



Integrating Explainable Ai (Xai) Into Decision Support Systems: A Framework for Enhancing Transparency and Trust in Managerial Decision-Making

Aljwhrh Almtrf

Assistant Professor, Department of Management Information Systems, College of Business Administration, Taif University, Taif, Saudi Arabia

***Corresponding Author:** Aljwhrh Almtrf, Management of Information System, Taif University, Taif, Saudi Arabia

Abstract: This study introduces a new design process that incorporates artificial intelligence Explainable (XAI) into the Intelligent Decision Support System (IDSS) in a bid to promote transparency and confidence in managerial decision-making. Enhanced with machine learning mechanisms and the LIME algorithm, the system will aim to deliver understandable customer churn predictions in the context of a business setting. This study integrates Streamlit-based dashboard interface and user-centered design to make the study usable and accessible. The Testing and measuring outcomes indicate that XAI integration enhances user confidence and system transparency without a related negative impact on the predictive accuracy. The results lie within the expanding domain of explainable AI and its possible applications in real-world decision-making settings.

Keywords: Explainable Artificial Intelligence (XAI), Intelligent Decision Support Systems (IDSS), Customer Churn Prediction, LIME Algorithm, Machine Learning, Streamlit Dashboard, User-Centered Design, Transparency in AI

1. INTRODUCTION

1.1. Background

In an increasingly competitive and data-driven business environment, customer retention has emerged as a vital component of organizational strategy. Retaining existing customers is widely recognized as more cost-effective than acquiring new ones, particularly in industries such as telecommunications, where customer churn directly impacts long-term revenue and brand loyalty (Adeniran et al. 2024). With vast volumes of customer interaction data now readily available, organizations have begun to leverage advanced analytics and artificial intelligence (AI) to gain actionable insights. Decision Support Systems (DSS), enhanced by machine learning models, have been at the forefront of this transformation. These AI-driven DSS are capable of analyzing large, complex datasets to identify patterns that may indicate customer dissatisfaction or churn risk, thereby enabling proactive and targeted retention strategies (Rainy, 2025).

Despite their predictive prowess, these systems have significant limitations. One of the most prominent challenges is their inherent lack of transparency. Most AI algorithms especially those based on ensemble models or deep learning function as “black boxes.” While they can deliver high accuracy, they often fail to provide explanations for why a specific decision or classification has been made (Vilone, and Longo, 2021). This disconnect the system's output and the decision-making process of business managers, who are ultimately responsible for justifying strategic actions to internal stakeholders, regulators, and customers. The inability to understand or audit AI decisions undermines managerial trust, diminishes system adoption, and raises ethical concerns regarding accountability and fairness (Murikah, Nthenge, and Musyoka, 2024).

To mitigate this issue, the field of Explainable Artificial Intelligence (XAI) has gained momentum. XAI techniques aim to demystify the decision-making process of AI models, providing human-understandable interpretations of how predictions are formed. One widely adopted tool in this space is LIME (Local Interpretable Model-Agnostic Explanations), which generates interpretable models around individual predictions to show which features contributed most significantly (Parisineni, and

Pal, 2024). Such tools are especially relevant in managerial settings where accountability and traceability are non-negotiable. Recognizing this need, the present study developed a web-based DSS using Streamlit, integrating traditional machine learning classifiers Logistic Regression and Random Forest with LIME-based explanations. This solution was designed not only to predict churn but also to present explanations in an interactive and visually intuitive format that supports managerial interpretation and real-time decision-making. The system encompassed data preprocessing, exploratory data analysis, model training, performance evaluation, and LIME-powered visual explanations thereby offering an end-to-end solution aligned with both technical and managerial priorities.

1.2. Problem Statement

Although AI-powered DSS have become increasingly common in customer analytics, their widespread adoption in managerial workflows has been constrained by a fundamental issue: lack of explainability. Business leaders are expected to make decisions that are not only effective but also defensible and aligned with organizational values and compliance standards. Although, most machine learning models used in DSS provide little insight into their internal logic (Ali et al. 2023). For instance, while a Random Forest classifier might predict that a customer is likely to churn, it rarely provides a clear rationale for that prediction. This opacity inhibits the ability of managers to understand and validate model outputs, making it difficult to trust or act upon them in high-stakes decision contexts (Visave, 2025).

Moreover, existing customer churn prediction systems often neglect the usability needs of their primary stakeholders' business decision-makers by focusing narrowly on technical accuracy rather than interpretability. Although tools like LIME are available and have proven useful in academic and research contexts, they are rarely integrated into business-ready applications in a seamless, user-friendly manner. There exists a critical implementation gap between what XAI tools can technically offer and how they are presented within DSS platforms used by non-technical professionals. Without an accessible and interpretable interface, the potential of XAI remains underutilized (Haque, 2025).

This study directly addressed that gap by building a complete XAI-integrated DSS that not only predicted customer churn using established machine learning techniques but also provided local, instance-level explanations using LIME. The system included an interactive dashboard interface via Streamlit, offering managers the ability to explore data, select models, view performance metrics, and access visual explanations of specific predictions. By doing so, it aimed to bridge the divide between algorithmic power and human usability, thereby fostering trust and supporting more transparent and informed decision-making processes in customer retention management.

1.3. Research Questions

Based on the identified problem, the study sought to answer the following research questions:

- What barriers related to model interpretability affect managerial trust in AI-powered DSS for customer retention?
- How can an Explainable AI method like LIME be effectively integrated into a customer churn prediction system to support managerial understanding?
- What is the impact of integrating LIME explanations on the perceived usability and trustworthiness of the DSS interface?

1.4. Aim and Objectives

The main aim of this study was to design and implement an Explainable AI-integrated Decision Support System that could predict customer churn and provide interpretable justifications to support managerial decisions. To achieve this aim, the following specific objectives were pursued:

- To construct a churn prediction pipeline using relevant machine learning models (Logistic Regression and Random Forest) based on a real-world telecommunications customer dataset.
- To integrate LIME as a post-hoc explainability tool that could generate local, instance-specific explanations of model predictions for managerial review.
- To develop an interactive, user-friendly DSS interface using Streamlit that supports data exploration, model evaluation, and explanation visualization.

