Volume 14, Issue 2, Page 11-25, 2025 https://doi.org/10.46810/tdfd.1529139 Research Article

BİNGÖL UNIVERSITY

Türk Doğa ve Fen Dergisi
Turkish Journal of Nature and Science
www.dergipark.gov.tr/tdfd

Comparative Analysis of XAI Techniques on Telecom Churn Prediction Using SHAP and Interpreted ML Partial Dependence

Cem ÖZKURT[1*]

[1*]Sakarya University of Applied Sciences, Computer Engineering, Sakarya, Türkiye
Cem ÖZKURT ORCID No: 0000-0002-1251-7715

*Corresponding author: cemozkurt@subu.edu.tr

(Received: 06.08.2024, Accepted: 24.03.2025, Online Publication: 27.06.2025)

Keywords
Explainable artificial ıntelligence,
SHAP,
Interpreted ML Partial Dependence,
Transparency and interpretability

Abstract: A comparative analysis of two prominent Explainable Artificial Intelligence (XAI) techniques, SHAP (SHapley Additive exPlanations) and Interpreted ML Partial Dependence, was conducted on a real-world Telecom Churn dataset consisting of 7,043 customer records. The objective of this study was to evaluate and compare the effectiveness of these techniques in enhancing the transparency and interpretability of machine learning models, specifically in telecom churn prediction. The study leveraged an XGBoost model, which achieved an accuracy of 94.12%, outperforming other machine learning models in the dataset. The methodology outlined the steps of data preprocessing and model training. Two separate analyses using SHAP and Interpreted ML Partial Dependence were conducted to evaluate their effectiveness in explaining model decisions and uncovering feature importance. The results of both techniques were discussed, highlighting their strengths and weaknesses, and providing valuable insights into interpretability and robustness. The SHAP analysis demonstrated that it is a powerful tool for identifying which features influence churn, particularly highlighting Contract type, Monthly Charges, and Tech Support as key drivers of customer churn. For instance, SHAP values revealed that customers with short-term contracts and no tech support were significantly more likely to churn, with SHAP values exceeding 0.6 in impact. This fine-grained analysis provided precise insights into individual predictions. On the other hand, the Interpreted ML Partial Dependence method showed the general effects of features, allowing for a broader perspective on model behavior. It illustrated how changes in Monthly Charges and Tenure affected churn probability across the dataset, showing that customers with longer tenure had lower churn probabilities on average. These results enhanced the transparency of model decisions, instilling trust in users and helping them understand how the model works. The key contribution of this study is the comparative evaluation of SHAP and Interpreted ML Partial Dependence in telecom churn prediction, offering a structured framework for selecting appropriate XAI techniques based on interpretability needs. SHAP provided instance-specific explanations, crucial for personalized customer retention strategies, while Partial Dependence offered a macroscopic view, useful for high-level decision-making and policy adjustments.This comparative analysis contributes to a deeper understanding of XAI methods and emphasizes the importance of selecting appropriate techniques to enhance transparency in telecom churn prediction models.

SHAP ve Interpreted ML Kısmi Bağımlılık Kullanarak Telekom İşten Ayrılma Tahmininde XAI Tekniklerinin Karşılaştırmalı Analizi

Anahtar Kelimeler
Açıklanabilir Yapay Zeka,
SHAP,
Yorumlanabilir ML'de Kısmi Bağımlılık,
Şeffaflık ve Yorumlanabilirlik

Öz: Önde gelen iki Açıklanabilir Yapay Zeka (XAI) tekniği olan SHAP (SHapley Additive exPlanations) ve Interpreted ML Kısmi Bağımlılık, 7.043 müşteri kaydından oluşan gerçek bir Telekom Churn veri kümesi üzerinde karşılaştırmalı bir analizle değerlendirildi. Bu çalışmanın amacı, bu tekniklerin makine öğrenimi modellerinin şeffaflığını ve yorumlanabilirliğini artırmadaki etkinliğini, özellikle telekom abonelik iptali (churn) tahmini bağlamında karşılaştırmaktır. Çalışmada, %94,12 doğruluk oranına ulaşarak diğer makine öğrenimi modellerinden üstün performans gösteren XGBoost modeli kullanılmıştır. Metodoloji kapsamında veri ön işleme ve model eğitimi adımları açıklanmış, SHAP ve Interpreted ML Kısmi Bağımlılık yöntemleriyle yapılan iki ayrı analiz karşılaştırılmıştır. Elde edilen sonuçlar, bu tekniklerin güçlü ve zayıf yönlerini ortaya koyarak model yorumlanabilirliği ve dayanıklılığı hakkında önemli içgörüler sunmuştur. SHAP analizi, Sözleşme Türü (Contract), Aylık Ücretler (Monthly Charges) ve Teknik Destek (Tech Support) gibi özelliklerin müşteri kaybı üzerindeki etkisini güçlü bir şekilde belirleyerek, modelin belirli tahminlerde bulunma nedenlerini açıklamada etkili bir yöntem olduğunu göstermiştir. Örneğin, SHAP değerleri, kısa vadeli sözleşmelere sahip ve teknik destek hizmeti almayan müşterilerin, 0,6'nın üzerinde SHAP etkisine sahip olarak, churn olasılığının önemli ölçüde yüksek olduğunu göstermiştir. Bu detaylı analiz, bireysel müşteri tahminlerini anlamak için kritik bilgiler sağlamıştır. Öte yandan, Interpreted ML Kısmi Bağımlılık yöntemi, Aylık Ücretler ve Abonelik Süresi (Tenure) gibi değişkenlerin churn olasılığı üzerindeki genel etkilerini göstererek daha geniş bir bakış açısı sunmuştur. Analiz sonuçları, özellikle uzun süreli abonelik süresine sahip müşterilerin daha düşük churn olasılığına sahip olduğunu ortaya koymuştur. Bu çalışmanın temel katkısı, telekom sektörü için SHAP ve Interpreted ML Kısmi Bağımlılık yöntemlerinin karşılaştırmalı bir değerlendirmesini sunarak, yorumlanabilirlik ihtiyaçlarına göre uygun XAI tekniklerinin seçilmesine yönelik yapılandırılmış bir çerçeve sunmasıdır. SHAP, bireysel tahminleri açıklamak için güçlü bir araç sunarken, Kısmi Bağımlılık yöntemi makro düzeyde analizler yapmak ve yüksek seviyeli kararlar almak için faydalıdır. Bu karşılaştırmalı analiz, XAI yöntemlerinin daha iyi anlaşılmasına katkı sağlamakta ve telekom abonelik iptali tahmin modellerinde şeffaflığı artırmak için uygun tekniklerin seçilmesinin önemini vurgulamaktadır.

1. INTRODUCTION

This study provides a comprehensive comparison of SHAP (SHapley Additive exPlanations) and Interpreted ML Partial Dependence for explainability in telecom churn prediction. By applying these techniques to a real-world dataset of 7,043 telecom customers, the research identifies Contract type, Monthly Charges, and Tech Support as key churn predictors, with SHAP values exceeding 0.6 in impact. The study demonstrates that SHAP is highly effective for customer-specific explanations, making it valuable for personalized retention strategies, while Partial Dependence Plots offer a broader understanding of feature influence, aiding in strategic decision-making. The XGBoost model, which achieved 94.12% accuracy, serves as the foundation for this analysis, highlighting the balance between performance and interpretability. The findings emphasize that SHAP excels in detailed, instance-based explanations, whereas Partial Dependence is more suitable for high-level business insights. By bridging the gap between model accuracy and transparency, this study offers practical guidance on selecting the most appropriate XAI technique for telecom churn analysis, contributing to more explainable and trustworthy AI-driven decision-making [25,27]. In today's data-driven world, where machine learning models increasingly dictate decision-making, the need for transparency and interpretability has reached unprecedented importance [25,27]. Nowhere is this more critical than in telecommunications, where accurate churn prediction is not just a technical goal but a strategic necessity for customer retention and service optimization [6,10]. This paper undertakes a comparative analysis of SHAP and Interpreted ML Partial Dependence, assessing their effectiveness in enhancing model transparency and interpretability in the context of telecom churn prediction [7,31]. The Methodology section details the dataset preprocessing and model training steps, focusing on XGBoost, a widely recognized model in this domain [10,17]. Recursive Feature Elimination (RFE) is applied to identify and retain only the most influential features, ensuring improved model performance and interpretability [31,19]. The Evaluation section provides a detailed comparison of SHAP and Partial Dependence, illustrating how each technique unveils feature importance and enhances interpretability in different ways [31,33]. The discussion highlights their respective strengths and weaknesses, offering insights into their applicability for various analytical needs [8,32]. Finally, this study contributes practical recommendations on selecting appropriate XAI techniques based on business objectives and interpretability needs. The Conclusion and Future Works section summarizes key findings and discusses their broader implications, emphasizing the importance of choosing the right XAI approach in telecom churn prediction models [25,26]. Additionally, avenues for future research are outlined, focusing on further refinement of interpretability techniques and their integration into real-world telecom decision-making frameworks [33,32].

