Implementasi Heuristic Miner Untuk Menganalisis Pola Perilaku Mahasiswa Dalam Mengerjakan Ujian Tengah Semester Pada Learning Management System

1st Fahlar Reva Fauzi
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
fahlarrevaa@students.telkomuniversity.ac.id

2nd Gede Agung Ary Wisudawan
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
degunk@telkomuniversity.ac.id

3rd Angelina Prima Kurniati
Fakultas Informatika
Universitas Telkom
Bandung, Indonesia
angelina@telkomuniversity.ac.id

***Abstrak* — *Learning Management System* merupakan sistem perangkat lunak yang membantu pembelajaran daring oleh mahasiswa dan juga dosen dalam mengelola kegiatan belajar mengajar secara mandiri di lembaga pendidikan. LMS adalah sistem manajemen pembelajaran Universitas Telkom yang dikembangkan berbasis Moodle. Penelitian ini melakukan *process mining* pada *event log* dari LMS pada mata kuliah Algoritma Pemrograman tahun ajaran 2021/2022 semester genap selama pengerjaan ujian tengah semester. Dilakukan proses *clustering* K-Means untuk mendapatkan kelompok varian dari pola perilaku mahasiswa pada *event log*. Metode *process mining* digunakan untuk mengetahui model proses pola perilaku mahasiswa pada *event log* berdasarkan varian yang ada selama pembelajaran daring. Dengan melihat pola mahasiswa pada model proses, dosen dapat mengetahui pola pengerjaan ujian mahasiswa dan kesiapan pengerjaan ujian selama perkuliahan menggunakan LMS. Penerapan algoritma *heuristic miner* menggunakan *process mining for python* (PM4Py). Setelah mendapatkan model proses, dilakukan *conformance checking* dengan hasil nilai *fitness* yang didapat memiliki nilai yang baik. Pada varian 1 dengan nilai *fitness* 0,98124, varian 2 dengan nilai 0,98753, varian 3 dengan nilai 0,98561, dan varian 4 dengan nilai 0,98814. Hal ini menunjukan bahwa proses model dapat memodelkan *event log* cukup baik.**

***Kata kunci*— learning management system (LMS), process mining, heuristic miner, conformance checking**

I. PENDAHULUAN

A. Latar Belakang

Seiring berjalannya waktu teknologi dan informasi di Indonesia berkembang dengan pesat dan menjadi topik bahasan yang menarik. Digitalisasi pun mulai merambat ke sektor - sektor pendukung pekerjaan manusia, salah satunya adalah sektor pendidikan. Pada awal tahun 2020 Indonesia resmi mengumumkan kasus pertama COVID-19 di Jakarta. Dengan adanya pengumuman tersebut, salah satu sektor yang terkena dampak buruknya adalah sektor pendidikan. Mayoritas lembaga pendidikan di Indonesia melakukan aktivitasnya secara tatap muka. Dengan adanya kewajiban untuk mengantisipasi penyebaran COVID-19 anjuran untuk tidak berkerumun merupakan hal yang perlu dilakukan pada saat itu.

Melalui surat edaran Telkom University pada tanggal 15 Maret 2020, pembelajaran tatap muka resmi dihentikan dan diganti secara daring. Pembelajaran secara daring ini dapat dilakukan melalui beberapa alat bantu pembelajaran salah satunya adalah *learning management systems* (LMS). LMS sendiri memiliki fitur yang mendukung proses pembelajaran secara daring seperti sistem manajemen pembelajaran yang memungkinkan mahasiswa untuk berkomunikasi dengan dosen, mengunduh materi pembelajaran yang dosen unggah, mengerjakan tugas yang diberikan dosen, dan juga mengerjakan kuis secara daring. Dalam penggunaannya LMS menyimpan banyak rekaman data dalam bentuk *event log* dengan jumlah besar dan akan terus bertambah seiring berjalan nya waktu. *Event log* ini berisi kejadian yang dilakukan dalam satu eksekusi proses. Salah satu kegunaan log ini dapat kita analisis sesuai kebutuhan.

B. Topik dan Batasannya.

Rekaman data berbentuk *event log* ini bisa sangat berguna karena saat pembelajaran daring informasi yang dosen peroleh akan lebih minim ketimbang pembelajaran tatap muka. Pola perilaku mahasiswa merupakan salah satu aspek penting dalam pembelajaran. Salah satu kegiatan yang umum dilakukan pada proses pembelajaran merupakan pengerjaan kuis. Dengan dilakukannya pembelajaran secara daring ini bentuk pola perilaku nya pun tidak bisa secara langsung diketahui oleh dosen. Dengan adanya *event log* ini dapat dianalisis menggunakan salah satu metode yaitu *Process mining*. Pada penelitian ini akan mencoba untuk menganalisis beberapa varian model proses pola perilaku mahasiswa dalam mengerjakan kuis secara daring dengan process mining.

Process mining merupakan suatu proses berfokus pada penggalian pengetahuan dari data yang dihasilkan dan disimpan dalam database sistem informasi. Tujuan nya adalah untuk mendapatkan informasi yang berguna dari *event log* dan digunakan sebagai panduan untuk proses bisnisnya.*Process mining* menggunakan algoritma *heuristic*

miner untuk mendapatkan suatu model proses karena algoritma *heuristic miner* dapat merepresentasikan suatu model dengan efektif yang ada di *event log* dengan *noise*, dan dapat menampilkan *main behavior* dari proses bisnis yang ada. Penerapan ini menggunakan *process mining for python* (PM4PY). Untuk melakukan penemuan proses model dan mengevaluasi model algoritma *heuristic miner* menggunakan*process mining conformance checking* (nilai *fitness*, nilai *precision*, dan nilai *generalization*) yang ada di PM4PY.

Maka dari itu terdapat rumusan masalah yang ada pada penelitian ini yaitu

1. Dapatkah metode *process mining* digunakan untuk menganalisis varian model proses pola perilaku mahasiswa tertentu selama pengerjaan kuis ujian tengah semester dalam sistem manajemen pembelajaran *moodle*?
2. Apakah varian model proses yang ada dapat dievaluasi dengan *process mining conformance checking*?

Adapun batasan pada penelitian ini seperti

1. Penelitian ini hanya dilakukan pada lingkup *learning management system* (LMS) Universitas Telkom di mata kuliah algoritma pemrograman yang melakukan pengerjaan ujian tengah semester pada LMS pada saat kuliah daring tahun ajaran 2021/2022 semester genap.
2. Penelitian ini hanya dilakukan sampai batas *process mining conformance checking*.
3. *Process Mining* hanya menggunakan bahasa python dengan beberapa *open source-library*.

C. Tujuan

Berdasarkan rumusan masalah diatas tujuan dari penelitian ini adalah

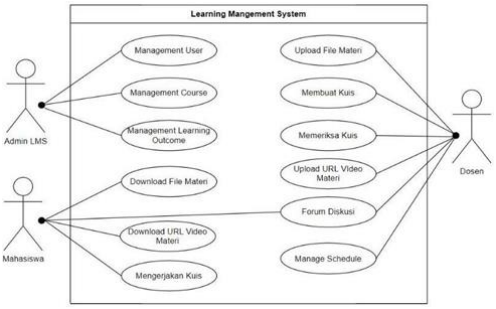
1. Mendapatkan varian pola perilaku mahasiswa saat mengerjakkan ujian tengah semester selama pembelajaran daring pada LMS yang digambarkan dengan model proses algoritma *heuristic miner.*

Melakukan implementasi *conformance checking* menggunakan algoritma *heuristic miner* pada pemodelan *event log* LMS Universitas Telkom berdasar nilai *fitness*, nilai *precision*, dan nilai *generalization*.