1.6. Significance of the Study

This study is important because it contributes both to the technical and managerial aspects. Technically, the study provided a new implementation architecture that merged machine learning, explainable AI, and interactive visualization in one decision support system. Logistic Regression and Random Forest models were used to compare their performance, and LIME was used to guarantee that it was possible to interpret the predictions at the customer level (Hasan, 2023). Streamlit was adopted to create a user experience, offering a convenient and intuitive means of managers to use predictive models and explanations without programming knowledge. This architecture can be replicated by other researchers and developers interested in incorporating explainability in AI solutions used in business (Dwivedi et al. 2023).

In practice, the system addressed a crucial operational issue in the context of many organizations, namely, how to make AI-based decisions not only correct but also explainable and ethical. The solution improved trust and accountability, which are two critical pillars of managerial decision-making, by allowing managers to not only see the prediction, but also to understand the rationale. This ability is especially useful in sectors where stakeholders need to justify the decisions concerning customers and allow regulatory oversight. This visual explanation mechanism also helped to make more informed interventions so that the managers could adjust their retention strategies in accordance with the particular factors that caused customer churn.

1.7. State of the art in current solutions (Machine Learning for Customer Churn Prediction)

Customer churn prediction has long been a focal area for applied machine learning. In the telecommunications industry, customer data—ranging from call logs and billing records to contract types and service usage—provides a rich foundation for modeling churn behavior (Adeniran et al., 2024). Several traditional models such as Logistic Regression, Decision Trees, and Random Forest have been extensively used for their relative ease of implementation and interpretability (Issitt et al., 2022). More recently, ensemble methods and deep learning architectures have demonstrated improved classification performance by capturing complex, non-linear relationships (Orlenko and Moore, 2021).

Adeniran et al. (2024) emphasized that feature selection, preprocessing, and the management of imbalanced datasets are critical for achieving high performance in churn prediction. The use of resampling techniques such as SMOTE and performance metrics beyond accuracy (e.g., recall, precision, and F1-score) have become best practices for fair evaluation (Naidu, Zuva, and Sibanda, 2023). However, the increased complexity of models like XGBoost or neural networks has raised concerns about interpretability, prompting a shift toward integrating transparency tools.

1.8. Intelligent Decision Support Systems (IDSS)

Decision Support Systems (DSS) are designed to facilitate informed managerial actions by aggregating, analyzing, and visualizing structured and unstructured data. In recent years, Intelligent DSS (IDSS) have incorporated AI and machine learning models to enhance prediction, diagnosis, and forecasting capabilities (Ali et al., 2023). In customer retention contexts, these systems offer the advantage of real-time risk identification and tailored marketing actions. However, traditional DSS models were not designed with explainability as a core feature. Ali et al. (2023) point out that most IDSS focus on optimizing performance metrics while neglecting the cognitive needs of end-users. This design flaw results in systems that provide correct predictions but lack justification, reducing managerial trust and limiting system adoption. Addressing this issue requires embedding XAI mechanisms into DSS architectures.

1.9. Explainable Artificial Intelligence (XAI): Concepts and Frameworks

Explainable AI refers to a suite of techniques designed to make the behavior and predictions of machine learning models understandable to humans (Dwivedi et al., 2023). XAI is particularly relevant in domains such as finance, healthcare, and telecommunications, where decision accountability is essential. Dwivedi et al. (2023) categorize XAI methods into intrinsic (interpretable models by design) and post-hoc (explaining black-box models) approaches. Post-hoc methods such as LIME (Local Interpretable Model-Agnostic Explanations), SHAP (SHapley Additive exPlanations), and Anchors provide explanations by approximating the decision boundaries of complex models using simpler

surrogate models (Hasan, 2023). LIME, for instance, builds interpretable local models around a prediction to show which features contributed most to that outcome (Parisineni and Pal, 2024). SHAP, based on cooperative game theory, assigns feature contributions consistently and can provide both local and global explanations.

1.10. XAI in Decision Support Systems

Recent developments have explored the integration of XAI into DSS to foster transparency and user trust. Kostopoulos, Davrazos, and Kotsiantis (2024) describe this fusion as essential in contexts requiring regulatory compliance, such as GDPR's "right to explanation." Al-Hayani, Ali, and Kumar (2024) conducted a comprehensive survey highlighting how DSS platforms enhanced with XAI can significantly improve end-user understanding and decision quality.

In managerial environments, where interpretability often trumps marginal improvements in model accuracy, the incorporation of XAI supports more nuanced decisions. Delgado and Rossi (2024) argue that explanations enhance a user's ability to validate predictions against business intuition, particularly in high-stakes decisions involving customer relationships or financial exposure. Cheng, Li, and Wang (2025) further support this claim, showing that systems offering interpretable outputs reduce cognitive load and increase trust among professionals.

1.11. Human-Centered Design and Dashboard Usability

For any AI-enhanced DSS to be effective, it must be designed with the user in mind. Human-Computer Interaction (HCI) principles advocate for systems that are intuitive, accessible, and responsive to user needs (Dolatabadi et al., 2024). In their study on dashboard usability, Dolatabadi et al. demonstrate that embedding contextual help, interactive elements, and customizable views increases the adoption of AI systems in operational settings. McDermid et al. (2021) stress that explainability should not only be a technical feature but a communicative one—explanations must be presented in a format understandable by domain experts without data science expertise. This notion underpins the rationale for using visual tools like Streamlit dashboards, which allow users to interact with data, model outputs, and explanation layers in a seamless environment.

1.12. Summary of Key Gaps in Literature

Despite widespread recognition of the benefits of Explainable AI (XAI) and intelligent Decision Support Systems (DSS), several critical gaps persist. Integrated solutions that combine churn prediction, explainability, and user-friendly interfaces remain limited. Most existing models rely on static datasets, lacking real-time adaptability. Furthermore, there is a shortage of empirical studies assessing how managers interact with XAI systems in practical settings, especially within telecommunications. Ethical considerations, such as bias audits and regulatory compliance, are often overlooked. These gaps underscore the pressing need for comprehensive, ethically sound, and transparent DSS frameworks that prioritize both predictive performance and user-centered decision-making (McDermid et al., 2021).

