Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.3, June 2026, pp. 2740~2748
e-ISSN: 2548-6861

Analysis of Customer Churn Classification for Sinarmas Syariah Lhokseumawe Insurance Services Using Deep Learning Tabnet and Explainable AI

Syarifah Muliana [1]*, Taufiq [2]*, Munirul Ula [3]*, Muhammad Sayuti [4]*, Nurdin [5]*
*Department of Information Technology, Universitas Malikussaleh, Lhokseumawe, Indonesia
syarifah.247110201013@mhs.unimal.ac.id [1], taufiq.te@unimal.ac.id [2], munirul.ula@unimal.ac.id [3], sayuti_m@unimal.ac.id [4], nurdin@unimal.ac.id [5]

Article Info

Article history:

Received 2026-04-26
Revised 2026-06-02
Accepted 2026-06-11

Keyword:

Deep Learning,
Tabnet, Churn,
Classification,
XAI,
SHAP.

ABSTRACT

Customer churn is a major challenge in the insurance industry because it directly affects customer retention, business sustainability, and company profitability. Early identification of customers at risk of churn is therefore essential for developing effective retention strategies. This study proposes an interpretable customer churn prediction framework for Sinarmas Syariah Lhokseumawe Insurance Services by integrating TabNet deep learning with Shapley Additive Explanations (SHAP). The dataset consists of 2,000 customer records containing demographic information, insurance transactions, premium payments, claims history, and customer interaction data. Due to the imbalanced class distribution, the Synthetic Minority Oversampling Technique (SMOTE) was applied exclusively to the training dataset to improve model learning while preventing data leakage. Model performance was evaluated using Accuracy, Precision, Recall, F1-Score, and ROC-AUC metrics. The experimental results demonstrate that the proposed approach achieved an accuracy of 98.77%, precision of 86.67%, recall of 92.86%, F1-score of 89.66%, and ROC-AUC of 0.995, indicating excellent classification performance. Furthermore, SHAP analysis revealed that communication, premi_2025, and reason_to_purchase were the most influential features affecting churn predictions. These findings highlight the importance of customer engagement, premium management, and purchasing motivations in customer retention. The proposed TabNet-SHAP framework provides both high predictive performance and model interpretability, making it a valuable decision-support tool for customer retention strategies in the insurance sector.

I. Introduction

Customer retention has become one of the most important challenges in the service industry, particularly in the insurance sector, where customer loyalty directly affects business sustainability and company profitability. Customer churn, which refers to customers discontinuing the use of insurance services or policies, can significantly reduce company revenue and increase operational costs due to the need to acquire new customers continuously. Therefore, the ability to identify potential churn customers at an early stage is essential for supporting effective customer retention strategies and improving long-term business performance [1], [2].

In the insurance industry, customer churn may be influenced by various factors, including service quality, communication effectiveness, premium costs, claim experiences, and customer satisfaction levels [3], [4]. In the context of sharia insurance, customer retention is even more critical because the relationship between customers and insurance providers is not solely based on financial transactions, but also on trust, transparency, and compliance with Islamic principles [5]. Based on preliminary information obtained from the operational division of Sinarmas Syariah Lhokseumawe Insurance Services, the annual customer churn rate is estimated to range between 15% and 20%, particularly in vehicle insurance products. This condition indicates the necessity of implementing a predictive system capable of detecting customers who are likely to discontinue their insurance services [6].

Along with the rapid advancement of information technology, machine learning and deep learning techniques have been widely applied to analyze customer behavior and predict churn patterns in various sectors, such as banking, telecommunications, and e-commerce [7], [8]. Previous studies have demonstrated that predictive models can assist companies in improving customer retention by identifying customers with high churn probability. Several machine learning approaches commonly used for churn prediction include Logistic Regression, Random Forest, Support Vector Machine, and XGBoost [9], [10], [11]. However, most existing studies mainly focus on improving predictive accuracy without providing sufficient explanation regarding the factors influencing the prediction results.

The lack of interpretability in predictive models creates challenges for decision-makers, especially in the financial and insurance sectors, where understanding the reasons behind model predictions is highly important [12]. To address this issue, Explainable Artificial Intelligence (XAI) has emerged as an important approach for improving transparency and interpretability in machine learning models [13], [14]. One of the most widely used XAI methods is SHAP (Shapley Additive Explanations), which is capable of explaining the contribution of each feature to the prediction results at both global and local levels [15], [16]. Through SHAP analysis, companies can better understand the dominant factors contributing to customer churn and use these insights to support strategic decision-making.

This study proposes the use of TabNet, a deep learning architecture specifically designed for tabular data classification tasks. Unlike conventional deep learning models, TabNet utilizes a sequential attention mechanism that performs automatic feature selection during the learning process. This mechanism enables the model to achieve strong predictive performance while maintaining interpretability. Compared to traditional machine learning methods such as Random Forest and XGBoost, TabNet offers the advantage of end-to-end learning and adaptive feature representation. Although TabNet has shown promising performance in several classification studies, its implementation in customer churn prediction within the sharia insurance sector remains limited [17].

