CUSTOMER CHURN PREDICTION IN TELECOM: ENHANCING CUSTOMER RETENTION STRATEGIES

SUPIRATHAN THANABORIPAT

A THEMATIC PAPER SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF MANAGEMENT COLLEGE OF MANAGEMENT MAHIDOL UNIVERSITY 2023

Thematic paper
entitled
CUSTOMER CHURN PREDICTION IN TELECOM: ENHANCING CUSTOMER RETENTION STRATEGIES

was submitted to the College of Management, Mahidol University
for the degree of Master of Management
on
28 April 2024

Supirathan Th.
....
Ms. Supirathan Thanaboripat
Candidate

....
Assoc. Prof. Randall Shannon,
Ph.D.
Advisor

....
Assoc. Prof. Astrid Kainzbauer,
Ph.D.
Chairperson

....
Assoc. Prof. Vichita Ractham,
Ph.D.
Dean
College of Management
Mahidol University

....
Asst. Prof. Manjiri Kunte,
Ph.D.
Committee member

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to all those who contributed to the completion of this thematic paper.

First and foremost, I extend my sincerest appreciation to Assoc. Prof. Randall Shannon and Dr. Murtaza Haider, my supervisors, whose guidance, encouragement, and invaluable insights significantly shaped the direction of this work. Their unwavering support and mentorship were instrumental in navigating the complexities of this project.

I am also grateful to the committee, Assoc. Prof. Astrid Kainzbauer and Asst. Prof. Manjiri Kunte, for their constructive feedback and suggestions. Their critical perspectives helped fulfill the completion of this paper.

Finally, I would like to thank all the friends at CMMU for their cheerful support along this journey until the completion of this paper.

Thank you to everyone who contributed to this endeavor in any capacity. Your support is deeply appreciated.

Supirathan Thanaboripat

CUSTOMER CHURN PREDICTION IN TELECOM: ENHANCING CUSTOMER RETENTION STRATEGIES

SUPIRATHAN THANABORIPAT 6549092

M.M. (MARKETING AND MANAGEMENT)

THEMATIC PAPER ADVISORY COMMITTEE: ASSOC. PROF. RANDALL SHANNON, Ph.D., ASSOC. PROF. ASTRID KAINZBAUER, Ph.D., ASST. PROF. MANJIRI KUNTE, Ph.D.

ABSTRACT

This thematic paper examines customer churn dynamics in the telecommunications industry, focusing on logistic regression as a predictive tool. Through an analysis of factors influencing churn and actionable insights for mitigation, the study aims to enhance customer retention and organizational resilience.

Key findings reveal the significance of factors like contract terms, internet service, and total charges in predicting churn. Promising model performance, with an accuracy of approximately 78.82%, underscores the importance of proactive retention strategies.

Recommendations include service improvements, promotion of long-term contracts, and targeted engagement initiatives to mitigate churn risk. While acknowledging limitations, the study suggests avenues for future research to refine models and explore customer motivations.

In conclusion, actionable insights provided can empower telecom companies to navigate churn challenges and drive business growth in a competitive landscape.

KEY WORDS: TELECOM INDUSTRY/ CUSTOMER CHURN/ LOGISTIC REGRESSION/ PREDICTIVE ANALYTICS/ RETENTION STRATEGIES

23 pages

CONTENTS

CONTENTS (Cont.)

LIST OF TABLES

LIST OF FIGURES

CHAPTER I
INTRODUCTION

1.1 Background and Context

The telecommunications industry stands as one of the most dynamic and rapidly evolving sectors globally, serving as a cornerstone of modern connectivity and communication. The telecom services market is expected to expand at a CAGR of 6.5% from 2023 to 2030, driven by the demand for advanced connectivity and next-generation technologies (Grandreviewresearch, 2022). With the advent of digitalization, the proliferation of mobile devices, and the ubiquitous presence of high-speed internet, telecom companies operate within a landscape marked by intense competition, technological disruption, and evolving consumer preferences (Moflix, 2020; World Economic Forum, 2016; Digital McKinsey, 2016). Customer churn, the phenomenon where subscribers terminate their relationship with a service provider and switch to a competitor, poses a significant threat to the sustainability and profitability of telecom companies (Actian Corporation, 2020).

Within this context, the quest to acquire and retain customers reigns supreme. Telecom companies are facing the challenge of working to fulfil the evolving connectivity needs of consumers with the balance of profitability (Deloitte, 2024).