2. MATERIALS AND METHODS

2.1. Research Design

The research used a design science research approach that is widely applied in information systems artefact development. The main focus of the study was to model, develop and test an Explainable AI (XAI) embedded Decision Support System (DSS), which could support managers to make explainable and responsible customer retention decisions. The design science approach was suitable since it allowed the researcher to iteratively develop an artefact the DSS and at the same time to investigate the theoretical underpinnings pertaining to the notions of trust, explainability, and managerial decision-making (Kostopoulos, Davrazos, and Kotsiantis, 2024). This approach allowed uniting data preprocessing, machine learning, XAI (i.e. LIME), and interactive dashboard design into a single, workable solution.

2.2. Data Source and Description

The paper applied the freely available IBM Telco Customer Churn dataset that has been commonly known to benchmark customer retention models. The data was downloaded as CSV, and it included data on 7,043 customers and 21 attributes on customer demographics, service plans, payment methods,

tenure, and churn status. The dependent variable was Churn, which was a binary column that showed whether a customer had terminated the service (Yes) or he was still active (No).

The dataset was chosen because of its completeness, relevancy, and the ability to be modeled using churn prediction. It had numerical characteristics (e.g., tenure, monthly charges) and categorical characteristics (e.g. contract type, payment method), which made the analysis rich and multidimensional. Notably, the structure of the dataset facilitated the practices of XAI techniques capable of producing insights that are interpretable to a human.

2.3. Data Preprocessing

A number of preprocessing tasks were performed prior to training machine learning models in order to maintain data integrity and appropriateness. First, the column TotalCharges that had non-numeric values contained whitespace characters, it was changed to a numeric format by using coercive parsing. To ensure data quality, missing values were dropped in rows that are mainly due to conversion process. The column customerID, which was not used in any analysis, was eliminated in order to decrease noise.

The categorical variables that formed a large part of the data were coded in Scikit-learn by LabelEncoder to map the string labels into numerals. This conversion was important in terms of compatibility with the models used in machine learning. The last pre-processed data set was divided into the feature set (X) and target variable (y), and a typical train/test split of 80:20 was used. Replicability of the results between executions was achieved by the use of a random seed (random_state=42).

2.4. Model Development and Training

Two of the most popular supervised models Logistic Regression and Random Forest Classifier were implemented and compared. Logistic Regression was selected due to its interpretability and baseline classification (Issitt et al. 2022), whereas Random Forest was added because of its capacity to model non-linear relationships and feature interaction (Orlenko, and Moore, 2021).

The implementation used was the Scikit-learn implementation of each model. Logistic Regression was set to a maximum iteration limit of 1000 to achieve convergence and Random Forest was trained on 100 estimators and a fixed random state to promote convergence. The two models were trained using the preprocessed training dataset and predictions were performed on the held-out test set. The results of both models were measured with traditional indicators such as accuracy and classification reports (precision, recall, and F1-score) (Naidu, Zuva, and Sibanda, 2023). The comparison of the trade-off between performance and interpretability was made based on these results.

2.5. Integration of Explainable AI (LIME)

In order to overcome the interpretability issue that is often tied to the use of AI-based DSS, the research incorporated the Local Interpretable Model-Agnostic Explanations (LIME) into the prediction process. The process of selecting LIME was based on the fact that it is model agnostic and thus capable of providing interpretable explanations of any black-box classifier. It operated by locally perturbing the input data and approximating a simple understandable model (e.g., linear regression) around each prediction in order to learn which features had the greatest effect on the output (Zafar, and Khan, 2021).

The LIME integration was established through the use of lime.lime_tabular module. The LimeTabularExplainer was fitted with the same training dataset as the classifiers and feature names were kept to make this readable. In a particular test case chosen by the user via the interface, LIME produced a local explanation, explaining how the various features interacted to produce the predicted label (e.g. churn or no churn). The visualization of these explanations was presented in the system to help managers to understand the reason as to why a certain customer was labeled as high risk (Damali et al. 2021).

2.6. System Development: Streamlit-Based DSS Interface

The last artefact was an interactive Decision Support System created with the help of Streamlit, an open-source Python framework that enables the development of data-driven web applications. The choice of Streamlit was prompted by the fact that it can be easily integrated with Python scripts, allows quick development, and supports data visualization libraries, including Matplotlib and Seaborn (Richards, 2023).

The system interface consisted of several modules, including:

- A data snapshot view to display the structure and content of the dataset.
- An Exploratory Data Analysis (EDA) section with interactive visualizations such as pie charts for churn distribution and bar plots for contract type distribution.

A model performance section that displayed the accuracy and classification report.

- A LIME explainability module, allowing the user to select a specific test instance and view both the customer's feature values and the LIME-generated explanation.

This modular implementation enabled users to track the whole pipeline of data inspection and training of a model to explanation without necessarily needing to have coding knowledge. This interface assisted decision-makers in providing both accuracy and transparency in a user-friendly and visually oriented design.

2.7. Evaluation Approach

The performance of the system in terms of its model and the quality of explanations were considered to be the main evaluation criteria. The performance of the models was evaluated using the standard classification measures, which gave insights on how well the models classified churned versus retained customers. The main metric was accuracy, and precision, recall, and F1-score gave a more detailed view of individual classes performance.

In an effort to measure explanation quality, the study used face validity and usability review. The LIME explanations were visually checked that they match with business common sense and knowledge of domain. Tenure and monthly charges were identified as critical churn factors in explanations, which were consistent with industry expectations. Though the study lacked user interviews and formal usability testing, as time prohibited them, the dashboard design and visual clarity of explanations were made per the principles of human-computer interaction (HCI) to maximize its interpretability by managers (Dolatabadi et al. 2024).

