Informatics Bulletin, Faculty of Computers and Artificial Intelligence, Helwan University
Published Online *Vol 7 Issue 1, January 2025*
(https://fcihib.journals.ekb.eg)

A Survey on Educational Process Mining

Deena M. Helal[1], Yehia M. Helmy[1], and Doaa S. ElZanfaly[3]
[1]Lecturer, Faculty of Commerce and Business Administration, Helwan University. Egypt.
dina.m.helal@commerce.helwan.edu.eg, ymhelmy@commerce.helwan.edu.eg
[3]Professor, Faculty of Computers and Artificial Intelligence, Helwan University, Egypt.
doaa.saad@fci.helwan.edu.eg

***Abstract*—Process Mining is an emerging discipline that utilizes techniques of data science to detect, verify and enhance processes. There has been a growing attention in the research area of educational process mining field due to the availability of students' log in our modern era and the advantage of its use to both learners and academic institutions. Educational Process Mining (EPM) is positioned between discovery, analysis, and visualization of students' learning patterns from educational records. Despite the availability of students' logs, there are few literature surveys that covered different areas of research in the EPM and the various techniques and algorithms of process mining that can be used within each respective topic. In this study we present a literature survey that covers the main areas of the research, methods, and techniques in EPM that helps in discovering the students' different aspects of behavior in the learning process and in predicting students' academic performance in an educational context, with the aim of enhancing final academic achievement.**

***Keywords*— Educational Process Mining, Prediction, Student Academic Performance, Process Discovery, Student Behavior.**

I. INTRODUCTION

E-learning is the method of delivering learning material over electronic resources such as the Internet. It has several advantages over conventional techniques, including enhanced active and autonomous learning, cost and time savings, and web-based information that is accessible at any time [1]. In such systems, learners produce a huge amount of data that can be seen as a source of significant information about the learning process and its outcomes. These data are utilized by Educational Data Mining (EDM) and Learning Analytics (LA) based researchers to enhance the education [2]. EDM is concerned with the development, study, and use of algorithms to find patterns in educational datasets that would otherwise be difficult or impossible to evaluate owing to their sheer bulk [3].

It has several advantages over conventional techniques, including enhanced active and autonomous learning, cost and time savings, and web-based information that is accessible at any time [1]. In such systems, learners produce a huge amount of data that can be seen as a source of significant information about the learning process and its outcomes. These data are utilized by Educational Data Mining (EDM) and Learning Analytics (LA) based researchers to enhance the education [2]. EDM is concerned with the development, study, and use of algorithms to find patterns in educational datasets that would otherwise be difficult or impossible to evaluate owing to their sheer bulk [3].

Today's E-learning platforms using Learning Management System (LMS) for administration, tracking, reporting, and delivery of courses and training programs. The learning content management system (LCMS) is a companion tool to the learning management system (LMS) that focuses more on the generation and administration of educational material. In certain advanced scenarios, a LCMS may automatically generate learning materials based on student profiles and learning styles using an adaptive system approach [4].

The advantages of this technology are dependent on the degree to which it is used in ways that are consistent with human cognitive abilities, in accordance with the principles of educational science research. Electronic gadgets, including personal computers, digital cameras, servers, and sensors, in this new era have taken their place in the educational process, with recorded educational data exceeding 4 zettabytes [5]. Student's data automatically collected in online learning environments are often not turned into meaningful information for teaching and learning [6] may be poorly utilized across the educational domain. With the emergence of e-learning, information systems may now record a variety of student activities and interactions, ranging from low-level events like mouse clicks and gestures to higher-level events like students' learning patterns and processes [7].

The current state of art in the educational field is increasingly focusing on the use of different techniques and approaches on the educational records of students in different formats to predict the students' performance and improve their academic achievement. Since machine and deep learning techniques have tackled numerous educational datasets in different ways where each had a unique purpose for a specific educational objective. There is a scarcity in the studies that tackle the student records from a process point of view where process mining may contribute to the generation of information and its integration within a holistic process framework considering the time as a significant factor in the study.

