CUSTOMER CHURN PREDICTION

[1]V. Nagesh, [2]A. Sannikha, [3]A. Jaswanth, [4]B. Bharadwaj, [5]B. Shreya

[1]Associate Professor,[2,3,4,5]Sreyas Institute of Engineering and Technology

Abstract: Customer churn, the termination of customer relationships with a business or service, is a critical metric that profoundly impacts a company's success. Effectively managing churn not only prevents revenue loss but also provides a competitive advantage by boosting customer retention rates. This underscores the significance of robust churn management strategies in the realm of customer service. In the contemporary landscape of data-driven decision-making, various algorithms aim to address customer churn prediction. One noteworthy approach is the Voting Classifier algorithm. This machine learning model leverages an ensemble of diverse models, combining their predictions to generate a final output or class. The crux of its functionality lies in aggregating the individual models' opinions and determining the class with the highest probability among them. The Voting Classifier excels in situations where diverse models contribute unique insights, mitigating biases inherent in individual algorithms. This diversity enhances the overall predictive power and generalizability of the model. By considering multiple perspectives, the Voting Classifier adapts well to intricate patterns in customer behaviour, offering a nuanced understanding of potential churn indicators. For companies invested in optimizing customer retention, the adoption of the Voting Classifier underscores a commitment to sophisticated data analytics. Its ability to synthesize the strengths of various models positions it as a powerful tool for assessing and predicting customer churn. In the dynamic landscape of business, where customer relationships are paramount, leveraging advanced machine learning techniques like the Voting Classifier becomes imperative for staying ahead in the competition and fostering long-term customer loyalty.

Keywords: Customer churn, Machine Learning, Voting classifier, Gradient Boosting, Logistic Regression, Adaboost.

1. INTRODUCTION

In the dynamic landscape of business applications fuelled by machine learning, few challenges are as pivotal and impactful as the prediction of customer churn. At its core, the churn rate encapsulates the departure of individuals or items from a collective group over a defined period, emerging as a critical metric with the inherent potential to significantly influence profits and overall business success. This challenge is particularly pronounced in industries that are abundant with high-quality, up-to-date data, where solving the churn prediction problem transcends operational necessity, evolving into a strategic avenue for profit maximization.

Customer churn, intricately entwined with customer satisfaction, serves as a pivotal differentiator between content, loyal patrons, and those opting to discontinue a service. Nowhere is this more salient than in highly competitive sectors such as telecommunications, where an annual churn rate ranging from 15-20 percent prevails. The ability to predict and effectively manage customer churn assumes strategic significance in this landscape. Foreseeing which customers are likely to leave empowers corporations to concentrate their retention efforts on these "high-risk" clients, thereby establishing a proactive stance in customer relationship management. However, the ultimate objective stretches beyond mere retention; it encompasses the expansion of coverage areas and the cultivation of enduring customer loyalty.

The financial implications of customer churn are profound, with retaining existing customers proving to be far more cost-effective than acquiring new ones. Addressing churn becomes not just a defensive strategy to preserve market position but a proactive catalyst for growth and thriving in competitive landscapes. The churn rate, intricately woven into customer lifetime value modelling, guides the estimation of net profit contributed over the entire future relationship with a customer. This modelling involves assessing the percentage of discontinuity in subscriptions within a given timeframe, providing a comprehensive perspective on a customer's journey and potential value to the business.

Effectively predicting customer churn necessitates the deployment of robust algorithms, and amidst various options, the Voting Classifier algorithm stands out as a beacon of innovation and effectiveness. Widely recognized for its ability to enhance model performance by leveraging an ensemble of diverse models, the Voting Classifier emerges as a strategic solution addressing the complexities inherent in customer churn prediction. The primary aim of this predictive approach is to classify customers into two categories: churn and no churn. By doing so, businesses gain the capacity to strategically tailor their efforts towards retaining customers identified within the churn category, thus optimizing resource allocation for maximum impact.

In the landscape where customer satisfaction is synonymous with success, the adoption of sophisticated algorithms like the Voting Classifier becomes not just beneficial but imperative for businesses navigating the dynamic terrain of competition and growth. As we delve deeper into the intricacies of customer churn prediction, it becomes evident that this analytical pursuit is not merely about data; it's a strategic imperative for businesses seeking not just to survive but to thrive in an ever-evolving market, ensuring that customer relationships remain at the heart of sustainable success.

2. RESEARCH PROBLEM

Customer churn poses a myriad of challenges for businesses, including lost revenue streams, increased customer acquisition costs, diminished profitability, potential damage to brand reputation, and disruptions to long-term growth prospects. It necessitates a focus on understanding the underlying causes, developing effective retention strategies, and employing predictive models to mitigate its adverse effects.

2.1 OBJECTIVES AND GOALS

The primary objectives and goals of customer churn prediction are to anticipate and mitigate customer attrition by utilizing historical data and predictive analytics to identify at-risk customers, allowing businesses to implement targeted retention strategies. This helps in preserving revenue, enhancing customer satisfaction, increasing customer lifetime value, and ensuring long-term sustainability and growth for the organization.

