DETEKSI DAN ANALISIS FRAUD EVENT DATA MENGGUNAKAN ALPHA ALGORITHM DAN ANOMALY DETECTION

TESIS

Oleh:
ANDI FEBRY PUTRA ADHITAMA
NIM. 210605220010

PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025

DETEKSI DAN ANALISIS FRAUD EVENT DATA MENGGUNAKAN ALPHA ALGORITHM DAN ANOMALY DETECTION

TESIS

Diajukan kepada:
Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Magister Komputer (M.Kom)

Oleh:
ANDI FEBRY PUTRA ADHITAMA
NIM. 210605220010

PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025

DETEKSI DAN ANALISIS FRAUD EVENT DATA MENGGUNAKAN ALPHA ALGORITHM DAN ANOMALY DETECTION

TESIS

Oleh:
ANDI FEBRY PUTRA ADHITAMA
NIM. 210605220010

Telah diperiksa dan disetujui untuk diuji
Tanggal: 7 November 2025

Pembimbing I,	Pembimbing II,
Dr. M. Ainul Yaqin, M.Kom	Prof. Dr. Suhartono, M.Kom
NIP. 19761013 200604 1 004	NIP. 19680519 200312 1 001

Mengetahui,
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Prof. Dr. Ir. Muhammad Faisal, S.Kom., M.T.
NIP. 19740510 200501 1 007

DETEKSI DAN ANALISIS FRAUD EVENT DATA MENGGUNAKAN ALPHA ALGORITHM DAN ANOMALY DETECTION

TESIS

Oleh:
ANDI FEBRY PUTRA ADHITAMA
NIM. 210605220010

Telah Dipertahankan di Depan Dewan Penguji
Dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Magister Komputer (M.Kom)
Tanggal: 12 November 2025

Susunan Dewan Penguji			**Tanda Tangan**
Penguji I	:	Prof. Dr. Ir. Muhammad Faisal, M.T NIP. 19740510 200501 1 007	( )
Penguji II	:	Dr. Yunifa Miftachul Arif, M.T NIP 19830616 201101 1 004	( )
Pembimbing I	:	Dr. M. Ainul Yaqin, M.Kom NIP. 19761013 200604 1 004	( )
Pembimbing II	:	Prof. Dr. Suhartono, M.Kom NIP. 19680519 200312 1 001	( )

Mengetahui dan Mengesahkan,
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Prof. Dr. Ir. Muhammad Faisal, S.Kom., M.T.
NIP. 19740510 200501 1 007

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama	: Andi Febry Putra Adhitama
NIM	: 210605220010
Program Studi	: Magister Informatika
Fakultas	: Sains dan Teknologi
Judul Thesis	: "Deteksi dan Analisis Fraud Event Data Menggunakan Alpha Algorithm dan Anomaly Detection"

Menyatakan dengan sebenarnya bahwa Thesis yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan Thesis ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 7 November 2025

Yang membuat pernyataan,

Andi Febry Putra Adhitama

NIM. 210605220010

HALAMAN MOTTO

"Lebih baik salah karena mencoba daripada stagnan karena takut analisis, terus belajar, terus beradaptasi, terus maju"

HALAMAN PERSEMBAHAN

ٱلْحَمْدُ لِلَّهِ رَبِّ ٱلْعَالَمِينَ

Puji syukur atas kehadirat Allah SWT
Shalawat serta salam kepada Rasulullah SAW

Dengan segenap hati, penulis mempersembahkan sebuah karya ini kepada:

Istri tercinta Vina Chandra Wijaya, S.Pd yang selalu menjadi penyemangat, dan selalu memberikan do'a dukungan, serta motivasi.

Orang tua penulis tercinta, Ayah Dr. Andi Bustan, M.Si dan Ibu Dra. Femmy, M.Pd yang selalu membimbing penulis, memberikan do'a, dukungan, serta motivasi yang tidak terhingga.

Saudara dan Saudari serta Ponakan tercinta dr. Andi Ferdi Saputra, S.Ked dan Andi Wira Gama Putra, S. IP dan dr. Indah Permata Bunda, S.Ked dan Andi Jacinda.

Dosen pembimbing Dr. M. Ainul Yaqin, M.Kom dan Prof. Dr. Suhartono, M.Kom yang telah membimbing penelitian ini dengan memberikan banyak pengarahan dan pengalaman yang berharga.

Segenap sivitas akademika Program Studi Magister Informatika, terutama seluruh dosen, terima kasih atas segenap ilmu dan bimbingannya.

Seluruh rekan-rekan mahasiswa Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang

Penulis ucapkan "*jazakumullah khairan katsiiraa"*. Semoga selalu diridhoi Allah SWT. Aamiin Ya Rabbal 'Alamiin.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Syukur alhamdulillah penulis hanturkan kehadirat Allah SWT yang telah melimpahkan Rahmat dan Hidayah-Nya, sehingga penulis dapat menyelesaikan studi di Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang sekaligus menyelesaikan Thesis ini dengan baik. Shalawat dan salam semoga senantiasa tercurahkan kepada junjungan Nabi Muhammad SAW yang senantiasa menjadi sumber inspirasi dan teladan terbaik begitu juga keluarga, para sahabat dan para pengikutnya seluruh umat Islam.

Penulis haturkan ucapan terima kasih seiring do'a dan harapan jazakumullah ahsanal jaza' kepada semua pihak yang telah membantu terselesaikannya Thesis ini. Ucapan terima kasih ini penulis sampaikan kepada:

1. Bapak Dr. M. Ainul Yaqin, M.Kom dan Bapak Prof. Dr. Suhartono, M.Kom selaku dosen pembimbing Thesis, yang telah banyak memberikan pengarahan dan pengalaman yang berharga.
2. Bapak Prof. Dr. M. Faisal, M.T dan Bapak Dr. Yunifa Miftachul Arif, M.T dan selaku dosen penguji Thesis, yang telah banyak memberikan pengarahan dan pengalaman yang berharga.
3. Keluarga tercinta yang telah banyak memberikan doa dan dukungan kepada penulis secara moril maupun materil hingga Thesis ini dapat terselesaikan.
4. Segenap Civitas Akademika Program Studi Magister Informatika, terutama seluruh Bapak dan Ibu dosen, terima kasih atas segenap ilmu dan bimbingannya
5. Semua pihak yang ikut membantu dalam menyelesaikan Thesis ini baik berupa materiil maupun moril yang tidak bisa penulis sebutkan satu persatu tanpa mengurasi rasa hormat dan terimakasih.

Penulis menyadari bahwa dalam penyusunan Thesis ini masih terdapat kekurangan dan penulis berharap semoga Thesis ini bisa memberikan manfaat kepada para pembaca khususnya bagi penulis secara pribadi. Aamiin Ya Rabbal 'Alamin.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Malang, 7 November 2025

Penulis

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

ABSTRAK

Adhitama, Andi Febry Putra. 2025. **Deteksi Dan Analisis Fraud Event Data Menggunakan Alpha Algorithm Dan Anomaly Detection.** Tesis. Program Studi Magister Informatika Fakultas Sains Dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. M. Ainul Yaqin, M.Kom (II) Prof. Dr. Suhartono, M.Kom.

Kata kunci: Analisis Fraud, Alpha Algorithm, Anomaly Detection.

Persediaan merupakan aset strategis yang rentan terhadap praktik kecurangan, khususnya pada aktivitas pergudangan yang melibatkan banyak transaksi dan pelaku operasional. Kecurangan dalam pengelolaan persediaan sering kali bersifat terselubung dan sulit dideteksi menggunakan metode audit konvensional. Penelitian ini bertujuan untuk mendeteksi dan menganalisis potensi fraud pada data event log sistem informasi gudang CV Putera Bumi dengan mengintegrasikan teknik process mining dan anomaly detection. Metode Alpha Algorithm digunakan untuk merekonstruksi alur proses bisnis aktual dan mengidentifikasi penyimpangan terhadap Standar Operasional Prosedur (SOP), sedangkan Isolation Forest diterapkan sebagai metode unsupervised learning untuk mendeteksi aktivitas anomali tanpa memerlukan data berlabel. Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen berbasis data dan mengadopsi kerangka kerja CRISP-DM. Hasil penelitian menunjukkan bahwa integrasi kedua metode tersebut efektif dalam mengidentifikasi deviasi proses, rare events, serta transaksi dengan skor anomali tinggi yang berpotensi mengindikasikan fraud, seperti manipulasi stok dan transaksi di luar jam operasional. Sekitar 0,86% dari total transaksi terdeteksi sebagai anomali signifikan. Temuan ini membuktikan bahwa pendekatan berbasis event log mampu memberikan deteksi dini terhadap potensi kecurangan serta mendukung penguatan pengendalian internal secara adaptif dan berbasis data.

ABSTRACT

Adhitama, Andi Febry Putra. 2025. **Detection and Analysis of Fraud Event Data Using Alpha Algorithm and Anomaly Detection**. Thesis. Master of Informatics Study Program, Faculty of Science and Technology, State Islamic University of Maulana Malik Ibrahim Malang. Supervisor: (I) Dr. M. Ainul Yaqin, M.Kom (II) Prof. Dr. Suhartono, M.Kom.

Keywords: Fraud Analysis, Alpha Algorithm, Anomaly Detection.

Inventory is a strategic asset vulnerable to fraudulent practices, particularly in warehousing activities involving numerous transactions and operational actors. Fraud in inventory management is often covert and difficult to detect using conventional audit methods. This study aims to detect and analyze potential fraud in event log data from CV Putera Bumi's warehouse information system by integrating process mining and anomaly detection techniques. The Alpha Algorithm method is used to reconstruct the actual business process flow and identify deviations from Standard Operating Procedures (SOPs), while Isolation Forest is applied as an unsupervised learning method to detect anomalous activity without requiring labeled data. This study uses a quantitative approach with a data-driven experimental design and adopts the CRISP-DM framework. The results show that the integration of the two methods is effective in identifying process deviations, rare events, and transactions with high anomaly scores that could potentially indicate fraud, such as stock manipulation and transactions outside of operating hours. Approximately 0.86% of the total transactions were detected as significant anomalies. These findings prove that an event log-based approach is capable of providing early detection of potential fraud and supporting the strengthening of internal controls in an adaptive and data-based manner.

الملخص

أدهيتاما، أندي فبراير بوترا. 2025. **الكشف والتحليل لبيانات أحداث الاحتيال باستخدام خوارزمية ألفا وكشف الشذوذ**. أطروحة. برنامج ماجستير المعلوماتية، كلية العلوم والتكنولوجيا، جامعة الدولة الإسلامية للمولانا مالك إبراهيم ملانغ. المشرف: (الأولى) الدكتورة م. عين اليقين، الممرضة (الثانية) البروفيسورة الدكتورة سوهارتونو، المتقدمة

الكلمات المفتاحية: تحليل الاحتيال، خوارزمية ألفا، اكتشاف الشذوذ.

يعد المخزون أصلا استراتيجيا عرضة للممارسات الاحتيالية، خاصة في أنشطة التخزين التي تشمل العديد من المعاملات والجهات التشغيلية. غالبا ما يكون الاحتيال في إدارة المخزون سريا ويصعب اكتشافه باستخدام طرق التدقيق التقليدية. تهدف هذه الدراسة إلى اكتشاف وتحليل الاحتيال المحتمل في بيانات سجلات الأحداث من نظام معلومات المستودع الخاص ب CV Putera Bumi من خلال دمج تقنيات التنقيب في العمليات واكتشاف الشذوذ. تستخدم طريقة خوارزمية ألفا لإعادة بناء تدفق العمليات التجارية الفعلية وتحديد الانحرافات عن إجراءات التشغيل القياسية (SOPs)، بينما تستخدم غابة العزل كطريقة تعلم غير خاضع للإشراف لاكتشاف النشاط الشاذ دون الحاجة إلى بيانات معنة. تستخدم هذه الدراسة نهجا كميا مع تصميم تجريبي قائم على البيانات وتعتمد إطار عمل CRISP-DM. تظهر النتائج أن دمج الطريقتين فعال في تحديد الانحرافات في العمليات، والأحداث النادرة، والمعاملات ذات درجات الشذوذ العالية التي قد تشير إلى احتيال، مثل التلاعب بالمخزون والمعاملات خارج ساعات العمل. تم اكتشاف حوالي 0.86٪ من إجمالي المعاملات كشذوذات ذات دلالة. تثبت هذه النتائج أن النهج القائم على سجل الأحداث قادر على الكشف المبكر عن الاحتيال المحتمل ودعم تعزيز الضوابط الداخلية بطريقة تكيفية وقائمة على البيانات.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Persediaan merupakan salah satu aset terpenting dalam perusahaan, khususnya bagi sektor manufaktur, distribusi, dan logistik. Nilai persediaan yang tinggi menjadikannya rentan terhadap risiko penyalahgunaan, baik berupa manipulasi data stok, penggelapan barang, maupun pelaporan persediaan yang tidak akurat. Sektor logistik dan penyimpanan secara global mengalami kerugian signifikan akibat kecurangan (*fraud*). Laporan Association of Certified *Fraud* Examiners (ACFE, 2022) mencatat bahwa rata-rata kerugian akibat *fraud* di sektor ini mencapai $150.000 per kasus, dengan sebagian besar baru terungkap setelah lebih dari 14 bulan berlangsung. Selain itu, sekitar 43% kasus kecurangan dilakukan oleh karyawan operasional yang memiliki akses langsung terhadap barang maupun sistem pencatatan (Albrecht et al., 2019).

Dalam lingkungan bisnis *modern*, hampir seluruh proses operasional, termasuk pengelolaan persediaan, dijalankan melalui sistem informasi yang menghasilkan *event* log. *Event log* adalah catatan kronologis aktivitas dalam sistem, seperti login pengguna, perubahan data, transaksi, dan eksekusi proses tertentu (Aalst, 2016). Data ini tidak hanya berfungsi sebagai dokumentasi aktivitas, tetapi juga menjadi sumber informasi penting untuk audit, evaluasi kinerja, dan deteksi penyimpangan.

Fraud dalam sistem informasi, termasuk yang terkait dengan persediaan, seringkali bersifat terselubung. Pelaku berupaya membuat aktivitasnya menyerupai perilaku normal sehingga sulit dideteksi dengan metode konvensional (Bolton & Hand, 2002). Inilah yang membuat pendekatan tradisional berbasis pemeriksaan manual atau aturan statis menjadi kurang efektif dalam mengidentifikasi aktivitas mencurigakan di *event* log.

Salah satu solusi yang berkembang pesat adalah *anomaly detection*. Teknik ini bertujuan mengidentifikasi pola yang menyimpang dari perilaku umum. Dalam konteks *event log*, *anomaly detection* memanfaatkan pembelajaran mesin, terutama unsupervised learning, untuk mempelajari pola normal dari data historis tanpa memerlukan label *fraud* secara eksplisit (Ahmed et al., 2016). Pendekatan ini mampu mendeteksi anomali yang jarang terjadi namun berdampak besar terhadap keamanan aset persediaan (Goldstein & Uchida, 2016).

Mengintegrasikan analisis *event log* dengan metode *anomaly detection* dapat meningkatkan deteksi dini terhadap aktivitas mencurigakan, mempercepat proses investigasi, serta meminimalkan potensi kerugian akibat *fraud*. Dengan besarnya kerugian yang ditimbulkan dan lamanya waktu deteksi dalam kasus-kasus *fraud* persediaan, penelitian mengenai penerapan *anomaly detection* pada data *event log* menjadi penting untuk dikembangkan dan diimplementasikan pada sektor logistik dan pergudangan.

aktor-faktor utama yang menyebabkan tingginya tingkat kecurangan di area gudang antara lain adalah lemahnya pengendalian internal, kurangnya pengawasan terhadap transaksi keluar-masuk barang, serta tidak digunakannya sistem pemantauan berbasis data yang mampu melakukan deteksi dini terhadap anomali dalam aktivitas operasional. Singleton et al. (2020) menegaskan bahwa pencatatan manual atau berbasis kertas masih lazim digunakan pada banyak perusahaan kecil dan menengah, sehingga membuka peluang terjadinya manipulasi data secara sistematis.

Pendekatan berbasis analisis *event log* menawarkan solusi strategis dalam mitigasi risiko *fraud* di gudang. *Event log* yang dihasilkan dari aktivitas pergudangan merekam setiap proses, mulai dari penerimaan barang, perpindahan antar lokasi, hingga pengeluaran barang. Dengan analisis berbasis *process mining*, perusahaan dapat mengidentifikasi pola aktivitas normal, mendeteksi penyimpangan, serta mengenali distribusi barang yang tidak wajar, misalnya pengeluaran di luar jam operasional atau perbedaan signifikan antara catatan sistem dengan hasil audit fisik (van der Aalst, 2016; Dumas et al., 2018).

