

Deteksi Anomali menggunakan Isolation Forest pada Permintaan Kebutuhan Farmasi Pasien di Rumah Sakit Mitra Sejati Medan

Farhan Novaldi¹, Syafrijon², Yeka Hendriyani³, Muhammad Anwar⁴

¹ Informatika, Universitas Negeri Padang

^{2,3,4} Elektronika, Universitas Negeri Padang

e-mail: farhannovaldi002@gmail.com

Abstrak

Rumah Sakit Mitra Sejati Medan menghadapi tantangan dalam mengelola volume permintaan farmasi yang tinggi, menyebabkan proses verifikasi manual menjadi tidak efisien dan berisiko. Penelitian ini bertujuan merancang dan mengimplementasikan sistem deteksi anomali untuk meningkatkan efektivitas pengelolaan permintaan. Metode yang digunakan adalah algoritma Isolation Forest dengan menerapkan metodologi *Cross-Industry Standard Process for Data Mining*. Data historis permintaan obat, barang medis habis pakai, dan alat kesehatan diolah menggunakan Python untuk melatih model secara kontekstual. Hasil penelitian menunjukkan dari 2.167.942 transaksi, model berhasil mengidentifikasi 13.503 (0,62%) permintaan sebagai anomali statistik. Sistem yang dikembangkan melalui aplikasi web ini terbukti berhasil menjadi alat bantu keputusan berbasis data untuk meningkatkan efisiensi operasional, akurasi stok, dan memberikan peringatan dini.

Kata kunci: *Deteksi Anomali, Isolation Forest, Data Mining*

Abstract

Mitra Sejati Hospital in Medan faced a significant challenge in managing its high volume of pharmacy requests, which rendered the manual verification process inefficient and risky. This study aimed to design and implement an anomaly detection system to improve the effectiveness of request management. The research utilized the Isolation Forest algorithm and applied the *Cross-Industry Standard Process for Data Mining* methodology. Historical data on drug, medical consumable, and medical device requests were processed using Python to train the model contextually. The results show that from 2,167,942 transactions, the model successfully identifies 13,503 (0.62%) requests as statistical anomalies. The developed web-based system proves to be a successful data-driven decision support tool that improves operational efficiency, enhances stock accuracy, and provides early warnings.

Keywords : *Anomaly Detection, Isolation Forest, Data Mining*

PENDAHULUAN

Transformasi digital di sektor kesehatan mendorong rumah sakit memanfaatkan teknologi informasi untuk meningkatkan mutu pelayanan agar lebih cepat, akurat, dan efisien. Salah satu penerapannya adalah Sistem Informasi Manajemen Rumah Sakit (SIMRS) yang dirancang untuk mengintegrasikan berbagai proses, mulai dari pendaftaran pasien, pengelolaan data medis, hingga distribusi obat dan alat kesehatan. Penerapan SIMRS terbukti mampu mendukung pelayanan kesehatan yang lebih baik dengan pengelolaan data yang selalu akurat dan up to date (Maryati & Utami, 2023).

Modul farmasi menjadi salah satu bagian penting dalam SIMRS karena berfungsi mencatat dan memproses permintaan obat, barang medis habis pakai (BMHP), dan alat kesehatan agar distribusi berjalan tepat waktu dan efisien. Namun, tingginya volume permintaan farmasi yang harus dikelola setiap bulan, seperti yang terjadi di Rumah Sakit Mitra Sejati Medan dengan rata-rata lebih dari 36.000 transaksi per bulan, membuat proses verifikasi manual menjadi tidak efektif.

Kondisi ini berpotensi memunculkan kesalahan pencatatan, pemborosan anggaran, serta risiko pada keselamatan pasien apabila permintaan yang tidak wajar tidak segera terdeteksi.

Kesulitan mendeteksi pola permintaan yang menyimpang secara cepat juga menambah beban kerja staf farmasi. Metode verifikasi konvensional yang sepenuhnya bergantung pada tenaga manusia sering kali tidak mampu mengimbangi volume data yang terus meningkat. Hal ini sejalan dengan temuan (Zulfikar et al., 2023) yang menyebutkan bahwa metode manual rawan terhadap human error dan memerlukan waktu yang panjang, sehingga berdampak pada efisiensi pengelolaan stok dan pelayanan farmasi di rumah sakit.