II. KAJIAN TEORI

A. Learning Management System

Learning management system merupakan sistem perangkat lunak pembelajaran daring yang digunakan untuk mengelola kursus atau pelatihan yang mendukung pembelajaran dengan metode *e-learning* seperti menyediakan materi pembelajaran serta membantu dalam pengelolaan proses pengajaran dan interaksi antar pengguna. Universitas Telkom menggunakan LMS *center for e-learning and open education* (CeLOE) yang merupakan sebuah platform pembelajaran daring bagi program studi reguler dan program studi pendidikan jarak jauh. Di dalam CeLOE ini setiap dosen dapat membuat konten digital untuk bahan pembelajaran mata kuliah dengan standar yang telah ditetapkan oleh Universitas Telkom yang dapat diakses oleh mahasiswa [2].



GAMBAR 1.
Use Case Diagram LMS

Seperti yang terlihat pada gambar 1 use case diagram learning management system admin bertugas untuk management user, management course, management learning outcome. Sedangkan untuk dosen bertugas untuk upload file materi, membuat kuis, memeriksa kuis, upload URL video materi, forum diskusi yang bisa dilakukan bersama mahasiswa, management schedule. Dan untuk mahasiswa bisa melakukan download file materi, forum diskusi yang bisa dilakukan bersama dosen, download URL video materi, mengerjakan kuis.

B. Event Log

Event log merupakan suatu proses dalam pencatatan *history* berupa transaksi atau *audit trail* pada *tools* sistem informasi yang digunakan. *Event log* terdiri dari kumpulan kasus, setiap kasus terdiri dari urutan kejadianyang dilakukan dalam satu eksekusi proses. *Event log* tersimpan dalam sistem informasi dan dapat dianalisa menggunakan *process mining*. Sebuah *event log* butuh informasi yang cukup seperti aktivitas yang dapat menjelaskan sebuah kejadian beserta waktu terjadinya untuk bisa dilakukan analisa *process mining*. Penelitian ini berfokus pada polaperilaku mahasiswa dalam melakukan pengerjaan ujian tengah sekmester pada LMS, dengan melakukan identifikasi pada pola perilaku mahasiswa dalam kegiatan pembelajaran yang terekam dalam *event log* LMS Universitas Telkom.

Case id	Event id	Properties				
		Timestamp	Activity	Resource	Cost	...
1	35654423	30-12-2010:11.02	register request	Pete	50	...
	35654424	31-12-2010:10.06	examine thoroughly	Sue	400	...
	35654425	05-01-2011:15.12	check ticket	Mike	100	...
	35654426	06-01-2011:11.18	decide	Sara	200	...
	35654427	07-01-2011:14.24	reject request	Pete	200	...
2	35654483	30-12-2010:11.32	register request	Mike	50	...
	35654485	30-12-2010:12.12	check ticket	Mike	100	...
	35654487	30-12-2010:14.16	examine casually	Pete	400	...
	35654488	05-01-2011:11.22	decide	Sara	200	...
	35654489	08-01-2011:12.05	pay compensation	Ellen	200	...
3	35654521	30-12-2010:14.32	register request	Pete	50	...
	35654522	30-12-2010:15.06	examine casually	Mike	400	...
	35654524	30-12-2010:16.34	check ticket	Ellen	100	...
	35654525	06-01-2011:09.18	decide	Sara	200	...
	35654526	06-01-2011:12.18	reinitiate request	Sara	200	...
	35654527	06-01-2011:13.06	examine thoroughly	Sean	400	...
	35654530	08-01-2011:11.43	check ticket	Pete	100	...
	35654531	09-01-2011:09.55	decide	Sara	200	...
	35654533	15-01-2011:10.45	pay compensation	Ellen	200	...

GAMBAR 2.
Contoh *Event Log*

Pada Gambar 2 Contoh *event log* terdiri dari sekumpulan *case* dan dibuat menjadi tabel khusus. Setiap *case* dalam *event log* diidentifikasi oleh sebuah *id*. Terlepas dari jenis analisis yang dilakukan pada *process mining*, masing - masing *event* tersebut diberi ciri (*case identifier*) berupa *case id* yang menginformasikan *case* yang berkaitan dengan *event*

tersebut dan sebuah label "*activity*" menjelaskan tindakan terkait. Timestamp merupakan kolom mengenai informasi waktu pada masing – masing aktivitas yang ada. Pada kolom resource ini merupakan sumber data diambil dari user. Cost yang ada ini merupakan nilai untuk suatu aktivitas tertentu pada sebuah *event log* [6].

C. Preprocessing

Data *preprocessing* merupakan tahap utama dan penting yang bertujuan untuk menghasilkan data yang berku- alitas dan berguna untuk dilakukan *process mining*. Saat mahasiswa terlibat dalam kegiatan pembelajaran, data berupa *log* akan tersimpan dalam sistem sebagai hasil kegiatan pembelajaran yang dilakukan dalam *database*. Pe- nerapan proses pemodelan diterapkan ketika *cleansing* terhadap *log* dilakukan kemudian diubah menjadi format yang dibutuhkan karena data pada suatu *log* memiliki kekurangan seperti atribut yang tidak lengkap, hilang atau perlu dihilangkan, dan data yang mengandung *error* atau *noise* [1].

Berikut metode-metode dalam melakukan *preprocessing*:

1. *Data Cleansing*. Pada tahap ini data yang mengandung *noise* dan data yang tidak konsisten atau data tidak relevan dihilangkan [6]
2. *Data Integration.* Pada tahap ini menggabungkan dua atau lebih data dari berbagai sumber *database* yang berbeda untuk mempermudah dalam proses analisa data. *Data integration* dilakukan pada atribut yang mengidentifikasi entitas unik seperti *id user, user*, dan lainnya [6].
3. *Data Reduction.* Pada tahap ini dilakukan pengurangan jumlah data tanpa menghilangkan kualitas pada data tersebut. Tujuannya yaitu mengatasi lamanya waktu yang dibutuhkan untuk menganalisis data kompleks dalam tiap *dataset* yang lengkap [6].
4. *Data Transformation.* Pada tahap ini data dipindahkan dari suatu struktur data ke struktur data lainnya dengan cara menghapus *noise* dari data, agregasi data, generalisasi data, dan normalisasi data [6].

D. *Clustering*

Clustering merupakan proses pengelompokan data ke dalam kelompok atau *cluster*, dimana setiap *cluster* memiliki data yang memiliki kesamaan tinggi dan antara *cluster* memiliki kesamaan yang rendah. Pengukuran jarak untuk mengukur kesamaan data memiliki peran yang sangat penting dalam kinerja algoritma *K-means*. Pada penelitian ini *clustering* algoritma *K-means* digunakan untuk menentukan variasi pola perilaku mahasiswa selama melakukan pengerjaan kuis secara daring [12].

Pada proses pengerjaan nya *clustering K-means* ini memerlukan jumlah *cluster* untuk menentukan hasil *cluster* yang diperlukan. Pada penelitian ini akan menggunakan metode *elbow method* sebagai penentuan *cluster* pada algoritma *K-means* ini. *Elbow method* adalah melakukan iterasi sebanyak k *cluster* pada algoritma *K-means*, dalam hal ini x *cluster*. Kemudian dari x iterasi tersebut ditentukan jumlah kuadrat dari jarak titik nilai aktivitas *event log* dalam *cluster* ke titik pusat *cluster*, kemudian dilakukan proses ploting dari x cluster terhadapa nilai jumlah kuadrat tersebut. Jika terdapat nilai jumlah kuadrat yang sangat berbeda (dalam plot terlihat patahan) maka cluster tersebut sudah optimal untuk menggambarkan jumlah *cluster* dari *event log* [17].