Integrasi *process mining* dengan teknik *anomaly detection* memungkinkan sistem mempelajari perilaku historis dan secara otomatis memberi peringatan ketika ditemukan pola yang menyimpang dari norma. Pendekatan ini tidak hanya membantu mendeteksi anomali yang sulit dilihat secara manual, tetapi juga memungkinkan analisis *root cause* untuk

memahami titik-titik rawan dalam proses pengelolaan gudang (Ahmed et al., 2016; Goldstein & Uchida, 2016).

Lebih lanjut, rekonstruksi proses berdasarkan *event log* dapat membantu dalam investigasi kejadian kehilangan barang (*event loss reconstruction*). Hasil analisis ini memberikan wawasan penting mengenai lokasi, pelaku potensial, serta mekanisme terjadinya *fraud*. Dengan demikian, perusahaan dapat mengimplementasikan langkah pencegahan yang tepat sasaran untuk meningkatkan keamanan dan akurasi pencatatan persediaan.

Sebagai landasan nilai dalam pengelolaan aset perusahaan secara jujur dan bertanggung jawab, maka prinsip kejujuran dan pencatatan transaksi secara adil juga telah ditegaskan dalam Al-Qur'an:

يَٰٓأَيُّهَا الَّذِينَ آمَنُوا إِذَا تَدَايَنتُم بِدَيْنٍ إِلَىٰ أَجَلٍ مُّسَمًّى فَاكْتُبُوهُ

Artinya: *"Wahai orang-orang yang beriman! Apabila kamu melakukan hutang-piutang untuk waktu yang ditentukan, maka hendaklah kamu menuliskannya."* (QS. Al-Baqarah: 282)

1. Perintah untuk Menulis Transaksi (كتابة الدين)

Ayat ini memberikan perintah tegas agar setiap transaksi hutang-piutang yang *bertenggang waktu* ditulis secara tertib.

Tujuannya:

- Mencegah kesalahpahaman antara pihak yang bertransaksi.
- Menghindari sengketa di kemudian hari.
- Menjaga hak masing-masing pihak.

Ulama menjelaskan bahwa penulisan ini adalah bentuk profesionalisme dan amanah dalam muamalah.

2. Pentingnya Kejelasan Akad

Dalam ayat ini ditegaskan konsep:

أَجَلٍ مُسَمًّى)jangka waktu yang jelas dan pasti)

Artinya:

- Hutang tidak boleh menggantung.
- Deadline harus jelas: hari, tanggal, atau waktu pasti.
- Untuk menghindari penundaan atau penipuan kelak.

3. Keadilan Dalam Pencatatan

Ayat ini juga memerintahkan agar penulisan transaksi dilakukan dengan **adil**, tidak memihak salah satu pihak.

Meskipun tidak terlihat dalam cuplikan, kelanjutan ayat menjelaskan:

- Adanya **pencatat (كاتب)** yang menuliskan dengan benar.
- Adanya **saksi (شهيدان)** untuk memastikan keabsahan.

Ini menjadi landasan hukum penting dalam fikih muamalah.

4. Hikmah Disyariatkannya Pencatatan Hutang

Para ulama menyebutkan beberapa hikmah:

- Merapikan administrasi
- Menjaga hubungan baik antara kedua pihak
- Menghindari lupa, zalim, atau manipulasi
- Menguatkan keadilan dalam sistem ekonomi Islam
- Ayat ini sering dijadikan rujukan untuk kontrak:
- jual-beli kredit

- sewa-menyewa
- akad bagi hasil
- pinjaman
- proyek bisnis

5. Hukum Menurut Para Ulama

Mayoritas ulama: penulisan hutang hukumnya *sunnah mu'akkadah* yang dianjurkan keras.

Namun sebagian berpendapat: *wajib* bila berpotensi menimbulkan sengketa atau ketidakterangan.

6. Relevansi Dengan Sistem Modern

Ayat ini menjadi dasar:

Kontrak bisnis modern

Perjanjian kredit

Bukti transaksi digital

Sistem pencatatan berbasis aplikasi atau database

Jadi walaupun teknologi berubah, esensinya tetap sama: transaksi harus dicatat**.**

Berdasarkan penjabaran di atas, maka peneliti tertarik untuk melakukan penelitian dengan judul "Deteksi Dan Analisis Fraud Event Data Menggunakan Alpha Algorithm dan Anomaly Detection"

1.2 Rumusan Masalah

Berdasarkan latar belakang, permasalahan penelitian tidak hanya terbatas pada kemampuan mendeteksi anomali, tetapi juga mencakup kualitas data event log, pemilihan metode, evaluasi hasil, serta bagaimana anomali tersebut diinterpretasikan sebagai indikasi fraud. Oleh karena itu, rumusan masalah dapat diperluas sebagai berikut:

1. Bagaimana menghasilkan dan memproses data *event log* agar layak digunakan untuk analisis anomali, yang mencakup normalisasi data, ekstraksi fitur, dan penanganan inkonsistensi data?
2. Bagaimana mendapatkan metode anomaly detection yang efektif untuk mendeteksi aktivitas tidak wajar serta mengidentifikasi anomali wajar (*benign anomalies*) dan anomali yang mengindikasikan potensi fraud (*fraud-oriented anomalies*)?
3. Bagaimana membedakan anomali wajar dan anomali fraud secara sistematis, mengevaluasi serta membandingkan kinerja metode *anomaly detection* dengan pendekatan konvensional, dan menyajikan hasilnya secara interpretatif bagi pengambil keputusan?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengembangkan suatu pendekatan berbasis *anomaly detection* yang dapat digunakan untuk mendeteksi dan menganalisis potensi *fraud* dalam *event log* sistem informasi. Secara lebih spesifik, tujuan dari penelitian ini adalah sebagai berikut:

1. Merancang dan membangun model deteksi anomali yang mampu mengenali perilaku tidak normal dalam data *event log* sistem informasi menggunakan metode machine learning, khususnya pendekatan *unsupervised learning*.
2. Mengidentifikasi dan menganalisis pola-pola *fraud* yang tersembunyi dalam data *log* berdasarkan hasil deteksi anomali, serta mengelompokkan jenis aktivitas yang berpotensi sebagai *fraud*.
3. Mengevaluasi kinerja model deteksi anomali dalam konteks deteksi *fraud*, dengan menggunakan metrik evaluasi yang sesuai seperti *precision, recall, F1-score,* dan/atau visualisasi hasil *clustering* atau *scoring*.
4. Menyediakan insight dan rekomendasi untuk pengembangan sistem keamanan berbasis *log* yang lebih adaptif dan proaktif dalam mendeteksi aktivitas mencurigakan.

1.4 Manfaat Penelitian

Berdasarkan seluruh pemaparan di atas, maka dilakukannya penelitian ini diharapkan dapat memberikan manfaat, sebagai berikut:

1. Manfaat Teoritis

   Penelitian ini diharapkan dapat menambah khazanah keilmuan di bidang data *mining* dan keamanan sistem informasi, khususnya dalam penerapan *anomaly detection* terhadap data *event log*.

2. Manfaat Praktis

Secara teoritis, penelitian ini berkontribusi dalam pengembangan kajian terkait pendeteksi *fraud* berbasis *log* sistem informasi

1.5 Batasan Masalah

Untuk membatasi ruang lingkup penelitian, penulis hanya akan fokus pada deteksi *fraud* yang berbasis anomali dalam data *event* log. Penelitian ini tidak mencakup investigasi lanjutan terhadap *fraud*, validasi hukum, atau integrasi dengan sistem forensik. Data yang digunakan berupa *log* dari sistem yang disimulasikan atau anonim, tanpa menyertakan informasi sensitif dari institusi tertentu.

1.6 Sistematika Penulisan

Adapun sistematika penulisan dalam tesis ini dibagi menjadi lima bab utama sebagai berikut:

1. Bab I Pendahuluan

   Berisi latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, dan sistematika penulisan.

2. Bab II Tinjauan Pustaka

   Membahas teori-teori yang relevan dengan topik penelitian, hasil penelitian terdahulu, serta kerangka teori yang digunakan.

3. Bab III Metodologi Penelitian

   Menjelaskan desain penelitian, sumber data, metode yang digunakan, serta tahapan penelitian.

4. Bab IV Hasil dan Pembahasan

   Berisi hasil implementasi metode, analisis terhadap data, serta pembahasan mengenai temuan penelitian.

5. Bab V Kesimpulan dan Saran

   Menyajikan kesimpulan dari penelitian dan saran untuk pengembangan selanjutnya.

BAB II

KAJIAN PUSTAKA

2.1 Penelitian Terdahulu

Penelitian sebelumnya merupakan upaya seorang peneliti untuk melakukan perbandingan dan kemudian menemukan motivasi baru untuk penelitian di masa mendatang. Lebih lanjut, penelitian masa lalu berfungsi untuk memposisikan dan mengilustrasikan orisinalitas penelitian. Pada bagian ini, peneliti mencantumkan dan meringkas temuan penelitian masa lalu yang berkaitan dengan penelitian yang direncanakan, baik yang telah dipublikasikan maupun yang belum dipublikasikan. Berikut adalah daftar penelitian masa lalu tentang subjek yang sedang diteliti.

1. Ribeiro et al. (2022)

Ribeiro et al. mengusulkan pendekatan *Isolation Forest* untuk deteksi anomali pada *event log* sistem bisnis dengan fokus pada data yang tidak memiliki label. Mereka mengadaptasi teknik ini karena kemampuannya untuk mengisolasi data outlier secara efisien tanpa perlu asumsi distribusi data. Dalam studi tersebut, skor anomali dihitung terhadap representasi trace berbasis frekuensi aktivitas dan durasi proses. Hasilnya menunjukkan bahwa *Isolation Forest* efektif dalam mengidentifikasi jejak yang menyimpang secara statistik.

Namun, penelitian Ribeiro tidak menggabungkan pendekatan ini dengan analisis struktur proses seperti yang ditawarkan oleh *process mining*. Pendekatan mereka lebih menekankan pada deteksi outlier dari

fitur numerik, tanpa mempertimbangkan konteks urutan aktivitas yang dapat memberikan informasi tambahan tentang deviasi proses. Ini menjadi salah satu celah yang ingin diisi dalam penelitian ini.

Dalam penelitiannya yang berjudul *"Detecting Deviations in Business Processes through Process Mining and Outlier Detection"*, Ribeiro menggabungkan *process mining* dengan algoritma *clustering* DBSCAN untuk mendeteksi outlier dalam proses bisnis. Mereka menganalisis durasi aktivitas dan urutan kejadian untuk menemukan penyimpangan proses.

- **Kelebihan:**

Pendekatan ini mampu mengidentifikasi outlier berbasis waktu proses dan membantu visualisasi proses abnormal secara intuitif.

- **Kelemahan:**

Tidak secara eksplisit dirancang untuk mendeteksi *fraud*. Outlier yang ditemukan belum tentu merupakan aktivitas *fraud*.

- **Relevansi:**

Penelitian ini menunjukkan pentingnya integrasi antara *process mining* dan *anomaly detection*, sebagaimana yang juga menjadi inti dari penelitian ini.

2. Conforti et al. (2015)

Dalam artikel *"Noise filtering of process execution logs based on outliers detection"*, Conforti mengembangkan metode untuk menyaring noise (aktivitas tidak relevan) dari *log* sistem, agar *process* model yang dihasilkan lebih akurat.

Conforti mengembangkan pendekatan berbasis *deviance mining* untuk mendeteksi penyimpangan perilaku dalam proses bisnis. Mereka memperkenalkan metode untuk mengidentifikasi perbedaan perilaku berdasarkan *performance metrics* dan *sequence patterns* dalam *event* log, serta membedakan antara proses normal dan abnormal berdasarkan *log* aktivitas yang telah ditandai.

Pendekatan ini menggunakan teknik supervised learning untuk membedakan jejak proses yang sesuai dan tidak sesuai dengan SOP, berdasarkan pelabelan sebelumnya. Meskipun metode ini sangat efektif dalam mendeteksi dan menjelaskan penyimpangan berbasis pola perilaku, namun ketergantungannya pada data berlabel menjadi salah satu keterbatasan utama, terutama dalam konteks *fraud* yang sulit diidentifikasi secara eksplisit.

Penelitian ini menjadi inspirasi bahwa penyimpangan proses dapat dianalisis lebih dalam dari sekadar deteksi numerik, namun juga menunjukkan perlunya pendekatan unsupervised yang tidak bergantung pada pelabelan manual—sesuatu yang diupayakan dalam penelitian ini melalui kombinasi *process mining* dan *Isolation Forest*.

- **Kelebihan:**

Memperkuat kualitas model proses dengan membersihkan noise yang dapat menurunkan akurasi.

- **Kelemahan:**

Tidak fokus pada *fraud*; lebih kepada *cleaning log* agar *process mining* bekerja lebih baik.

- **Relevansi:**

Menunjukkan bahwa kualitas *log* sangat berpengaruh terhadap kemampuan analisis dan deteksi deviasi.

3. Alsalman (2024)

Penelitian ini bertajuk *"A comparative study of anomaly detection techniques for IoT security using adaptive machine learning for IoT threats"*. Mereka membandingkan berbagai metode deteksi anomali seperti *Isolation Forest*, One-Class SVM, dan Autoencoder pada *event* log.

Bertoli, mengusulkan pendekatan kombinasi antara *process mining* dan teknik visualisasi berbasis *trace clustering* untuk mendeteksi deviasi proses yang berpotensi anomali. Mereka memanfaatkan algoritma *clustering* seperti DBSCAN untuk mengelompokkan jejak proses berdasarkan kemiripan urutan aktivitas, kemudian mengidentifikasi kelompok kecil yang menyimpang dari cluster mayoritas sebagai indikasi deviasi.

Penelitian ini cukup relevan karena mencoba mengurangi ketergantungan terhadap model referensi eksplisit seperti Petri Net, dan lebih menekankan pada *behavioral pattern analysis*. Hasilnya menunjukkan bahwa metode ini mampu mengungkap deviasi tidak biasa yang tidak tampak melalui perbandingan model statis.

Namun, meskipun Bertoli et al. menggunakan *clustering* sebagai basis deteksi, pendekatannya belum mempertimbangkan skoring individual dari trace atau integrasi langsung dengan algoritma deteksi anomali berbasis pohon seperti *Isolation Forest*. Di sisi lain, penelitian ini juga tidak sepenuhnya mengadopsi framework *process mining* klasik yang memungkinkan pemetaan aktivitas terhadap struktur proses formal.

Penelitian ini menginspirasi bahwa ada potensi besar dalam menggabungkan *pattern-based detection* dengan pendekatan *unsupervised anomaly scoring*, yang menjadi salah satu landasan utama dalam desain penelitian ini.

- **Kelebihan:**

Memberikan insight tentang performa berbagai algoritma dalam konteks *log event*. *Isolation Forest* ditemukan memiliki performa yang stabil di berbagai skenario.

- **Kelemahan:**

Tidak menggunakan *process mining*, jadi tidak mempertimbangkan urutan proses secara eksplisit.

- **Relevansi:**

Mendukung pemilihan *Isolation Forest* dalam penelitian ini sebagai algoritma *anomaly detection* yang cocok untuk *log* tanpa label.

Tabel 2.1 Perbandingan Penelitian Terdahulu

No	Penulis	Metode	Fokus
1	Ribeiro et al. (2020)	*Process Mining* + DBSCAN	Durasi & urutan
2	Conforti et al. (2016)	Filtering log	Precision model
3	Bertoli et al. (2021)	IF, SVM, Autoencoder	*Event log anomaly*

Posisi Penelitian Ini

Penelitian ini mengisi celah dengan:

- Menggabungkan *process mining* (*Alpha Algorithm*) dan *anomaly detection* (*Isolation Forest*)
- Fokus pada deteksi *fraud* berbasis *event* log
- Menghasilkan sistem deteksi yang otomatis, visual, dan berbasis data aktual proses

2.2 *Fraud*

Association of Certified Fraud Examiners (ACFE) mendefinisikan kecurangan sebagai tindakan yang melanggar hukum atau peraturan yang berlaku dan ditandai dengan perilaku tidak jujur, penggelapan, atau penyalahgunaan wewenang. Kejahatan ini tidak termasuk ancaman kekerasan, kekuatan fisik, atau kekuatan non-fisik yang digunakan oleh individu atau kelompok untuk keuntungan pribadi atau untuk membantu pihak tertentu yang melakukan kecurangan (ACFE, 2021).