Untuk mengatasi kendala tersebut, pendekatan cerdas dengan algoritma deteksi anomali diperlukan agar permintaan yang tidak wajar dapat diidentifikasi secara otomatis dan akurat. Isolation Forest merupakan salah satu algoritma yang banyak digunakan untuk mendeteksi outlier pada dataset besar karena prinsip kerjanya yang membangun pohon isolasi dan memisahkan data yang menyimpang dari pola normal (Dewa & Windarto, 2024). Penelitian sebelumnya menunjukkan bahwa Isolation Forest efektif mendeteksi anomali dalam berbagai domain, termasuk keuangan, jaringan, hingga farmasi rumah sakit (Wibawa & Karyawati, 2023).

Penelitian ini dirancang untuk membangun sistem deteksi anomali permintaan kebutuhan farmasi dengan menerapkan algoritma Isolation Forest. Model deteksi dikembangkan menggunakan Python dan pustaka pendukung seperti scikit-learn untuk pemrosesan data, sedangkan hasil deteksi akan disimpan ke dalam basis data MySQL dan divisualisasikan melalui antarmuka web berbasis PHP agar memudahkan staf farmasi melakukan verifikasi. Dengan pendekatan ini, rumah sakit diharapkan dapat mengurangi risiko pemborosan, menjaga ketersediaan stok, serta meningkatkan kualitas pelayanan pasien.

Implementasi sistem deteksi anomali ini diharapkan tidak hanya mendukung pengambilan keputusan berbasis data, tetapi juga menjadi contoh penerapan teknologi machine learning dalam pengelolaan farmasi di lingkungan rumah sakit. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi nyata dalam pengembangan SIMRS yang lebih cerdas dan adaptif terhadap tantangan volume data yang besar di era digital saat ini.

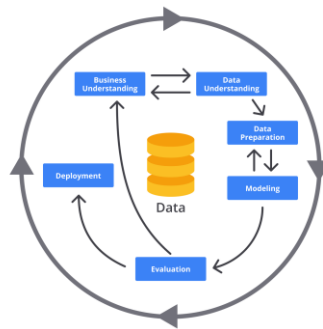
METODE

Jenis Penelitian, sumber data, dan variable penelitian

Penelitian ini merupakan penelitian terapan dengan pendekatan kuantitatif dan deskriptif. Pendekatan kuantitatif digunakan untuk menganalisis data permintaan farmasi dalam jumlah besar dengan menerapkan algoritma *Isolation Forest* sebagai metode deteksi anomali. Sementara itu, pendekatan deskriptif digunakan untuk menjelaskan hasil temuan pola permintaan yang menyimpang dan bagaimana sistem deteksi yang dibangun dapat mendukung pengambilan keputusan di Rumah Sakit Mitra Sehati Medan.

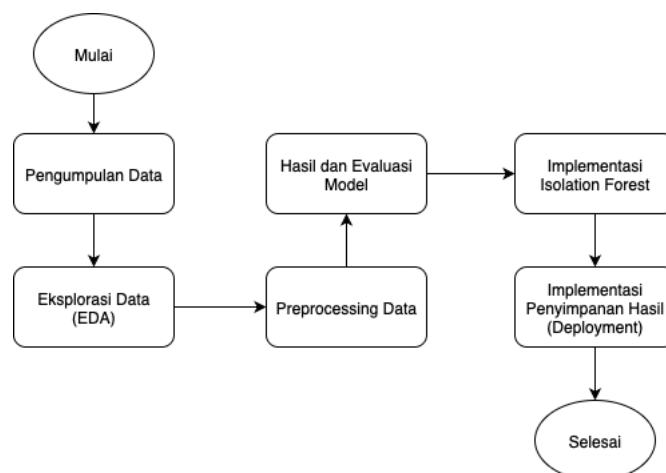
Data penelitian ini bersumber dari data sekunder yang diambil dari basis data SIMRS Rumah Sakit Mitra Sehati Medan, mencakup data permintaan obat, BMHP, dan alat kesehatan pada periode Maret 2024 hingga Juni 2025. Variabel yang digunakan terdiri dari variabel independen berupa tipe permintaan, nomor permintaan, nomor rekam medis, waktu, status permintaan, dan detail resep seperti nama item, dosis, jumlah, dan harga. Variabel dependen adalah klasifikasi hasil deteksi anomali, yaitu penentuan apakah suatu permintaan tergolong normal atau anomali berdasarkan hasil analisis *Isolation Forest* yang kemudian diverifikasi oleh pihak farmasi.