E. *Heuristic Miner*

Heuristic miner merupakan algoritma *process mining* yang dikembangkan oleh Dr. Ton Weitjers yang meng- gunakan pendekatan *heuristic* untuk mengatasi kekurangan pada algoritma *alpha*. Algoritma *heuristic miner* mengambil frekuensi peristiwa dan mengelola pencarian informasi dari perspektif proses. Maka hal yang diper- timbangkan adalah urutan aktivitas di dalam kasus serta penggunaan frekuensi membuat pendekatan ini jauh lebih kuat daripada kebanyakan pendekatan lainnya [6]. *Heruistic miner* dapat menangani *noise* dan dapat mengekspresikan kejadian utama yang terekam di dalam *event log* dengan baik. Algoritma *heuristic miner* membe- rikan pemodelan alur kerja ilmiah dengan pertimbangan urutan aktivitas dalam suatu kasus, dapat menerjemahkan *unstructed process* dan dapat menangani *incomplete logs* [1].

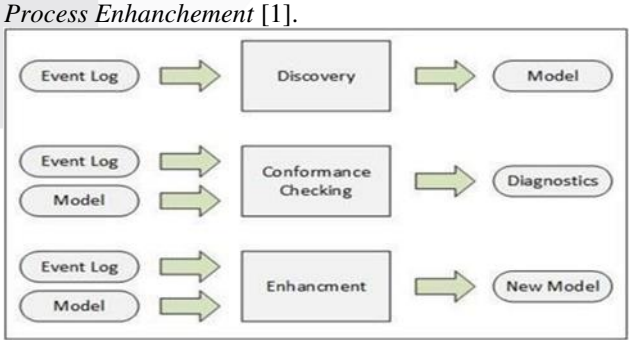
Heuristic miner memiliki tiga ambang batas yaitu:

1. *Dependency Threshold*, parameter yang mengacu nilai dari *dependency measure* antar aktivitas. Jika nilai *dependency measure* lebih besar dibanding nilai *dependency threshold*, maka *dependency relation* antar aktivitas tersebut diikut sertakan dalam model proses [6].
2. *Positive Observation Threshold*, parameter yang mengacu nilai frekuensi dari aktivitas. Jika nilai frekuensi- nya lebih besar dari nilai *positive observation threshold*, maka aktivitas tersebut dapat diikut sertakan dalam model proses [6].

Relative-to-best Threshold, parameter yang mengacu nilai dari selisih *dependency measure* aktivitas dengan *dependency measure* terbaik yang tercatat sebelumnya. Jika nilai selisihnya lebih kecil dibanding nilai *relative- to- best threshold*, maka aktivitas tersebut dapat diikutsertakan dalam model proses [6].

F. *Process Mining*

Process mining merupakan teknik yang menggunakan *evemt data* untuk mendapatkan suatu model proses, memeriksa compliance, menganalisis, membandingkan variasi proses dan dapat juga memberikan saran perbaikan. Pada penelitian ini *process mining* diterapkan untuk melihat model proses dari pola perilaku mahasiswa dalam mengerjakan kuis secara daring dengan algoritma *heuristic miner*. *Process mining* memiiki tiga proses utama yaitu *Process Discovery*, *Process Conformance Checking*, dan *Process Enhanchement* [1].



GAMBAR 3.
Tahapan pada *process mining*

Berikut proses - proses utama pada *process mining*:

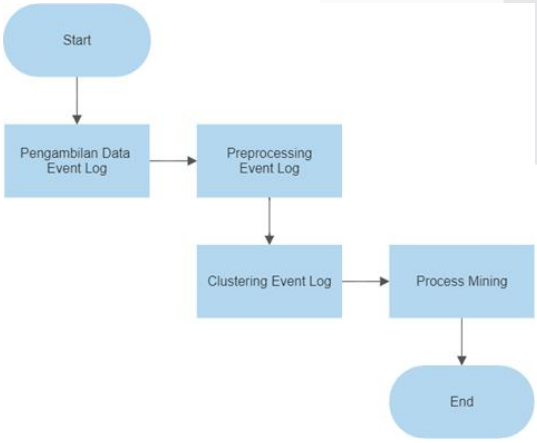
1. Process Discovery, pada proses ini dilakukan pembuatan suatu model berdasarkan *event log*

yang berasaldari sistem informasi yang digunakan informasi bersifat asumsi [6].

2. *Process Conformance Checking*, proses ini membandingkan model proses yang ada dengan event log dari proses yang sama. *Process conformance checking* digunakan untuk memeriksa kesesuaian kejadian nyata yang terjadi dengan model yang ada dan sebaliknya. Proses ini dapat digunakan untuk mengukur kinerja algoritma *process discovery* dan mendapatkan kesesuaian terhadap model yang telahdidapat sebelumnya. Hasil dari analisis yang dilakukan terhadap *process conformance checking* yaitu analisis*fitness*, *precision*, dan *generalization*. Analisis *fitness* berkaitan dengan pemeriksaan terhadap model proses yang dihasilkan mampu menggambarkan kembali seluruh urutan pada *event log* dan sebaliknya. Tujuan analisis nilai *fitness* adalah mendeteksi kesesuaian model proses dengan yang sebenarnya terekam dalam *event log*. Nilai *fitness* berkisar antara 0-1, dimana jika nilai semakin kecil mendekati 1 maka model menggambarkan catatan kejadian semakin akurat. Untuk nilai *precision* berkaitan dengan notasi overfitting dalam konteks data mining. Model proses dikatakan overfitting apabila model proses tersebut sangat spesifik dan berpatokan penuh pada contoh proses di event log. Nilai *precision* berkisar 0-1 semakin tinggi nilai presisi suatu model proses, maka semakin besar kecenderungan model tersebut dalam notasi overfitting. Pada nilai *generalization* berkaitan dengan notasi underfitting dalam konteks data mining. Model proses dikatakan underfitting apabila model proses tersebut menunjukkan proses yang cenderung berbeda dengan proses yang terlihat pada event log. Nilai *generalization berkisar 0-1* semakin tinggi nilai generalisasi model proses, maka semakin besar kecenderungan model tersebut dalam notasi underfitting [16].

Process Enhancement, memproses *event log* dengan model untuk mendapatkan model baru. Proses ini bertujuan memperluas atau memperbaiki model proses yang ada menggunakan informasi yang didapat dari *event log* [6].

III. METODE



GAMBAR 4.
Flowchart

Seperti yang terlihat pada gambar 4 penelitian ini akan dimulai dari pengambilan data event log pada LMS lalu data yang ada akan memasuki tahapan preprocessing. Setelah data selesai melalui tahapan preprocessing ini dilakukan clustering untuk menghasilkan varian dataset yang dibutuhkan untuk selanjutnya dilakukan tahapan process mining.

A. Pengambilan Data

Dalam penelitian ini objek data yang akan dikumpulkan dan digunakan adalah data dari event log LMS Universitas Telkom yang di dapatkan dari CeLOe. Didalam event log tersebut memuat data berupa rekaman kegiatan yang tercatat dalam mata kuliah. Data yang digunakan untuk penelitian ini adalah data event log dari mata kuliah Algortima Pemrograman tahun ajaran 2021/2022 semester genap jurusan Informatika. Dataset yang gunakan yaitu kelas IF-45-01, IF-45-02, IF-45-03, IF-45-04, IF-45-05, IF-45-06, IF-42-07, IF-45-08, IF-45-09,
IF-45-10, IF-45-11, IF-45-12, dan IF-45-INT. Didalam event log tersebut memiliki data mentah sebanyak 499.397 rows dari jumlah mahasiswa sebanyak 575. Data eventlog akan dibagi menjadi beberapa varian yang dihasilkan dari clustering sesuai pola perilaku selama pengerjaan ujian tengah semester.

Preprocessing

Pada penelitian ini, preprocessing yang dilakukan pada data mata kuliah Algoritma Pemrograman dengan menggunakan Jupyter Notebook yang bertujuan untuk mengurangi ukuran data, menghapus row yang dilakukan selain mahasiswa, menambah atribut atau kolom baru yang digunakan untuk process mining [8]. Dalam tahap preprocessing juga melakukan perubahan pada timestamp tanggal menjadi DD/MM/YY dan untuk waktu HH/MM. Kemudian Case ID yang ditentukan berdasarkan User Id. Tabel 2 merupakan deskripsi dari data event log yang didapatkan [1].

