CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

CUSTOMER BEHAVIOUR ANALYSIS AND CHURN PREDICTION IN COMMUNICATIONS AND AI

Er. Mandeep Singh[1], Er. Prabhjot Kaur[2], Dr. Inderpreet Kaur[3]
[1,2,3]University Institute of Computing Chandigarh University, Mohali, India
er.mandeep.it@gmail.com, prabjot906@gmail.com, inder5preet@gmail.com

Abstract **–** Customer churn has emerged as a key worry for the telecommunications industry, which is a crucial sector of the global economy. This study uses a multimodal method that incorporates churn predictors, account information implications, and demographic analysis to examine telecommunications customer turnover. Important discoveries highlight the complex relationship between client tenure and churn and show how contract modifications affect retention.

Our approach consists of a thorough review of the literature, a study of churn prediction techniques and indicators, and an examination of the architecture and governance of the telecom industry. Utilizing a wide range of Kaggle datasets, the research focuses on predictive modeling with XGBoost and achieves 79.4% test accuracy.

Tailored contract offers, focused client interaction, and incorporating predictive modeling into CRM are examples of practical consequences. Telecom firms may use these data to develop customer loyalty, reduce attrition, and successfully negotiate a changing sector.

It is imperative to recognize the limits of the study, including data constraints and model assumptions, in order to appropriately interpret the findings. Future research should investigate cross-industry comparative studies, dynamic modeling taking temporal trends into account, and qualitative findings. We hope to strengthen the telecom industry's agility and resilience as we continue to explore the intricacies of customer turnover.

Keywords **– BIG DATA ANALYTICS , CUSTOMER CHURN, CONTRACT VARIATIONS , DYNAMIC MODELING, PREDICTIVE MODELING, XGBOOST.**

I. INTRODUCTION

The telecommunications sector is vital to the world economy because of its fast technical advancement and dynamic nature. Customer churn is one of the biggest issues facing telecommunications firms as technology gets more and more ingrained in our daily lives. In order to sustain market share, profitability, and customer pleasure, telecom businesses now need to anticipate and understand customer attrition.

A. *The Value of Big Data Analysis in the Telecommunications Sector*

Big data has completely changed the way organization’s function, and the telecom industry is no exception. This industry generates a vast amount of data at a rapid pace, which poses both benefits and

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

challenges. Telecom businesses can use big data analytics to glean valuable insights from this abundance of data, which facilitates strategic planning and well-informed decision-making.

1) *The Issue: Churn Forecasting and Its Importance* The problem of customers quitting their services, or customer churn, is a serious threat to the long-term viability of telecom operations. The first steps in creating proactive retention tactics are determining the causes of churn and creating prediction models that work. Churn prediction *helps* businesses reduce customer attrition while also optimizing resource allocation, customizing marketing campaigns, and improving overall service quality.

B. *Questions and Objectives for the Research*

This study begins a thorough investigation of customer churn prediction in the telecom sector in light of the aforementioned difficulties. The main goals consist of:
Examining the Existing Literature: To create a solid base, synthesize knowledge from important research like Ahmad et al. [1], Huang et al. [2], Umayaparvathi and Iyakutti [3], Hashmi et al. [4], and Verbeke et al. [7].
2) Analyzing Methodologies and Metrics: According to Umayaparvathi and Iyakutti's survey [3], this involves looking into the datasets, methodologies, and metrics used in customer churn prediction.
3) *Examining Governance and Architecture:* As noted by Keshavarz et al. [12], Kastouni and Lahcen [9], and Zahid et al. [8], this section explains the governance and utility of big data analytics in the telecommunications sector.
4) Offering Useful information: By addressing the nuances of churn prediction and drawing on a variety of viewpoints and statistics, this section provides telecom firms with useful information [20].
5) By tackling these goals, the study aims to make a significant contribution to the area by offering a comprehensive perspective on customer churn prediction in the telecom industry.

C. Research Deficits

Customer churn, which is the loss of a customer to a rival, presents a great opportunity for businesses in a variety of sectors. Even though a lot of study has been done on consumer churn prediction, there are still important areas that need more investigation. This examination of objectives to fill in those deficiencies by specializing in:
Generalizability of Churn Models: An established method for predicting churn that is applicable to various sectors.
It is necessary to identify tactics that improve version overall performance when used in a variety of customer bases and corporate environments.