2. RELATED WORKS

The surge of machine learning applications across various sectors has been nothing short of transformative. Noteworthy contributions span diverse domains, showcasing the versatility and impact of these intelligent systems. In the realm of marketing, groundbreaking strides have been made with a coupon recommendation method, marrying machine learning prowess with simulated annealing algorithms [1]. This innovative approach not only enhances the order conversion rate but also fuels revenue growth. A parallel breakthrough emerges in the banking sector with DeepAFM, a sophisticated deep learning model tailored for precision marketing to predict potential credit card users [3].

Agriculture witnesses the fusion of technology and cultivation as machine learning techniques intricately map rice growth phases and bare land using Landsat-8 OLI imagery [2]. This not only demonstrates the accuracy of classification but also opens avenues for resource management applications. Analytical sciences delve into the intricacies of Raman and surface-enhanced Raman scattering experiments, where machine learning methods unravel chemical information from complex datasets [4].

Machine learning has made significant contributions in healthcare, such as disease classification developed a Support Vector Machine (SVM) classifier for diagnosing diabetes, showcasing the successful use of machine learning in medical datasets [5].

In the dynamic landscape of telecom, machine learning algorithms such as Prophet and XGBoost emerge as pivotal tools for forecasting network traffic [6]. The industry witnesses' cost and time reductions through automated test case generation, a testament to the efficiency brought by machine learning [7]. Anomaly detection in telecom operations sees a change in basic assumptions, resulting in a remarkable 90 percent reduction in team workload [9].

Within telecom, customer churn prediction stands out as a vital application, with machine learning models operating on big data platforms [10]. It is intriguing to note that non-machine learning AI approaches, such as propositional logic, find their place in certain tasks within telecom data analysis [8].

The literature extends its explorations into the constructive collaboration of machine learning across diverse domains. In the context of Cyber Manufacturing Systems, machine learning proves invaluable in detecting cyber-physical attacks [11]. The manufacturing industry witnesses a convergence of machine learning and optimization methods, capitalizing on advances in digitalization [12]. In radiological imaging, machine learning exhibits potential for recognizing and classifying complex patterns [14].

Furthermore, recent studies have delved into the integration of machine learning into the evaluation of physical security for cryptographic chips [15], the analysis of online transaction data [16], and its potential in medical imaging for disease detection at an early stage [18]. In the context of fault diagnosis, provided a survey of various machine learning algorithms applied in elevator systems [19]. Also specifically compared the performance of machine learning and deep learning algorithms in breast cancer prediction and diagnosis [17].

The literature reviewed demonstrates the wide-ranging applications of machine learning in diverse domains. This paper contributes to this body of knowledge by conducting a comparative analysis of Explainable Artificial Intelligence (XAI) techniques on telecom churn prediction, using SHAP and interpreted ML partial dependence.

3. MATERIAL AND METHOD

In this section, the procedures involved in data acquisition, pre-processing, and the implementation of our proposed methodology are outlined. The study commences with an explanation of the dataset used, followed by a description of the steps taken to preprocess the data. Subsequently, the proposed method is introduced, entailing the application of an XGBoost model for telecom churn prediction, augmented by Explainable AI techniques.

3.1. Dataset

The telecom churn dataset provides valuable insights into customer behaviour within the telecommunications sector, where customer loyalty plays a pivotal role. Understanding and mitigating customer subscription cancellations is critical for telecommunications companies to enhance decision-making processes. This dataset encompasses various customer attributes, including contract type, technical support status, and monthly charges, along with a binary "Churn" label indicating whether a customer terminated their subscription. It serves as a valuable resource for telecom companies seeking to analyse customer attrition patterns and formulate strategies to retain subscribers and improve overall service quality. Initially, all features of the available data set were not taken, but Recursive Feature Elimination (RFE) method was applied on the data set and the parameters obtained as output were used. Table 1 shows us the parameter ranking obtained as a result of RFE. MontlyCharges and tenure columns are reduced to values between 0-1 using Minmax Normalizer. TechSupport, Contract and OnlineSecurity columns were encoded as strings using the Label Encoding method. The churn value already contains the values 0 and 1 and shows whether the telecom service continues or not depending on the current parameters. "1" indicates that he left the service and "0" indicates that he did not leave the service.

Table 1. Feature Selection and Ranking Table

Feature	Selected	Ranking
gender	FALSE	4
Partner	FALSE	6
Dependents	FALSE	tab
tenure	TRUE	1
PhoneService	FALSE	12
InternetService	FALSE	5
OnlineSecurity	TRUE	1
OnlineBackup	FALSE	3
DeviceProtection	FALSE	8
TechSupport	TRUE	1
StreamingTV	FALSE	11
StreamingMovies	FALSE	10
Contract	TRUE	1
PaperlessBilling	FALSE	7
PaymentMethod	FALSE	2
MonthlyCharges	TRUE	1

Table 2. A piece of example from dataset

Monthly Charges	Tech Support	Contract	OnlineSecurity	tenure	Churn
0.115423	0	0	0	0.013889	0
0.385075	0	1	1	0.472222	0
0.354229	0	0	0	0.027778	1
0.239303	2	2	1	0.625	0
0.521891	0	0	0	0.027778	1

3.2. Proposed Model

Following the training of the Telecom Churn dataset using the XGBoost model predictions for customer churn were obtained. To gain deeper insights into the "why" and "how" behind these predictions leveraging interpretable artificial intelligence tools, namely InterpretML and SHAP were proposed. This article aims to explore and compare the explanations provided by these tools, shedding light on the factors influencing the model's decisions.

InterpretML and SHAP are chosen for their effectiveness in offering interpretability to complex machine learning models. InterpretML employs various interpretability techniques to provide a global understanding of model behaviour, while SHAP (SHapley Additive exPlanations) values offer a nuanced view of feature contributions for individual predictions.

In the proposed methodology, interpretability results obtained from InterpretML and SHAP will be analysed and contrasted to unravel the underlying dynamics of the XGBoost model. By addressing the "why" and "how" aspects of predictions, an aim is set to enhance the transparency and trustworthiness of the predictive model. This is expected to contribute to a more informed decision-making process in the context of telecom churn prediction.

In addition, this data set was created with data from 7043 people. An example of the data set is given in Table 2.

To enhance the reliability of the results obtained from SHAP and InterpretML, the performance and methodologies of these tools have been carefully evaluated and independently analysed. These analyses are crucial in ensuring the accuracy and reliability of the model's predictions.

3.2.1. Machine learning

Machine learning, a subset of artificial intelligence, has significantly transformed data analysis methodologies across various domains [20, 23]. This changes in basic assumptions is evident in the telecommunications industry, where machine learning plays a pivotal role, particularly in the prediction of customer churn. The application of machine learning algorithms empowers telecom companies to navigate vast datasets, discern intricate patterns, and make highly accurate predictions. Amidst the diverse array of machine learning algorithms, XGBoost emerges as a preferred choice due to its adaptability and robust performance [21].

XGBoost, denoting eXtreme Gradient Boosting, has evolved into a cornerstone within the realm of machine learning. Tailored to overcome the limitations of traditional gradient boosting algorithms, XGBoost excels in both classification and regression tasks. Its effectiveness stems from its adeptness in handling diverse data types, managing missing values, and mitigating overfitting. These attributes render XGBoost particularly well-suited for intricate prediction scenarios, such as telecom churn prediction. In the specific context of telecom churn prediction, XGBoost has emerged as a go-to solution for telecom companies. Leveraging its analytical prowess across various customer attributes, including contract type, technical support status, and monthly charges, XGBoost provides accurate forecasts regarding customer subscription cancellations. As the subsequent sections of this article are explored, focus will be shifted towards the interpretation of the predictions generated by the XGBoost model. This interpretative process involves the utilization of tools such as InterpretML and SHAP, enhancing the transparency of the model. The interpretability of the XGBoost model is deemed crucial for gaining profound insights into the factors influencing telecom churn. Ultimately, this knowledge assists businesses in making well-informed decisions aimed at customer retention and the enhancement of service quality. This narrative builds upon the foundation laid by various scholarly works in the field of machine learning [23,24], among others. The integration of insights from these works contributes to a comprehensive understanding of the intersection between machine learning and telecommunications, enriching the discourse on predictive analytics in this dynamic industry. Figure 1 gives a working example of machine learning.

