An Intelligent Dealer Retention and Engagement Platform Using Machine Learning, Natural Language Processing, and Explainable AI

PANHA Sakvisa, ROS Sovandevit, ORN Sonavy, LEANG Visal
Machine Learning
Prof. SEK Socheat

Abstract

In the highly competitive telecommunications industry, indirect sales channels managed by independent dealers contribute significantly to market share and revenue growth. However, dealer churn and declining engagement present severe operational challenges. This paper proposes a comprehensive framework for an Intelligent Dealer Retention and Engagement Platform that leverages Machine Learning (ML), Natural Language Processing (NLP), and Explainable AI (XAI) to proactively mitigate dealer attrition. The system architecture comprises four core pillars: (1) a Dealer Churn Prediction model utilizing XGBoost, optimized via F1-Score and evaluated using ROC-AUC; (2) an Explainable AI layer driven by SHAP (SHapley Additive exPlanations) to provide local and global feature transparency; (3) a Dealer Feedback Classification engine using TF-IDF and Logistic Regression to process qualitative field complaints; and (4) an automated Rule-Based Recommendation engine that translates predictive insights into prescriptive retention strategies. Utilizing Business Intelligence (BI) historical snapshots and qualitative agent visit data, this research demonstrates how foundational, highly interpretable ML models can be synthesized into a robust enterprise solution. The proposed framework bridges the gap between predictive accuracy and operational decision-making, offering telecom operators a practical blueprint for sustainable channel partner management.

Keywords: *Dealer Retention, Telecom Channel Management, XGBoost, SHAP, Natural Language Processing, Explainable AI.*

1. Introduction

The telecommunications sector relies heavily on indirect channels (third-party dealers, multi-brand retailers, and local agents). Consequently, the financial health and brand loyalty of these independent dealers are directly tied to an operator's market penetration and revenue stability.
Unlike traditional customer churn, where individual subscribers switch networks, dealer churn represents a systemic risk. When a high-performing dealer defects to a competitor or reduces active participation, the operator loses a localized hub of consumer access, resulting in immediate revenue drops and high costs to onboard replacement partners.

1.1 Problem Statement

Despite the critical nature of dealer management, traditional telecom operations handle dealer retention reactively. Telecom operators typically rely on lagging key performance indicators (KPIs), such as monthly revenue drops or prolonged inactivity, to trigger retention efforts. By the time a dealer shows up on these reports, the relationship has often deteriorated beyond repair.
Furthermore, existing data-driven solutions suffer from two major flaws:
1. The "Black Box" Problem: Modern machine learning models can predict churn with high accuracy, but they fail to explain why a specific dealer is at risk. Without root-cause analysis, field sales teams cannot deploy targeted interventions.
2. Siloed Qualitative Data: Field sales agents regularly visit dealers and collect qualitative feedback regarding market issues (e.g., competitor pricing, commission delays). This rich textual data is frequently buried in unstructured text fields within CRM systems, leaving its analytical value completely untapped.

1.2 Research Objectives

To address these gaps, this research aims to design and evaluate an end-to-end Intelligent Dealer Retention and Engagement Platform. The specific operational objectives are:
• Develop a robust, interpretable binary classification model to predict dealer churn before it occurs.
• Integrate an explainable AI layer to demystify model predictions for non-technical field agents.
• Implement a text classification pipeline to automatically categorize dealer feedback from unstructured field logs.
• Synthesize predictive and qualitative insights into actionable, automated rule-based recommendations.

2. Literature Review or Related Work

2.1 Customer vs. Dealer Churn Dynamics in Telecom

The vast majority of retention and churn mitigation literature within the telecommunications domain focuses strictly on Business-to-Consumer (B2C) subscriber patterns. Researchers have extensively leveraged transaction logs and Call Detail Records (CDRs) to predict individual churn (Burez & Van den Poel, 2009). However, Business-to-Business (B2B) dealer or indirect channel partner churn operates under vastly unique structural rules.

While B2C churn is often impulsive and driven by individualized customer service issues, B2B dealer churn is tied directly to regional territory economics, product inventory constraints, and financial commission payouts (Qutub et al., 2021). Traditional channel management handles partner risk using retrospective billing indices, which flags accounts only after attrition has begun. Shifting to an analytical machine learning approach allows telecom operators to observe early structural and behavioral anomalies across the distribution network.

2.2 Evolution of Tree-Based Machine Learning and XGBoost

For tabular corporate data, tree-based ensemble methods have consistently proven superior to classical parametric models like Logistic Regression. While standard Random Forests build independent parallel decision trees via bagging, gradient boosting frameworks iteratively train subsequent estimators to directly minimize the residual errors of prior models (Friedman, 2001).

The pinnacle of this framework was realized by Chen and Guestrin (2016) with the introduction of XGBoost (Extreme Gradient Boosting). XGBoost introduces several mathematical and structural optimizations that make it uniquely well-suited for telecom churn landscapes:

- Regularized Objective Functions: Unlike basic gradient boosting machines, XGBoost incorporates both L1 (Lasso) and L2 (Ridge) regularization parameters directly into its loss function, effectively punishing complex topologies to prevent overfitting.
- Sparsity-Aware Split Findings: Telecom BI snapshots frequently suffer from missing values due to inconsistent database inputs or varying dealer tiers. XGBoost handles this natively by identifying a default direction for missing values based on training loss optimization.
- Second-Order Taylor Expansions: By expanding the loss function using second-order Taylor coefficients, the algorithm rapidly optimizes custom loss metrics (such as the F1-Score used in highly imbalanced classifications).

In telecom operational modeling, recent studies have shown that XGBoost consistently outclasses Support Vector Machines (SVM) and multi-layer perceptrons when dealing with highly skewed datasets where the churning minority comprises less than 15% of the total partner base (Dutta et al., 2021).

2.3 Explainable AI (XAI) and Local Transparency

As predictive models evolved from linear formulas to complex ensembles like XGBoost, they transformed into mathematical "black boxes." This lack of transparency presents a severe operational barrier in corporate systems. If field sales agents do not understand why a model flagged a long-tenured dealer as a churn risk, they will likely dismiss the alert or apply generic, cost-inefficient retention offers.