II. LITERATURE REVIEW

In (2019) used device mastering techniques on large-scale information structures to anticipate customer attrition in the telecom sector. Their research, which was published in the Journal of Big Data,

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024
ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

emphasizes how big data, device science, and superior analytics all work together to forecast customer attrition. The analysis plays a pivotal role in molding our methodology, harmonizing with the state-of-the-art telecom analytics landscape, and directing our use of machine learning for potent churn prediction in a similar setting[1].
Customer churn prediction is paramount within the telecommunications sector for boosting customer retention strategies. Huang, Kechadi, and Buckley (2012) make contributions to this domain with the aid of using professional structures and programs. Their attention lies in leveraging predictive analytics to foresee patron churn, with an emphasis on algorithmic techniques. The have a look at, featured in Expert Systems with Applications,highlights the significance of accurate predictions in a dynamic telecom environment. The exploration of consumer behavior patterns and the software of professional structures provide insights precious for our research, guiding us within the selection of appropriate methodologies and predictive fashions[2].
This work serves as a valuable resource, providing insights into the various procedures and key issues necessary for knowledge and addressing patron churn within the telecom enterprise.

Umayaparvathi and Iyakutti's survey (2016) provides a thorough overview of customer churn prediction inside the telecom region. Published inside the International Research Journal of Engineering and Technology (IRJET), the survey appreciably covers datasets, methodologies, and metrics employed in existing studies on telecom consumer churn prediction.
Kastouni and Lahcen's (2022) latest artworks explore the realm of big data analytics within the telecom area. The authors of the paper "Big Data Analytics in Telecommunications: Governance, Architecture, and Use Cases," which was published in the King Saud University-Computer and Information Sciences Journal, carefully examine the governance, architectural issues, and practical use cases associated with big data analytics in the telecommunications industry. This research significantly advances our understanding of how to use big data analytics for improved performance and decision-making in the telecoms industry by addressing crucial elements including governance frameworks and providing real-world applications[9].
In "Big Data in Telecommunication Operators: Data, Platform, and Practices," the authors delve into the complex world of massive statistics projects within telecom companies. This picture, which was published in the Journal of Communications and Information Networks, offers insights into the various components of huge data, including the data itself, the frameworks supporting its analysis, and the useful applications in the telecommunications sector. Through the effective use of huge data, the study addresses these dimensions and offers insightful perspectives for enhancing operational performance and decision-making methodologies in the telecommunications industry[10].
The authors of the paper "The Value of Big Data Analytics Pillars in Telecommunication Industry" explore the importance of big data analytics pillars in the telecom sector. This study, which was published in Sustainability, highlights the critical role that several big data analytics pillars have in boosting sustainability in the telecommunications sector. The report offers significant insights for telecommunication companies looking to use big data for strategic decision-making and sustainable practices by examining the costs associated with these analytics pillars[11].Using massive facts, Huang

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

et al. (2015)'s work "Telco Churn Prediction with Big Data" provides a thorough investigation of churn prediction in the telecommunications industry. The research, which was published in the Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, focuses on using massive records to effectively predict churn in the telecom sector. With a focus on the use of huge information, the writers explore advanced strategies and procedures, offering invaluable insights to the field of patron churn prediction[12].In the field of telecommunications, the study "Intelligent Big Data Analysis to Design Smart Predictor for Customer Churn in Telecommunication Industry" significantly advances the use of astute big statistics analysis to forecast customer attrition. The study, which is included in the electronic book "Big Data and Smart Digital Environment" by Springer International Publishing, focuses on creating a sophisticated predictor for customer attrition through the application of sophisticated massive data analytics techniques. The authors contribute to the emerging subject of intelligent information evaluation for telecom agencies by providing insights and techniques to improve forecasting abilities in the context of customer churn[13].

