MOOCad: Visual Analysis of Anomalous Learning Activities in Massive Open Online Courses

Xing Mu[1], Ke Xu[1], Qing Chen[2], Fan Du[3], Yun Wang[4], Huamin Qu[1]

[1]The Hong Kong University of Science and Technology, Hong Kong, [2]Inria, France, [3]Adobe Research, America, [4]Microsoft Research, China

Abstract

The research on Massive Open Online Course (MOOC) has mushroomed worldwide due to the technical revolution and its unprecedented enrollments. Existing work mainly focuses on performance prediction, content recommendation, and learning behavior summarization. However, finding anomalous learning activities in MOOC data has posed special challenges and requires providing a clear definition of anomalous behavior, analyzing the multifaceted learning sequence data, and interpreting anomalies at different scales. In this paper, we present a novel visual analytics system, MOOCad, for exploring anomalous learning patterns and their clustering in MOOC data. The system integrates an anomaly detection algorithm to cluster learning sequences of MOOC learners into staged-based groups. Moreover, it allows interactive anomaly detection between and within groups on the basis of semantic and interpretable group-wise data summaries. We demonstrate the effectiveness of MOOCad via an in-depth interview with a MOOC lecturer with real-world course data.

CCS Concepts

• ***Human-centered computing*** → ***Visual analytics;*** • ***Applied computing*** → ***E-learning;***

1. Introduction

Massive Open Online Courses (MOOCs) have developed rapidly worldwide in recent years. Analyzing learning activities not only facilitates summarizing the common learning patterns but also assists in detecting anomalous learners who have special learning styles or patterns. These *anomalous learners* can be clustered into relatively small, underrepresented groups who have *anomalous learning activities* (i.e.similar activities but are obscured within the larger groups). By identifying and interpreting the anomalous groups, MOOC learners can make personalized plans by referring to the correlations between learning behaviors and outcomes of similar others. More importantly, MOOC instructors can gain insights for improving their course design and teaching strategy.

However, anomaly detection on MOOC data is a challenging task. Anomaly detection techniques [CBK09] face an inherent issue that there lacks a clear boundary to distinguish normal and abnormal cases, which makes the ground truth required for training and verifying models difficult to obtain. This challenge is further aggravated by the multifaceted types of MOOC data, including learner profiles, video clickstream interactions, and forum activities [GKR14]. Such data forms the large-scale learning sequences (i.e., ordering of learning activities or processes [BD12]) with rich context information and high individual variation. To cope with the volume and variety, an intelligent and interactive tool is required to help identify and interpret normal and abnormal MOOC learning behaviors.

Data visualization has been introduced to support the exploration and understanding of learning sequences [KGS*14, GKR14] by providing intuitive representations and rich interaction techniques. However, few of existing techniques are developed to address the anomaly detection problem in learning activities. Existing work on MOOC data analytics mainly falls into three categories based on their application domains: (1) MOOC prediction such as forecasting students' dropout rate [CCZ*16] and grades [JWS*14], (2) video and content recommendation like helping students decide which materials to learn next [CZBS18], and (3) pattern mining such as analyzing the clickstream data tracking how students watch lecture videos [SFCQ15]. These techniques are useful for summarizing general patterns rather than detecting anomalous learning activities. To develop an effective visual analytics system for detecting and interpreting anomalous learning activities, we need to deal with the large-scale and multifaceted MOOC data, and to capture both the temporal patterns (e.g., the evolution of learning activity) and content patterns (e.g., the frequent learning subsequence) in the data.

To fill this gap, we present MOOCad (MOOC anomaly detection), the first visual analytics system that helps MOOC instructors interactively detect anomalous groups and interpret their atypical learning activities with multiple perspectives and rich context information. We design an interactive framework with novel visualization designs which supports exploration and interpretation of abnormal groups and learning activities in three steps: stage summarization, within-stage learning pattern clustering, and individual sequence inspection. We further conduct a case study and discuss potential improvements with domain experts.

2. Related Work

In this section, we survey and discuss researches on analyzing and visualizing learning activities, anomaly detection and visualization.