Our research is motivated by the absence of literature reviews on the application of process mining in the field of education. Very few studies on process mining in other sectors, such as healthcare and other business areas, have been published. Also, regarding educational data and the application of data and process mining in the education sectors, research areas need to be examined. This could genuinely help in the areas of learning process such as early intervention and enhance students' performance by mitigating inefficiencies and enhancing the process,

The advantages of this technology are dependent on the degree to which it is used in ways that are consistent with human cognitive abilities, in accordance with the principles of educational science research. Electronic process optimization by streamlining the workflow, also by providing valuable data for pedagogical research and the development of evidence-based educational practices.

The rest of the paper is organized as Follows: in section 2, we describe the educational domain including the educational data used in analyzing learning process, in section 3, we discuss process mining with its types and relationship with students' performance in the educational domain. In section 4, we describe the discovered four main areas of research in the EPM domain. In section 5 we present the current challenges and the future work of the domain. In section 6 we present the conclusion and the outputs of the study.

II. Educational Domain

A. Educational Data

Educational data that is recorded in both the digital devices and offline may come in different structure that needs essentially preprocessing which is a crucial step in the research studies as they are considered one of the recent interests among educational scholars in learning analytics methodologies and Process Mining [6]. Educational data is a type of data that is extracted from any learning environment. The two most prevalent forms of educational data utilized in literature are learning log data and texts to examine learners' behavior and mostly predict their academic performance [8] Behavioral data explain how a student interacts with the learning system, including the number of times a module is visited each week, the number of posts, participating in discussion forums, the number of the seen videos, and the number of the submitted assignments [8] As researchers in [9],[10],[11] merge behavior data with demographic and academic transcripts. Text is the main form of data used in the literature, it could be extracted from video transcripts [12] ,[13] and student views [14],[15]. In addition, other research used sentiment analysis to extract student opinions from social media and LMS [16].

B. Learning Mode

There are several educational modes that serve different learning needs. The selection of an educational mode is often influenced by several variables, including personal preferences, educational objectives, availability of resources and technology, and particular criteria set out by the educational programs.

Zhang in [17] listed different modes and formats for education including Traditional Classroom-Based Education, Online Education, Blended or Hybrid Learning, Distance Learning, Self-Regulated Learning, Corporate Training, and Professional Development.

Classroom-based education is a widely used instructional approach in educational institutions, whereby students are required to be physically present in designated classrooms for their learning activities.

Online education encompasses a range of educational modalities, such as online courses, webinars, and virtual classrooms, that use internet connectivity and digital technology. Blended learning is an instructional approach that integrates conventional in-person teaching methods with online components, enabling students to engage in face-to-face classes while utilizing digital resources at other timeframes.

Distance Learning refers to a kind of education when people are geographically separated from the educational institution. This mode of learning is made possible via the utilization of many technologies such as postal mail, electronic mail, video conferencing, and web-based platforms. Self-Regulated Learning enables people to autonomously choose their educational endeavors, supported by various resources such as printed books, online courses, and instructional aids. Corporate training and professional development programs are designed to meet the specific requirements of employees and professionals by providing focused training and skill upgrading that is customized to their respective positions or industries.

C. Learning Analytics

In recent years, there has been a growing emphasis within the educational sector on using student records and data analytics to predict student performance and enhance their educational results. This has led to the arise of a domain that examines the learners' data and their performance metrics called Learning Analytics (LA). According to [18] Learning Analytics is defined as the collection and analysis of data pertaining to students' engagement with educational technology, coursework, and other educational resources, with the aim of acquiring valuable insights into their learning patterns and behaviors[19][32]. According to [20] there are three fundamental components of LA:

a) Data: the entity upon which analytical insights are based.
b) Analysis: the logical deployment of methodologies and tools.
c) Action: the insights in which the use of data analytics techniques and technologies to education-related topics [19].