III. DATA COLLECTION AND PREPROCESSING

A. Dataset Description:

The dataset used in this analysis, which was obtained from Kaggle, focuses on forecasting consumer behavior to enhance customer retention strategies. The data, which comes from IBM Sample Data Sets, includes several customer attributes listed in the column Metadata

B. About the Dataset:

Context:"Anticipate actions to retain customers. You are able to create customized patron retention packages by analyzing all relevant buyer data." [Model Data Sets from IBM]

C. *Content:*

Every row represents a buyer, while the columns represent different characteristics of the buyer, such as: Customers who departed in the last month (Churn column). Signed up for services included phone, multiple lines, internet, online security, online backup, device security, tech support, streaming TV, and movies.Customer account data, including contract specifics, pricing strategy, monthly costs, and overall spending, during the duration of their consumer dating. Demographic details, including gender, age range, and whether or not dependents and partners are present.

D. Inspiration:

The goal of the dataset is to make it easier to investigate client churn prediction models in the telecom sector. With an average monthly churn charge of 1.92%, churn is a significant challenge for this quarter. As such, the dataset is an invaluable tool for researching and devising strategies to address this challenge.[19]

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

E. Prediction and Exploratory Data Analysis (EDA):

NumPy, Pandas, Seaborn, and Matplotlib are among the libraries that the Python code makes use of in order to provide an overview of the facts exploration method. Loading the dataset, identifying information types, handling missing values, and formatting particular variables appropriately are the first steps.

F. Key Steps:

An analysis is conducted on the relationship between unique variables and client churn.
We investigate the distribution and association of the following demographics: gender, age, dependents, and associates.
In order to identify styles, customer account information are shown, including contract details and tenure.
The ability connections between factors and their consequences for customer attrition are somewhat illuminated by the EDA and prediction sections. This entire method sets the basis for later modeling and the identification of critical factors affecting turnover.

G. Exploratory Data Analysis (EDA):

1) Synopsis of Results from EDA:

   Many important conclusions were drawn from the exploratory information assessment (EDA) of the telecom customer dataset. The main observations are listed below:

2) Overview of Data:
The dataset comprises a multitude of client-related features, such as demographic information, specifics about used offerings, settlement information, and churn fame.
"Churn," the goal variable, indicates if a customer has churned or not.

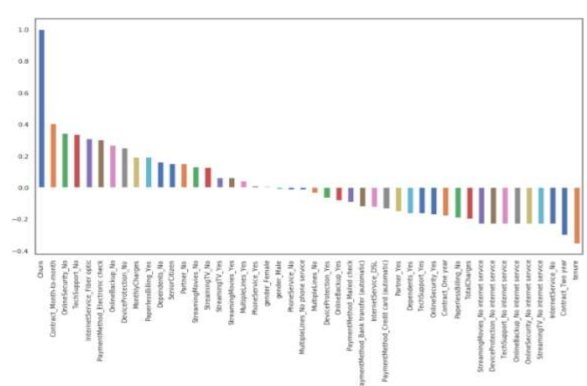

Figure 1: Correlation of Churn with other variables

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

3) Data Cleaning:
The "TotalCharges" column contained 11 missing values, which were eliminated from the dataset. Dummy variables were developed and categorical variables were transformed to numerical format.

4) Demographic Analysis: Gender Distribution: The dataset's male and female consumer distributions equal.

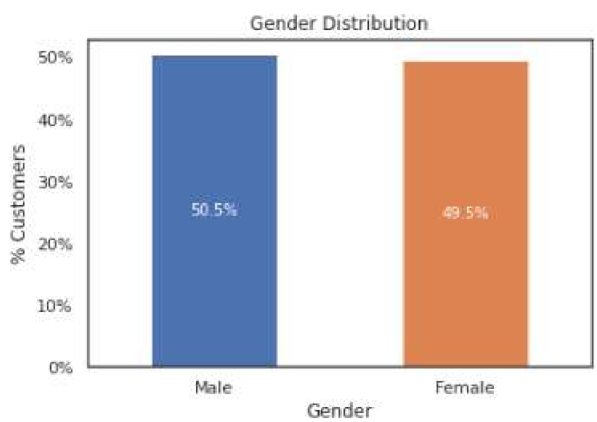

Figure 2: Gender Distribution in Data sets

Senior Citizens: The proportion of elderly clients is only 16%.
Partner and Dependents: Thirty percent of clients have dependents, and fifty percent have a partner. Subsequent investigation showed that only roughly half of clients who have partners also have dependents.

