



DECODING THE EXIT STRATEGY:

PREDICTING CUSTOMER DEPARTURES

Machine Learning Internship Project @ Mentorless

NAVIGATING CHURN:

ROADMAP FOR THE PRESENTATION

↓ **INTRODUCTION**

↓ **ABOUT THE DATASET**

↓ **ML : THE PROCESS**

↓ **EDA INSIGHTS**

↓ **MODEL DEVELOPMENT**

↓ **OVERCOMING CHALLENGES**

↓ **EVALUATION & RESULTS**

↓ **BENIFITS & IMPACTS**

↓ **FUTURE DIRECTIONS**

↓ **CONCLUSION**

Introduction

WHY STOPPING CUSTOMER DEFLECTION MATTERS?

Customer churn poses a significant **challenge to organizational growth and performance**. It directly impacts a company's bottom line through lost revenue and sales. Beyond financial implications, churn can also lead to negative customer sentiment and potentially **damage an organization's reputation**, making it easier for competitors to gain a foothold. Recognizing the critical role of customer retention, many companies face difficulties and complexities in **managing churn effectively**.



About the dataset



This dataset provides information about customer behavior and churn. It includes various customer attributes like demographics, service subscriptions and contract details. Additionally, it captures customer engagement details and financial information. The **target variable**, "Churn," indicates whether a customer discontinued their service (1) or remained active (0). This rich dataset allows us to explore factors influencing customer churn and build models to predict customer defection.



THE PROCESS

DATA PREPROCESSING

- DATA ENCODING
- DATA SCLING
- DATA BALANCING
- DATA SPLITTING

MODEL DEVELOPMENT

- TESTED 5 DIFFERENT ALGORITHMS
- RECORDED THEIR METRICS SCORES

FEATURE ENGINEERING

- CREATED NEW FEATURES TO UNDERSTAND CUSTOMER CHURNING PATTERNS

MODEL EVALUATION

- EVAULATED 3 BEST PERFORMING MODELS USING GRID AND RANDOMIZEDSEARCHCV

UNDERSTANDING THE DATA

- COLUMNS
- SHAPE
- TARGET VARIABLE
- DATA INFORMATION

EDA

- THOROUGH DATA ANALYSIS ON DIFFERENT FEATURES.

FINALIZING THE MODEL

- FINALIZED THE BEST MODEL AND TRAINED WHOLE DATA ON IT.

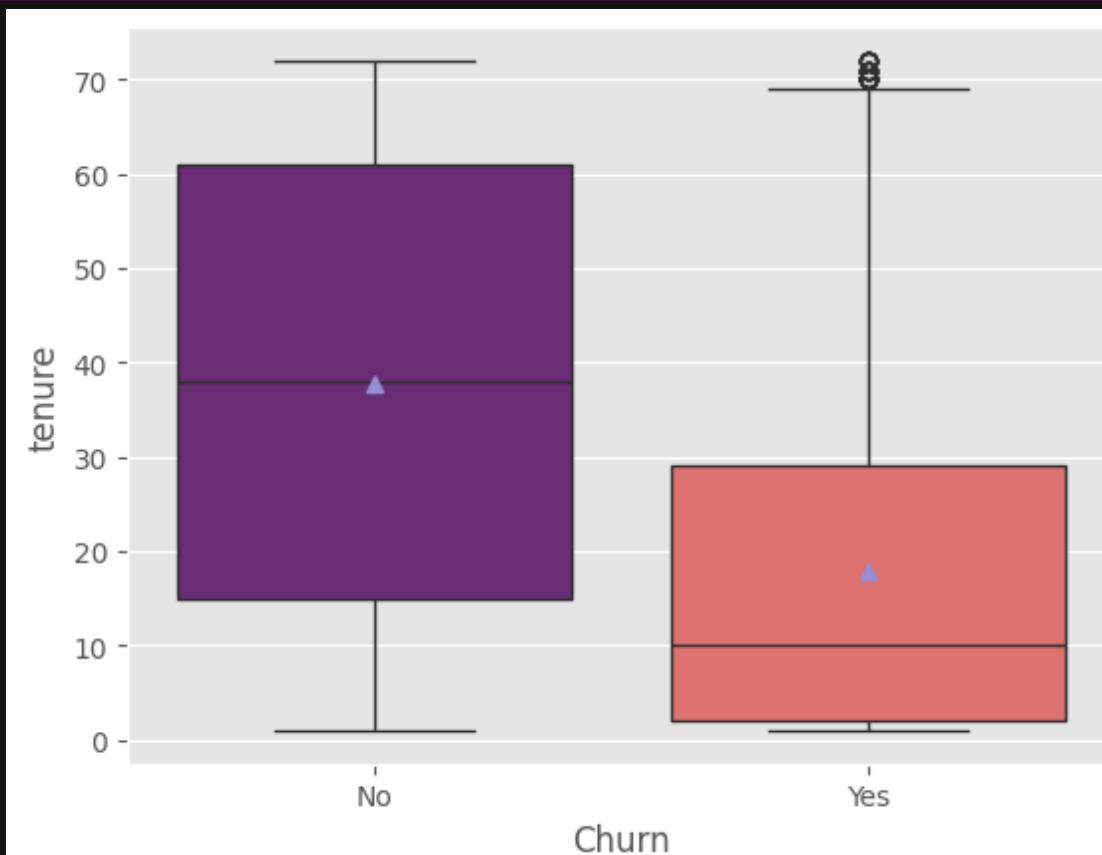
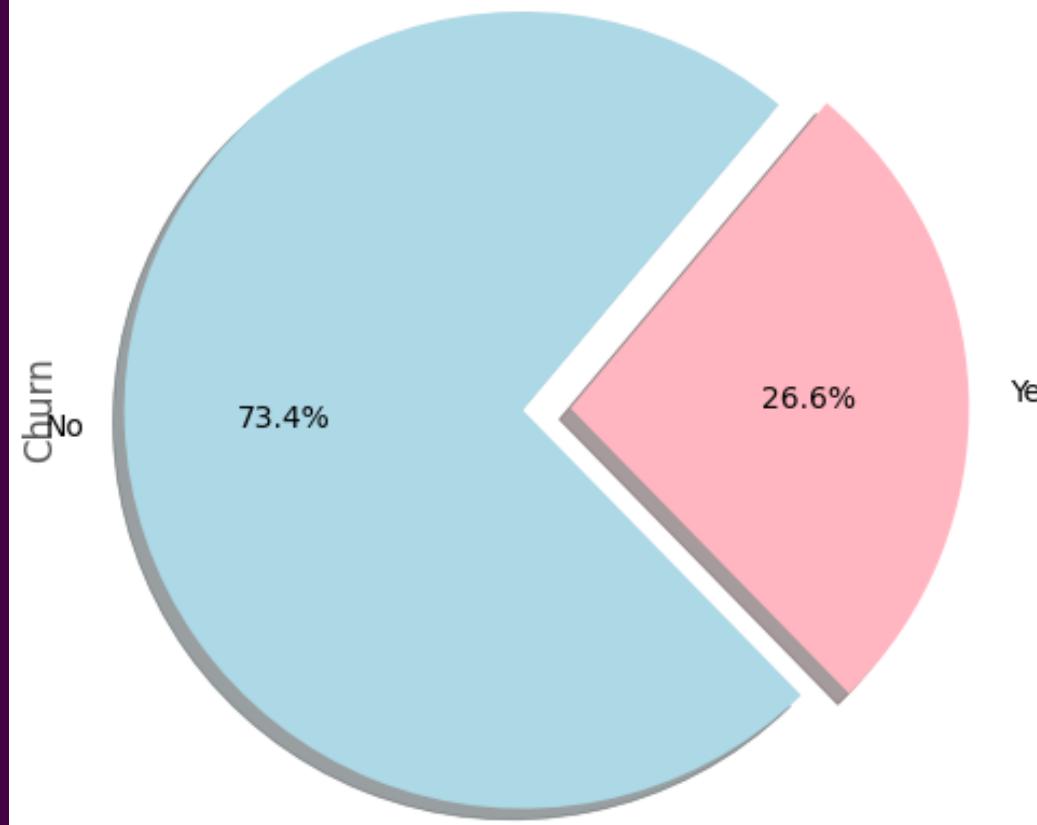