The use of learning analytics data facilitates decision-making based on empirical evidence, allowing educators to make well-informed decisions on the most efficacious teaching approaches. The use of learning analytics amplifies the capacities of EPM, therefore empowering educators, and institutions to make well-informed choices and enhance the entire educational experience for students. Many researchers have articulated early predictions as a fundamental task that can surely enhance students' performance by predicting either a category of learning outcome or by predicting a final score [21]. These predictions would be of more value if it could be done as early as possible.

The early prediction, as defined in [22], is the utilization of predictive techniques incorporated with performance metrics that accurately predict student outcome as early as possible. This early prediction often depends on the examination of diverse data sources, including student records, patterns of interaction, and past performance data and it is very critical as it affects the student's outcome related to their engagement in the course, engagement with the LMS.

Predicting the category of learning outcome is often a binary classification problem that distinguishes learners into "passed" and "failed" to estimate the likelihood of ensuring academic accomplishments in the future [22] [23]. Since predictive learning is a significant advantage of machine learning technology, it is frequently used to train the learning performance prediction model using a straightforward way [24].

Although this form of predictor can produce accurate predictions, it has several drawbacks. Due to low generalizability and high computing complexity, a huge amount of e-learning behavior data of different dimensions is acquired and recorded throughout the e-learning process. Some of the predictors will evaluate the integrated influence of learning behavior data (i.e., perform feature fusion processing) on the same type of learning behaviors data and then utilize it for training.

Finally, crucial learning behavior markers are not standardized, and those discovered by various investigations vary. This field of study has been unable to find important behavioral variables that can accurately predict learning performance. The objective is to better fulfil students' requirements by providing personalized learning pathways, adaptive evaluations and suggestions, or adaptive and just-in-time feedback [25]. However, a deeper knowledge of how learning processes are connected to and may be captured by the data accessible in contemporary digital learning environments is necessary [26], [27]. Different papers have focused on the students' records using different data mining point of view disregarding the time as significant feature in their studies leaving very few studies that tackled student sequential behavior pattern from a process mining point of view which is the focus in our study [28][29].

III. Process Mining

Process mining is a process centric discipline method that focuses on the examination of process execution records where it automatically constructs process patterns that accurately represent a business process by collecting relevant information from easily available event logs. Thus, it is a revolutionary way based on process model-driven methods and data mining offering complete tools for generating fact-based insights and supporting process changes [30], [31] [32].

Process mining involves three basic processes: defining an accurate process model, running a conformance check to reconcile event logs to a specified process model, and updating and extending the model. This emerging technology is utilized in numerous industries, including healthcare, manufacturing, education, and finance The main drawbacks of currently used methods include the inability to distinguish between some of the behavioral patterns and the failure to consider the natural process of learning activities (existing methods established in educational data mining and learning analytics primarily based on clustering and classification), where new analytical processes and approaches using process mining can be created to solve these issues [33].

Given that process mining is the approach that appears to hold the most promise in the body of available research. In a Process Mining Manifesto [34], the IEEE CIS Task Force on Process Mining identified three primary categories of process mining: Discovery, Conformance, and Enhancement. Discovery employs an event log to generate a model without prior knowledge, such as the alpha algorithm. Conformance checks involve comparing an existing process model with the event log of the same process, detecting, and measuring deviations. Enhancement attempts to extend or enhance an existing process model by utilizing actual process data recorded in an event journal. Extension adds a new perspective by cross correlating the model with the log. Enhancement involves modifying the model to better reflect reality. Using performance data as an example, a process model can be extended to display bottlenecks, service levels, throughput times, and frequencies. These process mining techniques aid in the comprehension and enhancement of processes [35] In the past terms like Workflow Mining and Automated Business Process Discovery (ABPD) were used as it has been noticeably clear that the process discovery category has constructed more than 85% of the relevant literature studies.

