Modeling Consistency Using Engagement Patterns in Online Courses

JIANING ZHOU, Department of Computer Science, University of Illinois Urbana-Champaign, USA
SUMA BHAT, Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign, USA

Consistency of learning behaviors is known to play an important role in learners' engagement in a course and impact their learning outcomes. Despite significant advances in the area of learning analytics (LA) in measuring various self-regulated learning behaviors, using LA to measure consistency of online course engagement patterns remains largely unexplored. This study focuses on modeling consistency of learners in online courses to address this research gap. Toward this, we propose a novel unsupervised algorithm that combines sequence pattern mining and ideas from information retrieval with a clustering algorithm to first extract engagement patterns of learners, represent learners in a vector space of these patterns and finally group them into groups with similar consistency levels. Using clickstream data recorded in a popular learning management system over two offerings of a STEM course, we validate our proposed approach to detect learners that are inconsistent in their behaviors. We find that our method not only groups learners by consistency levels, but also provides reliable instructor support at an early stage in a course.

CCS Concepts: • **Applied computing → Education**; • **Information systems → Data mining**; • **Computing methodologies → Machine learning**.

Additional Key Words and Phrases: cluster, behavior modeling, consistency analysis

1 INTRODUCTION

With continued access to online courses and a corresponding increase in enrollments, there is a need to understand how students apply self-regulated learning strategies (SRL) [42] so as to support their engagement with the course and their eventual academic success within the online environment [13]. SRL behaviors, such as goal setting, self-monitoring and effort regulation are closely linked with the *consistency* of study habits that control these behaviors. This is because study habits (e.g., doing a set of activities at specified times in a week in a regular manner) can be interpreted to also control regulation of effort. This study concerns itself with modeling study habits that are associated with being consistent in online courses.

Prior studies have shown that low SRL skills and poor time management are leading factors of unsuccessful learning, impacting both retention and learning outcomes in online course environments (e.g., [3, 13, 50, 55]). Accordingly, literature on modeling regularity in learner engagement has focused on measuring time-related behaviors by looking into the "blackbox" of learning behaviors from clickstream data [7]. For example, Borougeni et al. [11] and Park et al. [41] study ways of measuring temporal aspects of regularity (e.g., studying on similar weekdays over the course duration) and relating these measures of regularity to learning outcomes. Other studies have used temporal behaviors to identify learners who may be disengaged [25], relate activity levels with grades [37], or infer studying habits (cramming or regular) [24].

Time-related behaviors constitute only one facet of consistency related to time management. Another important facet of consistency relates to *doing* a set of course-related activities in a regular manner. Knowing what course-related activities are associated with consistency can be used as support mechanisms to encourage persistence in disengaged learners, by suggesting them what to do next. This aspect of consistency remains largely under-explored in recent literature that uses learning analytics (LA) for modeling learner behaviors. Recent studies, including one by Sher and colleagues [48], are beginning explorations in this direction, studying engagement with one specific course component (e.g., discussion forum participation). The current study centrally addresses this research gap and aims to study patterns of engagement with multiple course components associated with consistency and relates them to learning outcomes (grades).

A broad array of empirical research on LA has demonstrated how several SRL behaviors can be measured from the clickstream captured by a learning management system (LMS) hosting online courses [52]. The ability to detect such SRLs, including consistency in learning habits, is useful for gaining deeper insights about the nature of learner participation in online courses and other learning environment, thereby informing instructional design. Beyond informing instructional design, it has the potential to serve as an early-warning signal to instructional staff about learners struggling in a course [25]. Scoping out the state-of-LA for measuring SRL behaviors, Viberg et al. [52] have underscored an urgent need to apply LA to not only measure SRL but also to find ways in which LA can facilitate teaching/learning support. This is the second research gap this study aims to fill—to inform instructional design and help instructors by using measurements derived from learners' course related activities to detect students who may be inconsistent early on in the course.

Activity sequences corresponding to the order in which learners carry out course-related activities can be automatically constructed from the clickstream using the timing and eventing information. We define engagement patterns to mean subsequences (e.g., visit to the course page followed by posting to the discussion forum) that are automatically extracted from these activity sequences. We propose an algorithm that first extracts engagement patterns of learners that frequently co-occur through the duration of a course, represents the learners' activities in terms of these engagement patterns by embedding learners in a vector space whose dimensions correspond to these engagement patterns and finally groups learners into consistency groups based on the regularity of their engagement behaviors. Our proposed method of measuring consistency makes no LMS-specific assumption, but we validate the utility of our proposed algorithm on a course hosted on Moodle, a popular LMS. We demonstrate how the proposed method can serve as a lens using which the engagement patterns associated with consistency of learners can be inferred after a course has ended. Additionally, we show how our approach can be used as a tool to provide the instructor an warning signal early on in the course and flag learners that are being inconsistent in their learning habits. To the best of our knowledge this is the first study to tackle this joint problem of measurement and support.

To summarize, available measures of SRL are limited because they may not provide a holistic view of the consistency of engagement and have not been evaluated for their ability to serve as early-warning signals. In this work, we make the following contributions:

(1) We propose an unsupervised algorithm that extracts engagement patterns from learners' activity sequences from LMS logs and uses them to group learners into consistency-based groups;
(2) Using an online STEM course offered in a popular LMS, we validate our algorithm and empirically demonstrate its reliability to jointly group learners by their consistent behaviors after a course has ended, and provide early-warning support to the instructor to detect learners that are inconsistent in their engagement.

2 RELATED WORK

2.1 Modeling Learner Behaviors

Modeling learner behaviors focuses on measuring, quantifying and modeling learners' engagement in courses and also explores the relationship between the engagement and learning outcomes relying on clickstream data as the source for empirical studies. Prior studies have sought to understand the relation between learner engagement in online courses and learning outcomes [30, 45], including SRL behaviors in online course settings [29, 49, 51, 52]. Other studies have focused on predicting dropout [18, 26], persistence [19], learning outcomes [31] and even broader demographic affiliation, such as underrepresented groups in STEM [12], using learning behaviors. Detecting changes in engagement patterns and relating them to course performance have been the focus of many studies (e.g., [40, 48]), while engagement with specific course components, such as the number of videos watched [16], the number of forum posts [9] and the number of forum views [8] and relating them to learning outcomes has also been studied.

A primary paradigm of studying engagement behaviors of students has been representing them as time series of the number of clicks over specified periods [15, 30, 44], with the primary goal of grouping learners by similar behaviors [37, 43], or associating them with successful or at-risk behaviors [25].

However, representing engagement aggregated by the total number of clicks ignores the multivariate and sequential nature of learning behaviors and provides a limited view of the complex nature of engagement and learning [21]. Because of the rich sequential features underlying the behavior sequences, patterns of engagement are better analysed as sequences of events taking into account not just the frequency of occurrence but also the relative order in which they occur. This has been the goal of [20, 22, 27] that harness the sequential nature of engagement primarily relying on hidden Markov models to understand, summarize and visualize them, or others that clustered learners based on their activity sequences [10, 29, 48]. All these methods have analyzed learners' behaviors based on all the activities, which may be redundant and noisy.

