Real-time Anomaly Detection in Community Pharmacies Prescription Processing Using Machine Learning

Suprito Biswas[a], Sina Ansari[a], Vahid Alizadeh[b], Ateequr Rahman[c,*]

[a]*Driehaus College of Business, DePaul University, Chicago, IL, USA 60604*
[b]*College of Computing and Digital Media, DePaul University, Chicago, IL, USA 60604*
[c]*College of Pharmacy, Rosalind Franklin University, North Chicago, IL, USA 60064*

Abstract

Community pharmacies play a vital role in the U.S. healthcare system by ensuring timely, accurate medication delivery and essential patient counseling. However, increasing workloads, chronic understaffing, and the impact of COVID-19 have disrupted pharmacy workflows, leading to inefficiencies, delays, and prescription errors. Identifying these anomalies in real-time is critical for maintaining patient safety and operational efficiency. This study introduces a machine learning (ML)-based anomaly detection framework designed to optimize pharmacy workflows and reduce errors. Using data from a community pharmacy in Illinois, we developed phase-based detection models targeting two key stages of the prescription process: entry and verification. Algorithms such as Isolation Forest, Local Outlier Factor (LOF), and K-Nearest Neighbors (KNN) were employed to identify irregularities linked to pending verifications, staff workload fluctuations, and patient visit frequencies. Our findings indicate that anomalies significantly correlate with extended processing times. Implementing real-time anomaly detection allows pharmacies to proactively address workflow bottlenecks, optimize resource allocation, and improve prescription accuracy.

1. Introduction

Community pharmacists are vital in the healthcare system, offering medication therapy management, patient education, and prescription services. However, increasing workload and time constraints have placed significant pressure on pharmacists and impacted patient care [1, 2]. Reports indicate that over 68% of pharmacists experience job-related stress and role overload [3], with 76.8% facing burnout [4]. In 2020, independent community pharmacies dispensed an average of 57,678 prescriptions (RX) annually (approximately 185 per day), up from 57,414 in 2019 [5]. With 19,397 independent community pharmacies in the U.S., representing a $67.1 billion market [5], operational inefficiencies—exacerbated by the COVID-19 pandemic and staffing shortages—have led to notable disruptions in pharmaceutical care [6, 4, 7, 8].

A primary factor contributing to these inefficiencies is the presence of undetected process anomalies within pharmacy workflows. An anomaly in prescription processing refers to any deviation from expected operational patterns, such as irregular handling times, data entry errors, or discrepancies in patient or staff activities. Undetected anomalies in healthcare can result in serious outcomes, such as medical errors, compromised patient safety, and increased costs [9, 10]. Detecting these anomalies is crucial, as they can lead to workflow delays, prescription errors, and risks to patient safety. This study aims to enhance workflow efficiency in community

*Corresponding author: ateequr.rahman@rosalindfranklin.edu

pharmacies by leveraging AI for real-time anomaly detection. While anomaly detection (also known as outlier or novelty detection) has been widely explored in AI and analytics literature [11, 12], its application in healthcare remains limited. The complexity of healthcare data and the potential consequences of undetected anomalies present unique challenges in this domain. Effective anomaly detection in healthcare settings can improve patient outcomes, reduce operational costs, and enhance overall efficiency.

This study explores anomaly detection in prescription processing workflows using data from a community pharmacy in Illinois. We focus on identifying irregular patterns that signal inefficiencies or delays in the workflow. We found that anomalies often occur during two critical phases: prescription entry and final verification. High volumes of prescriptions pending verification, fluctuations in staff workload, and irregular patient visit frequencies were key indicators of workflow disruptions. These anomalies were associated with significant delays, particularly in the verification stage, highlighting bottlenecks that can compromise patient safety and operational efficiency. Our results suggest that real-time anomaly detection can help pharmacies proactively identify and address inefficiencies, optimize staff allocation, and reduce prescription errors, ultimately enhancing patient care and satisfaction.

