Kokoç, M., Akçapınar, G., & Hasnine, M. N. (2021). Unfolding Students' Online Assignment Submission Behavioral Patterns using Temporal Learning Analytics. *Educational Technology & Society*, *24* (1), 223-235.

Unfolding Students' Online Assignment Submission Behavioral Patterns using Temporal Learning Analytics

Mehmet Kokoç[1], Gökhan Akçapınar[2*] and Mohammad Nehal Hasnine[3]
[1]School of Applied Sciences, Trabzon University, Turkey // [2]Faculty of Education, Hacettepe University, Turkey // [3]Research Center for Computing and Multimedia Studies, Hosei University, Japan //
kokoc@trabzon.edu.tr // gokhana@hacettepe.edu.tr // nehal.hasnine.79@hosei.ac.jp
[*]Corresponding author

ABSTRACT: This study analyzed students' online assignment submission behaviors from the perspectives of temporal learning analytics. This study aimed to model the time-dependent changes in the assignment submission behavior of university students by employing various machine learning methods. Precisely, clustering, Markov Chains, and association rule mining analysis were used to analyze students' assignment submission behaviors in an online learning environment. The results revealed that students displayed similar patterns in terms of assignment submission behavior. Moreover, it was observed that students' assignment submission behavior did not change much across the semester. When these results are analyzed together with the students' academic performance at the end of the semester, it was observed that students' end-of-term academic performance can be predicted from their assignment submission behaviors at the beginning of the semester. Our results, within the scope of precision education, can be used to diagnose and predict students who are not going to submit the next assignments as the semester progresses as well as students who are going to fail at the end of the semester. Therefore, learning analytics interventions can be designed based on these results to prevent possible academic failures. Furthermore, the findings of the study are discussed considering the development of early-warning intervention systems for at-risk students and precision education.

Keywords: Precision education, Temporal learning analytics, Educational data mining, Assignment submission behavior, Learning performance

1. Introduction

A deeper understanding of online learning experiences is required for learning designers and researchers. Studies on theory and practice on how students learn individually or in groups in online environments by analyzing students' trace data have increased in recent years (Yang et al., 2020). In the last decade, learning analytics (LA) studies employing machine learning methods have been carried out to gain actionable insights such as at-risk students' detection, learning outcome assessment, and drop-out detection for improving the teaching quality and learning process. Precision education is known to be a relatively new discipline in higher education that uses the core philosophies of LA and data-driven methods. Precision education is, as addressed by (Yang, 2019), a new challenge for conventional LA, machine learning, and artificial intelligence for solving critical aspects in online education such as spotting at-risk, drop-out, low-engaged students as early as possible by analyzing online learning behaviors (for instance, assignment submission pattern, and engagement with learning materials). Precision education contributes towards maximizing students' online learning experiences and value proposition, and therefore, it uses data from the latest learning technology and integrates student support processes to ensure the highest quality teaching (Wilson & İsmaili, 2019). One of the goals of precision education is to predict students' learning performance by analyzing their online learning behaviors and providing timely intervention for supporting their learning process (Lu et al., 2018). Furthermore, precision education can be leveraged to uncover various critical aspects of education including behavioral, cognitive and emotional.

While precision education emphasizes employing artificial intelligence and other data-driven methods on large-scale datasets collected from technology-enhanced learning environments (i.e., learning management systems, digital textbooks), data about assignment submission behavior can be explored more within the scope of precision education. Students' online assignment submission behavior is a meaningful part of online learning experiences (Akçapınar & Kokoç, 2020) and has a relationship with procrastination (Yang et al., 2020). Students' online assignment submission behavior could reveal information about learning behavior such as how students' behavioral patterns of online assignment submission change over time or relationships between students' online assignment submission behaviors and their learning performance. These insights on learning behavior are crucial for teachers to monitor their students' learning progress, particularly to spot at-risk or inattentive students as early as possible. Therefore, modeling students' online learning behaviors hidden in the learning traces is an important LA contribution for precision education. Thus, modeling students' online assignment submission behavior using temporal analysis techniques can provide important insights into the online learning process and help teachers to plan timely interventions for procrastinators and/or at-risk students for precision education. Furthermore, the temporal aspect of online assignment submission behavior in precision education has much to offer in diagnosing students' learning behavior, however, not been explored much.

As of now, much effort has given to explore learning behavior patterns and predict learning performances based on interaction data, far too little attention has been paid to analyzing temporal and sequential aspects of trace data of students (Chen, Knight, Wise, 2018; Olsen, Sharma, Rummel, & Aleven, 2020). Several studies have focused mainly on aggregated data (e.g., the total number of events) without considering temporal aspects of online learning behaviors (Juhaňák, Zounek, & Rohlíková, 2019). To the best of our knowledge, a limited number of studies detect patterns in students' online assignment submission behaviors using temporal analysis techniques. Considering the importance of temporal analytics in precision education for diagnosing students' learning, behavioral patterns, and learning performance prediction, this study explored students' online assignment submission behavior patterns by using clustering, Markov Chains, and association rule mining analysis. With these analyses, this study aimed to contribute precision education literature to investigate whether these patterns can be used to diagnose and predict at-risk students (e.g., who are not going to submit next assignment and low-performers) as early as possible. Our study addressed the following research questions:

RQ1. What are the students' behavioral patterns of online assignment submission?
RQ2. How do students' behavioral patterns of online assignment submission change over time?
RQ3. What are the association rules between students' online assignment submission behaviors and their learning performance that can be used to predict at-risk students as early as possible?

This paper aims to employ educational data mining methods for precision education to uncover a core focus of precision education, namely, understanding learning behavior while the semester progresses. More precisely, this paper aimed to diagnose and predict at-risk students based on their online assignment submission behavior over time using temporal LA. Students' online submission behavioral data were collected from Moodle and analyzed with regards to- how students' assignment submission behavior changes over the period of time, finds association between their assignment submission and final score, and analyzes the factors affecting students' learning performance at the end of the semester; and visualizes students' assignment submission patterns so that the teacher can get an early insight about the students. Thus, the predictive models and the findings of this study contribute to the core of precision analytics.

2. Background and literature review

2.1. Precision education

Employing artificial intelligence and machine learning techniques in education and psychology has led to significant developments in related fields such as educational intelligence, self-regulated learning, and precision education. Depending on the developments in information and communication technologies, a paradigm shift in learning and teaching has occurred and new pedagogical models have emerged. One of the new educational models considering personalized learning is precision education. Precision education can be defined as a new challenge of applying artificial intelligence, machine learning, and LA for improving teaching quality and learning performance (Yang, 2019). Precision education aims to analyze educational and learner data, predict students' performance and provide timely interventions based on learner profiles for enhancing learning (Lu et al., 2018). For effective learning design in precision education, LA has contributed not only to the dashboards and intervention tools but also as the conceptual frameworks guiding research experiences.

