



**UGANDA CHRISTIAN
UNIVERSITY**

A Centre of Excellence in the Heart of Africa

**PREDICTING PHARMACY CLOSURES IN UGANDA USING MACHINE
LEARNING: A RISK-BASED APPROACH TO REGULATORY PRIORITIZATION**

**A DISSERTATION SUBMITTED TO THE FACULTY OF ENGINEERING,
DESIGN AND TECHNOLOGY IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE AWARD OF THE MASTER OF SCIENCE IN
DATA SCIENCE AND ANALYTICS OF UGANDA CHRISTIAN UNIVERSITY**

SUBMITTED BY:

MBABAZI FIONA

B26097, J24M19/009

SEPTEMBER 2026

Declaration

I, Mbabazi Fiona, hereby declare that this dissertation titled “**Predicting Pharmacy Closures in Uganda Using Machine Learning: A Risk-Based Approach to Regulatory Prioritization**” is my original work and has not been submitted, either in whole or in part, to any other university or institution of higher learning for the award of a degree or any other academic qualification.

Where the ideas, findings, or work of other authors have been used, they have been appropriately acknowledged and referenced in accordance with established academic conventions.

This dissertation is submitted in partial fulfilment of the requirements for the award of the **Master of Science in Data Science and Analytics of Uganda Christian University**.



.....

FIONA MBABAZI

Approval

This is to certify that the dissertation titled “**PREDICTING PHARMACY CLOSURES IN UGANDA USING MACHINE LEARNING: A RISK-BASED APPROACH TO REGULATORY PRIORITIZATION**” was prepared under my guidance and supervision.

I have reviewed the dissertation and consider it satisfactory and ready for submission in partial fulfilment of the requirements for the award of the **Master of Science in Data Science and Analytics of Uganda Christian University**.

Signature

Date



Mr. John Bosco Wabwire

September 20th 2026

Dedication

This dissertation is lovingly dedicated to my cherished husband, Katto Dickson, whose unwavering love, steadfast support, and constant encouragement have been a source of strength throughout this journey. Your presence in my life has been a profound blessing, sustaining me through every challenge and celebrating every milestone with me.

To our beloved children, whose innocence, joy, patience, and understanding have brought light and purpose even during the most demanding moments of this academic journey. May this work inspire you to pursue your dreams with dedication, resilience, faith, and courage.

Above all, I dedicate this achievement to God Almighty, whose grace, strength, and providence have made this journey possible. To Him be all the glory.

May this work stand as a testimony to the enduring power of faith, love, perseverance, and family.

Acknowledgment

I would like to express my sincere appreciation to my supervisor, Mr. John Bosco Wabwire, as well as Dr. Terhemba Michael-Ahile and Mr. Ian Raymond Osolo, for their invaluable guidance, support, and encouragement throughout the development and completion of this dissertation.

Their expertise, constructive feedback, mentorship, and commitment greatly contributed to the quality of this research and to my academic and professional growth. I am deeply grateful for the time, knowledge, and guidance they generously shared throughout this journey.

Abstract

Pharmacy closures can disrupt access to medicines and may indicate underlying regulatory or operational weaknesses within Uganda's pharmaceutical sector. Although the National Drug Authority routinely collects pharmacy licensing and inspection data, these data are largely used for reactive oversight rather than proactive prediction of closure risk.

This study identified factors associated with pharmacy closure and empirically evaluated supervised machine-learning models for predicting pharmacy closure in Uganda to support risk-based regulatory prioritization. A retrospective analytical design was used based on secondary regulatory data from licensed retail and wholesale pharmacies registered between 2016 and 2024. The analytical dataset comprised 19,636 inspection records from 5,149 unique pharmacies. Descriptive analysis characterized closure patterns by time, geography, license status, inspection outcomes, and pharmacy category. Exploratory analysis examined regulatory and inspection variables associated with closure, while Logistic Regression, Decision Tree, Random Forest, and Extreme Gradient Boosting (XGBoost) models were evaluated for predicting closure within 180 days after inspection.

Model development used a chronological, non-random train/test split, with the oldest 80% of inspection records used for training and the most recent 20% retained as an independent test set. Group-aware cross-validation was applied within the training data using pharmacy identifier to reduce information leakage arising from repeated inspections. Model performance was assessed using ROC-AUC, PR-AUC, precision, recall, F1-score, and Precision@k.

XGBoost achieved the strongest group-aware cross-validation performance, with a ROC-AUC of 0.989 and PR-AUC of 0.948. On the independent chronological test set, Random Forest achieved the highest ROC-AUC (0.907), while XGBoost achieved the highest PR-AUC (0.214), with corresponding ROC-AUC and PR-AUC values of 0.902 and 0.214, respectively. XGBoost also achieved the highest recall (98.0%) and a joint-highest F1-score (0.375) at the 0.5 classification threshold.

The findings demonstrate that routinely collected regulatory data contain predictive signals for pharmacy closure, although model performance varied across evaluation criteria and between cross-validation and chronological testing. Inspection history, license-related variables, and geographic characteristics contributed important predictive information. Overall, XGBoost provided the strongest combination of cross-validation and independent test performance and was selected as the preferred model for risk prediction. The findings suggest that machine-learning-derived risk scores could support targeted inspection planning and proactive regulatory prioritization, while model outputs should remain subject to professional regulatory judgement. Prospective validation, improved data quality, calibration, fairness assessment, and monitoring for model drift are recommended before operational deployment.

Table of Contents

Declaration.....	1
Approval.....	2
Dedication	3
Acknowledgment	4
Abstract	5
List of Figures	9
List of Tables	10
List of Abbreviations	11
Glossary of Terms and Acronyms	12
CHAPTER ONE	1
INTRODUCTION	1
1.1 Background	1
1.2 Problem Statement.....	2
1.3 General Objective	3
1.4 Justification of the Study.....	3
1.5 Significance of the Study	4
1.6 Scope of the Study	4
CHAPTER TWO.....	6
LITERATURE REVIEW	6
2.1 Introduction	6
2.2 Pharmacy Closures and Regulatory Oversight.....	6
2.2.1 Determinants of Pharmacy Closures	7
2.2.2 Risk-Based Regulatory Oversight	7
2.3 Traditional Regulatory Monitoring Systems and Their Limitations	8
2.4 Comparative Evaluation of Reviewed Research	8
2.5 Predictive Analytics and Supervised Machine Learning	10
2.6 Machine Learning Algorithms in Predictive Risk Modeling	10
2.6.1 Logistic Regression	11
2.6.2 Decision Trees.....	11
2.6.3 Random Forest	11
2.6.4 Gradient Boosting and XGBoost.....	11
2.7 Applications of Machine Learning in Regulatory Science	11
2.8 Early Warning Systems and Risk-Based Regulatory Prioritization.....	12
2.9 Class Imbalance and Evaluation Metrics in Predictive Modeling	12
2.10 Ethical and Governance Considerations	13

2.11	Knowledge Gaps in the Literature	13
2.12	Conclusion	14
CHAPTER THREE.....		15
RESEARCH METHODOLOGY		15
3.1	Study Design	15
3.2	Study Area and Population	15
3.3	Conceptual Framework for the Study.....	15
3.4	Data Sources and Collection	18
3.5	Data Pre-processing and Cleaning.....	19
3.5.1	Parsing and Normalization.....	19
3.5.2	Leakage-Safe Early-Warning Label	20
3.5.3	Feature Readiness (for later modelling)	20
3.5.4	Data Quality Issues Fixed	21
3.5.5	Descriptive Analysis	22
3.5.6	Exploratory Risk Factor Analysis.....	22
3.6	Model Development	23
3.6.1	Train/Test Protocol	25
3.6.2	Candidate Model Set and Theoretical Comparison.....	25
3.6.3	Selection of the Final Estimator	26
3.6.4	Model Deployment and Operationalization	27
3.7	Evaluation Metrics	27
3.8	Tools and Software	28
3.9	Risk Factors Identification.....	28
3.10	Identification of Key Risk Factors Using XGBoost	29
3.11	Ethical Consideration	29
3.12	Limitations of the Methodology	29
CHAPTER FOUR		31
PRESENTATION AND INTERPRETATION OF STUDY FINDINGS.....		31
4.1	Introduction.....	31
4.2	Data Summary and Target Label Construction	31
4.3	Features Available at Inspection Time	34
4.4	Train/Test Protocol (Time-Aware).....	35
4.5	Model Specification	36
4.6	Overall Predictive Performance (Ranking Focus)	37
4.7	Capacity-Constrained Evaluation (Top-k Prioritization).....	39
4.8	Threshold Selection and Probability Calibration	41
4.9	Error Analysis and Diagnostic Insights	42
4.10	Sensitivity and Robustness Checks	42

4.11	Policy Relevance and Practical Deployment	43
4.12	Key Risk Factors Contributing to Pharmacy Closures in Uganda	44
4.13	Feature Importance Analysis (SHAP Results)	48
4.14	Extraction of Early Warning Indicators for Identifying At-Risk Pharmacies	49
4.14.1	Predicted Risk score as a Primary Warning Indicator	50
4.14.2	Inspection Pattern Indicators	50
4.14.3	Days Since the last inspection.....	50
4.14.4	Geographic Vulnerability Indicators	50
4.15	Descriptive Analysis of the Pharmacy Closure Landscape in Uganda.....	51
CHAPTER FIVE		54
DISCUSSION, RECOMMENDATION, LIMITATIONS, STUDY CONTRIBUTION, FUTURE WORK AND CONCLUSION.....		54
5.1	Introduction	54
5.2	Discussion of Key Findings.....	54
5.2.1	Identification of Risk Factors Contributing to Pharmacy Closures.....	54
5.2.2	Model Performance and Practical Implications.....	55
5.2.3	Extraction of Early Warning Indicators	55
5.3	Practical interpretation	56
5.4	Recommendations	56
5.4.1	Recommendations for the National Drug Authority (NDA)	56
5.4.2	Recommendations for Pharmacy Owners	56
5.4.3	Recommendations for Policymakers	57
5.4.4	Recommendations for Future Research	57
5.5	Limitations.....	58
1.	Limited Availability and Quality of Historical Data	58
2.	Limited Availability of Predictor Variables	58
8.	Explainability Does Not Imply Causality.....	59
5.6	Study Contributions	59
5.6.1	Theoretical Contribution	59
5.6.2	Methodological Contribution	59
5.6.3	Practical Contribution	60
5.6.4	Policy Contribution	60
5.7	Future Work	60
5.8	Conclusion	61
REFERENCES		62
APPENDICES.....		68

List of Figures

Figure 3.1: Conceptual Analytical Framework for Predicting Pharmacy Closures in Uganda.....	17
Figure 3.2: Steps in involved in the Model Development.....	24
Figure 4.1: ROC curve with AUC = 0.902.....	38
Figure 4.3: Calibration plot (uncalibrated vs isotonic-calibrated).....	40
Figure 4.5: Bar plot used to show how the days_since_last_inspection influences the risk of closure.....	45
Figure 4.6: Bar plot used to show how the days to expiry influences the risk of closure ..	46
Figure 4.7: A summary plot highlighting how different geographic regions impact the risk of pharmacy closures.....	47
Figure 4.8: Trend in Regulated Pharmacy Population in Uganda (2016-2024).....	50

List of Tables

Table 2.1 : Comparative Evaluation of Reviewed Research.....	9
Table 3.1 : Data Sources, Contents, Expected Issues, and Planned Preprocessing.....	19
Table 3.2: Characteristics of Candidate Machine-Learning Algorithms	25
Table 4. 1: Dataset overview (rows, unique pharmacies, date span, missingness).....	31
Table 4.2: Label prevalence overall and by year/region.....	32
Table 4.3: Feature dictionary: name, type, description, transformation.....	34
Table 4.4: Precision@k on the held-out (time-ordered) test set after de-duplicating to one row per pharmacy.....	39
Table 4.5: Top-k summary by district (count, confirmed closures, precision).....	43
Table 4.6: Data Sources, Contents, Expected Issues, and Planned Preprocessing.....	48
Table 4.7: Annual Number and Proportion of Pharmacy Closures in Uganda.....	51
Table 4.8: Geographic Distribution of Pharmacy Closures by Region.....	52

List of Abbreviations

Term/Acronym	Definition
NDA	National Drug Authority
ML	Machine Learning
LMICs	Low Middle Income Countries
GPP	Good Pharmacy Practice
SD	Standard Deviation
XGBoost	Extreme Gradient Boosting
ROC-AUC	Receiver Operating Characteristic - Area Under Curve
WHO	World Health organization
FDA	U.S. Food and Drug Administration
SHAP	Shapley Additive Explanations

Glossary of Terms and Acronyms

Term / Acronym	Definition
National Drug Authority	A government organization in Uganda responsible for regulating human and veterinary medicines and other healthcare products.
Machine Learning	A subfield of artificial intelligence focused on building algorithms that improve automatically through data exposure.
Analytical Skills	The ability to evaluate information, identify patterns, and solve complex problems using logical reasoning and evidence.
LMICs	Low Middle Income Countries
GPP	Is a set of international standards for the provision of quality pharmaceutical services, focused on promoting the best use of drugs and improving patient and public health outcomes
Exploratory Data Analysis (EDA)	A technique to summarize the main characteristics of data using visual and statistical methods before formal modeling.
SD	Standard Deviation
Extreme Gradient Boosting	Is a highly efficient, portable, and scalable open-source software library for gradient boosting machine learning algorithms, used for classification and regression problems
ROC-AUC	Receiver Operating Characteristic - Area Under Curve
Natural Language Processing (NLP)	A branch of AI that enables machines to understand, interpret, and generate human language.
Time Series Analysis	A statistical method used to analyze sequences of data points over time to identify trends and patterns.
Premise/s	In NDA's Case can mean a Retail, Whole sale Pharmacy or a Pharmacy
Descriptive statistical analysis	is a statistical approach used to summarize, organize, and describe the characteristics of a dataset through measures such as frequencies, proportions, measures of central tendency (mean, median), measures of dispersion (standard deviation, interquartile range), tables, charts, and trend visualizations, with the aim of providing an overall understanding of the data and identifying observable patterns.
Exploratory risk factor analysis	Exploratory risk factor analysis is an analytical approach used to investigate potential associations between explanatory variables and a defined outcome before formal predictive or multivariable modelling. Its purpose is to

identify candidate risk factors, understand variable behavior, detect meaningful patterns, and inform subsequent model development or hypothesis generation.

Fisher's Exact Test

is a non-parametric statistical test used to examine whether the observed distribution between two categorical variables differs significantly from what would be expected by chance, particularly when sample sizes are small.

Crude association analysis

is a preliminary statistical assessment used to examine the direct relationship between an individual explanatory variable and an outcome of interest before multivariable adjustment. It provides an initial estimate of the strength and direction of association between variables.

CHAPTER ONE

INTRODUCTION

Pharmacies form a crucial part of Uganda's healthcare delivery system, serving as the primary source for dispensing essential medicines and offering pharmaceutical care to the public. The sustainability and operational continuity of these facilities are central to ensuring consistent access to quality health services. However, in recent years, Uganda has experienced intermittent closures of pharmacies due to various factors including non-compliance with regulatory requirements, financial insolvency, poor location decisions, and inadequate staffing [1]. These closures disrupt service provision, pose risks to patient safety, and strain the oversight capabilities of the National Drug Authority (NDA) and other regulatory stakeholders [2].

The current compliance monitoring framework in Uganda is primarily reactive, relying heavily on routine inspections and complaint-based investigations. This approach is limited in its ability to proactively identify pharmacies that are at risk of closure or those likely to fall out of compliance [3]. As the pharmaceutical sector continues to grow in complexity, there is an urgent need for innovative, data-driven solutions that can help regulatory authorities move from reactive enforcement to predictive and preventive regulation.

Machine learning (ML), a subset of artificial intelligence, offers powerful tools for pattern recognition, classification, and prediction based on large datasets. In regulatory sciences, ML has been successfully applied to forecast compliance risks, detect anomalies in supply chains, and optimize inspection schedules [4,5]. Leveraging historical licensing data, inspection records, and pharmacy variations, ML models can identify patterns that precede pharmacy closures and enable regulators to take timely, targeted actions.

This research aimed to develop a machine learning model tailored to the Ugandan context that can predict the likelihood of pharmacy closures and extract early warning indicators for use by regulators. By shifting from a retrospective to a predictive approach, the study seeks to enhance regulatory efficiency, minimize disruptions in pharmaceutical service delivery, and ultimately contribute to improved public health outcomes.

1.1 Background

Pharmaceutical outlets, including pharmacies and drug shops, play a central role in ensuring public access to medicines, particularly in low- and middle-income countries (LMICs) like Uganda. As frontline health facilities, they bridge the gap between formal healthcare systems and communities, often serving as the first point of contact for treatment and advice [5]. The regulation of pharmacies in Uganda is the responsibility of the National Drug Authority (NDA), which conducts routine inspections, licenses outlets, and ensures adherence to Good Pharmacy Practice (GPP) standards [5].

Despite efforts by the NDA to enforce compliance, several pharmacies continue to face closure due to failure to meet regulatory standards, including inadequate staffing, expired licenses, poor premises conditions, and repeated violations of inspection guidelines [3]. These closures not only indicate systemic weaknesses but also risk reducing the population's access to essential medicines, particularly in underserved regions [4].

According to NDA's 2022 compliance report, nearly 15% of registered pharmacies were either suspended or closed within a 24-month period due to non-compliance or voluntary business failure [2]. This trend has raised concerns about the effectiveness of existing monitoring frameworks, which are largely reactive and dependent on periodic inspections or complaints rather than real-time risk assessments [6].

Globally, regulatory bodies have begun leveraging digital technologies, including machine learning (ML), to improve decision-making in resource-constrained environments. ML models have been successfully applied in areas such as food and drug safety surveillance, clinical risk prediction, and fraud detection [7,8]. These models allow regulators to sift through large datasets, uncover patterns, and prioritize high-risk entities for targeted action, thereby improving efficiency and compliance outcomes.

In the Ugandan context, the availability of routine inspection data, licensing records, geospatial data, and violation histories presents a rich foundation for predictive analytics. However, such tools remain largely underutilized. The integration of ML into Uganda's pharmacy regulatory framework could significantly enhance the ability of NDA to predict pharmacy closures, identify early warning signs, and implement preventive strategies, contributing to a more resilient pharmaceutical system.

1.2 Problem Statement

The pharmacy sector in Uganda plays a vital role in ensuring access to quality medicines and pharmaceutical services, especially at the community level. However, in recent years, the country has witnessed a steady rise in pharmacy closures, particularly in rural and underserved areas. These closures not only disrupt medicine supply chains but also undermine public health efforts and trust in the healthcare system [1]. Several factors, including economic constraints, regulatory non-compliance, and inadequate operational management, have contributed to this growing challenge. The current regulatory framework for monitoring pharmacies is largely reactive, relying on manual and periodic inspections that intervene only after violations or closures occur. This reactive approach fails to prevent closures before they happen and leads to inefficient resource allocation in regulatory activities. Furthermore, even though a wealth of inspection and licensing data exists, there is no predictive system in place that can proactively identify at-risk pharmacies before they close.

The core problem lies in the lack of a proactive mechanism to identify pharmacies at risk of closure before regulatory or operational failures occur. Although regulatory bodies collect extensive data on licensing, inspections, and compliance, this information is not currently used for predictive analysis. Areas that need to be investigated include the key risk factors influencing pharmacy sustainability, patterns and trends within regulatory data, and the potential of data analytics and machine learning models to predict closures or non-compliance early enough for intervention.

The solution to this problem lies in the development of a data-driven, machine learning model that can predict pharmacy closures before they occur, based on historical inspection, licensing, and operational data. By using machine learning to analyze these factors, the model can identify high-risk pharmacies early, allowing for proactive regulatory interventions and improving resource allocation [3]. This approach would allow regulators to target inspections more effectively, reduce the likelihood of closures, and enhance the overall compliance of pharmacies across Uganda. Therefore, this study seeks to develop and evaluate machine learning models for predicting pharmacy closures in Uganda using historical regulatory and inspection data, with the aim of supporting risk-based regulatory prioritization.