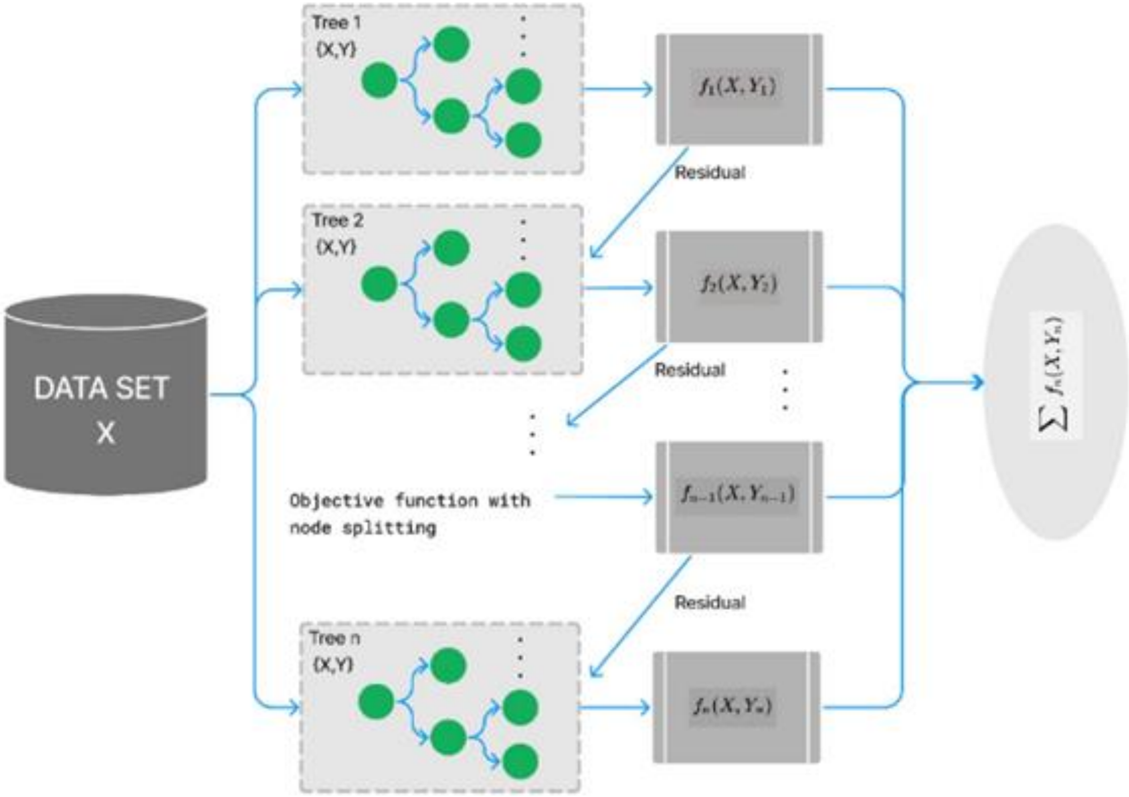

Figure 1. Working of machine learning

3.2.2. Explainable AI

Explainable AI, often abbreviated as XAI, represents a paradigm shift in artificial intelligence, emphasizing the transparency and interpretability of machine learning models [25]. As AI systems become increasingly intricate, understanding their decision-making processes becomes paramount. Explainable AI aims to demystify the inner workings of these models, providing comprehensible insights into how and why specific predictions are made [26]. In the realm of telecom churn prediction, where the stakes are high, Explainable AI plays a pivotal role [27]. Identifying the key parameters influencing churn is crucial for telecom companies seeking to retain customers effectively. With the application of machine learning models, particularly XGBoost in our case, a plethora of features contribute to predicting churn [28]. However, discerning which factors carry the most weight can be challenging without the aid of Explainable AI [29]. In the telecom churn landscape, parameters such as contract type, technical support status, and monthly charges emerge as influential factors [27]. Machine learning models analyse these features, but the "why" behind a prediction remains a black box [29]. A schematic visual presenting the Explainable AI method is given in Figure 2 also Figure 3 shows a comparative XAI method with machine learning algorithms.