Tahapan Penelitian



Gambar 1. Metode CRISP-DM

Penelitian ini mengadopsi metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM), yang merupakan kerangka kerja standar dalam proyek data mining. Metodologi ini terdiri dari enam fase utama yang diimplementasikan secara berurutan untuk memastikan proses penelitian berjalan sistematis, mulai dari pengumpulan data hingga implementasi hasil. Alur tahapan penelitian yang dilakukan diadaptasi dari kerangka CRISP-DM dan disajikan dalam urutan seperti Gambar 2.



Gambar 2. Alur Tahapan Penelitian

1. Pengumpulan Data

Tahap awal ini berfokus pada pengumpulan data transaksional permintaan farmasi yang mencakup obat, BMHP, dan alkes. Data bersumber langsung dari basis data Sistem Informasi Manajemen Rumah Sakit (SIMRS) RS Mitra Sehati Medan. Proses ini dilakukan secara terprogram menggunakan Python dengan bantuan pustaka seperti Pandas dan SQLAlchemy untuk berinteraksi langsung dengan basis data. Data dari periode 1 Maret 2024 hingga waktu penelitian ditarik menggunakan query SQL yang menggabungkan tiga kategori permintaan menjadi satu DataFrame tunggal untuk dianalisis.

2. Eksplorasi Data (EDA)

Setelah data terkumpul, dilakukan tahap eksplorasi untuk memahami struktur, kualitas, serta distribusi data. Pada tahap ini, dilakukan pemeriksaan terhadap tipe data, ada tidaknya nilai kosong (*missing values*), serta distribusi nilai pada setiap kolom. Dari hasil pemeriksaan, ditemukan 36.266 nilai kosong pada kolom tipe. Visualisasi data juga dilakukan untuk mempermudah pemahaman terhadap frekuensi permintaan berdasarkan jenisnya.

3. *Preprocessing Data*

Tahap ini bertujuan membersihkan dan mempersiapkan dataset agar siap digunakan untuk pemodelan. Tiga kegiatan utama dilakukan:

- a) Penanganan Nilai Kosong
Sebanyak 36.266 nilai kosong pada kolom tipe diisi dengan kategori baru, yaitu farmasi. Kategori ini dipilih untuk mewakili transaksi yang tidak terikat pada jenis layanan spesifik seperti penjualan bebas.
- b) Konversi Tipe Data
Kolom-kolom yang berisi data kategorikal seperti *jenis_permintaan*, *nama_item*, dan *tipe* diubah menjadi tipe data *category* agar lebih rapi dan siap untuk di-encode.
- c) Pembersihan Data Tidak Valid
Sebanyak 28.707 baris data dengan kuantitas (*qty*) bernilai nol atau negatif dihapus karena dianggap sebagai noise atau kesalahan data.

4. Implementasi *Isolation Forest*

Tahap ini merupakan fase inti dan paling krusial dari penelitian, di mana algoritma *Isolation Forest* diimplementasikan secara praktis untuk mendeteksi anomali pada data permintaan farmasi. Implementasi dilakukan secara terprogram menggunakan pustaka *Scikit-learn* di Python, yang menjadi mesin utama untuk analisis. Pendekatan pemodelan yang digunakan bersifat iteratif dan kontekstual. Artinya, model tidak dilatih pada keseluruhan data sekaligus. Sebaliknya, proses pemodelan dilakukan dengan melakukan perulangan (*looping*) melalui setiap nama_item unik, kemudian dilanjutkan dengan perulangan lagi untuk setiap tipe layanan yang terkait dengan item tersebut. Pendekatan ini dipilih untuk memastikan model dapat mempelajari pola kewajaran kuantitas untuk setiap item dalam konteks spesifiknya, sehingga deteksi anomali menjadi lebih akurat dan relevan secara operasional. Di dalam setiap iterasi, langkah-langkah berikut dijalankan:

- a) Seleksi dan Normalisasi Fitur
Fitur tunggal yang menjadi fokus analisis adalah kuantitas. Fitur ini kemudian dinormalisasi menggunakan *StandardScaler* dari *Scikit-learn* untuk mengubah skala data agar memiliki rata-rata 0 dan standar deviasi 1, yang dapat meningkatkan stabilitas dan kinerja model.
- b) Pelatihan Model
Model *Isolation Forest* diinisialisasi dan dilatih pada data kuantitas yang telah dinormalisasi. Parameter *contamination* diatur ke 0.01, yang mengindikasikan ekspektasi bahwa sekitar 1% dari data dalam setiap sub-kelompok dianggap sebagai anomali.
- c) Prediksi dan Pelabelan
Setelah dilatih, metode *fit_predict* dijalankan untuk menghasilkan label prediksi (-1 untuk anomali dan 1 untuk normal). Selain itu, *decision_function* digunakan untuk menghitung skor anomali untuk setiap titik data, di mana skor yang lebih rendah (negatif) menunjukkan tingkat penyimpangan yang lebih tinggi. Hasil prediksi dan skor ini kemudian ditambahkan kembali ke dalam *DataFrame* untuk dianalisis pada tahap evaluasi

5. Hasil dan Evaluasi Model

Kinerja model yang telah dilatih dievaluasi secara kuantitatif dan kualitatif.

- a) Analisis Kuantitatif
Dari total 2.167.942 data transaksi, model berhasil mengidentifikasi 13.503 permintaan (0,62%) sebagai anomali statistik, hasil ini sejalan dengan parameter *contamination* yang ditetapkan.
- b) Analisis Kualitatif
Evaluasi dilakukan dengan menganalisis distribusi skor anomali yang menunjukkan pemisahan jelas antara data normal dan anomali. Selain itu, studi kasus visual pada item spesifik seperti 'Hepagusan Inj' dan 'Bactigras' membuktikan kemampuan model dalam mendeteksi *outlier* ekstrem maupun penyimpangan yang lebih halus.

6. Implementasi Penyimpanan Hasil (*Deployment*)

Tahap akhir adalah mengimplementasikan model ke dalam sebuah sistem prototipe yang fungsional untuk menjembatani model analitis dengan antarmuka praktis. Hasil deteksi

anomali disimpan ke dalam basis data dengan pendekatan idempoten, yang berarti skrip dapat dijalankan berulang kali tanpa menimbulkan duplikasi data.

Sebuah API berbasis *Flask* dibangun untuk menyediakan endpoint yang memungkinkan antarmuka web memicu skrip deteksi anomali. Hasilnya kemudian disajikan dalam antarmuka web untuk divalidasi oleh Staf dan Kepala Farmasi.

HASIL DAN PEMBAHASAN

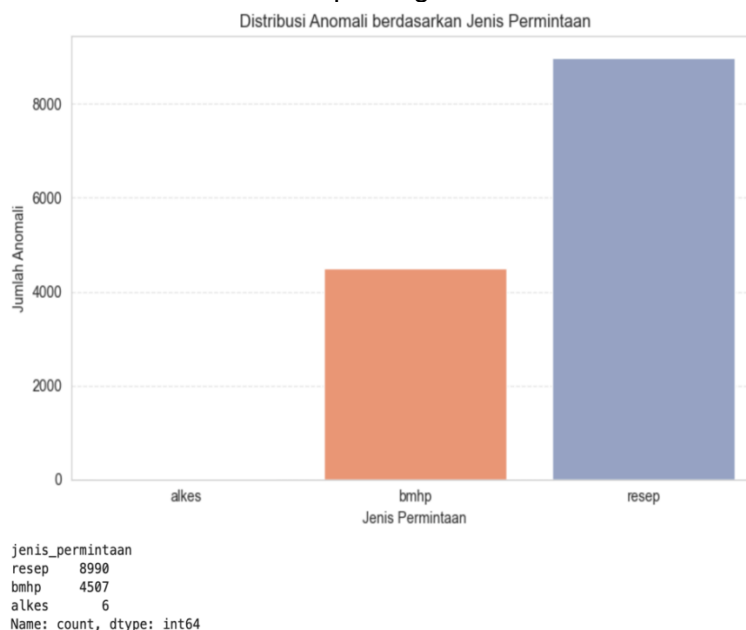
Implementasi model deteksi anomali pada data permintaan farmasi di Rumah Sakit Mitra Sejati Medan dilakukan dengan pendekatan kontekstual menggunakan algoritma *Isolation Forest*. Model dilatih secara terpisah untuk setiap kombinasi unik antara nama item dan tipe layanan (rawat inap, rawat jalan, IGD, dan farmasi) dengan fokus utama pada anomali kuantitas (*qty*) permintaan. Algoritma ini dipilih karena efisiensinya dalam menangani data berskala besar dan kemampuannya mengisolasi outlier secara efektif tanpa memerlukan data berlabel (*unsupervised*), sejalan dengan temuan pada penelitian oleh Zulfikar et al. (2023) dan Dewa & Windarto (2024). Dengan parameter *contamination* yang diatur sebesar 0.01 (1%), model ini dirancang untuk mengasumsikan sekitar 1% dari data dalam setiap sub-kelompok merupakan anomali statistik.

