

Discovering Student Learning Paths: An Educational Process Mining Approach in Moodle

Supriyanto¹, Taufiq Ismail², Fariz Setyawan³

^{1,2}Informatics Department, Universitas Ahmad Dahlan, Indonesia

³Mathematics Education Department, Universitas Ahmad Dahlan, Indonesia

Article Info

Article history:

Received Jun 29, 2025

Revised Oct 21, 2025

Accepted Nov 12, 2025

Keywords:

educational process mining

heuristic miner

inductive miner

student learning path

ABSTRACT

E-learning platforms like Moodle are critical to modern education, with their effectiveness deeply reliant on fostering optimal student engagement. A thorough understanding of how students interact with these platforms is therefore essential for enhancing the learning experience. This study aimed to analyze student interaction patterns within Moodle by employing educational process mining techniques. The core objective was to uncover hidden behavioral patterns and gain valuable insights into the underlying learning processes. To achieve this, we utilized both heuristic miner and inductive miner algorithms to analyze Moodle's extensive event log data. The effectiveness of various student activity variants was rigorously assessed through fitness checking. This study presents a novel, integrated analytical approach combining frequent sequential pattern mining, user behavior analysis, feature engineering, and process mining to comprehensively evaluate student learning effects in Moodle. While applying both Heuristic Miner and Inductive Miner algorithms to extensive Moodle event logs, we not only generated precise process models highlighting effective and ineffective student activity sequences but also uncovered unique challenges, such as the Inductive Miner's inability to accurately model the 'Tugas' (assignment) component's complex activity patterns. These findings offer distinct, actionable insights for refining Moodle course design and delivery, moving beyond general observations to pinpoint specific pedagogical interventions. Ultimately, our work advances the understanding of student behavior and academic performance within the Moodle ecosystem by providing a granular, data-driven methodology for optimization.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Supriyanto,

Informatics Department,

Universitas Ahmad Dahlan, Yogyakarta, Indonesia.

Email: supriyanto@tif.uad.ac.id

1. INTRODUCTION

The rapid evolution of online learning platforms has revolutionized the education landscape, enabling widespread access to educational resources and facilitating remote learning. E-learning, introduced to improve the quality of education and learning, provides a more accessible and flexible way of acquiring knowledge and skills [1], [2]. Its goals have become increasingly diverse [3], ranging from accessible communication [4] and speedy delivery of material to overcoming distance and time constraints [5] and fostering greater student interaction.

Learning Management Systems like Moodle are now integral to both fully online and hybrid learning environments [6]. Moodle, in particular, enhances collaborative and adaptive learning and facilitates online assessments [7]. It provides an array of learning components such as materials, assignments, quizzes, forums, and questionnaires. These e-learning systems generate vast amounts of data from student interactions, offering valuable insights into student behavior, learning patterns, and performance. Analyzing this data can improve instructional design,

personalize learning experiences, and enhance educational outcomes. The transition to online learning, particularly following the declaration of the COVID-19 pandemic in March 2020, accentuated the critical need for educators to cultivate and implement robust strategies for the effective design, development, and delivery of online educational resources. Crucially, during this transformative period, educational institutions that already possessed established Learning Management Systems were able to execute a seamless and rapid transition to online learning modalities [8], [9].

The application of learning analytics to Moodle data can be highly beneficial in enhancing the quality of learning [10]. Educational Data Mining (EDM) [11], and learning analytics can significantly improve student engagement and learning outcomes. These technologies are essential tools for institutions committed to enhancing the educational experience [12]. Personalized learning can be achieved by analyzing student activity patterns when accessing learning components, allowing teachers to identify the most effective approaches. To ensure transparent and interpretable results, the collected data must be well-structured and thoroughly evaluated [13]. Selecting an appropriate analytical approach is crucial for the success of EDM [14].

This work aims to analyze student activity patterns, using process mining techniques, to provide instructors with insights that enable them to improve the learning environment. By analyzing the process models, instructors can see which learning activities or resources students engage with most frequently and identify areas where students may be struggling [15]. This facilitates a data-driven approach to resource allocation and the refinement of pedagogical strategies. While previous studies have explored individual facets of Moodle data analysis or applied process mining in educational contexts, this work introduces a distinctive analytical approach. Our methodology uniquely integrates frequent sequential pattern mining, detailed user behavior analysis, and feature engineering alongside process mining. This comprehensive framework provides a holistic and granular evaluation of student learning effects, specifically by critically linking process mining outcomes with numerical performance evaluations. This integrated perspective allows for a more comprehensive understanding of complex learning dynamics than typically achieved by isolated analytical techniques, thereby addressing a crucial gap in current educational process mining literature [16].

Educational Process Mining stands as a significant research domain, strategically positioned at the intersection of Process Mining and Educational Data Mining. By employing process mining techniques, EPM is capable of dissecting educational data and subsequently enhancing learning processes. The capacity to derive profound insights into student behavior, learning strategies, and overall performance is a key attribute, ultimately fostering improvements in the quality of education [17]. One of EPM's strengths lies in its ability to construct comprehensive and concise educational process models. These models facilitate the verification of observed behaviors, uncover previously overlooked implicit knowledge, and provide feedback to stakeholders [18]. This feedback benefits students, teachers, researchers, and institutions by facilitating a nuanced understanding of the educational process and offering recommendations for improvement.

Process mining is a data mining subfield that extracts insights from event logs using sophisticated techniques. It enables the identification of inefficiencies, optimization of process performance, and data-driven decision-making. PM transforms data analysis, providing a deeper understanding of processes for growth and improvement [19]. It uncovers hidden patterns in various processes, while EDM produces models that provide educators with information to organize and direct students toward a better learning path. Therefore, this study applies process mining techniques to analyze student activity patterns using Moodle event log data. By visualizing student learning paths [20], and identifying bottlenecks, process mining offers actionable insights for improving the design and delivery of online courses. Huge amounts of data are generated by learners' activities through an LMS and can be used to develop the learning environment and improve the overall learning experience [21].