EDA INSIGHTS

WHAT THE DATA
REVEALED:
KEY DRIVERS OF
CHURN

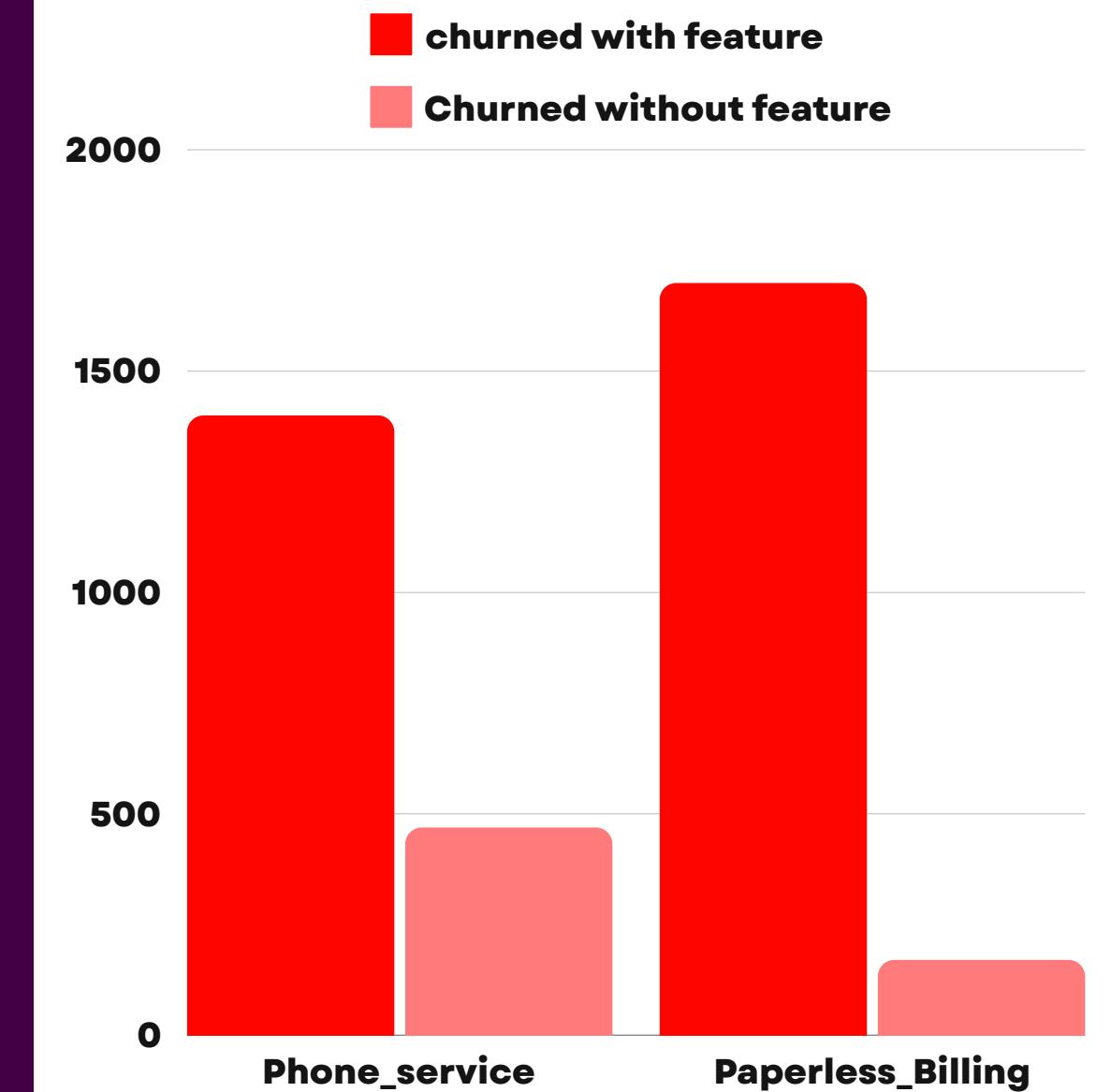
The analysis revealed that the churn rate of **26.6%** sits within the average range for the telecom industry (30-35%). While not ideal, it suggests we're on the right track with customer engagement.