Focusing on specific events and leveraging their sequential structure many studies have sought to model learners' behaviors using sequential pattern mining methods [2, 28, 35, 38]. In line with these prior studies, we seek to study learners' engagement patterns by leveraging the sequential structure derived from the clickstream and focus on latent engagement patterns. In this sense our study is similar in spirit to [10], [29], and [48]. However, we go beyond identifying engagement patterns to show how our approach can be used to model consistency in learning habits, both after course completion (inference) as well as early on in the course (prediction) using the case study of two offerings of a course in a popular LMS.

2.2 Consistency Modeling of Learning Behaviors

Consistency modeling of learning behaviors aims to analyze how stable students are in terms of learning behaviors and the underlying patterns, which is also similar to change detection in students' behaviors. Consistency modeling and change-point detection techniques have been studied for time-series behavioral data [37, 40, 48, 49], and some have even grouped learners by their behaviors and consistency [10, 48]. This work aims to extend this latter line of prior work on consistency modeling.

Again, we deviate from these prior studies by accounting for the richness of the features available in the activity sequences, while also utilizing the underlying frequency of these different activities.

More importantly, most studies have focused on an offline analysis (inferring behaviors after a course is complete) and LA methods that support early stage detection of lapse in learning behaviors are severely lacking [52]. Because measuring learning behaviors and supporting them are both desired objectives, this study jointly addresses them both by proposing a method to model consistency and detect inconsistencies at an early stage in a course.

2.3 Early Warning Prediction

Early warning prediction aims to predict learning outcomes accurately at an early stage and provide appropriate support. Several studies have used static variables, including student demographics, self-report data and prior educational reports to predict learning outcomes at an early stage [5, 6, 47]. Despite the popularity of relying on static data for early warning

Table 1. Moodle Events Collected

Activity	Meaning	Code
Quiz Attempt	Quiz attempt started	QA
Quiz Attempt Review	Quiz attempt reviewed	QAR
Quiz Attempt Summary View	Quiz attempt summary viewed	QAS
Quiz Submission	Quiz attempt submitted	QS
Forum Post	Replying to or initiating a post	FP
Forum View	Discussion Viewed	FV
Grade View	Grade user report viewed	GV
Course View	Course Viewed	CV
User View	User profile or list viewed	UV

prediction, relying on static data ignores students' learning dynamics during the course and thus may be inadequate for early warning support.

As a solution, count-based methods using frequencies of different online learning activities for early warning prediction have been studied [4, 14, 17, 53]. However, they suffer from the same shortcomings of other count-based methods mentioned before and using the sequential structure has the potential to provide a more holistic support.

This prompts us to consider the richer sequential features available in the activities sequences. In addition, we argue that the approach of predicting learning outcomes for the purpose of intervening may be ineffective while relying on the activity sequences early in the course. This is due to the sparsity of the data that may result in unreliable predictions of learning outcomes. As a more reliable alternative, our study aims to provide early warnings by detecting whether learners are inconsistent in their learning behaviors because it may be the case that the limited data are adequate for detecting inconsistency.

3 DATA

The data used for this study were collected from the clickstream of students from a fully online for-credit undergraduate STEM course hosted on the Moodle learning management system (LMS) at the University of Illinois, Urbana-Champaign, USA. All data came from offerings of the same course over two semesters (termed C1 and C2 henceforth), and were available for our analyses only after the course had ended and students' grades had been posted. The eight-week course was composed of eight assigned readings (lectures)[1], eight assignments (homework), weekly participation in the discussion forum on a weekly basis (viewing and posting, except Week 8), working on a video as a group on a pre-selected topic (project), and a final exam.

The course materials were released on a weekly basis. Each assignment could be attempted up to 3 times and was accepted before the (hard) deadline. The final score was computed as the weighted average of the scores on the weekly assignments, required forum activities of viewing or replying to a post, a pre-class activity (offered before week 1), group project activities, and a final exam.

The dataset includes five categories of events, describing learners' interactions with (1) the assignments (attempts and submits), (2) the discussion forum topics (view or post), (3) the grade page (to check course grades as they became available), (4) the course landing page (to check the week's tasks and announcements), and (5) the profiles of other learners (used while forming a group as part of the course project requirement). The events we used are listed in Table 1.

[1]The LMS did not log events related to lectures for C1 and hence we exclude those events from our study.

Algorithm 1: Consistency Model

Input: **B** = $\{[S_{1,1}, S_{1,2}, ..., S_{1,J}], ..., [S_{I,1}, S_{I,2}, ..., S_{I,J}]\}$
$S_{i,j}$: Learner i's behavior sequence in week j.
Output: **L** = $\{[l_{1,1}, l_{1,2}, ..., l_{1,J}], ..., [l_{I,1}, l_{I,2}, ..., l_{I,J}]\}$,
O = $[O_1, O_2, ..., O_I]$
$l_{i,j}$: Learner i's behavior pattern in week j,
O_i: Number of times learner i changed behavior pattern.

Patents = **Apriori**(**B**)
for ***all*** $S_i, j \in \boldsymbol{B}$ **do**
  $V_{i,j}$ = [TF-IDF(pattern 1), ... ,TF-IDF(pattern K)]
end
L = **Cluster**(**V**)
for ***all*** $L_i \in \boldsymbol{L}$ **do**
  $L_i = [l_{i,1}, l_{i,2}, ..., l_{i,J}]$
  for ***all*** $l_{i,j} \in L_i$ **do**
    if $l_{i,j} \neq l_{i,j+1}$ **then**
      O_i += 1
    end
  end
end

Data preprocessing: In order to analyze learners' study patterns, we consider the 'window of observation' to overlap with weekly assessment periods, starting from just past midnight of a Sunday (when the previous assessment was due) to the midnight of the following Sunday. This constitutes our unit of time. For each student we obtain their weekly activity sequence by concatenating the native LMS events (representing interactions with course components over that week, of which, grade view is an example, as shown in Table 1) in the temporal order of occurrence over each assessment period. Because we wanted to study engagement patterns over the duration of the course and associate them with consistency and learning outcomes, we excluded students who dropped the course to avoid wrongly biasing the consistency detector. Following these data preprocessing steps, the final dataset used in this study contains interaction sequences of 202 learners in C1 and 279 learners in C2, both over eight assessment periods (weeks). This yielded a total of 82416 activities for C1 with an average of 51 activities per student per week; for C2, the total number of activities was 104,904 with an average of 47 activities per student per week. In addition to the clickstream data from both the courses, we had access to students' weekly assessment scores, their final exam scores as well as their final grades only for C1[2].