The proliferation of data in modern learning environments necessitates the adoption of advanced analytical techniques such as process mining. Traditional statistical methods often prove inadequate in capturing the dynamic and sequential nature of learning processes. Process mining

fills this gap by offering a process-centric perspective on learning, empowering educators and researchers to understand how students learn, rather than simply what they learn. This approach is particularly critical in online learning environments, where student interactions are predominantly digital and readily trackable [22].

Recent research has showcased the effectiveness of process mining across diverse educational contexts. For instance, studies have employed process mining to compare the learning paths of successful and unsuccessful students, revealing notable differences in their activities throughout a course [23]. Furthermore, researchers have explored the application of process mining, clustering, and conformance-checking techniques to analyze learning processes in MOOCs [24], pinpointing patterns associated with favorable outcomes [25]. The Inductive Miner algorithm has been used to analyze student's self-regulated learning processes. Specifically, the Inductive Miner algorithm has shown promise in educational data analysis, as it can discover optimal models in terms of fitness [26].

The imperative to personalize learning experiences and enhance student outcomes in an increasingly digital world underscores the urgency of this work. Beyond the general application of process mining techniques, this study makes several specific contributions to the expanding body of knowledge on EPM. We offer novel, data-driven practical insights by rigorously analyzing student interaction patterns within various Moodle components, including the identification of specific algorithmic limitations in modeling certain learning activities. These findings are translated into precise recommendations for refining the design and delivery of online courses, particularly within the Moodle environment. The goal is to leverage data-driven insights to create more effective, efficient, and engaging learning experiences for all students.

Previous investigations have explored Moodle-based process mining, with a focus on the preprocessing phase and the development of tools to streamline early analysis stages [27]. These efforts facilitate the efficient utilization of EPM on Moodle event log data by teachers and researchers [28]. Among the process mining algorithms, Heuristic Miner and Inductive Miner are of particular interest. Heuristic Miner is a procedurally based algorithm that has been used successfully for process discovery [29]. Inductive Miner is a newer algorithm that has shown promising results in the educational domain [26]. In fact, in one study, the Inductive Miner algorithm produced the best results in terms of fitness when tested on a dataset of student interactions on the Moodle 2.0 platform [26]. Studies also suggest the need for improvement of the Heuristics Miner to increase the quality of the resulting process model [30]. While the Heuristic Miner and Inductive Miner algorithms have been applied in prior educational process mining studies, including those demonstrating Inductive Miner's superior fitness, our research extends this line of inquiry by performing an in-depth, component-specific analysis within Moodle. We do not merely apply these algorithms; instead, we critically evaluate their effectiveness across diverse Moodle activities, identifying nuanced performance differences and revealing specific instances where even advanced algorithms like Inductive Miner encounter significant modeling challenges, such as with the 'Tugas' component. This granular evaluation and subsequent derivation of actionable, component-level pedagogical interventions represent a key departure from and advancement upon existing research.

2. RESEARCH METHOD

We extracted the event log for this study from a Moodle Learning Management System course. The dataset encompasses student activity data collected across three semesters of lectures, totaling 360,042 recorded activities. Specifically, this data was sourced from Object Oriented Programming course over the aforementioned four academic semesters, spanning from September 2020 to March 2023. The participants in this study consisted of 200 undergraduate students enrolled in the Informatics Department at Universitas Ahmad Dahlan. These students were primarily in their second year of studies during the data collection period. The selection of these courses was based on their consistent use of Moodle. Furthermore, the Object-Oriented

Programming course, being fundamental to the curriculum, provided a robust dataset for analyzing diverse student interaction patterns.

These activities are associated with various learning components accessed by the students, which serve as the attributes for analysis. Table 1 provides a comprehensive list of the learning components available in the course.

Table 1. The raw data of component.

No	Component	Description
1	Sistem	The Learning Management System (LMS) includes all its features and student login conditions.
2	Tugas	Learning component in the LMS contains student assignments.
3	Forum	Learning components in the LMS for student discussion rooms.
4	Questionnaire	Components in the LMS for student satisfaction surveys.
5	Kuis	Components in the LMS such as Assignment Components for college assessments.
6	Presensi Kehadiran	Components in the LMS are used for student attendance.
7	Pelajaran	Components in the LMS that represent general lecture materials.
8	Berkas	The components in the LMS are used as a complement to the learning materials.
9	Pengiriman Teks Daring	A feature used by students to write answer texts for Assignment and Quiz components.
10	Pengiriman Berkas	A feature used by students to link files when working on Assignments and Quizzes components.
11	User Report	Components in the LMS to show student learning outcomes.

Figure 1 illustrates the Educational Process Mining stages implemented in this research, aligning with a typical EPM workflow that includes Data Collection, Data Preprocessing, and Process Discovery. As noted in previous research, raw data from Moodle requires preprocessing to ensure its quality and suitability for process mining techniques. In Moodle, student learning activities automatically generate data, which is then stored as logs in the database. These logs are prepared using data preprocessing techniques to ensure data quality and usefulness for subsequent analysis.

