JOIV : Int. J. Inform. Visualization, 8(3-2): IT for Global Goals: Building a Sustainable Tomorrow - November 2024 1861-1868

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Predicting and Explaining Customer Response to Upselling in Telecommunications: A Malaysian Case Study

Railey Shahril Abdullah [a], Nur Shaheera Shastera Nulizairos [b], Nor Hapiza Mohd Ariffin [c], Deden Witarsyah [d], Ruhaila Maskat [b,*]

[a] Petronas Digital Sdn Bhd, Kuala Lumpur, 50088, Malaysia
[b] College of Computing, Informatics and Mathematics, Malaysia, Universiti Teknologi MARA, Shah Alam, 40450, Malaysia
[c] MIS Department, Sohar University, 311, Oman
[d] School of Industrial and System Engineering, Telkom University, Bandung, Indonesia
Corresponding author: *ruhaila256@uitm.edu.my

Abstract*— This research explores the predictive capabilities of XGBoost (XGB) and Random Forest (RF) models for customer upsell responses, emphasizing the use of Explainable Artificial Intelligence (XAI) techniques to gain insights. Initially trained without hyperparameter tuning, both models were later optimized using 5-fold cross-validation. While RF consistently achieved high accuracy (0.99), XGB exhibited lower accuracy (0.85) yet demonstrated superior precision and recall. Post-tuning, XGB maintained its competitive edge despite a slight decrease in ROC-AUC scores (0.76 and 0.75 versus RF's 0.67 and 0.72), indicating proficiency in classifying positive cases. XAI techniques complemented XGB's prediction, revealing significant predictors such as inactive duration in days, race (Chinese), total communication count, age, and active period in days. Lesser predictive value was attributed to factors such as race (Indian), gender (female), and region (northern). While the feature importance plot provided a broad overview, it did not detail specific attribute relationships to predictions. To address this, a summary violin plot was employed to illustrate how feature importance varies with actual values, enhancing the understanding of each feature's impact. Results indicated that longer inactivity periods negatively influenced predictions, while non-Chinese ethnicity, higher communication frequency, and younger age were associated with positive outcomes. Dependence plots further elucidated these relationships, highlighting how older non-Chinese customers and those with shorter inactive periods and frequent communication were more likely to accept offers. Local explanations using Shapley's force plot and LIME offered deeper insights into specific instances. Overall, the study underscores the complementary use of XAI techniques to understand a model's predictions.

***Keywords*— Explainable Artificial Intelligence (XAI); predictive analytics; upselling response; Shapley Additive Explanations (SHAP).**

Manuscript received 18 Jun. 2024; revised 22 Aug. 2024; accepted 4 Sep. 2024. Date of publication 30 Nov. 2024.

I. INTRODUCTION

An "upsell" is a sales tactic that aims to convince existing clients to upgrade from a lower-priced service or product to a higher-priced one. This could enhance customer interactions with the company [1], hence increasing sales and adding value. Multiple studies have shown that the act of upselling, which involves upgrading existing clients, is more financially helpful compared to the process of recruiting new consumers. Existing literature indicates that improving customer happiness can lead to an increase in revenue [1], [2], [3]. Historically, semi-manual upselling techniques entailed sales representatives regularly engaging with consumers to explore their preferences, difficulties, and new product options, potentially leading to an upsell [4]. This could potentially yield inaccurate recommendations, hence eroding confidence, diminishing client happiness, and ultimately jeopardizing customer retention. This incentivizes all salespeople and stakeholders to meticulously coordinate their upselling endeavors and strive to pinpoint optimal prospects for upselling in order to secure the sustainability of the telecommunications companies.

Utilizing machine learning algorithms to forecast the reaction of prospective clients can significantly save the time, expense, and exertion involved in addressing those who have little or no interest. Nevertheless, while prediction has demonstrated its advantages, augmenting information to improve customer satisfaction may present certain difficulties. The endeavor is hindered by ethical concerns, privacy issues, and the need for consent. Since the output is critical in many domains, It raises some concerns because it can have some drawbacks that can lead to bias and unfair decisions [5], [6], [7], [8], [9]. Business in general would no longer want to be spoon-fed results; instead, they want to absorb and understand the information offered to them [10]. A good explanation helps the user to gain a thorough understanding of a model, which is required for further improvement or addressing flaws as well as to increase trust among all users on why and how predictions are made [11]. Therefore, the decisions that form the basis of a model's forecast are essential to gain a deeper understanding of the aspects that impact the decision-making process. This is done by using existing data, which can either fully eliminate the need for other data or, at the very least, delay the need of it.

Explainable Artificial Intelligence (XAI) refers to the approaches used to explain predictions. Popular techniques include Feature Importance approaches, Partial Dependence Plots (PDP), Individual Conditional Expectation (ICE) plots, Shapley Additive Explanations (SHAP), and Local Interpretable Model-agnostic Explanations (LIME). The objective of XAI techniques is to provide understandable explanations on how a machine learning model arrives at a prediction. Certain machine learning models are often perceived as black-box, exhibiting a level of complexity that makes the models difficult to comprehend [12]. Black-box ML models incorporate a complex mathematical function or require an in-depth grasp of the distance function. [13]. Some samples of ML that use complex mathematical functions are SVM, XGB, and Neural Network (NN) while the models that use distance functions are KNN and K-Means. Black-box model reasoning is complicated to understand [9], [13], [14], [15] and frequently requires professional knowledge of actual applications. Black-box models that predict without explanation are deemed problematic for a variety of reasons, not just because of their lack of transparency, but also because they conceal potential biases inside the system [16]. This is different compared to white-box models like decision trees and linear regression. White-box models are inherently transparent, offering data scientists explicit justifications for the predicted outcomes [17]. The white-box models' algorithm is more neutral, and the decision structure of the models is straightforward to understand [18]. Therefore, the evaluation of XAI is an expanding field of research that is worth investigating.