DOI: 10.2312/evs.20191176

Learning Analytics and Visualization. MOOC data consists of web logs which can be analyzed as event sequence data [GXZ*18, HLDF13]. Many research efforts have been devoted to detecting learning sequence and cluster student activities by building learner models and utilizing automated or semi-automated methods [HW10, DB12]. Some recommender systems are also designed to help individual learners obtain resources in e-learning [ACKP95,Bru03]. For example, Salehi et al. [SK13] applied a hybrid approach to recommend learning materials based on the sequential pattern of the accessed material and the learner's preference. Besides, the prediction of learning activities has also gained attention. Sinha et al. [SLJD14] built a system focusing on drop-outs prediction by discovering hidden structural configurations in the learning sequence of MOOC data. Basic visualizations are commonly used for showing general patterns, such as scatter-plots, heatmaps, and node-link diagrams [DTO*13, HPNG13, KGS*14]. Coffrin et al. [CCdBK14] provided a series of combined bar charts and line charts to help understand learning behaviors. Recent work presents comprehensive visualizations and new functionalities to explore learner behaviors. For example, VisMOOC [SFCQ15] designs an interactive system based on stacked graphs for showing temporal patterns of video clickstream data. Most previous techniques aim at analyzing learning sequence data across all records while few focus on detecting underrepresented groups in MOOC data to help uncover irregular patterns. In this paper, we extend previous research and design a novel visual exploration system, MOOCad, for summarizing user activities and identifying the anomalous learning path with multifaceted perspectives. **Anomaly Detection and Visualization.** Anomaly detection has been widely studied over the past decades [CBK09]. These methods for anomaly detection within different situations can be broadly categorized into tensor-based algorithms [CZC*15], statistics-based algorithms [RL05], classification-based algorithms [HHWB02, LTZ08], and distance-based algorithms [BKNS00]. Two major challenges for anomaly detection are the fuzzy boundary between normality and abnormality and the absence of labeled data collected for training and verifying models [XXM*18]. Thus, more efficient tools with intuitive visualizations are required to encourage experts' involvement in refining results and supporting decision making [CLZ*18, XGC*18]. Recently, plenty of visual analysis systems have been proposed to detect anomalies in sequential data. For example, Phong et al. [NTA*18] presented a visual analytics approach that helps understand unusual behavior through action sequences. Zhao et al. developed #FluxFlow [ZCW*14] to analyze the anomalous information processes in social media. Although the above-mentioned systems help detect abnormal points in sequential data, few approaches are capable of detecting inconsistent patterns in MOOC data by considering both the temporal pattern and content pattern with human knowledge. To overcome this challenge, we design a novel progression anomaly detection system which clearly represents the detailed and comprehensive anomalous learning activities and helps conduct large-scale data analysis and exploration.

3. System Requirements and Design

To inform the design of MOOCad, we conducted a series of interview with one MOOC instructor, two domain experts (in developing MOOC analytics systems), and five MOOC learners. A list of requirements was derived from the interviewees' comments and guided the development of our system.

R1. Consider the entire learning sequence. Our goal is to allow MOOC instructors and learners to detect anomalous learning activities with consideration of the entire learning sequence of learners and all types of MOOC data.

R2. Facilitate anomaly detection and reasoning. The system should detect anomalous groups and provide intuitive designs to help users discover and understand abnormal activities.

R3. Allow individual learning path analysis. The domain experts mentioned their interests in inspecting individual learner's behaviors in addition to the group behaviors.

R4. Provide interactivity and diversity. It is necessary to incorporate flexible interactions that help users quickly navigate a large-scale learning sequence data, as well as understand and compare the anomalies with multiple perspectives.

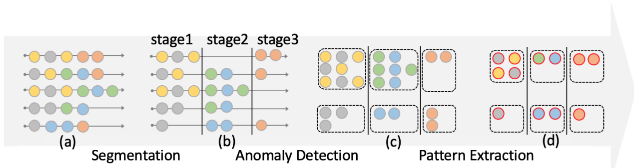

Figure 1: *The visual anomaly detection of learning sequence data.*

The MOOCad architecture consists of three modules: (1) data prepossessing module that transforms the entire learning activities of each learner into a learning sequence (**R1**), (2) anomaly detection module that detects groups with anomalous learning activities within each learning stages (**R2**), and (3) visualization module that serves as an interface for MOOC instructors to explore the learning sequences data, and understand the detected anomalies (**R3, R4**).