5) Details about Customer Accounts:
Tenure: According to the tenure distribution, a large number of consumers only have brief interactions with the telecom provider, but some have been doing so for up to 72 months. Contracts: The majority of clients choose month-to-month agreements, which are followed by one- and two-year agreements.

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

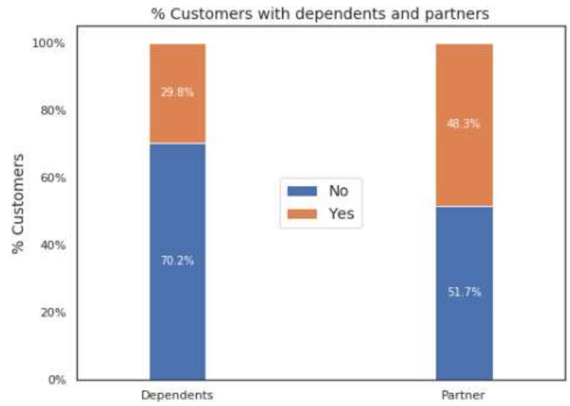

Figure 3: percentage of customers with dependents and partners.

Tenure vs. Contract: Longer-term customers typically have more tenure, indicating greater commitment.

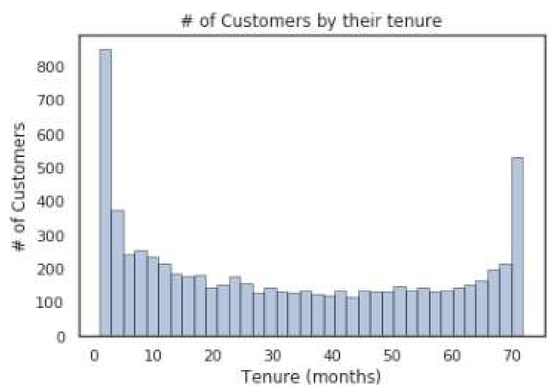

Figure 4: percentage of customers by tenure

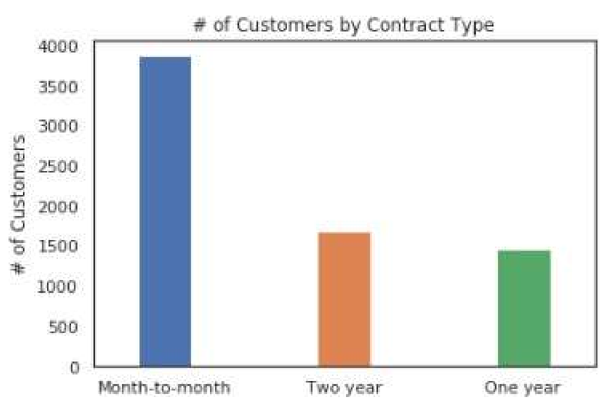

Figure 5: Customers by different contact types

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

1) Distribution of Services:

Phone Service, Multiple Lines, Internet Service, and other services were analyzed to understand their prevalence among customers.

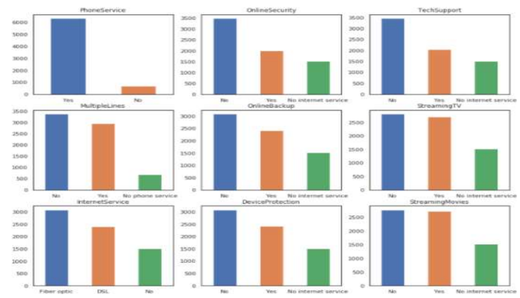

Figure 6: different types of services used by customers

2) Monthly and Total Charges:

A positive correlation was observed between Monthly Charges and Total Charges, indicating that as monthly bills increase, the total charges also increase.

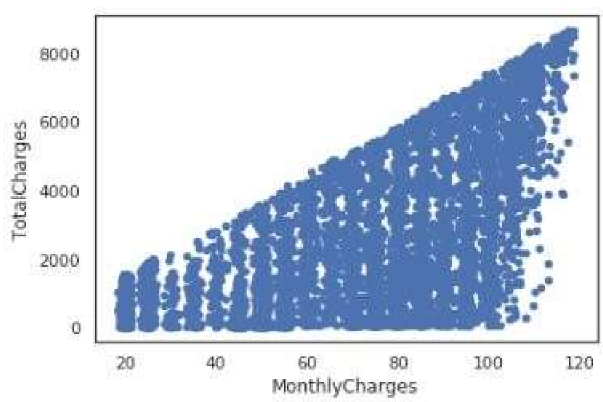

Figure 7: Relation ship between monthly and total charges

3) Churn Analysis:

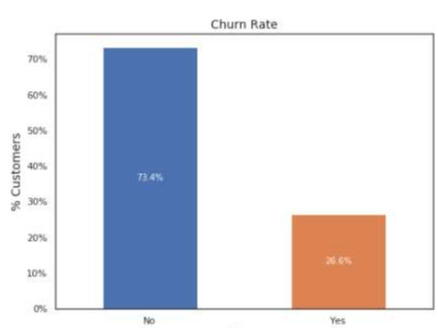

Figure 8: Churn Rate Distribution: Percentage of Customers Churning and Not Churning

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

"Correlation of Variables with Churn"
Churn Rate: The dataset's distribution was skewed because 74% of the consumers did not churn.
Tenure vs. Churn: Telecom customers with longer tenures seem to be less likely to churn.
9) Churn vs. Monthly/Total rates: Customers that discontinue service frequently incur higher monthly rates than their non-churning counterparts.

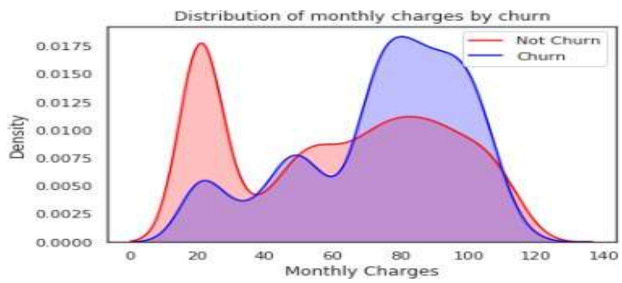

Figure 9: Distribution of monthly charges by churn

Plots of kernel density showing the monthly and total charge distribution for both churn and non-churn consumers.
These results offer insightful information about the traits and behavior of customers that could affect churn. In the telecom sector, churn rates can be decreased by developing strategies and identifying relevant predictors with the aid of more data and modeling.

H. Data Preprocessing:

To guarantee the accuracy and dependability of the data utilized in the analysis, data preprocessing is an essential step. The dataset was cleaned and prepared using the following steps:

1. Managing Absent Values:
2. found and examined the dataset's missing values. imputation or removal techniques were used, depending on the kind and volume of missing data.
3. made ensuring that any imputation techniques used were reported in a transparent manner.

2) Data Type Conversions:
Verified that the data types of the variables were compatible with the tools and techniques used for analysis.
carried out the requisite conversions to category or numeric data types confirmed that the data's meaning and integrity were preserved throughout the alterations.

3) Encoding:
Use appropriate encoding techniques to address categorical variables. Used techniques such as label

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

encoding or one-hot encoding in accordance with the variables' properties. analyzed the effects of encoding decisions on model performance and downstream analysis. Through careful attention to detail, data type conversions, and appropriate encoding techniques, the dataset was cleaned up and prepared for further study. The effectiveness and dependability of any insights obtained from the data are enhanced by these preprocessing procedures.

IV. RESULTS:

A. Analysis of Demographics:

1) Analysis of Gender
2) Upon doing a thorough analysis of the gender distribution of consumers, no noteworthy distinctions were found in the proportions of customers who had partners and dependents. Gender-wise, the distribution did not change. Moreover, there was no appreciable gender difference in senior citizen status according to the analysis. There were no discernible gender variations in the distribution of older citizens.

2) Customer Account Information:

a) Tenure Analysis: A broad range was shown by the customer tenure histogram, with a concentration of customers with tenures of less than one month. On the other hand, several of them had been in place for about 72 months. The diversity of client contracts may be the cause of this tenure variation, indicating that contract conditions may have an impact on customer retention.

3) Analysis of Predictor Variable (Churn):
a) Churn Distribution: 74% of consumers did not churn, according to an analysis of the churn variable. The data showed a bias in favor of non-churn cases, highlighting the necessity of addressing skewness in modeling to prevent false negative results.

b) Variation of Turn by Important Variables:
I. CHURN VS. TENURE:
The churn vs. tenure plot revealed that non-churning clients typically had longer tenures with the telecom provider. This finding raises the possibility of a relationship between client retention and attrition, which calls for more research.

