Instal : Jurnal Komputer

E-ISSN: 2808-683X
Edisi : Volume 17 Nomor 08 | September 2025

Available online at https://journalinstal.cattleyadf.org/index.php/Instal/index

Applying Process Mining to Analyze Anomalies in Student Registration

Gede Agung Ary Wisudiawan[1], Indra Lukmana Sardi[2]
[1,2]Fakultas Informatika Universitas Telkom

ARTICLE INFO

Article history:

Received : 11 September 2025
Revised : 15 September 2025
Accepted : 30 September 2025

Keywords:

process mining,
conformance checking,
academic registration,
anomali detection

ABSTRACT

This study applies process mining to analyze the student course registration process using activity log data from the SIRAMA application. A process model was discovered and evaluated through conformance checking by comparing it with the standard operating procedures (SOP) to detect anomalies. Four key metrics were employed: fitness, precision, generalization, and simplicity. The evaluation results reveal a fitness score of 0.846, indicating that most events in the log are represented by the model, although some transitions remain uncaptured. The relatively low precision score of 0.69531 suggests that the model is overly permissive, allowing many potential paths that never occurred in practice. In contrast, the generalization score of 0.99992 demonstrates that the model is highly robust, capable of representing unseen but valid cases without overfitting. Further audit and analysis identified three anomalous transitions: "Start" directly followed by "DELETE," "Start" followed by "SIAP REGISTRASI" without prior "ADD," and "ADD" directly followed by "END." Expert validation confirmed that these anomalies were not user-driven but caused by technical issues such as delayed log writing (non-atomic transactions) and timestamp overflow errors. To mitigate such anomalies, this study recommends enforcing ACID principles—particularly atomicity—along with strict timestamp validation and automatic correction of invalid log formats. The findings highlight that process mining is not only effective for modeling real-world academic processes but also serves as a diagnostic tool for detecting systematic deviations and improving system reliability.

Corresponding Author:
Gede Agung Ary Wisudiawan
Universitas Telkom
Email: degunk@telkomuniversity.ac.id

INTRODUCTION

SIRAMA (Sistem Registrasi Akademik Mahasiswa) is the platform used by Telkom University students to perform course registration. The application is directly integrated with the iGracias and SITU database systems. Its users include students, academic advisors, and heads of study programs (HoPs). Through this platform, students can select courses and complete their KRS

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

(Kartu Rencana Study). Academic advisors are responsible for approving their students' registration, while HoPs are authorized to approve registrations for all students in their respective programs and to adjust course quotas if necessary. In addition, administrators manage student data, both for those who are about to register and those who have already completed the registration process. The academic registration process in SIRAMA involves three main activities: selecting courses, deleting previously selected courses, and finalizing the registration with the status "Ready to Register" (SIAP REGISTRASI). Once registration is completed, it must be validated and approved by the academic advisor or head of the study program. The process concludes with the issuance of the KSM (Kartu Study Mahasiswa). All user activities in SIRAMA are automatically recorded and stored as system-generated event logs.

These event logs provide an opportunity for research, particularly in analyzing user activities through process mining. Process mining is an event data-driven analysis approach used to understand, monitor, and improve processes[1]. By utilizing activity data recorded in information systems, process mining enables process mapping, anomaly detection, and conformance analysis[2]. Based on this background, the following research questions are formulated How can process mining be applied to academic registration processes in SIRAMA? How can the resulting process models of SIRAMA be evaluated? How can anomalies in student registration processes be identified?

The objectives of this study are to develop a process model of academic registration in SIRAMA using event log data, to evaluate the resulting model against the SOP, and to identify anomalies within the student registration process.

Related studies show that process mining can be used to assess conformance between event logs recorded in large-scale information systems such as Enterprise Resource Planning (ERP) and Standard Operating Procedures (SOP), thereby identifying anomalies in process flows[3]. Six types of anomalies or fraudulent behaviors may occur in business processes[4]: (1) skipped activity, where required activities are omitted; (2) wrong throughput time, where activities are performed too early or too late compared to SOP-defined time limits; (3) wrong resource, where activities are performed by unauthorized actors; (4) wrong duty, where one person or system executes multiple activities that should be handled by different parties; (5) wrong pattern, where the sequence of activities deviates from the defined process model; and (6) wrong decision, where decisions are inconsistent with SOP-based rules. Each type represents a deviation from the SOP in terms of activity sequence, timing, resource assignment, or decision-making.

Based on these findings, process mining can be applied to detect anomalies in academic registration processes. In this study, process mining is used to discover process models from SIRAMA event logs, which are then evaluated against the official SOP of academic registration.

METHODS

This section describes the Process Mining Project Methodology (PM2), which serves as the research framework in this study, as illustrated in Figure 1. The methodology begins with a general overview of PM2, followed by a detailed discussion of each stage. PM2 provides a systematic guideline for both researchers and educational institutions in conducting process mining projects, particularly in the context of process behavior analysis and anomaly detection [5].

The objectives of a process mining project can be concrete—such as detecting registration patterns that deviate from standard procedures—or exploratory, such as understanding students' behavioral tendencies during the registration process. Through the PM2 framework,

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

these objectives are formulated into research questions that can be addressed using student activity log data.

