Predicting Customer Churn in Indonesian ISPs with Multilayer Perceptron and Marketing Intelligence

Gema Persada Arihta[1]*, Tanika D Sofianti[2], Win Sukardi[3]
[1]Swiss German University
[1]The Prominence Tower Alam Sutera, Tangerang, 15143, Indonesia
Email*: gema.arihta@student.sgu.ac.id

ARTICLE INFORMATION

Received dd-mmmm-yyyy
Revised dd-mmmm-yyyy
Accepted dd-mmmm-yyyy

Keywords:
Customer Churn
Multilayer Perceptron
Marketing Intelligence
Predictive Modelling
Telecommunications Industry

A B S T RACT

Customer churn is a major challenge in the highly competitive Indonesian Internet Service Provider (ISP) market, where companies face significant customer turnover rates impacting profitability and sustainability. This study integrates multilayer perceptron (MLP) neural networks with marketing intelligence to predict and mitigate churn effectively. The methodology includes data preparation, exploratory data analysis (EDA), and model development. EDA plays a critical role in identifying key features for churn prediction, ensuring meaningful insights into customer behavior. The model uses the Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance, improving prediction performance. The final model achieved an area under the curve (AUC) of 99%, a metric that measures how well the model distinguishes between churned and non-churned customers, and an F1 score of 97%, which balances the model's precision (accuracy of positive predictions) and recall (identification of all true churners). These findings provide actionable insights for ISPs to tailor customer retention strategies and improve business performance.

1. Introduction

The Internet Service Provider (ISP) industry in Indonesia has experienced exponential growth due to increasing internet adoption and digital customer engagement (Christiadi & Sule, 2018). Indonesia has become one of the fastest-growing telecommunications markets globally, with the number of internet users reaching 196.7 million in 2019/2020, an 8.9% increase from the previous year (Salma & Aprianingsih, 2021).This surge in internet users has driven significant investments in network capacity to meet growing demands for both mobile and fixed networks (Gaivoronski et al., 2017). However, this growth also intensifies market competition, with over 155 ISP companies in operation as early as 2001, nearly doubling by 2016 (Christiadi & Sule, 2018). Such fierce competition has led to heightened challenges in customer retention. Churn rates, which measure the percentage of customers discontinuing services, can reach as high as 40% annually among leading ISPs (Salma & Aprianingsih, 2021). Customer churn directly impacts a company's profitability and long-term sustainability, making it one of the most pressing issues in the Indonesian telecommunication industry.

This study addresses a critical issue faced by PT XYZ, one of Indonesia's largest ISPs, which has been experiencing a consistent increase in customer turnover over the past five years (Figure 1). Despite efforts to attract new customers, churn rates have steadily risen, driven by external market pressures and challenges in implementing effective retention strategies. This trend underscores the urgent need for a data-driven approach to managing churn. This study provides actionable solutions to address PT XYZ's churn problem by integrating predictive analysis through MLP models with marketing intelligence.

In the competitive Indonesian telecommunications industry, customer churn is a critical indicator of business performance. It reflects the rate at which customers discontinue using a company's services, directly impacting revenue and long-term sustainability. For ISPs like PT XYZ, rising customer churn rates have led to significant financial losses, as acquiring new customers often costs more than retaining existing ones. This underscores the need for accurate churn forecasting and tailored retention strategies. To address this, it is essential to understand the underlying factors driving customer churn, including service quality, pricing, customer support, and market competition. By leveraging predictive analytics and targeted marketing efforts, companies can mitigate these challenges and improve customer loyalty (Aulia Triyafebrianda & Windasari, 2022; Ribeiro et al., 2024; Salma & Aprianingsih, 2021).

Identifying churn-risk customers is crucial for ISPs to maintain existing clients and gain a competitive edge in the dynamic telecommunications market. Advanced methods, such as machine learning, have become indispensable in churn prediction and analysis. These techniques enable companies to uncover patterns in customer behavior, identify at-risk customers, and develop targeted strategies to improve retention. For PT XYZ, the application of Multilayer Perceptron (MLP) neural networks offers a novel approach to churn prediction by providing precise and actionable insights. This reflects the broader industry trend of leveraging data-driven solutions to reduce customer attrition and enhance competitive advantage (Edwine et al., 2022a).

To improve customer loyalty and reduce churn, it is essential for telecommunications companies to go beyond churn prediction and address its root causes. By combining machine learning predictions with an understanding of key churn drivers such as customer support, pricing, and service quality, companies can develop more effective retention strategies. For PT XYZ, this approach provides actionable insights to tailor customer engagement efforts and enhance loyalty programs. Integrating predictive analytics with marketing intelligence offers a comprehensive strategy, enabling ISPs to address customer demands and adapt to market dynamics effectively (Edwine et al., 2022; Ribeiro et al., 2024; Sana et al., 2022).

Customer churn prediction strategies have evolved significantly, transitioning from traditional statistical methods to advanced machine learning techniques. Early approaches, such as survival analysis and logistic regression, provided a solid foundation but struggled to address the complexities of non-linear data in modern datasets (Geiler et al., 2022a). The advent of machine learning, particularly artificial neural networks (ANN), has revolutionized churn prediction. Among these, Multilayer Perceptron (MLP) models stand out for their ability to handle large datasets and detect intricate patterns. With a deep, layered architecture, MLP models excel in adapting to new data, making them highly effective for precise churn projections. This study leverages MLP's predictive power to address the specific challenges faced by PT XYZ, demonstrating its applicability in a real-world, competitive ISP environment (Bogaert & Delaere, 2023a; Fareniuk et al., 2022a).

The adoption of machine learning, particularly MLP, has transformed churn prediction in the telecommunications industry. As a feedforward artificial neural network, MLP is adept at modeling complex, non-linear data interactions, making it ideal for identifying patterns in customer behavior. This is exemplified by its success in previous studies, such as SyriaTel's telecom data analysis, where MLP effectively processed extensive datasets to predict customer attrition (Ahmad et al., 2019a). For PT XYZ, MLP's ability to uncover subtle trends in large datasets offers a powerful tool to address rising churn rates. By leveraging these insights, PT XYZ can implement targeted retention strategies to enhance customer satisfaction and loyalty, leading to reduced turnover rates (Lalwani et al., 2022a). Despite its strengths, MLP requires careful calibration to avoid overfitting, as its complex architecture can increase computational demands. Balancing model complexity with practical efficiency is crucial to achieving reliable results. This study addresses these challenges by optimizing MLP's design and incorporating techniques like SMOTE to enhance its performance. The findings provide actionable insights that support PT XYZ's customer retention goals while showcasing the broader applicability of MLP in the telecommunications sector (Lalwani et al., 2022a).

In the ISP sector, marketing intelligence has emerged as a critical tool for improving decision-making by leveraging consumer and market data (Plangger et al., 2022). It plays a key role in evaluating customer behavior, understanding retention and loyalty trends, supporting targeted marketing strategies, and designing effective campaigns. For PT XYZ, marketing intelligence complements predictive analytics by translating churn predictions into actionable strategies. By integrating insights from Multilayer Perceptron (MLP) models with marketing data, PT XYZ can develop targeted campaigns that address customer needs more effectively. Previous studies have highlighted the benefits of combining marketing intelligence with machine learning techniques to handle customer churn. For instance, (Jahan & Farah Sanam, 2022) demonstrated how incorporating Exploratory Data Analysis (EDA) enhanced machine learning models, achieving remarkable prediction accuracy. Similarly, this study integrates EDA with MLP and marketing intelligence, enabling PT XYZ to align predictive insights with customer engagement strategies. This approach underscores the growing importance of combining technological and data-driven strategies to reduce churn and improve retention (Bogaert & Delaere, 2023a; Faritha Banu et al., 2022).