IV. Related Studies

This research focuses implementations of process mining techniques in the field of education and the main areas of research. There are sufficient reviews of the use of data mining in a variety of educational domains, but there were no studies that compiled and analyzed real cases in which process mining was applied to the educational domain.

Using an unbiased search strategy, the objective of the search is to locate different research papers with case studies that are related to our proposed work as feasible. We searched papers in four well-known general databases: IEEE, Emerald, Springer and Google Scholar. Some extremely broad search terms have been used to include all potential case studies. Iterative searches of these databases have been conducted between 2017 and 2023.The titles, abstracts, and keywords of each of the papers that resulted from the search for articles on the application of process mining in the field of education have been reviewed. The abstracts of each article were evaluated using the inclusion and exclusion criteria to determine which papers would be included in the review. Articles that have been examined in this area of research have been categorized into four main areas which are:

a) Discovering the student' behavior.
b) Predicting academic performance.
c) Discovering the learning paths and strategies.

This analysis seeks to accomplish the following: Compile current case studies on educational data process extraction. Then, evaluate those real case studies to identify their most pertinent sector areas in the domain with the distinguished techniques and algorithms of process mining that had been used. The purpose of this literature review is to summarize current developments in the EPM and to propose potential additional topics for further studying.

A. Discovering the Students' Behavior

Several studies have focused on students' behavior discovery through process mining to detect students' behavioral patterns while interacting with LMS or online MOOC environment such as Coursera or Udemy, or to understand their SRL (Self-Regulated Learning) behavior while engaging in these online environments.

In the context of education research, nine significant studies have been discovered, all focused on using process mining techniques to detect how students behave in the learning process. These studies vary in their methods, including fuzzy mining and social mining, often combined with machine learning or statistical approaches. Their main goal is to uncover significant sequence of actions, like visiting web pages, participating in online forums, and watching lecture videos, that influence student performance in both online and blended learning settings. Notable among these studies, Cenka et al.in [36] examined in blended learning and found that successful students followed more complex paths through fuzzy miner and manual clustering.

Yu Wang et al. in [37] created their own online modeling platform and used advanced techniques like feature engineering and sequential mining. Bernardi et al. in [38] explored coding behavior, Libor et al. in [39] detected students' behavioral pattern in quiz submission and Macak et al. in [52] inspected software development behavioral patterns in git logs. Meanwhile, Ghorbel et al. in [32], Dolak et.al in [40],Mukala et al. in [41], Ven den Beemt et al. in [42], Arpasat et al. in [43], explored different aspects of student behavior, shedding light on how high-achieving students tend to follow more organized sequences in activities like video watching, quiz submissions, and web page visits, behavior on social networks and. These studies collectively enhance our understanding of how students learn, offering valuable insights for educators and researchers.

As for detecting SRL (Students Regulated Learning) of students' behavior researchers in [44], [45],[30] researchers aimed to mine students' logs and extract the significant SRL activities based on Zimmerman's famous SRL and to contribute to the educational field by discovering the students behavioral pattern showing that high academic achieving students reflected a sequential pattern rather than those who followed unstructured behavior.

B. Predicting the Academic Performance

Prediction of student academic performance has been an important task in process mining in the educational domain where it has been tackled from different perspectives such as early prediction of student performance or prediction of trainee performance in each course.

Up to our knowledge, there are two attempts in the literature that covered the area of prediction students' performance using process mining techniques.

Umer et al. in [46] proposed an early prediction model where they used both Process discovery , conformance checking and machine learning techniques where Naïve Bayes and logistic regression showed higher prediction accuracy as classifiers on the generated dataset using conformance checking.

While Appice et al. in [47] wanted to predict the trainee outcome according to their behavior in a virtual learning environment noting that the utmost level of accuracy attained in prediction with J48 algorithm over KNN, LOG and RF resulting in the highest level of performance.