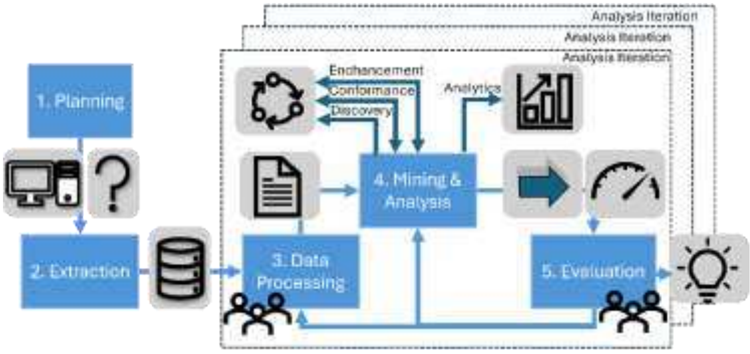

Figure 1. The Process Mining Project Methodology (PM2) (adapted from van Eck, Lu, Leemans, & van der Aalst [6])

The methodology is carried out iteratively, starting from planning to evaluation, ensuring that each phase contributes to refining the process model and addressing the research questions. Ultimately, the application of PM2 in this study is expected to generate insights for identifying anomalies and formulating recommendations to improve the academic registration system. The following subsections provide a detailed explanation of the PM2 methodology adopted in this research.

1. Planning
   The planning phase aims to initiate the process mining project in a structured manner by defining the project objectives and formulating the research questions to be answered through process data analysis[6]. In general, process mining projects are usually driven by two main motivations: improving business process performance or evaluating process compliance with established rules and policies.
   In this phase, the existing business process within the organization serves as the input, while the output consists of clearly formulated research questions aligned with the project objectives, as well as a list of information systems that store the execution data to be analyzed. There are three main activities in the planning phase[7]:
    a. Identifying the business process to be analyzed.
    b. Formulating the research questions in accordance with the project's objectives.
    c. Establishing the research team responsible for carrying out the process mining project.
   The second activity involves the formulation of research questions[8]. Organizational objectives are translated into research questions that can be addressed using event logs. These questions may relate to aspects such as process duration, quality, resource utilization, or costs. Research questions often begin at an abstract level and are then refined through exploratory analysis, ultimately leading to concrete findings.
   The third activity is the formation of the project team. A successful project team should consist of members with diverse backgrounds and expertise, including:
    a. Business owners, who hold responsibility for the business process,
    b. Business experts, who understand the operational and contextual aspects of the process,
    c. System experts, who are familiar with the technical details of the system and event log data,
    d. Process analysts, who specialize in process analysis and process mining techniques.
    e. Collaboration between business experts and process analysts is particularly crucial to ensure that the insights generated are both relevant and actionable.

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

In conclusion, the planning phase serves as a critical foundation for any process mining project. Projects that lack well-defined research questions, a suitable process for analysis, or an appropriately structured team risk producing irrelevant or unusable results. Careful planning ensures that process mining outcomes effectively support decision-making and lead to meaningful improvements within the organization.

2. Extraction.

This phase aims to collect event data from the information system supporting the selected business process. The inputs for this phase include the research questions and the information system where process data is stored, while the outputs consist of structured event data and, if available, an initial process model (e.g., workflow model from an ERP system). The main activities in this phase are: determining the scope, extracting event data, and transferring process knowledge.

The first activity, determining scope, involves defining the level of detail, time range, attributes to be collected, and correlations among different data sources. Once the scope is established, the second activity, extracting event data, is performed. In this step, raw data is retrieved from the system and consolidated into a structured format, typically as a table or CSV file [9]. Each row should represent a single event and contain essential attributes such as case_id, activity, timestamp, user, and other relevant information.

The third activity is transferring process knowledge, which bridges the gap between technical data and business context. This step requires close collaboration with business experts, who provide domain knowledge to process analysts [10]. Knowledge transfer can be achieved through interviews, workshops, or documentation such as SOPs and process flow diagrams. The objective is to ensure that analysts fully understand the process context, enabling them to correctly interpret and prepare the data.

In conclusion, the extraction phase is a critical bridge between research questions and raw system data. By carefully defining the scope, performing structured data extraction, and embedding business knowledge, analysts can ensure that the resulting event data is valid, well-structured, and ready for subsequent process mining activities.

3. Data Processing.

The purpose of the data processing phase is to transform event data into an event log that is ready for analysis. At this stage, the log is refined and optimized for process mining and subsequent analyses. In some cases, an existing process model can also be utilized to guide or enhance data processing [11]. The four main activities in this phase are: creating views, aggregating events, enriching logs, and filtering.

The first activity, creating views, involves constructing the event log from raw event data by defining cases, i.e., grouping events that belong to the same process instance[12]. Once a coherent event log is obtained, the second activity, aggregating events, is performed to simplify the log and reduce complexity in the mining results. Event aggregation can take several forms:

a. Is-a aggregation, where different activities are grouped into a common category (e.g., "ADD initial" and "ADD revised" are merged into "ADD") without reducing the number of events.
b. Part-of aggregation, where multiple sub-activities are merged into a single event, or the reverse (specialization), where events are decomposed into finer levels.
c. Hierarchical aggregation, which organizes activities based on attributes such as location (e.g., city → province → country).

The third activity, enriching logs, consists of adding supplementary information to the event log, either from external sources or derived internally from the log itself. Enrichment improves the analytical power of the log by providing additional contextual attributes.

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

The final activity is filtering, which aims to reduce complexity and focus the analysis on relevant aspects of the data. Three main types of filtering are commonly applied:

a. Slice-and-dice (attribute filtering): filtering events or cases based on attributes such as activity type, user, or timestamp (e.g., selecting only students with more than five activities).
b. Variance-based filtering: grouping event logs based on behavioral similarity, often supported by clustering techniques. This is particularly useful for complex and highly variable processes.
c. Compliance-based filtering: removing or isolating cases and events that do not conform to established rules or standard process models. This method can flexibly target normal or exceptional (anomalous) cases.