4. Algorithm

The anomaly detection algorithm consists of three sequential steps (Fig. 1): (1) Stage analysis of learning sequence. With the learning sequences of different learners as input data, the first step is to segment these sequences into meaningful progression stages based on a word embedding method [GJG*18]. The stages are extracted purely from the data based on their semantic meaning. The learning sequence data is often collected in large volumes and contains important progression patterns, showing the evolution of an entity over different stages (e.g., the beginning phase of a course and final exam period). Thus an anomaly analysis based on the temporal pattern of learning activities is more interpretable. (2) Identification of anomalous groups in each stage. Within each stage, the sequence segments (i.e., sub-sequences) of different learners that belongs to this stage are grouped into clusters based on their anomaly degree [LGG*18]. (3) Extraction of frequent patterns. For each of the group obtained in the last step, we extract its frequent patterns to help the understanding of this group [LW08]. In general, the output is the frequent sub-sequences of learners that are grouped within different stages. For more details, please refer to the supplemental material [Xu19].

5. Description of MOOCad

In this section, we give an overview of MOOCad's user interface and workflow (Fig. 2), and then the details of each interface component.

5.1. Tasks and User Interface

In response to the general requirements mentioned in Section 3, we discussed with the experts about the more specified visualization and interaction tasks in MOOCad to interpret the results detected by the analytic algorithm with multiple perspectives.

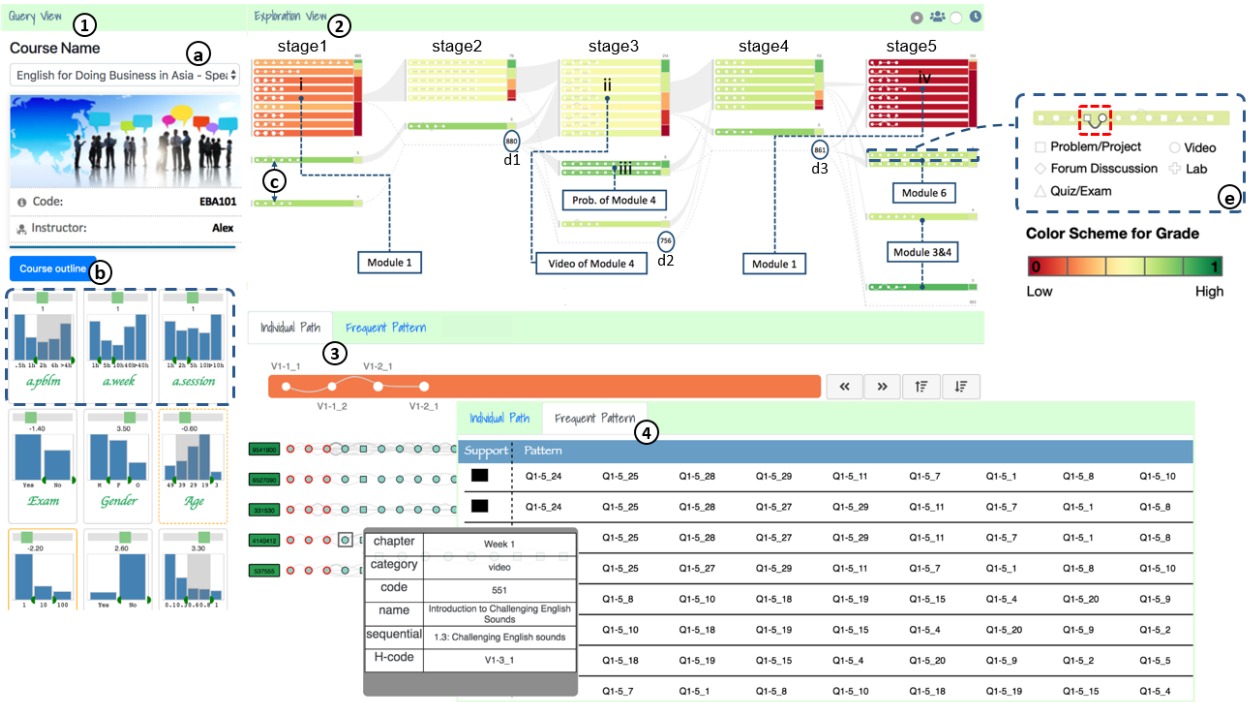

Figure 2: *The user interface of MOOCad contains four coordinated views: (1) query view, (2) exploration view, (3) individual path view,and (4) frequent pattern view. Users can select a course in (a) and retrieve the detailed course outline in (b).*

T1. Allow custom queries before visualization.
T2. Display the progression pattern of learning sequence data.
T3. Identify the anomalous groups within a given stage.
T4. Indicate the frequent subsequences which characterize an anomalous group.
T5. Provide supporting information for stage-based analysis.
T6. Flexible data exploration and comparison.

