ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

KOMPLEKSITAS PROSES DAN NON-KONFORMITAS SEBAGAI FAKTOR KETERLAMBATAN PADA ALUR KERJA *PURCHASE-TO-PAY* LENGKAP: PENDEKATAN MACHINE LEARNING

Lucia Sri Istiyowati✉, Pratiwi Rachmadi, Nani Krisnawaty Tachjar
Sistem Informasi, Institut Keuangan-Perbankan dan Informatika Asia Perbanas, Jakarta, Indonesia
Email: lucia.istyowati@perbanas.id

DOI: https://doi.org/10.46880/jmika.Vol10No2.pp801-809

ABSTRACT

Invoice processing is a crucial component of a financial information system, particularly the accounts payable module integrated with an Enterprise Resource Planning (ERP) system. Delays in invoice processing are common, leading to financial penalties. This study employs a data-driven approach by integrating process mining and machine learning using the XGBoost algorithm to analyze the factors contributing to delays in the purchase-to-pay (P2P) workflow. Event log data from the BPI Challenge 2019 public dataset and refined to 3,471 complete cases. Process discovery and performance analysis were performed using the Disco tool, while an XGBoost binary classification model was developed with 16 process-level features extracted from the event logs. The model achieved an AUC-ROC of 0.6290, with a 5-fold cross-validation score of 0.6326 and a standard deviation of 0.0218. Comparisons with baseline models (Logistic Regression AUC 0.6291, Random Forest AUC 0.6374) reveal performance convergence at an AUC-ROC of approximately 0.63, indicating that the performance ceiling is driven by the dataset's inherent signal rather than the choice of algorithm. Feature importance analysis identified conformity to the "happy path" (0.329), process complexity (0.215), and resource diversity (0.080) as the dominant factors contributing to delays, whereas document completeness played an insignificant role. These findings suggest that in a complete P2P workflow, the cause of delay lies in the characteristics of process execution. The results of this study provide direction for developing accounts payable information systems with a focus on complexity reduction and real-time compliance monitoring.

Keywords: Accounts Payable, Invoice Processing, Machine Learning, Process Mining.

ABSTRAK

Pemrosesan faktur merupakan komponen penting dari sistem informasi keuangan, khususnya modul utang usaha yang terintegrasi dengan sistem Perencanaan Sumber Daya Perusahaan (ERP). Keterlambatan dalam pemrosesan faktur sering terjadi, yang menyebabkan denda finansial. Studi ini menggunakan pendekatan berbasis data dengan mengintegrasikan penambangan proses dan pembelajaran mesin menggunakan algoritma XGBoost untuk menganalisis faktor-faktor yang berkontribusi terhadap keterlambatan dalam alur kerja pembelian hingga pembayaran (the purchase-to-pay/P2P). Data log peristiwa dari dataset publik BPI Challenge 2019 disaring menjadi 3.471 kasus lengkap. Penemuan proses dan analisis kinerja dilakukan menggunakan alat Disco, sementara model klasifikasi biner XGBoost dikembangkan dengan 16 fitur tingkat proses yang diekstrak dari log peristiwa. Model tersebut menghasilkan AUC-ROC sebesar 0,6290 dengan validasi silang 5-fold sebesar 0,6326 dan deviasi standar sebesar 0,0218. Perbandingan dengan baseline models (Logistic Regression AUC 0,6291, Random Forest AUC 0,6374) menunjukkan konvergensi kinerja pada ~0,63 AUC-ROC, mengindikasikan bahwa batas kinerja disebabkan oleh sinyal inheren dataset, bukan pemilihan algoritma. Analisis kepentingan fitur mengidentifikasi konformitas terhadap happy path (0,329), kompleksitas proses (0,215), dan keragaman resource (0,080) sebagai faktor dominan. sementara kelengkapan dokumen memberikan kontribusi yang tidak signifikan. Temuan ini menunjukkan bahwa dalam alur kerja P2P yang lengkap, penyebab penundaan terletak pada karakteristik eksekusi proses. Hasil penelitian ini memberikan arahan untuk mengembangkan sistem informasi utang usaha dengan fokus pada pengurangan kompleksitas dan pemantauan kepatuhan secara real-time.

Kata Kunci: Utang usaha, Pembelajaran Mesin, Pemrosesan Faktur, Penambangan Proses.

PENDAHULUAN

Pemrosesan faktur merupakan komponen penting dari sistem informasi akuntansi, khususnya manajemen utang usaha. Alur kerja pemrosesan faktur yang efisien memainkan peran vital dalam manajemen arus kas, pembayaran tepat waktu kepada pemasok,

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

menjaga hubungan pembeli-pemasok, dan mengendalikan keuangan organisasi(Flynn & Li, 2023; Sánchez-Rebull et al., 2020). Meskipun sistem *Enterprise Resource Planning* (ERP) telah diadopsi secara luas, banyak organisasi masih mengalami keterlambatan dalam pemrosesan faktur akibat kompleksitas alur kerja, ketidaksesuaian dokumen, dan kebutuhan penanganan pengecualian (Berti & Van Der Aalst, 2022; W. Van Der Aalst, 2016; W. M. P. Van Der Aalst & Carmona, 2022).