Pie Chart of Customers

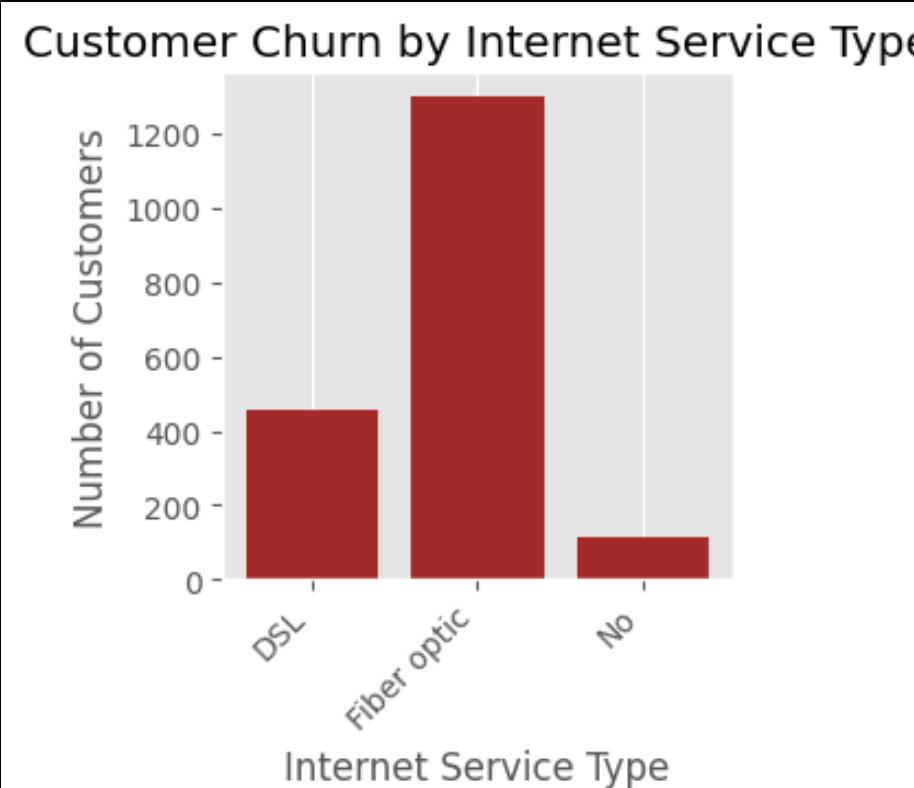


The analysis reveals a clear distinction in customer lifespans. Churned customers typically stay with us for an average of **20 years**, while loyal customers remain subscribed for an average of **40 years**. This significant difference highlights the value of our loyal customer base.

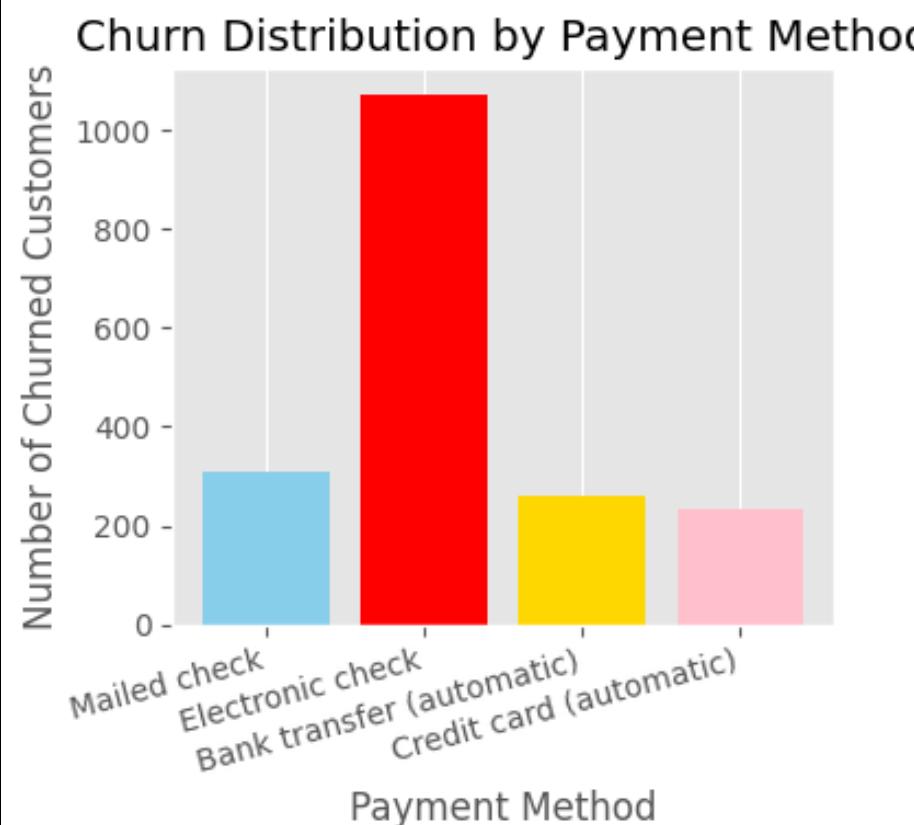
In the analysis, I found a predominance of **paperless billing** and **phone service** usage among customers who churned, making them the most contributing factors in churn