Guided by the aforementioned tasks, we design the user interface of MOOCad as shown in Fig. 2. The system consists of five major views with the primary *exploration view* surrounded by a set of coordinated views which support anomaly detection of learning activities from different perspectives. In particular, once selecting the course in Fig. 2(a), the user interface begins with the *query view* (Fig. 2(1)) which uses a set of bar charts to display different distributions of learners based on certain attributes (e.g., average learning time per week), and these charts allow users to query the learning sequences with attributes-based constraints (**T1**) by brushing. Then the selected data would be analyzed by the anomaly detection algorithm, and results would be displayed in the *exploration view* (Fig. 2(2)), which shows the groups of frequent learning subsequences within each stage (**T3**) and to present the evolution pattern of these groups across the different stages (**T2**). If a frequent pattern (thread) from a given group in the *exploration view* is selected, the *individual path view* (Fig. 2(3)) will illustrate the detailed individual learning sequence of learners having this pattern (**T5**). The *frequent pattern view* (Fig. 2(4)) provides all the frequent patterns extracted from a selected group (**T5**) in the *exploration view*. The color ranging from red to yellow, and to green is consistently used to encode the grades which are normalized from 0 (red) to 1 (green). Different learning events are represented by different symbols (Fig. 2(e)). All the views are linked together to help users identify corresponding elements like the events, and zooming is supported in most views.

5.2. Stage-based Anomaly Detection

To improve the scalability and interpretability, we design the *exploration view* by combing a flow-based approach [GS14] to display the high-level overview of the stage segmentation results, and a matrix-based approach [ZGC*16] to indicate the content pattern of each anomalous groups within the stage. As illustrated in Fig. 2 (2), each stage is represented by a list of frequent activity patterns shown as horizontal threads. The background color of the bar encodes the average grade of students with this pattern. Patterns with a red background are associated with lower grades while patterns with a green background are associated with higher grades. The numbers of white symbols inside the bar represent the lengths of the patterns. The dashed line at bottom of each stage represents students who are inactive in this stage.

Group Analysis within Stage

The patterns within a stage are clustered to the corresponding anomalous group and the groups are ordered according to the number of students related with the group. The vertical distance (Fig. 2(c)) between two group is computed based on the similarity. To assess the similarity value of two groups, we first calculate the Levenshtein distance of each pair of patterns between the two groups. The similarity between two groups equals to the average distance between all possible pattern pairs of the two groups. In this view, users can identify anomalous learning activities by analyzing how similar patterns lead to different grades within a group or by comparing different activity patterns between different groups.

Each thread (Fig. 2(e)) in a group represents a frequent activity

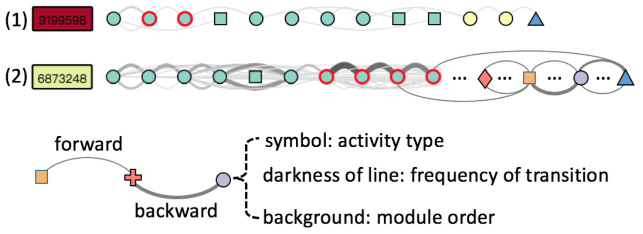

Figure 3: *The visual design of the individual view and two examples.*

pattern shown by the group's students, which is a segmented subsequence extracted from their learning sequences. The white symbols in the thread mean specific learning activities. Five types of learning activities typically exist throughout a learning sequence and are encoded using symbols of different shapes: Video (circle), Problem or Project (square), Quiz or Exam (triangle), Forum Discussion (diamond), and Lab (cross). The size of the symbol encodes either the number of learners involved in it or the average duration that learners conducting this learning activity. The sequence in a thread is connected by smooth curves in a zig-zag manner. The degree of curvature represents the average interval time that learners act between two learning activities. On the right side of each group, there is a stacked bar chart, which shows the distribution of learners in this group according to their grades. The grades are normalized and divided into five bins ($[0, 0.2, 0.4, 0.6, 1.0]$) with 1.0 as the best grade. The bars in the chart represent different grade bins in the descending order. By default, the stacked bar chart illustrates the distribution of the whole group learners. When hovering on a specific thread, the chart will update and show the distribution of the learners involved in the focused frequent pattern.