In addition, research combining TabNet with SHAP-based Explainable Artificial Intelligence for customer churn analysis in sharia insurance services has rarely been explored in previous studies. Most churn prediction research in the insurance sector still emphasizes prediction performance rather than interpretability and practical business implementation. Therefore, there remains a research gap regarding the development of an interpretable deep learning framework capable of supporting customer retention strategies in sharia insurance companies [18].

Based on these considerations, this study aims to develop an interpretable customer churn classification model for Sinarmas Syariah Lhokseumawe Insurance Services using TabNet deep learning integrated with SHAP-based Explainable Artificial Intelligence. The proposed approach is expected not only to achieve high classification performance but also to provide meaningful insights into the factors influencing customer churn behavior. Furthermore, the findings of this study are expected to support insurance companies in designing more effective, data-driven, and targeted customer retention strategies.

II. METHODOLOGY

A. Research Stage

The research workflow was designed systematically to ensure that the model development process could be conducted in a structured and reproducible manner. The overall stages of this study are illustrated in Figure 1.

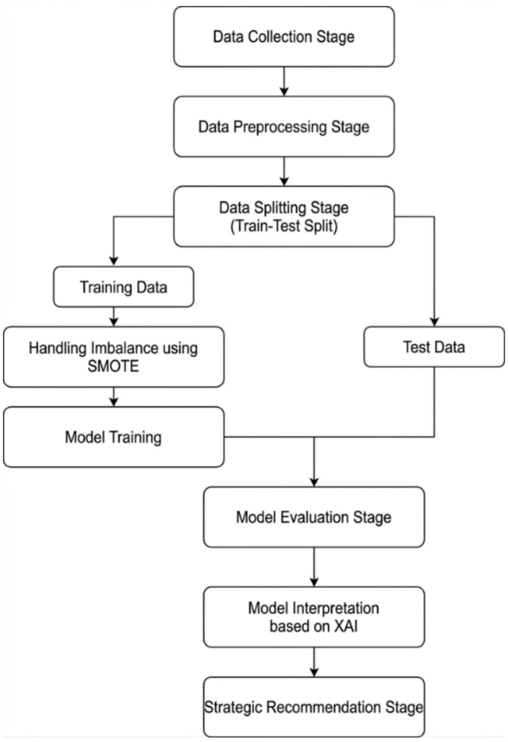

Figure 1. Research stages process

Figure 1 illustrates the overall research framework used in this study. The process begins with the data collection stage, where customer data are gathered from the insurance database, followed by data preprocessing to clean and prepare the data for analysis. Subsequently, the dataset is divided into training and testing sets using a train-test split approach. To address the class imbalance problem, the Synthetic Minority Oversampling Technique (SMOTE) is applied exclusively to the training data before model training. The trained model is then evaluated using the testing data to assess its predictive performance. Furthermore, Explainable Artificial Intelligence

(XAI) techniques are employed to interpret the model's predictions and identify the most influential factors contributing to customer churn. Finally, the insights obtained from the interpretation stage are used to formulate strategic recommendations for improving customer retention and reducing churn rates.

B. Dataset Characteristics

The dataset used in this study consists of 2,000 historical customer records obtained from Sinarmas Syariah Lhokseumawe Insurance Services. Each record represents an individual customer and contains information related to customer profiles, insurance transactions, premium payments, claims history, and customer interactions. The target variable in this study is customer churn status, which is categorized into churn and non-churn customers. Table 1 presents the distribution of customer churn status in the dataset.

TABLE I
DISTRIBUTION OF CUSTOMER CHURN STATUS

Class	**Number Of Customers**	**Percentage (%)**
Churn	128	6.4
Non-Churn	1872	93.6
Total	2.000	100.0

Table I shows that the dataset is highly imbalanced, with only 128 customers (6.4%) classified as churn customers and 1,872 customers (93.6%) classified as non-churn customers. This distribution indicates that the majority class significantly dominates the minority class, which may cause the classification model to be biased toward the majority class and reduce its ability to accurately identify churn customers. To address this issue, imbalance handling techniques were applied during the model development process. Furthermore, the dataset was divided into training and testing sets prior to the oversampling stage. The Synthetic Minority Oversampling Technique (SMOTE) was applied exclusively to the training data, while the testing data remained unchanged. This strategy was implemented to prevent data leakage, which could artificially inflate model performance and produce overly optimistic evaluation results. By ensuring that information from the testing dataset was not exposed during model training, the evaluation process provides a more realistic assessment of the model's generalization capability.

In addition, this study utilized customer data obtained from internal company records. To protect customer privacy, all personally identifiable information (PII) was anonymized before the analysis process. The research only used attributes relevant to customer churn prediction and excluded sensitive personal identifiers such as customer names, addresses, telephone numbers, and other confidential information. All data processing activities were conducted solely for academic research purposes while maintaining data confidentiality and privacy principles. Furthermore, future implementation of predictive analytics systems in the insurance sector should consider stronger data governance, access control, and security frameworks to ensure compliance with customer data protection regulations and industry best practices.