In summary, the data processing phase serves as a crucial bridge between raw extracted data and the mining stage. Decisions made in this phase—such as case definition, aggregation, enrichment, and filtering—significantly influence the quality and relevance of the final process mining results.

4. Mining and Analysis.

This phase represents the core of process mining, where analytical techniques are applied to the event log in order to answer the research questions and derive insights into both process performance and compliance with organizational rules [13]. The main objective is twofold: (i) to provide answers to exploratory or specific research questions, and (ii) to evaluate process execution in terms of performance indicators such as duration and efficiency, as well as its alignment with normative models or standard operating procedures (SOPs).

When research questions are still abstract, the analysis typically begins with an exploratory process discovery, producing an initial process model from the event log. This model serves as the foundation for further analysis. Once the general patterns are identified, more focused analysis can be conducted to answer concrete questions. The input to this phase consists of event logs and, if available, documented process models or discovery results. The outputs are findings that address performance, compliance, and potential areas for improvement.

The main activities in this phase are: process discovery, conformance checking, enhancement, and process analytics [14].

a. Process Discovery: Automatically generates a process model from the event log without relying on prior knowledge. This provides a baseline representation of the actual process execution.
b. Conformance Checking: Compares the actual process recorded in the event log with the prescribed process model (or SOP). Its goal is to detect deviations (anomalies), such as skipped activities, incorrect ordering, or timing inconsistencies.
c. Enhancement: Improves or enriches the process model based on real execution data. This may involve adding performance attributes such as time, cost, or frequency to each activity, or updating outdated models to better reflect actual practice.

Process Analytics: Uses the enriched and validated models to derive insights into performance, bottlenecks, resource usage, and other aspects of process behavior.

A central aspect of conformance checking is the use of quantitative metrics to evaluate the alignment between event logs (L) and process models (M). Three key metrics are employed [15]:

a. Fitness – measures the extent to which the process model (M) can reproduce the behavior recorded in the event log (L). A high fitness score indicates that the model successfully explains most observed traces, while low fitness reveals that many real-world cases cannot be captured by the model.

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

b. Precision – evaluates how restrictive the process model (M) is in representing real behavior. A precise model avoids overgeneralization, i.e., allowing paths that never occur in the event log. High precision ensures that the model only permits observed behavior, while low precision allows excessive, unrealistic variants.
c. Generalization – reflects the ability of the model to accommodate valid but unseen cases, ensuring that the model is not overly specific (overfitting). A well-generalized model strikes a balance between faithfully representing observed cases and flexibly accounting for potential new cases.

Together, these metrics provide a rigorous assessment of model quality and support the second objective of this research, namely evaluating the discovered process models against real-world execution.

$$fitness = \frac{|L \cap M|}{|L|} \quad (1)$$

$$precision = \frac{|L \cap M|}{|M|} \quad (2)$$

The mining and analysis phase provides concrete answers to the research questions formulated at the beginning of the study. Through the combination of process discovery, conformance checking, model enhancement, and additional analytics, researchers are able to comprehensively understand the actual process execution and derive data-driven conclusions.

5. Evaluation.

The evaluation phase aims to link the analytical results obtained in the previous stage with ideas for process improvement, ensuring that the main objectives of the project can be achieved[16]. This stage serves as a crucial bridge between analytical findings and actionable steps in the development of the analyzed business process or system. The inputs of this phase include the discovered process model and the results of performance and conformance analysis, while the outputs consist of process improvement ideas and/or new research questions for the next iteration. The key activities in the evaluation stage are diagnosis, verification, and validation [6].

Diagnosis involves accurately interpreting the analysis results, such as understanding the process flows represented in the model, identifying unusual patterns or deviations (e.g., frequent violations of standard procedures), and refining or formulating new research questions arising from unexpected findings. Verification refers to checking whether the results are consistent with the original data or with how the system actually operates. This activity is typically carried out by analysts through technical comparisons with logs, system scripts, or other data sources. Validation, on the other hand, assesses whether the analytical findings align with real-world conditions from the stakeholders' perspective. This can be conducted through discussions, interviews, or direct observation of process participants (e.g., academic advisors or students).

One of the main challenges in process mining projects is the knowledge gap between process analysts and business experts. Therefore, involving process experts is essential in this phase to correctly interpret findings, ensure relevance and applicability, and formulate realistic and impactful improvement ideas. The evaluation stage is thus a crucial moment where analytical results are transformed into actionable insights. Through diagnosis, verification, and validation, process mining moves beyond mere findings to strategic steps aimed at improving and refining the analyzed process.

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

RESULTS AND DISCUSSION

This chapter presents the findings of process mining applied to the event log data of student academic registration activities in the SIRAMA application. The primary objective of the analysis is to construct the actual process model based on system-recorded data, evaluate its conformance with the Standard Operating Procedures (SOP), and identify potential anomalies within the registration process. The discussion begins with the initial process modeling obtained through the process discovery stage, followed by conformance checking to assess how well the constructed model reflects real-world events. Subsequently, an audit is conducted on minority paths that deviate from the SOP, including investigations into potential causes of anomalies such as timestamp overflow errors and illogical activity sequences. This chapter also discusses technical approaches and corrective measures, such as enforcing atomicity in transaction logging and implementing input time validation, to enhance the integrity and reliability of the academic registration process in the future. All analytical results are supported by process model visualizations, filtering outcomes, and expert validation to ensure accurate interpretation of each potential anomaly identified.