Penambangan proses (*process mining*) telah berkembang menjadi pendekatan berbasis data untuk menganalisis proses bisnis melalui data event log yang dihasilkan sistem informasi. Pada alur kerja purchase-to-pay (P2P), penambangan proses memungkinkan organisasi menemukan perilaku proses yang sebenarnya terjadi, mengidentifikasi bottleneck, dan menganalisis konformitas terhadap prosedur standar (Berti et al., 2025; Emamjome et al., 2019). Berbagai penelitian sebelumnya menunjukkan nilai penambangan proses untuk analisis kinerja P2P, khususnya dalam mengidentifikasi varian yang menyimpang dari alur kerja standar (Badakhshan et al., 2022).

Sejalan dengan itu, teknik *machine learning* telah diterapkan pada monitoring proses prediksi untuk memprediksi hasil proses seperti durasi kasus dan kemungkinan keterlambatan. Hasil studi serupa menunjukkan bahwa nilai AUC-ROC berkisar antara 0,70 hingga 0,85 untuk *outcome-oriented process monitoring* pada berbagai proses administratif(Teinemaa et al., 2019). Namun demikian, studi yang menggabungkan kasus pada tahap penyelesaian yang berbeda dapat menghadapi tantangan dalam mengisolasi faktor keterlambatan dari sinyal progres kasus.

Penelitian ini mengatasi gap tersebut dengan menganalisis faktor keterlambatan tingkat proses secara khusus pada kasus P2P lengkap. Dengan menyaring dataset dengan hanya menyertakan kasus yang telah mencapai tahap penyelesaian faktur (*invoice clearance)*, penelitian ini menghilangkan efek pembaur (*confounding effect*) dari ketidaklengkapan kasus terhadap klasifikasi keterlambatan. Berdasarkan pemahaman konteks tersebut, tujuan penelitian ditetapkan sebagai berikut: (1) mengidentifikasi faktor tingkat proses yang paling berkontribusi terhadap keterlambatan pada kasus P2P lengkap, dan (2) merumuskan implikasi manajerial berbasis pola feature importance yang dapat menjadi arahan bagi optimalisasi proses accounts payable.

Kontribusi penelitian ini terdiri dari tiga hal. Pertama, mengusulkan kerangka analisis tersaring (*filtered analysis framework*) yang mengisolasi kasus lengkap dari populasi campuran dalam dataset penambangan proses. Kedua, penelitian menyediakan peringkat empiris *feature importance* yang membedakan faktor keterlambatan pada alur kerja yang sudah lengkap secara operasional. Ketiga, penelitian menerjemahkan pola *feature importance* menjadi arahan manajerial yang dapat ditindaklanjuti untuk optimalisasi proses utang usaha (*accounts payable*).

TINJAUAN PUSTAKA

Penambangan Proses pada Alur Kerja *Purchase-to-Pay*

Penelitian penambangan proses pada alur kerja P2P sebagian besar berfokus pada analisis deskriptif dan identifikasi bottleneck. Badakhshan dkk. menunjukkan bahwa penambangan proses dapat memperbaiki *throughput time* pada proses P2P dengan mengungkap varian tersembunyi dan rework loops (Badakhshan et al., 2022). Berti dkk. mengusulkan *object-centric process mining* sebagai perluasan yang menangkap data P2P dari beragam perspektif, termasuk faktur, pesanan pembelian, dan tanda terima barang (Berti et al., 2025). Pada konteks utang usaha secara khusus, Jessen dkk. menunjukkan bagaimana penambangan proses berorientasi pengguna (*user-oriented process mining)* mendukung identifikasi pola keterlambatan pembayaran (Jessen et al., 2025).

Meskipun berbagai kemajuan telah dicapai, sebagian besar studi penambangan proses P2P menghasilkan wawasan deskriptif tanpa kapabilitas prediktif. Identifikasi fitur umumnya bergantung pada analisis manual varian proses, bukan pada metode atribusi fitur berbasis data. Hal ini membuka peluang bagi pendekatan *machine learning* untuk melengkapi temuan penambangan proses dengan analisis *feature importance* kuantitatif.

Machine Learning untuk *Predictive Process Monitoring*

Pemantauan proses prediktif *(Predictive process monitoring*/PPM menerapkan *machine learning* untuk memprediksi hasil proses seperti sisa waktu, aktivitas berikutnya, atau luaran kasus. PPM merupakan perluasan dari teknik penambangan proses yang bersifat prospektif/prediktif. Prediksi ini mencakup alur peristiwa berikutnya (*next event*), hasil akhir proses, hingga kinerja seperti durasi waktu. Namun, pendekatan berbasis urutan (*sequence-based approaches*) umumnya memerlukan volume data yang lebih besar dan menghasilkan keluaran yang relatif lebih sulit diinterpretasikan dibandingkan dengan

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

pendekatan berbasis fitur agregat. (Ceravolo et al., 2024).

Teinemaa dkk. memberikan benchmark komprehensif untuk pendekatan PPM berorientasi *outcome* dan melaporkan nilai AUC-ROC pada rentang 0,70 hingga 0,85 pada beberapa dataset (Teinemaa et al., 2019). Studi terkini mengeksplorasi arsitektur *deep learning* berbasis sekuens, termasuk jaringan *Long Short-Term Memory* (LSTM) (Tax et al., 2016; Weinzierl et al., 2024) dan model transformer (Ni et al., 2023), untuk prediksi aktivitas berikutnya dan luaran.