TABEL 2.
Deskripsi kolom Event Log

No	ribut/Kolom Data Mentah	Keterangan
1	*Time*	Waktu dimana user akses LMS
2	*User Full Name*	Nama user yang melakukan akses di LMS
3	*Affected User*	User lain yang terlibat dalam aktivitas user pertama
4	*Event Context*	melakukan akses terhadap menu didalam LMS
5	*Component*	nana user melakukan akses terhadap menu didalam LMS
6	*Event Name*	s dari user ketika melakukan akses padamenu yang ada didalam LMS
7	*Description*	user saat menjalankan menu yang ada didalam LMS
8	*Origin*	Media yang digunakan user untuk mengakses LMS

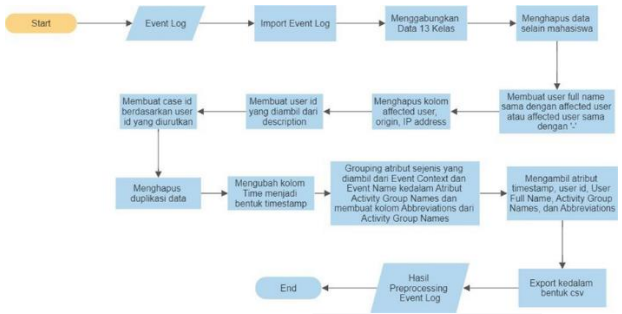
9	*IP Address*	Alamat IP dari user saat mengakses LMS

Berikut ini adalah contoh data mentah event log Learning Management System Universitas Telkom

TABEL 3.
Contoh Event Log Algoritma Pemrograman

No	Nama Atribut/Kolom Data Mentah	Contoh Isi Dari Data Mentah
1	*Time*	30/06/22, 13:14
2	*User Full Name*	AMURA MAULIDI FACHRY
3	*Affected User*	-
4	*Event Context*	File: Pdf materi sorting
5	*Component*	File
6	*Event Name*	Course module viewed
7	*Description*	The user with id '56693' viewed the 'resource' activity with course module id '2126219'.
8	*Origin*	web
9	*IP Address*	223.255.224.102

Data mentah tersebut tidak memenuhi kriteria pemodelan dalam process mining dan memerlukan preprocessing. Berikut pada gambar 5 merupakan flowchart untuk melakukan preprocessing.



GAMBAR 5.
Flowchart Preprocessing

Pada penelitian ini *preprocesing* dilakukan menggunakan Jupyter Notebook. Tahap pertama yang dilakukan adalah *data integration* yaitu menggabungkan data dari 13 kelas yang berbeda setelah *data integration* selesai selanjutnya melakukan data *cleaning* dari data *event log* diolah dengan menghapus aktivitas yang bukan dilakukan oleh mahasiswa. Membuat user full name sama dengan affected user agar tidak ada data ganda. Seteleh itu hapus beberapa kolom yang tidak dibutuhkan seperti kolom *affected user, origin, IP address.* Membuat *user id* berdasarkan kolom *description* yang dilakukan oleh mahasiswa dalam menggunakan LMS, setelah mendapatkan *user id,* membuat *case id* berdasarkan kolom *user id* yang diurutkan. Selanjutnya menghapus duplikasi data. Mengubah *format time* menjadi *timestamp* DD/MM/YY untuk tanggal dan untuk waktu HH:MM. Mengambil kolom yang dibutuhkan untuk tahap *process mining* seperti kolom *Timestamp, user_id, User, Event context, case_id, Component*. Dilakukan proses *grouping* atribut sejenis yang diambil dari *Event Context* dan *Event Name* kedalam Atribut Activity Group Names dan membuat kolom Abbreviations dari *Activity Group Names*. Berikut ini tabel pada tahap *grouping* atribut *Activity Group Names*.

TABEL 4.
Grouping *Atribut Activity Group Names*

No	*Event Context*/Name	*Activity Group Names*	*Abbreviations*
1	*Assignment*	activity_viewed	av
2	*Forum*	activity_viewed	av
3	*Quiz*	activity_viewed	av
4	*Quiz*: 2. SESI PILIHAN GANDA DAN TRACING/ISIAN	course_module_viewed	cmv
5	*Quiz*: 3. SESI ESAI	course_module_viewed	cmv
6	*Course*: ALGORITMA PEMROGRAMAN IF-45-XX [YYY]	course_viewed	cv
7	*File*	material_viewed	mv
8	*Folder*	material_viewed	mv
9	*URL*	material_viewed	mv
10	*Other*	other_course_viewed	ocv
11	*Quiz Attempt Started*	quiz_attempt_started	qas
12	*Quiz Attempt Submitted*	quiz_attempt_submitted	qasub
13	*Quiz Attempt Summary Viewed*	quiz_attempt_summary_viewed	qasv
14	*Quiz Attempt Viewed*	quiz_attempt_viewed	qav

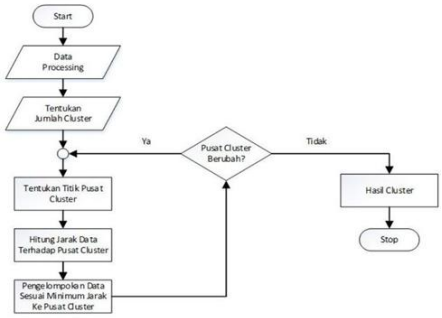
Data yang telah melewati *preprocessing* merupakan kumpulan dataset untuk mata kuliah algoritma pemrograman, seperti terlihat pada tabel 5.

TABEL 5.
Contoh Isi Dataset Setelah *Preprocessing*

No	Nama Atribut	Isi Dataset Setelah *Preprocessing*
1	*Time*	4/12/2022 13:19
2	*User Full Name*	ADIANDRA HAIKAL
3	user_id	'58268'
4	*Activity Group Names*	material_viewed
5	*Abbreviations*	mv

A. Clustering

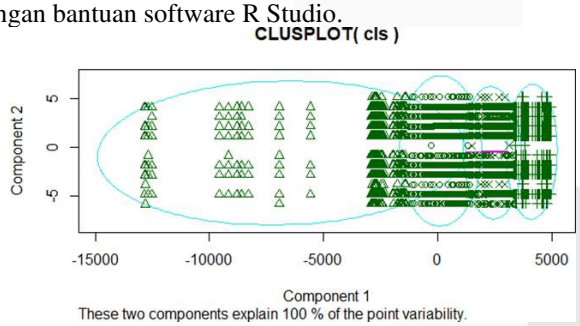
Pada penelitian ini dilakukan proses pengelompokan dengan menggunakan analisis *clustering* untuk mendapatkan hasil *interpretasi* pola perilaku mahasiswa. Analisis *clustering* ini dapat memetakan kesamaan pola perilaku mahasiswa. Proses *clustering* ini menggunakan algoritma *K-Means*. Algoritma ini bekerja dengan menggunakan prinsip ukuran pemusatan untuk menentukan titik awal dalam proses pengelompokannya [17]. Berikut merupakan proses algoritma *K-Means* yang dilakukan terhadap tahap pembentukan varian – varian.



GAMBAR 6.
Proses Algoritma K-Means Pembentukan Varian

Pada proses ini penentuan jumlah cluster dilakukan sebagai bentuk evaluasi cluster menggunakan metode elbow method sebagai penentuan cluster pada algoritma K-means ini. Elbow method ini melakukan iterasi sebanyak k cluster pada algoritma K-means, dalam hal ini 10 cluster. Kemudian dari 10 iterasi tersebut ditentukan jumlah kuadrat dari jarak titik nilai aktivitas event log dalam cluster ke titik pusat cluster, kemudian dilakukan proses ploting dari 10 cluster terhadap nilai jumlah kuadrat tersebut. Jika terdapat nilai jumlah kuadrat yang sangat berbeda (dalam plot terlihat patahan) maka cluster tersebut sudah optimal untuk menggambarkan jumlah varian dari pola perilaku mahasiswa [17]

Dari proses algoritma K-means diatas dapat ditentukan varian - varian yang akan terbentuk. Hasil dari tiap varian-varian ini akan berisi beberapa user_id yang sudah dikerjakan oleh algoritma K-means berdasarkan event log yang sudah ditransformasi menjadi nilai numerik. Berikut visualisasi dari hasil proses clustering algoritma K-Means seperti yang terlihat pada CLUSPOT (Cluster Plot), terdapat empat cluster (varian) yang terbentuk. Proses clustering ini digunakan dengan bantuan software R Studio.

