Process Mining to Improve the Development of the Course Syllabus

Gede Agung Ary Wisudiawan
(School of Computing, Telkom University, Bandung, Indonesia,
degunk@telkomuniversity.ac.id)

ABSTRACT

This paper focused on how to improve the development of course syllabus using process mining. Course syllabus developers need more input when making improvements. In this paper, the improvement of course syllabus development is based on enriching the input process of course syllabus development. The input based on the facts from student activities when student use LMS. LMS has an event log that keeps logs of student activities while using the LMS. The event log will be mined using a process mining, the algorithm used in this process mining is heuristic algorithm. Process mining is proven to be able to enrich the input in the development of course syllabus. The first finding is that there are two topics that are carried out simultaneously. The topic is the topic 7th and 8th as well as the 12th and 13th. The second finding is that there are jumping week that can be seen at the 2nd topic, directly to the 6th topic and at the 9th topic, continued to the 11th topic. These findings can be used to enrich input during development of the course syllabus. Regarding the evaluation of the model process from student learning activities, database modeling for one semester at the LMS that was formed is as follows: the fitness value is 99%, the generalization value is 99%, and the precision value is 45%.

Keywords: **course syllabus, heuristic miner, process mining, learning management system, process modelling.**

INTRODUCTION

Each course must have a course syllabus during one semester. The course syllabus is a schedule of tasks that lecturers and students will complete during class lectures. For preparing the new semester, lecturer must develop the course syllabus that they lead in (Julianto et al., 2022). The name of the study program, the name of the subject code, credits, and the lecturer's name must at the very least be included in the course syllabus identity (Ilmiani et al., 2020). Course syllabus generally contains learning outcomes, learning indicators, assessment instrument, learning methods, times, assessment rating, and reference (Nurdin, 2019). Learning outcomes are developed from internalization of attitudes, competences, knowledge, skills, and cumulative work experience results that every student must achieved. Learning indicators consist of learning content and competency of behavior that students must achieved.

The use of LMS (Learning Management System) in universities has become common (Deepak, 2017). Some universities and colleges use LMS to support the learning process. The use of LMS is very helpful for students in following the learning process. One of the conveniences obtained by students is that learning materials can be accessed without time and place restrictions. One of the open-source applications commonly used as an LMS is Moodle (Febliza & Okatariani, 2020). Moodle is an acronym for Modular Object-Oriented Dynamic Learning Environment. Moodle is easy to obtain and develop which makes Moodle widely used throughout the world (Zabolotniaia et al., 2020). Moodle can contain multiple courses. Each course can contain users like student, teacher, and administrator roles (Wisudiawan & Kurniati, 2022). Moodle has several learning materials such as files, quizzes, assignments, forums, glossaries, interactive videos. The lecturer must have plan what learning material should be available to the student. So, the student can achieve the learning outcomes. Every learning material can find in the course syllabus that have been prepare at the beginning of the semester. Course syllabus have been developed by the lecturer. Each learning material can be accessed by students that enrolled in their course. Moodle has log data that records every user activity (Rotelli & Monreale, 2021). When students access learning material, Moodle will record it into the log. Log data can be accessed by teacher in Logs report feature at course administration. Log data can be filtered by course, participant, day, and activity. Log data contain user activity report like time, user full name, affected user, event context, component, event name, description, origin, and Ip address. The example value for Moodle log file can see at table 1. Column time inform when the user does the activity. User full name contain user full name that registered in Moodle. Affected user contain username who affected by the activity that user does. Event context contains type of module followed by its name. Components contain module type. Event name gives a brief description of the action taken. The description includes a description of the event together with the relevant user and module IDs. The origin contains the sources like CLI, Restore, Web, Web service. Ip address contain location that user from. In table 1 was describe the example for each log columns. Before process mining will conduct, log data from Moodle that contain 10 columns must eliminate some columns. In this study we use column time, user full name, event context and description to do process mining.

