LearnerExp: Exploring and Explaining the Time Management of Online Learning Activity

Huan He
Key Laboratory of Intelligent Networks and Network Security, Ministry of Education
Xi'an Jiaotong University
Xi'an, Shaanxi, China
hehuan@mail.xjtu.edu.cn

Qinghua Zheng
School of Electronic and Information Engineering
Xi'an Jiaotong University
Xi'an, Shaanxi, China
qhzheng@mail.xjtu.edu.cn

Bo Dong
School of Continuing Education
National Engineering Lab for Big Data Analytics
Xi'an Jiaotong University
Xi'an, Shaanxi, China
dong.bo@mail.xjtu.edu.cn

ABSTRACT

How do learners schedule their online learning? This issue is concerned by both course instructors and researchers, especially in the context of self-paced online learning environment. Many indicators and methods have been proposed to understand and improve the time management of learning activities, however, there are few tools of visualizing, comparing and exploring the time management to gain intuitive understanding. In this demo, we introduce the **LearnExp**, an interactive visual analytic system designed to explore the temporal patterns of learning activities and explain the relationships between academic performance and these patterns. This system will help instructors to comparatively explore the distribution of learner activities from multiple aspects, and to visually explain the time management of different learner groups with the prediction of learning performance.

CCS CONCEPTS

• **Information systems** → **Data analytics**; • **Human-centered computing** → *Visualization systems and tools*; • **Applied computing** → E-learning.

KEYWORDS

Online learning; time management; visualization

ACM Reference Format:

Huan He, Qinghua Zheng, and Bo Dong. 2019. LearnerExp: Exploring and Explaining the Time Management of Online Learning Activity. In *Proceedings of the 2019 World Wide Web Conference (WWW '19), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3308558.3314140

1 INTRODUCTION

Accessing world class educational resources without geographical and social boundaries has become the daily behavior of online learners. With the advantages of learning anywhere and anytime, learners can schedule their learning at their own pace. However, this self-scheduled learning process also challenges their time management [1, 14]. In the less supervised online learning environment, the inappropriate scheduling of study time (e.g., absence, procrastination, etc.) may lead to reduced learning performance and even exam failures [16], which in turn would result in reduced completion rate and course rating [1]. Therefore, understanding the time management pattern of online learning activities has been an important issue for both learners and instructors.

With the support of learning management system (LMS), the detailed learning activities can be tracked via logs (e.g., viewing videos, reading materials, posting threads, etc.). Existing studies have proposed many indicators and models based on these logs to measure the time management features, such as the interval time between two tasks [1], the time that students spent on a certain task and LMS [4, 9], the remaining time between the deadlines [7], the aggregated daily activity count [14], etc. By using these indictors, recent studies examined the learners' behavioral patterns [12], self-regulated learning strategic [10, 11, 15], the influences on learning performance [16], visualization tools for supporting self-regulated learning [7], etc. These indicators and methods provide basis for measuring and improving the time management. In addition, many visualization tools have been developed to further investigate the patterns of student engagement, including SST [6], moocRP [13], DropoutSeer [3], PeakVisor [2] and iForum [5], etc.

Theses systems demonstrate the student engagement from different perspectives. However, the instructor has limited tools to customize student groups and explore each student's complete study history freely. To address this issue, we introduce the LearnerExp system for the instructor to explore the time management intuitively. First, we propose the activity vector to represent daily learning activities based on well-studied indicators and existing models. Moreover, we embed the activity vector in the calendar-like activity matrix to facilitate users understand the learning process intuitively. Second, we use grade point prediction model based on convolutional neural network to explain the relationships between learning performance and time management. Third and last, we develop a visual analytic system for course instructors to explore the temporal activity patterns of different learner groups. Main contributions of this work are summarized as follows.

- A learner activity model for describing and analyzing the time management in learning activities, which not only captures the temporal aspect of learning activities, but also supports visual exploration.
- A web-based interactive visual analytic system for exploring and explaining the time management of online learning activity based on the proposed learner activity model.

WWW '19, May 13–17, 2019, San Francisco, CA, USA

ACM ISBN 978-1-4503-6674-8/19/05.
https://doi.org/10.1145/3308558.3314140

2 LEARNER ACTIVITY MODEL

In this section, we introduce the learner activity representation and the prediction model for analyzing the time management.