2. Background

Technological advancements in community pharmacies have aimed to reduce workload and enhance productivity through tools such as decision-support systems for inventory management [13, 14], automated drug inventory control systems [15], and data analytics for improving medication management and minimizing errors [16, 17]. However, these solutions primarily focus on inventory and administrative tasks, leaving gaps in addressing the complexities of prescription processing workflows. Specifically, detecting anomalies—irregularities that can disrupt operations—remains underexplored in community pharmacy settings.

When applied to healthcare data, traditional anomaly detection methods face challenges, which are high-dimensional, dynamic, and heterogeneous [18]. The vast amount of healthcare data, computational limitations, and the scarcity of labeled anomaly data further complicate effective detection [19, 20]. Recent developments in machine learning offer promising solutions by effectively managing complex, high-dimensional data and learning patterns without the need for extensive labeled datasets [21, 22]. Despite their success in other domains, the application of machine learning-based anomaly detection in community pharmacies remains largely unexplored. Given the unique characteristics of pharmacy data—including its interdependencies and heterogeneity—there is a significant opportunity to apply these advanced techniques to improve prescription processing and error detection in this critical healthcare sector.

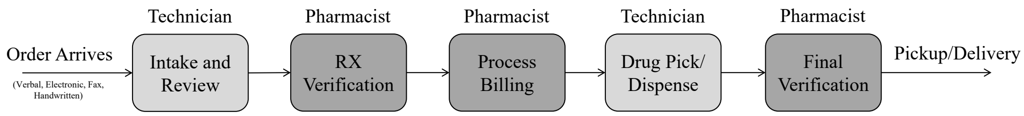
3. Method

This section discusses the data, our Anomaly Detection framework, and how we interpret and Validate the results.

3.1. Data Collection and Dataset Overview

The dataset was sourced from a family-owned community pharmacy in Illinois that became fully operational in 2020. The pharmacy employs three pharmacists (two full-time, one part-time) and two technicians (one full-time, one part-time), processing 80-100 prescriptions on slow days and up to 140 on busy days. Figure 1 illustrates the collaborating pharmacy's workflow. The process begins with a technician performing an *Intake and Review*, followed by *RX (Prescription) Verification* by the pharmacist. Once verified, billing processes are conducted, leading to drug dispensation by the technician. A pharmacist then completes a *Final Verification* before the medication is prepared for *Pickup/Delivery* to the patient.

Figure 1: High-level Workflow for the Community Pharmacy



Data collected from January to August 2023 includes 16,610 prescription records, excluding prescriptions containing only COVID-19 test kits or vaccines were excluded. The dataset tracks various aspects of prescription processing, staff activity, patient details, and insurance information, reflecting standard community pharmacy operations, summarized in Table **??**.

Table 1: Description of Prescription Process Variables

Category	**Description and Variables**
Timestamp Variables	Records the date and time at various phases of the prescription process (e.g., *RxEntered-DateTime, RxPhVerifDate, RxDatePickup*).
Staff Variables	Identifies the staff members involved in different phases of the prescription process (e.g., *RxEnteredBy, RxPhVerifUser, RxDPVerifUser*).
Patient Variables	Includes patient identifiers and the doctors prescribing the medication (e.g., *RxPatientNo, RxDoctorname*).
Prescription Variables	Details about the medications prescribed, their National Drug Code (NDC) status, and current status (e.g., *RxDrugName, RxDrugNDC, RxStatus, RxRefillStatus*).
Insurance Variables	Information about the insurance carriers involved in the prescription process (e.g., *InsuranceCarrierName, InsuranceEntityName*).

3.2. Anomaly Detection Framework

Our approach focuses on phase-based detection models targeting specific stages of prescription processing, including the entry and verification phases. We engineered phase- and non-phase-dependent features, capturing temporal, staff-related, patient, insurance, and medication-specific variables. Machine learning models employed include Isolation Forest, Local Outlier Factor (LOF), and K-Nearest Neighbors (KNN). These models are chosen based on their effectiveness in handling high-dimensional healthcare data, as reported in the benchmark study by [23].

To integrate the outputs from these models, we apply two strategies: the *Majority Votes* approach, which labels anomalies identified by at least two models, and the *Conservative* approach, which only considers anomalies consistently flagged by all three models. This dual strategy balances sensitivity and precision, ensuring robust anomaly detection.