C. Discovering Learning Paths and Strategies

Process mining provides a process-centric methodology for comprehending the learning patterns of students and determining effective tactics that promote their academic output. In such a context researcher used different process mining techniques to detect and verify the students' learners' strategies. Four studies in this area of study were found to navigate students' activities, paths and to validate the learning paths.

Crosslin et al.[48] used pMiner R tool to extract models based on First-Order Markov Model (FOMM) and cluster analysis by using the Expectation Maximization (EM) algorithm to examine learning paths. Patel et al. in [49] developed graph-based discovery algorithm to detect frequent learning pathways from groups of students over 3 million anonymized students as a big dataset.

Some studies have proposed heuristic miner such as in [50] to validate the proposed learning trajectory with other techniques available in ProM framework and PM4Py library or such as Intayaod et al. in [51] to aid students to discover their own paths.

Through the exploration of student paths and strategies, educational institutions have the potential to augment the potential of education they provide, enhance student involvement and retention, and contribute to the overall enhancement of the learning experience.

V. Challenges and Future work

The limitation or the challenge of the process mining can vary according to the perspective or the direction of the study that has been tackled from.

From our perspective, there were few studies covered the area of process mining in the educational domain in general and specifically in prediction of student academic performance from a process mining point of view. So, we recommend the application of more studies in the learning domain using process mining and specifically in predicting the student academic performance since it is vital in timely intervention for students who need support or might be at risk.

Moreover, this area of the future work based on the presented study could be beneficial to students in their academic achievement as well as the institution as it helps also in timely data driven decision that could save them cost and time.

VI. Conclusion

The application of process mining in the field of education can provide stakeholders (teachers, instructors, students, etc.) with insights into the actual processes that are executed. Various forms of process mining can be used to uncover learning process models, verify compliance with predefined principles, and identify opportunities to enhance educational processes.

In this study as we explored the literature and found main areas in the educational domain that have been tackled through process mining such as detecting the students' patterns, predicting the academic performance, or detecting students' paths and strategies. Those different areas had gained a great advantage from different process mining techniques and algorithms throughout process mining three popular phases: process discovery, conformance checking, and enhancement. In addition to several machine learning, statistical techniques that have been implemented in different areas of EPM with many aspects left unexplored. The analysis of literature review outcomes shed light on the methodologies used by researchers and emphasized the potential benefits of this emerging field of study. We have highlighted prospective concerns and future challenges that need more investigation.

VII. Acknowledgment

We acknowledge the contributions of the researchers whose work has served as a foundation for this paper. Their dedication and advancements in the field of EPM have been instrumental in shaping our study.

References

[1] Crystal Fulton and Claire McGuinness, *Digital detectives: Solving information dilemmas in an online world.* . 2016.

[2] A. Charitopoulos, M. Rangoussi, and D. Koulouriotis, "On the Use of Soft Computing Methods in Educational Data Mining and Learning Analytics Research: a Review of Years 2010–2018," *Int J Artif Intell Educ*, vol. 30, no. 3, pp. 371–430, Oct. 2020, doi: 10.1007/s40593-020-00200-8.

[3] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 40, no. 6. pp. 601–618, Nov. 2010. doi: 10.1109/TSMCC.2010.2053532.

[4] J.-L. Sierra-Rodrígues, International Symposium on Computers in Education 16 2014.11.12-14 Logroño, and SIIE 16 2014.11.12-14 Logroño, *2014 International Symposium on Computers in Education (SIIE) 12-14 Nov. 2014, Logroño, La Rioja, Spain*.

[5] john Wiley & sons., *Clark, R. C., & Mayer, R. E. (2016). E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning.* . 2016.

[6] H. Drachsler and W. Greller, "Confidence in Learning Analytics," 2010.

[7] A. Bogarín, R. Cerezo, and C. Romero, "Discovering learning processes using inductive miner: A case study with learning management systems (LMSs)," *Psicothema*, vol. 30, no. 3, pp. 322–329, 2018, doi: 10.7334/psicothema2018.116.

