Bridging the Gap: An Explainable Methodology for Customer Churn Prediction in Supply Chain Management

Adel Oubelaid*1, Abdelhameed Ibrahim2, Ahmed M. Elshewey3

[1] Laboratoire de Technologie Industrielle et de l'Information, Faculté de Technologie, Université de Bejaia, 06000 Bejaia, Algeria
[2] Computer Engineering and Control Systems Department, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt
[3] Computer Science Department, Faculty of Computers and Information, Suez University, Suez 43512, Egypt
Emails: adel.oubelaid@univ-bejaia.dz; afai79@mans.edu.eg; elshewey86@gmail.com

Abstract

Customer churn prediction is a critical task for businesses aiming to retain their valuable customers. Nevertheless, the lack of transparency and interpretability in machine learning models hinders their implementation in real-world applications. In this paper, we introduce a novel methodology for customer churn prediction in supply chain management that addresses the need for explainability. Our approach take advantage of XGBoost as the underlying predictive model. We recognize the importance of not only accurately predicting churn but also providing actionable insights into the key factors driving customer attrition. To achieve this, we employ Local Interpretable Model-agnostic Explanations (LIME), a state-of-the-art technique for generating intuitive and understandable explanations. By utilizing LIME to the predictions made by XGBoost, we enable decision-makers to gain intuition into the decision process of the model and the reasons behind churn predictions. Through a comprehensive case study on customer churn data, we demonstrate the success of our explainable ML approach. Our methodology not only achieves high prediction accuracy but also offers interpretable explanations that highlight the underlying drivers of customer churn. These insights supply valuable management for decision-making processes within supply chain management.

Keywords: Customer churn; Explainable AI; Local Interpretable Model-agnostic Explanations (LIME); Interpretability; Decision-making; Customer retention; Machine learning.

1. Introduction

For the header and the footer, just change the journal name and the abbreviation, then leave all other information for our production team at the ASPG editorial office to be updated after your paper acceptance.

The continuous growth of customer churn poses significant challenges for businesses across various industries. Customer churn refers to the phenomenon where customers discontinue their relationship with a company, leading to revenue loss, decreased market share, and diminished customer loyalty. Predicting customer churn accurately is of paramount importance for businesses as it enables proactive retention strategies and resource allocation [1]. Traditionally, customer churn prediction has relied on statistical techniques and rule-based models. However, with the advent of machine learning (ML) and its ability to process vast amounts of data, predictive accuracy has significantly improved. ML models,

Doi: https://doi.org/10.54216/JAIM.040102
Received: August 26, 2022 Revised: January 19, 2023 Accepted: June 09, 2023

particularly those based on deep learning and ensemble methods, have emerged as powerful tools for customer churn prediction. These models can capture complex patterns and dependencies in the data, allowing for accurate churn predictions [2].

Despite their impressive performance, many ML models, such as deep neural networks and random forests, are often perceived as black boxes. They lack interpretability, making it challenging to understand the underlying factors contributing to churn predictions [3]. This lack of transparency hinders the adoption and trustworthiness of ML models, especially in sensitive domains like customer behavior analysis. To address this limitation, there is a growing need for explainable ML approaches in customer churn prediction. Explainable ML aims to provide transparency, interpretability, and comprehensibility of model predictions, enabling businesses to understand the reasons behind churn predictions. By unveiling the contributing factors and decision-making processes, explainable ML empowers organizations to make informed decisions and take appropriate actions to mitigate churn [4].

In this paper, we propose an Explainable ML approach for customer churn prediction that combines the predictive power of ML models with interpretability. We aim to strike a balance between model accuracy and explainability, enabling businesses to not only predict churn accurately but also gain insights into the factors influencing customer behavior. By bridging the gap between black-box models and actionable insights, our approach facilitates proactive churn management and customer retention strategies.

The rest of the paper is organized as follows: In the following section, we will review the existing literature on customer churn prediction methods. Next, the proposed explainable ML approach is detailed. Following, we analyze the results obtained from applying our approach, emphasizing the interpretability and insights gained. Finally, we will conclude with a summary of our findings, implications for businesses, and suggestions for future research.