Table 1. Example value for moodle log file

Column Name	Example
Time	13 September 2022, 7:39 AM
User full name	GEDE AGUNG ARY WISUDIAWAN
Affected user	GEDE AGUNG ARY WISUDIAWAN
Event context	Quiz: Quiz Review 07
Component	Quiz
Event name	Quiz report viewed
Description	The user with id '6' viewed the report 'overview' for the quiz with course module id '1905605'.
Origin	Web
IP address	103.233.100.228

Process mining is a field that analyzes processes using scientific methods (Van Der Aalst et al., 2012). The log data is the input for the mining process. Users' actions within an application or information system generate log data. The log data is analyzed using process mining to identify patterns in user activity and create models based on that behavior. Process mining examines log data to find trends and patterns in process execution. Process mining can provide answers to inquiries about business processes, such as how and why those processes were implemented in the past, what may occur in the future, when and why those implementations differ from procedures, how to control those implementations, and how to redesign those implementations to perform better. The three main types of analysis in process mining are process discovery, conformance checking, and enhancement. Using process mining methods, process discovery processes log data to create process model visualizations. A specific notation can be used to display the process model like using BPMN, PetriNet, state transition diagram, fuzzy net, and Heuristic Net. The outcomes of the analysis in process mining can be used to track and boost process quality, verify compliance to procedures, find activities that disrupt the process, and identify potential issues that might occur during execution. Process mining can be used in the field of education to examine the learning process, particularly the computer-based learning process controlled by a learning management system (LMS).

Other study that related to this study is from Angelina, that demonstrate the data analytics on Telkom University LMS depending on various user roles (Kurniati & Wisudiawan, 2021). From that study we know, the Telkom University LMS can be used to analyze learning process performance. The algorithm that used is heuristic miner

and the tools that was used is ProM and Disco from Fluxicon. Other study that related is from Wisudiawan. The result from that study can produce lecture and student process model. The process model can help the lecturer to improve the course plan. The course plan contains sequence of learning material that available in the LMS. The algorithm that uses in that study are fuzzy algorithm, and heuristic miner. The tool that uses in that study is Disco from Fluxicon to develop the process model.

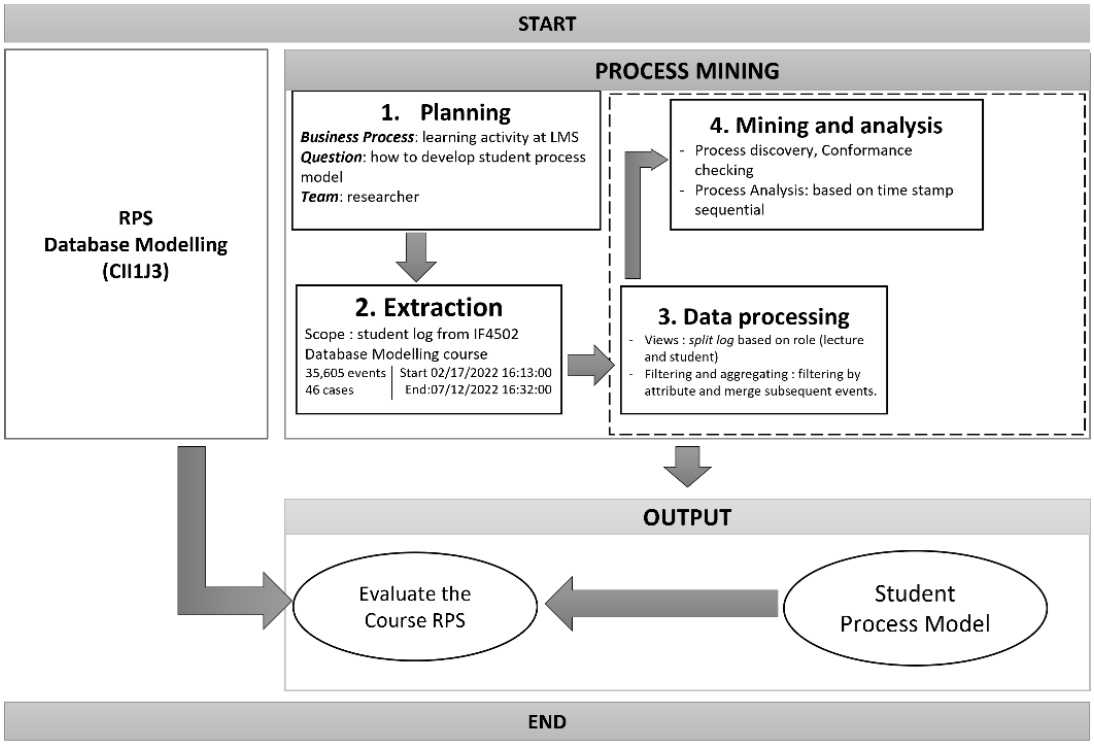
In this paper, the application is Moodle based LMS. Log data can be thought of as describing actual learning activities because log data is created automatically in the LMS based on user activities in real-time. The log data from Moodle based LMS can be processed by process mining to produce a process. The process model describes the learning path of the learning process carried out by students when they are accessing available learning media. Learning path contain sequence activity of student learning process during they are access the LMS based on Moodle. From the learning path the lecturer knows the behavior of the student activities during they access the course material in LMS (Kadoic & Oreski, 2018). Student behavior can describe how the student learn their lesson each course topics. The lecturer can make course syllabus based on the student behavior.

