

Pembentukan dan Evaluasi Model Proses Berdasarkan Data Proses pada Sistem Informasi Karaoke di PT. XYZ

Afina Lina Nurlaili^{1*}, Agung Mustika R², Eristya Maya S³, Kartika Maulida H⁴

^{1,2}Informatika, UPN "Veteran" Jawa Timur

²Agung.mustika.if@upnjatim.ac.id

³Sistem Informasi, UPN "Veteran" Jawa Timur

³maya.si@upnjatim.ac.id

⁴Sains Data, UPN Veteran Jawa Timur

⁴kartika.maulisa.ds@upnjatim.ac.id

*Corresponding author email: ¹afina.lina.if@upnjatim.ac.id

Abstrak— Efisiensi merupakan salah satu focus utama dalam sebuah proses bisnis. Dalam meningkatkan efisiensi, sistem informasi memiliki peran untuk mengatur proses bisnis di dalamnya sesuai dengan Standar Operasional dan Prosedur (SOP) untuk mencapai efisiensi yang maksimum. Namun dalam praktiknya, terdapat perbedaan dari SOP dengan proses bisnis yang dijalankan. Hal ini menyebabkan model proses yang terbentuk dengan SOP berbeda. Maka dari itu, perlu adanya evaluasi agar tidak berdampak negatif terhadap proses bisnis utama. Pembentukan model proses dilakukan dengan menggunakan metode process mining. Pada penelitian ini, menggunakan metode Alpha dan Algoritma Genetika. Hasil evaluasi menunjukkan bahwa model yang terbentuk menggunakan algoritma Alpha memiliki fitness 0,94 sedangkan Algoritma Genetika 0.98. Hal ini karena model yang terbentuk berasal dari algoritma Alpha tidak dapat menangani kasus perulangan pendek dan aktivitas tersembunyi.

Kata Kunci— process mining, discovery, alpha, algoritma genetika, fitness.

I. PENDAHULUAN

Efisiensi merupakan isu terkini dalam perkembangan teknologi. Berbagai macam peningkatan teknologi dilakukan dengan tujuan efisiensi terutama pada aspek waktu dan biaya. Salah satu dari teknologi yang sering digunakan adalah sistem informasi [1]. Sistem informasi memungkinkan peningkatan efisiensi dengan menggabungkan aspek teknologi informasi dan aktivitas pengguna.

Sistem informasi sering digunakan perusahaan untuk mendukung pengambilan keputusan serta untuk mengatur proses bisnis sesuai dengan Standar Operasional dan Prosedur (SOP) [2]. Dalam hal ini, sistem informasi memuat aktivitas-aktivitas yang harus dijalankan pengguna dari proses bisnis tersebut dan dicatat dalam data proses atau *event log* [3]. Event log yang tercatat dalam sistem informasi atau aktual dengan SOP. Namun dalam implementasinya, seringkali terjadi proses bisnis yang terdapat dalam event log berbeda dengan proses bisnis SOP. Hal ini dapat membuka kesempatan untuk berbuat curang (*fraud*) [4]. Oleh karena itu, diperlukan suatu metode untuk mengetahui kesesuaian proses bisnis aktual dengan SOP dengan menggunakan *process mining*.

Process mining merupakan gabungan dari analisis di bidang (*data mining* dan *machine learning*) dengan analisis pada model proses [5]. Beberapa metode dalam bidang data mining seringkali digunakan untuk menemukan model proses dan juga menganalisis kualitas model proses [6]. Urutan jalannya process mining dapat dilihat pada Gbr.1. Pada Gbr.1 **Error! Reference source not found.** event log dibentuk menjadi model proses memiliki tiga tahap, yaitu *Discovery* (Penemuan Model Proses), *Conformance Checking* (Pemeriksaan Kualiatas), dan *Enhancement* (Peningkatan Kualitas).