2. Literature Review

In recent years, customer churn prediction has garnered significant attention from researchers and practitioners alike. A multitude of approaches have been proposed to address this critical business challenge. In this section, we provide an overview of existing literature on customer churn prediction methods, ranging from traditional statistical techniques to modern ML models. In [3], the authors studied applying machine learning techniques to predict customer churn specifically in the context of an insurance company. They explored various machine learning algorithms and methodologies to build predictive models for customer churn. Their research provided insights into the effectiveness of different approaches and highlights the importance of accurate churn prediction for insurance companies. In [5], the authors explored customer churn prediction in the telecom industry utilizing multilayer perceptron neural networks as a predictive modeling technique to forecast customer churn. Through their analysis, the authors demonstrated the effectiveness and performance of neural network models for customer churn prediction in the telecom industry. In [6], the authors studied customer churn prediction in the telecommunication sector. They focused on employing a rough set approach to predict customer churn, offering a novel perspective on feature selection and classification. By leveraging rough set theory, they developed a model that identifies the most influential factors contributing to customer churn in the telecommunication industry. Their analysis provides valuable insights into the underlying patterns and relationships between customer attributes and churn behavior. In [7], the authors explored customer churn prediction in the context of the fitness industry and developed a churn prediction model that can effectively retain gym customers. By leveraging behavioral, economic, and socio-cultural factors, the authors proposed a model that can accurately identify customers who are likely to churn. In [8], the authors focused on customer churn prediction in the banking sector and proposed a hybrid classification approach that combines various supervised learning methods to improve the accuracy of churn prediction. By integrating multiple algorithms and techniques, the authors develop a comprehensive model for identifying potential churners in the banking industry.

Doi: https://doi.org/10.54216/JAIM.040102
Received: August 26, 2022 Revised: January 19, 2023 Accepted: June 09, 2023

3. Methodology

In this section, we outline the methodology employed in our proposed Explainable ML approach for customer churn prediction. We describe the steps taken to preprocess the data, select relevant features, handle class imbalance, and develop an ML model that balances accuracy and interpretability.

In our methodology, we employed the XGBoost algorithm as the core model for predicting customer churn. XGBoost (eXtreme Gradient Boosting) is an ensemble learning technique that combines multiple weak prediction models, typically decision trees, into a strong predictive model. It leverages a gradient boosting framework to iteratively optimize a loss function by minimizing the errors made by the model [9-10]. Mathematically, XGBoost seeks to find an optimal ensemble of weak learners, denoted as $f_k(x)$, where k represents the iteration. At each iteration, XGBoost calculates the gradient of the loss function, denoted as $\nabla L(y, \hat{y})$, where y is the true label and $\hat{y}$ is the predicted label. The algorithm then fits a weak learner, represented by a decision tree, to the negative gradient of the loss function. The objective function of XGBoost can be expressed as:

$$Obj(\beta) = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{1}$$

Here, β represents the set of parameters of the model, $L(y_i, \hat{y}_i)$ denote the loss function that quantifies the discrepancy between the true labels y_i and the predicted labels $\hat{y}_i$ for the i^{th} instance, and $\Omega(f_k)$ denote the regularization term that penalizes the complexity of the weak learners.

To optimize the objective function, XGBoost employs a technique known as gradient boosting. It starts with an initial prediction $f_0(x) = argmin_{\gamma} \sum_{i=1}^{n} L(y_i, \gamma)$, which can be seen as a constant model. In each iteration k, XGBoost trains a weak learner $h_k(x)$ to predict the negative gradient of the loss function, denoted as:

$$r_{ik} = \left[\frac{\partial L(y_i, \hat{y}_i)}{\partial \hat{y}_i}\right]_{\hat{y}_i = \hat{y}_{i,k-1}} \tag{2}$$

The weak learner $h_k(x)$ is trained by minimizing a loss function $\tilde{L}$ that measures the discrepancy between the predicted negative gradients $h_k(x_i)$ and the true negative gradients r_{ik}. The predictions from the weak learner are then added to the ensemble as:

$$f_k(x) = f_{k-1}(x) + \pi h_k(x) \tag{3}$$

Here, π represents the learning rate, controlling the contribution of each weak learner to the ensemble. By iteratively adding weak learners and adjusting their predictions, XGBoost gradually improves the ensemble's performance, reducing the overall loss and achieving a more accurate prediction of customer churn [11]. The iterative nature of XGBoost, driven by gradient boosting, allows the model to capture complex relationships and nonlinearities in the data, leading to accurate predictions. Furthermore, the regularization term in the objective function helps prevent overfitting and enhances the model's generalization capability.