4 METHOD

We propose an unsupervised algorithm to group learners into different consistency groups by using students' engagement patterns derived from their activity sequences. Despite users' overall differences in learning habits (i.e., different activity sequences), our method of extracting engagement patterns for the purpose of representing learners permits a comparison of the learners. We use such a learner representation to group students based on their consistent behaviors across different weeks.

Our approach relies on the key assumption that if a learner is consistent, their engagement patterns do not change drastically from one observation period to another. This suggests that if we were to group learners by their engagement

[2]Because of data curation slippages, the data dump for C2 did not include information on learners' scores

Table 2. An example of a daily activity sequence with the engagement pattern CV FV corresponding to visiting a course page followed by a discussion forum view (color added to aid identifying the pattern)

User ID	Week	Date	Activity Sequence
2b2dadf247	1	Monday	CV FV UV FP FV CV FV
2b2dadf247	1	Wednesday	CV FP UV CV FV CV FV CV FV
2b2dadf247	1	Friday	CV CV

patterns over a reasonable observation period of the course (e.g., an assessment period of approximately a week), consistent learners would remain in the same group for each of the observation periods of the course. This also suggests that for inconsistent learners, owing to their inconsistent engagement patterns, their group membership would change through the duration of the course. An important component of this study that sets it apart from prior related studies is that it not only permits analyzing the behavior of students over the duration of a course, but also demonstrates how to use the proposed approach to detect students with learning patterns unfavorable to learning early on in the course. In this sense, we study how our approach serves as an *offline* behavior model for the instructor to infer *after* the course has ended, as well as an *online* behavior model for the instructor to use while the course is underway.

As is shown in Algorithm 1, our algorithm takes as input the weekly time-stamped activity sequences of learners, transforms them into a bag of engagement patterns, where each pattern serves as a dimension of a vector space to embed a learner into. The result is a set of vectors, one for each student for each week. The algorithm proceeds by using an appropriate clustering algorithm to group the vectors into an optimal number of clusters yielding groups of learners with similar engagement behaviors in each week. By taking as input nothing other than the learners' activity sequences, our approach permits tracking learners' patterns over different units of time.

Our unsupervised algorithm consists of two major steps: (1) Representation of learners using their individual activity sequences, and (2) Clustering learners into consistency groups. We next describe these steps in detail.

4.1 Learner Representation via Bag-of-Patterns

For each time unit (one week in our study), we split the activity sequence over the week into subsequences corresponding to the days of the week. From these subsequences, we use sequence mining methods to extract smaller subsequences of varying lengths that correspond to consecutive actions with the requirement that they frequently co-occur on a daily basis, which we believe to be more meaningful than a considering a single activity. This step yields a set of 'action phrases' or engagement patterns.

Next, each learner's weekly activity sequence is represented as a bag-of-patterns using the engagement patterns identified, where we borrow the idea of a bag-of-words model that is popular in information retrieval (IR) [34]. Here, instead of using words (i.e., single events) we use engagement patterns. A compromise is that, even though a bag-of-words model disregards word order and only considers words' occurrence while representing a document, our bag-of-patterns model takes local sequential information into consideration by accepting a phrase of two or more consecutive events that frequently co-occur.

We denote by $V_{i,j} = [p_1, p_2, ..., p_K]$, the vector representing learner i's behaviors in week j, where each dimension p_k captures the occurrence of an engagement pattern k in the week and K is the number of engagement patterns. We design p_k to be the number of occurrences of the engagement pattern k in the learner's weekly sequence weighted by the pattern's term frequency-inverse document frequency (tf-idf) value [34]. Considering that some of the short patterns may be sub-sequences of some longer patterns and to avoid counting them twice, we count the patterns by first listing them in descending order of their lengths (i.e.,the number of activities in the engagement patterns). The tf-idf value captures the importance of pattern k in a learner's weekly sequence and normalizes its count relative to another pattern that may not be unique to a given learner's activity sequence. p_k in $V_{i,j}$ is calculated as follows:

$$p_k = \frac{f_{k,i,j}}{\sum_{k'} f_{k',i,j}} \cdot log(\frac{N}{|\{b \in B : phrase_k \in b\}|}),$$

where $f_{k,i,j}$ is the number of occurrences of pattern k in learner i's behaviors in week j. N is the number of all the data entries (here, number of weeks x number of learners) in B, where B is the set of activity sequences of all the students over all the weeks. Here, we make a simplifying assumption that given a learner, their weekly activities are independent, rendering the vectors $V_{i,j}$ independent. The result of this step is a set of vectors $V_{i,j}$ collectively called $\mathbf{V}$.

4.2 Clustering

After representing learners as vectors, we cluster the resulting vectors $\mathbf{V}$ using a clustering algorithm. As a result of clustering, each data point $V_{i,j} \in \mathbf{V}$ is assigned a label $l_{i,j}$ to mean that based on learner i's behaviors in week j the learner was assigned to cluster $l_{i,j}$, yielding one label per week for every learner.

$$l_i = [l_{i,1}, l_{i,2}, ..., l_{i,J}],$$

where J is the total number of time units. We note that our approach permits considering any granularity of events and time units for representation, potentially yielding events at a finer granularity than what we had, or coarser (by grouping events into suitable categories). We leave experiments using these variations for future work.

The whole process of our algorithm is also provided in Algorithm 1.

5 CONSISTENCY MODELING

Our algorithm can be used to group learners by consistency of behaviors both in a post-course setting (offline) as well as during the course (online). We describe each of these settings below.

5.1 Offline Consistency Modeling

Understanding learner behaviors through the duration of a course after it has ended (offline) offers important highlights to instructional designers and course staff. More specifically, it may be of interest to know the extent to which learners were consistent and what their engagement patterns were.

Based on our assumptions, the course-related activities of a consistent learner should result in similar learning-related activities, through the weeks. In terms of the activity sequence, this means that a consistent learner's weekly activity sequence should include the same set of engagement patterns for all the course weeks. Therefore, a consistent learner's vectors should be placed in the same cluster for all the weeks. In contrast, those of irregular learners will appear in different clusters over the weeks depending on the degree of irregularity.

This suggests that we can quantify the extent to which a learner is consistent by detecting how often that learner changed their behavior patterns, which is revealed by how many times the cluster label of the learner changed over all the weeks. For instance, if $l_{i,j-1} \neq l_{i,j}$, then learner i's behavior patterns in week $j-1$ are different from those in week j. This provides a method to explicitly quantify each learner's trajectory of consistency through the duration of a course as the number of changes of the cluster label. We then assign students into consistency groups based on the number of times the cluster label of the learner changed over the duration of the course, with 0 denoting the most consistent learner group and the increasing numbers corresponding to decreasing consistency.

5.2 Online Consistency Detection

Similar to offline consistency detection, our goal here is to group learners into consistency groups while the course is in session, for example in first few weeks. This will serve as an early warning signal for intervention. Toward this end, we first use the bag-of-patterns model to represent the learners, then cluster the learners based on these features to finally determine learners' consistency level after the results of clustering.