International Auditing and Assurance Standard Board (IAASB) sebagai bagian dari Internasional Federation of Accountants (IFAC, penipuan didefinisikan sebagai tindakan yang disengaja oleh satu atau lebih orang yang

berada dalam posisi berwenang, seperti manajemen, staf, atau pihak ketiga, untuk memperoleh keuntungan yang melanggar hukum dan tidak jujur (IFAC, 2020).

Penipuan didefinisikan sebagai setiap perilaku atau tindakan ilegal yang sengaja melanggar undang-undang yang berlaku untuk menipu individu atau organisasi dan memperoleh keuntungan bagi pelaku. Dalam konteks ini, muncul tiga istilah penting: perilaku yang melanggar hukum, disengaja, dan menimbulkan kerugian atau kerusakan pada orang lain (Tjahjono, 2013).

Menurut Rozali (2017), suatu bisnis dapat menghadapi risiko penipuan dalam berbagai cara.

1. Penipuan internal: Ini merujuk pada penipuan yang terjadi di dalam lembaga itu sendiri dan dapat dilakukan oleh staf. Karyawan yang memiliki akses ke dokumen akuntansi lebih rentan terhadap pencurian, terutama jika kontrol internal longgar.
2. Penipuan eksternal: Berbeda dengan penipuan internal, penipuan eksternal merujuk pada aktivitas curang yang dilakukan oleh orang-orang di luar lembaga atau organisasi, seperti perampokan, pencurian, penipuan, atau peretasan komputer. Hal ini biasanya disebabkan oleh sistem keamanan aset yang kurang memadai.
3. Kolusi: Kolusi terjadi ketika dua pihak atau lebih berkolaborasi, baik di dalam maupun di luar lembaga.

Menurut Tjahjono (2013) dilihat dari sudut pandang pemeriksa *fraud* dan hukum, ada 4 karakteristik utama yang menunjukan terjadinya *fraud*:

1. Tindakan yang bersifat material dan keliru.
2. Adanya kesepakatan/sepengetahuan bahwa tindakan tersebut keliru ketika dilakukan.
3. Adanya keyakinan atau pengakuan dari pelaku akan tindakan yang salah tersebut.
4. Adanya kerugian yang diderita oleh pihak lain.

2.3 Anomaly Detection

Suatu kumpulan data terdiri dari sejumlah titik data, yang masing-masing memiliki serangkaian atribut yang unik. Kumpulan data ini dapat dianalisis untuk memberikan informasi yang bermakna menggunakan teknik penambangan data seperti klasifikasi, pengelompokan, dan sebagainya. Tidak jarang kita menemukan data yang sangat berbeda dari sifat-sifat data pada umumnya.

Menurut Rahayu et al (2022), data yang memiliki kualitas tersebut kadang-kadang disebut sebagai anomali, outlier, atau noise. Outlier adalah data yang menyimpang dari norma, sedangkan inlier adalah data yang berada dalam kisaran normal. Keberadaan outlier dalam pengelompokan dapat menghasilkan hasil yang tidak memuaskan. Untuk menghasilkan hasil terbaik, penghapusan outlier sering dilakukan sebagai bagian dari persiapan kumpulan data. Dalam beberapa keadaan, hal ini dapat membantu menemukan kondisi data yang aneh. Dalam bab ini, kata-kata anomali data, outlier, dan noise digunakan secara sinonim.

Outlier umumnya didefinisikan sebagai objek/item data dengan jumlah yang sangat kecil jika dibandingkan dengan data reguler lainnya. Misalnya, peluang kemunculannya adalah satu banding seribu, tetapi dapat meningkat menjadi seribu jika data mencapai satu juta. Oleh karena itu, deteksi outlier dalam data yang menyimpang sangat penting untuk berbagai aplikasi dalam penambangan data.

Di dunia nyata, perbedaan data dapat diamati saat menyimpan data usia manusia. Usia manusia berkisar dari nol hingga sembilan puluh tahun. Ini tidak mengesampingkan potensi manusia hidup lebih dari 90 tahun, seperti 100 atau 110 tahun. Anomali, outlier, dan noise adalah istilah yang digunakan untuk menggambarkan data manusia yang usianya berbeda dari data usia secara keseluruhan. Data semacam itu sering disebut sebagai pengecualian data.

Aplikasi deteksi anomali dapat ditemukan di sejumlah bidang, termasuk deteksi penipuan, deteksi intrusi, gangguan ekosistem, kesehatan masyarakat, dan kedokteran. Berikut ini menjelaskan penggunaan deteksi anomali di bidang-bidang tersebut:

1. Deteksi penggelapan (*fraud detection*)

   Transaksi kartu kredit yang dilakukan secara online telah mengakibatkan munculnya penipu kartu kredit yang memanfaatkan detail kartu kredit konsumen untuk membeli barang dengan jenis, jumlah, dan pola pembelian yang berbeda dari yang dimaksudkan oleh klien. Tanpa upaya penyedia kartu kredit untuk mendeteksi anomali dalam pola transaksi kartu

kredit setiap pelanggan secara teratur, penipu pasti akan mencuri saldo kartu kredit, sehingga merugikan konsumen mereka.

2. Deteksi penyusupan (intrusion *detection*)

   Akses ke sumber data di suatu instansi, baik komputer maupun jaringan, tidak dapat ditolak hanya karena berasal dari tempat umum (internet), tetapi akses yang biasanya ditujukan, misalnya, untuk mematikan fungsi server atau merusak data yang tersimpan, akan menunjukkan perilaku yang berbeda, sebagaimana dibuktikan oleh isi header protokol atau isi pesan spesifik yang disisipkan di dalamnya. Diharapkan bahwa penggabungan teknologi deteksi anomali ke dalam protokol jaringan seperti firewall atau router akan meningkatkan sistem keamanan jaringan.

3. Gangguan ekosistem (ecosystem disturbances)

   Terdapat beberapa ritme kehidupan di Bumi, seperti musim dan pola kehidupan hewan, yang seringkali menyimpang. Perbedaan ini dapat menghasilkan kejadian alam yang aneh. Misalnya, variasi pola musiman dapat menyebabkan perubahan ritme kehidupan spesies tertentu. Mendeteksi variasi dalam keadaan alam sangat penting untuk peringatan dini terhadap kemungkinan risiko seperti banjir, kebakaran, gempa bumi, polusi, dan sebagainya.

4. Kesehatan masyarakat (public health)

   Di Indonesia, terdapat Puskesmas di setiap kecamatan atau kabupaten. Rekam medis yang disimpan di setiap Puskesmas mencatat kasus-kasus yang ditangani di sana, yang umumnya berasal dari daerah terdekat. Pola

penyakit di antara pasien di seluruh Puskesmas dalam suatu wilayah dapat diamati untuk mengidentifikasi berbagai pola penyakit yang ditangani di Puskesmas, termasuk kesalahan vaksinasi masyarakat (untuk anak-anak).

2.4 Isolation Forest

Salah satu teknik terbaru untuk mendeteksi anomali dari kumpulan data adalah algoritma *Isolation Forest*. Algoritma *Isolation Forest* dibangun berdasarkan algoritma decision trees. Algoritma ini memilih fitur secara acak dan memilih nilai secara acak antara nilai minimum dan maksimum dari fitur yang dipilih. *Isolation forest* merupakan sebuah metode yang diusulkan oleh Liu et al. (Liu, Ting, & Zhou, 2008) untuk mendeteksi data anomali. Metode *Isolation Forest* dibangun berdasarkan asumsi kuantitatif sebagai berikut: (1) data anomali berasal dari kategori minoritas atau kategori yang memiliki proporsi jumlah data lebih kecil dari proporsi jumlah data kategori lain, dan (2) data anomali memiliki nilai atribut sangat berbeda dari data yang lain.

Secara umum, langkah penting dari metode ini adalah membangun sebuah ensemble dari beberapa *isolation* tree (berupa binary tree) untuk memisahkan data input yang diberikan. Didalam sebuah *isolation* tree, data yang normal cenderung akan terisolasi pada ujung yang lebih dalam dari *isolation* tree; sedangkan, data anomali cenderung akan terisolasi lebih dekat ke akar *isolation* tree.

Jika diberikan n buah data $X = \{X_1, X_2, \ldots, X_n\}$ dimana setiap data memiliki d-dimensi (atribut). *Isolation* tree dibangun dengan membagi data secara rekursif dengan langkah-langkah sebagai berikut. Pertama, sebuah

atribut (q) dan nilai pemisah atau nilai split (p) dipilih secara random sampai salah satu kriteria berikut dipenuhi:

1) Tinggi (depth) *isolation* tree mencapai sebuah nilai maksimum;

2) Elemen data X tinggal satu atau |X| = 1;

3) Semua elemen data X memiliki nilai yang sama.

Sebuah *isolation* tree merupakan pohon biner (binary tree), di mana setiap node hanya memiliki maksimum 2 anak (children). Dengan asumsi semua elemen data memiliki nilai yang berbeda, setiap data akan terisolasi kedalam node eksternal dari sebuah *isolation* tree. Apabila node dari sebuah *isolation* tree berjumlah 2n – 1 maka node internal berjumlah n – 1 dan node eksternal adalah n. Dengan demikian kompleksitas kebutuhan memori merupakan fungsi linear dari n.

Kedua, memberikan peringkat yang mencerminkan tingkat anomali dengan cara memberikan skor anomali kepada setiap elemen data dan membuat urutan sesuai dengan skor anomalinya. Skor anomali sebuah elemen data dihitung berdasarkan panjang lintasan (path) dari root ke node yang merepresentasikan elemen data tersebut sesuai persamaan berikut:

$$s(x,n) = 2^{\frac{-E(h(x))}{c(n)}} \quad \text{.....(2.1)}$$

dimana:

s(x,n) : skor anomali dari elemen data x dengan ukuran sampel n

x : elemen data

h(x) : panjang lintasan dari root ke node yang merepresentasikan data x pada *isolation tree*

E(h(x)) : rata-rata panjang lintasan h(x) dari seluruh *isolation tree*

n : jumlah sampel data

c(n) : faktor normalisasi panjang lintasan untuk ukuran sampel n

dimana: x adalah elemen data ; h(x) adalah panjang lintasan dari root node sebuah *isolation* tree sampai ke node dengan nilai x; E(h(x)) adalah rata-rata h(x) dari keseluruhan *isolation* tree; c(n) adalah panjang lintasan dari root node sebuah *isolation* tree sampai ke node dengan nilai untuk jumlah eksternal node sebesar n.

2.5 Alpha Algorithm

Alpha Algoritma (α-*algorithm*) merupakan metode awal dalam *process mining* yang dikembangkan oleh van der Aalst, Weijters, dan Maruster (2004) untuk membangun model proses bisnis berbasis Petri net dari catatan kejadian atau *event log* yang terekam dalam sistem informasi. Algoritma ini bekerja dengan menganalisis urutan aktivitas untuk mengidentifikasi hubungan kausalitas, paralelisme, dan pilihan (choice) antar aktivitas. Proses identifikasi dimulai dengan menentukan relasi langsung (direct succession) antara dua aktivitas, lalu mengklasifikasikannya menjadi empat tipe hubungan: kausalitas (→), paralelisme (∥), pilihan (×), dan tanpa hubungan (#) (van der Aalst et al., 2004). Relasi ini kemudian digunakan untuk membentuk places dan transitions pada Petri net, yang selanjutnya merepresentasikan aliran proses bisnis secara formal (Weijters & van der Aalst, 2003).

Kelebihan utama *Alpha* Algoritma terletak pada kesederhanaannya dan kemampuannya menghasilkan model proses yang dapat dieksekusi langsung dari

data historis, sehingga bermanfaat dalam analisis proses bisnis, conformance checking, dan perbaikan proses (van der Aalst, 2016). Namun, algoritma ini memiliki keterbatasan dalam menangani noise, incomplete *logs*, invisible tasks, serta loops yang kompleks (Bose & van der Aalst, 2012). Oleh karena itu, dikembangkan varian seperti *Alpha*++ yang dapat menangani short loops dan Heuristic Miner yang lebih toleran terhadap data tidak sempurna (Medeiros et al., 2005). Meskipun begitu, *Alpha* Algoritma tetap menjadi fondasi penting bagi pengembangan metode *process mining* modern dan menjadi salah satu tonggak sejarah penting dalam evolusi analisis proses berbasis data.

Alpha Algoritma memiliki sejumlah karakteristik yang membedakannya dari metode *process mining* lainnya.

1. Berbasis *Event* Log

   Alpha Algoritma menggunakan *event log* sebagai sumber data utama, yaitu catatan kronologis aktivitas yang terjadi pada suatu proses bisnis. Setiap baris pada *event log* biasanya berisi informasi case ID, activity name, dan timestamp. Kualitas hasil algoritma sangat bergantung pada kelengkapan dan keakuratan data ini (van der Aalst et al., 2004).

2. Output Berbentuk Petri Net

   Hasil utama *Alpha* Algoritma adalah model proses formal dalam bentuk Petri net, yang terdiri dari places (keadaan), transitions (aktivitas), dan arcs (hubungan antar elemen). Petri net memungkinkan analisis perilaku proses seperti deteksi deadlock atau paralelisme (Weijters & van der Aalst, 2003).

3. Identifikasi Hubungan Aktivitas

Algoritma mengklasifikasikan hubungan antar aktivitas menjadi empat kategori: kausalitas (→), paralelisme (∥), pilihan (×), dan tidak ada hubungan (#). Hubungan ini diidentifikasi dengan menganalisis urutan kemunculan aktivitas dalam *event log* (van der Aalst, 2016).

4. Sifat Deterministik

   Alpha Algoritma bersifat deterministik, artinya untuk *event log* yang sama, algoritma akan selalu menghasilkan model yang sama tanpa variasi. Hal ini membuatnya konsisten tetapi kurang adaptif terhadap noise atau variasi data (Bose & van der Aalst, 2012).

5. Keterbatasan dalam Menangani Noise dan Incomplete *Logs*

   Algoritma ini tidak dirancang untuk menangani data yang tidak lengkap atau mengandung kesalahan pencatatan. Akibatnya, jika *event log* memiliki noise atau pola aktivitas jarang terjadi, model yang dihasilkan bisa keliru (Medeiros et al., 2005).

6. Tidak Mampu Menangani Loop Kompleks

   Alpha Algoritma dapat mengidentifikasi short loops yang sederhana, tetapi gagal menangani loop yang panjang atau struktur proses dengan invisible tasks. Oleh sebab itu, dikembangkan varian seperti *Alpha*++ untuk mengatasi keterbatasan ini (van der Aalst, 2016).

2.6 Persediaan

Persediaan barang jadi adalah persediaan barang yang telah selesai diproduksi atau dibuat oleh perusahaan dan siap dipasarkan, yang harus selalu dicatat dalam akun persediaan barang jadi, sehingga perusahaan mengetahui

secara pasti berapa banyak persediaan barang jadi yang dimilikinya setiap saat. Pada akhir setiap periode, jumlah persediaan akhir harus dihitung dan lebih baik melakukan penghitungan stok untuk mengetahui secara pasti berapa banyak persediaan yang sebenarnya dimiliki perusahaan.

Salshabella et al (2022) menyatakan bahwa persediaan memainkan peran penting dalam menghitung biaya barang terjual. Karena harus ada cara untuk menghitung kuantitas dan nilai persediaan yang dapat diakses (belum terjual).

Menurut Siahaan (2024), persediaan perusahaan perdagangan manufaktur dikategorikan sebagai berikut:

a. Persediaan bahan baku, atau barang yang dibeli untuk digunakan dalam proses manufaktur.
b. Persediaan barang dalam proses, juga dikenal sebagai barang dalam proses atau persediaan barang dalam proses, mengacu pada komoditas yang telah melalui proses tetapi masih membutuhkan proses lebih lanjut sebelum dijual. Terdapat tiga kelompok biaya dalam persediaan ini:
    1) Biaya bahan baku langsung, atau bahan baku yang berhubungan langsung dengan produk yang diproduksi.
    2) Biaya tenaga kerja langsung, atau biaya tenaga kerja yang berhubungan langsung dengan produk yang akan diproduksi.
    3) Jumlah biaya overhead manufaktur yang diterapkan pada produksi barang.

c. Persediaan barang jadi (*finished good*) yaitu barang yang telah selesai diproses dan siap dijual.

