Customer Churn Analysis Report

Overview

This report presents an analysis of customer churn based on a dataset containing various features related to customer activity, consumption, and pricing. The objective is to identify the key drivers of churn and provide insights to improve customer retention strategies.

Key Findings

Churn Statistics

- **Churn Rate**: 10% of the total customers have churned.
- **Gas Service Impact**: Customers who do not buy gas are 2% more likely to churn compared to those who do.

Influential Features

- **Net Margin and 12-Month Consumption**: These are top drivers for churn, indicating that customers with lower net margins and consumption over the past 12 months are more likely to churn.
- **Margin on Power Subscription**: This is another significant factor, suggesting that customers with lower margins on their power subscriptions are at higher risk of churning.
- **Time-Related Factors**: The number of months a customer has been active, their overall tenure, and the number of months since their last contract update are influential in determining churn. Customers with shorter tenures and more recent contract updates are more prone to churn.
- **Price Sensitivity Features**: While these features are scattered throughout the dataset, they are not the primary drivers for churn.

Model Performance

- **True Negatives**: Out of all the negative cases (customers who did not churn), the model correctly identified 3282 out of 3286, indicating excellent performance in predicting non-churners.
- **False Negatives**: The model predicted 348 customers as non-churners when they actually churned, highlighting a significant area for improvement.
- **False Positives**: There were only 4 cases where the model predicted churn incorrectly, which is a positive outcome.

- **True Positives**: The model correctly identified only 18 out of 366 churners, showing poor performance in predicting actual churners.

Model Metrics

- **Accuracy Score**: While the accuracy score is high, it is misleading due to the imbalance in the dataset (high number of non-churners).
- **Precision Score**: The precision score is 0.82, which is decent but could be improved. This score indicates the proportion of predicted churners who are actual churners.
- **Recall Score**: The recall score is very low, indicating the model's poor ability to identify positive samples (actual churners). This is the main concern that needs to be addressed in model improvement efforts.

Recommendations

1. **Enhance Model to Reduce False Negatives**: Focus on improving the model's ability to identify actual churners to reduce the number of false negatives.
2. **Target Customers with Low Net Margins and Consumption**: Develop strategies to retain customers with lower net margins and lower consumption over the past 12 months.
3. **Improve Engagement for Newer Customers**: Implement programs to engage customers with shorter tenures and those who recently updated their contracts.
4. **Leverage Price Sensitivity Features**: Although not primary drivers, price sensitivity features should still be considered in customer retention strategies.

Conclusion

This analysis highlights the importance of net margin, consumption, and time-related factors in understanding customer churn. While the model performs well in predicting non-churners, significant improvements are needed to accurately identify churners. By addressing the highlighted areas and implementing targeted strategies, the company can improve customer retention and reduce churn rates.