Telecom Customer Churn Prediction

BAX-452: Machine Learning

Avantika Goyal
Charles Wang
Shalagha Mundepi

TABLE OF CONTENTS

Section 1: Executive Summary

1.1 Introduction

The telecommunications industry is fiercely competitive, with companies constantly striving to retain customers and mitigate churn. Predicting churn trends is paramount for product development and devising effective retention strategies. By understanding customer behavior and anticipating churn patterns, telecom companies can tailor their offerings and services to meet customer needs and preferences more effectively.

1.2 Data Cleaning and Exploratory Data Analysis

Our data preparation involved encoding categorical variables and addressing missing values to ensure the dataset's integrity. Through Exploratory Data Analysis, we uncovered valuable insights into customer behavior, service usage patterns, and potential churn indicators. These insights lay the foundation for our subsequent model development and analysis.

1.3 Machine Learning Model Building and Evaluation

Using the processed data, we developed machine learning models to predict customer churn. By leveraging advanced algorithms and techniques, we aimed to accurately identify factors influencing churn and assess the effectiveness of various predictive models. Through rigorous evaluation and validation, we assessed the performance of each model and identified the most effective approach for churn prediction.

1.4 Business Questions/Solutions & Conclusions

Sections 2 and 3 will address key business questions and current industry solutions. Sections 4 and 5 will delve into model analysis, followed by recommendations in Section 6.

Section 2: Key Business Challenges

1. How can telecom companies gain a deeper understanding of their customers and prevent them from leaving by analyzing data on service usage patterns, customer demographics, and churn rates?

2. What effective pricing strategies can be implemented to accommodate diverse customer preferences and service usage across different customer segments, thereby reducing churn and increasing customer satisfaction?

3. How can companies incorporate contract strategies to encourage customers to sign up for longer-term plans and stay with the company longer, thus minimizing churn?

Section 3: Traditional Approaches to Customer Churn Mitigation

Customers' Behavior & Churn Prevention

Telecom companies analyze recent market research and individual customer data to predict churn. They segment customers based on demographics, subscription plans, tenure, and service usage. Insights help identify potential churners and loyal customers. Companies offer promotions to retain customers, such as personalized discounts and special offers based on usage patterns and preferences.

Effective Pricing Strategies

Telecom firms implement dynamic pricing models tailored to individual customers' behavior and value. Prices vary based on usage, subscription history, and willingness to pay, ensuring competitive pricing while maximizing revenue. Additionally, bundling services or offering loyalty rewards can incentivize customers to stay.

Long-Term Contract Strategy

To reduce churn, telecom companies incentivize customers with rewards and discounts for longer contract commitments. By offering benefits like discounted rates or exclusive perks for extended contracts, companies encourage customers to stay committed, fostering long-term relationships and reducing churn rates.

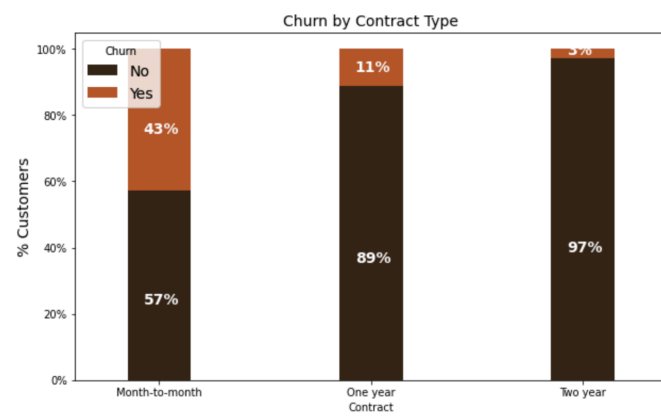## **Section 4: Exploratory Data Analysis**

This section will be performing our Exploratory Data Analysis (The following are the 4 most important EDAs, the rest of the EDA will be in the appendix).