1.2 Significance of Customer Churn Analysis

The significance of customer churn analysis in the telecom industry cannot be overstated. Churn not only represents a loss of revenue but also entails substantial costs associated with customer acquisition (Singh et al. 2024), service provisioning, and marketing efforts. Moreover, in an era characterized by heightened competition and commoditization of services, retaining existing customers assumes paramount importance as a means of sustaining profitability and fostering long-term relationships (Singh et al. 2024). Furthermore, acquiring new customers' demands more significant marketing resources compared to retaining existing ones, rendering it a more challenging task (Brandusoiu, 2013; Forbes Business Council, 2022).

In today's dynamic economy, minimizing customer churn is crucial for success as it indicates the health of businesses amidst increasing disruption and saturation (Phumchusri & Amornvetchayakul, 2024). Recognizing the significance of anticipating churn, telecom companies emphasize robust customer relationship management (CRM) strategies (Hemalatha & Amalanathan, 2019) to enable tailored marketing campaigns and retention strategies to retain at-risk customers, fostering loyalty and reducing churn rates. Ultimately, churn analysis acts as a proactive tool for preserving revenue, strengthening customer relationships, and enhancing competitiveness in the telecom sector.

1.3 Objectives of the Study

Against this backdrop, this thematic paper endeavors to delve into the intricacies of customer churn within the telecom industry, with a particular focus on leveraging logistic regression as a predictive analytics tool. The overarching objectives of this study are delineated as follows:

1. Investigate Factors Contributing to Customer Churn: To explore the significant factors contributing to customer churn within the telecom sector, encompassing both intrinsic (e.g., service quality, pricing) and extrinsic (e.g., competitive offerings, market dynamics) determinants.

2. Provide Insights and Recommendations for Churn Mitigation: To distill actionable insights and recommendations for telecom operators aimed at mitigating churn, enhancing customer retention strategies, and fostering sustainable growth in a competitive marketplace.

By interrogating the nuances of customer churn and harnessing the predictive power of logistic regression, this study seeks to illuminate pathways toward improved customer retention and organizational resilience within the telecom industry.

CHAPTER II
LITERATURE REVIEW

2.1 The Definition of Customer Churn

Customer churn, a crucial metric in subscription-based industries, refers to customers leaving a company's services, including unsubscribing or switching to competitors (Başarslan et al., 2023). Wagh et al. (2024) describe it as the number of current customers expected to discontinue their association within a specific period. Understanding churn's nuances is vital across sectors like telecommunications, insurance, and banking.

Churn goes beyond service termination, involving changes in behavior or declining Customer Lifetime Value (CLV) (Amornvetchayakul & Phumchusri, 2020). In telecom, churn includes contract terminations, cancellations, or migration to competitors, often due to dissatisfaction or competitive offers from competing providers (Wanchai, 2017). Churn is a complex phenomenon studied by researchers and experts, influenced by various factors.

This research defines churn as customers ending their relationship with a company through actions like unsubscribing or switching to competitors.

2.2 The factors leading to customer churn in telecom sector

In a study by Wagh et al. (2024), examining features affecting churn in telecom, several key factors emerged. Firstly, the Contract-Month-To-Month feature was identified as having the highest importance score (0.517) in the decision tree model and the second highest score (0.1245) in the random forest model. This suggests that customers with month-to-month contracts are more prone to churn compared to those with longer-term contracts such as those who subscribe for the 2 year's contract. Secondly, Total Charges ranked prominently, with the second highest importance score (0.104) in the decision tree model and the fourth highest score (0.07) in the random forest model.

Higher total charges were associated with a greater likelihood of churn. Thirdly, the Tenure Group feature was highlighted, indicating that customers with shorter tenure are more predisposed to churn than those with longer tenure. The authors categorized tenure into six groups ranging from 1 to 72 months, similar to the finding in the research of Wanchai (2017) that found customers with less than 2 years' tenure were more likely to churn. Lastly, Internet Service emerged as a significant predictor, with the fourth highest importance score (0.0517) in the decision tree model and the third highest score (0.07693) in the random forest model. Specifically, customers with fiber optic internet service were more inclined to churn compared to those with DSL or no internet service.