2.1 The Learner Activity Representation

In online courses, the time-stamped learner-generated events are recorded by the LMS and saved in log files on a daily basis to describe the learning activity of each day, which consist of event timestamp, learner identifier, course information, operation and other data recorded by JavaScript script embedded in the LMS. With these log files, we count the day activity of each learner based on well-studied indicators. As shown in Figure 1(a), the day activity is measured by a set of indicators, including daily interaction counts of learner-course, learner-learner and learner-system, video viewing time, session time [4, 8, 14] etc.

Based on the day activity, the history of whole activities in a semester can be represented as the activity vector (Figure 1(b)). In terms of notation, we let $\boldsymbol{x}_i$ be the activity vector for learner i, where $i = 1, \ldots, N$. The dimensionality of activity vector $\boldsymbol{x}_i$ is $D \times K$, where K is the number of indicators and D is the number of days in the semester. The activity vector $\boldsymbol{x}_i$ is a complete representation of a learner's detailed activities on a daily basis, which captures the temporal aspect of learning activities. It can be further used in predicting learning performance to analyze the patterns of time management. However, this representation is not suitable for users to gain an intuitive understanding.

In order to visualize the activity vector, we first compress the $D \times K$-dimensional activity vector $\boldsymbol{x}_i$ into $D \times 1$-dimensional vector, and then embed this compressed vector in the calendar-based $W \times 7$-dimensional activity matrix by date, where W represents the number of weeks in a semester (usually $W < 28$). As shown in Figure 1(c), the compressed day activity is represented as a color-encoded cell (the darker the higher, indicates that the activity on the corresponding date is more frequent). There are many ways to compress the day activity vector, such as linear weighted aggregation, PCA-based selection, etc. In this demo, we use a weighted sum of all indicators to compress day activity according to the suggestions from instructors.

This calendar-like representation of activity matrix is consistent with the calendar that users often use. By adding some auxiliary graphical elements (e.g., segment lines, week numbers, etc.), the user can clearly understand learners' study histories throughout the semester. As shown in Figure 1(d), the learning activity history of five learners are fully recorded by the activity matrix, and different patterns of time management can be intuitively presented.

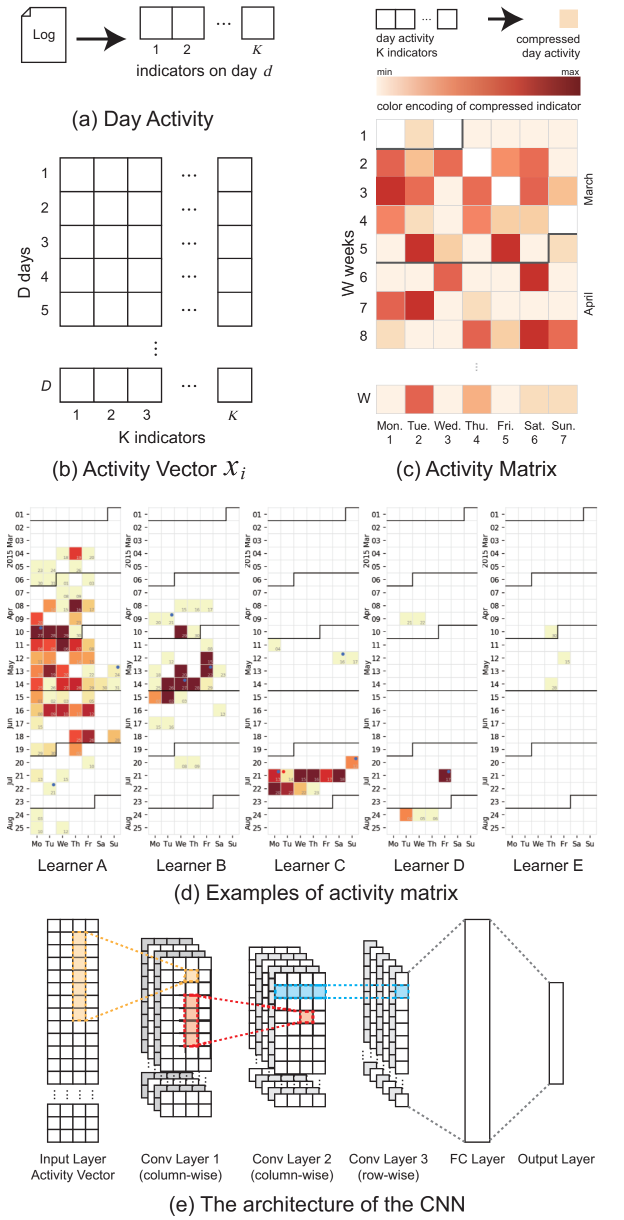

Figure 1: The learner activity representation and the prediction model based on convolutional neural network