Planning

The three main activities in the Planning stage are: selecting the business process to be analyzed, formulating research questions, and assembling the project team.
The process mining project begins with the selection of the business process to be investigated. This study focuses on the student academic registration process. The registration process is carried out through an information system called SIRAMA, which can be accessed at https://sirama.telkomuniversity.ac.id
. The primary users of SIRAMA are academic advisors, students, and heads of study programs. The registration period is announced according to the academic calendar, which is officially provided by the Academic Service Standards Division (BSLA) of Telkom University. This information is publicly accessible at https://baa.telkomuniversity.ac.id/kalender-akademik-2-2/. The overall process of student academic registration is illustrated in Figure 2.

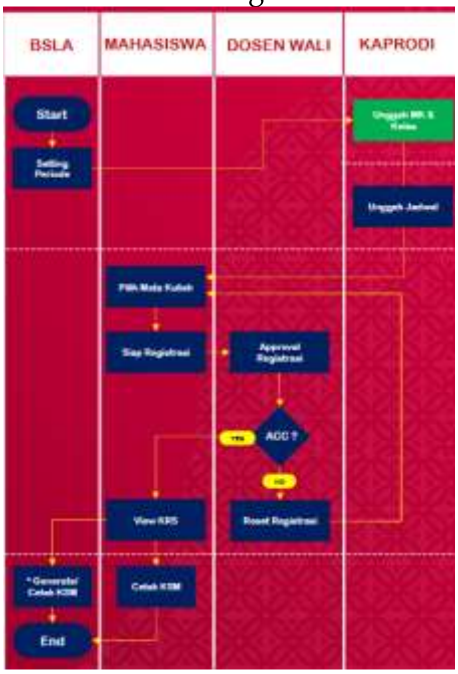

Figure 1. Academic Registration Business Process of Telkom University Students [20]

Based on the selected business process, the next step is to formulate the research questions. The research questions in this study are as follows:

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

a. How can process mining be applied to the academic registration process in the SIRAMA application?
b. How can the process model of academic registration in SIRAMA be evaluated?
c. How can anomalies in the student registration process in SIRAMA be identified?

This study involves academic advisors as business experts, who possess practical experience in the context and operations of the student registration process, and the researcher as the process analyst, who provides expertise in process analysis and process mining techniques.

Extraction

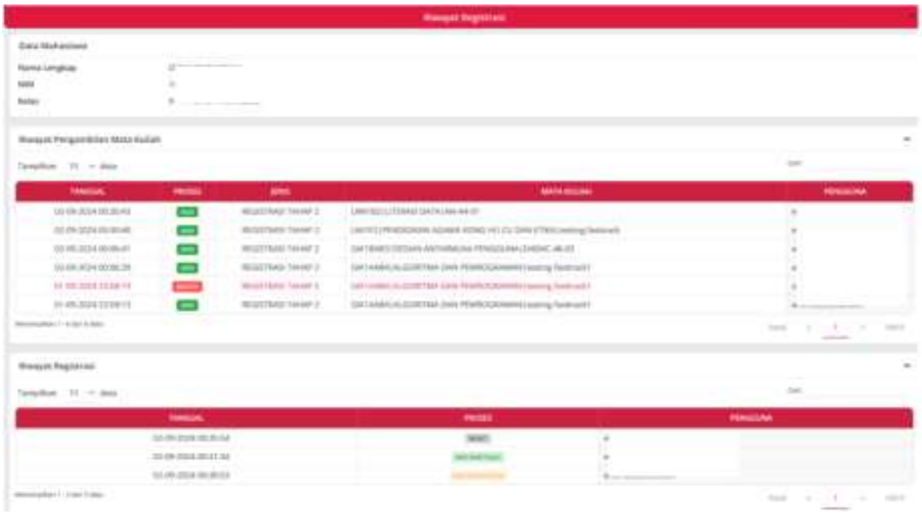

Figure 3. Academic Registration Activity Data in the SIRAMA Application [21]

In this study, academic advisors serve as business experts. The SIRAMA application records user activities (logs) for every interaction performed within the system. Based on interviews with business experts, it was confirmed that activity data can be accessed by academic advisors. Furthermore, the interviews revealed that SIRAMA is only accessible during the official academic registration period. Figure 3 presents a screenshot of one of SIRAMA's features, which displays the recorded activity data.

SIRAMA logs capture all actions performed by students when interacting with the system. From interviews with business experts and direct observation of the application, it was identified that the activity data includes the following attributes:

a. "Tanggal" (date): records the date and time when an activity is performed,
b. "Proses" (process): specifies the type of activity (e.g., ADD for adding a course, DELETE for dropping a course, SIAP REGISTRASI indicating that the student has finalized course selection),
c. "Jenis" (type): indicates the registration type,
d. "Mata kuliah" (course): specifies the course being added or deleted,
e. "Pengguna" (user): records the user performing the activity.

Given these characteristics, the data meets the requirements for process mining, which requires three essential attributes: timestamp, case, and event.

This study focuses on extracting event logs from the Spring 2024/2025 academic registration period, capturing all accessible data during that timeframe. To align with the research objectives, the data processing stage filters the logs to include only student users. The required attributes for process mining are: tanggal (timestamp), pengguna (case), and jenis (event). Each row in the log represents one event, recording when a user performed an activity in SIRAMA.