C. Data Preprocessing

Data preprocessing was conducted to ensure that the dataset was suitable for the modeling process and to reduce potential bias caused by inconsistent data [19], [20]. The initial stage involved identifying missing values, incomplete records, and duplicate data entries. Records that were considered irrelevant or inconsistent were removed from the dataset.

Categorical variables were subsequently transformed into numerical representations using label encoding [21]. This transformation was required because the TabNet model processes numerical input data in tabular format.

In addition, feature normalization was applied to reduce differences in feature scales and to support a more stable training process. After preprocessing, the dataset was re-evaluated to ensure that all variables were consistent and ready for the classification stage.

D. Handling Imbalanced Data Using SMOTE

The customer churn dataset used in this study exhibited class imbalance, where the number of non-churn customers was substantially higher than churn customers. Such conditions may affect classification performance and cause the model to be biased toward the majority class.

To address this issue, the Synthetic Minority Oversampling Technique (SMOTE) was applied to the training dataset. SMOTE works by generating synthetic samples from the minority class to produce a more balanced class distribution.

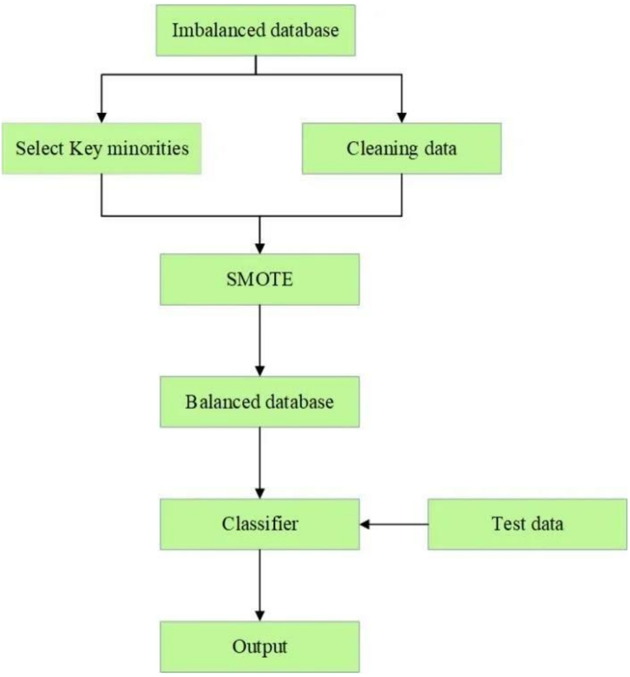

Figure 2. Imbalance Class Handler

Figure 2 illustrates the workflow of handling imbalanced data using the Synthetic Minority Oversampling Technique (SMOTE). The process begins with an imbalanced dataset, where the minority class is significantly underrepresented compared to the majority class. Before applying SMOTE, the data undergoes a cleaning process to remove inconsistencies and improve data quality, while minority class instances are identified as the target for oversampling. SMOTE then generates synthetic samples based on the characteristics of existing minority-class data, resulting in a more balanced dataset. The balanced data are subsequently used to train the classification model, whereas the test data remain unchanged to ensure an objective evaluation. Finally, the trained classifier produces prediction outputs that can be used to identify customer churn behavior more accurately.

E. Deep Learning (TabNet)

This study employs TabNet as the primary classification model for customer churn prediction. TabNet is a deep learning architecture specifically designed for tabular data and utilizes a sequential attention mechanism to perform automatic feature selection. Unlike conventional neural networks, TabNet selects relevant features at each decision step, allowing the model to focus on the most informative attributes during the learning process. This approach not only improves predictive performance but also enhances model interpretability.

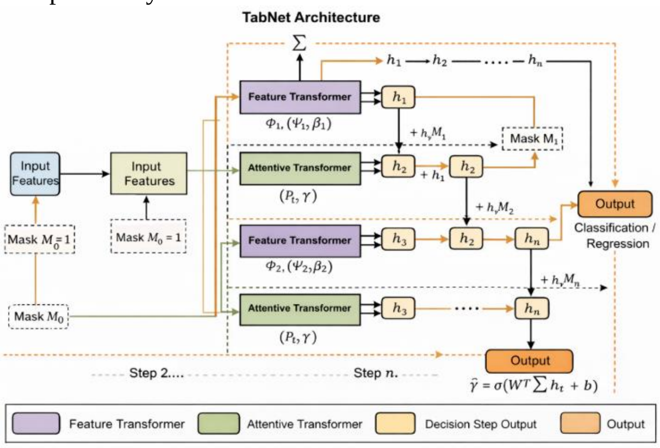

Figure 3. TabNet Architecture Diagram

The TabNet architecture consists of two main components: the Feature Transformer and the Attentive Transformer. The Feature Transformer is responsible for extracting meaningful feature representations from the input data, while the Attentive Transformer determines feature importance through an attention-based masking mechanism at each decision step.