Within the scope of this study, the assessment of XAI in several sources of literature has revealed deficiencies, highlighting the importance of this study. Prior research has not examined the use of XAI in Malaysian telecommunication firms specifically for upsell models. There is a scarcity of research in the current literature that examines XAI approaches using data particularly obtained from Malaysian telecommunications companies. It is important to address this gap, as there may be differences in data and results between Malaysia and other locations. Recently, the primary focus of research in the telecommunications field has been on churned customers. This study seeks to shift the current emphasis on churn and instead investigate the use of XAI in the context of upselling, acknowledging the significance of upselling tactics in the telecoms business. Finally, there is a significant opportunity to improve the assessment methods used to explain outputs and assess the processes used to provide these explanations. Existing literature emphasizes the need to improve the current evaluation process by comparing the results produced by XAI methodologies and including experts in the field. In addition, they emphasized deficiencies in the assessment of the techniques employed for constructing explanations, which have not been resolved in the literature.

Therefore, the objective of this work is to determine the most effective black-box model for Malaysian telecommunications data in predicting upsell and utilize model-agnostic XAI approaches (LIME and SHAP) on that model. Then analyze and interpret the global and local explanations generated by both LIME and SHAP techniques, tailored explicitly to the context of upselling practices within the telecommunications industry.

II. Materials and Method

A. Explainable Artificial Intelligence (XAI)

XAI can be categorized into two primary groups: model-specific and model-agnostic. Model-specific approaches exploit the inherent structures of the model being used. This approach entails utilizing the internal mechanics of the model to clarify the process by which the algorithm arrives at judgments. Although model-specific techniques provide a more comprehensive understanding of a model's internal mechanisms and facilitate the development of personalized and interpretable models, their drawback is the need to reconstruct the entire structure of the model, which adversely affects both duration and performance. This implies that any alterations in the structure of the model will require substantial revisions to the process or minor adjustments to the hyperparameters of the algorithm [19].

On the other hand, model-agnostic techniques are specifically designed to be applicable in various circumstances [5]. These techniques are not dependent on the model's structure and can be applied to any machine learning algorithm [20]. They operate by perturbing and modifying the input data, while observing the effect of these modifications on the model's performance compared to the original data. Contrary to model-specific methods, these techniques do not impact the model's performance as they do not necessitate retraining. Notable post-hoc model agnostic techniques employed in XAI for explaining existing models include SHAP and LIME [20], [21], [22], [23].

The explanations are classified into two distinct levels: local explainability and global explainability. Global explainability pertains to understanding the complete functioning of the model or obtaining a comprehensive view of its operations. The explanations provide a comprehensive account of the model's reasoning process during training, which can be helpful for users who are looking to gain insights into the entire dataset [24]. Conversely, local explainability focuses on individual model predictions and provides an explanation of how the model's output changes when specific attributes are varied. Local explanations are usually obtained from models either directly or through closely related local models [22]. Their purpose is to comprehend the reasoning behind the model's application to specific instances and to address specific queries [24]. This study employs a model-agnostic approach.

B. Differences Between LIME and SHAP

Local Interpretable Model Agnostic Explanation (LIME) is used to explain a model based on Local explanation, which means the explanation is crafted close to the observation, and model agnosticism, where it may provide an explanation for any supervised learning model by considering it as a black box separately [25]. To make the word easier to understand, user can think of LIME as an algorithm that any ML model can use to explain its predictions. It does this by making a local approximation of the model around a specific prediction and then explaining how the model behaves in that area. The findings were obtained by imposing a regularization constraint on the linear regression model to access a single input feature that matches a line of linearity. Essentially, LIME attempts to provide a local linear approximation of the model's behavior by developing local surrogate models that are trained to replicate the ML model's predictions locally. This framework was suggested by [26] to address the problem of understanding the reasoning behind a black-box model.

The advantages of LIME are users can still use the same local interpretable model for explanation even if the underlying ML model is replaced. It can provide a human-friendly explanation and is able to interact with various types of data, including tabular, text, and even image files. The primary disadvantage of LIME is that users must experiment with different kernel settings and determine whether the offered explanation makes complete sense [25]. Another issue highlighted by [27] in a study was that LIME provided different explanations after repeating the same process, making it difficult to trust the explanations.

Shapley proposed SHapley Additive exPlanation (SHAP) in 1953 as a method for explaining the outputs of the ML model based on game theory. It provides a contribution value to each feature, indicating its significance in establishing the final prediction. SHAP is introduced as an alternative kernel-based estimation method for Shapley values influenced by the local surrogate model [25]. KernelSHAP is a broad method for calculating SHAP values that can provide more precise and robust explanations. However, it necessitates a large amount of computation power, making it slow to generate[28].

In the regression model, the weighting of instances distinguishes SHAP from LIME. While LIME weights instances based on their proximity to the original instances, SHAP weights sampled instances based on their weight in coalition to estimate Shapley value. SHAP gained popularity in XAI-related research where it has been utilized in a variety of ML model explanation use cases. In recent years, there have been intuitive and interesting representations of several aspects of model explainability for this library

C. Upsell Prediction

There are several studies have been conducted to improve the accuracy or conversion rate for the upselling models [1], [4], [3], [29], [30]. Based on the results of the upsell prediction studies in Table 1, three out of five studies demonstrate that the black-box model is the most accurate model, and there are not many studies that focus on the explanation of the model output. This is one of the gaps detected in upsell prediction, as well as the rationale for why the use case of upsell is selected in this study.