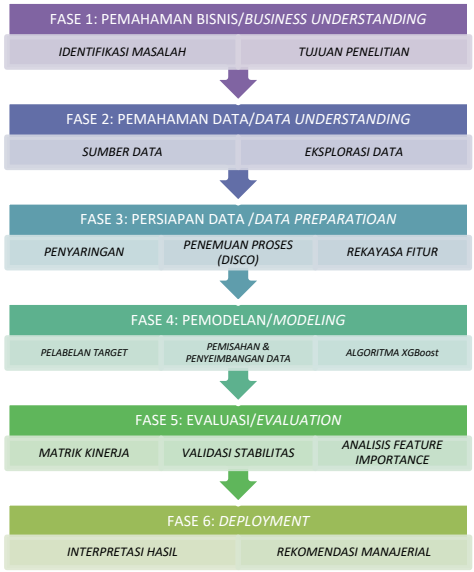
Dalam konteks manajerial yang menuntut wawasan operasional dan rekomendasi yang dapat ditindaklanjuti, fitur agregat pada tingkat kasus yang diturunkan dari artefak penambangan proses menawarkan tingkat interpretabilitas yang lebih tinggi, meskipun disertai kemungkinan penurunan tertentu pada kinerja prediktif. Metode *gradient boosting*, seperti XGBoost, telah menunjukkan kinerja yang kuat dalam berbagai tugas klasifikasi berbasis data tabular serta mampu menghasilkan peringkat kepentingan fitur (*feature importance*) yang dapat digunakan untuk mendukung interpretasi hasil model (Chen & Guestrin, 2016).

Posisi Penelitian

Penelitian ini mengadopsi pendekatan fitur agregat dengan XGBoost pada dataset kasus P2P lengkap. Filtering hanya pada kasus yang telah mencapai penyelesaian faktur mengorbankan ukuran sampel untuk memperoleh kualitas label yang lebih baik dan menghasilkan pola *feature importance* yang lebih dapat ditafsirkan (*interpretable*) untuk pengambilan keputusan manajerial.

METODE PENELITIAN

Penelitian ini menggunakan kerangka kerja *Cross-Industry Standard Process for Data Mining* (CRISP-DM), yang terdiri atas enam tahap, yaitu pemahaman bisnis (*Business Understanding*), pemahaman data (*Data Understanding*), persiapan data (*Data Preparation*), pemodelan (*Modeling*), evaluasi (*Evaluation*), dan penerapan (*Deployment*) (Chapman, Kerber, et al., 2000). Kerangka kerja ini dipilih karena CRISP-DM telah diakui sebagai standar dalam pelaksanaan proyek data mining dan analisis data terapan (Schröer et al., 2021). Selain itu, CRISP-DM menyediakan tahapan yang sistematis, mulai dari pemahaman konteks dan permasalahan bisnis hingga penyusunan rekomendasi penerapan hasil analisis (Schröer et al., 2021). Keseluruhan alur metodologi penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Metodologi Penelitian

Pemahaman Bisnis *(Business Understanding)*

Fase ini bertujuan untuk mengidentifikasi masalah dan menetapkan tujuan penelitian dari perspektif kebutuhan bisnis. Konteks yang menjadi fokus penelitian ini adalah pemrosesan faktur pada modul accounts payable yang terintegrasi dalam sistem *Enterprise Resource Planning* (ERP). Keterlambatan pada alur kerja *purchase-to-pay* (P2P) merupakan permasalahan operasional yang berdampak langsung pada arus kas organisasi, hubungan dengan vendor, dan kepatuhan finansial. Meskipun sistem ERP telah menyediakan otomatisasi pada berbagai tahap alur kerja P2P, keterlambatan pemrosesan tetap sering terjadi dan sulit diidentifikasi akar penyebabnya (Kermani et al., 2024).

Pemahaman Data *(Data Understanding)*

Fase pemahaman data meliputi eksplorasi awal terhadap sumber data yang akan dianalisis. Penelitian ini menggunakan dataset event log publik BPI Challenge 2019 yang berisi data proses purchase-to-pay dari perusahaan manufaktur multinasional yang berkantor pusat di Belanda (van Dongen, 2019). Pemilihan dataset ini didasarkan pada ketersediaan publik, cakupan yang komprehensif terhadap seluruh alur kerja P2P, dan penggunaan yang luas dalam literatur penambangan proses sebagai benchmark data (Augusto et al., 2019; van Dongen, 2019).

Persiapan Data *(Data Preparation)*

Fase Data Preparation mencakup tiga kegiatan utama, yaitu penyaringan dataset, penemuan proses, dan rekayasa fitur. Pada tahap penyaringan dataset yang bertujuan menghilangkan kontaminasi label yang

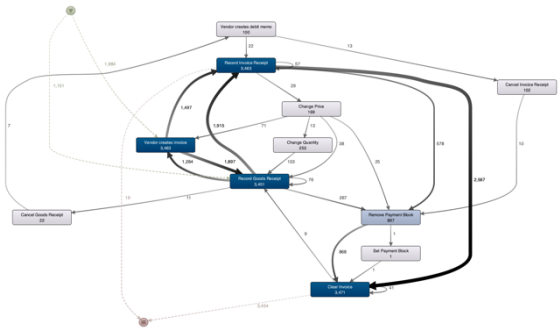
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

berasal dari kasus tidak lengkap. Proses penyaringan mengurangi dataset dari 4.500 menjadi 3.471 kasus lengkap dengan tingkat retensi 77,1 persen. Kasus yang dihapus memiliki durasi pendek (median 5,0 hari) dibandingkan kasus lengkap (median 65,7 hari), yang mengindikasikan bahwa kasus-kasus tersebut terhenti secara prematur dalam jendela observasi.