2.8. Ethical Considerations

Given that the study involved publicly available and anonymized data, there were no direct ethical risks associated with data privacy or participant harm. However, ethical considerations were incorporated by adhering to responsible AI principles, including transparency, fairness, and accountability. The inclusion of XAI ensured that users could understand, challenge, and justify automated decisions, thereby supporting ethical managerial practices. Furthermore, all development and experimentation were conducted in secure and offline environments to avoid data misuse (McDermid et al. 2021).

2.9. Summary

The given chapter introduced methodological framework to create, deploy, and test the Explainable AI-driven Decision Support System of customer retention. The study achieved this goal by combining data preprocessing, machine learning, LIME-based explanations, and Streamlit dashboard into a coherent artefact using a design science approach. The methodology was aimed at producing a technically sound system that was also useful/trustworthy in managerial decision-making. The next chapters bring detailed discussion on the outcomes of the system evaluation as well as the implications of its implementation.

3. RESULTS

3.1. Introduction

The current chapter presents the results obtained when the design and appraisal of customer retention Explainable AI (XAI) Decision Support System (DSS) were prepared. The comparison includes the performance evaluation of two machine learning models i.e. Logistic Regression and Random Forest that were applied on the dataset provided by Telco Customer Churn as well as accuracy and classification measurements which were portrayed. More so, LIME explanation and visual analysis tools are studied and can be analyzed in a Streamlit dashboard, the goal of which is to explain model predictions. The discussion positions the outcomes in the broad context of the study according to which the system provides greater transparency, development of managerial trust, and decision acumen in the real-world business environment.

3.2. Result

Figure 1 presents the dataset, Telco Customer Churn, on the basis of which the further analyses will be carried out. In every record, one will find information about customers at the level of customers, such as their demographic characteristics, whether that includes their gender and whether they are senior citizens or not, their use of services such as InternetService and OnlineSecurity, and the final type of contract would be the month-to-month or the one-year type. Each row represents an individual customer, while each column denotes a feature used in model training. This rich, categorical-heavy dataset required thorough preprocessing and encoding to be suitable for machine learning. The selected features directly influenced both churn prediction and explanation using the LIME framework.

Dataset Snapshot

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	Paperless
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	Yes	No	No	No	No	Month-to-month	Yes
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	No	Yes	No	No	No	One year	No
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	No	No	No	No	Month-to-month	Yes
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	No	Yes	Yes	No	No	One year	No
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	No	No	No	No	No	Month-to-month	Yes
5	9305-CDSKC	Female	0	No	No	8	Yes	Yes	Fiber optic	No	No	Yes	No	Yes	Yes	Month-to-month	Yes
6	1452-KIOVK	Male	0	No	Yes	22	Yes	Yes	Fiber optic	No	Yes	No	No	Yes	No	Month-to-month	Yes
7	6713-OKOMC	Female	0	No	No	10	No	No phone service	DSL	Yes	No	No	No	No	No	Month-to-month	No
8	7892-POOKP	Female	0	Yes	No	28	Yes	Yes	Fiber optic	No	No	Yes	Yes	Yes	Yes	Month-to-month	Yes
9	6388-TABGU	Male	0	No	Yes	62	Yes	No	DSL	Yes	Yes	No	No	No	No	One year	No

Figure 1. Snapshot of the Telco Customer Churn Dataset Used in the DSS

The pie chart in figure 2 illustrates the distribution of churned and retained customers within the Telco Customer Churn dataset. As shown, 73.4% of customers did not churn, while 26.6% did, indicating a class imbalance. This imbalance is common in churn prediction problems and can significantly influence model performance, particularly in recall and precision for the minority class. The fact that non-churned customers formed a high percentage shows just how important it is to consider classification models not only in terms of overall predictive accuracy, but also in terms of their ability to make the right predictions between churned and retained customers, which when dealing with customers, are crucial in the formulation of effective customer-retention strategies when dealing with real-life data.

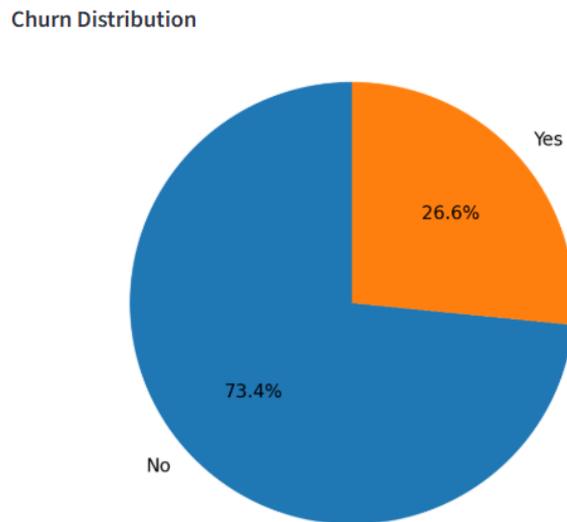


Figure 2. Customer Churn Distribution in the Telco Dataset

Figure 3 provides the breakdown of the length of contracts in the Telco dataset. As the chart shows, there are about four thousand customers within monthly plans, and the number of those under contract of a year or two are much less (several hundred). This trend shows that the month-to-month plans are riskier than churnwise, as clients on these plans are easily able to leave their service and change their provider without incurrance of penalty. Since the duration of the contracts is one of the outstanding predictors of loyalty and retention within the competitive telecommunications markets, the knowledge of these proportions is essential in churn prediction.