Additionally, there are more of categorical features to predict customer churn in the research of Wagh et al. (2024), the paper found that Online Security emerged as one of the most crucial features for predicting customer churn, showcasing a high negative correlation with churn rates. This categorical attribute signifies whether customers avail online security services, with those utilizing such services exhibiting lower churn probabilities compared to those without. In the Random Forest classifier model, Online Security attained an importance score of 0.03977, ranking seventh among all features, while in the Decision Tree classifier model, it achieved a score of 0.0795, placing it fourth in importance. Payment Method, another categorical attribute indicating how customers pay for services, also demonstrated significance in churn prediction. Notably, customers utilizing electronic checks displayed a higher likelihood of churn compared to other payment methods, indicating a moderate positive correlation with churn rates. In the Random Forest classifier model, Payment Method attained an importance score of 0.02819, ranking fourteenth among all features, whereas in the Decision Tree classifier model, it garnered a score of 0.0419, positioning it sixth in importance. Tech Support, indicating the availability of technical assistance services, emerged as another critical predictor of churn. Customers with access to tech support exhibited lower churn rates, highlighting a high negative correlation with churn probability. In the Random Forest classifier model, Tech Support achieved an importance score of 0.02491, ranking sixteenth among all features, while in the Decision Tree classifier model, it obtained a score of 0.00696, placing it ninth in importance. Finally, Paperless Billing, reflecting customer preferences for electronic billing methods, displayed a lower importance in churn prediction compared to other features. While customers opting for paperless billing showed a slightly higher churn likelihood, this feature exhibited a low positive correlation with churn rates. In the Random Forest classifier model, Paperless Billing attained an importance score of 0.001289, ranking nineteenth among all features, and in the Decision Tree classifier model, it garnered a score of 0.004727, marking it as the least important feature.

These findings underscore the multifaceted nature of churn prediction in the telecom industry, highlighting the varied impacts of features on customer churn behavior. To summarize, to predict customer churn in the telecom sector, this research will use these features as followed

Contract: Month-to-Month (highest importance), One Year, Two Year

Internet Service: Fiber Optic, DSL

Online Security: No, Yes

Payment Method: Electronic Check

Tech Support: No

Paperless Billing

Total Charges

Tenure

2.3 Methodologies in Customer Churn Analysis

The literature review highlights various methodologies used in previous studies for customer churn prediction in the telecom sector. Wagh et al. (2024) employed Random Forest, KNN, and Decision Tree Classifier alongside survival analysis and Cox proportional hazard model, achieving high accuracy with Random Forest. Wanchai (2017) utilized decision trees, logistic regression, and neural networks, with decision trees performing the best, followed by logistic regression which was ranked the second in the accuracy. Phumchusri & Amornvetchayakul (2024) employed logistic regression, support vector machine, decision tree, and random forest. Amornvetchayakul & Phumchusri (2020) applied logistic regression, support vector machine, decision tree, and random forest. Dalvi et al. (2016) proposed decision trees and logistic regression to predict churn behavior. Jain et al. (2020) compared Logistic Regression and Logit Boost. Considering the approaches of the similar previous studies, logistic regression is widely chosen due to its interpretability, simplicity, and reasonable effectiveness in churn prediction. However, the result of each research showed inferior performance of logistic regression. Hence, this research will predict customer churn in the telecom sector using relevant empirical features by optimizing logistic regression, considering its widespread use and acceptable level of performance, to enhance its accuracy.

2.4 Application of Logistic Regression in Churn Prediction

Among the myriad methodologies employed in churn prediction, logistic regression stands out as a widely adopted and versatile approach (Vafeiadis et al., 2015). Logistic regression is a statistical technique used to model the probability of a binary outcome, making it well-suited for predicting churn events (Vafeiadis et al., 2015; Radosavljevik et al., 2010; Gursoy, 2010). By estimating the likelihood of churn based on a set of predictor variables, such as customer demographics, usage patterns, and satisfaction metrics, logistic regression enables telecom operators to identify at-risk subscribers and prioritize retention efforts accordingly (Brandusoiu, 2013). Here are the research papers adopting logistic regression for churn analysis.

Previous research extensively supports the use of logistic regression in predicting customer churn in the telecommunications industry. Dalvi et al. (2016) utilized logistic regression alongside decision trees, employing data preprocessing techniques and maximum likelihood estimation. Jain et al. (2020) compared logistic regression with Logit Boost, finding Logit Boost superior in certain metrics. However, both studies emphasize logistic regression's significance due to its interpretability and proven performance in churn prediction tasks.