6 EMPIRICAL EVALUATION

In this section, we present our research questions that guide the validation of our proposed algorithm. We validate our algorithm's ability to carry out offline consistency modeling in the context of our first and second research questions: **RQ1:** Can we quantitatively evaluate different sequence mining methods and clustering algorithm combinations for the purpose of grouping learners by consistency levels? **RQ2:** How do we use the bag-of-patterns model in combination with a clustering algorithm to group students based on their consistency levels?

Unlike offline detection, where the activity sequences of learners over the entire duration were available, for online detection the activity sequences that are available are only from the first few weeks. Under this constraint, generating the list of engagement patterns based on limited data may be unreliable. This leads us to explore two conditions in which we generate the list of engagement patterns. In the first, we utilize the list generated from a previous course offering using the offline detection method. We hypothesize here that students exhibit similar engagement patterns across course offerings when the course structure remains unchanged. In the second, we extract engagement patterns using only the first two weeks, where we make the simplifying assumption that the total number of activities will be proportionately reduced compared to the total number of activities over the entire duration. Online consistency modeling then uses this list of engagement patterns to group learners as in the offline setting. This leads us to our third research question: **RQ3:** To what extent are the engagement patterns (of the bag-of-pattern model) similar across course offerings?

7 EXPERIMENTAL SETUP

In this section we provide the details used in our experiments for answering the research questions.
Analysis window: So as to better analyze the consistency of learners' behaviors, we drop the data from the first two weeks and start with the third week. This is done so as to let the learners settle into a habit and reduce the noise in the data coming from the first two weeks of learners' activities.
Sequential pattern mining algorithms: For a broader view of the effect of the pattern extraction method we explored the following popular sequential pattern mining algorithms: Apriori [1], PrefixSpan [23] and N-gram [36]. For the Apriori algorithm, each transaction was the original *daily* activity sequence. To extract the most frequently co-occurring and representative engagement patterns, we set the thresholds of support, confidence and lift (parameters of the pattern mining algorithms) to 0.45, 0.4, and 1 respectively. Just like the Apriori algorithm, the PrefixSpan algorithm takes as input the original daily activity sequence and we select heuristically select the top 150 patterns from the output. For the N-gram algorithm too we heuristically select the top 150 patterns with 2-, 3-, 4- and 5-grams.
Clustering algorithms: We also explored the use of three clustering algorithms: Spectral Clustering [39], Agglomerative Clustering [54] and K-means [32]. Because our learner representations are unit vectors the three clustering

algorithms used the Euclidean distance to calculate the similarity between two vectors. The other parameters were set to their default values. The optimal number of clusters was obtained using the "Silhouette statistic" [46]. The Silhouette statistic measures the degree of confidence in a particular clustering result and lies in the interval [-1,1], with values close to 1 for well-clustered observations and values close to -1 for poorly clustered observations.

Evaluation of Consistency Grouping: In the absence of ground truth information on the level of consistency of the students, we use reports of the positive association between consistent learning behaviors and course grade (learning outcomes) [48] to select the best (sequence mining,clustering algorithm) pair to provide the optimal consistency grouping. Toward this, we use the mean of total weekly scores calculated as $\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{J} Score_{i,j}$, mean grade variance $\frac{1}{N}\sum_{i=1}^{N} Var([Score_{i,1}, ..., Score_{i,J}])$ and mean final exam score $\frac{1}{N}\sum_{i=1}^{N} FScore_i$, where $Score_{i,j}$ represents student i's weekly assessment score in week j and $FScore_i$ represents student i's final exam score. We expect the optimal consistency grouping to yield a set of clusters where the more consistent groups are associated with higher learning outcomes. Thus, the larger the number of cluster changes exhibited by the learners the lesser consistent they are. Additionally, we expect that learners with an increase in the number of cluster changes are associated with lower mean of total weekly scores and lower mean final exam score, but higher mean grade variance.

8 EXPERIMENTAL RESULTS AND ANALYSES

We now answer each of the research questions in the context of our experimental results.

RQ1: Can we quantitatively evaluate the different combinations of sequence mining methods (for phrase extraction) and clustering algorithms for the purpose of grouping learners by consistency levels using the proposed approach?

For the data from C1, we perform offline consistency modeling using different clustering algorithms (Spectral, agglomerative and k-means) combined with the different pattern mining algorithms (Apriori, PrefixSpan and N-gram). The results are summarized in Figure 1, which compare the different clustering algorithms (Spectral, agglomerative and k-means going from the first row to the third row of the plots respectively) when used with the different pattern mining algorithms (Apriori, PrefixSpan and N-gram). Each plot shows the number of cluster changes after the clustering process along the x-axis and the learning outcome related aspects (mean of total weekly scores, mean final exam score and mean grade variance) along the y-axis (going from the first column to the final column of plots).

From figure 1, we notice that for mean of total weekly scores and mean final exam score, the results of agglomerative clustering (Figures 2-a and 2-b) and K-means (Figure 3-a and 3-b) do not show the correspondence between the number of cluster changes and the mean of total weekly scores that is expected. For instance, we see that the associations barely exist for agglomerative clustering results regardless of the sequence pattern mining approach, whereas those of spectral clustering with the Apriori algorithm more closely approximate our intuition. Likewise, for mean grade variance, only the results of Spectral clustering show to the desired association. Based on these visualizations, we choose the Apriori pattern mining algorithm with Spectral clustering as the best performing combination for offline consistency modeling.

RQ2: How do we use the proposed approach to group students based on their consistency level in an offline setting?

For C1, the 202 students' activity sequences yielded 147 engagement patterns, which were then used for the bag-of-patterns representation of the students. Using the best silhouette statistic yielded the solution with six clusters. These resulted in five consistency groups of learners after grouping learners by the number of cluster changes.