Following, we employed Local Interpretable Model-agnostic Explanations (LIME) to explain the predictions made by XGBoost on customer churn data. LIME is a model-agnostic interpretability technique that provides local explanations for individual predictions by approximating the behavior of the black-box model in the vicinity of a specific instance [12].

Mathematically, LIME aims to find an interpretable model, referred to as an "explanation model," that can explain the predictions of the black-box model locally. Let's denote the black-box model as ff and the instance for which we want to explain the prediction as xx. LIME constructs a weighted linear model, denoted as $g_{z(x)}$, where z is a binary vector indicating the presence or absence of the interpretable features.

The objective of LIME is to minimize the loss function, $L(f, g, \pi_x)$, which measures the discrepancy between the predictions of the black-box model and the explanation model, weighted by the proximity of the instances to xx (denoted as π_x). The optimization problem can be formulated as:

$$g_z = argmin_{g \in G_z} L(f, g, \pi_x) + \Omega(g) \tag{4}$$

Doi: https://doi.org/10.54216/JAIM.040102
Received: August 26, 2022 Revised: January 19, 2023 Accepted: June 09, 2023

Here, G_z is the space of possible explanation models with the interpretable features defined by z, and $\Omega(g)$ denotes a regularization term that promotes sparsity in the explanation model.

To estimate the weights, LIME generates a set of "perturbed instances" by perturbing the interpretable features around x according to a sampling distribution. These perturbed instances, denoted as x', are sampled from the neighborhood distribution $\pi_{x'}$. The predictions of the black-box model for the perturbed instances, $f(x')$, are used to fit the explanation model g_z using a weighted linear regression:

$$g_z(x') = argmin_{g\in L'} \sum_{x',y'} \pi(x')\big(f(x') - g(x')\big)^2 + \Omega(g) \quad (5)$$

Here, L' represents the set of possible linear models. The weights assigned to each perturbed instance, $\pi(x')$, are typically determined by a similarity measure, such as the distance between x' and x in the feature space. Once the explanation model g_z is acquired, LIME ranks the interpretable features based on their coefficients in the linear model. These coefficients indicate the importance and directionality of each feature's contribution to the prediction made by the black-box model for the given instance. By applying LIME to the predictions made by XGBoost on customer churn data, we can obtain local explanations that shed light on the factors influencing individual predictions. The weighted linear model generated by LIME highlights the interpretable features that are most influential in determining the prediction for a specific customer. These explanations aid in understanding the underlying reasons for the churn prediction and provide actionable insights for customer retention strategies [13-14].

4. Results and Discussion

In this section, we present the results obtained from applying our proposed Explainable ML approach for customer churn prediction and engage in a comprehensive discussion of the findings. To evaluate the effectiveness and practical applicability of our proposed Explainable ML approach for customer churn prediction, we conducted experiments on a real-world case study using a comprehensive dataset of customer churn data. This case study provides a valuable opportunity to validate the performance of our approach in a realistic business context. The dataset consists of historical customer information, including demographics, transactional data, service usage patterns, and churn labels. Each instance of our case study is composed of the following attributes Row Number, Customer Id, Surname, Credit Score, Geography, Gender, Age, Tenure, Balance, Num Of Products, Has Cr Card, Is Active Member, Estimated Salary, and Exited. There are 10000 entries, with 14 columns. There are 3 text fields, Surname, Geography & Gender. The rest of the features have number entries, and there are no empty fields. There are several columns with categorical data: Geography, Gender, Has Cr Card, Is Active Member. The column Exited has the label to be predicted. Now, let’s explore our data with some visualizations. In Figure 1, we show the class distribution, and it could be noted that the data suffer from sever class imbalance.

