IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

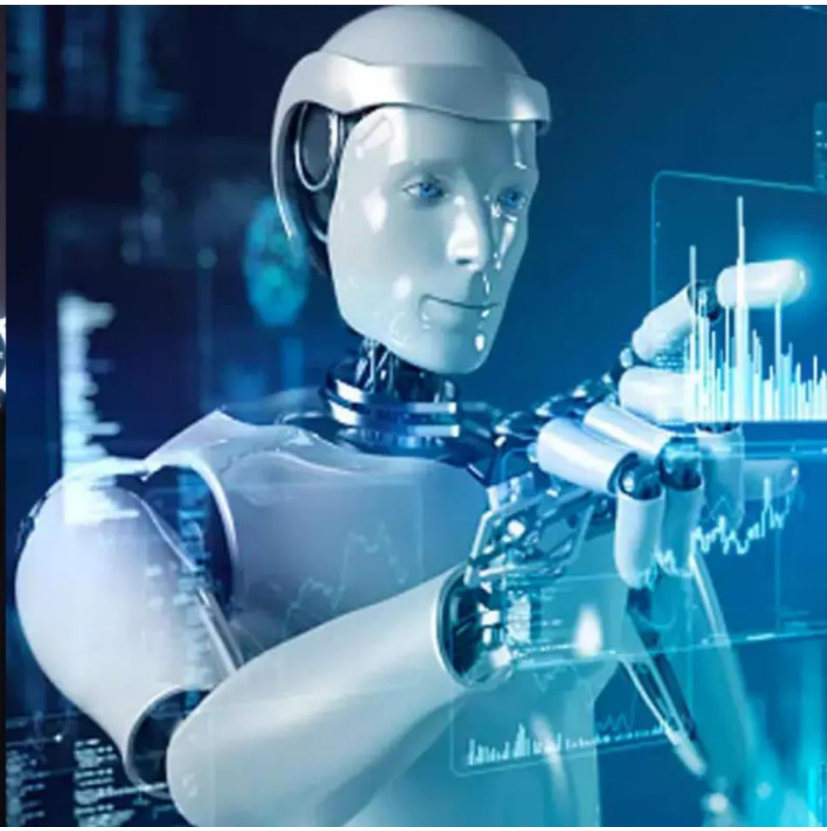
ISSN(O): 2319-8753
ISSN(P): 2347-6710



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 9.935

Volume 15, Issue 11, November 2026

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

A Unified Data Engineering Framework for Monitoring and Detection of Stock Discrepancies and Leakage in Medicine Supply Chains

Sharanabasavaraddi B, Yashawanth B, Purushotham M, Sharath Kumar H, Deekshith P

Assistant Professor, Department of Computer Science & Engineering, SJM Institute of Technology, Chitradurga, Karnataka, India

B.E. Student, Department of Computer Science & Engineering, SJM Institute of Technology, Chitradurga, Karnataka, India

B.E. Student, Department of Computer Science & Engineering, SJM Institute of Technology, Chitradurga, Karnataka, India

B.E. Student, Department of Computer Science & Engineering, SJM Institute of Technology, Chitradurga, Karnataka, India

B.E. Student, Department of Computer Science & Engineering, SJM Institute of Technology, Chitradurga, Karnataka, India

ABSTRACT: Pharmaceutical supply chains are growing in size and complexity, which makes transparency, efficiency, and accountability harder to maintain. Medicine leakage, stock discrepancies, and expiry-related wastage increase the cost of healthcare delivery and can affect patient care. This paper presents a unified data engineering framework, the Medicine Supply Chain Leakage Detection System, that integrates heterogeneous records on procurement, dispatch, inventory, consumption, and expiry. Records from different sources are aligned using standardized Hospital_ID and Medicine_ID keys, and an expected stock level is computed for every facility–medicine combination and compared with the reported stock. Discrepancies are examined using statistical thresholds, rule-based checks, and an Isolation Forest anomaly detector to identify possible diversion, undocumented loss, and irregularities in the chain of custody. The prototype is implemented with Python (Pandas), MySQL storage, a FastAPI backend, and a React-based dashboard. The development dataset combines available operational records with synthetically generated records that include injected leakage scenarios; the results therefore demonstrate the behaviour of the framework on that dataset and are not a field validation. Within this setting, the framework reconciled stock for 50 hospital–medicine combinations, flagged 964 anomalies with an ensemble of rule-based, Isolation Forest, and Z-score detection, identified expired stock held in inventory, and improved visibility across the supply chain stages.