Despite the growing prominence of alternative algorithms in churn prediction within the telecommunications sector, logistic regression remains a viable and valuable tool for predictive modeling due to its simplicity, interpretability, and reasonable accuracy (Vafeiadis et al., 2015; Radosavljevik et al., 2010).

CHAPTER III
METHODOLOGY

3.1 Data Collection and Preprocessing

The dataset used for this study was sourced from Kaggle, specifically the Telco Customer Churn dataset provided by Blastchar (Kaggle, 2024). This dataset contained information on telecom customers including demographics, services subscribed to, contract details, billing information, and churn status. The dataset comprised both numerical and categorical variables in the forms of integer, string, and Boolean, providing a comprehensive basis for churn analysis.

The analysis included 30 features due to the dataset's inherent complexity, which comprises approximately 20 primary features (columns or factors). Each feature encompassed various conditions, such as binary options like yes/no, which necessitated the expansion to 30 distinct features for comprehensive analysis.

Prior to analysis, the dataset underwent preprocessing to ensure data quality and suitability for logistic regression modeling. This included handling missing values, encoding categorical variables, and scaling numerical features as necessary. Additionally, outliers and irrelevant variables were identified and addressed to enhance the model's predictive performance.

3.2 Feature Selection

Based on the literature review and previous research findings, relevant features for predicting customer churn in the telecom sector were identified. These included

1. Contract type (Month-to-Month, One Year, Two Year) (string)
2. Internet Service (Fiber Optic, DSL) (string)
3. Online Security (Yes, No) (string)
4. Payment Method (Electronic Check) (string)

5. Tech Support (Yes, No) (string)
6. Paperless Billing (True, False) (Boolean)
7. Total Charges (integer)
8. Tenure (integer)

These features encompassed both categorical and numerical variables known to influence telecom customer churn behavior (True, False; Boolean) which acted as the target variable (dependent variable) in this research. However, all of the provided features of the dataset were utilized for the analysis.

3.3 Logistic Regression Modeling

Logistic regression was chosen as the primary predictive modeling technique due to its interpretability, simplicity, and established effectiveness in churn prediction tasks within the telecommunications industry. The logistic regression model estimated the probability of churn based on the selected features, allowing for the identification of at-risk subscribers. The data analysis was done with Python programming language in Jupyter Notebook.

The logistic regression model was trained on the preprocessed dataset, with appropriate evaluation metrics such as accuracy, precision, recall, and F1-score used to assess its performance.

3.4 Model Evaluation and Interpretation

Once the logistic regression model was trained and evaluated, its results were interpreted to gain insights into the factors driving customer churn. Coefficients associated with each feature provided information on their impact on churn probability, enabling telecom operators to prioritize retention efforts accordingly. Furthermore, a visualization such as feature importance plot was utilized to communicate the model's findings effectively.

Overall, the methodology outlined in this chapter provided a systematic approach to investigating and predicting customer churn within the telecom industry using logistic regression modeling. Through data-driven analysis and interpretation, valuable insights and recommendations could be derived to support informed decision-making and enhance customer retention strategies.

CHAPTER IV
DATA ANALYSIS AND RESULTS

In this section, the author presents the analysis of the churn prediction model based on logistic regression. The dataset consists of 7,043 records, with features such as gender, senior citizenship, tenure, services subscribed, contract terms, and payment methods, among others. The target variable, "churn," denotes whether a customer has terminated their service ("Yes") or not ("No").

4.1 Logistic Regression Model Performance

Upon training and testing the logistic regression model, the obtained results are as followed:

Accuracy: The accuracy of the model in predicting churn is approximately 78.82%. This indicates that the model correctly predicts churn status for nearly 79% of the customers in the test dataset.

4.1.1 Confusion Matrix:

Table 4.1 Confusion Matrix

Churn	Predicted No Churn	Predicted Churn
Actual No	916	117 (False Positive)
Actual Yes	181 (False Negative)	193

The confusion matrix illustrates the model's performance in predicting churn and non-churn instances. Out of 1407 instances, the model correctly classified 1109 cases (916 no churn, 193 churn) and misclassified 298 cases (117 actual churn predicted as no churn, and 181 actual no churn predicted as churn).