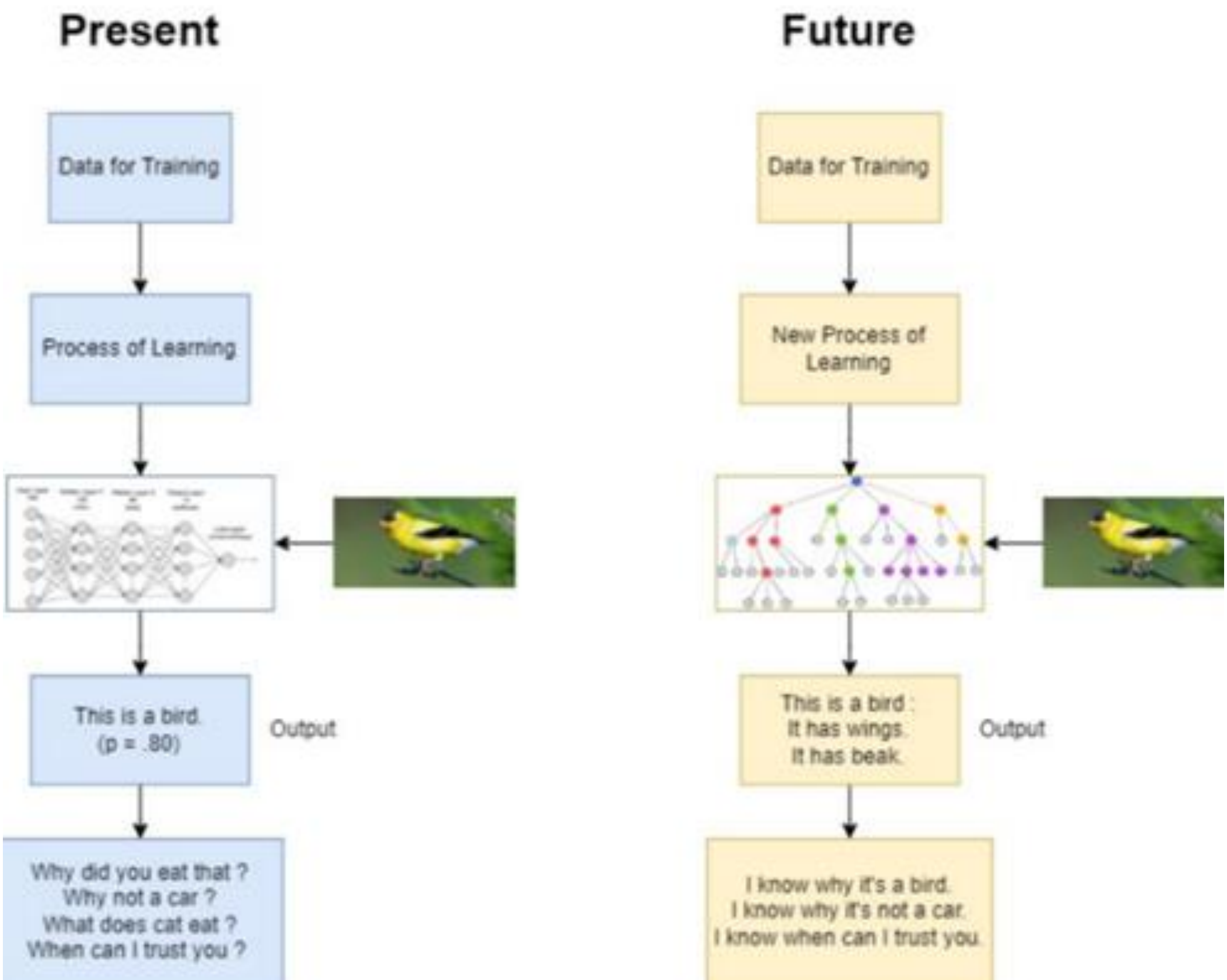

Figure 2. Explainable AI schema

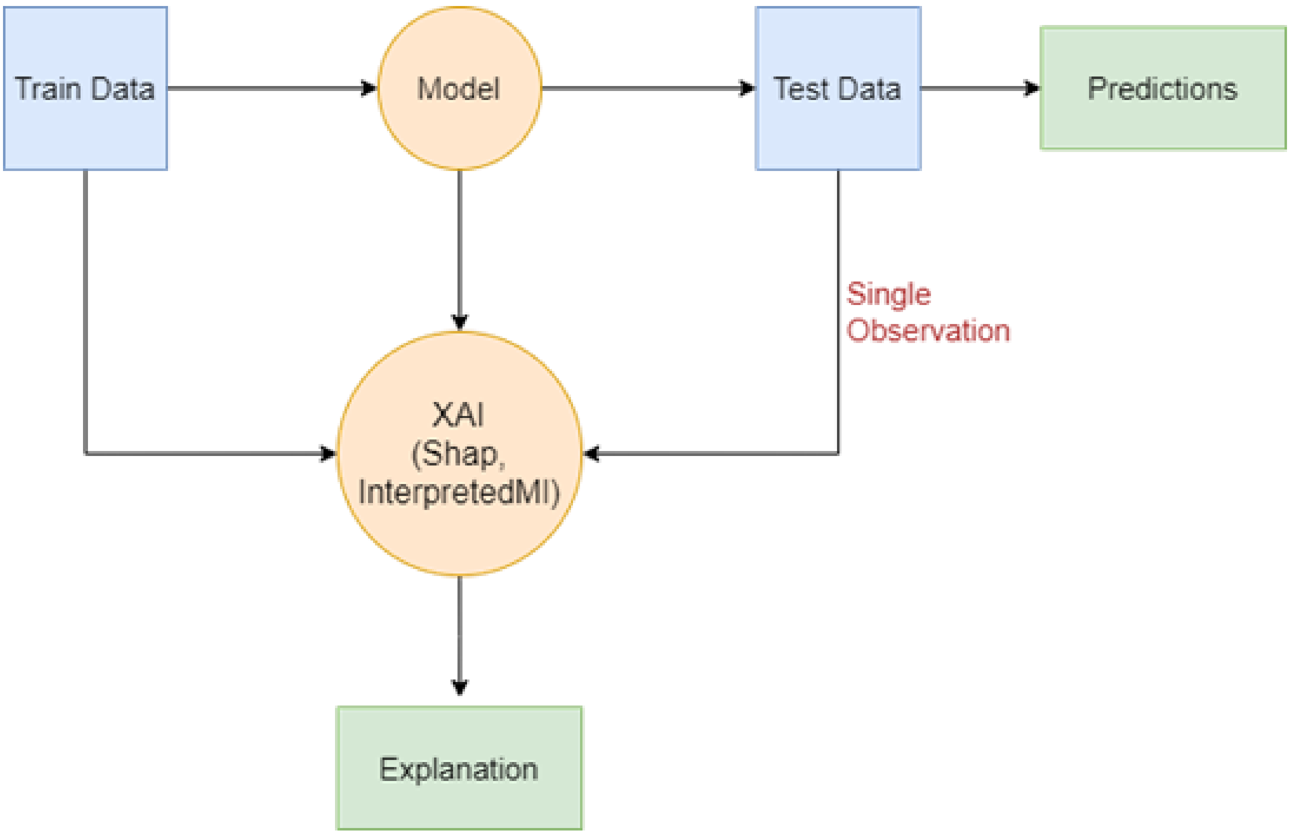

Figure 3. Comparative of XAI and Machine Learning

Explainable AI tools, such as InterpretML and SHAP, step in to shed light on the decision rationale, offering a transparent view of how these parameters contribute to the prediction of churn. After employing machine learning techniques like XGBoost to predict telecom churn, the next logical step is to leverage Explainable AI. InterpretML and SHAP, for instance, provide methodologies for interpreting and visualizing complex model outputs. By applying these tools, businesses can uncover the significance of each feature in the churn prediction process. This not only enhances the credibility of the predictions but also empowers telecom companies with actionable insights. Decision-makers can now understand not only that a customer is likely to churn but also the specific reasons driving that prediction, enabling them to implement targeted retention strategies and improve overall customer satisfaction.

3.3. SHAP (Shapley Additive exPlanations)

SHAP values are derived from cooperative game theory, specifically Shapley values, which distribute the contribution of each player in a coalition [32]. Applied to machine learning, SHAP values quantify the impact of each feature on a model's output [31]. For a given prediction f(x) of the model, the SHAP value ϕ_i for feature i can be expressed as:

$$\phi_i(f) = \sum_{S \subseteq N\{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \quad (1)$$

Here, N is the set of all features, and S represents a subset of features excluding [31]. This formula computes the contribution of feature by considering all possible subsets of features.

$$f_x(S) = E[f(x)|X_s] \quad (2)$$

$f_x(S)$ is the conditional expectation of f(x) given the information set X_s,

E[.] denotes the expectation operator,

f(x) is a random variable, and

X_s is an information set.

In Equation 2, the conditional expectation $f_x(S)$ is determined by the expected value of the random variable f(x) given the information set X_s. The information set X_s conditions the expectation, indicating that the anticipated value of f(x) is contingent upon the knowledge encapsulated in X_s. This formulation is a standard representation in probability theory and statistics, where the conditional expectation provides a way to model the expected outcome of a random variable in the presence of specific information or conditions specified by the information set. The equation serves as a concise expression encapsulating the relationship between the conditional expectation and the underlying random variable in a given informational context.

In the equation,

$$M + \sum_{i=1}^{n} \beta_i z_i = bias + \sum_{i=1}^{n} \beta_i \cdot Feature\ Contribution \mid x_i \quad (3)$$

where M represents an intercept term, and β_i denotes the coefficients associated with the variables z_i and x_i. The variables z_i and x_i are the features used in the model. The left side of the equation is the sum of the intercept $z_i$and the product of each coefficient $\beta_i$with its corresponding variable z_i.

The right side of the equation consists of the bias term and the sum of the product of each coefficient β_i with its corresponding feature x_i. This can be interpreted as the sum of the individual contributions of each feature x_i to the overall prediction.

Overall, the equation represents a linear relationship between the features and the predicted outcome, with each feature's contribution weighted by its respective coefficient.

The XAI output in Figure 4 is a SHAP (SHapley Additive exPlanations) value plot, which is a way of explaining the predictions of a machine learning model by looking at how much each feature contributes to the output. In this case, the model is predicting revenue, and the XAI output shows how much each of the features in the plot contributes to the predicted revenue [30]. Also Figure 5 shows the average impact of our features on model output.