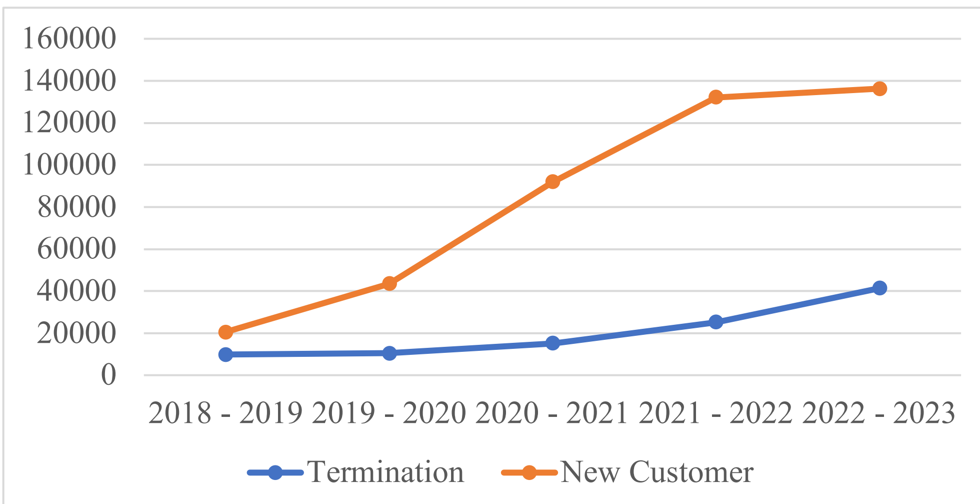


Figure 1. Termination and new customer number at PT XYZ

This study addresses a critical issue faced by PT XYZ, one of Indonesia's largest ISPs, which has been experiencing a consistent increase in customer turnover over the past five years (Figure 1). Despite efforts to attract new customers, churn rates have steadily risen, driven by external market pressures and challenges in implementing effective retention strategies. This trend underscores the urgent need for a data-driven approach to managing churn. This study provides actionable solutions to address PT XYZ's churn problem by integrating predictive analysis through MLP models with marketing intelligence. The integration of Exploratory Data Analysis (EDA) enables the identification of key churn drivers, while marketing intelligence translates these insights into targeted retention strategies. By leveraging these insights, PT XYZ can implement targeted customer engagement strategies, optimize loyalty programs, and tailor marketing campaigns to reduce churn effectively and enhance customer satisfaction. This comprehensive approach not only addresses the specific challenges faced by PT XYZ but also serves as a model for other ISPs operating in similarly competitive markets.

Building on the identified challenges, this study seeks to address the following research questions: (a) what factors and patterns drive customer churn at PT XYZ? (b) How effectively can Multilayer Perceptron (MLP) models predict customer attrition? (c) how can insights from MLP predictions and marketing data be integrated to develop targeted retention strategies? (d) what specific actions can PT XYZ implement to improve customer engagement and reduce churn?

The research questions lead to the following research objectives: (a) identify the key patterns and causes of customer churn at PT XYZ, (b) develop an effective churn prediction model using Multilayer Perceptron (MLP), (c) integrate MLP's predictive insights with marketing intelligence to design targeted retention strategies, (d) provide actionable recommendations to enhance customer engagement, reduce churn, and improve customer satisfaction, tailored to the unique characteristics of the Indonesian ISP market.

2. Literature review

2.1. *Customer Churn*

In the dynamic and competitive telecommunications business, particularly among Indonesian ISPs, customer churn remains a critical indicator of business performance. Defined as the rate at which customers discontinue using a company's services, churn directly impacts revenue and long-term profitability. The financial burden of customer attrition highlights the necessity of accurate predictions and effective management strategies (Peng et al., 2023). Research has identified several factors influencing churn, including customer loyalty, service quality, pricing structures, customer satisfaction and dissatisfaction, customer support, complaint management, switching costs, multi-brand attitudes, and customer demographics (Aulia Triyafebrianda & Windasari, 2022; Ribeiro et al., 2024; Salma & Aprianingsih, 2021). Recent advancements in churn prediction models, particularly leveraging data transformation techniques like Weight-of-Evidence (WOE) and feature selection, have significantly improved model performance (Sana et al., 2022). These methods have been shown to enhance customer churn prediction in the telecommunication industry, offering valuable insights into retention strategies. These advancements highlight the growing necessity for ISPs to adopt efficient approaches to identify churn-risk customers and leverage advanced analytics for retention strategies.

The need for an efficient approach to identifying churn-risk customers is critical for ISPs in the competitive telecommunications industry. Retaining existing customers not only reduces operational costs but also enhances competitive advantage. Machine learning and advanced data analytics have emerged as essential tools in churn prediction and assessment (Edwine et al., 2022). Techniques such as Random Forest, Support Vector Machines (SVM), and K-Nearest Neighbors have demonstrated notable prediction accuracy, offering valuable insights into customer behavior. These advancements reflect a broader industry trend toward data-driven strategies, enabling telecom firms to proactively monitor and minimize customer churn (Edwine et al., 2022). While these machine learning methods provide powerful tools for predicting churn, their effectiveness is maximized when combined with an understanding of the underlying drivers of customer behavior.

To enhance customer loyalty and reduce churn rates, telecommunications companies must go beyond merely estimating churn and address its underlying causes. By integrating machine learning predictions with an understanding of key churn drivers—such as service quality, pricing, and customer support—ISPs can design more targeted and effective retention strategies (Edwine et al., 2022). This comprehensive approach combines traditional churn research with advanced data analytics, providing ISPs with deeper insights into customer behavior. Such a strategy underscores the importance of blending technological advancements with a nuanced understanding of customer needs and market dynamics, enabling telecom firms to remain competitive in an evolving industry (Edwine et al., 2022; Ribeiro et al., 2024; Sana et al., 2022).

2.2 *Churn Prediction Models*

Globally, the telecommunications sector has increasingly adopted machine learning for churn prediction. For example, models like logistic regression and neural networks have been used to analyze customer behavior and improve retention rates in competitive markets (Geiler et al., 2022). Indonesian ISPs can draw from these

advancements to enhance their own predictive capabilities. The advent of advanced computing systems and machine learning techniques marked a pivotal shift, enabling the development of more sophisticated models.
Artificial Neural Networks (ANN), including decision trees, Support Vector Machines (SVM), and particularly Multilayer Perceptrons (MLP), have emerged as powerful tools for churn prediction. MLPs, with their deep, layered architecture, excel at processing massive datasets and detecting intricate patterns. Their adaptability to new data has been widely demonstrated, making them indispensable for precise churn predictions in dynamic sectors like telecommunications (Bogaert & Delaere, 2023; Fareniuk et al., 2022).
Comparative analyses reveal a clear trend: while machine learning models, particularly complex neural networks like MLPs, often outperform traditional statistical approaches, basic statistical models remain valuable for their simplicity and interpretability. Ensemble techniques, which combine multiple models to enhance robustness and predictive accuracy, exemplify the potential of integrating diverse approaches (Bogaert & Delaere, 2023).
While Random Forest and SVM are effective for churn prediction, MLP stands out for its ability to model complex, non-linear relationships within large, high-dimensional datasets. Random Forest excels in classification tasks but struggles with capturing intricate, non-linear patterns that are common in churn data. SVM, while powerful for binary classification, can become computationally intensive with large datasets, requiring extensive parameter tuning. In contrast, MLP's deep learning architecture can capture these complexities, making it better suited for churn prediction in dynamic and large-scale datasets like those in PT XYZ.

Recent advancements in feature selection and data transformation techniques, such as Weight-of-Evidence (WOE), have further enhanced the predictive power of neural networks and logistic regression models. Specifically, WOE transforms raw data into structured and predictive indicators, enabling MLP models to effectively analyze complex patterns in telecom churn prediction. These developments suggest that the future of churn prediction lies in the strategic combination of traditional statistical methods and modern machine learning techniques, offering a balanced approach to accuracy and efficiency (Sana et al., 2022).