The goal of this study is to make suggestions for improving the development of course syllabus using process mining. In the other side developing the course syllabus usually based on experience from the lecturer (Julianto et al., 2022). Therefore, lecturers need more input when making improvements based on the facts from student activities. In this paper, the student activities that are the focus are student learning process when student use LMS. The improvement process is how to enrich the input of course syllabus development. Enrich the input based on the fact of student behavior when they are going through to the learning process that supported with LMS based on Moodle.

From the background in this research the first research question is how to use process mining method to analyze learning process that provided by student activities in the LMS. Research question number two is what the finding of student process model in database modelling courses subject based on Moodle LMS. Research question number three is how to enrich the input of course syllabus development using student behavior in LMS based on moodle LMS.

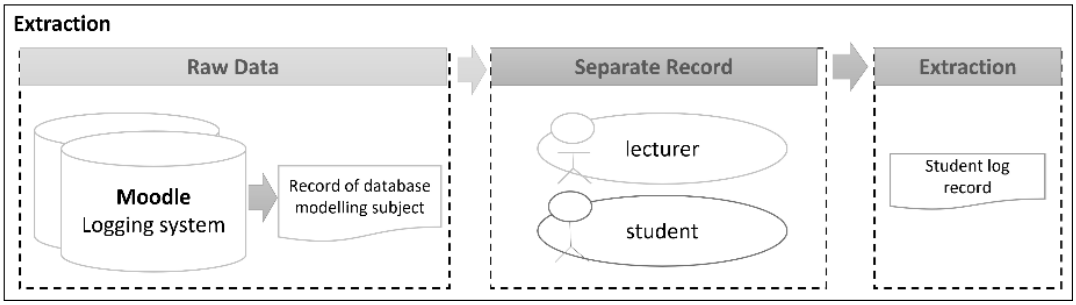
METHODS

This research consists of 5 phases, which are: (1) Planning, (2) Extraction, (3) Data processing, (4) Mining and analysis, and (5) Evaluation. The output of this study is evaluating the courses course syllabus and develop student process model. Picture 1 describes how this research did. The completely describe will explain with next sub section.