KEYWORDS: Medicine Supply Chain, Leakage Detection, Anomaly Detection, Isolation Forest, Data Integration, Inventory Management, Healthcare Logistics, Supply Chain Transparency.

I. INTRODUCTION

The healthcare sector depends on a coordinated supply chain to keep essential medicines available when patients need them. Medicines pass through several stages, including sourcing, storage, dispatch, dispensing, and consumption, and the performance of the whole chain depends on how well each stage is managed. As healthcare infrastructure expands and demand for services grows, managing this flow becomes more difficult. Existing systems often work in isolation with limited connection between stages, which gives poor visibility into how medicines actually move.

Three problems are commonly reported. The first is **stock mismatch**, where recorded inventory differs from the physical quantity because of manual entry errors, weak tracking, or diversion of stock. The second is **wastage**, where medicines expire because usage is poorly monitored or forecast; unusually high usage may also indicate misuse. The

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

third is **fragmented data**: dispatch, inventory, consumption, and prescription records are held in separate systems that are rarely linked, so irregularities are noticed only after losses have become large.

The **Medicine Supply Chain Leakage Detection System** presented in this paper addresses these problems through a data-engineering approach. It collects data from several sources, cleans and aligns it, calculates the stock that should exist at each facility, compares it with the stock that is reported, and raises alerts when the difference is significant. The paper describes the motivation, design, and implementation of the framework, together with its evaluation and limitations. The system is intended to support administrators and auditors in prioritizing investigations; a flagged discrepancy is an indication that needs verification and is not proof of theft or misconduct.

II. LITERATURE SURVEY

[1] Anomaly Detection in Electronic Health Records Across Hospital Networks (Niu, X., et al., 2025)
This work combines machine learning with graph-based methods to detect anomalies in electronic health records across hospital networks. The approach improves detection accuracy on large healthcare datasets. However, it works on clinical records and does not consider supply chain or logistics data.

[2] A Systematic Review of Blockchain Applications in Pharmaceutical Supply Chain Traceability and Security (Chaabaoui, A., et al., 2024)
This review explains how decentralization and immutability of blockchain records improve transparency and security in pharmaceutical supply chains. It highlights protection against counterfeit products and data tampering. The reviewed solutions focus on traceability and do not apply anomaly detection to find discrepancies in stock.

[3] Blockchain in Pharmaceutical Supply Chain Management (Nimsadkar, S., et al., 2024)
This study applies blockchain to the pharmaceutical supply chain to reduce counterfeit drugs and prevent tampering with transaction data. It increases trust among stakeholders. It does not provide analytical methods for detecting leakage or inefficiency.

[4] Enhancing Pharmaceutical Supply Chain Visibility Using IoT and Blockchain (Fallah, M., et al., 2024)
This paper proposes a framework in which IoT sensors collect live data on temperature, location, and storage conditions, and the data is secured using blockchain. It improves real-time monitoring and data integrity. It does not include intelligent analysis of the collected data.

[5] Machine Learning-Enhanced Anomaly Detection in Healthcare Monitoring: A Survey (Various Authors, 2024)
This survey shows that machine learning models are effective for analysing healthcare data and automating anomaly detection. Most of the surveyed applications target specific monitoring tasks rather than the movement and stock of medicines across a supply chain.

[6] A Secure Blockchain-based Pharmaceutical Supply Chain Management System: Traceability and Detection of Counterfeit COVID-19 Vaccines (Dashtizadeh, M., et al., 2022)
This work uses a permissioned blockchain to secure the vaccine supply chain and detect counterfeit vaccines. It improves traceability and security. It does not address quantity-level stock discrepancies or leakage between dispatch and consumption.

[7] Blockchain Technology in Pharmaceutical Supply Chain Management (Pathak, A., et al., 2022)
This study uses smart contracts to automate transactions and improve traceability in the pharmaceutical supply chain. It increases automation and transparency but includes no method for discovering leakage or process inefficiency.

[8] Enhancing Traceability in Pharmaceutical Supply Chain using IoT and Blockchain (Botcha, R., et al., 2019)
This paper shows that combining IoT tracking with blockchain can provide end-to-end traceability of pharmaceutical products. It improves real-time visibility, but it does not analyse the recorded data to identify losses or abnormal movement.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

The reviewed work shows that blockchain improves transparency and security, IoT improves real-time monitoring, and machine learning helps analyse complex datasets. These techniques are mostly used independently. No holistic approach integrates dispatch, inventory, consumption, and expiry data to detect medicine leakage, stock discrepancies, and process inefficiency. This research gap motivates the present work.