2.3 *Machine Learning in Churn Prediction*

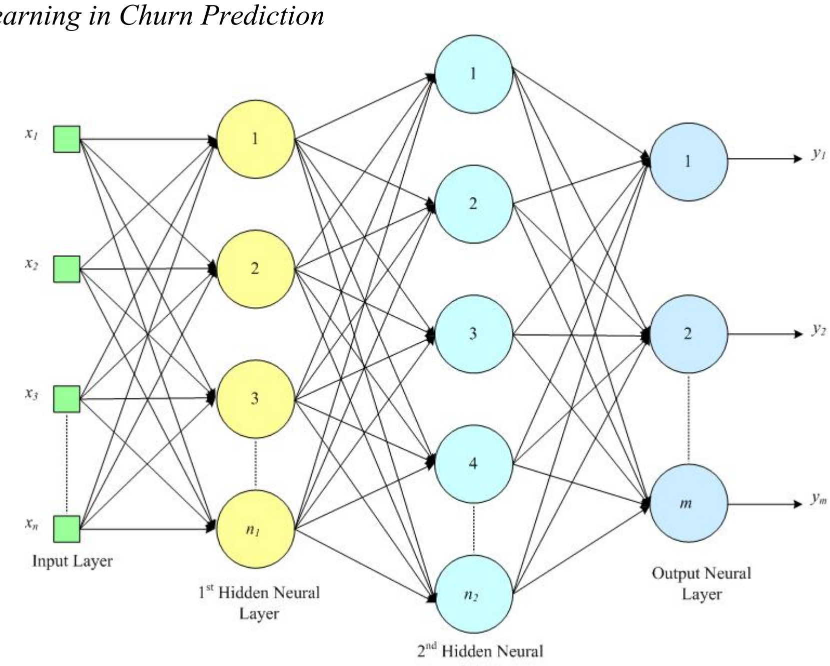


Figure 2. Multiple-Layer Perceptron Architectures (da Silva et al., 2017)

Machine learning has revolutionized churn prediction, transforming client retention strategies in the telecommunications sector. Among these techniques, Multilayer Perceptrons (MLP), a type of feedforward artificial neural network, excel at modeling complex, non-linear data interactions. This capability is evident in telecommunications case studies, such as those conducted on SyriaTel's telecom data, where MLP effectively processed large datasets to predict customer churn with high accuracy (Ahmad et al., 2019).

The MLP architecture typically consists of input, hidden, and output layers, enabling it to analyze customer behavior and usage trends comprehensively. In the context of telecom churn prediction, the input layer can process variables such as usage patterns, customer complaints, subscription tenure, payment history, and demographic data. These features are analyzed in the hidden layers to uncover nuanced trends, such as the interplay between service usage and dissatisfaction levels, which contribute to churn. By leveraging these insights, telecom companies can proactively identify potential churners, implement targeted interventions, and enhance customer satisfaction and loyalty, ultimately reducing churn rates (Lalwani et al., 2022).

Despite its high precision and adaptability, MLP has limitations. It demands significant computational resources and careful calibration to avoid overfitting. Consequently, a balance must be achieved between model complexity and practical application efficiency (Lalwani et al., 2022). These constraints notwithstanding, MLP exemplifies the convergence of technology and customer relationship management, offering a powerful tool for competitive advantage in the telecommunications sector (Ahmad et al., 2019; Lalwani et al., 2022).

3. Method

The research methodology consisted of four main steps: data preparation and preprocessing, exploratory data analysis (EDA), MLP design, and model evaluation (da Silva et al., 2017; Lalwani et al., 2022). All experiments were conducted using Python programming to implement the machine learning process in MLP. The dataset consisted of 36 attributes and 84,605 records, including both quantitative and qualitative features.

Data preparation was a critical step to ensure the validity and reliability of the dataset for machine learning. Missing values were imputed using the mean for numerical variables, such as tenure and monthly usage, and the mode for categorical variables, such as package type. Outliers were identified through descriptive statistics and visualizations, including boxplots, and were addressed to prevent skewed results. Categorical features, such as packagename and category, were converted into numerical representations using one-hot encoding, ensuring compatibility with machine learning algorithms. The dataset was split into training and testing subsets using a 70-30 split ratio, a commonly used standard to balance model training and evaluation. Additionally, Synthetic Minority Oversampling Technique (SMOTE) was applied to address class imbalances in the target variable, improving the model's ability to predict churn effectively.

Exploratory Data Analysis (EDA) helps understand the data and guides feature selection. We used descriptive statistics like mean and standard deviation to summarize numerical features, and correlation analysis to identify relationships between variables. For example, subscription tenure showed a strong positive correlation with churn, while billing complaints had a negative correlation. We also used visualizations such as histograms and scatter plots to explore the distribution of numerical features, and bar charts for trends in categorical features like package type. Time-series analysis revealed patterns in customer subscriptions over time, helping us understand churn dynamics. Based on these insights, we selected key features like age, tenure, billing complaints, and package type for the MLP model, while excluding irrelevant features. We also applied transformations like one-hot encoding for categorical variables and scaling for numerical ones, ensuring compatibility with the MLP model and improving its performance.

The MLP neural network was designed to handle the complexities of churn prediction. The architecture consisted of an input layer, two hidden layers, and an output layer. The hidden layers contained 64 and 32 nodes, respectively, optimized through hyperparameter tuning. ReLU (Rectified Linear Unit) activation functions were used in the hidden layers due to their ability to handle non-linear patterns without vanishing gradient issues. The sigmoid activation function was applied in the output layer to enable binary classification (churn vs. no churn), providing probabilistic outputs. Hyperparameter tuning was carried out using grid search to optimize the performance of the MLP model. The key hyperparameters optimized included the number of hidden layers, with two layers consisting of 64 and 32 nodes, chosen based on cross-validation performance. A learning rate of 0.001 was selected to balance training speed with model accuracy, while a batch size of 32 was chosen to optimize memory usage and ensure model stability during training. The model was trained for 50 epochs to allow for sufficient learning without overfitting. These hyperparameters were carefully tuned to strike a balance between model complexity and training efficiency, and the resulting configuration achieved the highest F1 score and AUC.

The model's performance was evaluated using a combination of metrics and validation techniques. Key metrics, including precision, recall, accuracy, F1 score, and Area Under the Curve (AUC), were calculated to assess the model's predictive capabilities comprehensively. A confusion matrix was employed to analyze class-specific predictions, such as true positives and false negatives, offering insights into the model's handling of class imbalances. Additionally, 5-fold cross-validation was applied to ensure robustness across varying dataset splits, with the k-value chosen to balance computational efficiency and reliable performance estimates. By following these steps, the methodology provided a robust and reliable approach to customer churn prediction, enabling actionable insights to enhance customer retention strategies.

4. Results and Discussion

4.1. *Exploratory Data Analysis*

Following data preparation and preprocessing, a subset of 73,704 records was utilized for Exploratory Data Analysis (EDA). EDA plays a pivotal role in transforming raw data into actionable marketing insights, guiding strategic decision-making processes. By leveraging EDA, businesses can ensure that decisions are informed by robust, data-driven evidence. EDA not only highlights trends and patterns contributing to customer churn but also supports key tasks such as feature selection, identifying the most impactful predictors to include in the model. Additionally, it helps resolve issues like class imbalance and visualizes the relationships between variables, both of which are essential for developing effective predictive models (Riaz Sadia et al., 2021). Below are the features that were highly related to affect customer churn.

4.1.1. *Cstmr_age*

The dataset's variable cstmr_age exhibited a wide range of values, including anomalies such as 0 and 124, which are likely data entry errors or outliers. These discrepancies do not accurately represent the typical customer base. To address this issue, customers outside the age range of 17 to 80 were omitted from the analysis. While cstmr_age was excluded as a predictor in the churn prediction model to avoid potential biases, it was retained for generating marketing insights and supporting research. The distribution of this variable within the specified age range provides valuable information for designing targeted marketing campaigns. Figure 3 illustrates the distribution of churn across different age groups. The plot highlights trends and variability in churn rates, revealing that the 25–34 age group experiences the highest churn levels. This trend suggests that customers in this demographic are particularly vulnerable to churn, likely due to higher dissatisfaction levels or unmet expectations. In contrast, older age groups display significantly lower churn rates, indicating greater loyalty or satisfaction. These findings underscore the importance of tailoring marketing and retention strategies to specific age groups. For instance, retention efforts should prioritize addressing the concerns of younger customers in the 25–34 demographic, while targeted campaigns can focus on leveraging loyalty within older age brackets. By aligning strategies with these insights, businesses can effectively reduce churn and enhance customer satisfaction across diverse age groups.