The ultimate goals of using LA are to increase student success and improve students' online learning experience (Pardo & Dawson, 2016). Studies in LA and precision education literature have provided new findings based on multimodal data and actionable knowledge to increase the learning/teaching context's effectiveness. There have been several attempts (e.g., Azcona, Hsiao, & Smeaton, 2019; Tsai et al., 2020) to explore students' interaction and behavioral patterns, to predict students' learning performance based on their online learning behaviors, to develop early-warning systems for at-risk students, to support students and teachers decision-making processes, and to investigate effects of interventions and LA dashboards. The results of the aforementioned studies indicate that LA provides important clues about students' online learning experiences and LA tools offer personalized recommendations to students by visualizing and analyzing their trace data to optimize and improve learning. It is clear that LA and employing educational data mining methods in educational studies contributes to our understanding of learning.

While many studies have been carried out on profiling learners and prediction learning performances based on interaction data, less attention has been paid to analyzing temporal and sequential aspects of trace data of students (Chen, Knight, & Wise, 2018; Juhaňák, Zounek, & Rohlíková, 2019). Rather than modeling the frequency of clicks and interaction of students in an online learning environment, students' learning paths need to be modeled based on time and probability (Cerezo, Sánchez-Santillán, Paule-Ruiz, & Núñez, 2016). Thus, there is an important gap in the relevant field in terms of behavior modeling. To overcome this gap, event logs reflecting students' learning experiences have been modeled using temporal analysis and temporal LA approach (Knight, Wise, & Chen, 2017). The following section is about temporal LA and its implementation in the educational context. In precision education, the diagnosis of online learning behavior patterns for using predictive student modeling is vital to provide students real-time intervention.

2.2. Temporal LA and its role in the educational context

Literature in the educational contexts indicates that both individual and collaborative learning do not happen in one moment (Knight, Wise, & Chen, 2017). In general, learning happens over a period, which is referred to as a process. Temporal characteristics of students' learning data contain valuable insights about the time period or process of occurrence of particular events (Mahzoon et al., 2018). Thus, analyzing time-related data rather than just frequencies gives more information about the learning process (Knight, Wise, & Chen, 2017). The temporal analysis of students' learning data provides a more in-depth insight into individual and collaborative learning processes (Nguyen, Huptych, & Rienties, 2018; Olsen et al., 2020). What makes temporal analysis vital in online and blended learning is that modeling transitions between different students' actions considering temporal changes enhance our understanding of online learning behavioral patterns. Also, temporal analysis supports a more robust prediction model of students' learning performance to make timely interventions for precision education.

In the temporal analysis, various techniques are employed for modeling students' behaviors extracted from their trace data include process mining, sequential pattern mining, Markov chains, and hidden Markov models. While process mining discovers a process model from the students' activity sequences, sequential pattern mining finds the most frequent patterns through a range of action sequences. Markov chains aggregates sequences of students' actions into transition models and hidden Markov models have been used for discovering students' behavioral patterns considering transitions over time (Boroujeni & Dillenbourg, 2019). There is a significant difference between time-series analysis and temporal LA. While time-series analysis typically looks for recurring patterns within a time period for numeric features (Mahzoon et al., 2018), temporal analytics methods help researchers analyze dynamic student data and mode student behaviors over time at different levels of granularity.

There is an increasing trend of temporal analytics methods being used to diagnose students' online learning behavior patterns and predict their learning performance based on temporal data for planning timely interventions (Cheng et al., 2017; Juhaňák, Zounek, & Rohlíková, 2019; Matcha et al., 2019). Previous studies have shown that temporal analytics is beneficial to predict students' learning performance (Papamitsiou & Economides, 2014), to diagnose of learning patterns and behaviors (Boroujeni & Dillenbourg, 2019), to identify at-risk learners (Mahzoon et al., 2018), to detect learning tactics and strategies (Matcha et al., 2019) and to explore the relationship between students' timing of engagement and learning design (Nguyen, Huptych, & Rienties, 2018). While the importance of analyzing students' temporal trace data in online and blended learning has great potential in improving educational practice, applying temporal analytics to student data is less explored in educational research (Chen, Knight, & Wise, 2018; Knight, Wise, & Chen, 2017). To date, the temporal analysis of trace data has been mostly employed in modeling students' online behaviors in the LA field (Juhaňák, Zounek, & Rohlíková, 2019). These studies highlight the critical role of temporal analysis of trace data in diagnosing online learning behaviors and predicting students' further actions. Although temporal analysis has been used to unlock students' online learning behaviors such as quiz-taking, content navigation, e-book reading, and video viewing, few studies have paid attention to exploring online assignment submission behavior patterns. Therefore, in our study, we intended to use the temporal LA method to model students' online learning behavior patterns, specifically students' trace data while engaging in online assignment activities.

2.3. Online assignment submission behaviors

There is an increasing demand for online assignments to assess the learning process and evaluate learning performance. Submission of online assignments is one of the most performed online learning activities by students (Cerezo et al., 2016). In addition, assignment activity is a commonly used LMS component in blended learning environments and fully online courses (Azcona, Hsiao, & Smeaton, 2019). Moreover, several studies have shown that number of submitted online assignments, assignment scores, and interaction with assignments are predictors of students' learning performances (Lu et al., 2018; Zacharis, 2015). According to a study that modeled LMS-generated interaction data, students' interaction with assignments and learning tasks are vital parts of their learning experiences (Kokoç & Altun, 2019). Since online assignments play a meaningful role both in evaluating to what extent students understand the course subjects and practicing a course topic (Tila & Levy, 2020), online assignment submission behavior can have crucial consequences for learning process assessment. Thus, the diagnosis of students' online assignment submission behaviors has been the subject of much attention in the literature. Previous studies indicated that students who uploaded their assignments previous to the submission deadline had been better online learning experiences and higher course performance (Akçapınar & Kokoç, 2020; Paule-Ruiz, Riestra-González, Sánchez-Santillán, & Pérez-Pérez, 2015).

One of the key educational aspects that makes online assignment submission times vital for precision education is the early identification of students with procrastination tendencies (Yang et al., 2020). Students' online assignment submission times have been added to the LA indicators as a proxy measure of academic procrastination for identifying students at risk of failure (Cormack, Eagle, & Davies, 2020). For example, Yang et al. (2020) predicted students' academic performance through submission pattern data reflecting their procrastination behaviors with an accuracy of 97%. Additionally, previous studies showed that delaying online assignment submission as a procrastination behavior resulted in lower grades (Cerezo, Esteban, Sánchez-Santillán, & Núñez, 2017; Cormack, Eagle, & Davies, 2020). This indicates the importance of analyzing online assignment submission behavior to identify at-risk and procrastinator students for precision education.