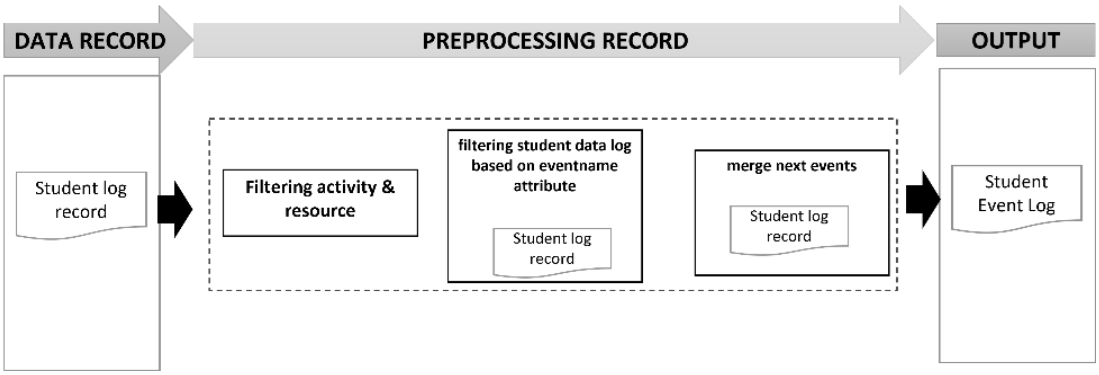
Picture 1. Research methodology

Planning, which determines the scope of the mining process, research questions, and the composition of the research team, is the first step in the mining process. The focus of this study is on the learning activities that students taking a Relational Database subject course at Telkom University will engage in during the odd-numbered semester of 2021–2022. The goal of this study is to construct a model of the student learning process. To comprehend business processes in the Telkom University learning management system, a research team of authors, lecturers, and Telkom University learning management system administrator is working with.



Picture 2. Extraction process from Moodle

Picture 2 depicts the student extraction phases from the Moodle logging system during the extraction phase. The research data for this study came from event logs from Telkom University's learning management system, which are stored in the moodle logging system. This event log records all user activity for the odd semester database modeling course during the academic year 2021–2022, including those of lecturers, student's activities. The log data extracted from Logs feature in LMS at course administration. No need filter to extract the log data from the LMS. In this study we use four columns that is time, user full name, event context and description to do process mining. "Time" is mapping to timestamp (pattern: 'dd/MM/yy, HH:mm'), "user full name" is mapping to case ID, "event context" is mapping to activity, and "description" is mapping to the other. Filter will use at disco from Fluxicon application. From the disco application we use filter by Case ID at filter feature. Prior to preprocessing, the event log is divided into student and lecturer categories, with 3.385 records of lecturer activities and 50.576 records of student activities, respectively



Picture 3. Data Preprocessing

Data preprocessing is the third phase of mining process according to picture 3. Data preprocessing is done to ensure that the existing event log is ready to be processed

by the heuristic miner algorithm. The processes carried out on data processing are setting column header to meet process mining requirement, filtering student log record based on eventname attribute, merging subsequent events in student log data. Setting column header to meet process mining requirement is the process to selecting the student log record column that used to process mining. There are three column that need to meet the process mining requirement there are case ID, Activity, time stamp. From the student log record then select the "user full name" to be case ID, "event context" to be activity and "time" to be time stamp.

After getting the event log of student meet the process mining requirement, the next step is the filtering process based on eventname attributes. This step is mandatory because this is Moodle behavior when saving the user activity. Moodle always record the user activity every time user accesses the activities. For example, when user can take the quiz, Moodle will record the user activity with several eventname. Eventname that will use to describe the user activity when access the quiz like course module viewed, quiz attempt started, quiz attempt submitted, quiz attempt summary viewed, and quiz attempt viewed.

For filtering the eventname based on Wisudiawan paper to select the eventname from the Moodle log. This filtering step do for all activity from student log record. The activity from the student log record are quiz, forum, assignment, and file. The filtering process is assisted by the disco application from Fluxicon. The relationship between activity and the chosen eventname showed with table 2.

Table 2. Activity and chooses eventname

No	Activity	Eventname
1	Quiz	Quiz attempt submitted
2	Forum	Some content has been posted
3	Assignment	A submission has been submitted
4	File	Course module viewed

The next step for preprocessing the student record data is merge the subsequence. This merge is important to make sure there are no redundancy record when conduct the process mining step. The merging subsequence step is using prom tools application and merge subsequent events (AAABB->AB) plugin. After all the preprocessing steps are finished, the student log record processing to process mining step. From the preprocessing step, number of records decreased from 53,961 records to 1,509 records. The student log metadata showed in table 3.