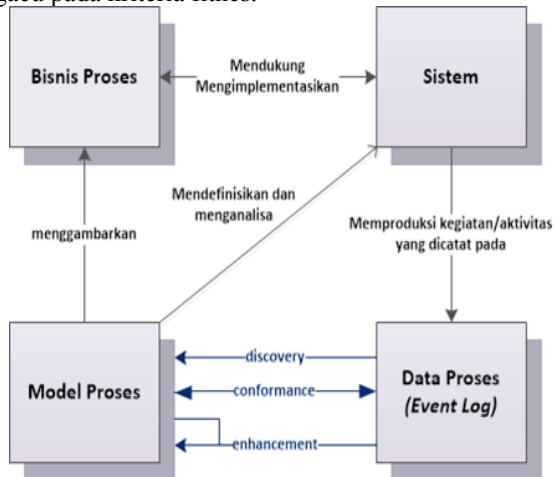
Discovery merupakan pembentukan model proses dari penggalan informasi *event log* yang terekam [7]. Hal ini dilakukan untuk memudahkan pengguna mengetahui proses bisnis yang berjalan. Selain itu, *discovery* juga dapat mempercepat proses pendeteksian kesesuaian antara model proses yang terbentuk dengan SOP. Dalam penemuan model proses, salah satu permasalahan adalah banyaknya jumlah aktivitas dan setiap aktivitas memiliki relasi dengan banyak aktivitas lain sehingga model proses yang terbentuk menjadi rumit. Inilah yang disebut dengan *spaghetty process*. Ada berbagai macam algoritma *discovery* diantaranya Alpha, Alpha +, dan Alpha ++ [8]. Kemudian untuk mengatasi permasalahan *spaghetty process*, dikembangkan algoritma Heuristic, Fuzzy, dan Genetic Miner [9]. Akan tetapi algoritma tersebut memiliki kompleksitas tinggi.

Tahap kedua adalah *conformance checking* yang bertujuan untuk mengukur kualitas model proses yang terbentuk [10]. Pengukuran kualitas model proses diukur menggunakan empat kriteria, yaitu fitness, presisi, simplicity dan generalisasi [11]. Fitness mengukur seberapa besar proses pada *event log* dapat tergambarkan di model proses. Presisi mengukur seberapa besar trace yang terbentuk dari model proses tergambarkan dalam event log. Simplicity mengukur kesederhanaan model proses tanpa menghilangkan proses dari event log. Sedangkan, generalisasi menyatakan bahwa suatu model proses seharusnya menunjukkan generalisasi dari contoh proses yang terlihat pada *event log*.

Tahap terakhir adalah *enhancement* (peningkatan kualitas) [12]. Pada tahap ini, dilakukan peningkatan kualitas model proses yang telah dibentuk oleh algoritma *discovery*. Kegiatan yang dilakukan pada tahap ini dibedakan menjadi dua, yaitu

memperbaiki model proses dan yang kedua menambahkan informasi yang terkait dengan model proses.

Paper ini membandingkan algoritma penemuan model proses menggunakan Alpha dan Algoritma Genetika (Genetic Miner). Kasus yang digunakan untuk pembandingan adalah kasus yang memiliki aktivitas berulang dan aktivitas tersembunyi. Pengukuran kualitas model proses diukur dengan mengacu pada kriteria fitness.



Gbr. 1 Tahap-tahap dalam Process Mining

II. TINJAUAN PUSTAKA

Tinjauan Pustaka menjelaskan istilah yang terkait dengan penelitian ini.

A. Data Proses

Data proses atau *event log* merupakan kumpulan kegiatan atau aktivitas-aktivitas yang berhubungan dalam suatu proses bisnis[13]. Data proses tercatat atau terekam dalam sistem informasi sehingga dapat dilakukan analisis. Contoh dari data proses dapat dilihat pada Gbr. 2. Di dalam data proses terdapat beberapa istilah yang sering digunakan untuk dijadikan sebagai bahan acuan analisis, yaitu case, trace, activity, time, resource. Case adalah setiap data proses mencatat semua rangkaian aktivitas yang berjalan dan setiap rangkaian aktivitas tersebut dimulai dari aktivitas awal sampai aktivitas akhir. Trace adalah kumpulan case yang memiliki proses bisnis yang sama. Activity adalah aktivitas yang dilaksanakan pada proses bisnis. Time adalah waktu pelaksanaan yang dicatat pada setiap aktivitas. Resource adalah orang yang bertugas menyelesaikan setiap aktivitas.