TABLE I
LATEST RESEARCH CONDUCTED ON UPSELL PREDICTION

References	Domain	Focus	Result	XAI
[4]	Telco	Accuracy	LR	No
[30]	Airline	Accuracy	GBM	Yes
[3]	Health	Accuracy	GBM	No
[29]	Telco	Accuracy	SVM	No
[1]	Telco	Accuracy	LR	No

D. Application of XAI in Telecommunication Industries

Duval [31] in his study utilized six XAI libraries, namely PDP, ICE, LIME, SHAP, ALE (Accumulated Local Effects), and GSM (Global Surrogate Model), to analyze a telecommunication customer churn prediction model that employed XGBoost (XGB) and Artificial Neural Network (ANN). The research primarily focused on explainable data sciences, employing a rigorous theoretical approach to showcase the implementation, interpretation, and amalgamation methods. The aim was to achieve a comprehensive understanding of real-world issues. Duval emphasized the importance of employing various methodologies, with survival analysis being particularly noteworthy. The evaluation of libraries encompassed six primary elements: summary statistics of features, visualization capabilities, internal workings of models, inherently interpretable models, text-based explanations, and individual data points. The study also stressed on the significance of focusing on multiple machine learning algorithms and data sources in future research to create more sophisticated explainable artificial intelligence (XAI) systems that can provide user-friendly explanations for different models.

Leung et al. [24] improved the clarity of explanations by utilizing churn data and an RF algorithm. The study implemented a two-tiered strategy, consisting of an abstraction layer that connects predictive models with business operations. The ML model and explanation library were hosted in the back end, while the front end displayed SHAP explanations and graphical information. The RF machine learning model, in conjunction with SHAP, produced four categories of explanations: an overview of the model, global explanations, local explanations, and recommendations. The data included fundamental demographic information such as age, gender, marital status, income, and profession. They determined that their method provided a versatile solution, advocating for the incorporation of interactive visualizations in XAI to improve understanding.

Nkolele and Wang [32] demonstrated the application of machine learning model explanation techniques using churn data. The study employed three models, namely Decision Tree (DT), Random Forest (RF), and Light Gradient Boosting (LGB), along with SHAP and LIME for attribute evaluation and visualization of explanations. The importance of conducting explainability analysis in machine learning for customer churn was emphasized, with a focus on the valuable insights that can be obtained from a decision tree. The need for further research to investigate counterfactual explanations was also emphasized.

Ullah et al. [20] investigated the use of Layer-wise Relevance Propagation (LRP) to improve the clarity of explanations in XAI. Originally developed for computer vision deep models, LRP has been utilized for analyzing categorical and numeric datasets in the context of credit card fraud detection and telecommunication churn data. When comparing LIME and SHAP, it was found that LRP is particularly effective in terms of computational efficiency. The study emphasized the significance of examining the suitability of LRP for structured datasets and its comparability with other techniques such as LORE, MAPLE, and L2X for potential real-world applications.

Meanwhile, this study discovered a dominant pattern in XAI research in the telecommunications industry, where there is a strong emphasis on studying churn as the main case study. Significantly, it is evident that there is a lack of Malaysian research investigating the application of XAI for upsell prediction. This research also highlights the widespread use of XGB and RF as the primary models in XAI implementation, consistently demonstrating superior performance compared to other algorithms. These findings provided the foundation for the methodology used in this study, strengthening the justification for the chosen approaches [20].

This study's methodology encompasses two core layers: prediction and explanation. Figure 1 illustrates the components within each layer and the following subsections discuss them.

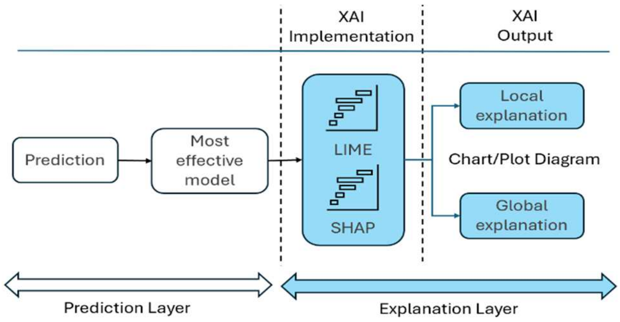


Fig. 1 Methodology details

A. Prediction Layer

The goal of this layer is to generate a set of predictions using the most effective model, which will then be passed on to the explanation layer for interpretation. The prediction pipeline was based on the CRISP-DM approach.

1) Dataset: The dataset consists of 27 attributes and 1,572,503 entries collected from a recent marketing campaign conducted by a telecommunications company in Malaysia. The target audience of this campaign was specifically active consumers who have been living in Malaysia and continue to leverage the telco's services. The target variable is the Last Upsell Campaign Result, which is represented by a value of 1 if the customer accepts the upsell offer and 0 if they do not. The discovery uncovers a notable imbalance in the data, as only 0.4% (6,100 occurrences) of the complete dataset agreed to the proposal. This observation is consistent with previous research that shows a common pattern where three out of five studies using upsell data have consistently shown a low rate of customer acceptance, leading to high imbalance.