4.1.2 Classification Report:

Table 4.2 Classification Report

	Precision	**Recall**	**F1-Score**	**Support**
No Churn	0.84	0.89	0.86	1033
Churn	0.62	0.52	0.56	374
Overall	0.78	0.79	0.78	1407

The classification report presents the performance metrics of the logistic regression model in predicting customer churn within the telecom sector.

4.1.3 Precision

No Churn (Class "No"): The precision for predicting customers who do not churn is 0.84. This indicates that out of all instances predicted as not churn, 84% were correctly classified, while the remaining 16% were falsely predicted as churn.

Churn (Class "Yes"): The precision for predicting churn is 0.62. This suggests that out of all instances predicted as churn, 62% were correctly classified as churn, while the remaining 38% were falsely predicted as not churn.

4.1.4 Recall (Sensitivity)

No Churn (Class "No"): The recall for predicting customers who do not churn is 0.89. This implies that 89% of actual non-churn instances were correctly identified by the model, while the remaining 11% were incorrectly classified as churn.

Churn (Class "Yes"): The recall for predicting churn is 0.52. This indicates that 52% of actual churn instances were correctly identified by the model, while the remaining 48% were incorrectly classified as not churn.

4.1.5 F1-Score

No Churn (Class "No"): The F1-score for predicting customers who do not churn is 0.86. This metric considers both precision and recall, providing a balanced measure of the model's performance for the "No" class.

Churn (Class "Yes"): The F1-score for predicting churn is 0.56. This indicates the harmonic mean of precision and recall for the "Yes" class, reflecting the model's overall performance in predicting churn instances.

4.1.6 Support

No Churn (Class "No"): The support for the "No" class is 1033, indicating the number of instances of non-churn in the test dataset.

Churn (Class "Yes"): The support for the "Yes" class is 374, reflecting the number of instances of churn in the test dataset.

In summary, our logistic regression model demonstrates promising performance in predicting customer churn, with an accuracy of approximately 78.82%. However, there is room for improvement, particularly in enhancing the model's ability to correctly classify churn instances.

4.2 Feature Importance

Table 4.3 Feature Importance

Features	Coefficient
InternetService_Fiber optic	1.038682
TotalCharges	0.688224
StreamingMovies_Yes	0.348088
StreamingTV_Yes	0.329858
PaperlessBilling	0.274834
MultipleLines_Yes	0.273244
PaymentMethod_Electronic check	0.26265
SeniorCitizen	0.259502
MultipleLines_No phone service	0.203975
Partner	0.049663
DeviceProtection_Yes	0.030609
gender	-0.03175

Table 4.3 Feature Importance (Cont.)

Features	Coefficient
PaymentMethod_Mailed check	-0.06433
OnlineBackup_Yes	-0.07126
PaymentMethod_Credit card (automatic)	-0.13919
OnlineSecurity_No internet service	-0.14156
TechSupport_No internet service	-0.14156
StreamingMovies_No internet service	-0.14156
StreamingTV_No internet service	-0.14156
DeviceProtection_No internet service	-0.14156
InternetService_No	-0.14156
OnlineBackup_No internet service	-0.14156
PhoneService	-0.2038
Dependents	-0.21749
MonthlyCharges	-0.30802
TechSupport_Yes	-0.37082
OnlineSecurity_Yes	-0.39656
Contract_One year	-0.75681
Contract_Two year	-1.2929
tenure	-1.4442

Selecting the most significant features from a logistic regression model involves considering the magnitude (absolute value) of the coefficients assigned to each feature. Features with larger absolute coefficients typically have a greater impact on the prediction of the target variable (in this case, churn). With 0.5 coefficient threshold to show the important features, Figure 4.1 displays feature importance, highlighting coefficients higher than 0.5 based on the analysis.