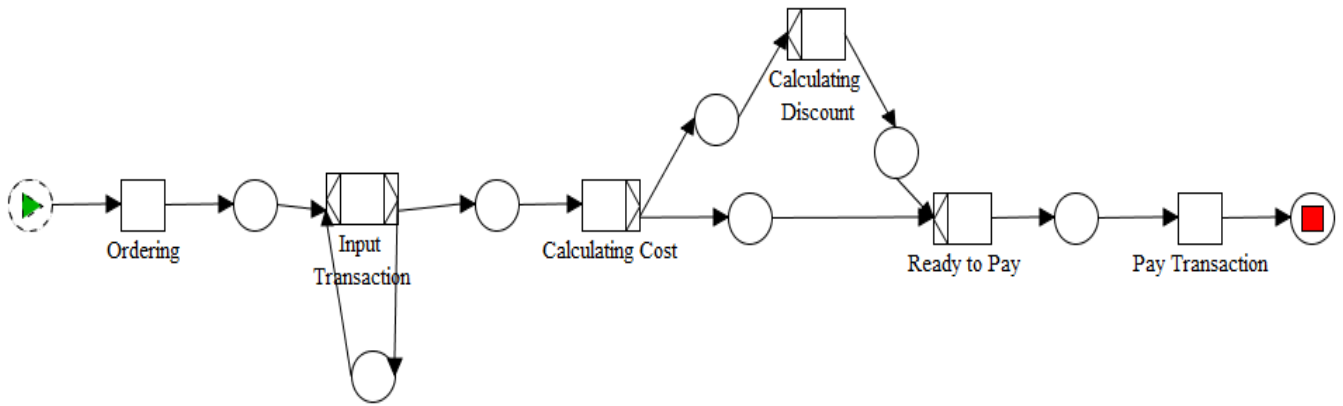
Event log mencatat setiap perlakuan dalam proses bisnis berupa transaksi atau *audit trail* dari sebuah sistem informasi. Event log yang baik adalah jika setiap aktivitas dalam proses bisnis terisi dan tidak menyimpang dari proses bisnis utama.

Trace ID	Case ID	Activity	Time	Resource
1	1	ordering	13/12/2015 08.00	Sisca
1	1	input_transaction	13/12/2015 08.01	Desi
1	1	calculating_cost	13/12/2015 08.02	Desi
1	1	calculating_discount	13/12/2015 08.03	Desi
1	1	ready_to_pay	13/12/2015 08.04	Desi
1	1	pay_transaction	13/12/2015 08.05	Sisca
2	2	ordering	13/12/2015 08.06	Paul
2	2	input_transaction	13/12/2015 08.07	Desi
2	2	input_transaction	13/12/2015 08.08	Desi
2	2	calculating_cost	13/12/2015 08.09	Desi
2	2	ready_to_pay	13/12/2015 08.10	Desi
2	2	pay_transaction	13/12/2015 08.11	Paul
1	3	ordering	13/12/2015 08.00	Jesica
1	3	input_transaction	13/12/2015 08.01	Gani
1	3	calculating_cost	13/12/2015 08.02	Gani
1	3	calculating_discount	13/12/2015 08.03	Gani
1	3	ready_to_pay	13/12/2015 08.04	Gani
1	3	pay_transaction	13/12/2015 08.05	Jesica

Gbr. 2 Contoh Event Log

TABEL I
RELASI DALAM PETRI NET

Nama Relasi	Notasi	Bentuk Relasi
Sequence	Causal (\rightarrow), $A \rightarrow B$	
AND Split	Parallel (\parallel), $B \rightarrow C$, $B \rightarrow D$, dan $C \parallel D$	
AND Join	Parallel (\parallel), $C \rightarrow E$, $D \rightarrow E$, dan $C \parallel D$	
XOR Split	Unrelated ($\#$), $E \rightarrow F$, $E \rightarrow G$, dan $F \# G$	
XOR Join	Unrelated ($\#$), $F \rightarrow H$, $G \rightarrow H$, dan $F \# G$	