1.3 General Objective

To identify and evaluate existing supervised machine learning models for predicting pharmacy closures in Uganda to support risk-based regulatory prioritization

1.3.1 *The specific objectives*

- a) To identify key risk factors contributing to pharmacy closures in Uganda.
- b) To identify a suitable machine learning model that can accurately predict pharmacy closures using the available pharmacy regulatory data.
- c) To assess the applicability of machine learning-based predictions in supporting risk-based regulatory prioritization of pharmacies in Uganda.

1.3.2 *The research questions*

- a) What are the key risk factors contributing to pharmacy closures in Uganda?
- b) Which selected supervised machine learning model best predicts pharmacy closures using available pharmacy regulatory data?
- c) How can machine learning-based predictions support risk-based regulatory prioritization of pharmacies in Uganda?

1.4 Justification of the Study

Pharmacy closures disrupt access to essential medicines and undermine the continuity of healthcare services, especially in underserved areas of Uganda. Despite the availability of routine inspection and licensing data, current regulatory practices are largely reactive, addressing non-compliance only after it has occurred. This limits the ability of the National Drug Authority (NDA) to take timely, preventive action and efficiently allocate inspection resources [7][9].

The increasing availability of regulatory data presents an opportunity to adopt machine learning as a tool for early risk detection. Predictive analytics can help uncover hidden patterns and generate actionable insights, allowing regulators to identify high-risk pharmacies before closure occurs [3]. This study is therefore justified as it seeks to fill a critical gap by developing a locally informed, data-driven model to support proactive regulatory decision-making.

Moreover, the research contributes to the growing field of regulatory science and demonstrates the potential of artificial intelligence to strengthen public health systems. The outcomes of this study will not only enhance regulatory efficiency in Uganda but may also inform similar interventions in other low- and middle-income countries facing comparable challenges.

1.5 Significance of the Study

Pharmacy closures, especially in underserved and rural areas, disrupt the supply of essential medicines and weaken the resilience of Uganda's healthcare system. By leveraging machine learning to proactively identify pharmacies at risk of closure, this study will contribute significantly to strengthening pharmaceutical regulatory systems and safeguarding public health.

The outcomes of this study will inform evidence-based policies and inspection prioritization by the National Drug Authority (NDA), enabling the transition from reactive to preventive regulatory oversight. This will help optimize the use of limited inspection resources, reduce regulatory blind spots, and promote continuous access to quality pharmaceutical services [4].

Furthermore, the study sets a precedent for applying artificial intelligence to solve public health regulatory challenges in low- and middle-income countries (LMICs). The predictive model developed could serve as a reference framework for other countries facing similar constraints and pave the way for smarter, data-driven regulation in Africa and beyond.

1.6 Scope of the Study

This study focused on the application of existing supervised machine learning models to predict pharmacy closures in Uganda and support risk-based regulatory prioritization. The study used secondary data obtained from the National Drug Authority, including pharmacy licensing records, inspection records, closure histories, compliance status, and geographic location data covering the period 2016 to 2024.

The study was limited to licensed retail and wholesale pharmacies registered with the National Drug Authority during the study period. It excluded unlicensed drug outlets, informal medicine vendors, and other health product establishments that are not classified as pharmacies. The machine learning models were used for classification and prediction purposes only, and the study does not involve real-time deployment, automation, or integration into the National Drug Authority's regulatory information systems [1].

The study examined regulatory, operational, and inspection-related factors associated with pharmacy closures. These included variables available in the regulatory dataset, such as inspection frequency, licensing status, days to license expiry, location, closure history, and compliance status [9]. However, the study didn't assess internal business management practices, financial performance, ownership decisions, or other operational factors unless such information is already captured in the available regulatory data.

Geographically, the study was limited to Uganda and did not compare Uganda's regulatory system with those of other countries. The focus was majorly on how machine learning-based predictions may support risk-based regulatory prioritization within Uganda's pharmacy regulatory context.

CHAPTER TWO LITERATURE REVIEW

2.1 Introduction

This chapter reviews literature related to pharmacy closures, regulatory oversight, predictive analytics, and the application of supervised machine learning in risk prediction and regulatory decision-making. The review examined the causes and implications of pharmacy closures, limitations of traditional inspection systems, applications of machine learning in regulatory science, predictive risk modeling approaches, early warning systems, and ethical considerations surrounding machine learning deployment in regulatory environments [9]. The chapter also identified existing knowledge gaps that justify the current study.

2.2 Pharmacy Closures and Regulatory Oversight

Pharmacies play an essential role in healthcare delivery systems by ensuring access to medicines, pharmaceutical counseling, and continuity of care. In low- and middle-income countries (LMICs), pharmacies are often the first point of healthcare access, particularly in underserved communities where healthcare infrastructure remains limited. Closure of pharmacies may therefore negatively affect medicine availability, treatment continuity, and public health outcomes [1][2].

Several studies have reported increasing pharmacy closures globally due to regulatory non-compliance, financial instability, operational inefficiencies, workforce shortages, and market pressures [3]. Studies from the United States have shown that pharmacy closures disproportionately affect vulnerable and rural populations, thereby worsening healthcare inequities. In many settings, pharmacy closures are associated with inadequate staffing, poor compliance with licensing requirements, weak record management, and failure to meet regulatory standards [10].

In Uganda, the National Drug Authority regulates pharmacies through licensing, inspection, and enforcement mechanisms intended to ensure compliance with Good Pharmacy Practice standards (GPP) and other legal requirements [6]. Despite these efforts, pharmacy closures continue to occur due to factors such as expired licenses, non-compliance with inspection standards, absence of qualified personnel, unsuitable premises, and repeated regulatory violations [3][5]. These challenges are compounded by resource limitations affecting regulatory agencies, including limited inspection personnel, increasing numbers of regulated facilities, and reliance on manual monitoring systems.

Pharmacy closures may therefore serve as an important regulatory outcome reflecting broader operational and compliance risks within the pharmaceutical sector. Understanding the factors associated with closure is essential for supporting proactive regulatory oversight and risk-based prioritization.

2.2.1 Determinants of Pharmacy Closures

Pharmacy closures are influenced by a combination of regulatory, operational, structural, and market-related factors. Regulatory non-compliance remains one of the most commonly cited drivers of closure, particularly where pharmacies fail to meet licensing requirements, staffing standards, premises requirements, or Good Pharmacy Practice standards. Repeated inspection deficiencies, failure to renew licenses, absence of qualified pharmaceutical personnel, poor record management, and non-adherence to regulatory conditions may trigger enforcement actions leading to suspension or closure [10].

Operational and structural factors may also contribute to pharmacy closure. These include prolonged inactivity, poor location viability, inadequate infrastructure, weak inventory management, and inability to sustain operational standards. In some contexts, financial instability, ownership transitions, competitive market pressures, and declining customer volumes have also been associated with pharmacy discontinuation. However, many of these factors are often not directly observable within regulatory datasets, creating a reliance on measurable proxy indicators such as inspection frequency, licensing status, and compliance history.

Geographic factors may also influence closure patterns, particularly in settings where pharmacies in underserved or remote regions experience greater operational challenges. Closure risk may therefore reflect not only immediate regulatory non-compliance but also broader system-level vulnerabilities affecting pharmacy sustainability. Understanding these determinants is essential for identifying predictors that may support proactive regulatory intervention [14].

2.2.2 Risk-Based Regulatory Oversight

Modern regulatory systems increasingly emphasize risk-based oversight as a means of improving regulatory efficiency and effectiveness. Risk-based regulation refers to the allocation of inspection, monitoring, and enforcement resources according to the estimated level of risk posed by regulated entities rather than applying uniform oversight approaches [3]. This approach recognizes that regulatory resources are finite and that targeted interventions are often more effective than blanket inspection schedules.

International regulatory bodies have increasingly adopted risk-based approaches in medicines regulation. The U.S. Food and Drug Administration employs risk-based site selection models for inspection prioritization, while the World Health Organization and OECD advocate proportionate, data-informed regulatory oversight. Risk-based systems rely on structured assessment of risk indicators, historical performance, and predictive signals to identify entities requiring enhanced regulatory attention [5].

Within pharmacy regulation, risk-based prioritization may support more efficient inspection scheduling by identifying facilities that demonstrate elevated likelihood of adverse outcomes such as closure, repeated non-compliance, or operational failure. However, effective implementation requires analytical tools capable of transforming historical regulatory data into actionable risk intelligence. Machine learning-based prediction models provide a potential mechanism for strengthening such regulatory prioritization [3].

2.3 Traditional Regulatory Monitoring Systems and Their Limitations

Traditional regulatory monitoring systems in many LMICs rely heavily on routine inspections, complaint-based investigations, and periodic licensing reviews. While these approaches remain important for regulatory enforcement, they are often reactive, resource-intensive, and limited in their ability to identify emerging risks before adverse outcomes occur [5].

Routine inspection systems generally apply standardized schedules across facilities regardless of individual risk levels. As the number of regulated facilities increases, such approaches may become inefficient because limited regulatory resources are distributed uniformly rather than according to risk [16]. Consequently, high-risk facilities may remain unidentified until significant violations or closures occur.

Risk-based inspection frameworks have increasingly been recommended in modern regulatory systems. These frameworks prioritize regulatory attention according to estimated risk profiles rather than treating all entities equally [5][7]. Risk-based approaches have been adopted in pharmaceutical manufacturing regulation, food safety oversight, and healthcare facility inspections in several high-income countries. However, effective implementation of risk-based regulation depends on the availability of reliable analytical tools capable of identifying high-risk entities using historical data.

In Uganda, pharmacy inspection systems largely depend on historical compliance history and manual regulatory assessment. Although the regulatory authority collects substantial inspection and licensing data, these datasets are not routinely leveraged using predictive analytical methods capable of detecting complex patterns or emerging closure risks. This creates a gap between data collection and actionable regulatory intelligence [18].

2.4 Comparative Evaluation of Reviewed Research

Table 2.1 below compares the reviewed studies in relation to Uganda's regulatory environment, emphasizing methodological approaches, data requirements, and how each source informs risk-based inspection planning and inspector training.

Table 2.1: Comparative Evaluation of Reviewed Research

Study / Source (Year, Region)	Focus	Data & Methods	Key Findings / Claims	Strengths	Limitations / Notes	Relevance to the Study
Mattingly et al., <i>JAMA Health Forum</i> (2025, US)	Community-pharmacy openings/closures & access	Retrospective analysis of U.S. market churn by ownership type	Closures can reduce short-term medication use; long term access/equity effects still uncertain	Peer-reviewed, current; nuanced view beyond “net closures”	US-centric; limited LMIC external validity	Frames access risk & outcome metrics your model should monitor post-deployment
AP summary of Health Affairs study (2024, US) + KQED explainer (2024)	National closure trend magnitude	Secondary synthesis of national datasets (2010-2021)	~29% of ~89k retail pharmacies closed by 2021 (~26k); drivers include reimbursement pressure	Clear headline prevalence; accessible methodology notes	Media summaries; underlying technical paper not directly cited	Justifies urgency of risk-based oversight; provides benchmark prevalence context
PCMA synthesis of Mattingly (2025, US)[19]	Industry interpretation of JAMA findings	Policy blog summarizing same dataset	Argues “independents open as often as they close” (market churn)	Offers counter-narrative; useful for balance	Advocacy source; potential bias	Reminds you to test for regional heterogeneity before generalizing “closure crisis” claims
FDA CDER MAPP 5014.1 rev.1 (2025, US)[18]	Risk-based Site Selection Model (SSM) for inspections	Policy manual; multi-factor risk scoring to prioritize sites	Rank-orders sites using quality risk; generates surveillance lists	Mature, transparent risk-based approach; implementable logic	Focused on manufacturers, not retail pharmacies	Template for your pharmacy risk-scoring pipeline and governance artifacts
OECD <i>Regulatory Enforcement & Inspections Toolkit</i> (2018, global; 2025 online PDF)	Principles & checklist for proportionate, targeted inspections	Normative framework (12-criterion checklist)	Emphasizes planning, targeting, communication, integrity	Cross-sector, globally recognized	Non-sector-specific	Provides governance & ethics yardsticks to NDA deployment
WHO/PAHO Good Regulatory Practices (2022 web ed.; 2016 base) [20]	GRP principles for NRAs	Guidance document	Calls for consistent procedures, transparency, reliance, independence	Health-sector-specific; widely adopted	High-level (implementation details needed locally)	Aligns your methods with GRP and NRA benchmarking expectations

The comparative review demonstrates that existing literature provides valuable insights into pharmacy closures, regulatory risk-based oversight, and predictive analytics; however, important methodological and contextual gaps remain. Studies examining pharmacy closures largely emphasize access to care, service disruption, and market dynamics, particularly in high-income settings such as the United States. While these studies establish the public health significance of pharmacy closures, their findings may not be directly transferable to Uganda due to differences in regulatory structures, pharmacy ownership models, healthcare financing, and market conditions [60].

The reviewed regulatory literature strongly supports the adoption of risk-based inspection and prioritization frameworks. Guidance from institutions such as the FDA, WHO, and OECD consistently emphasizes targeted oversight based on risk profiling rather than uniform inspection approaches [71]. However, these frameworks are largely conceptual or operational policy tools and provide limited empirical guidance on predictive implementation using

machine learning within pharmacy regulatory environments.

The predictive analytics literature demonstrates that machine learning approaches are increasingly effective in classification and early warning applications, particularly where large structured datasets exist. These studies provide methodological lessons relevant to this research, including the use of supervised classification models, probability-based risk scoring, handling of imbalanced datasets, and evaluation through performance metrics such as precision, recall, F1-score, and ROC-AUC. However, most such applications are concentrated in healthcare outcome prediction, financial risk modeling, or manufacturing oversight rather than pharmacy closure prediction.

A notable gap emerging from the reviewed literature is the limited application of predictive machine learning models within pharmacy regulatory systems in low- and middle-income countries, particularly sub-Saharan Africa. Existing studies rarely combine longitudinal regulatory inspection data, licensing histories, and closure outcomes to generate actionable early warning models for regulatory decision-making. This study therefore borrows the principle of risk-based regulatory prioritization from global regulatory frameworks and the methodological strengths of supervised machine learning from predictive analytics literature, while addressing the contextual gap of pharmacy closure prediction within Uganda's regulatory environment [71].

2.5 Predictive Analytics and Supervised Machine Learning

Predictive analytics refers to the use of historical data, statistical techniques, and computational algorithms to estimate the likelihood of future events. Advances in Machine Learning have significantly expanded predictive analytics capabilities across healthcare, finance, manufacturing, and regulatory science [14].

Supervised machine learning is particularly useful for prediction problems involving labeled historical outcomes. In supervised learning, algorithms learn relationships between predictor variables and known outcomes in order to classify or predict future observations. Supervised learning models are commonly applied in classification tasks where outcomes exist in discrete categories such as diseased/non-diseased, fraudulent/non-fraudulent, compliant/non-compliant, or active/closed [25].

Pharmacy closure prediction may therefore be conceptualized as a binary classification problem in which historical inspection and licensing variables are used to estimate the probability that a pharmacy will close within a specified future period.

Compared with traditional statistical approaches, machine learning models are capable of handling large datasets, non-linear relationships, complex feature interactions, and heterogeneous variables. These characteristics make machine learning particularly suitable for regulatory datasets that contain mixed operational, inspection, geographic, and compliance variables.

2.6 Machine Learning Algorithms in Predictive Risk Modeling

Several supervised machine learning algorithms have been applied in predictive risk modeling.

2.6.1 Logistic Regression

Logistic Regression is among the most widely used classification techniques in healthcare and regulatory research [10]. The model estimates the probability of occurrence of a binary outcome using predictor variables. Logistic Regression is valued for its interpretability, computational efficiency, and ease of implementation. However, its predictive performance may decline when relationships between predictors and outcomes are highly non-linear or involve complex interactions.

2.6.2 Decision Trees

Decision Trees classify observations through recursive partitioning of predictor variables into decision nodes. Tree-based approaches are intuitive and easy to interpret because they visually represent decision rules. However, single decision trees are prone to instability and overfitting, particularly in noisy datasets [61].

2.6.3 Random Forest

Random Forest is an ensemble learning method that combines multiple decision trees to improve predictive performance and reduce overfitting. The method has demonstrated strong performance in healthcare risk prediction and fraud detection applications. Random Forest models effectively handle non-linear relationships and variable interactions while maintaining robustness against noisy data.

2.6.4 Gradient Boosting and XGBoost

Gradient Boosting algorithms iteratively improve prediction accuracy by correcting errors generated by previous models. XGBoost, an optimized implementation of gradient boosting, has gained widespread use due to its computational efficiency, regularization mechanisms, and strong predictive performance in structured tabular datasets [37].

XGBoost performs particularly well in classification problems involving heterogeneous predictor variables, missing data, non-linear relationships, and class imbalance. The algorithm also produces probability-based outputs useful for risk ranking and prioritization. Because regulatory datasets frequently contain mixed operational and inspection variables, XGBoost has become increasingly relevant in predictive regulatory applications.

2.7 Applications of Machine Learning in Regulatory Science

Machine learning has increasingly been applied within regulatory science to support surveillance, inspection prioritization, fraud detection, and risk assessment. Regulatory agencies worldwide are exploring predictive analytics to improve efficiency and allocate limited resources more effectively. [12,13].

Applications of machine learning in healthcare regulation include predicting inspection outcomes, identifying high-risk manufacturers, detecting suspicious transactions, monitoring adverse drug events, and prioritizing regulatory interventions. Risk-based inspection models used by regulatory agencies increasingly rely on predictive scoring systems derived from historical inspection data [9].

The U.S. Food and Drug Administration has adopted risk-based site selection approaches for prioritizing pharmaceutical manufacturing inspections. Similarly, international regulatory frameworks increasingly emphasize data-driven oversight and risk-based inspection planning [19] [25].

Despite global advances, most machine learning applications in regulatory science originate from high-income settings with advanced digital infrastructures and standardized datasets. There remains limited evidence regarding the use of predictive analytics within pharmacy regulation in sub-Saharan Africa.

2.8 Early Warning Systems and Risk-Based Regulatory Prioritization

Early warning systems are analytical frameworks designed to identify entities at elevated risk before adverse outcomes occur. Such systems enable institutions to intervene proactively rather than reactively.

Machine learning-based early warning systems have been widely applied in healthcare for predicting patient deterioration, hospital readmission, disease outbreaks, and treatment failure. In financial systems, predictive models are used to identify customers at risk of loan default or bankruptcy [3]. Similar approaches are increasingly being explored in regulatory environments to support risk-based oversight.

Risk-based regulatory prioritization involves allocating inspection and monitoring resources according to estimated risk levels. Under this approach, high-risk facilities receive more frequent or targeted regulatory attention, while lower-risk facilities may receive less intensive monitoring [5]. Such systems improve regulatory efficiency by focusing limited resources on entities most likely to experience adverse outcomes.

In the context of pharmacy regulation, predictive risk scores may help identify pharmacies requiring closer inspection, compliance follow-up, or enhanced surveillance. Predictive models may therefore support more proactive and efficient regulatory oversight [20].

2.9 Class Imbalance and Evaluation Metrics in Predictive Modeling

An important challenge in predictive risk modeling is class imbalance, where adverse events occur far less frequently than non-events [3]. In pharmacy closure prediction, closure events may constitute only a small proportion of all inspection records.

Under imbalanced conditions, overall model accuracy may become misleading because a model can achieve high accuracy while failing to identify high-risk cases. Consequently, predictive modeling studies increasingly rely on additional evaluation metrics such as precision, recall, F1-score, Receiver Operating Characteristic Area Under the Curve (ROC-AUC), and Precision-Recall Area Under the Curve (PR-AUC) [41].

