.w. |
| Word count in Mahara | Upper 10% | Upper 20% | Upper 30% | Upper 40% | Upper 50% | o.w. |

Table 2: Course information.

| Course No. | Day | Period | Teacher | # of Students |
|------------|----------|--------|---------|---------------|
| 1 | Monday | 3 | A | 204 |
| 2 | Friday | 3 | B | 206 |
| 3 | Monday | 3 | C | 188 |
| 4 | Friday | 3 | A | 217 |
| 5 | Monday | 4 | D | 201 |
| 6 | Thursday | 5 | E | 171 |
| 7 | Monday | 4 | C | 178 |
| 8 | Tuesday | 3 | F | 172 |
| 9 | Thursday | 3 | B | 170 |
| 10 | Tuesday | 3 | G | 207 |
| 11 | Thursday | 3 | H | 142 |
| 12 | Tuesday | 3 | E | 135 |
| 13 | Tuesday | 4 | F | 171 |
| 14 | Thursday | 3 | I | 151 |
| 15 | Tuesday | 4 | J | 180 |

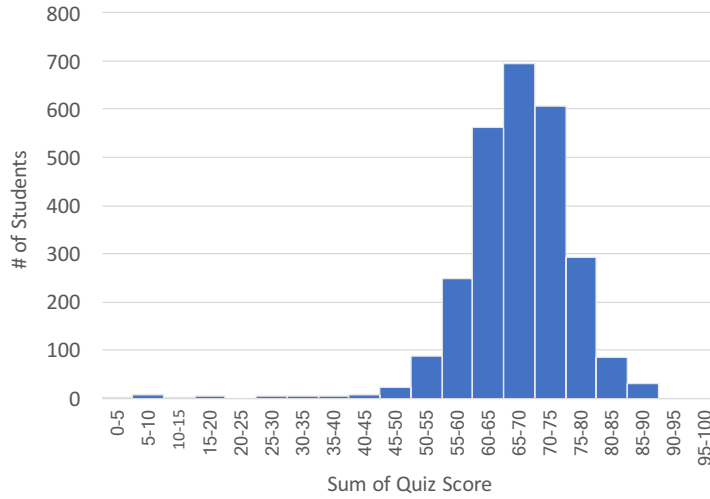


Figure 1: Histogram of the total score of the quizzes.

3 METHOD

3.1 Recurrent Neural Network

A recurrent neural network (RNN) is a variant of neural networks that handles time series data. In Figure 2 (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 2 (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.

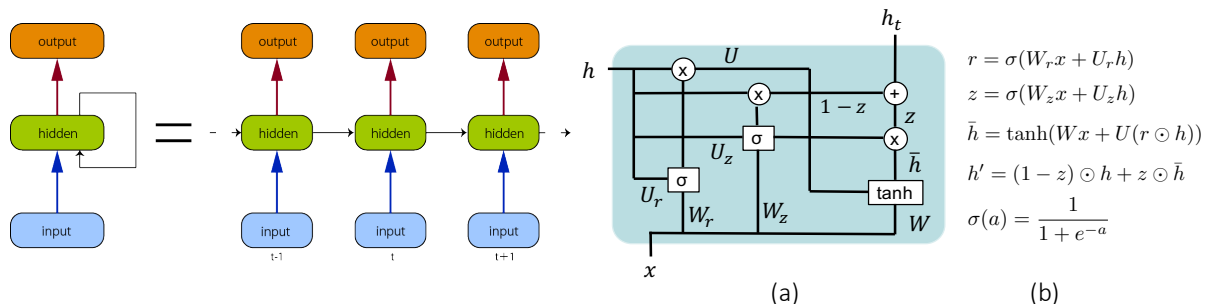


Figure 2: Structure of recurrent neural network.

Figure 3: Gated Recurrent Unit (GRU).

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. In Figure 3 (a), a graphical illustration of a structure of GRU is illustrated. A hidden layer of an RNN consists of n GRUs, where n is predefined. A GRU calculates intermediate values r , z from the input value and a value h of a hidden layer of the previous time. Then, the output value h_t is calculated from the value obtained by multiplying the input data by W , and the

intermediate value z , and the value h of the hidden layer of the previous time. Through these processes, with the formulas shown in Figure 3 (b), it can be decided whether to emphasize the input data or past data, which is then reflected in the output.