2) Data pre-processing: To address the problem of imbalanced data, the technique of random oversampling was employed. This study chose not to use undersampling because the dataset is smaller than the one used by Melidis [3], which could potentially affect the performance of the modelling. This choice was made to prevent the loss of important data and to maximize the accuracy of predicting uptake. The evaluation of dataset correlation necessitated the application of Pearson correlation on the oversampled data. Attributes that showed a correlation higher than 90% were removed to enhance the performance of the model during the modelling phase. The choice of a 90% threshold was made to reduce the impact of multicollinearity, which could otherwise lead to unstable models, inaccurate coefficient estimates, difficulties in interpreting the effects of individual variables, and reduced predictive accuracy. Eliminating variables with strong correlations reduced redundancy in the model. The decision to not adopt lower correlation thresholds was based on the understanding that variables with lower percentages may still have significance for the model. Removing these variables could potentially reduce the predictive accuracy. Considering the relatively small number of attributes examined in this study, eliminating any additional attributes could potentially harm the performance of the model. Furthermore, the dataset was subjected to preprocessing procedures that included eliminating records with missing values and outliers, leading to a decreased dataset size of 1,572,497 entries. Due to high correlation and null values, a total of 7 attributes were removed from the dataset, leaving only 20 attributes.

3) Modelling: At the point of writing, based on the investigation of most recent literature, the statistics (Figure 2) show that RF has the most utilization in the studies, followed by Logistic Regression (LR), Gradient Boosting Machine (GBM), Support Vector Machine (SVM), and XGB. Since the goal of this study is to apply XAI techniques to the most effective black-box model, the Linear Regression model cannot be used in this experiment because LR is a white-box model.

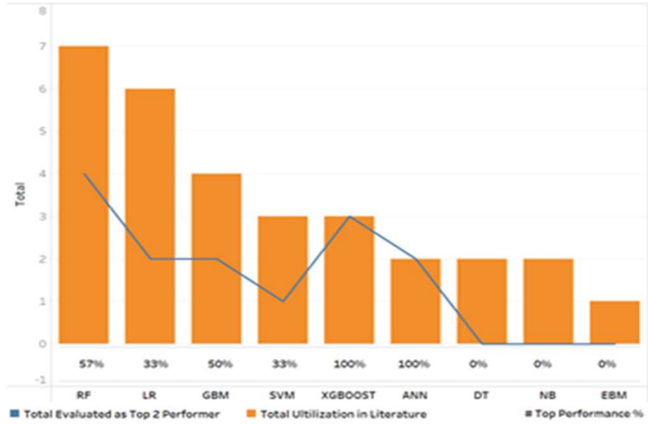


Fig. 2 Performance of Models from Literature

XGB exhibits superior performance, as it ranks among the top two performers in all three use cases, with ANN and RF following suit. RF and XGB were selected for this study based on their results, as both algorithms offer a favorable balance between utilization and performance. To enhance the model's performance, the hyperparameters were fine-tuned by employing a 5-fold cross-validation approach to identify the optimal hyperparameters for both models.

4) Performance measurement: During this stage, the performance of the model was evaluated using metrics such as F1 score, specificity, sensitivity, and Receiver Operating Characteristic (ROC) score. Considering the unequal distribution of the data and the main goal of maximizing accurate positive predictions, the ROC AUC (Area Under the Curve) was chosen as the primary evaluation metric. The ROC curve is a widely used metric that is especially useful for evaluating datasets with imbalanced classes. It provides a concise summary of the performance of a classifier by considering different rates of true positive and false positive results. Additionally, it assists in determining the most effective threshold for maximizing outcomes [33]. The AUC, which provides a single value for comparing multiple models, varies from zero to one. A value less than 0.5 signifies an impractical categorization [34]. The model with the highest ROC AUC will be combined with XAI libraries to achieve the goals of the project.

B. Explanation Layer

The optimal model from the modelling layer is subsequently integrated with XAI libraries to achieve the primary goal of this project: applying model-agnostic XAI techniques (LIME and SHAP) to the best-performing black-box model. Initially, all necessary objects for calculating explanation values were created in Jupyter Notebook using the LIME and SHAP libraries. These objects serve as the foundation for generating various explanation mediums, such as diagrams and tables. The SHAP values computed during the XAI application were used to produce global explanation plots using the SHAP library.

From the test datasets, observations were selected for local explanations using SHAP, generating visualization plots. These same observations were then used as a basis for constructing local explanations with LIME. This approach ensures a valid comparison by contrasting similar explanatory instances. The data was thoroughly analysed. The global and local explanations were discussed in detail, providing a comprehensive understanding of how the attributes influence the prediction outcomes.

III. Results and Discussion

A. Most Effective Model Chosen

For this study, XGB and RF models were first trained without adjusting hyperparameters in order to establish the initial performance level. Afterwards, a 5-fold cross-validation was conducted using the GridSearchCV library in Python to determine the best hyperparameter values for each model, with the goal of improving their performance. The optimal hyperparameters obtained are presented in Table 2, and the models were subsequently retrained based on these hyperparameters.

TABLE II
Optimal Hyperparameter

Algorithm	n_estimator	max depth	learning_rate	min_sample_split
XGB	50	8	0.3	N/A
RF	10	8	N/A	2

According to the analysis in Table 3, the RF model achieves the highest accuracy at 0.99, both with and without hyperparameter tuning, followed by both the tuned and untuned XGB models, each with an accuracy of 0.85. However, precision and recall metrics reveal that XGB outperforms RF, showing higher precision (0.012) and recall (0.53 for untuned, 0.46 for tuned) compared to RF, which has a precision of 0.0 for the untuned model and slightly improved values with tuning (0.02 and 0.01). This discrepancy is reflected in the ROC-AUC scores, where XGB models have higher values (0.76 and 0.75) compared to RF (0.67 and 0.72), indicating superior performance in classifying positive cases.