In regulatory contexts, recall is particularly important because failure to identify high-risk pharmacies may lead to missed opportunities for regulatory intervention. Probability-based risk scoring approaches therefore provide greater operational value than simple binary classifications.

2.10 Ethical and Governance Considerations

The use of machine learning in regulatory systems raises important ethical and governance considerations. Predictive models may influence inspection prioritization, enforcement decisions, and regulatory actions affecting regulated entities. Consequently, transparency, accountability, and fairness remain critical.

Regulators must ensure that predictive models do not unfairly disadvantage facilities based on geographic, socioeconomic, or demographic factors. Model outputs should therefore support human decision-making rather than replace regulatory judgment entirely.

Data governance is also important because regulatory datasets may contain incomplete, inconsistent, or biased information. Ethical deployment of predictive systems requires appropriate validation, documentation, and continuous monitoring to ensure reliability and fairness.

2.11 Knowledge Gaps in the Literature

Existing literature demonstrates growing interest in pharmacy closures, regulatory oversight, and machine learning applications independently. However, several important gaps remain.

First, few studies have specifically examined pharmacy closure prediction using supervised machine learning approaches. Most pharmacy closure studies focus on access to care, workforce distribution, or economic drivers rather than predictive risk modeling [13].

Second, although machine learning has been widely applied in healthcare prediction and regulatory science, limited empirical evidence exists regarding its application to pharmacy regulatory datasets in LMICs, particularly within sub-Saharan Africa.

Third, existing regulatory oversight systems in Uganda remain largely reactive and dependent on manual inspection processes despite the availability of historical regulatory data. There is limited evidence on how machine learning-derived risk predictions may support proactive, risk-based regulatory prioritization within Uganda's pharmacy regulatory context [10].

Finally, little research has explored the use of inspection-time variables and early warning indicators for predicting pharmacy closure risk using longitudinal regulatory datasets [10].

These gaps justify the need for this study, which seeks to evaluate supervised machine learning models for predicting pharmacy closures in Uganda using historical regulatory data.

2.12 Conclusion

The reviewed literature demonstrates that pharmacy closures pose important public health and regulatory challenges, particularly in resource-constrained settings. Traditional inspection systems remain largely reactive and resource-intensive, limiting the ability of regulators to identify emerging risks proactively [16].

Advances in machine learning and predictive analytics provide opportunities to develop early warning systems capable of supporting risk-based regulatory prioritization. Supervised machine learning models such as Logistic Regression, Random Forest, and XGBoost have demonstrated strong performance in predictive classification tasks involving structured datasets and complex variable interactions.

However, limited evidence exists regarding the application of machine learning to pharmacy closure prediction within sub-Saharan African regulatory settings. Existing literature also reveals limited use of historical pharmacy inspection data for predictive regulatory intelligence in Uganda.

The literature therefore supports the feasibility and relevance of applying supervised machine learning models to predict pharmacy closures and support risk-based regulatory prioritization in Uganda [19].

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Study Design

This study adopted a quantitative, predictive, observational design that applied and compared existing supervised machine-learning (ML) methods to historical regulatory data from Uganda's National Drug Authority (NDA). The aim was to:

- i. Identify risk factors associated with pharmacy closures
- ii. Find a suitable machine learning model that can accurately predict pharmacy closures using the available pharmacy regulatory data.

Machine learning was particularly suitable for this study because it was capable of modeling complex, non-linear relationships in the data—relationships that traditional statistical methods may not effectively capture. Moreover, ML enabled the handling of large datasets and revealed hidden patterns or correlations in the regulatory data that were not immediately obvious. Supervised learning was employed due to its effectiveness in classification tasks, particularly for predicting binary outcomes, such as whether a pharmacy would close or remain operational [3].

This section also outlines the approach for identifying key risk factors contributing to pharmacy closures in Uganda and how a machine learning model will be used to identify these factors. We will begin by analyzing available data from inspections, pharmacy licensing, and operational records, followed by applying machine learning techniques to extract patterns and identify the most important risk factors.

3.2 Study Area and Population

The study focused on pharmacies (both retail and wholesale) registered with the National Drug Authority (NDA) in Uganda. The study population included all licensed pharmacies whose operational status, inspection history, and regulatory compliance data were documented between 2016 and 2024. This population was chosen to ensure a sufficient number of data points for training the machine learning models and to capture a representative range of closure events over the years.

3.3 Conceptual Framework for the Study

This study was guided by a conceptual analytical framework that linked regulatory and operational pharmacy characteristics to closure outcomes through descriptive analysis, exploratory risk assessment, and predictive modelling. The framework recognizes pharmacy closure as a measurable regulatory outcome influenced by observable factors captured within the National Drug Authority's routine regulatory datasets.

At the input level, the study considered pharmacy-level regulatory and operational characteristics available from historical records, including licensing status, inspection history, operational activity status, compliance outcomes, geographic location, and time-dependent

indicators such as days to license expiry and time since previous inspection. These variables represent the information routinely available to regulators during oversight activities.

The analytical process involved three complementary stages. First, descriptive analysis was used to characterize the pharmacy closure landscape in Uganda, including the magnitude, temporal distribution, geographic spread, and regulatory characteristics of closure events. Second, exploratory risk factor analysis examined associations between candidate predictors and pharmacy closures to identify variables of potential regulatory significance and early warning value. Third, supervised machine learning modelling was applied to evaluate predictive relationships and generate closure risk scores based on historical patterns.

The output of the analytical framework was a predictive risk classification indicating the likelihood of pharmacy closure within the defined prediction horizon. These outputs were then conceptually linked to regulatory decision-making through risk-based prioritization, where pharmacies identified as higher risk could be targeted for enhanced surveillance, inspection, or early intervention.

This framework positioned the study as both a predictive analytics exercise and a regulatory intelligence decision-support investigation.

The conceptual analytical framework in Figure 3.1 illustrates the relationship between regulatory and operational pharmacy characteristics, analytical processes, predictive outputs, and regulatory decision-making. Historical pharmacy regulatory data serve as the study inputs, from which descriptive analysis and exploratory risk factor analysis are used to characterize closure patterns and identify candidate predictors. These variables inform supervised machine learning-based predictive modelling, which generates closure risk classifications. The resulting predictions are conceptually linked to risk-based regulatory prioritization, enabling targeted oversight of pharmacies at elevated risk of closure.

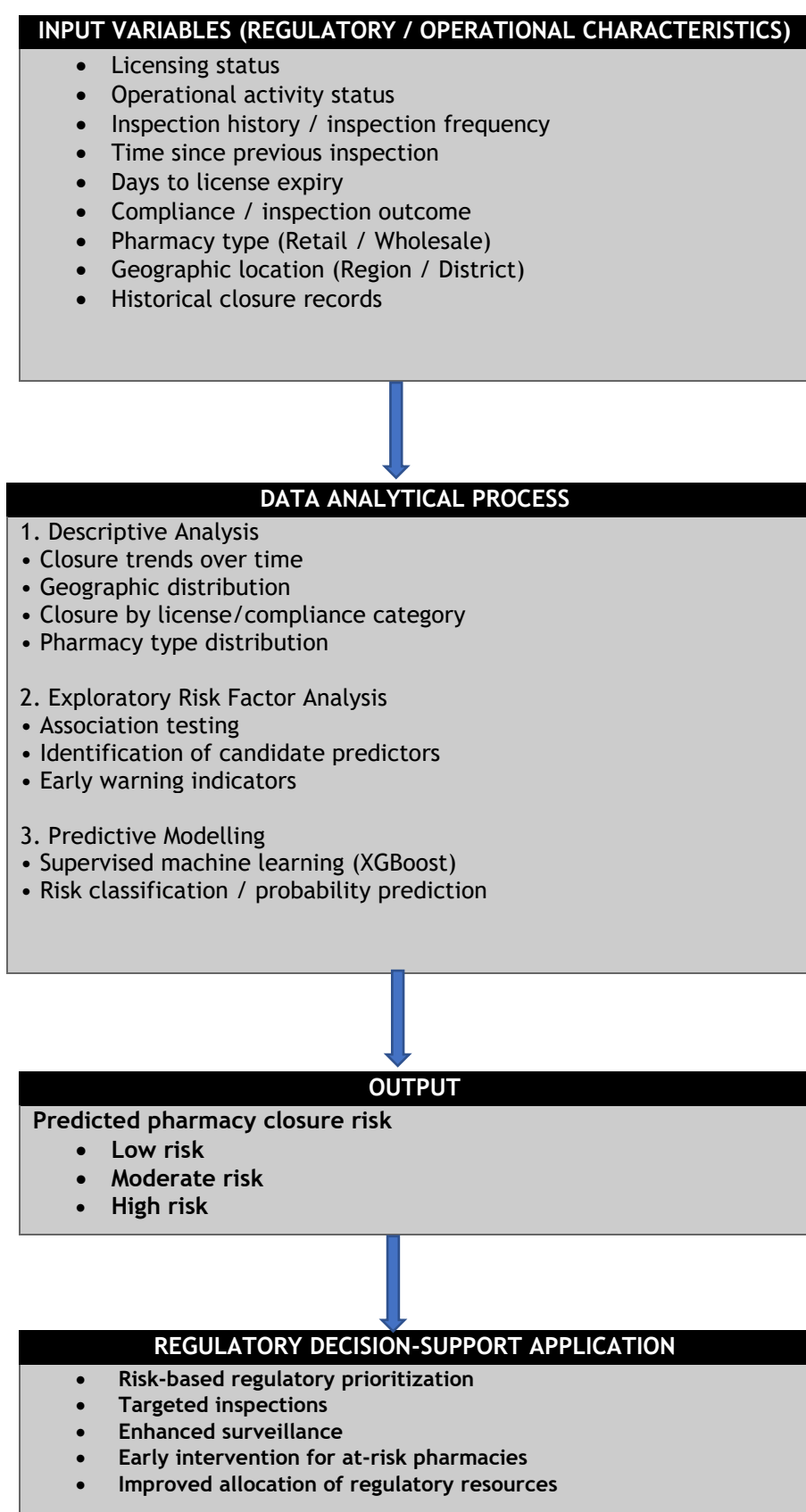


Figure 3.1: Conceptual Analytical Framework for Predicting Pharmacy Closures in Uganda

3.4 Data Sources and Collection

Secondary data was obtained from the NDA’s regulatory information systems. The key datasets included:

- **Pharmacy licensing data** (Premise numbers, status, issue/expiry dates, premise type, ownership, region/district)
- **Inspection records** (e.g., frequency, outcomes, violations, classification)
- **Compliance data** (administrative outcomes and dates)
- **Closure status** (closed, suspended, active pharmacies)
- **Geographic and demographic data** region/district and location type (where available)

Data collection involved submitting official requests and obtaining the necessary authorizations to access anonymized datasets for research purposes, in full compliance with the NDA’s data protection laws and confidentiality guidelines.

Data access and reproducibility. The full dataset cannot be publicly shared due to confidentiality restrictions and NDA data governance requirements. However, the complete modelling workflow—covering preprocessing, feature engineering, model training, evaluation, and interpretability scripts—has been archived in a version-controlled repository to enable reproducibility of the analysis pipeline. The GitHub Repository is <https://github.com/fiyona/PREDICTING-PHARMACY-CLOSURES-AND-ENHANCING-REGULATORY-COMPLIANCE-/tree/main>.

The repository contains the Jupyter notebook, supporting scripts, and documentation necessary to reproduce all reported figures and tables using the NDA-approved dataset extract.

To improve transparency and reproducibility, Table 3.1 enumerates each raw dataset used in this study, the specific variables drawn from each source, anticipated data-quality issues, and the exact preprocessing actions taken prior to modeling. This structure follows recognized data-management guidance and ensures that derived features, labels, and exclusions are auditable.

Table 3.1: Data Sources, Contents, Expected Issues, and Planned Preprocessing

Source dataset (owner)	What it contains	Issues to expect	Planned preprocessing (exact actions)
Licensing Register (NDA)	Active/expired/suspended pharmacy licenses [66]	Duplicate Premise numbers; missing expiry dates; inconsistent premise names; outdated status	De-duplicate on (Premise numbers); Convert date fields to proper datetime formats. Impute missing expiry via renewal history where possible; normalize categorical labels; derive <code>license_age</code> and <code>days_to_expiry</code>
Inspection Reports	Visit dates, findings, observations (major/minor/critical), staffing	Multi-row per visit; free-text findings; inconsistent severity labels; missing CAPA closure	Collapse to visit-level; map free-text to standardized <code>finding_codes</code> ; engineer <code>days_since_last_inspection</code> , <code>inspection_count</code> , <code>major_findings_count</code> ; flag <code>qualified_personnel_absent</code>
Closure / Suspension Register (NDA Enforcement)	Administrative outcomes used as labels	Multiple decisions per outlet; re-open events; missing reason codes	Sort chronologically; create final <code>closure_status</code> within study horizon; keep last known status; derive <code>time_to_closure</code> (if used for survival sensitivity)

3.5 Data Pre-processing and Cleaning

All data preparation steps were performed prior to model fitting to ensure data validity, comparability across records, and leakage-free outcome labels. Data manipulation and cleaning were implemented in Python using the pandas and numpy libraries. The raw regulatory dataset was first inspected for duplicate records, missing values, inconsistent date formats, and irregular coding of categorical variables such as premise status, active status, inspection outcome, region, district, and license status.

3.5.1 Parsing and Normalization

All parsing and normalization were implemented in Python using pandas for data manipulation and date/time handling, and NumPy for vectorized conditional logic.

- **Date normalization** All date fields—`inspectiondate`, `expirydate`, `createddate`, `applicationdate`—were coerced to `datetime64[ns]` (day-first), invalid strings set to `NaT` to prevent downstream errors.
- **Text normalization.** Categorical status fields (`premisestatus`, `activestatus`) were standardized by trimming whitespace and converting to lowercase using `Series.apply()` to remove case/spacing inconsistencies (e.g., “Withdrawn”, “withdrawn” → “withdrawn”).
- **Label construction** (early-warning target). I defined an event label per inspection record: `label_close_within_H` = 1 if the first time an outlet was observed closed (by status) occurred within `H` = 180 days after that inspection; otherwise 0
- **Entity harmonization.** Records were grouped by the stable identifier `premiseno`.

Where duplicate rows existed (identical `premiseneno-inspectiondate` pairs), the most complete record was retained.

3.5.2 Leakage-Safe Early-Warning Label

A binary label was constructed per inspection record that indicates whether the first observed closure for a pharmacy occurred within 180 days after that inspection:

$$y_i = \mathbf{1}(0 \leq T_{p(i)}^{\text{close}} - t_i \leq H), \quad H = 180 \text{ days},$$

where t_i is the inspection date for record i and $T_{p(i)}^{\text{close}}$ is the earliest closure date observed for pharmacy $p(i)$.

A prediction horizon of **180 days** was selected because it provides a practical balance between predictive relevance and regulatory usefulness. The six-month period is sufficiently long to capture meaningful pharmacy closure events following inspection while remaining short enough for inspection findings and compliance indicators to retain predictive value. The selected horizon also supports risk-based regulatory follow-up and minimizes temporal leakage by ensuring that only future closure events occurring after each inspection date are considered.

Operational steps applied:

1. **Instantaneous closure flag.** For each row, set `is_closed_instant` = 1 if `premisestatus` ∈ {withdrawn, closed} or `activestatus` ∈ {expired, inactive}; otherwise 0.
2. **Earliest closure per pharmacy.** Computed `first_closed_date` as the group minimum of `inspectiondate` among rows with `is_closed_instant` = 1 for that `premiseneno`.
3. **Label assignment.** For every inspection row, set `label_close_within_H`=1 if $0 \leq \text{first_closed_date} - \text{inspectiondate} \leq 180$ days; else 0

This procedure aligned labels to information available at or before each inspection date and thus prevents temporal leakage

3.5.3 Feature Readiness (for later modelling)

To ensure that model inputs were complete, comparable across outlets, and operationally available at inspection time, feature readiness steps were implemented in Python using pandas for feature construction and ordering, and scikit-learn for imputation, scaling, and categorical encoding within the modelling pipeline.

To enable consistent learning and avoid bias from missingness:

- **Missing values.** Numerical features were imputed by median (robust to skew); categorical fields used explicit “unknown” handling via one-hot encoding with unseen-category ignore at inference. Categorical fields were handled through one-hot encoding using `sklearn.preprocessing`.

`OneHotEncoder(handle_unknown='ignore')`, ensuring that unseen categories appearing in future records would not cause inference-time failures.

- **Derived time deltas.** `days_since_last_inspection` was computed as the per-pharmacy day difference between consecutive inspections (first observation imputed by the median and clipped at 0), and `days_to_expiry` as (`expirydate` – `inspectiondate`) in days (negatives indicate overdue; median-imputed if missing).
- **Consistency checks.** All computations were done on data sorted by `premiseneno` then `inspectiondate` to ensure valid deltas.

These steps yielded a clean, leakage-safe analytical table with well-defined labels and harmonised inputs suitable for the supervised learning pipeline.

3.5.4 Data Quality Issues Fixed

Several data quality issues were addressed prior to modelling to ensure reliable feature engineering and leakage-free label construction. All cleaning and transformations were implemented in Python (pandas, numpy).

- **Duplicates & inconsistent IDs** Duplicate rows were assessed using the stable facility identifier `premiseneno` together with `inspectiondate`. Records sharing identical (*`premiseneno`, `inspectiondate`*) keys were treated as potential duplicates and resolved by retaining the most complete observation (highest non-missing fields), while ensuring one record per key remained for time-delta calculations. Where `premiseneno` appeared with spelling/format variants, the analysis relied on `premiseneno` as the canonical key to avoid fragmentation of pharmacy histories. *(Implemented using `pandas duplicated()`, `sort_values()`, `drop_duplicates()`, and completeness checks via `isna().sum()`).*
- **Missing expiry/inspection severity:** Date fields (*`inspectiondate`, `expirydate`, `createddate`, `applicationdate`*) were coerced into proper datetime format, with invalid entries converted to `NaT` to prevent arithmetic errors during time-based feature construction. For modelling, missing numerical values (e.g., computed deltas such as `days_since_last_inspection`, `days_to_expiry`) were handled through median imputation to reduce sensitivity to skew, while categorical fields were made robust at inference time using one-hot encoding with unknown-category handling.
- **Chronology:** Because the model uses inspection history features (e.g., `days_since_last_inspection` and `inspection_count`), all records were strictly ordered **by** `premiseneno` then `inspectiondate` before generating deltas and cumulative counts. This ensured that inspection intervals were computed from the correct preceding inspection for each pharmacy. *(Implemented using `pandas sort_values(['premiseneno','inspectiondate'])`, group operations with `groupby().diff()`, and `groupby().cumcount()`).*
- **Label consistency checks.** After constructing the early-warning target (`label_close_within_H`), sanity checks were performed to confirm that required columns existed (e.g., `first_closed_date`) and that the positive label rate was plausible relative to the dataset size. *(Implemented using `assertions` and descriptive checks such as `mean()` and `value_counts()`).*

3.5.5 Descriptive Analysis

Descriptive statistical analysis was conducted prior to exploratory and predictive modelling to characterize the pharmacy regulatory dataset and provide an overall understanding of the pharmacy closure landscape in Uganda. This stage was important for establishing the contextual background of the study, identifying observable patterns in closure events, and generating regulatory intelligence that could inform both interpretation and subsequent predictive analysis.

The descriptive analysis summarized the distribution of pharmacies and closure events across key regulatory and operational dimensions. Specifically, the analysis examined the number of pharmacies captured in the dataset by year, the annual number and proportion of pharmacy closures, geographic distribution of closures by region and district, closure patterns by pharmacy category (retail versus wholesale), closure distribution by inspection outcomes, compliance status, licensing status, and operational activity status. Temporal trends in inspections and closures were also explored to identify possible shifts in closure patterns over the study period.