Setelah melalui tahap pra-pemrosesan data, model diterapkan pada total 2.167.942 transaksi yang valid. Dari jumlah tersebut, model berhasil mengidentifikasi 13.503 permintaan sebagai anomali, atau setara dengan 0,62% dari keseluruhan data. Persentase ini menunjukkan bahwa model bekerja sesuai ekspektasi tanpa mengklasifikasikan data secara berlebihan.

Tabel 1. Hasil Kuantitatif Deteksi Anomali

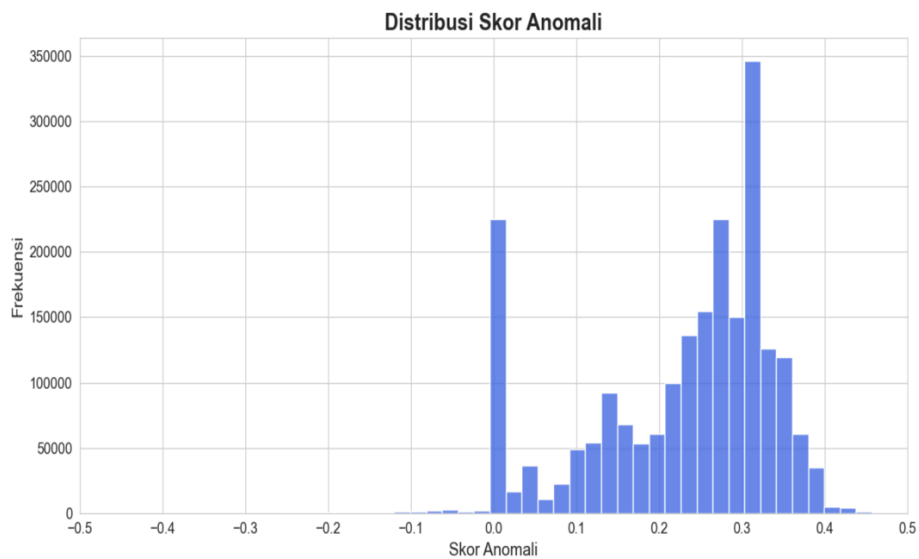
Keterangan	Jumlah
Total Data Permintaan	2.167.942
Total Anomali Terdeteksi	13.503
Persentase Anomali	0,62%

Untuk memberikan konteks lebih dalam terhadap temuan tersebut, anomali yang terdeteksi didistribusikan berdasarkan jenis permintaannya. Hasilnya menunjukkan bahwa anomali terbanyak ditemukan pada kategori resep (8.990) dan BMHP (4.507). Jumlah ini sebanding dengan volume transaksi pada masing-masing kategori, yang mengindikasikan bahwa model bekerja secara proporsional dan tidak memiliki bias terhadap kategori tertentu.