Between-Stage Anomaly Detection. Links are drawn between two adjacent stages to reveal the transition of students from the current group in the current stage to the group they belong to after progressing to the next stage. The width of each end of the link encodes the percentage of students. Thereby, users can identify anomalous learning activities from the perspective that considers the general progression pattern of the entire learning sequences. The design is mostly inspired by Egolines [ZGC*16], a matrix-based method for event sequence data. However, the two designs are different in several aspects: (1) the data type is different; (2) we have the stage transition information but Egolines extends along the full time span; (3) we group the sub-sequence within each stage.

5.3. Individual Inspection

As shown in Fig. 3, in the *individual path view*, each row shows the learning sequence of a student and each symbol inside the row is a learning activity. We use a categorical color scheme to encode the background of each activity which is ordered according to the syllabus, with each color representing a distinct lecture module in the syllabus. Lines on the upper half represents incoming transition (from the left activity to the right activity) and lines on the bottom half represent outgoing activity (from right to left). The darkness encodes the frequency. In addition, a rectangle is placed on the left side of each row, of which the text information represents the student ID and the background color encodes his/her grade.

6. Case study

We evaluated MOOCad on a language course with 6 modules/weeks' material and 1,684 learners in the same offering round. A senior lecturer who serves as the instructor of this MOOC was invited to participate in this study. We report the findings as follows.

Five stages are generated and shown in the *exploration view*(Fig. 2(2)). The first feature that attracted the instructor's attention is the bottom dashline in each stage and the number of inactive students (Fig. 2(d1-3)). The instructor said this result validates his concern about the attendance rate of the course: most learners skip certain classes and just finish the quiz and exam to get the credits. Then he took a detailed inspection into each stage to analyze abnormal learning activities (**T3&4**) and got some meaningful findings. First, the most obvious anomaly caught by him is the first group in stage-5 which has an extremely red background (Fig. 2(iv)) compared with the other groups in this stage, indicating the students in this group having very bad performance on average. Then he checked the frequent patterns in this group and found that most learning activities are related with module 1, which means learners in the first group just start watching the module 1's videos until they have to take exam (Fig. 3(1)), and then skip most the following lectures. Thereby, the average grade of this group is lower than the others. Secondly, he noticed that the learning sequences regarding module 1 (Fig. 2(i&iv)) is less consistent than other modules. For example, the curvature of lines between events in the major (1st) group from module 3 (stage-2) to module 6 (stage-4) is smooth, which means the time interval between two activities is similar. On the contrary, the line curvature between events regarding module 1 is more variant. He mentioned that this may be caused by the discontinuity of the teaching materials in the first module, and the detection result was useful to help modify the materials related to module 1, making them less separated in his following teaching. Finally, he found that the second group in stage-4 has many triangles (Fig. 2(iii)), which means students are frequently doing practice. By contrast, the first group in this stage mainly watches videos of module 4 (Fig. 2(ii)). The second group has a better grade because the practice of this module has the largest weight in the final grade.

7. Conclusion and Future Work

In this paper, we introduce the visual analysis system, MOOCad, that enables MOOC instructors to interactively detect anomalous learning activities. The system incorporates an integrated algorithm to facilitate the identification and reasoning of the anomalous groups based on their learning sequence data. We also propose multiple coordinated views and flexible interactions to support the exploration of learning progression between stages, group comparison within stages and individual path inspection. We evaluate MOOCad through a case study on real datasets with a MOOC instructor which demonstrated the capability of our system in detecting anomalous learning activities. After discussions with domain experts, we found that the semantic meaning of some stage groups was not clear to the experts at the beginning. However, some groups were interpreted later by the experts after exploration. This change shows an advantage of our unsupervised learning method as it can discover the unexpected. While the visual designs were highly praised by instructors and the interactions were considered easy to use, there are still some limitations with system designs. In the future, we plan to enhance the visualizations by designing more advanced alternatives to present information such as individual quiz scores as well as other properties of learners.