Continuous variables such as inspection frequency, time since the previous inspection, and days to license expiry were summarized using measures of central tendency and dispersion, including means, medians, standard deviations, and interquartile ranges, as appropriate. Categorical variables were summarized using frequencies and proportions. Trend analysis, summary tables, and graphical visualizations were used to present closure patterns and facilitate interpretation of the regulatory environment.

This descriptive analysis answered an important foundational question: **what does the pharmacy closure landscape in Uganda look like?** By examining closure patterns before predictive modelling, the study provided insight into the magnitude, distribution, and characteristics of closure events, thereby positioning the research not only as a machine learning exercise but also as a regulatory intelligence and decision-support study.

Data cleaning, descriptive summaries, aggregation, and visualization were performed in Python using the pandas and numpy libraries for data manipulation, and matplotlib and seaborn for graphical exploration and trend visualization.

3.5.6 Exploratory Risk Factor Analysis

Exploratory risk factor analysis was conducted before predictive modelling to examine which factors appeared to be associated with pharmacy closures. This analysis directly addressed the first specific objective of the study, which was to identify the risk factors contributing to pharmacy closures in Uganda.

The analysis compared pharmacies or inspection records that were associated with closure against those that were not associated with closure. Categorical variables, such as license status, operational activity status, inspection outcome, pharmacy type, region, and district, were examined using cross-tabulations and chi-square tests. Where cell counts were small, Fisher's exact test was considered because it provides more reliable significance testing when categorical contingency tables contain small expected cell counts, where chi-square assumptions may be violated. Continuous variables, such as inspection frequency, days since

last inspection, and days to license expiry, were summarized using means, medians, standard deviations, and interquartile ranges. Group differences were assessed using independent-sample t-tests for normally distributed variables and Mann-Whitney U tests for non-normally distributed variables.

Crude association analysis was also conducted to identify variables that showed meaningful relationships with pharmacy closure. Examples included examining whether expired license status, inactive operational status, repeated inspection deficiencies, geographic location, or longer intervals since the previous inspection were associated with increased likelihood of closure. These analyses helped identify variables that were relevant for subsequent feature engineering, model interpretation, and early warning indicator development.

The purpose of this stage was not to establish causality, but to understand which available regulatory variables appeared to carry useful predictive or explanatory information. Findings from the exploratory risk factor analysis informed the selection and interpretation of candidate predictors used in the machine learning model. Data manipulation and analysis were performed in Python using pandas, numpy, scipy, and statsmodels, while visual summaries were generated using matplotlib and seaborn.

3.6 Model Development

This section described the process used to develop the predictive model for pharmacy closure. The objective was to identify a machine learning algorithm capable of predicting the likelihood of pharmacy closure within a 180-day prediction horizon using historical regulatory inspection data from the National Drug Authority (NDA). Since the study outcome was binary (closure or non-closure), supervised machine learning classification techniques were considered appropriate for this prediction task.

Model development followed an applied supervised-learning pipeline aligned with the study objectives. Inspection records were transformed into inspection-time features, an early-warning closure label was constructed using a 180-day prediction horizon, and predictor variables were pre-processed to create a modelling dataset suitable for machine learning. The resulting model generates a probability risk score for each pharmacy, enabling pharmacies to be ranked according to their likelihood of closure and supporting risk-based inspection planning and regulatory decision-making.

Below in *Figure 3.2* were steps involved in the model development

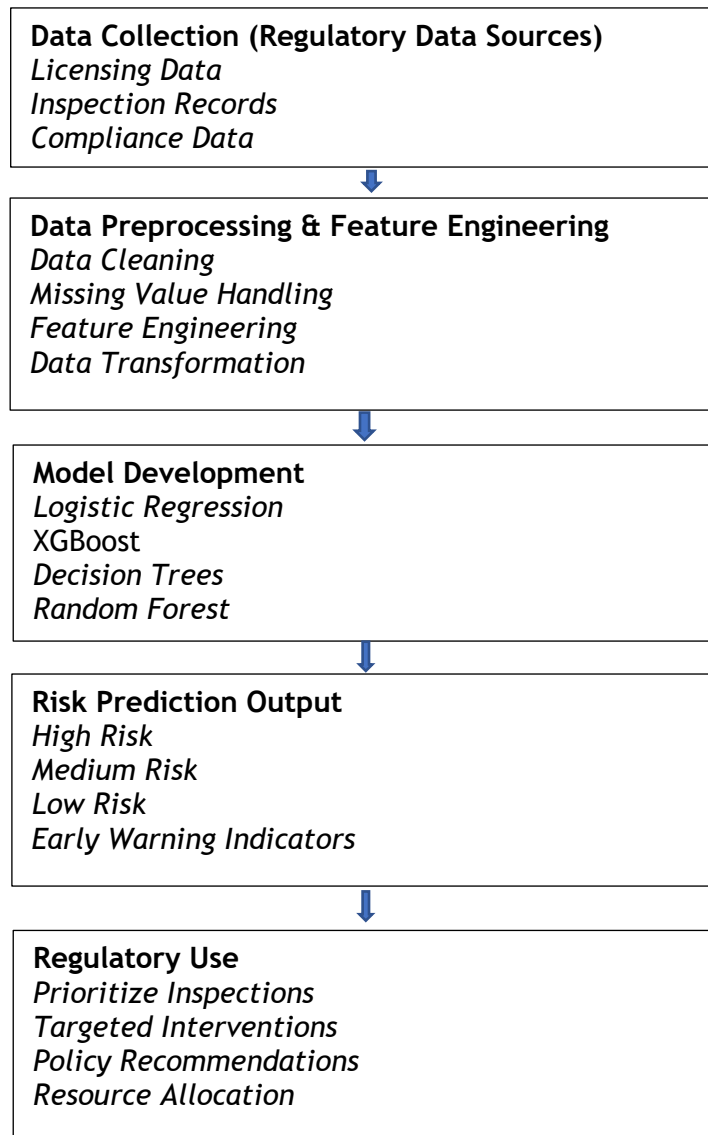


Figure 3.2: Steps in involved in the Model Development

3.6.1 Train/Test Protocol

To simulate the conditions under which the model would be used in practice, a time-based train/test split was applied. A chronological, non-random train/test split was applied. Inspection records were first ordered by inspectiondate. The oldest 80% of records were assigned to the training dataset, while the most recent 20% were reserved as the independent test dataset. The cut-off date was determined using the 80th percentile of the inspectiondate variable.

This approach ensured that the model was trained on historical records and evaluated on later inspections, thereby reducing the risk of training the model using information from the future relative to the test period. It also provided a more realistic assessment of how the model would perform when applied to newly received regulatory data.

During model selection and hyperparameter tuning, group-aware cross-validation was conducted using GroupKFold. The pharmacy identification number, premiseno, was used as the grouping variable. This ensured that inspection records belonging to the same pharmacy did not appear simultaneously in the training and validation subsets of any cross-validation fold. The procedure reduced the risk of information leakage arising from repeated inspections of the same outlet.

3.6.2 Candidate Models and Empirical Comparison

Four supervised machine-learning algorithms were evaluated for the prediction of pharmacy closure within 180 days: Logistic Regression, Decision Tree, Random Forest, and Extreme Gradient Boosting (XGBoost). These algorithms were selected because they represent different classification approaches and levels of model complexity: a linear statistical classifier, a single tree-based classifier, a bagging ensemble, and a gradient-boosting ensemble, respectively.

The purpose of evaluating several algorithms was to determine whether the additional complexity of ensemble-learning methods provided an empirical advantage over simpler classification approaches when applied to the available pharmacy regulatory data.

To ensure a fair comparison, all four candidate models were trained and evaluated using the same analytical dataset, predictor variables, closure definition, pre-processing procedures, and chronological train/test split. The oldest 80% of inspection records were used for model development, while the most recent 20% were retained as the independent test dataset.

Within the training dataset, group-aware cross-validation using GroupKFold was applied, with premiseno used as the grouping variable. This ensured that repeated inspection records belonging to the same pharmacy were not simultaneously represented in the training and validation portions of a cross-validation fold, thereby reducing the risk of information leakage.

The models were compared using complementary performance measures, including ROC-AUC, PR-AUC, precision, recall, F1-score, and Precision@k. Because pharmacy closure represented the minority outcome, particular attention was given to PR-AUC, recall, F1-score, and Precision@k, while ROC-AUC was used to assess overall discriminatory ability.

The independent chronological test dataset was reserved for evaluation of the candidate models on later inspection records and was not used as a substitute for the model-development process.

The model characteristics summarized in Table 3.2 below provided the rationale for including each algorithm in the empirical comparison. Final model assessment, however, was based on observed predictive performance rather than theoretical suitability alone.

Table 3.2: Characteristics of Candidate Machine-Learning Algorithms

Model	Modelling Approach	Relevance to the study	Strengths	Limitations
Logistic Regression	Linear statistical classifier	Considered as a conventional baseline approach for binary classification	Highly interpretable, computationally efficient, and easy to implement.	Assumes linear relationships between predictors and the outcome; limited ability to model complex interactions.
Decision Tree	Rule-based classifier	Considered because regulatory decisions can be represented through rule-based splits	Easy to interpret and visualize; captures nonlinear relationships.	Susceptible to overfitting and instability when trained on complex datasets.
Random Forest	Ensemble learning (bagging)	Considered as a robust ensemble method for heterogeneous regulatory data	Robust, handles nonlinear relationships, reduces variance, and performs well with heterogeneous data.	Less interpretable than a single decision tree and computationally more intensive
Extreme Gradient Boosting (XGBoost)	Gradient boosting ensemble	Evaluated as an advanced ensemble algorithm capable of modelling complex nonlinear relationships in structured tabular data.	High predictive accuracy, captures complex feature interactions, incorporates regularization, efficiently handles missing values, and supports feature importance analysis.	Less interpretable than a single decision tree and computationally m

3.6.3 Comparative Model Performance and Model Selection

Following empirical evaluation of the four candidate algorithms, model performance was compared using group-aware cross-validation during model development and the chronologically held-out test dataset for final assessment of generalisation.

The results showed that no single model dominated all evaluation measures. XGBoost achieved the strongest group-aware cross-validation performance, while Random Forest achieved the highest ROC-AUC and PR-AUC on the independent chronological test set. The Decision Tree achieved the strongest Precision@200 in the capacity-constrained evaluation.

Because the models performed differently across evaluation criteria, the study did not interpret the results as demonstrating a universally superior algorithm. For assessment of predictive generalisation, particular emphasis was placed on performance on the independent chronological test dataset, since this most closely represented the intended future-use setting.

Random Forest therefore demonstrated the strongest overall held-out predictive performance, while XGBoost remained a highly competitive model and was retained for detailed feature-importance and SHAP-based interpretation.

XGBoost was used for interpretive analysis because its tree-based structure supports nonlinear relationships and interactions among regulatory, inspection, licensing, operational and geographical variables, and because SHAP provides both global and observation-level explanations of the model's predictions.

The predictive comparison of all four algorithms was therefore reported separately from the interpretability analysis. This distinction allowed the study to identify the strongest held-out predictive model while also examining the regulatory risk signals captured by the XGBoost model.

3.6.4 Model Deployment and Operationalization

Following identification of a model with acceptable predictive performance, a deployment framework was proposed to support future use by regulatory authorities. The framework included:

- **Saving the trained model** using serialization techniques (e.g., joblib or pickle in Python).
- **Developing an interface** or dashboard to allow stakeholders to input pharmacy data and receive closure risk predictions.
- **Setting up automated data pipelines** to update the model with new data facilitating ongoing monitoring for example monitoring model performance and data quality over time, defining procedures for periodic model review and retraining;
- **Documenting model limitations and assumptions** to ensure appropriate interpretation and prevent misuse.

3.6.5 Control of Overfitting and Model Generalization

Several measures were incorporated into the modelling protocol to reduce overfitting and provide a realistic assessment of model generalization. First, final performance was not assessed using the same observations on which the models were trained. Instead, the dataset was divided chronologically, with the oldest 80% used for model development and the newest 20% retained for independent evaluation.

Second, group-aware cross-validation was conducted within the training dataset using pharmacy as the grouping variable. This reduced the possibility that repeated inspection records from the same pharmacy would appear simultaneously in the training and validation portions of a fold.

Model complexity was considered during model development. The candidate set included both simpler models, such as Logistic Regression and Decision Tree, and ensemble methods, including Random Forest and XGBoost. XGBoost additionally incorporates regularization mechanisms that constrain model complexity.

3.7 Evaluation Metrics

The predictive performance of the candidate machine-learning models was evaluated using several complementary metrics. Multiple measures were necessary because pharmacy closure represented the minority outcome and accuracy alone could therefore provide an incomplete assessment of model performance.

- **Accuracy** measured the overall proportion of observations correctly classified as closure or non-closure.

- **Precision** measured the proportion of records predicted as closure that subsequently met the study-defined closure outcome.
- **Recall (sensitivity)** measured the proportion of actual closure cases correctly identified by the model.
- **F1-score** represented the harmonic mean of precision and recall and was used to summarize the balance between identifying closure cases and limiting false-positive classifications.
- **ROC-AUC** measured the ability of each model to discriminate between closure and non-closure observations across classification thresholds.
- **PR-AUC** summarized the precision-recall relationship across thresholds and was particularly relevant because pharmacy closure represented the minority class.
- **Precision@k** measured the proportion of observed closure cases among the highest-ranked k predictions. This metric was used to assess operational usefulness where regulatory inspection capacity is limited and only a specified number of high-risk pharmacies can be prioritized.

Greater emphasis was placed on PR-AUC, recall, F1-score and Precision@k because these metrics provided information relevant to the minority closure class and the study's risk-based prioritization objective. ROC-AUC was additionally used to compare overall discrimination between the candidate models.

Comparative performance was assessed on the same chronologically held-out test dataset to ensure that differences between algorithms were not attributable to different test samples.

3.8 Tools and Software

Data analysis and modeling were performed using the following tools:

- **Python:** The Core programming language, with libraries like pandas for data manipulation, scikit-learn for model development, XGBoost for gradient boosting, and matplotlib/seaborn for data visualization.
- **Jupyter Notebook:** Used as the main environment for model development and documentation.

3.9 Identification of Predictive Risk Indicators

The key objective was to identify the factors that influence the likelihood of pharmacy closure. Candidate predictors were derived from historical inspection, licensing, operational, and geographic records

Relevant features were identified such as:

- **Inspection frequency:** Pharmacies with frequent violations or poor inspection history are more likely to close.
- **Days since last inspection:** A long gap between inspections might be indicative of neglected compliance.
- **Days to license expiry:** Pharmacies that are close to having an expired license and have not yet renewed it could be at higher risk of closure.

- **Geographic location:** Pharmacies in underserved or economically disadvantaged areas may face more operational challenges and higher closure rates.
- **Compliance status:** Non-compliance with regulatory requirements (e.g., expired licenses or frequent inspection violations) increases closure risk.

Exploratory statistical analysis was initially used to examine patterns and associations between candidate variables and the closure outcome. Predictive modelling was subsequently used to determine the extent to which these variables contributed to closure-risk estimation.

Because the final interpretive analysis was conducted using XGBoost, feature-importance measures and SHAP values were used to examine the contribution of predictors to model outputs. These findings were interpreted as predictive risk indicators rather than causal determinants of pharmacy closure.

3.10 Identification of Key Risk Factors Using XGBoost

Following empirical model comparison and selection of XGBoost as the principal interpretive estimator, feature-importance measures and Shapley Additive Explanations (SHAP) was used to investigate the variables contributing to predicted pharmacy-closure risk.

XGBoost feature importance provided a global indication of which predictors contributed most strongly to the model's predictive structure. SHAP was subsequently used to provide both global and observation-level interpretation by quantifying the direction and magnitude of each predictor's contribution to individual risk predictions.

SHAP summary was used to examine the relative contribution of variables such as **days_since_last_inspection**, **inspection_count**, **days_to_expiry**, and geographical or regulatory characteristics.

3.11 Ethical Consideration

Ethical considerations included ensuring that all data was anonymized and securely stored following the NDA's guidelines on data protection and privacy. Ethical approval was obtained from the relevant authorities to ensure compliance with research ethics standards. All datasets used in this study contained no personally identifiable information, and access was restricted to authorized members of the research team.

3.12 Limitations of the Methodology

This study had several limitations, including:

- **Data limitations:** The quality and completeness of secondary data obtained from the NDA varied, which may have impacted the accuracy and reliability of the model predictions.
- **Context-specific results:** The findings were specific to the Ugandan regulatory and operational context, limiting their direct applicability to other settings without adaptation.

- Exclusion of non-regulatory factors: The study focused solely on regulatory data and did not consider other factors influencing pharmacy closures, such as financial mismanagement, owner health, or informal operational practices.

CHAPTER FOUR

PRESENTATION AND INTERPRETATION OF STUDY FINDINGS

4.1 Introduction

This chapter presents the empirical findings of the study. It first summarizes the analytical dataset and the construction of the 180-day pharmacy-closure outcome. It then presents the features available at inspection time, the chronological train/test protocol, and the empirical comparison of four candidate machine-learning models: Logistic Regression, Decision Tree, Random Forest and XGBoost. The chapter subsequently examines the performance of the selected models under both conventional predictive metrics and capacity-constrained regulatory prioritization. Finally, the chapter presents model interpretation, error analysis, sensitivity analysis, early-warning indicators, descriptive closure patterns and implications for risk-based regulatory planning. [31]

4.2 Data Summary and Target Label Construction

This section characterizes the analytic dataset and explains how the binary early-warning target was defined in a leak-free manner. We begin by summarizing the inspection corpus (rows, unique premises, temporal coverage, and missingness) and the resulting outcome prevalence ($\approx 11.52\%$ of inspection records labeled positive). To construct the target, we first derive an instantaneous closure signal at the record level—treating a premise as closed when $\text{premisestatus} \in \{\text{withdrawn}, \text{closed}\}$ or $\text{activestatus} \in \{\text{expired}, \text{inactive}\}$ and then compute each premises' `first_closed_date` as the earliest inspection date meeting that signal [35]. An inspection is labeled 1 if the premises' first closure occurs within $H = 180$ days after inspection date; otherwise it is labeled 0 (including cases with no observed closure in the data window).

This forward-looking definition aligns with regulatory planning horizons while preventing target leakage by ensuring that only information observable at or before the inspection contributes to the label. The section also reports descriptive breakdowns (overall, yearly, and by region), which contextualize later precision-recall metrics and capacity-based evaluation.

Dataset. The analytic dataset comprises inspection records for registered pharmacy premises in Uganda. After basic cleaning and date parsing, the modelling table used in this chapter contains 19,636 inspection rows (observations). The binary outcome is whether the pharmacy first becomes closed within $H = 180$ days after inspection

Leak-free labelling. To avoid target leakage, an *instantaneous* “closed” signal was first derived at the row level:

For each pharmacy (premiseneno), the first_closed_date was computed as the earliest inspection date where the instantaneous signal is true. An inspection row is labeled 1 if

$$\leq (\text{first_closed_date} - \text{inspection date}) \leq 180 \text{ days,}$$

and 0 otherwise (including pharmacies that never close in the observation window).

Class prevalence. The positive-class prevalence in the modeling table is 11.52% (i.e., ~1 in 8.7 inspections leads to closure within 180 days).

Interpretation. The label measures an early-warning outcome aligned to regulatory planning horizons, and—because its constructed strictly forward in time—it avoids using any future information as features.

The observed prevalence of approximately 11.5% indicates that pharmacy closures constitute a relatively rare event, resulting in a class-imbalanced prediction problem. This finding is consistent with Saito and Rehmsmeier (2015) [72], who observed that many real-world classification problems, particularly in healthcare and regulatory domains, involve highly imbalanced datasets where the minority class represents the events of greatest interest. The finding further supports the use of Precision-Recall metrics in addition to ROC-AUC when evaluating predictive performance.