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

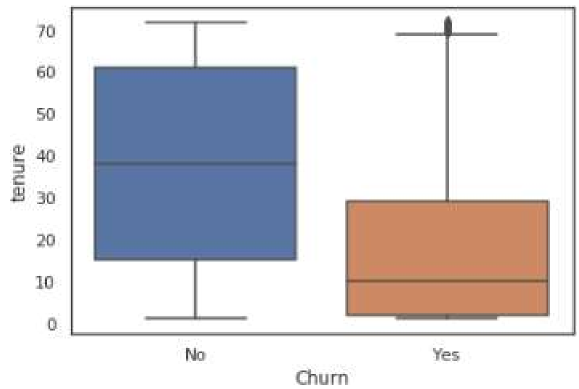

Figure 12: Box plots showing the tenure and churn relationship.

4) Performance of Machine Learning Models:
Hyperparameter optimization was used to fine-tune the XGBoost model, leading to the optimal set of parameters: {'subsample': 0.8, 'n_estimators': 200,'min_child_weight': 2,'max_depth': 3.'gamma': 0, 'learning_rate': 0.1, 'colsample_bytree': 0.8}.
The final model showed a test accuracy of 0.794, demonstrating its ability to reasonably anticipate client attrition.

V. IMPLICATIONS:

For telecom businesses looking to improve customer retention and lower attrition, the study's findings have important real-world ramifications. The following observations provide doable tactics to deal with the issues raised:
A. Customizing Contract Offerings:
The significance of customizing contract offerings is underscored by the observed variation in client tenure, which is impacted by different contract terms. By optimizing contract arrangements to better suit client preferences, telecom businesses may be able to increase customer loyalty and lower the risk of customer attrition.

B. Targeted Customer Engagement:
Telecom companies can concentrate on targeted customer engagement tactics because of the association between tenure and churn. Longer client loyalty and lower churn rates may result from initiatives to improve customer connections, offer individualized services, and attend to specific demands over time.

C. Integration of Predictive Modeling:
The XGBoost model that was constructed has a test accuracy of 79.4%, making it a useful tool for churn prediction. Telecom businesses can proactively identify customers who are at risk of leaving by incorporating predictive modeling into their CRM systems. Through focused retention efforts, this

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024
ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

enables early intervention.
D. Customer Communication Strategies: A key component of customer loyalty and satisfaction is communication. Telecom businesses can improve their communication tactics by making sure that contract conditions are clear, advertising more services, and quickly responding to client complaints. By fostering a great customer experience, effective communication helps to mitigate variables that lead to client attrition.

E. Offering Incentives for Long-Term Commitments:
Telecom companies may offer incentives for long-term commitments in order to promote customer retention. A win-win situation can be created by loyalty programs, special offers, and advantages linked to longer contracts. This will increase customer loyalty and lower the risk of churn.

VI. LIMITATIONS:

Although this study offers insightful information, it is important to recognize the following restrictions, since they could affect how the results are interpreted and how broadly applicable they are:

A. Data Restrictions:
The study is dependent on the available dataset, and the features of the samples it includes determine how representative the dataset is. The conclusions may not be as broadly applicable due to variations in demographic parameters, geographic areas, or particular market situations that are not sufficiently covered in the dataset.

B. Temporal Dynamics:
Market trends, regulatory changes, and technology breakthroughs all have a quick impact on the telecom industry's dynamics. Our dataset's temporal structure might not adequately reflect recent advancements, which could limit the applicability of our conclusions to the state of the telecom industry today.

C. Modeling Assumptions:
During the modeling process, assumptions are established that affect how well the predictive model performs. Even though our XGBoost model shows good predicting abilities, it's important to understand that all predictive models simplify complex reality by nature and might not take into consideration all the factors impacting consumer behavior.

D. Inadequate Context:
We mainly examined quantitative variables in our investigation, omitting to look at qualitative factors. Although they are outside the purview of this study, customer attitudes, subjective experiences, and unobservable factors may have an impact on churn. The lack of qualitative data makes it more difficult to get a thorough grasp of customer motivations.