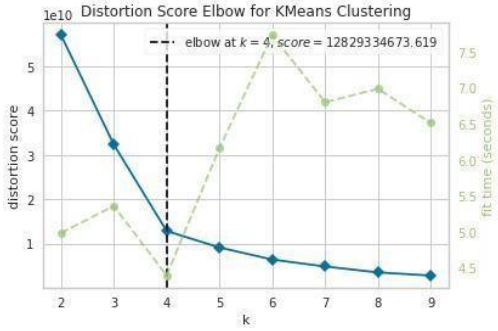


GAMBAR 7.
Visualisasi hasil proses Clustering algoritma K-Means

TABEL 6.
Jumlah Mahasiswa Tiap Cluster

No	Varian	Jumlah Mahasiswa
1	1	209
2	2	188
3	3	63
4	4	100

Setelah didapat hasil *cluster* ini dilakukan proses evaluasi *cluster* untuk memastikan bahwa jumlah cluster yang dihasilkan merupakan jumlah yang terbaik menggunakan *elbow method*. Berikut hasil gambaran dari *elbow method* yang dilakukan:



GAMBAR 8.
Visualisasi Hasil Proses Evaluasi *Clustering Elbow Method*

Dari 4 cluster (varian) yang dihasilkan berdasarkan titik pusat cluster yang ditentukan oleh algoritma *K- Means*. Dari CLUSPLOT dihasilkan nilai *component* 1 dan *component* 2 yang merupakan standarisasi dari user_id dan *event log* yang sudah di transformasi sebelumnya. Terjadi *overlapping* di tiap *cluster*, hal ini diakibatkan tiap *cluster* memiliki *event log* yang sama namun terdapat perbedaan pada *event log* tertentu. Berikut diinterpretasikan kriteria hasil tiap *cluster* (varian) dilihat berdasarkan *trace* plotnya:

Varian pertama didapat contoh trace seperti berikut

GAMBAR 9.
Varian 1 dari hasil proses clustering

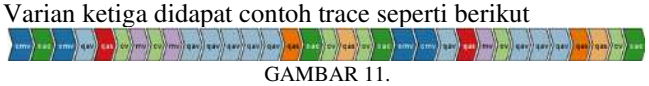
Pada trace ini terlihat mahasiswa mengerjakan ujian tengah semester tanpa ada kegiatan lain selain membuka aktivitas ujian. Dapat disimpulkan bahwa pola pada varian 1 merupakan tipe mahasiswa yang sudah mempersiapkan materi ujian tengah semester sebelum kegiatan ujian berlangsung.

Varian kedua didapat contoh trace seperti berikut

GAMBAR 10.
Varian 2 dari hasil proses clustering

Pada trace ini terlihat mahasiswa melakukan aktivitas lain sebelum memulai mengerjakan ujian tengah semester, tipe pada varian ini pola mahasiswa cenderung untuk membaca modul terlebih dahulu. Kemudian mengakses aktivitas lain sebelumnya kemudian baru melakukan aktivitas Ujian. Pada varian ini juga merupakan tipe mahasiswa yang mempersiapkan bahan ujian sebelum ujian dimulai.

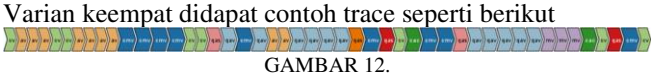
Varian ketiga didapat contoh trace seperti berikut



GAMBAR 11.
Varian 3 dari hasil proses clustering

Pada trace ini terlihat mahasiswa melakukan aktivitas lain setelah memulai mengerjakan ujian tengah semester. Tipe pada varian ini pola mahasiswa cenderung untuk membaca material terlebih dahulu (multitasking), ditengah melakukan aktivitas Ujian.

Varian keempat didapat contoh trace seperti berikut

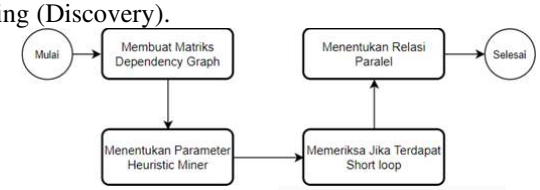


GAMBAR 12.
Varian 4 dari hasil proses clustering

Pada trace ini terlihat mahasiswa melakukan aktivitas lain sebelum memulai. Tipe pada varian ini pola mahasiswa cenderung untuk multitasking yang artinya mahasiswa tersebut sebelum dan sesudah mengerjakan ujian selalu membaca material terlebih dahulu. Pada varian ini terindikasi bahwa mahasiswa-mahasiswa ini tidak siap ketika ujian. Performansi clustering tidak dapat ditentukan secara spesifik seperti berdasrkan akuras, hal ini dikarenakan clustering merupakan model unsupervised atau yang disebut tanpa ada pengawasan secara langsung oleh variabel dependen [17].
Hasil yang didapat dari varian pola perilaku mahasiswa ini dapat dicocokan sesuai aturan yang berlaku selama pengerjaan ujian berjalan. Pada kasus ini aturan yang ditetapkan merupakan bentuk ujian open book dimana mahasiswa diperbolehkan mengakses catatan.

B. Process Mining (Discovery)

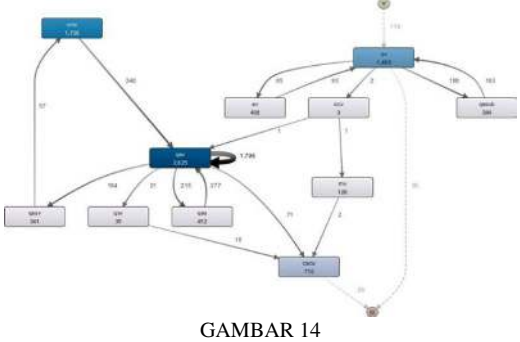
Process discovery merupakan sebuah proses ekstraksi data event log untuk mendapatkan model proses menggunakan PM4PY dengan menerapkan algoritma Heuristic Miner. Pada tahap process mining (discovery), setelah melakukan tahap preprocessing terhadap event log dan menghasilkan file CSV. Selanjutnya pada tahap process mining (discovery) dilakukan penerapan algoritma heuristic miner menggunakan process mining for python (PM4Py). Gambar 13 merupakan flowchart untuk melakukan process mining (Discovery).



GAMBAR 13.
Flowchart process mining

Tahapan pertama yang dilakukan pada process discovery dengan algoritma Heuristic Miner adalah membut matriks dependency graph untuk menyimpan jumlah relasi ketergantungan antar dua event. Lalu menentukan parameter heuristic miner seperti Dependency threshold, And threshold dan Loop Two threshold untuk dasar pemilihan relasi ketergantungan yang akan ditampilkan dalam model proses, setelah itu memeriksa jika terdapat short loop (length-one-loop atau length-two-loop), Menentukan relasi parallel (XOR atau AND) antar event/aktivitas. Setelah semua tahap dilakukan model proses algortima heuristic miner dapat terbentuk

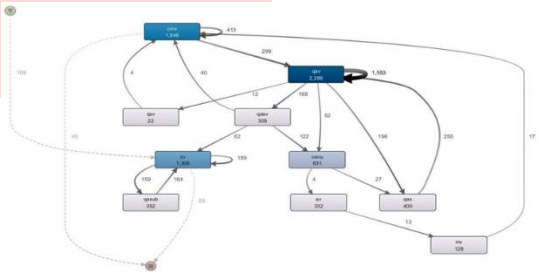
Process discovery membutuhkan data input event log untuk membuat model proses yang mewakili proses aktual yang sedang berjalan. Dibawah ini gambar menunjukkan model proses belajar sederhana mahasiswa dari masing – masing varian menggunakan DISCO tools.