8. Acknowledgements

Ke Xu is the corresponding author. This work is supported by HK ITF Grant ITS/388/17FP.

References

[ACKP95] ANDERSON J. R., CORBETT A. T., KOEDINGER K. R., PELLETIER R.: Cognitive tutors: Lessons learned. *The journal of the learning sciences 4*, 2 (1995), 167–207. 2

[BD12] BINGHAM C. B., DAVIS J. P.: Learning sequences: Their existence, effect, and evolution. *Academy of Management Journal 55*, 3 (2012), 611–641. 1

[BKNS00] BREUNIG M. M., KRIEGEL H.-P., NG R. T., SANDER J.: Lof: Identifying density-based local outliers. In *ACM SIGMOD Record* (2000), vol. 29, ACM, pp. 93–104. 2

[Bru03] BRUSILOVSKY P.: Developing adaptive educational hypermedia systems: From design models to authoring tools. In *Authoring Tools for Advanced Technology Learning Environments*. Springer, 2003, pp. 377–409. 2

[CBK09] CHANDOLA V., BANERJEE A., KUMAR V.: Anomaly detection: A survey. *ACM Computing Surveys (CSUR) 41*, 3 (2009), 15. 1, 2

[CCdBK14] COFFRIN C., CORRIN L., DE BARBA P., KENNEDY G.: Visualizing patterns of student engagement and performance in moocs. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge* (2014), ACM, pp. 83–92. 2

[CCZ*16] CHEN Y., CHEN Q., ZHAO M., BOYER S., VEERAMACHANENI K., QU H.: Dropoutseer: Visualizing learning patterns in massive open online courses for dropout reasoning and prediction. In *IEEE Conference on Visual Analytics Science and Technology (VAST)* (2016), IEEE, pp. 111–120. 1

[CLZ*18] CAO N., LIN C., ZHU Q., LIN Y.-R., TENG X., WEN X.: Voila: Visual anomaly detection and monitoring with streaming spatiotemporal data. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 23–33. 2

[CZBS18] COOPER M., ZHAO J., BHATT C., SHAMMA D. A.: Moocex: Exploring educational video via recommendation. In *Proceedings of the ACM on International Conference on Multimedia Retrieval* (2018), ACM, pp. 521–524. 1

[CZC*15] CHEN H., ZHANG S., CHEN W., MEI H., ZHANG J., MERCER A., LIANG R., QU H.: Uncertainty-aware multidimensional ensemble data visualization and exploration. *IEEE Transactions on Visualization and Computer Graphics 21*, 9 (2015), 1072–1086. 2

[DB12] DESMARAIS M. C., BAKER R. S.: A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction 22*, 1-2 (2012), 9–38. 2

[DTO*13] DERNONCOURT F., TAYLOR C., OÂĂŹREILLY U.-M., VEERAMACHANENI K., WU S., DO C., HALAWA S.: Moocviz: A large scale, open access, collaborative, data analytics platform for moocs. In *Proceedings of NIPS Workshop on Data-Driven Education* (2013). 2

[GJG*18] GUO S., JIN Z., GOTZ D., DU F., ZHA H., CAO N.: Visual progression analysis of event sequence data. *IEEE Transactions on Visualization and Computer Graphics (Early Access)* (2018). 2

[GKR14] GUO P. J., KIM J., RUBIN R.: How video production affects student engagement: An empirical study of mooc videos. In *Proceedings of the First ACM Conference on Learning@ Scale Conference* (2014), ACM, pp. 41–50. 1

[GS14] GOTZ D., STAVROPOULOS H.: Decisionflow: Visual analytics for high-dimensional temporal event sequence data. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1783–1792. 3

[GXZ*18] GUO S., XU K., ZHAO R., GOTZ D., ZHA H., CAO N.: Eventthread: Visual summarization and stage analysis of event sequence data. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 56–65. 2