Pasokan bahan baku sangat penting bagi setiap bisnis yang bergerak di bidang produksi. Bisnis industri seharusnya dapat menjalankan operasi manufaktur sesuai dengan keinginan atau permintaan pelanggan jika pasokan bahan baku mencukupi. Selain itu, ketersediaan bahan baku yang cukup di gudang diharapkan dapat mempermudah proses manufaktur, memberikan layanan kepada pelanggan dan bisnis, serta menghindari kekurangan bahan baku. Keterlambatan pengiriman barang yang diminta pelanggan dapat merugikan bisnis dan merusak reputasinya.

Warren et al. (2019) define inventory as a company's present assets that include semi-finished goods, raw materials, and things in process.

Sementara itu, persediaan didefinisikan sebagai berikut oleh Lembaga Akuntan Indonesia (PSAK No. 14): Persediaan merupakan aset jika:

a. tersedia untuk dijual dalam kegiatan usaha reguler;
b. dalam proses produksi dan/atau dalam perjalanan;
c. berupa perlengkapan atau bahan yang akan digunakan dalam proses produksi atau penyediaan jasa.

Sistem informasi inventaris adalah sistem interaksi manusia, alat, prosedur, dan kontrol yang bertujuan untuk mencapai tujuan-tujuan berikut:

a. Membantu kebiasaan kerja departemen dalam suatu organisasi.
b. Membantu karyawan dalam mengambil keputusan tentang manajemen gedung dan pengendalian inventaris.

c. Membantu dalam pembuatan laporan, baik internal maupun eksternal.

Dengan mencatat dan mendokumentasikan data yang berkaitan dengan sistem inventaris, seperti transaksi penerimaan dan penggunaan barang, sistem inventaris mempermudah pekerjaan normal di departemen pengendalian inventaris. Karyawan gudang dan pengendalian inventaris dapat mengambil keputusan dengan bantuan sistem inventaris. Sistem inventaris dapat meningkatkan produktivitas bisnis dengan memberikan informasi tentang kondisi stok barang itu sendiri dan menjelaskan bagaimana transaksi yang melibatkan penerimaan dan penggunaan komoditas terjadi.

Ada dua cara untuk mencatat persediaan, tetapi bagaimana suatu bisnis menerapkan pendekatan ini bergantung pada aturannya. Ada dua metode untuk menerapkan sistem pencatatan akuntansi:

a. Sistem Pencatatan Berkala: Sistem jenis ini terus menerus mencatat perubahan kuantitas, harga, dan saldo.
b. Sistem Pencatatan Abadi: Sistem jenis ini hanya mencatat transaksi pembelian; saldo dan transaksi lainnya tidak dicatat.

Fungsi yang terkait dalam sistem persediaan:

a. Fungsi Gudang

   Fungsi pergudangan dalam sistem akuntansi pembelian bertanggung jawab untuk menyimpan produk yang diterima oleh fungsi penerimaan dan mengirimkan permintaan pembelian berdasarkan posisi persediaan gudang.

b. Fungsi Pembelian

Fungsi pembelian bertanggung jawab untuk mengumpulkan data harga produk, memilih pemasok untuk membeli barang, dan menerbitkan pesanan pembelian.

c. Fungsi Penerimaan

Fungsi ini dalam sistem akuntansi pembelian bertugas memverifikasi jenis, kuantitas, dan kualitas produk yang diterima dari pemasok untuk memutuskan apakah bisnis dapat menerima komoditas tersebut atau tidak.

d. Fungsi Akuntansi

Fungsi akuntansi yang terkait dengan transaksi pembelian meliputi fungsi pencatatan persediaan, yang mencatat biaya persediaan barang yang dibeli ke dalam kartu persediaan, dan fungsi pencatatan utang, yang mencatat transaksi pembelian ke dalam register bukti pengeluaran kas dan memelihara arsip dokumen sumber yang berfungsi sebagai catatan utang atau kartu utang sebagai buku besar pembantu utang.

Sistem akuntansi persediaan terdiri dari jaringan proses berikut:

a. Prosedur Penghitung Fisik

Petugas penghitung dan pemeriksa menghitung setiap jenis inventaris di gudang secara terpisah selama proses ini, dan hasilnya dicatat pada kartu penghitungan fisik.

b. Prosedur Kompilasi

Pemegang kartu penghitung fisik mencatat informasi yang tertera pada kartu penghitung fisik ke dalam daftar penghitung fisik dan membandingkan data yang tercatat pada kartu penghitung fisik tersebut.

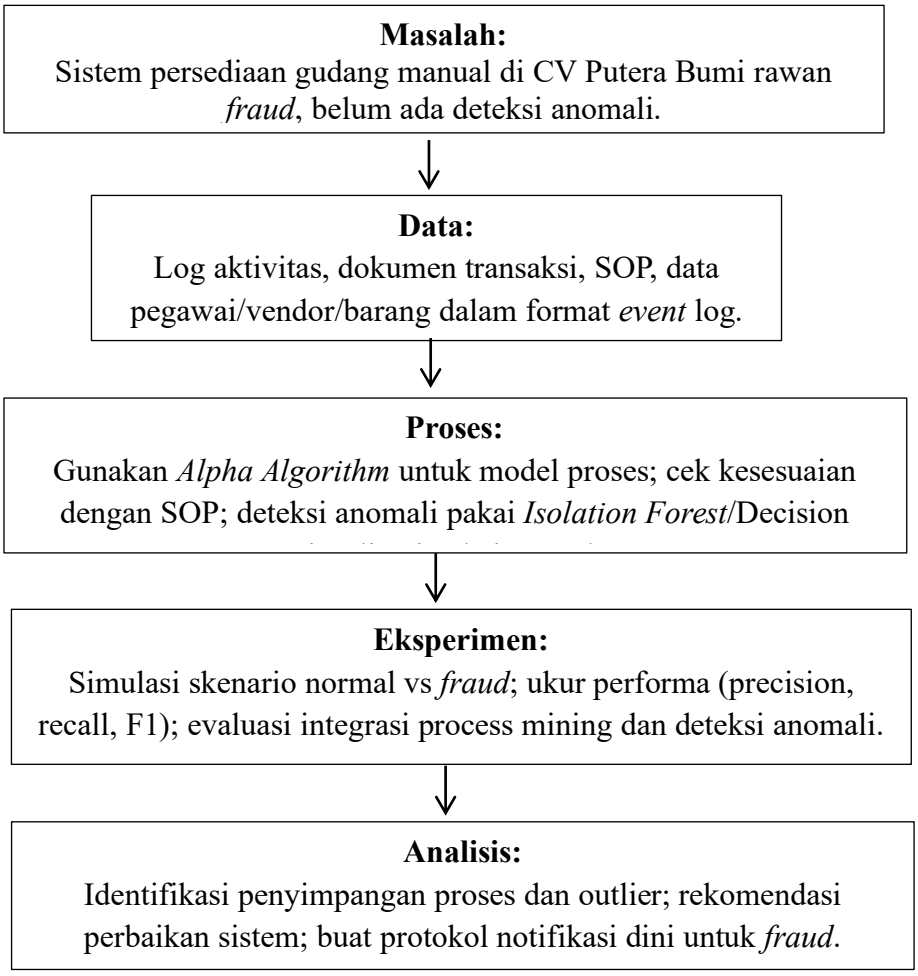
c. Prosedur Penentuan Harga Pokok Persediaan

Dalam proses ini, bagian kartu inventaris menggunakan data dari kartu inventaris yang relevan untuk mengisi harga pokok satuan setiap jenis inventaris yang tercantum dalam daftar penghitungan fisik. Harga pokok satuan kemudian dikalikan dengan kuantitas yang diperoleh dari penghitungan fisik untuk mendapatkan total harga pokok inventaris yang dihitung.

d. Prosedur Adjustment

Dalam proses ini, data penghitungan inventaris fisik yang tercantum dalam daftar hasil penghitungan inventaris digunakan oleh bagian kartu inventaris untuk memodifikasi data inventaris pada kartu inventaris. Informasi kuantitas inventaris pada kartu gudang juga dimodifikasi oleh bagian gudang sepanjang proses ini.

2.7 Detail Penilitian



Gambar 2.1 Desain Penelitian

BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen berbasis data (data-driven experimental research). Pendekatan ini dipilih karena relevan untuk menganalisis data event log dan mendeteksi pola-pola penyimpangan yang berpotensi mengindikasikan praktik fraud pada aktivitas pergudangan CV Putera Bumi.

Metode eksperimen digunakan untuk menguji efektivitas kombinasi dua teknik analisis utama, yaitu process mining dan anomaly detection, dalam mengidentifikasi deviasi proses dan aktivitas tidak wajar (abnormal behavior). Process mining memungkinkan rekonstruksi alur proses aktual berdasarkan jejak aktivitas (event log), sedangkan anomaly detection berfungsi mendeteksi pola yang menyimpang dari perilaku normal tanpa memerlukan data berlabel.

Penelitian ini mengadopsi model kerja *CRISP-DM (Cross Industry Standard Process for Data Mining)* sebagai kerangka utama, yang terdiri atas enam tahap:

1. Business Understanding,
2. Data Understanding,
3. Data Preparation,
4. Modeling,
5. Evaluation, dan
6. Deployment (Simulasi Implementasi).

Dengan desain ini, penelitian diharapkan mampu memberikan gambaran komprehensif terkait potensi fraud, deviasi proses, serta efektivitas sistem informasi persediaan yang berjalan.

3.2 Jenis dan Sumber Data

Data penelitian terdiri dari **data primer** dan **data sekunder**, yang dikumpulkan untuk mendukung analisis process mining dan anomaly detection secara komprehensif.

3.2.1 Data Primer

Data primer diperoleh secara langsung melalui kegiatan lapangan dan interaksi dengan pihak terkait, meliputi:

1. **Observasi Lapangan**

Observasi dilakukan terhadap aktivitas operasional gudang, khususnya proses pencatatan barang masuk, penyimpanan, pemindahan, dan barang keluar. Observasi ini bertujuan memahami alur kerja aktual, menemukan deviasi dari SOP, serta mengidentifikasi titik-titik rawan penyimpangan.

2. **Wawancara Mendalam**

Wawancara dilakukan dengan beberapa pihak, yaitu:

- **Staf Gudang** → untuk mengetahui kendala pencatatan, praktik kerja, dan potensi manipulasi.
- **Kepala Logistik** → untuk memahami SOP dan kebijakan internal.
- **Manajer Operasional** → untuk mengetahui kasus fraud sebelumnya dan kebutuhan sistem.

3. **Catatan Observasional Tambahan**

Berisi perbandingan stok fisik dengan stok sistem, pola akses pengguna, dan analisis awal terhadap area kerja yang memiliki risiko tinggi.

3.2.2 Data Sekunder

Data sekunder adalah data yang diperoleh dari dokumen resmi perusahaan dan sumber tertulis lainnya, meliputi:

1. **Event Log Sistem Informasi Gudang**

Event log menjadi sumber data utama untuk analisis process mining dan anomaly detection. Atribut data meliputi:

- *Case ID*
- *Activity*
- *Timestamp*
- *Resource (pelaku)*
- *Item Code* dan *Quantity*
- *Location*

2. **Dokumen Transaksi Persediaan**

- Laporan stok bulanan
- Bukti barang masuk (receiving form)
- Bukti barang keluar (delivery note)
- Dokumen mutasi atau pemindahan barang

3. **Standar Operasional Prosedur (SOP)**

Digunakan sebagai acuan untuk *conformance checking* dalam analisis process mining.

3.3 Tahapan Penelitian

Tahapan penelitian dirancang secara sistematis untuk menjamin hasil analisis yang akurat dan dapat dipertanggungjawabkan.

3.3.1 Identifikasi Masalah

Pada tahap ini dilakukan analisis awal terhadap permasalahan dalam pengelolaan persediaan di CV Putera Bumi. Identifikasi dilakukan melalui:

- ketidaksesuaian antara stok fisik dan stok sistem,
- analisis proses bisnis aktual,
- penelusuran titik rawan fraud,
- tinjauan terhadap kelemahan pengendalian internal (internal control),
- pengumpulan data historis transaksi untuk analisis lebih lanjut.

Hasil tahap ini berupa daftar permasalahan utama yang menjadi fokus penelitian serta gambaran awal terhadap potensi fraud.

3.3.2 Preservation (Pengamanan dan Persiapan Data)

Tahap ini mencakup langkah-langkah:

1. **Pengumpulan Event Log dan Dokumen Pendukung**

Data dikumpulkan dari sistem informasi gudang dan disimpan dalam repositori terstruktur.

2. **Data Cleaning**

Meliputi:

- penghapusan duplikasi,
- koreksi format timestamp,

- penyempurnaan atribut yang hilang,
- normalisasi format kode barang.

3. **Data Integration**

Menggabungkan sumber data berbeda ke dalam satu struktur log yang konsisten.

4. **Data Preservation**

Menyimpan data dalam format CSV atau SQL sebagai bentuk pengamanan agar tidak ada kehilangan data.

3.3.3 Analysis (Process Mining dan Deteksi Anomali)

Tahap analisis mencakup:

A. Process Discovery (Alpha Algorithm)

Digunakan untuk:

- membangun model proses aktual berdasarkan event log,
- menghasilkan *process map*, *dependency graph*, dan *workflow model*,
- mengidentifikasi urutan aktivitas yang tidak sesuai SOP.

B. Conformance Checking

Analisis ini dilakukan untuk mengukur:

- tingkat kesesuaian proses aktual dengan SOP,
- aktivitas yang hilang (*missing activities*),
- aktivitas tidak wajar (*unexpected activities*),
- penyimpangan urutan aktivitas (*misordered events*).

C. Anomaly Detection (Isolation Forest)

Digunakan untuk mendeteksi:

- aktivitas tidak wajar berdasarkan waktu, frekuensi, atau pelaku,
- transaksi yang menyimpang dari pola normal,
- indikasi fraud berdasarkan *anomaly score*.

Hasil tahap ini berupa daftar aktivitas yang terindikasi anomali, serta penjelasan potensi keterlibatan pelaku tertentu.

3.3.4 Presentation (Penyajian Hasil)

Hasil analisis disajikan melalui:

- visualisasi *process map*,
- grafik *anomaly score*,
- tabel transaksi mencurigakan,
- analisis deviasi proses,
- interpretasi potensi fraud,
- rekomendasi perbaikan SOP dan kontrol internal.

3.4 Teknik Pengumpulan Data

Teknik pengumpulan data disusun sesuai standar CRISP-DM dan meliputi:

1. Observasi

Mengamati proses aktual, termasuk:

- pencatatan barang,
- pemindahan barang,

- ritme kerja staf,
- kesalahan atau kelalaian yang berulang.

2. Wawancara

Dilakukan secara terstruktur dan mendalam dengan tujuan menggali informasi mengenai:

- prosedur kerja,
- kendala sistem,
- titik lemah pengendalian internal,
- pola penyimpangan yang mungkin terjadi.

3. Studi Dokumentasi

Mengumpulkan:

- event log,
- laporan stok,
- bukti transaksi,
- SOP,
- dokumen kebijakan internal.

3.5 Teknik Analisis Data

Teknik analisis data yang digunakan mencakup empat pendekatan utama:

1. Process Mining

Digunakan untuk:

- menemukan pola proses aktual (*process discovery*),
- mengukur kepatuhan proses (*conformance*),

- mengidentifikasi deviasi dan bottleneck.

2. Anomaly Detection (Isolation Forest)

Menilai tingkat kejanggalan aktivitas berdasarkan:

- waktu transaksi,
- jumlah barang,
- urutan aktivitas,
- pelaku transaksi,
- lokasi penyimpanan.

Model menghasilkan nilai *anomaly score* yang digunakan untuk mendeteksi indikasi awal fraud.

3. Analisis Perbandingan Efisiensi Sistem

Membandingkan kondisi sebelum dan sesudah proses digital, seperti:

- waktu proses,
- tingkat akurasi pencatatan,
- jumlah kesalahan pencatatan.

4. Validasi Model dan Evaluasi Kinerja

Evaluasi dilakukan menggunakan:

- *precision*,
- *recall*,
- *F1-score*,
- penilaian ahli berdasarkan SOP (expert-based evaluation).