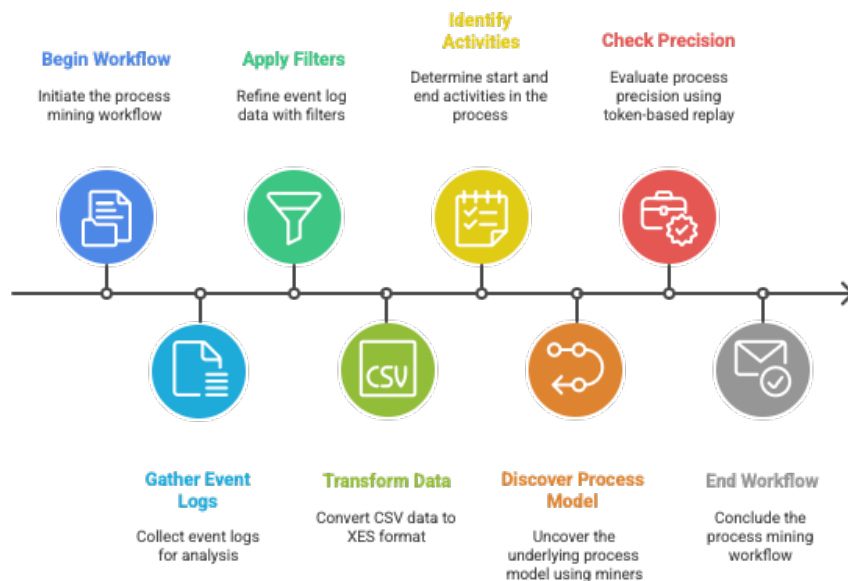


Figure 1. Education Process Mining

The Case ID is a critical attribute for effective data analysis and sorting. We sort activities based on their timestamps to identify the corresponding Case IDs. Furthermore, it is beneficial to align the user's name with that of the lecturer teaching the subject, facilitating a more accurate assessment of each user's status. Given the exclusive focus on students in this research, we should remove any data pertaining to lecturer status. Removing duplicate lines and users with lecturer status ensures data accuracy.

To generate a data frame, the attributes are defined as follows: 'user' as Case ID, 'component' as Activity Key, and 'timestamp' as Timestamp Key. The data frame is then saved in XES format using the PM4Py library [31]. The Extensible Event Stream format is a standard for

event logs, capable of describing the contents of relationships where a log consists of multiple traces, and each trace contains several events. Objects such as logs, traces, or events can also have attributes. This feature is valuable, as data attributes that remain constant for a log or a trace can be stored at that level. The XES standard offers a more seamless approach to retaining such information.

This research identified six start activities and four end activities, resulting in 244 variants of student activities. Table 2 presents a summary of variants in the series of student activities, categorized by the start activity. We then apply process mining techniques to derive student activity patterns for each start activity. Process mining can discover, monitor, and improve real processes (learning processes) by extracting knowledge from event logs to understand how the learning process occurs.

This study leverages process mining methodologies, specifically the Heuristic Miner and Inductive Miner algorithms, to extract insights and patterns from the collected data [32]. HM is recognized for its ability to manage short loops and detect infrequent event logs; however, it does not guarantee the correctness of the resulting process model [33]. HM employs the Directly Follows Graph to construct a heuristic net, which is subsequently transformed into a Petri net graph for process analysis and optimization. DFGs are graphical representations where events are depicted as nodes connected by directed edges, labeled with frequency and performance metrics. The frequency in a DFG indicates how often one event directly follows another, while performance measures, such as the average time interval between events, provide insights into the efficiency of transitions [34].

Inductive Miner excels at handling infrequent transitions, skipping, and looping sections within a process model [35]. IM employs splitting techniques—such as sequential, parallel, detailed, and prewired splits—on the event log to generate sub-logs, which are recursively processed until the desired model scenario is achieved [35]. The process tree generated by the Inductive Miner algorithm can be converted into a Petri net. This algorithm is designed to create robust process models capable of handling incomplete and noisy data [36]. According to [37], the inductive miner algorithm can create a model based on the event log well.

Conformance checking is a widely used technique in process mining that involves comparing a process model with its corresponding event log. This method requires two main inputs: the event log and the process model and generates a diagnostic report as output. The report presents a detailed analysis of the discrepancies and similarities between the model and the log, aiding in the identification of potential issues and improving the overall efficiency and effectiveness of the process [38]. The primary goal of conformance checking is to ascertain the degree of concordance between the process model and the event log [39]. Token-Based Replay is a method utilized to assess the suitability of a process. During the evaluation of a process instance, a trace and a Petri net model are utilized to identify which transitions have been executed and where tokens are either absent or extra. Token-based replay allows for the identification of errors in process execution by detecting instances where tokens are missing or unexpectedly present in the Petri net model, indicating deviations from the expected process flow [40]. If the final marking is considered necessary, a trace may be deemed appropriate if it successfully arrives at the final marking without any instances of missing or remaining tokens. This information is crucial in identifying any issues or inefficiencies in the process and can be used to optimize the process flow.

3. RESULTS AND DISCUSSION (11 PT)

The experimental setup involved a detailed examination of the initial data preprocessing steps, the specifics of which are outlined in Table 2. We then conducted process discovery for each identified start activity, leveraging the pm4py discovery library.

When applying process mining, the visual representations of student activities, such as Petri nets, serve as critical tools. These diagrams illustrate the sequence and flow of student interactions within the Moodle platform.

Table 2. Start-End activity and variants identification.