F. Hyperparameter Configuration

The TabNet model was configured using several key hyperparameters that control the network architecture, feature selection process, and learning behavior during training. These hyperparameters were determined experimentally through multiple trials to achieve a balance between predictive performance and model generalization. Table 2 summarizes the hyperparameter settings used in this study.

TABLE II
TABNET HYPERPARAMETER CONFIGURATION

Hyperparameter	**Value**
n_d	16
n_a	16
n_steps	5
gamma	1.5
lambda_sparse	1×10^{-4}
Learning Rate	0.02
Optimizer	Adam

As shown in Table II, the values of *n_d* and *n_a* were both set to 16 to provide a balanced representation capacity for the decision and attention prediction layers. The number of decision steps (*n_steps*) was set to 5, allowing the model to perform sequential feature selection while maintaining computational efficiency. The *gamma* parameter was set to 1.5 to encourage the model to explore diverse features across different decision steps, whereas *lambda_sparse* was set to 1×10^{-4} to promote sparse feature selection and improve model interpretability.

Furthermore, the model was trained using the Adam optimizer with a learning rate of 0.02, which provided stable convergence during the optimization process. To prevent overfitting and improve generalization performance, an early stopping mechanism was implemented during training. Training was automatically terminated when no significant improvement was observed in the validation performance over a predefined number of epochs.

G. Model Evaluation

The performance of the proposed TabNet model was evaluated using four widely adopted classification metrics, namely Accuracy, Precision, Recall, and F1-Score. These metrics were selected to provide a comprehensive assessment of the model's predictive capability in distinguishing between churn and non-churn customers. Furthermore, the use of multiple evaluation metrics is particularly important when dealing with imbalanced datasets, as relying solely on accuracy may not adequately reflect the model's performance on the minority class.

1. The formula for calculating the accuracy equation for classifying customer churn at SinarMas Lhokseumawe Syariah insurance is as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

2. The formula for calculating the precision equation for customer churn classification for SinarMas Lhokseumawe Syariah insurance is as follows:

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

3. Recall measures the model's ability to effectively detect churn data. The recall calculation formula for the churn classification of SinarMas Syariah Lhokseumawe insurance customers is as follows:

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

4. F1-Score is an evaluation metric used to measure the balance between precision and recall. The F1-Score is calculated as the harmonic mean of precision and recall, providing a more comprehensive picture of model performance, particularly when handling imbalanced data. The formula for calculating the F1-Score for customer churn classification in the SinarMas Lhokseumawe Syariah insurance service is as follows:

$$F1 - Score = 2\text{x}\frac{presisi \text{ x } recall}{presisi + recall} \quad (4)$$

III. RESULTS AND DISCUSSION

This section presents the experimental results obtained from the implementation of the TabNet deep learning model for customer churn classification at Sinarmas Syariah Lhokseumawe Insurance Services. The discussion includes dataset distribution analysis, preprocessing results, model evaluation, comparison with baseline models, SHAP interpretation, and business implications derived from the prediction results.

A. *Dataset Distribution Analysis*

The dataset used in this study consists of 2,000 customer records containing customer demographic information, policy data, payment history, claims information, and customer interaction variables. Before applying the balancing technique, the dataset showed an imbalanced class distribution where non-churn customers significantly outnumbered churn customers. The imbalance condition can potentially affect classification performance because the model may become biased toward the majority class. Therefore, balancing techniques were required to improve the model's capability in detecting churn customers accurately. The distribution of churn and non-churn classes before applying SMOTE is illustrated in Figure 4.

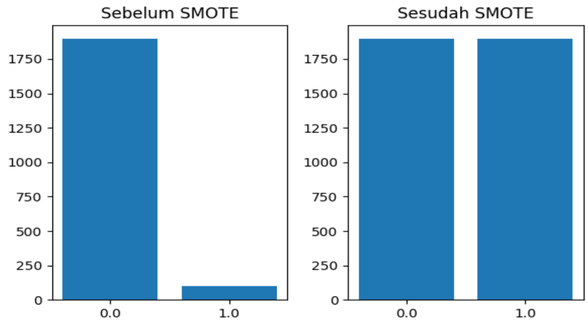

Figure 4. Comparison of Dataset Before and After Applying the SMOTE Method

As shown in Figure 4, the original dataset exhibited a significant class imbalance, where the number of non-churn customers substantially exceeded the number of churn customers. This condition can negatively affect the learning process of machine learning models because the model tends to focus on the majority class while overlooking patterns associated with the minority class. Consequently, the model may achieve high overall accuracy but perform poorly in identifying customers who are likely to churn, which is the primary objective of this study. To address this issue, the Synthetic Minority Oversampling Technique (SMOTE) was applied to generate synthetic samples for the minority class, resulting in a more balanced class distribution. By providing a more representative dataset during training, the balancing process enables the model to learn churn-related patterns more effectively and improves its ability to classify churn customers accurately.