1. Evaluation of the total churn rate in the dataset
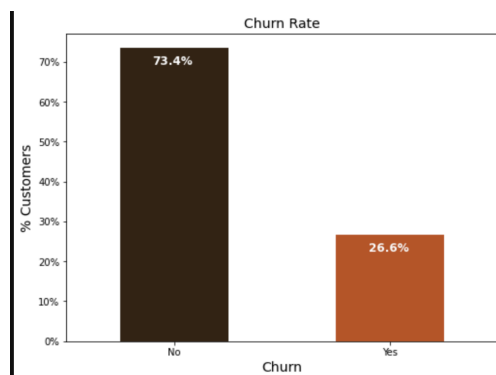
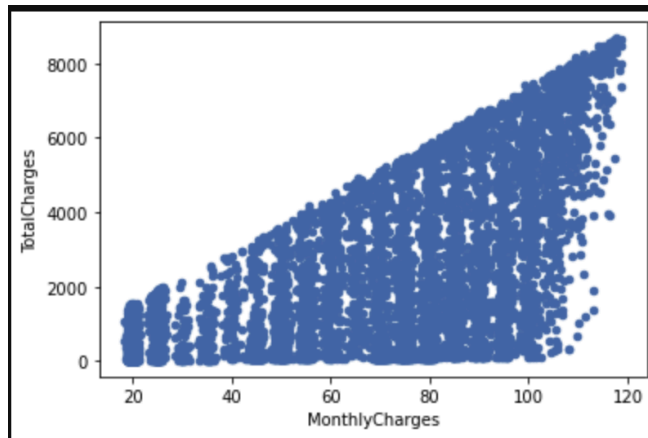

Approximately 73.4% of customers choose not to churn, while 26.6% opt to churn.
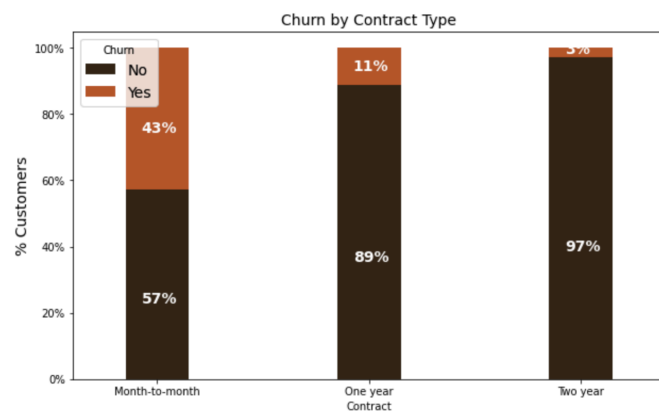
2. Churn Rate by Contract Type

Most customers are on month-to-month contracts, with similar numbers for 1-year and 2-year contracts.

3. Relationship Between Monthly and Total Charges:

The total charges increase as the monthly bill for a customer increases. This matches the customer behavior and bill charges that monthly bill charges and total charges are positively correlated.

4. Churn Rate by Contract Type:

Matching the result above, the customers who have a month-to-month type contract are more likely to churn with a 43% churning rate. For a 2-year type contract, they're not likely to churn with only a 3% churn rate.

Section 5: Model Analyses

1. Logistic Regression model

```
Classification Report:
              precision    recall  f1-score   support

           0       0.86      0.89      0.87      1549
           1       0.66      0.58      0.62       561

    accuracy                           0.81      2110
   macro avg       0.76      0.74      0.75      2110
weighted avg       0.80      0.81      0.81      2110


Confusion Matrix:
[[1380  169]
 [ 234  327]]
```

Key Metrics: Test Accuracy: 80.90%, Precision: 66%, Recall: 58%.

Strengths: Good at predicting non-churn.

2. Decision Tree model

```
Classification Report:
              precision    recall  f1-score   support

           0       0.82      0.81      0.82      1549
           1       0.50      0.52      0.51       561

    accuracy                           0.73      2110
   macro avg       0.66      0.67      0.66      2110
weighted avg       0.74      0.73      0.74      2110

Confusion Matrix:
[[1255  294]
 [ 267  294]]
```

*T*est Accuracy: 73.41%, Precision: 50%, Recall: 52%.

Strengths: Provides a clear decision-making path, which could be useful for interpreting how different factors lead to churn.