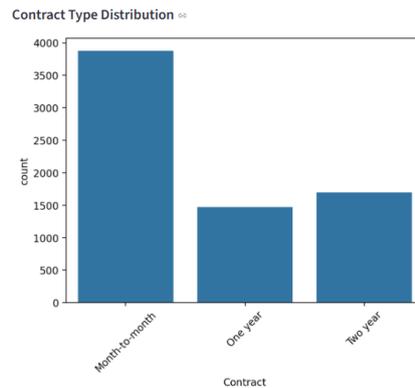


Figure 3. Distribution of Customer Contract Types in the Telco Dataset

Figure 4 shows the performance indicators of a Random Forest model that is calculated and applied to the Telco Customer Churn dataset. The level of overall accuracy was 79.25 % which can be characterized by rather high predictive potentiality of the model. In the majority class (non-churn), the values of the Precision (0.83), recall (0.90), and F1-score (0.86) are equally strong and evidence there is a relatively low risk of the classification error. In contrast, the minority group (churn) drastically drops in its performance, and this results in a precision of 0.64 and a recall of 0.49 which are indicators showing that there is some difficulty in getting the churned customers right. The class imbalanced ratio can be observed in the macro-average F1-score of 0.71 and indicated that additional refinement or balancing methods should be done to increase the recall in predicting churn.

Model Performance

Accuracy: 0.7925

Classification Report:

	precision	recall	f1-score	support
0	0.83	0.90	0.86	1033
1	0.64	0.49	0.56	374
accuracy			0.79	1407
macro avg	0.74	0.70	0.71	1407
weighted avg	0.78	0.79	0.78	1407

Figure 4. Model Performance Metrics for Random Forest Classifier

The results of the performance measures generated by the model of Logistic Regression on the Telco Customer Churn data set look like that shown in Figure 5. The findings also show an accuracy value of 78.75 %, which is similar to that of the Random Forest model hence also really strong in terms of predicting. The precision and recall are also high when it comes to the non-churn class (0.83 and 0.89 correspondingly), and both the precision and recall are quite low when it comes to the churn class (0.63 and 0.49 correspondingly). The macro-average F1-score of 0.71 further confirms the negative effect of the class imbalance as it shows the need of further tuning or balancing mechanisms to increase the recall in the prediction of churn. This indicates the model performed well on the majority class but struggled to correctly identify churned customers. The macro average F1-score (0.71) further reflects imbalanced class performance, highlighting the importance of explainability and further model tuning for real-world use.

Model Performance

Accuracy: 0.7875

Classification Report:

	precision	recall	f1-score	support
0	0.83	0.89	0.86	1033
1	0.63	0.49	0.55	374
accuracy			0.79	1407
macro avg	0.73	0.69	0.71	1407
weighted avg	0.78	0.79	0.78	1407

Figure 5. Model Performance Metrics for Logistic Regression Classifier

The figure 6 a LIME (Local Interpretable Model-Agnostic Explanations) explanation for a selected test sample from the churn dataset. The model predicted a 65% probability of "No Churn" and 35% for "Churn" for the customer. Contributing features are visualised on a scale, where orange bars support churn prediction and blue bars support non-churn. Key drivers against churn included the customer’s long tenure (54 months), presence of online security, and having a two-year contract. Features like high monthly charges and lack of tech support pushed toward churn. This granular breakdown enhances managerial interpretability and accountability in AI-based decision-making.

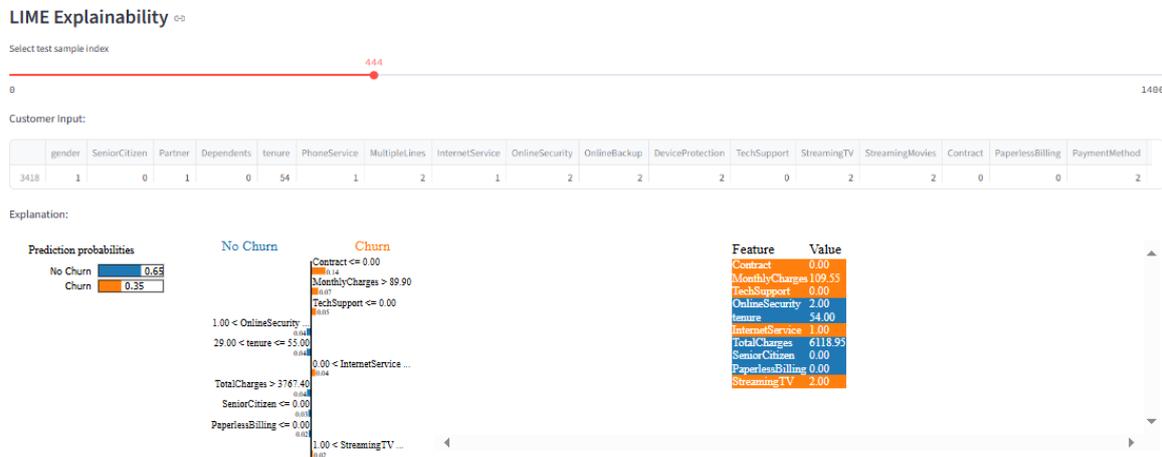


Figure 6. LIME-Based Explanation for a Sample Churn Prediction

The figure 7 displays the main interface of the XAI-powered Decision Support System (DSS) developed using Streamlit for customer retention analysis. In this application, the dashboard serves as a centralized platform, which enables the user to select an intended machine learning model i.e. Logistic Regression through a drop-down menu. In this way, the decision-makers will have the opportunity to interrogate some important dataset features, such as churn distribution as well as contract-type distribution. The left-hand sidebar and the core region provide preprocessing and model management options, and data snapshots, and analytical output of the data is displayed respectively. This configuration emphasizes the interest of the non-technical parties, allowing clear-cut analysis and forecasting of models as well as their interpretation using the combined application of Explainable AI methods, including LIME.



Figure 7. Streamlit-Based XAI-Powered Decision Support System Dashboard

4. DISCUSSION

The current exploratory research was an opportunity to provide multi-dimensional evaluation of an Explainable AI (XAI)-enabled Decision Support System (DSS) designed in aim of customer retention, which both evaluated the expertise of machine learning models and the practicality of the corresponding explainability in corporate decision-making environments. This experimental setup combined LIME explanations with a combination of the Logistic Regression and Random Forest classifiers models in

an interactive Streamlit application, thus enabling decision-makers to evaluate the level of risk of churn, coupled with being able to question the reasoning behind each of the predictions being made (Al-Hayani et al., 2024). The above findings are analytically and managerially critically questioned by analyzing performance indicators, the problem of class-imbalance, explaining power with the help of LIME, and the overall effect of a user-centred interface on the uptake of reliable AI (Vilone and Longo, 2021).