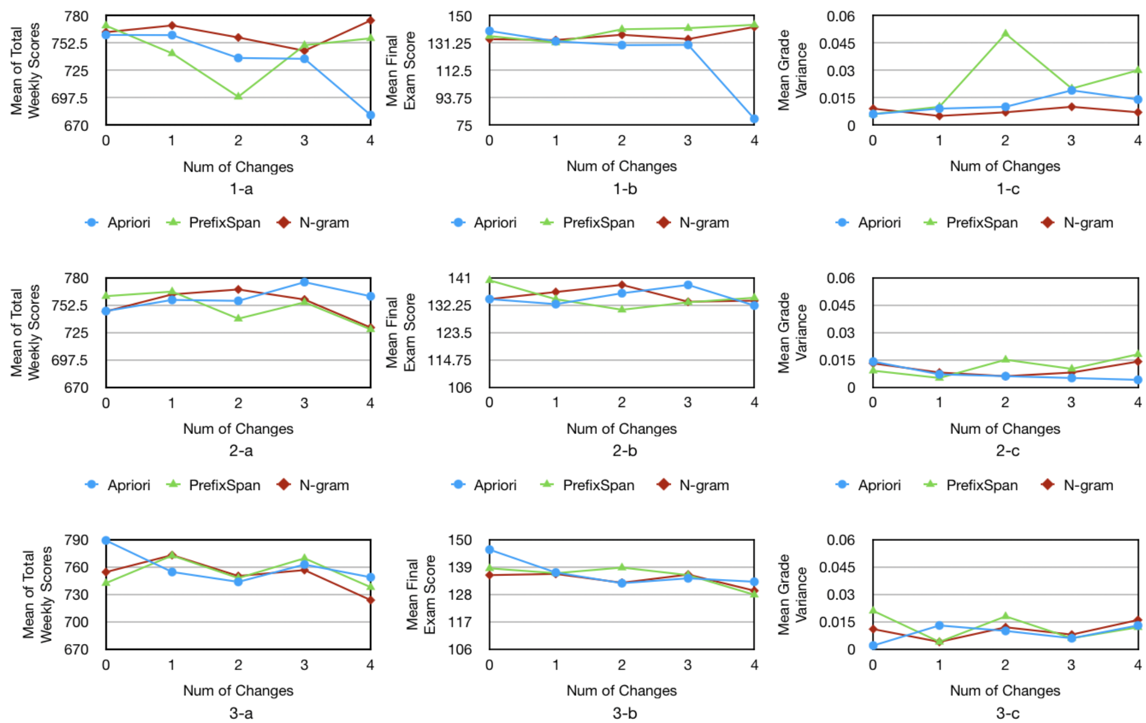

Fig. 1. Comparison between different clustering algorithms with different pattern mining algorithms. 1: Spectral Clustering. 2: Agglomerative Clustering. 3: K-means. a. Relationship between number of changes and mean of total weekly scores. b. Relationship between number of changes and mean final exam score. c. Relationship between number of changes and mean grade variance.

Table 3. Grade-related statistics for the different consistency-related groups obtained by our model for learners in C1.

Group	Num of Changes	Mean of Total Weekly Scores	Mean Grade Var	Mean Final Exam Score	Learners
1	0	760.49	0.006	139.42	98
2	1	760.3	0.009	132.35	77
3	2	737.35	0.010	129.72	15
4	3	736.52	0.019	139.93	10
5	4	680.04	0.014	79.295	2

For each group of learners, we calculated their mean of total weekly scores, mean grade variance and mean final exam score as defined in Section 7. The results are shown in Table 3. We notice that the first group of learners never change their behavior patterns making them the most consistent learners. From the table we also see that this group of most consistent learners has the highest mean of total weekly scores, the highest mean final exam score and lowest mean grade variance. From group 2 to group 5, the number of changes gradually increases and we see a corresponding decrease in the mean of total weekly scores and the mean final exam score, whereas the mean grade variance increases.
Analysis of the association between group assignment and final exam scores. In order to further evaluate our model and better analyze the difference between groups, we performed a pairwise significance test. We observed

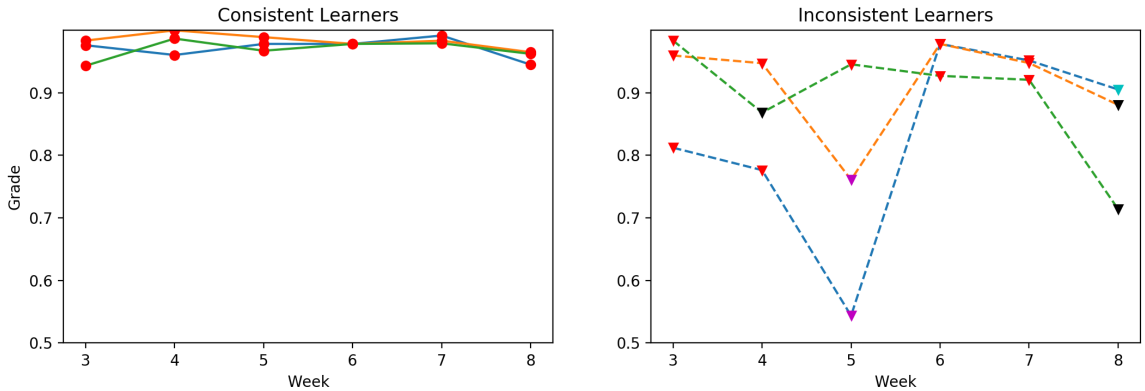

Fig. 2. Relationship between weekly scores and weekly behavior patterns for consistent learners and inconsistent learners. For each week, points with different colors represent different clusters and different behavior patterns. Different lines denote different learners.

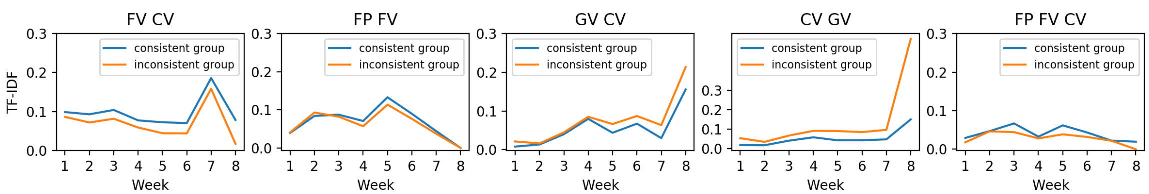

Fig. 3. Average TF-IDF values in each week for students in C1

that the final exam scores were not normally distributed and so used the Wilcoxon rank sum test with continuity correction to compare the groups in a pairwise manner. The pairwise comparison of groups with respect to the final grade revealed that the differences between groups (2,3) and (3,4) were not significant (all p greater than 0.05). Moreover, group 5 with only two students was merged with the group 4. Only groups 1 and 2 had differences in mean scores that were statistically significant. This is also in line with our expectation—students in Group 1 who never change their behavior patterns should be consistent and students in the other groups who changed their behavior patterns should be inconsistent. We use these results to merge the groups into classes as follows.

- Class I - **Consistent** (N = 98, 48.5%): This class constitutes the consistent group of learners in Group 1 in Table 3, who do not change their behavior patterns over the duration of the course. We also notice that this group of learners has the highest mean of total weakly scores (760.49) and mean final exam grade (139.42) and lowest mean grade variance (0.006) compared to the other groups.
- Class II - **Not Consistent** (N = 104, 51.5%): This class consists of Groups 2, 3, 4 and 5 in Table 3. They change their behavior pattern at least once. With the decrease of consistency, we can also see the decrease of mean of total weekly scores and mean final exam score and the increasing of mean grade variance. Associated with their inconsistency, are their lower mean of total weekly scores (751.35) and mean final exam grade (131.68) and higher mean grade variance (0.01) compared to Class I.