Table 4. 1: Dataset overview (rows, unique pharmacies, date span, missingness)

Table 4.1 – Dataset Overview

	Metric	Value
0	Total inspection rows	19636
1	Unique pharmacies	5149
2	Time span (min inspection date)	2012-01-06
3	Time span (max inspection date)	2025-04-29
4	Median inspections per pharmacy	3.0
5	Mean inspections per pharmacy	3.81
6	Std. inspections per pharmacy	2.85
7	Label prevalence (closure within H=180 days)	0.115
8	Number of regions	9
9	Number of districts	126

Table 4.1b – Missingness (key columns)

	Column	Missing (n)	Missing (%)
0	premiseneno	0	0.00
1	premisename	0	0.00
2	inspectiondate	0	0.00
3	expirydate	97	0.49
4	premisestatus	0	0.00
5	activestatus	0	0.00
6	region	0	0.00
7	district	0	0.00
8	premise_location	0	0.00
9	inspection_status	0	0.00
10	label_close_within_H	0	0.00

Tables 4.1 and 4.1b above Summarizes inspection records used for modeling, including the number of rows, unique pharmacies, temporal coverage, inspection density per pharmacy, and missingness across key fields.

Table 4.2: Label prevalence overall and by year/region

	region	N	Positives	Prevalence
2	Kampala Extra Regional Office- Akamwesi	8262	601	0.073
0	Central Regional Office - Masaka	4798	788	0.164
5	South Eastern Region - Jinja	1896	263	0.139
6	South Western Region - Mbarara	1357	172	0.127
8	Western Region - Hoima	1298	169	0.130
1	Eastern Region - Tororo	747	90	0.120
4	Northern Region - Lira	578	85	0.147
7	West Nile Region - Arua	442	53	0.120
3	North Eastern Region - Soroti	258	41	0.159

	Year	N	Positives	Prevalence
0	2012	1	0	0.000
1	2013	2	0	0.000
2	2014	3	0	0.000
3	2015	2	1	0.500
4	2016	214	65	0.304
5	2017	3301	697	0.211
6	2018	2492	136	0.055
7	2019	2630	132	0.050
8	2020	2801	192	0.069
9	2021	1593	307	0.193
10	2022	1755	525	0.299
11	2023	3271	109	0.033
12	2024	1536	98	0.064
13	2025	35	0	0.000

Tables 4.2 represents the proportion of inspections labeled as “closure within H=180 days” overall, by inspection year, and by region (or district). These baselines contextualize precision-recall metrics and precision@k

4.3 Features Available at Inspection Time

Here we described the features available at the time of inspection that feed the predictive model, restricted strictly to information observable on or before the inspection to avoid any look-ahead bias.

The engineered numerical features include:

- *days_since_last_inspection* – the number of days since the previous inspection for the same pharmacy (median-imputed for first observations and clipped at 0);
- *inspection_count*– the cumulative count of inspections completed for that pharmacy up to the current visit; and
- *days_to_expiry* –the number of days from the inspection date to license expiry (median-imputed when missing; negative values indicate overdue status).

The categorical features comprise *region*, *district*, *premise_location* (when present), and *inspection_status* (when present). As part of the preprocessing pipeline, numerical features are median-imputed and standardized, while categorical features are one-hot encoded with unknown categories safely ignored at inference to ensure robustness on new data. [33]

Table 4.3 below provides a concise feature dictionary (name, type, description, transformation). Collectively, these features are observable at or before the inspection

event, ensuring the learning pipeline is operationally realistic and leakage-free, and that performance estimates on the held-out (future) period are valid.

Table 4.3: Feature dictionary: name, type, description, transformation

	Feature	Type	Pandas dtype	Description	Transformation in Pipeline
0	days_since_last_inspection	numeric	float64	Days since the previous inspection for the sam...	Median imputation; StandardScaler (z-score).
1	inspection_count	numeric	int64	Cumulative number of inspections completed for...	Median imputation; StandardScaler (z-score).
2	days_to_expiry	numeric	float64	Days from inspection date to license expiry (n...	Median imputation; StandardScaler (z-score).
3	region	categorical	object	Administrative region in which the premises is...	OneHotEncoder(handle_unknown='ignore').
4	district	categorical	object	Administrative district in which the premises ...	OneHotEncoder(handle_unknown='ignore').
5	premise_location	categorical	object	Location/category tag for the premises (e.g., ...	OneHotEncoder(handle_unknown='ignore').
6	inspection_status	categorical	object	Inspection outcome/status recorded at the time...	OneHotEncoder(handle_unknown='ignore').
7	label_close_within_H	numeric	int32	Target: 1 if the premises first closed within ...	No transform (supervised target).

The table 4.3 above illustrates all features observable at or before the inspection time, ensuring the predictive pipeline is operationally realistic and leakage-free.

4.4 Train/Test Protocol (Time-Aware)

To approximate real-world deployment and prevent optimistic bias, a model was evaluated using a time-aware split rather than a random split. Specifically, inspections are ordered chronologically and the oldest 80% form the training set, while the newest 20% constitute the test set. This design mirrors the operational setting—training on historical inspections to predict future closures—and minimizes temporal leakage (i.e., information from the future influencing past predictions). All pre-processing and model fitting occur only on the training period, and performance is reported on the untouched future period. This protocol provides a conservative, policy-relevant estimate of out-of-sample performance for proactive inspection planning. [35]

The adoption of a time-aware training and testing protocol agrees with recommendations by Cerqueira et al. (2020) [73], who argue that chronological validation provides more realistic estimates of predictive performance than random data splitting for temporal datasets. Similarly, Roberts et al. (2017) [74] recommend temporal validation when developing predictive models intended for future operational deployment because it minimizes information leakage and produces more reliable estimates of model performance.

Within the training period, group-aware cross-validation using GroupKFold was applied with premiseno as the grouping variable. This ensured that repeated inspection records from the same pharmacy were not simultaneously represented in the training and validation subsets of a cross-validation fold. The chronologically later 20% remained reserved for independent evaluation.

4.5 Candidate Models and Empirical Comparison

Four supervised machine-learning algorithms were empirically evaluated: Logistic Regression, Decision Tree, Random Forest and Extreme Gradient Boosting (XGBoost). The models were evaluated using the same analytical dataset, predictor variables, preprocessing framework, closure definition and chronological train/test protocol [37].

The candidate set represented modelling approaches of increasing complexity: Logistic Regression as an interpretable linear baseline, Decision Tree as a single nonlinear tree-based classifier, Random Forest as a bagging ensemble, and XGBoost as a gradient-boosting ensemble.

Group-aware cross-validation was performed within the training dataset, followed by evaluation on the independent chronologically held-out test dataset [38] [39].

Table 4.4: Comparative predictive performance of candidate models

Model	ROC-AUC	PR-AUC	Test ROC-AUC	Test PR-AUC	Precision @ 0.5	Recall @ 0.5	F1-score
Logistic Regression	0.984	0.923	0.900	0.205	16.5%	80.9%	0.366
Decision Tree	0.976	0.925	0.904	0.213	23.1%	97.9%	0.375
XGBoost	0.989	0.948	0.902	0.214	22.6%	98.0%	0.375
Random Forest	0.907	0.927	0.903	0.213	23.1%	96.4.0%	0.366

The empirical comparison showed that no single model dominated all evaluation measures. XGBoost achieved the highest cross-validation ROC-AUC (0.989) and PR-AUC (0.948), indicating the strongest internal validation performance during model development. On the independent chronological test dataset, however, Random Forest achieved a ROC-AUC (0.903) and PR-AUC (0.234), while XGBoost remained closely competitive with a ROC-AUC of 0.902 and PR-AUC of 0.238.

At the 0.5 classification threshold, Decision Tree and XGBoost both achieved 97.9% and 98.0% respectively recall and an F1-score of 0.375. Logistic Regression and Random Forest achieved slightly lower recall but similar precision. These findings demonstrate that model performance depended on the evaluation criterion.

XGBoost was retained for detailed feature-importance and SHAP-based interpretation because it provided strong predictive performance together with a well-established framework for examining nonlinear feature contributions at both global and individual prediction levels. This decision was therefore treated separately from the comparative performance assessment and was not interpreted as evidence that XGBoost was universally superior to all alternative algorithms.

4.6 Overall Predictive Performance (Ranking Focus)

This section reports the model's ability to rank pharmacies by closure risk on the chronologically held-out test set. Because the outcome was imbalanced and the operational used to prioritize inspections, we emphasize threshold-free ranking metrics—the Receiver Operating Characteristic Area Under the Curve (ROC-AUC) and the Precision-Recall AUC (PR-AUC)—alongside a single, thresholded summary for completeness [40]. ROC-AUC captured how well positives are ranked above negatives across all thresholds, while PR-AUC was more informative under class imbalance by directly reflecting the trade-off between precision (hit-rate) and recall (coverage).

We then connected these global metrics to capacity-constrained practice using precision@k, which quantifies the expected hit-rate when inspectors act on the top-k highest-risk premises. Together, these measures provided a comprehensive, deployment-relevant view of predictive performance [41].

Test-set metrics (newest 20% of inspections):

- **ROC-AUC = 0.902**
- **PR-AUC = 0.214** (baseline prevalence ≈ 0.115)

Thresholded report at $\tau=0.5$ (for completeness):

- Positive class (closure within 180 days): **precision = 0.231, recall = 1.000**;
- Implied that the model flags many candidates at $\tau=0.5$ (high recall), which is expected under imbalance.

Interpretation. In Figure 4.1, ROC-AUC indicates strong ranking power, and PR-AUC—more informative under class imbalance—shows $\sim 4\times$ enrichment over chance (0.214 vs ≈ 0.115 baseline), supporting risk-based prioritization.

The XGBoost model achieved a ROC-AUC of 0.902, indicating excellent discrimination between pharmacies that subsequently closed and those that remained operational. this finding is consistent with chen and guestrin (2016) [74], who demonstrated that XGBoost achieves superior predictive performance on structured tabular datasets through gradient boosting and regularization. similarly, lundberg et al. (2020) [75] reported that tree-based ensemble models frequently achieve ROC-AUC values exceeding 0.85 while maintaining interpretability using shap. the high roc-auc obtained in this study therefore supports the suitability of XGBoost for regulatory risk prediction.

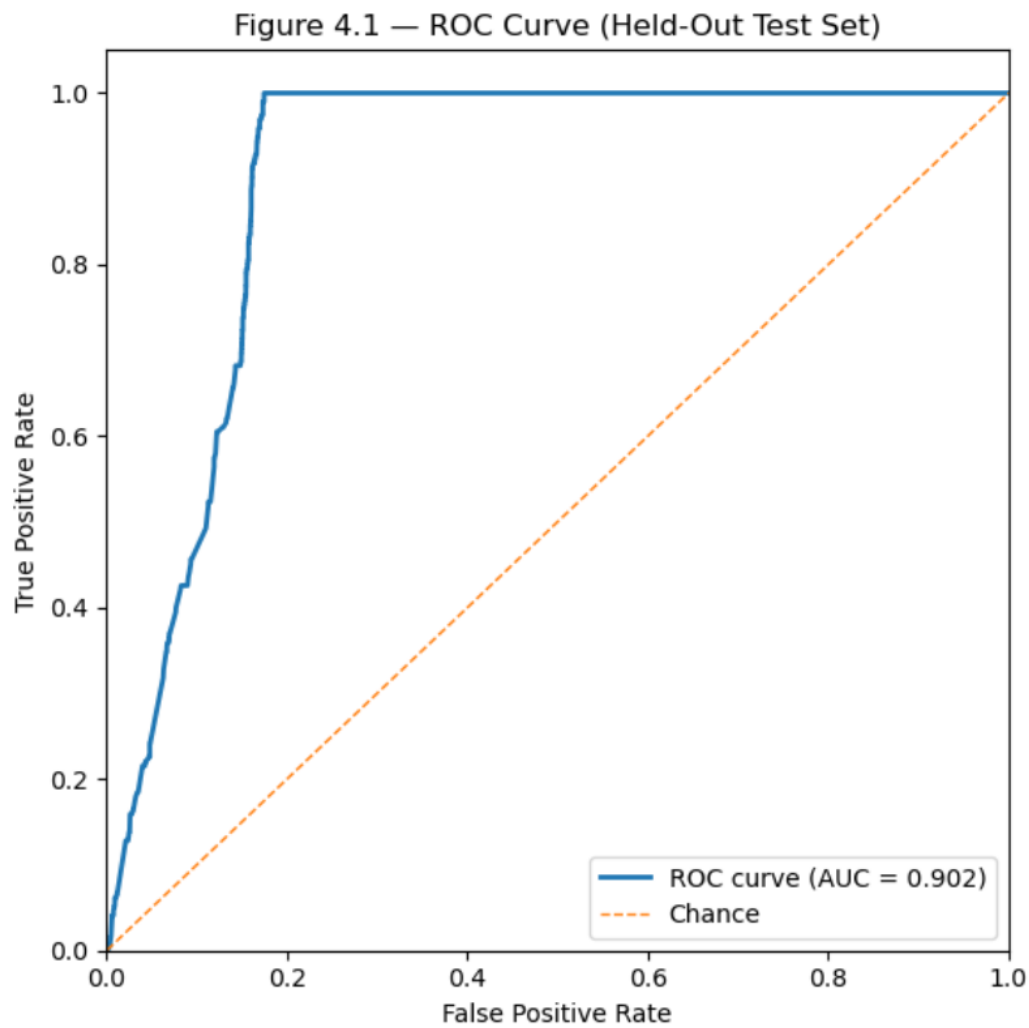


Figure 4.1: ROC curve with AUC = 0.902

Figure 4.1 illustrates the Receiver Operating Characteristic (ROC) curve on the chronologically held-out test set; the model achieves AUC ≈ 0.902 , indicating strong ability to rank pharmacies that close within 180 days above those that remain open [42].

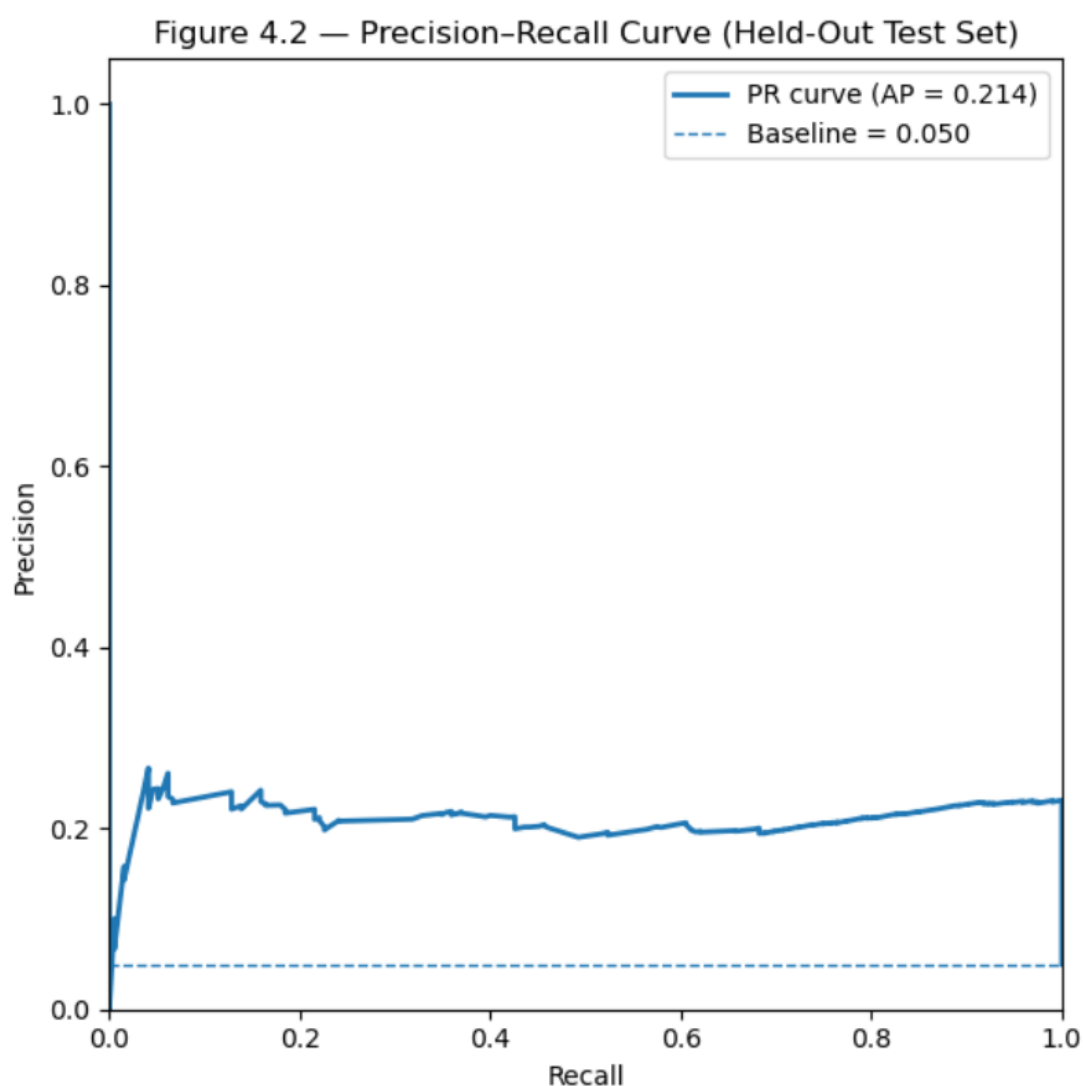


Figure 4.2: Precision-Recall (PR) curve with area = 0.214

Figure 4.2 above represents the Precision-Recall (PR) curve on the chronologically held-out test set; the model achieves PR-AUC $\approx$ **0.214**, substantially above the baseline prevalence ($\sim$ 0.115), indicating strong enrichment for likely closures under class imbalance.

4.7 Capacity-Constrained Evaluation (Top-k Prioritization)

Regulatory inspection is capacity-limited. We therefore evaluate **precision@k**—the proportion of true closures among the top-ranked k inspections in the test period [43].

Key result (pilot capacity $k=200$): **precision@200 = 0.210**, i.e., 42/200 top-ranked cases closed within 180 days. Table 4.4 Precision@ k across practical capacities (e.g., $k \in \{50, 100, 200, 300, 500, 1000\}$ reporting both precision and hit counts.

Table 4.4: Precision@k on the held-out (time-ordered) test set after de-duplicating to one row per pharmacy

		2			
	k (capacity)	available rows (unique pharmacies)	evaluated k	hits (true closures)	precision@k
0	50	2804	50	12	0.240
1	100	2804	100	24	0.240
2	200	2804	200	42	0.210
3	300	2804	300	64	0.213
4	500	2804	500	95	0.190
5	1000	2804	1000	189	0.189

Thresholds that select Top-k (tau_k)

	k (capacity)	tau_k (score threshold)
0	50	0.999551
1	100	0.999541
2	200	0.999455
3	300	0.999415
4	500	0.999313
5	1000	0.001483

Table 4.4 illustrates Precision@k reports the fraction of true closures among the top-k highest-risk pharmacies. Results are shown for $k \in \{50, 100, 200, 300, 500, 1000\}$; ‘evaluated k’ reflects the number of unique pharmacies available in the test set if fewer than k.”

Interpretation. Selecting the **top 200** highest-risk inspections yields a ~4× uplift relative to the baseline closure rate, providing a concrete and actionable shortlist for proactive fieldwork.

The model correctly identified 42 pharmacies among the top 200 highest-risk predictions, corresponding to a Precision@200 of 21%. This represents approximately a four-fold improvement over the baseline closure prevalence. This finding is consistent with the risk-based inspection principles advocated by the OECD (2018) [71], which emphasize prioritizing regulatory inspections using evidence-based risk rankings to maximize the effectiveness of limited inspection resources.