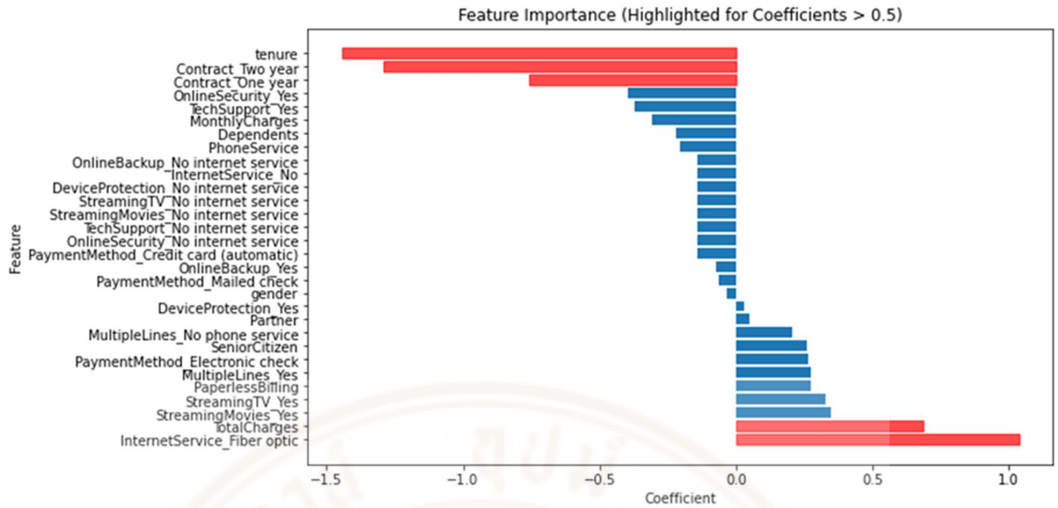

Figure 4.1 Feature Importance (Highlighted for Coefficients > 0.5)

According to Figure 4.1, the features with the largest absolute coefficients are

- InternetService_Fiber optic: Customers subscribed to fiber optic internet service are more likely to churn, as indicated by the positive coefficient value of 1.038682.
- TotalCharges: Higher total charges are associated with a higher likelihood of churn, as reflected by the positive coefficient of 0.688224.
- Contract_One year and Contract_Two year: Customers with longer contract terms (one year and two years) are less likely to churn, with coefficients of -0.756808 and -1.292903, respectively.
- Tenure: Longer tenure is associated with lower churn probability, as reflected by the negative coefficient of -1.444198.

4.3 Insights

- The availability of fiber optic internet service contributes significantly to churn, suggesting potential dissatisfaction with service quality or pricing.
- Customers with higher total charges are more likely to churn than those with lower charges.

- Contract duration and tenure play crucial roles in reducing churn, highlighting the importance of long-term customer relationships and commitment.

4.4 Recommendations

- Service Improvement: Addressing issues related to internet service quality can help reduce churn rates.
- Promote Long-term Contracts: Encourage customers to opt for longer contract terms to foster loyalty and reduce churn risk.
- Targeted Retention Strategies: Implement targeted retention strategies for customers with higher total charges by offering personalized incentives, privileges, proactive support, and value-added services to enhance their satisfaction and loyalty, thereby reducing churn risk and preserving long-term revenue.

By leveraging these insights, telecom companies can develop proactive strategies to mitigate churn and foster long-term customer relationships, ultimately driving business growth and profitability.

CHAPTER V
DISCUSSION

The findings of this research align with previous studies in several key aspects. Firstly, the identification of significant predictors such as contract terms, total charges, internet service type, and customer tenure resonates with existing literature on churn prediction in the telecom sector (Wagh et al., 2024; Wanchai, 2017). The emphasis on these factors underscores their consistent impact on customer retention dynamics.

Moreover, the application of logistic regression as a predictive modeling technique reflects a broader trend observed in prior research (Dalvi et al., 2016; Jain et al., 2020). Moreover, our study's results reaffirm logistic regression's utility, albeit with room for improvement in predicting churn cases.

In the realm of churn prediction, the significance of false negatives cannot be overstated. False negatives occur when the model incorrectly predicts customers with churn intentions as non-churners. This is particularly consequential because it hampers the company's ability to leverage the predictive power of machine learning models effectively. If customers who are likely to churn are erroneously classified as non-churners, the company misses the opportunity to intervene and implement targeted retention strategies, such as sending personalized marketing campaigns or offering incentives to prevent churn, negatively impacting the long-term customer retention rate, resulting in revenue loss.

5.1 Novel Insights

While reaffirming established findings, this study also offers novel insights into churn dynamics within the telecom sector. The identification of fiber optic internet service as one of the significant predictors shows fluctuation in customer preferences and behaviors. These insights provide valuable guidance and further observation for telecom companies seeking to refine their retention strategies in response to changing market dynamics, as well as to oversee their internet service quality and enhance it.