3. LITERATURE SURVEY

Gupta M, Jain R [1]: "Ensemble Techniques in Customer Churn Prediction" Explored the effectiveness of ensemble techniques in predicting customer churn and Investigated various ensemble models, such as Random Forest and Gradient Boosting, for improved accuracy.

Li C, Sun J, [2]: "Deep Learning Approaches for Customer Churn Prediction" Investigated the application of deep learning approaches, including neural networks, in customer churn prediction. Explored how deep learning models handle complex patterns in customer behaviour data.

Smith A, Brown B [3]: "Temporal Aspects in Customer Churn Prediction" Focused on the temporal dynamics of customer churn. Explored how time-related features and trends impact the accuracy of churn prediction models.

Chen Y, Wang F [4]: "Customer Churn Prediction in E-commerce: A Case Study" Conducted a case study on customer churn prediction in the e-commerce sector. Analysed the unique challenges and opportunities for predicting churn in online retail environments.

Kim H, Kim G [5]: "Social Network Analysis for Customer Churn Prediction" Explored the incorporation of social network analysis in predicting customer churn. Investigated how social connections and interactions influence churn patterns.

Wang L, Zhang Y [6]: "Feature Engineering in Customer Churn Prediction" Investigated the role of feature engineering techniques in enhancing customer churn prediction models. Explored novel approaches for extracting meaningful features from diverse datasets.

Jiang P, Li T [7]: "Explainable AI in Customer Churn Prediction" Explored the importance of explainability in customer churn prediction models. Investigated methods to make machine learning models interpretable for business stakeholders.

Yang J, Liu Q [8]: "Cross-Industry Insights for Customer Churn Prediction" Conducted a cross-industry analysis of customer churn prediction strategies. Investigated commonalities and variations in churn patterns across different business sectors.

Xu W, Li H [9]: "Customer Churn Prediction in the Telecom Industry: A Comprehensive Review" Provided a comprehensive review specifically focused on customer churn prediction in the telecom industry. Analysed the unique challenges and solutions within the telecommunications sector.

Zhang S, Wu Z [10]: "Dynamic Models for Customer Churn Prediction" Explored dynamic models that adapt to changing customer behaviour over time. Investigated the effectiveness of models that account for evolving patterns in churn prediction.

Liao H, Chen Z [11]: "Machine Learning Pipelines for Scalable Customer Churn Prediction" Explored scalable machine learning pipelines for handling large-scale datasets in customer churn prediction. Investigated techniques for efficient feature extraction, model training, and deployment.

Zhu M, Guo Y [12]: "Customer Churn Prediction with Imbalanced Datasets" Addressed the challenge of imbalanced datasets in customer churn prediction. Investigated strategies for handling class imbalance and improving predictive performance.

3.1 EXISTING SYSTEM

Companies currently rely on the basic demographic data to predict the customer churn. The basic demographic data has its limitations in predicting churn. This may not accurately reflect the customer behaviour as it may not capture the nuances of individual customer behaviour or the dynamic factors influencing churn. The existing systems make no use of customer behaviour data.

3.1.1 DRAWBACKS:

- Rules for existing systems need to be manually defined and may not adapt to changing customer behavior.
- Existing systems do not leverage customer behavior data to provide accurate predictions of churn.
- Current systems have limited accuracy in predicting churn, leading to loss of revenue.

3.2 PROPOSED SYSTEM

The proposed system collects and analyse customer profile, usage, and engagement data to identify important factors that lead to churn. The voting classifier algorithm is used in the proposed system. The voting classifier algorithm is a machine learning model that trains on an ensemble of numerous models and predicts an output (class) based on their highest probability of chosen class as the output. The proposed system enables early identification of potential churners by analyzing their behavior data. The system automatically adapts to changing customer behaviors, leading to improved accuracy over time. The proposed system is simpler and easier to implement than rule-based systems.

In customer churn prediction, various algorithms are used as base models within the Voting Classifier ensemble to improve predictive performance. The Voting Classifier combines the predictions of multiple individual models to produce a final prediction.

Here are the algorithms that are used in the ensemble for voting classifier algorithm in the context for customer churn prediction are:

3.2.1 Gradient Boosting Algorithm:

Gradient Boosting, a powerful machine learning algorithm, operates through sequential training of weak learners, often decision trees, to progressively correct errors and enhance predictive accuracy. The algorithm minimizes a loss function that quantifies the disparity between predicted and actual values. Emphasizing instances with higher errors, it assigns weights to data points and employs gradient descent to update the model in the direction opposite to the gradient of the loss function. This iterative process, guided by a learning rate, results in an ensemble model that excels in capturing intricate patterns and relationships within the data, making Gradient Boosting particularly effective for customer churn prediction.

3.2.2 Logistic Regression:

Logistic Regression is a statistical method used for binary classification problems, such as predicting whether an email is spam or not, or whether a customer will churn or stay. Despite its name, logistic regression is a classification algorithm rather than a regression one. It models the probability that a given instance belongs to a particular category using the logistic function, also known as the sigmoid function. The logistic regression algorithm estimates the coefficients for the features in the input data, and the sigmoid function transforms the linear combination of these coefficients and features into a value between 0 and 1, representing the probability of belonging to the positive class. A decision threshold is then applied to classify instances into one of the two classes. Logistic Regression is widely used due to its simplicity, interpretability, and efficiency in binary classification tasks.