2.2 The Grade Point Prediction Model

Since the grade point is often used in management to evaluate the learning performance, we convert exam scores into grade points from 0 to 4, which is consistent with our school practice (0 indicates unqualified, 1 to 4 respectively indicates from passing to excellent). To capture the temporal aspects of learner activity representation, we use a grade point prediction model based on convolutional neural network to analyze the relationship between time management patterns and grade point.

As shown in Figure 1(e), the model consists of three convolutional layers using the rectified linear unit (ReLU) as activation function to capture the temporal features of the activity vector. Specifically, the first two column-wise convolutional layers take convolutions that capture the short-term (e.g., 1 cell = 7 days) and long-term (e.g., 1 cell = 4 weeks) patterns of learning activity, respectively. The third row-wise convolutional layer takes convolutions that capture the pattern of activity indicators at each timepoint. Then, the fully-connected layer extracts latent representation from the convolutional layer via a sigmoid function. Finally, the last layer outputs the probabilities of each grade point by a softmax function.

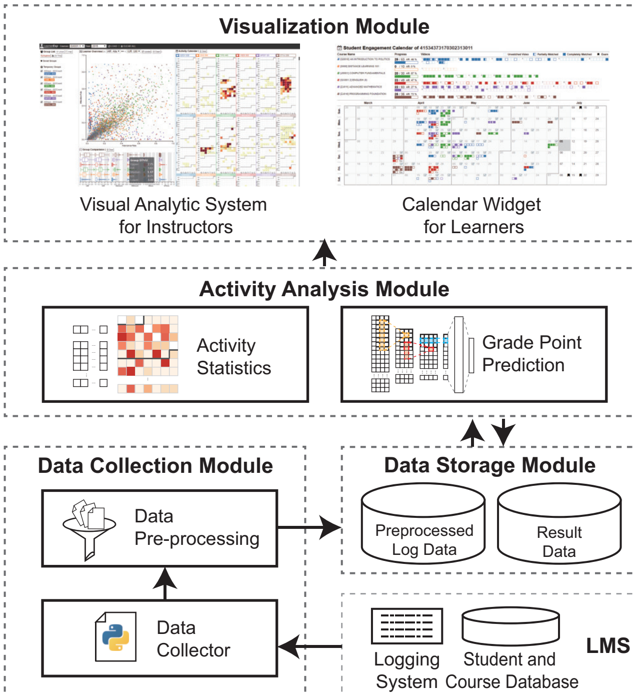

Figure 2: The architecture of the LearnerExp system

We trained this model on the dataset from the online undergraduate program of our university, which consists of 10,529 students and their learning logs in the first semester (14,942,964 lines). The average accuracy of current trained model is 72.2% and we will improve it in future research. The prediction results will be used in the exploration of the relationship between time management patterns and grade points in the LearnerExp system.

3 THE LEARNEREXP SYSTEM

3.1 The Architecture

The architecture of LearnerExp system is shown in Figure 2. It consists of four major components: *data collection module*, *data storage module*, *activity analysis module* and *visualization module*. Since we separate the different functions into modules, each module can be executed independently or integrated into the subsystem of the LMS. For example, the calendar widget in visualization module is integrated in learner's dashboard of the LMS.

Data collection and data storage module. In data collection module, we use Python scripts to collect course data and raw learner logs from the LMS. After filtering irrelevant information and filling missing fields, the preprocessed learner log data is saved in the data storage module. The personal identification information in log data is removed and will be not saved for privacy concerns. In addition, to facilitate access from the visualization module, the intermediate results of statistics and predictions are also saved in the data storage module.

Activity analysis module. This module includes two submodules: the activity statistics module measures the activity vector of each learner and generates the activity matrix defined in Section 2.1. Then, the grade point prediction module predicts the probability of each learner's grade points based on the pre-trained model mentioned in Section 2.2. The learners with a higher probability of grade 0 (i.e., failing the exam) will be grouped by the probability value for further analysis in the visualization module.