The extraction stage was carried out with the assistance of business experts (academic advisors). Through interviews and observation, the scope of data was defined, event data was

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

extracted and stored in a structured flat file (CSV), and the process context was clarified. This ensured that the researcher obtained valid, well-structured data ready for the subsequent data processing stage.

Data Processing

The purpose of data processing is to transform raw event data into an event log that is ready for analysis, optimize the log for process mining, and ensure its suitability for subsequent analysis. Existing process models can also be leveraged to support data preparation. In the case of SIRAMA, the user activity data contains both the course enrollment history and the registration history. This study specifically uses the course enrollment history data.

Before applying process mining, the dataset must be filtered based on user roles. The filtering step removes users with non-student roles. To determine which users were categorized as students, the researcher conducted interviews with the business expert (academic advisor).

The next step in data processing is the selection of relevant attributes from the raw dataset. The original data contains five attributes: tanggal (date), proses (process), jenis (type), mata kuliah (course), and pengguna (user). However, process mining requires three core attributes: case, event, and timestamp. Through observation and consultation with the business expert, the appropriate attributes for process mining were determined. A summary of the mapping between raw data attributes and process mining requirements is presented in the following table.

After completing the data processing stage, the prepared event log is ready for the next phase, namely process mining, which will be described in the subsequent section.

Table 1. Attributes Utilized in Process Mining

Dataset Attribute	Process Mining Attribute
Tanggal (Date)	*Time Stamp*
Pengguna (User)	*Case_id*
Proses (Process)	*Event*

Mining and Analysis

The mining and analysis stage represents the core of process mining, where analytical techniques are applied to the event log to answer the research questions and gain insights into process performance and compliance. In this study, the mining stage was conducted using the Disco software developed by Fluxicon. Aligned with the first research objective—"To develop the actual process model of academic registration in the SIRAMA application"—the main output of this stage is the process model derived from student registration event logs.

Figure 4 presents the initial process model obtained from the academic registration data. Overall, the discovered process model is consistent with the sequence of activities defined in the official Standard Operating Procedures (SOP). In the model, activities are represented by rectangles, while transitions between them are shown as arrows. The primary activities observed in the registration process are ADD, DELETE, and READY FOR REGISTRATION (SIAP REGISTRASI).

Process Model Description

a. Start
   The entry point of the process. Out of 36 cases, 34 directly continued to the ADD activity.
b. ADD (Course Addition)

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

This is the initial and most dominant activity, with a total of 400 occurrences. It reflects students adding courses to their study plan.

- A total of 253 self-loops (ADD → ADD) were recorded, indicating that students frequently explored and added multiple courses in a single registration session.
- From ADD to DELETE, 105 transitions were observed, suggesting that students often removed previously added courses.
- From ADD to READY FOR REGISTRATION, 34 transitions occurred, reflecting students who finalized their registration immediately after course selection.

c. DELETE (Course Deletion)
This activity occurred 143 times. From DELETE to ADD, 112 transitions were recorded, highlighting iterative behavior where students revised and adjusted their course selections.

d. READY FOR REGISTRATION
This activity signals that students completed course selection. A total of 46 events were recorded, with 35 transitions directly leading to END, indicating that most students finalized their registration after declaring readiness.

e. END
The termination point of the process. All 36 cases successfully reached this endpoint.

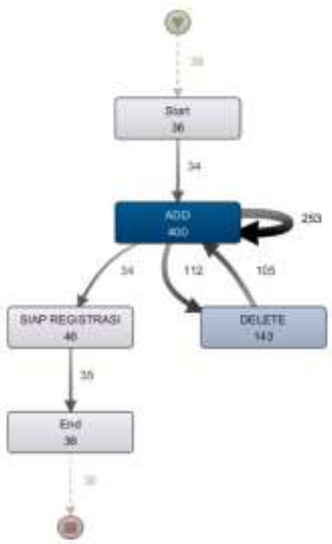

Figure 4. Student Academic Registration Initial Process Model

The analysis stage employed conformance checking, a technique used to compare the actual process (event log) with the ideal process as defined by the Standard Operating Procedure (SOP). The objective is to identify deviations such as skipped activities, incorrect ordering, or discrepancies in execution time.
Conformance checking in this study was evaluated using three key metrics:

a. Fitness – measures the extent to which the behavior recorded in the event log can be explained by the process model. A maximum value of 1 indicates that all events in the log are fully represented by the model.
b. Precision – assesses the alignment of the model with the actual event log by focusing on how many paths in the model are truly executed in reality. The fewer “empty” or unused paths, the higher the precision.
c. Generalization – indicates the model’s ability to represent valid future cases not yet observed in the log. This reflects whether the model is too specific (overfitted) or sufficiently flexible.

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

This conformance checking procedure supports the second objective of the research: "to evaluate the discovered process model."

Based on the results, the process model achieved a Fitness value of 0.846, indicating that most events in the log are captured by the model. However, the presence of deviations suggests that certain activities or transitions were either missing or anomalous in the registration process. The Precision value of 0.69531 reflects relatively low precision, meaning that the model was overly permissive, incorporating many paths that did not actually occur in student behavior. This highlights the need for refinement to reduce irrelevant paths and increase model accuracy. In contrast, the Generalization value of 0.99992 is nearly perfect, demonstrating that the model has strong capability to accommodate new but valid cases while avoiding overfitting. This indicates that the model is robust and can be reliably applied to future analyses of similar data.