Table 1. Metadata of student event log

Column name	Details
Cases	46
Events	1509
Event classes	38
Activity	4 (Quiz, File, Forum, and Assignment)
Start	02/17/2022 22:51:00
End	07/11/2022 15:12:00
Average usage	17.8 weeks
Min usage	56 days 22 hours.
Max usage	139 days 1 hours

The mining process uses a heuristic miner algorithm assisted by the ProM tools application. The stages of the mining and analysis process will be explained in the Results and discussion section FINDING AND DISCUSSION.

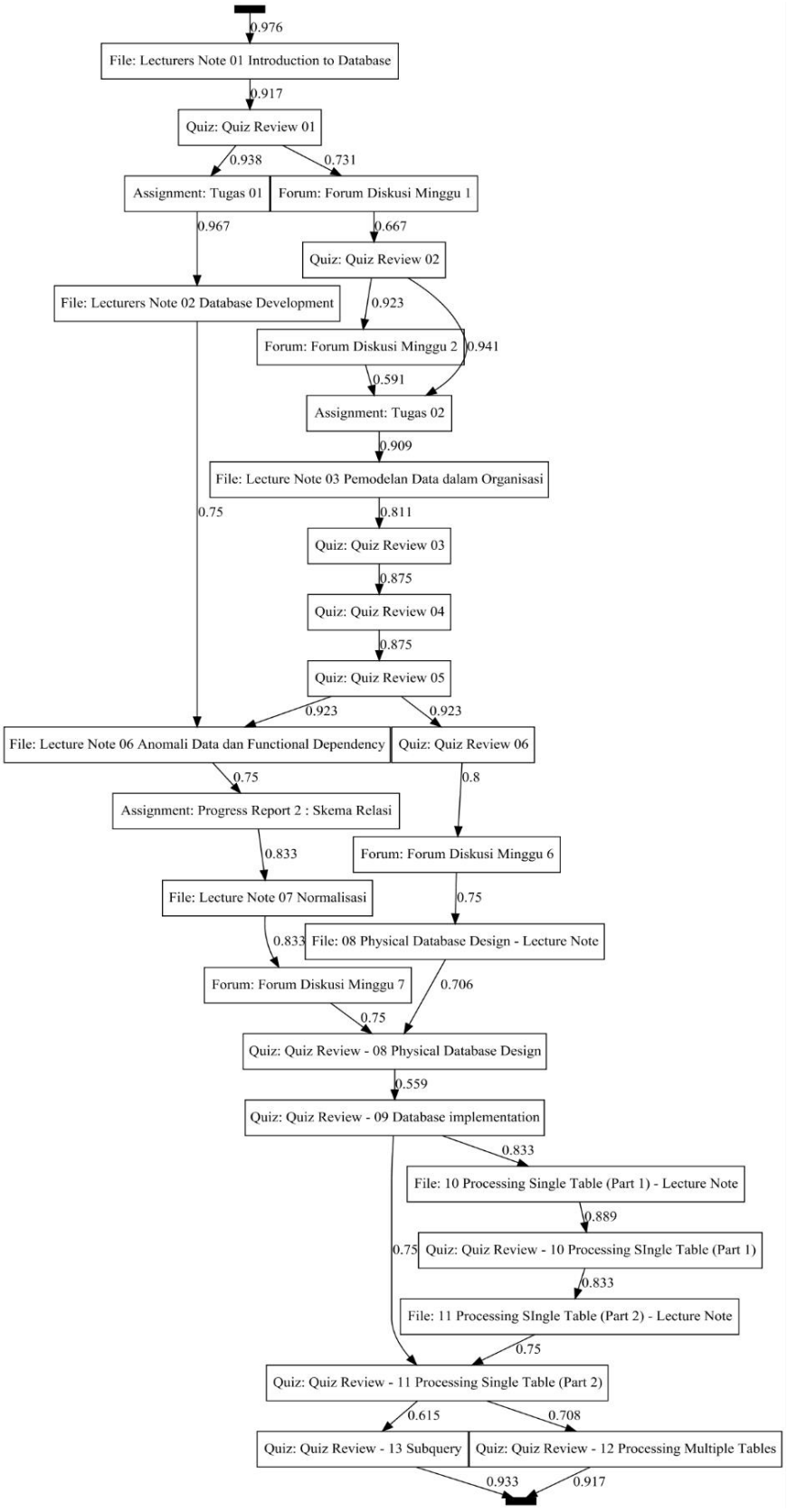
This study uses student event logs obtained from the learning management system owned by Telkom University. For the case study, the selected log data came from the database modeling course. In the mining and analysis phase, there is a discovery process that uses a heuristic miner algorithm assisted by the ProM tools application that produces a process model. The process model will be converted into a petri net form so that the conformance checking process can be carried out which is also assisted by the ProM tools application.