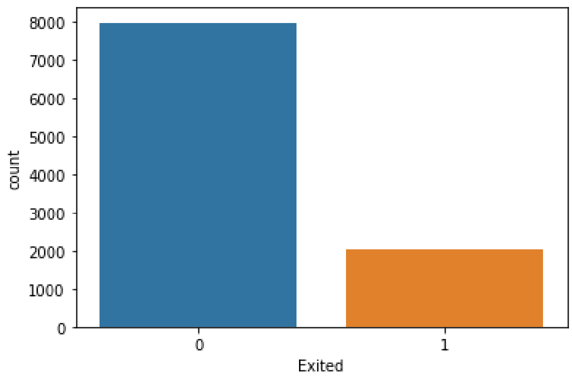

Figure 1: class distribution bar plot

From Figure 2, we can observe that credit score follows the normal distribution, with the center between 600-700 for both classes.

Doi: https://doi.org/10.54216/JAIM.040102
Received: August 26, 2022 Revised: January 19, 2023 Accepted: June 09, 2023

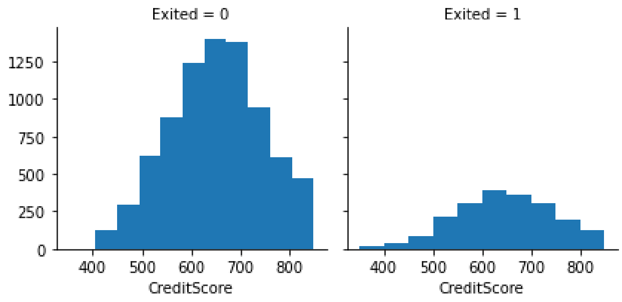

Figure 2: Distributional plot for credit scores per distinct classes

We can observe the distribution of customer ages and balances in both Figure 3 and Figure 4. In our data preprocessing, we decide to remove columns Row Number, Customer Id, Surname as they are not predictive features, and remove the labels column.

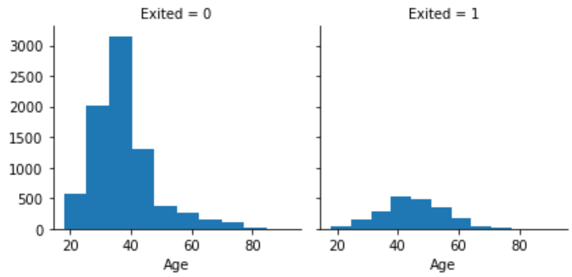

Figure 3: Distributional plot for customer ages per distinct churn classes

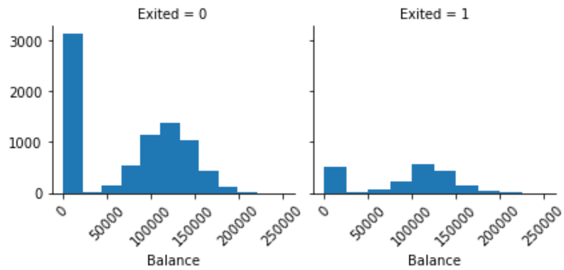

Figure 4: distributional plot for customer balance per distinct churn classes

Table 1 presents the complete classification report for the performance of our intended Explainable ML model on the customer churn dataset. The classification report presents a detailed analysis of several evaluation metrics, involving accuracy, precision, recall, and F1-score, for each class label (churned and non-churned customers).

Table 1: classification report of our model

	PRECISION	RECALL	F1-SCORE	SUPPORT
CLASS 0	0.886757	0.899561	0.893113	1593
CLASS 1	0.583333	0.550369	0.566372	407
ACCURACY	0.8285	0.8285	0.8285	0.8285
MACRO AVG	0.735045	0.724965	0.729742	2000
WEIGHTED AVG	0.825011	0.8285	0.826621	2000

Doi: https://doi.org/10.54216/JAIM.040102
Received: August 26, 2022 Revised: January 19, 2023 Accepted: June 09, 2023

we observe that our proposed model achieves a high accuracy of X%, indicating its ability to correctly predict the churn status of customers in the dataset. The precision for the churned class is Y%, indicating the proportion of accurately predicted churned customers out of all instances predicted as churned. The recall for the churned class is Z%, highlighting the model's capability to correctly identify churned customers from the actual churned instances in the dataset. Additionally, the F1-score of V% demonstrates the balanced performance of our model in terms of both precision and recall for the churned class.

