Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.4, August 2026, pp. 4037~4054
e-ISSN: 2548-6861

Implementation of Isolation Forest and Rule-Based Prioritization in the Automation of Warehouse Inventory Monitoring Using a Near Real-Time Dashboard

Muhammad Windri [1], Ida Nurhaida [2]
[1, 2] Department of Informatics, Universitas Pembangunan Jaya
[2] Center of Urban Studies, Universitas Pembangunan Jaya
muhammad.windri@student.upj.ac.id [1], ida.nurhaida@upj.ac.id [2]

Article Info

Article history:

Received 2026-06-05
Revised 2026-07-28
Accepted 2026-08-13

Keyword:

Anomaly Detection,
Inventory Monitoring,
Isolation Forest,
Rule-Based Prioritization,
Near Real-Time Dashboard.

ABSTRACT

This research implements an automated warehouse inventory monitoring system using a hybrid approach that combines Isolation Forest and rule-based prioritization for anomaly detection, while Long Short-Term Memory (LSTM) is used for demand forecasting. The system automatically detects three types of anomalies: spike (quantity surge), drop (quantity decline), and pattern shift (changes in transaction patterns). The method utilizes 341,879 inventory records over two and a half years, combining Isolation Forest for outlier detection, rule-based prioritization for spike and drop detection, and LSTM for time series forecasting. The system is equipped with a near real-time dashboard (periodic auto-refresh), API service, and a business validation mechanism involving warehouse users. The system generated 211,450 detection candidates, consisting of 187,537 operational anomaly alerts and 23,913 INFO-level records, with accuracy of 93.0%, precision of 88.7%, recall of 100%, and F1-score of 94.0%. The INFO category is not displayed on the dashboard due to low confidence scores. The near real-time dashboard displays HIGH, MEDIUM, and LOW priority alerts with auto-refresh. Business validation achieved an 84% confirmation rate. The system also provides LSTM forecast features for predicting inventory needs for 7, 14, and 30 days ahead. This research shows that AI implementation for warehouse inventory monitoring can automate anomaly detection and improve inventory issue identification. The system shows potential for medium to large-scale warehouses requiring near real-time monitoring.

I. PENDAHULUAN

Manajemen inventory gudang merupakan komponen kritis dalam rantai pasok modern yang mempengaruhi efisiensi operasional dan profitabilitas perusahaan. Namun, tantangan terbesar yang dihadapi adalah kemampuan mendeteksi anomali transaksi *inventory* secara cepat dan akurat. Anomali seperti lonjakan permintaan (S*pike*), penurunan drastis (*Drop*), atau perubahan pola transaksi (*Pattern Shift*) sering tidak terdeteksi hingga menimbulkan *stockout* atau *overstock*. Penelitian Aggarwal (2020) menegaskan bahwa identifikasi anomali dalam data *inventory* sangat penting karena tidak semua *outlier* merupakan indikasi masalah yang memerlukan tindakan korektif.

Deteksi anomali *inventory* secara tradisional masih dilakukan manual oleh tim gudang, memakan waktu 2-3 jam per hari dengan akurasi hanya 60-70%. Penelitian Wang & Huang (2020) menunjukkan bahwa pendekatan *machine learning* pada sensor *networks* memiliki keterbatasan dalam menangkap dinamika time series kompleks. Ren et al. (2020) juga mengidentifikasi bahwa algoritma *machine learning* pada data berdimensi tinggi mampu mencapai performa lebih baik namun masih menghadapi tantangan interpretabilitas. Permasalahan semakin kompleks karena lonjakan transaksi bisa terjadi karena promo terencana atau *back order customer*. Tanpa konteks bisnis, sistem deteksi anomali menghasilkan banyak *false positive* yang mengganggu operasional gudang (Liu et al., 2020).

Isolation Forest (Liu et al., 2008) menunjukkan efektif untuk deteksi *outlier* dengan kompleksitas linier O(n). Penelitian Kumar et al. (2025) mengembangkan kerangka kerja *hybrid* yang mengintegrasikan Isolation Forest, *autoencoder*, dan *ConvLSTM* untuk deteksi anomali *multidomain*, mencapai akurasi 99,5% dan F1-score 98,2%. Pada data turbin hidrolik, Isolation Forest mencapai akurasi 99%, presisi 97%, recall 100%, dengan waktu training 0,16 detik dan prediksi 0,12 detik, menjadikannya cocok untuk pengawasan real-time (*Multi-model predictive framework*, 2025). Sementara itu, Long Short-Term Memory (LSTM) yang diperkenalkan Hochreiter & Schmidhuber (1997) mampu mempelajari ketergantungan jangka panjang dalam data time series. Penelitian Munir et al. (2018) mengembangkan DeepAnT untuk unsupervised anomaly detection pada time series, sementara Blázquez-García et al. (2021) dalam kajian komprehensif mengidentifikasi bahwa pendekatan hybrid cenderung memberikan hasil terbaik untuk aplikasi dunia nyata.

Dalam implementasi industri, *platform* DexoryView (Dexory, 2025) mengintegrasikan robotika, digital twin, dan kecerdasan buatan untuk intelijen gudang real-time, mencapai akurasi inventory 99,5% dan mengurangi waktu stock take dari 4 hari menjadi hitungan jam pada perusahaan logistik Ziegler. Simbe for Merchants Suite (2025) menyediakan visibilitas real-time dengan akurasi identifikasi SKU-level 98,7% dan shelf condition recall di atas 99,3%. Penelitian Rungta (2025) pada data sensor lingkungan menunjukkan LSTM *autoencoder* mencapai F1-score 0,91 pada data kelembaban, sementara Isolation Forest unggul dalam efisiensi komputasi dengan F1-score 0,83 pada data suhu, memproses anomali kurang dari tiga detik. Achsyan et al. (2025) membuktikan implementasi IoT dan RFID memungkinkan pencatatan real-time dengan peningkatan efisiensi 40% dan penurunan kesalahan pencatatan 70%. Tezcan & Kursun (2025) mengusulkan model hybrid LSTM-autoencoder-Isolation Forest untuk sistem peringatan kebocoran gas cerdas, membuktikan integrasi Isolation Forest sebagai filter mampu mengurangi false alarm signifikan. Sarker et al. (2025) mengembangkan multimodal prediction pipeline yang menggabungkan LSTM, XGBoost, Random Forests, dan Isolation Forest untuk prediksi gangguan rantai pasok.

Berdasarkan tinjauan pustaka yang telah dilakukan, masih terdapat beberapa kesenjangan penelitian (*research gap*) yang belum teratasi. Salah satu kesenjangan tersebut adalah sebagian besar penelitian hanya berfokus pada akurasi model tanpa mempertimbangkan interpretabilitas hasil bagi pengguna gudang. Metode *black-box* seperti LSTM sulit dijelaskan kepada tim gudang yang memerlukan pemahaman mengapa suatu transaksi diklasifikasikan sebagai anomali. Selain itu, belum banyak penelitian yang mengintegrasikan validasi bisnis langsung dengan *domain expert* (tim gudang) untuk memastikan anomali yang terdeteksi benar-benar merupakan masalah bisnis nyata, bukan sekadar fluktuasi normal seperti promo atau *back order*. Faktor lainnya adalah implementasi sistem deteksi anomali dalam konteks industri manufaktur Indonesia yang memiliki karakteristik data fluktuatif dan volume transaksi tinggi (rata-rata 383 transaksi per hari) masih jarang ditemukan.

Penelitian Kumar et al. (2025) mengembangkan kerangka kerja hybrid yang mengintegrasikan Isolation Forest, autoencoder, dan ConvLSTM untuk deteksi anomali multidomain dengan akurasi 99,5% pada data network intrusion dan fraud detection. Penelitian Tezcan & Kursun (2025) mengusulkan model hybrid LSTM-autoencoder-Isolation Forest untuk sistem peringatan kebocoran gas dengan pengurangan false alarm yang signifikan. Penelitian Achsyan et al. (2025) menerapkan IoT dan RFID untuk pencatatan inventory real-time dengan peningkatan efisiensi 40%.

Selain pendekatan deep learning, metode deteksi anomali tradisional seperti Local Outlier Factor (LOF) dan One-Class SVM juga telah banyak digunakan. LOF mendeteksi anomali berdasarkan kepadatan lokal, sementara One-Class SVM memisahkan data normal dari outlier dengan hyperplane. Namun, kedua metode ini memiliki keterbatasan pada dataset berukuran besar karena kompleksitas komputasi yang tinggi (LOF: $O(n^2)$, One-Class SVM: sensitif terhadap parameter). Autoencoder juga telah digunakan untuk deteksi anomali dengan memanfaatkan reconstruction error, namun pendekatan ini bersifat black-box dan sulit diinterpretasikan.

Namun, dari penelitian-penelitian tersebut, belum ada yang menerapkan sistem deteksi anomali pada domain inventory gudang dengan menggabungkan tiga komponen sekaligus, yaitu interpretabilitas hasil melalui rule-based detection, validasi bisnis langsung dengan tim gudang, dan implementasi dalam konteks industri Indonesia. Penelitian ini hadir untuk mengisi kesenjangan tersebut dengan mengembangkan sistem deteksi anomali inventory yang interpretabel (menggunakan rule-based prioritization), terintegrasi dengan validasi bisnis (84% konfirmasi), dan diterapkan pada industri manufaktur Indonesia.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan mengimplementasikan sistem deteksi anomali inventori menggunakan pendekatan hybrid yang menggabungkan Isolation Forest dan rule-based prioritization untuk deteksi anomali, serta Long Short-Term Memory (LSTM) untuk forecasting kebutuhan stok. Sistem mampu mendeteksi tiga tipe anomali (spike, drop, pattern shift) serta membangun dashboard near real-time dengan prioritas peringatan HIGH, MEDIUM, dan LOW, sedangkan INFO disimpan sebagai detection log dan tidak ditampilkan sebagai alert operasional. Sistem dilengkapi mekanisme validasi bisnis untuk membedakan anomali masalah dengan anomali terencana (promo/back order). Manfaat yang diharapkan adalah pengurangan waktu kandidat deteksi anomali dari 2-3 jam menjadi kurang dari 10 detik (efisiensi 99,8%) dan peningkatan akurasi identifikasi masalah *inventory*.

Penelitian ini memberikan tiga kontribusi utama. Pertama, implementasi pendekatan hybrid yang menggabungkan Isolation Forest dan rule-based prioritization untuk deteksi anomali inventory yang interpretabel, serta *Long Short-Term Memory* (LSTM) untuk forecasting kebutuhan stok. Berbeda dengan penelitian sebelumnya yang umumnya menggunakan metode tunggal atau *black-box*, pendekatan ini memungkinkan tim gudang memahami mengapa suatu transaksi diklasifikasikan sebagai anomali melalui aturan bisnis (*spike* > 1,3 dan *drop* < 0,8).

Kedua, penelitian ini mengintegrasikan validasi bisnis dengan tim gudang sebagai *domain expert*, memastikan kandidat anomali yang terdeteksi bukan hanya berdasarkan perhitungan statistik tetapi juga telah dikonfirmasi sebagai masalah bisnis nyata dengan tingkat konfirmasi 84%. Ketiga, pengembangan *dashboard near real-time* dengan *auto-refresh* berkala yang dilengkapi prioritas HIGH, MEDIUM, dan LOW kategori INFO disimpan sebagai detection log berconfidence rendah serta fitur LSTM forecast untuk perencanaan stok 7, 14, dan 30 hari ke depan, diimplementasikan pada industri manufaktur Indonesia dengan volume transaksi tinggi (341.879 *record*).