Pada tahap penemuan proses menggunakan aplikasi Disco (Fluxicon). Proses ini menghasilkan process map yang memvisualisasikan alur aktivitas aktual, frekuensi transisi antar aktivitas, dan performance metrics seperti durasi rata-rata dan median per jalur proses. Analisis ini digunakan untuk mengidentifikasi happy path standar (*Vendor Creates Invoice, Record Goods Receipt, Record Invoice Receipt)* yang menjadi dasar penghitungan *conformance score*. Peta proses yang dihasilkan dari process discovery pada 3.471 kasus lengkap ditunjukkan pada Gambar 2.

Pada tahapan *feature engineering* menghasilkan enam belas fitur tingkat proses dari data event log dan dikelompokkan dalam empat kategori yaitu kompleksitas, konformitas, rework, dan kelengkapan dokumen. Fitur has_clear yang menandai pencapaian Clear Invoice sengaja dikeluarkan karena akan menjadi konstan pada dataset tersaring dan tidak memberikan nilai diskriminatif.



Gambar 2. Peta Proses Hasil Discovery dari 3.471 Kasus P2P Lengkap

Pemodelan *(Modeling)*

Fase Modeling meliputi pelabelan target, pemisahan data, penyeimbangan kelas, dan pelatihan model. Target klasifikasi biner ditentukan berdasarkan total durasi kasus. Kasus dengan durasi melebihi median 65,7 hari diberi label sebagai delayed (delay = 1), sedangkan kasus dengan durasi di bawah atau sama dengan ambang batas diberi label sebagai on-time (delay = 0). Ambang batas ini menghasilkan distribusi kelas mendekati seimbang, yaitu 50,0 persen delayed dan 50,0 persen on-time.

Dataset dibagi menjadi set pelatihan (70 persen, n = 2.429) dan set pengujian (30 persen, n = 1.042) menggunakan stratified random sampling dengan random_state = 42 untuk memastikan reprodusibilitas.

Untuk menghindari pemilihan hyperparameter secara arbitrer, dilakukan *hyperparameter tuning (Florek & Zagdański, 2023)* menggunakan GridSearchCV dengan 5-fold cross-validation pada data pelatihan. Grid search dilakukan pada tiga parameter utama: n_estimators (100, 200, 300), max_depth (4, 6, 8), dan *learning_rate* (0,05, 0,10). Kombinasi optimal adalah n_estimators=100, max_depth=4, dan learning_rate=0,05, dengan best CV AUC-ROC sebesar 0,6411.

Algoritma XGBoost dipilih sebagai model klasifikasi karena kinerjanya yang telah terbukti kuat pada tugas klasifikasi data tabular dan tersedianya kapabilitas feature importance yang dapat diinterpretasi (Chen & Guestrin, 2016). Hyperparameter dikonfigurasi sebagai berikut: jumlah estimator = 100, learning rate = 0,05, kedalaman maksimum pohon = 4 (hasil GridSearchCV), subsample = 0,8, column subsample per pohon = 0,8, dan minimum child weight = 3.

Perbandingan Dasar *(Baseline Comparison)*

Untuk memvalidasi bahwa kinerja moderate model bukan artifak dari pemilihan algoritma, dilakukan perbandingan dengan dua baseline models: *Logistic Regression* (dengan StandardScaler) dan *Random Forest* (n_estimators=300, max_depth=10). Kedua baseline dievaluasi dengan protokol yang sama seperti XGBoost.

Analisis Sensitivitas Ambang Batas

Untuk memvalidasi robustness terhadap pemilihan ambang batas keterlambatan, dilakukan analisis sensitivitas dengan tiga ambang batas berdasarkan persentil durasi kasus: P50 (65.7 hari, delay 50%), P60 (76.1 hari, delay 40%), dan P75 (92.0 hari, delay 25%). Model XGBoost dilatih ulang dengan setiap ambang batas untuk membandingkan stabilitas kinerja.

Evaluasi *(Evaluation)*

Fase Evaluasi mengukur kinerja model dan menganalisis *feature importance*. Kinerja model dievaluasi menggunakan lima metrik standar klasifikasi biner, yaitu *accuracy, precision, recall, F1-score, dan area under the receiver operating characteristic curve* (AUC-ROC). Untuk menilai stabilitas generalisasi model di berbagai partisi data, dilakukan *5-fold stratified cross-validation* pada dataset tersaring, dengan pelaporan mean dan standar deviasi AUC-ROC.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

Analisis *feature importance* dilakukan pada dua tingkat, pertama ranking fitur individual dan kedua agregasi berdasarkan kategori konseptual (kompleksitas, konformitas, rework, kelengkapan dokumen).

***Deployment* (Rekomendasi Implementasi)**

Fase *Deployment* berupa perumusan rekomendasi manajerial berbasis pola *feature importance* untuk mengurangi keterlambatan, yang disajikan secara detail pada Bagian Pembahasan.