Start Activities	End Activities	Variant
{'Tugas': 8}	{'Sistem': 6, 'Kuis': 1, 'Tugas': 1}	8
{'Forum': 7}	{'Sistem': 5, 'Kuis': 2}	7
{'Questionnaire': 5}	{'Sistem': 5}	5
{'Presensi Kehadiran': 2}	{'Kuis': 1, 'Tugas': 1}	2
{'Halaman': 3}	{'Sistem': 2, 'Kuis': 1}	3
{'Tugas': 8}	{'Sistem': 6, 'Kuis': 1, 'Tugas': 1}	8

A significant observation was the presence of "noise" in the resulting process models, particularly those generated by the Heuristic Miner, as illustrated in Figure 2. Within the context of Petri nets for educational data, noise represents irregular or uncommon student activity flows, indicating inconsistencies in how students interact with Moodle activities and deviations from clear, structured processes. For instance, the Heuristic Miner models exhibited substantial noise, signifying a wide and often unpredictable array of student behaviors. This 'noisy' behavior suggests that students are not always following a linear or expected path through the learning materials. This "noisy" behavior suggests that students are not always following a linear or expected path through the learning materials. This could imply:

- Exploratory Learning: Students might be navigating the platform in a non-linear fashion, exploring resources in a way that doesn't conform to a predefined instructional design.
- Confusion or Difficulty: High noise levels could also signal areas where students are struggling, frequently switching between activities without clear progress, or returning to previous steps repeatedly due to misunderstanding or difficulty.
- Inefficient Pathways: Some "noisy" paths might represent less efficient learning strategies, where students spend more time or take more steps than necessary to complete a task.

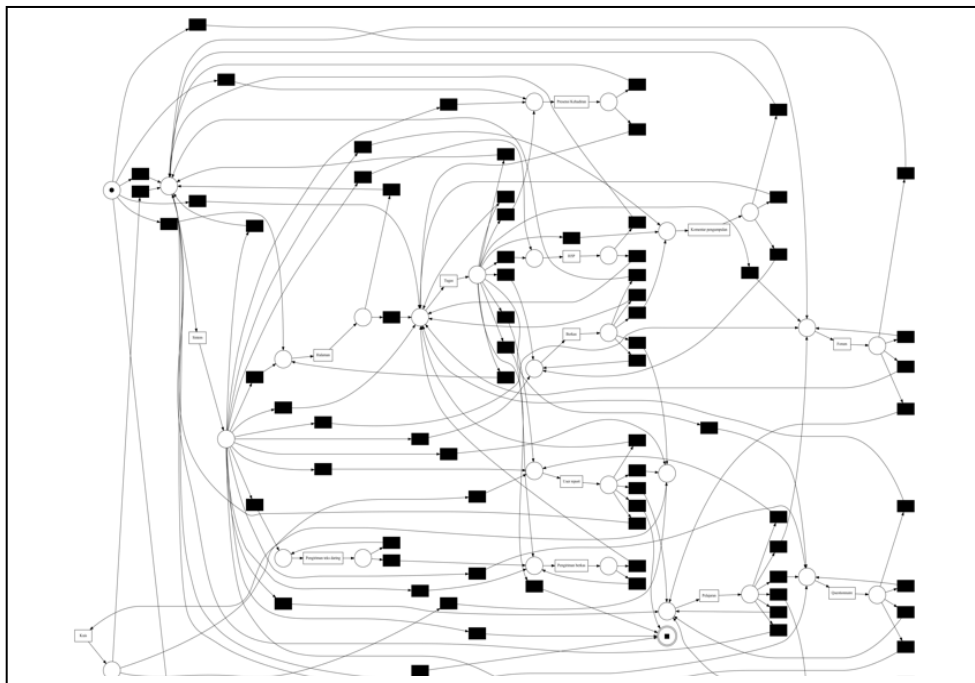


Figure 2. Petri net visualization of heuristic miner process discovery. ▲

While the Inductive Miner generally generated process models with more structured and orderly activity flows as illustrated in Figure 3, a notable exception was the 'Tugas' component. Even with the Inductive Miner, the 'Tugas' model displayed considerable noise. This finding is

highly significant because it highlights a specific challenge in modeling student engagement with assignments and offers crucial insights into their learning behavior [16]. It suggests that:

- Varied Engagement with Assignments: Students use the 'Tugas' activity in diverse ways, making it difficult to capture a single, dominant process flow [Current Document]. This variety could stem from different approaches to problem-solving, varying levels of preparation, or individual preferences in how they tackle assignments.
- Impact on 'Assignment' Component: The document notes that frequent switching between the assignment activity and other Moodle resources without completing the assignment negatively influences engagement with the 'Assignment' component [Current Document]. This highlights a critical behavioral pattern: "noise" in this context might reveal students initiating an assignment, getting sidetracked or overwhelmed, and then engaging with other resources (or simply leaving the assignment incomplete) before returning later, or not at all. This lack of a coherent, direct path to completion is a key insight derived from the noisy Petri net.