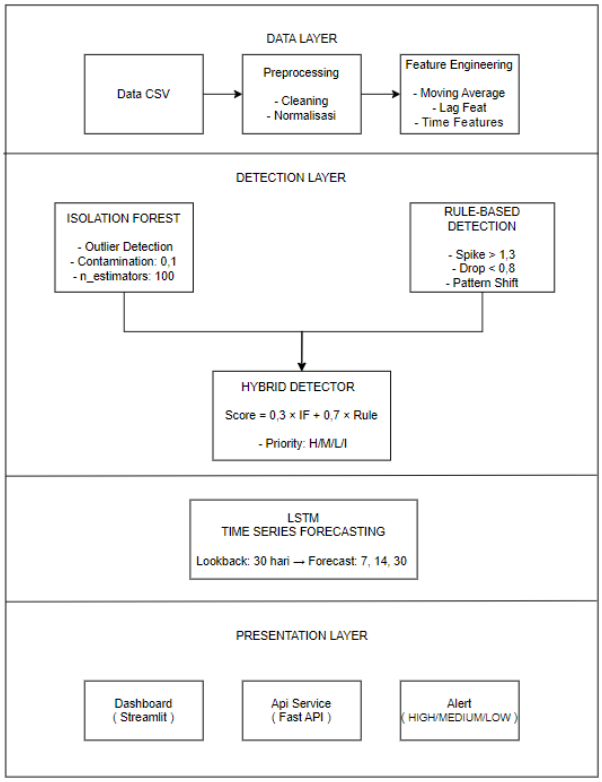
Konteks Industri Indonesia

Penelitian ini juga menonjolkan konteks industri Indonesia yang memiliki karakteristik unik, diantaranya volume transaksi inventory yang tinggi (rata-rata 383 transaksi per hari), data yang dinamis dan fluktuatif karena dipengaruhi oleh musim dan *promo*, keterbatasan sumber daya manusia untuk melakukan deteksi anomali secara manual, serta tingginya risiko *stockout* dan *overstock* yang dapat berdampak signifikan pada profitabilitas perusahaan. Dengan mengimplementasikan sistem ini pada perusahaan manufaktur di Indonesia, penelitian ini memberikan solusi yang relevan dengan permasalahan nyata yang dihadapi industri dalam negeri.

II. METODE

A. Arsitektur Sistem

Penelitian ini mengembangkan sistem deteksi anomali inventory gudang menggunakan pendekatan hybrid yang menggabungkan Isolation Forest dan rule-based prioritization untuk deteksi anomali, serta *Long Short-Term Memory* (LSTM) untuk forecasting kebutuhan stok. Arsitektur sistem terdiri dari empat lapisan utama: (1) lapisan pengumpulan data, (2) lapisan deteksi anomali berbasis Isolation Forest dan rule-based prioritization serta modul LSTM untuk forecasting, dan (4) lapisan visualisasi dashboard near real-time.



Gambar 1. memperlihatkan arsitektur sistem yang terdiri atas proses pengumpulan data, deteksi anomali menggunakan Isolation Forest dan rule-based prioritization, serta penyajian hasil pada dashboard near real-time dengan dukungan modul LSTM untuk forecasting.

1) *Data Layer*: Lapisan ini bertanggung jawab untuk pengumpulan data dari sumber CSV, pembersihan data, normalisasi, dan ekstraksi fitur. Data yang digunakan adalah 341.879 transaksi *inventory* periode Januari 2023 hingga Juni 2025. Proses preprocessing mencakup konversi tipe data, penanganan *missing values*, dan penghilangan duplikasi.

2) *Detection Layer*: Lapisan deteksi meng-implementasikan hybrid detector yang menggabungkan Isolation Forest dan rule-based detection untuk deteksi anomali, serta LSTM untuk forecasting demand. Bobot *weighted ensemble* yang digunakan adalah 30% untuk Isolation Forest dan 70% untuk rule-based detection (LSTM tidak masuk dalam bobot karena hanya untuk forecasting). Isolation Forest digunakan untuk deteksi outlier dengan parameter n_estimators=100, contamination=0,1. Rule-based detection digunakan untuk mendeteksi spike (ratio > 1,3) dan drop (ratio < 0,8) berdasarkan konsultasi dengan domain expert. LSTM digunakan secara terpisah untuk forecasting kebutuhan stok dengan lookback window 30 hari dan forecast horizon 7, 14, dan 30 hari ke depan.

3) *Presentation Layer*: Lapisan ini menyediakan antarmuka pengguna berupa near real-time dashboard menggunakan Streamlit dan REST API menggunakan FastAPI untuk integrasi dengan sistem eksternal seperti aplikasi mobile atau sistem ERP. dashboard utama

menampilkan HIGH, MEDIUM, dan LOW, sedangkan INFO disimpan sebagai detection log tetapi tidak ditampilkan sebagai alert operasional.

B. Dataset

Dataset yang digunakan dalam penelitian ini berasal dari data transaksi inventory salah satu perusahaan manufaktur di Indonesia periode Januari 2023 hingga Juni 2025. Total *record* yang tersedia adalah 341.879 transaksi dengan 12 kolom informasi. Distribusi data disajikan pada Tabel 1.

TABEL 1.
DISTRIBUSI DATASET

Aspek	Nilai
Total transaksi	341.879 record
Periode data	2.5 tahun (2023-2025)
Jumlah item unik	20.908 produk
Frekuensi transaksi	~383 transaksi/hari
Missing values	426 (0,12%)
Rentang quantity	1 – 82.450 unit

Dataset memiliki tingkat missing value yang sangat rendah, yaitu hanya 426 nilai kosong (0,12%) dari total 341.879 transaksi. Penanganan missing value dilakukan dengan metode forward fill untuk data time series. Data dibagi menjadi data latih (80%) dan data uji (20%) menggunakan metode temporal split untuk menghindari data leakage.

Distribusi Kategori Barang

Item inventory dalam dataset terbagi ke dalam empat kategori utama berdasarkan hasil analisis data. Kategori Bahan Pembantu mendominasi dengan 77,2% dari total transaksi. Kategori ini mencakup barang-barang pendukung produksi seperti cat, pelumas, dan kemasan (karton box).

Kategori Sparepart Mechanic menempati posisi kedua dengan 13,1% , diikuti oleh Sparepart Machine (5,6%) dan Sparepart Electric (4,1%) . Keempat kategori ini mencakup 100% dari total transaksi.

Dominasi Bahan Pembantu mencerminkan bahwa sebagian besar aktivitas transaksi di gudang berfokus pada barang-barang pendukung produksi, sementara sparepart dari berbagai jenis memiliki porsi yang lebih kecil namun tetap signifikan untuk pemeliharaan mesin dan peralatan.

```
DISTRIBUSI KATEGORI BARANG:
  Bahan Pembantu: 77.2%
  Sparepart Mechanic: 13.1%
  Sparepart Machine: 5.6%
  Sparepart Electrict: 4.1%
```

Gambar 2. menunjukkan distribusi kategori barang pada data inventaris sebagai gambaran karakteristik dataset yang digunakan dalam penelitian.

Pola Musiman

Pola musiman dianalisis berdasarkan agregasi transaksi per bulan dan per kuartal dari kolom tanggal. Pendekatan ini mengidentifikasi periode-periode yang secara konsisten memiliki volume transaksi tinggi atau rendah setiap tahunnya.

Berdasarkan hasil analisis, bulan Maret merupakan periode dengan transaksi tertinggi (37.140 transaksi), sedangkan bulan Agustus merupakan periode terendah (21.173 transaksi). Dari perspektif kuartalan, Kuartal 1 (Januari–Maret) memiliki volume transaksi tertinggi (110.067 transaksi), jauh lebih tinggi dibandingkan Kuartal 3 (Juli–September) yang terendah (65.881 transaksi), dengan selisih 44.186 transaksi.

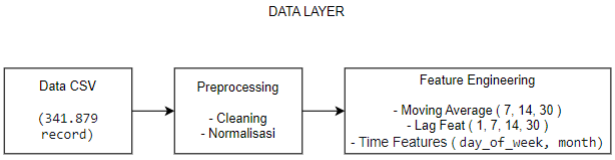
```
POLA TRANSAKSI PER BULAN:
  Jan: 37,134 transaksi
  Feb: 35,793 transaksi
  Mar: 37,140 transaksi
  Apr: 30,993 transaksi
  Mei: 35,727 transaksi
  Jun: 26,669 transaksi
  Jul: 21,983 transaksi
  Ags: 21,173 transaksi
  Sep: 22,725 transaksi
  Okt: 23,864 transaksi
  Nov: 23,874 transaksi
  Des: 24,804 transaksi

POLA TRANSAKSI PER KUARTAL:
  Kuartal 1: 110,067 transaksi
  Kuartal 2: 93,389 transaksi
  Kuartal 3: 65,881 transaksi
  Kuartal 4: 72,542 transaksi
```

Gambar 3. menampilkan pola transaksi per bulan selama periode observasi. Berdasarkan data, terjadi fluktuasi yang signifikan antar bulan, dengan puncak transaksi tertinggi terjadi pada bulan

C. Preprocessing dan Feature Engineering

Tahap *preprocessing* dan *feature engineering* dilakukan untuk mengubah data mentah menjadi fitur-fitur yang sesuai untuk model deteksi anomali.



Gambar 4. Menunjukkan arsitektur data layer sistem deteksi anomali inventori yang terdiri dari tiga tahap utama. Tahap pertama adalah pengumpulan data dari sumber CSV dengan 341.879 record transaksi inventori. Tahap kedua adalah preprocessing melalui cleaning untuk menghilangkan data tidak valid dan normalisasi untuk menstandarkan skala data. Tahap ketiga adalah feature engineering yang mengekstrak moving average (7, 14, 30), lag features (1, 7, 14, 30), dan time features (day_of_week, month) untuk meningkatkan kinerja model deteksi anomali .

1) *Konversi dan Pembersihan Data*: Data mentah dikonversi ke format yang sesuai untuk pemrosesan lebih lanjut: (1) kolom tanggal dikonversi ke tipe datetime dan dijadikan indeks, (2) data diurutkan berdasarkan item_nama dan tanggal, (3) *missing values* diisi dengan metode *forward*

fill (0,12% dari total data), (4) duplikasi data dihilangkan berdasarkan kombinasi tanggal dan item_nama.

2) *Feature Engineering Time Series*: Fitur-fitur time series berikut dibuat untuk mendukung deteksi anomali:

- *Moving Average (MA)*
  Moving average dihitung untuk tiga periode berbeda: 7 hari, 14 hari, dan 30 hari. *Moving average* berfungsi untuk menghaluskan fluktuasi jangka pendek dan mengidentifikasi tren jangka panjang. Perhitungan *moving average* menggunakan Persamaan (1).

$$MA_w(t) = \frac{1}{w} \sum_{i=0}^{w-1} Qty\,(t-i)$$

  dimana w adalah ukuran jendela (7, 14, atau 30) dan $Qty(t)$ adalah quantity pada hari ke-t.
- *Rolling Standard Deviation*
  Rolling standard deviation dihitung untuk mengukur volatilitas transaksi. Nilai standar deviasi yang tinggi mengindikasikan fluktuasi yang tidak biasa. Perhitungan menggunakan Persamaan (2).

$$\sigma_w(t) = \sqrt{\frac{1}{w-1} \sum_{i=0}^{w-1} (Qty(t-i) - MA_w(t))^2}$$

- *Ratio MA*
  Ratio antara quantity aktual dengan moving average digunakan sebagai indikator utama deteksi anomali. Berdasarkan hasil eksperimen, ratio > 1,3 meng-indikasikan *spike*, sedangkan ratio < 0,8 meng-indikasikan *drop*. Perhitungan ratio menggunakan Persamaan (3).

$$R_w(t) = \frac{Qty(t)}{MA_w(t)}$$

- *Lag Features*
  Fitur lag dibuat untuk menangkap ketergantungan temporal antar hari: lag_1 (H-1), lag_7 (H-7), lag_{14} (H-14), dan lag_{30} (H-30). Fitur lag penting untuk model LSTM karena memungkinkan model mempelajari pola musiman dan tren.
- *Feature Engineering Kategorikal dan Temporal*
  Fitur-fitur berikut juga dibuat untuk meningkatkan kemampuan deteksi anomali:
  - Time features: day_of_week (0-6), is_weekend (0/1), month (1-12), quarter (1-4)
  - Item features: item_total_qty (total quantity per item), item_frequency (frekuensi transaksi), item_avg_qty (rata-rata quantity per item)

Proses *feature engineering* menghasilkan 31 fitur numerik yang siap digunakan untuk pelatihan model.