HASIL DAN PEMBAHASAN

Kinerja Model

Tabel 1 menyajikan metrik kinerja model pada data pengujian yang terdiri atas 1.042 kasus terpisah (*held-out cases*). Pengklasifikasi XGBoost memperoleh akurasi sebesar 60,08%, presisi 58,43%, recall 69,87%, F1-score 63,64%, dan AUC-ROC 0,6290. Validasi silang lima lipatan menghasilkan rata-rata AUC-ROC sebesar 0,6326 dengan simpangan baku 0,0218. Hasil tersebut menunjukkan bahwa model memiliki kinerja generalisasi yang relatif stabil pada berbagai pembagian data.

Tabel 1. Metrik Kinerja Model

Metrik	Nilai
Akurasi	60,08%
Presisi	58,43%
Recall	69,87%
F1-Score	63,64%
AUC-ROC	0,6290
5-*Fold CV Mean*	0,6326 ± 0,0218

Matriks kebingungan (confusion matrix) menunjukkan bahwa dari 1.042 kasus pengujian, model berhasil mengklasifikasikan dengan benar sebanyak 364 kasus terlambat (true positive) dan 262 kasus tepat waktu (true negative). Namun, model juga menghasilkan 259 positif palsu (false positive) dan 157 negatif palsu (false negative). Nilai recall sebesar 69,87% menunjukkan bahwa model mampu mengidentifikasi dengan benar sekitar tujuh dari sepuluh kasus keterlambatan. Kemampuan ini relevan secara operasional untuk mendukung pengelolaan utang usaha (accounts payable) secara proaktif, meskipun tingkat akurasi keseluruhan model masih tergolong moderat pada 60,08%.

Nilai AUC-ROC yaitu 0,6290 termasuk dalam kategori yang moderat, mencerminkan kompleksitas dalam mengklasifikasikan keterlambatan pada kasus P2P yang telah selesai hanya dengan menggunakan fitur agregat tingkat proses. Tanpa informasi pada tingkat konten, seperti nilai faktur, karakteristik pemasok, atau konteks organisasi, model hanya mengandalkan pola pelaksanaan proses. Dengan demikian, model ini dapat dipandang sebagai tolok ukur awal yang konservatif, tetapi tetap memiliki interpretabilitas yang memadai untuk mengidentifikasi faktor-faktor yang berkaitan dengan keterlambatan.

Tabel 2. Perbandingan Baseline

Model	Accuracy	Precision	Recall	F1	AUC-ROC	CV Mean
Logistic Regression	59,88%	59,74%	60,65%	60,19%	0,629	0,634 ± 0,037
Random Forest	60,84%	59,22%	69,67%	64,02%	0,637	0,623 ± 0,026
XGBoost (tuned)	60,08%	58,43%	69,87%	63,64%	0,629	0,632 ± 0,021

Ketiga algoritma konvergen ke AUC-ROC sekitar 0,63 dengan selisih kurang dari 0,01, menunjukkan bahwa kinerja moderate model bukan disebabkan oleh pemilihan algoritma, melainkan mencerminkan batas inheren dari sinyal proses tingkat kasus dalam dataset.

Analisis Feature Importance

Tabel 3 menyajikan sepuluh fitur dengan tingkat kepentingan tertinggi berdasarkan hasil perhitungan model XGBoost. conformance_score muncul sebagai prediktor tunggal terkuat dengan skor kepentingan sebesar 0,329, diikuti oleh kompleksitas proses (n_unique_activities) sebesar 0,215 dan keragaman resource (n_resources) sebesar 0,080. Ketiga fitur ini secara kumulatif menyumbang sekitar 62% dari total tingkat kepentingan fitur.

Tabel 3. Peringkat Tingkat Kepentingan Fitur (*Feature Importance)*

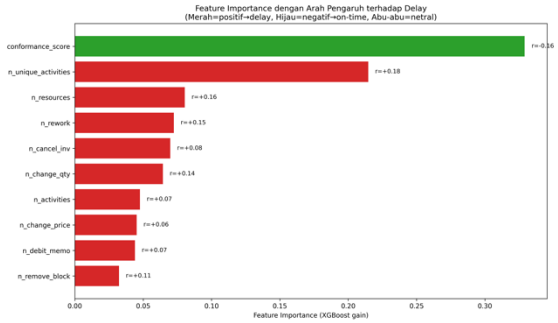
Fitur	Nilai
conformance_score	0,329

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

Fitur	Nilai
n_unique_activities	0,215
n_resources	0,080
n_rework	0,072
n_cancel_inv	0,070
n_change_qty	0,064
n_activities	0,048
n_change_price	0,045
n_debit_memo	0,044
n_remove_block	0,032

Sebanyak enam fitur menunjukkan nilai kepentingan sebesar nol pada model akhir, yaitu doc_completeness, has_vendor_invoice, has_gr, has_inv_receipt, n_cancel_gr, dan n_set_block. Pola ini memiliki makna teoretis yang penting. Karena seluruh kasus dalam dataset hasil penyaringan merupakan kasus lengkap, fitur keberadaan dokumen (has_vendor_invoice, has_gr, dan has_inv_receipt) menjadi hampir konstan sehingga tidak memberikan sinyal diskriminatif. Demikian pula, fitur kelengkapan dokumen (doc_completeness), yang mengagregasikan ketiga indikator dokumen tersebut, tidak memberikan kontribusi terhadap kemampuan prediksi model.



Gambar 3. Feature Importance dengan Arah Pengaruh Terhadap Keterlambatan.