Previous studies indicated that the late completion of an online assignment was associated with lower academic performances and procrastination tendencies (Cormack, Eagle, & Davies, 2020; Yang et al., 2020). Whereas online assignment submission behavior is essential for the prediction of students' learning performance and understanding their online learning experiences, little is still known about it from temporal LA perspectives. To the best of our knowledge, only one study by Akçapınar and Kokoç (2020) analyzed students' online assignment submission behaviors and found that three clusters emerged based on submission behaviors and most of the students who did not submit the assignment failed in the blended course. Although this study provides valuable results on the assignment submission behavior process, more LA research is needed to expand our understanding of online assignment submission behavior in an online and blended learning environment, especially following temporal analysis and modeling (Azcona, Hsiao, & Smeaton, 2020; Yang et al., 2020). Understanding the process of students' online assignment submission behavior can provide important insights into an effective personalized/adaptive learning environment and help teachers to plan timely interventions for procrastinators and/or at-risk students for precision education. Thus, our study aims to better understand students' online assignment submission transition behaviors by visualizing the patterns and predicting their further assignment behaviors in a blended learning course. We hope that the study sheds some light on online assignment submission behavioral patterns and provides actionable knowledge to design timely interventions for improving learning.

3. Method

In order to answer the research questions, students' assignment submission data were analyzed using state-of-the-art educational data mining techniques including clustering, Markov Chains, and association rule mining. Markov models and clustering and predictive analysis are commonly used in precision education research as they can generate easy-to-understand models to diagnose and predict at-risk students on time by analyzing their behavioral data collected from the educational learning environments (Boroujeni & Dillenbourg, 2019). These methods can also help researchers to understand the transition probabilities of different students' behaviors that can be valuable to plan further interventions to prevent possible academic failures. The employed combined method allows us to obtain interpretable models to understand the students' assignment submission behavior, its relation with the academic performance, and changes that happened over time. The data collection and data analysis processes are explained in detail in the following sections.

3.1. Participants and context

The data were collected from an Operating Systems course offered by a public university in Turkey. A total of sixty-nine students participated in the study. In this course, Moodle was actively used as a part of the lecture delivery together with face-to-face lessons. The students' activities in Moodle can be summarized as following the course resources, participating in the discussions, and doing assignments. The assignments included open-ended questions related to the weekly topics. The purpose of the assignments was to make the students come prepared for the class. Students are given five-six days before the class to complete the assignments. The starting time of the class was set as the deadline for the assignment of the last week. During the semester, 10 assignments were given to the students. In this study, the data related to the assignment given to the students in the 4th, 6th, 8th, and 10th week were analyzed. These assignments are chosen because they are directly related to course objectives. The instructor prepares questions in quizzes to promote students' use of higher-order thinking skills such as remembering, understanding, applying, analyzing, revising, and creating. An example of a question related to the disk scheduling topic is given below. In order to answer this question, the students must know how the disk scheduling algorithms work and apply them to the given context.

Example Question: Let's take an example where the queue has the following requests with cylinder numbers as follows: 90, 198, 27, 112, 16, 104, 69, and 60. Assume the head is initially at cylinder 50. Sort incoming requests according to the SSTF (shortest-seek-time-first) algorithm.

The students submitted their assignments through the Quiz module in Moodle. Among 69 students, 48 students submitted the first assignment, 57 students submitted the second assignment, 50 students submitted the third assignment, and 48 students submitted the fourth assignment. The events that students can perform in the assignment submission process are presented in Table 1. All the activities related to these events were logged in Moodle's database with a time stamp.

Table 1. Activities that the students can perform in the assignment submission process

Event	Description
Assignment viewed	The student viewed the assignment module, saw the assignment description, but did not open the questions.
Attempt started	This is only the case when the student views the assignment for the first time, and this does not happen again on subsequent visits.
Question viewed	The student's displaying each question in the assignment is logged in this way. Displaying the question also means recording the text in the answer field.
Assignment submitted	This happens when the student completes the assignment. The student can submit the assignment once and then cannot change the answers.
Question reviewed	If the student displays the assignment after the deadline, it will be labeled as a review. At this stage, the student can view the answer s/he gave or see the grade if the assignment is graded.

Within the scope of RQ3, the final grades of the students for the Operating Systems course were considered as an indicator of academic performance. Students took two written exams (i.e., first in the midterm and second in the final exam) during the semester. Apart from that, they received assignments regularly in Moodle during the semester. The students' final grades were calculated by taking 25% of the midterm exam, 25% of their assignment scores in Moodle, and 50% of the final exam. The final score was used in the data analysis by categorizing it as "Passed" and "Failed." The grades were categorized as "Failed" (n = 30, final score < 50) and "Passed" (n = 39, final score ≥ 50) considering the indicators in the undergraduate regulations of the university.

3.2. Data pre-processing and feature extraction

A total of 9633 activities of 69 students who submitted their assignments before the deadline are exported from Moodle's database. The log sequence for a student can include all the events given in Table 1. Also, *Assignment viewed*, *Question viewed*, and *Question reviewed* events can take place more than once in a log sequence. Among the examined records, the shortest log sequence contains only 4 records, while the longest log sequence consists of 268 records. While an average log consists of 45 records, the median value is 39. An example of a log sequence consisting of 14 records of a student is as follows: *Assignment viewed -> Attempt started -> Question viewed -> Question viewed -> Question viewed -> Question viewed ->* Question viewed *-> Question viewed -> Assignment submitted –> Assignment viewed -> Question reviewed-> Question reviewed-> Question reviewed-> Question reviewed*. During the data pre-processing, Moodle log records are processed and features were extracted for each student. This operation was repeated four times for each assignment. Description of the extracted features are given in Table 2. These features were selected in the light of existing literature (Akçapınar & Kokoç, 2020; Cerezo et al., 2017; Stiller & Bachmaier, 2019). For example; time-related features (e.g., Duration, Time taken) were selected since previous studies showed that time spent on a task is an important feature while identifying at-risk students as well as understanding their motivation and competencies in metacognitive learning strategies (Stiller & Bachmaier, 2019). Features related to procrastination behavior (e.g., Started on, Completed) were also found to be effective while clustering students based on their assignment submission behaviors (Akçapınar & Kokoç, 2020) and predicting their academic achievements (Cerezo et al., 2017).