III. METHODOLOGY

1) System Architecture

The Medicine Supply Chain Leakage Detection System is a data-driven monitoring solution that tracks medicines from central warehouses and distribution centres to hospitals, pharmacies, and patients. In this paper, leakage means any situation in which medicines are diverted, misreported, stolen, expired without documentation, or otherwise lost between dispatch and final consumption. The architecture consists of five main components:

- **Data Sources:** Procurement, dispatch, inventory, consumption, and expiry records from healthcare facilities and supply chain entities.
- **Data Engineering Layer (Python, Pandas):** Cleans, aligns, and merges the records and derives the expected stock for each facility and medicine.
- **Database (MySQL):** Stores raw and processed data, including medicine details, transaction records, and inventory status.
- **Detection Engine:** Applies threshold rules, statistical checks, and Isolation Forest to flag anomalies and assign severity levels.
- **Backend and Dashboard:** A FastAPI backend serves aggregated results through a REST API to a React-based dashboard that shows alerts, reports, and charts.

These components work together to give an organized view of medicine movement, stock status, and potential leakage.

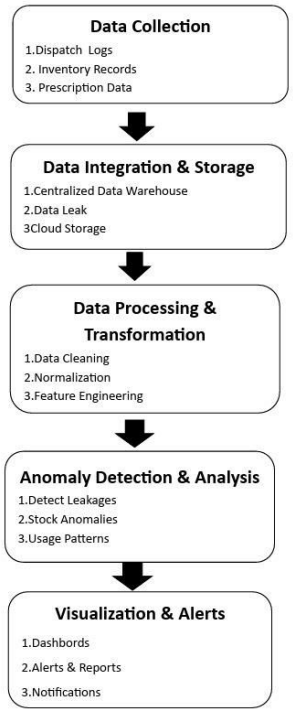


Fig.1. System Architecture and Operational Pipeline of the Medicine Supply Chain Leakage Detection System

2) Data Sources

The system works with real-world and synthetically generated datasets that mirror the operational data kept by healthcare institutions. Where available, procurement records, pharmacy dispensing logs, and warehouse dispatch

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

manifests are used. Synthetic data is generated to fill gaps and to simulate seasonal demand, occasional human-error entries, and deliberate leakage scenarios for validation. Each record carries at least a Hospital_ID, a Medicine_ID, a date or timestamp, and a quantity. The data categories are summarized in Table I.

Table I. Data Sources and Their Description

Data Source	Description
Procurement Records	Details of medicines ordered, received, and invoiced at each facility.
Dispatch Records	Information on medicines dispatched from central stores and warehouses.
Inventory Logs	Stock-level records maintained at hospitals, clinics, and pharmacies.
Consumption Records	Data on medicines administered to patients or dispensed from wards.
Expiry & Waste Records	Logs of medicines discarded due to expiry, damage, or spoilage.
Hospital & Medicine IDs	Unique identifiers that link records across all datasets.

3) Data Preprocessing
Raw data from multiple sources is rarely clean or consistent, so the preprocessing stage performs the operations summarized in Table II.

Table II. Summary of Data Preprocessing Techniques

Technique	Description
Missing Values	Records with missing quantities, unidentifiable medicine codes, or absent facility identifiers are imputed using contextual averages, or flagged and excluded with a log entry.
Duplicate Removal	Duplicates caused by double entry or synchronization errors are identified using the composite key (Hospital_ID + Medicine_ID + Date + Quantity) and removed.
Date Standardization	All dates and timestamps are converted to one format so that records from different systems align chronologically.
Unit Normalization	Quantities recorded in different units (tablets, strips, bottles, millilitres) are converted to one standard unit per medicine using a normalization map.
Outlier Flagging	Statistically extreme values, such as an unrealistically large single dispatch, are flagged for manual review before analysis.

4) Data Alignment and Processing
Data alignment is one of the most important steps. All datasets are aligned using two mandatory fields: Hospital_ID, which ensures that records of one facility are never matched with those of another, and Medicine_ID, which ensures that quantities are compared for the same medicine across datasets. A unified master record is created for each Hospital_ID and Medicine_ID combination within a defined time window. Records that cannot be matched are kept in an unmatched-records register, which is itself a source for investigation of leakage or data gaps.