3.2.3 Adaboost:

AdaBoost, short for Adaptive Boosting, is an ensemble learning algorithm that combines the predictions of multiple weak learners to create a strong classifier. It is particularly effective in binary classification problems. AdaBoost assigns weights to data points and focuses on the misclassified instances in each iteration, giving them higher weights to correct their classification in the next round. It sequentially trains a series of weak learners, often decision trees with limited depth, and combines their predictions by assigning different weights to each learner's output. The final model aggregates the weak learners into a robust and accurate classifier. AdaBoost's strength lies in its ability to adapt and improve its performance by giving more emphasis to the challenging instances, ultimately leading to a powerful ensemble model.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

3.3 SYSTEM ARCHITECTURE:

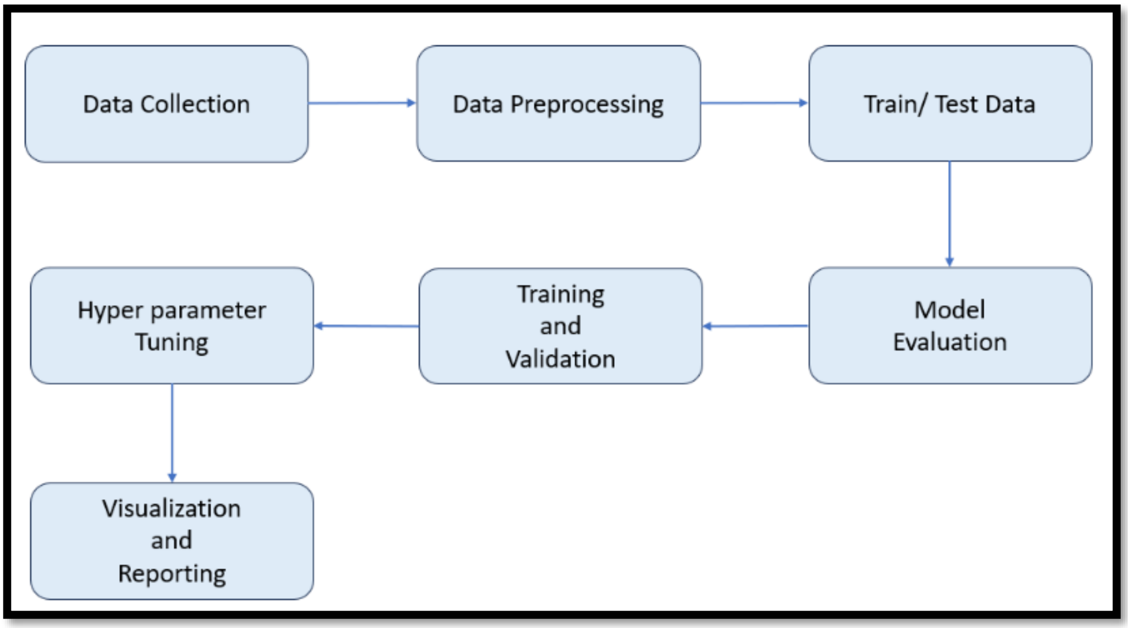

Fig-1: System Architecture

3.3.1 Data Collection: Here the data is collected from varioius data sources such as databases CRM systems, external data providers.

3.3.2 Data Preprocessing: Once the data is collected the data is cleaned and prepared for analysis. Common preprocessing tasks such as handling missing values, dealing with outliers, and encoding categorical variables are performed.

3.3.3 Train/Test Data: The dataset is typically split into two subsets- training and testing data. The training dataset is used to train the machine learning model, while the testing dataset is used to evaluate its performance. The data split represents the overall distribution of data.

3.3.4 Hyper Parameter Tuning: Hyperparameter tuning is the process of optimizing the hyperparameters of a machine learning model to improve its performance. Hyperparameters are settings or configurations that are not learned from the data but are set prior to training the model. Properly tuning these hyperparameters can significantly impact the model's performance and generalization.

3.3.5 Training and Validation: Model training is an iterative process. The training dataset is used to train the machine learning model, assess model performance and mitigate overfitting.

3.3.6 Model Evaluation: The trained model is evaluated using the testing dataset to assess its performance. Common evaluation metrics for churn prediction include accuracy, precision, recall, F1-score, and ROC AUC.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

3.3.7 Visualization and Reporting: Visualization tools and techniques can help in interpreting and communicating the results. You can create visualizations such as ROC curves, confusion matrices, and feature importance plots. A final report summarizing the model's performance, insights, and recommendations for reducing churn should be generated to guide business decisions.