Gbr. 3 Proses Bisnis Penyewaan Ruang Karaoke

B. Kakas Bantu

Kakas bantu yang digunakan pada penelitian ini, diantaranya adalah Yet Another Workflow Language (YAWL) dan Process Mining framework (ProM). YAWL adalah sebuah kakas bantu yang digunakan untuk pembuatan model proses dan mensimulasikan workflow dari proses bisnis. Kakas bantu YAWL juga dapat menghasilkan data proses dari simulasi yang telah dijalankan dari suatu proses bisnis untuk selanjutnya dilakukan analisis melalui ProM. Kakas bantu ProM digunakan untuk pembuatan model proses secara otomatis dan melakukan penilaian terhadap kualitas model proses. ProM menyediakan platform bagi pengguna dan pengembang algoritma process mining yang mudah digunakan dan mudah dikembangkan.

C. Algoritma Process Mining

Dalam pengerjaan penelitian ini, digunakan dua metode sebagai pembandingan, yaitu:

1) *Algoritma Alpha*: adalah algoritma yang paling umum digunakan dalam penemuan model proses dan menjadi dasar untuk proses analisis yang menghasilkan pola aktivitas pada sebuah case. Algoritma Alpha dalam penggunaan event log hanya menggunakan urutan aktivitas sedangkan atribut lain seperti timestamp dan resource tidak diperhatikan.

Alpha dalam penggambaran model proses menggunakan Petri net. Petri net adalah grafik dua arah yang terdiri dari place, transisi, dan garis panah yang menghubungkan keduanya. Fungsi dari place adalah sebagai penanda input atau output suatu transisi sedangkan transisi merupakan aktivitas yang ada pada proses bisnis. Beberapa relasi pada Petri net dapat dilihat pada Tabel I yang terdiri dari beberapa relasi sebagai berikut.

- Follows ($>$) apabila terdapat tuple (D,E) saling terhubung
- Causal (\rightarrow) apabila terdapat aktivitas $D > E$ dan $E! > D$
- Parallel (\parallel) apabila terdapat aktivitas $D > E$ dan $E > D$
- Unrelated ($\#$) apabila terdapat aktivitas $B! > B$ dan $B! > A$

Tahap pengerjaan pembentukan Petri net dari event log sebagai berikut:

- Membuat kelompok transisi (aktivitas) pada Petri net (T_N)

$$T_N = \{ t \in T \mid \exists \sigma \in w, t \in \sigma \}$$

- Membuat kelompok transisi output dari source place pada Petri net (T_I)

$$T_I = \{ t \in T \mid \exists \sigma \in w, t \in \text{first}(\sigma) \}$$

- Membuat kelompok transisi input dari sink place pada Petri net (T_O)

$$T_O = \{ t \in T \mid \exists \sigma \in w, t \in \text{last}(\sigma) \}$$

- Proses penentuan place dari Petri net terjadi pada tahap keempat dan kelima. Transisi yang berelasi secara causal disusun pada tahap keempat. Maka untuk setiap tuple (D,E) transisi di kelompok A akan berelasi secara *causal* pada semua transisi di kelompok B serta tidak ada relasi *follows* ($>$) atau antara tuple (A,B) tidak saling terhubung.

$$\begin{aligned} X_N &= \{ (D,E) \mid D \subseteq T_N \wedge E \\ &\subseteq T_N \wedge \forall d \in D \forall e \in E d \rightarrow_N e \\ &\wedge \forall d_1, d_2 \in D d_1 \#_N d_2 \\ &\wedge \forall e_1, e_2 \in E e_1 \#_N e_2 \end{aligned}$$

- Menentukan banyaknya secara pasti places yang dimiliki oleh Petri net

$$\begin{aligned} Y_N &= \{ (D,E) \in X_N \mid \forall D', E' \in X_N D \subseteq D' \wedge E \subseteq E' \\ &\Rightarrow (D,E) = (D',E') \} \end{aligned}$$