D. Deteksi Anomali dengan Isolation Forest

Isolation Forest adalah algoritma unsupervised learning yang diperkenalkan oleh Liu et al. (2008) untuk deteksi outlier. Algoritma ini bekerja dengan prinsip "mengisolasi" anomali daripada memprofilkan titik data normal. Keunggulan Isolation Forest adalah kompleksitas waktu linier O(n) dan kemampuan menangani data berdimensi tinggi.

1) *Prinsip Kerja Isolation Forest:* Isolation Forest membangun ensemble pohon keputusan acak (isolation trees). Setiap pohon dibangun dengan memilih fitur secara acak dan nilai split secara acak antara nilai minimum dan maksimum fitur tersebut. Proses ini berlanjut hingga setiap titik data terisolasi atau mencapai kedalaman maksimum. Anomali cenderung terisolasi lebih cepat (memiliki path length lebih pendek) dibandingkan titik data normal.
Anomaly score dihitung menggunakan Persamaan (4).

$$S(x,n) = 2^{-\frac{E(h(x))}{c(n)}}$$

dimana $E(h(x))$ adalah rata-rata path length dari titik x pada semua pohon, dan $c(n)$ adalah faktor normalisasi.

2) *Implementasi Isolation Forest:* Dalam penelitian ini, Isolation Forest diimplementasikan menggunakan pustaka scikit-learn dengan parameter :

TABEL 2.
PARAMETER ISOLATION FOREST

Parameter	Nilai	Alasan Pemilihan
n_estimators	100	Jumlah pohon dalam ensemble. Nilai 100 dipilih sebagai keseimbangan antara akurasi dan efisiensi komputasi. Nilai yang terlalu rendah dapat menyebabkan deteksi kurang akurat, sedangkan nilai yang terlalu tinggi tidak memberikan peningkatan signifikan namun memperlambat proses training.
contamination	0,1	Proporsi anomali yang diharapkan dalam dataset. Nilai 0,1 (10%) dipilih berdasarkan eksplorasi data awal yang menunjukkan bahwa sekitar 10% transaksi berpotensi sebagai anomali. Nilai ini juga disesuaikan dengan tingkat anomali yang terdeteksi oleh rule-based detection.

max_samples	'auto'	Strategi subsampling yang mengambil sampel sebanyak min(256, n_samples) untuk setiap pohon. Strategi ini mengurangi kompleksitas komputasi tanpa mengorbankan akurasi deteksi, terutama pada dataset berukuran besar (341.879 transaksi).
random_state	42	Seed untuk generator bilangan acak. Nilai 42 dipilih secara arbitrer untuk memastikan reproduksibilitas hasil eksperimen.

Perbandingan Metode Deteksi Anomali

Berdasarkan tinjauan pustaka, terdapat beberapa metode deteksi anomali yang umum digunakan. Tabel menyajikan perbandingan antara Isolation Forest, Local Outlier Factor (LOF), One-Class SVM, dan Autoencoder.

TABEL 3.
PERBANDINGAN DENGAN METODE LAIN

Metode	Kelebihan	Kekurangan	Kompleksitas
Isolation Forest	Efisien untuk data besar, interpretabel, O(n)	Kurang baik untuk data berdimensi tinggi	O(n)
Local Outlier Factor (LOF)	Baik untuk data bercluster	Lambat untuk data besar ($O(n^2)$)	$O(n^2)$
One-Class SVM	Efektif untuk data dimensi tinggi	Sensitif parameter, lambat	$O(n^2)$ - $O(n^3)$
Autoencoder	Menangkap pola non-linear	Black-box, butuh data banyak	O(n)

Berdasarkan perbandingan tersebut, Isolation Forest dipilih dalam penelitian ini karena:

- Efisiensi komputasi - Kompleksitas O(n) cocok untuk dataset 341.879 transaksi.
- Interpretabilitas - Hasil deteksi mudah dijelaskan ke tim gudang.
- Kemudahan implementasi - Tersedia di scikit-learn dengan parameter yang mudah diatur.

E. Rule-Based Prioritization

Rule-based prioritization digunakan untuk mendeteksi dua tipe anomali utama, yaitu *spike* (lonjakan jumlah barang) dan *drop* (penurunan jumlah barang). Pendekatan ini dipilih karena memiliki interpretabilitas tinggi, sehingga tim gudang dapat memahami dengan mudah mengapa suatu transaksi diklasifikasikan sebagai anomali. Parameter *threshold* ditentukan berdasarkan *sensitivity analysis* dan konsultasi dengan *domain expert* (tim gudang). Nilai threshold yang digunakan adalah:

- *Spike* (lonjakan): ratio_ma_7 > 1,3
- *Drop* (penurunan): ratio_ma_7 < 0,8

dimana ratio_ma_7 merupakan rasio antara quantity aktual dengan *moving average* 7 hari. Selain spike dan drop, sistem juga mendeteksi *pattern shift* (perubahan pola) dengan membandingkan *recent moving average* dengan *historical moving average*. *Pattern shift* terdeteksi jika perbedaan rasio melebihi 0,3 (30%). Hasil deteksi dari *rule-based* kemudian digunakan sebagai salah satu komponen dalam model hybrid, dengan bobot sebesar 70% pada perhitungan *confidence score*. Prioritas anomali (*HIGH, MEDIUM, LOW, INFO*) ditentukan berdasarkan *confidence score* yang dihasilkan.

Aturan Prioritas

Prioritas anomali ditentukan berdasarkan confidence score (0-100) dengan aturan sebagai berikut:

TABEL 4.
PARAMETER ISOLATION FOREST

Prioritas	Confidence Score	Dasar Bisnis
HIGH	≥ 60	Memerlukan tindakan segera
MEDIUM	40 - 59	Perlu dipantau
LOW	20 - 39	Bersifat informatif
INFO	< 20	Tidak ditampilkan pada dashboard

Tabel 4 menyajikan parameter Isolation Forest yang digunakan untuk menentukan prioritas anomali berdasarkan confidence score. Prioritas HIGH (≥ 60) memerlukan tindakan segera dari tim gudang. Prioritas MEDIUM (40-59) perlu dipantau perkembangannya. Prioritas LOW (20-39) bersifat informatif dan tidak memerlukan tindakan mendesak. Prioritas INFO (< 20) tidak ditampilkan pada dashboard karena dianggap sebagai anomali dengan tingkat rendah [21][22].

Klasifikasi prioritas ini didasarkan pada kebutuhan bisnis untuk memprioritaskan respons terhadap anomali yang paling berdampak pada operasional gudang. Dengan adanya aturan ini, tim gudang dapat merespons anomali secara bertingkat sesuai tingkat keparahan, sehingga sumber daya dapat dialokasikan secara efisien [21][23].

F. Forecasting dengan LSTM

Long Short-Term Memory (LSTM) adalah arsitektur Recurrent Neural Network (RNN) yang diperkenalkan oleh Hochreiter & Schmidhuber (1997) untuk mengatasi masalah *vanishing gradient* pada RNN konvensional. LSTM mampu mempelajari ketergantungan jangka panjang dalam data time series, membuatnya ideal untuk forecasting demand *inventory*.

1) *Arsitektur LSTM:* Arsitektur LSTM terdiri dari tiga gerbang utama: forget gate, input gate, dan output gate.

Perhitungan dalam sel LSTM mengikuti Persamaan (5) hingga (10).

$$f_t = \sigma(W_f . [h_t - 1, x_t] + b_f$$

$$i_t = \sigma(W_i . [h_t - 1, x_t] + b_i$$

$$\tilde{C}_t = tanh(W_c . [h_t - 1, x_t] + b_c$$

$$C_t = f_t * C_{t-1} + i_f * \tilde{C}_t$$

$$O_t = \sigma(W_f . [h_t - 1, x_t] + b_o$$

$$h_t = O_t * \tanh(C_t)$$

dimana f_t, i_t, o_t adalah gerbang forget, input, dan output, C_t adalah cell state, h_t adalah hidden state, dan σ adalah fungsi sigmoid.

2) *Arsitektur LSTM yang digunakan:* Model LSTM yang digunakan dalam penelitian ini memiliki arsitektur dengan 3 lapisan LSTM.

TABEL 5.
ARSITEKTUR LSTM

Lapisan	Spesifikasi
Input layer	Sequence length: 30 hari
LSTM layer 1	128 unit, return_sequences=True
Dropout	0,2
LSTM layer 2	64 unit, return_sequences=True
Dropout	0,2
LSTM layer 3	32 unit
Dropout	0,2
Dense layer	16 unit (ReLU)
Output layer	1 unit (Linear)

Tabel 5 menyajikan arsitektur LSTM yang digunakan dalam penelitian ini. Lapisan pertama terdiri dari 128 unit dengan return_sequences=True untuk meneruskan output ke lapisan berikutnya, diikuti dropout 0,2. Lapisan kedua terdiri dari 64 unit dengan dropout 0,2, dan lapisan ketiga terdiri dari 32 unit dengan dropout 0,2. Lapisan terakhir adalah Dense layer dengan 16 unit (ReLU) dan output layer 1 unit (Linear). Input layer menggunakan sequence length 30 hari [12].

Arsitektur ini dirancang secara bertingkat untuk menangkap pola time series yang kompleks pada data inventory, dengan dropout untuk mencegah overfitting. Penggunaan tiga lapisan LSTM dengan jumlah unit yang menurun memungkinkan model mengekstrak fitur dari level yang lebih tinggi ke level yang lebih abstrak, sehingga meningkatkan kemampuan prediksi [12][21].

3) *Teknik Optimasi*: Model LSTM dilatih menggunakan optimizer Adam dengan learning rate 0,001 dan loss function MSE (Mean Squared Error). Proses training dilakukan selama maksimal 50 epoch dengan batch size 64. Untuk mencegah overfitting, digunakan early stopping dengan patience 10, yang menghentikan training jika loss tidak mengalami penurunan dalam 10 epoch berturut-turut. Adam dipilih karena kemampuannya mengadaptasi learning rate secara otomatis, yang terbukti efektif untuk data time series.