FINDINGS AND DISCUSSION

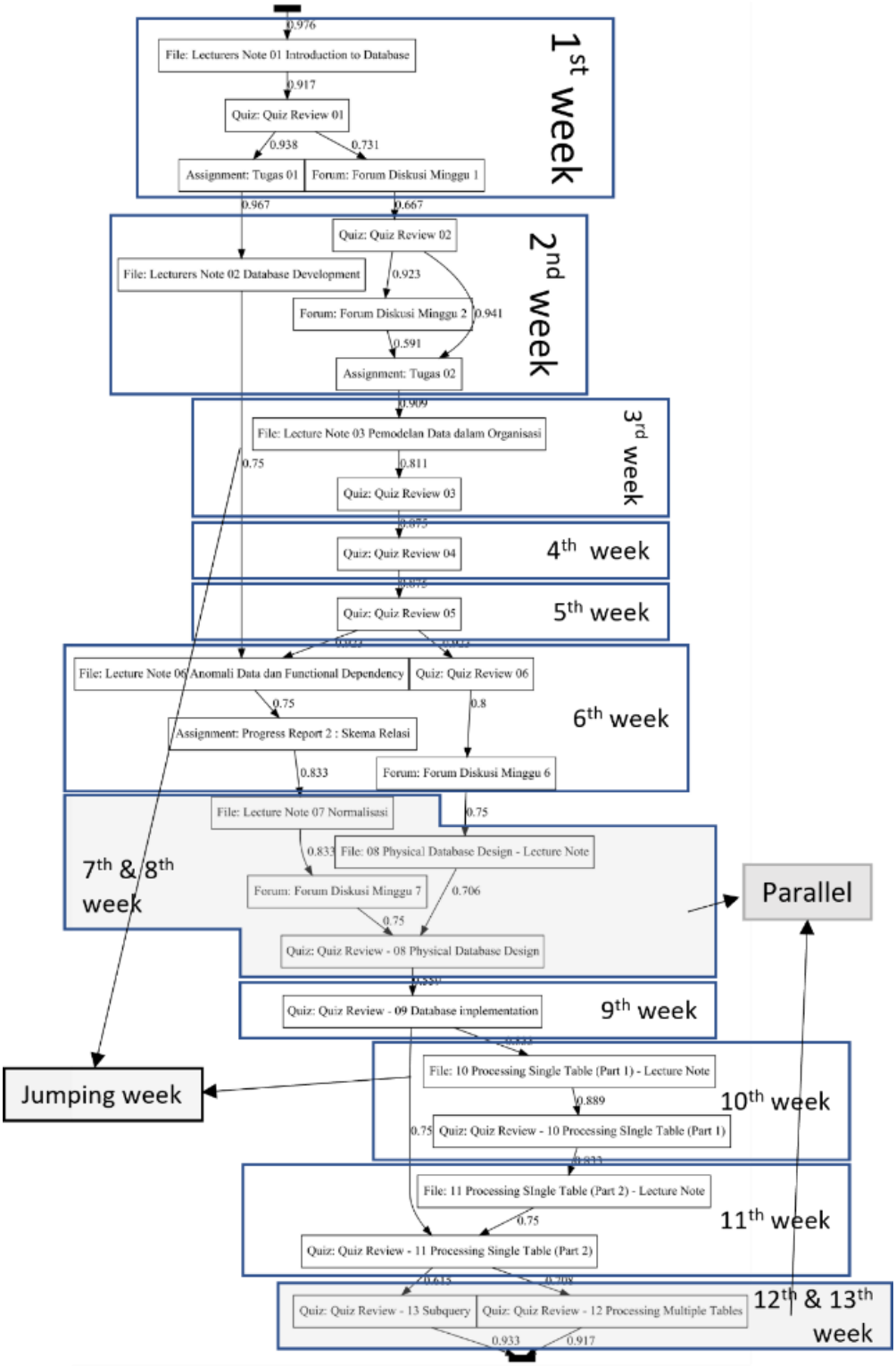
$$a \Rightarrow_w b = \left(\frac{|a>_w b| - |b>_w a|}{|a>_w b| + |b>_w a| + 1}\right) \quad (1)$$