Visualization module. To make full use of the analysis results, we developed a web-based visual analytic system that enables course instructors to explore the engagement distribution and the temporal patterns of learners' activities from various aspects (Figure 3). In addition, we developed a web-based calendar widget to help learners understand their own learning activity history (Figure 2). This widget is integrated in the dashboard of the LMS, learners can check their activity history immediately after login.

3.2 Visual Design

In this section, we illustrate the visual design of the LearnerExp system and the interactions among the views. As shown in Figure 3, there are four coordinated views to show the analysis results from different perspectives: the Group List (Figure 3(a) and (b)), the Learner Overview (Figure 3(c)), the Group Comparison view (Figure 3(d)) and the Activity Calendar view (Figure 3(e)).

The Learner Overview and The Group List. To provide course instructors with intuitive impression on the overall activity distribution of all learners, we use scatter plot to present the statistics on the entire semester (Figure 3(c)). The values of each indicator over the entire semester are aggregated to represent learner's overall performance. For example, as shown in Figure 3(c), the *attendance rate* represents the ratio of the total number of daily videos watched to the total number of course videos. In this chart, each gray dot represents a learner, while each colored dot represents a learner who belongs to a specific learner group. Both the x-axis and the y-axis can be set to a specified indicator to see the distribution of same population under different indicators.

There are two types of learner groups in our system: the prediction groups generated by the model described in Section 2.2, and the temporary groups customized by selecting learners on the Leaner Overview. We developed an interactive selector to make it easier for users to intuitively select specific learners of interest (Figure 3(f)). By drawing a closed shape on the scatter plot with mouse, users can create a group of the selected dots within the shape. The newly created group will be automatically named and displayed in the "Temporary Groups" of the Group List (Figure 3(b)). To facilitate identification of learner groups in different views, the color-encoding used in the Group List is also used in other views.

The Group Comparison View. This view presents a set of box charts to simultaneously compare the distribution differences in learner groups across multiple indicators. As shown in Figure 3(d), from left to right, each box chart displays an indicator, in which each series represents a learner group from top to bottom. The color of each series is consistent with the color displayed in the Group List. When the mouse is over the series, the detailed statistics of the group on this indicator will be displayed.

The Activity Calendar View. In this view, the activity matrix defined in Section 2.1 is displayed to explore the details of learner's learning activity history. This view displays the learner groups in columns from left to right, in which the activity matrix of each