GAMBAR 14
Model Proses Varian 1.

TABEL 7.
Aktivitas Frekuensi Varian 1

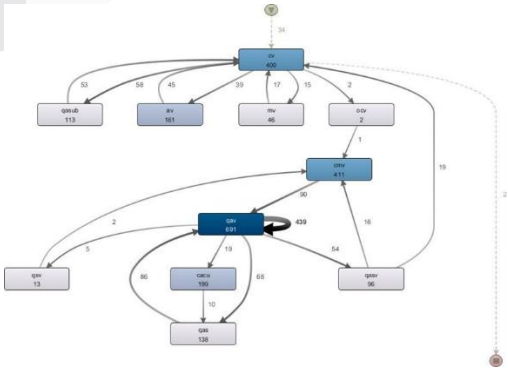
Aktivitas	Frekuensi
quiz_attempt_viewed	31.93%
course_module_viewed	21.14%
course_viewed	17.03%
course_activity_completion_updated	8.64%
quiz_attempt_started	5.5%
activity_viewed	4.96%
quiz_attempt_submitted	4.67%
quiz_attempt_summary_viewed	4.15%
material_viewed	1.53%
other_course_viewed	0.04%



GAMBAR 15
Model Proses Varian 2

TABEL 8.
Aktivitas Frekuensi Varian 2

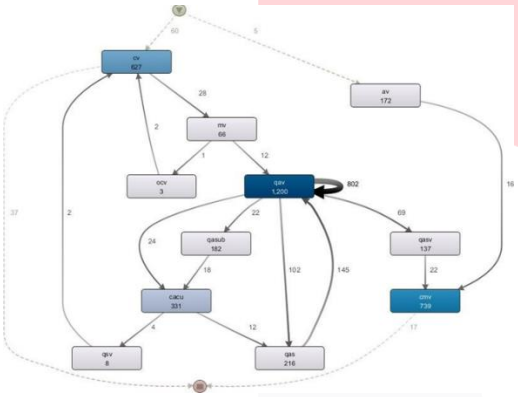
Aktivitas	Frekuensi
quiz_attempt_viewed	31.28%
course_module_viewed	21.17%
course_viewed	17.85%
course_activity_completion_updated	8.62%
quiz_attempt_started	5.47%
quiz_attempt_submitted	4.81%
activity_viewed	4.54%
quiz_attempt_summary_viewed	4.22%
material_viewed	1.75%



GAMBAR 16
Model Proses Varian 3

TABEL 9.
Aktivitas Frekuensi Varian 3

Aktivitas	Frekuensi
quiz_attempt_viewed	30.44%
course_module_viewed	18.11%
course_viewed	17.62%
course_activity_completion_updated	8.77%
activity_viewed	7.09%
quiz_attempt_started	6.08%
quiz_attempt_submitted	4.98%
quiz_attempt_summary_viewed	4.23%
material_viewed	2.03%
other_course_viewed	0.09%



GAMBAR 17
Model Proses Varian 4

TABEL 10.
Aktivitas Frekuensi Varian 4

Aktivitas	Frekuensi
quiz_attempt_viewed	32.6%
course_module_viewed	20.08%
course_viewed	17.03%
course_activity_completion_updated	8.99%
quiz_attempt_started	5.87%
quiz_attempt_submitted	4.94%
activity_viewed	4.67%
quiz_attempt_summary_viewed	3.72%
material_viewed	1.79%
other_course_viewed	0.08%

C. Conformance Checking

Setelah process discovery dilakukan dan hasil model proses telah didapat, selanjutnya dilakukan tahap conformance checking. Conformance checking ini dilakukan untuk memverifikasi apakah model yang dihasilkan sebelumnya sesuai dengan event log yang ada dan sebaliknya. Pada tahap ini pun dilakukan transformasi dari model proses berbentuk heuristic net menjadi petri net menggunakan PM4PY. Proses setelah berbentuk petri net dianalisis dengan conformance checker, hasilnya berupa nilai fitness, precision, generalization yang menunjukan kecocokan antara event log dan proses model juga sebaliknya.

IV. HASIL DAN PEMBAHASAN

Bagian ini berisi dua sub-bagian, yaitu Hasil Pengujian dan Analisis Hasil Pengujian. Pengujian dan analisis yang dilakukan selaras dengan tujuan TA sebagaimana dinyatakan dalam Pendahuluan.

A. Conformance Checking

Untuk mendapatkan hasil yang terbaik dari model proses dalam proses discovery, dilakukan beberapa pengujian conformance. Model proses diuji dengan menetapkan beberapa nilai pada parameter untuk menemukannilai fitness yang terbaik. Parameter yang digunakan pada algoritma heuristic miner di PM4PY ini adalah dependency threshold, and threshold, dan loop two threshold. Nilai ketiga parameter ini ditetapkan pada PM4PY untuk algoritma heuristic miner adalah 0,5 untuk dependency threshold, 0,65 untuk and threshold, dan 0,5 untuk loop two threshold. Nilai – nilai ini ditetapkan berdasar default pada PM4PY. Lalu pada conformance ini salah satu nilai parameter yang digunakan akan diganti secara acak dan nilai parameter lainnya tetap sesuai dengan default yang ada dan bergantian diubah. Hasil dari gabungan parameter ini akan ditampilkan menggunakan tabel dan dianalisis pengaruh ketiga parameter tersebut terhadap nilai fitness yang dihasilkan terhadap model proses.

B. Analisis Hasil Pengujian Conformance Checking

Pengujian digunakan untuk menentukan model proses terbaik menggunakan heuristic miner, dengan melakukan perbandingan perubahan hasil nilai fitness terhadap parameter dependency threshold (DT), andthreshold (AT), dan loop two threshold (LTT). Pengujian ini dilakukan menggunakan PM4PY.

C. Analisis Parameter Dependency Threshold (DT)

TABEL 11.
Hasil pengujian terhadap parameter Dependency Threhold (DT)

Varian	DT	AT	LTT	Fitness
1	0,1	0,65	0,5	0.9033183791412172
	0,2	0,65	0,5	0.9468705490029147
	0,3	0,65	0,5	0.9515948076011398
	0,4	0,65	0,5	0.9741568356420658
	0,5	0,65	0,5	0.9812462225869050
2	0,1	0,65	0,5	0.9079756098094653
	0,2	0,65	0,5	0.9617337447034288
	0,3	0,65	0,5	0.9778924847576309
	0,4	0,65	0,5	0.9875396106034748
	0,5	0,65	0,5	0.9875396106034748
3	0,1	0,65	0,5	0.9114631536774891
	0,2	0,65	0,5	0.9574485758587239
	0,3	0,65	0,5	0.9719227320776307
	0,4	0,65	0,5	0.9827777248724099
	0,5	0,65	0,5	0.9856123418619721
4	0,1	0,65	0,5	0.8896632922491345
	0,2	0,65	0,5	0.9720846060325632
	0,3	0,65	0,5	0.9844023305587621
	0,4	0,65	0,5	0.9848820846721071
	0,5	0,65	0,5	0.9881454735905697

Berdasar hasil pengujian seperti pada tabel 11, mahasiswa dengan varian 1 memiliki nilai fitness 0,90331-0,98124, varian 2 memiliki nilai fitness 0,90797-0,98753, varian 3 memiliki nilai fitness 0,91146- 0,98561, dan varian 4 memiliki nilai fitness 0,88966-0,98814. Hasil pengujian tersebut menunjukan nilai fitness yang dihasilkan cukup baik untuk menggambarkan event log pada model proses. Dapat dikatakan bahwa semakin tinggi nilai parameter dependency threshold ini maka semakin tinggi nilai fitness yang dihasilkan.