Analysis of the engagement patterns. In Figure 3 we show the TF-IDF values of some engagement patterns in each week averaged over the learners in each class (Class I is the consistent group and Class II is the inconsistent group). For each pattern, we observe the difference between the consistent group and inconsistent group. We can see

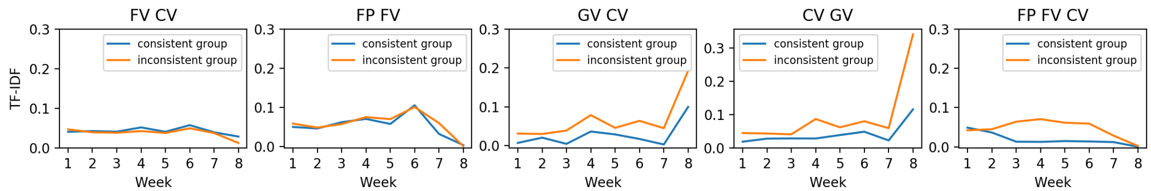

Fig. 4. Average TF-IDF values in each week for students in C2

Table 4. The different consistency-based groups for learners in C2

Group	1	2	3	4	5	6
Num of Changes	0	1	2	3	4	5
Number of Learners	29	65	57	85	30	13

that for pattern “FV CV”, “FP FV” and “FP FV CV”, the TF-IDF values of the consistent group are always higher than those of the inconsistent group, which indicates that the learners in the consistent group exhibit these patterns more frequently. And for pattern “GV CV” and “CV GV”, the TF-IDF values of the inconsistent group are higher than those of the consistent group. Going by the activity captured by these patterns, it appears that the learners in the inconsistent group are more concerned about checking their grades (GV= grade view) and pay less attention to viewing their course material or participating in the discussion forum compared to the consistent learners.

Analysis of the relationship between behavior patterns and grades. We illustrate the relationship between the change of behavior patterns and the change of weekly scores using a random sample of learners from each consistency group. In Figure 2, lines with different colors represent different learners and points with different colors represent different clusters. For three consistent learners (plot on the left), we can see that they are assigned to the same cluster through the course duration. We also notice that their scores do not vary much. However, for three inconsistent learners (plot on the right), with the change of cluster there is always a significant change of their weekly scores.

RQ3: To what extent are the engagement patterns of the bag-of-patterns model similar across course offerings?

We answer this question in two analysis settings: 1) By performing an offline consistency modeling on C2; and 2) by performing online consistency detection based on fewer number of weeks in C2. For the first setting, we compare the results of offline consistency modeling on C2 with that of C1. For the second setting, we check whether using patterns extracted from C1 also leads to a good performance of online consistency detection in C2, and whether the results are comparable to the performance of the detector using patterns extracted from C2 alone.

For offline consistency modeling on C2, 142 engagement patterns were extracted for 279 students. Of these, 130 engagement patterns (91.5%) overlapped with the patterns extracted from C1. Moreover, the clustering solution with six clusters was found to be optimal. The six groups of learners were grouped by their consistency levels as shown in Table 4, where we see that 29 of the learners were consistent and 250 were inconsistent (albeit to different degrees).

We also show the TF-IDF values of some engagement patterns in each week averaged by learners each group in Figure 4. For each engagement pattern, we observe that the differences between the consistent group and the inconsistent group engagement are similar to those in C1 (Figure 3). For instance, we can see that the trends in participation for the

Table 5. Summary of results for online detection for learners in C1

Group	Num of Changes	Mean Total Weekly Scores	Mean Grade Var	Mean Final Exam Score	Learners
Consistent	0	762.63	0.003	134.86	107
Inconsistent	1	738.58	0.015	134.29	95

patterns “FV CV” and “FP FV” are comparable between C1 and C2. Likewise, as in C1, for the pattern “GV CV” and “CV GV”, the TF-IDF values of inconsistent group are also higher than the TF-IDF values of consistent group.

For online consistency detection, the results from the offline setting were used as ground truth owing to the absence of actual ground truth information on consistency. We used two methods to get the engagement patterns. In the first method, we use the patterns extracted from a previous offering (C1) for detection in a subsequent offering (C2). Here we refer to the results above (offline for C2), where we saw that there were 29 consistent learners and 250 inconsistent learners. With this as a reference, using the first two weeks' data (weeks 3 and 4), we found that 166 learners were correctly detected as being inconsistent and 18 learners were correctly detected as being consistent.

The second method was to use the patterns extracted from the first two weeks of a given semester. For learners in C1, we found that there were 98 consistent learners and 104 inconsistent learners. Using the first two weeks' data, we were able to correctly detect 88 of these 98 learners as consistent and 93 of these 104 learners as inconsistent. Table 5 summarizes these results for the consistent and the inconsistent groups. For learners detected as consistent, we can see that their mean of total weekly scores and mean final exam score are both higher than learners detected as inconsistent. The mean grade variance of consistent learners is much lower than that of inconsistent learners, which is also expected. For learners in C2, there were 191 learners correctly detected as inconsistent and 18 learners correctly detected as consistent, which is slightly better than the first method.

To quantify the online detection performance, we evaluate how our method performs when detecting inconsistent student groups because our focus is to provide early intervention. Toward this, we consider detecting some consistent students as inconsistent to be acceptable as long as most of the inconsistent students can be correctly detected as inconsistent. Therefore, for a reliable detector, we are interested in the number of inconsistent students (positive class) that are correctly detected as being inconsistent (true positives, TP) and the number of inconsistent students that are wrongly detected as being consistent (false negative, FN). We use the Recall score, defined as $R = \frac{TP}{TP+FN}$. The recall of inconsistent student detection for C1 is 0.83. The recall of inconsistent student detection for C2 using the first method is 0.80 and the recall of inconsistent student detection for C2 using the second method is 0.92, indicating that using the second method has a better performance. The high recall scores for both course offerings, suggests the reliability of our method for detecting inconsistent students at an early stage.

Overall, although both methods show the feasibility of reliably detecting inconsistent learners early on in a course, we find that using the patterns extracted from the first two weeks of the course under consideration has a slightly more reliable performance.

9 DISCUSSION

Our experimental results showed that learners can be grouped into groups based on the consistency of their behavior patterns. This suggests that the proposed approach to group learners can serve as a way of measuring consistency of behaviors after a course has ended. In addition, the experimental results on online consistency modeling also suggest the feasibility and reliability of using our proposed approach to detect students showing inconsistent behavior patterns and potentially in need of instructional support to persist in the course and learn in an effective manner. Our results also showed how learning analytic measures of consistency are strongly associated with learners' academic performance adding further support to the consistency detection algorithm.

Interpretation of the engagement patterns: The engagement patterns we discover from our sequence mining step do not have explicit interpretations, and in this sense are unlike those proposed in previous studies (e.g., Maldonado-Mahauad and colleagues sought to ascribe theory-driven SRL behaviors to such patterns [33]). Nevertheless, we have shown that the engagement patterns play an important role both in terms of offline and online consistency detection. We leave it to future work to study consistency detection using interpretable engagement patterns.