3.3. Interpretation and Validation

To interpret the detected anomalies, we employ association mining and the Frequent Pattern Outlier Factor (FPOF) method to uncover patterns leading to workflow inefficiencies. Additionally, Welch's T-tests are conducted to compare processing times (e.g., turnaround time) between anomalous and non-anomalous cases, determining the statistical significance of detected anomalies. Rigorous cross-validation procedures are implemented to ensure the reliability of the models, with evaluation metrics assessing accuracy, precision, and recall.

4. Results

In this section, we present the results of the two phase-based anomaly detection models, which target distinct stages of the prescription workflow: Phase 1 (prescription entry) and Phase 2 (prescription verification). By isolating these phases, we identified specific operational patterns and inefficiencies contributing to workflow anomalies.

4.1. Anomaly Detection Performance

Table 2 summarizes the Isolation Forest, LOF, and KNN performance across both phases.

Table 2: Summary of the Phase 1 and Phase 2 anomaly detection models features and performance.

Attributes	**Phase 1 Model**	**Phase 2 Model**
List of Features	Temporal: RxPendingVerificationCount_AtEntry PendingPickupCount_AtEntry DayOfWeek_AtEntry DayOfMonth_AtEntry PeakHoursIndicator_AtEntry Staff-related: PrescriptionsPerStaff_AtEntry StaffVolume_AtEntry Patient-related: PatientVisitFrequency IsNewPatient Insurance-related: InsuranceType Medication-related: AdministrationRoute DEASchedule	Temporal: EntryToVerificationTime RxPendingVerificationCount_AtVerification PendingPickupCount_AtVerification DayOfWeek_AtVerification DayOfMonth_AtVerification PeakHoursIndicator_AtVerification Staff-related: PrescriptionsPerStaff_AtVerification StaffVolume_AtVerification Patient-related: PatientVisitFrequency IsNewPatient Insurance-related: InsuranceType Medication-related: AdministrationRoute DEASchedule
% of Observations labeled Anomalies	Labeled by three methods (conservative): 1.28% Labeled by two methods (majority votes): 5.64% Labeled by one method: 14.85%	Labeled by three methods (conservative): 0.55% Labeled by two methods (majority votes): 4.98% Labeled by one method: 18.22%
Silhouette Coefficient	Isolation Forest: 0.1471 LOF: 0.1568 KNN: 0.2395	Isolation Forest: 0.1325 LOF: 0.1229 KNN: 0.2493
IREOS Criterion	Isolation Forest: 2.2679 LOF: 2.5142 KNN: 2.4314	Isolation Forest: 2.1820 LOF: 2.3243 KNN: 2.3691

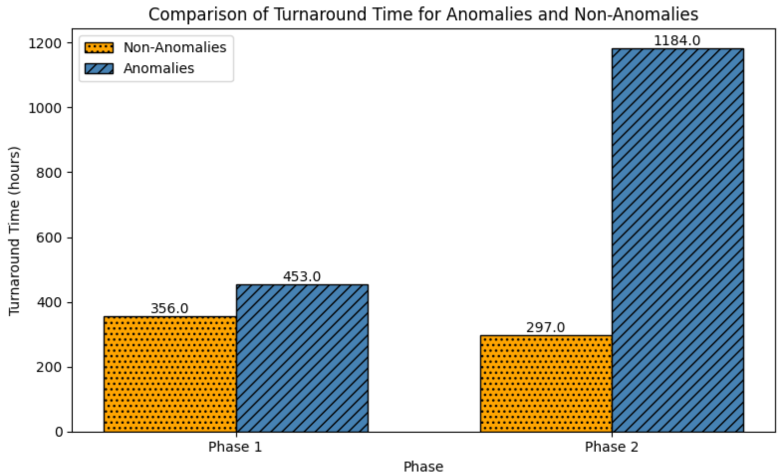
Notes. This table outlines the attributes used, the percentage of observations labeled as anomalies by different methods, and evaluation metrics, including the Silhouette Coefficient and IREOS Criterion across Isolation Forest, LOF, and KNN models.