[8] A. Yunita, H. B. Santoso, and Z. A. Hasibuan, "Research Review on Big Data Usage for Learning Analytics and Educational Data Mining: A Way Forward to Develop an Intelligent Automation System," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Jun. 2021. doi: 10.1088/1742-6596/1898/1/012044.

[9] A. F. Gontzis, S. Kotsiantis, C. T. Panagiotakopoulos, and V. S. Verykios, "A predictive analytics framework as a countermeasure for attrition of students," *Interactive Learning Environments*, vol. 30, no. 3, pp. 568–582, 2022, doi: 10.1080/10494820.2019.1674884.

[10] K. Mouri, F. Okubo, A. Shimada, and H. Ogata, "Bayesian network for predicting students' final grade using e-book logs in university education," in *Proceedings - IEEE 16th International Conference on Advanced Learning Technologies, ICALT 2016*, Institute of Electrical and Electronics Engineers Inc., Nov. 2016, pp. 85–89. doi: 10.1109/ICALT.2016.27.

[11] H. Waheed, S. U. Hassan, N. R. Aljohani, J. Hardman, S. Alelyani, and R. Nawaz, "Predicting academic performance of students from VLE big data using deep learning models," *Comput Human Behav*, vol. 104, Mar. 2020, doi: 10.1016/j.chb.2019.106189.

[12] D. Dessì, G. Fenu, M. Marras, and D. Reforgiato Recupero, "Bridging learning analytics and Cognitive Computing for Big Data classification in micro-learning video collections," *Comput Human Behav*, vol. 92, pp. 468–477, Mar. 2019, doi: 10.1016/j.chb.2018.03.004.

[13] Nen-Fu Huang, "VideoMark: A video-based learning analytic technique for MOOCs," 2017.

[14] R. K. Jena, "Sentiment mining in a collaborative learning environment: capitalising on big data," *Behavior and Information Technology*, vol. 38, no. 9, pp. 986–1001, 2019, doi: 10.1080/0144929X.2019.1625440.

[15] V. Dhanalakshmi, "Opinion mining from student feedback data using supervised learning algorithms," 2016.

[16] R. K. Jena, "Sentiment mining in a collaborative learning environment: capitalizing on big data," *Behavior and Information Technology*, vol. 38, no. 9, pp. 986–1001, 2019, doi: 10.1080/0144929X.2019.1625440.

[17] J.-P. Zhang, "Hybrid Learning and Ubiquitous Learning."

[18] Society for Learning Analytics and Research, H. and W. Association for Computing Machinery. Special Interest Group on Hypertext, SIGCHI (Group : U.S.), and Association for Computing Machinery, *LAK 2014 : fourth International Conference on Learning Analytics and Knowledge : 24-28, March 2014, Indianapolis, IN, USA*.

[19] A. Korhonen and J. Multisilta, "Learning analytics," in *New Ways to Teach and Learn in China and Finland: Crossing Boundaries with Technology*, Peter Lang AG, 2017, pp. 301–310. doi: 10.1177/0002764213498851.

[20] J. Wong *et al.*, "Educational theories and learning analytics: From data to knowledge: The whole is greater than the sum of its parts," in *Utilizing learning analytics to support study success*, Springer International Publishing, 2019, pp. 3–25. doi: 10.1007/978-3-319-64792-0_1.

[21] R. T. Osguthorpe and C. R. Graham, "Blended Learning Environments: Definitions and Directions. Quarterly Review of Distance Education," *Blended Learning Environments: Definitions and Directions. Quarterly Review of Distance Education*, 2003.

[22] S. Joksimović *et al.*, "How do we model learning at scale? A systematic review of research on moocs," *Rev Educ Res*, vol. 88, no. 1, pp. 43–86, Feb. 2018, doi: 10.3102/0034654317740335.