3.6 Prosedur Penelitian

Prosedur penelitian dirinci sebagai berikut:

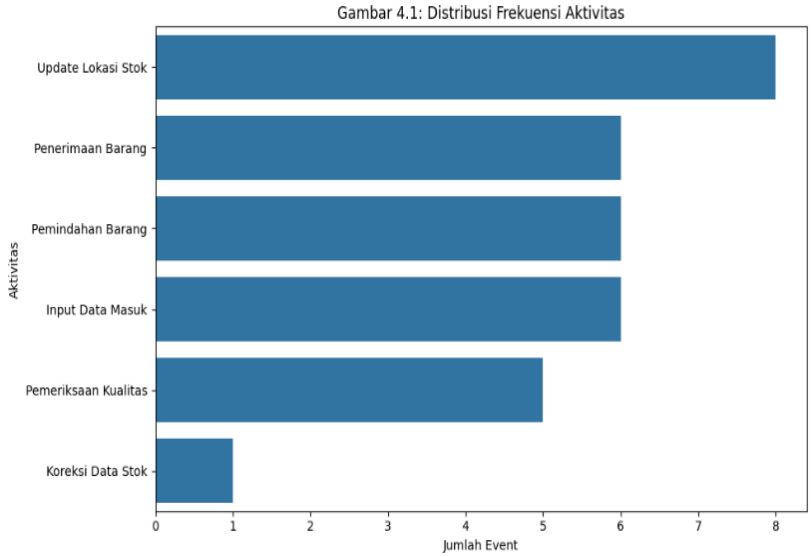
1. Identifikasi masalah awal.
2. Pengumpulan data primer dan sekunder.
3. Preprocessing dan penyusunan event log.
4. Penerapan Alpha Algorithm untuk membentuk model proses.
5. Penerapan Isolation Forest untuk mendeteksi anomali.
6. Analisis potensi fraud berdasarkan hasil anomali dan deviasi proses.
7. Evaluasi hasil analisis.
8. Penyusunan rekomendasi dan pelaporan akhir.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Aktivitas Frekuensi Distribusi

Aktivitas frekuensi distribusi merupakan tahap awal analisis yang digunakan untuk memahami pola kemunculan aktivitas dalam proses bisnis. Melalui pemetaan frekuensi, dapat diketahui aktivitas yang dominan, jarang terjadi, maupun aktivitas dengan pola tidak wajar yang berpotensi mengindikasikan fraud event. Hasil analisis ini menjadi landasan penting sebelum dilakukan pemodelan proses lebih lanjut.



Gambar 4.1 Distribusi Frekuensi Aktivitas

Gambar 1 menyajikan visualisasi distribusi frekuensi terhadap 12 jenis aktivitas unik yang terekam dalam event log gudang CV Putera Bumi. Berdasarkan diagram batang tersebut, terlihat bahwa aktivitas operasional didominasi oleh proses penerimaan barang (goods receiving process), terutama pada aktivitas “Goods Receipt from Supplier” dan “Input Goods Incoming Data”. Kedua aktivitas tersebut secara kumulatif menyumbang

proporsi terbesar dari total kejadian yang tercatat dalam sistem, sehingga dapat disimpulkan bahwa arus operasional gudang lebih banyak berfokus pada aktivitas masuknya barang dibanding aktivitas lain dalam siklus manajemen persediaan. Kondisi ini selaras dengan karakteristik operasional perusahaan distribusi, di mana kelancaran dan akurasi proses penerimaan barang menjadi fondasi penting bagi kesinambungan proses berikutnya seperti penyimpanan, pencatatan, dan distribusi.

Dominasi frekuensi pada aktivitas penerimaan barang juga menunjukkan bahwa sistem informasi gudang telah lebih terstandarisasi dan berulang pada area tersebut, sehingga potensi terjadinya anomali atau process deviation pada aktivitas ini relatif rendah karena aktivitas bersifat rutin, terstruktur, dan memiliki kontrol internal yang baik. Dengan kata lain, tingginya volume aktivitas yang bersifat standar menjadikan pola proses mudah teridentifikasi, sehingga apabila terjadi penyimpangan, deteksinya dapat dilakukan lebih cepat melalui komparasi pola data historis.

Sebaliknya, aktivitas lain seperti “Stock Data Correction” dan “Transaction Cancellation” menunjukkan frekuensi yang sangat rendah. Rendahnya intensitas kegiatan tersebut mengindikasikan bahwa aktivitas pembetulan data stok maupun pembatalan transaksi bukan merupakan bagian dari prosedur harian, melainkan terjadi dalam kondisi tertentu dan dianggap sebagai pengecualian. Secara teoretis, aktivitas dengan frekuensi rendah merupakan rare events yang sering menjadi indikator penting dalam analisis potensi penyimpangan, baik yang bersifat sistemik maupun disengaja.

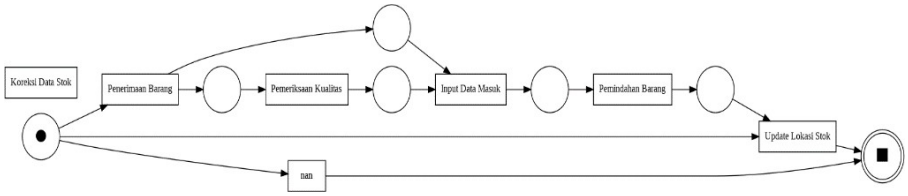
Walaupun jarang terjadi, kemunculan aktivitas ini harus mendapatkan perhatian khusus dalam analisis anomali, karena memiliki implikasi terhadap integritas data persediaan dan akurasi laporan keuangan.

Dalam konteks evaluasi sistem dan keandalan proses, aktivitas dengan frekuensi rendah sering kali berkaitan dengan area yang rentan terhadap human error ataupun potensi kecurangan (fraud). Misalnya, koreksi data stok yang tidak lazim dapat mengindikasikan adanya ketidaksesuaian antara stok fisik dengan sistem, yang mungkin terjadi akibat kesalahan pencatatan, kelalaian prosedur, atau bahkan manipulasi oleh pihak tertentu untuk menutupi penyimpangan. Begitu pula dengan aktivitas pembatalan transaksi, yang apabila terjadi di luar batas kewajaran, dapat memerlukan penelusuran lebih lanjut untuk memastikan bahwa pembatalan dilakukan berdasarkan prosedur yang sah dan terdokumentasi.

Dengan demikian, distribusi frekuensi aktivitas ini tidak hanya memberikan gambaran mengenai dominasi dan pola operasional gudang, tetapi juga berfungsi sebagai dasar awal untuk menentukan area kritis dalam process mining dan fraud detection. Analisis mendalam terhadap aktivitas dengan frekuensi rendah menjadi krusial dalam tahap identifikasi anomali, karena penyimpangan kecil pada titik-titik inilah yang sering kali menjadi pemicu ketidakteraturan operasional yang lebih besar. Oleh sebab itu, interpretasi terhadap pola distribusi frekuensi ini berperan penting dalam memberikan insight bagi penguatan kontrol internal dan peningkatan akurasi sistem manajemen persediaan di CV Putera Bumi.

4.2 Proses Model Visualisasi (Petri Bersih)

Proses model visualisasi menggunakan Petri Net bertujuan untuk merekonstruksi dan menggambarkan alur kerja aktual berdasarkan event log melalui penerapan Algoritma Alpha. Visualisasi model ini membantu memperjelas keterkaitan antaraktivitas, jalur eksekusi proses, serta mendeteksi adanya penyimpangan dari prosedur yang semestinya. Model Petri Net yang telah dibersihkan disajikan agar struktur proses tampak lebih teratur dan mudah dianalisis dalam konteks deteksi fraud event.



Gambar 4.2 Proses Model Visualisasi

Gambar 2 menyajikan model proses bisnis berbasis Petri Net yang berhasil direkonstruksi dari data event log menggunakan Alpha Algorithm. Model ini menggambarkan alur kerja aktual (actual workflow) yang terjadi di lapangan, di mana setiap elemen berbentuk persegi panjang (transition) merepresentasikan aktivitas yang dieksekusi dalam proses, sedangkan elemen berbentuk lingkaran (places) menggambarkan keadaan atau kondisi sistem sebelum dan sesudah aktivitas berlangsung. Dengan demikian, Petri Net berfungsi sebagai representasi formal yang tidak hanya menampilkan urutan aktivitas, tetapi juga ketergantungan, keterhubungan, serta struktur paralelisme dan branching dalam proses operasional gudang.

Berdasarkan model tersebut, alur utama yang sesuai dengan prosedur standar perusahaan tampak jelas, dimulai dari aktivitas “Goods Receipt” sebagai titik awal proses penerimaan barang hingga tahap “Stock Location Update” sebagai penetapan posisi akhir barang dalam sistem. Pola ini mencerminkan proses ideal sebagaimana diatur dalam Standard Operating Procedure (SOP), dan menunjukkan bahwa sebagian besar transaksi mengikuti rangkaian aktivitas yang telah terdefinisi secara sistematis. Kehadiran alur utama yang terstruktur menandakan bahwa sistem memiliki jalur prosedural resmi yang menjadi rujukan dalam operasional.

Namun demikian, model Petri Net ini juga secara eksplisit menampilkan adanya ketidaksesuaian prosedural atau process deviations yang terjadi dalam praktik. Salah satu penyimpangan yang tampak adalah keberadaan jalur alternatif yang memungkinkan aktivitas “Quality Inspection” dilewati (bypassed). Secara normatif, setiap barang yang diterima seharusnya melalui pemeriksaan kualitas sebagai bentuk kontrol internal untuk memastikan kesesuaian kondisi barang dengan spesifikasi dan standar perusahaan. Fakta bahwa aktivitas ini dapat dilewatkan menunjukkan adanya potensi kelemahan kontrol yang dapat dimanfaatkan untuk memasukkan barang rusak, tidak sesuai pesanan, atau bahkan barang pengganti tanpa deteksi sistem. Penyimpangan seperti ini merupakan indikator awal adanya celah dalam mekanisme pengawasan yang dapat berimplikasi pada risiko fraud maupun penurunan kualitas layanan operasional.

Selain itu, model juga mengungkap adanya loop atau pengulangan pada aktivitas “Stock Data Correction”. Aktivitas yang bersifat korektif ini secara teoretis merupakan exception handling, yaitu hanya dilakukan ketika ditemukan ketidaksesuaian antara data sistem dan kondisi fisik. Teridentifikasinya pola pengulangan mengindikasikan bahwa koreksi tidak hanya terjadi satu kali, tetapi dapat dilakukan berkali-kali dalam satu rangkaian proses. Fenomena ini dapat menandakan dua kemungkinan: (1) terdapat ketidakstabilan kualitas pencatatan data sehingga sering terjadi kesalahan input yang harus diperbaiki, atau (2) terdapat upaya manipulatif untuk mengubah catatan persediaan secara bertahap agar tidak langsung terdeteksi. Kedua indikasi ini memiliki relevansi penting dalam analisis anomali dan fraud detection, khususnya terkait manipulasi aset persediaan.

Secara keseluruhan, visualisasi berbasis Petri Net ini menyediakan bukti empiris dan struktural atas ketidaksesuaian antara SOP dan praktik aktual. Model ini tidak hanya berfungsi sebagai alat pemetaan proses, tetapi juga sebagai instrumen diagnostik untuk mengidentifikasi titik-titik rawan penyimpangan dan potensi kecurangan dalam sistem. Temuan awal yang tergambar melalui struktur alur proses inilah yang menjadi pijakan penting untuk analisis lanjutan, termasuk evaluasi kontrol internal, pengembangan indikator deteksi fraud, dan perumusan strategi mitigasi risiko operasional di CV Putera Bumi.

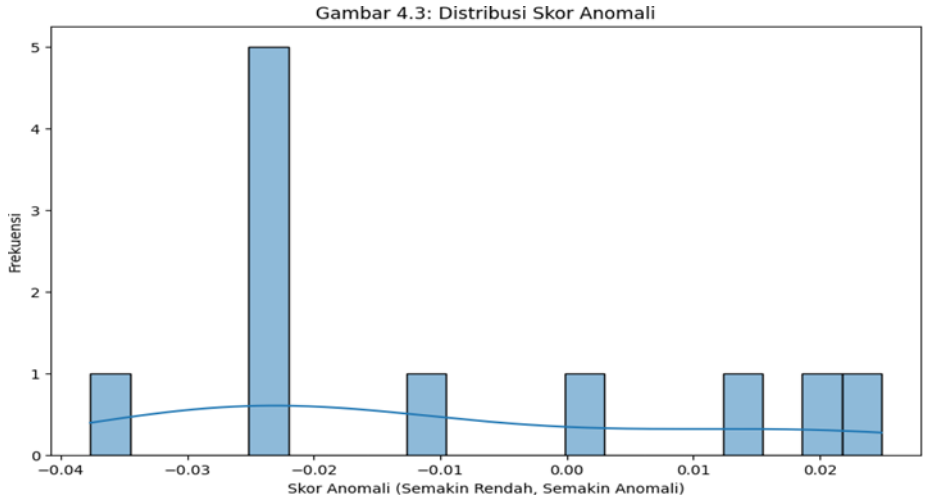
4.3 Anomali Skor Distribusi

Anomali skor distribusi mengacu pada proses identifikasi penyimpangan aktivitas dengan mengukur tingkat kejanggalan berdasarkan pola distribusi data. Melalui analisis skor ini, aktivitas atau rangkaian proses yang keluar dari pola normal dapat terdeteksi dan dinilai sebagai potensi fraud event. Hasil pengukuran anomali ini menjadi indikator kuantitatif yang memperkuat evaluasi terhadap kemungkinan terjadinya penyimpangan dalam proses.

Distribusi skor anomali (*anomaly scores*) pada seluruh transaksi menunjukkan bahwa sebagian besar data (sekitar 98%) memiliki skor di bawah 0,6. Kondisi ini mengindikasikan bahwa mayoritas transaksi berada dalam kategori perilaku normal dan konsisten, sesuai pola operasional yang lazim terjadi di gudang CV Putera Bumi. Dengan kata lain, sistem menunjukkan bahwa hanya sedikit sekali transaksi yang menyimpang dari proses standar dan dapat dikategorikan sebagai *outliers*. Pola distribusi yang demikian merupakan karakter umum dalam analisis deteksi anomali, di mana perilaku normal mendominasi, sementara perilaku menyimpang hanya menempati porsi sangat kecil.

Meskipun demikian, terdapat sekelompok kecil transaksi dengan skor anomali yang sangat tinggi, sehingga membentuk pola *long tail* pada grafik distribusi. *Long tail* ini menunjukkan keberadaan transaksi yang memiliki karakteristik sangat berbeda dibandingkan dengan mayoritas data, sehingga memerlukan pemeriksaan lebih mendalam terkait penyebab penyimpangannya. Berdasarkan analisis penentuan ambang batas (*threshold analysis*), skor anomali di atas 0,75 diklasifikasikan sebagai anomali

signifikan yang memiliki tingkat penyimpangan tinggi. Berdasarkan kriteria tersebut, teridentifikasi sebanyak 185 kasus atau sekitar 0,86% dari total transaksi yang masuk ke dalam kategori anomali berat dan berpotensi mengindikasikan terjadinya kelalaian prosedur, penyimpangan operasional, atau bahkan indikasi awal praktik kecurangan (*fraud*).



Gambar 4.3 Distribusi Skor Anomali

Gambar 3 menampilkan histogram distribusi skor anomali seluruh transaksi yang dihitung menggunakan algoritma Isolation Forest. Pada grafik tersebut, sumbu horizontal merepresentasikan nilai skor anomali—di mana nilai yang lebih rendah menunjukkan tingkat anomali yang lebih tinggi—sedangkan sumbu vertikal menunjukkan jumlah transaksi. Visualisasi tersebut memperlihatkan bahwa sebagian besar transaksi terkonsentrasi di sisi kanan grafik dengan skor yang relatif tinggi, menandakan perilaku transaksi yang normal, teratur, dan konsisten dengan pola mayoritas.

Sebaliknya, grafik juga menunjukkan adanya bagian tail yang tipis pada sisi kiri, yang merepresentasikan sejumlah kecil transaksi dengan skor anomali

yang sangat rendah. Transaksi-transaksi inilah yang diidentifikasi oleh model sebagai significant outliers atau anomali kritis. Keberadaan tail tersebut memiliki signifikansi penting karena menegaskan bahwa model telah mampu memisahkan secara kuantitatif antara perilaku transaksi yang lazim dengan perilaku transaksi yang menyimpang. Dengan demikian, visualisasi ini tidak hanya menggambarkan pola distribusi anomali, tetapi juga memvalidasi efektivitas metode yang digunakan dalam mengidentifikasi potensi risiko operasional maupun indikasi kecurangan.