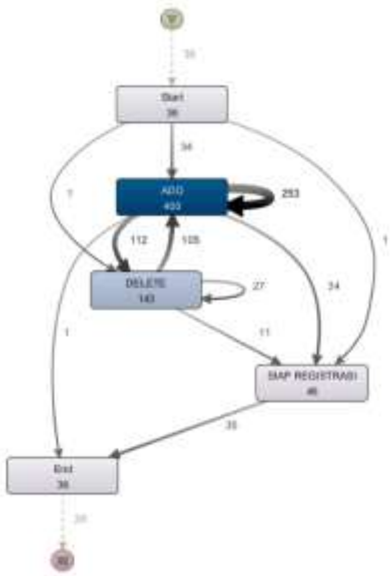

Figure 5. Second Process Model of Student Academic Registration

Evaluation

The results of conformance checking in the mining and analysis stage across the three main metrics—Fitness, Precision, and Generalization—indicate that Precision and Fitness values were not optimal. This suggests that certain activities or transitions between activities were not fully captured by the discovered process model, even though the main activities such as "ADD", "DELETE", "SIAP REGISTRASI", "Start", and "END" were aligned with the event log. The limitation is likely due to incomplete or unrecognized transitions that the discovery algorithm failed to model according to the SOP. To verify which transitions were excluded, further analysis of both the log and the process model is required. Consequently, an evaluation stage focusing on process auditing was conducted to address the third research objective: "to identify potential anomalies in the student registration process."

Visualization of the initial model (Figure 4) revealed an inconsistency in the number of cases at the "Start" point: 36 cases were recorded, yet only 34 proceeded directly to the "ADD" activity. This indicates that two cases deviated from the standard flow, raising suspicion of anomalies. To investigate these cases, a process audit involving business experts was performed to determine whether the deviations represented genuine anomalies or acceptable variations.

Given these initial anomaly indications, a second round of process mining was conducted to construct a more comprehensive model. The threshold was lowered to 50% to reveal both

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

major and minor transitions previously hidden in the visualization. The refined model (Figure 5) revealed three additional transitions:

a. A path from "Start" directly to "DELETE" (1 case),
b. A path from "Start" directly to "SIAP REGISTRASI" without "ADD" (1 case), and
c. A path from "ADD" directly to "END" (1 case).

These findings indicate the presence of anomalies in the student registration process. A detailed analysis of the dataset, supported by insights from business experts, was then performed to classify the anomalies, explain their mechanisms, and propose preventive measures to ensure such deviations do not recur in future registration cycles.

	Activity	Date	Time
1	Start	05.01.2025	07:58:00
2	DELETE	05.01.2025	07:58:00
3	ADD	05.01.2025	07:58:00
4	DELETE	05.01.2025	08:00:00
5	ADD	05.01.2025	08:01:00
6	ADD	05.01.2025	08:01:00
7	ADD	05.01.2025	08:02:00
8	ADD	05.01.2025	08:02:00
9	ADD	05.01.2025	08:03:00

Figure 6. Event Log Data for the First Case

The first anomaly was identified in cases where the process started directly with the "DELETE" activity. This indicates that a student attempted to remove a course at the very beginning of the registration process. According to the SOP, however, the "DELETE" activity can only occur after the student has performed an "ADD" activity, making this sequence formally invalid.

To examine this anomaly, the dataset was filtered to isolate the case with the transition "Start" → "DELETE". The filtered log (Figure 6) shows that the first recorded activity is indeed "DELETE", immediately following "Start". However, closer inspection reveals that the "ADD" activity appears as the third entry, sharing the same timestamp (Date: 05.01.2025, Time: 07:58:00) as both "Start" and "DELETE".

This suggests that the anomaly does not originate from user behavior, but from the system's data logging mechanism. Verification with the business expert confirmed that such a sequence can occur due to logging errors in the SIRAMA application. Specifically, the expected order ("ADD" before "DELETE") may have been disrupted by system-level issues such as delayed synchronization, slow server response, or batch logging processes that lack atomicity. As a result, "DELETE" was recorded before "ADD", even though in reality the user's actions followed the SOP. Such anomalies highlight the importance of considering system-related factors—such as logging reliability and network latency—when interpreting process mining results.

	Activity	Date	Time
1	Start	05.01.2025	07:58:00
2	ADD	05.01.2025	07:58:00
3	DELETE	05.01.2025	07:58:00
4	DELETE	05.01.2025	08:00:00
5	ADD	05.01.2025	08:01:00
6	ADD	05.01.2025	08:01:00
7	ADD	05.01.2025	08:02:00
8	ADD	05.01.2025	08:02:00
9	ADD	05.01.2025	08:03:00

Figure 7. Event Log of the First Case after Correction

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

When the sequence "Start" → "DELETE" is corrected according to the adjusted ordering shown in Figure 7, the anomaly disappears. The resulting sequence aligns with the SOP for the academic registration process, indicating that the observed irregularity stemmed from a logging error rather than an actual deviation in student behavior.

This finding implies that such anomalies can be prevented by ensuring that the system adheres to the principles of transactional integrity, specifically the ACID properties (Atomicity, Consistency, Isolation, and Durability). Among these, atomicity plays a crucial role. Atomicity ensures that a series of operations within a transaction is executed completely or not at all. In other words, partial or out-of-order execution cannot occur, even in the presence of system delays or failures.

By enforcing atomic transactions in the logging mechanism, the SIRAMA application would guarantee that activities such as "ADD" and "DELETE" are recorded in their correct order. This would maintain data consistency and prevent the appearance of spurious anomalies caused by incomplete or delayed log entries. Consequently, improving atomicity in system logging is a critical step toward enhancing the reliability of process mining analyses and ensuring that models accurately reflect real-world student behavior.