Warna merah pada gambar 3 menunjukkan fitur yang memprediksi delay (korelasi positif), warna hijau menunjukkan fitur yang memprediksi on-time (korelasi negatif). Analisis arah pengaruh mengungkap pola konsisten: dari sepuluh fitur teratas, sembilan berkorelasi positif dengan keterlambatan (nilai tinggi memprediksi delay), sementara hanya conformance_score yang berkorelasi negatif (r=-0,157, nilai tinggi memprediksi on-time). Temuan ini menunjukkan bahwa kepatuhan terhadap alur proses standar (*happy path*) merupakan satu-satunya faktor protektif terhadap keterlambatan pada dataset ini, sementara semua bentuk deviasi (variasi aktivitas, keragaman *resource*, pembatalan, perubahan) merupakan faktor risiko.

Temuan ini menunjukkan bahwa pada kasus P2P yang telah selesai, faktor utama yang membedakan faktur terlambat dan tepat waktu terletak pada cara proses tersebut dijalankan, bukan pada jenis dokumen yang tersedia. Variasi kompleksitas proses, penyimpangan terhadap alur proses standar, dan aktivitas pengerjaan ulang (rework) muncul sebagai sinyal keterlambatan yang paling bermakna.

Analisis Tingkat Kategori

Untuk memberikan gambaran yang lebih umum mengenai pola tingkat kepentingan fitur, seluruh fitur dikelompokkan ke dalam empat kategori berdasarkan makna konseptualnya. Tabel 4 menyajikan hasil agregasi tingkat kepentingan fitur pada setiap kategori. Fitur yang berkaitan dengan pengerjaan ulang (rework), yaitu n_cancel_inv, n_rework, n_change_price, n_change_qty, n_remove_block, dan n_debit_memo, secara kumulatif memberikan kontribusi sebesar 0,328. Fitur kompleksitas (n_activities dan n_unique_activities) memberikan kontribusi sebesar 0,263. Sementara itu, fitur konformitas dan keragaman sumber daya masing-masing berkontribusi sebesar 0,329 dan 0,080. Meskipun lebih kecil, kontribusi kedua kategori tersebut tetap memiliki nilai yang substantif.

Tabel 4. Tingkat Kepentingan Fitur (*Feature Importance)* Berdasarkan Kategori

Kategori	Total Tingkat Kepentingan	Proporsi
Rework	0,328	32,8%
Kompleksitas	0,263	26,3%
Konformitas	0,329	32,9%
Resource	0,080	8,0%

Keempat kategori tersebut secara keseluruhan mencakup seluruh fitur yang memiliki nilai kepentingan lebih dari nol dalam model. Dominasi fitur rework dan kompleksitas dibandingkan fitur yang berkaitan dengan dokumen menegaskan suatu temuan operasional penting: pada kasus P2P yang telah selesai, keterlambatan terutama dipengaruhi oleh variasi dalam pelaksanaan proses, bukan oleh kualitas atau kelengkapan dokumen.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Analisis Sensitivitas

Analisis sensitivitas menunjukkan bahwa kinerja model relatif stabil di berbagai threshold (AUC 0,63-0,68), dengan conformance_score konsisten muncul sebagai top feature. Hasil ini mengkonfirmasi robustness temuan terhadap pemilihan ambang batas keterlambatan.

Tabel 5. Analisis sensitivitas

Threshold	Cut-off (hari)	% Delay	AUC-ROC	Top 3 Features
P50 (median)	65,7	50,0%	0,629	conformance_score, n_unique_activities, n_resources
P60	76,1	40,0%	0,678	conformance_score, has_gr, doc_completeness
P75	92,0	25,0%	0,670	conformance_score, n_unique_activities, n_cancel_inv

Pembahasan

Temuan penelitian ini memberikan implikasi teoretis dan praktis bagi pengembangan kajian penambangan proses serta praktik pengelolaan utang usaha (*accounts payable)*. Secara teoretis, hasil penelitian mendukung hipotesis keterlambatan yang didorong oleh variabilitas (*variability-driven delay hypothesis*) dalam manajemen operasi. Ketika kelengkapan dokumen dikendalikan dengan membatasi analisis hanya pada kasus yang telah selesai, varians residual dalam durasi kasus terutama dijelaskan oleh karakteristik pelaksanaan proses, yaitu jumlah aktivitas berbeda yang dilakukan, penyimpangan alur kerja dari jalur standar, serta frekuensi tindakan korektif seperti pembatalan dan perubahan. Temuan ini sejalan dengan literatur mengenai variabilitas proses yang menyatakan bahwa ketidaksesuaian terhadap proses standar dan aktivitas pengerjaan ulang merupakan sumber utama variasi waktu penyelesaian pada proses bisnis yang terstruktur.

Secara praktis, pola tingkat kepentingan fitur memberikan arahan yang jelas bagi perbaikan proses pengelolaan utang usaha. Organisasi yang berupaya mengurangi keterlambatan pemrosesan faktur perlu memprioritaskan intervensi pada tiga area. Pertama, pengurangan kompleksitas melalui standardisasi alur kerja dan konsolidasi aktivitas dapat mengatasi sumber tunggal terbesar dalam tingkat kepentingan fitur, yaitu n_unique_activities. Kedua, pemantauan konformitas melalui peringatan penyimpangan secara waktu nyata dapat membantu mendeteksi kasus yang menyimpang dari happy path sebelum keterlambatan terakumulasi. Ketiga, analisis akar penyebab terhadap pembatalan dan perubahan faktur dapat mengatasi kategori rework yang mencakup 32,8% dari keseluruhan tingkat kepentingan fitur, hampir setara dengan kontribusi konformitas (32,9%) sebagai faktor protektif utama.