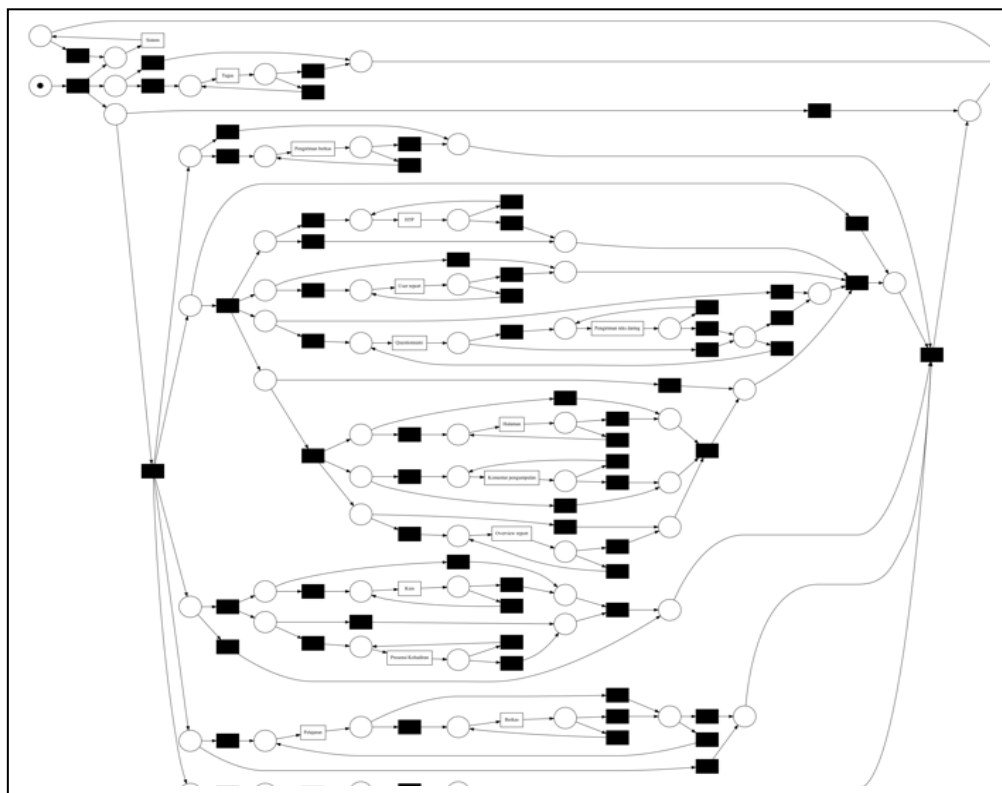


Figure 3. Petri net visualization of inductive miner process discover. ▲

The token-based replay measurements shown in Table 3 revealed that the Heuristic Miner process model consistently exhibited lower fitness percentages compared to the Inductive Miner. Token-based replay is a technique for assessing the fitness of a process model by simulating the execution of event logs on it and checking for discrepancies. The higher fitness percentages observed for the Inductive Miner generally suggest a better fit for the log data. However, the unique behavior around the 'Tugas' component reveals a significant and previously underexplored challenge for the Inductive Miner in contexts of high process variability, especially when students engage with critical activities like assignments in highly diverse and non-linear ways. This granular insight into the challenges of modeling the 'Tugas' component, which is central to student assessment and learning, offers a unique contribution to EPM by highlighting the need for more adaptive or specialized process mining techniques for such complex educational interactions.

To gain a more granular understanding of model accuracy and its implications, we utilized token-based replay. This technique assesses the "fitness" of a process model by comparing it with the actual event log data. Fitness Percentage: The fitness percentage reflects the degree of concordance between the process model and the event log. A higher fitness percentage means the model accurately represents the observed student behavior, validating the discovered learning path. Conversely, a lower fitness percentage indicates significant discrepancies, meaning the model does not fully capture the reality of student actions.

Table 3. Token-Based Replay Fitness Comparison per Component.

Attribute	Heuristic PFT	Inductive PFT
Sistem	1.80%	100%
Tugas	0.00%	0.00
Forum	14.28%	100%
Questionnaire	0.00%	100%
Halaman	0.00%	100%
Presensi Kehadiran	0.00%	100%
Sistem	1.80%	100%

The token-based replay measurements shown in Table 3 revealed that, while the inductive miner generally showed higher fitness percentages, the start activity of the 'Tugas' component yielded a 0% fit for traces and an average trace fitness value below 1. This 0% fit is a profound indicator that none of the observed student activity traces perfectly matched the generated model for the 'Tugas' component. The low average trace fitness further emphasizes large differences between the model and actual student actions. This low fitness for the 'Tugas' component strongly suggests:

- **Unpredictable Assignment Engagement:** Student behavior around assignments is highly varied and often deviates significantly from any idealized or expected path. There is no single, consistent "way" students approach or complete assignments within the Moodle environment, making it challenging to model their behavior accurately.
- **Instructional Design Implications:** For educators, this indicates that current instructional designs or platform functionalities for assignments might not be aligning with how students actually interact. It highlights a need to investigate why students are not following more structured or anticipated learning paths for assignments. This could lead to refining assignment instructions, breaking down complex tasks, or providing more guided support to help students engage more effectively with assignment components.

The findings from our process mining analysis offer several significant pedagogical implications for educators and course designers:

- **Identification of Learning Difficulties:** The "noise" in Petri nets and low fitness scores for the 'Tugas' component serve as strong indicators of where students might be experiencing difficulties or confusion. Unstructured, inconsistent, or highly varied activity patterns suggest that students are not following optimal learning pathways, potentially leading to frustration or incomplete understanding.
- **Revealing Suboptimal Learning Strategies:** The identified "negative influence" of frequently switching between assignment activities and other resources without completion highlights a suboptimal learning strategy. This could manifest as procrastination, distraction, or a lack of self-regulation. Recognizing these patterns allows educators to intervene proactively.
- **Mismatched Expectations:** A significant divergence between the expected (modelled) and actual student behavior (event log), particularly evident in low fitness, implies a mismatch between instructional design intent and student execution. This prompts a re-evaluation of how tasks are presented, supported, and integrated within the Moodle environment.
- **Need for Adaptive Support:** The diverse nature of student engagement, especially around complex tasks like assignments, underscores the need for more adaptive and personalized support mechanisms. A "one-size-fits-all" approach may not be sufficient when student learning paths are highly varied.

The insights derived from process mining can be directly translated into actionable strategies for improving instructional design:

Refine Assignment Design and Scaffolding:

- **Break Down Complex Tasks:** For components like 'Tugas' that exhibit high noise or low fitness, consider breaking down large assignments into smaller, more manageable sub-tasks [Current Document]. Each sub-task could have its own clear set of instructions and submission points, reducing the likelihood of students becoming overwhelmed and switching away.
- **Clearer Instructions and Rubrics:** Enhance the clarity and explicitness of assignment instructions and rubrics. Ambiguity can lead to varied, and often inefficient, student approaches.
- **Provide Intermediate Milestones:** Introduce interim deadlines or check-ins to guide students through the assignment process and prevent prolonged periods of disengagement.