3.2 Prediction of Students' Test Score

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 student's quiz score from 0 to 100 is regressed as the output. Let the number of GRUs included in a hidden layer be 32. The time series data of the vectors of nine kinds of points is fed into the RNN, and in each time, the quiz score 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 the optimal parameters.

4 EXPERIMENTS AND DISCUSSION

To evaluate the prediction performance of the proposed method, we applied cross validation for each course, that is, the data of each course was selected as the test data and the data of the rest 14 courses were treated as the training data for an RNN. For each week, a total score of quizzes were regressed for each student in the test course. We calculated the error between the predicted value and the actual quiz score. In Figure 4, the values of errors of each course for each week are summarized. Thick line represents the average of all courses. Although there is a difference depending on course, in general, we can see that the accuracy of prediction is rising as the course progresses. Looking at the average value, the error at the end of the first week is about 5.87 points, and at the end of the 8th week, it is 3.51 points. Then, in order to compare the accuracy of the case where the same teacher is included in both training data and test data with the case of not so, we extract the courses 8, 10, 11, 14, 15 by the teachers who taught in just one course and calculate the average of these courses. The result is summarized in Figure 5.

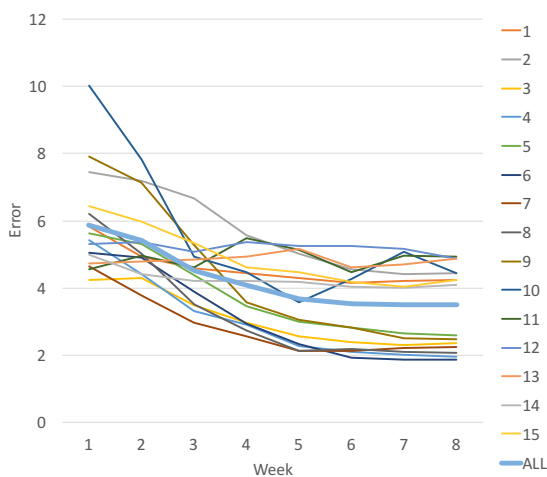


Figure 4: Difference between the predicted values and the actual data of each week.

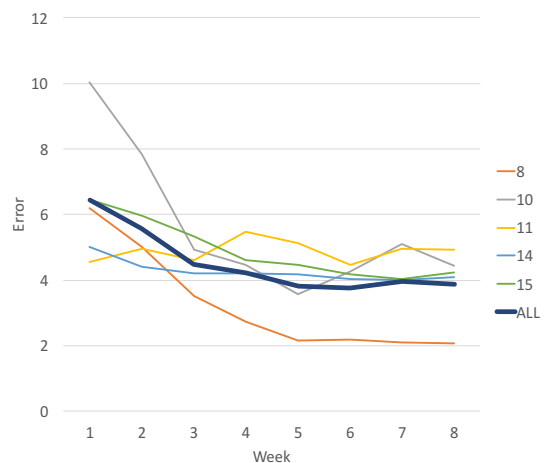


Figure 5: Difference between the predicted values and the actual data for the courses by the teachers who taught in just one course.

We also consider predicting the score of quiz from some particular items in the ALP. We selected attendance, course views and the number of actions in BookRoll that can collect from all students naturally when using the system in the course. Then, we made predictions using these three items for the courses 8, 10, 11, 14, 15. The result is summarized in Figure 6. In Figure 7, the three cases are summarized, that is, (i) the average of error of prediction using the ALP of all courses, (ii) one using the ALP of the courses 8, 10, 11, 14, 15 by the teachers who taught in just one course, and (iii) one using the data of attendance, course views and actions in BookRoll in the courses 8, 10, 11, 14, 15. Comparing the cases (i) with (ii), although the difference is not so large, it is found that the accuracy of prediction is higher in the case (i) than the case (ii) in most weeks. In the case for using the selected courses, comparing the case (iii) of using only three items in the ALP with the case (ii) of using all items in the ALP, the accuracy of (iii) is lower than (ii) in the first week, but reverses at the 4th week. At the 8th week, there is a difference of 0.67 point between (iii) and (ii). From this result, it is suggested that since there is little information at the beginning of a course, it is better to construct a prediction model including various items of learning logs, however, in the latter half, it is better to update the prediction model by continuing to collect selected important information for a long term.