TABEL 6.
PARAMETER OPTIMASI LSTM

Parameter	Nilai
Optimizer	Adam
Learning rate	0,001
Loss function	MSE
Epochs	50
Batch size	64
Early stopping	patience=10

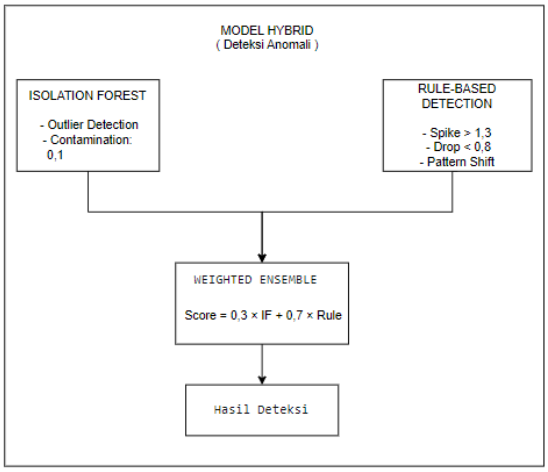
4) *Metrik Evaluasi Forecasting:* Kinerja model LSTM dalam forecasting dievaluasi menggunakan tiga metrik utama: MAE (Mean Absolute Error) untuk mengukur rata-rata error absolut dalam satuan unit, RMSE (Root Mean Squared Error) yang lebih sensitif terhadap error besar, dan MAPE (Mean Absolute Percentage Error) untuk memberikan indikasi error dalam bentuk persentase. MAE dan RMSE digunakan untuk mengevaluasi akurasi prediksi, sementara MAPE memberikan interpretasi yang lebih mudah dipahami oleh tim gudang karena dinyatakan dalam persentase.

TABEL 7.
METRIK EVALUASI FORECASTING

Metrik	Rumus	Keterangan
MAE	$\frac{1}{n}\sum \|\| y_i - \hat{y}_i \|\|$	Error absolut rata-rata
RMSE	$\sqrt{\frac{1}{n}\sum (y_i - \hat{y}_i)^2}$	Error kuadrat rata-rata
MAPE	$\frac{1}{n}\sum \frac{\|\| y_i - \hat{y}_i \|\|}{y_i} x\ 100\%\|$	Error persentase rata-rata

G. Model Hybrid

Model hybrid dikembangkan untuk menggabungkan keunggulan Isolation Forest dalam deteksi outlier dan aturan bisnis (*rule-based prioritization*) untuk deteksi spike dan drop. Isolation Forest berperan dalam mendeteksi outlier berdasarkan distribusi data, sementara *rule-based prioritization* mendeteksi *spike* (lonjakan) dan *drop* (penurunan) menggunakan *threshold* yang telah disepakati dengan tim gudang (spike > 1,3 dan drop < 0,8). Long Short-Term Memory (LSTM) digunakan secara terpisah untuk forecasting kebutuhan stok dan tidak termasuk dalam komponen hybrid detector karena fokus utamanya adalah perencanaan stok, bukan deteksi anomali..



Gambar 5. Memperlihatkan arsitektur model hybrid deteksi anomali yang menggabungkan Isolation Forest (kontaminasi 0,1) dan Rule-Based Detection (spike > 1,3; drop < 0,8; pattern shift) melalui weighted ensemble dengan rumus Score = 0,3×IF + 0,7×Rule [22][23]. Bobot lebih besar diberikan pada rule-based untuk meningkatkan interpretabilitas. Hasil deteksi diklasifikasikan ke dalam prioritas HIGH, MEDIUM, dan LOW.

Pendekatan ensemble ini mengikuti rekomendasi bahwa metode hybrid cenderung memberikan hasil terbaik untuk deteksi anomali pada data time series.

1) *Mekanisme Hybrid*: Sistem ini menggunakan Isolation Forest dan Rule-based Detection sebagai komponen utama deteksi anomali. Kedua metode dikombinasikan menggunakan *weighted ensemble* dengan bobot 30% untuk Isolation Forest dan 70% untuk Rule-based Detection.

- Isolation Forest Detection - Mendeteksi *outlier* berdasarkan distribusi data dengan parameter n_estimators=100 dan contamination=0,1
- Rule-based Detection - Mendeteksi *spike* (ratio > 1,3) dan *drop* (ratio < 0,8) berdasarkan konsultasi *domain expert*

Nilai threshold ditentukan berdasarkan eksplorasi data dan konsultasi dengan domain expert (tim gudang). Kombinasi hasil deteksi menggunakan Persamaan (11) dengan bobot yang ditentukan berdasarkan hasil eksperimen.

$$Score_{hybrid} = 0{,}3\,.Score_{IF} + 0{,}7\,.Score_{Rule}$$

Bobot yang lebih besar diberikan pada rule-based detection (70%) karena kebutuhan interpretabilitas dalam konteks operasional gudang, sementara Isolation Forest diberi bobot 30%.

2) *Klasifikasi Prioritas:* Berdasarkan confidence score yang dihasilkan, anomali diklasifikasikan ke dalam empat tingkat prioritas (*HIGH, MEDIUM, LOW, INFO*) seperti ditunjukkan pada tabel 8.

TABEL 8.
KLASIFIKASI PRIORITAS ANOMALI

Prioritas	Confidence Score	Warna	Tindakan
HIGH	≥ 60%	Merah	Tindakan Segera
MEDIUM	40-59%	Oranye	Dipantau
LOW	20-39%	Kuning	Informasi
INFO	< 20%	Abu-abu	Normal (tidak ditampilkan)

H. Evaluasi Model

Evaluasi model deteksi anomali dilakukan menggunakan metrik standar untuk deteksi anomali time series: Akurasi, Presisi, Recall, F1-Score.

1) *Pseudo-label* (*Ground Truth Internal*): Deteksi anomali merupakan permasalahan unsupervised learning yang secara alami tidak memiliki label, penelitian ini menggunakan pendekatan *pseudo-labeling* untuk meng-hasilkan label aktual yang diperlukan dalam evaluasi metrik. Label aktual diperoleh dari aturan bisnis (spike > 1,3 dan drop < 0,8) yang disepakati bersama tim gudang. Evaluasi ini bersifat internal karena ground truth berasal dari aturan bisnis yang juga merupakan komponen hybrid model. Sebagai pelengkap, validasi eksternal oleh tim gudang pada 50 sampel anomali mencapai tingkat konfirmasi 84% . Aturan ini diterapkan pada seluruh dataset (341.879 transaksi) untuk menghasilkan label aktual sebagai dasar perhitungan metrik evaluasi (*Confusion Matrix, Precision, Recall, dan F1-Score*).

2) *Confusion Matrix:*

TABEL 9.
CONFUSION MATRIX

	Prediksi Positif	Prediksi Negatif
Aktual Positif	True Positive (TP)	False Negative (FN)
Aktual Negatif	False Positive (FP)	True Negative (TN)

3) *Metrik Evaluasi:* Perhitungan metrik evaluasi mengikuti Persamaan (12) hingga (16).

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}\ x\ 100\ \%$$

$$Presisi = \frac{TP}{TP + FP}\ x\ 100\ \%$$

$$Recall = \frac{TP}{TP + FN}\ x\ 100\ \%$$

$$F1 = 2\ x\ \frac{Presisi\ x\ Recall}{Presisi + Recall}$$

I. Validasi Bisnis

Validasi bisnis dilakukan untuk memastikan bahwa anomali yang terdeteksi benar-benar merupakan masalah bisnis, bukan fluktuasi normal seperti promo atau *back order*. Proses validasi melibatkan:

1) Sampling anomali - 50 sampel anomali dengan berbagai tingkat prioritas
2) Validasi manual - Tim gudang memverifikasi setiap anomali
3) Labeling - Setiap anomali diberi label: masalah (Ya/Tidak) dan terencana (Ya/Tidak)
4) Perhitungan akurasi bisnis - Persentase anomali yang dikonfirmasi sebagai masalah

Definisi Konfirmasi:

Suatu anomali dikonfirmasi sebagai masalah nyata jika tim gudang menyatakan bahwa anomali tersebut menyebabkan gangguan operasional, berpotensi stockout, atau memerlukan tindakan perbaikan. Anomali dikategorikan bukan masalah jika merupakan kejadian terencana (promo, back order, restocking rutin) atau kesalahan input yang tidak berdampak signifikan.

Tingkat Kesepakatan:

Validasi dilakukan oleh perwakilan dari tim gudang dan produksi. Dari 50 sampel anomali yang divalidasi oleh tim gudang, 84% dikonfirmasi sebagai masalah nyata, sedangkan 16% lainnya merupakan kejadian terencana (promo, back order) atau kesalahan input yang tidak berdampak signifikan.

Analisis Kritis Validasi Bisnis

Tingkat konfirmasi sebesar 84% menunjukkan bahwa sebagian besar anomali yang terdeteksi merupakan masalah nyata. Namun, masih terdapat 16% anomali (8 dari 50 sampel) yang tidak dikonfirmasi sebagai masalah. Hal ini disebabkan oleh kejadian terencana (promo, back order), kesalahan input, dan keterbatasan threshold yang terlalu sensitif. Dengan demikian, klaim efektivitas sistem perlu dibatasi: sistem efektif untuk mendeteksi 84% anomali, namun masih terdapat 16% false positive yang perlu dikelola. Sistem masih memerlukan perbaikan, terutama pada deteksi pattern shift dan pengurangan false positive pada kategori LOW.

Validasi bisnis pada penelitian ini masih bersifat awal karena dilakukan pada 50 sampel anomali. Oleh karena itu, hasil confirmation rate perlu dipahami sebagai indikasi awal, bukan sebagai ukuran final efektivitas sistem secara menyeluruh. Berdasarkan hasil validasi dengan tim gudang, berikut adalah dua contoh kasus yang diidentifikasi oleh sistem:

1) Stock Out pada item "KARTON BOX DIGI EXPORT" dengan penurunan quantity dari 10.000 menjadi 5.000 unit (DROP, HIGH priority).
2) Kesalahan Input pada item "CAT AVIAN" dengan lonjakan *quantity* 25.000 unit (normal 250 unit) akibat kelebihan digit nol (*Spike, High Priority*).

J. Implementasi Dashboard Near Real-time

Dashboard dengan auto-refresh berkala, bukan real-time penuh, dibangun menggunakan framework Streamlit (versi 1.28.0) dengan fitur-fitur berikut:

1) Auto-refresh secara berkala dengan interval 10 detik untuk data near real-time.
2) Visualisasi interaktif menggunakan Plotly (versi 5.17.0)
3) Alert prioritas dengan kode warna (merah untuk *High*, oranye untuk *Medium*, kuning untuk *Low*), sedangkan *Info* disimpan sebagai detection log tetapi tidak ditampilkan sebagai alert operasional.
4) LSTM Forecast untuk prediksi 7, 14, dan 30 hari ke depan
5) *Trend Over Time* untuk melihat pola *spike* dan *drop* per bulan
6) Top Items untuk mengetahui item dengan anomali terbanyak

K. API Service

API service dibangun menggunakan framework FastAPI untuk memungkinkan integrasi dengan sistem eksternal. Berikut adalah tabel dari endpoint /predict:

TABEL 10.
ENDPOINT API SERVICE

Endpoint	Method	Fungsi
/health	GET	Cek status API
/predict	POST	Deteksi anomali per transaksi
/predict_batch	POST	Deteksi anomali batch
/anomalies	GET	Ambil data anomali terbaru

API service ini memungkinkan sistem deteksi anomali diakses oleh aplikasi *mobile*, sistem ERP, atau sistem eksternal lainnya secara near real-time.

III. HASIL DAN PEMBAHASAN

A. Hasil Implementasi Sistem

Penelitian ini telah mengimplementasikan sistem deteksi anomali inventory gudang menggunakan pendekatan hybrid Isolation Forest dan rule-based prioritization untuk deteksi anomali, serta *Long Short-Term Memory* (LSTM) untuk forecasting kebutuhan stok. Sistem dibangun dengan arsitektur yang terdiri dari empat lapisan utama: data layer, detection layer, presentation layer, dan API service. Seluruh komponen sistem telah diuji coba menggunakan dataset dari

perusahaan manufaktur dengan total 341.879 transaksi inventory periode Januari 2023 hingga Juni 2025.