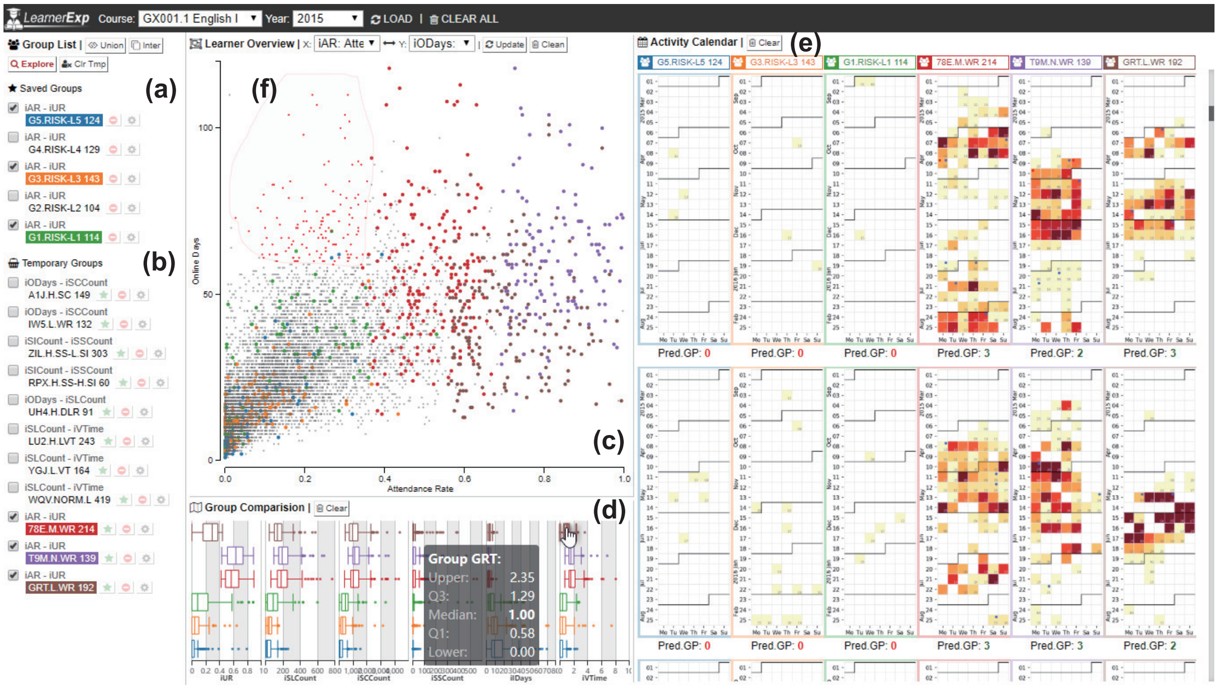

Figure 3: Screenshot of the LearnerExp system. (a) The saved group list. (b) The temporary group List. (c) The Learner Overview. (d) The Group Comparation View. (e) The Activity Calendar view. (f) A group of learners selected by the interactive selector.

learner is shown from top to bottom. When the mouse is over each activity matrix, the corresponding dot in the Learner Overview will be enlarged, and the corresponding location in each box chart will be marked to show the learner's relative performance in the whole population and in the group.

4 DEMONSTRATION

During the demonstration the attendee will evaluate LearnerExp [1] with a real dataset from our university, which is the same dataset as we mentioned in Section 2.2 but without personal privacy information. Additionally, we provide an exemplified scenario in which the course instructor wants to find the learners who may fail the posttest exam [2]. The demonstration scenario is the following.

(1) The user starts by loading data. After selecting the course in the "Course" drop-down list and clicking the "Load" button, the corresponding dataset will be loaded into the web page. All the learners of this course will be shown in the Learner Overview, and the prediction groups with different probability of failing the exam will be listed in the Group List. Figure 3(a) shows five groups, which are listed by the probability of failure of the exam from high to low.

(2) Next, the user can select groups and click the "Explore" button. Then, the selected groups will be colored, and the details will be displayed in each view. As shown in Figure 3, the learner dots belong to each group are colored and enlarged to show their relative position in the Leaner Overview (Figure 3(c)), and their study histories are listed in the Activity Calendar View (Figure 3(e)). For example, learners in the red, purple, and brown groups show regular time management characteristics and their predicted score points are higher than the blue, orange, and green groups.

(3) In addition to the prediction groups, the user can customize groups of interested learners for comparative exploration with the interactive selector (Figure 3(f)). After creating new groups, the user can select these newly created groups from the "Temporary Groups" and click the "Explore" button to continue exploring.

(4) Finally, after repeatedly exploring different groups, the user may be interested in a few customized groups and plan to analyze them in the future. By clicking the star-icon button next to the group label in the Group List, the corresponding temporary group will be saved and moved to the "Saved Groups".

5 CONCLUSION AND FUTURE WORK

In this demo, we present an interactive visual analytic system, LearnerExp, for visually exploring the patterns of online learning engagement and time management based on our proposed learner activity model. In future research, we will further expand the features of the LearnerExp and investigate the impact of different time management patterns on learning performance.

ACKNOWLEDGMENTS

This research was partially supported by "The Fundamental Theory and Applications of Big Data with Knowledge Engineering" under the National Key Research and Development Program of China with Grant No. 2018YFB1004500, the MOE Innovation Research Team No. IRT17R86, the National Science Foundation of China under Grant Nos. 61721002 and 61532015, and Project of China Knowledge Center for Engineering Science and Technology.

[1]Screencast for LearnerExp: https://youtu.be/j_mtLp-8v0U
[2]Live demo with source code: https://hehuan2112.github.io/LearnerExp

REFERENCES

[1] J. Broadbent and W. L. Poon. 2015. Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education* 27 (2015), 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007

[2] Qing Chen, Yuanzhe Chen, Dongyu Liu, Conglei Shi, Yingcai Wu, and Huamin Qu. 2016. PeakVizor: Visual Analytics of Peaks in Video Clickstreams from Massive Open Online Courses. *IEEE Transactions on Visualization and Computer Graphics* 22, 10 (Oct. 2016), 2315–2330. https://doi.org/10.1109/TVCG.2015.2505305