Table 2. Features used in the study and their descriptions

Feature	Description
Attempt count	The number of time student view the questions.
Duration	The amount of time a student spends on an assignment (in minutes).
Started on	The difference between the date and time the assignment was started and the due date (in hours).
Completed	The difference between the date and time the assignment was submitted and the due date (in hours).
Time taken	The amount of time it took the student to start and submit the assignment (in hours).

3.3. Data analysis

The study used cluster analysis to group the students according to similar assignment submission behaviors. As a temporal analysis, Markov Chains were conducted to model transition behaviors of online assignment submission, and association rule mining was used to build predictive rules based on the students' behaviors and academic performances. Since the contents, question types, and the numbers of the questions are varied in different assignments, the students' assignment submission behaviors are clustered independently for each assignment. To map the clusters in different assignments, each assignment should have the same number of clusters and features. The clustering process was carried out with categorical data. Hence, all features were categorized into three levels using the equal interval method. Data analysis and visualizations were performed using the R data mining tool (R Core Team, 2017). Specifically, cluster analysis was carried out using the K-Modes algorithm with the help of the R package named *klaR*. Markov Chains analysis was performed using the *Markov Chain* package and the association rule mining analysis was performed using the *arules* package.

4. Results

4.1. What are the students' behavioral patterns of online assignment submission? (RQ1)

Within the scope of the second research question, it was investigated whether the students' homework submission behavior changed over time. For this purpose, students were divided into three clusters for each assignment independently. The number of clusters determined to be three due to the high interpretability of having *high*, *medium*, and *low* engaged clusters. Whether the three clusters solution fits the data is validated visually using the Elbow method. The scaled cluster centers' distributions formed after the cluster analysis are presented in Figure 1 for each assignment. The cluster centers showed that students displayed similar patterns in all four assignments. For example, the students in the second cluster in Assignment1 and the students in the first cluster in Assignment2, the students in the third cluster in Assignment3, and the students in the first cluster in Assignment4 displayed the same pattern. The prominent features of these students are- they start the assignment at the last moment (*StartedOn*), spent less time to complete the assignment (*Duration*), and the number of questions displayed (*AttemptCount*) is less. In other words, the students in these clusters submitted the assignment, but they gave a minimum effort for the assignment. Similarly, the students in Cluster3 in Assignment1, the students in Cluster2 in Assignment2, the students in Cluster1 in Assignment3, and the students in Cluster2 in Assignment4 also displayed a similar behavioral pattern. The prominent features of these students are- they started the assignment much earlier than the given deadline (*StartedOn*), spent more time to complete the assignment (*Duration*), there is a significant difference between the start and end time of the assignment (*TimeTaken*), and the number of question views (*AttemptCount*) is much higher than the other students. Although most of the students in these clusters complete their assignment submission on the last day, they start working on the assignment much earlier than the other students and they make much more effort to complete the assignment. Finally, it is observed that the students in Cluster1 in Assignment1, in Cluster3 in Assignment2, in Cluster1 in Assignment3, and in Cluster3 in Assignment4, exhibit similar assignment submission patterns. Like the students in the first group, these students start their assignment submission near the deadline (*StartedOn*), but they spend more time completing the assignment than the first group.

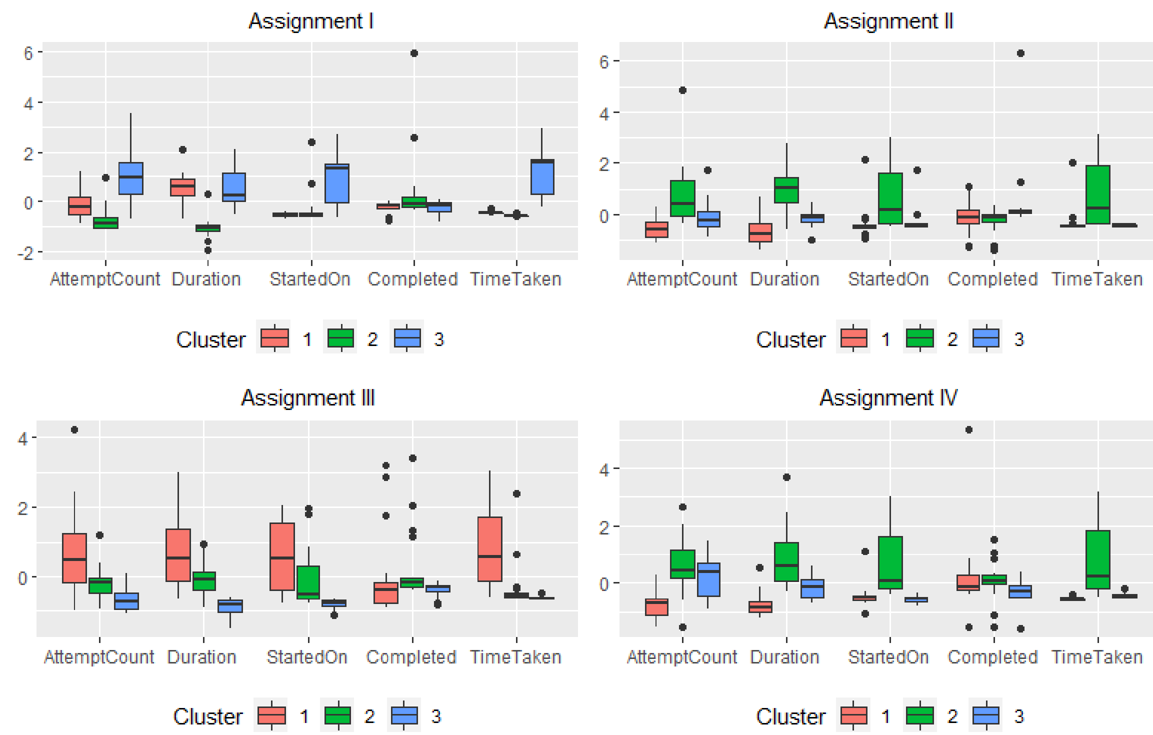

Figure 1. Box plots of features in different clusters for each assignment

In further analysis, similar clusters in each assignment were labeled as *High*, *Medium*, and *Low* in order to analyze students who followed a similar assignment submission pattern. Students who did not submit their assignments are labeled as *None*. Regarding this analysis, Cluster3 in Assignment I, Cluster2 in Assignment II, Cluster1 in Assignment III, and Cluster2 in Assignment IV are mapped to the High group. Cluster1 in Assignment I, Cluster3 in Assignment II, Cluster2 in Assignment III, and Cluster3 in Assignment IV are mapped to the Medium group. Cluster2 in Assignment I, Cluster1 in Assignment II, Cluster3 in Assignment III, and Cluster1 in Assignment IV are mapped to the Low group. Students who did not submit their assignments were manually assigned to the None group. The distribution of students in each group for all assignments are shown in Table 3.