B. *Data Preprocessing*

The data preprocessing stage includes key processes such as handling missing values, encoding categorical features, and data cleaning. These steps improve the quality of the dataset, improve its structure, and make it ready for the modeling phase. Figure 5 shows a visualization of the data ready for the Tabnet modeling phase.

	id_nasabah	nama	usia	jenis_kelamin	pekerjaan	produk	lama_langganan_tahunan	metode_bayar	telat_bayar	jumlah_tunggakan
0	0	69	56.0	1	21	5	0	0	0.0	0.0
1	1	927	64.0	0	29	5	0	0	0.0	0.0
2	2	884	51.0	0	29	5	0	0	0.0	0.0
3	3	923	32.0	0	9	5	4	0	0.0	0.0
4	4	677	57.0	0	29	5	3	0	0.0	0.0

5 rows × 27 columns

Figure 5. Preprocessing Results

C. *Model Evaluation Results*

The performance of the proposed TabNet model was evaluated using several classification metrics, including Accuracy, Precision, Recall, F1-Score, and ROC-AUC (Receiver Operating Characteristic – Area Under the Curve). These metrics were selected to provide a comprehensive assessment of the model's effectiveness in distinguishing between churn and non-churn customers. While Accuracy, Precision, Recall, and F1-Score evaluate the classification performance from different perspectives, ROC-AUC measures the model's overall ability to discriminate between the two classes across various decision thresholds.

The ROC curve represents the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) at different threshold settings. A higher ROC-AUC value indicates a better capability of the model to correctly separate churn customers from non-churn customers. The evaluation results of the proposed TabNet model are presented in Table III.

TABLE III
PERFORMANCE EVALUATION RESULTS OF THE PROPOSED TABNET MODEL

No	Evaluation Metric	Value	Percentage	Description
1	Accuracy	0.9876	98.77%	Overall model accuracy
2	Precision	0.8667	87%	Accuracy of churn predictions
3	Recall	0.9286	93%	Ability to detect churn
4	F1-Score	0.8966	90%	Balance between precision and recall
5	ROC-AUC	0.9950	99.50%	Ability to distinguish churn and non-churn

To provide a clearer visualization of the model's performance, the evaluation results presented in Table III are further illustrated using graphical representations. Figure 6 presents a bar chart comparing the values of Accuracy, Precision, Recall, F1-Score, and ROC-AUC obtained by the proposed TabNet model. The graphical representation facilitates a more intuitive interpretation of the model's performance across different evaluation metrics.

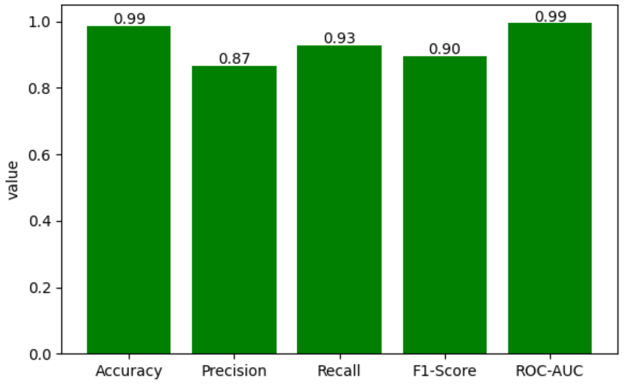

Figure 6. Comparison Of Model Evaluation Metrics

Furthermore, the discriminative capability of the model is analyzed using the Receiver Operating Characteristic (ROC) curve, as shown in Figure Y. The ROC curve illustrates the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR) at various classification thresholds. The corresponding Area Under the Curve (AUC) value provides an overall measure of the model's ability to distinguish between churn and non-churn customers, where higher AUC values indicate superior classification performance.

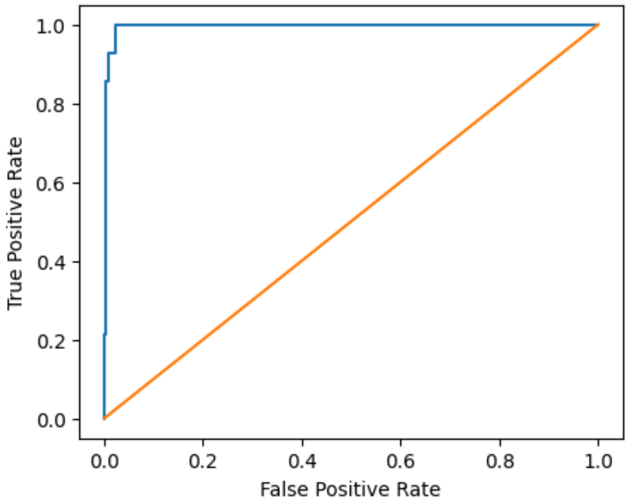

Figure 7. ROC curve

The results presented in Table III indicate that the proposed TabNet model achieved strong classification performance across all evaluation metrics. The accuracy of 98.77% demonstrates that the model was able to correctly classify the majority of customer records. Moreover, the precision of 87% and recall of 93% indicate that the model effectively identified churn customers while maintaining a relatively low number of false predictions. The higher recall compared to precision suggests that the model prioritizes detecting potential churn customers, which is beneficial in customer retention scenarios where missing actual churn customers may result in business losses.