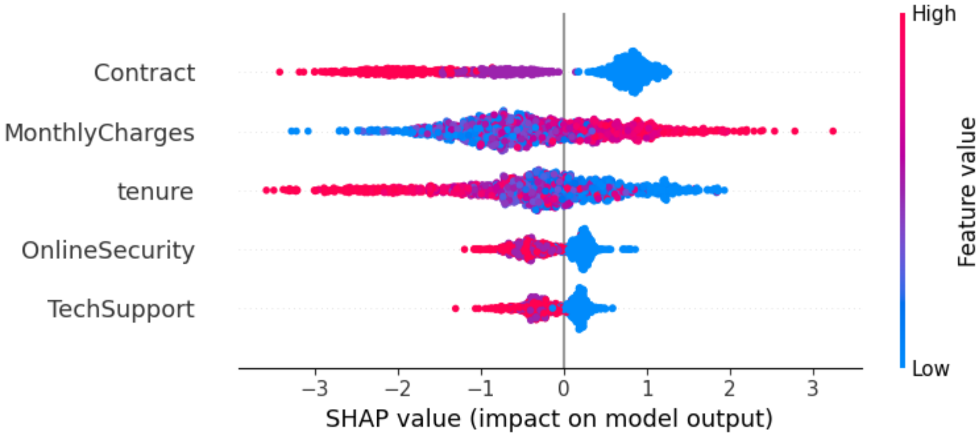

Figure 4. SHAP results for dataset

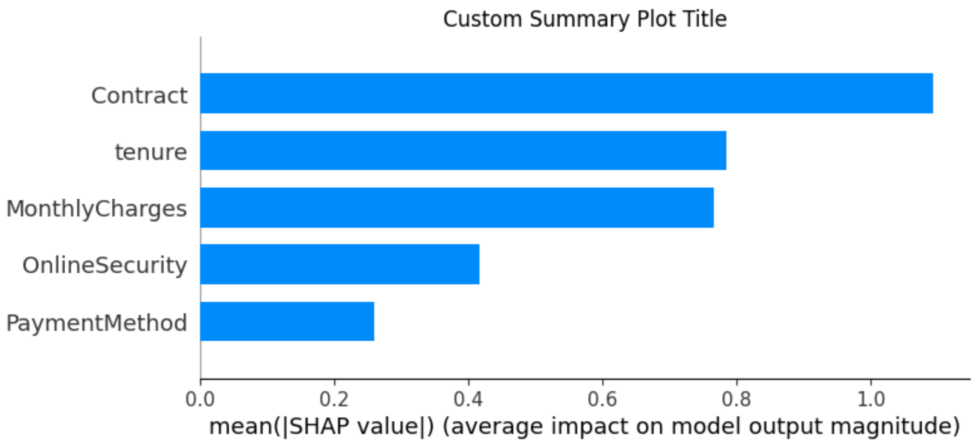

Figure 5. SHAP results for our data set shows average impact for each feature

The features in the plot are contract, monthlyCharges, tenure, onlineSecurity, and techSupport .The SHAP value for each feature is shown on the y-axis, and the features are arranged on the x-axis from lowest to highest SHAP value. The SHAP value for a feature tells you how much the model's prediction would change if the value of that feature were to change. For example, the SHAP value for contract is positive, which means that increasing the value of contract will increase the predicted revenue. Conversely, the SHAP value for tenure is negative, which means that increasing the value of tenure will decrease the predicted revenue [33].

The absolute value of the SHAP value tells you how important the feature is to the model's prediction. The features with the largest absolute SHAP values are the most important to the model. In this case, the most important features are contract, monthlyCharges, and onlineSecurity. The XAI output also shows how the SHAP values are distributed across different data points. The grey line in the plot shows the average SHAP value for each feature and the blue and red dots show the SHAP values for individual data points. The spread of the blue and red dots shows how much the SHAP values can vary for different data points. Overall, the XAI output tells you that the value of a contract, monthly charges, online security, and tech support has a significant impact on the predicted revenue. The higher the value of these features, the higher the predicted revenue. Tenure also has an impact on the predicted revenue, but it is a negative impact. The higher the value of tenure, the lower the predicted revenue.

3.4. Interpreted ML Partial Dependence

Partial dependence plots offer insights into the relationship between a specific feature and the model's predictions while holding other features constant. Interpreted ML Partial Dependence extends this concept to make partial dependence more accessible for interpretation. For a model M with input features X and output Y, the partial dependence $PDx_i(x_i)$ for a specific feature X_i at a certain value x_i is calculated as follows:

$$PDx_i(x_i) = E_{X \sim p(X)}[M(x)|X_i = x_i] \tag{4}$$

This equation represents the expected prediction of the model M given a fixed value x_i for the feature X_i. It provides a more intuitive understanding of how changes in a single feature influence the model's output.

Partial dependence plots (PDPs) are invaluable tools in the realm of interpretability for machine learning models. As illustrated in Figure 6 of the InterpretML framework the nuanced changes in the average response value of the model as we vary the tenure feature across the 0-1 range on both the x and y axes were observed. These plots provide a visual representation of how the feature's values influence the model predictions.

$$fs(s) = Exc[f(xs, Xc)] = \int f(xs, Xc)dP(Xc) \quad (5)$$

The function fs is defined as the expected value, denoted as Exc, of the conditional function $f(x_s, X_C)$, where x_s represents the features of interest and X_c denotes the remaining features treated as random variables in the machine learning model f. This expectation is obtained by integrating the function $f(x_s, X_C)$ with respect to the probability distribution $P(X_c)$.

In simpler terms, $f_s(s)$ signifies the average predicted outcome based on the features x_s of interest, considering the variability in the remaining features X_c according to their probability distribution $P(X_c)$. The integral reflects the process of taking into account the contribution of each combination of x_s and X_c values weighted by their respective probabilities in the overall prediction.

$$f_s(x_S) = \frac{1}{n}\sum_{i=1}^{n} f^{\wedge}(x_S, x_C^{(i)}) \quad (6)$$

Equation (7) expresses the partial dependence function $f_s(x_S)$, which is determined by calculating the average of the predicted values from the machine learning model $f^{\wedge}$ for the features of interest x_S. This average is obtained by summing the predicted values for each instance i and then dividing the sum by the total number of instances n. The feature values $x_C^{(i)}$ represent the actual features from the dataset that are not of interest, emphasizing the marginalization over these features in the calculation of the partial dependence. In passive voice, the equation conveys that the partial dependence function is computed by summing the predicted values for each instance and subsequently dividing the sum by the total number of instances.

$$I(x_S) = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}((f^{\wedge}(x_S^{(k)}) - \frac{1}{K}\sum_{k=1}^{K}(f^{\wedge}(x_S^{(k)}))^2} \quad (7)$$

The equation $I(x_S)$ is formulated to calculate a measure of variability or dispersion associated with the partial dependence estimates $f^{\wedge}(x_S^{(k)})$ for a set of feature values x_S. This measure is expressed as the square root of the average squared difference between each individual estimate and the overall average estimate.

In more detail, the expression involves obtaining K partial dependence estimates for the features in set S, denoted as $f^{\wedge}(x_S^{(k)})$ for $k = 1$ to K. The quantity inside the square root signifies the average squared deviation of each individual estimate from the mean estimate, which is the average of all K estimates. The fraction $\frac{1}{K-1}$ scales the summation to provide an unbiased estimate of the variance.

The resulting value $I(x_S)$ represents the standard deviation of the partial dependence estimates, offering insights into the spread or consistency of the model's predictions for different instances of x_S. This measure can be useful in assessing the stability and reliability of the partial dependence estimates for the specified set of features.

$$(x_S) = \frac{max_k\left(Sf\left(x_S^{(k)}\right)\right) - min_k\left(Sf\left(x_S^{(k)}\right)\right)}{4} \quad (8)$$

The formulation of the expression (x_S) provides a standardized representation for the features of interest x_S, achieved by calculating the difference between the maximum and minimum values of the partial dependence function Sf evaluated at different instances $x_S^{(k)}$, and subsequently dividing this difference by four. This normalization step ensures that the range of values resulting from the partial dependence function is adjusted, facilitating a more interpretable understanding of the features' impact on the model's predictions.

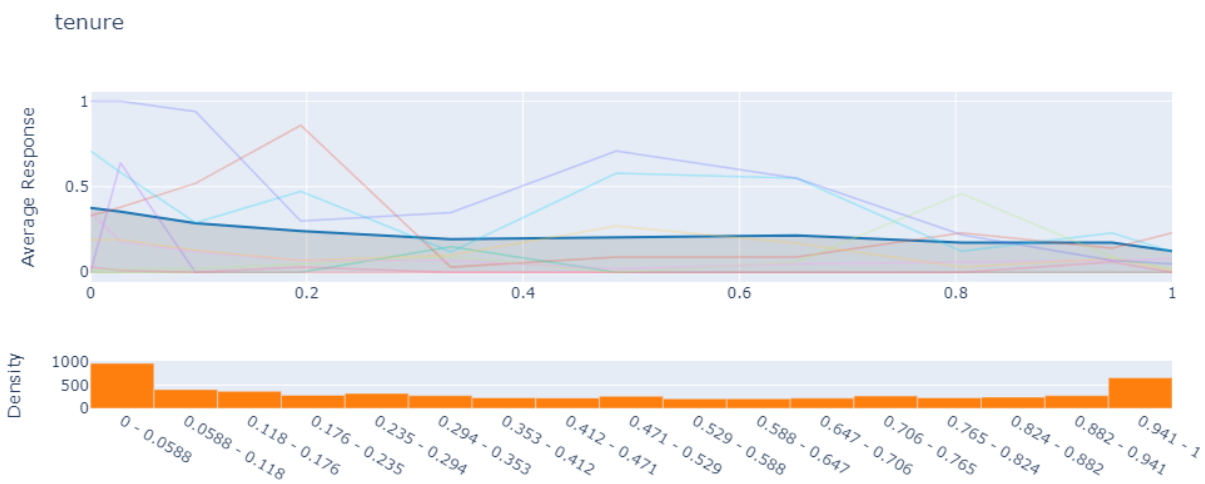

Figure 6. Tenure results InterpretML from dataset

In transitioning to the subsequent paragraph, the focus shifts to Figure 6, which specifically examines the tenure feature. Within the specified range, this figure elucidates the average response values, with the x and y axes capturing variations in the feature and its impact on the model's output, respectively. The plot's density visually encapsulates the concentration of the tenure feature's influence on the overall model.