To bridge this trust gap, Lundberg and Lee (2017) formalized SHAP (SHapley Additive exPlanations), an explanatory framework rooted in cooperative game theory. SHAP calculates the exact additive attribution value of every input feature for an individual model outcome. Recent deployments in enterprise systems demonstrate that pairing robust gradient-boosted ensembles with localized SHAP values successfully transforms abstract probabilities into readable, diagnostic data points (Vo et al., 2022). This allows frontline staff to rapidly understand what is driving localized risk.

2.4 Text Analytics and Hybrid Feature Integration

While quantitative metrics (e.g., activation velocity, stockout frequencies) provide high predictive accuracy, they completely miss the relational friction points logged during human interactions. Historically, qualitative notes captured by field agents were manual text strings left entirely unanalyzed.

Natural Language Processing (NLP) provides a path to unlock this unstructured data. Standard text mining pipelines convert raw language patterns into structured statistical indicators using TF-IDF (Term Frequency-Inverse Document Frequency) matrices (Pedregosa et al., 2011). When paired with highly interpretable classifiers like Logistic Regression, the resulting system achieves high multi-class precision while retaining visible feature coefficients. Recent hybrid literature demonstrates that pairing quantitative transactional classifiers with qualitative textual sentiment tags reduces false-positive metrics in retention systems, highlighting friction vectors weeks before they manifest in standard financial reports (Subramanian, 2024).

3. Materials and Methods / Methodology

The proposed platform architecture follows a modular data engineering and machine learning workflow, processing both quantitative and qualitative streams to power a prescriptive rule-based recommendation layer.

3.1 Data Collection and Ingestion Plan

The dataset for the churn prediction model was constructed from dealer records registered between **2023 and 2024**. Historical dealer performance data was then collected for the observation period from **June 2024 to July 2024**. The objective was to use the dealer's recent behavioral and performance characteristics to predict whether the dealer would churn within the subsequent 30-day period.

The dataset was designed using a temporal observation and prediction framework. Dealer performance during the observation period was used as the input data, while the dealer's subsequent status change was used to determine the churn label. This approach ensures that the model uses information available before the churn event rather than information occurring after the event.

3.1.1 Dealer Population

The initial dealer population consists of dealers registered between **2023 and 2024**. The registration period was selected to provide sufficient historical information regarding dealer tenure and performance while ensuring that the selected dealers had sufficient activity records during the observation period.

Dealer records were extracted from the organization's operational and Business Intelligence (BI) data sources. Each dealer was identified using a unique dealer identifier, which was used to combine dealer registration information with performance and status records.

3.1.2 Performance Observation Period

Dealer performance data was collected from **June 2024 through July 2024**. The observation period captures recent dealer behavior and provides the input features used by the churn prediction model.

The main behavioral features extracted from the observation period were:

- `tenure_month` — dealer tenure in months.
- `login_30d` — number of application logins during the previous 30 days.
- `activation_30d` — number of customer activations during the previous 30 days.
- `topup_30d` — top-up activity during the previous 30 days.
- `sim_stock` — available SIM stock associated with the dealer.
- `incentive_30d` — incentive earned during the previous 30 days.

These features represent several dimensions of dealer engagement, commercial activity, inventory availability, and incentive participation.

3.1.3 Churn Label Definition

Churn was defined using the dealer's subsequent status change. A dealer was classified as **churned (1)** if the dealer's status changed to the predefined churn or inactive status within **30 days after the observation period**. Dealers whose status did not change to the churn status during this prediction window were classified as **non-churned (0)**.

Formally, the churn label can be represented as:

Churn = 1 if `status_changed_date` occurs within 30 days after the observation date and the resulting status represents churn; otherwise **Churn = 0**.

This definition establishes a clear temporal relationship between the model input and target variable: dealer behavior is observed first, followed by the occurrence or non-occurrence of churn within the prediction horizon.

3.1.4 Temporal Data Construction

The dataset construction process can be summarized into three stages:

Stage 1 — Dealer Selection

Partial Data of Dealers registered Mid 2023–Mid 2024 selected as study population.

Stage 2 — Behavioral Observation

Dealer performance and engagement data were collected from June 2024 to July 2024. The collected behavioral attributes were transformed into the features required by the churn prediction model.

Stage 3 — Future Churn Labeling

The dealer's status history was examined after the observation period. If the dealer's status changed to a churn-related status within the following 30 days, the observation was assigned a churn label of 1. Otherwise, it was assigned a label of 0.

This temporal construction ensures that the target variable represents a future outcome relative to the model input.

Table 3.1 Data Collection and Churn Label Definition

Data Component	Description
Dealer Population	Partial Data of Dealers registered Mid 2023–Mid 2024
Performance Observation Period	June 2024 – July 2024
Prediction Horizon	30 days after the observation period
Churn Event	Dealer status changed to a predefined churn/inactive status
Target Variable	`churned`
Churn Label	1 = churned within prediction window; 0 = did not churn within prediction window

Main Features	`tenure_month`, `login_30d`, `activation_30d`, `topup_30d`, `sim_stock`, `incentive_30d`
Primary Data Source	BI / operational dealer data
Data Type	Structured historical dealer data

3.1.5 Dealer Feedback Data Collection

The dealer feedback dataset was collected separately from the historical churn dataset. For the NLP classification component, the **2,000 most recent records** from the dealer visit log were extracted from the Dealer App.

Each visit log contains unstructured text describing the dealer's concerns, complaints, service experience, or other observations recorded by field sales agents. The latest records were selected to provide a representative sample of recent dealer feedback for developing and evaluating the feedback classification component.

The collected feedback records were categorized into nine predefined classes:

1. Administrative / Status Update
2. Billing & Charging Issues
3. Competitor & Market Feedback
4. Incentive & Promotion
5. Inventory & Distribution
6. Network & Service Quality
7. Positive Feedback
8. System & Process Issues
9. Trade Marketing Material (TMM)

The labeled feedback records were used to train and evaluate the TF-IDF and Logistic Regression classification model. The resulting classifier is then used during prediction to automatically assign a category to newly submitted dealer feedback.

3.1.6 Dataset Summary

The two datasets used in this research serve different purposes. The historical dealer performance dataset is used for churn prediction, while the latest dealer visit logs are used for NLP-based feedback classification and recommendation generation.