Weaknesses: Prone to overfitting

3. Support Vector Machine

```
              precision    recall  f1-score   support

           0       0.83      0.93      0.87      1052
           1       0.66      0.42      0.51       355

    accuracy                           0.80      1407
   macro avg       0.74      0.67      0.69      1407
weighted avg       0.78      0.80      0.78      1407
```
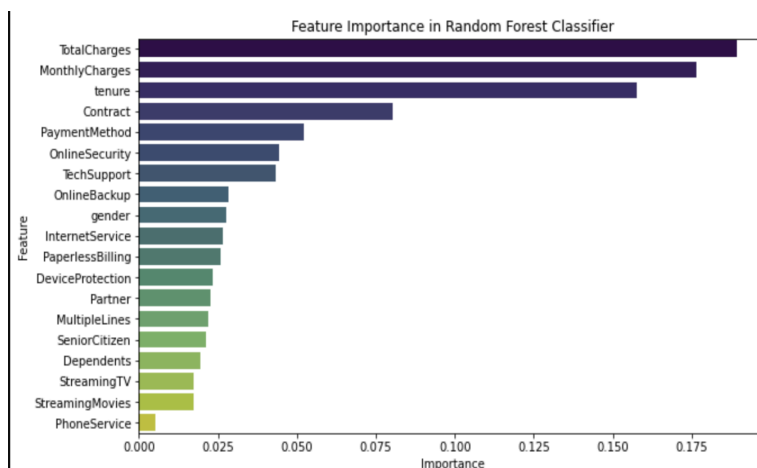
Testing Accuracy: 79.95%, Precision: 66.22%

Strengths: Reliable predictions for non-churn.

Weaknesses: Less accurate for churn.

4. Random Forest Classifier model
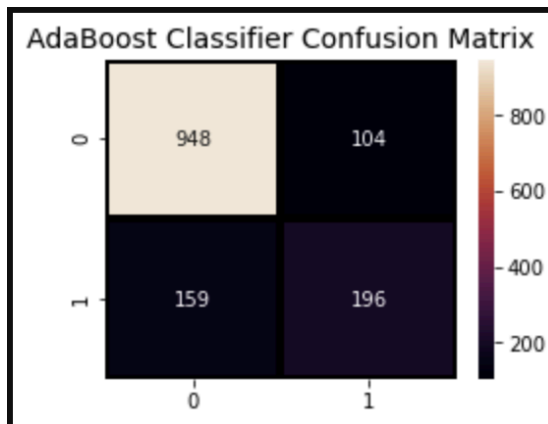

Feature Importance in Random Forest Classifier

Testing Accuracy: 81.16%

Strengths: Identifies key predictors like total charges, monthly charges & tenure.

Weaknesses: Complexity may lead to overfitting.

5. ADABoost model

Key Metrics: Accuracy: 81.31%, Precision: 65.33%, Recall: 55.21%.

Strengths: Highest accuracy and good ability to differentiate churn and non-churn customers.
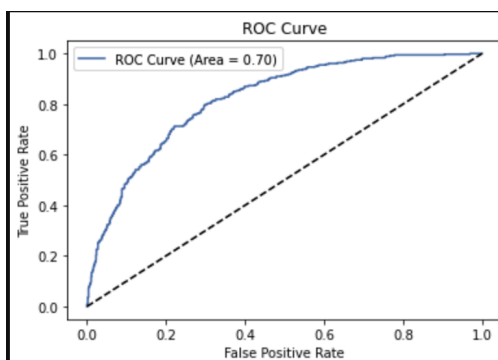
Weaknesses: None significant.

6. Gradient Boosting Classifier model

Key Metrics: Accuracy: 80.67%, Precision: 65%, Recall: 51%.

Strengths: Robust to overfitting and effective in capturing complex patterns in the data.

Weaknesses: Can be computationally intensive and challenging to tune due to the numerous hyperparameters.

7. XG Boost model

Key Metrics: Accuracy: 79.25%, Precision: 60.40%, Recall: 51.55%.

Strengths: Handles missing data well.

Weaknesses: Requires careful parameter tuning.

8. <u>Voting Classifier Model</u>

Testing Accuracy: 81.25% (approximately)

Strengths: Excellent at identifying non-churn, Combines multiple model strengths; potentially more accurate.