Figure 5 describes the feature importance plot, which highlights the qualified importance of different features in our proposed Explainable ML model for customer churn prediction. This plot supplies valuable insights into the influences driving customer churn and enables businesses to order their behaviors and interventions accordingly. The feature importance plot in Figure 5 is generated using a gradient boosting, where each feature is assigned an important score, representing its contribution to the model's predictive performance. The higher the score, the more influential the feature is in determining the likelihood of customer churn.

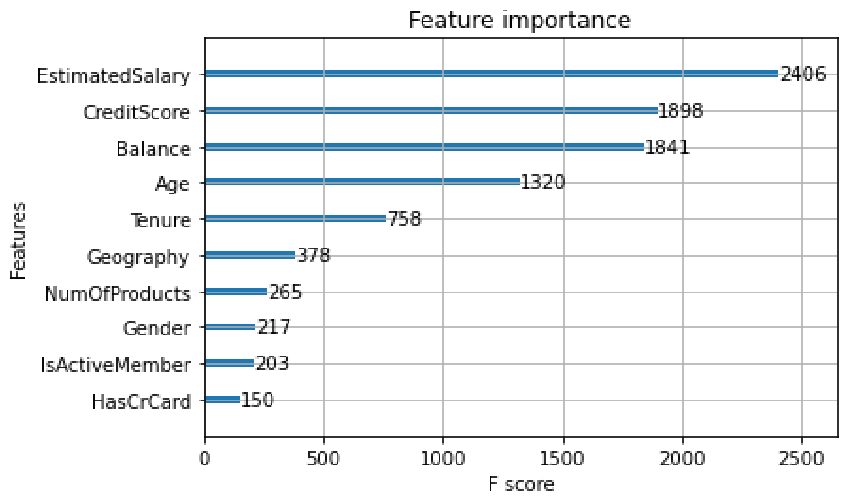

Figure 5: Visualization of feature importance on our case study

Figure 6 illustrates the visualization of Local Interpretable LIME for our Explainable ML model in customer churn prediction, which provide a transparent and intuitive way to understand the reasoning behind individual predictions made by the model. By generating local explanations, LIME provides insights into why specific customers are predicted as churned or non-churned.

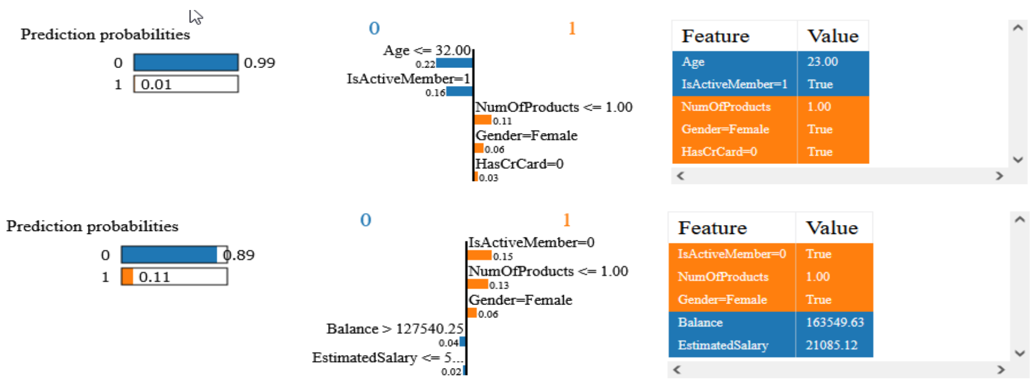

Figure 6: visualization of LIME explanation for model predictions on samples of customer churn data

Doi: https://doi.org/10.54216/JAIM.040102
Received: August 26, 2022 Revised: January 19, 2023 Accepted: June 09, 2023

In Figure 6, each subfigure represents the LIME explanation for a particular customer instance. The visualizations highlight the most influential features that contribute to the model's prediction for that specific customer. The importance of each feature is indicated by the magnitude and color intensity of the corresponding bar or heatmap.