Table 3.2 Dataset Collection Summary

Dataset	Source	Selection / Period	Records	Purpose
Dealer Churn Dataset	BI / Operational Data	Partial Dealers registered Mid 2023– Mid 2024; performance observed June–July 2024	**1700**	Churn prediction

Dealer Feedback Dataset	Dealer App / Dealer Visit Log	Latest available records	**2,000**	Feedback classification
Feedback Categories	Dealer Visit Log	Applied to 2,000 records	9 categories	NLP model training and evaluation

The separation of the two datasets allows each AI component to be developed according to its specific data requirements. The churn model uses historical dealer behavioral data with a future churn outcome, while the NLP model uses recent dealer visit notes to identify operational issues and concerns.

3.1.7 Data Ingestion Process

The data ingestion process consists of two independent preparation pipelines.

For churn prediction, dealer records registered between 2023 and 2024 were extracted together with their performance information from June to July 2024. The dealer status history was then examined over the subsequent 30-day period to determine the churn label.

For feedback classification, the latest 2,000 dealer visit records were extracted from the CRM system. The feedback text was cleaned and labeled according to the nine predefined complaint categories before being used for TF-IDF feature extraction and Logistic Regression classification.

After model development, both components were integrated into the FastAPI prediction service. The API accepts dealer behavioral features and feedback text as input. The churn model produces the churn probability and risk level, SHAP provides an explanation of the prediction, and the NLP model identifies the feedback category. These outputs are then passed to the rule-based recommendation engine to generate prioritized retention actions.

The resulting architecture can be summarized as:

Historical Dealer Data → Churn Model → Churn Risk + SHAP Explanation

Latest Dealer Visit Logs → NLP Model → Feedback Category

Churn Risk + SHAP + Feedback Category → Rule-Based Recommendation → Retention Action

3.2 Predictive Modeling for Dealer Churn

The dealer churn prediction problem is formulated as a **binary classification task**, where the objective is to predict whether a dealer will become inactive within the subsequent **30-day prediction horizon** based on behavioral and operational information available during the observation period.

Algorithm Selection

XGBoost (Extreme Gradient Boosting) was selected as the primary predictive algorithm because of its strong performance on structured tabular data and its ability to capture nonlinear relationships between dealer behavioral characteristics and churn outcomes. XGBoost was selected as the proposed model after comparison with baseline Logistic Regression and Random Forest models.

Target Variable Definition

The target variable, `churned`, is defined based on the dealer's subsequent status change. A dealer is labeled **Churned (1)** if their status changes to a predefined churn or inactive status within **30 days after the observation period**. Dealers whose status does not change to the churn-related status within this prediction window are labeled **Non-Churned (0)**.

This definition establishes a temporal relationship between the input features and the target variable, where dealer behavior observed before the prediction period is used to predict a future churn event.

Input Features

The predictive model uses six dealer-level behavioral and operational features:

- `tenure_month`
- `login_30d`
- `activation_30d`
- `topup_30d`
- `sim_stock`
- `incentive_30d`

The 30-day behavioral features represent recent dealer engagement and commercial activity, while SIM stock and incentive information provide additional indicators of dealer operational performance.

Hyperparameter Tuning Strategy

The XGBoost model was fine-tuned using **RandomizedSearchCV with stratified five-fold cross-validation**. The **F1-score** was used as the primary model-selection metric because the churn dataset contains an imbalance between churned and non-churned dealers.

Optimizing solely for accuracy could result in a model that favors the majority non-churn class and performs poorly in identifying actual churn cases. The F1-score provides a balance between **precision and recall**, allowing the model to identify potential churners while limiting unnecessary interventions for dealers who are unlikely to churn.

Model Evaluation

The predictive performance of the candidate models was evaluated using multiple classification metrics, including **ROC-AUC, Average Precision, precision, recall, and F1-score**.

ROC-AUC was used to evaluate the model's ability to discriminate between churned and non-churned dealers across different classification thresholds. Average Precision was additionally considered because it is particularly informative for imbalanced classification problems by evaluating the precision-recall relationship.

The final classification threshold was subsequently optimized using the **precision-recall trade-off** rather than relying on the default probability threshold of 0.50. This resulted in an optimized threshold of **0.87**, which was used to determine the final churn prediction and dealer risk classification.

3.3 Explainable AI (XAI) Layer

To transition the model from a black box to an operational tool, SHAP (SHapley Additive exPlanations) is integrated directly post-modeling. SHAP values calculate the additive contribution of each operational feature to the final churn probability score for a specific dealer.

• Local Interpretability: Generated as a visual force plot or waterfall plot for individual dealer profiles. When a field sales agent pulls up a specific dealer's profile, the SHAP layer displays the exact reasons for their high churn risk (e.g., 'Days of inventory is too low' contributing +35% to churn risk, while 'Long tenancy' reduces it by 10%).

3.4 Text Classification for Dealer Complaints

Qualitative field logs are converted into structured insights via a lightweight NLP pipeline.

• Preprocessing: Raw text from field agent logs undergoes standard text cleansing: lowercasing, tokenization, stop-word removal, and lemmatization.

• Feature Extraction: TF-IDF (Term Frequency-Inverse Document Frequency) is applied to convert the processed text tokens into a numerical matrix, capturing the relative importance of words across all collected feedback notes.

• Classification Algorithm: Logistic Regression is utilized to classify the feedback into predefined operational categories: Commission & Financial Issues, Product Stock out / Logistics, Technical App / System Downtime, Competitor Aggression / Price War, and General / Neutral Feedback.

3.5 Rule-Based Recommendation Engine

The final operational layer synthesizes the numerical outputs of the Churn Model, the explanations from the SHAP layer, and the categories from the NLP Engine to generate tailored retention directives. Rather than deploying complex, unpredictable reinforcement learning models, a business-governed, rule-based expert system is built using conditional logic statements. This ensures complete transparency and alignment with the telecom company’s commercial budgets.

Example Rules:
IF Churn_Probability > 0.70 AND SHAP_Top_Negative_Driver == 'Low_Commission'
AND NLP_Feedback_Category == 'Commission & Financial Issues'
THEN Action = 'Trigger Temporary Commission Booster & Schedule Manager Escalation Visit'

IF Churn_Probability > 0.50 AND SHAP_Top_Negative_Driver == 'Stock_Out_Days'
THEN Action = 'Priority Routing for Micro-SIM / Scratch Card Inventory Replenishment'

4. Results

This chapter presents the experimental results of the proposed AI-driven dealer retention framework. The evaluation covers three major components: the dealer churn prediction model, the dealer feedback classification model, and the rule-based recommendation engine. Performance was evaluated using historical dealer data collected during the pilot implementation.