3.4 RESULTS

3.4.1 Dataset

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection	TechSupp
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	
5	9305-CDSKC	Female	0	No	No	8	Yes	Yes	Fiber optic	No	...	Yes	
6	1452-KIOVK	Male	0	No	Yes	22	Yes	Yes	Fiber optic	No	...	No	
7	6713-OKOMC	Female	0	No	No	10	No	No phone service	DSL	Yes	...	No	
8	7892-POOKP	Female	0	Yes	No	28	Yes	Yes	Fiber optic	No	...	Yes	
9	6388-TABGU	Male	0	No	Yes	62	Yes	No	DSL	Yes	...	No	

10 rows × 21 columns

Fig-2: Sample dataset

3.4.2 Churn Distribution:

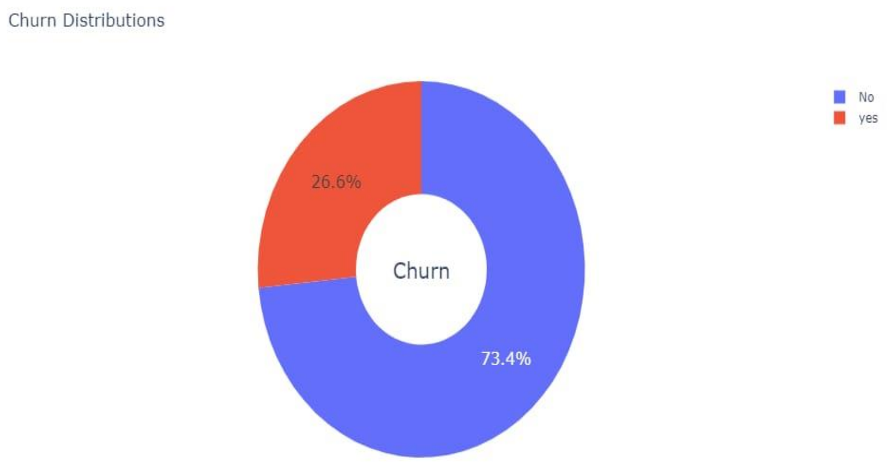

Fig-3: 26.6 % of customers switched to another firm.

3.4.3 Churn Distribution with respect to Gender:

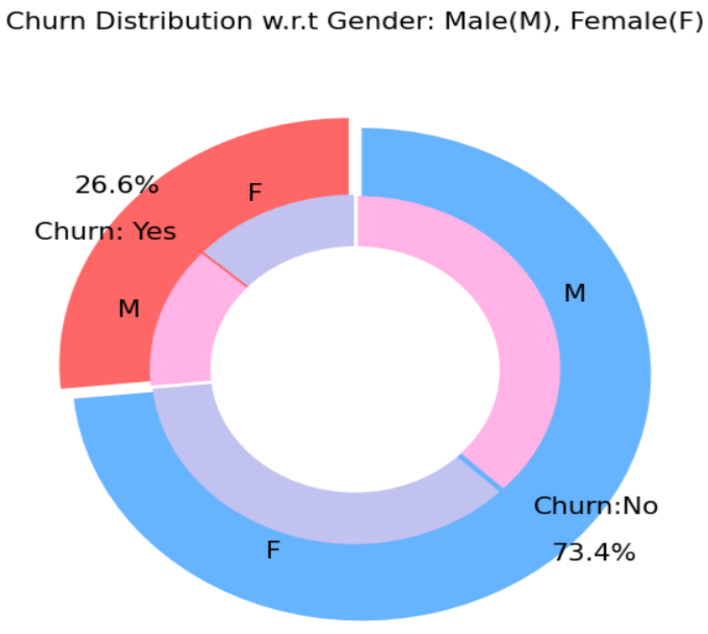

Fig-4: Churn distribution w.r.t Gender

There is negligible difference in customer percentage/count who changed the service provider. Both genders behaved in similar fashion when it comes to migrating to another service provider.

3.4.4 Customer Contract Distribution:

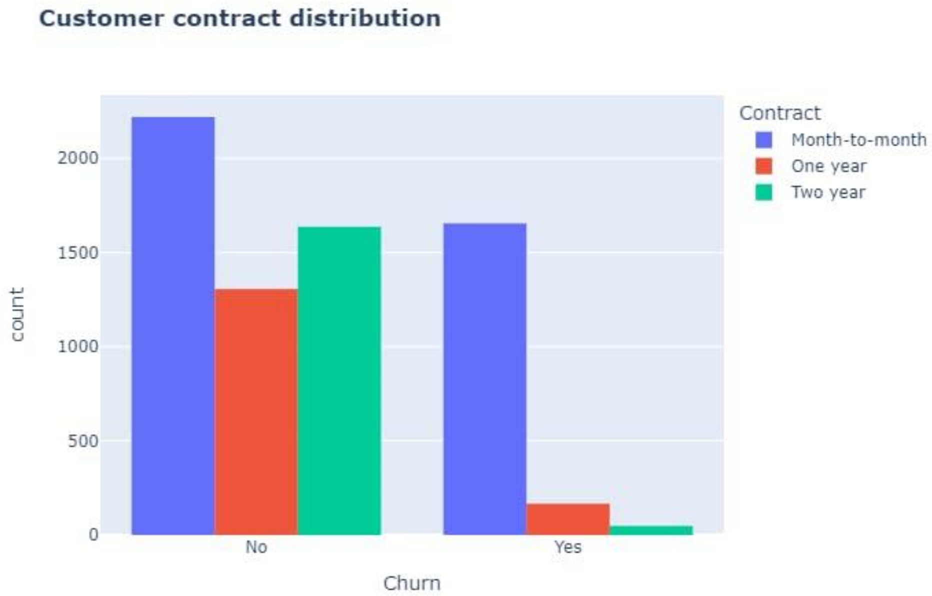

Fig-5: Customer contract distribution

About 75% of customer with Month-to-Month Contract opted to move out as compared to 13% of customers with One Year Contract and 3% with Two Year Contract