After alignment, the processing stage performs dataset merging into a single analytical dataset, derived-field calculation (expected remaining stock), time-window grouping into monthly periods for trend analysis, and validation checks that confirm required fields are present and quantities are non-negative.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

5) Algorithm Flow

The detection workflow follows six steps.

- **Step 1 – Input Datasets:** Records are collected from manufacturers, distributors, and retailers. Each record contains batch ID, quantity, timestamp, and location.
- **Step 2 – Data Alignment:** Duplicates are removed, missing values handled, and formats standardized. Records are aligned using batch ID and time sequence so the flow of each batch can be traced.
- **Step 3 – Calculate Expected Stock:** For each stage, Expected Stock = (Previous Stock + Incoming) − Outgoing. This gives a continuous record of the quantity that should be available at each checkpoint.
- **Step 4 – Compare with Actual Stock:** The reported or real-time stock is retrieved and compared with the expected stock to find discrepancies.
- **Step 5 – Flag Anomalies:** If the difference between expected and actual stock exceeds a predefined threshold, it is marked as an anomaly. Possible causes include leakage or theft, data-entry errors, and counterfeit or unauthorized stock.
- **Step 6 – Generate Alerts:** Dashboard notifications and SMS or email alerts are produced with the location, time, batch details, and severity level to support quick investigation.

6) Anomaly and Leakage Detection

For each Hospital_ID and Medicine_ID combination, the system performs three comparisons: dispatched quantity against received quantity (a shortfall suggests possible diversion in transit); received quantity against consumed plus remaining quantity (a mismatch indicates undocumented loss); and consumed quantity against patient records (consumption that consistently exceeds the expected clinical dosage may indicate false reporting).

Not every small discrepancy is leakage. Configurable threshold rules log differences within a tolerance range as informational, and classify those above the threshold as potential leakage. An Isolation Forest model is used alongside the statistical and rule-based checks to isolate unusual patterns that fixed thresholds may miss. Each detected anomaly is assigned one of four severity levels:

- **Critical:** large discrepancy in a high-value medicine with no justification.
- **High:** moderate discrepancy that recurs.
- **Medium:** discrepancy indicating systemic inefficiency.
- **Low:** minor discrepancy most likely caused by a data-entry error.

7) Implementation

The system is implemented in Python. The Pandas library loads data from manufacturers, distributors, and retailers into DataFrames and performs cleaning, alignment on batch ID and timestamps, and expected-versus-actual stock calculations. All raw and processed data is stored in a MySQL database with tables for medicine details, transaction records, and inventory status, which provides structured storage and SQL-based querying. Python scripts fetch data from MySQL, apply the detection algorithm, and identify anomalies; they can run periodically as scheduled tasks or as event-based processing. The visualization layer is a React web application using Recharts for bar, pie, and line charts across five dashboard pages (overview, stock reconciliation, anomaly detection, expiry monitoring, and pipeline status). It communicates with a FastAPI backend through a REST API, and all aggregations are performed server-side with Pandas. The architecture is decoupled, which keeps the stack modular and focused on the data engineering pipeline. Access to the dashboard is controlled through sign-in and registration pages.

The final outputs are automated alerts (containing facility, medicine, period, dispatched quantity, expected and actual remaining quantity, and discrepancy value), analytical reports (total leakage events, breakdown by facility and medicine category, trends, and aggregate value of discrepancies), and a visual dashboard with summary statistics, a filterable table of leakage events, time-series charts, and a facility-level comparison.

8) Evaluation Methodology

The framework is evaluated on its ability to track medicine movement, detect discrepancies, and report them clearly. Because the dataset combines available records with synthetic records that contain deliberately injected leakage scenarios (stock shortages, excess stock, missing transactions, and delays), the evaluation checks whether these scenarios are detected and how the flagged events are distributed by severity, hospital, medicine, and detection method. Accuracy, precision, recall, and F1-score are used to assess the anomaly detector, and expiry status and wastage are reported for each medicine. As the data is partly synthetic and no field audit was carried out, the results demonstrate the behaviour of the framework on this dataset and do not establish detection accuracy in a live healthcare setting.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

IV. EXPERIMENTAL RESULTS

The Medicine Supply Chain Leakage Detection System was executed end to end on the development dataset of 5,900 supply-chain records, covering ten medicines and five hospitals (Apollo, Fortis, AIIMS, Manipal, and a District Government Hospital), which gives 50 hospital–medicine combinations for reconciliation. The outputs are presented through the five pages of the React dashboard: Overview, Stock Reconciliation, Anomaly Detection, Expiry Monitoring, and Pipeline Status (Figs. 3–7), which are reached through the sign-in and registration pages (Fig. 2). The headline figures are summarized in Table III.