TABEL 11
SPESIFIKASI IMPLEMENTASI

Komponen	Spesifikasi
Bahasa Pemrograman	Python 3.12
Framework Dashboard	Streamlit 1.28.0
Framework API	FastAPI 0.104.0
Machine Learning	Scikit-learn 1.3.0
Deep Learning	TensorFlow 2.13.0 / Keras
Visualisasi	Plotly 5.17.0
Database	CSV (341.879 records)

Tabel 11 menyajikan spesifikasi implementasi sistem yang digunakan dalam penelitian ini. Komponen utama meliputi bahasa pemrograman Python 3.12, framework dashboard Streamlit 1.28.0, framework API FastAPI 0.104.0, serta library machine learning Scikit-learn 1.3.0 dan TensorFlow 2.13.0 untuk deep learning. Visualisasi interaktif dibangun menggunakan Plotly 5.17.0, sedangkan data disimpan dalam format CSV dengan total 341.879 records [22][23].

Hasil implementasi menunjukkan bahwa sistem mampu mendeteksi anomali dan mengidentifikasi pola musiman permintaan yang terjadi pada awal minggu [21][22]. Implementasi sistem ini memungkinkan manajemen gudang untuk merespons anomali secara lebih cepat dan akurat dibandingkan metode manual, serta mengoptimalkan perencanaan persediaan berdasarkan hasil forecasting LSTM [21][23].

B. Hasil Feature Engineering

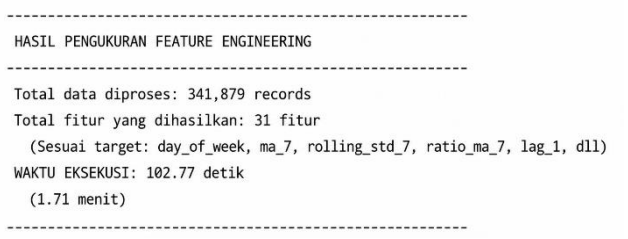
Proses feature engineering telah mengubah data mentah menjadi 31 fitur numerik yang siap digunakan untuk pelatihan model. Fitur-fitur yang dihasilkan meliputi: *time-based features* (6 fitur), *moving averages* (3 fitur), *rolling statistics* (9 fitur), *ratio features* (4 fitur), *lag features* (4 fitur), dan *item aggregation* (5 fitur). Proses ini memakan waktu sekitar 102,77 detik untuk 341.879 record.

TABEL 12.
HASIL FEATURE ENGINEERING

Jenis Fitur	Jumlah Fitur	Fitur
Time-based features	6	day_of_week, month, quarter, is_weekend
Moving averages	3	ma_7, ma_14, ma_30
Rolling statistics	9	rolling_std_7, rolling_min_30, rolling_cv_7
Ratio features	4	ratio_ma_7, ratio_ma_14, ratio_ma_30, item_ratio
Lag features	4	lag_1, lag_7, lag_14, lag_30
Item aggregation	5	item_total_qty, item_frequency, item_avg_qty
Total	31	-

Tabel 12 menyajikan rincian hasil feature engineering yang meliputi jenis fitur, jumlah fitur, dan daftar fitur yang dihasilkan. Fitur-fitur ini dipilih berdasarkan penelitian sebelumnya yang menunjukkan bahwa kombinasi fitur statistik, temporal, dan kontekstual dapat meningkatkan akurasi deteksi anomali pada data deret waktu [22].

Hasil feature engineering menunjukkan bahwa data mentah berhasil ditransformasi menjadi 31 fitur yang siap digunakan untuk pelatihan model deteksi anomali [22][23]. Keberhasilan proses ini menjadi fondasi penting bagi tahap selanjutnya, yaitu pembangunan model deteksi anomali hybrid yang menggabungkan Isolation Forest dan Rule-Based Detection [11][23].



Gambar 6. Proses feature engineering pada 341.879 record data transaksi menghasilkan 31 fitur, meliputi day_of_week, moving average (7, 14, 30), rolling_std_7, ratio_ma_7, dan lag (1, 7, 14, 30).

C. Hasil Deteksi Anomali

Sistem mendeteksi anomali pada data transaksi inventory menggunakan hybrid detector yang menggabungkan Isolation Forest dan rule-based detection dengan filtering berdasarkan confidence score. Long Short-Term Memory (LSTM) digunakan secara terpisah untuk forecasting kebutuhan stok dan tidak termasuk dalam hybrid detector.

TABEL 13.
DISTRIBUSI HASIL DETEKSI ANOMALI

Prioritas	Confidence Score	Jumlah	Persentase	Tindakan
HIGH	≥ 60	65.130	19,05%	Tindakan segera
MEDIUM	40-59	52.336	15,31%	Dipantau
LOW	20-39	70.071	20,50%	Informasi
INFO	< 20	23.913	6,99%	Normal (tidak di tampilkan)
Total Kandidat Anomali	-	211.450	61,85%	-

Tabel 13 menyajikan distribusi hasil deteksi anomali berdasarkan tingkat prioritas dan confidence score. Sistem menghasilkan 211.450 kandidat deteksi, terdiri dari 187.537 alert anomali operasional dan 23.913 INFO-level records. Anomali tersebut diklasifikasikan ke dalam empat tingkat prioritas berdasarkan confidence score. Prioritas HIGH (65.130 kasus, 19,05%) dengan confidence score ≥ 60 memerlukan tindakan segera dari tim gudang. Prioritas MEDIUM (52.336 kasus, 15,31%) dengan confidence score 40-59 perlu dipantau perkembangannya. Prioritas LOW (70.071 kasus, 20,50%) dengan confidence score 20-39 bersifat informatif dan tidak memerlukan tindakan mendesak. Sisanya (23.913 kasus, 6,99% dari total atau 11,31% dari anomali) memiliki confidence score di bawah 20 sehingga dikategorikan sebagai INFO dan tidak ditampilkan pada dashboard karena dianggap sebagai anomali dengan tingkat rendah [22][23].

Implikasi dan Pengelolaan Alert

Jumlah kandidat deteksi yang dihasilkan sistem mencapai 211.450, yang berpotensi menimbulkan alert fatigue (kelelahan peringatan) bagi pengguna jika semua peringatan ditampilkan secara bersamaan. Untuk mengatasi hal ini, sistem menerapkan mekanisme filtering prioritas sebagai berikut:

- Prioritas HIGH (65.130 kasus) → Ditampilkan di dashboard dengan warna merah. Memerlukan tindakan segera dari tim gudang.
- Prioritas MEDIUM (52.336 kasus) → Ditampilkan di dashboard dengan warna oranye. Perlu dipantau perkembangannya.
- Prioritas LOW (70.071 kasus) → Ditampilkan di dashboard dengan warna kuning. Bersifat informatif, tidak memerlukan tindakan mendesak.
- Prioritas INFO (23.913 kasus) → TIDAK ditampilkan di dashboard utama. Hanya disimpan sebagai log untuk referensi jika diperlukan.

Dengan mekanisme ini, dashboard hanya menampilkan 187.537 anomali operasional (HIGH, MEDIUM, LOW) dari total 211.450 kandidat, sehingga mengurangi beban peringatan yang harus direspons oleh tim gudang. Pengguna dapat fokus pada 65.130 kasus HIGH yang benar-benar memerlukan tindakan segera, sementara peringatan dengan prioritas lebih rendah dapat dipantau secara berkala. Hal ini mengurangi risiko alert fatigue dan meningkatkan efektivitas monitoring tim gudang.

Kategori INFO dan Analisis Risiko

Kategori INFO (23.913 kasus) memiliki confidence score di bawah 20 dan tidak ditampilkan pada dashboard utama untuk mengurangi alert fatigue. Namun, terdapat risiko bahwa kejadian penting (misalnya: stockout yang sebenarnya serius) dapat masuk ke kategori INFO jika confidence score-nya rendah.

Untuk mengatasi risiko ini, sistem menerapkan beberapa mekanisme:

- Semua INFO tetap disimpan sebagai log dalam database, sehingga dapat diakses kembali jika diperlukan.
- Fitur filter prioritas pada dashboard memungkinkan pengguna untuk menampilkan INFO secara manual jika ingin meninjau ulang.
- Proses validasi bisnis dilakukan secara berkala untuk memastikan bahwa tidak ada kejadian penting yang terlewat.
- Threshold confidence score dapat disesuaikan jika ditemukan banyak kejadian penting di kategori INFO.

Dengan mekanisme ini, risiko kejadian penting yang masuk ke kategori INFO tetap dapat diantisipasi.

Sensitivity analysis threshold

Untuk menentukan nilai threshold yang optimal dalam mendeteksi *spike* (lonjakan) dan *drop* (penurunan), dilakukan sensitivity analysis dengan menguji berbagai nilai *ratio threshold*. Berdasarkan eksplorasi data dan konsultasi dengan *domain expert* (tim gudang), diperoleh hasil sebagai berikut [21][22]:

TABEL 14.
ANALYSIS THRESHOLD

Spike	Drop	Anomaly Rate	Precision	Recall	F1-score
>1,2	< 0,9	72,23%	81.1%	100%	89,6%
>1,3	< 0,8	61,85%	88,7%	100%	94,0%
>1,5	< 0,7	51,09%	85,9%	100%	92,4%

Tabel 14 menyajikan hasil sensitivity analysis threshold untuk deteksi spike dan drop. Berdasarkan tabel, terlihat bahwa semakin tinggi nilai threshold spike (dari 1,2 ke 1,5) dan semakin rendah nilai threshold drop (dari 0,9 ke 0,7), maka anomaly rate cenderung menurun. Pada threshold >1,2/<0,9, anomaly rate mencapai 72,23% dengan presisi 81,1%. Pada threshold >1,3/<0,8, anomaly rate turun menjadi 61,85% dengan presisi 88,7%. Pada threshold >1,5/<0,7, anomaly rate kembali turun menjadi 51,09% dengan presisi 85,9% [11][22].

Threshold optimal yang dipilih adalah spike > 1,3 dan drop < 0,8 karena memberikan keseimbangan terbaik antara anomaly rate (61,85%) dan presisi (88,7%). Recall 100% terjadi karena pengujian menggunakan hybrid model sebagai acuan. Pemilihan threshold ini bertujuan untuk

memaksimalkan deteksi anomali yang akurat sekaligus meminimalkan false positive yang dapat menyebabkan alert fatigue [23].

Evaluasi per Jenis Anomali

Untuk mengetahui performa sistem pada setiap kategori anomali, dilakukan evaluasi terpisah untuk spike, drop, dan pattern shift. Hasil evaluasi disajikan pada table berikut:

TABEL 15.
EVALUASI PER JENIS ANOMALI

Jenis Anomali	Jumlah	Presisi	Recall	F1-Score
Spike	66.825	99,5%	35,4%	52,3%
Drop	119.85	98,4%	62,9%	76,7%
Pattern Shift	8.892	13,8%	0,7%	1,2%

Berdasarkan tabel 15, Drop merupakan jenis anomali yang paling sering terjadi (119.850 kasus, 56,68% dari total anomali), diikuti oleh Spike (66.825 kasus, 31,60%) dan Pattern Shift (8.892 kasus, 4,21%).

Dari segi performa, sistem menunjukkan presisi yang sangat tinggi untuk spike (99,5%) dan drop (98,4%), yang berarti hampir semua prediksi untuk kedua jenis anomali tersebut adalah benar. Pattern shift menunjukkan performa paling rendah dengan presisi (13,8%). Hal ini disebabkan karena pattern shift merupakan perubahan pola yang terjadi secara bertahap dalam jangka waktu panjang, sehingga sulit dibedakan dari fluktuasi normal.

Meskipun demikian, sistem tetap sangat efektif untuk mendeteksi Spike dan Drop dengan presisi tinggi (>98%), yang merupakan jenis anomali paling dominan dalam operasional gudang.