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

E. Inference of Causation:
Our study mainly identifies links between variables; correlation does not imply causation. Even if we find correlations, determining the causes necessitates more investigation and taking into account outside variables that were not included in this analysis.

F. Sampling Bias:
There is a chance that some customer categories in the dataset were overrepresented or underrepresented due to sampling bias. This might affect how broadly applicable our findings are, particularly if we try to apply our advice to a larger and more varied group of clients.

G. External influences:
Our study does not specifically take into account external influences like world events, the state of the economy, or disruptions unique to a particular industry. These elements may contribute to unexplained fluctuations in churn rates and have a considerable impact on consumer behavior.

VII. FUTURE WORK:

A. Investigating Qualitative Insights: In order to gather customer opinions, reviews, and sentiments, future research may focus on qualitative issues by conducting surveys, interviews, or sentiment analysis. Comprehending the affective factors influencing attrition should offer a more sophisticated viewpoint.

B. Dynamic Modeling: Because the telecom industry is dynamic, improving the accuracy of churn predictions could involve integrating time-collection evaluation and predictive modeling that takes changing market dynamics and temporal patterns into account. This can entail creating models that adjust to changing patterns of consumer behavior.

B. External Environmental Factors: Researching how external factors, such as changes in the economy, technical advancements, or disruptions specific to a business, affect customer attrition should yield more detailed data. Subsequent research endeavors may delve into the extent to which macro-level influences influence variations in turnover costs.

C.

D. Examination of Segmentation:
By segmenting the customer base according to demographics, usage patterns, or other pertinent criteria, a more detailed study could be conducted, revealing hidden patterns and variances in turnover behavior within excellent customer organizations. This should enable customized retention plans for exceptional groups.

VIII. CONCLUSION

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

The purpose of this study was to determine whether the machine learning algorithm XGBoost could be used to forecast subscriber upside in the telecom sector. With an acceptable test accuracy of 79.4%, the model identified the time frame, contract type, and monthly cost as the main factors influencing customer turnover. These findings support current studies on client retention tactics in telecommunications and offer insightful information.
There are restrictions, though. The particular data set that was employed may have limited the model's generalizability to other tasks. Additionally, the study concentrated on data that was readily quantitative, which would have overlooked the significance of qualitative criteria like customer happiness.
There are a few directions that future study should go in order to overcome these constraints and enhance the field of churn prediction studies. Examining the mechanisms that improve model generalizability across many industries is crucial first. Secondly, greater insights on churn behavior can be obtained by adding rich customer-firm interaction data, such as sentiment analysis from social media or customer service encounters. Third, strong assessment standards created especially for unbalanced churn datasets must be established. By filling up this research void, it might be possible for later research to create a customer churn forecasting model that is more accurate, scalable, and understandable. In the end, this will make it possible for companies in a variety of sectors to actively retain their clientele and guarantee sustained success.

REFERENCES

[1] A. Singh, N. Thakur, and A. Sharma, "A review of supervised machine learning algorithms," in 2016 IEEE 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 1310-1315, 2016.

[2] Ray and Susmita, "A quick review of machine learning algorithms," in 2019 IEEE International conference on machine learning, big data, cloud and parallel computing (COMITCon), pp. 35-39, 2019.

[3] Burrell and Jenna, "How the machine 'thinks': Understanding opacity in machine learning algorithms," Big data & society 3, no. 1, 2016.

[4] Gelman and Andrew, "Exploratory data analysis for complex models," Journal of Computational and Graphical Statistics 13, no. 4, pp. 755-779, 2004.

[5] Behrens and T. John, "Principles and procedures of exploratory data analysis," Psychological Methods 2, no. 2, pp. 131, 1997.

[6] Jebb, T. Andrew, S. Parrigon, and S. E. Woo, "Exploratory data analysis as a foundation of inductive research," Human Resource Management Review 27, no. 2, pp. 265-276, 2017.

[7] Io, H. N., and C. B. Lee, "Chatbots and conversational agents: A bibliometric analysis," in 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 215-219, IEEE, 2017.

[8] Adam, Martin, Michael Wessel, and Alexander Benlian, "AI-based chatbots in customer service and their effects on user compliance," Electronic Markets 31, no. 2, pp. 427-445, 2021.