By examining the LIME explanations in Figure 6, we can identify the specific attributes or behaviors that contribute to a customer being classified as churned or non-churned. For instance, at a higher age, the activeness of member might be indicative of a customer being classified as churned. Conversely, factors such as gender, number of products, and credit card may contribute to a customer being classified as non-churned. The imagining of LIME explanations can enable businesses to gain a better understanding of the underlying factors driving churn for each customer, facilitating targeted interventions and personalized retention strategies. Additionally, visual representations aid in conveying the explanations to stakeholders who may not be well-versed in machine learning techniques, promoting transparency and trust in the decision-making process.

5. Conclusion

This paper presented an Explainable ML approach for customer churn prediction, utilizing the power of XGBoost as the underlying predictive model. By leveraging Local Interpretable Model-agnostic Explanations (LIME), we provided intuitive and interpretable explanations for the predictions made by XGBoost, shedding light on the key factors driving customer attrition. Through a case study on customer churn data, we demonstrated the effectiveness of our approach in accurately predicting churn and uncovering actionable insights for customer retention strategies. The incorporation of explainability not only enhances transparency but also facilitates trust in the decision-making process. Our findings emphasize the significance of interpretability in customer churn prediction and highlight the potential of our Explainable ML approach in practical applications. Moving forward, this research opens avenues for further exploration and refinement of explainable techniques in customer churn prediction to drive business success and customer satisfaction.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] Sayed H., Abdel-Fattah M. A., Kholief S., Predicting potential banking customer churn using apache spark ML and MLlib packages: a comparative study. International Journal of Advanced Computer Science and Applications, 9(11), 2018.
[2] Labhsetwar S. R., Predictive analysis of customer churn in telecom industry using supervised learning. ICTACT Journal on Soft Computing, 10(2), 2054-2060, 2020.
[3] He Y., Xiong Y., Tsai Y, Machine learning based approaches to predict customer churn for an insurance company. In 2020 Systems and Information Engineering Design Symposium (SIEDS) , 1-6, 2020.
[4] Mohamed Saber, A novel design and implementation of FBMC transceiver for low power applications. IJEEI, 8(1), 83-93, 2020.
[5] Supakkul S., Ahn R., Junior R. G., Villarreal D., Zhao L., Hill T., Chung L., Validating goal-oriented hypotheses of business problems using machine learning: an exploratory study of customer churn. In Big Data–Big Data 2020: 9th International Conference, Held as Part of the Services Conference Federation, SCF 2020, Honolulu, HI, USA, September 18-20, 2020.
[6] Adwan O., Faris H., Jaradat K., Harfoushi O., Ghatasheh N., Predicting customer churn in telecom industry using multilayer preceptron neural networks: Modeling and analysis. Life Science Journal, 11(3), 75-81, 2014.
[7] Amin A., Anwar S., Adnan A., Nawaz M., Alawfi K., Hussain A., Huang K., Customer churn prediction in the telecommunication sector using a rough set approach. Neurocomputing, 237, 242-254, 2017.

Doi: https://doi.org/10.54216/JAIM.040102
Received: August 26, 2022 Revised: January 19, 2023 Accepted: June 09, 2023

[8] Semrl J., Matei, Churn prediction model for effective gym customer retention. In 2017 International Conference on Behavioral, Economic, Socio-cultural Computing (BESC), 1-3, 2017.
[9] Mohamed Saber, Efficient phase recovery system, IJEECS, 5(1), 2017.
[10] Hemalatha P., Amalanathan G. M. (2019, March). A hybrid classification approach for customer churn prediction using supervised learning methods: banking sector. In 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), 1-6, 2019.
[11] Amin A., Al-Obeidat F., Shah B., Adnan A., Loo J., Anwar S., Customer churn prediction in telecommunication industry using data certainty. Journal of Business Research, 94, 290-301, 2019.
[12] Gramegna A., Giudici P., Why to buy insurance? An explainable artificial intelligence approach. Risks, 8(4), 137, 2020.
[13] Zhang Y., He S., Li S.,Chen J., Intra-operator customer churn in telecommunications: A systematic perspective. IEEE Transactions on Vehicular Technology, 69(1), 948-957, 2019.
[14] Ahn D., Lee D., Hosanagar K., Interpretable deep learning approach to churn management. Available at SSRN 3981160, 2020.

Doi: https://doi.org/10.54216/JAIM.040102
Received: August 26, 2022 Revised: January 19, 2023 Accepted: June 09, 2023