[23] S. B. Kotsiantis, C. J. Pierrakeas, and P. E. Pintelas, "Preventing student dropout in distance learning using machine learning techniques," in *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, Springer Verlag, 2003, pp. 267–274. doi: 10.1007/978-3-540-45226-3_37.

[24] N. Gedik, E. Kiraz, and M. Y. Özden, "The Optimum Blend: Affordances and Challenges of Blended Learning for Students," 2012.

[25] A. P. Rovai and H. M. Jordan, "Blended learning and sense of community: A comparative analysis with traditional and fully online graduate courses," *International Review of Research in Open and Distance Learning*, vol. 5, no. 2, 2004, doi: 10.19173/irrodl.v5i2.192.

[26] Y. Zhu, W. Au, and G. Yates, "University students' self-control and self-regulated learning in a blended course,"

Internet and Higher Education, vol. 30, pp. 54–62, 2016, doi: 10.1016/j.iheduc.2016.04.001.

[27] A. Jokhan, B. Sharma, and S. Singh, “Early warning system as a predictor for student performance in higher education blended courses,” *Studies in Higher Education*, vol. 44, no. 11, pp. 1900–1911, Nov. 2019, doi: 10.1080/03075079.2018.1466872.

[28] L. , R. E. , V. J. , & B. R. van der Aalst, “The Engineer of the Future. In K. Mertins, F. Benaben, R. Poler, and J. Bourrieres,” *Springer-Berlin*, vol. 7, pp. 13–28, 2014.

[29] O. Viberg, M. Hatakka, O. Bälter, and A. Mavroudi, “The current landscape of learning analytics in higher education,” *Computers in Human Behavior*, vol. 89. Elsevier Ltd, pp. 98–110, Dec. 01, 2018. doi: 10.1016/j.chb.2018.07.027.

[30] J. Maldonado-Mahauad, M. Pérez-Sanagustín, R. F. Kizilcec, N. Morales, and J. Munoz-Gama, “Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses,” *Comput Human Behav*, vol. 80, pp. 179–196, Mar. 2018, doi: 10.1016/j.chb.2017.11.011.

[31] K. Engelmann and M. Bannert, “Analyzing temporal data for understanding the learning process induced by metacognitive prompts,” *Learn Instr*, vol. 72, Apr. 2021, doi: 10.1016/j.learninstruc.2019.05.002.

[32] W. Hachicha, L. Ghorbel, R. Champagnat, C. A. Zayani, and I. Amous, “Using process mining for learning resource recommendation: A Moodle case study,” in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 853–862. doi: 10.1016/j.procs.2021.08.088.

[33] W. Van Der Aalst, “Process mining,” *Commun ACM*, vol. 55, no. 8, pp. 76–83, 2012, doi: 10.1145/2240236.2240257.

[34] W. M. P. van der Aalst, *Process Mining*. Springer Berlin Heidelberg, 2011. doi: 10.1007/978-3-642-19345-3.

[35] W. M. P. Aalst, *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, vol. 136, no. 2. 2011. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/18487736

[36] B. A. N. Cenka, H. B. Santoso, and K. Junus, “Analyzing student behavior in a learning management system using a process mining approach,” *Knowledge Management and E-Learning*, vol. 14, no. 1, pp. 62–80, Mar. 2022, doi: 10.34105/j.kmel.2022.14.005.

[37] Y. Wang, T. Li, C. Geng, and Y. Wang, “Recognizing patterns of student’s modeling behavior patterns via process mining,” *Smart Learning Environments*, vol. 6, no. 1, Dec. 2019, doi: 10.1186/s40561-019-0097-y.

[38] IEEE Computational Intelligence Society and Institute of Electrical and Electronics Engineers, *Ardimento, P., Bernardi, M. L., Cimitile, M., & De Ruvo, G. (2019, June). Learning analytics to improve coding abilities: a fuzzy-based process mining approach. In 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (pp. 1-7). IEEE.*

[39] L. Juhaňák, J. Zounek, and L. Rohlíková, “Using process mining to analyze students’ quiz-taking behavior patterns in a learning management system,” *Comput Human Behav*, vol. 92, pp. 496–506, Mar. 2019, doi: 10.1016/j.chb.2017.12.015.