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

[9] Nuruzzaman, Mohammad, and O. K. Hussain, "A survey on chatbot implementation in customer service industry through deep neural networks." in 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE), pp. 54-61. IEEE, 2018.

[10] Parab, Archana, S. Palkar, S. Maurya, and S. Balpande, "An Intelligent Career Counselling Bot A System for Counselling," International Research Journal of Engineering and Technology (IRJET) 4, no. 03, pp. 2325-2330, 2017.

[11] Nguyen, D. Cuong, N. H. D. Dinh, C. Pham-Nguyen, T. L. Dinh, and N. H. L. Nam, "ITCareerBot: A Personalized Career Counselling Chatbot," in Asian Conference on Intelligent Information and Database Systems, pp. 423-436. Springer, Singapore, 2022.

[12] L. H. Su, T. Dang-Huy, T. Thi-Yen-Linh, N. Thi-Duyen-Ngoc, L. Bao-Tuyen, and N. Ha-Phuong-Truc, "Development of an AI Chatbot to support admissions and career guidance for universities," International Journal of Emerging Multidisciplinary Research 4, no. 2, pp. 11-17, 2020.

[13] Zahour, Omar, A. Eddaoui, H. Ouchra, and O. Hourrane, "A system for educational and vocational guidance in Morocco: Chatbot E-Orientation," Procedia Computer Science 175, pp. 554-559, 2020.

[14] Smutny, Pavel, and P. Schreiberova, "Chatbots for learning: A review of educational chatbots for the Facebook Messenger," Computers & Education, vol. 151, 2020.

[15] Kohli, Bhaumik, T. Choudhury, S. Sharma, and P. Kumar, "A platform for human-chatbot interaction using python," in 2018 IEEE Second International Conference on Green Computing and Internet of Things (ICGCIoT), pp. 439-444, 2018.

[16] Yalçınkaya, Abdullah, İ. G. Balay, and B. Şenoğlu, "A new approach using the genetic algorithm for parameter estimation in multiple linear regression with long-tailed symmetric distributed error terms: An application to the Covid-19 data," Chemometrics and Intelligent Laboratory Systems, vol. 216, 2021.

[17] Patel, H. Harsh and P. Prajapati, "Study and Analysis of Decision Tree-Based Classification Algorithms," International Journal of Computer Sciences and Engineering, vol. 10, pp. 74-78, 2018.

[18] Schonlau, Matthias, and R. Y. Zou, "The random forest algorithm for statistical learning," The Stata Journal, vol. 20, no. 1, pp. 3-29, 2020.

[19] Cervantes, Jair, F. G. Lamont, L. R. Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," Neurocomputing, vol. 408, pp. 189-215, 2020.

[20] B. Tamizharasi, L. J. Livingston, and S. Rajkumar, "Building a medical chatbot using support vector machine learning algorithm," In Journal of Physics: Conference Series, vol. 1716, no. 1, p. 012059, IOP Publishing, 2020.

[21] Lee, Terri, T. Zhu, S. Liu, L. Trac, Z. Huang, and Y. Chen, "CASExplorer: A Conversational Academic and Career Advisor for College Students," in The Ninth International Symposium of Chinese CHI, pp. 112-116. 2021.

CAHIERS MAGELLANES-NS
Volume 06 Issue 2
2024

ISSN:1624-1940
DOI 10.6084/m9.figshare.2632599
http://magellanes.com/

[22] G. Dongre, F. Vadraiwala, J. Upadhyaya, D. Ghosal, "Web-Based Career Counselling System With Chatbot," International Research Journal of Modernization in Engineering Technology and Science, volume 03, pp. 571-579, 2021.

[23] V. Shreeram, N., and A. Muthukumaravel, "Student Career Prediction Using Machine Learning Approaches." In I3CAC 2021: Proceedings of the First International Conference on Computing, Communication and Control System, I3CAC 2021, 7-8 June 2021, Bharath University, Chennai, India, p. 444, European Alliance for Innovation, 2021.

[24] Deepa, Ashlin, P. Karlapati, M. R. Mulagondla, P. Amaranayani, and A. P. Toram, "An Innovative Emotion Recognition and Solution Recommendation Chatbot," in 2022 IEEE 8th International Conference on Advanced Computing and Communication Systems (ICACCS), vol. 1, pp. 1100-1105, 2022.