In Phase 1, the Majority Votes Approach detected more anomalies, mainly linked to early workflow indicators such as RXs-pending verification and staff workload measures like staff volume at the entry. This observation highlights the sensitivity of the entry-phase to prescription inflow and staff performance fluctuations. Conversely, Phase 2 focused on verification-related variables like RXs and pickups pending verification, identifying anomalies tied to bottlenecks in the final verification stage. The Conservative Approach across both phases produced fewer anomalies but with higher precision, reinforcing the reliability of consensus-based detection. Evaluation metrics, including the Silhouette Coefficient and IREOS Criterion, confirmed the models' effectiveness, with Isolation Forest excelling in identifying outliers broadly, while LOF and KNN captured more localized irregularities.

4.2. Feature Importance and Anomaly Insights

Using association mining and FPOF, we identify the key features influencing anomaly detection. In Phase 1, a high number of RXs-pending verification and a low patient visit frequency were significant predictors of anomalies, suggesting that bottlenecks at the entry stage and unfamiliar patients contribute to workflow disruptions. Similarly, in Phase 2, anomalies were strongly associated with high pending verification, pickup counts, and prolonged processing times.

Figure 2: Turnaround times of anomalies and non-anomalies in phases 1 and 2



As shown in Figure 2, the comparison of turnaround times revealed significant differences between anomalous and non-anomalous cases. In Phase 1, anomalous prescriptions had an average turnaround time of 453 hours compared to 356 hours for non-anomalous cases. Similarly, Entry-To-Verification times were longer in anomalous cases (343 hours) than in non-anomalous ones (226 hours). Welch's T-tests confirmed these differences were statistically significant ($p < 0.001$). In Phase 2, anomalous cases exhibited extreme delays, with turnaround times averaging 1184 hours compared to 297 hours for non-anomalous cases ($p < 0.001$).

5. Discussion and Conclusion

This study developed an ML-based, phase-based anomaly detection framework to improve workflow efficiency in community pharmacies. The system provided real-time insights into operational inefficiencies by integrating features specific to each stage of the prescription process—entry, verification, and pickup—as well as staff workload variables. The two-phase detection approach enabled precise identification of anomalies, particularly during peak hours and staff transitions, significantly reducing errors and delays in prescription processing.

Phase 1 (prescription entry) anomalies were linked to high pending verification counts and staff workload imbalances during prescription entry. In contrast, Phase 2 (prescription verification) anomalies were associated with bottlenecks in the final verification and pickup stages. Temporal factors, such as day-of-week patterns, also influenced anomaly occurrence. Our time-based analyses highlight a clear association between workflow anomalies and processing delays across both prescription entry and verification stages. Anomalous cases consistently exhibited prolonged processing times compared to non-anomalous ones, emphasizing the impact of inefficiencies on overall workflow performance. These findings reinforce the importance of real-time anomaly detection in identifying and addressing delays to improve pharmacy operations and patient care.

These findings offer actionable insights for optimizing pharmacy operations. Real-time anomaly alerts can help managers redistribute workloads, address verification backlogs promptly, and prioritize prescriptions at risk of delays. Balancing verification queues and monitoring staff workload fluctuations can prevent cascading inefficiencies, improving patient satisfaction and operational reliability.

References

[1] K. Gibson, "After pharmacists walk out, CVS vows to improve working conditions," https://www.cbsnews.com/news/cvs-pharmacy-strike-2023-working-conditions/, (Accessed on 09/07/2023).

[2] N. Goodkind, "Walgreens walkout: Your pharmacy might be closed next week," https://www.cnn.com/2023/10/06/business/walgreens-pharmacists-walkouts-protest-working-conditions/index.html, (Accessed on 09/07/2023).

[3] D. A. Mott, W. R. Doucette, C. A. Gaither, C. A. Pedersen, and J. C. Schommer, "Pharmacists' attitudes toward worklife: results from a national survey of pharmacists," *Journal of the American Pharmacists Association*, vol. 44, no. 3, pp. 326–336, 2004.

[4] J. C. Hayden and R. Parkin, "The challenges of covid-19 for community pharmacists and opportunities for the future," *Irish Journal of Psychological Medicine*, vol. 37, no. 3, pp. 198–203, 2020.