Table 3. The number of students in each cluster after mapping

Cluster	Assignment I	Assignment II	Assignment III	Assignment IV
High	14	18	18	19
Medium	18	14	19	11
Low	16	25	13	18
None	21	12	19	21
Total	69	69	69	69

4.2. How do students' behavioral patterns of online assignment submission change over time? (RQ2)

Within the scope of the second research problem, it was investigated whether the homework submission behavior of the students changed over time. For this purpose, firstly, the transition between the sets in which the students took part in different assignments is visualized in Figure 2. As seen in the graph, there are transitions between High-Medium, Medium-High, Medium-Low, Low-Medium, Low-None, and None-Low states. On the other hand, it is also noticed that there are limited transitions between High-Low, Low-High, High-None, None-High, Medium-None, and None-Medium states. Markov Chains analysis was used to analyze the transitions between different states in more detail. In this way, the student's probabilities of transition from None, Low, Medium, or High status in one assignment to None, Low, Medium, or High status in another assignment were calculated. The values calculated for Assignment1-Assignment2, Assignment2-Assignment3, and Assignment3-Assignment4 transitions are presented in Figure 3.

As stated earlier, we clustered students in High, Medium, Low, and None after mapping their assignment submission behavior. Hence, the Markov Chain analysis in Figure 3 shows the actual transition probabilities between the groups across the semester.

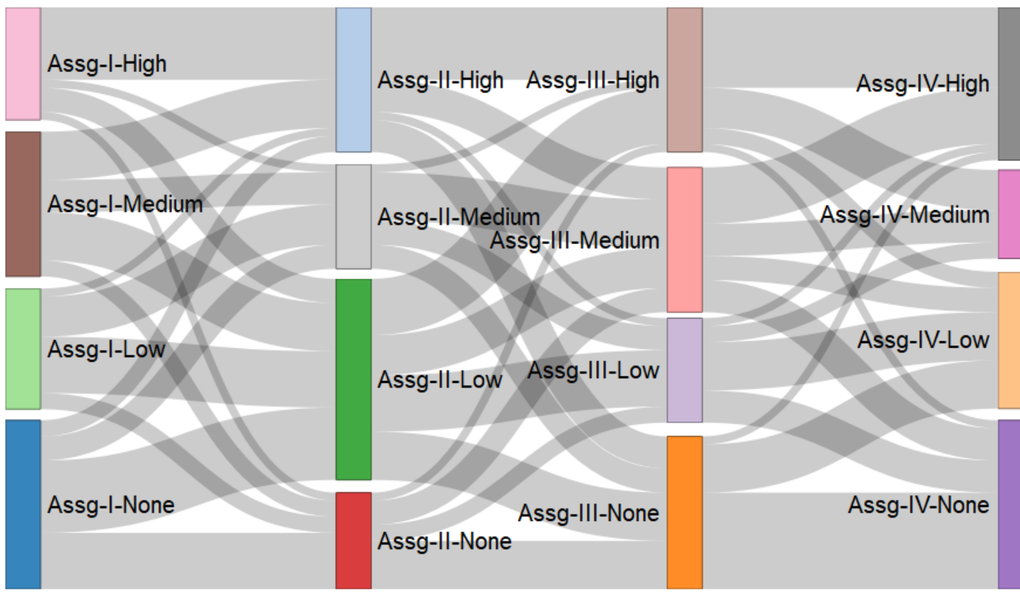

Figure 2. The students' assignment submission behaviors over time

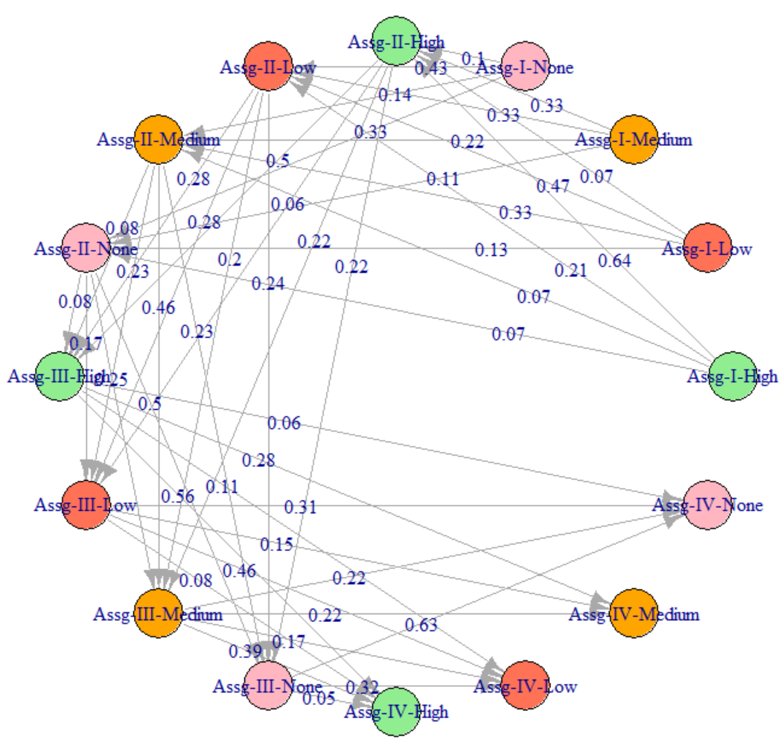

Figure 3. Transition probabilities among different assignments

The arrow between the groups indicates the direction of the transition and the numerical values represent the probability of the transition between each group. The highest probability of each transition is 1 (that is, 100%). Our Markov Chains analysis uncovered some important assignment submission behaviors of the students; therefore, we elaborate four key transitions, namely High-to-None, High-to-Low, None-to-High, and Low-to-High. For High-to-None transition, the Markov Chains analysis indicates that- students in the High cluster who submitted Assignment I have the transition probability of 0.07 to be in the None cluster in their Assignment II submission. This means the High-to-None cluster transition is like this that only 7 out of 100 students will not submit their Assignment II who belonged to the High cluster in their Assignment I submission. Consequently, for High-to-Low transition, students in the High cluster who submitted Assignment I will have a 0.21 (i.e., 21 students out of 100) transition probability to be in the Low cluster in their Assignment II submission. Similarly, for the None-to-High transition, the probability is 0.1. This means, only 10 out of 100 students who belonged to the None cluster in their Assignment I submission will be in the High cluster in their Assignment II submission. In the case of Low-to-High transition behavior, we found that the transition probability of assignment submission is 0.07 (7 out of 100 students) between Assignment I's Low cluster and Assignment II's High cluster.