Generalizability of the engagement patterns: We also notice that the manner in which the engagement patterns are expressed may not necessarily be generalizable from one course offering to the other. For example, we can see in Figure 3 and Figure 4 that there are differences in how the same pattern, for example "FP FV CV", are differently exhibited by learners in C1 and C2. Despite the difference in the way individual patterns are expressed from one course offering to another, their collective use in detecting consistency is what we highlight and demonstrate by showing comparable results for online consistency detection. Additionally, it is also important to note that the engagement patterns themselves are course/LMS specific and may not generalize from a course in one domain to another (e.g., physics to mathematics). Although we make no LMS specific assumptions in the way we extract the patterns, the list of engagement patterns naturally depends on the inherent differences in LMSs and the native events gathered by the LMS. Finally, course-specific activities and grading schemes custom-created by instructional designers or instructors for each course (e.g., watching lecture videos may be part of one course, and forum participation may not be a gradable activity in a course) will affect the engagement patterns that are extracted.

10 CONCLUSION AND FUTURE WORK

This paper studied a novel unsupervised approach combining sequence mining methods with ideas from information retrieval and clustering, to detect consistency of learners' behaviors using engagement patterns extracted from their activity sequences. Using an online STEM course offered in a popular LMS, we empirically demonstrated the algorithm's reliable performance in detecting inconsistent learners at an early stage in a course.

A primary limitation of this study was the use of only one course to validate our method. We leave it to future work to explore the generalizability of our approach to a wider set of courses and LMSs. A second direction for future exploration could be verifying online inconsistency detection performance with ground truth information based on instructor and learner input in an explicit experimental setting.

ACKNOWLEDGMENTS

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education through Grant number R305A180211 to the Board of Trustees of the University of Illinois. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

REFERENCES

[1] Rakesh Agarwal, Ramakrishnan Srikant, et al. 1994. Fast algorithms for mining association rules. In *Proc. of the 20th VLDB Conference*. 487–499.

[2] Rakesh Agrawal and Ramakrishnan Srikant. 1995. Mining sequential patterns. In *Proceedings of the eleventh international conference on data engineering*. IEEE, 3–14.

[3] Nora'ayu Ahmad Uzir, Dragan Gašević, Wannisa Matcha, Jelena Jovanović, and Abelardo Pardo. 2020. Analytics of time management strategies in a flipped classroom. *Journal of Computer Assisted Learning* 36, 1 (2020), 70–88.

[4] Abdullah Alsheddy and Mohamed Habib. 2017. On the application of data mining algorithms for predicting student performance: A case study. *Int. J. Comput. Sci. Netw. Secur* 17, 10 (2017), 189–197.

[5] Raheela Asif, Saman Hina, and Saba Izhar Haque. 2017. Predicting student academic performance using data mining methods. *International Journal of computer science and network security* 17, 5 (2017), 187–191.

[6] Azwa Abdul Aziz, Nur Hafieza Ismail, Fadhilah Ahmad, and Hasni Hassan. 2015. A FRAMEWORK FOR STUDENTSâ€™ ACADEMIC PERFORMANCE ANALYSIS USING NAÃ VE BAYES CLASSIFIER. *Jurnal Teknologi* 75, 3 (2015).

[7] Rachel Baker, Di Xu, Jihyun Park, Renzhe Yu, Qiujie Li, Bianca Cung, Christian Fischer, Fernando Rodriguez, Mark Warschauer, and Padhraic Smyth. 2020. The benefits and caveats of using clickstream data to understand student self-regulatory behaviors: opening the black box of learning processes. *International Journal of Educational Technology in Higher Education* 17 (2020), 1–24.

[8] Yoav Bergner, Deirdre Kerr, and David E Pritchard. 2015. Methodological Challenges in the Analysis of MOOC Data for Exploring the Relationship between Discussion Forum Views and Learning Outcomes. *International Educational Data Mining Society* (2015).

[9] Fernanda Bonafini, Chungil Chae, Eunsung Park, and Kathryn Jablokow. 2017. How much does student engagement with videos and forums in a MOOC affect their achievement? *Online Learning Journal* 21, 4 (2017).

[10] Mina Shirvani Boroujeni and Pierre Dillenbourg. 2018. Discovery and temporal analysis of latent study patterns in MOOC interaction sequences. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. 206–215.

[11] Mina Shirvani Boroujeni, Kshitij Sharma, Łukasz Kidziński, Lorenzo Lucignano, and Pierre Dillenbourg. 2016. How to quantify student's regularity?. In *European conference on technology enhanced learning*. Springer, 277–291.

[12] Nigel Bosch, R Wes Crues, Genevieve M Henricks, Michelle Perry, Lawrence Angrave, Najmuddin Shaik, Suma Bhat, and Carolyn J Anderson. 2018. Modeling key differences in underrepresented students' interactions with an online STEM course. In *Proceedings of the Technology, Mind, and Society*. 1–6.

[13] Jim Broadbent and Walter 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.

[14] Michael Geoffrey Brown, R Matthew DeMonbrun, Steven Lonn, Stephen J Aguilar, and Stephanie D Teasley. 2016. What and when: the role of course type and timing in students' academic performance. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. 459–468.

[15] Sahan Bulathwela, María Pérez-Ortiz, Aldo Lipani, Emine Yilmaz, and John Shawe-Taylor. 2020. Predicting Engagement in Video Lectures. *arXiv preprint arXiv:2006.00592* (2020).

[16] Yunfan Chen and Ming Zhang. 2017. Mooc student dropout: Pattern and prevention. In *Proceedings of the ACM Turing 50th Celebration Conference-China*. 1–6.

[17] Evandro B Costa, Baldoino Fonseca, Marcelo Almeida Santana, Fabrísia Ferreira de Araújo, and Joilson Rego. 2017. Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior* 73 (2017), 247–256.

[18] Scott Crossley, Luc Paquette, Mihai Dascalu, Danielle S McNamara, and Ryan S Baker. 2016. Combining click-stream data with NLP tools to better understand MOOC completion. In *Proceedings of the sixth international conference on learning analytics & knowledge*. 6–14.

[19] R Wes Crues, Nigel Bosch, Michelle Perry, Lawrence Angrave, Najmuddin Shaik, and Suma Bhat. 2018. Refocusing the lens on engagement in MOOCs. In *Proceedings of the fifth annual ACM conference on learning at scale*. 1–10.

[20] Louis Faucon, Lukasz Kidzinski, and Pierre Dillenbourg. 2016. Semi-Markov Model for Simulating MOOC Students. *International Educational Data Mining Society* (2016).

[21] Ed Fincham, Alexander Whitelock-Wainwright, Vitomir Kovanović, Srećko Joksimović, Jan-Paul van Staalduinen, and Dragan Gašević. 2019. Counting clicks is not enough: Validating a theorized model of engagement in learning analytics. In *Proceedings of the 9th international conference on learning analytics & knowledge*. 501–510.