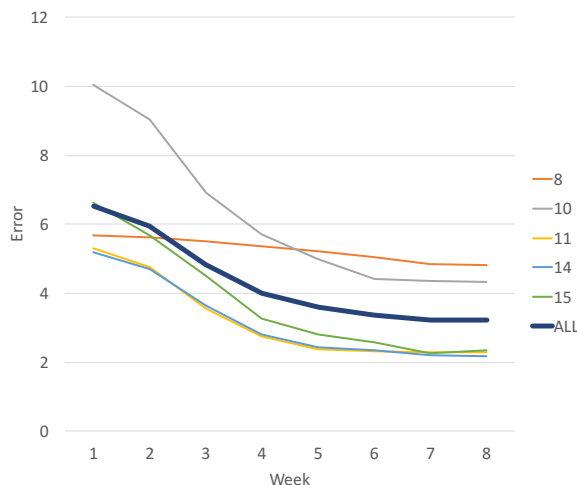


Figure 6: Difference between the predicted values by using the three items and the actual data.

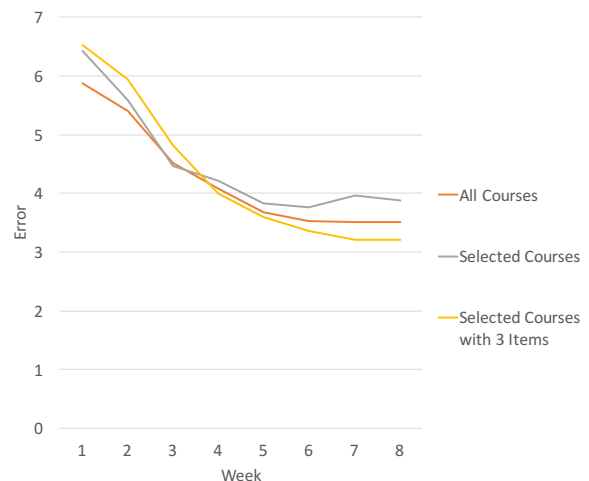


Figure 7: Difference between the predicted values the actual data for the three cases; all courses, selected courses, and selected courses with the data of attendance, course views, the number of actions.

5 CONCLUSION

In this paper, we explored the factor for improving the performance of prediction of students' quiz scores by using a Recurrent Neural Network (RNN). For this sake, the learning logs from 2693 students were collected. The nine selected learning logs stored in the 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 and the total scores of quizzes are treated

with input and output of RNN. The data of each course was selected as the test data and the data of the rest 14 courses were treated as the training data for RNN. Then, we calculated the error between the predicted value and the actual quiz score for each week. From the results, in the case where the same teacher is not included in both training and test data, the accuracy of prediction slightly lower. Next, we confirmed whether the accuracy of prediction do not become lower when using only selected items in ALP, that is, attendance, course views, and actions. The result suggests that it is important to investigate the method which enables us to select the optimal items of learning logs and to construct a prediction model at each time, automatically. Note that, in this paper, even though the teachers are different, training of RNN and prediction were carried out on multiple courses with the same syllabus. Hence, it is a future work to verify whether similar results can be obtained when predicting by using log data of completely different courses for training and test data.

The proposed method is shown to have high performance to predict students' quiz scores, however, to enhance students' performance using the results of prediction, the method and the timing of feedback may be an important problem to be investigated.

ACKNOWLEDGMENTS

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

- Baradwaj, B. and Pal, S. (2011). Mining Educational Data to Analyze Student's Performance, *International Journal of Advanced Computer Science and Applications*, vol. 6, 2, pp. 63-69.
- Marbouti, F., H. A. Diefes-Dux and 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. and 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. and 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. and 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., Yamashita, T., Shimada, A. and Konomi, S. (2017a). Students' Performance Prediction Using Data of Multiple Courses by Recurrent Neural Network, *Proc. ICCE2017*, pp.439-444.
- Okubo, F., Yamashita, T., Shimada, A. and Ogata, H. (2017b). A Neural Network Approach for Students' Performance Prediction, *proc. LAK2017*, pp.598-599., *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.