3.4.5 Payment Methods:

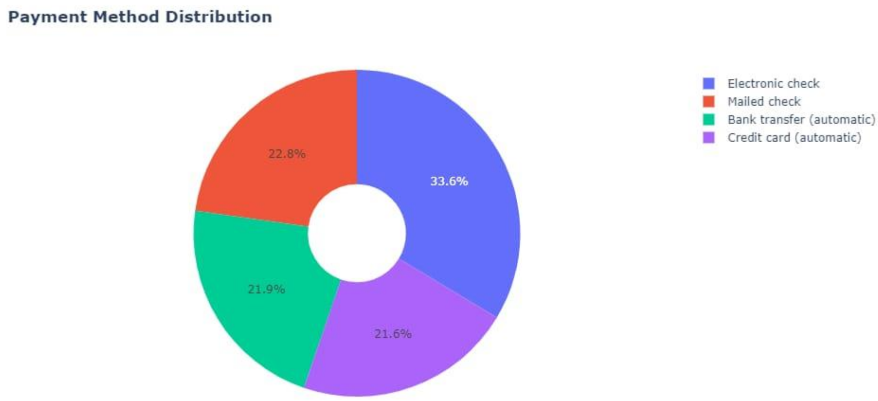

Fig-6: Payment Method Distribution

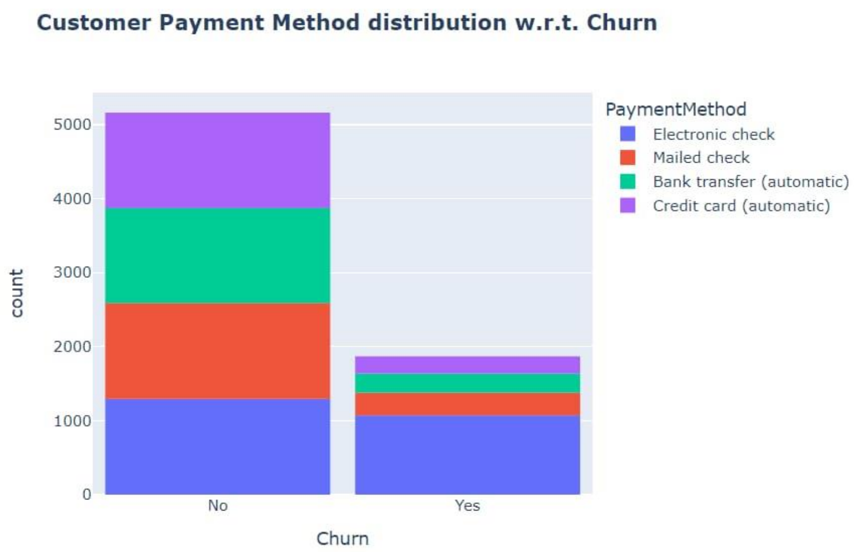

Fig-7: Customer Payment Method Distribution w.r.t Churn

Major customers who moved out were having Electronic Check as Payment Method. Customers who opted for Credit-Card automatic transfer or Bank Automatic Transfer and Mailed Check as Payment Method were less likely to move out.