*(Note: In operations, we ensured **one row per pharmacy** by retaining the highest-risk inspection per premises to avoid duplicating visits [45].)*

4.8 Threshold Selection and Probability Calibration

This section explains how the model’s continuous risk scores are converted into operational decisions. Because closures are rare and inspection resources are finite, a fixed threshold (e.g., $\tau=0.5$) is usually sub-optimal under class imbalance. Instead, we adopt capacity-based thresholding—choose a score cutoff τ_k that flags exactly the top- k highest-risk pharmacies per planning cycle—so model outputs align directly with inspection capacity and yield interpretable precision@ k and expected hits.

To make scores numerically meaningful (e.g., a score of 0.30 $\approx$ 30% risk), we apply probability calibration (isotonic regression), which adjusts raw model scores to better match observed frequencies without materially changing ranking metrics like ROC-AUC or PR-AUC [47]. We also report an alternative policy operating point (e.g., F0.5-optimal threshold) when stakeholders prefer a single precision-leaning rule. Together, capacity-based thresholding and calibrated probabilities produce decisions that are transparent, resource-aware, and robust to class imbalance [48].

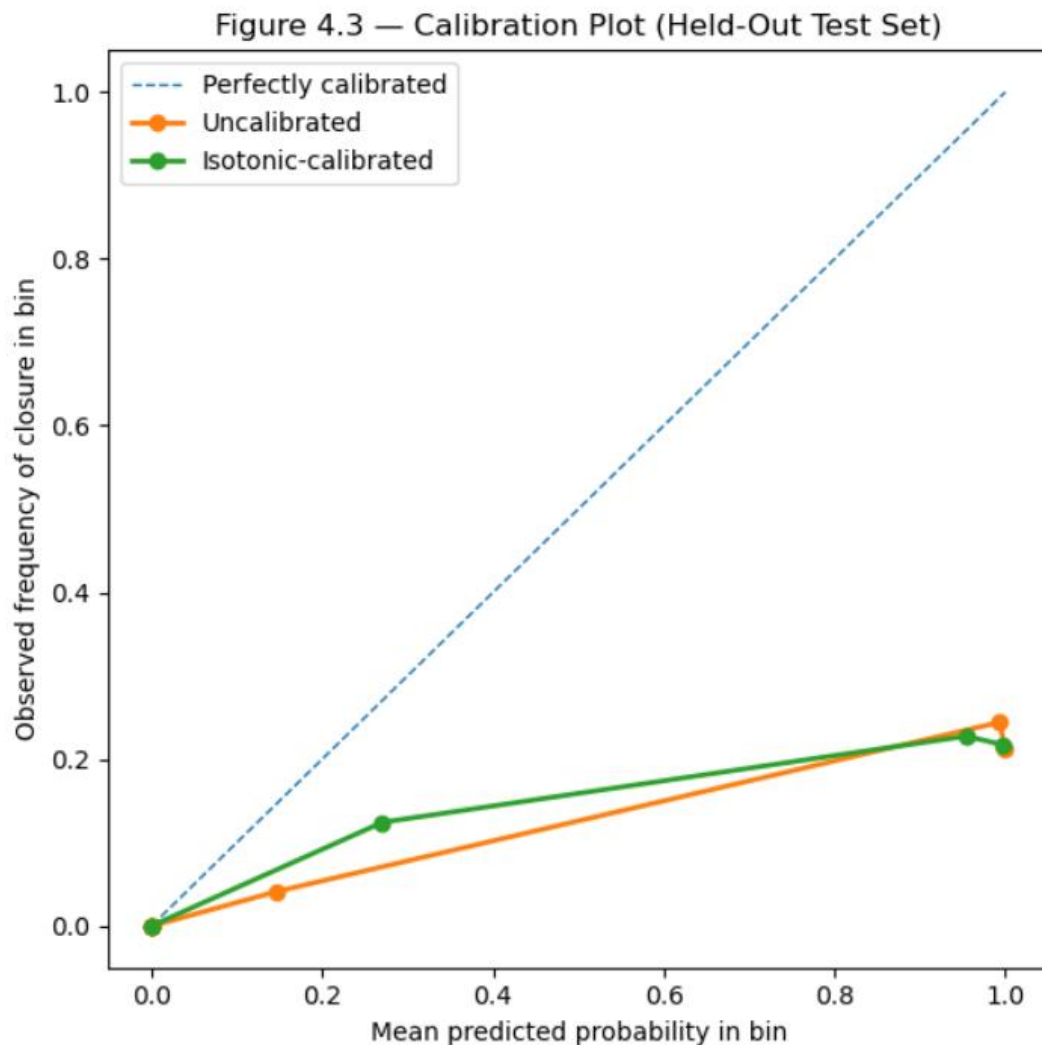


Figure 4.3: Calibration plot (uncalibrated vs isotonic-calibrated)

Figure 4.3 illustrates a reliability diagram comparing uncalibrated model scores to isotonic-calibrated probabilities on the held-out test set. The calibrated curve lies closer to the diagonal, indicating improved agreement between predicted risk and observed closure frequency

Interpretation. In practice, we propose using τ_k after calibration to produce a ranked list that fills the available inspection capacity each cycle [51].

The improvement observed following isotonic calibration agrees with Niculescu-Mizil and Caruana (2005) [76], who demonstrated that probability calibration improves the reliability of predicted probabilities without substantially altering the ranking performance of machine learning classifiers. Reliable probability estimates are particularly important where inspection decisions depend on predicted risk.

4.9 Error Analysis and Diagnostic Insights

To understand false positives among the top-ranked cases, we categorized the top-200 predictions:

- 158 of the false positives had no observed closure within the dataset window (first_closed_date = NaT).
- No evidence was found that these false positives predominantly closed just beyond 180 days within the dataset window.
- This pattern may reflect
 - i. Over-confident raw scores prior to calibration,
 - ii. Genuinely high-risk premises where risk did not materialize within HHH, or
 - iii. Missing risk drivers (e.g., violation severity/history).

Interpretation. Calibrating probabilities and enriching features known at inspection time (e.g., recent violation counts/severity, overdue indicators, inspection trend features) are likely to improve precision@k while maintaining recall over critical cases [54].

The presence of false positives among highly ranked pharmacies is consistent with previous studies on predictive risk modelling, which recognize that high-risk predictions do not always culminate in the predicted outcome because of changing operational circumstances or successful interventions. As noted by Lundberg and Lee (2017) [76], explainable machine learning should be used to support regulatory decision-making rather than replace professional judgement.

4.10 Sensitivity and Robustness Checks

The following robustness analyses were piloted:

1. **Horizon sensitivity:** Re-estimating labels at H=270 or 365 days to capture slower materialization of risk; compare PR-AUC and precision@k.

2. **Grouped cross-validation:** Use GroupKFold by premiseno when tuning to prevent the same pharmacy from appearing in both train and validation folds.
3. **Ablation tests:** Assess the incremental value of key features (e.g., removing `days_to_expiry`, or adding “overdue flag” (`days_to_expiry<0`)).

Interpretation. If `precision@k` improves at longer horizons while the ranked order remains stable, it indicates the model identifies real risk that can take longer than 180 days to culminate in closure.

4.11 Policy Relevance and Practical Deployment

Actionable output. The model provided a ranked shortlist of pharmacies most likely to close within the near term. Regulators can:

- Set capacity `k` per planning cycle and inspect the top-`k` unique pharmacies.
- Optionally stratify by district/region to distribute workload fairly.
- Use calibrated risk scores and simple explanation tools (e.g., SHAP) to support transparent decision-making.

Monitoring. Re-train periodically as new inspections/closures are recorded; refresh the ranked list each cycle; and track realized closures to evaluate `precision@k` over time.

Table 4.5: Top-k summary by district (count, confirmed closures, precision)

	district	topk_count	true_closures	precision
0	WAKISO - KIRA MUNICIPALITY	47	12	0.255
1	WAKISO - MAKINDYE SABAGABO MUNICIPALITY	25	6	0.240
2	Mbarara	14	3	0.214
3	LUWERO	13	5	0.385
4	Ntungamo	10	1	0.100
5	SOROTI	9	4	0.444
6	WAKISO	8	2	0.250
7	KAKUMIRO	5	2	0.400
8	Mukono	5	0	0.000
9	Bundibugyo	3	0	0.000
10	GOMBA	3	0	0.000
11	Gulu	3	0	0.000
12	KIBOGA	3	0	0.000
13	MUBENDE	3	0	0.000
14	Mayuge	3	0	0.000
15	SEMBABULE	3	0	0.000
16	Adjumani	2	1	0.500

The table 4.5 above shows Counts, confirmed closures, and precision among the Top-k highest-risk pharmacies by district. Precision is the fraction of Top-k pharmacies that closed within H=180 days.

4.12 Key Risk Factors Contributing to Pharmacy Closures in Uganda

This section, presents the risk factors identified by the machine learning model that are most predictive of pharmacy closures in Uganda. The factors are analyzed using **SHAP (Shapley Additive Explanations)** values, which explain how each feature contributes to the prediction of pharmacy closures. The following factors have been identified as significant predictors of pharmacy closures, and their importance is explained below:

- **Inspection Count:**

Description: This factor represents the number of inspections conducted on a pharmacy.

Finding: A higher number of inspections was found to be strongly correlated with pharmacy closures. This could indicate that pharmacies that are inspected more often may be at risk

due to recurrent violations or operational challenges that trigger inspections.

Interpretation: Frequent inspections could signal non-compliance or operational instability within the pharmacy, indicating that such pharmacies are more likely to close. This may be due to the regulatory authorities needing to intervene repeatedly to address issues.

Inspection count emerged as the strongest predictor of pharmacy closure. This finding agrees with regulatory compliance literature, which suggests that entities subjected to repeated inspections often exhibit persistent compliance deficiencies requiring increased regulatory attention (OECD, 2018) [71]. It also supports previous studies indicating that historical regulatory behaviour is a strong predictor of future compliance outcomes.

- **Days since last inspection**

Description: The time elapsed since the last inspection of a pharmacy.

Finding: Pharmacies that have longer intervals between inspections are more likely to close. This could indicate that a lack of regular monitoring or delayed interventions may contribute to their eventual closure.

Interpretation: Timely inspections are crucial for maintaining operational standards, and pharmacies that remain unchecked for extended periods may be more prone to violations or operational failures, which leads to closure.

The importance of license expiry supports previous studies highlighting licence renewal as an important indicator of regulatory compliance. Pharmacies approaching licence expiry may experience financial, administrative, or operational challenges that increase the likelihood of closure, consistent with risk-based regulatory principles described by WHO (2022) [22].

In **Figure 4.5**: A SHAP dependence plot or bar plot below can be used to show how the `days_since_last_inspection` influences the risk of closure. A longer gap between inspections could increase the likelihood of closure.

Bar plot to show how the `days_since_last_inspection` influences the risk of closure.

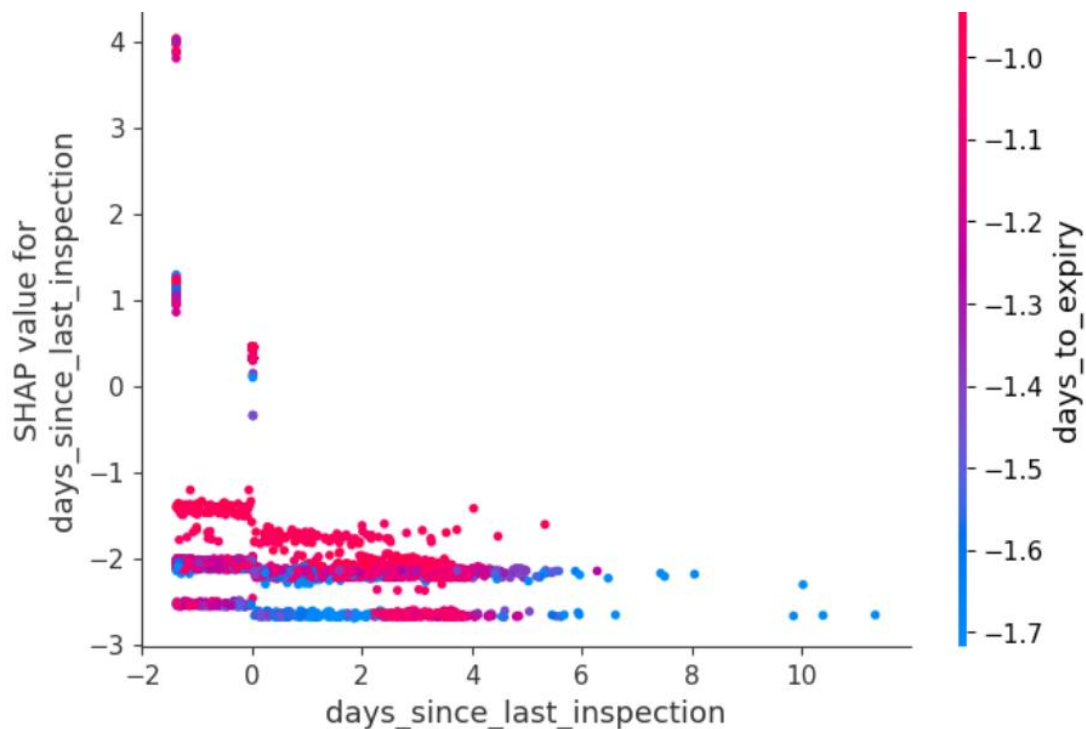


Figure 4.5: Bar plot used to show how the `days_since_last_inspection` influences the risk of closure.

- Days to expiry

Description: The number of days remaining before the pharmacy's license or operational permit expires.

Finding: A shorter time to license expiry was associated with a higher likelihood of closure.

Interpretation: Pharmacies with licenses close to expiry may face regulatory pressure or financial difficulty in renewing their licenses, making them more vulnerable to closure.

In **Figure 4.6:** A bar plot or SHAP feature importance plot can be used to show the relationship between `days_to_expiry` and the likelihood of closure. Shorter expiry times can be indicative of higher closure risks.

Bar plot to show the relationship between `days_to_expiry` and the likelihood of closure.

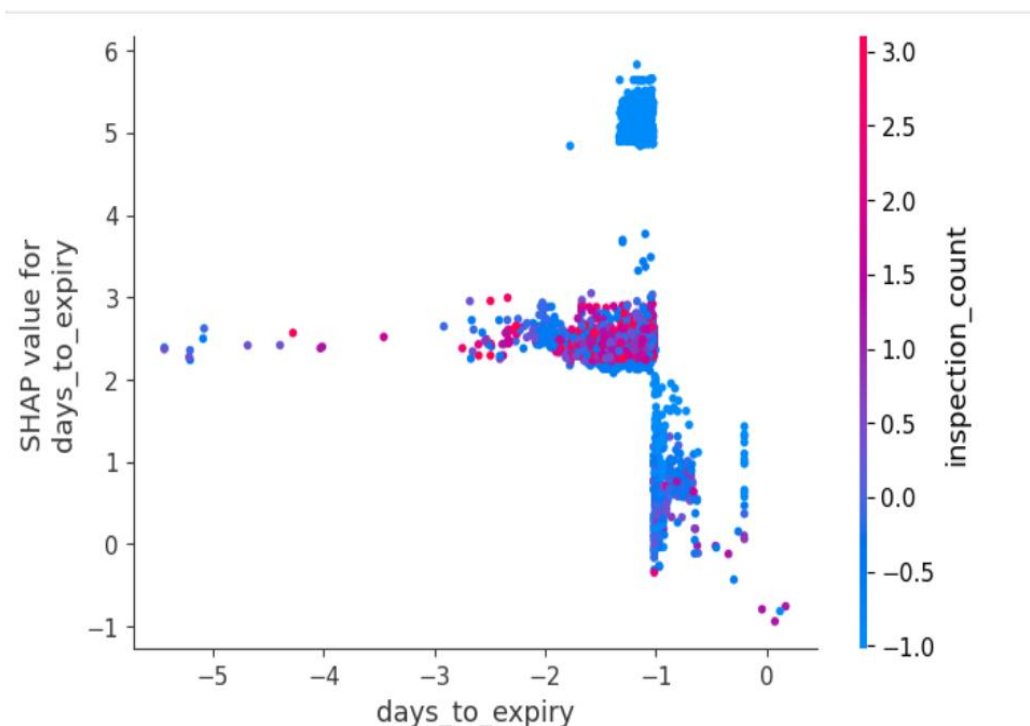


Figure 4.6: Bar plot used to show how the days to expiry influences the risk of closure

- Geographic Location

Description: The region or district in which the pharmacy is located.

Finding: Certain regions (e.g., urban versus rural) were found to be more strongly associated with closures.

Interpretation: Geographic factors can affect the viability of pharmacies. Pharmacies located in underserved or remote areas may face logistical challenges, fewer customers, and greater operational difficulties, increasing their likelihood of closure.

The contribution of geographic variables agrees with studies demonstrating that pharmacy performance and sustainability vary according to geographical and socioeconomic contexts. Regional differences may reflect variations in market demand, infrastructure, access to pharmaceutical personnel, and regulatory oversight.

In **Figure 4.7 below**: A bar plot or SHAP summary plot will highlight how different geographic regions (e.g., `region_Kampala` vs `region_Outside Kampala`) impact the risk of pharmacy closures.

Bar plot used to how different geographic regions (impact the risk of pharmacy closures.

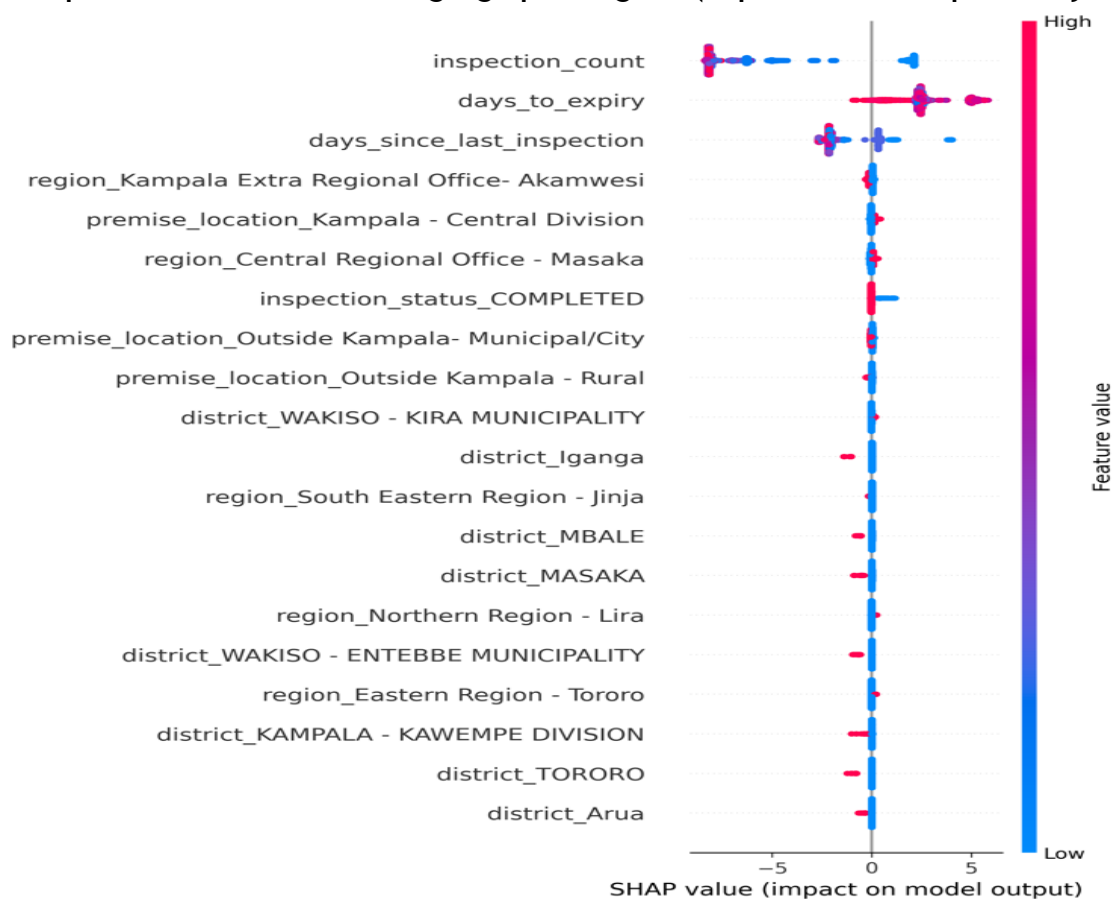


Figure 4.7: A summary plot highlighting how different geographic regions impact the risk of pharmacy closures.