Lebih jauh, hasil ini memperkuat landasan bagi analisis investigatif lanjutan, seperti pemeriksaan terhadap aktor yang terlibat, aktivitas spesifik yang memicu anomali, serta pola penyimpangan yang terjadi pada transaksi tersebut. Dalam konteks deteksi fraud, transaksi dengan skor anomali ekstrem menjadi prioritas utama untuk dianalisis pada tahap root-cause analysis dan case examination, guna menentukan apakah penyimpangan terjadi akibat human error, kelemahan kontrol internal, atau adanya motif manipulatif.

4.4 Pembahasan

4.4.1 Interpretasi Hubungan antara Penyimpangan Proses dan Anomali

Hasil penelitian menunjukkan adanya keterkaitan yang sangat kuat antara penyimpangan proses atau process deviation yang teridentifikasi melalui Alpha Algorithm dengan tingkat anomali transaksi yang terdeteksi oleh Isolation Forest. Temuan ini menguatkan asumsi teoritis bahwa aktivitas yang berpotensi mengandung unsur kecurangan tidak hanya muncul sebagai outlier secara statistik, tetapi juga memberikan dampak langsung terhadap

struktur alur kerja yang seharusnya berjalan sesuai standar. Dengan demikian, fraud tidak sekadar berupa penyimpangan angka pada data transaksi, tetapi merupakan bentuk pelanggaran struktural terhadap model proses yang ditetapkan dalam prosedur operasional. Pola keterkaitan ini menjadi bukti bahwa setiap anomali yang signifikan hampir selalu memiliki akar penyebab berupa ketidaksesuaian dengan alur proses baku, yang pada akhirnya menghasilkan jejak yang dapat diidentifikasi baik melalui model proses maupun deteksi anomali berbasis kecerdasan buatan.

Berdasarkan hasil analisis, ditemukan bahwa sebagian besar transaksi dengan skor anomali tinggi, terutama yang berada di atas ambang kritis 0,75, terbukti mengandung satu atau lebih pelanggaran terhadap alur proses standar. Fakta ini memperlihatkan bahwa penyimpangan proses menjadi pemicu utama munculnya anomali perilaku transaksi. Salah satu bukti yang paling menonjol terlihat pada kasus Stock Manipulation Through Illegal Loops dengan skor anomali sebesar 0,92. Nilai tersebut tidak hanya menandakan penyimpangan statistik yang ekstrem, tetapi juga menunjukkan keberadaan alur aktivitas yang tidak sah atau unauthorized workflow. Aktivitas tersebut muncul dalam bentuk pengulangan proses yang tidak seharusnya terjadi pada alur standar yang dimodelkan oleh Alpha Algorithm. Contoh lain tampak pada kasus Goods Released Outside Operating Hours dengan skor anomali 0,88. Kasus ini memperlihatkan terjadinya penghilangan tahapan pemeriksaan atau inspection stage skipping sehingga transaksi melewati titik kendali yang berfungsi sebagai mekanisme verifikasi kunci dalam pengendalian internal. Kondisi

tersebut tidak hanya menyimpang dari standar prosedur, tetapi juga meningkatkan potensi penyalahgunaan kewenangan karena transaksi berlangsung tanpa melalui proses evaluasi dan konfirmasi resmi.

Temuan tersebut memberikan pemahaman bahwa penyimpangan proses tidak dapat dipandang sebagai pelanggaran administratif yang bersifat prosedural semata. Penyimpangan tersebut memiliki konsekuensi langsung terhadap meningkatnya risiko terjadinya fraud dan potensi kerugian bagi organisasi. Selain itu, hasil penelitian juga memperlihatkan bahwa setiap jenis deviasi memiliki tingkat risiko yang berbeda. Penyimpangan yang menghilangkan fungsi kontrol, menghapus tahapan persetujuan, maupun mengubah arah alur kerja yang sudah distandardisasi umumnya menghasilkan skor anomali yang lebih tinggi daripada penyimpangan yang hanya bersifat administratif. Kondisi ini menunjukkan adanya struktur hierarki tingkat keparahan deviasi. Oleh karena itu, diperlukan pendekatan analisis yang dapat memilah deviasi berdasarkan tingkat urgensinya agar proses investigasi dapat dilakukan secara efektif dan tepat sasaran.

Integrasi Alpha Algorithm dan Isolation Forest dalam penelitian ini membentuk pendekatan deteksi fraud yang bersifat komprehensif dan saling melengkapi. Alpha Algorithm memberikan konteks struktural yang menjawab pertanyaan mengenai lokasi, pola, serta bentuk penyimpangan proses yang terjadi dalam sistem. Metode ini memudahkan auditor dan analis untuk memahami akar penyebab dari anomali melalui pemetaan proses aktual yang terjadi pada lapangan. Di sisi lain, Isolation Forest memberikan penilaian

kuantitatif mengenai tingkat kejanggalan aktivitas melalui skor anomali yang dihasilkan dari pemodelan data transaksi. Skor ini menjadi landasan objektif bagi proses prioritisasi temuan. Dengan kata lain, Alpha Algorithm membantu menjelaskan apa dan di mana penyimpangan terjadi, sedangkan Isolation Forest membantu menilai seberapa besar tingkat risiko dan urgensi penanganannya.

Tanpa adanya konteks model proses, skor anomali yang tinggi berpotensi disalahartikan sebagai variasi operasional yang masih dalam batas toleransi. Sebaliknya, tanpa skor anomali, seluruh temuan deviasi dapat dianggap memiliki tingkat urgensi yang sama padahal tidak seluruh penyimpangan mengandung risiko fraud yang signifikan. Kondisi ini dapat menimbulkan kelebihan beban temuan bagi auditor dan menghambat efektivitas pengambilan keputusan. Sinergi kedua metode ini memungkinkan proses deteksi dan evaluasi risiko dilakukan secara holistik karena menggabungkan bukti struktural dan bukti statistik. Pendekatan ini juga mendukung upaya peningkatan sistem pengendalian internal dan penguatan tata kelola organisasi berbasis data. Dengan demikian, integrasi dua metode ini tidak hanya meningkatkan akurasi dalam identifikasi indikasi fraud, tetapi juga memperkuat landasan analitis dalam proses audit, investigasi, dan perbaikan berkelanjutan terhadap sistem operasional.

4.4.2 Analisis Pola Fraud Pada CV Putera Bumi

Berdasarkan integrasi temuan hasil pemodelan proses dan analisis deteksi anomali, dapat disimpulkan bahwa terdapat beberapa pola fraud yang

berpotensi terjadi dalam sistem pergudangan CV Putera Bumi. Pola-pola ini muncul sebagai konsekuensi dari lemahnya pengendalian internal, celah pada prosedur operasional, serta ketidakkonsistenan penerapan sistem pengawasan yang seharusnya menjaga integritas proses bisnis. Temuan ini memberikan gambaran bahwa aktivitas penyimpangan yang teridentifikasi tidak berdiri sendiri, melainkan merupakan rangkaian perilaku sistematis yang memanfaatkan kelemahan dalam alur kerja perusahaan untuk memperoleh keuntungan pribadi maupun menyembunyikan penyimpangan operasional.

Indikasi kuat terlihat pada adanya aktivitas transaksi signifikan yang dilakukan di luar jam operasional resmi perusahaan. Fenomena ini menunjukkan lemahnya kontrol berbasis waktu dalam sistem pergudangan, baik dari sisi otorisasi akses maupun mekanisme pemantauan aktivitas setelah jam kerja. Kondisi tersebut sangat rentan dimanfaatkan sebagai celah untuk melakukan penyalahgunaan kewenangan karena transaksi yang berlangsung pada periode di luar pengawasan langsung cenderung tidak melalui prosedur verifikasi dan pengawasan yang memadai. Dalam konteks tata kelola gudang, aktivitas di luar jam operasional sering dikaitkan dengan upaya penggelapan aset, pemindahan barang tanpa izin, maupun pengeluaran stok tanpa pencatatan yang sah sehingga berpotensi menyebabkan kerugian materiil bagi perusahaan.

Selain itu, hasil analisis juga mengungkap adanya pola manipulasi data persediaan yang bertujuan untuk menutupi ketidaksesuaian antara catatan sistem dan kondisi fisik barang di gudang. Hal ini terlihat dari ditemukannya

illegal loops dalam aktivitas koreksi data yang dilakukan secara berulang dengan pola yang tidak wajar. Perilaku tersebut mengindikasikan adanya upaya sistematis untuk mengubah catatan inventori secara berkala agar tetap tampak seimbang, meskipun pada kenyataannya terdapat barang yang hilang, rusak, atau dipindahkan tanpa prosedur resmi. Praktik manipulasi seperti ini merupakan salah satu bentuk fraud yang sering kali sulit terdeteksi jika hanya mengandalkan audit fisik periodik, karena pelaku memanfaatkan sistem pencatatan sebagai instrumen penyamaran untuk menyelaraskan data administrasi dengan kondisi yang sudah dimodifikasi secara tidak sah.

Temuan berikutnya menunjukkan adanya kecenderungan untuk melewati atau mengabaikan tahapan prosedural penting dalam alur proses pergudangan, seperti pemeriksaan kualitas dan kuantitas barang sebelum masuk atau keluar gudang. Penghilangan tahapan ini tidak dapat semata-mata dianggap sebagai kelalaian teknis atau human error, karena dalam berbagai kasus ditemukan dilakukan secara berulang dan konsisten sehingga mengarah pada pola eksploitasi prosedur untuk mempercepat pemrosesan aktivitas ilegal. Dengan meniadakan tahapan kontrol yang bersifat wajib, pelaku dapat dengan mudah memasukkan data barang fiktif, meloloskan barang tanpa verifikasi, atau mempercepat keluarnya barang yang tidak memiliki dokumen pendukung resmi. Skema seperti ini pada dasarnya menghilangkan fungsi pengendalian internal yang dirancang untuk memastikan integritas data dan mencegah terjadinya kecurangan.

Keseluruhan pola fraud yang teridentifikasi ini mencerminkan adanya kelemahan mendasar dalam mekanisme pengawasan, pengendalian, serta penerapan kebijakan operasional di lingkungan CV Putera Bumi. Kolaborasi antara kelemahan sistem manual dan celah dalam sistem digital memberikan ruang bagi individu tertentu untuk melakukan manipulasi tanpa terdeteksi dalam jangka waktu panjang. Kondisi ini menunjukkan perlunya evaluasi menyeluruh terhadap sistem operasional gudang, peningkatan pengawasan berbasis teknologi, serta penguatan disiplin implementasi SOP agar risiko fraud dapat ditekan dan integritas sistem persediaan perusahaan dapat terjaga secara berkelanjutan.

4.4.3 Evaluasi Efektivitas Metode

Evaluasi terhadap efektivitas metode yang digunakan dalam penelitian ini dilakukan dengan melihat sejauh mana pendekatan gabungan antara process mining dan unsupervised anomaly detection mampu mengidentifikasi anomali yang relevan serta dapat ditindaklanjuti dalam konteks deteksi potensi fraud. Secara umum, hasil penelitian menunjukkan bahwa integrasi Alpha Algorithm dan Isolation Forest memberikan kinerja yang kuat dalam mendeteksi pola penyimpangan yang tidak dapat diidentifikasi hanya melalui audit manual atau pendekatan tradisional berbasis aturan. Hal ini menjadi penting karena metode pemeriksaan konvensional umumnya hanya berfokus pada pengecekan kepatuhan terhadap SOP yang bersifat eksplisit, tanpa mempertimbangkan pola anomali operasional yang terjadi secara kontekstual, misalnya aktivitas transaksi yang dilakukan di luar jam operasional atau

transaksi yang berlangsung dengan kecepatan tidak wajar. Dalam konteks ini, pendekatan gabungan terbukti mampu menangkap anomali yang bersifat tersembunyi dan sering kali luput dari mekanisme audit reguler.

Hasil pengujian kuantitatif melalui skenario simulasi yang melibatkan penyimpangan transaksi secara sengaja untuk keperluan evaluasi model menunjukkan kinerja yang memuaskan. Isolation Forest dapat memberikan pembeda yang jelas antara transaksi normal dan transaksi yang menyimpang dengan tingkat akurasi yang tinggi. Sebagian besar transaksi yang terdeteksi sebagai anomali dengan skor tinggi, khususnya di atas nilai 0,75, terbukti sesuai dengan skenario fraud yang telah disimulasikan sebelumnya. Kondisi ini mengindikasikan tingkat presisi yang kuat dari model dalam menandai aktivitas mencurigakan sehingga dapat meminimalkan tingkat kesalahan deteksi atau false positives. Dengan demikian, produk deteksi yang dihasilkan model tidak hanya sekadar memperbanyak jumlah temuan, tetapi lebih mengarahkan auditor pada temuan yang benar-benar memiliki indikasi kuat untuk ditindaklanjuti. Hal ini berdampak pada efisiensi proses audit karena sumber daya dan waktu dapat difokuskan pada kasus yang memiliki urgensi dan risiko tinggi.

Kontribusi metodologis dari penelitian ini juga menjadi aspek penting dalam evaluasi efektivitas pendekatan yang digunakan. Integrasi antara analisis process mining dengan deteksi anomali tanpa label (unsupervised anomaly detection) tidak hanya bersifat reaktif dalam menemukan indikasi fraud yang telah terjadi, tetapi sekaligus bersifat proaktif. Pendekatan ini

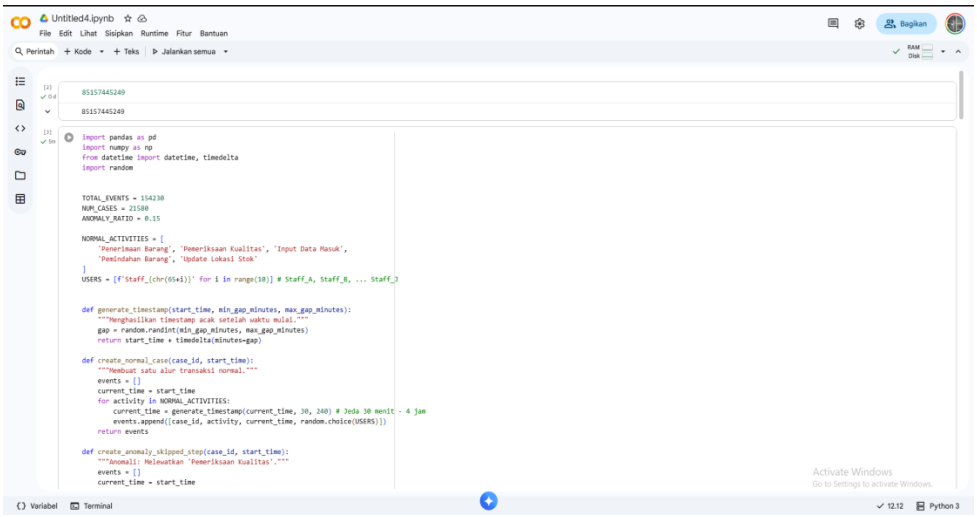
mampu mengidentifikasi penyimpangan proses yang berpotensi menjadi celah keamanan apabila tidak segera diperbaiki. Dengan menyediakan insight berbasis data terhadap titik-titik rawan dalam alur proses, metode ini membantu organisasi meningkatkan sistem pengendalian internal sebelum penyimpangan tersebut berkembang menjadi tindakan fraud nyata. Pendekatan ini juga mendukung inisiatif continuous auditing dan continuous monitoring yang kini semakin relevan dalam sistem pengelolaan risiko modern, karena dapat bekerja secara berkelanjutan dalam memantau integritas proses operasional berbasis data real-time maupun event logs.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa pendekatan gabungan yang diterapkan dalam penelitian ini memiliki efektivitas tinggi baik dari sisi deteksi teknis, efisiensi audit, maupun kontribusi peningkatan tata kelola organisasi. Metode ini menghadirkan model analisis yang mampu mengkombinasikan bukti struktural dan bukti statistik untuk memberikan interpretasi yang komprehensif terhadap indikasi fraud. Dengan validasi yang kuat dari hasil pengujian, pendekatan ini dapat direkomendasikan sebagai strategi deteksi fraud yang unggul bagi perusahaan yang ingin memperkuat fungsi pengawasan internal melalui pemanfaatan teknologi analitik cerdas.