The high performance achieved by the model can be attributed to the combination of SMOTE and TabNet. The SMOTE technique helped address the class imbalance problem by generating additional minority-class samples, enabling the model to learn churn-related patterns more effectively. Meanwhile, TabNet's sequential attention mechanism allowed the model to focus on the most relevant features for churn prediction. This is further supported by the ROC-AUC value of 99.50%, indicating an excellent ability to distinguish between churn and non-churn customers across different classification thresholds. These results suggest that the proposed approach can serve as a reliable decision-support tool for identifying customers at risk of churn and supporting proactive retention strategies in the insurance sector.

D. Feature Importance

TabNet model has the advantage of interpretability because it uses an attention mechanism that can indicate the importance of each feature at each stage of the decision-making process. To find the features deemed important by TabNet, see Table 4.

TABLE IV
FEATURE IMPORTANCE SCORES USING TABNET MODEL

No	Feature	Importance Value
1	communication	0.52
2	purchase_reason	0.08
3	character	0.08
4	claims_2022	0.07
5	premi_2025	0.07
6	claim_value	0.06
7	premi_2023	0.05
8	premi_2022	0.04
9	age	0.04
10	product	0.02

Table 5 presents the feature importance scores obtained from the TabNet model, revealing the relative contribution of each feature to customer churn prediction. The results indicate that communication has the highest importance score (0.52), substantially exceeding the contribution of other variables. This finding suggests that communication-related factors are the most influential determinants of customer churn, indicating that the quality, frequency, and effectiveness of interactions between customers and the insurance provider play a critical role in customer retention. The dominance of this feature implies that insufficient engagement or ineffective communication may increase the likelihood of customers discontinuing their insurance services.

The remaining features contribute to churn prediction with considerably lower importance scores. Variables such as purchase_reason, character, claims_2022, and premi_2025 exhibit moderate influence, suggesting that customer motivations, behavioral characteristics, claim history, and premium-related factors also affect customer retention decisions. Meanwhile, age and product show relatively low importance values, indicating that demographic and product-related attributes have a weaker impact on churn prediction compared to customer interaction and behavioral factors. These findings suggest that customer churn in the insurance sector is driven primarily by engagement and service-related experiences rather than demographic characteristics alone. Therefore, improving communication strategies and strengthening customer relationship management may be effective approaches for reducing customer churn and enhancing long-term customer retention.

E. SHAP (Shapley Additive Explanations)

In addition to TabNet feature importance, the study also applied the XAI SHAP (Shapley Additive Explanations) method to describe the contribution of features to predictions, both globally and locally [22]. SHAP is based on game theory, with each feature as a "player" that contributes to the final predicted outcome.

1. Global SHAP Interpretation Stage, Global interpretation aims to identify the most influential features overall in the model. The results are visualized in Figure 8 through the absolute average SHAP value.

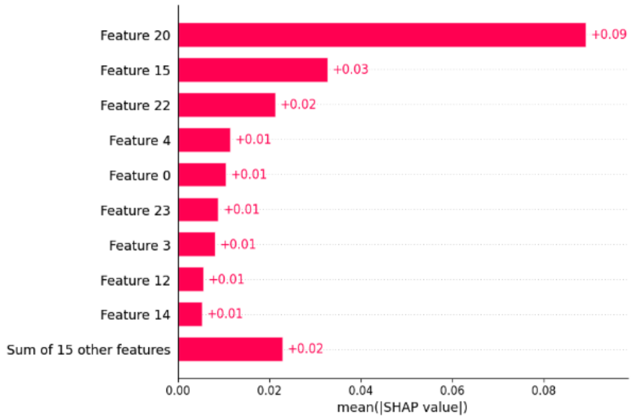

Figure 8. Global SHAP Interpretation.

Based on the global SHAP analysis results, several features exhibit a more dominant influence on the model's predictions than others. The findings indicate that Communication (Feature 20) has the highest SHAP value, followed by Premi_2025 (Feature 15) and Reason_to_Purchase (Feature 22). This result suggests that customer communication, premium-related factors, and purchasing motivations are the most influential variables in determining customer churn behavior. Features with higher SHAP values contribute more substantially to the model's decision-making process, either by increasing or decreasing the probability of churn.

2. SHAP distribution, In addition to the feature importance analysis, an examination of the SHAP value distribution was also carried out via the summary plot in Figure 9, which illustrates the distribution of feature contributions to the overall data.