[22] Chase Geigle and ChengXiang Zhai. 2017. Modeling MOOC student behavior with two-layer hidden Markov models. In *Proceedings of the fourth (2017) ACM conference on learning@ scale*. 205–208.

[23] Jiawei Han, Jian Pei, Behzad Mortazavi-Asl, Helen Pinto, Qiming Chen, Umeshwar Dayal, and Meichun Hsu. 2001. Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth. In *proceedings of the 17th international conference on data engineering*. IEEE Washington, DC, USA, 215–224.

[24] Lauren C Hensley, Christopher A Wolters, Sungjun Won, and Anna C Brady. 2018. Academic probation, time management, and time use in a college success course. *Journal of College Reading and Learning* 48, 2 (2018), 105–123.

[25] Jui-Long Hung, Morgan C Wang, Shuyan Wang, Maha Abdelrasoul, Yaohang Li, and Wu He. 2015. Identifying at-risk students for early interventions—A time-series clustering approach. *IEEE Transactions on Emerging Topics in Computing* 5, 1 (2015), 45–55.

[26] Ali Shariq Imran, Fisnik Dalipi, and Zenun Kastrati. 2019. Predicting Student Dropout in a MOOC: An Evaluation of a Deep Neural Network Model. In *Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence*. 190–195.

[27] Hogyeong Jeong and Gautam Biswas. 2008. Mining student behavior models in learning-by-teaching environments. In *Educational data mining 2008*.

[28] John S Kinnebrew, Kirk M Loretz, and Gautam Biswas. 2013. A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining* 5, 1 (2013), 190–219.

[29] René F Kizilcec, Chris Piech, and Emily Schneider. 2013. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the third international conference on learning analytics and knowledge*. 170–179.

[30] Andrew S Lan, Christopher G Brinton, Tsung-Yen Yang, and Mung Chiang. 2017. Behavior-Based Latent Variable Model for Learner Engagement. *International Educational Data Mining Society* (2017).

[31] Xiao Li, Ting Wang, and Huaimin Wang. 2017. Exploring n-gram features in clickstream data for MOOC learning achievement prediction. In *International Conference on Database Systems for Advanced Applications*. Springer, 328–339.

[32] Stuart Lloyd. 1982. Least squares quantization in PCM. *IEEE transactions on information theory* 28, 2 (1982), 129–137.

[33] Jorge Maldonado-Mahauad, Mar Pérez-Sanagustín, René F Kizilcec, Nicolás Morales, and Jorge Munoz-Gama. 2018. Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior* 80 (2018), 179–196.

[34] Christopher D Manning, Hinrich Schütze, and Prabhakar Raghavan. 2008. *Introduction to information retrieval*. Cambridge university press.

[35] R Martinez, Kalina Yacef, Judy Kay, A Al-Qaraghuli, and Ahmed Kharrufa. 2011. Analysing frequent sequential patterns of collaborative learning activity around an interactive tabletop. In *4th International Conference on Educational Data Mining, EDM 2011*. CEUR-WS, 111–120.

[36] Hirokazu Masataki and Yoshinori Sgisaka. 1996. Variable-order N-gram generation by word-class splitting and consecutive word grouping. In *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, Vol. 1. IEEE, 188–191.

[37] Ewa Młynarska, Derek Greene, and Pádraig Cunningham. 2016. Time series clustering of Moodle activity data. In *24th Irish Conference on Artificial Intelligence and Cognitive Science (AICS'16), University College Dublin, Dublin, Ireland, 20-21 September 2016*.

[38] John C Nesbit, Mingming Zhou, Yabo Xu, and P Winne. [n.d.]. Advancing log analysis of student interactions with cognitive tools.

[39] Andrew Y Ng, Michael I Jordan, and Yair Weiss. 2002. On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*. 849–856.

[40] Jihyun Park, Kameryn Denaro, Fernando Rodriguez, Padhraic Smyth, and Mark Warschauer. 2017. Detecting changes in student behavior from clickstream data. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. 21–30.

[41] Jihyun Park, Renzhe Yu, Fernando Rodriguez, Rachel Baker, Padhraic Smyth, and Mark Warschauer. 2018. Understanding Student Procrastination via Mixture Models. *International Educational Data Mining Society* (2018).

[42] Paul R Pintrich. 2000. The role of goal orientation in self-regulated learning. In *Handbook of self-regulation*. Elsevier, 451–502.

[43] C Pradana, SS Kusumawardani, and AE Permanasari. 2020. Comparison Clustering Performance Based on Moodle Log Mining. In *IOP Conference Series: Materials Science and Engineering*, Vol. 722. IOP Publishing, 012012.

[44] Arti Ramesh, Dan Goldwasser, Bert Huang, Hal Daume, and Lise Getoor. 2018. Interpretable Engagement Models for MOOCs using Hinge-loss Markov Random Fields. *IEEE Transactions on Learning Technologies* (2018).

[45] Arti Ramesh, Dan Goldwasser, Bert Huang, Hal Daumé III, and Lise Getoor. 2014. Learning latent engagement patterns of students in online courses. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*. 1272–1278.

[46] Peter J Rousseeuw. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics* 20 (1987), 53–65.

[47] Amjad Abu Saa et al. 2016. Educational data mining & students' performance prediction. *International Journal of Advanced Computer Science and Applications* 7, 5 (2016), 212–220.

[48] Varshita Sher, Marek Hatala, and Dragan Gašević. 2020. Analyzing the consistency in within-activity learning patterns in blended learning. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. 1–10.

[49] Atsushi Shimada, Yuta Taniguchi, Fumiya Okubo, Shin'ichi Konomi, and Hiroaki Ogata. 2018. Online change detection for monitoring individual student behavior via clickstream data on E-book system. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. 446–450.

[50] Jordan Thibodeaux, Aaron Deutsch, Anastasia Kitsantas, and Adam Winsler. 2017. First-year college students' time use: Relations with self-regulation and GPA. *Journal of Advanced Academics* 28, 1 (2017), 5–27.

[51] Nora'ayu Ahmad Uzir, Dragan Gašević, Jelena Jovanović, Wannisa Matcha, Lisa-Angelique Lim, and Anthea Fudge. 2020. Analytics of time management and learning strategies for effective online learning in blended environments. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. 392–401.

[52] Olga Viberg, Mohammad Khalil, and Martine Baars. 2020. Self-regulated learning and learning analytics in online learning environments: a review of empirical research. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. 524–533.

[53] Richard Joseph Waddington, SungJin Nam, Steven Lonn, and Stephanie D Teasley. 2016. Improving early warning systems with categorized course resource usage. *Journal of Learning Analytics* 3, 3 (2016), 263–290.

[54] Joe H Ward Jr. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association* 58, 301 (1963), 236–244.

[55] Saijing Zheng, Mary Beth Rosson, Patrick C Shih, and John M Carroll. 2015. Understanding student motivation, behaviors and perceptions in MOOCs. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 1882–1895.