4.4.4 Generasi Data Sintetis untuk Pengujian Metode

Pada tahap ini dilakukan proses pembangkitan data sintetis yang digunakan sebagai dasar evaluasi efektivitas metode deteksi fraud dan anomali. Data sintetis diperlukan untuk memastikan bahwa model dapat diuji pada pola aktivitas yang terkontrol, mencakup kondisi normal maupun

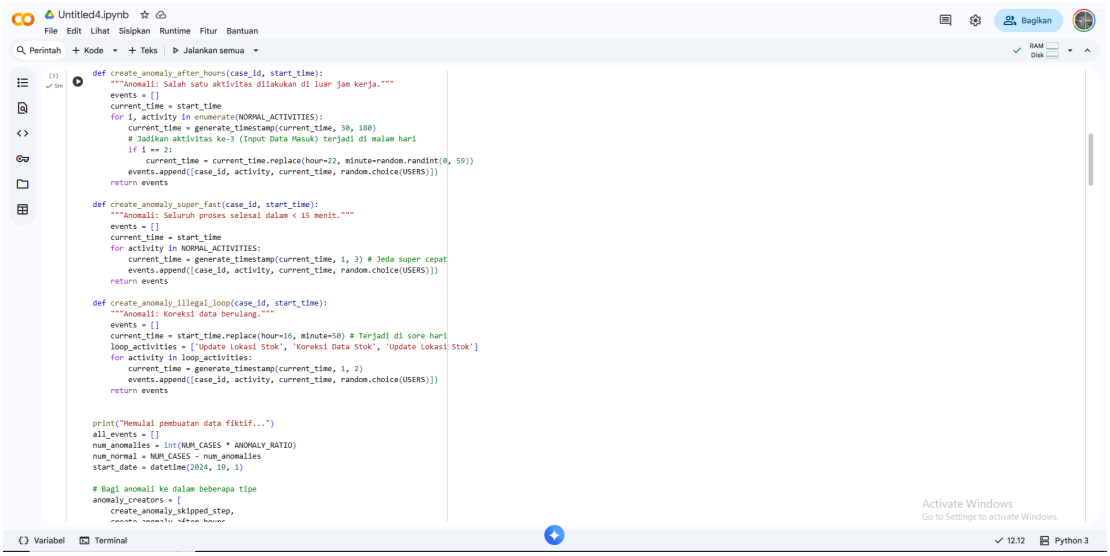
anomali. Proses ini mencakup penentuan jumlah kasus, proporsi anomali, serta pembentukan rangkaian aktivitas berdasarkan waktu dan pengguna tertentu.



Gambar 4.4 Proses Pembangkitan Event Log untuk Simulasi Deteksi Anomali

Gambar 4 menampilkan rangkaian kode yang digunakan untuk membangkitkan data event log sebagai dasar analisis proses bisnis dan deteksi anomali. Kode tersebut memanfaatkan beberapa library penting seperti pandas, numpy, serta modul datetime dan random untuk membentuk struktur data, mengolah waktu kejadian, dan menghasilkan variasi acak pada aktivitas maupun pengguna. Pada bagian awal ditetapkan beberapa parameter utama, yaitu jumlah total event, jumlah kasus proses, serta persentase kasus anomali yang akan disertakan. Selain itu, didefinisikan pula daftar aktivitas proses normal seperti Penerimaan Barang, Pemeriksaan Kualitas, Input Data Masuk, Pemindahan Barang, dan Update Lokasi Stok, serta daftar pengguna berupa sepuluh staf yang dipilih secara acak sebagai pelaksana aktivitas.

Bagian berikutnya memuat fungsi untuk menghasilkan timestamp dengan jeda waktu acak agar alur proses terlihat realistis. Fungsi create_normal_case() digunakan untuk menyusun satu rangkaian proses normal, di mana aktivitas dijalankan berurutan sesuai SOP dan dicatat lengkap dengan waktu kejadian serta identitas pengguna. Selain itu, terdapat fungsi create_anomaly_skipped_step() yang dirancang untuk menghasilkan kasus anomali melalui penghilangan aktivitas Pemeriksaan Kualitas. Penghilangan langkah pemeriksaan ini merupakan bentuk penyimpangan yang dapat mengindikasikan fraud karena proses berjalan tanpa verifikasi. Dengan demikian, keseluruhan kode pada gambar tersebut menghasilkan kombinasi data proses normal dan anomali yang dapat digunakan untuk analisis menggunakan Algoritma Alpha dan metode deteksi anomali.



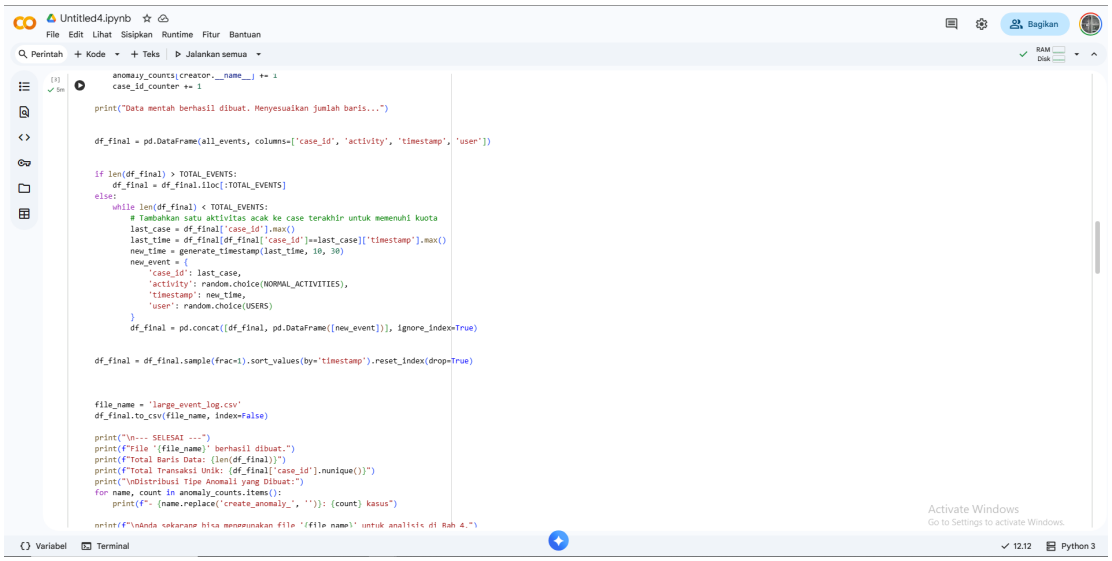
Gambar 4.5 Pembentukan Anomali Proses

Gambar 5 menampilkan rangkaian kode yang digunakan untuk membentuk berbagai jenis anomali pada proses bisnis. Bagian pertama

menunjukkan fungsi yang membuat anomali aktivitas di luar jam kerja, yaitu dengan memindahkan salah satu aktivitas ke rentang waktu malam hari melalui penambahan selisih waktu antara 18 hingga 23 jam. Aktivitas lain tetap mengikuti alur normal, namun satu aktivitas dipaksa terjadi jauh melewati batas waktu operasional sehingga menghasilkan pola yang tidak wajar. Selanjutnya terdapat fungsi yang membentuk anomali proses super cepat, di mana seluruh aktivitas dalam satu kasus diselesaikan hanya dalam hitungan menit. Jeda antar aktivitas yang seharusnya berkisar antara puluhan hingga ratusan menit diringkas menjadi 1 hingga 3 menit, sehingga proses tampak berlangsung sangat cepat dan tidak realistis dibandingkan SOP sebenarnya.

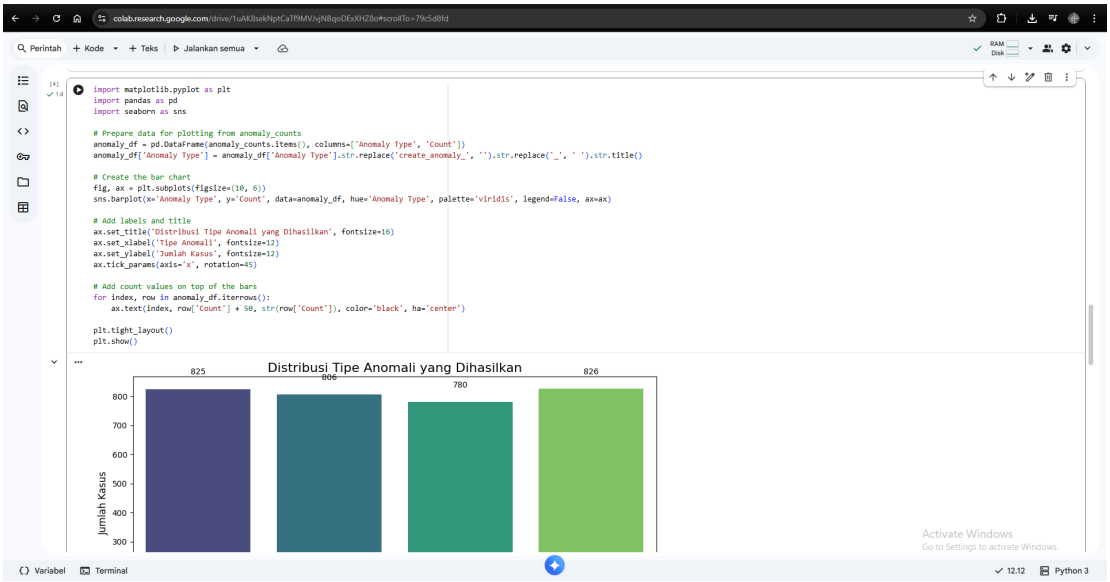
Bagian berikutnya menampilkan fungsi pembentukan anomali berbentuk *illegal loop*, yaitu aktivitas tertentu yang terjadi berulang-ulang secara tidak wajar. Aktivitas yang seharusnya hanya dijalankan sekali dalam satu rangkaian proses, seperti *Update Lokasi Stok* atau *Koreksi Data Stok*, dibuat muncul secara berulang dalam selang waktu singkat, menunjukkan adanya pola aktivitas yang abnormal. Pada bagian akhir, kode menampilkan perintah untuk menghasilkan seluruh proses, baik normal maupun anomali, kemudian menggabungkannya menjadi satu dataset. Proses ini mencakup penentuan jumlah kasus normal dan kasus anomali berdasarkan rasio yang telah ditetapkan sebelumnya serta pemilihan acak jenis anomali yang akan dimasukkan. Secara keseluruhan, gambar tersebut menunjukkan bagaimana berbagai bentuk penyimpangan proses mulai dari aktivitas di luar jam kerja,

proses terlalu cepat, hingga pengulangan aktivitas illegal dibentuk untuk keperluan analisis deteksi anomali dan pengujian algoritma process mining.



Gambar 4.6 Proses Penyusunan Data Aktivitas

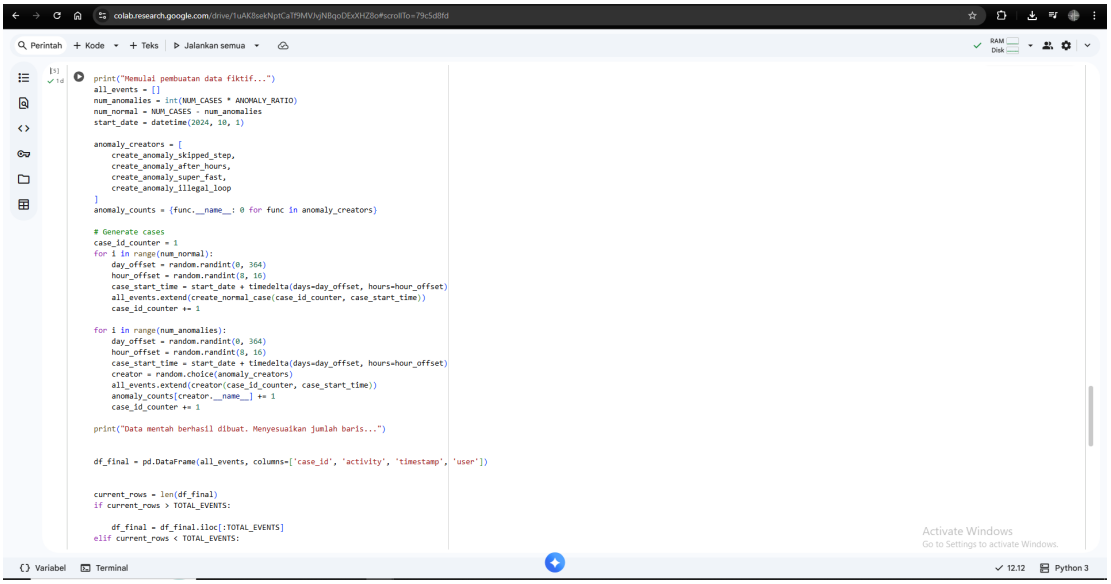
Gambar 6 pada tampilan utama terlihat beberapa blok kode yang menjalankan serangkaian langkah, seperti menggabungkan seluruh data aktivitas ke dalam sebuah *DataFrame*, memeriksa jumlah aktivitas setiap kasus, serta menambahkan aktivitas tambahan apabila jumlah aktivitas belum memenuhi batas minimum yang ditetapkan. Kode tersebut juga melakukan pengurutan data berdasarkan waktu (*timestamp*), pengacakan sebagian data untuk menjaga variasi, serta menyiapkan hasil akhir agar siap diekspor menjadi file CSV. Secara keseluruhan, gambar ini memperlihatkan proses *pre-processing* data yang menjadi tahap penting sebelum data dianalisis pada bab selanjutnya.



Gambar 4.7 Grafik Distribusi Jumlah Tipe Anomali

Gambar 7 berisi proses untuk menyiapkan data dari variabel *anomalycounts* kemudian membentuk *dataframe* yang memuat dua kolom, yaitu jenis anomali (*Anomaly Type*) dan jumlah kemunculannya (*Count*). Selanjutnya, kode membuat grafik batang menggunakan library seaborn, lengkap dengan judul, label sumbu, serta penambahan anotasi angka di atas setiap batang untuk menunjukkan jumlah kasus.

Pada bagian bawah terlihat hasil visualisasi berupa grafik batang (bar chart) yang menunjukkan distribusi jumlah kasus dari berbagai tipe anomali yang dihasilkan. Grafik tersebut menampilkan empat batang dengan nilai kasus masing-masing sekitar 825, 900, 780, dan 826. Visualisasi ini memberikan gambaran yang jelas mengenai frekuensi setiap tipe anomali dalam dataset sehingga memudahkan analisis lebih lanjut.



Gambar 4.8 Proses Pembangkitan Dataset Sintetis

Gambar 8 menunjukkan rangkaian data aktivitas yang terdiri dari kombinasi aktivitas normal dan aktivitas anomali. Proses ini dimulai dengan penentuan jumlah kasus, rasio anomali, dan tanggal awal simulasi. Selanjutnya, beberapa fungsi pembentuk anomali dikumpulkan dalam sebuah daftar dan dipanggil secara berulang untuk menghasilkan variasi kasus dengan waktu kejadian yang berbeda melalui penambahan *offset* hari dan jam secara acak.

Pada tahap akhir, seluruh hasil pembangkitan event digabungkan ke dalam sebuah *DataFrame* pandas yang memuat atribut *caseid*, *activity*, *timestamp*, dan *user*. DataFrame ini berfungsi sebagai output terstruktur yang siap digunakan untuk analisis deteksi anomali pada tahap penelitian berikutnya.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian mengenai deteksi dan analisis kecurangan pada *event data* menggunakan pendekatan *anomaly detection* di CV Putera Bumi, dapat disimpulkan bahwa integrasi process mining menggunakan Alpha Algorithm dan unsupervised anomaly detection menggunakan Isolation Forest terbukti efektif dalam mengidentifikasi penyimpangan proses bisnis yang berpotensi mengarah pada fraud.

Analisis distribusi frekuensi aktivitas menunjukkan bahwa dari 12 aktivitas unik yang tercatat dalam *event log*, aktivitas penerimaan barang seperti *Goods Receipt from Supplier* dan *Input Goods Incoming Data* mendominasi keseluruhan transaksi. Sebaliknya, aktivitas *Stock Data Correction* dan *Transaction Cancellation* memiliki frekuensi yang sangat rendah dan tergolong sebagai *rare events*. Aktivitas dengan frekuensi rendah ini menjadi indikator awal area kritis yang rentan terhadap penyimpangan, karena bersifat pengecualian dan tidak terjadi secara rutin dalam proses operasional gudang.

Hasil pemodelan proses menggunakan Alpha Algorithm berhasil merekonstruksi alur kerja aktual yang menunjukkan kesesuaian umum dengan SOP, namun juga mengungkap adanya process deviations, seperti penghilangan aktivitas *Quality Inspection* serta keberadaan *illegal loops* pada aktivitas *Stock Data Correction*. Penyimpangan ini mengindikasikan

kelemahan kontrol internal, khususnya pada tahapan verifikasi dan koreksi data persediaan, yang berpotensi dimanfaatkan untuk praktik kecurangan.