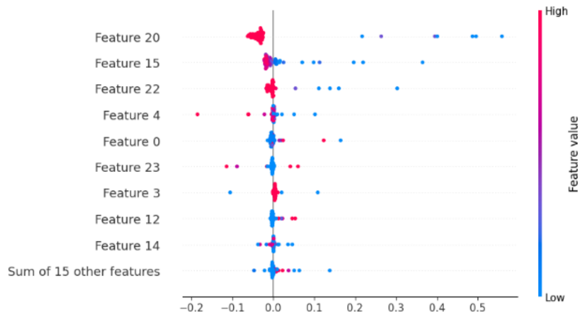

Figure 9. SHAP Distribution Analysis

In the graph shown in Figure 9, each point represents a single customer's data, while the color indicates the feature value, with red indicating high values and blue indicating low values. Based on the analysis, it appears that key features, such as Feature 20, have a wide distribution of SHAP values, indicating that this feature has a significant influence on the model.

3. Local Interpretation of SHAP: Local interpretation aims to describe the model's decisions on specific individual

customer data. The analysis results are presented in Figure 10 via a waterfall plot.

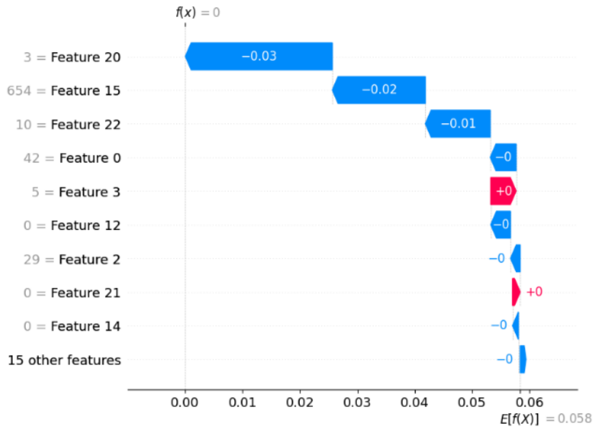

Figure 10. Local SHAP Interpretation

Based on the results shown in Figure 10, it can be seen that several features contribute dominantly to determining the prediction results. Features in red indicate contributions that increase the likelihood of churn, while those in blue indicate contributions that decrease the likelihood of churn. This analysis provides a deeper understanding of the rationale behind the model's classification for each individual customer.

Overall, the SHAP analysis consistently indicates that communication, premi_2025, and reason_to_purchase are the most influential factors affecting customer churn predictions. The results obtained from the global feature importance analysis, SHAP distribution plot, and local interpretation demonstrate that these features contribute significantly to the model's decision-making process, both at the dataset level and for individual customer predictions. The dominance of communication suggests that customer interaction and engagement play a critical role in retention decisions, while premium-related factors and purchasing motivations reflect customers' perceptions of value and service suitability.

These findings are not only relevant to Sinarmas Syariah Lhokseumawe Insurance Services but may also provide insights for the broader insurance industry. Similar customer-related factors are frequently associated with churn behavior in insurance and financial service organizations, where communication quality, customer experience, and policy characteristics influence customer loyalty. However, the relative importance of these features may vary across insurance companies due to differences in customer profiles, product portfolios, business strategies, and market conditions. Therefore, further studies involving datasets from other insurance providers are recommended to evaluate the generalizability of the proposed TabNet-SHAP framework and to determine whether similar churn patterns can be observed in different insurance environments.

IV. CONCLUSION

This study demonstrates that the integration of TabNet and SMOTE provides an effective approach for customer churn prediction in the insurance sector. The proposed model achieved strong predictive performance, with an accuracy of 98.77%, precision of 86.67%, recall of 92.86%, F1-score of 89.66%, and ROC-AUC of 0.995. These results indicate that the model is capable of accurately distinguishing between churn and non-churn customers while maintaining a high ability to detect customers at risk of leaving. The application of SMOTE successfully addressed the class imbalance problem, enabling the model to learn churn-related patterns more effectively and improving its overall classification performance.

SHAP-based explainability analysis revealed that communication, premi_2025, and reason_to_purchase were the most influential factors affecting churn predictions. These findings provide actionable insights for insurance companies, highlighting the importance of customer engagement, premium management, and understanding customer purchasing motivations in retention strategies. Although the proposed framework demonstrated promising results on the Sinarmas Syariah Lhokseumawe dataset, future research should evaluate its robustness using datasets from other insurance companies and compare its performance with established models such as Random Forest and XGBoost to further assess its generalizability across different business environments.

REFERENCES

[1] X. Liu, G. Xia, and X. Zhang, "Customer Churn Prediction Model Based on Hybrid Neural Networks," *Sci. Rep.*, vol. 14, p. 30707, 2024.

[2] T. P. Quinn and others, "A Primer on the Use of Machine Learning to Distil Knowledge from Data in Biological Psychiatry," *Mol. Psychiatry*, vol. 29, no. 2, 2024.