- Places yang teridentifikasi pada tahap kelima dibuat

$$P_N = \{ p_{D,E} \mid (D,E) \in Y_N \} \cup \{ i_N, o_N \},$$

- Setiap places dihubungkan dengan transisi input atau transisi output yang sesuai

$$\begin{aligned} F_N &= \{ (d, p_{(D,E)}) \mid (D,E) \in Y_N \wedge d \\ &\in D \} \cup \{ (p_{(D,E)}, e) \mid (D,E) \in Y_N \wedge e \\ &\in E \} \cup \{ (i_N, t) \mid t \\ &\in T_1 \} \cup \{ (t, o_N) \mid t \in T_0 \} \end{aligned}$$

2) *Algoritma Genetika*: adalah metode pencarian adaptif yang meniru proses evolusi[14]. Algoritma ini dimulai dengan populasi awal individu dapat dilihat pada Gbr. 4 dan Tabel II. Setiap individu dihitung nilai fitnessnya untuk diketahui kualitasnya. Dalam process mining, individu adalah

kemungkinan model proses dan fitness berfungsi mengevaluasi seberapa baik perilaku individu dalam event log. Populasi berkembang biak dengan memilih individu yang paling cocok dan menghasilkan individu baru dengan menggunakan operator genetic miner yaitu *crossover* (menggabungkan dua atau lebih individu) dan *mutation* (memodifikasi individu secara acak). Selain itu, pada Algoritma Genetic Miner juga dilakukan penyalinan sejumlah individu terbaik dalam populasi saat ini (elit) ke populasi berikutnya. Dengan cara ini, dapat dipastikan bahwa individu terbaik yang ditemukan akan disimpan untuk populasi berikutnya.

D. EVALUASI KUALITAS MODEL PROSES

Dalam penelitian ini menggunakan kriteria fitness sebagai pembandingan kualitas model proses. Fitness dapat dinyatakan sebagai

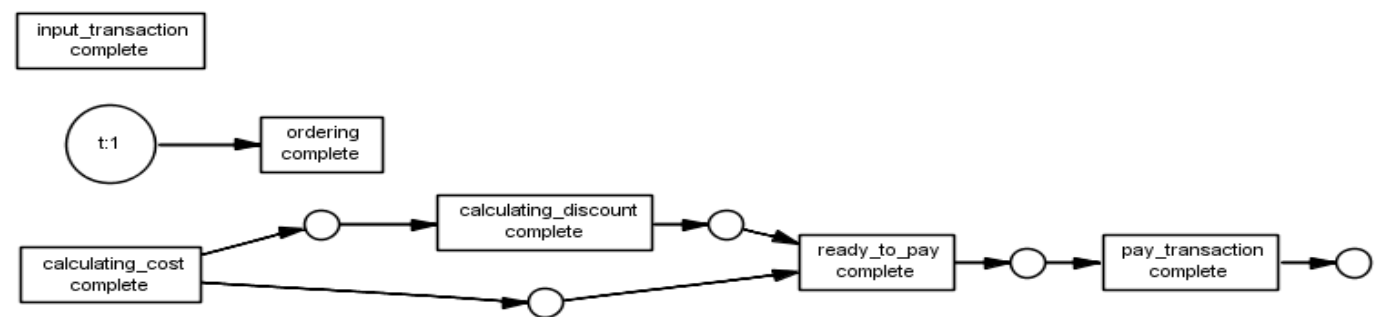
$$X_f = \frac{C_{EL}}{L_{EL}} \quad (1)$$

dimana C_{EL} adalah jumlah case dalam *event log* yang dapat dibentuk menjadi model proses sedangkan L_{EL} adalah jumlah case di dalam *event log*. Sebuah model proses dikatakan memiliki *fitness* sempurna jika semua jalannya proses bisnis (*trace*) di *event log* dapat diwakili oleh model proses dari awal sampai akhir. Sebaliknya, jika banyak *trace* di *event log* tidak dapat diwakili oleh proses model dari awal sampai akhir, maka proses model disebut memiliki fitness yang buruk.