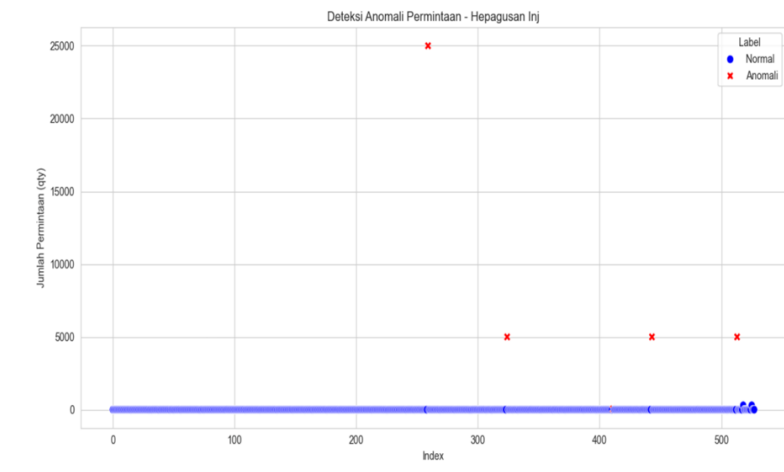
Gambar 3. Distribusi Data Anomali Berdasarkan Jenis Permintaan

Analisis visual terhadap distribusi skor anomali semakin memperkuat evaluasi kemampuan model dalam memisahkan data. Terlihat adanya pemisahan yang jelas antara data normal yang terkonsentrasi pada skor positif dan data anomali (pencilan) yang tersebar pada skor negatif, yang menandakan keberhasilan proses isolasi.



Gambar 4. Distribusi Skor Anomali

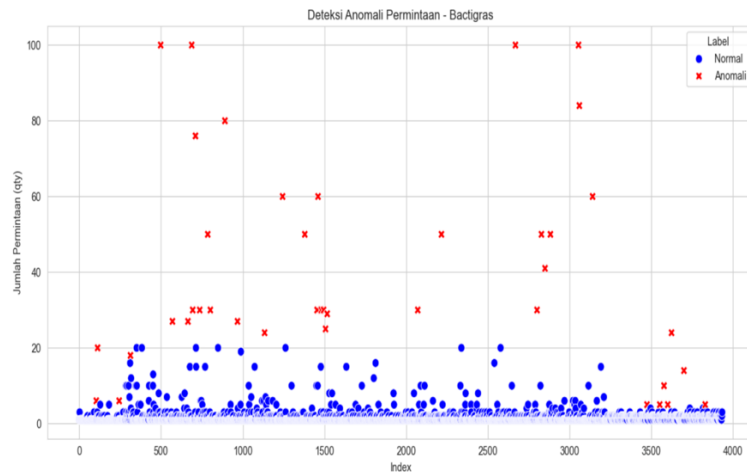
Validasi kualitatif melalui studi kasus pada item spesifik semakin memperkuat temuan ini. Studi kasus pertama pada item Hepagusan Inj menunjukkan keandalan model dalam mendeteksi point outlier yang sangat jelas.



Gambar 5. Deteksi Anomali Hepagusan Inj

Seperti terlihat pada Gambar 5, model berhasil mengidentifikasi permintaan dengan kuantitas ekstrem (5.000 hingga 25.000) sebagai anomali, yang sangat kontras dengan pola permintaan normalnya yang cenderung seragam pada 1 satuan.

Selanjutnya, untuk menguji sensitivitas model, dilakukan analisis pada item Bactigras yang memiliki fluktuasi permintaan lebih dinamis.



Gambar 6. Deteksi Anomali Bactigras

Pada Gambar 5, model tidak hanya menandai permintaan dengan kuantitas sangat tinggi (>60) sebagai anomali, tetapi juga mampu mendeteksi penyimpangan yang lebih halus pada rentang kuantitas menengah (20-40). Temuan ini membuktikan fleksibilitas model dalam menangani berbagai jenis penyimpangan, sekaligus menegaskan pentingnya validasi ahli untuk interpretasi lebih lanjut.

Secara keseluruhan, keberhasilan model ini menunjukkan bahwa pendekatan analitis berbasis machine learning dapat menjadi fondasi yang kuat untuk sistem pendukung keputusan di lingkungan farmasi. Meskipun efektif, perlu digarisbawahi bahwa anomali yang terdeteksi bersifat statistik dan tidak secara otomatis berarti sebuah kesalahan atau penyalahgunaan (fraud). Sebuah permintaan bisa jadi anomali karena kebutuhan klinis yang langka atau faktor situasional lain yang valid. Oleh karena itu, kontribusi utama dari penelitian ini adalah pengembangan sebuah model analitis yang terbukti andal dalam menyaring dan menandai potensi ketidakwajaran dari volume data yang sangat besar. Langkah krusial berikutnya adalah validasi domain oleh Staf dan Kepala Farmasi, yang perannya adalah untuk menginterpretasi konteks di balik setiap temuan statistik dan menerjemahkannya menjadi tindakan operasional yang tepat guna.

SIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma *Isolation Forest* berhasil diimplementasikan untuk mendeteksi anomali pada permintaan kebutuhan farmasi di Rumah Sakit Mitra Sejati Medan, dengan mengidentifikasi 0,62% data sebagai anomali statistik dari total 2.167.942 transaksi. Sistem prototipe yang dirancang menggunakan metodologi *CRISP-DM* dengan arsitektur *Python*, *MySQL*, dan *frontend PHP* terbukti fungsional untuk mendukung validasi oleh ahli domain. Disarankan bagi rumah sakit untuk mengintegrasikan hasil validasi staf sebagai *feedback loop* untuk melatih ulang model secara *semi-supervised*, sehingga akurasi deteksi terus meningkat sesuai konteks operasional. Untuk penelitian selanjutnya, disarankan mengeksplorasi algoritma lain dan memperkaya fitur untuk meningkatkan akurasi model.

DAFTAR PUSTAKA

- Dewa, R. P., & Windarto. (2024). *Deteksi Anomali Jaringan Menggunakan Isolation Forest pada Log Wazuh dengan Pemberitahuan WhatsApp di PT XYZ Network Anomaly Detection Using Isolation Forest on Wazuh Logs with WhatsApp Notifications at PT XYZ*. 4, 208–216.
- Maryati, W., & Utami, Y. T. (2023). OPTIMALISASI MUTU PELAYANAN KESEHATAN DI KLINIK DENGAN IMPLEMENTASI SISTEM INFORMASI REKAM MEDIS BERBASIS WEB. *LINK*, 19(1), 14–18. <https://doi.org/10.31983/link.v19i1.9387>
- Wibawa, I. M. S. T., & Karyawati, A. A. I. N. E. (2023). *Isolation Forest dengan Exploratory Data Analysis pada Anomaly Detection untuk Data Transaksi*.
- Zulfikar, A., Rahmani, F. A., Azizah, N., Perbendaharaan, D. J., Keuangan, K., & Pinang, P. (2023). *Deteksi Anomali Menggunakan Isolation Forest Belanja Barang Persediaan*.

- Konsumsi Pada Satuan Kerja Kepolisian Republik Indonesia. *Jurnal Manajemen Perbendaharaan*, 4(1), 1–15. <https://doi.org/10.33105/jmp.v4i1.435>
- Xu, H., Pang, G., Wang, Y., & Wang, Y. (2023). Deep Isolation Forest for Anomaly Detection. *IEEE Transactions on Knowledge and Data Engineering*, 35(12), 12591–12604. <https://doi.org/10.1109/TKDE.2023.3270293>
- Sunarto, M. W., Kurniawan, D., Siswanto, E., & Huda, H. I. (2023). Deteksi Anomali Menggunakan Extended Isolation Forest (Eif). *Teknik: Jurnal Ilmu Teknik Dan Informatika*, 1(2), 96–111. <https://doi.org/10.51903/teknik.v1i2.324>
- Bakumenko, A., & Elragal, A. (2022). Detecting Anomalies in Financial Data Using Machine Learning Algorithms. *Systems*, 10(5). <https://doi.org/10.3390/systems10050130>
- Dharmawan, D. (2023). *Deteksi Anomali pada Dataset BPJS Tahun 2020*. 1–141.
- Chen, H., Ma, H., Chu, X., & Xue, D. (2020). Anomaly detection and critical attributes identification for products with multiple operating conditions based on isolation forest. *Advanced Engineering Informatics*, 46(July). <https://doi.org/10.1016/j.aei.2020.101139>
- Lesouple, J., Baudoin, C., Spigai, M., & Tournet, J.-Y. (2021). Generalized isolation forest for anomaly detection. *Pattern Recognition Letters*, 149, 109–119. <https://doi.org/10.1016/j.patrec.2021.05.022>