Optimize Resource Integration and Navigation:

- **Contextual Resource Recommendations:** Leverage the observation that students often return to the 'Sistem' component after assignments [Current Document]. Instead of a generic dashboard, design the platform to offer contextual recommendations for related resources or follow-up activities, creating a more interconnected learning journey [Current Document].
- **Guided Pathways for Key Activities:** For critical learning paths, consider implementing more structured sequences, perhaps using Moodle's "restrict access" features to ensure students complete prerequisite activities before moving on, thereby reducing "noise" in core processes.
- **Visual Cues and Progress Indicators:** Implement visual progress bars or checklists within the LMS to help students track their advancement through modules or assignments, potentially reducing feelings of being sidetracked or lost.

Proactive Interventions and Feedback Mechanisms:

- **Automated Alerts for Deviant Paths:** Future integrations could involve systems that identify when a student's activity path deviates significantly from effective patterns (e.g., spending excessive time without progress, frequent task switching without completion). These could trigger automated nudges or alerts for instructors.
- **Targeted Feedback:** Use process mining insights to inform more targeted feedback. If the models show students struggle with a particular sequence, instructors can provide specific guidance on how to navigate that part of the course more effectively.

Continuous Course Improvement:

- **Iterative Design Cycle:** Process mining should be integrated into an iterative course design cycle. After implementing changes based on previous findings, new event log data can be analyzed to evaluate the effectiveness of these pedagogical adjustments, fostering continuous improvement.

Figure 4 presents a comparison of the heuristic and inductive miner algorithms, which indicates that the inductive miner generally produces more accurate process models for analyzing activity log data in Moodle. This aligns with existing research suggesting that the Inductive Miner is a newer algorithm that not only yields superior results compared to other algorithms but also generates valid models. However, the observed inconsistency in the 'Tugas' component suggests a need to refine the inductive mining approach to improve its accuracy and effectiveness in identifying process models for specific learning activities. For example, the inconsistency modeling of the 'Tugas' component might be improved by using enhanced fitness checking. The integration of the EPM framework with machine learning techniques is a promising avenue for future research. For instance, by identifying common successful learning pathways through process mining, machine learning algorithms could be trained to predict student performance based on their activity patterns. This approach could enable the creation of tailored learning recommendations, such as suggesting specific resources or activities to students who are deviating from successful pathways.

These recommendations would be grounded in process models and customized to individual student learning styles.

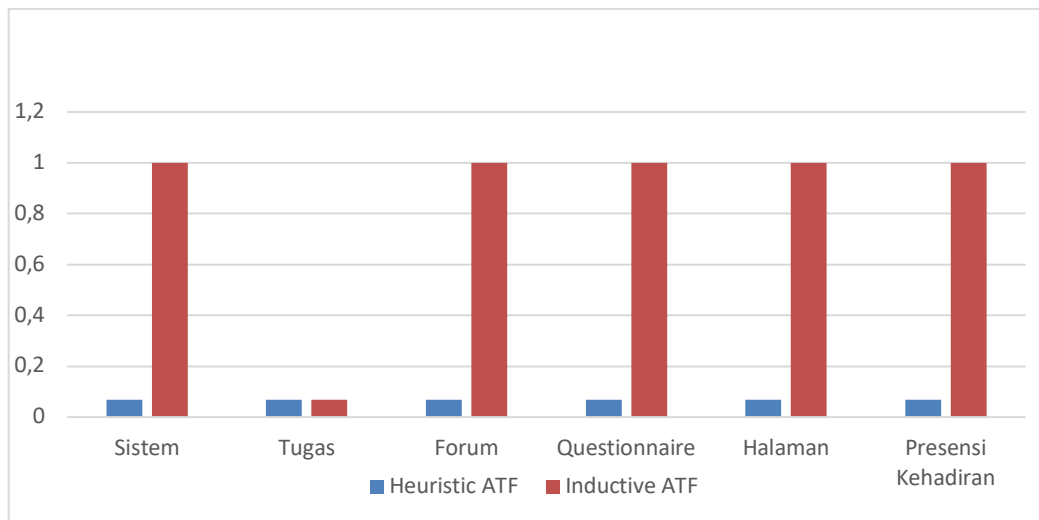


Figure 4. TBR ATF Comparison.

By interpreting the noise in Petri nets and the fitness results from token-based replay, we gain actionable insights into student learning behavior. These insights help educators understand where students might be struggling, identify suboptimal learning strategies, and refine the design of learning materials and Moodle environments to cultivate a more engaging and effective online learning experience. For example, the high frequency of students returning to the 'Sistem' component after completing assignments suggests an opportunity to integrate learning materials and activities more effectively, perhaps by recommending related resources rather than simply returning to the dashboard.

4. CONCLUSION

Crucially, our study uniquely identified a negative correlation between specific sequences of student activities within the Moodle LMS and their performance in the 'Assignment' component. This granular identification of detrimental behavioral patterns, beyond general activity levels, provides educators with unprecedented, actionable insights to refine the design and delivery of learning materials and pedagogical strategies, directly targeting areas for improved student outcomes. These findings provide educators with actionable insights to refine the design and delivery of learning materials, optimizing student outcomes. By applying advanced process mining techniques, particularly the inductive miner, to Moodle event log data, this study offers a comprehensive and efficient analysis of student activity patterns, contributing to a deeper understanding of learning processes within the platform. The results of this analysis can be leveraged to enhance students' academic performance through data-driven interventions and pedagogical adjustments.

Future research could explore integrating process mining with learning resource recommendations to provide personalized learning experiences. By analyzing student activity patterns, educators can identify resources that are most effective for different types of learners and recommend them accordingly.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support provided by Lembaga Penelitian dan Pengabdian Universitas Ahmad Dahlan, which made this research possible. Their commitment to fostering research is deeply appreciated.