TABLE III
Model Performance

	RF(Normal)	XGB(Normal)	RF(Tuned)	XGB(Tuned)
Accuracy	0.99	0.83	0.99	0.85
Precision	0.0	0.012	0.02	0.012
Recall	0.0	0.53	0.01	0.46
F1 Score	0.01	0.023	0.02	0.24
ROC-AUC	0.67	0.76	0.72	0.75

Nevertheless, the ROC-AUC score for XGB decreases by 0.1 after hyperparameter tuning, implying that while the model makes more reliable overall predictions, it may increase the number of false negatives. Achieving optimal model performance necessitates balancing accuracy and recall, as improving one metric often adversely affects the other. This balance is crucial depending on the specific task, especially when identifying true positives is more important than reducing false positives.

Further analysis using confusion matrices confirms these findings (Figure 3). The untuned RF model fails to predict any true positives, whereas the untuned XGB correctly predicts 969 out of 1,830 positive cases. Hyperparameter tuning slightly improves RF's performance, predicting two true positives. Conversely, the tuned XGB model predicts fewer true positives (839, down from 969). Despite this reduction, the tuned XGB's precision increases, predicting that 68,501 customers would accept the offer with a true positive rate of 1.22%, compared to the untuned XGB's 80,768 predictions with a 1.2% true positive rate. These results indicate that the tuned XGB model is optimal for predicting upsell propensity, effectively balancing precision and recall identifying actionable leads, which is critical for campaign targeting and minimizing effort and revenue loss.

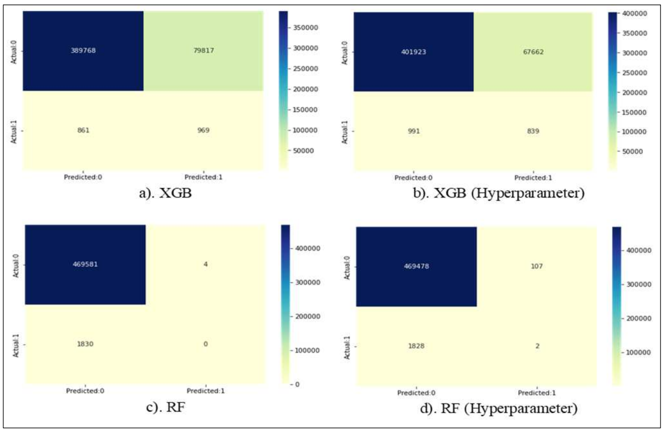


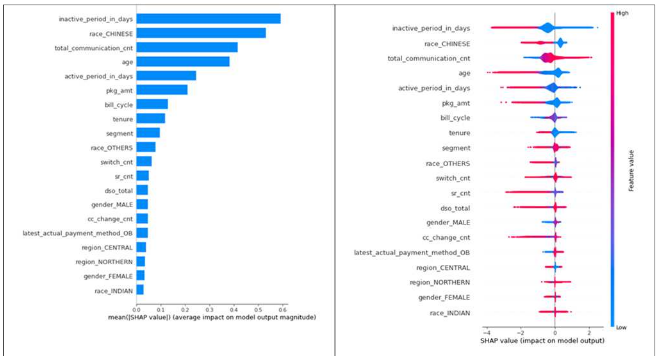
Fig. 3 Confusion Matrix for All Models

B. Explaining Predictions

SHAP and LIME were utilized to elucidate the predictions of the black-box model. Both are model-agnostic techniques, with SHAP providing explanations at both global and local levels, while LIME offers explanations only at the local level.

1) Output of SHAP Global Explanation: Global explanations were generated using a feature importance plot to provide an overarching understanding of how the XGB model makes predictions without focusing on specific inputs or outputs. Illustrated in Figure 4 (a), this plot highlights the significance of various attributes based on their mean Shapley values, identifying inactive duration in days as the most influential predictor, followed by race (Chinese), total communication count, age, and active period in days. Attributes such as race (Indian), gender (female), and region (northern) were found to have lesser predictive value. However, while this plot offers a general overview, it does not provide detailed insights into the specific relationships between these attributes and the predicted outcomes.

To better understand how feature importance varies with actual values, a summary violin plot from the SHAP library was created, as shown in Figure 4 (b). This plot combines feature importance and feature effects, with each point representing a Shapley value for a specific instance and feature. The y-axis position corresponds to the feature, while the color gradient indicates the feature's value from high to low. This plot provides a more detailed view of how each feature's importance changes across different values, enabling a deeper analysis of their impact on predictions. It allows users to visualize the distribution of Shapley values for each feature, thus offering insights into how the model's decisions vary based on the specific characteristics of each instance.

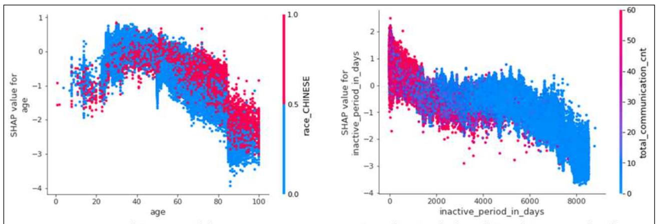


a). Feature Importance Plot

b). Violin Summary Plot

Fig. 4 SHAP Global Explanation

The results indicate that the duration of inactivity significantly influences predictions, with longer inactive periods leading to more negative scores and shorter periods resulting in more positive scores. This correlation is expected, as active customers are more likely to respond positively to offers. Non-Chinese customers also showed a higher likelihood of accepting offers, and increased communication with the service provider positively affected prediction results. Additionally, younger customers were more likely to accept offers, contrary to initial expectations regarding the active period in days, which negatively affected predictions. Dependence plots from the SHAP library (Figure 5) were also reviewed to further understand model behavior, revealing relationships between different attributes. For example, older non-Chinese customers had a more pronounced impact on predicting non-take-up compared to Chinese customers of similar age, and customers with shorter inactive periods and higher communication frequency were more likely to accept offers. These insights highlight the importance of considering various attributes and their interrelationships when interpreting model predictions.