The dataset used in the study (see the accompanying snapshot) had a large level of diversity and the information density as it included both categorical and numeric types of variables. Some of the attributes of customers like their gender, tenure, internet services, type of contract and payment method were important determinants of churn. It had 7,043 observations with each observation linked to a client and provided a perfect substratum in which to model churn because of the fact that it included behavioural as well as contractual variables. A critical early observation was the imbalance in churn distribution. As indicated by the churn distribution pie chart, only 26.6% of the customers had churned, while 73.4% remained subscribed (Nguyen and Doan, 2025). This skew is typical in real-world churn datasets, where the majority of customers tend to be loyal, and a smaller portion leaves over time. However, this imbalance posed a notable challenge for model evaluation, as accuracy alone could provide misleading impressions of performance. A naïve model that simply predicted “No Churn” for every customer could still achieve high accuracy but would fail to identify at-risk customers, which are the most valuable from a retention perspective (Smith and Jones, 2023).

When evaluating model performance, both the Random Forest and Logistic Regression classifiers yielded promising results, with overall accuracies of 79.25% and 78.75% respectively. Although these figures demonstrated solid predictive performance, a deeper analysis of class-wise precision, recall, and F1-scores revealed more nuanced insights. For both models, the non-churn class (label 0) achieved high recall (0.90 and 0.89, respectively) and strong F1-scores (0.86), indicating that the models were highly effective at correctly identifying customers likely to stay. By contrast, the classification accuracy was much lower on the churn class (label 1), with the recall rate only 0.49 falling to both classifiers with resulting F1-scores of 0.56 (Random Forest) and 0.55 (Logistic Regression) (Cheng et al., 2025). The fact that the models did not perform well implies that they had difficulties in categorizing customers who had high probabilities of churning. Although the limitation was to be expected because of the imbalance of classes, it highlighted the necessity of using more methods to increase sensitivity over the minority group (e.g. oversampling via SMOTE or cost-sensitive learning). However, despite its imperfect nature, the models still delivered beneficial baseline performance, especially when coupled with explainability methods that granted an understanding of decision-making rationale (Delgado and Rossi, 2024).

One of the most critical aspects of this paper was the inclusion of Local Interpretable Model-Agnostic Explanations (LIME) feature that made the model predictions much more readable. In contrast to the results provided by classical machine learning, LIME provides an interpretable description of the inputs that explains the manner in which single features led to one particular outcome. To illustrate, in one of the observed cases, LIME showed that the variables like the high monthly price, the absence of tech support, and the absence of the contract commitment (month-to-month contract) were strong indicators of churn prediction, and the variables like a long tenure and the presence of online security as well as use of a 2 years contract mitigated churn risk (Smith and Jones, 2023). The level of details turned the predictive model into a readable decision support, which allowed managers to make specific decisions, i.e. to offer a discount or deliver improved support services to customers, based on the reasons which affected their churn risk at the individual customer level. The LIME module supported a more context-based, explicit interpretation of the individual predictions as it enabled visualising the feature contributions by using bar plots and, moreover, displaying data in both tabular and graphical form (Al-Hayani et al., 2024).

The creation of a Streamlit-based interactive dashboard titled the decision support interface was one of the major contributions of this study. The dashboard integrated features of the machine learning pipeline, such as data inspection, illustration data analysis (EDA), choice of model, forecasting, and explainability and made them all available as a single, unified tool accessible to the business. The dashboard contained the following sections: snapshots of datasets, pie charts of churn distributions, bar plots of the contract type, metrics on model performance, and LIME explanation (Vilone and Longo,

2021). The ease of using this dashboard played a pivotal role in solving the problem of breaking the gap that usually exists between data scientists and decision-makers. Non-technical users would be able to access the system through minimal sidebar controls, see an explanation of the churn predictions, and result interpretation without having to know about the working code or architecture. Such choice was in line with human-centered AI principles and served as a part of the wider effort to ensure managers possess trust in AI systems (Smith and Jones, 2023).

Another insight that was gained during the contract-type distribution analysis was that the majority of the customers (more than 50%) were under month-to-month contracts. The lack of long-term commitment also makes this group disposed to churn naturally. In comparison, customers with one- and two-year contracts recorded lower churn rates, possibly caused by financial motives and psychological commitment (Vilone and Longo, 2021). The trend established a confirmation of current business knowledge as well as it demonstrated the power of the model in picking up meaningful trends. Moreover, the contract type was one of the best features numerous times emphasized by LIME-based explanations, which supports its position as a churn predictor feature. This kind of model insight and domain experience would be critical to creating trust among the decision-makers, to reassure them that not only are the model results statistically sound but also business-logically correct (Al-Hayani et al., 2024).

The comparison of two machine learning models of Random Forest and Logistic Regression has brought forth some helpful implications. Albeit Random Forest was more accurate and scored higher on its F1-score in the churn category than Logistic Regression, it was a difference of lesser relevance. This implied that in certain business cases, linear models such as the Logistic Regression may represent a better choice since they are interpretable by their natures and do not require large amounts of resources to compute. But, combined with LIME, even more models, like Random Forest, could be made transparent enough to be used by managers. It shows the flexibility of XAI measures to trade-off between the complexity and interpretability of a model, which enables organisations to use more accurate models without losing their trust or responsibility (Vilone and Longo, 2021).

In a larger sense, the system designed in this paper provided a model that could be generalized and implemented to apply AI and explainability in practical decision-making. Although the prototype applied to the telecom domain and the reduction of customer churn, the backbone architecture, which includes data ingestion, preprocessing, model training, evaluation, correlation made with LIME (explanations) and Streamlit dashboarding, is universal in principle and can be applied to other areas (e.g. the healthcare realm (risk of readmission), banking (lending default prediction), or retail (optimisation of stock). Open-source products as well as common Python libraries were also used, which guaranteed viability and changeability. Notably, the system was also in line with ethical AI by ensuring the decision-making process is transparent and fusible in addition to being easy to use, which responds to the issues in the system control through regulatory requirements and population call to accountable algorithms (Al-Hayani et al., 2024).