	Activity	Date	Time
1	Start	05.01.2025	00:02:00
2	SIAP REGISTRASI	05.01.2025	00:02:00
3	ADD	05.01.2025	08:05:00
4	ADD	05.01.2025	08:06:00
5	ADD	05.01.2025	08:08:00
6	ADD	05.01.2025	08:08:00
7	ADD	05.01.2025	08:09:00
8	ADD	05.01.2025	08:10:00
9	DELETE	05.01.2025	08:13:00
10	ADD	05.01.2025	08:16:00
11	ADD	05.01.2025	08:22:00
12	DELETE	05.01.2025	08:22:00
13	ADD	05.01.2025	08:23:00
14	DELETE	05.01.2025	08:25:00
15	SIAP REGISTRASI	05.01.2025	09:02:00
16	ADD	05.01.2025	10:17:00
17	End	05.01.2025	10:17:00

Figure 8. Event Log Data for the Second and Third Cases

The second and third cases, which both belong to the same user, reveal distinct deviations from the expected academic registration sequence. In the second case, the transition "Start" → "SIAP REGISTRASI" indicates that the student confirmed readiness for registration without first performing "ADD". In the third case, the path "ADD" → "END" shows that the student added a course and immediately completed the process without executing "SIAP REGISTRASI".

To investigate these anomalies, the suspected traces were filtered for closer inspection. The filtered log, shown in Figure 8, records 17 activities in chronological order. A detailed analysis confirmed that these anomalies did not originate from user actions but were instead caused by a timestamp overflow error. This finding was validated by consultation with the business expert. A timestamp overflow occurs when the system processes invalid or non-standard time values, such as 24:01:00, which exceeds the valid 24-hour format (23:59:59). Ideally, the system should interpret 24:01:00 as 00:01:00 of the following day, but not all systems can perform such adjustments automatically.

To correct this issue, the timeline of the anomalous activity was recalibrated. Specifically, the second occurrence of "SIAP REGISTRASI", originally recorded as Date: 05.01.2025, Time: 00:02:00, was adjusted to Date: 06.01.2025, Time: 00:02:00. After this correction, the sequence of student activities aligned with the SOP, as illustrated in Figure 9.

This case highlights the importance of robust timestamp handling mechanisms to avoid misinterpretations that may lead to false anomaly detection in process mining analyses.

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

	Activity	Date	Time
1	Start	05.01.2025	00:02:00
2	ADD	05.01.2025	08:05:00
3	ADD	05.01.2025	08:06:00
4	ADD	05.01.2025	08:08:00
5	ADD	05.01.2025	08:08:00
6	ADD	05.01.2025	08:09:00
7	ADD	05.01.2025	08:10:00
8	DELETE	05.01.2025	08:13:00
9	ADD	05.01.2025	08:16:00
10	ADD	05.01.2025	08:22:00
11	DELETE	05.01.2025	08:22:00
12	ADD	05.01.2025	08:23:00
13	DELETE	05.01.2025	08:25:00
14	SIAP REGISTRASI	05.01.2025	09:02:00
15	ADD	05.01.2025	10:17:00
16	SIAP REGISTRASI	06.01.2025	00:02:00
17	End	06.01.2025	00:02:00

Figure 9. Event Log Data for the Second and Third Cases after Adjustment

The adjustment of anomalous traces confirmed that the previously detected deviations were primarily caused by system-level recording errors (e.g., timestamp overflow and non-atomic log storage), rather than by user behavior. Once corrected, the activity sequences aligned with the Standard Operating Procedure (SOP) of academic registration, eliminating false anomalies. Preventive measures for the second and third cases—namely the transitions "Start" → "SIAP REGISTRASI" and "ADD" → "END"—include stricter input validation (ensuring timestamps do not exceed 23:59:59), automated correction scripts to reassign invalid times (e.g., 24:01:00 to 00:01:00 on the next day), and the inclusion of seconds and time zones in timestamp attributes to ensure transparency and auditability.

The conformance checking results further highlight discrepancies between the discovered process model and the event log, particularly in Fitness (0.846) and Precision (0.69531), while Generalization remained high (0.99992). This indicates that although core activities ("ADD", "DELETE", "SIAP REGISTRASI", "Start", and "END") were consistent, several transitions were either incompletely captured or incorrectly recorded. Lowering the mining threshold revealed three minor but significant paths ("Start" → "DELETE", "Start" → "SIAP REGISTRASI", "ADD" → "END"), which initially suggested anomalies. However, in-depth analysis and domain expert validation confirmed that these were technical errors rather than genuine process deviations.

This study underscores the importance of integrating technical validation and domain expertise in process mining projects. Beyond anomaly detection, the evaluation stage transforms findings into actionable improvements, ensuring both methodological rigor and practical relevance for future applications.

CONCLUSION

This study successfully applied a process mining approach to reconstruct, evaluate, and audit the academic registration process of students based on event log data extracted from the SIRAMA application. The key findings can be summarized as follows:

1. Process Model Reconstruction.
   The discovered process model was able to represent the core flow of academic registration, including the main activities "ADD", "DELETE", "SIAP REGISTRASI", "Start", and "END". Nevertheless, several minor paths were also detected, which deviated from the prescribed Standard Operating Procedure (SOP).
2. Model Evaluation through Conformance Checking.