Nilai kepentingan fitur kelengkapan dokumen yang mendekati nol menunjukkan bahwa pada kasus yang secara operasional telah selesai, intervensi terhadap kualitas dokumen hanya memberikan nilai tambah yang terbatas dalam mengurangi keterlambatan. Temuan ini menantang asumsi umum bahwa pengelolaan dokumen merupakan faktor utama dalam meningkatkan efisiensi utang usaha, setidaknya pada kasus yang telah mencapai tahap penyelesaian faktur.

Penting untuk ditegaskan bahwa hasil analisis feature importance dan arah pengaruh dalam penelitian ini menggambarkan hubungan prediktif, bukan hubungan kausal. Misalnya model mengidentifikasi korelasi kuat antara frekuensi *rework* dan keterlambatan, namun temuan ini tidak dapat diinterpretasikan sebagai bukti bahwa aktivitas *rework* merupakan penyebab langsung dari keterlambatan. Hubungan sebaliknya juga sama valid untuk dipertimbangkan, yaitu bahwa proses yang telah mengalami hambatan operasional cenderung memicu aktivitas *rework* sebagai respons korektif. Validasi klaim kausal secara metodologis memerlukan desain eksperimen atau pendekatan *causal inference* dengan asumsi kausal yang eksplisit, yang berada di luar cakupan penelitian ini dan direkomendasikan sebagai arah penelitian selanjutnya.

Konvergensi kinerja tiga algoritma yang berbeda kompleksitas (*Logistic Regression* 0,6291, *Random Forest* 0,6374, XGBoost 0,6290) ke rentang AUC-ROC yang sama (~0,63) memberikan bukti empiris kuat bahwa keterbatasan kinerja bukan disebabkan oleh *underfitting* atau algoritma yang tidak optimal. Sebaliknya, hal ini mencerminkan batas inheren informasi yang dapat diekstrak dari fitur proses tingkat kasus tanpa integrasi fitur konten (misalnya nilai faktur, karakteristik vendor) atau representasi berbasis sekuens.

Penelitian ini memiliki sejumlah keterbatasan. Pertama, analisis hanya menggunakan satu dataset publik yang berasal dari sebuah perusahaan manufaktur

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

di Belanda sehingga generalisasi hasil pada industri dan konteks organisasi lainnya masih terbatas. Kedua, ambang batas keterlambatan ditetapkan secara statistik berdasarkan nilai median, bukan berdasarkan perjanjian tingkat layanan (*service level agreement*) yang spesifik terhadap kebutuhan bisnis. Oleh karena itu, batas tersebut belum tentu sesuai dengan definisi operasional keterlambatan pada setiap organisasi. Ketiga, pendekatan fitur agregat tidak dapat menangkap pola urutan aktivitas yang mungkin mengandung sinyal prediktif tambahan. Keempat, kinerja model belum dievaluasi menggunakan jejak proses parsial (*partial process traces*) sehingga penerapan temuan dalam skenario operasional secara waktu nyata masih terbatas. Kelima, meski SHAP awalnya direncanakan untuk analisis arah pengaruh yang lebih rigorous, masalah kompatibilitas versi teknis mengharuskan penggunaan pendekatan alternatif berbasis korelasi Pearson dan perbandingan mean per class. Analisis SHAP lengkap direkomendasikan untuk penelitian selanjutnya.

KESIMPULAN

Penelitian ini menganalisis faktor-faktor tingkat proses yang berkaitan dengan keterlambatan pemrosesan faktur pada kasus *purchase-to-pay* yang telah selesai dengan menggunakan dataset publik BPI Challenge 2019. Dengan menyaring dataset agar hanya mencakup kasus yang telah mencapai tahap penyelesaian faktur, penelitian ini mengisolasi faktor pendorong keterlambatan dalam alur kerja yang secara operasional telah selesai dari pengaruh perancu akibat ketidaklengkapan kasus.

Pengklasifikasi XGBoost menghasilkan AUC-ROC sebesar 0,6290 dengan kinerja *5-fold cross-validation* yang stabil, yaitu rata-rata 0,6326 dan simpangan baku 0,0218. Analisis tingkat kepentingan fitur mengidentifikasi kompleksitas proses, konformitas terhadap *happy path*, dan frekuensi pembatalan faktur sebagai faktor dominan yang berkaitan dengan keterlambatan. Fitur terkait konformitas memberikan kontribusi sebesar 32,9%, diikuti oleh rework sebesar 32,8% dan kompleksitas sebesar 26,3%. Fitur kelengkapan dokumen memberikan kontribusi yang sangat kecil. Temuan ini menunjukkan bahwa pada kasus yang telah selesai, faktor keterlambatan lebih berkaitan dengan karakteristik pelaksanaan proses daripada kualitas atau kelengkapan dokumen.