From the equation (1) a, b is activity that a followed by b and w is weight (frequency) of a followed by b. Based on that equation, the dependency graph is developed. The dependency graph is a visual representation that contains the relationship between activities. The relationship is formed based on the dependency matrix, the length-one loop dependency, and the length-two loop dependency. The discovery process uses a heuristic miner algorithm and is assisted by ProM tools. This study uses three thresholds there are dependency threshold, conditions threshold and frequency threshold. The values from each threshold are for dependency threshold =

0.9, conditions threshold = 0.5, and frequency threshold = 0.1. Which maps the process model obtained from process discovery based on the student's event log can be seen in picture 4. The process model obtained from process discovery based on student event logs can be seen in picture 5. The discovery process uses a heuristic miner algorithm and is assisted by ProM tools. Fitness of the model is 55.6% and the precision is 54.9%. The course syllabus contains the sequence of topics that are learned by students. The sequence is separated into multiple sessions, with each meeting typically lasting one week. There are also topics that serve as prerequisites for others. Similar to a subject can be studied if the preceding subject has been studied. The process mining forms a student learning model process for one semester in the database modeling course based on the learning activity log on the LMS.



Picture 4. Student Process Model



Picture 5. Result Mapping Each Topic with Course syllabus to The Student Process Model.

In this study, the mapping process of each topic in the course syllabus with a model process from students in the LMS for one semester. This mapping process

involves the lecturer concerned who teaches database modeling courses. The results of the process of mapping each topic in the course syllabus to the student process model can be seen in picture 5. These results are valid based on interviews with the lecturer concerned who teaches database modeling courses. This database modeling course consists of 13 topics. The 13 topics were given to students for 13 weeks of meetings. According to result from mapping each topic in the course syllabus to the student process model (see picture 5), there are non-sequential student activities (both parallel and jumping week). Parallel means students access different topics at the same time. Jumping week means that students access the material by jumping from the order that has been designed in the course syllabus. The following are the results of the student activity model methodology: In the seventh and eighth topics, students' access both concurrently. Some students after studying the lecture notes on the topic of the 3rd meeting, then immediately studied the lecture notes on the topic of the 6th meeting. Regarding the ninth meeting's quiz, several students immediately took the eleventh meeting's quiz. On the topic quiz, the 12th and 13th graders simultaneously accessed the two quizzes on that topic.

CONCLUSION

In this paper, the use of the mining process based on heuristic algorithm is proven to be able to enrich the input in the development of course syllabus. This can be seen in the findings section in this paper. The first finding is that there are two lecture topics that are carried out simultaneously. The topic is the topic of the 7th and 8th weeks as well as the 12th and 13th weeks. This can happen because at the 7th and 8th meeting topics there is a material connection, as well as at the 12th and 13th topic meetings. The second finding is that there is a sequence learning model from students where the order of topics that should be sequential every week but instead becomes a week jump. The findings of this jumping week can be seen at the 2nd meeting, directly to the 6th meeting and at the 9th meeting, continued to the 11th meeting. These two findings can be used to enrich input during the development of the course syllabus or the next database modeling course. So that at the time of making the course syllabus there were inputs based on the student's learning activities in the previous semester, especially topics 7th and 8th that could be discussed simultaneously, as well as at the 12th and 13th meetings. For the 6th meeting topics, the material on the topics of the