3.4.6 Internet Services:

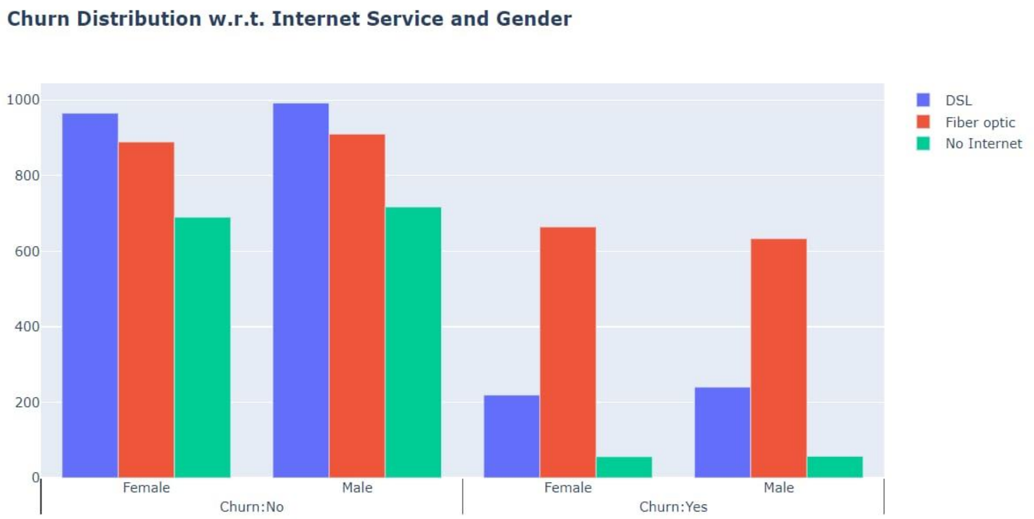

Fig-8: Churn distribution w.r.t Internet Service and Gender

3.4.7 Dependent distribution:

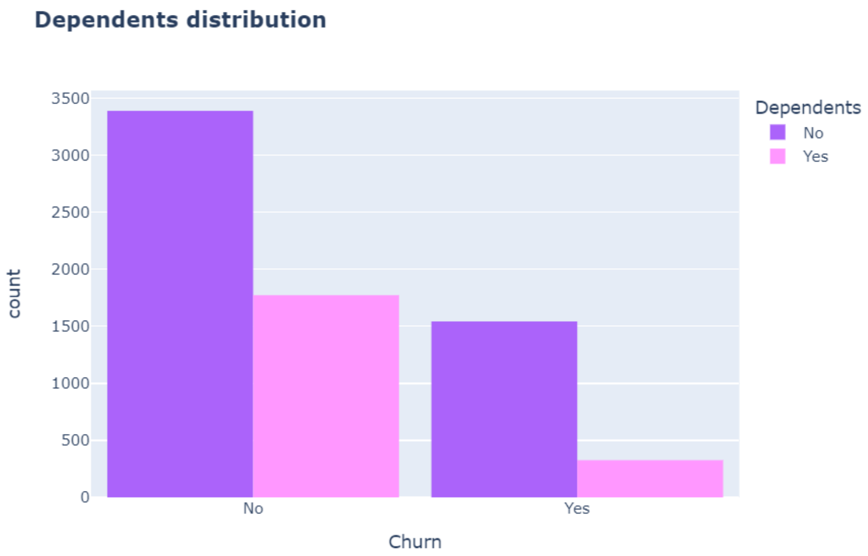

Fig-8: Dependents distribution

3.4.8 Online Security:

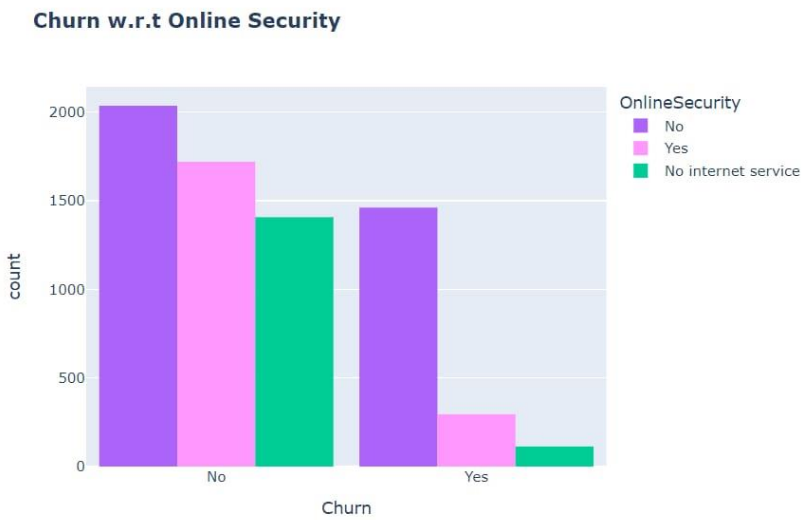

Fig-9: Online Security

3.4.9 Senior Citizen:

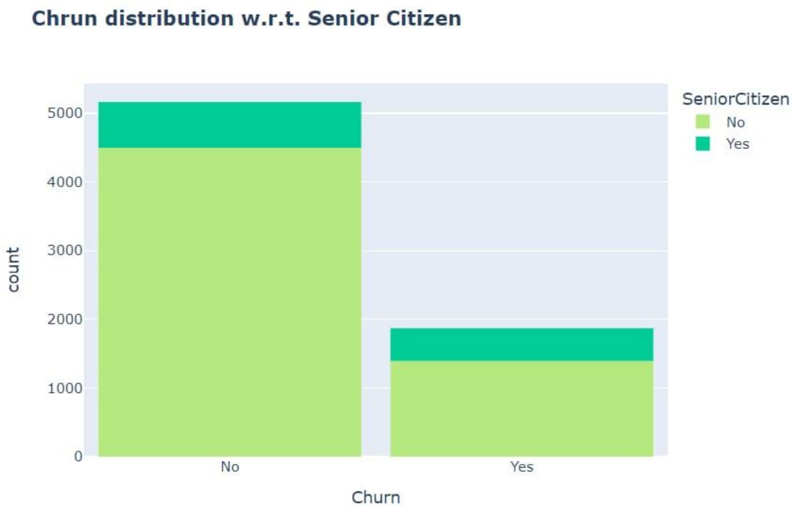

Fig-10: Churn distribution w.r.t Senior Citizen

3.4.10 Paperless billing:

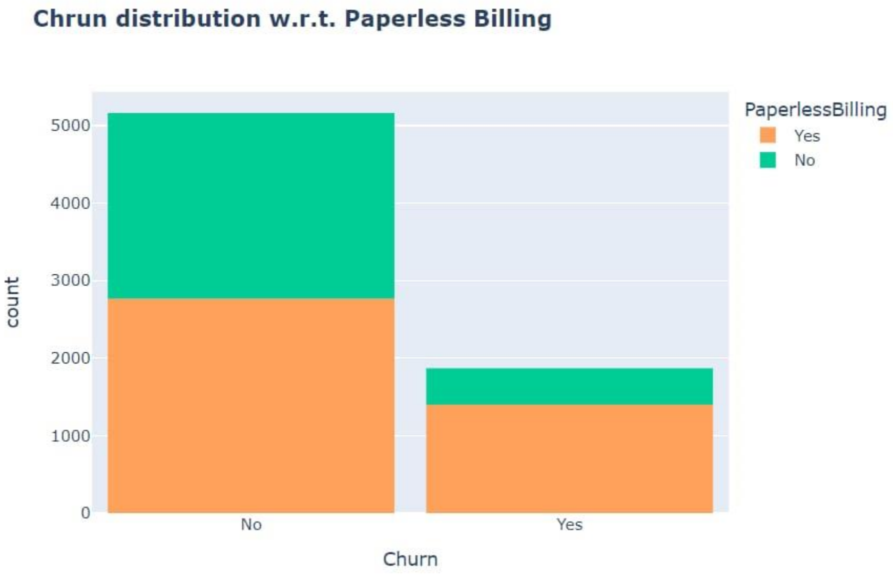

Fig-11: Churn distribution w.r.t Paperless Billing

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

3.4.11 Tech Support:

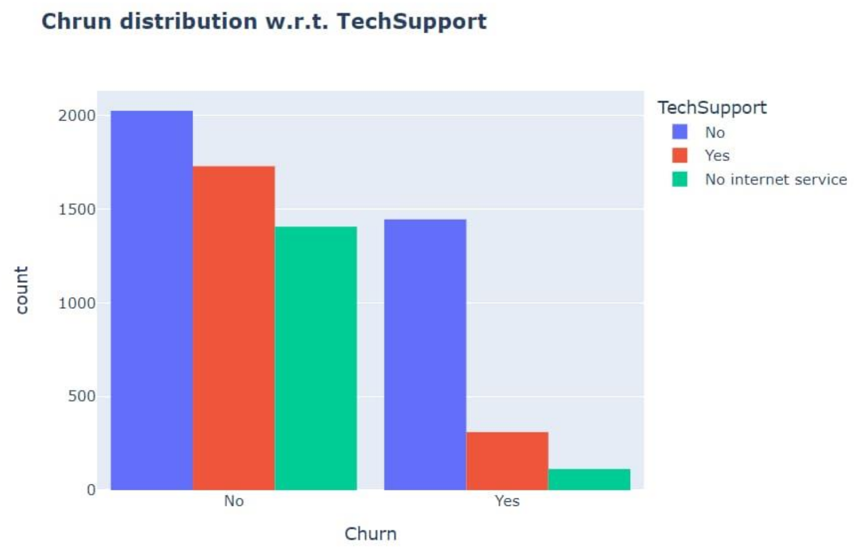

Fig-12: Churn distribution w.r.t TechSupport

3.4.12 Distribution w.r.t Charges and Tenure:

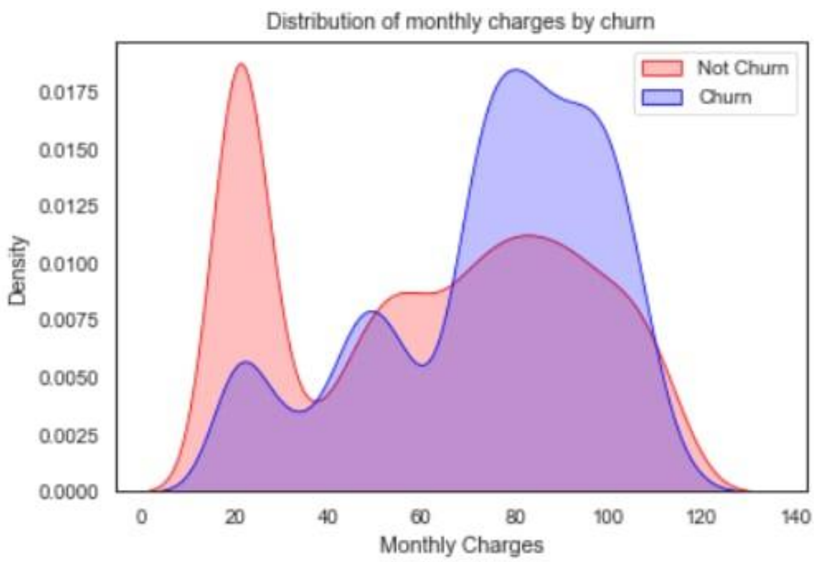

Fig-13: Churn distribution w.r.t Charges and Tenure

3.4.13 Comparison of results of various Algorithms:

Table-1: Comparison of results of various Algorithms

Algorithm	ROC AUC Mean	ROC AUC STD	Accuracy Mean	Accuracy STD
Voting Classifier	84.93	1.39	80.23	1.89
Gradient boost classifier	84.72	1.42	79.72	1.95
Adaboost	84.55	1.25	80.09	1.77
Logistic Regression	84.39	1.47	74.38	1.94
SVC	82.99	2.07	79.11	2.01
Random Forest	82.75	2.01	78.67	1.98
Gaussian NB	82.32	1.28	75.38	1.23
Kernel SVM	79.65	2.12	79.26	1.67
KNN	77.14	1.43	75.90	2.01
Decision Tree Classifier	66.67	1.07	73.73	1.12

3.5 CONCLUSION

In summary, the implementation of the customer churn prediction project, leveraging the Voting Classifier algorithm, has proven to be a strategic asset for businesses seeking to optimize customer retention strategies. The algorithm's effectiveness in amalgamating diverse machine learning models contributes significantly to its superior predictive accuracy, surpassing alternative algorithms by achieving an accuracy rate of approximately 87%, compared to the standard 82%. This enhanced