5.2 Limitations

Despite its contributions, this study is subject to several limitations that warrant acknowledgment. Firstly, the reliance on a single predictive modeling technique (logistic regression) may limit the generalizability of the findings. Future research could explore alternative modeling approaches to gain a more comprehensive understanding of churn dynamics.

Additionally, the use of secondary data sources may introduce inherent biases and limitations inherent in the dataset. While efforts were made to mitigate these issues through rigorous preprocessing and validation procedures, the potential for residual biases cannot be entirely eliminated.

Furthermore, the scope of this study was limited to a specific geographic region or telecom provider, potentially limiting the generalizability of the findings to broader contexts. Future research could adopt a multi-provider or cross-national approach to enhance the external validity of the findings.

5.3 Future Studies

Building on the insights and limitations identified in this study, a few avenues for future research emerge

5.3.1 Mitigate False Negatives through Model Refinement

Future studies should prioritize model refinement efforts aimed at reducing false negatives in churn prediction. This may involve fine-tuning algorithm parameters, optimizing feature selection processes, or exploring alternative modeling techniques that prioritize sensitivity over specificity. By prioritizing the identification of potential churners, companies can enhance their ability to proactively intervene and retain valuable customers.

5.3.2 Qualitative Research

Incorporating qualitative research methods such as interviews or focus groups can complement quantitative analyses by providing richer insights into customer motivations and preferences. By combining quantitative and qualitative approaches, researchers can develop more holistic models of churn behavior that capture both rational and emotional factors.

5.4 Recommendations

Building on these insights, the following recommendations are proposed for telecom companies seeking to mitigate churn and foster long-term customer relationships.

5.4.1 Service Improvement

Addressing issues related to internet service quality is significant in reducing churn rates. Telecom companies should prioritize investments in infrastructure upgrades and internet service enhancements to ensure competitive offerings that meet or exceed customer expectations.

5.4.2 Promote Long-term Contracts

Encouraging customers to opt for longer contract terms through incentives and rewards can foster loyalty and reduce churn risk. By highlighting the benefits of extended commitments, such as discounted rates or exclusive perks, telecom companies can incentivize customers to maintain long-term relationships with the brand.

5.4.3 Targeted Retention Strategies

Implement targeted retention strategies for customers with higher total charges by offering personalized incentives, privileges, proactive support, and value-added services to enhance their satisfaction and loyalty, thereby reducing churn risk and preserving long-term revenue.

5.4.4 Invest in Customer Engagement Strategies

Telecom companies should invest in proactive customer engagement strategies aimed at preempting churn. This may involve implementing targeted outreach campaigns, offering personalized incentives, or providing proactive customer support to address potential churn drivers proactively. By fostering meaningful interactions with customers and addressing their needs and concerns promptly, companies can enhance overall satisfaction and loyalty, mitigating the risk of churn and maximizing long-term revenue potential.

5.5 Conclusion

This research contributes to the understanding of churn prediction in the telecom sector by identifying key predictors and reaffirming the utility of logistic regression. The significance of minimizing false negatives in churn prediction is emphasized, highlighting its impact on revenue and customer retention. Novel insights into churn dynamics, particularly regarding fiber optic internet service and senior citizens, underscore the need for tailored retention strategies. While limitations exist, future research avenues are identified. Overall, the recommendations provided offer actionable strategies for telecom companies to mitigate churn and drive business growth.

REFERENCES

Actian. (2020). Reducing customer churn in the telecom industry with connected data. Retrieved from https://www.actian.com/blog/data-analytics/reducing-customer-churn-in-the-telecom-industry-with-connected-data/

Amornvetchayakul, P., & Phumchusri, N. (2020, April). Customer churn prediction for a software-as-a-service inventory management software company: A case study in Thailand. In *2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA)* (pp. 514-518). IEEE. https://doi.org/10.1109/ICIEA49774.2020.9102099

Başarslan, M., Aslıhan, Ü. N. A. L., & KAYAALP, F. (2023). A novel robust LSTM model on Churn and comparative analysis versus deep learning, machine learning and ensemble learning: Telecommunications Case. https://doi.org/10.21203/rs.3.rs-3421473/v1

Blastchar. (n.d.). Telco Customer Churn [Data set]. Kaggle. Retrieved from https://www.kaggle.com/datasets/blastchar/telco-customer-churn

Brandusoiu, I. (2013). Predicting Churn in Mobile Telecommunications Industry. *Acta Technica Napocensis Electronics and Telecommunications, 54*(3), 7-12.