4.3. What are the association rules between students' online assignment submission behaviors and their learning performance that can be used to predict at-risk students as early as possible? (RQ3)

RQ3 was answered using Association Rule Mining (ARM) analysis. The rules related to passing and failing the course were filtered among the found rules. As a result, 20 rules for students who passed the course and 14 rules for students who failed the course were obtained. The list of rules obtained and Support, Confidence, and Lift values for each rule are presented in Table 4.

Table 4. The list of the association rules extracted

No	LHS		RHS	Support	Confidence	Lift
1	{Assg-IV-Medium}	=>	{Passed}	0.16	1.00	1.77
2	{Assg-II-High,Assg-IV-High}	=>	{Passed}	0.16	1.00	1.77
3	{Assg-III-High,Assg-IV-High}	=>	{Passed}	0.14	1.00	1.77
4	{Assg-I-High,Assg-II-High}	=>	{Passed}	0.13	1.00	1.77
5	{Assg-II-High,Assg-III-High}	=>	{Passed}	0.13	1.00	1.77
6	{Assg-I-High,Assg-III-High}	=>	{Passed}	0.12	1.00	1.77
7	{Assg-I-Medium,Assg-III-High}	=>	{Passed}	0.12	1.00	1.77
8	{Assg-I-Medium,Assg-IV-High}	=>	{Passed}	0.10	1.00	1.77
9	{Assg-I-High,Assg-II-High,Assg-IV-High}	=>	{Passed}	0.10	1.00	1.77
10	{Assg-II-High,Assg-III-High,Assg-IV-High}	=>	{Passed}	0.10	1.00	1.77
11	{Assg-IV-High}	=>	{Passed}	0.26	0.95	1.68
12	{Assg-III-High}	=>	{Passed}	0.25	0.94	1.67
13	{Assg-II-High}	=>	{Passed}	0.25	0.94	1.67
14	{Assg-I-High}	=>	{Passed}	0.19	0.93	1.64
15	{Assg-I-High,Assg-IV-High}	=>	{Passed}	0.14	0.91	1.61
16	{Assg-I-Medium}	=>	{Passed}	0.22	0.83	1.47
17	{Assg-II-Medium}	=>	{Passed}	0.13	0.64	1.14
18	{Assg-III-Medium}	=>	{Passed}	0.16	0.58	1.02
19	{Assg-I-Low}	=>	{Passed}	0.13	0.56	1.00
20	{Assg-III-Low}	=>	{Passed}	0.10	0.54	0.95
21	{Assg-II-None,Assg-IV-None}	=>	{Failed}	0.13	1.00	2.30
22	{Assg-I-None,Assg-III-None,Assg-IV-None}	=>	{Failed}	0.13	1.00	2.30
23	{Assg-I-None,Assg-II-None}	=>	{Failed}	0.10	1.00	2.30
24	{Assg-II-Low,Assg-IV-None}	=>	{Failed}	0.10	1.00	2.30
25	{Assg-I-None,Assg-IV-None}	=>	{Failed}	0.20	0.93	2.15
26	{Assg-I-None,Assg-III-None}	=>	{Failed}	0.17	0.92	2.12
27	{Assg-II-None}	=>	{Failed}	0.16	0.92	2.11
28	{Assg-III-None,Assg-IV-None}	=>	{Failed}	0.16	0.92	2.11
29	{Assg-IV-None}	=>	{Failed}	0.28	0.90	2.08
30	{Assg-I-None}	=>	{Failed}	0.28	0.90	2.08
31	{Assg-I-None,Assg-II-Low}	=>	{Failed}	0.12	0.89	2.04
32	{Assg-III-None}	=>	{Failed}	0.22	0.79	1.82
33	{Assg-IV-Low}	=>	{Failed}	0.14	0.56	1.28
34	{Assg-II-Low}	=>	{Failed}	0.19	0.52	1.20

Rule 1 can be interpreted as- students belonging to the Medium cluster who submitted Assignment 4 on time will pass at the end of the semester. The confidence of this rule is found to be high (Confidence = 1.0, Support = 0.16, Lift = 1.77). Rule 2 also has a high confidence rate as equal (Confidence = 1.0, Support = 0.16, Lift = 1.77) as Rule 1, where it is established that- students in the High cluster who submitted both Assignment 2 and Assignment 4 are likely to pass at the end of the semester. Rules 3 to 10 generated by the association rule mining analysis have the same confidence (Confidence = 1.0) and lift (Lift = 1.77); however, the support values vary.

Rules 11 to 20 that represent the rules for the students who passed at the term-end differ much concerning each rule's confidence, support, and lift.

Rule 21 to 34 are for those students who are likely to fail at the end of the semester. For instance, Rule 21 suggests that- the students in the None cluster who had not submitted Assignment 2 and Assignment 4 are likely to fail in this course. Here, the confidence of our analysis is high (Confidence = 1.0) which means all the students who are following this pattern failed the course. The rule 22, 23, and 24 for the failed students have equal confidence (Confidence = 1.0) as Rule 21 revealed that- those students had not submitted Assignment 1-2 & 4, Assignment 1 & 2, and Assignment 2 & 4, respectively.

5. Discussion, conclusion, and limitations

Precision education aims to use artificial intelligence, LA, data analytics, text analytics, image analytics, and machine learning methods to solve complex educational problems that are yet to uncover in higher education. Along with LA, precision education also improves teaching quality and learning performance by identifying inattentive students in the classroom, at-risk students, potential drop-outs, and predicting final scores. By doing this, precision education aims to assist teachers in re-designing pedagogy, provide special care to those students in need, and provide timely feedback. A student's assignment submission is a complex aspect that has always been crucial for teachers to understand in order to provide timely feedback. In recent days, students are asked to submit their assignments using online platforms such as Moodle, Blackboard, and Google classroom. Teachers often find it difficult to understand how well a given assignment is prepared and submitted while using an online platform. In addition, it is difficult for the teachers to understand a student's learning process and assess the learning outcome just by looking at the logs. Therefore, we need to analyze these logs using precision education guidelines to reveal more insightful learning patterns such as how a student's online assignment submission behavior changes as the semester progress or find the association between students' online assignment submission behaviors and their final score. Finding these insightful learning patterns are important for teachers to provide quality education. To date, studies in precision education primarily emphasized online learning behaviors such as quiz-taking, content navigation, e-book reading, and video viewing. Therefore, most of the predictive models in the precision education literature are about identifying at-risk and drop-out using online interaction data such as reading behavior, content viewing behavior, slide navigation behavior, and related. However, a few studies have been found that analyzed online assignment submission behavior. In addition, to analyze the online submission behavior, most of the studies have overlooked the temporality (that is, the temporal analysis of learning interaction data). As mentioned earlier, temporal LA in precision education can bring new insightful information from online assignment submission behavioral patterns.