In summary, these Partial Dependence plots serve as powerful tools for model interpretation, providing a clear and intuitive understanding of how individual features contribute to the model's behaviour. The visualizations not only depict the average response values but also convey the intensity of the feature's impact, enhancing our grasp of the model's inner workings. Incorporating such visual aids in high-level articles and discussions can significantly contribute to the accessibility and interpretability of complex machine learning models.
In our implementation, these Explainable AI techniques are applied post-training the XGBoost model. SHAP values provide a detailed breakdown of feature contributions, while Interpreted ML Partial Dependence offers insightful plots illustrating the impact of individual features on predictions.

4. RESULTS

Table 3 presents the results of experiments conducted with various machine learning models on the same dataset. Upon careful examination of the table, it becomes evident that the XGBoost model outperforms other machine learning models, exhibiting higher accuracy, precision, recall and F1 score. The superior performance of the XGBoost model underscores its effectiveness in capturing and generalizing patterns within the given dataset. This heightened accuracy, precision, recall, and F1 score not only emphasize the model's predictive prowess but also its robustness in consistently delivering reliable and high-quality predictions compared to alternative machine learning approaches. The model's capacity to adapt to intricate patterns and make accurate predictions contributes to its superior performance metrics across multiple evaluation criteria. The nuanced evaluation metrics, including precision, recall, and F1 score, shed light on the model's ability to strike a balance between correctly identifying positive instances and minimizing false positives and false negatives.

Table 3. Machine learning algorithm results

Machine Learning Methods	Accuracy	Precision	Recall	F1-Score
XGBoost	0.94125	0.94	1	0.97
Adaboost	0.93933	0.94	1	0.97
KNN	0.93933	0.94	1	0.97
CatBoost	0.93933	0.94	1	0.97
Logistic Regression	0.93933	0.94	1	0.97
SVM	0.93933	0.94	1	0.97
TFX	0.93835	0.94	1	0.97
Tensorflow Extends BBO	0.93835	0.94	1	0.97
Extra Trees	0.93737	0.94	1	0.97
Decision Tree	0.93542	0.94	0.99	0.97
Random Forest	0.93542	0.94	0.99	0.97
LightGBM	0.93346	0.94	1	0.97
NaiveBayes	0.43639	1	0.4	0.57

In this study, interpretable artificial intelligence tools, namely Shap and InterpretML-Partial Dependence, were employed to elucidate the XGBoost model trained on the telecom churn dataset. The primary goal was to explain the predictions made by the model, with a particular focus on two selected customers—one continuing with the service (churn=0) and the other leaving the company (churn=1). By comparing Shap outputs for these customers, it was observed that the 734th customer had a longer tenure than the 544th customer, as depicted in Figure 7. Additionally, Figure 7 revealed that the MonthlyCharges for the 544th customer were lower than those for the 734th customer.


```
Churn 1 Example Features:
      MonthlyCharges  TechSupport  Contract  OnlineSecurity    tenure
544        0.271642            2         0               2  0.180556

Churn 0 Example Features:
      MonthlyCharges  TechSupport  Contract  OnlineSecurity  tenure
734        0.715423            0         0               0   0.375
```

Figure 7. Customers parameters

Examining Figure 8 and Figure 9, the feature distributions influencing predictions for these selected customers were illustrated. These visualizations supported the interpretation that the 734th customer did not churn, while the 544th customer did. Notably, Figure 11 and Figure 12 indicated that the Contract feature played a crucial role in predicting churn for both customers. However, a discrepancy arose between Monthly Charges and tenure, with Figure 9 showing higher values for the non-churning customer, whereas Figure 8 revealed a significant impact of these features, both exceeding 0.6.