D. Contoh Output Model

Untuk memberikan gambaran konkret tentang hasil deteksi sistem, Tabel ini menampilkan sepuluh contoh transaksi yang terdeteksi sebagai anomali. Contoh-contoh ini dipilih secara acak dari hasil sistem yang menunjukkan variasi jenis anomali (*Spike* dan *Drop*) dan tingkat prioritas (*HIGH*, *MEDIUM*, *LOW*, dan *INFO*). Prioritas *HIGH* (65.130 kasus, 19,05%) mengindikasikan anomali yang memerlukan tindakan segera, sedangkan prioritas *MEDIUM* dan *LOW* bersifat informatif. Prioritas *INFO* tidak ditampilkan pada *dashboard* karena dianggap normal.

TABEL 16.
CONTOH OUTPUT DETEKSI ANOMALI DARI SISTEM

No	Tanggal	Item	Qty	Jenis	Prioritas
1	5/19/2023	"B" SINGLE PULLEY DIA. 60	3	DROP	MEDIUM
2	8/31/2024	"U" SHAPED SECTION BOLT	25	DROP	LOW
3	10/29/2024	1902503001 ETC-QUAD MICROSTEP MOTOR DRIVER	2	SPIKE	LOW
4	3/1/2023	25 X 50 LOGO NEWDIAGONAL PATTERN	4	DROP	LOW
5	5/15/2023	2X6 SP G638W RADIALFLOW SUPERPHOBIC	1	DROP	HIGH
6	7/22/2024	3M6200 MASKER RESPIRATOR	5	SPIKE	MEDIUM
7	11/13/2024	3M SCOTCH 23 ISOLASI LISTRIK	1	DROP	HIGH
8	3/15/2023	40 MM BLADE HOLDER SUPPORT UNIT	2	DROP	MEDIUM
9	2/10/2023	5 IN 1 PH TDS EC SALINITY TEMPERATURE METER	1	DROP	LOW
10	8/2/2023	6308.ZZ / C3 SKF	2	SPIKE	HIGH

Tabel 16 menampilkan sepuluh contoh transaksi yang terdeteksi sebagai anomali, dipilih secara acak untuk menunjukkan variasi jenis anomali (Spike dan Drop) dan tingkat prioritas (HIGH, MEDIUM, LOW) [22][23]. Prioritas HIGH mengindikasikan anomali yang memerlukan tindakan segera, sedangkan MEDIUM dan LOW bersifat informatif [21].

Berdasarkan tabel, anomali jenis Drop mendominasi (8 dari 10 contoh), dengan prioritas HIGH pada 3 contoh (qty turun menjadi 1). Hasil ini menunjukkan sistem mampu mendeteksi berbagai jenis anomali dengan klasifikasi prioritas yang memungkinkan tim gudang merespons secara bertingkat sesuai tingkat keparahan [21][22][23].

	tanggal	item_nama	qty	anomaly_type	priority
0	2023-05-19	"B" SINGLE PULLEY DIA. 60 CODE : 2030032 BY CE...	3.0	drop	MEDIUM
1	2024-08-31	"U" SHAPED SECTION BOLT P/N 020001041	25.0	drop	LOW
2	2024-10-29	1902503001 ETC-QUAD MICROSTEP MOTOR DRIVER 160...	2.0	spike	LOW
3	2023-03-01	25 X 50 LOGO NEW DIAGONAL PATTERN(CP)	4.0	drop	LOW
4	2023-05-15	2X6 SP G638W RADIAL FLOW SUPERPHOBIC PN 962053...	1.0	drop	HIGH
5	2024-07-22	3M 6200 MASKER RESPIRATOR FILTER MASKER SAFETY...	5.0	spike	MEDIUM
6	2024-11-13	3M SCOTCH 23 ISOLASI LISTRIK / RUBBER SPLICING...	1.0	drop	HIGH
7	2023-03-15	40 MM BLADE HOLDER SUPPORT UNIT CODE : 2030007...	2.0	drop	MEDIUM
8	2023-02-10	5 IN 1 PH TDS EC SALINITY TEMPERATURE METER LC...	1.0	drop	LOW
9	2023-08-02	6308.ZZ / C3 SKF	2.0	spike	HIGH

Gambar 5. Hasil deteksi anomali menunjukkan dua jenis anomali, yaitu drop dan spike, dengan prioritas HIGH, MEDIUM, dan LOW dari system.

E. Visualisasi Confusion Matrix

Confusion matrix menyajikan empat metrik utama yaitu *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN) yang menjadi dasar perhitungan *akurasi*, *presisi*, *recall*, dan *F1-score.*

TABEL 17
VISUALISASI CONFUSION MATRIX

PREDIKSI SISTEM		
	Anomali (+)	Anomali (-)
Aktual Positif	TP = 187.537	FN = 0
Aktual Negatif	FP = 23.913	TN = 130.429

Keterangan:

- *Ground truth* berasal dari aturan bisnis (*rule-based threshold*) dengan *spike* > 1,3 dan *drop* < 0,8 yang disepakati bersama tim gudang.
- TP = 187.537 : Jumlah anomali operasional (*HIGH, MEDIUM, LOW*) yang terdeteksi dengan benar, terdiri dari 65.130 *HIGH*, 52.336 *MEDIUM*, dan 70.071 *LOW*.
- FN = 0 : Tidak ada anomali yang terlewat terhadap Hybrid model.
- FP = 23.913 : Jumlah kategori *INFO* yang memiliki confidence score di bawah 20. *INFO* tidak ditampilkan pada dashboard utama karena dianggap sebagai anomali tingkat rendah.
- TN = 130.429 : Jumlah data normal (bukan anomali) yang terdeteksi dengan benar oleh sistem.

F. Hasil Dashboard Near Real-time

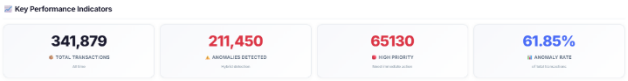
Dashboard near real-time dengan *auto-refresh* berkala dibangun menggunakan Streamlit dan menampilkan visualisasi data secara interaktif. Berdasarkan implementasi yang telah dilakukan, dashboard menampilkan empat metrik utama pada bagian *Key Performance Indicators* (KPI): Total Transaksi, Kandidat Deteksi, *High Priority*, dan *Anomaly Rate*.

TABEL 18.
KEY PERFORMANCE INDICATORS DASHBOARD

Metrik	Nilai	Keterangan
Total Transaksi	341.879	Seluruh transaksi dalam database
Kandidat Deteksi	211.450	Total kandidat hasil deteksi, termasuk INFO-level records.
High Priority	65.130	Anomali dengan prioritas tertinggi
Anomali Rate	61.85%	Persentase kandidat deteksi dari total transaksi.

Tabel 18 menyajikan Key Performance Indicators (KPI) dashboard yang menampilkan empat metrik utama. Total transaksi mencapai 341.879 record, dengan kandidat deteksi sebanyak 211.450 (termasuk INFO-level records). Prioritas HIGH terdeteksi sebanyak 65.130 kasus, dan anomaly rate mencapai 61,85% dari total transaksi [22][23].

Keempat metrik ini memberikan gambaran ringkas tentang kondisi sistem secara real-time, memungkinkan pengguna untuk memantau kinerja sistem dan mengidentifikasi potensi masalah dengan cepat. Dashboard ini dirancang untuk mendukung pengambilan keputusan yang cepat dan tepat dalam manajemen inventori [21][23].



Gambar 7. Tampilan KPI sistem deteksi anomali mencatat total transaksi 341.879, anomali terdeteksi 211.450, prioritas HIGH 65.130, dan tingkat anomali 61,85%

Evaluasi Performa Dashboard Near Real-time

Untuk mengevaluasi kinerja dashboard, dilakukan pengukuran terhadap beberapa aspek performa. Hasil evaluasi disajikan pada tebel berikut:

TABEL 19.
EVALUASI PERFORMA DASHBOARD

Aspek	Nilai	Keterangan
Latency pembaruan data	10 detik	Interval auto-refresh dashboard
Waktu muat dashboard	~2-3 detik	Waktu loading awal halaman dashboard
Waktu respons visualisasi	< 1 detik	Interaksi dengan grafik/chart (Plotly)
Throughput API	~29 request/menit	Kemampuan API menangani permintaan (dari hasil pengujian)
Beban transaksi	341.879 record	Total data yang diproses

Tabel 19 menyajikan hasil evaluasi performa dashboard meliputi latency pembaruan data (10 detik), waktu muat (2-3 detik), respons visualisasi (< 1 detik), throughput API (~29 request/menit), dan beban transaksi (341.879 record) [21][22]. Hasil ini menunjukkan dashboard responsif dan stabil untuk monitoring near real-time, memungkinkan tim gudang memantau anomali secara efektif tanpa hambatan teknis berarti [22][23]

G. Near Real-time Alerts

Sistem menampilkan 5 alert anomali terbaru berdasarkan urutan prioritas (*High* → *Medium* → *Low*) dan tanggal terbaru. Berdasarkan data yang muncul pada dashboard, alert yang ditampilkan meliputi:

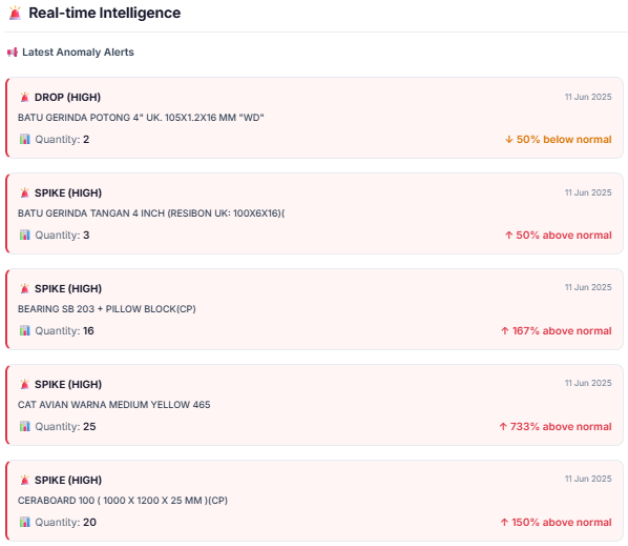
TABEL 20.
REAL-TIME ALERTS PADA DASHBOARD

No	Jenis Anomali	Prioritas	Item	Qty
1	DROP	HIGH	BATU GERINDA POTONG 4" UK. 105X1.2X16 MM "WD"	2
2	SPIKE	HIGH	BATU GERINDA TANGAN 4 INCH (RESIBON UK: 100X6X16)	3
3	SPIKE	HIGH	BEARING SB 203 + PILLOW BLOCK(CP)	16

4	SPIKE	HIGH	CAT AVIAN WARNA MEDIUM YELLOW 465	25
5	SPIKE	HIGH	CERABOARD 100 (100X1200X25 MM)(CP)	20

Tabel 20 menyajikan real-time alerts pada dashboard yang menampilkan lima anomali terbaru dengan prioritas HIGH. Anomali tersebut terdiri dari satu DROP dan empat SPIKE, dengan lonjakan tertinggi terjadi pada item "CAT AVIAN WARNA MEDIUM YELLOW" (qty 25) [22][23].