D. Analisis Parameter And Threshold (AT)

TABEL 12.
Hasil pengujian terhadap parameter And Threhold (AT)

Varian	DT	AT	LTT	Fitness
1	0,5	0,2	0,5	0.9795635415856236
	0,5	0,3	0,5	0.9795635415856236
	0,5	0,4	0,5	0.9796847433305755
	0,5	0,5	0,5	0.9796847433305755
	0,5	0,65	0,5	0.9812462225869050
2	0,5	0,2	0,5	0.9826493441383093
	0,5	0,3	0,5	0.9823305597264316
	0,5	0,4	0,5	0.987210572550344
	0,5	0,5	0,5	0.987210572550344
	0,5	0,65	0,5	0.9875396106034748
3	0,5	0,2	0,5	0.9767872881972564
	0,5	0,3	0,5	0.9852908280659153
	0,5	0,4	0,5	0.9856123418619721
	0,5	0,5	0,5	0.9857144818140995
	0,5	0,65	0,5	0.9856123418619721
4	0,5	0,2	0,5	0.9856196267689117
	0,5	0,3	0,5	0.9856234731335456
	0,5	0,4	0,5	0.9874042774419872
	0,5	0,5	0,5	0.9874042774419872
	0,5	0,65	0,5	0.9881454735905697

Berdasar hasil pengujian seperti pada tabel 12, mahasiswa dengan varian 1 memiliki niali fitness 0,97956-0,98124, varian 2 memiliki nilai fitness 0,98264-0,98753, varian 3 memiliki nilai fitness 0,97678- 0,98561, dan varian 4 memiliki nilai fitness 0,98561-0,98814. Hasil pengujian tersebut menunjukan nilai fitness yang dihasilkan cukup baik untuk menggambarkan event log pada model proses. Dapat dikatakan bahwa seluruh aktivitas yang memiliki frequency rendah akan menyebabkan nilai fitness yang dihasilkan menjadi rendah. Sebaliknya jika nila and threshold tinggi akan meningkatnya nilai fitness.

E. Analisis Parameter Loop Two Threshold (LTT)

TABEL 13.
Hasil pengujian terhadap parameter Loop Two Threhold (LTL)

Varian	DT	AT	LTT	Fitness
1	0,5	0,65	0,5	0.981246222586905
	0,5	0,65	0,6	0.981246222586905
	0,5	0,65	0,7	0.9788072652480403
	0,5	0,65	0,8	0.9713482802859063
	0,5	0,65	0,9	0.9535817200137933
	0,5	0,65	0,5	0.9875396106034748
	0,5	0,65	0,6	0.982393214514546
	0,5	0,65	0,7	0.9792106165114258
2	0,5	0,65	0,8	0.9664890886137959
	0,5	0,65	0,9	0.9515418612054771
3	0,5	0,65	0,5	0.9856123418619721
	0,5	0,65	0,6	0.9671287741934005
	0,5	0,65	0,7	0.9549328061465827
	0,5	0,65	0,8	0.9549328061465827
	0,5	0,65	0,9	0.9522563107013013
4	0,5	0,65	0,5	0.9881454735905697
	0,5	0,65	0,6	0.96745092871858
	0,5	0,65	0,7	0.9666736230610073
	0,5	0,65	0,8	0.9652574284678201
	0,5	0,65	0,9	0.9586396302913842

Berdasar hasil pengujian seperti pada tabel 13, mahasiswa dengan varian 1 memiliki niali fitness 0,95358-0,98124, varian 2 memiliki nilai fitness 0,95154-0,98753, varian 3 memiliki nilai fitness 0,95225- 0,98561, dan varian 4 memiliki nilai fitness 0,95863-0,98814. Hasil pengujian tersebut menunjukan nilai fitness yang dihasilkan cukup baik untuk menggambarkan event log pada model proses. Dapat dikatakan bahwa semakin tinggi parameter loop two threshold maka semakin rendah nilai fitness.

F. Evaluasi Hasil Pengujian Conformance Checking

Conformance Checking melakukan analisis hubungan antara aktivitas yang tergambarkan dalam model proses dan event log. Penelitian inipun menguji kesesuaian model proses mahasiswa dengan event log yang ada di LMS dan menganalisis kinerja algoritma heuristic miner berdasar nilai–nilai seperti nilai fitness, precision, dan generalization.

G. Hasil Pengujian Conformance Checking

Pengujian yang dilakukan pada perubahan parameter DT, AT, dan LTT tersebut menghasilkan model proses yang memiliki nilai fitness cukup baik. Pada tabel 14 ini menunjukan hasil keseluruhan untuk parameter DT, AT, dan LTT dengan nilai fitness dihasilkan terbaik dari varian mahasiswa yang ada.

TABEL 14.
Hasil pengujian Conformance Checking

Varian	DT	AT	LTT	Fitness
1	0,5	0,65	0,5	0.9812462225869050
2	0,5	0,65	0,5	0.9875396106034748
3	0,5	0,65	0,5	0.9856123418619721
4	0,5	0,65	0,5	0.9881454735905697

Dari hasil pengujian ini, nilai fitness terbaik yang dihasilkan oleh varian 1 dengan nilai 0,98124, varian2 dengan nilai 0,98753, varian 3 dengan nilai 0,98561, dan varian 4 dengan nilai 0,98814.

H. Evaluasi Hasil Conformance Checking

Hasil pengujian conformance checking ini terdiri dari tiga bagian yaitu nilai fitness, precision, dan generalization menggunakan PM4PY seperti tabel berikut ini.

TABEL 15.
Evaluasi hasil Conformance Checking

Varian	Fitness	Precision	Generalization
1	0,98124	0.2610093825665859	0.7918239375551215
2	0,98753	0.22306411801848347	0.7880959737376002

3	0,98561	0.22625443487075525	0.6967753389957655
4	0,98814	0.20957573692281484	0.7430874529734255

Evaluasi nilai fitness yang terdapat pada Tabel 15 mahasiswa varian 1 mendapatkan nilai fitness 0,98124, mahasiswa varian 2 dengan nilai 0,98753, mahasiswa varian 3 dengan nilai 0,98561, dan mahasiswa varian 4 dengan nilai 0,98814. Hasil ini menunjukan bahwa proses model dapat memodelkan event log yang ada dengan baik, dilihat dari nilai sempurna yang dihasilkan jika seluruhan log secara benar dan sempurna adalah bernilai 1. Nilai fitness yang cukup baikpun memiliki arti bahwa model proses yang dibuat sangat baik dan menunjukkan bahwa semua trace atau case pada event log dapat diwakili oleh model prosesnya dengan baik.

Evaluasi nilai precision yang terdapat pada tabel 14 mahasiswa varian 1 mendapatkan nilai 0,26100, mahasiswa varian 2 dengan nilai 0,22306, mahasiswa varian 3 dengan nilai 0,22625, dan mahasiswa varian 4 dengan nilai 0,20957. Hasil ini menunjukan bahwa yang didapatkan belum sempurna dikarenakan model yang dihasilkan terlalu mengeneralisasi contoh behavior event log. Model memungkinkan menampilkan behavior berbeda dari apa yang terlihat pada event log.

Evaluasi nilai generalization yang terdapat pada tabel 14 mahasiswa varian 1 mendapatkan nilai 0,79182,mahasiswa varian 2 dengan nilai 0,78809, mahasiswa varian 3 dengan nilai 0,69677, dan mahasiswa varian 4 dengan nilai 0,74308. Hasil ini menunjukan bahwa yang didapatkan cukup baik. Generalization dikatakan baik jika model yang dihasilkan tidak overfitting. Model dikatakan overfitting dikarenakan model proses yang dihasilkan terlalu spesifik dan mengikuti contoh pada event log dimana menunjukan bahwa model tidak cukup mengeneralisasi suatu proses.