[HHWB02] HAWKINS S., HE H., WILLIAMS G., BAXTER R.: Outlier detection using replicator neural networks. In *International Conference on Data Warehousing and Knowledge Discovery* (2002), Springer, pp. 170–180. 2

[HLDF13] HUANG Z., LU X., DUAN H., FAN W.: Summarizing clinical pathways from event logs. *Journal of Biomedical Informatics 46*, 1 (2013), 111–127. 2

[HPNG13] HUANG J., PIECH C., NGUYEN A., GUIBAS L.: Syntactic and functional variability of a million code submissions in a machine learning mooc. In *AIED 2013 Workshops Proceedings Volume* (2013), p. 25. 2

[HW10] HSIEH T.-C., WANG T.-I.: A mining-based approach on discovering courses pattern for constructing suitable learning path. *Expert Systems with Applications 37*, 6 (2010), 4156–4167. 2

[JWS*14] JIANG S., WILLIAMS A., SCHENKE K., WARSCHAUER M., O'DOWD D.: Predicting mooc performance with week 1 behavior. In *Educational Data Mining* (2014). 1

[KGS*14] KIM J., GUO P. J., SEATON D. T., MITROS P., GAJOS K. Z., MILLER R. C.: Understanding in-video dropouts and interaction peaks inonline lecture videos. In *Proceedings of the First ACM Conference on Learning@ Scale Conference* (2014), ACM, pp. 31–40. 1, 2

[LGG*18] LIN H., GAO S., GOTZ D., DU F., HE J., CAO N.: Rclens: Interactive rare category exploration and identification. *IEEE Transactions on Visualization and Computer Graphics 24*, 7 (2018), 2223–2237. 2

[LTZ08] LIU F. T., TING K. M., ZHOU Z.-H.: Isolation forest. In *Eighth IEEE International Conference on Data Mining* (2008), IEEE, pp. 413–422. 2

[LW08] LI C., WANG J.: Efficiently mining closed subsequences with gap constraints. In *Proceedings of the SIAM International Conference on Data Mining* (2008), SIAM, pp. 313–322. 2

[NTA*18] NGUYEN P. H., TURKAY C., ANDRIENKO G., ANDRIENKO N., THONNARD O., ZOUAOUI J.: Understanding user behaviour through action sequences: from the usual to the unusual. *IEEE transactions on visualization and computer graphics* (2018). 2

[RL05] ROUSSEEUW P. J., LEROY A. M.: *Robust regression and outlier detection*, vol. 589. John wiley & sons, 2005. 2

[SFCQ15] SHI C., FU S., CHEN Q., QU H.: Vismooc: Visualizing video clickstream data from massive open online courses. In *IEEE Pacific Visualization Symposium (PacificVis)* (2015), IEEE, pp. 159–166. 1, 2

[SK13] SALEHI M., KAMALABADI I. N.: Hybrid recommendation approach for learning material based on sequential pattern of the accessed material and the learnerâĂŹs preference tree. *Knowledge-Based Systems 48* (2013), 57–69. 2

[SLJD14] SINHA T., LI N., JERMANN P., DILLENBOURG P.: Capturing" attrition intensifying" structural traits from didactic interaction sequences of mooc learners. *arXiv preprint arXiv:1409.5887* (2014). 2

[XGC*18] XU K., GUO S., CAO N., GOTZ D., XU A., QU H., YAO Z., CHEN Y.: Ecglens: Interactive visual exploration of large scale ecg data for arrhythmia detection. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2018), ACM, p. 663. 2

[Xu19] XU K.: Supplementary material of moocad. `lukeluker.github.io/MOOCad/suppl.pdf`, 2019. 2

[XXM*18] XU K., XIA M., MU X., WANG Y., CAO N.: Ensemblelens: Ensemble-based visual exploration of anomaly detection algorithms with multidimensional data. *IEEE Transactions on Visualization and Computer Graphics (Early Access)* (2018). 2

[ZCW*14] ZHAO J., CAO N., WEN Z., SONG Y., LIN Y.-R., COLLINS C.: # fluxflow: Visual analysis of anomalous information spreading on social media. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1773–1782. 2

[ZGC*16] ZHAO J., GLUECK M., CHEVALIER F., WU Y., KHAN A.: Egocentric analysis of dynamic networks with egolines. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2016), ACM, pp. 5003–5014. 3, 4