accuracy translates to a more precise identification of at-risk customers, empowering businesses to proactively address potential churners and thereby minimize revenue loss. One of the key strengths of the Voting Classifier algorithm lies in its scalability and adaptability across various industries. Its ability to synthesize insights from different models not only improves accuracy but also provides a versatile solution that can be tailored to the specific characteristics of different businesses and customer bases. The proactive identification of high-risk customers allows companies to allocate resources more efficiently, focusing retention efforts where they are needed most. Furthermore, the algorithm's commitment to continuous data analysis ensures its ongoing relevance in the dynamic landscape of customer behavior and preferences. By adapting to changing customer behaviors over time, the system remains robust and capable of capturing intricate patterns that might indicate potential churn. This adaptability is crucial in today's fast-paced business environment, where customer expectations and market conditions evolve rapidly. Ultimately, the customer churn prediction project with the Voting Classifier algorithm underscores the pivotal role of predictive analytics in fostering customer retention and ensuring long-term business sustainability. The insights derived from this model not only contribute to the preservation of existing market positions but also open doors for growth and prosperity. As businesses increasingly recognize the importance of customer-centric approaches, the Voting Classifier algorithm stands out as a valuable tool in their arsenal, enabling them to stay ahead of the competition and cultivate enduring customer loyalty.

3.6 REFERENCES

[1] A. Keramati and S. M. S. Ardabili, "Churn analysis for an Iranian mobile operator," Telecommunications Policy, vol. 35, no. 4, pp. 344–356, 2011.