REFERENCES

- [1] E. Elihami, M. Musdalifah, and H. Hasan, "Increasing the Higher of Think through Moodle Application: The Effects of Learning Procees in the Higher Education System," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Feb. 2021. doi: 10.1088/1742-6596/1783/1/012054.
- [2] M. Alenezi, "Deep dive into digital transformation in higher education institutions," *Educ Sci (Basel)*, vol. 11, no. 12, Dec. 2021, doi: 10.3390/educsci11120770.
- [3] A. Çebi, "How e-learning readiness and motivation affect student interactions in distance learning?," *Educ Inf Technol (Dordr)*, vol. 28, no. 3, pp. 2941–2960, 2023, doi: 10.1007/s10639-022-11312-0.
- [4] N. C. Eli-Chukwu *et al.*, "Challenges confronting e-learning in higher education institutions in Nigeria amid Covid -19," *Journal of Applied Research in Higher Education*, vol. 15, no. 1, pp. 238–253, 2023, doi: 10.1108/JARHE-09-2021-0346.
- [5] E. E. Egorov, M. P. Prokhorova, T. E. Lebedeva, O. A. Mineeva, and S. Y. Tsvetkova, "Moodle LMS: Positive and Negative Aspects of Using Distance Education in Higher Education Institutions," *Propósitos y Representaciones*, vol. 9, no. SPE2, 2021, doi: 10.20511/pyr2021.v9nspe2.1104.
- [6] O. K. Xin and D. Singh, "Development of Learning Analytics Dashboard based on Moodle Learning Management System," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, Apr. 2021, doi: 10.14569/IJACSA.2021.0120793.
- [7] S. H. P. W. Gamage, J. R. Ayres, and M. B. Behrend, "A systematic review on trends in using Moodle for teaching and learning," Dec. 01, 2022, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1186/s40594-021-00323-x.
- [8] S. N. Sato *et al.*, "Navigating the New Normal: Adapting Online and Distance Learning in the Post-Pandemic Era," *Educ Sci (Basel)*, vol. 14, no. 1, p. 19, Dec. 2023, doi: 10.3390/educsci14010019.
- [9] D. J. Camacho and J. M. Legare, "Pivoting to online learning—The future of learning and work," *The Journal of Competency-Based Education*, vol. 6, no. 1, Mar. 2021, doi: 10.1002/cbe2.1239.
- [10] F. A. Simanca Herrera, R. G. Crespo, L. R. Baena, and D. Burgos, "A Solution to Manage the Full Life Cycle of Learning Analytics in a Learning Management System: AnalyTIC," *Revista Iberoamericana de Tecnologías del Aprendizaje*, vol. 14, no. 4, pp. 127–134, Nov. 2019, doi: 10.1109/RITA.2019.2950148.
- [11] P. Bachhal, S. Ahuja, and S. Gargrish, "Educational Data Mining: A Review," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Aug. 2021. doi: 10.1088/1742-6596/1950/1/012022.
- [12] H. Aldowah, H. Al-Samarraie, and W. M. Fauzy, "Educational data mining and learning analytics for 21st century higher education: A review and synthesis," Apr. 01, 2019, *Elsevier Ltd*. doi: 10.1016/j.tele.2019.01.007.
- [13] H. E. Abdelkader, A. G. Gad, A. A. Abohany, and S. E. Sorour, "An Efficient Data Mining Technique for Assessing Satisfaction Level With Online Learning for Higher Education Students during the COVID-19," *IEEE Access*, vol. 10, pp. 6286–6303, 2022, doi: 10.1109/ACCESS.2022.3143035.
- [14] R. S. Baker, "Challenges for the Future of Educational Data Mining: The Baker Learning Analytics Prizes," *Journal of Educational Data Mining*, vol. 11, no. 1, pp. 1–17, 2019, doi: <https://doi.org/10.5281/zenodo.3554745>.
- [15] R. Cerezo, A. Bogarín, M. Esteban, and C. Romero, "Process mining for self-regulated learning assessment in e-learning," *J Comput High Educ*, vol. 32, no. 1, pp. 74–88, Apr. 2020, doi: 10.1007/s12528-019-09225-y.
- [16] Y. Wang, T. Li, C. Geng, and Y. Wang, "Recognizing patterns of student's modeling behaviour patterns via process mining," *Smart Learning Environments*, vol. 6, no. 1, Dec. 2019, doi: 10.1186/s40561-019-0097-y.