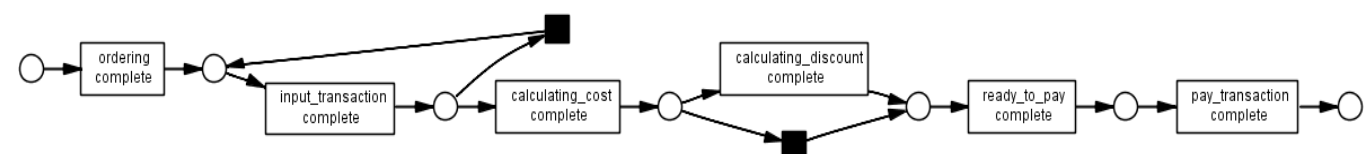
III. HASIL DAN PEMBAHASAN

A. Data Proses

Dalam penelitian ini menggunakan data proses dari Proses Reservasi Room Karaoke yang terdiri dari 3 *trace* dan 30 case setiap case terdiri dari 6 aktivitas. Relasi yang ada dalam event log adalah *sequence* dan *parallel*. Untuk lebih jelas dapat dilihat pada Tabel I.



Gbr. 5 Model Proses Terbentuk melalui Algoritma Alpha

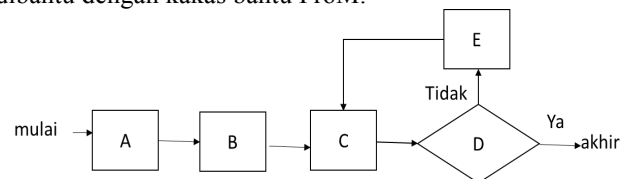


Gbr. 6 Model Proses Terbentuk melalui Algoritma Genetika (Genetic Miner)

B. Pengujian

Dalam pengujian ini menggunakan dua algoritma discovery sebagai pembandingan. Proses bisnis utama yang digunakan dapat dilihat pada Gbr. 3. Pada Gbr. 3 terlihat bahwa proses bisnis diawali dengan aktivitas *ordering* dilanjutkan dengan aktivitas *input_transaction* dimana pada aktivitas ini merupakan aktivitas perulangan pendek. Aktivitas selanjutnya adalah *calculating_cost* yang memiliki relasi XOR pada aktivitas *calculating_discount* dan aktivitas *ready_to_pay*. Aktivitas terakhir adalah aktivitas *pay_transaction*. Gbr. 5 menunjukkan hasil dari Alpha sedangkan Gbr. 6 menunjukkan hasil dari Genetic Miner. Model proses yang terbentuk pada Gbr. 6 memiliki relasi *sequence*, *parallel XOR* dan juga memiliki perulangan pendek dan aktivitas tersembunyi.

Setelah dilakukan penemuan model proses, dalam penelitian ini dilakukan evaluasi *conformance checking* dengan membandingkan *fitness* pada model proses oleh algoritma Alpha dan Algoritma Genetika. Pengukuran fitness dibantu dengan kaskas bantu ProM.



Gbr. 4 Tahap Genetic Miner

TABEL II
DESKRIPSI PROSES GENETIC MINER

Kode	Deskripsi
A	Memindai event log
B	Membuat populasi awal
C	Menghitung fitness dari setiap individu dalam populasi
D	Berhenti dan ambil individu terbaik
E	Membuat populasi berikutnya menggunakan operator genetika dan elit

Parameter fitness pada (1) digunakan sebagai evaluasi kinerja. Hasil dari perbandingan kedua algoritma dapat dilihat pada Tabel III. Berdasarkan hasil simulasi, algoritma genetika memiliki nilai fitness lebih tinggi dibanding algoritma alpha. Hasil ini disebabkan oleh kemampuan algoritma genetika untuk menemukan perulangan pendek (*short loop*) dan aktivitas tersembunyi (*hidden task*).

TABEL III
HASIL PENGUKURAN KRITERIA FITNES

Alpha	Genetic Miner
0.94	0.98

IV. KESIMPULAN

Algoritma genetic miner terbukti memiliki performa fitness yang lebih baik dibandingkan algoritma alpha dengan selisih fitness 0,04. Performa ini disebabkan oleh kemampuan algoritma genetic miner untuk menangani perulangan pendek dan aktivitas tersembunyi dimana kemampuan ini tidak dimiliki oleh algoritma alpha.