1) User Authentication

Access to the dashboard is controlled through sign-in and registration pages (Fig. 2). The sign-in page asks for an email address and a password and links to account creation. The create-account page asks for a full name, an email address, a password of at least 8 characters, and a password confirmation. After sign-in, the user name and role (for example, Analyst) are shown in the sidebar together with a Logout button, as seen in Figs. 3–7.

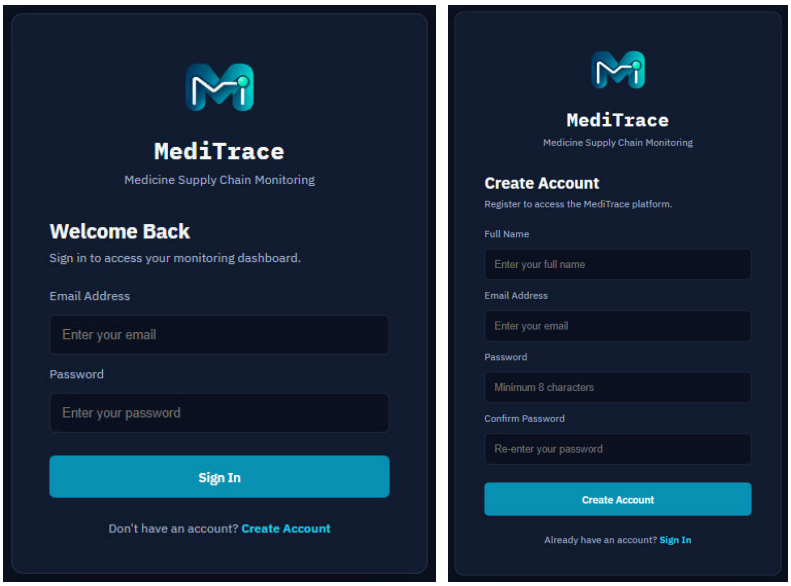


Fig.2. Authentication pages of MediTrace: (a) sign-in page; (b) create-account page.

Table III. Summary of Key Outputs Produced by the System

Output Measure	Value	Source (Dashboard Page)
Total dispatched quantity	589,651 units	Overview
Total stock discrepancy	3,976 units	Overview / Stock Reconciliation
Leakage cases (flagged dispatches)	81	Overview
Reconciliation combinations by risk	48 Low, 2 Medium (50 total)	Stock Reconciliation
Anomalies detected (ensemble)	964	Anomaly Detection
Accuracy / precision / recall / F1	90.6% / 41.8% / 40.7% / 41.3%	Anomaly Detection
Expired in stock / expiring in 90 days / units wasted	311 / 91 / 107,051	Expiry Monitoring

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

2) Overview Dashboard

The Overview page (Fig. 3) combines the main indicators in one view: 589,651 units dispatched, 3,976 units of total discrepancy, 81 leakage cases, 311 batches that expired while still in stock, and 964 anomalies flagged by the ensemble detector. Amoxicillin 250mg was the most dispatched medicine at 74,589 units across all hospitals. Anomalies were spread fairly evenly over the five hospitals, with Fortis Hospital and Apollo Hospital each accounting for about 22%, followed by the District Government Hospital (20%), Manipal Hospital (19%), and AIIMS (16%); Fortis Hospital had the highest count at 208 anomalies, so it is the first candidate for a supply-chain audit.



Fig.3. Overview dashboard: summary indicators, dispatch volume by medicine, anomalies by hospital, discrepancy by medicine, and anomaly type breakdown.

The discrepancy chart ranks medicines from the highest to the lowest total discrepancy; it peaks at 600 units for Omeprazole 20mg and falls to roughly 130 units for Atorvastatin 10mg. In the anomaly type breakdown, Expired_In_Stock is the most common type with 311 occurrences, followed by over-prescription and usage-related anomalies.