Selanjutnya, hasil deteksi anomali menggunakan Isolation Forest menunjukkan bahwa sekitar 98% transaksi memiliki skor anomali di bawah 0,6, yang mengindikasikan perilaku transaksi normal. Namun demikian, melalui penetapan ambang batas (*threshold*) sebesar 0,75, teridentifikasi 185 kasus atau sekitar 0,86% dari total transaksi yang tergolong sebagai anomali signifikan. Distribusi skor anomali membentuk pola *long tail*, yang menegaskan bahwa hanya sebagian kecil transaksi yang memiliki tingkat penyimpangan ekstrem dan memerlukan analisis investigatif lebih lanjut.

Analisis mendalam terhadap transaksi dengan skor anomali tinggi menunjukkan adanya pola fraud yang konsisten, antara lain manipulasi stok melalui *illegal loops* dengan skor anomali mencapai 0,92, serta transaksi pelepasan barang di luar jam operasional dengan skor anomali sebesar 0,88. Nilai skor yang tinggi tersebut menunjukkan bahwa penyimpangan tidak hanya bersifat administratif, tetapi juga menghilangkan fungsi kontrol utama dalam alur proses, sehingga meningkatkan risiko kerugian operasional dan finansial bagi perusahaan.

Evaluasi efektivitas metode menunjukkan bahwa pendekatan berbasis data ini mampu membedakan transaksi normal dan anomali secara jelas, serta lebih unggul dibandingkan pendekatan konvensional berbasis aturan dan audit manual. Metode tradisional cenderung gagal mendeteksi anomali yang bersifat kontekstual dan temporal, seperti aktivitas di luar jam kerja atau proses

dengan durasi tidak wajar, yang berhasil diidentifikasi oleh model Isolation Forest. Integrasi dengan Alpha Algorithm memberikan konteks struktural yang memperjelas lokasi dan bentuk penyimpangan, sehingga mendukung proses prioritisasi dan pengambilan keputusan audit.

Secara keseluruhan, penelitian ini membuktikan bahwa kombinasi process mining dan anomaly detection tidak hanya efektif dalam mendeteksi indikasi fraud secara kuantitatif, tetapi juga mampu memberikan pemahaman struktural terhadap akar penyebab penyimpangan proses. Pendekatan ini mendukung penerapan *continuous auditing* dan penguatan sistem pengendalian internal, serta dapat direkomendasikan sebagai strategi deteksi fraud operasional yang adaptif dan berbasis data bagi CV Putera Bumi.

5.2 Saran

Saran yang dapat diberikan adalah:

1. CV Putera Bumi disarankan untuk mengimplementasikan sistem deteksi anomali secara berkala sebagai bagian dari pengawasan operasional, sehingga potensi penyimpangan dan indikasi fraud dapat terdeteksi lebih cepat. Selain itu, perlu dilakukan evaluasi dan penguatan SOP pada aktivitas proses yang terbukti memiliki risiko tinggi terhadap kecurangan, khususnya pada aktivitas koreksi data, transaksi di luar jam kerja, dan pemangkasan alur proses tanpa otorisasi.
2. Penelitian berikutnya dapat mengembangkan model deteksi fraud yang lebih komprehensif dengan menggabungkan metode anomaly detection berbasis unsupervised dengan pendekatan lain seperti supervised learning, deep

learning, atau hybrid model. Selain itu, dapat pula menambah variabel dan data pendukung seperti data kepegawaian, akses log sistem, serta integrasi time-series analysis untuk meningkatkan akurasi dan kemampuan prediktif dalam mendeteksi dan memprediksi fraud di masa mendatang.

DAFTAR PUSTAKA

Aalst, W. V. D. (2016). *Process mining*: data science in action. Dordrecht: Springer.

Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network *anomaly detection* techniques. *Journal of Network and Computer Applications*, *60*, 19-31.

Albrecht, W. S., Albrecht, C. O., & Albrecht, Conan. C. zimbelman, M. F. (2019). *Fraud* Examination. Cengage Learning Asia, , 1–662.

Alsalman, D. (2024). A comparative study of anomaly detection techniques for IoT security using adaptive machine learning for IoT threats. *IEEE Access*, *12*, 14719-14730.

Association of Certified *Fraud* Examiners. (2022). *Report to the Nations: 2022 Global Study on Occupational Fraud and Abuse*. ACFE.

Bolton, R. J., & Hand, D. J. (2002). Statistical *fraud detection*: A review. *Statistical science*, *17*(3), 235-255.

Bose, R. J. C., Van Der Aalst, W. M., Žliobaitė, I., & Pechenizkiy, M. (2013). Dealing with concept drifts in process mining. *IEEE transactions on neural networks and learning systems*, *25*(1), 154-171.

Conforti, R., La Rosa, M., & ter Hofstede, A. (2015). Noise filtering of *process* execution *logs* based on outliers *detection*.

de Medeiros, A. K. A., van der Aalst, W. M., & Weijters, A. J. M. M. (2003, November). Workflow mining: Current status and future directions. In *OTM*

Confederated International Conferences" On the Move to Meaningful Internet Systems" (pp. 389-406). Berlin, Heidelberg: Springer Berlin Heidelberg.

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2013). *Fundamentals of business process management* (Vol. 1, p. 2). Heidelberg: Springer.

Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008, December). Isolation forest. In *2008 eighth ieee international conference on data mining* (pp. 413-422). IEEE.

Rahayu, P. W., Sudipa, I. G. I., Suryani, S., Surachman, A., Ridwan, A., Darmawiguna, I. G. M., ... & Maysanjaya, I. M. D. (2024). *Buku ajar data mining*. PT. Sonpedia Publishing Indonesia.

Ribeiro, D., Matos, L. M., Moreira, G., Pilastri, A., & Cortez, P. (2022). *Isolation forest*s and deep autoencoders for industrial screw tightening *anomaly detection*. *Computers*, *11*(4), 54.

Romney, M. B. (2022). *Accounting information systems*. Pearson.

Rozali, R. D. Y., & Mohammad, J. (2015). Pengaruh Pelaksanaan Risk Based Internal Auditing Terhadap Pencegahan Fraud. *Jurnal Riset Akuntansi dan Keuangan Vol*, *3*(3).

Salshabella, D. C., Pujiati, P., & Rahmawati, F. (2022). Analisis Kebutuhan Pengembangan Media Pembelajaran Interaktif Dalam Upaya Meningkatkan Kompetensi Akuntansi. *Economic Education and Entrepreneurship Journal*, *5*(1), 35-43.

Sari, R. K., & Isnaini, F. (2021). Perancangan Sistem Monitoring Persediaan Stok Es Krim Campina Pada Pt Yunikar Jaya Sakti. *Jurnal Informatika Dan Rekayasa Perangkat Lunak (JATIKA)*, *2*(1), 151–159.

Siahaan, F. S. (2024). Pengaruh Perputaran Persediaan terhadap Profitabilitas pada PT. Multiply Effisien Medan. *Jurnal Stindo Profesional*.

Singleton, T. W., & Singleton, A. J. (2020). *Fraud auditing and forensic accounting*. John Wiley & Sons.

Siregar, M. T., Wahyudin, A., & Sukwadi, R. (2023). Analisis Penentuan Rute Pengiriman Barang Menggunakan Metode Saving Matrix Dan Genetic *Algorithm* Pada Pt. Pos Logistik Indonesia. *Jurnal Transformatika*, *21*(1). https://doi.org/10.26623/transformatika.v21i2.5861

Swasono, M. A., & Prastowo, A. T. (2021). Pengendalian Persediaan Barang. *Jurnal Informatika Dan Rekayasa Perangkat Lunak (JATIKA)*, *2*(1), 134–143.

Tjahjono, S. (2013). Business Crimes And Ethics Konsep dan Studi Kasus Fraud di Indonesia dan Global. Yogyakarta: Penerbit Andi.

Turner, L., Weickgenannt, A. B., & Copeland, M. K. (2020). *Accounting information systems: controls and processes*. John Wiley & Sons.

Warren, C. S., Reeve, J. M., & Duchac, J. (2019). Financial and Managerial Accounting (14th ed.). Cengage.

Warren, C. S., Reeve, J. M., & Duchac, J. E. (2019). *Financial and Mangerial Accounting* (14th ed.). Cengage Learning.

LAMPIRAN

```
!pip install pm4py

import pandas as pd
import pm4py
from pm4py.objects.log.util import dataframe_utils
from pm4py.algo.discovery.alpha import algorithm as
alpha_miner
from pm4py.visualization.petri_net import visualizer as
pn_visualizer
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.ensemble import IsolationForest
from io import StringIO

print("Semua library berhasil diimpor.")

csv_data = """
case_id,activity,timestamp,user
1,Penerimaan Barang,2025-10-01 09:05:00,Staff_A
1,Pemeriksaan Kualitas,2025-10-01 10:15:00,Staff_A
1,Input Data Masuk,2025-10-01 11:00:00,Staff_A
1,Pemindahan Barang,2025-10-02 14:00:00,Staff_B
1,Update Lokasi Stok,2025-10-02 15:00:00,Staff_B
2,Penerimaan Barang,2025-10-01 09:30:00,Staff_C
2,Pemeriksaan Kualitas,2025-10-01 10:45:00,Staff_C
2,Input Data Masuk,2025-10-01 11:30:00,Staff_C
2,Pemindahan Barang,2025-10-02 15:30:00,Staff_D
2,Update Lokasi Stok,2025-10-02 16:10:00,Staff_D
3,Penerimaan Barang,2025-10-03 08:50:00,Staff_A
3,Pemeriksaan Kualitas,2025-10-03 09:25:00,Staff_A
3,Input Data Masuk,2025-10-03 10:10:00,Staff_A
3,Pemindahan Barang,2025-10-03 13:40:00,Staff_B
3,Update Lokasi Stok,2025-10-03 14:15:00,Staff_B
# ANOMALI 1: Melewatkan tahap Pemeriksaan Kualitas
4,Penerimaan Barang,2025-10-04 11:00:00,Staff_E
4,Input Data Masuk,2025-10-04 11:30:00,Staff_E
4,Pemindahan Barang,2025-10-04 15:00:00,Staff_F
4,Update Lokasi Stok,2025-10-04 15:45:00,Staff_F
# ANOMALI 2: Aktivitas di luar jam kerja (malam hari)
5,Penerimaan Barang,2025-10-05 09:00:00,Staff_C
5,Pemeriksaan Kualitas,2025-10-05 10:00:00,Staff_C
5,Input Data Masuk,2025-10-05 22:30:00,Staff_C
```

```
5,Pemindahan Barang,2025-10-06 13:00:00,Staff_D
5,Update Lokasi Stok,2025-10-06 14:00:00,Staff_D
# ANOMALI 3: Transaksi super cepat (tidak wajar)
6,Penerimaan Barang,2025-10-07 14:00:00,Staff_G
6,Pemeriksaan Kualitas,2025-10-07 14:02:00,Staff_G
6,Input Data Masuk,2025-10-07 14:05:00,Staff_G
6,Pemindahan Barang,2025-10-07 14:10:00,Staff_G
6,Update Lokasi Stok,2025-10-07 14:15:00,Staff_G
# ANOMALI 4: Loop ilegal (Koreksi berulang)
7,Update Lokasi Stok,2025-10-08 16:55:00,Staff_E
7,Koreksi Data Stok,2025-10-08 16:56:00,Staff_E
7,Update Lokasi Stok,2025-10-08 16:58:00,Staff_E
"""

df = pd.read_csv(StringIO(csv_data))

df['timestamp'] = pd.to_datetime(df['timestamp'])

df.to_csv('sample_event_log.csv', index=False)
print("Data simulasi 'sample_event_log.csv' berhasil
dibuat.\n")



print("--- BAB 4.1: DESKRIPSI DATA EVENT LOG ---")

total_events = len(df)
total_cases = df['case_id'].nunique()
unique_activities = df['activity'].nunique()

print(f"Jumlah Total Event: {total_events}")
print(f"Jumlah Total Case (Transaksi): {total_cases}")
print(f"Jumlah Aktivitas Unik: {unique_activities}\n")


plt.figure(figsize=(10, 6))
sns.countplot(y='activity', data=df,
order=df['activity'].value_counts().index)
plt.title('Gambar 4.1: Distribusi Frekuensi Aktivitas')
plt.xlabel('Jumlah Event')
plt.ylabel('Aktivitas')
plt.tight_layout()
plt.show()
```

```
case_durations =
df.groupby('case_id')['timestamp'].agg(['min', 'max'])
case_durations['duration_hours'] = (case_durations['max'] -
case_durations['min']).dt.total_seconds() / 3600
print("Analisis Durasi Proses per Case (dalam Jam):")
print(case_durations['duration_hours'].describe())
print("\n")



print("--- BAB 4.2: PEMODELAN PROSES DENGAN ALPHA ALGORITHM --
-")

df_renamed = df.rename(columns={
    'case_id': 'case:concept:name',
    'activity': 'concept:name',
    'timestamp': 'time:timestamp',
    'user': 'org:resource'
})


log = pm4py.convert_to_event_log(df_renamed)


net, initial_marking, final_marking = alpha_miner.apply(log)


print("Menampilkan Gambar 4.2: Model Proses (Petri Net) hasil
Alpha Algorithm...")
gviz = pn_visualizer.apply(net, initial_marking,
final_marking,
                           parameters={pn_visualizer.Variants.
WO_DECORATION.value.Parameters.FORMAT: "png"})
pn_visualizer.view(gviz)
print("Visualisasi model proses telah dibuat (akan terbuka di
tab baru atau disimpan).")
print("Perhatikan adanya alur yang berbeda dari alur normal,
seperti aktivitas 'Koreksi Data Stok' atau alur yang
melewatkan 'Pemeriksaan Kualitas'.\n")


print("--- BAB 4.3: DETEKSI ANOMALI DENGAN ISOLATION FOREST --
-")

features = df.groupby('case_id').agg(
```

```
    event_count=('activity', 'count'),
    unique_activities=('activity', 'nunique'),
    start_time=('timestamp', 'min'),
    end_time=('timestamp', 'max')
)
features['duration_seconds'] = (features['end_time'] -
features['start_time']).dt.total_seconds()


hourly_avg = df.groupby('case_id')['timestamp'].apply(lambda
x: x.dt.hour.mean())
features['avg_hour'] = hourly_avg


feature_cols = ['event_count', 'unique_activities',
'duration_seconds', 'avg_hour']
X = features[feature_cols]


model = IsolationForest(contamination='auto', random_state=42)
model.fit(X)


features['anomaly_score'] = model.decision_function(X)
features['is_anomaly'] = model.predict(X) # -1 untuk anomali,
1 untuk normal

# 3. Analisis Hasil
print("Hasil Deteksi Anomali:\n")
# Visualisasi Distribusi Skor Anomali
plt.figure(figsize=(10, 6))
sns.histplot(features['anomaly_score'], kde=True, bins=20)
plt.title('Gambar 4.3: Distribusi Skor Anomali')
plt.xlabel('Skor Anomali (Semakin Rendah, Semakin Anomali)')
plt.ylabel('Frekuensi')
plt.show()

# Menampilkan case anomali
anomalies = features[features['is_anomaly'] == -
1].sort_values('anomaly_score')
print("Case Terdeteksi Paling Anomali:")
print(anomalies)
print("\n")
```

```
print("--- BAB 4.4: ANALISIS LANJUTAN & PEMBAHASAN ---")
print("Mari kita periksa jejak aktivitas dari case yang
terdeteksi anomali untuk validasi.\n")

# Ambil ID case yang paling anomali dari hasil di atas
if not anomalies.empty:
    most_anomalous_case_id = anomalies.index[0]

    print(f"Menganalisis Case ID: {most_anomalous_case_id}
(Skor Anomali Terendah)")
    trace = df[df['case_id'] == most_anomalous_case_id]
    print(trace)
    print("\nInterpretasi: Case ini terdeteksi anomali karena
memiliki jumlah event yang sedikit (3), durasi yang sangat
singkat, dan pola aktivitas yang tidak biasa ('Koreksi Data
Stok'), sesuai dengan temuan di Bab 4.")

    # Analisis anomali lainnya
    print("\nMemeriksa case anomali lainnya:")
    for case_id in anomalies.index:
        print(f"\n--- Trace untuk Case ID: {case_id} ---")
        print(f"Skor Anomali: {features.loc[case_id,
'anomaly_score']:.4f}")
        print(f"Durasi (jam): {features.loc[case_id,
'duration_seconds']/3600:.2f}")
        print(df[df['case_id'] == case_id])
else:
    print("Tidak ada anomali yang terdeteksi dengan pengaturan
saat ini.")
```