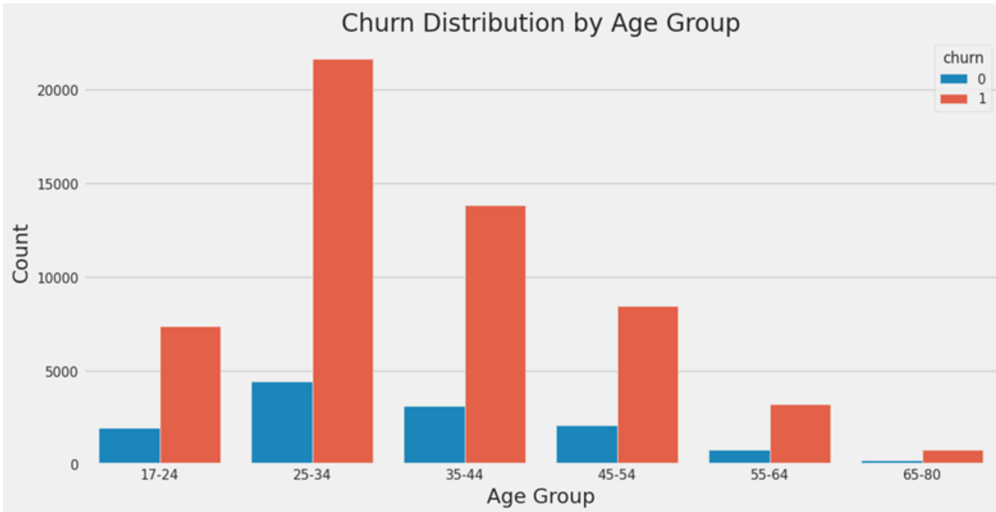


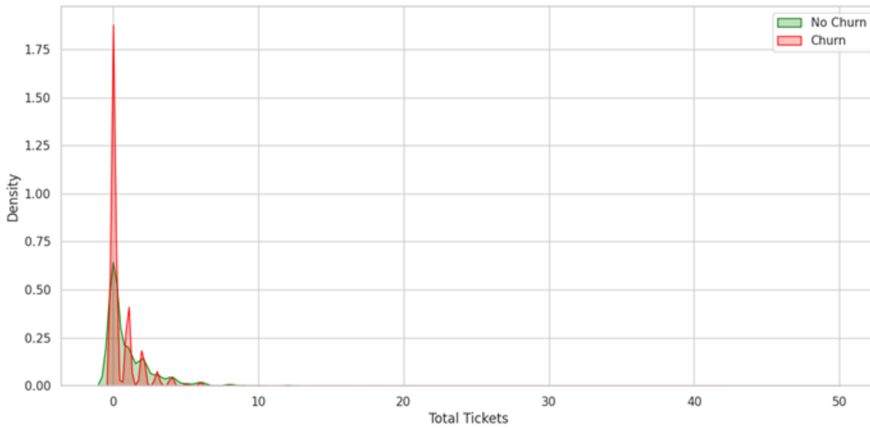
Figure 3. Churn Distribution by Age Group

4.1.2. *Packagename*

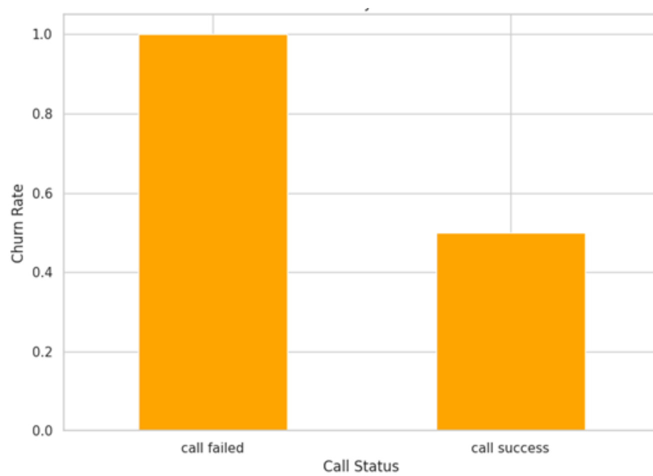
Analysis of the packagename variable revealed that the top three most popular packages are "Biznet Home Internet 1B Rent," "Biznet Home Internet 1C Rent," and "Biznet Home Internet 1C." However, churn rates do not always align with a package's popularity. While some popular packages exhibit lower churn rates, others have unexpectedly high churn rates. This observation suggests that factors beyond popularity, such as service quality, pricing, or customer expectations, significantly influence customer retention. Interestingly, packages with a smaller subscriber base tend to experience higher churn rates. This trend may indicate customer dissatisfaction or a mismatch between the features of these packages and customer expectations. For instance, niche packages may fail to meet the specific needs of their target audience, resulting in greater customer turnover. A chi-square test ($\chi^2 = 3207.558$, $p < 0.01$) confirmed the statistical significance of the relationship between packagename and churn. These results highlight the need for PT XYZ to delve deeper into package performance and customer preferences. Investigating why certain packages have higher churn rates—despite their popularity or niche appeal—can uncover critical insights for product improvement. To address these findings, PT XYZ should refine underperforming packages by conducting surveys or focus groups to identify gaps between customer expectations and the features offered. Additionally, targeted marketing campaigns can highlight the strengths of high-retention packages to attract new subscribers while addressing pain points in underperforming packages. Finally, continuous monitoring of package-level churn trends will enable PT XYZ to adapt product offerings dynamically to meet customer needs.

4.1.3. *Call_status, total_tickets, category*

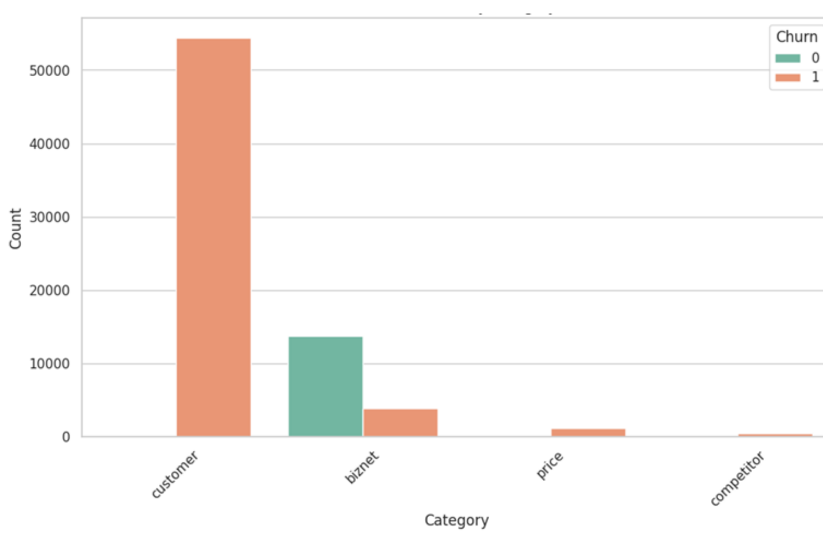
In the domain of communication and service interaction, the features total tickets (the total number of service tickets raised by a customer), call_status (the outcome of service calls), and category (the major category of customer retention call results) were found to have a significant impact on customer churn, as illustrated in Figures 4(a-c).