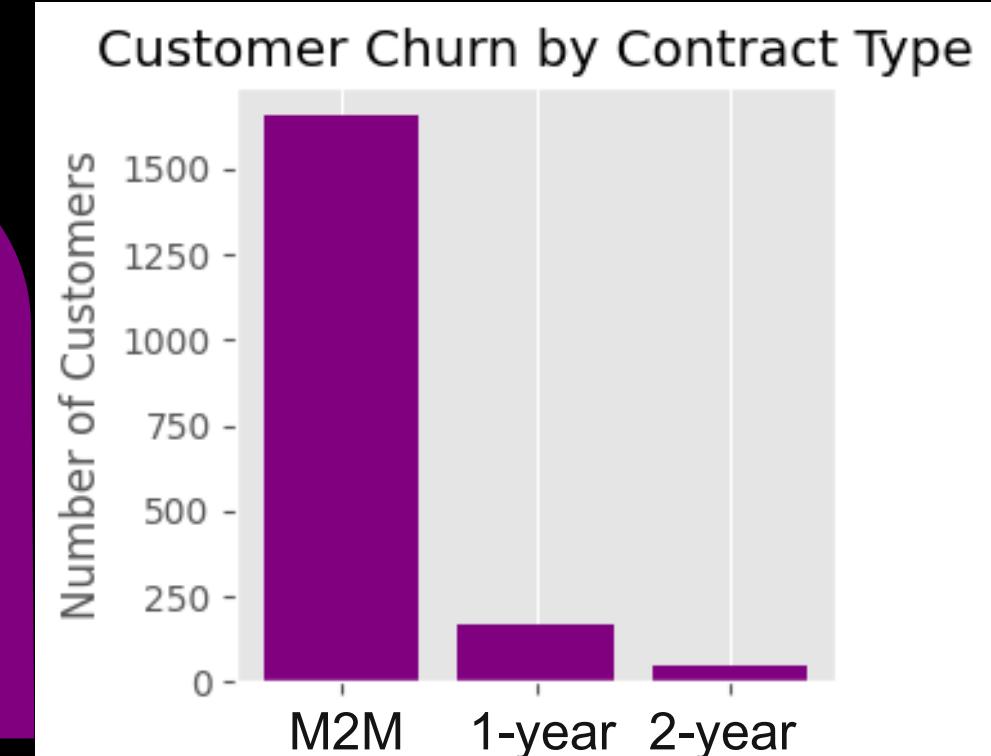
Other features contributing in churn:



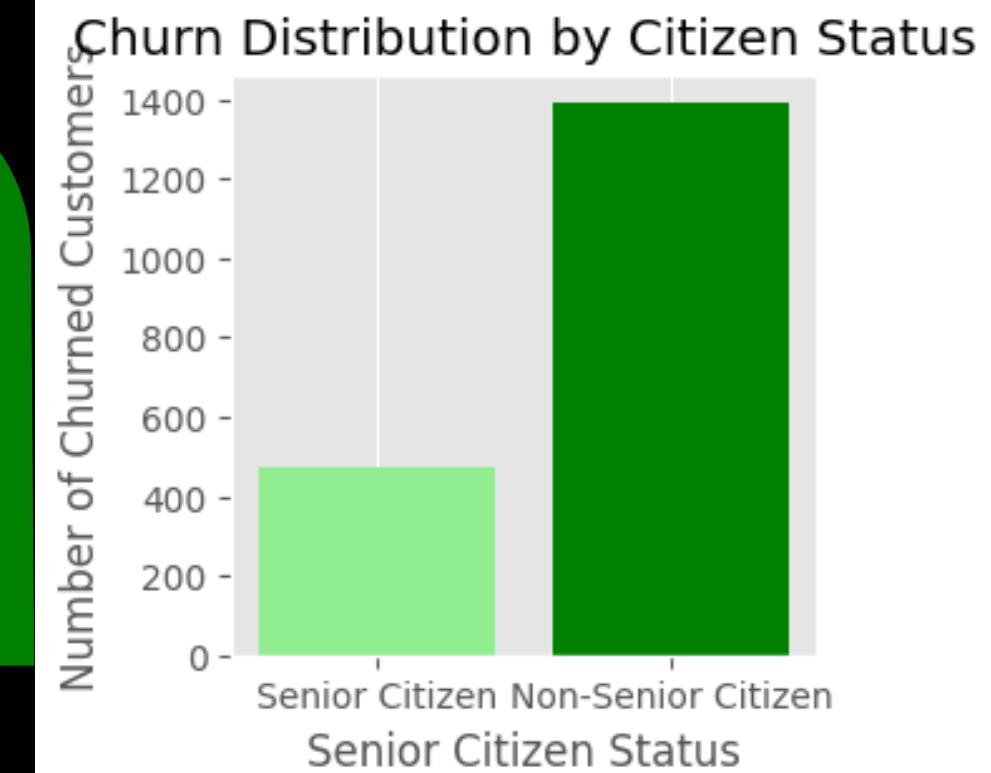
“Fiber optic” internet users exhibited the highest churn rate compared to other internet options within our customer base.