2nd meeting could be discussed. Likewise, the topic of the 11th meeting can repeat the explanation of the 9th meeting topic. Regarding the evaluation of the model process from student learning activities, database modeling for one semester at the LMS that was formed is as follows: the fitness value is 99%, the generalization value is 99%, and the precision value is 45%.

REFERENCES

Deepak, K. C. (2017). Evaluation of Moodle Features at Kajaani University of Applied Sciences-Case Study. *Procedia Computer Science*, *116*, 121–128. https://doi.org/10.1016/j.procs.2017.10.021

Febliza, A., & Okatariani, O. (2020). The Development of Online Learning Media by Using Moodle for General Chemistry Subject. *Journal of Educational Science and Technology (EST)*, *6*(1), 40–47. https://doi.org/10.26858/est.v6i1.12339

Ilmiani, A. M., Hamidah, H., Wahdah, N., & Mubarak, M. R. (2020). The Implementation of Semester Learning Plan (RPS) on Maharah Istima' Course by Using Cartoon Story Maker. *LISANIA: Journal of Arabic Education and Literature*, *4*(1), 1–22. https://doi.org/10.18326/lisania.v4i1.1-22

Julianto, P., Samin, S., & Faizin, F. (2022). EVALUASI PENERAPAN RENCANA PEMBELAJARAN SEMESTER (RPS) PADA PROGRAM STUDI MANAJEMEN IAIN KERINCI. *EQUILIBRIUM: JURNAL PENELITIAN PENDIDIKAN DAN EKONOMI*, *19*(02), 274–280.

Kadoic, N., & Oreski, D. (2018). Analysis of student behavior and success based on logs in Moodle. *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2018 - Proceedings*, 654–659. https://doi.org/10.23919/MIPRO.2018.8400123

Kurniati, A. P., & Wisudiawan, G. A. A. (2021). Analisis Kesiapan Penerapan Process Mining pada Sistem Manajemen Pembelajaran Universitas Telkom. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *8*(6), 1227. https://doi.org/10.25126/jtiik.2021863875

Nurdin, S. (2019). Pengembangan Kurikulum dan Rencana Pembelajaran Semester (RPS) Berbasis KKNI di Perguruan Tinggi. *Murabby: Jurnal Pendidikan Islam*, *1*(2), 140–147. https://doi.org/10.15548/mrb.v1i2.305

Rotelli, D., & Monreale, A. (2021). *Making Sense of Moodle Log Data*. *2019*, 1–14. http://arxiv.org/abs/2106.11071

Van Der Aalst, W., Adriansyah, A., De Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., Bose, J. C., Van Den Brand, P., Brandtjen, R., Buijs, J., Burattin, A., Carmona, J., Castellanos, M., Claes, J., Cook, J., Costantini, N., Curbera, F., Damiani, E., De Leoni, M., … Wynn, M. (2012). Process mining manifesto. *Lecture Notes in Business Information Processing*, *99 LNBIP*(PART 1), 169–194. https://doi.org/10.1007/978-3-642-28108-2_19

Wisudiawan, G. A. A., & Kurniati, A. P. (2022). Process Mining on Learning Activities in a Learning Management System. *International Conference on*

Advanced Communication Technology, ICACT, *2022-Febru*, 476–482. https://doi.org/10.23919/ICACT53585.2022.9728903

Zabolotniaia, M., Cheng, Z., Dorozhkin, E. M., & Lyzhin, A. I. (2020). Use of the LMS Moodle for an effective implementation of an innovative policy in higher educational institutions. *International Journal of Emerging Technologies in Learning*, *15*(13), 172–189. https://doi.org/10.3991/ijet.v15i13.14945