(a)



(b)



(c)

Figure 4. Churn rate status based on (a) total tickets, (b) call status, and (c) category

The analysis of total_tickets reveals that customers with zero service tickets are more likely to churn. This trend suggests low engagement or a lack of awareness of available support resources, which can contribute to dissatisfaction. The relationship between total_tickets and churn is statistically significant (p-value = 0.0), emphasizing the importance of customer engagement in retention strategies. Despite minor fluctuations in the graph, the findings confirm the critical role of total_tickets in predictive modeling. This insight highlights the need to actively engage customers by promoting support channels and enhancing service accessibility to mitigate churn risks.

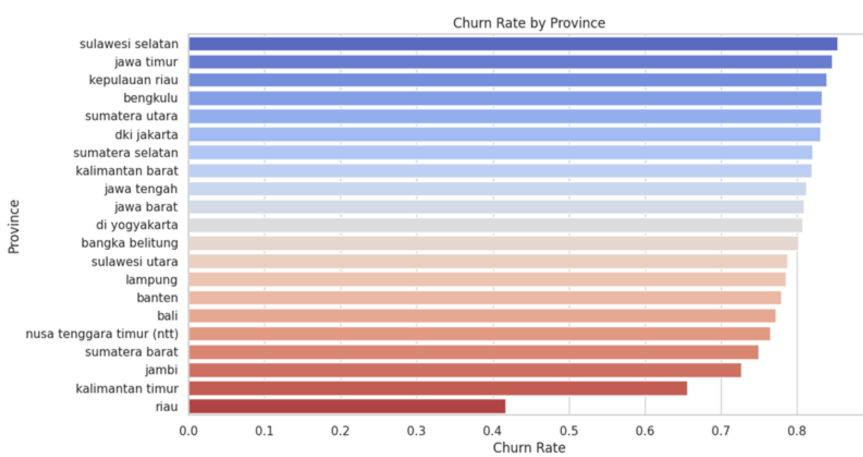
The visualization of call_status indicates a marked disparity in churn rates between successful and unsuccessful service calls. Customers with unsuccessful service calls exhibit a significantly higher churn rate, underscoring the importance of effective communication. Statistical analysis (chi-square = 28,400.75, p-value = 0.0) validates the strong association between call outcomes and churn. Improving communication strategies, such as ensuring prompt and effective resolution of customer issues, could substantially enhance customer satisfaction and retention.

The category variable, which reflects the outcome of customer interactions, also demonstrates a strong association with churn (chi-square = 53,692.48, p-value = 0.0). The data highlights recurring issues related to ineffective communication and trouble contacting customers. These findings suggest that strategies developed to address call_status issues—such as improving outreach efforts and maintaining accurate customer data—could similarly benefit category-related challenges. This alignment underscores the need for a cohesive strategy to enhance communication and service quality across all touchpoints.

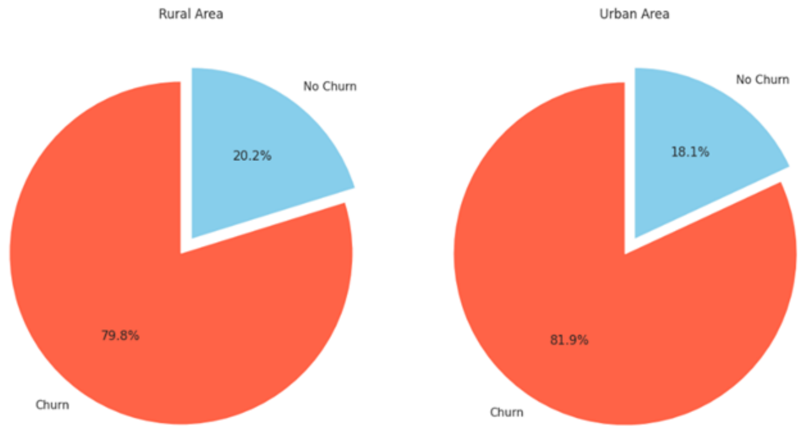
The overlapping themes in call_status and category highlight the critical need for improved outreach strategies, strengthened communication channels, and accurate customer data. Ensuring timely and effective interactions with customers while maintaining up-to-date contact information could minimize failed interactions. Additionally, expanding service accessibility and providing proactive solutions to customer issues would further enhance customer satisfaction and retention. By addressing these communication-related challenges, PT XYZ can better understand and predict churn while implementing targeted strategies to improve customer retention..

4.1.4. *Sa_province, territory, area_type*

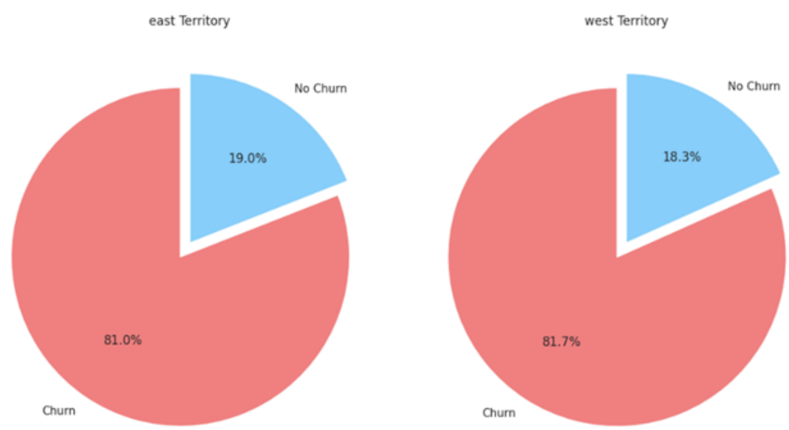
The geographic analysis indicates that customer churn is significantly influenced by physical location, with notable differences observed across provinces, territories, and urban/rural areas (Figures 5a–c). These findings underscore the importance of region-specific strategies for managing churn and meeting customer needs effectively.



(a)



(b)



(c)

Figure 5. Churn rate status based on (a) province, (b) area type (rural/urban), and (c) territory

At the province level, churn rates vary substantially, with Riau experiencing the highest churn rate of approximately 78% (Figure 5a). Conversely, provinces such as Sulawesi Selatan and Jawa Timur exhibit relatively lower churn rates, indicating stronger customer retention. These disparities may be attributed to differences in service quality, the availability of competitive alternatives, or regional customer expectations. PT XYZ should prioritize improving service quality and tailoring retention programs for high-churn provinces like Riau and Kalimantan Timur. In terms of area type, Figure 5b illustrates churn rates in urban and rural areas. Urban areas exhibit slightly higher churn rates (81.9%) compared to rural areas (79.8%). This difference suggests that urban customers may have greater access to alternative service providers or higher expectations of service quality. PT XYZ can address this challenge by offering competitive pricing and enhanced customer support in urban regions while leveraging loyalty in rural areas.

The territory-level analysis (Figure 5c) reveals similar churn rates across east and west territories, with the west territory exhibiting a marginally higher churn rate (81.7%) compared to the east territory (81.0%). While the differences are minimal, they reflect consistent trends across the company's service areas, emphasizing the need for a uniform but adaptable retention strategy. These geographic findings highlight the need for tailored strategies. For high-churn provinces, such as Riau and Kalimantan Timur, PT XYZ should introduce loyalty programs and address specific service complaints to improve retention. Urban customers require competitive pricing plans and faster service resolutions to counter competition and dissatisfaction. Meanwhile, consistent service quality across all territories can help maintain overall customer satisfaction. By aligning marketing and retention efforts with geographic insights, PT XYZ can address churn more effectively and enhance customer satisfaction across its service regions.

4.1.5. *Cust_seg*

The 'cust_seg' attribute, which differentiates individual and business customers, revealed variations in churn rates between the two segments. Although the Chi-Square test showed no statistically significant relationship between 'cust_seg' and churn (chi-square = 0.99, p-value = 0.3196), further analysis highlights segment-specific trends that provide valuable insights for PT XYZ's retention strategies. Figure 6 demonstrates that individual customers exhibit a slightly higher churn rate (82.8%) compared to business customers (81.3%). These differences may stem from varying levels of engagement, satisfaction, and service expectations across the two segments. Individual customers are more likely to prioritize affordability, ease of use, and personalized interactions, while business customers tend to value service reliability, tailored packages, and long-term partnerships. These findings underscore the importance of adopting segment-specific approaches to effectively address churn in both groups.