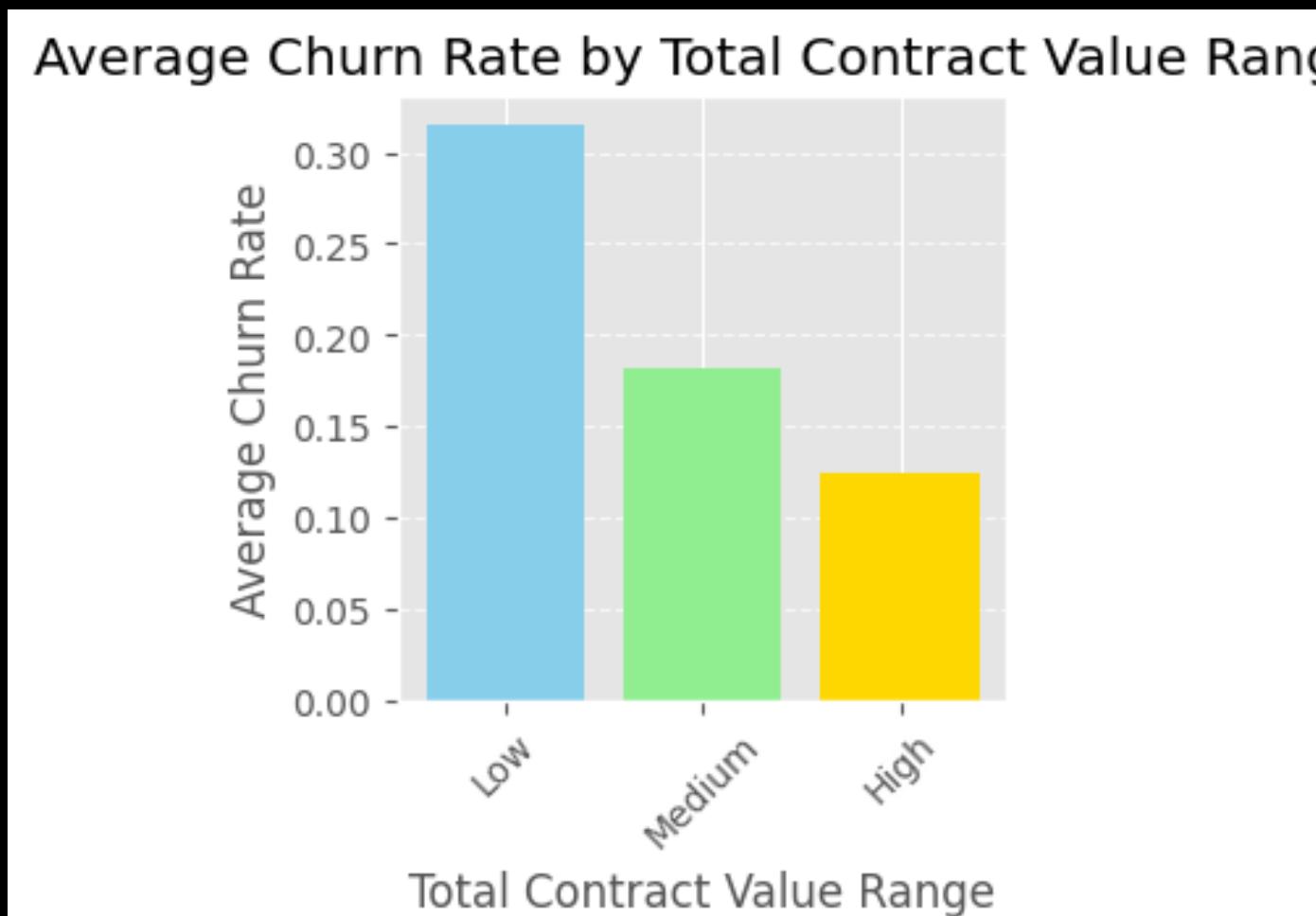
Customers using **“Electronic Check”** as their payment method exhibited the highest churn rate as compared to others