[40] Dolak Radim, “Using Process Mining Techniques to Discover Student’s Activities, Navigation Paths, and Behavior in LMS Moodle,” L. Rønningsbakk, T.-T. Wu, F. E. Sandnes, and Y.-M. Huang, Eds., in Lecture Notes in Computer Science, vol. 11937. Cham: Springer International Publishing, 2019, pp. 129–138. doi: 10.1007/978-3-030-35343-8.

[41] P. , Mukala, Buijs, J. C. A. M., and W. M. P. & Van Der Aalst, “Exploring students’ learning behaviour in moocs using process mining techniques. BPM reports, 1510.”.

[42] A. van den Beemt, J. Buys, and W. van der Aalst, “Analysing structured learning behavior in Massive Open Online Courses (MOOCs): An approach based on process mining and clustering,” *International Review of Research in Open and Distance Learning*, vol. 19, no. 5, pp. 38–60, 2018, doi: 10.19173/irrodl.v19i5.3748.

[43] P. Arpasat, N. Premchaiswadi, P. Porouhan, and W. Premchaiswadi, “Applying process mining to analyze the behavior of learners in online courses,” *International Journal of Information and Education Technology*, vol. 11, no. 10, pp. 436–443, Oct. 2021, doi: 10.18178/ijiet.2021.11.10.1547.

[44] Mohamad Hasbullah Abu Bakar, “ Process Mining Approach to Understand Self Regulated- Learning in Moodle Environment,” *https://www.researchgate.net/journal/International-Journal-of-Advanced-Trends-in-Computer-Science-and-Engineering-2278-3091*, pp. 74–80, Jul. 2019.

[45] M. Bannert, P. Reimann, and C. Sonnenberg, “Process mining techniques for analyzing patterns and strategies in students’ self-regulated learning,” *Metacogn Learn*, vol. 9, no. 2, pp. 161–185, 2014, doi: 10.1007/s11409-013-9107-6.

[46] A. Appice *et al.*, “Training in a Virtual Learning Environment : A Process Mining Approach,” in *In 2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS*, 2020, pp. 1–8.

[47] R. Umer, T. Susnjak, A. Mathrani, and S. Suriadi, “On predicting academic performance with process mining in learning analytics,” *Journal of Research in Innovative Teaching & Learning*, vol. 10, no. 2, pp. 160–176, Jul. 2017, doi: 10.1108/jrit-09-2017-0022.

[48] M. Crosslin, K. Breuer, N. Milikić, and J. T. Dellinger, “Understanding student learning pathways in traditional online history courses: utilizing process mining analysis on clickstream data,” *Journal of Research in Innovative Teaching and Learning*, vol. 14, no. 3, pp. 399–414, Nov. 2021, doi: 10.1108/JRIT-03-2021-0024.

[49] N. Patel, C. Sellman, and D. Lomas, “Mining Frequent Learning Pathways from a Large Educational Dataset,” May 2017, [Online]. Available: http://arxiv.org/abs/1705.11125

[50] E. Machado Real, P. Pimentel, L. Vieira De Oliveira, J. C. Braga, and I. Stiubiener, “Educational Process Mining for Verifying Student Learning Paths in an Introductory Programming Course.”

[51] Intayoad, W., Kamyod, C., & Temdee, P. (2018, February). Process mining application for discovering student learning paths. In 2018 International Conference on Digital Arts, Media and Technology (ICDAMT) (pp. 220-224). IEEE.

[52] Macak, M., Kruzelova, D., Chren, S., & Buhnova, B. (2021). Using process mining for git log analysis of projects in a software development course. Education and information technologies, 26(5), 5939-5969.