Age and Race = Chinese

Inactive Period and Total Communication

Fig. 5 SHAP Global Explanation (Dependence Plot)

2) Output of SHAP Local Explanation: In XAI, local explanation entails elucidating the reasoning behind a model's decision for a specific instance. Shapley's force plot serves as a tool to explicate the underlying logic behind predictions for two chosen instances, namely the 12th and 58th observations, as depicted in Figures 6 and 7, respectively. These instances have contrasting predictions. In the case of Observation 12, the model forecasts a minimal probability of -3.51 for accepting the upsell offer. This subdued likelihood primarily stems from this particular customer's Chinese ethnicity and advanced age, nearing 60 years, which corresponds to the broader explanation provided earlier at the global level. Furthermore, the total amount of subscription (pkg_amt) of this customer is high thus adversely favoring any upsell. This confirms our earlier global explanation visualized by the violin plot in Figure 4 (b).

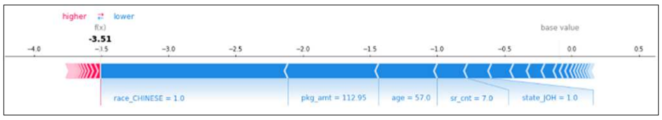


Fig. 6 SHAP Local Explanation (Force Plot for 12th Observation)

On the contrary, in the 58th observation (depicted in Figure 7), there is a notably higher predicted probability, standing at 1.86, suggesting a greater likelihood of accepting an upsell offer. This inclination predominantly stems from the relatively brief period of inactivity (inactive_period_in_days) observed for this customer, indicating an active engagement. Furthermore, this customer is newly onboarded with a tenure of only one year with the company, hence, exhibited a propensity to engage with upsell offers. This aligns with the global explanation of the model (refer Figure 4 (b)). Observation 58 indicates having Indian ethnicity, yet this very feature negatively contributes to the acceptance of upsells. Evidently from this observation, the XGB model has discerned a trend wherein customers not of Chinese descent generally display a higher tendency to accept upsell offers, potentially influenced by historical data indicating elevated acceptance rates among this demographic. However, within this non-Chinese cohort, the model has detected a lower inclination among Indian customers to accept such offers compared to their non-Chinese counterparts (e.g., Malay, others), which could be ascribed to cultural inclinations, past consumer behavior, or other factors specific to the Indian demographic. Interestingly, both observations indicate the lack of enthusiasm among residents of the Johor state towards accepting any upsell.

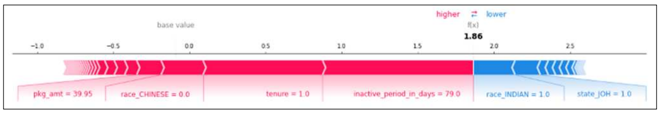


Fig. 7 SHAP Local Explanation (Force Plot for 58th Observation)

3) LIME Global Explanation: LIME is not recommended for providing a global explanation as it's designed to be local, and its output might be misleading or inconsistent. Technically, it is possible to use LIME to provide global explanations, but it would require generating explanations for many instances and then aggregating the results. This would be a time-consuming process and may not provide a clear overall picture of the model's behavior.

4) Output of LIME Local Explanation: The LIME technique provides additional local-level explanations when compared to SHAP. Figure 8 displays the results for the 12th observation, while Figure 9 shows the results for the 58th observation. LIME's output for the 58th observation reveals a high probability of 0.97 for not accepting the upsell and a low probability of 0.03 for accepting it. In contrast to SHAP, LIME's prediction for this customer was greatly influenced by the recorded long period of inactivity. This maps more closely to the earlier global explanation. Furthermore, LIME shows XGB's ability to learn from the training set that resident from the central and southern regions of Malaysia, particularly Johor and Selangor, are not inclined to the offered upsell, however, people in the Northern region states does. Another notable result is this upsell is generally more appealing to male customers instead of female customers, although with this specific customer the gender neither pushes the prediction towards nor away from the predicted class.

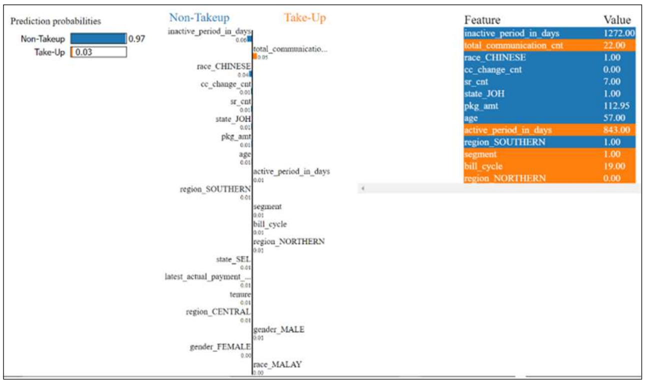


Fig. 8 LIME Local Explanation (Details for 12th Observation)

Observation 58, as depicted in Figure 9, has a high likelihood of accepting the upsell, with a probability score of 0.87. The primary factor contributing to this prediction is the customer's short inactivity period. The second most significant factor is race, where not being ethnically Chinese leads the model to predict a positive response to the offer. Other influential features include the total number of credit card changes (cc_change_cnt), the method of the last payment (last_actual_payment_mtd), the account subtype (acc_sub_type), the customer's credit risk (credit_class), and the combined attributes of gender, region, and state, specifically states located in northern Malaysia. Although Observation 58 does not reside in these northern regions, the XGB model still utilizes these features to forecast the outcome.