To improve retention, PT XYZ should develop distinct strategies tailored to each customer segment. For individual customers, personalized communication, competitive pricing, and loyalty programs can help foster stronger connections and improve satisfaction. For business customers, strategies should focus on ensuring high-quality, reliable service, offering tailored solutions that align with their operational needs, and rewarding long-term commitments through partnership incentives. Conducting regular customer surveys can further uncover the unique needs and preferences of each segment, enabling PT XYZ to refine its customer engagement initiatives. By aligning customer retention efforts with the specific requirements of individual and business segments, PT XYZ can improve customer satisfaction, strengthen relationships, and reduce churn across its client base.

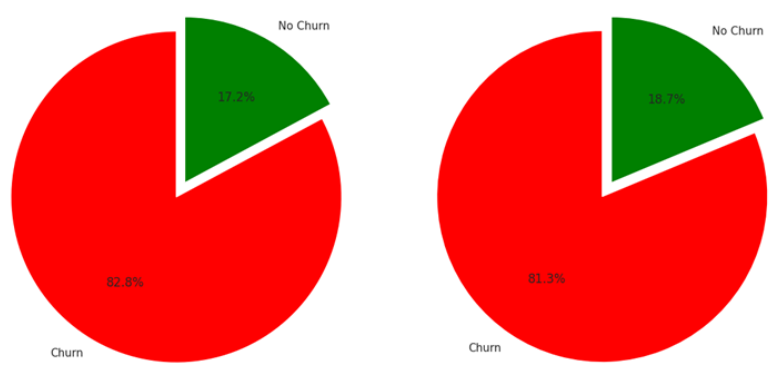


Figure 6. Churn rate status based on customer segment (business [left] and individual customers [right]) shown on the left.

4.1.6. *Packageprice_idr & subscription_tenure*

The analysis of subscription details revealed a relationship between package price and subscription tenure, both of which significantly impact customer churn. These insights offer valuable opportunities for PT XYZ to refine its pricing strategies and early-stage engagement initiatives. The packageprice_idr data suggests that customers who start with higher-priced packages are more likely to churn, as observed in the lower quartile (25th percentile, Q1) of churned customers (Figure 7). This trend indicates that customers who opt for premium packages may have elevated expectations and are more likely to cancel if these expectations are not met. Interestingly, when comparing churned and non-churned customers, the middle and higher ranges of package prices exhibit notable similarities. This finding suggests that most customers, whether they churn or not, tend to choose packages within a specific price range, which likely represents the company's most popular or standard offerings.

The pricing similarity across customer groups indicates the presence of a broadly appealing price tier. This suggests that PT XYZ's pricing strategy successfully caters to a wide consumer base, emphasizing the importance of maintaining this competitive pricing structure. However, additional research into the churn behavior of high-paying customers could help tailor premium services to better align with their expectations. Figure 8 illustrates that churn rates are highest among customers in the early stages of their subscription. This pattern suggests that newly onboarded customers are more likely to cancel their service, highlighting the critical importance of early engagement. During this initial period, creating a positive onboarding experience and offering compelling introductory incentives can play a pivotal role in building customer loyalty. Engagement initiatives such as personalized welcome communications, targeted promotions, and responsive customer support could significantly reduce churn among new customers.

These findings highlight several actionable strategies for PT XYZ. First, the company should refine its premium offerings by conducting surveys to better understand the needs and expectations of high-paying customers, ensuring that premium services align with their value perceptions. Second, strengthening onboarding processes with tailored engagement initiatives during the first six months can help address potential dissatisfaction early, fostering stronger relationships with new customers. Lastly, PT XYZ should continue offering its popular mid-range pricing packages while exploring additional bundles that appeal to specific customer segments. By addressing these key aspects of package pricing and subscription tenure, PT XYZ can enhance customer satisfaction and retention while building a stronger connection with its diverse consumer base.

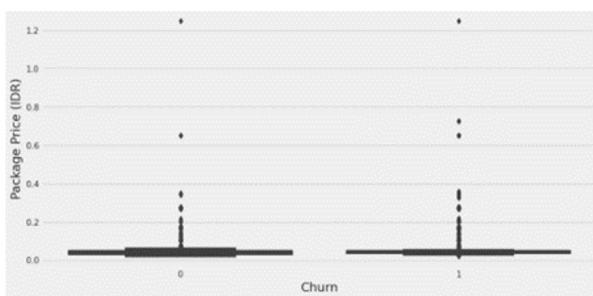


Figure 7. Boxplot of package price and churn

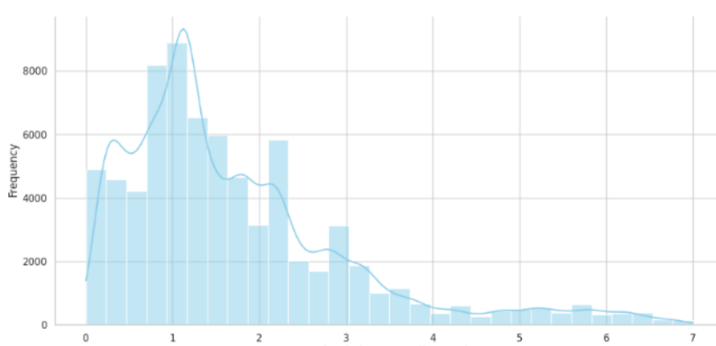


Figure 8. Distribution of subscription tenure

4.1.7. *Churn*

The analysis of the target variable, churn, revealed a critical issue for PT XYZ, with 81.4% of customers discontinuing services—far exceeding industry benchmarks of 10%–30% globally and 50%–60% regionally. This high churn rate threatens revenue stability, increases customer acquisition costs, and weakens competitiveness. Addressing this requires proactive engagement, targeted retention strategies, and enhancements in service quality and pricing. Predictive analytics can identify at-risks customers early, enabling data-driven interventions to reduce churn and improve satisfaction.

4.2. *MLP Model Development*

Table 1. Feature Selection for MLP Modeling from EDA analysis.

Features Group	Features
Categorical data	packagename, call_status, category, sa_province, territory, cust_seg, area_type
Numerical data	packageprice_idr, subscription_tenure, total_tickets
Target	churn

Based on the results of Exploratory Data Analysis (EDA), ten features were selected for MLP modeling (Table 1). These features include both categorical and numerical data, along with the target variable churn. The selected features were tested for their significance in predicting churn using a correlation matrix and statistical tests to validate the EDA findings. While most features demonstrated strong associations with churn, the cust_seg attribute, which differentiates individual and business customers, showed weaker predictive power. The Chi-Square test result (chi-square = 0.99, p-value = 0.3196) indicated no statistically significant relationship with churn, leading to its exclusion to streamline the MLP model. However, analysis revealed segment-specific trends, with individual customers exhibiting a slightly higher churn rate (82.8%) than business customers (81.3%). These insights highlight the need for PT XYZ to adopt targeted strategies, such as personalized communication for individuals and tailored solutions for business clients, to complement the model's predictions. The final set of features was carefully chosen to ensure relevance and predictive accuracy. The combination of these features reflects a balanced inclusion of categorical and numerical data, enabling the MLP model to capture complex patterns in customer churn effectively.

The MLP model architecture consisted of one input layer with 9 neurons, two hidden layers with 64 and 32 neurons respectively, and one output layer with a single neuron for binary classification. This architecture was carefully designed to balance complexity and computational efficiency, ensuring the model effectively captured non-linear patterns in the data. To determine the most optimal model, the MLP was trained and evaluated under three different scenarios. First, the model was trained using the original dataset without any modifications. Second, class weights were applied to adjust the loss function, making the model more sensitive to errors in the minority class (Prabadevi et al., 2023). Finally, the Synthetic Minority Over-sampling Technique (SMOTE) was used to balance the class distribution by creating synthetic samples for the minority class. This approach aimed to improve the model's ability to generalize across a broader range of data (Gu et al., 2020).