[3] Y. Chen, Q. Chen, Mingqian Zhao, S. Boyer, K. Veeramachaneni, and H. Qu. 2016. DropoutSeer: Visualizing learning patterns in Massive Open Online Courses for dropout reasoning and prediction. In *2016 IEEE Conference on Visual Analytics Science and Technology (VAST)*. 111–120. https://doi.org/10.1109/VAST.2016.7883517

[4] Gianni Fenu, Mirko Marras, and Massimiliano Meles. 2017. A Learning Analytics Tool for Usability Assessment in Moodle Environments. *Journal of e-Learning and Knowledge Society* 13, 3 (Sept. 2017). https://doi.org/10.20368/1971-8829/1388

[5] Siwei Fu, Jian Zhao, Weiwei Cui, and Huamin Qu. 2017. Visual Analysis of MOOC Forums with iForum. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan. 2017), 201–210. https://doi.org/10.1109/TVCG.2016.2598444

[6] Diego Alonso Gómez-Aguilar, Ángle Hernández-García, Francisco J. García-Peñalvo, and Roberto Therón. 2015. Tap into visual analysis of customization of grouping of activities in eLearning. *Computers in Human Behavior* 47 (June 2015), 60–67. https://doi.org/10.1016/j.chb.2014.11.001

[7] Kalle Ilves, Juho Leinonen, and Arto Hellas. 2018. Supporting Self-Regulated Learning with Visualizations in Online Learning Environments. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18)*. ACM, New York, NY, USA, 257–262. https://doi.org/10.1145/3159450.3159509

[8] Srećko Joksimović, Dragan Gašević, Thomas M. Loughin, Vitomir Kovanović, and Marek Hatala. 2015. Learning at distance: Effects of interaction traces on academic achievement. *Computers & Education* 87 (Sept. 2015), 204–217. https://doi.org/10.1016/j.compedu.2015.07.002

[9] Ayaan M. Kazerouni, Stephen H. Edwards, and Clifford A. Shaffer. 2017. Quantifying Incremental Development Practices and Their Relationship to Procrastination. In *Proceedings of the 2017 ACM Conference on International Computing Education Research*. ACM, 191–199. https://doi.org/10.1145/3105726.3106180

[10] René F. Kizilcec, Mar Pérez-Sanagustín, and Jorge J. Maldonado. 2016. Recommending Self-Regulated Learning Strategies Does Not Improve Performance in a MOOC. In *Proceedings of the Third (2016) ACM Conference on Learning @ Scale (L@S '16)*. ACM, New York, NY, USA, 101–104. https://doi.org/10.1145/2876034.2893378

[11] René F. Kizilcec, Mar Pérez-Sanagustín, and Jorge J. Maldonado. 2017. Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education* 104 (Jan. 2017), 18–33. https://doi.org/10.1016/j.compedu.2016.10.001

[12] Mohammad Javad Mahzoon, Mary Lou Maher, Omar Eltayeby, Wenwen Dou, and Kazjon Grace. 2018. A Sequence Data Model for Analyzing Temporal Patterns of Student Data. *Journal of Learning Analytics* 5, 1 (April 2018), 55–74. https://doi.org/10.18608/jla.2018.51.5

[13] Zachary A. Pardos and Kevin Kao. 2015. moocRP: An Open-source Analytics Platform. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale*. ACM Press, 103–110. https://doi.org/10.1145/2724660.2724683

[14] Jihyun Park, Renzhe Yu, and Fernando Rodriguez. 2018. Understanding Student Procrastination via Mixture Models. In *Proceedings of the 11th International Conference on Educational Data Mining*. 11.

[15] Masanori Yamada, Yoshiko Goda, Takeshi Matsuda, Yutaka Saito, Hiroshi Kato, and Hiroyuki Miyagawa. 2016. How does self-regulated learning relate to active procrastination and other learning behaviors? *Journal of Computing in Higher Education* 28, 3 (Dec. 2016), 326–343. https://doi.org/10.1007/s12528-016-9118-9

[16] Tzu-Chi Yang, Meng Chang Chen, and Sherry Y. Chen. 2018. The influences of self-regulated learning support and prior knowledge on improving learning performance. *Computers & Education* 126 (Nov. 2018), 37–52. https://doi.org/10.1016/j.compedu.2018.06.025