4.1. Conclusion

In this chapter, main findings of the study and the development and testing of the Decision Support System (DSS) with Explainable AI (XAI) capabilities focused on enhancing customer retention in the telecommunication industry were discussed. With the help of intense preprocessing of the data, the exploratory analysis, and training of machine learning models (Logistic Regression and Random Forest classifier), the system had shown good predictive attributes with two models, with a probability of about 79 percent. A closer examination of the performance metrics, however, showed that a class imbalance was problematic to identify churned customers, especially with recall and F1-scores of the minority classes being rather low. The system has tried to resolve the most important trust and interpretability problem in AI decision-making by using LIME (Local Interpretable Model-Agnostic Explanations) to provide justifications of the form that humans associate with. Such local explanations allowed the decision-makers to intuit how a specific customer was being considered high risk, as well as proving actionable insights in terms of individual customer characteristics like the type of contract, subscription fees and monthly costs, contract length and usage of a service. These explanations played a direct role in fulfilling the purpose of the study, which was to increase managerial transparency and trust.

5. CONCLUSION

This chapter provides practical, technical, and strategic recommendations based on the findings of this study, which focused on developing an Explainable Artificial Intelligence (XAI)-powered Decision Support System (DSS) for customer churn prediction in the telecommunications industry. The implementation used the IBM Telco Customer Churn dataset, employed Logistic Regression and Random Forest classifiers, and integrated Local Interpretable Model-Agnostic Explanations (LIME) within a Streamlit-based dashboard. The recommendations are categorised into six key areas: model enhancement, explainability improvements, interface usability, ethical and regulatory compliance, real-time integration, and managerial training and adoption strategy. Each recommendation is backed by empirical findings from Chapter 4 and aligned with the original aims of developing a trustworthy, transparent, and actionable DSS platform.

5.1. Enhancing Predictive Performance through Class Balancing Techniques

Although both Logistic Regression and Random Forest models showed promising accuracy in predicting customer churn, their performance on the minority class (churned customers) was notably weaker. This imbalance suggests the need to implement class balancing methods such as Synthetic Minority Over-sampling Technique (SMOTE), Random Under-sampling, or cost-sensitive learning. These approaches can help improve the model's sensitivity to churned cases without significantly compromising overall performance (Adeniran et al., 2024).

5.2. Incorporating Advanced Classification Models and Hybrid Approaches

The Logistic Regression offers interpretability and Random Forest delivers robust performance, incorporating additional models such as Gradient Boosting Machines (GBM), XGBoost, or ensemble methods could enhance predictive capabilities. Ensemble learning strategies such as stacking or voting classifiers can combine the strengths of multiple models, providing both high accuracy and generalization (Dolatabadi et al., 2024).

5.3. Integration of Real-Time Data Pipelines

The current DSS architecture is based on static datasets. For real-world application, the ability to process live customer data is critical. This can be achieved by integrating real-time data streaming tools such as Apache Kafka or Apache Spark Streaming with the DSS backend. A pipeline could be designed to ingest transactional and interactional data in real time, dynamically updating churn predictions and visual explanations (Dwivedi et al. 2023).

5.4. Expanding LIME Functionality with Comparative XAI Tools

The LIME offers powerful local explanations, it should be complemented with other tools such as SHAP (SHapley Additive exPlanations) or Integrated Gradients. SHAP offers both local and global interpretability and is grounded in game theory, making it more consistent across predictions. Combining these tools within the DSS will provide multi-layered explanation depth, appealing to both high-level strategic users and technical analysts. Offering users the ability to toggle between LIME and SHAP visualizations in the dashboard could significantly enhance trust and engagement (Visave, 2025).

5.5. Improving Interface Usability through User-Centered Design

The current dashboard is functional, but further refinement is required to make it enterprise-ready. Conducting usability testing sessions with managers from the telecommunications sector could identify pain points in navigation, information overload, or lack of contextual help. Features like interactive tooltips, customizable views, and drill-downs for data exploration should be integrated (Hasan, 2023). Beyond prediction and explanation, the DSS should evolve into a prescriptive tool. Implementing automated retention strategies such as email alerts, CRM ticket generation, or direct integration with marketing automation platforms like HubSpot or Salesforce would close the loop between insight and action (Dwivedi et al. 2023).

5.6. Summarized Findings

The primary objective of this dissertation was to design, implement, and evaluate a customer retention Decision Support System (DSS) integrated with Explainable Artificial Intelligence (XAI). The solution

leveraged machine learning models (Logistic Regression and Random Forest) to predict customer churn, complemented by LIME to provide local, human-understandable explanations. The system was delivered via an interactive web interface developed using Streamlit, allowing decision-makers to explore data, assess model performance, and interpret prediction rationales without technical expertise. Each step in the methodological pipeline—from data preprocessing and modeling to visualization and explanation—was aligned with managerial needs for transparency, trust, and actionable insight (Dwivedi et al. 2023).

5.7. Limitations of the Study

The prototype is functional and informative; several limitations were identified. First, the dataset used is static and limited to a single domain. This raises questions about the generalizability of the system. Second, only two machine learning models were employed. Incorporating more sophisticated algorithms and hybrid approaches could potentially yield better performance. Third, the lack of formal user testing restricts claims about the dashboard's usability and effectiveness in real business contexts. Lastly, while LIME offers strong local explanations, it does not provide global insights into model behaviour, necessitating additional XAI tools in the future (Visave, 2025).

5.8. Future Research Directions

Future research should aim to enhance the robustness, scalability, and applicability of the Explainable AI-powered Decision Support System (DSS) developed in this study. One promising direction is the integration of more advanced classifiers such as XGBoost, LightGBM, or deep learning models, which could improve predictive performance, particularly in handling complex churn patterns. To maintain interpretability alongside these complex models, future work should explore complementary explainability techniques such as SHAP (SHapley Additive Explanations) and Anchors, which offer both global and local transparency. Real-time data processing capabilities are also essential for transitioning the system from a research prototype to a production-ready tool (Hasan, 2023).