The outcomes of these scenarios, as shown in Table 2, highlight distinct strengths and trade-offs. While training the model on the original data showed strong overall performance, it struggled with class imbalance, resulting in a higher number of false negatives. The AUC of 99% demonstrates the model's ability to effectively distinguish between churned and non-churned customers, allowing PT XYZ to prioritize high-risk customers and reduce churn. The F1 score of 97% indicates a good balance between precision and recall, enabling PT XYZ to accurately target churners while avoiding unnecessary spending on customers who are not at risk. Applying class weights improved precision to 1.00, but this came at the expense of slightly lower recall, reflecting a trade-off between sensitivity and specificity. Meanwhile, the SMOTE-augmented model offered a balanced improvement in precision, recall, and AUC, making it the most robust approach for handling class imbalance.

Table 2. MLP Model Performance Results.

Model Performance Metrics	Performance Result		
	Using Original Data	Applying Class Weight	Using Smote
Loss	0.11	0.15	0.11
Precision	0.99	1.00	1.00
Recall	0.96	0.94	0.95
AUC	0.98	0.98	0.99
F1 Score	0.97	0.97	0.97
Confusion Matrix	TP: 8636.0 FP: 98.0 TN: 1946.0 FN: 376.0	TP: 8477.0 FP: 5.0 TN: 2039.0 FN: 535.0	TP: 8542.0 FP: 12.0 TN: 2032.0 FN: 470.0

Both class weights and SMOTE proved essential in addressing the challenges posed by imbalanced datasets. Class weights adjusted the loss function to reduce errors in the minority class, enhancing precision. SMOTE, on the other hand, increased data diversity and variability, allowing the model to generalize better to unseen data. These findings highlight the importance of incorporating such techniques into predictive modeling for churn analysis. By evaluating these scenarios, PT XYZ gains valuable insights into effective strategies for improving churn predictions. The adoption of SMOTE or similar techniques can enhance the model's accuracy and reliability, ultimately supporting better decision-making in customer retention efforts.

The model trained on the original data showed high accuracy and recall, leading to a strong F1 score and high AUC, indicating excellent performance. However, challenges with class imbalance were evident, as True Positives (TP) and True Negatives (TN) highlighted areas for improvement in minority class predictions. This limitation suggested the need for alternative techniques to address class imbalance and enhance model effectiveness.

When class weights were applied, the model exhibited a slight increase in loss, while accuracy remained high. Despite this, there was a noticeable reduction in recall, meaning the model became less effective at identifying all positive instances (churned customers). On the other hand, the increase in TN reduced false positives (FP), enhancing the model's ability to predict non-churned customers accurately. These results indicated a trade-off between improving specificity and reducing sensitivity, necessitating further exploration of class balancing techniques.

Using SMOTE-augmented data, the model achieved similar loss values to the original data scenario (0.11) while showing improved precision and recall compared to the class weight approach. Notably, the model with SMOTE obtained the highest F1 score (0.973) and AUC, indicating the best balance between specificity and sensitivity.

Despite the strong performance of the MLP model, there are several limitations to consider. First, issues with data quality, like outliers and data entry errors, could still affect the predictions. Some anomalies, such as unrealistic values in features like cstmr_age, may not have been fully addressed, which could distort the results. Additionally, feature engineering could be improved to enhance accuracy. Incorporating more features, such as customer service interactions or usage patterns, could provide better insights into churn. Another limitation is the computational complexity of the MLP model. Its high resource demand may make it impractical for smaller ISPs, which might benefit more from simpler models like logistic regression or decision trees, although these may not capture as many patterns. Lastly, even though SMOTE was used to balance the data, the model still had higher false negatives (FN), meaning some high-risk customers were missed. This could lead to lost revenue opportunities. Future work should focus on improving feature selection and recall to better identify customers at risk of churn.

The findings from this study provide valuable insights for PT XYZ in addressing customer churn. By accurately identifying high-risk customers, PT XYZ can prioritize retention efforts, offering tailored promotions or special offers to retain these individuals. The model's ability to accurately predict non-churned customers (True Negatives) also allows for better resource allocation, avoiding unnecessary interventions for customers unlikely to leave. Although there are concerns with false negatives, improving recall can help PT XYZ capture more at-risk customers. By segmenting customers based on churn likelihood, PT XYZ can deliver personalized communication and engagement strategies to both individual and business customers. Additionally, continuing to monitor customer satisfaction and integrating feedback with churn prediction models will further refine retention strategies. Leveraging the MLP model alongside marketing intelligence tools can help PT XYZ develop more targeted, data-driven strategies to enhance customer satisfaction, improve loyalty, and reduce churn.

While the model delivered stronger predictions for TN, it also recorded a higher number of False Negatives (FN) than in the original data scenario. The relatively higher false negatives (470) in the SMOTE-augmented model result from the synthetic data generation process, which can introduce slight noise. This trade-off emphasizes the importance of prioritizing sensitivity (recall) to identify as many at-risk customers as possible, which is crucial for minimizing churn. However, this comes with a risk of misclassifying some churned customers as non-churned. Balancing sensitivity and specificity is essential for PT XYZ—focusing on recall helps target high-risk customers proactively while ensuring efficient use of retention resources.

To further validate this outcome, the SMOTE-augmented MLP model was tested using a confusion matrix with 9,012 test data points. As shown in Figure 9, the model accurately identified 8,542 churned customers (TP), demonstrating excellent sensitivity. It also correctly classified 2,032 non-churned customers (TN), highlighting strong specificity. Additionally, only 12 non-churned customers (FP) were misclassified as churned, reflecting exceptional precision. However, 470 churned customers (FN) were misclassified as non-churned, which, while larger than FP, remains relatively small compared to the number of TP. This reinforces the overall effectiveness of the SMOTE-augmented model.

The SMOTE-augmented model achieved the highest F1 score and AUC, providing the best balance between specificity and sensitivity. It demonstrated high sensitivity and specificity, accurately identifying churned and non-churned customers while maintaining excellent precision. Although the model recorded a slightly higher number of false negatives compared to the original data scenario, this trade-off was acceptable given the overall improvements in predictive performance. These findings validate the use of SMOTE for churn prediction, enabling PT XYZ to prioritize at-risk customers effectively while maintaining reliable predictions for non-churned clients.

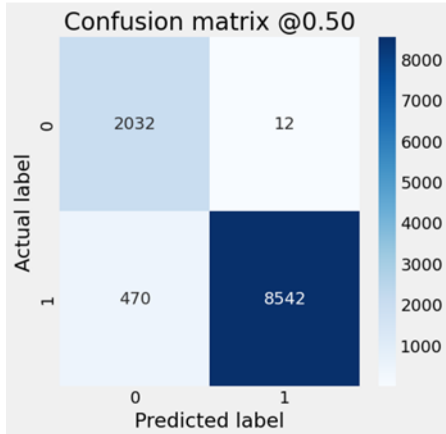


Figure 9. SMOTE confusion matrix result

The business implications derived from the predictions made by the MLP model are as follows:

a. True positives (TP)

This group consists of customers who have been accurately identified as likely to churn. Prompt and efficient interventions can help retain these customers, thereby maintaining revenue and strengthening customer relationships. The model's ability to identify a large number of true positives is advantageous for focusing retention efforts on the right individuals.

b. False positives (FP)

This group represents customers who are incorrectly predicted to churn. While targeting these customers with incentives or discounts could lead to unnecessary spending, the impact is minimal due to the extremely low number of false positives predicted by the model. This makes the cost of inaccuracy negligible in practical terms.

c. True negatives (TN)

This group comprises customers who are correctly predicted not to churn. Since no intervention is required for this group, the company can avoid unnecessary retention activities, allowing resources to be allocated more effectively. The high number of true negatives predicted by the model demonstrates its ability to reduce redundant efforts.

d. False negatives (FN)

This group includes customers who churned but were incorrectly predicted to stay. This is a critical area because it directly correlates with lost revenue due to customer attrition. Retention strategies should prioritize this group to prevent further losses and ensure higher accuracy in predicting churn risks.