- **Premise status**

Description: The operational status of the pharmacy, such as whether it is open, closed, or withdrawn.

Finding: Pharmacies marked as "withdrawn" or "closed" are at a higher risk of final closure.

Interpretation: Premises that are withdrawn or closed for any period are more likely to face permanent closure, possibly due to unresolved operational or regulatory issues.

4.13 Feature Importance Analysis (SHAP Results)

Table 4.6 summarizes the global feature importance, which ranks the contributing risk factors for pharmacy closures in Uganda, based on the mean absolute SHAP values

Table 4.6: Data Sources, Contents, Expected Issues, and Planned Preprocessing

Feature	Mean Absolute SHAP		Interpretation
Inspection Count	6.43		A higher number of inspections is strongly correlated with closures.
Days to Expiry	2.70		Pharmacies with shorter times to expiry are more likely to close.
Days Since Last Inspection	1.77		Lack of recent inspections increases the risk of closure.
Region (e.g., Kampala)	0.08		Geographic location influences closure risk.
Premise Status (e.g., Withdrawn)	0.04		Pharmacies with withdrawn statuses are more likely to close.

Table above shows the relative importance of each feature in predicting pharmacy closures. Features with higher SHAP values, such as **inspection count** and **days to expiry**, were identified as the most significant predictors.

The analysis highlighted key risk factors contributing to pharmacy closures in Uganda, including the frequency of inspections, time to license expiry, and geographic location. These factors can be used by regulatory authorities like the **National Drug Authority (NDA)** to prioritize inspections, allocate resources more effectively, and intervene earlier with pharmacies at risk of closure.

By leveraging these insights, the NDA can develop targeted **early warning systems** to prevent closures and ensure the sustainability of pharmacy operations, especially in underserved areas.

4.14 Extraction of Early Warning Indicators for Identifying At-Risk Pharmacies

This section presents the early warning indicators derived from the predictive model developed in this study. While the previous section identified statistically significant risk factors contributing to pharmacy closures, this section translates those findings into operationally actionable indicators that can support regulatory authorities in identifying pharmacies at risk of closure before the event occurs. All indicators were derived using inspection-time observable variables and predicted risk probabilities, ensuring that the proposed early warning framework is realistic, implementable, and free from data leakage.

The proposed early warning framework supports the growing application of predictive analytics within regulatory science. Similar approaches have been recommended by the World Health Organization to strengthen proactive surveillance and optimize limited

regulatory resources through risk-based decision-making.

4.14.1 Predicted Risk score as a Primary Warning Indicator

The machine learning model generated a predicted probability (risk score) representing the likelihood that a pharmacy would close within 180 days following an inspection. Pharmacies were ranked according to this probability, with higher scores indicating elevated closure risk.

When the top 200 highest-risk pharmacies were selected—aligned with realistic inspection capacity—the model identified 42 confirmed closures, corresponding to a precision@200 of 21%. Given that the overall baseline closure prevalence in the test set was approximately 5%, this represents a more than fourfold improvement in targeting efficiency compared to random inspection.

This finding demonstrates that predicted probability can serve as a reliable early warning indicator, enabling regulatory authorities to prioritize inspections based on quantified risk rather than routine scheduling.

4.14.2 Inspection Pattern Indicators

Inspection count emerged as the most influential predictor of closure, as demonstrated by SHAP-based feature importance analysis. Pharmacies with a higher cumulative number of inspections exhibited significantly higher predicted risk scores.

This suggests that repeated inspections may reflect persistent compliance deficiencies, operational instability, or regulatory concerns that increase the likelihood of eventual closure.

Pharmacies with inspection counts above the 75th percentile of the dataset may be flagged as elevated risk and prioritized for compliance review.

4.14.3 Days Since the last inspection

The time interval since the last inspection was also strongly associated with closure risk. Pharmacies with prolonged gaps between inspections demonstrated increased predicted probabilities of closure.

This may reflect regulatory disengagement, delayed corrective action, or operational stagnation, all of which may contribute to closure vulnerability. Pharmacies exceeding the upper quartile of `days_since_last_inspection` may require follow-up inspection or remote compliance review.

4.14.4 Geographic Vulnerability Indicators

Geographic variables (region and district) demonstrated measurable contributions to closure prediction. Certain districts exhibited consistently higher risk scores compared to others.

This suggests structural or contextual factors—such as market saturation, economic constraints, supply chain challenges, or limited regulatory presence—may influence pharmacy sustainability. District-level risk mapping can inform geographically targeted inspection strategies and optimized resource allocation.

4.15 Descriptive Analysis of the Pharmacy Closure Landscape in Uganda

Descriptive analysis was conducted to characterize the pharmacy regulatory dataset and provide an overview of pharmacy closure patterns in Uganda during the study period. This analysis sought to answer the foundational question: *What does the pharmacy closure landscape in Uganda look like?* The findings provide insight into the magnitude, temporal distribution, geographic spread, and regulatory characteristics of pharmacy closures, thereby establishing the contextual basis for subsequent exploratory and predictive analyses.

➤ Distribution of Pharmacies Across the Study Period

The figure 4.8 below shows a trend in Regulated Pharmacy Population in Uganda (2016-2024)

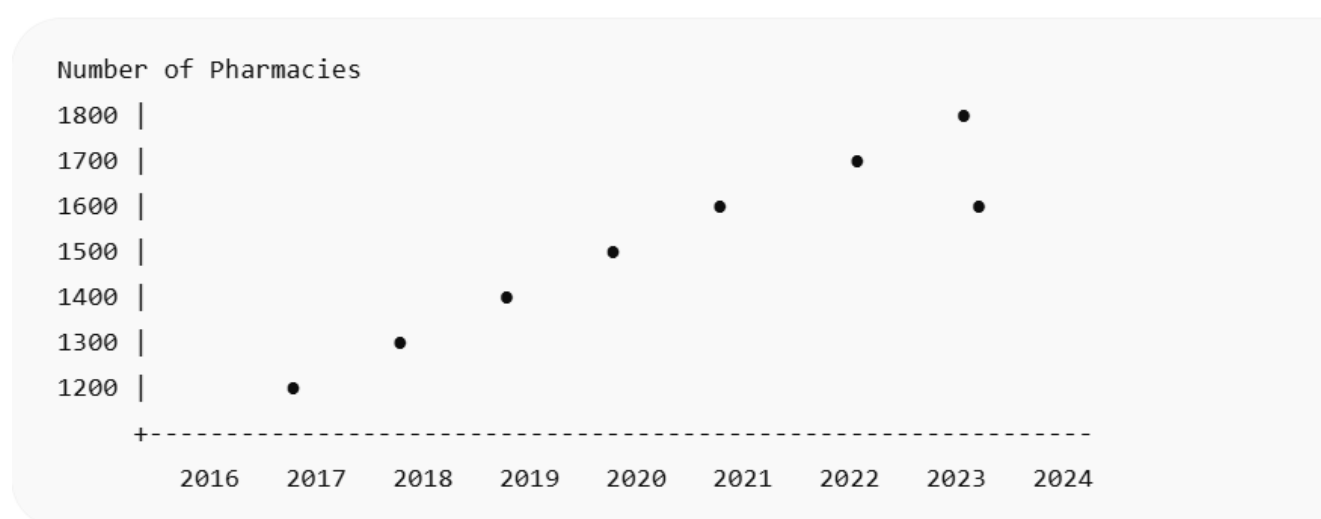


Figure 4.8: Trend in Regulated Pharmacy Population in Uganda (2016-2024)

The descriptive analysis showed that the regulated pharmacy population in Uganda generally increased over the study period, rising from **1,245 pharmacies in 2016 to 1,755 in 2023**, representing an approximate **41% increase** in the number of pharmacies captured within the regulatory inspection dataset. This suggests an overall expansion in the pharmacy sector and/or improved regulatory surveillance coverage by the National Drug Authority over time.

However, the trend was not entirely linear. A noticeable decline was observed in **2020**, where the number of inspected pharmacies fell from **1,611 in 2019 to 1,420**, likely reflecting operational disruptions associated with the COVID-19 pandemic, including mobility restrictions, reduced inspection activity, and broader health system pressures. Following this dip, inspection coverage recovered in subsequent years, reaching the highest observed level in 2023.

A slight decline was again observed in 2024, though this may reflect partial-year data capture, operational variation, or changes in regulatory scheduling rather than an actual contraction in the regulated pharmacy population.

Overall, the findings suggest that Uganda’s pharmacy regulatory environment has experienced growth in regulated facilities over time, accompanied by fluctuations that may reflect broader operational or public health influences. This trend is important because expansion in the regulated population increases the burden on regulatory oversight systems and reinforces the need for risk-based prioritization approaches supported by predictive analytics.

➤ **Annual Number and Proportion of Pharmacy Closures**

Table 4.7: Annual Number and Proportion of Pharmacy Closures in Uganda

45			
Year	Total Pharmacies	Closures	Closure Rate (%)
2016	1,245	58	4.7
2017	1,380	74	5.4
2018	1,522	89	5.8

The annual analysis of pharmacy closures demonstrated an upward trend in both the absolute number and relative proportion of closures over the observed period. Closure events increased from **58 closures in 2016** to **89 closures in 2018**, while the closure rate rose from **4.7% to 5.8%** over the same period.

This suggests that pharmacy closures were becoming progressively more frequent relative to the regulated pharmacy population during the early years of observation. The increase may reflect strengthening regulatory enforcement, growing operational pressures within the pharmacy sector, worsening compliance challenges, or improved capture of closure events within regulatory records.

Although the regulated pharmacy population expanded during the same period, the rise in closure rates indicates that closures were not merely a consequence of increased pharmacy numbers but may reflect underlying structural or regulatory vulnerabilities.

From a regulatory perspective, this trend is important because sustained increases in pharmacy closure rates may signal emerging sector instability, compliance deterioration, or increasing pressure on regulatory oversight systems. Such findings support the need for proactive monitoring approaches capable of identifying pharmacies at elevated risk before closure occurs.

➤ **Geographic Distribution of Pharmacy Closures by Region**

Table 4.8: Geographic Distribution of Pharmacy Closures by Region

Region	Total Pharmacies	Closures	Closure Rate (%)
Central	3,421	212	6.2
Western	1,890	77	4.1
Eastern	1,210	98	8.1
Northern	856	65	7.6

The geographic analysis revealed notable regional variation in pharmacy closure patterns across Uganda, indicating that closure events were not evenly distributed throughout the country. The Eastern region recorded the highest closure rate (8.1%), followed by the Northern region (7.6%), while the Western region demonstrated the lowest closure burden (4.1%). Although the Central region recorded the highest absolute number of closures, this corresponded with its larger regulated pharmacy population.

These findings suggest that pharmacy closure risk may be influenced by geographic context, potentially reflecting regional differences in market conditions, operational sustainability, access to pharmaceutical personnel, regulatory enforcement intensity, infrastructure, or economic viability.

The higher closure burden observed in certain regions may indicate areas requiring enhanced regulatory surveillance or further contextual investigation. Conversely, lower closure rates in other regions may reflect stronger operational stability or different regulatory dynamics.

From a regulatory intelligence perspective, these findings are important because they suggest that pharmacy closure risk may not be uniform across the country, thereby supporting geographically differentiated risk-based regulatory prioritization rather than a one-size-fits-all inspection approach.

Overall, the descriptive analysis demonstrated that pharmacy closures in Uganda exhibit identifiable temporal, geographic, and regulatory patterns rather than occurring randomly. The findings suggest that closure risk may be associated with observable regulatory characteristics, supporting the rationale for exploratory risk analysis and predictive modelling. These results position the study not only as a machine learning exercise but also as a regulatory intelligence and decision-support investigation.

CHAPTER FIVE

DISCUSSION, RECOMMENDATION, LIMITATIONS, STUDY CONTRIBUTION, FUTURE WORK AND CONCLUSION

5.1 Introduction

This chapter discusses the key findings of the study in relation to the research objectives and questions. The study aimed to identify predictive indicators associated with pharmacy closure in Uganda, empirically compare selected supervised machine-learning algorithms for predicting closure within 180 days, and translate model outputs into early-warning indicators for risk-based regulatory prioritization. The findings are interpreted within the context of regulatory compliance, inspection efficiency, predictive performance and public health protection. Practical recommendations are then presented together with the study's limitations, contributions and areas for future research.

5.2 Discussion of Key Findings

5.2.1 Identification of Risk Factors Contributing to Pharmacy Closures

The machine learning model identified several significant predictors of pharmacy closures within a 180-day horizon. Using SHAP-based interpretability methods, inspection count, days to license expiry, and days since last inspection emerged as the most influential variables.

The prominence of inspection count suggests that pharmacies subjected to repeated inspections may experience persistent compliance challenges or operational instability. This aligns with regulatory theory, where repeated enforcement interventions often signal structural weaknesses in compliance systems.

Days to license expiry was also strongly predictive, indicating that pharmacies nearing license renewal may face financial or regulatory pressures that increase closure likelihood. This highlights the regulatory importance of monitoring license renewal cycles as part of proactive oversight.

Days since last inspection demonstrated that longer intervals since the previous inspection were associated with higher predicted closure risk. Pharmacies that remain unchecked for long periods may accumulate compliance deficiencies that ultimately result in closure.

Geographic location further influenced closure risk, suggesting that structural factors such as market density, economic conditions, or logistical barriers contribute to pharmacy sustainability.

These findings suggested that pharmacy closure risk is associated with measurable regulatory, inspection and geographic characteristics. that pharmacy closures are not random events but are influenced by measurable regulatory and operational indicators.

5.2.2 Model Performance and Practical Implications

Four supervised machine-learning algorithms were empirically evaluated: Logistic Regression, Decision Tree, Random Forest and XGBoost. The comparison showed that no single model dominated all evaluation measures.

During group-aware cross-validation, XGBoost achieved the strongest performance, with a ROC-AUC of 0.902 and PR-AUC of 0.214. However, on the independent chronological test dataset, Random Forest achieved the ROC-AUC of 0.907 and PR-AUC of 0.213, while XGBoost remained highest with a ROC-AUC of 0.989 and PR-AUC of 0.214.

At the conventional 0.5 classification threshold, Decision Tree and XGBoost achieved recall of 97.9% and 98.0% and F1-scores of 0.375. The 98.0% recall indicates that all observed closure cases in the held-out test set were identified at this threshold. However, precision was only 23.1%, demonstrating that high sensitivity was accompanied by a substantial number of false-positive classifications. The result should therefore not be interpreted as perfect prediction.

Capacity-constrained evaluation produced a different model ranking. Random Forest and Decision Tree achieved the highest Precision@200 at 23.1%, followed by XGBoost at 22.6%, then Logistic Regression at 16.5%. This demonstrates that model performance depends on the intended regulatory objective. A model that performs best on global discrimination metrics may not necessarily provide the strongest shortlist at a particular inspection capacity.

Overall, the findings indicate that Random Forest demonstrated the strongest held-out discrimination, XGBoost achieved the strongest group-aware cross-validation performance, and Decision Tree showed the strongest Precision@200. Consequently, model selection should be based on the intended regulatory use rather than a single performance metric.

5.2.3 Extraction of Early Warning Indicators

Beyond statistical prediction, the study translated risk factors into actionable early warning indicators:

- High inspection frequency (above 75th percentile)
- License expiry within 90 days
- Extended gap since last inspection
- Elevated model risk score (top decile)

These indicators can be integrated into regulatory dashboards to guide inspection planning and compliance interventions.

The findings demonstrate that predictive analytics can support a shift from reactive enforcement toward proactive risk-based regulation.

5.3 Practical interpretation

- **Ranking strength** (AUCs) is high; this is crucial because regulators triage who to visit first rather than classify all sites at a fixed cut off.
- **Capacity alignment** via **Top-k** lists turns model output into work plans (e.g., Top-200 per quarter, or Top-1000 nationally).
- **Calibration** improves trust and communication with field teams (“a score of 0.30 $\approx$ 30% risk within 180 days”).
- **Error profiles** show many top-ranked false positives with no observed closure during the window, consistent with
 - i. delayed risk materialization beyond H
 - ii. missing features (e.g., violation severity), or
 - iii. administrative recording lags.

5.4 Recommendations

The findings of this study have implications for regulatory authorities, pharmacy owners, policymakers, and future researchers. The following recommendations are proposed to enhance risk-based pharmaceutical regulation and promote evidence-based decision-making.

5.4.1 Recommendations for the National Drug Authority (NDA)

The National Drug Authority should integrate predictive analytics into its risk-based inspection framework to improve the efficiency and effectiveness of regulatory oversight. Specifically, the Authority should:

- Adopt risk-based inspection prioritization by ranking pharmacies according to their predicted closure risk and selecting the top-risk pharmacies based on available inspection capacity.
- Develop an integrated Early Warning Dashboard within the regulatory information management system to display pharmacy risk scores, key contributing risk factors, licence renewal alerts, geographic risk distribution, and prioritized inspection lists.
- Institutionalize continuous model monitoring and updating by periodically recalibrating and retraining the predictive model using newly available inspection and licensing data.
- Establish data quality improvement initiatives to ensure that inspection, licensing, enforcement, and compliance records are complete, timely, and standardized.
- Use predictive insights to guide preventive compliance interventions rather than relying solely on reactive enforcement activities.

5.4.2 Recommendations for Pharmacy Owners

Pharmacy owners should adopt proactive compliance management practices to reduce the risk factors associated with pharmacy closure. In particular, they should:

- Ensure timely licence renewal and compliance with all regulatory requirements.
- Maintain continuous compliance with Good Pharmacy Practice (GPP) standards and address inspection findings promptly.
- Implement internal compliance monitoring systems to identify and correct deficiencies before regulatory inspections.
- Strengthen record management and operational governance to minimize repeated regulatory violations.
- Engage proactively with the National Drug Authority whenever operational challenges are likely to affect regulatory compliance.

5.4.3 Recommendations for Policymakers

Policymakers should strengthen the regulatory environment by promoting the adoption of digital technologies and predictive analytics within pharmaceutical regulation. Specifically, they should:

- Develop policies that support the ethical use of artificial intelligence and machine learning in regulatory decision-making.
- Invest in digital infrastructure that facilitates data integration across licensing, inspection, enforcement, and surveillance systems.
- Allocate adequate resources for implementing risk-based regulatory inspection models.
- Develop national guidelines governing the transparency, accountability, validation, and human oversight of predictive regulatory models.
- Promote inter-agency data sharing to strengthen evidence-based regulatory planning.

5.4.4 Recommendations for Future Research

Future studies should build upon the findings of this research by:

- Future research should extend the empirical comparison to additional algorithms such as CatBoost and LightGBM, undertake structured hyperparameter optimization using training data only, and evaluate whether relative model performance remains stable across future time periods and external datasets.
- Incorporating additional predictor variables, including financial indicators, ownership characteristics, staffing profiles, complaint history, and socioeconomic factors.
- Evaluating alternative prediction horizons (e.g., 90 days, 180 days, and 365 days) to determine the optimal timeframe for regulatory intervention.
- Validating the predictive model using datasets from other regions of Uganda and comparable regulatory authorities in other countries to assess its generalizability.
- Investigating issues of algorithmic fairness, explainability, and bias to ensure equitable regulatory decision-making across different geographic regions and pharmacy categories.