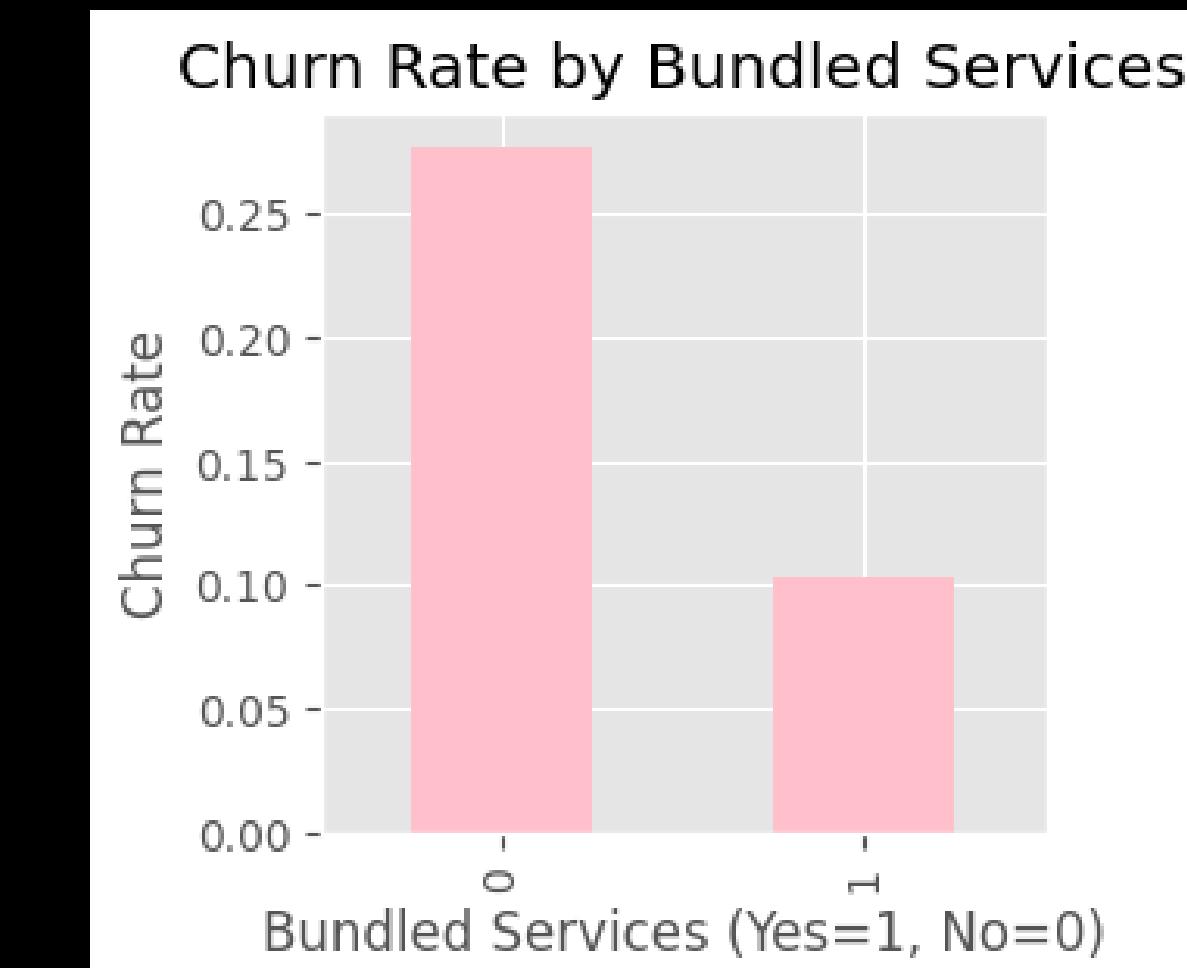
“Non-Senior Citizens” exhibited the highest churn rate compared to Senior Citizens within our customer base.



Feature Engineering identified valuable insights:



Customers with lower total contract values exhibit a higher risk of churn compared to those with higher spending. This insight can be used to develop targeted retention strategies.



Customers who subscribe to bundled services exhibit a significantly lower churn rate compared to those without bundled services. This suggests bundling can be a powerful tool for customer retention.

MODEL
DEVELOPMENT

From data
to decisions:

CHOOSING THE RIGHT
MACHINE LEARNING MODEL



To effectively predict customer churn, I compared the performance of five different machine learning algorithms using metrics like accuracy, precision, recall, and F1 score. Through this rigorous evaluation process, the Random Forest algorithm emerged as the strongest performer. This data-driven selection ensures our churn prediction model has the best possible foundation for accuracy and actionable insights.

By considering multiple metrics, I can be confident that the Random Forest model will not only identify churn risk effectively but also provide a balanced view of true positives and negatives.

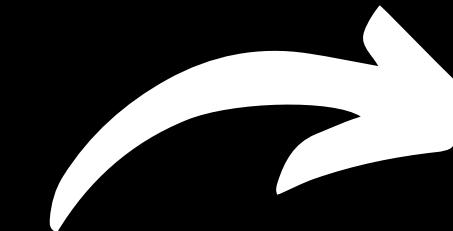




model evaluation & results

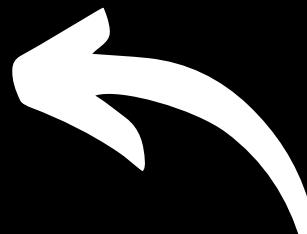
Testing the waters
MEASURING
THE MODELS'
PERFORMANCE

Having identified the Random Forest algorithm as the strongest initial performer, I further optimized its effectiveness through hyperparameter tuning. I employed two established techniques: GridSearchCV and RandomizedSearchCV, utilizing stratified k-fold cross-validation for robust evaluation. By comparing these approaches across the three most promising models, Random Forest consistently emerged as the top contender. This comprehensive optimization process ensures our churn prediction model is fine-tuned for the best possible results.



EVALUATION SCORES FROM GRIDSEARCHCV

	model	best_score	best_params
0	svm	0.806901	{'C': 100, 'kernel': 'rbf'}
1	random_forest	0.832567	{'max_depth': 20, 'n_estimators': 500}
2	xgboost	0.830630	{'learning_rate': 0.1, 'max_depth': 10, 'n_est...}



	model	best_score	best_params
0	svm	0.890756	{'kernel': 'rbf', 'C': 21.54434690031882}
1	random_forest	0.907460	{'n_estimators': 400, 'max_depth': 16}
2	xgboost	0.907184	{'n_estimators': 200, 'max_depth': 15, 'learni...}

EVALUATION SCORES FROM RANDOMIZEDSEARCHCV

CHALLENGES FACED

FROM STUMBLING BLOCKS
TO STEPPING STONES:

*My
Solutions*



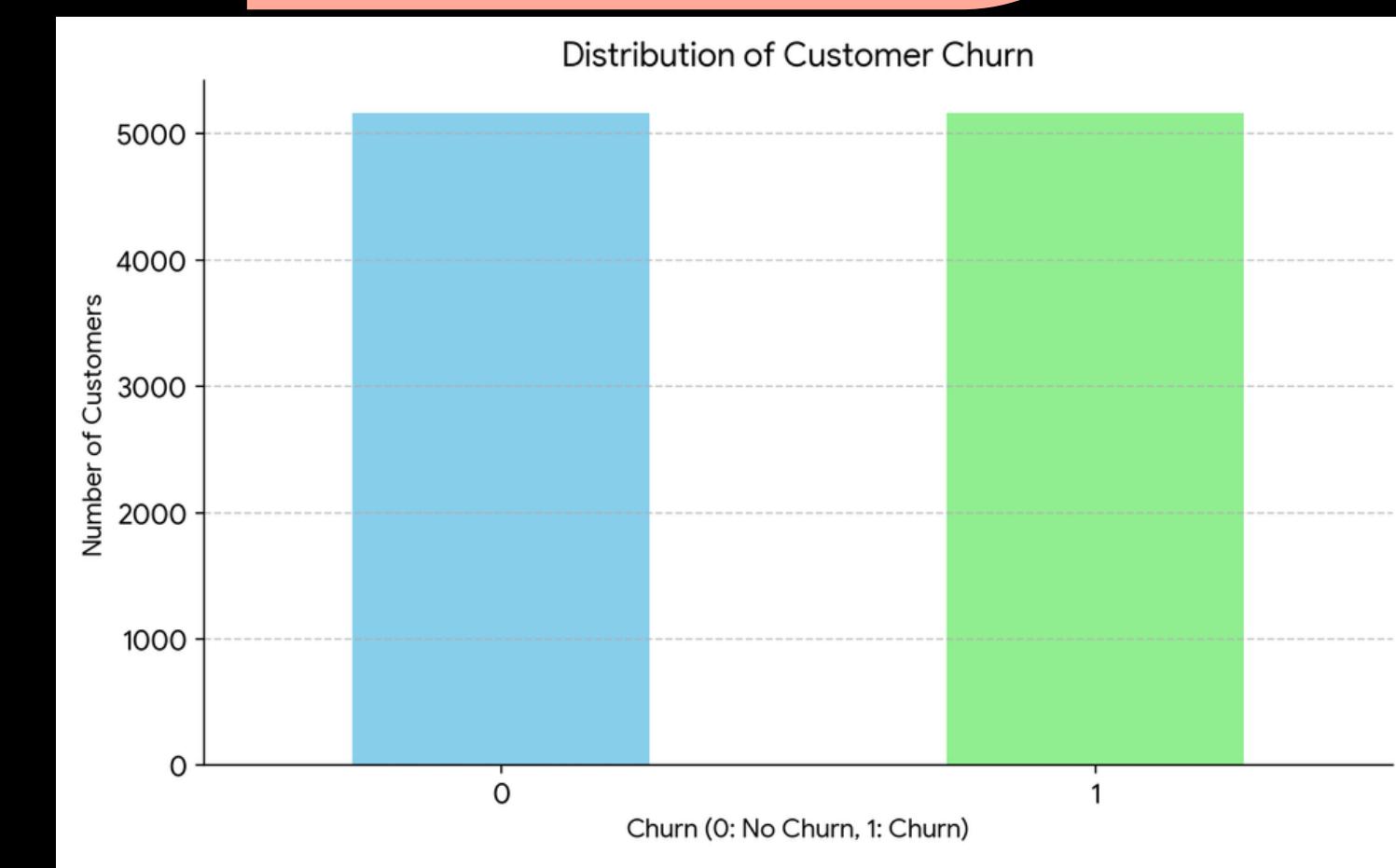
CHALLENGE : 1

UNBALANCED
DATASET



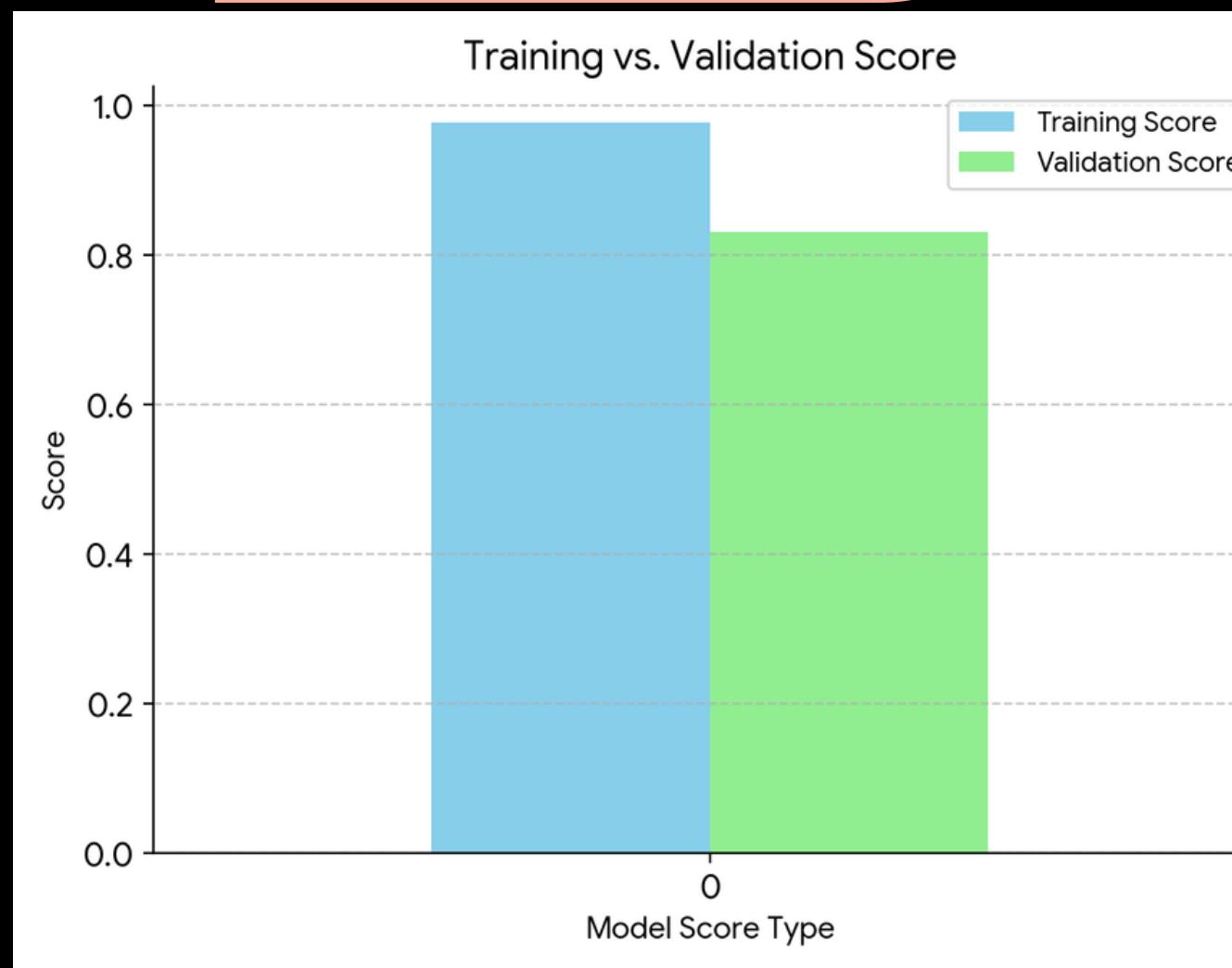
MY SOLUTION:

BALANCED
DATA
VIA SMOTE



CHALLENGE : 12

MODEL
OVERFITTING



MY SOLUTION:

FURTHER
TUNED THE
MODEL





BENEFITS & IMPACTS

***How this
model helps
business***

Our churn prediction model offers significant advantages by leveraging machine learning to proactively implement targeted retention strategies before customers cancel their service. By focusing efforts on improving our services, we can improve customer satisfaction, reduce churn rate, and generate long-term value.

Additionally, the insights gleaned from the model can inform product development and marketing campaigns, further enhancing customer engagement and loyalty, ultimately leading to increased customer lifetime value and revenue growth.



FUTURE DIRECTIONS

**BEYOND PREDICTIONS:
PROACTIVE STRATEGIES
FOR RETENTION**





Improve Customer Experience:

- **Enhance Phone Service:** Invest in improved phone service infrastructure to reduce wait times, dropped calls, and improve customer interaction experience. Consider offering multi-channel support options (chat, email) to cater to diverse customer preferences.
- **Streamline Paperless Billing:** Make the paperless billing process user-friendly and secure. Offer clear instructions on how to enroll and manage paperless billing online. Provide multiple payment options for paperless billing to ensure convenience.

Promote Value and Retention:

- **Develop Compelling Bundled Services:** Analyze customer needs and usage patterns to design attractive bundled services that offer significant value. Consider bundling internet, phone, and streaming services at competitive prices. Highlight the convenience and cost-savings benefits of bundled packages.
- **Incentivize Long-Term Contracts:** Offer attractive discounts for customers who choose 1 and 2-year contract options. This provides revenue predictability and promotes customer loyalty.
- **Enhance Fiber Optic Services:** Invest in infrastructure upgrades to ensure consistent high speeds and reliable performance for fiber optic internet customers. Consider offering exclusive benefits or priority service for fiber optic users.