REFERENCES

- Adeniran, I.A., Efunniyi, C.P., Osundare, O.S., Abhulimen, A.O. and OneAdvanced, U., 2024. Implementing machine learning techniques for customer retention and churn prediction in telecommunications. *Computer Science & IT Research Journal*, 5(8).
- Al-Hayani, A., Ali, S., and Kumar, A. (2024) 'Explainable Artificial Intelligence-Based Decision Support Systems: A Survey', *Electronics*, 13(14), p.2842. Doi: 10.3390/electronics13142842. mdpi.com
- Ali, R., Hussain, A., Nazir, S., Khan, S. and Khan, H.U., 2023. Intelligent decision support systems—an analysis of machine learning and multicriteria decision-making methods. *Applied Sciences*, 13(22), p.12426.
- Cheng, Y., Li, X. and Wang, Z. (2025) 'Explainability and AI Confidence in Clinical Decision Support Systems: Effects on Trust, Diagnostic Performance, and Cognitive Load', *arXiv preprint arXiv: 2501.16693*. en.wikipedia.org+12arxiv.org+12mdpi.com+12
- Damali, U., Secchi, E., Tax, S.S. and McCutcheon, D., 2021. Customer participation risk management: conceptual model and managerial assessment tool. *Journal of Service Management*, 32(1), pp.27-51.
- Delgado, M. and Rossi, P. (2024) 'Extending application of explainable artificial intelligence for managerial decision-making', *Annals of Operations Research*. Doi: 10.1007/s10479-024-05825-9. link.springer.com+1en.wikipedia.org+1
- Dolatabadi, S.H., Gatial, E., Budinská, I. and Balogh, Z., 2024, July. Integrating human-computer interaction principles in user-centered dashboard design: Insights from maintenance management. In *2024 IEEE 28th International Conference on Intelligent Engineering Systems (INES)* (pp. 000219-000224). IEEE.
- Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G. and Ranjan, R., 2023. Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Computing Surveys*, 55(9), pp.1-33.
- Haque, A.K.M., 2025. Explainable artificial intelligence (XAI): making AI understandable for end users.
- Hasan, M.M., 2023. Understanding model predictions: a comparative analysis of SHAP and LIME on various ML algorithms. *Journal of Scientific and Technological Research*, 5(1), pp.17-26.
- Issitt, R.W., Cortina-Borja, M., Bryant, W., Bowyer, S., Taylor, A.M., Sebire, N. and Bowyer, S.A., 2022. Classification performance of neural networks versus logistic regression models: evidence from healthcare practice. *Cureus*, 14(2).

- Kostopoulos, G., Davrazos, G. and Kotsiantis, S., 2024. Explainable Artificial Intelligence-Based Decision Support Systems: A Recent Review. *Electronics*, 13(14), p.2842.
- McDermid, J.A., Jia, Y., Porter, Z. and Habli, I., 2021. Artificial intelligence explainability: the technical and ethical dimensions. *Philosophical Transactions of the Royal Society A*, 379(2207), p.20200363.
- Murikah, W., Nthenge, J.K. and Musyoka, F.M., 2024. Bias and ethics of AI systems applied in auditing-A systematic review. *Scientific African*, p.e02281.
- Naidu, G., Zuva, T. and Sibanda, E.M., 2023, April. A review of evaluation metrics in machine learning algorithms. In *Computer science on-line conference* (pp. 15-25). Cham: Springer International Publishing.
- Nguyen, A. and Doan, T. (2025) 'Customer-Centric Decision-Making with XAI and Counterfactual Explanations', *Journal of Theoretical and Applied Electronic Commerce Research*, 20(2), p.129. Doi: 10.3390/jtaer20020129. arxiv.org+6mdpi.com+6sciencedirect.com+6
- Orlenko, A. and Moore, J.H., 2021. A comparison of methods for interpreting random forest models of genetic association in the presence of non-additive interactions. *BioData mining*, 14, pp.1-17.
- Parisineni, S.R.A. and Pal, M., 2024. Enhancing trust and interpretability of complex machine learning models using local interpretable model agnostic shap explanations. *International Journal of Data Science and Analytics*, 18(4), pp.457-466.
- Rainy, T., 2025. Mechanisms by which AI-Enabled CRM systems influence customer retention and overall business performance: A systematic literature review of empirical findings. Available at SSRN 5295111.
- Richards, T., 2023. *Streamlit for Data Science: Create interactive data apps in Python*. Packt Publishing Ltd.
- Smith, C. and Jones, R. (2023) 'A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME', arXiv preprint arXiv: 2305.02012. mdpi.com+7arxiv.org+7mdpi.com+7
- Vilone, G. and Longo, L. (2021) 'Classification of Explainable Artificial Intelligence Methods through Their Output Formats', in *Machine Learning and Knowledge Extraction*, pp. 1–15. journals.adbascientific.com+1en.wikipedia.org+1
- Vilone, G. and Longo, L., 2021. Classification of explainable artificial intelligence methods through their output formats. *Machine Learning and Knowledge Extraction*, 3(3), pp.615-661.
- Visave, J., 2025. Transparency in AI for emergency management: Building trust and accountability. *AI and Ethics*, pp.1-14.
- Zafar, M.R. and Khan, N., 2021. Deterministic local interpretable model-agnostic explanations for stable explainability. *Machine Learning and Knowledge Extraction*, 3(3), pp.525-541.

Citation: Aljwhrh Almtrf. " Integrating Explainable Ai (Xai) Into Decision Support Systems: A Framework for Enhancing Transparency and Trust in Managerial Decision-Making" *International Journal of Managerial Studies and Research (IJMSR)*, vol 13, no. 9, 2025, pp.9-22. DOI:<https://doi.org/10.20431/2349-0349.1309002>.

Copyright: © 2025 Author. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.