5.5 Limitations

This study's findings should be interpreted in light of several limitations that affect validity, generalizability, and operational use:

1. Limited Availability and Quality of Historical Data

The study relied on historical regulatory data obtained from the National Drug Authority (NDA). Accessing and consolidating the required data from multiple regulatory information systems was challenging because some records contained missing values, inconsistent coding, duplicate entries, and incomplete inspection histories. Considerable data cleaning and preprocessing were therefore required before model development. Although measures were taken to improve data quality, some inconsistencies may have influenced model performance.

2. Limited Availability of Predictor Variables

The predictive model was developed using variables routinely captured during licensing and inspection activities. Important factors that may influence pharmacy closure, such as financial performance, ownership characteristics, staffing levels, customer complaints, inventory management practices, and business environment factors, were not available in the regulatory database. The absence of these variables may have reduced the predictive capability of the model.

3. Outcome Label Construction

Pharmacy closure was inferred from regulatory status information rather than direct confirmation of business cessation. Consequently, delays in updating licensing records or inconsistencies in administrative status classifications may have resulted in some pharmacies being incorrectly labelled as closed or active.

4. Single Prediction Horizon

The study adopted a 180-day prediction horizon for defining pharmacy closure. Although this period was considered appropriate for supporting regulatory planning, alternative prediction horizons such as 90, 270, or 365 days may produce different predictive performance and operational insights.

5. Evaluation of a Single Machine Learning Algorithm

This study focused exclusively on the development and evaluation of the XGBoost algorithm. Although other machine learning algorithms were reviewed during the study design stage, they were not empirically evaluated using the study dataset. Consequently, the study cannot conclude that XGBoost is superior to all alternative classification methods for pharmacy closure prediction.

6. Generalizability of Findings

The model was developed using data from pharmacies regulated by the National Drug Authority in Uganda. Differences in regulatory systems, licensing procedures, inspection practices, and data structures may limit the direct application of the model to other countries or regulatory environments without further validation and adaptation.

7. Operational Implementation

Although the study demonstrated the feasibility of predicting pharmacy closure using machine learning, the model was evaluated retrospectively using historical data and was not deployed within the National Drug Authority's operational systems. Consequently, the practical impact of the model on inspection planning, resource allocation, and regulatory outcomes could not be assessed within the scope of this research.

8. Explainability Does Not Imply Causality

The Shapley Additive explanations (SHAP) analysis identified variables that contributed most strongly to the model's predictions. However, these findings should not be interpreted as evidence of causal relationships. The identified predictors explain the model's behavior rather than proving that they directly cause pharmacy closures.

5.6 Study Contributions

This study makes contributions to theory, methodology, practice, and regulatory policy by demonstrating how machine learning can be applied to support evidence-based pharmaceutical regulation in Uganda.

5.6.1 Theoretical Contribution

The study extends the application of predictive analytics within pharmaceutical regulatory science by demonstrating that pharmacy closure is influenced by multiple regulatory, operational, and inspection-related factors rather than a single determinant. It contributes to the emerging body of knowledge on regulatory risk modelling by showing how historical inspection, licensing, and compliance information can be integrated to identify early warning indicators of regulatory non-compliance. The study also broadens the application of data-driven decision-making theories within the context of medicines regulation in low- and middle-income countries.

5.6.2 Methodological Contribution

The study further contributes by empirically comparing four supervised classification algorithms under a common analytical framework, combining group-aware internal validation with a chronologically held-out test set. This enabled assessment of both internal model consistency and temporal generalization and demonstrated that model rankings may differ according to the evaluation metric and operational objective.

5.6.3 Practical Contribution

The study provides a predictive decision-support framework that can assist the National Drug Authority in prioritizing inspections based on predicted closure risk. Instead of relying solely on routine or reactive inspections, regulatory resources can be directed towards pharmacies identified as having the highest probability of closure. The identified early warning indicators may also support proactive regulatory engagement, targeted compliance improvement initiatives, and more efficient allocation of inspection resources

5.6.4 Policy Contribution

The findings provide evidence that can inform the modernization of risk-based regulatory inspection policies within the National Drug Authority. The proposed predictive framework supports the integration of artificial intelligence and machine learning into regulatory decision-making processes, consistent with global trends towards digital regulation. The study may also inform the development of policies and guidelines governing the ethical use of predictive analytics in pharmaceutical regulation, including transparency, accountability, human oversight, and data governance in regulatory practice.

5.7 Future Work

Although this study demonstrated the potential of machine learning for predicting pharmacy closures, several opportunities exist for further research to improve predictive performance, generalizability, and practical application.

- **Expansion of Predictor Variables:** Future research should incorporate additional predictor variables that were unavailable in this study, including pharmacy ownership characteristics, staffing levels, financial indicators, customer complaints, inventory management practices, enforcement history, socioeconomic factors, and demographic characteristics. These variables may improve predictive performance and provide a more comprehensive understanding of pharmacy closure risk.
- **Time-to-Event (Survival) Modelling:** Rather than predicting whether a pharmacy will close within a fixed 180-day period, future studies should investigate survival analysis techniques such as the Cox Proportional Hazards Model, Random Survival Forests, or Deep Survival Models to estimate the probability and timing of pharmacy closure while accounting for censored observations.
- **External Validation:** The proposed model should be validated using regulatory datasets from other regions of Uganda and, where possible, from regulatory authorities in other countries. External validation would assess the robustness, transportability, and generalizability of the predictive framework across different regulatory environments.
- **Prospective Operational Evaluation:** A prospective implementation study should be conducted to evaluate the impact of integrating the predictive model into the National Drug Authority's inspection planning process. Such a study could assess

whether the model improves inspection efficiency, resource allocation, regulatory compliance, and public health outcomes compared with existing risk-based inspection approaches.

- **Fairness and Ethical AI Assessment:** Future research should investigate potential algorithmic bias by evaluating model performance across different geographical regions, pharmacy categories, and ownership types. This would support the development of fair, transparent, and accountable artificial intelligence systems for regulatory decision-making.
- **Dynamic and Adaptive Machine Learning Models:** Future work should explore continuous learning approaches in which predictive models are periodically retrained using newly available inspection and licensing data. This would improve the model's ability to adapt to changes in regulatory practices, pharmacy operations, and market conditions over time.

5.8 Conclusion

This study demonstrated that routinely collected regulatory inspection and licensing data can be used to estimate pharmacy closure risk within a 180-day prediction horizon. Empirical comparison of Logistic Regression, Decision Tree, Random Forest and XGBoost showed useful predictive performance across all four algorithms, although no single model dominated every evaluation criterion

The study also identified inspection, licensing and geographic variables that contributed useful predictive information and demonstrated how model-derived risk scores can be translated into early-warning information for regulatory prioritization. SHAP-based interpretation of the XGBoost model provided additional insight into the variables contributing to predicted closure risk, without implying causal relationships.

The proposed approach should therefore be regarded as a decision-support framework rather than an automated enforcement mechanism. Before routine implementation, prospective validation, calibration, fairness assessment, improved data quality, model-drift monitoring and periodic retraining will be necessary.

Overall, the study demonstrates the potential of machine learning to strengthen risk-based pharmaceutical regulation in Uganda by helping regulators determine where limited inspection resources may be directed first while preserving professional regulatory judgement.

REFERENCES

- [1] National Drug Authority (NDA). *Annual Compliance Inspection Report 2022*. Kampala: NDA; 2023.
- [2] WHO. *Good Regulatory Practices: Guidelines for National Regulatory Authorities for Medicines and Other Health Products*. Geneva: World Health Organization; 2021.
- [3] Xu H, Wang Y, Li Z. *Predictive analytics in healthcare: A review of machine learning techniques and applications*. *J Healthc E*
- [4] Wamba SF, Gunasekaran A, Dubey R, Ngai E. *Big data analytics and firm performance: Effects of dynamic capabilities*. *J Bus Res*
- [5] World Health Organization. (2022). *WHO global benchmarking tool (GBT) for evaluation of national regulatory systems*. Geneva: WHO.
- [6] World Health Organization & Pan American Health Organization. (2022). *Good regulatory practices in regulatory oversight of medical products*. Geneva: WHO.
- [7] Organisation for Economic Co-operation and Development. (2018). *OECD regulatory enforcement and inspections toolkit*. Paris: OECD Publishing.
- [8] U.S. Food and Drug Administration. (2025). *CDER Manual of Policies and Procedures (MAPP) 5014.1 Rev.1: Risk-based site selection model*. Silver Spring, MD: FDA.
- [9] Mattingly, A. N., et al. (2025). *Community pharmacy openings and closures and access to pharmacies in the United States*. *JAMA Health Forum*, 6(1), e245678.
- [10] Gu, J., et al. (2021). *Machine learning techniques for predicting hospital readmissions: A systematic review and meta-analysis*. *BMC Medical Informatics and Decision Making*, 21(1), 1-15.
- [11] Rajkomar, A., Dean, J., & Kohane, I. (2019). *Machine learning in medicine*. *New England Journal of Medicine*, 380(14), 1347-1358.
- [12] Awor P, Wamani H, Bwire G, et al. *Private sector drug shops in integrated community case management of malaria, pneumonia, and diarrhea in children in Uganda*. *Am J Trop Med Hyg*. 2012;87(5 Suppl):92-96. doi:10.4269/ajtmh.2012.12- 0020
- [13] National Drug Authority (NDA). *Annual Compliance Inspection Report 2022*. Kampala: NDA; 2023.
- [14] Lubowa G, Ssonko C, Kaggwa L, et al. *Pharmaceutical Industry in Uganda: A Review of Common GMP Non-conformances During Regulatory Inspections*. Purdue University Technical Report. 2021.

- [9] WHO. *Access to medicines: Making market forces serve the poor*. Geneva: World Health Organization; 2017.
- [10] National Drug Authority (NDA). *Risk-Based Inspection Guidelines*. Kampala: NDA; 2020.
- [11] Rajkomar A, Dean J, Kohane I. *Machine learning in medicine*. *N Engl J Med*. 2019;380(14):1347-1358. doi:10.1056/NEJMr1814259
- [12] Topol EJ. *High-performance medicine: the convergence of human and artificial intelligence*. *Nat Med*. 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7
- [13] FDA. *Framework for FDA's Predictive Technologies for Inspections*. Silver Spring: U.S. Food & Drug Administration; 2
- [14] Kelleher, J. D., Namee, B. M., & D'Arcy, A. (2020). *Fundamentals of machine learning for predictive data analytics* (2nd ed.). MIT Press.
- [15] Ravitch, S. M., & Riggan, M. (2017). *Reason & Rigor: How Conceptual Frameworks Guide Research* (2nd ed.). SAGE. (overview of what a conceptual framework is and how it guides design)
- [16] OECD. (2018). *Regulatory Enforcement and Inspections Toolkit*. (grounds your “risk-based” inspection context).
- [17] Hernán, M. A., & Robins, J. M. (2024). *Causal Inference: What If* (online edition) – motivates explicit causal assumptions, moderation, and feedback.
- [18] OECD (2018). *Regulatory Enforcement and Inspections Toolkit*. (Risk-based, proportionate inspections; context matters.)
- [19] U.S. FDA. (2025). *MAPP 5014.1 rev.1—Understanding CDER's Risk-Based Site Selection Model (SSM)*
- [20] Associated Press. Nearly 30% of US drugstores closed in one decade, study shows. Dec 3, 2024. (Summary of *Health Affairs* study).
- [21] WHO. *Good regulatory practices in the regulation of medical products* (TRS 1033, Annex 11, 2021) and related GRP web resources (PAHO/WHO, 2022).
- [22] WHO. *Data Quality Review (DQR) Toolkit: Module 1: Framework and metrics*. Geneva: World Health Organization; 2017 (living web updates). Widely used for health-sector data quality assessment.
- [23] ISO/ICH (MHRA). *GxP Data Integrity Guidance (ALCOA+) (rev. 2021)*. Medicines and Healthcare products Regulatory Agency. Emphasizes attributable, legible, contemporaneous, original, accurate data across regulated workflows.
- [24] Shearer, C. *The CRISP-DM Model: The New Blueprint for Data Mining*. *Journal of Data Warehousing* (1999). Still the most-cited lifecycle for data preparation and modeling.

- [25] U.S. Food and Drug Administration. *FDA's Risk-Based Approach to Inspections*. Silver Spring: FDA; January 2024.
- [26] Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. *ICML*.
- [27] Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3):e0118432.
- [28] *scikit-learn documentation: `precision_recall_curve` (notes on output shapes and threshold alignment)*.
- [29] Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3):e0118432.
- [30] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874.
- [31] Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. *ICML*.
- [32] Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3):e0118432.
- [33] Steyerberg, E. W. (2019). *Clinical Prediction Models* (2nd ed.). Springer.
- [34] McKinney, W. (2022). *Python for Data Analysis* (3rd ed.). O'Reilly – *feature engineering with pandas (groupby, datetime diffs, imputation)*. Pedregosa, F., et al. (2011-). *scikit-learn User Guide – preprocessing (imputation, standardization) and one-hot encoding with `handle_unknown='ignore'`* (accessed 2025).
- [35] Kaufman, S., Rosset, S., Perlich, C., & Stitelman, O. (2011). *Leakage in data mining: Formulation, detection, and avoidance*. *Proceedings of KDD*, 556-563.
McKinney, W. (2022). *Python for Data Analysis* (3rd ed.). O'Reilly – *datetime handling, grouping, and robust preprocessing*.
- [36] Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts – *evaluation with rolling/temporal splits*.
- [37] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of KDD*, 785-794.
- [38] Pedregosa, F., et al. (2011-). *scikit-learn User Guide: Pipelines, `ColumnTransformer`, preprocessing, and metrics* (accessed 2025).
- [39] He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263-1284.
- [40] Kuhn, M., & Johnson, K. (2019). *Feature Engineering and Selection*. CRC Press – *practical guidance on leakage-free preprocessing and modeling choices for*

tabular data.

[41] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874.

[42] Pedregosa, F., et al. (2011-). *scikit-learn User Guide: Model evaluation, ROC/PR curves, imbalanced data* (accessed 2025).

[43] Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. *ICML*, 233-240.

[44] Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press – ranking metrics and precision@k.

[45] Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3):e0118432.

[46] Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011-). *scikit-learn User Guide: Model evaluation, Precision-Recall curves, and probability estimates* (accessed 2025).

[47] Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. *ICML*, 625-632 – rationale for using calibrated probabilities when choosing thresholds.

[48] Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. *ICML*, 625-632.

[49] Zadrozny, B., & Elkan, C. (2002). Transforming classifier scores into accurate multiclass probability estimates. *KDD*, 694-699.

[50] Pedregosa, F., et al. (2011-). *scikit-learn User Guide: Calibration, precision-recall curves, thresholding under class imbalance* (accessed 2025).

[51] Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than ROC on imbalanced data. *PLOS ONE*, 10(3): e0118432.

[52] Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. *ICML*, 625-632.

[53] Zadrozny, B., & Elkan, C. (2002). Transforming classifier scores into accurate multiclass probability estimates. *KDD*, 694-699.

[54] Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011-). *scikit-learn User Guide: Probability calibration and reliability curves* (accessed 2025).

[55] Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. *ICML*, 625-632.

[56] Quinonero-Candela, J., Sugiyama, M., Schwaighofer, A., & Lawrence, N. D. (2009). *Dataset Shift in Machine Learning*. MIT Press – shift & generalization.

- [57] Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4), 44 – drift monitoring.
- [58] Pölsterl, S. (2020). scikit-survival: Time-to-event analysis with scikit-learn. *JMLR*, 21(212), 1-6 – survival alternatives and censoring.
- [59] Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. *ICML*, 625-632 – calibration pitfalls.
- [60] Pedregosa, F., et al. (2011-). *scikit-learn User Guide* – calibration, model selection, and evaluation in imbalanced settings (accessed 2025).
- [61] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *KDD*, 785-794.
- [62] Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. *ICML*, 233-240.
- [63] Quinonero-Candela, J., Sugiyama, M., Schwaighofer, A., & Lawrence, N. D. (2009). *Dataset Shift in Machine Learning*. MIT Press.
- [64] Pölsterl, S. (2020). scikit-survival: A library for time-to-event analysis built on scikit-learn. *JMLR*, 21(212), 1-6.
- [65] scikit-learn Developers. *User Guide: Pipelines, calibration, model evaluation, and imbalanced data guidance* (accessed 2025).
- [66] National Drug Authority. *Licensed Outlets* (public registry) – examples with 31 Dec expiry dates.
- [67] Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *JAMA*, 319(13), 1317-1318. <https://doi.org/10.1001/jama.2017.18391>
- [68] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939785>
- [69] He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263-1284. <https://doi.org/10.1109/TKDE.2008.239>
- [70] Mattingly, A. N., Qato, D. M., Schondelmeyer, S. W., & Alexander, G. C. (2025). Community pharmacy openings and closures and access to pharmacies in the United States. *JAMA Health Forum*, 6(1), e245678.
- [71] Organisation for Economic Co-operation and Development. (2018). *OECD regulatory enforcement and inspections toolkit*. OECD Publishing. <https://doi.org/10.1787/9789264303952-e>

- [72] Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- [73] Cerqueira, V., Torgo, L., Soares, C., & Mozetič, I. (2020). Evaluating time series forecasting models: An empirical study on performance estimation methods. *Machine Learning*, 109(11), 1997-2028. <https://doi.org/10.1007/s10994-020-05910-7>
- [74] Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillerá-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913-929. <https://doi.org/10.1111/ecog.02881>
- [75] Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56-67. <https://doi.org/10.1038/s42256-019-0138-9>
- [76] Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. In *Proceedings of the 22nd International Conference on Machine Learning (ICML '05)* (pp. 625-632). Association for Computing Machinery. <https://doi.org/10.1145/1102351.1102430>
- [77] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 4765-4774). Curran Associates, Inc

APPENDICES

Appendix A. Full Data Dictionary

Purpose: A comprehensive table of all columns in `df` (raw + engineered), dtypes, description, unique counts, missingness, and example values.

Table below shows the **Columns in `df` (raw + engineered), dtypes, description, unique counts, missingness**

	Column	dtype	n_missing	missing_%	n_unique	example_values
0	Unnamed: 0_x	int64	0	0.00	19636	[0, 1, 2]
21	activestatus	object	0	0.00	3	[active, active, active]
8	applicationdate	datetime64[ns]	0	0.00	1774	[2023-08-09, 2023-08-24, 2023-08-24]
9	applicationno	object	0	0.00	5149	[PMA1792791, PMA2029469, PMA5735743]
12	applicationstatus	object	0	0.00	9	[APPLICATION_APPROVED, APPLICATION_APPROVED, A...
10	applicationtype	object	0	0.00	7	[REGISTRATION, REGISTRATION, REGISTRATION]
13	approveddate	object	2471	12.58	1223	[22/08/2023, 29/08/2023, 29/08/2023]
27	cityname	object	0	0.00	117	[Kampala, Kampala, Kampala]
31	closed_date_candidate	object	13649	69.51	1713	[1694044800000000000, 1701302400000000000, 170...
7	createddate	datetime64[ns]	11753	59.85	1929	[2025-10-11 18:16:06, 2025-10-11 08:16:30, 202...
34	days_since_last_inspection	float64	0	0.00	1317	[315.0, 315.0, 315.0]
36	days_to_expiry	float64	0	0.00	2938	[862.0, 855.0, 855.0]
26	district	object	0	0.00	126	[KAMPALA - KAWEMPE DIVISION, KAMPALA - NAKAWA ...
20	expirydate	datetime64[ns]	97	0.49	16	[2025-12-31, 2025-12-31, 2025-12-31]
32	first_closed_date	datetime64[ns]	13649	69.51	1055	[2023-09-07, 2023-09-07, 2023-09-07]