PTarget Specific Customer Segments:

- **Cater to Streaming Enthusiasts:** Analyze popular streaming services and create targeted promotions or bundled packages that include streaming subscriptions at discounted rates. Offer high-bandwidth internet options to optimize streaming experiences.
- **Customer Segmentation and Insights:** Leverage customer data to segment your customer base and develop targeted marketing campaigns tailored to specific needs and preferences. Use churn prediction models to identify at-risk customers and implement proactive retention strategies.
- **Track and Measure Results:** Continuously monitor and track the effectiveness of your initiatives. Analyze key performance indicators (KPIs) like customer satisfaction, churn rate, and revenue to assess the success of your implemented changes.

Overall, focus on:

- **Customer Centricity:** Prioritize improving customer experience and satisfaction across all touchpoints.
- **Value Proposition:** Clearly communicate the value proposition of your services and how they benefit customers.
- **Data-Driven Strategies:** Utilize data analytics to inform decision-making and measure the impact of your initiatives.

By implementing these suggestions, you can attract new customers, retain existing ones, and ultimately drive business growth and profitability.



Conclusion



In conclusion, this project has been a truly insightful journey for me. Delving into customer churn data and discovering actionable insights that can benefit our business has been incredibly rewarding. I'm excited about the positive impact these findings can have on our customer retention strategies.

Thank you all for your time and attention.

Beyond the results, this project has been a valuable learning experience. I've tackled various technical problems, gaining a deeper understanding of the entire machine learning process, from data exploration and model selection to optimization and evaluation. This experience has ignited a passion within me to continue learning and applying machine learning to real-world problems. I'm eager to contribute to society by tackling similar challenges in the future.