3) Stock Reconciliation

The Stock Reconciliation page (Fig. 4) applies the expected-stock rule, Expected Stock = Opening Stock + Received Quantity − Dispensed Quantity, to every hospital–medicine combination and compares the result with the reported stock. Of the 50 combinations, 48 were classified as Low risk and 2 as Medium risk, and the total discrepancy was 3,976 units. The worst case was Amoxicillin 250mg at the District Government Hospital, with a discrepancy of 304 units (Medium risk). The table also gives the risk level of every combination, so a flagged case can be traced to its facility. Several combinations show negative values, for example Metformin 500mg at Apollo Hospital (−24 units), which means that the reported stock is higher than the expected stock; such cases point to data-entry inconsistency or unrecorded receipts and not to loss.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

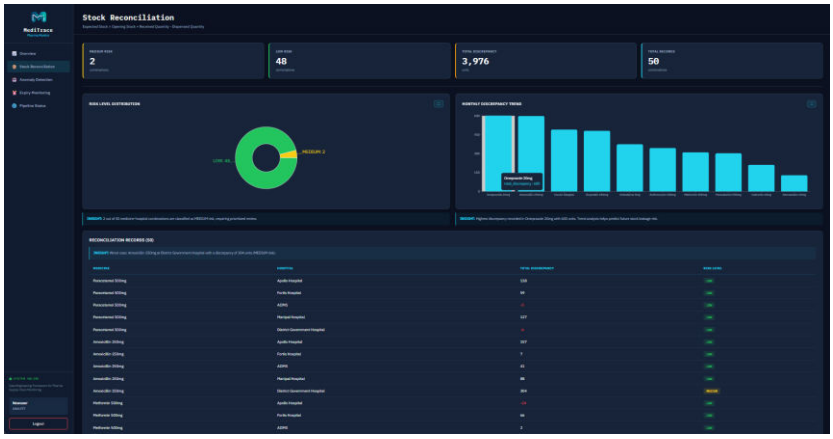


Fig.4. Stock reconciliation page: risk-level distribution, discrepancy by medicine, and the reconciliation records table with risk levels.

4) Anomaly Detection

The Anomaly Detection page (Fig. 5) reports the output of the ensemble of rule-based checks, Isolation Forest, and Z-score detection. In total 964 anomalies were found among the 5,900 processed records. The rule-based component detected the largest share (586 anomalies), and the Isolation Forest and Z-score components contributed the remaining detections, which indicates that the statistical and machine-learning methods add detections beyond fixed rules. By severity, High is the most frequent class with 523 cases, followed by Medium, while Critical cases are the fewest; High-severity cases are the ones flagged for immediate intervention. The detailed table lists the medicine, hospital, anomaly type, severity, detection method, and confidence score (90% for the rule-based Medicine_Leakage records shown).

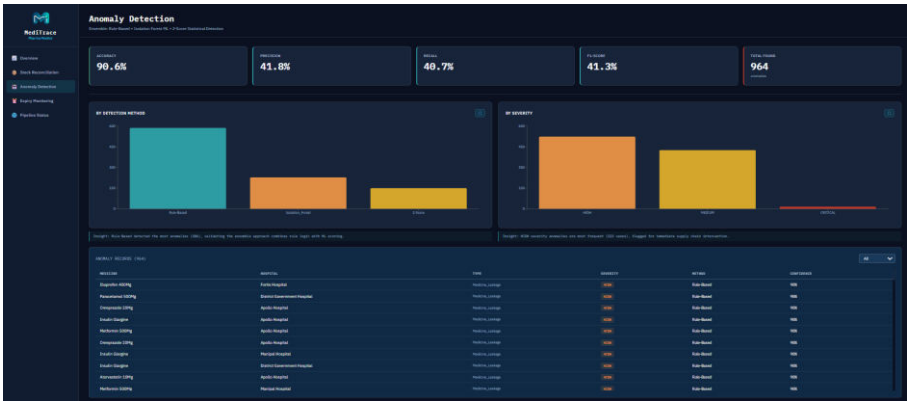


Fig.5. Anomaly detection page: performance metrics, anomalies by detection method and by severity, and the anomaly records table.

The detection metrics on the dataset are an accuracy of 90.6%, a precision of 41.8%, a recall of 40.7%, and an F1-score of 41.3%. The high accuracy partly reflects the fact that anomalous records are a minority of all records, so precision and recall give a more realistic picture: only about two in five flagged records are correct, and about two in five anomalies are found. These values were obtained on the partly synthetic development dataset and not against an audited ground truth. They show that the ensemble is useful for ranking records for review but not for automatic decisions.