[2] A. Sharma and P. Kumar Panigrahi, "A neural network-based approach for predicting customer churn in cellular network services," International Journal of Computer Application, vol. 27, no. 11, pp. 26–31, 2011.

[3] S. A. Qureshi, A. S. Rehman, A. M. Qamar, and A. Kamal, "Telecommunication subscribers' churn prediction model using machine learning," in Proceedings of the 8th International Conference on Digital Information Management (ICDIM '13), pp. 131–136, Islamabad, Pakistan, September 2013.

[4] A. Rodan, A. Fayyoumi, H. Faris, J. Alsakran, and O. Al-Kadi, "Negative correlation learning for customer churn prediction: a comparison study," $e Scientific World Journal, vol. 2015, 2015.

[5] A. K. Ahmad, A. Jafar, and K. Aljoumaa, "Customer churn prediction in telecom using machine learning in big data platform," Journal of Big Data, vol. 6, no. 1, pp. 28–24, 2019.

[6] U. Paschen, C. Pitt, and J. Kietzmann, "Artificial intelligence: building blocks and an innovation typology," Business Horizons, vol. 63, no. 2, pp. 147–155, 2020.

[7] D. Paulraj, "A gradient boosted decision tree-based sentiment classification of twitter data," International Journal of Wavelets, Multiresolution and Information Processing, vol. 18, no. 4, 2020.

[8] M. Jiang, F. Qiang, L. D. Xu, B. Zhang, Y. Sun, and H. Cai, "Multilingual interoperation in cross-country industry 4.0 system for one belt and one road," Information Systems Frontiers, 2021.

[9] C. Ramalingam and P. Mohan, "An efficient applications cloud interoperability framework using I-anfis," Symmetry, vol. 13, no. 2, p. 268, 2021.

[10] S. Neelakandan, M. A. Berlin, S. Tripathi, V. B. Devi, I. Bhardwaj, and N. Arulkumar, "IoT-based traffic prediction and traffic signal control system for smart city," Soft Computing, vol. 25, no. 18, pp. 12241–12248, 2021, https://doi.org/10.1007/s00500-021-05896-x.