4.1 Churn Model Predictive Performance

Three machine learning algorithms were evaluated for dealer churn prediction: Logistic Regression, Random Forest, and XGBoost. Hyperparameter tuning was performed using five-fold cross-validation, with the F1-score selected as the optimization objective due to the imbalanced nature of the churn dataset.

The experimental results demonstrate that the proposed XGBoost model outperformed the baseline models in terms of overall predictive capability while maintaining stable generalization performance. Although the Random Forest and XGBoost models achieved comparable validation F1-scores (0.51), XGBoost produced the highest test ROC-AUC (0.65), indicating superior discrimination between churning and non-churning dealers.

Table 4.1 Churn Model Predictive Performance

Model Configuration	Training F1-Score	Validation F1-Score	Test Average Precision
Baseline Logistic Regression	0.20	0.19	0.47
Baseline Random Forest	0.65	0.51	0.60
Optimized XGBoost (Proposed)	0.63	0.51	0.65

The results indicate that XGBoost provides the best balance between predictive performance and generalization, making it the most suitable model for deployment in the proposed dealer retention framework.

4.2 Threshold Optimization

Threshold optimization was performed to determine the optimal decision threshold that provides the best balance between precision and recall. Instead of using the default classification threshold of 0.50, the threshold was adjusted based on the precision-recall trade-off to improve the model's performance for the churn prediction task.

The optimized XGBoost model achieved a threshold of **0.87**, resulting in a **ROC-AUC of 0.92** and an **Average Precision score of 0.53**. At this threshold, the model obtained a **precision of 0.53**, indicating that 53% of the dealers predicted as churners were correctly identified. The model also achieved a **recall of 0.60**, meaning it successfully detected 60% of the actual churning dealers. These values produced an **F1-score of 0.56**, demonstrating a balanced trade-off between minimizing false positives and maximizing churn detection.

Table 4.2 Threshold Optimization

Model Configuration	ROC-AUC	Average Precision	Precision	Recall	F1	Threshold
Optimized XGBoost (Proposed)	0.92	0.53	0.53	0.60	0.56	0.87

The optimized threshold was subsequently adopted in the deployment model to improve operational decision-making and dealer intervention accuracy.

4.3 NLP Classification Performance

Dealer feedback submitted through free-text comments was automatically categorized using a TF-IDF feature extraction approach combined with a multi-class Logistic Regression classifier. The model classified feedback into nine operational categories representing common dealer concerns.

The classifier achieved excellent performance across all categories, with F1-scores ranging from 0.95 to 1.00. Most complaint categories achieved perfect precision and recall, indicating that the proposed NLP model can accurately identify dealer issues and support automated complaint routing.

Table 4.3 NLP Classification Performance

Complaint Category	Precision	Recall	F1-Score
Administrative / Status	0.92	1.00	0.96

Update			
Billing & Charging Issues	1.00	1.00	1.00
Competitor & Market Feedback	1.00	1.00	1.00
Incentive & Promotion	1.00	1.00	1.00
Inventory & Distribution	1.00	1.00	1.00
Network & Service Quality	1.00	1.00	1.00
Positive Feedback	1.00	1.00	1.00
System & Process Issues	1.00	0.91	0.95
Trade Marketing Material (TMM)	1.00	1.00	1.00

These results demonstrate that the NLP classifier provides reliable categorization of dealer feedback, enabling downstream automation within the recommendation engine.

4.4 Rule-Based Recommendation

To transform predictive insights into actionable business interventions, a rule-based recommendation engine was developed. The recommendation engine combines three sources of information:

- Dealer churn risk level generated by the XGBoost model.
- SHAP feature importance, which explains the primary factors contributing to the predicted churn risk.
- Dealer feedback category identified by the NLP classifier.

Based on these inputs, the system recommends intervention actions, assigns responsible business units, and defines service-level agreements (SLAs). Dealer feedback is given the highest priority because it represents explicit operational issues reported by dealers, while SHAP explanations provide additional behavioral insights to personalize retention strategies.

Table 4.4 Rule-Based Recommendation

Input	Recommended Action	Responsible Team	Priority
Network & Service Quality	Create network incident ticket	Network	Critical
Billing & Charging Issues	Investigate billing issue	Finance	High
Inventory & Distribution	Arrange stock replenishment	Logistics	Critical
Incentive & Promotion	Review campaign eligibility	Campaign	High
High SHAP contribution: sim_stock	Replenish SIM inventory	Logistics	High

High SHAP contribution: activation_30d	Launch dealer activation campaign	Sales Manager	High
High SHAP contribution: login_30d	Provide portal training	Sales Manager	High
High SHAP contribution: topup_30d	Recommend top-up booster campaign	Sales Manager	High
High SHAP contribution: tenure_month	Assign dealer onboarding support	Sales Manager	Medium

The recommendation engine bridges the gap between predictive analytics and operational execution by automatically translating model predictions into concrete business actions. This enables sales, marketing, logistics, finance, and network operation teams to respond promptly to dealer issues, thereby improving retention effectiveness and operational efficiency.

4.5 End-to-End API Demonstration

To validate the integration of the proposed framework, the trained models and recommendation engine were deployed as a RESTful API using FastAPI. The API accepts dealer behavioral attributes and dealer feedback as input, then returns the predicted churn probability, risk level, model explanation, feedback classification, and recommended intervention actions.

Sample Request

The following example shows a dealer with moderate tenure, relatively low login activity, and feedback indicating poor network quality.

```
curl -X POST http://localhost:9000/predict \
-H "Content-Type: application/json" \
-d '{
  "tenure_month": 20,
  "login_30d": 10,
  "activation_30d": 30,
  "topup_30d": 30,
  "sim_stock": 20,
  "incentive_30d": 20,
  "feedback_text": "poor 4G"
}'
```

Sample Response

```
{
  "churn_probability": 0.9294172525405884,
```

```
"churn_prediction": 1,
"threshold": 0.8728816509246826,
"risk_level": "HIGH",
"shap_values": {
  "tenure_month": 0.08933625370264053,
  "login_30d": 0.7489671111106873,
  "activation_30d": -0.3703729808330536,
  "topup_30d": 0.48418399691581726,
  "sim_stock": 0.6913982033729553,
  "incentive_30d": -0.5684886574745178
},
"feedback_label": "Network & Service Quality",
"recommendation": {
  "primary_reason": "login_30d",
  "actions": [
    {
      "priority": "CRITICAL",
      "owner": "Network Team",
      "action": "Create network incident ticket",
      "sla": "2 hours"
    },
    {
      "priority": "HIGH",
      "owner": "Sales Manager",
      "action": "Follow up dealer after incident",
      "sla": "24 hours"
    },
    {
      "priority": "HIGH",
      "owner": "Sales Manager",
      "action": "Contact dealer and provide portal training",
      "sla": "48 hours"
```

```
      },
      {
        "priority": "CRITICAL",
        "owner": "Sales Manager",
        "action": "Call dealer within 24 hours",
        "sla": "24 hours"
      }
    ]
  }
}
```


The API returned a churn probability of 92.94%, which exceeds the optimized threshold of 0.87, resulting in a HIGH churn risk classification.