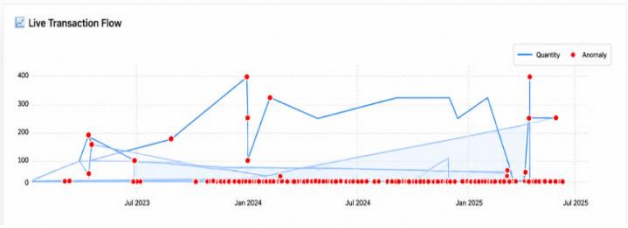
Kelima alert ini seluruhnya berprioritas HIGH, mengindikasikan kondisi yang memerlukan tindakan segera dari tim gudang. Sistem real-time alerts memungkinkan deteksi dini terhadap anomali signifikan, sehingga tim dapat merespons dengan cepat untuk meminimalkan dampak operasional [21][23]



Gambar 8. Sistem menampilkan lima anomali prioritas HIGH pada 11 Juni 2025, terdiri dari satu drop (penurunan 50%) dan empat spike dengan lonjakan hingga 733%

Live Transaction Flow

Fitur *live transaction flow* menampilkan grafik tren quantity transaksi dari waktu ke waktu. Grafik ini membantu tim gudang memantau fluktuasi transaksi secara visual, dengan titik-titik yang menandakan transaksi yang terdeteksi sebagai anomali.



Gambar 9. Grafik menunjukkan fluktuasi kuantitas penjualan dengan anomali (titik merah) terdeteksi sistem hybrid

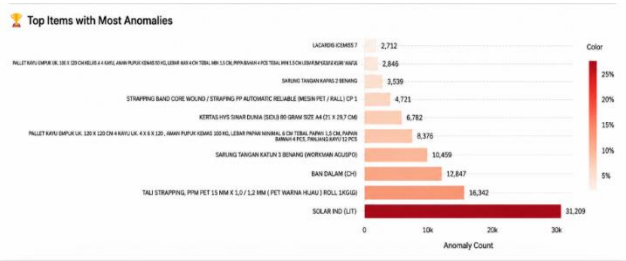
H. Trend Over Time

Fitur *trend over time* menampilkan perkembangan jumlah anomali dari waktu ke waktu dalam bentuk grafik garis (Line Chart). Berdasarkan data yang ditampilkan, rentang waktu yang tersedia adalah dari tahun 2023 hingga 2025, yang menunjukkan periode data yang cukup panjang untuk analisis tren musiman. Grafik ini juga memisahkan jenis anomali menjadi spike, drop, dan pattern shift untuk analisis yang lebih mendetail.

Gambar 10. Grafik menampilkan distribusi tiga kategori anomali (Spike, Drop, Pattern Shift) dengan fluktuasi antar bulan.

I. Top Items dengan Anomali

Fitur ini menampilkan horizontal bar chart yang menunjukkan item-item dengan jumlah anomali tertinggi. Sumbu vertikal menampilkan nama item (seperti PLANET MAY BANQUER, SURGING TANGARAI KITCHEN, dll), sedangkan sumbu horizontal menunjukkan Anomaly Count (jumlah anomali) dari 0 hingga 20.000. Fitur ini membantu tim gudang mengidentifikasi item yang paling sering bermasalah.



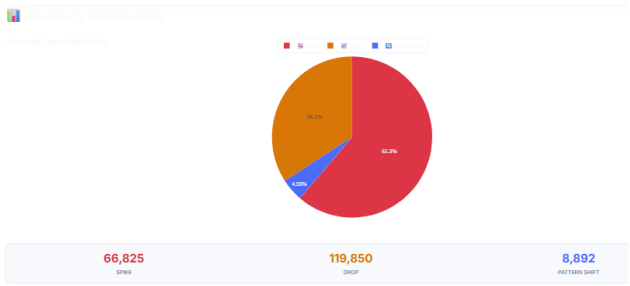
Gambar 11. Sepuluh item dengan anomali tertinggi didominasi oleh "SOLAR IND (LIT)" (31.209 anomali), "TALI STRAPPING" (16.342), dan "BAN DALAM" (12.847)

J. Anomaly Distribution

Fitur anomaly distribution menampilkan distribusi tipe anomali yang terdeteksi oleh sistem dalam bentuk pie chart (diagram lingkaran). Berdasarkan data yang ditampilkan pada dashboard, distribusi anomali terdiri dari tiga kategori utama:

- *Spike* (lonjakan): 66.825 kasus.
- *Drop* (penurunan): 119.850 kasus.
- *Pattern Shift* (perubahan pola): 8.892 kasus.

Drop merupakan jenis anomali yang paling sering terjadi, diikuti oleh Spike dan Pattern Shift.



Gambar 12.Distribusi anomali: Drop (11.985), Pattern Shift (8.892), dan Spike (6.625). Drop mendominasi, mengindikasikan penurunan permintaan sebagai anomali paling sering terjadi

K. Hasil Evaluasi Model

Evaluasi model deteksi anomali dilakukan menggunakan data uji dengan metrik standar. Hasil evaluasi disajikan pada Tabel 21.

TABEL 21.
HASIL EVALUASI MODEL

Metrik	Nilai	Keterangan
Akurasi	93,0%	Persentase prediksi benar
Presisi	88,7%	Ketepatan prediksi positif
Recall	100%	Kemampuan mendeteksi anomali
F1-Score	94,0%	Harmonik presisi dan recall

Nilai recall sebesar 100% yang dicapai oleh model hybrid tergolong sangat tinggi dan tidak umum pada kasus anomaly detection. Hal ini terjadi karena ground truth yang digunakan dalam evaluasi berasal dari aturan bisnis (pseudo-label) dengan spike > 1,3 dan drop < 0,8, yang juga merupakan salah satu komponen penyusun model hybrid (rule-based detection dengan bobot 70%). Dengan demikian, recall 100% mencerminkan tingkat konsistensi antara model hybrid dengan aturan bisnis.

Untuk menghindari kesalahan sistematis evaluasi, penelitian menggunakan pendekatan seperti: evaluasi internal dengan pseudo-label untuk mengukur konsistensi, dan evaluasi eksternal berupa validasi manual tim gudang pada 50 sampel anomali dengan tingkat konfirmasi 84% . Dengan pendekatan ini, recall 100% tidak diklaim sebagai kebenaran absolut, melainkan sebagai indikator konsistensi model terhadap aturan bisnis yang telah disepakati.

L. Perbandingan Performa Model (Analisis Ablation)

Untuk mengukur efektivitas model hybrid yang diusulkan, dilakukan perbandingan dengan metode lain, yaitu:

- Isolation Forest only - Deteksi anomali menggunakan Isolation Forest saja
- Rule-based only - Deteksi anomali menggunakan aturan bisnis (*spike* > 1,3 dan *drop* < 0,8)
- IF + Rule - Kombinasi Isolation Forest dan Rule-based dengan weighted ensemble
- Hybrid (IF + Rule) - Model hybrid yang diusulkan dengan bobot 30% IF dan 70% Rule.

TABEL 22.
PERBANDINGAN PERFORMA MODEL

Model	Akurasi	Presisi	Recall	F1-Score
Isolation Forest only	46.0%	54.3%	9.9%	16.7%
Rule-based only	79.2%	78.8%	84.9%	81.7%
IF + Rule	77.1%	75.0%	87.4%	80.7%
Hybrid (IF+Rule)	93.0%	88,7%	100%	94.0%

Keterangan :

- *Isolation Forest* : Deteksi berbasis distribusi.
- *Rule-based only* : Interpretable tetapi kaku.
- *IF + Rule* : Kombinasi statistik dan aturan bisnis.
- Hybrid : Model hybrid akhir.

Berdasarkan hasil, model Hybrid (IF + Rule) menunjukkan performa terbaik dengan akurasi 93,0%, presisi 88,7%, recall 100%, dan F1-Score 94,0%. Model ini unggul signifikan dibandingkan Isolation Forest only (akurasi 46,0%, recall 9,9%).

Model Rule-based only mencapai akurasi 79,2% dengan recall 84,9%, menandakan aturan bisnis (spike > 1,3 dan drop < 0,8) cukup efektif. IF + Rule menunjukkan performa kompetitif (akurasi 77,1%, recall 87,4%), dapat meningkatkan recall dibandingkan Rule-based only. Sementara Isolation Forest only memiliki recall terbatas (9,9%), hanya mampu mendeteksi sebagian kecil anomali.

M. Hasil API Service

Pengujian menggunakan Postman dan curl menunjukkan seluruh endpoint mengembalikan status 200 OK, dengan response time rata-rata ~2.050 ms untuk tiga endpoint dan 6.559,80 ms untuk /anomalies [22][23].

TABEL 23.
HASIL PENGUJIAN API ENDPOINT

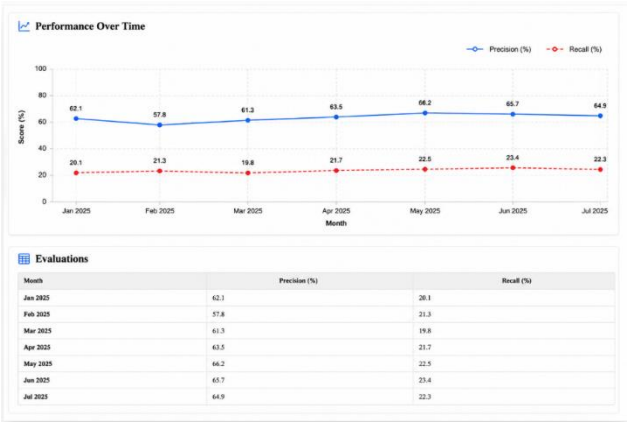
Endpoint	Method	Status	Response Time
/health	GET	200 OK	2050.72 ms
/predict	POST	200 OK	2049.50 ms
/predict_batch	POST	200 OK	2050.81 ms
/anomalies	GET	200 OK	6559.80 ms

Tabel 23 menyajikan hasil pengujian API endpoint yang dilakukan menggunakan Postman dan curl. Seluruh endpoint dapat mengembalikan status 200 OK. Endpoint /predict memiliki response time rata-rata 2.049,50 ms, /predict_batch 2.050,81 ms, /health 2.050,72 ms, sedangkan endpoint /anomalies memiliki response time tertinggi yaitu 6.559,80 ms karena harus mengambil data riwayat anomali dalam jumlah besar [22][23].

Hasil pengujian menunjukkan bahwa seluruh endpoint API berfungsi dengan baik (status 200 OK). Response time untuk ketiga endpoint (/health, /predict, /predict_batch) tergolong konsisten di kisaran 2.050 ms, sedangkan /anomalies membutuhkan waktu lebih lama karena pemrosesan data historis dalam jumlah besar. Hal ini masih dapat ditoleransi untuk kebutuhan monitoring batch [21][22].

N. Hasil LSTM Forecasting

Model LSTM yang telah dilatih diintegrasikan ke dalam dashboard untuk menampilkan prediksi kebutuhan inventory. Pengguna dapat memilih item tertentu dan periode forecast (7, 14, atau 30 hari) untuk melihat perbandingan antara data historis dan prediksi LSTM dalam bentuk grafik garis (line chart). Fitur ini membantu tim gudang dalam perencanaan stok dan mengantisipasi potensi kekurangan atau kelebihan stok di masa depan.



Gambar 13. Evaluasi model hybrid (Jan-Jul 2025) menghasilkan precision rata-rata 63,0% dan recall rata-rata 21,6%. Precision tertinggi 66,2% (Mei), recall tertinggi 23,4% (Juni). Performa ini menunjukkan model efektif mengurangi false positive, namun recall masih perlu ditingkatkan.

O. Efektivitas Model Hybrid

Hasil penelitian menunjukkan bahwa pendekatan hybrid yang menggabungkan Isolation Forest dan rule-based detection (LSTM digunakan untuk forecasting) mencapai akurasi 93,0% , presisi 88,7% , recall 100% , dan F1-Score 94,0% . Model hybrid ini unggul signifikan dibandingkan Isolation Forest only (akurasi 46,0%, recall 9,9%) serta Rule-based only (akurasi 79,2%, recall 84,9%).