5) Expiry Monitoring

The Expiry Monitoring page (Fig. 6) tracks expired and near-expiry batches. The system found 311 batches that had already expired but were still recorded in active inventory, a further 91 batches expiring within 90 days, and 107,051 units wasted in total. The status distribution shows 50% of the monitored records as Active, 11% as Expiring Soon, and

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

39% as Expired. Metformin 500mg had the highest wastage at 12,964 units, followed by Ibuprofen 400mg and Amoxicillin 250mg, which suggests over-procurement or a weak stock-rotation policy for these medicines.

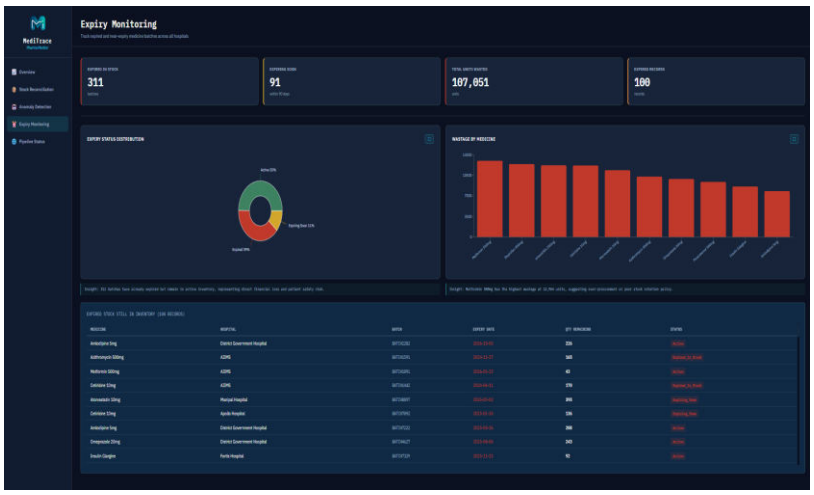


Fig.6. Expiry monitoring page: expiry status distribution, wastage by medicine, and the table of expired stock still in inventory.

6) Pipeline Status

The Pipeline Status page (Fig. 7) shows the result of the last run of the seven-task workflow. The pipeline finished with the status SUCCESS, processed 5,900 records, and found 964 anomalies. Three of the seven tasks were fully verified: data ingestion into MySQL, anomaly detection, and report generation (anomaly CSV and ML-metrics JSON). Health check and stock reconciliation were marked Partial, data quality check was marked Attention because one inventory dataset failed a quality check, and alert generation was still Pending because alert-delivery integration has not yet been verified.

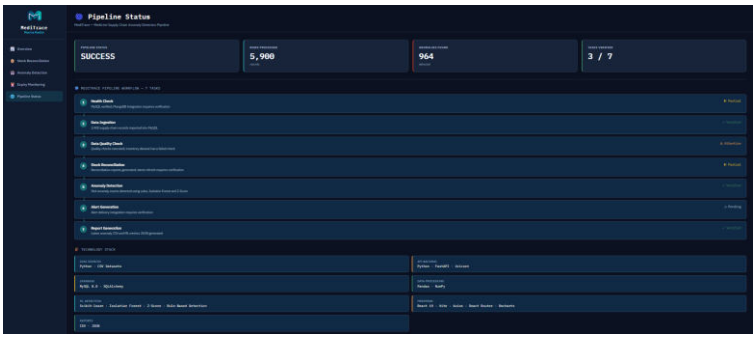


Fig.7. Pipeline status page: workflow task status, technology stack, and the pipeline report.

V. PLANNED IMPROVEMENTS

The dashboard outputs in Section IV show several points that need correction before the final version of the system. The hospital chart lists AIIMS and Aims as separate categories; some records show an Active status although their expiry dates are in the past; the discrepancy charts group values by medicine and not by date; the pipeline health check refers to MongoDB although the technology stack lists only MySQL; the precision, recall, and F1-score of the anomaly detector are moderate; and only 3 of the 7 pipeline tasks are verified, with alert generation still pending. Table IV lists the planned improvements and the result expected from each.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