By focusing on these interpretations, businesses can optimize their resource allocation, targeting customers most likely to churn while avoiding unnecessary expenditures on those likely to stay. This strategic focus can improve retention efficiency and profitability. The integration of the MLP model with marketing intelligence in this study demonstrates the potential for predictive analytics to inform business decision-making. These findings align with the study by Sana et al. (2022), which emphasizes the importance of combining predictive modeling with marketing intelligence. This integrated approach offers a strong foundation for building comprehensive client retention strategies.

Insights from the EDA, including activity patterns and demographic trends, further enhance the ability to develop tailored engagement and retention strategies. By concentrating efforts on the most vulnerable customers, the predictive accuracy of the model supports effective resource allocation and improved churn control, ensuring better outcomes for PT XYZ's customer retention initiatives.

5. Conclusion

This study demonstrated the successful integration of Exploratory Data Analysis (EDA) and Multilayer Perceptron (MLP) modeling to predict and address customer churn at PT XYZ. The SMOTE-augmented MLP model performed exceptionally well, with an AUC of 99% and an F1 score of 97%, offering strong insights for churn prediction. EDA identified ten key churn-related factors, offering valuable insights for retention strategies. To effectively address churn, PT XYZ should develop tailored retention strategies based on segmentation by age, geography, and tenure. For instance, loyalty programs targeting new customers (0-2 years) and personalized promotions for high-risk groups will encourage long-term customer loyalty. Additionally, targeted communication based on customer behavior insights should be implemented through PT XYZ's mobile application, allowing personalized offers and increasing customer engagement. Regular satisfaction surveys and tracking metrics like payment timeliness will also help refine these strategies and ensure continuous improvement. While the MLP model showed strong performance, addressing false negatives remains important for future improvements. Future work should focus on enhancing feature engineering and exploring additional algorithms to improve recall. For smaller ISPs, simpler models like logistic regression or decision trees can be employed as alternatives, with cloud-based solutions enabling cost-effective predictive analytics.

In conclusion, this study provides actionable, data-driven insights to enhance customer retention strategies at PT XYZ. By adopting the model's predictions and implementing these tailored strategies, PT XYZ can improve customer loyalty, reduce churn, and maintain a competitive advantage in the Indonesian ISP market. A hybrid approach combining the simplicity of statistical models with the power of machine learning will provide robust and practical insights for both technical and business teams.

6. Acknowledgements

The main contributors to this research are Gema Persada Arihta, Tanika D. Sofianti, and Win Sukardi. The authors would like to express their deepest gratitude to Swiss German University for providing the tools and atmosphere that enabled our research to succeed.

References

Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. Journal of Big Data, 6(1). https://doi.org/10.1186/s40537-019-0191-6

Aulia Triyafebrianda, H., & Windasari, N. A. (2022). Factors influence customer churn on internet service providers in Indonesia. TIJAB (The International Journal of Applied Business), 6(2), 134–144.

Bogaert, M., & Delaere, L. (2023). Ensemble Methods in Customer Churn Prediction: A Comparative Analysis of the State-of-the-Art. Mathematics, 11(5), 1137. https://doi.org/10.3390/math11051137

Christiadi, H., & Sule, E. T. (2018). The Influence of Distinctive Capability and Innovation Management Towards the Performance of ISPs in Indonesia. Journal of Advanced Research in Law and Economics, 1212(34), 1212–1221. https://doi.org/10.14505/jarle.v9.4(34).06

da Silva, I. N., Hernane Spatti, D., Andrade Flauzino, R., Liboni, L. H. B., & dos Reis Alves, S. F. (2017). Artificial Neural Networks (First Edition). Springer International Publishing. https://doi.org/10.1007/978-3-319-43162-8

Edwine, N., Wang, W., Song, W., & Ssebuggwawo, D. (2022). Detecting the Risk of Customer Churn in Telecom Sector: A Comparative Study. Mathematical Problems in Engineering, 2022, 1–16. https://doi.org/10.1155/2022/8534739

Fareniuk, Y., Zatonatska, T., Dluhopolskyi, O., & Kovalenko, O. (2022). Customer churn prediction model: a case of the telecommunication market. ECONOMICS, 10(2), 109–130. https://doi.org/10.2478/eoik-2022-0021

Faritha Banu, J., Neelakandan, S., Geetha, B. T., Selvalakshmi, V., Umadevi, A., & Martinson, E. O. (2022). Artificial Intelligence Based Customer Churn Prediction Model for Business Markets. Computational Intelligence and Neuroscience, 2022, 1–14. https://doi.org/10.1155/2022/1703696

Gaivoronski, A. A., Nesse, P. J., & Erdal, O. B. (2017). Internet service provision and content services: paid peering and competition between internet providers. NETNOMICS: Economic Research and Electronic Networking, 18(1), 43–79. https://doi.org/10.1007/s11066-017-9114-x

Geiler, L., Affeldt, S., & Nadif, M. (2022). A survey on machine learning methods for churn prediction. International Journal of Data Science and Analytics, 14(3), 217–242. https://doi.org/10.1007/s41060-022-00312-5

Gu, X., Angelov, P. P., & Soares, E. A. (2020). A self-adaptive synthetic over-sampling technique for imbalanced classification. International Journal of Intelligent Systems, 35(6), 923–943. https://doi.org/10.1002/int.22230

Jahan, I., & Farah Sanam, T. (2022). An Improved Machine Learning Based Customer Churn Prediction for Insight and Recommendation in E-commerce. 2022 25th International Conference on Computer and Information Technology (ICCIT), 1–6. https://doi.org/10.1109/ICCIT57492.2022.10054771

Lalwani, P., Mishra, M. K., Chadha, J. S., & Sethi, P. (2022). Customer churn prediction system: a machine learning approach. Computing, 104(2), 271–294. https://doi.org/10.1007/s00607-021-00908-y

Peng, K., Peng, Y., & Li, W. (2023). Research on customer churn prediction and model interpretability analysis. PLOS ONE, 18(12), e0289724. https://doi.org/10.1371/journal.pone.0289724

Plangger, K., Grewal, D., de Ruyter, K., & Tucker, C. (2022). The future of digital technologies in marketing: A conceptual framework and an overview. Journal of the Academy of Marketing Science, 50(6), 1125–1134. https://doi.org/10.1007/s11747-022-00906-2

Prabadevi, B., Shalini, R., & Kavitha, B. R. (2023). Customer churning analysis using machine learning algorithms. International Journal of Intelligent Networks, 4, 145–154. https://doi.org/10.1016/j.ijin.2023.05.005

Riaz Sadia, Mushtaq Arif, & Kaur Maninder Jeer. (2021). Information Visualization : Perception and Limitations for Data-Driven Designs. In PredictiveAnalysis (1st ed., p. 23). CRCPress.

Ribeiro, H., Barbosa, B., Moreira, A. C., & Rodrigues, R. G. (2024). Determinants of churn in telecommunication services: a systematic literature review. Management Review Quarterly, 74(3), 1327–1364. https://doi.org/10.1007/s11301-023-00335-7

Salma, N., & Aprianingsih, Ph. D. A. (2021). Customer Churn Analysis: Analyzing Customer Churn Determinants on an ISP Company in Indonesia. Buletin Pos Dan Telekomunikasi, 29–40. https://doi.org/10.17933/bpostel.2021.190103

Sana, J. K., Abedin, M. Z., Rahman, M. S., & Rahman, M. S. (2022). A novel customer churn prediction model for the telecommunication industry using data transformation methods and feature selection. PLOS ONE, 17(12), e0278095. https://doi.org/10.1371/journal.pone.0278095