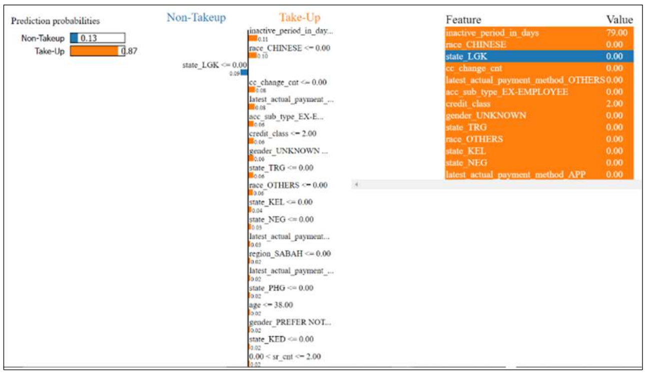


Fig. 9 LIME Local Explanation (Details for 58th Observation)

IV. Conclusion

In conclusion, the study utilized global explanations with a feature importance plot to elucidate the predictive mechanisms of the XGB model. This analysis revealed that the most influential predictors included inactive duration in days, race (Chinese), total communication count, age, and active period in days. Lesser predictive value was attributed to factors such as race (Indian), gender (female), and region (northern). While the feature importance plot provided a broad overview, it did not detail specific attribute relationships to predictions. To address this, a summary violin plot was employed to illustrate how feature importance varies with actual values, enhancing the understanding of each feature's impact.

Results indicated that longer inactivity periods negatively influenced predictions, while non-Chinese ethnicity, higher communication frequency, and younger age were associated with positive outcomes. Dependence plots further elucidated these relationships, highlighting how older non-Chinese customers and those with shorter inactive periods and frequent communication were more likely to accept offers. Local explanations using Shapley's force plot and LIME offered deeper insights into specific instances. For Observation 12, advanced age and high subscription amount were primary factors for a negative prediction, whereas Observation 58 showed a high likelihood of acceptance due to a short inactivity period and non-Chinese ethnicity. LIME results corroborated these findings, emphasizing regional and demographic influences on predictions. Overall, the study underscores the importance of considering multiple attributes and their interrelationships in model interpretations, providing a comprehensive understanding of how the XGB model makes predictions.

Acknowledgment

The authors would like to acknowledge the support by the Universiti Teknologi MARA Shah Alam, Selangor, Malaysia and Sohar University, Oman.

References

[1] S. M. A. M. Manchanayake et al., "Potential Upselling Customer Prediction Through User Behavior Analysis Based on CDR Data," 2019 14th Conference on Industrial and Information Systems (ICIIS), pp. 46–51, Dec. 2019, doi: 10.1109/iciis47346.2019.9063278.

[2] B. Denizci Guillet, "Online upselling: Moving beyond offline upselling in the hotel industry," International Journal of Hospitality Management, vol. 84, p. 102322, Jan. 2020, doi:10.1016/j.ijhm.2019.102322.

[3] A. Melidis, "Personalized marketing campaign for upselling using predictive modeling in the health insurance sector," 2020, [Online]. Available: https://run.unl.pt/handle/10362/99076

[4] N. Dookeram, Z. Hosein, and P. Hosein, "A Recommender System for the Upselling of Telecommunications Products," 2022 24th International Conference on Advanced Communication Technology (ICACT), pp. 66–72, Feb. 2022, doi:10.23919/icact53585.2022.9728818.

[5] V. Belle and I. Papantonis, "Principles and Practice of Explainable Machine Learning," Frontiers in Big Data, vol. 4, Jul. 2021, doi:10.3389/fdata.2021.688969.

[6] B. Dimanov, U. Bhatt, M. Jamnik, and A. Weller, "You shouldnat trust me: Learning models which conceal unfairness from multiple explanation methods," *Frontiers in Artificial Intelligence and Applications*, vol. 325, no. 2019, pp. 2473–2480, 2020, doi:10.3233/FAIA200380.

[7] S. Mohseni, N. Zarei, and E. D. Ragan, "A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems," ACM Transactions on Interactive Intelligent Systems, vol. 11, no. 3–4, pp. 1–45, Sep. 2021, doi: 10.1145/3387166.

[8] M. Nazar, M. M. Alam, E. Yafi, and M. M. Su'ud, "A Systematic Review of Human–Computer Interaction and Explainable Artificial Intelligence in Healthcare With Artificial Intelligence Techniques," IEEE Access, vol. 9, pp. 153316–153348, 2021, doi:10.1109/access.2021.3127881.

[9] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," Nature Machine Intelligence, vol. 1, no. 5, pp. 206–215, May 2019, doi:10.1038/s42256-019-0048-x.

[10] N. Burkart and M. F. Huber, "A Survey on the Explainability of Supervised Machine Learning," Journal of Artificial Intelligence Research, vol. 70, pp. 245–317, Jan. 2021, doi: 10.1613/jair.1.12228.

[11] R. Confalonieri, L. Coba, B. Wagner, and T. R. Besold, "A historical perspective of explainable Artificial Intelligence," WIREs Data Mining and Knowledge Discovery, vol. 11, no. 1, Oct. 2020, doi:10.1002/widm.1391.

[12] A. Barredo Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," Information Fusion, vol. 58, pp. 82–115, Jun. 2020, doi: 10.1016/j.inffus.2019.12.012.