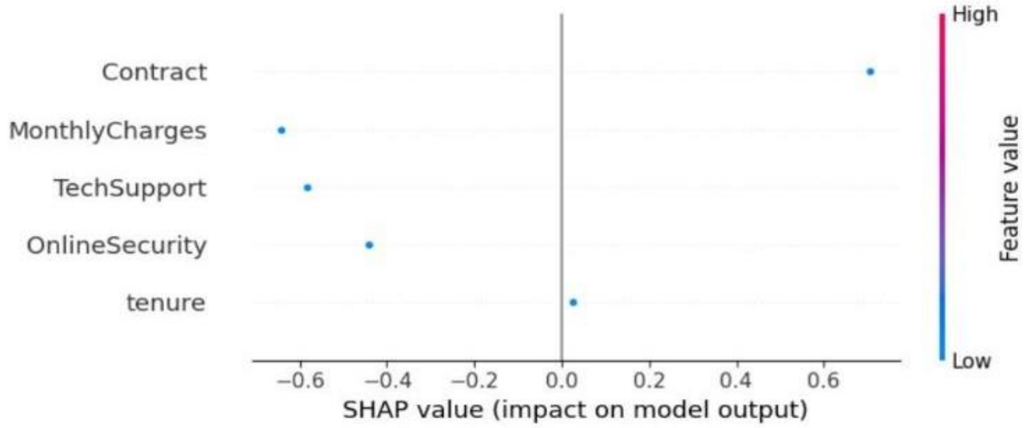

Figure 8. 544’th Customers SHAP values

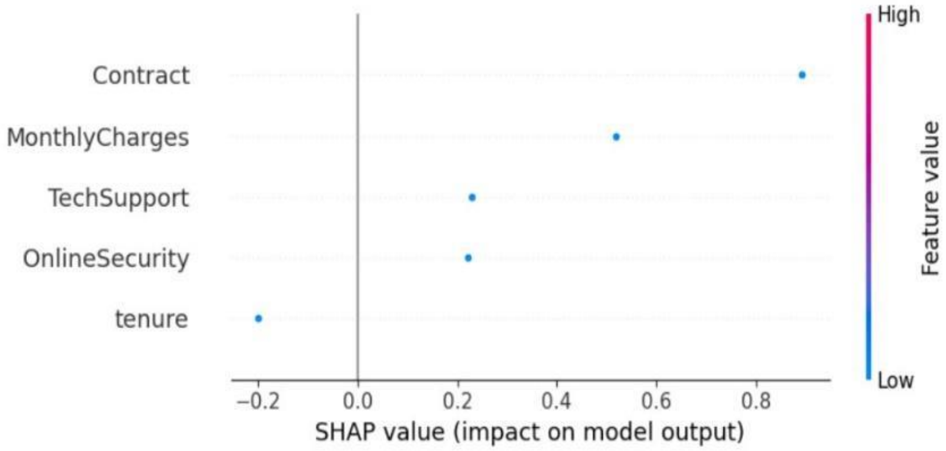

Figure 9. 734’th Customers SHAP values

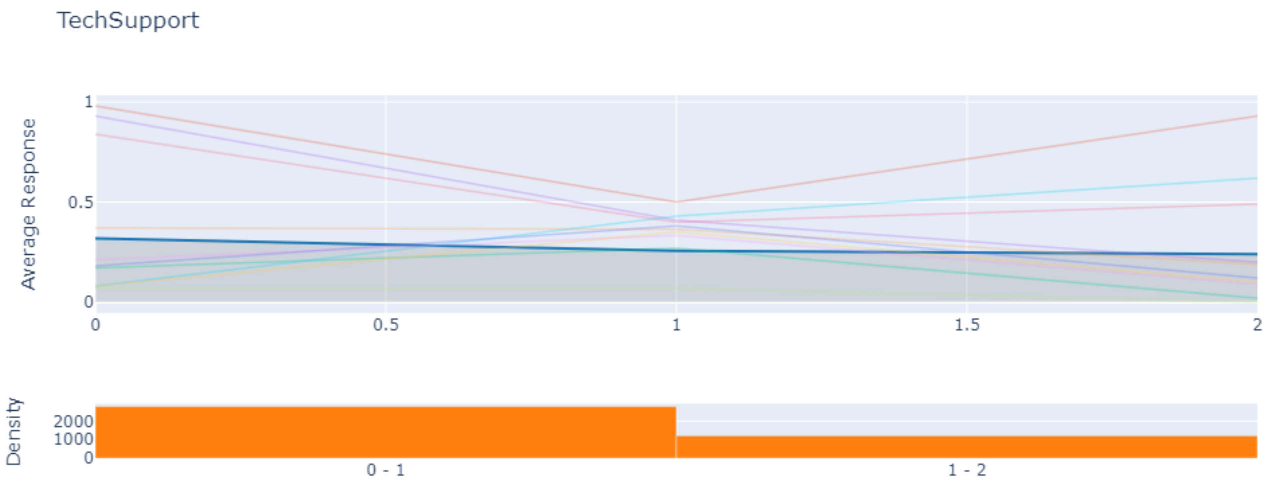

Figure 10. InterpretML results for TechSupport

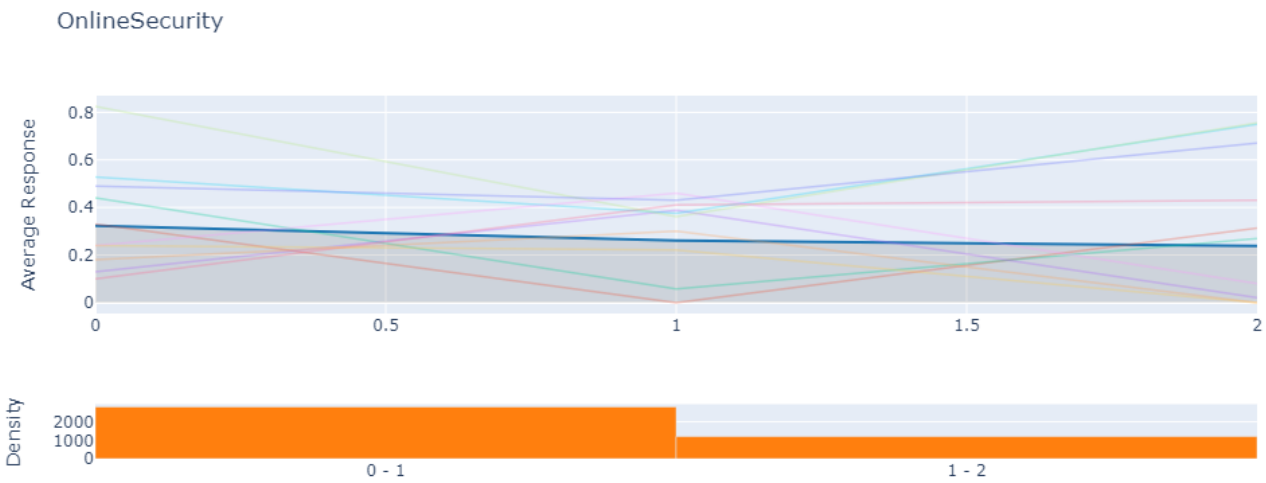

Figure 11. InterpretML results for OnlineSecurity

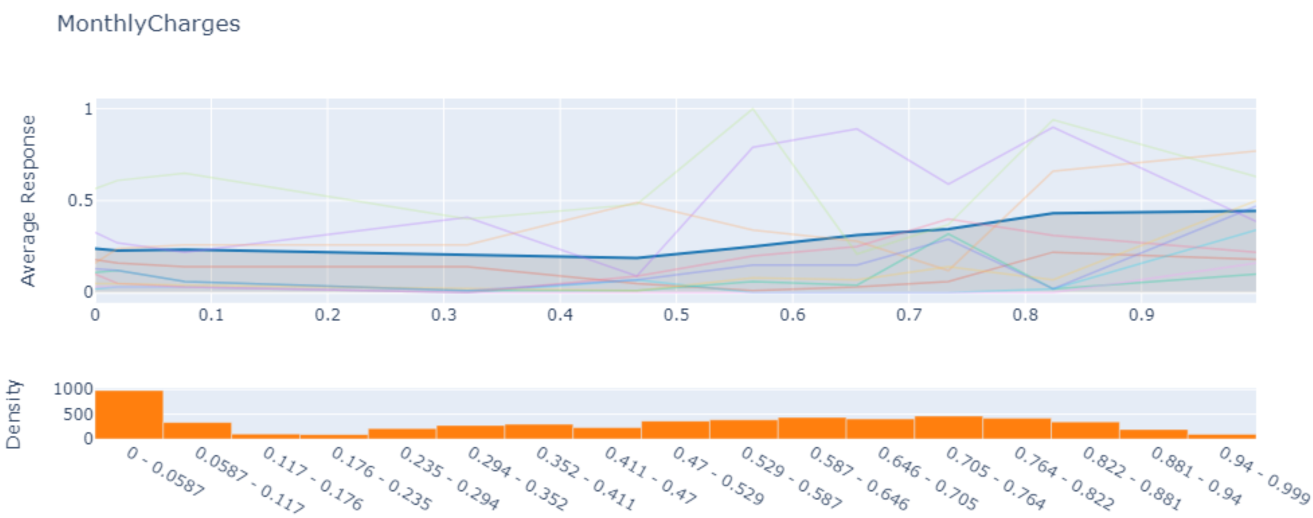

Figure 12. InterpretML results for MontlyCharges

Furthermore, limitations were encountered in delivering detailed explanations for the reasons behind specific customer churn or retention when employing InterpretML-Partial Dependence. Unlike Shap, InterpretML-Partial Dependence offers insights into how the model's prediction changes as a single feature varies while keeping all other features constant. It provides a partial dependence plot highlighting the average response of the prediction and the density of these features' influence on the prediction, as depicted in Figure 10.

The InterpretML partial dependence graphics shown at Figure 10,11 and 12 for the monthly charges, contract type, and tech support features offer a nuanced understanding of how these variables collectively influence the predictions, particularly in the context of OnlineSecurity. These graphics provide insightful visualizations that depict the relationship between the mentioned features and the predicted outcomes. By examining the partial dependence plots, one can discern the impact of monthly charges, contract type, and the availability of tech support on the likelihood of having OnlineSecurity services. The partial dependence graphics illustrate how changes in the values of monthly charges, various contract types, and the presence of tech support correlate with shifts in the model's predictions regarding the probability of having OnlineSecurity. This visual representation aids in identifying trends, patterns, and potential nonlinear relationships between these features and the target variable. Consequently, stakeholders can glean valuable insights into the factors that significantly contribute to the presence or absence of OnlineSecurity services, enabling informed decision-making and a deeper understanding of the model's behaviour in relation to these key predictors.

5. DISCUSSION AND CONCLUSION

The results obtained from this study highlight the effectiveness of SHAP and Interpreted ML Partial Dependence in enhancing the interpretability of telecom churn prediction models. The findings indicate that SHAP is particularly effective in identifying key features that influence individual customer churn decisions, whereas Interpreted ML Partial Dependence provides a more global perspective on feature importance and trends. Several previous studies have explored Explainable AI techniques in the context of customer churn prediction, primarily focusing on improving predictive accuracy rather than model interpretability. While prior research demonstrated the effectiveness of machine learning models such as XGBoost in telecom forecasting, they often lacked a comparative evaluation of explainability techniques. Our study builds on this by integrating SHAP and Interpreted ML Partial Dependence to bridge the gap between model performance and transparency. One of the key advantages of the proposed approach is its ability to provide improved transparency, enabling telecom companies to understand and trust AI-driven predictions. The SHAP analysis offers actionable insights by pinpointing individual-level risk factors, allowing for more personalized retention strategies. At the same time, Interpreted ML Partial Dependence presents a high-level overview of feature importance, which is useful for strategic decision-making. Additionally, the XGBoost model used in this study achieved a high predictive accuracy of 94.12%, reinforcing its reliability in telecom churn prediction. However, there are some limitations to consider. SHAP computations can be computationally expensive, especially for large datasets, making real-time analysis challenging. Moreover, Partial Dependence Plots do not fully capture feature interactions, limiting their ability to reflect complex dependencies within the data. There is also a trade-off between interpretability and performance, as the use of interpretability techniques sometimes necessitates a compromise on model complexity and predictive power. Despite these limitations, the study provides a valuable framework for applying and evaluating Explainable AI techniques in telecom churn prediction. Future research should focus on integrating additional interpretability methods, such as counterfactual explanations or LIME, to further enhance the transparency and usability of machine learning models in the telecom sector.