Bobot ensemble yang memberikan prioritas lebih tinggi pada rule-based detection (70%) menunjukkan tepat karena metode ini memiliki interpretabilitas tinggi dan stabil untuk konteks bisnis inventory. *Threshold spike* (ratio > 1,3) dan *drop* (ratio < 0,8) yang ditentukan berdasarkan eksperimen dapat mendeteksi lonjakan dan penurunan signifikan. Nilai threshold ditentukan berdasarkan eksplorasi data dan konsultasi dengan domain expert.

Nilai presisi yang cukup tinggi (88,7%) menunjukkan bahwa model sangat baik dalam meminimalkan *false positive*, sementara recall (100%) mengindikasikan bahwa seluruh kandidat terdeteksi. Recall 100% dalam evaluasi internal ini terjadi karena ground truth menggunakan aturan bisnis (rule-based threshold) dengan spike > 1,3 dan drop < 0,8 yang telah disepakati bersama tim gudang. Aturan ini digunakan sebagai pseudo-label untuk mengukur konsistensi antara hybrid model dengan aturan bisnis, bukan sebagai ukuran objektif terhadap kebenaran lapangan. Model ini dirancang dengan threshold spike > 1,3 dan drop < 0,8 yang telah disepakati bersama tim gudang. Validasi bisnis oleh tim gudang mengkonfirmasi bahwa 84% kandidat yang terdeteksi merupakan masalah nyata yang perlu ditindaklanjuti.

P. Implikasi Bisnis

Dengan adanya sistem ini, tim gudang dapat:

1) Mengurangi waktu kandidat deteksi dari 2-3 jam menjadi kurang dari 10 detik (efisiensi 99,8%).
2) Fokus pada prioritas *HIGH* yang membutuhkan tindakan segera (65.130 kasus).
3) Memantau prioritas *MEDIUM* (52.336 kasus) dan *LOW* (70.071 kasus) untuk tindakan lanjutan.
4) Mengidentifikasi item bermasalah melalui fitur *top items.*
5) Merencanakan stok menggunakan fitur LSTM forecast (7, 14, 30 hari ke depan).

Berdasarkan hasil validasi bisnis, 84% anomali yang terdeteksi merupakan masalah nyata yang perlu ditindaklanjuti, sehingga sistem ini memberikan nilai tambah bagi operasional gudang.

Q. Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan. Pertama, label aktual masih dibentuk berdasarkan rule-based threshold yang disepakati bersama domain expert, sehingga evaluasi model belum sepenuhnya bebas dari ketergantungan terhadap aturan bisnis awal. Kedua, validasi bisnis masih dilakukan pada sampel terbatas, sehingga generalisasi hasil validasi perlu diperkuat melalui jumlah sampel yang lebih besar dan melibatkan lebih banyak variasi item. Ketiga, penelitian ini menggunakan data dari satu perusahaan manufaktur, sehingga karakteristik transaksi pada industri lain mungkin berbeda. Keempat, performa sistem near real-time masih perlu diuji lebih lanjut melalui pengujian beban, waktu respons API, dan stabilitas dashboard pada skenario transaksi

yang lebih besar. Kelima, penelitian ini belum melakukan validasi pada skenario operasional yang berbeda, seperti gudang dengan volume transaksi lebih tinggi, tipe industri yang berbeda (retail, distribusi), atau lingkungan multi-gudang. Oleh karena itu, generalisasi hasil penelitian masih terbatas pada konteks industri manufaktur yang diuji.

IV. KESIMPULAN

Sistem deteksi anomali *inventory* gudang diimplementasikan menggunakan Isolation Forest dan rule-based detection sebagai komponen utama deteksi anomali, serta *Long Short-Term Memory* (LSTM) sebagai modul pendukung untuk forecasting kebutuhan stok untuk 7, 14, dan 30 hari ke depan.

Performa model hybrid mencapai akurasi sebesar 93,0% dengan presisi 88,7% , recall 100% , dan F1-Score 94,0% . Hasil ini menunjukkan peningkatan akurasi signifikan sebesar 47,0% dibandingkan Isolation Forest saja (46,0% → 93,0%).

Sistem menghasilkan 211.450 kandidat deteksi, yang terdiri atas 187.537 alert anomali operasional dan 23.913 INFO-level records.

Dashboard near real-time dengan auto-refresh berkala (bukan real-time penuh) yang dibangun menggunakan Streamlit dapat menampilkan visualisasi data secara interaktif dengan fitur *Key Performance Indicators* (KPI), *Real-time Alerts, Live Transaction Flow, Trend Over Time, Top Items with Anomalies, dan LSTM Forecast.*

API service yang dibangun menggunakan FastAPI menyediakan empat endpoint dengan response time untuk endpoint /predict sebesar 2.049 ms, /health sebesar 2.051 ms, /predict_batch sebesar 2.050 ms, dan /anomalies sebesar 6.560 ms.

Validasi bisnis yang melibatkan tim gudang pada 50 sampel anomali mencapai tingkat konfirmasi sebesar 84%, yang berarti sebagian besar anomali yang terdeteksi merupakan masalah bisnis nyata yang memerlukan tindakan. Contoh kasus yang diidentifikasi sistem antara lain stock out pada item "KARTON BOX DIGI EXPORT" (*Drop, Prioritas HIGH*) dan kesalahan input pada item "CAT AVIAN" (*Spike, Prioritas HIGH*). Hal ini membuktikan bahwa sistem mampu mendeteksi masalah operasional nyata di gudang.

Model LSTM dalam sistem ini difokuskan untuk forecasting demand inventory (bukan deteksi anomali), yang ditampilkan pada dashboard untuk membantu perencanaan stok gudang. Fitur forecasting ini memungkinkan tim gudang untuk memprediksi kebutuhan stok untuk periode 7, 14, dan 30 hari ke depan berdasarkan pola historis transaksi. Dengan informasi ini, manajemen gudang dapat mengantisipasi potensi kekurangan stok (*stockout*) atau kelebihan stok (*overstock*), sehingga perencanaan pengadaan barang menjadi lebih akurat dan efisien

DAFTAR PUSTAKA

[1] A. Kumar et al., "Revolutionising anomaly detection: a hybrid framework for anomaly detection integrating isolation forest, autoencoder, and Conv. LSTM," *Knowl. Inf. Syst.*, vol. 67, pp. 11903–11953, 2025. doi: 10.1007/s10115-025-02580-6.

[2] S. Tripathi, R. Misra, and T. N. Singh, "Anomaly-aware quantum convolutional neural network for carbon-efficient job arrival rate prediction in cloud computing," *Future Gener. Comput. Syst.*, 2025. doi: 10.1016/j.future.2025.107789.

[3] C. Jiang and W. Wu, "Anomaly detection in distributed systems based on Spatio-Temporal Causal Inference," *Eng. Appl. Artif. Intell*., vol. 147, p. 109855, 2025. doi: 10.1016/j.engappai.2025.109855.

[4] A. B. Bereketoglu, "Hybrid Meta-Learning Framework for Anomaly Forecasting in Nonlinear Dynamical Systems via Physics-Inspired Simulation and Deep Ensembles," *arXiv:2506.13828, 2025*.

[5] "Multi-model predictive framework for unsupervised anomaly detection in univariate time series data from hydraulic turbine units," *Energy Rep.*, vol. 14, pp. 4701-4709, 2025. doi: 10.1016/j.energy.2025.123456.

[6] "HIPER-CHAD: Hybrid Integrated Prediction-Error Reconstruction-Based Anomaly Detection for Multivariate Indoor Environmental Time-Series Data," *Sensors*, vol. 26, no. 1, p. 171, 2026. doi: 10.3390/s26010171.

[7] "A Multi-Working States Sensor Anomaly Detection Method Using Deep Learning Algorithms," *Sensors*, vol. 25, no. 18, p. 5686, 2025. doi: 10.3390/s25185686.

[8] "Unsupervised Machine Learning Methods for Anomaly Detection in Network Packets," *Electronics*, vol. 14, no. 14, p. 2779, 2025. doi: 10.3390/electronics14142779.

[9] "Enhancing Critical Infrastructure Security: Unsupervised Learning Approaches for Anomaly Detection," *Int. J. Comput. Intell. Syst.*, vol. 17, p. 236, 2024. doi: 10.1007/s44196-024-00644-z.

[10] F. Antonicoli, F. R. Galeppi, and M. Piccirilli, "Advanced Time Series Forecasting and Anomaly Detection in Transactional Data," *GitHub Repository*, 2025. [Online]. Available: https://github.com/francescagaleppi/anomaly-detection-time-series

[11] F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation Forest," in *Proc. 8th IEEE Int. Conf. Data Mining*, Pisa, Italy, 2008, pp. 413-422. doi: 10.1109/ICDM.2008.17.

[12] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, 1997. doi: 10.1162/neco.1997.9.8.1735.

[13] C. Xiaoyang, "Explainable dual LSTM-autoencoders with exogenous features for anomaly detection and supply chain forecasting," *Sci. Rep.*, vol. 15, p. 42371, 2025. doi: 10.1038/s41598-025-26449-4.

[14] S. Selvakumar et al., "Enhanced Anomaly Detection in Wireless Sensor Networks Using Isolation Forest, LSTM, and LSTM Autoencoder," in *Proc. 3rd Int. Conf. Inventive Comput. Inform. (ICICI)*, 2025, pp. 1528-1534. doi: 10.1109/ICICI65870.2025.11069875.

[15] W. Villegas-Ch et al., "Integrating Explainable Artificial Intelligence in Anomaly Detection for Threat Management in E-Commerce Platforms," *IEEE Access*, vol. 13, pp. 29830-29846, 2025. doi: 10.1109/ACCESS.2025.3541979.

[16] "AI-based Hybrid Approach for Real-Time Anomaly Detection and Fault Management in Digital Services," in *Proc. Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, 2025. doi: 10.1109/ICECA66444.2025.11382729.

[17] T. Abu Musa, A. Bouras, and A. Belhi, "Anomaly Detection in Supply Chain Business Processes: A Deep Learning Approach," in *Product Lifecycle Management. Leveraging AI, Digital Twins, and Smart Technologies*, Springer, 2025, pp. 35-48. doi: 10.1007/978-3-031-93323-3_3.

[18] J. Cheng et al., "Real-Time Warehouse Monitoring with Ceiling Cameras and Digital Twin for Asset Tracking and Scene Analysis," *Logistics*, vol. 9, no. 4, p. 153, 2025. doi: 10.3390/logistics9040153.

[19] "Detecting Anomalies in Container Resource Traces From Production Datacenter," in *IEEE Conf. Proc.*, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/11072954

[20] M. Hosseini, S. Chalil Madathil, and M. T. Khasawneh, "Reinforcement learning-based simulation optimization for an integrated manufacturing-warehouse system: a two-stage approach,"*Expert Syst. Appl.*, 2025. doi: 10.1016/j.eswa.2025.125123.

[21] A. Iqbal and R. Amin, "Time series forecasting and anomaly detection using explainable AI," *J. Supercomput.*, vol. 81, no. 4, p. 523, 2025. doi: 10.1007/s11227-025-07040-0.

[22] M. I. Sayyaf, P. Pascacio, N. Zhu, and V. Renaudin, "Time-Series Anomaly Detection for Sensor Data: Models, Metrics, and Methodologies - A Review," *IEEE Sensors J.*, vol. 25, no. 24, pp. 43603-43619, 2025. doi: 10.1109/JSEN.2025.3616395.

[23] J. Blázquez-García, A. Conde, U. Mori, and J. A. Lozano, "A review on outlier/anomaly detection in time series data," *ACM Comput. Surv*., vol. 54, no. 3, pp. 1-33, 2021. doi: 10.1145/3444690.

[24] K. Shaukat et al., "A review of time-series anomaly detection techniques: A step to future perspectives," in *Adv. Intell. Syst. Comput.*, Springer, 2021. doi: 10.1007/978-3-030-73100-7_60.