Table IV. Planned Improvements and Expected Results

No.	Improvement	What Will Be Implemented	Expected Result
1	Hospital name standardization	Normalize AIIMS/Aims and duplicate names	Correct five-hospital chart
2	Expiry validation	Correct status using expiry dates	No expired medicine incorrectly shown as Active
3	Discrepancy trend	Add chronological date-based aggregation	Actual discrepancy-over-time chart
4	Database health check	Remove incorrect MongoDB verification if MySQL is the only database	Accurate pipeline health status
5	ML evaluation	Investigate precision, recall, and F1-score	Reliable anomaly evaluation
6	Alert generation	Implement and test notification functionality	Working alert generation
7	Pipeline verification	Verify all seven pipeline tasks	Accurate pipeline completion status
8	Research reporting	Generate consistent experimental results	Reproducible research evidence

The expected changes in each part of the final MediTrace implementation are summarized in Table V.

Table V. Expected Changes in the Final MediTrace Implementation

Module	Planned Changes
Overview Dashboard	Five correctly identified hospitals; correct anomaly count; an actual chronological discrepancy trend; consistent KPIs across all pages.
Stock Reconciliation	Correct medicine–hospital combinations; accurate risk classification; verified discrepancy calculations.
Anomaly Detection	Correct evaluation methodology with a confusion matrix, precision, recall, accuracy, and F1-score; comparison of Rule-Based, Isolation Forest, and Z-Score detection; better detection validated against known labels.
Expiry Monitoring	Correct expired, active, and expiring-soon classification; accurate expired stock quantity; correct wastage calculations.
Pipeline Status	Verification of the MySQL health check, dataset validation, data ingestion, reconciliation, anomaly detection, alert generation, and report generation.
Alert System	Notification generation and testing of actual delivery, with a clear distinction between generated alerts and successfully delivered notifications.

VI. CONCLUSION

This paper presented a unified data engineering framework for detecting stock discrepancies and leakage in medicine supply chains. The framework integrates dispatch, inventory, consumption, and expiry data; computes expected stock; and combines threshold rules, statistical checks, and Isolation Forest to flag and prioritize discrepancies through alerts, reports, and a dashboard. On the development dataset, it flagged 964 anomalies in 5,900 records, reached an accuracy of 90.6% with a precision of 41.8% and a recall of 40.7%, and showed where discrepancies and expired stock were concentrated by hospital and medicine.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 9.935

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 11, November 2026

|DOI: 10.15680/IJIRSET.2026.1511033|

The framework offers a shift from manual record-keeping to data-driven monitoring, faster detection with less manual checking, and a scalable design suitable for small hospitals as well as larger healthcare networks. Possible applications include hospitals, clinics, pharmacies, government drug distribution programmes, pharmaceutical distributors, hospital inventory systems, and healthcare analytics platforms.

The current work has limitations. The dataset is partly synthetic, detection quality depends on data completeness and threshold settings, and the precision and recall obtained so far are moderate and have been computed on the development dataset and not against audited ground truth. The planned improvements in Section V address the hospital-name, expiry-classification, discrepancy-trend, anomaly-evaluation, and alert-delivery issues identified in the dashboard outputs. Future work also includes validation with real, ethically obtained data; machine-learning-based demand forecasting; real-time monitoring with instant alerts; blockchain integration for tamper-proof records; and scaling to national-level supply chains. The system is a decision-support tool, and flagged events must be verified by authorized personnel before any action is taken.

REFERENCES

1. Niu, X., et al., “Anomaly Detection in Electronic Health Records Across Hospital Networks,” 2025.
2. Chaabaoui, A., et al., “A Systematic Review of Blockchain Applications in Pharmaceutical Supply Chain Traceability and Security,” 2024.
3. Nimsadkar, S., et al., “Blockchain in Pharmaceutical Supply Chain Management,” 2024.
4. Fallah, M., et al., “Enhancing Pharmaceutical Supply Chain Visibility Using IoT and Blockchain,” 2024.
5. “Machine Learning-Enhanced Anomaly Detection in Healthcare Monitoring: A Survey,” 2024.
6. Dashtizadeh, M., et al., “A Secure Blockchain-based Pharmaceutical Supply Chain Management System: Traceability and Detection of Counterfeit COVID-19 Vaccines,” 2022.
7. Pathak, A., et al., “Blockchain Technology in Pharmaceutical Supply Chain Management,” 2022.
8. Botcha, R., et al., “Enhancing Traceability in Pharmaceutical Supply Chain using IoT and Blockchain,” 2019.



INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details