Weaknesses: Struggles with churn prediction, Increases complexity; difficult interpretation.

Section 6: Recommendations and Business Solutions

The *ADABoost mode*l demonstrates the highest testing accuracy of 81.30% among all models analyzed. This indicates its superior performance in predicting customer churn, making it a valuable tool for the telecom company.

Recommendations for customer retention strategies:

1. ***Focus on Contract Length***: Customers with month-to-month contracts are more prone to churn. Consider incentivizing longer-term contracts through discounts or added benefits, which could enhance customer loyalty and reduce churn.

2. ***Target High-Risk Segments***: Utilize the predictive power of AdaBoost and Random Forest models to identify customers at high risk of churn. Tailor specific retention strategies for these segments, such as personalized offers or proactive customer service outreach.

3. ***Enhance Online Security and Tech Support***: Customers lacking online security and tech support services show higher churn rates. Enhancing these services could improve customer satisfaction and retention.

4. **Revise Pricing Strategies**: High monthly charges are associated with increased churn. Review and potentially adjust your pricing strategy, considering competitive pricing and value-added services to offer more attractive packages to customers.

5. **Improve Customer Engagement**: Engage with customers regularly through feedback surveys, updates on service enhancements, and personalized communication. This can build a stronger relationship and potentially reduce churn.

6. **Optimize Payment Method Options**: Customers using electronic checks show higher churn rates. Encourage more stable payment methods like automatic bank transfers or credit card payments through incentives or convenience features.

7. **Special Attention to Senior Citizens**: This demographic shows a higher tendency to churn. Tailor services and support to meet their specific needs, enhancing their overall customer experience.

8. **Monitor and Respond to Usage Patterns:** Analyze customer usage patterns and proactively offer customized plans or services that better align with their usage, preventing potential churn due to dissatisfaction with existing plans.

Implementation in Business Processes:

- **Integration with CRM Systems**: Integrate churn prediction models with customer relationship management (CRM) systems to flag high-risk customers, enabling targeted intervention

strategies.

- **Employee Training**: Train customer-facing employees on the key findings and strategies derived from churn analysis to empower them to address customer concerns effectively.

- **Iterative Testing and Optimization:** Continuously test and refine retention strategies in a controlled environment, learning from the outcomes to optimize future approaches.

Section 7: Summary & Conclusion

In conclusion, the telecommunication industry is a competitive market nowadays, it tightly connects with everyone's life to help people connect easily like never before. However, in a competitive market, customers have a large amount of choice between companies and competitors. It's easy for a company to cause customer churn if others offer better choices. Therefore, telecom companies must implement various machine learning models to predict potential customer churn. These models will study past data from customers, such as their internet services, contract type, internet service, senior citizen or not, and how many bills they're paying, to predict customer behavior in the future of churn choice.

We conclude here with the ADABoost model showing the highest accuracy in predicting churn for our testing portion data, and a high ROC AUC score has proven its strong predictive power. We also suggest the company employ a random forest classifier model, with the 2nd highest prediction accuracy score and provides feature importance insights. By the insights of feature importance, the company would easily know which feature in their service would prevent customers from churning, leading to better strategies and pricing changes.

The ultimate goal of the telecom company is to gain revenue from its service. Customer churn will lead to revenue decrease and market share loss due to competitors having more customers. The company must implement machine learning models for analysis of customers' churn predictions**.**

References

1. Atindra Bandi, "Telecom Churn Prediction," Kaggle, [Online]. Available:

https://www.kaggle.com/code/bandiatindra/telecom-churn-prediction/input.

2. Blastchar, "Telco Customer Churn," Kaggle, [Online]. Available:

https://www.kaggle.com/datasets/blastchar/telco-customer-churn.

3.Sharmila K. Wagh a, Aishwarya A. Andhale, "Customer churn prediction in telecom sector

using machine learning techniques" in "Science Direct". Available:

https://www.sciencedirect.com/science/article/pii/S2666720723001443.

4. Abdelrahim Kasem Ahmad, Assef Jafar & Kadan Aljoumaa , "Customer churn prediction in

telecom using machine learning in big data platform," in "Journal of Big Data". Available:

https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0191-6.