This study is conducted to tackle the abovementioned aspects of precision education. In this study, at first, we employed cluster analysis to profile students based on their online assignment submission behaviors; after that, we performed the Markov Chains analysis to investigate whether their patterns of online assignment submission behaviors change over time; and lastly, we applied the association rule mining method to examine the relationship between students' online submission behaviors and their course success. Although numerous studies use educational data mining methods such as clustering, regression, and classification to diagnose students' online assignment submission behaviors (Yang et al., 2020), temporal analysis has been rarely employed in educational research (Olsen et al., 2020). Thus, the study combined exploratory methods and temporal LA to extract actionable knowledge for learning designers and instructors. Our predictive models contribute to the precision education literature in terms of a deeper understanding of students' online assignment submission behavior's temporal patterns and establish the relation of these temporal patterns with their learning performances.

The first research question concerns profiling the students based on their online assignment submission behaviors. It was revealed that the students were clustered into three groups according to similar assignment submission behaviors. This result is consistent with Akçapınar and Kokoç (2020) findings, where it was found that the students' assignment submission data yielded three different clusters. Our results indicate that most of the students in cluster low and medium started their assignment submissions just before the due date. This result is likely to be related to academic procrastination behaviors. Procrastination involves delaying an assignment submission and learning task as long as possible (Yang et al., 2020). It is implied that most of the students had high procrastination tendencies based on their assignment submission behaviors. Our results are supported by previous studies indicating that time-related indicators reflected students' procrastination behaviors in online learning (Cerezo et al., 2017; Paule-Ruiz et al., 2015; You, 2016). The clusters based on the students' behaviors can be used as input to online learning environments to prevent procrastination behaviors. This predictive model can be applied to detect students' procrastination behavior from their online assignment submission behavioral data and inform the course instructor about the group of students using procrastination. Hence, our predictive model would help the instructor in planning an early intervention for those who are using procrastination regularly in an assigned learning task or a given assignment.

The second research question showed us whether the student followed the same pattern while submitting their assignments throughout the term. As a result, we found that the probability of shifting between the High and Low groups was less than 10%. We yield the conclusion that students in the High group have a low probability of going to the Low or None group. Likewise, students in the Low or None group during the beginning of the semester have a relatively low probability of going to the high group as the semester progresses. As a result, students in the Low and None group are at-risk of failing the course at the end of the semester. Nonetheless, these results support the idea that using temporal analytics provides exciting possibilities to move towards a new paradigm of assessment that replaces current point-in-time evaluations of learning states (Molenaar & Wise, 2016).

The third research question examined the relationship between students' assignment submission behavior and academic performance. The relationship was modeled using association rule mining. A total of 34 rules were generated which are related to academic performance. In practice, these rules can be used by the instructors or system designers to understand students' assignment submission patterns while the semester is in progress and to plan necessary interventions to prevent possible academic failures. Regarding the early prediction of students' end of year academic performance following rules can be used. For example, based on Rule 14 it can be speculated that if a student belongs to the High interaction group in the first assignment s/he will pass the course with a probability of 0.93. However, if s/he is in the Low group the probability of passing the course decreases to 0.56 (Rule 19). On the other hand, if s/he does not submit the first assignment (Assg-I-None) s/he will fail the course with a probability of 0.90 (Rule 30). Predictive models can be developed by using these rules to provide teachers with actionable insights to support their decision-making processes (Romero & Ventura, 2020). Thus, these rules can also be used to develop a rule-based intervention engine to prevent at-risk students, which is a core focus of precision education. The rules found in the study could be used as an input for student models in LA dashboards and intervention engines. Furthermore, researchers can use the rules to design automatic early interventions for increasing students' performance for precision education. Similarly, Tsai et al. (2020) concluded that the dropout prediction model in their study could provide early warnings and interventions to at-risk students for achieving precision education. It can be mentioned that identifying at-risk students is a key concern of precision education. Therefore, a LA intervention is required to help them to change their behaviors. By using these association rules that we generated to address RQ3, the instructor can spot the at-risk students by using the data from the first assignment (around the 4^{th} week).

In conclusion, the main contribution of the study is unfolding students' online assignment submission behavior using temporal LA. Leveraging online assignment submission behavioral data, this study aims to contribute to precision education literature in various ways, namely by early detection of procrastination behavior, detection of at-risk students (students in Low and in None group in Figure 2), and generation of association rules for building a rule-based intervention for the course teacher. Obtained rules can be used to predict students' end-of-term academic performance from their assignment submission behaviors at the beginning of the semester. These predictive models are primarily for instructors, but students can also get benefited from them. By using the simple visualizations that have been generated by our predictive models, students can take control of their assignment submissions. For instance, if a student finds him/herself in a Low or None group in the first few weeks of the semester, s/he can step-up and quickly submit the assignment. Also, a student can control his/her procrastination behavior. Students can also find their peers who have similar behavior. The study opens up the space for future studies as well as the design and development of intervention tools based on temporal features of online assignment submission behaviors. Moreover, the study asks whether clustering analysis, temporal analysis, and association rule mining analysis could be used to explore specific patterns of assignment submission behavior. Our results indicate that temporal analysis can be used to detect the students' online assignment submission behavior patterns and transitions between related actions. This study also proves that combining various analytic methods including clustering, Markov Chains, and association rule mining is useful for modeling temporal patterns of online assignment submission behaviors. This is a methodological contribution of the study for further studies in precision education, which provides us with deeper insights into students' behavior.

This study has some limitations that need to be discussed. First, the small sample size has decreased the generalizability of our results. To overcome this, a large-scale study in the future may be conducted in the context of open online courses. Second, this study used a data-driven approach for temporal analysis of students' behaviors. Apart from online assignment submission behavioral features, other features such as the quality of assignments, learning achievement, gender, and device students used to complete learning tasks are not analyzed. In developing predictive models, it is important to analyze LMS data combining with multimodal data to understand the learning process and predictive studies (Olsen et al., 2020). Thus, in future behavior modeling studies, researchers may collect different data types from different time periods. Third, the present study did not compare procrastination tendencies, self-regulation skills, and cognitive differences of the students who have the same sequential behavioral patterns in the learning process. Therefore, future studies regarding the student modeling of online assignment submission behaviors in precision education would consider these variables.