[5] NCPA, "2021 ncpa digest: A roadmap for independent community pharmacies," https://ncpa.org/newsroom/news-releases/2021/10/11/ncpa-releases-2021-digest-report, (Accessed on 09/07/2023).

[6] E. Magola, S. C. Willis, and E. I. Schafheutle, "Community pharmacists at transition to independent practice: Isolated, unsupported, and stressed," *Health & Social Care in the Community*, vol. 26, no. 6, pp. 849–859, 2018.

[7] C. M. Bookwalter, "Challenges in community pharmacy during covid-19: the perfect storm for personnel burnout," *US Pharm*, vol. 46, no. 5, pp. 28–31, 2021.

[8] S. G. Francis, "Pharmacists' perceptions about the effect of work environment factors on patient safety in large-chain retail pharmacies," *Journal of Pharmacy Technology*, vol. 38, no. 6, pp. 376–378, 2022.

[9] V. Snorovikhina and A. Zaytsev, "Unsupervised anomaly detection for discrete sequence healthcare data," 2021.

[10] O. K. Odukoya, J. A. Stone, and M. A. Chui, "How do community pharmacies recover from e-prescription errors?" *Research in Social and Administrative Pharmacy*, 2014.

[11] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.

[12] C. C. Aggarwal and C. C. Aggarwal, *An introduction to outlier analysis.* Springer, 2017.

[13] R. Gustriansyah, D. I. Sensuse, and A. Ramadhan, "Decision support system for inventory management in pharmacy using fuzzy analytic hierarchy process and sequential pattern analysis approach," in *2015 3rd International Conference on New Media (CONMEDIA).* IEEE, 2015, pp. 1–6.

[14] F. Silva-Aravena, I. Ceballos-Fuentealba, and E. Álvarez-Miranda, "Inventory management at a chilean hospital pharmacy: Case study of a dynamic decision-aid tool," *Mathematics*, vol. 8, no. 11, p. 1962, 2020.

[15] T. Awaya, K.-i. Ohtaki, T. Yamada, K. Yamamoto, T. Miyoshi, Y.-i. Itagaki, Y. Tasaki, N. Hayase, and K. Matsubara, "Automation in drug inventory management saves personnel time and budget," *Yakugaku zasshi*, vol. 125, no. 5, pp. 427–432, 2005.

[16] A. Hermansyah, E. Sainsbury, and I. Krass, "Multiple policy approaches in improving community pharmacy practice: the case in indonesia," *BMC Health Services Research*, vol. 18, pp. 1–14, 2018.

[17] U. Athiyah, C. D. Setiawan, G. Nugraheni, E. Zairina, W. Utami, and A. Hermansyah, "Assessment of pharmacists' knowledge, attitude and practice in chain community pharmacies towards their current function and performance in indonesia," *Pharmacy Practice (Granada)*, vol. 17, no. 3, 2019.

[18] V. G. Monakhov, V. Thambawita, P. Halvorsen, and M. A. Riegler, "Gridhtm: Grid-based hierarchical temporal memory for anomaly detection in videos," *Sensors*, 2023.

[19] R. A. Ariyaluran Habeeb, F. H. Nasaruddin, A. Gani, M. Hashem, E. Ahmed, and M. Imran, "Real-time big data processing for anomaly detection: A survey," *International Journal of Information Management*, 2019.

[20] C. P. Ngo, A. A. Winarto, C. K. Khor Li, S. J. Park, F. Akram, and H. K. Lee, "Fence gan: Towards better anomaly detection," 2019.

[21] G. Pang, C. Shen, L. Cao, and A. v. den Hengel, "Deep learning for anomaly detection," *ACM Computing Surveys*, 2021.

[22] C. Aytekin, X. Ni, F. Cricri, and E. Aksu, "Clustering and unsupervised anomaly detection with l_2 normalized deep auto-encoder representations," in *2018 International Joint Conference on Neural Networks (IJCNN).* IEEE, 2018, pp. 1–6.

[23] S. Han, X. Hu, H. Huang, M. Jiang, and Y. Zhao, "Adbench: Anomaly detection benchmark," *Advances in Neural Information Processing Systems*, vol. 35, pp. 32 142–32 159, 2022.