Secara metodologis, penelitian ini menunjukkan nilai penting analisis terfilter terhadap kasus lengkap dalam pemantauan proses prediktif. Secara empiris, penelitian ini memberikan bukti berdasarkan tingkat kepentingan fitur yang dapat mengarahkan upaya perbaikan pengelolaan utang usaha pada pengurangan kompleksitas, pemantauan konformitas, dan pencegahan pembatalan faktur. Penelitian selanjutnya perlu memvalidasi temuan ini menggunakan berbagai dataset, mengevaluasi pendekatan pemodelan yang mempertimbangkan urutan aktivitas, serta mengeksplorasi integrasi fitur tingkat konten, seperti nilai faktur dan karakteristik pemasok.

DAFTAR PUSTAKA

Augusto, A., Leno, V., & Reissner, D. (2019). *BPI Challenge 2019 Report: a Purchase-to-Pay Process Analysis*.

Badakhshan, P., Wurm, B., Grisold, T., Geyer-Klingeberg, J., Mendling, J., & Vom Brocke, J. (2022). Creating business value with process mining. *The Journal of Strategic Information Systems*, *31*(4), 101745. https://doi.org/10.1016/j.jsis.2022.101745

Berti, A., Jessen, U., Park, G., Rafiei, M., & Van Der Aalst, W. M. P. (2025). Analyzing interconnected processes: using object-centric process mining to analyze procurement processes. *International Journal of Data Science and Analytics*, *20*(2), 475–497. https://doi.org/10.1007/s41060-023-00427-3

Berti, A., & Van Der Aalst, W. (2022). *OC-PM: Analyzing Object-Centric Event Logs and Process Models*. https://doi.org/10.1007/s10009-022-00668-w

Ceravolo, P., Comuzzi, M., De Weerdt, J., Di Francescomarino, C., & Maggi, F. M. (2024). Predictive process monitoring: concepts, challenges, and future research directions. *Process Science*, *1*(1), 2. https://doi.org/10.1007/s44311-024-00002-4

Chapman, P., Khabaza, T., Kerber, R., & Clinton, J. (2000). *Step-by-step data mining guide*. DaimlerChrysler.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Emamjome, F., Andrews, R., & Ter Hofstede, A. H. M. (2019). A Case Study Lens on Process Mining in Practice. In H. Panetto, C. Debruyne, M. Hepp, D. Lewis, C. A. Ardagna, & R. Meersman (Eds.), *On the Move to Meaningful Internet Systems: OTM 2019 Conferences* (Vol. 11877, pp. 127–145). Springer International Publishing. https://doi.org/10.1007/978-3-030-33246-4_8

Florek, P., & Zagdański, A. Z. (2023). *Benchmarking state-of-the-art gradient boosting algorithms for classification*. https://arxiv.org/pdf/2305.17094

Flynn, A., & Li, Q. (2023). Determinants of supplier payment times before and during the pandemic:

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

Empirical evidence from UK firms. *Journal of Purchasing and Supply Management*, *29*(4), 100850. https://doi.org/10.1016/j.pursup.2023.100850

Jessen, U., Sroka, M., & Berti, A. (2025). Towards User-Oriented Process Mining: A Collaborative Approach to Minimize Late Payments in Accounts Payable Process. In J. Vom Brocke, J. Mendling, & M. Rosemann (Eds.), *Business Process Management Cases Vol. 3* (pp. 119–131). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-80793-0_9

Kermani, M. A. M. A., Maghsoudi, M., & Darzi, E. (2024). Enhancing Procurement Performance in Project-Oriented Organizations: A Process Analysis Approach. *IEEE Access*, *12*, 134340–134354. https://doi.org/10.1109/ACCESS.2024.3462852

Ni, W., Zhao, G., Liu, T., Zeng, Q., & Xu, X. (2023). Predictive Business Process Monitoring Approach Based on Hierarchical Transformer. *Electronics*, *12*(6), 1273. https://doi.org/10.3390/electronics12061273

Sánchez-Rebull, M.-V., Ferrer-Rullan, R., Hernández-Lara, A.-B., & Niñerola, A. (2020). Six Sigma for improving cash flow deficit: a case study in the food can manufacturing industry. *International Journal of Lean Six Sigma*, *11*(6), 1105–1126. https://doi.org/10.1108/IJLSS-12-2018-0137

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science*, *181*, 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Tax, N., Verenich, I., La Rosa, M., & Dumas, M. (2016). *Predictive Business Process Monitoring with LSTM Neural Networks*. https://doi.org/10.48550/ARXIV.1612.02130

Teinemaa, I., Dumas, M., Rosa, M. La, & Maggi, F. M. (2019). Outcome-Oriented Predictive Process Monitoring: Review and Benchmark. *ACM Transactions on Knowledge Discovery from Data*, *13*(2), 1–57. https://doi.org/10.1145/3301300

Van Der Aalst, W. (2016). *Process Mining: Data Science in Action*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49851-4

Van Der Aalst, W. M. P., & Carmona, J. (Eds.). (2022). *Process Mining Handbook* (Vol. 448). Springer International Publishing. https://doi.org/10.1007/978-3-031-08848-3

van Dongen, B. (2019). *BPI Challenge 2019*. 4TU.Centre for Research Data. https://doi.org/10.4121/UUID:D06AFF4B-79F0-45E6-8EC8-E19730C248F1

Weinzierl, S., Zilker, S., Dunzer, S., & Matzner, M. (2024). Machine learning in business process management: A systematic literature review. *Expert Systems with Applications*, *253*, 124181. https://doi.org/10.1016/j.eswa.2024.124181

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.