References

Akçapınar, G., & Kokoç, M. (2020). Analyzing the relationship between student's assignment submission behaviors and course achievement through process mining analysis. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 11*(2), 386-401. doi:10.16949/turkbilmat.711683

Azcona, D., Hsiao, I., & Smeaton, A. F. (2019). Detecting students-at-risk in computer programming classes with learning analytics from students' digital footprints. *User Modeling and User-Adapted Interaction, 29*, 759–788. doi:10.1007/s11257-019-09234-7

Boroujeni, M. S., & Dillenbourg, P. (2019). Discovery and temporal analysis of MOOC study patterns. *Journal of Learning Analytics, 6*(1), 16-33. doi:10.18608/jla.2019.61.2

Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., & Núñez, J. C. (2016). Students' LMS interaction patterns and their relationship with achievement: A case study in higher education. *Computers & Education, 96*, 42-54. doi:10.1016/j.compedu.2016.02.006

Cerezo, R., Esteban, M., Sánchez-Santillán, M., & Núñez, J.C. (2017). Procrastinating behavior in computer-based learning environments to predict performance: A case study in Moodle. *Frontiers in Psychology, 8*, 1403. doi:10.3389/fpsyg.2017.01403

Chen, B., Knight, S., & Wise, A. (2018). Critical issues in designing and implementing temporal analytics. *Journal of Learning Analytics, 5*(1), 1-9. doi:10.18608/jla.2018.53.1

Cheng, H. N., Liu, Z., Sun, J., Liu, S., & Yang, Z. (2017). Unfolding online learning behavioral patterns and their temporal changes of college students in SPOCs. *Interactive Learning Environments, 25*(2), 176-188. doi:10.1080/10494820.2016.1276082

Cormack, S. H., Eagle, L. A., & Davies, M. S. (2020) A Large-scale test of the relationship between procrastination and performance using learning analytics. *Assessment & Evaluation in Higher Education, 45*(7), 1046-1059. doi:10.1080/02602938.2019.1705244

Juhaňák, L., Zounek, J., & Rohlíková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Computers in Human Behavior, 92*, 496-506. doi:10.1016/j.chb.2017.12.015

Knight, S., Wise, A. F., & Chen, B. (2017). Time for change: Why learning analytics needs temporal analysis. *Journal of Learning Analytics, 4*(3), 7-17. doi:10.18608/jla.2017.43.2

Kokoç, M., & Altun, A. (2019). Building a learning experience: What do learners' online interaction data imply? In D. G. Sampson, D. Ifenthaler, J. M. Spector, P. Isaias, & S. Sergis (Eds.), *Learning technologies for transforming teaching, learning and assessment at large scale* (pp. 55-70). New York, NY: Springer. doi:10.1007/978-3-030-15130-0_4

Lu, O., Huang, A., Huang, J., Lin, A., Ogata, H., & Yang, S. J. H. (2018). Applying learning analytics for the early prediction of students' academic performance in blended learning. *Educational Technology & Society, 21*(2), 220-232.

Mahzoon, M. J., Maher, M. L., Eltayeby, O., Dou, W., & Grace, K. (2018). A sequence data model for analyzing temporal patterns of student data. *Journal of Learning Analytics, 5*(1), 55-74. doi:10.18608/jla.2018.51.5

Matcha, W., Gašević, D., Uzir, N. A. A., Jovanović, J., & Pardo, A. (2019, March). Analytics of learning strategies: Associations with academic performance and feedback. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 461-470). New York: ACM. doi:10.1145/3303772.3303787

Molenaar, I., & Wise, A. (2016). Grand challenge problem 12: Assessing student learning through continuous collection and interpretation of temporal performance data. In J. Ederle, K. Lund, P. Tchounikine, & F. Fischer (Eds.), *Grand challenge problems in technology-enhanced learning II: MOOCs and beyond* (pp. 59–61). New York, NY: Springer. doi:10.1007/978-3-319-12562-6_13

Nguyen, Q., Huptych, M., & Rienties, B. (2018). Using temporal analytics to detect inconsistencies between learning design and student behaviours. *Journal of Learning Analytics, 5*(3), 120-135. doi:10.18608/jla.2018.53.8

Olsen, J. K., Sharma, K., Rummel, N., & Aleven, V. (2020). Temporal analysis of multimodal data to predict collaborative learning outcomes. *British Journal of Educational Technology, 51*(5), 1527-1547. doi:10.1111/bjet.12982

Papamitsiou, Z., & Economides, A. A. (2014). Temporal learning analytics for adaptive assessment. *Journal of Learning Analytics, 1*(3), 165-168. doi:10.18608/jla.2014.13.13

Pardo, A., & Dawson, S. (2016). Learning analytics: How can data be used to improve learning practice? In P. Reimann, S. Bull, M. Kickmeier-Rust, R. Vatrapu, & B. Wasson (Eds.), *Measuring and Visualizing Learning in the Information-Rich Classroom* (pp. 41-55). New York, NY: Routledge.

Paule-Ruiz, M. P., Riestra-González, M., Sánchez-Santillán, M., & Pérez-Pérez, J. R. (2015). The procrastination related indicators in e-learning platforms. *Journal of Universal Computer Science, 21*(1), 7-22. doi:10.3217/jucs-021-01-0007

R Core Team (2017). *R: A Language and environment for statistical computing* [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Available from https://cran.r-project.org/

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*(3), e1355. doi:10.1002/widm.1355

Stiller, K., & Bachmaier, R. (2019). Using study times for identifying types of learners in a distance training for trainee teachers. *Turkish Online Journal of Distance Education, 20*(2), 21-45. doi:10.17718/tojde.557728

Tila, D., & Levy, D. (2020). Revising online assignments and the impact on student performance at a community college. *Community College Journal of Research and Practice, 44*(3), 163-180. doi:10.1080/10668926.2018.1564089

Tsai, S.-C., Chen, C.-H., Shiao, Y.-T., Ciou, J.-S., & Wu, T.-N. (2020). Precision education with statistical learning and deep learning: A Case study in Taiwan. *International Journal of Educational Technology in Higher Education, 17*(1), 12. doi:10.1186/s41239-020-00186-2

Wilson, M. S., & Ismaili, P. B. (2019). Toward maximizing the student experience and value proposition through precision education. *Business Education Innovation Journal, 11*(2), 119-124.

Yang, S. J. H. (2019). Precision education: New challenges for AI in education [conference keynote]. In *Proceedings of the 27th International Conference on Computers in Education (ICCE)* (pp. XXVII-XXVIII). Kenting, Taiwan: Asia-Pacific Society for Computers in Education (APSCE).

Yang, Y., Hooshyar, D., Pedaste, M., Wang, M., Huang, Y.-M., & Lim, H. (2020). Prediction of students' procrastination behaviour through their submission behavioural pattern in online learning. *Journal of Ambient Intelligence and Humanized Computing*. doi:10.1007/s12652-020-02041-8

You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *Internet and Higher Education, 29*, 23-30. doi:10.1016/j.iheduc.2015.11.003

Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education, 27*, 44-53. doi:10.1016/j.iheduc.2015.05.002