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

The evaluation of the process model using three standard metrics—Fitness (0.846), Precision (0.695), and Generalization (0.999)—revealed that although the model captured the overall registration process, discrepancies remained. Specifically, certain transitions were either not fully captured or inaccurately represented in the initial model, reducing its precision.

3. Anomaly Identification and Diagnosis.
   Three anomalous transitions were identified: (i) direct transition from "Start" → "DELETE", (ii) transition from "Start" → "SIAP REGISTRASI" without an "ADD" activity, and (iii) transition from "ADD" → "END" without "SIAP REGISTRASI". Subsequent auditing and validation revealed that these anomalies were not caused by user behavior but stemmed from system-level issues, namely timestamp recording errors (overflow or synchronization delay) and non-atomic transaction logging.
4. Role of Business Experts.
   Collaboration with business experts proved essential for diagnosis and validation. Their domain knowledge ensured that deviations classified as anomalies were correctly distinguished between genuine process violations and acceptable operational variations.

In conclusion, the study demonstrates that process mining not only provides an effective means for visualizing and evaluating academic registration processes but also serves as a diagnostic tool for detecting both technical and operational anomalies.

REFERENCES

[1] P. Vu, H. Son, and N. V. Hung, "Develop a data-driven approach under the integration of 4D visualization and process mining to simulate , diagnose and predict real-world construction execution," pp. 6147–6169, 2024.

[2] J. Kurniasih, Z. A. Abas, S. A. Asmai, and A. B. Wibowo, "LogForum PROCESS MINING AND SIMULATION MODELING IN PRODUCTION," *LogForum*, vol. 20, no. 4, pp. 429–453, 2024.

[3] B. F. A. Hompes, *Analyzing Process Performance Bart Hompes*, vol. 1. Proefschrift: Eindhoven University of Technology, 2025. [Online]. Available: http://dx.doi.org/978-90-386-6256-5

[4] Y. G. and S. Q. W. Guan, J. Cao, H. Zhao, "Survey and Benchmark of Anomaly Detection in Business Processes," *IEEE Trans. Knowl. Data Eng.*, vol. 37, no. 1, pp. 493–512, 2025, [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10738505&isnumber=10786487

[5] A. P. Kurniati and G. A. A. Wisudiawan, "Analisis Kesiapan Penerapan Process Mining pada Sistem Manajemen Pembelajaran Universitas Telkom," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 6, p. 1227, 2021, doi: 10.25126/jtiik.2021863875.

[6] M. L. Van Eck, X. Lu, S. J. J. Leemans, and W. M. P. Van Der Aalst, "PM 2 : a Process Mining Project Methodology".

[7] A. Furquim, E. D. F. Rocha, and E. A. Portela, "ScienceDirect Application of the PM² methodology in the analysis of assembly processes," *Procedia CIRP*, vol. 132, pp. 159–164, 2025, doi: 10.1016/j.procir.2025.01.027.

[8] R. Hobeck, L. Pufahl, and I. Weber, *Process Mining on Curriculum-Based Study Data : A Case Study at a German*. Springer Nature Switzerland, 2023. doi: 10.1007/978-3-031-27815-0.

[9] J. Brock, S. Von Enzberg, K. Arno, and P. Roman, "ScienceDirect Process Mining Data Canvas : A method to identify data and process knowledge for data collection and preparation in process mining projects," vol. 119, pp. 602–607, 2023, doi: 10.1016/j.procir.2023.03.114.

[10] Vavpotič D; Bala S; Mendling J; Hovelja T, "Software Process Evaluation from User

Page : 509-524
DOI : https://doi.org/10.54209/jurnalinstall.v17i08.422

Perceptions and LogData," *J. Softw. Evol. Process*, vol. 34, no. 4, 2022.
[11] G. A. A. W. and G. P. K. A. P. Kurniati, "Potentials of Clinical Pathway Analysis Using Process Mining on the Indonesia National Health Insurance Data Samples: an Exploratory Data Analysis," in *International Conference on Data Science and Its Applications (ICoDSA)*, Bandung, 2022, pp. 294–299.
[12] L. Uden and D. L. Eds, *Learning Technology for Education Challenges*. 2024.
[13] and P. A. B. Tim May, Malcolm Williams, Richard Wiggins, "Pembelajaran Menggunakan Learning Management Systemberbasis Moodle pada Masa Pandemi Covid-19.," *Indones. J. Teach. Educ.*, vol. 2 (1), no. 1996, p. 6, 2021.
[14] F. Kurniati, Angelina Prima and Ary Wisudiawan, Gede Agung and Kusuma, Guntur Prabawa and Saadah, Siti and Osman, Nurul Aida and Zulhelmy and Wan, Wan Nur Syahidah Binti and Hafidz, "Patient Clustering to Improve Process Mining for Disease Trajectory Analysis Using Indonesia Health Insurance Dataset," in *7th International Conference on Artificial Intelligence and Big Data (ICAIBD)*, Chengdu, 2024, pp. 88–93.
[15] W. M. P. Van Der Aalst *et al.*, *Process Mining Handbook*. Barcelona, Spain: Springer Cham, 2022. doi: 10.1007/978-3-031-08848-3.
[16] E. Alves and P. Santos, "Framework for Process Mining in Cargo Transportation Safety – PM 2 RCS Framework para Mineração de Processos na Segurança do," *Gestão & Produção*, vol. 32, no. e1524, pp. 1–40, 2025, [Online]. Available: https://doi.org/10.1590/1806-9649-2025v32e1524