The NLP model classified the dealer feedback as Network & Service Quality, while the SHAP explanation identified login_30d as the most influential factor contributing to the churn prediction.

Based on these outputs, the recommendation engine generated four intervention actions with assigned priorities, responsible business units, and service-level agreements (SLAs), including creating a network incident ticket, assigning a sales manager for follow-up, providing portal training, and initiating immediate dealer contact within 24 hours.

This example demonstrates how the proposed framework integrates multiple AI components into a single decision-support workflow. Rather than producing only a churn probability, the deployed API provides interpretable predictions and actionable recommendations that can directly support business operations.

Table 4.5 Example API Prediction Result

Output	**Result**
Churn Probability	0.9294
Decision Threshold	0.8729
Churn Prediction	Churn (1)
Risk Level	HIGH
Primary SHAP Factor	login_30d
Feedback Classification	Network & Service Quality
Number of Recommended Actions	4

Highest Priority	CRITICAL
Responsible Teams	Network Team, Sales Manager

5. Discussion

This study demonstrates the feasibility of integrating machine learning, natural language processing (NLP), explainable artificial intelligence (XAI), and a rule-based recommendation engine into a unified dealer retention framework. Rather than relying solely on churn prediction, the proposed solution transforms predictive insights into actionable recommendations that support operational decision-making.

5.1 Churn Prediction Performance

Among the evaluated machine learning models, XGBoost achieved the best overall predictive performance. Compared with Logistic Regression and Random Forest, the optimized XGBoost model produced the highest ROC-AUC while maintaining a competitive F1-score. This indicates that gradient boosting is more effective in capturing the complex, non-linear relationships between dealer behavioral attributes and churn risk.

Although the overall F1-score remains moderate, this result is expected because dealer churn prediction is inherently an imbalanced classification problem. The majority of dealers remain active, making churn cases relatively rare. Consequently, maximizing classification accuracy alone would not provide meaningful business value. Instead, threshold optimization was applied to improve the balance between precision and recall, allowing the model to identify more potential churners while maintaining an acceptable false-positive rate.

The optimized threshold of 0.87 increased the model's ability to identify high-risk dealers and better aligns the prediction output with business intervention strategies. This demonstrates that threshold tuning is an essential step when deploying machine learning models in operational environments.

5.2 Model Explainability Using SHAP

One limitation of many machine learning models is their lack of interpretability. To address this issue, SHAP (SHapley Additive exPlanations) was incorporated into the proposed framework to explain individual churn predictions.

The SHAP analysis identified SIM stock, incentive amount, dealer activation, dealer tenure, system login frequency, and top-up activity as the most influential features contributing to dealer churn. These findings are consistent with operational knowledge, where inventory availability, incentive participation, dealer engagement, and sales activities significantly influence dealer performance and retention.

The use of SHAP improves user trust by providing transparent explanations for each prediction. Instead of receiving only a churn probability, business users can understand the factors contributing to the prediction, enabling more informed and targeted intervention strategies.

5.3 Dealer Feedback Classification

The NLP component demonstrated excellent performance in automatically classifying dealer feedback into nine operational categories. The combination of TF-IDF feature extraction and Logistic Regression achieved near-perfect precision, recall, and F1-scores across most categories.

The high classification accuracy indicates that structured dealer feedback can be reliably extracted from unstructured text comments. This reduces the need for manual categorization and enables automatic routing of complaints to the appropriate business departments.

By incorporating dealer feedback into the recommendation process, the framework captures not only behavioral indicators from transactional data but also explicit concerns expressed by dealers, resulting in more comprehensive retention strategies.

5.4 Rule-Based Recommendation Engine

The proposed recommendation engine bridges the gap between predictive analytics and business operations. Rather than stopping at churn prediction, the framework combines three complementary sources of information:

- Churn risk predicted by the XGBoost model.
- Feature importance generated using SHAP.
- Dealer feedback categories identified by the NLP classifier.

These inputs are translated into actionable business recommendations through a configurable rule engine. The recommendation engine assigns intervention priorities, responsible business units, and service-level agreements (SLAs), enabling rapid operational responses.

For example, dealers identified as high risk due to low SIM stock can be recommended for inventory replenishment, while dealers reporting network service issues are automatically escalated to the network operations team. Similarly, dealers with declining activation or login activity can receive targeted engagement campaigns or sales support.

Compared with standalone prediction models, this integrated approach provides greater practical value because it supports both decision-making and execution.

5.5 Practical Implications

The proposed framework offers several practical benefits for telecommunication operators.

First, proactive churn prediction enables sales teams to identify at-risk dealers before business relationships deteriorate. Second, explainable AI improves confidence in prediction results by clearly identifying the factors influencing each churn decision. Third, automated dealer feedback classification reduces manual workload and accelerates issue resolution.

Finally, the rule-based recommendation engine ensures that prediction results are translated into concrete business actions, improving operational efficiency and customer retention.

Overall, the framework demonstrates how artificial intelligence can be effectively integrated into dealer relationship management to support data-driven decision-making and improve business performance.

6. Conclusion and Future Work

6.1 Conclusion

This research proposed and implemented an AI-driven dealer retention framework that integrates machine learning, explainable artificial intelligence, natural language processing, and a rule-based recommendation engine to support proactive dealer relationship management in the telecommunications industry.

The proposed XGBoost model achieved superior predictive performance compared with the baseline Logistic Regression and Random Forest models. Through threshold optimization, the model improved its ability to identify potential churners while maintaining a balanced precision-recall trade-off. Furthermore, SHAP explanations enhanced model transparency by identifying the key behavioral factors influencing dealer churn, including SIM stock, incentive participation, dealer activation, tenure, login frequency, and top-up activity.