Dalvi, P. K., Khandge, S. K., Deomore, A., Bankar, A., & Kanade, V. A. (2016). Analysis of customer churn prediction in telecom industry using decision trees and logistic regression. In *2016 Symposium on Colossal Data Analysis and Networking (CDAN)* (pp. 1-4). DOI: 10.1109/CDAN.2016.7570883.

Deloitte. (n.d.). *2024 telecom industry outlook: Five key trends shaping the telecom landscape.* Retrieved from https://www2.deloitte.com/us/en/pages/technology-media-and-telecommunications/articles/telecommunications-industry-outlook.html

Forbes Business Council. (2022, December 12). Customer retention versus customer acquisition. Retrieved from https://www.forbes.com/sites/forbesbusinesscouncil/2022/12/12/customer-retention-versus-customer-acquisition/?sh=7ea430211c7d

REFERENCES (Cont.)

Grand View Research. (n.d.). Telecom Services Market Size, Share & Trends Analysis Report by Service Type (Mobile Data Services, Machine-To-Machine Services), By Transmission (Wireline, Wireless), By End-use, By Region, And Segment Forecasts, 2023-2030 [PDF]. Retrieved from https://www.grandviewresearch.com/industry-analysis/global-telecom-services-market#

Gursoy, U.T. (2010). Customer churn analysis in telecommunication sector. *Istanbul University Journal of the School of Business, 39*, 35-49.

Hemalatha, P., & Amalanathan, G. (2019). A hybrid classification approach for customer churn prediction using supervised learning methods: Banking sector. In *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)* (pp. 1-6). https://doi.org/10.1109/ViTECoN.2019.8899692

Jain, H., Khunteta, A., & Srivastava, S. (2020). Churn prediction in telecommunication using logistic regression and logit boost. *Procedia Computer Science, 167*, 101-112. https://doi.org/10.1016/j.procs.2020.03.187

McKinsey & Company. (n.d.). How telecom companies can win in the digital revolution. Retrieved from https://www.mckinsey.com/~/media/McKinsey/Industries/Technology%20Media%20and%20Telecommunications/Telecommunications/Our%20Insights/How%20telecom%20companies%20can%20win%20in%20the%20digital%20revolution/How-telecom-companies-can-win-in-the-digital-revolution.pdf

MoFlix Group. (n.d.). Digital disruption in the telecom industry. Retrieved from https://www.moflixgroup.com/blog/digital-disruption-in-the-telecom-industry

Phumchusri, N., & Amornvetchayakul, P. (2024). Machine learning models for predicting customer churn: A case study in a software-as-a-service inventory management company. International *Journal of Business Intelligence and Data Mining, 24*, 74-106.

REFERENCES (Cont.)

Radosavljevik, D., van der Putten, P., & Larsen, K. K. (2010). The impact of experimental setup in prepaid churn prediction for mobile telecommunications: What to predict, for whom and does the customer experience matter? *Transactions on Machine Learning and Data Mining, 3*(2), 80–99.

Singh, P. P., Anik, F. I., Senapati, R., Sinha, A., Sakib, N., & Hossain, E. (2024). Investigating customer churn in banking: A machine learning approach and visualization app for data science and management. *Data Science and Management, 7*(1), 7-16. https://doi.org/10.1016/j.dsm.2023.09.002

Vafeiadis, T., Diamantaras, K., Sarigiannidis, G., & Chatzisavvas, K. (2015). A Comparison of Machine Learning Techniques for Customer Churn Prediction. *Simulation Modelling Practice and Theory, 55*. https://doi.org/10.1016/j.simpat.2015.03.003

Wagh, S. K., Andhale, A. A., Wagh, K. S., Pansare, J. R., Ambadekar, S. P., & Gawande, S. H. (2024). Customer churn prediction in telecom sector using machine learning techniques. *Results in Control and Optimization, 14,* 100342. https://doi.org/10.1016/j.rico.2023.100342

Wanchai, P. (2017). Customer churn analysis : A case study on the telecommunication industry of Thailand. 2017 12th International Conference for Internet Technology and Secured Transactions (ICITST), 325-331.

World Economic Forum. (2016). Why digital innovation in telecommunications needs to be about more than just connectivity. Retrieved from https://www.weforum.org/agenda/2016/09/why-digital-innovation-in-telecommunications-needs-to-be-about-more-than-just-connectivity/