[3] P. C. Verhoef and others, "Digital Transformation and Customer Engagement in Service Industries," *J. Bus. Res.*, vol. 122, pp. 889–901, 2021.

[4] R. T. Rust, "The Future of Service Management and Customer Retention," *J. Serv. Res.*, vol. 24, no. 1, pp. 3–10, 2021.

[5] S. S. Alam and others, "Customer Loyalty and Retention in Islamic Insurance Services," *J. Islam. Mark.*, vol. 13, no. 4, pp. 1010–1028, 2022.

[6] M. Bahri and others, "Machine Learning Applications for Insurance Customer Retention," *Expert Syst. Appl.*, vol. 223, p. 119845, 2023.

[7] Y. Huang and J. Chen, "Customer Churn Prediction in Financial Services Using Machine Learning Techniques," *Appl. Soft Comput.*, vol. 122, p. 108839, 2022.

[8] M. Ahmad and others, "Deep Learning Approaches for Customer Churn Prediction: A Systematic Review," *IEEE Access*, vol. 11, pp. 23451–23470, 2023.

[9] H. Jabbar and R. Khan, "Comparative Analysis of Machine Learning Models for Churn Prediction," *Inf. Sci. (Ny).*, vol. 608, pp. 1110–1128, 2022.

[10] C. Molnar, "Interpretable Machine Learning in Finance and Business Analytics," *Mach. Learn. Knowl. Extr.*, vol. 4, no. 2, pp. 245–268, 2022.

[11] A. Khaidar, Nurdin, Fajriana, Taufiq, and D. Hamdhana, "Classification Analysis of Single Tuition Fees Using the Random Forest Method with K-Fold Cross Validation," *J. Appl. Informatics Comput.*, vol. 10, no. 1, pp. 125–133, Feb. 2026, doi: 10.30871/jaic.v10i1.11798.

[12] S. M. Lundberg and others, “From Local Explanations to Global Understanding with Explainable AI for Trees,” *Nat. Mach. Intell.*, vol. 2, no. 1, pp. 56–67, 2020.

[13] L. Merrick and A. Taly, “The Role of Explainable AI in Financial Decision Support Systems,” *AI Mag.*, vol. 43, no. 2, pp. 125–138, 2022.

[14] S. O. Arik and T. Pfister, “TabNet: Attentive Interpretable Tabular Learning,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, pp. 6679–6687.

[15] S. Wa, X. Lu, and M. Wang, “Stable and Interpretable Deep Learning for Tabular Data: Introducing InterpreTabNet,” *arXiv Prepr. arXiv2310.02870*, 2023.

[16] J. Si, W. Y. Cheng, and M. Cooper, “InterpreTabNet: Distilling Predictive Signals from Tabular Data by Salient Feature Interpretation,” *arXiv Prepr. arXiv2406.00426*, 2024.

[17] A. F. Tjahjono, H. Hasan, R. P. Putera, D. M. P. Indranto, and A. T. Hermawan, “Klasifikasi URL Phishing untuk SIEM: Perbandingan Model Machine Learning XGBoost dan Deep Learning TabNet dalam Deteksi Ancaman Siber,” *Sains Data J. Stud. Mat. dan Teknol.*, vol. 3, no. 2, pp. 62–71, 2025, doi: 10.52620/sainsdata.v3i2.227.

[18] K. Perawatan, P. Pada, D. Tabular, E. Ismanto, A. Fadlil, and A. Yudhana, “Jurnal Computer Science and Information Technology (CoSciTech) Analisis Perbandingan Model Fully Connected Neural Networks (FCNN) dan TabNet Untuk,” *J. Comput. Sci. Inf. Technol.*, vol. 5, no. 3, pp. 526–532, 2024.

[19] S. E. Awan, M. Bennamoun, F. Sohel, F. Sanfilippo, and G. Dwivedi, “A Reinforcement Learning-Based Approach for Imputing Missing Data,” *Neural Comput. Appl.*, vol. 34, pp. 9701–9716, 2022, doi: 10.1007/s00521-022-06958-3.

[20] F. Pargent, F. Pfisterer, J. Thomas, and B. Bischl, “Regularized Target Encoding Outperforms Traditional Methods in Supervised Machine Learning with High Cardinality Features,” *Comput. Stat.*, vol. 37, pp. 2671–2692, 2022, doi: 10.1007/s00180-022-01207-6.

[21] H. Kim, J. Lee, and others, “Data Preprocessing Combination to Improve the Performance of Quality Classification in the Manufacturing Process,” *Electronics*, vol. 11, no. 3, p. 477, 2022, doi: 10.3390/electronics11030477.

[22] M. Safii, Husain, Ika Okta Kirana, Sasha Aiko Leana, and Yuli Indahwati Gultom, “Model Prediksi Penjadwalan Produksi Energi Terbarukan dengan Algoritma XGBoost dan Analisis Interpretatif Menggunakan SHAP,” *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 4, no. 4, pp. 794–801, 2025, doi: 10.53513/jursi.v4i4.11443.