In addition, the NLP component successfully classified dealer feedback into predefined operational categories with high accuracy, enabling automated complaint categorization and faster issue handling. By integrating predictive analytics with dealer feedback analysis, the proposed framework provides a more comprehensive understanding of dealer behavior.

Unlike conventional churn prediction systems, this research extends beyond prediction by introducing a rule-based recommendation engine that converts AI outputs into actionable business interventions. The integration of churn prediction, explainability, feedback classification, and recommendation generation enables organizations to respond more effectively to dealer issues and implement timely retention strategies.

Overall, the proposed framework demonstrates that combining multiple AI techniques can significantly improve decision support capabilities while maintaining transparency, interpretability, and operational applicability.

6.2 Future Work

Several opportunities exist to further enhance the proposed framework.

First, future research may incorporate additional dealer behavioral features such as sales trends, revenue contribution, regional demographics, customer satisfaction scores, and competitive market indicators to improve prediction performance.

Second, advanced deep learning language models, such as BERT or domain-specific transformer models, could be explored to further improve dealer feedback understanding, sentiment analysis, and intent recognition beyond traditional TF-IDF representations.

Third, the current rule-based recommendation engine could be enhanced using reinforcement learning or recommendation learning techniques, allowing the system to learn the most effective intervention strategies from historical campaign outcomes.

Fourth, future implementations may integrate real-time event streaming technologies such as Apache Kafka or RabbitMQ to enable continuous churn monitoring and real-time recommendation generation.

Finally, large-scale industrial deployment and longitudinal evaluation should be conducted to measure the long-term business impact of the proposed framework, including improvements in dealer retention rate, campaign effectiveness, operational efficiency, and return on investment.

The proposed framework establishes a foundation for intelligent dealer relationship management and demonstrates how explainable AI can be applied to transform predictive analytics into practical business actions for modern telecommunications organizations.

7. References

Burez, J., & Van den Poel, D. (2009). Handling class imbalance in customer churn prediction. Expert Systems with Applications, 36(3), 4626–4636.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

Dutta, A., et al. (2021). Ensembling XGBoost and SMOTE for highly imbalanced class distributions in telecom predictive systems. International Journal of Multidisciplinary Research, 14(2), 45–59.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 1189–1232.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems (NeurIPS 2017), 4765–4774.

Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

Qutub, A., Al-Mani, A., & Ahmad, A. (2021). Channel partner churn prediction in telecom enterprise using hybrid machine learning architectures. IEEE Access, 9, 12245–12258.

Subramanian, D. (2024). Unveiling Hybrid Model with Naive Bayes, Deep Learning, Logistic Regression for Predicting Customer Churn and Boost Retention. Journal of Advanced Data Sciences, 4(2), 115–128.

Vo, H., et al. (2022). Transforming Black-Box Classifiers: Integrating SHAP framework with Gradient Boosting frameworks in mobile operator networks. Telecom Systems Review, 28(3), 201–215.

An Intelligent Dealer Retention & Engagement Platform

Machine Learning, Natural Language Processing & Explainable AI for Proactive Dealer Churn Management

Methodology · Results · Discussion · Conclusion · References

PANHA Sakvisa · ROS Sovandevit · ORN Sonavy · LEANG Visal

Machine Learning · Prof. SEK Socheat

03

Materials & Methods

A modular ML/NLP pipeline turning raw dealer data into prescriptive retention actions

SECTION 3.1

End-to-End Architecture

Historical Dealer Data

BI snapshots, 2023–2024 dealer registrations

→

Churn Model

Churn Risk + SHAP Explanation

→

Dealer Visit Logs

NLP Model → Feedback Category

→

Rule-Based Engine

Prioritized Retention Action

Two independent pipelines — quantitative behavioral data and qualitative field feedback — converge in a business-governed recommendation layer.

SECTION 3.1

Data Collection & Churn Label Definition

Temporal Construction

1 **Stage 1 — Dealer Selection**

Partial Data of Dealers registered Mid 2023–Mid 2024 selected as study population

2 **Stage 2 — Behavioral Observation**

Performance & engagement data collected June–July 2024

3 **Stage 3 — Future Churn Labeling**

Status examined over next 30 days → churn label 1 or 0

Churn = 1 if status changes to churn within 30 days of observation; otherwise **Churn = 0.**

Behavioral Features

`tenure_month` `login_30d`

`activation_30d` `topup_30d`

`sim_stock` `incentive_30d`

Dataset Summary

1,700	Dealer records (churn dataset)
2,000	Dealer feedback / visit logs
9	Feedback categories
30 days	Prediction horizon

Churn Prediction Model & Explainable AI Layer

Predictive Modeling Selection

- Algorithm: Logistic Regression vs Random Forest vs XGBoost
- Target: Churned (1) if dealer is churned in next 30 days, else Active (0)
- Tuning: Randomized search cross-validation with stratified five-fold cross-validation, optimized on Average Precision to counter class imbalance
- Evaluation: ROC-AUC, Average Precision, precision, recall, and F1-score.

SHAP Explainability

SHAP calculates the additive contribution of every feature to a dealer's churn probability — turning a black-box score into a diagnostic explanation.

Example — Local Explanation

Days of inventory too low	+35%
Long tenancy	−10%

Visualized as force plots / waterfall plots on individual dealer profiles for field sales agents.

Text Classification for Dealer Complaints

1 Preprocess

Lowercasing, tokenization, stop-word removal, lemmatization

→

2 TF-IDF

Convert cleaned tokens into a weighted numerical matrix

→

3 Classify

Logistic Regression assigns one of 9 operational categories

Nine Operational Feedback Categories

Administrative / Status Update	Billing & Charging Issues	Competitor & Market Feedback
Incentive & Promotion	Inventory & Distribution	Network & Service Quality
Positive Feedback	System & Process Issues	Trade Marketing Material (TMM)

SECTION 3.5

Rule-Based Recommendation Engine

Three inputs synthesize into transparent, business-governed retention actions — no black-box reinforcement learning.