[13] O. Loyola-Gonzalez, "Black-Box vs. White-Box: Understanding Their Advantages and Weaknesses From a Practical Point of View," IEEE Access, vol. 7, pp. 154096–154113, 2019, doi:10.1109/access.2019.2949286.

[14] S. Lockey, N. Gillespie, D. Holm, and I. A. Someh, "A Review of Trust in Artificial Intelligence: Challenges, Vulnerabilities and Future Directions," Proceedings of the 54th Hawaii International Conference on System Sciences, 2021, doi: 10.24251/hicss.2021.664.

[15] C. C. Yang, "Explainable Artificial Intelligence for Predictive Modeling in Healthcare," Journal of Healthcare Informatics Research, vol. 6, no. 2, pp. 228–239, Feb. 2022, doi: 10.1007/s41666-022-00114-1.

[16] A. M. Antoniadi et al., "Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical Decision Support Systems: A Systematic Review," Applied Sciences, vol. 11, no. 11, p. 5088, May 2021, doi: 10.3390/app11115088.

[17] P. L. Fung et al., "Evaluation of white-box versus black-box machine learning models in estimating ambient black carbon concentration," Journal of Aerosol Science, vol. 152, p. 105694, Feb. 2021, doi:10.1016/j.jaerosci.2020.105694.

[18] T. Rieg, J. Frick, H. Baumgartl, and R. Buettner, "Demonstration of the potential of white-box machine learning approaches to gain insights from cardiovascular disease electrocardiograms," PLOS ONE, vol. 15, no. 12, p. e0243615, Dec. 2020, doi:10.1371/journal.pone.0243615.

[19] A. Das and P. Rad, "Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey," pp. 1–24, 2020, [Online]. Available: http://arxiv.org/abs/2006.11371

[20] I. Ullah, A. Rios, V. Gala, and S. Mckeever, "Explaining Deep Learning Models for Tabular Data Using Layer-Wise Relevance Propagation," Applied Sciences, vol. 12, no. 1, p. 136, Dec. 2021, doi:10.3390/app12010136.

[21] J. Duell, X. Fan, B. Burnett, G. Aarts, and S.-M. Zhou, "A Comparison of Explanations Given by Explainable Artificial Intelligence Methods on Analysing Electronic Health Records," 2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), pp. 1–4, Jul. 2021, doi: 10.1109/bhi50953.2021.9508618.

[22] G. Plumb, D. Molitor, and A. Talwalkar, "Model agnostic supervised local explanations," *Adv Neural Inf Process Syst*, vol. 2018-Decem, no. NeurIPS, pp. 2515–2524, 2018.

[23] D. Slack, S. Hilgard, E. Jia, S. Singh, and H. Lakkaraju, "Fooling LIME and SHAP," Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, pp. 180–186, Feb. 2020, doi:10.1145/3375627.3375830.

[24] C. K. Leung, A. G. M. Pazdor, and J. Souza, "Explainable Artificial Intelligence for Data Science on Customer Churn," 2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA), pp. 1–10, Oct. 2021, doi: 10.1109/dsaa53316.2021.9564166.

[25] C. Molnar, *Interpretable Machine Learning*. 2022. [Online]. Available: https://christophm.github.io/interpretable-ml-book

[26] M. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, 2016, doi:10.18653/v1/n16-3020.

[27] D. Alvarez-Melis and T. S. Jaakkola, "On the Robustness of Interpretability Methods," no. Whi, 2018, [Online]. Available: http://arxiv.org/abs/1806.08049

[28] P. S. R. Aditya and M. Pal, "Local Interpretable Model Agnostic Shap Explanations for machine learning models," no. c, 2022.

[29] L.-Y. Zhou, D. M. Amoh, L. K. Boateng, and A. A. Okine, "Combined Appetency and Upselling Prediction Scheme in Telecommunication Sector Using Support Vector Machines," International Journal of Modern Education and Computer Science, vol. 11, no. 6, pp. 1–7, Jun. 2019, doi: 10.5815/ijmecs.2019.06.01.

[30] N. A. Emadi, S. Thirumuruganathan, D. R. Robillos, and B. J. Jansen, "Will You Buy It Now?: Predicting Passengers that Purchase Premium Promotions Using the PAX Model," Journal of Smart Tourism, vol. 1, no. 1, pp. 53–64, Mar. 2021, doi: 10.52255/smarttourism.2021.1.1.7.

[31] A. Duval, "Explainable Artificial Intelligence (XAI) Explainable Artificial Intelligence (XAI) by Alexandre Duval MA4K9 Scholarly Report Submitted to The University of Warwick Mathematics Institute," no. April, p. 58, 2019, doi: 10.13140/RG.2.2.24722.09929.

[32] R. Nkolele and H. Wang, "Explainable Machine Learning: A Manuscript on the Customer Churn in the Telecommunications Industry," 2021 Ethics and Explainability for Responsible Data Science (EE-RDS), pp. 1–7, Oct. 2021, doi: 10.1109/ee-rds53766.2021.9708561.

[33] K. Roshan and A. Zafar, "Utilizing XAI Technique to Improve Autoencoder based Model for Computer Network Anomaly Detection with Shapley Additive Explanation(SHAP)," International journal of Computer Networks & Communications, vol. 13, no. 6, pp. 109–128, Sep. 2021, doi: 10.5121/ijcnc.2021.13607.

[34] A. Tharwat, "Classification assessment methods," Applied Computing and Informatics, vol. 17, no. 1, pp. 168–192, Jul. 2020, doi:10.1016/j.aci.2018.08.003.