V. KESIMPULAN

Berdasarkan penelitian yang dilakukan dapat disimpulkan bahwa selama pengerjaan ujian tengah semester secara daring melalui LMS ini terdapat banyak pola perilaku yang dilakukan mahasiswa selama ujian. Pada kasus ini aturan yang diterapkan selama ujian merupakan bentuk ujian open book . Pada penelitian ini terdapat 4 varian yang ada pada event log yang ada berdasarkan proses clustering. Varian – varian yang dihasilkan memperlihatkan terdapat aktivitas selain pengerjaan ujian tengah semester selama waktu pengerjaan ujian. Namun terdapat perbedaan waktu aktivitas selain ujian tersebut, dan varian inipun dapat menggambarkan berbagai macam pola perilaku mahasiswa akan kesiapan untuk melakukan ujian tengah semester ini.

Pada penelitian ini penerapan algoritma heuristic miner ini dapat memodelkan event log pada model proses secara baik berdasar nilai fitness yang didapat. Nilai fitness yang didapat memiliki nilai yang baik. Pada varian 1 mendapatkan nilai fitness 0,98124, varian 2 dengan nilai 0,98753, varian 3 dengan nilai 0,98561, dan varian 4 dengan nilai 0,98814. Hal ini menunjukan bahwa proses model dapat memodelkan event log cukup baik. Nilai fitness yang cukup baik ini berarti model proses yang dibuat sangat baik dan menunjukan semua trace atau case pada event log ini dapat diwakili dengan model prosesnya. Meski nilai precision yang didapat kurang baik yaitu mahasiswa varian 1 mendapatkan nilai 0,26100, varian 2 dengan nilai 0,22306, varian 3 dengan nilai 0,226525, dan varian 4 dengan nilai 0,20957. Hal ini menunjukan bahwa hasil yang didapat belum sempurna atau underfitting, behavior dari model proses yang dihasilkan tidak menggambarkan behavior yang ada pada event log secara umum dengan baik karena model proses yang memiliki cukup banyak jalur untuk merepresentasikan urutan task yang berlangsung sesuai event log nya. Kemudian untuk nilai generalization memberikan hasil yang cukup baik terhadap keempat varian yang ada seperti mahasiswa varian 1 mendapatkan nilai 0,79182, varian 2 dengan nilai 0,78809, varian 3 dengan nilai 0,69677, dan mahasiswa varian 4 dengan nilai 0,74308. Nilai tersebut merupakn hasil yang cukup baik untuk nilai generalization.

Dari hasil parameter yang didapat penelitian ini akan dibatasi pada proses conformance checking karena nilai yang dihasilkan menunjukan hasil pemodelan yang akurat terhadap event log yang ada sehingga tidak diperlukan lanjutan proses enhancement.

Setelah melakukan penelitian ini saran yang diberikan dapat melakukan penelitian process mining menerapkan algoritma ataupun metode lain yang berbeda agar model proses yang dihasilkan dapat dijadikan perbandingan dengan penelitian ini. Penerapan process mining berdasar event log Learning Management System Universitas Telkom ini dapat ditingkatkan di penelitian selanjutnya dengan seluruh mata kuliah yang ada untuk memberikan peningkatan kualitas selama proses pembelajaran.

REFERENSI

[1] R. Andreswari I. A. Fitransah and M. A. Hasibuan. Analisis dan implementasi process mining pada aktivitas mahasiswa berdasarkan event log penggunaan aplikasi bimbingan tugas akhir. 2019.

[2] M. Y. Chkouri M. Ouadoud, A Nejjari and K. E. El-Kadiri. Learning management system and the underlying learning theories. Lect. Notes Networks Syst., 37:732–744, 2018.

[3] Zhao, YanPing & Zhou, XiaoLai. (2021). K-means Clustering Algorithm and Its Improvement Research. Journal of Physics: Conference Series. 1873. 012074. 10.1088/1742-6596/1873/1/012074.

[4] Pradana, Chesa & Kusumawardani, Sri & Permanasari, A. (2020). Comparison Clustering Performance Based on Moodle Log Mining. IOP Conference Series: Materials Science and Engineering. 722. 012012. 10.1088/1757-899X/722/1/012012.

[5] Singh, Satya. (2021). Study of K-Means and Enhanced K-Means Clustering Algorithm.. International Journal of Research in Computer Science. Vol. 4 Issue 10. p103-107. 5p.

[6] S. Chanifah, R. Andreswari and R. Fauzi, "Analysis of Student Learning Pattern in Learning Management System (LMS) using Heuristic Mining a Process Mining Approach," 2021 3rd International Conference on Electronics Representation and Algorithm (ICERA), Yogyakarta, Indonesia, 2021, pp. 121-125, doi: 10.1109/ICERA53111.2021.9538654..

[7] Agusriandi, Agusriandi & Elihami, Elihami & Syarif, Irman & Samad, Ita. (2022). Model Analisis Aktivitas Tutor Dalam Learning Management System Berdasarkan Data Log Menggunakan K-Means Dan

Deteksi Outlier. Jurnal Teknologi Informasi dan Ilmu Komputer. 9. 709. 10.25126/jtiik.2022934764.

[8] Nicholson C, Beattie L, Beattie M, Razzaghi T, Chen S (2022) A machine learning and clustering-based approach for county-level COVID-19 analysis. PLOS ONE 17(4): e0267558.

[9] Valensia, Larasati & Andreswari, Rachmadita & Fauzi, Rokhman. (2021). Implementation of Process Mining to Discover Student Learning Patterns using Fuzzy Miner Algorithm (Case Study: Learning Management System (LMS) Telkom University). 126-131. 10.1109/ICERA53111.2021.9538789.

[10] De Cnudde, Sofie & Claes, Jan & Poels, Geert. (2014). Improving the quality of the Heuristics Miner in ProM 6.2. Expert Systems with Applications. 41. 10.1016/j.eswa.2014.05.055.

[11] Dongen, Boudewijn & Medeiros, Ana & Verbeek, H. & Weijters, A. & Aalst, Wil. (2005). The ProM Framework: A New Era in Process Mining Tool Support. Lecture Notes in Computer Science. 3536. 444- 454. 10.1007/11494744_25

[12] Sarasvananda, Ida Bagus Gede, Retantyo Wardoyo, and Anny Kartika Sari. "The k-means clustering algorithm with semantic similarity to estimate the cost of hospitalization." IJCCS (Indonesian Journal of Computing and Cybernetics Systems) 13.4 (2019): 313-322.

[13] J. Carmona, B. v. Dongen, A. Solti dan M. Weidlich, "“Relating Processes and Models” dalam Conformance Checking , Springer," vol. XIV, no. 270, 2018.

[14] Atif M, Leisch F (2022) clusTransition: An R package for monitoring transition in cluster solutions of temporal datasets. PLOS ONE 17(12): e0278146.

[15] Libor Juhaňák, Jiří Zounek, Lucie Rohlíková, Using process mining to analyze students' quiz-taking behavior patterns in a learning management system, Computers in Human Behavior, Volume 92, 2019, Pages 496-506, ISSN 0747-5632, https://doi.org/10.1016/j.chb.2017.12.015.

[16] Yutika Amelia Effendi, Fitri Retrialisca, Nania Nuzulita (2019). Conformance Checking to Evaluate Business Process Models using Modified Time-based Heuristics Miner Algorithm. 2019 IEEE Conference on Sustainable Utilization and Development in Engineering and Technologies (CSUDET)

Umargono, Edy, Jatmiko Endro Suseno, and SK Vincensius Gunawan. "K-means clustering optimization using the elbow method and early centroid determination based on mean and median formula." The 2nd International Seminar on Science and Technology (ISSTEC 2019). Atlantis Press, 2020..