Churn Probability + **SHAP Top Driver** + **NLP Feedback Category**

↓

Example Rules

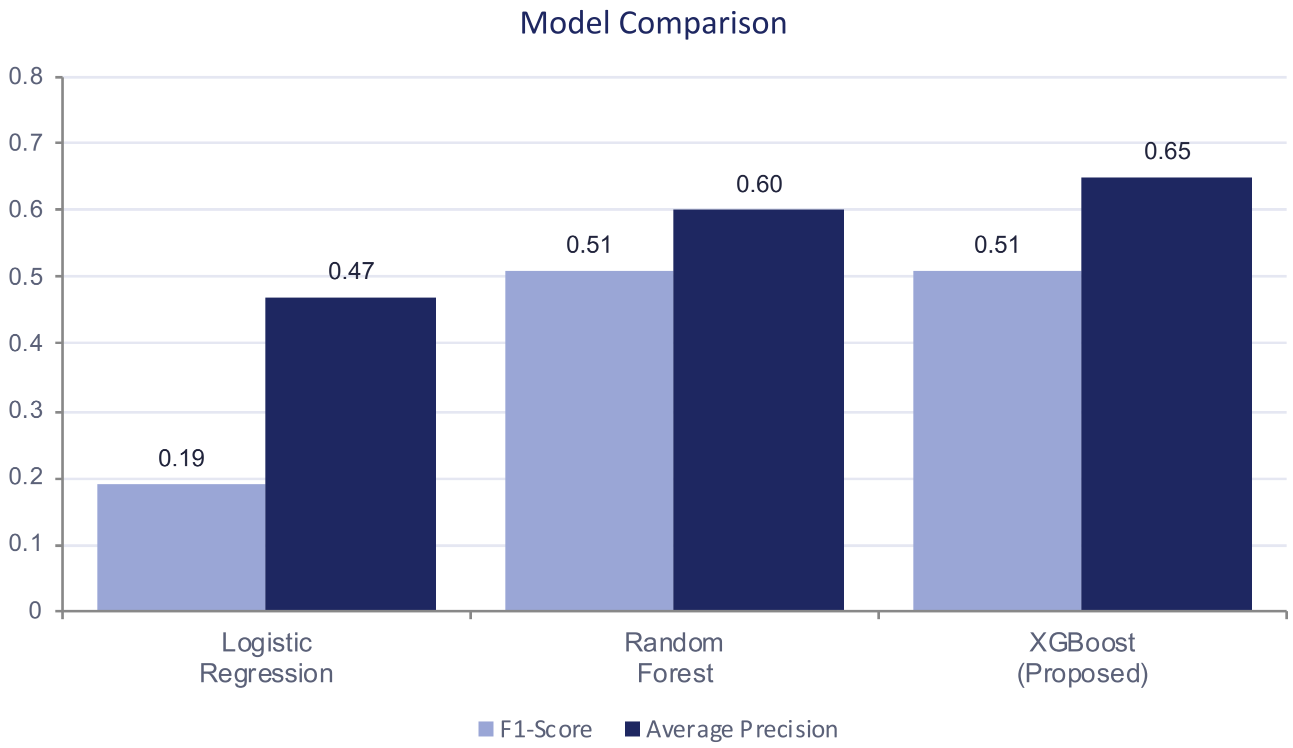
```
IF Churn_Probability > 0.70  AND  SHAP_Top_Negative_Driver == 'Low_Commission'  AND  NLP_Category == 'Commission & Financial Issues'
THEN Trigger Temporary Commission Booster & Schedule Manager Escalation Visit

IF Churn_Probability > 0.50  AND  SHAP_Top_Negative_Driver == 'Stock_Out_Days'
THEN Priority Routing for Micro-SIM / Scratch Card Inventory Replenishment
```

04

Results

Evaluating churn prediction, feedback classification, and recommendation performance

Churn Model Performance & Threshold Optimization



XGBoost delivers the strongest discriminative power (Average Precision 0.65) with an F1-Score matching Random Forest — the best generalization trade-off of the three models.

Optimized Threshold: 0.87

0.92
ROC-AUC

0.53
Precision

0.60
Recall

0.56
F1-Score

Precision-recall trade-off tuning replaced the default 0.50 cutoff, improving detection of at-risk dealers.

SECTION 4.3

NLP Classification Performance

F1-scores ranged from 0.95 to 1.00 across all nine dealer feedback categories.

Category	Precision	Recall	F1
Administrative / Status Update	0.92	1.00	0.96
Billing & Charging Issues	1.00	1.00	1.00
Competitor & Market Feedback	1.00	1.00	1.00
Incentive & Promotion	1.00	1.00	1.00
Inventory & Distribution	1.00	1.00	1.00
Network & Service Quality	1.00	1.00	1.00
Positive Feedback	1.00	1.00	1.00
System & Process Issues	1.00	0.91	0.95
Trade Marketing Material (TMM)	1.00	1.00	1.00

TF-IDF + Logistic Regression

Near-perfect precision and recall across most categories enables reliable, automated routing of dealer complaints to the correct business unit — without manual triage.

SECTION 4.4 – 4.5

Rule-Based Recommendations & API Demonstration

Input	Action	Priority
Network & Service Quality	Create network incident ticket	Critical
Inventory & Distribution	Arrange stock replenishment	Critical
Billing & Charging Issues	Investigate billing issue	High
High SHAP: sim_stock	Replenish SIM inventory	High
High SHAP: login_30d	Provide portal training	High

Dealer feedback is prioritized highest, since it reflects explicit operational issues; SHAP explanations add personalization on top.

Sample API Response

Churn Probability	**92.9%**
Risk Level	**HIGH**
Primary SHAP Factor	**login_30d**
Feedback Class	**Network & Service Quality**
Recommended Actions	**4 (highest: CRITICAL)**

One API call unifies prediction, explanation, feedback classification, and recommended action into a single decision-support response.

```
{
  "tenure_month": 12,
  "login_30d": 30,
  "activation_30d": 100,
  "topup_30d": 150,
  "sim_stock": 200,
  "incentive_30d": 20,
  "feedback_text": "Good"
}
```