- [17] F. Liu, L. Zhao, J. Zhao, Q. Dai, C. Fan, and J. Shen, "Educational Process Mining for Discovering Students' Problem-Solving Ability in Computer Programming Education," *IEEE Transactions on Learning Technologies*, vol. 15, no. 6, pp. 709–719, 2022, doi: 10.1109/TLT.2022.3216276.
- [18] O. Esteban Bajo, I. Amarasinghe, N. F. Gutiérrez-Páez, and D. Hernández-Leo, "Using Process Mining Techniques to Discover the Collective Behaviour of Educators in a Learning Community Platform," in *Collaboration Technologies and Social Computing*, vol. 13632, L.-H. Wong, Y. Hayashi, C. A. Collazos, C. Alvarez, G. Zurita, and N. Baloian, Eds., Cham: Springer International Publishing, 2022, pp. 175–189. [Online]. Available: https://link.springer.com/10.1007/978-3-031-20218-6_12
- [19] F. Wafda, T. Usagawa, and E. Mahendrawathi, "Systematic Literature Review on Process Mining in Learning Management System," in *2022 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, IEEE, 2022, pp. 160–166. doi: 10.1109/IAICT55358.2022.9887428.
- [20] J. Costa, A. Azevedo, and L. Rodrigues, "Educational Process Mining based on Moodle courses: a review Educational Process Mining based on Moodle courses: a review of literature of literature," in *Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI'2020)*, 2020, pp. 1–10. [Online]. Available: <https://aisel.aisnet.org/capsi2020>
- [21] M. Salihoun, "State of Art of Data Mining and Learning Analytics Tools in Higher Education," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 21, p. 58, Apr. 2020, doi: 10.3991/ijet.v15i21.16435.
- [22] T. H. A. N. D. L. Mehnen, "Cross-course Process Mining of Student Clickstream Data – Aggregation and Group Comparison," *arXiv (Cornell University)*, Apr. 2024, doi: 10.48550/arxiv.2409.14244.
- [23] E. M. Real, E. Pinheiro Pimentel, L. V. De Oliveira, J. Cristina Braga, and I. Stiubiener, "Educational Process Mining for Verifying Student Learning Paths in an Introductory Programming Course," in *2020 IEEE Frontiers in Education Conference (FIE)*, IEEE, 2020, pp. 1–9. doi: 10.1109/FIE44824.2020.9274125.
- [24] A. Van Den Beemt, J. Buijs, and W. Van Der Aalst, "Analysing Structured Learning Behaviour in Massive Open Online Courses (MOOCs): An Approach Based on Process Mining and Clustering," 2018.
- [25] N. Ahmad Uzir *et al.*, "Discovering Time Management Strategies in Learning Processes Using Process Mining Techniques," in *Transforming Learning with Meaningful Technologies*, vol. 11722, M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, and J. Schneider, Eds., Cham: Springer International Publishing, 2019, pp. 555–569. [Online]. Available: http://link.springer.com/10.1007/978-3-030-29736-7_41
- [26] R. Cerezo, A. Bogarín, M. Esteban, and C. Romero, "Process mining for self-regulated learning assessment in e-learning," *J Comput High Educ*, vol. 32, no. 1, pp. 74–88, 2020, doi: 10.1007/s12528-019-09225-y.
- [27] P. Nafasa, I. Waspada, N. Bahtiar, and A. Wibowo, "Implementation of Alpha Miner Algorithm in Process Mining Application Development for Online Learning Activities Based on MOODLE Event Log Data," in *2019 3rd International Conference on Informatics and Computational Sciences (ICICoS)*, IEEE, 2019, pp. 1–6. doi: 10.1109/ICICoS48119.2019.8982384.
- [28] 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.
- [29] J. De Smedt, J. De Weerd, and J. Vanthienen, "Multi-paradigm Process Mining: Retrieving Better Models by Combining Rules and Sequences," in *Lecture notes in computer science*, Springer Science+Business Media, 2014, pp. 446–453. doi: 10.1007/978-3-662-45563-0_26.
- [30] Sofie De Cnudde, "Improving the quality of the Heuristics Miner in ProM 6.2," Apr. 2014. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417414003509>

- [31] A. Berti, S. van Zelst, and D. Schuster, "PM4Py: A process mining library for Python[Formula presented]," *Software Impacts*, vol. 17, Sep. 2023, doi: 10.1016/j.simpa.2023.100556.
- [32] A. Stefanini, D. Aloini, E. Benevento, R. Dulmin, and V. Mininno, "A process mining methodology for modeling unstructured processes," *Knowledge and Process Management*, vol. 27, no. 4, pp. 294–310, Oct. 2020, doi: 10.1002/kpm.1649.
- [33] R. D. Ridwanah, R. Andreswari, and R. Fauzi, "Analysis and Implementation of TELKOM University Lecture Business Processes Evaluation on Heuristic Miner Algorithm: A Process Mining Approach," in *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*, IEEE, Jan. 2022, pp. 243–248. doi: 10.1109/ISMODE53584.2022.9742973.
- [34] S. Chanifah, R. Andreswari, and R. Fauzi, "Analysis of Student Learning Pattern in Learning Management System (LMS) using Heuristic Mining a Process Mining Approach," in *2021 3rd International Conference on Electronics Representation and Algorithm (ICERA)*, IEEE, 2021, pp. 121–125. doi: 10.1109/ICERA53111.2021.9538654.
- [35] 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.
- [36] H. Alqaheri and M. Panda, "An Education Process Mining Framework: Unveiling Meaningful Information for Understanding Students' Learning Behavior and Improving Teaching Quality," *Information (Switzerland)*, vol. 13, no. 1, Jan. 2022, doi: 10.3390/info13010029.
- [37] M. W. Wibisono, A. P. Kurniati, and G. A. A. Wisudiawan, "Process Mining using Inductive Miner Algorithm to Determine the actual Business Process Model," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 4, p. 1128, Aug. 2022, doi: 10.30865/jurikom.v9i4.4769.
- [38] I. Nuritha and E. R. Mahendrawathi, "Structural Similarity Measurement of Business Process Model to Compare Heuristic and Inductive Miner Algorithms Performance in Dealing with Noise," *Procedia Comput Sci*, vol. 124, pp. 255–263, 2017, doi: 10.1016/j.procs.2017.12.154.
- [39] V. Naderifar, S. Sahran, and Z. Shukur, "A review on conformance checking technique for the evaluation of process mining algorithms," *TEM Journal*, vol. 8, no. 4, pp. 1232–1241, Nov. 2019, doi: 10.18421/TEM84-18.
- [40] A. Berti, "Reviving Token-based Replay: Increasing Speed While Improving Diagnostics," *ATAED@ Petri Nets/ACSD*, 2019, [Online]. Available: files/2294/Berti - Reviving Token-based Replay Increasing Speed Whil.pdf