UCAPAN TERIMA KASIH

Terimakasih kami ucapkan kepada segenap pihak yang memberikan dukungan penuh dalam penyelesaian naskah Seminar Nasional Informatika Bela Negara (SANTIKA). Semoga naskah ini berkontribusi dalam pengembangan keilmuan khususnya dibidang teknologi informasi dan pendidikan.

REFERENSI

- [1] C. S. Wahyuni, N. Y. Setiawan, and I. Aknuranda, "Pemodelan dan Evaluasi Proses Bisnis Berdasarkan Hasil Ekstraksi Event Log dengan Menerapkan Process Mining pada PT . Kutai Timber Indonesia Kota Probolinggo," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 9, pp. 3087–3094, 2018.
- [2] H. Rizqifaluthi and M. A. Yaqin, "Process Mining Akademik Sekolah menggunakan RapidMiner," *Matics*, vol. 10, no. 2, p. 47, 2019, doi: 10.18860/mat.v10i2.5158.
- [3] R. Sarno, W. A. Wibowo, Y. A. Effendi, and K. R. Sungkono, "Determining Process Model Using Time-Based Process Mining and Control-Flow Pattern," vol. 14, no. 1, 2016, doi: 10.12928/TELKOMNIKA.v14i1.3257.
- [4] D. Rahmawati, M. A. Yaqin, and R. Sarno, "Fraud detection on event logs of goods and services procurement business process using Heuristics Miner algorithm," in *2016 International Conference on Information Communication Technology and Systems (ICTS)*, 2016, pp. 249–254, doi: 10.1109/ICTS.2016.7910307.
- [5] A. Rozinat and W. M. P. van der Aalst, "Conformance checking of processes based on monitoring real behavior," *Inf. Syst.*, vol. 33, no. 1, pp. 64–95, 2008, doi: 10.1016/j.is.2007.07.001.
- [6] R. Sarno, Y. A. Effendi, and F. Haryadita, "Modified Time-Based Heuristics Miner for Parallel Business Processes," *IRECOS (International Rev. Comput. Software)*, vol. 11, no. 3, pp. 249–260, 2016, doi: <https://doi.org/10.15866/irecos.v11i3.8717>.
- [7] A. Vahedian Khezerlou and S. Alizadeh, "A new model for discovering process trees from event logs," *Appl. Intell.*, vol. 41, no. 3, pp. 725–735, 2014, doi: 10.1007/s10489-014-0564-7.
- [8] A. Rozinat, R. S. Mans, M. Song, and W. M. P. van der Aalst, "Discovering colored Petri nets from event logs," *Int. J. Softw. Tools Technol. Transf.*, vol. 10, no. 1, 2008, doi: 10.1007/s10009-007-0051-0.
- [9] M. Van Eck, "Alignment-based Process Model Repair and its Application to the Evolutionary Tree Miner," no. December, 2013.
- [10] A. Adriansyah, B. F. Van Dongen, and W. M. P. Van Der Aalst, "Conformance checking using cost-based fitness analysis," *Proc. - IEEE Int. Enterp. Distrib. Object Comput. Work. EDOC*, pp. 55–64, 2011, doi: 10.1109/EDOC.2011.12.
- [11] J. C. A. M. Buijs, B. F. Van Dongen, and W. M. P. Van Der Aalst, "On the role of fitness, precision, generalization and simplicity in process discovery," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7565 LNCS, no. PART 1, pp. 305–322, 2012, doi: 10.1007/978-3-642-33606-5_19.
- [12] W. M. P. Van Der Aalst, *Process Mining Discovery, Conformance and Enhancement of Business Processes*. Germany: Springer, 2011.
- [13] J. C. a M. Buijs, "Flexible Evolutionary Algorithms for Mining Structured Process Models," 2014, p. 345, 2014, doi: 10.6100/IR780920.
- [14] W. M. P. Van Der Aalst, A. K. A. De Medeiros, and A. J. M. M. Weijters, *Genetic process mining*, vol. 3536. 2005.