This study presents a comparative analysis of SHAP (SHapley Additive exPlanations) and Interpreted ML Partial Dependence to enhance the interpretability of telecom churn prediction models. By analyzing a real-world telecom dataset of 7,043 customers, the findings reveal that Contract type, Monthly Charges, and Tech Support are the most significant factors affecting customer churn, with SHAP values exceeding 0.6 in impact. The XGBoost model, achieving 94.12% accuracy, demonstrated high predictive performance, while the applied explainability techniques provided valuable insights into customer retention strategies. SHAP proved particularly effective in offering individual-level explanations, making it an ideal tool for personalized churn intervention strategies, highlighting that customers on short-term contracts and those without tech support are at higher risk of leaving. Conversely, Interpreted ML Partial Dependence offered a macroscopic view of feature impact, illustrating general trends such as the inverse relationship between tenure and churn, making it more suitable for strategic decision-making at a higher level. These insights suggest that telecom companies should focus on contract renewal incentives, enhanced customer support, competitive pricing strategies, and loyalty programs to mitigate churn risk. Future research should explore advanced feature engineering to refine customer behavior insights, hyperparameter tuning for improved model accuracy, and integration of additional behavioral metrics such as customer service interactions, payment history, and network usage patterns to strengthen predictive capabilities. Additionally, evaluating the real-world impact of XAI techniques in telecom operations and incorporating other interpretability methods like counterfactual explanations or LIME could provide a more comprehensive understanding of churn drivers. By applying context-specific XAI approaches, telecom providers can make data-driven decisions to enhance customer retention, improve pricing strategies, and optimize service offerings, ultimately ensuring greater transparency and trust in AI-driven predictive models.

Data Availability

The kaggle link of dataset is "https://www.kaggle.com/datasets/moe5998/telecom-customer-churn".

Conflict of Interest

The author declare that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical Considerations

This research adheres to ethical principles and guidelines in conducting the comparative analysis of Explainable Artificial Intelligence (XAI) techniques, specifically SHAP (SHapley Additive exPlanations) and Interpreted ML Partial Dependence, on a Telecom Churn dataset.

REFERENCES

[1] Y. Qi, Jun Wang, Xin Ma, Na Dong, Hong-Mei Yuan, "Coupon Recommendation System Based on Machine Learning and Simulated Annealing Algorithm," Other Conferences, 2023.

[2] Fadhlullah Ramadhani, Reddy Pullanagari, Gabor Kereszturi, Jonathan Procter, "Mapping of Rice Growth Phases and Bare Land Using Landsat-8 OLI with Machine Learning," International Journal of Remote Sensing, 2020.

[3] Yue Qiu, Jianan Fang, "Prediction of Potential Credit Card Users of Bank Based on Deep Learning," Neural Networks, Information and Communication Engineering, 2022.

[4] Félix Lussier, Vincent Thibault, Benjamin Charron, Gregory Q. Wallace, Jean-Francois Masson, "Deep Learning and Artificial Intelligence Methods for Raman and Surface-enhanced Raman Scattering," Trends in Analytical Chemistry, 2020.

[5] V Kirankumar, Somula Ramasubbareddy, G Kannayaram, K Nikhil Kumar, "Classification Of Diabetes Disease Using Support Vector Machine," Journal of Computational and Theoretical Nanoscience, 2019.

[6] Garima Jain, Rajeev Ranjan Prasad, "Machine Learning, Prophet and XGBoost Algorithm: Analysis of Traffic Forecasting in Telecom Networks with Time Series Data," 2020 8th International Conference on Reliability, Infocom ..., 2020.

[7] K. Kikuma, Takeshi Yamada, Koki Sato, K. Ueda, "Preparation Method in Automated Test Case Generation Using Machine Learning," Proceedings of the 10th International Symposium on ..., 2019.

[8] Sandro Skansi, Kristina Sekrst, Marko Kardum, "A Different Approach for Clique and Household Analysis in Synthetic Telecom Data Using Propositional Logic," 2020 43rd International Convention on Information, ..., 2020.

[9] Ryan Mukai, Zaid Towfic, Monika Danos, Mazen Shihabi, David Bell, "MSL Telecom Automated Anomaly Detection," 2020 IEEE Aerospace Conference, 2020.

[10] Abdelrahim Kasem Ahmad, Assef Jafar, Kadan Aljoumaa, "Customer Churn Prediction in Telecom Using Machine Learning in Big Data Platform," Journal of Big Data, 2019.

[11] Mingtao Wu, Zhengyi Song, Young B. Moon, "Detecting Cyber-physical Attacks in CyberManufacturing Systems with Machine Learning Methods," Journal of Intelligent Manufacturing, 2019.

[12] Dorina Weichert, Patrick Link, Anke Stoll, Stefan Rüping, Steffen Ihlenfeldt, Stefan Wrobel, "A Review of Machine Learning for The Optimization of Production Processes," The International Journal of Advanced Manufacturing ..., 2019.

[13] Raffaele Cioffi, Marta Travaglioni, Giuseppina Piscitelli, Antonella Petrillo, Fabio De Felice, "Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions," Sustainability, 2020.

[14] Zhenwei Zhang, Ervin Sejdic, "Radiological Images And Machine Learning: Trends, Perspectives, And Prospects," arXiveess.IV, 2019.

[15] Jiajing Liu, Congming Wei, Shengjun Wen, An Wang, "Design and Implementation of A Physical Security Evaluation System for Cryptographic Chips Based on Machine Learning," Conference on Intelligent Computing and Human-Computer, 2023.

[16] Kailai Chen, "On A Machine Learning Based Analysis of Online Transaction," Conference on Machine Learning and Computer Application, 2023.

[17] W. A. W. A. Bakar, M. Zuhairi, M. Man, J. A. Jusoh, N. L. N. Josdi, "Deep Learning Algorithm Vs XGBoost Using Wisconsin Breast Cancer Diagnosis," Conference on Computer Science and Communication Technology, 2022.

[18] Angona Biswas, MD Abdullah Al Nasim, Md Shahin Ali, Ismail Hossain, Dr. Md Azim Ullah, Sajedul Talukder, "Active Learning on Medical Image," arXiv-eess.IV, 2023.

[19] Jun Gong, Yueyi Zhang, Siji Chen, Jingnan Liu, "Survey on The Application of Machine Learning in Elevator Fault Diagnosis," International Conference on Applied Mathematics, Modelling, 2023.

[20] Johnson, J. M. (2019). Survey on Deep Learning with Class Imbalance.

[21] Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., Bennamoun, M. (2019). Deep Learning For 3D Point Clouds: A Survey.

[22] Chen, J., Ran, X. (2019). Deep Learning With Edge Computing: A Review.

[23] Hüllermeier, E., Waegeman, W. (2019). Aleatoric And Epistemic Uncertainty In Machine Learning: An Introduction To Concepts And Methods.

[24] Dargan, S., Kumar, M., Ayyagari, M. R., Kumar, G. (2019). A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning.

[25] Duan, Y., Edwards, J. S., Dwivedi, Y. K. (2019). Artificial Intelligence for Decision Making in The Era of Big Data - Evolution, Challenges and Research Agenda. International Journal of Information Management, IF: 7.

[26] Long, D., Magerko, B. (2020). What Is AI Literacy? Competencies and Design Considerations. Proceedings of the 2020 CHI Conference on Human Factors in..., IF: 6.

[27] Shrestha, Y. R., Ben-Menahem, S. M., von Krogh, G. (2019). Organizational Decision-Making Structures in The Age of Artificial Intelligence. California Management Review, IF: 5.

[28] Shen, Y., Song, K., Tan, X., Li, D., Lu, W., Zhuang, Y. (2023). HuggingGPT: Solving AI Tasks with ChatGPT and Its Friends in Hugging Face. arXiv-CS.CL, IF: 5.

[29] Ma, S., Lei, Y., Wang, X., Zheng, C., Shi, C., Yin, M., Ma, X. (2023). Who Should I Trust: AI or Myself? Leveraging Human and AI Correctness Likelihood to Promote Appropriate Trust in AI-Assisted Decision-Making. arXiv-CS.HC, IF: 3.

[30] Carlos Fernández-Loría; Foster Provost; Xintian Han; "Explaining Data-Driven Decisions Made By AI Systems: The Counterfactual Approach", ARXIV-CS.LG, 2020.

[31] Remah Younisse; Ashraf Ahmad; Q. A. Al-Haija; "Explaining Intrusion Detection-Based Convolutional Neural Networks Using Shapley Additive Explanations (SHAP)", BIG DATA COGN. COMPUT., 2022.

[32] Heewoo Jun; Alex Nichol; "Shap-E: Generating Conditional 3D Implicit Functions", ARXIV-CS.CV, 2023.

[33] T. Pianpanit; Sermkiat Lolak; Phattarapong Sawangjai; Thapanun Sudhawiyangkul; Theerawit Wilaiprasitporn; "Parkinson's Disease Recognition Using SPECT Image and Interpretable AI: A Tutorial", IEEE SENSORS JOURNAL, 2021.