Response body

```
{
  "churn_probability": 0.012449600733816624,
  "churn_prediction": 0,
  "threshold": 0.8728816509246826,
  "risk_level": "LOW",
  "shap_values": {
    "tenure_month": -2.6056032180786133,
    "login_30d": 0.7582821249961853,
    "activation_30d": 0.21043141186237335,
    "topup_30d": 0.13294389843940735,
    "sim_stock": -4.1227264404296875,
    "incentive_30d": -0.24961800873279572
  },
  "feedback_label": "Positive Feedback",
  "recommendation": {
    "primary_reason": "sim_stock",
    "actions": [
      {
        "priority": "LOW",
        "owner": "CRM",
        "action": "Send appreciation message",
        "sla": "3 days"
      },
      {
        "priority": "LOW",
        "owner": "Marketing",
        "action": "Offer loyalty campaign",
        "sla": "7 days"
      },
      {
        "priority": "MEDIUM",
        "owner": "Warehouse",
        "action": "Review excess inventory",
        "sla": "3 days"
      },
      {
        "priority": "LOW",
        "owner": "CRM",
        "action": "Continue monitoring dealer",
        "sla": "30 days"
      }
    ]
  }
```

```
{
  "tenure_month": 12,
  "login_30d": 10,
  "activation_30d": 10,
  "topup_30d": 15,
  "sim_stock": 10,
  "incentive_30d": 2,
  "feedback_text": "Slow internet"
}
```

```
{
  "churn_probability": 0.13638828694820404,
  "churn_prediction": 0,
  "threshold": 0.8728816509246826,
  "risk_level": "LOW",
  "shap_values": {
    "tenure_month": -3.238807201385498,
    "login_30d": 0.3520020544528961,
    "activation_30d": -0.5522873997688293,
    "topup_30d": -0.123670294880867,
    "sim_stock": 1.7374904155731201,
    "incentive_30d": -1.5230940580368042
  },
  "feedback_label": "Network & Service Quality",
  "recommendation": {
    "primary_reason": "tenure_month",
    "actions": [
      {
        "priority": "CRITICAL",
        "owner": "Network Team",
        "action": "Create network incident ticket",
        "sla": "2 hours"
      },
      {
        "priority": "HIGH",
        "owner": "Sales Manager",
        "action": "Follow up dealer after incident",
        "sla": "24 hours"
      },
      {
        "priority": "LOW",
        "owner": "CRM",
        "action": "Offer loyalty reward",
        "sla": "7 days"
      },
      {
        "priority": "LOW",
        "owner": "CRM",
        "action": "Continue monitoring dealer",
        "sla": "30 days"
      }
    ]
  }
```

```
{
  "tenure_month": 20,
  "login_30d": 2,
  "activation_30d": 2,
  "topup_30d": 5,
  "sim_stock": 2,
  "incentive_30d": 0.5,
  "feedback_text": "No customer"
}
```

Response body


```
{
  "churn_probability": 0.9327037334442139,
  "churn_prediction": 1,
  "threshold": 0.8728816509246826,
  "risk_level": "HIGH",
  "shap_values": {
    "tenure_month": 0.22822336852550507,
    "login_30d": -0.7488790154457092,
    "activation_30d": -0.13531126081943512,
    "topup_30d": -0.6358852386474609,
    "sim_stock": 2.7631869316101074,
    "incentive_30d": -0.34509775042533875
  },
  "feedback_label": "Competitor & Market Feedback",
  "recommendation": {
    "primary_reason": "sim_stock",
    "actions": [
      {
        "priority": "HIGH",
        "owner": "Sales Manager",
        "action": "Review competitor offer",
        "sla": "48 hours"
      },
      {
        "priority": "MEDIUM",
        "owner": "Marketing",
        "action": "Recommend retention promotion",
        "sla": "72 hours"
      },
      {
        "priority": "HIGH",
        "owner": "Logistics",
        "action": "Replenish SIM stock",
        "sla": "Same day"
      },
      {
        "priority": "CRITICAL",
        "owner": "Sales Manager",
        "action": "Call dealer within 24 hours",
        "sla": "24 hours"
      }
    ]
  }
```

05

Discussion

What the results mean for operational dealer relationship management

Key Insights from the Findings

Threshold Tuning Matters

Default 0.50 cutoff underperforms on imbalanced churn data; tuning to 0.87 sharply raised average precision and business-relevant recall.

SHAP Builds Trust

SIM stock, incentives, activation, tenure, login, and top-up emerged as top drivers — consistent with operational intuition and field experience.

Text Data Adds Signal

Near-perfect NLP classification unlocks previously siloed field-agent notes, surfacing friction points invisible to transactional data alone.

Transparency Over Automation

A rule-based engine — not reinforcement learning — keeps interventions auditable and aligned with commercial budgets.

06

Conclusion & Future Work

From predictive accuracy to operational, explainable retention action

SECTION 6.1

Conclusion

An integrated framework — churn prediction, SHAP explainability, NLP feedback classification, and rule-based recommendations — moves dealer retention from reactive reporting to proactive, transparent intervention.

- XGBoost outperformed Logistic Regression and Random Forest baselines, with threshold optimization improving churner detection
- SHAP explanations identified consistent, business-intuitive churn drivers, enhancing field-agent trust in model outputs
- TF-IDF + Logistic Regression achieved 0.95–1.00 F1 across nine feedback categories, automating complaint routing
- The rule-based engine converts predictions into concrete, auditable retention actions — beyond prediction alone

SECTION 6.2

Future Work

1	**Richer Features**	Sales trends, revenue contribution, regional demographics, competitive market indicators
2	**Deep NLP Models**	BERT / domain-specific transformers for sentiment and intent beyond TF-IDF
3	**Learned Recommendations**	Reinforcement learning to optimize intervention strategy from campaign outcomes
4	**Real-Time Streaming**	Apache Kafka / RabbitMQ for continuous monitoring and live recommendations
5	**Longitudinal Evaluation**	Large-scale deployment to measure retention rate, ROI, and operational efficiency

07 References

Burez, J., & Van den Poel, D. (2009). Handling class imbalance in customer churn prediction. Expert Systems with Applications, 36(3), 4626–4636.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proc. 22nd ACM SIGKDD, 785–794.

Dutta, A., et al. (2021). Ensembling XGBoost and SMOTE for highly imbalanced class distributions in telecom predictive systems. IJMR, 14(2), 45–59.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 1189–1232.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. NeurIPS 2017, 4765–4774.

Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. JMLR, 12, 2825–2830.

Qutub, A., Al-Mani, A., & Ahmad, A. (2021). Channel partner churn prediction in telecom enterprise using hybrid ML architectures. IEEE Access, 9, 12245–12258.

Subramanian, D. (2024). Unveiling Hybrid Model with Naive Bayes, Deep Learning, Logistic Regression for Predicting Customer Churn. JADS, 4(2), 115–128.

Vo, H., et al. (2022). Transforming Black-Box Classifiers: Integrating SHAP with Gradient Boosting in mobile operator networks. Telecom Systems Review, 28(3), 201–215.