Customer Churn Prediction Report

📌 Conclusion and Recommendations

🎯 Project Objective

The primary goal of this project was to predict and analyze customer churn in order to support data-driven retention strategies and foster sustainable business growth.

📊 Conclusion

• The overall churn rate in the dataset is 27%, highlighting a significant portion of customers at risk.

• Analysis revealed key high-risk segments:

- 🗓️ Customers on month-to-month contracts

- 🌐 Customers using fiber optic internet service

- 💳 High-paying customers who may not perceive sufficient value

---

1. Executive Summary

This report presents an analysis of customer churn within our service. Key findings include:

• 27% overall churn rate

• Higher churn observed among customers who pay more, are older, and use month-to-month contracts

• Most churn occurs within the first six months

The report highlights risk factors, customer lifetime value (CLV) impact, and actionable business recommendations.

---

2. Customer Overview

Metric	Value
Total Customers	7,043
Churn Rate	27%
New Customers	6%
Average Tenure of Churned Customers	18 months

---

3. Customer Profile by Status

Metric	Churned	Joined	Stayed
Average Age (years)	49.7	42.9	45.6
Number of Dependents	0.12	0.36	0.62
Average Tenure (months)	17.98	1.72	41.04
Monthly Charges (R)	73.35	42.78	61.74
Average Monthly Data Usage (GB)	22.18	15.28	20.36

---

4. Key Drivers of Churn

- Customers on month-to-month contracts have higher churn rates.
- Fiber optic users show increased churn compared to other internet types.
- Customers without online security are more likely to churn.
- Slightly higher churn observed among older customers.

---

5. Customer Risk Analysis

Metric	Value
At-Risk Customers	2,448
Active Churned Customers	1,869
High-Risk Churn Rate	48.2%

Risk scores indicate most churn occurs in customers with moderate to high risk levels.

---

6. Customer Lifetime Value (CLV)

Metric	Value
Total CLV	$21,371,131.69
CLV Lost Due to Churn	$3,684,459.82 (17.24%)
CLV Retained	$17,632,392.12
Estimated Recoverable CLV (20% of Lost)	$736,891.96

---

7. Retention Probability at Tenure Milestones

Tenure (Months)	Retention Probability
6	88.5%
12	84.3%
24	78.9%
36	74.9%
60	66.5%

Retention probability declines steadily over time, emphasizing the need for ongoing engagement.

---

8. Business Recommendations

• Improve onboarding and early engagement to reduce churn within the first six months.

• Consider introductory promotions (e.g., 20% discount for first 1–2 months) to attract and retain new customers.

• Focus retention efforts on high-paying, month-to-month fiber optic users without online security.

• Tailor communication and support for older customers who show higher churn risk.

• Use risk scores to prioritize customer retention interventions effectively.

• Develop campaigns to recover lost revenue based on CLV insights.

---

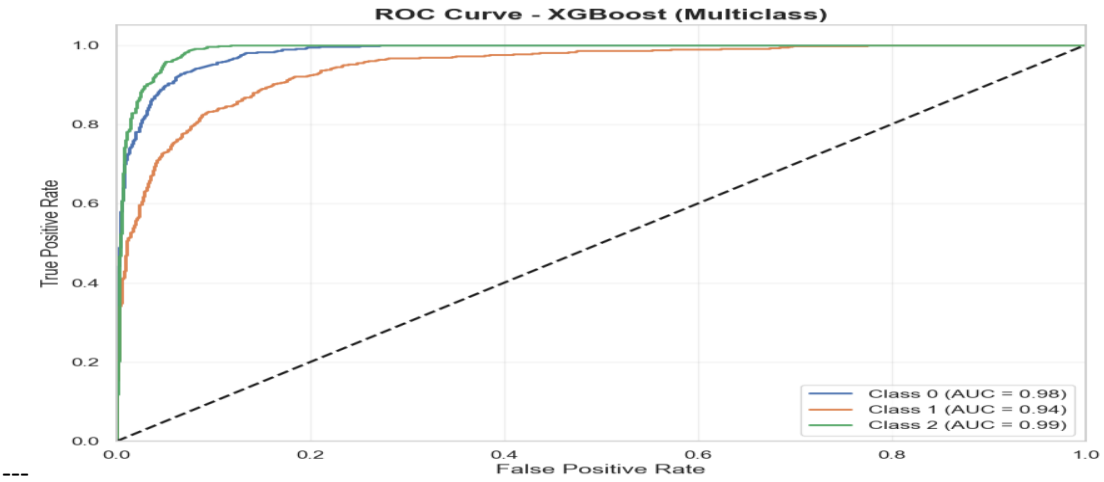
9. Next Steps

• Implement recommended strategies and monitor churn impact.

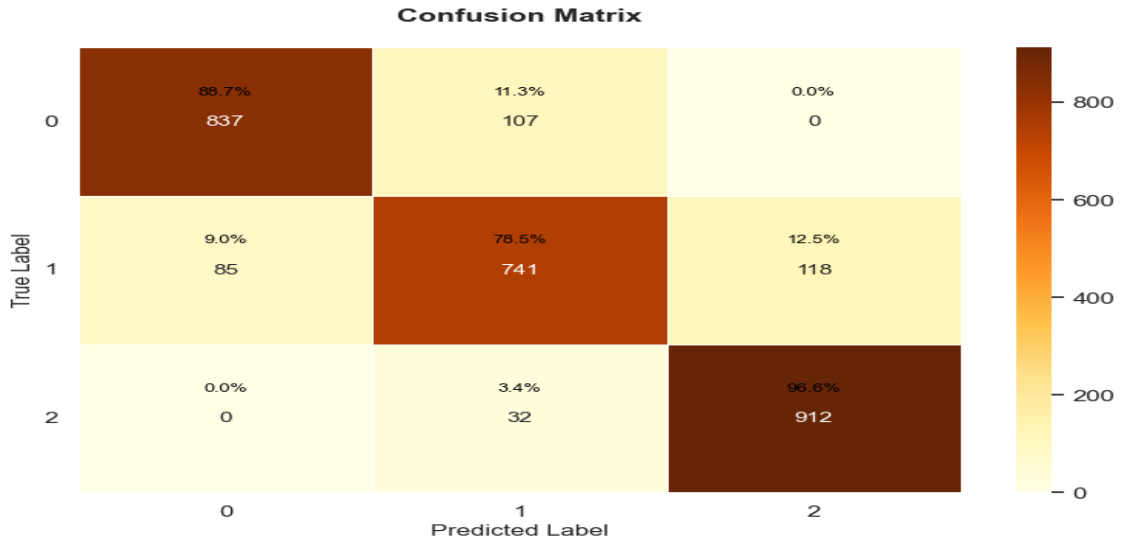
• Prepare for the next phase: building machine learning models to predict churn and enable proactive retention.

Model Performance Highlight

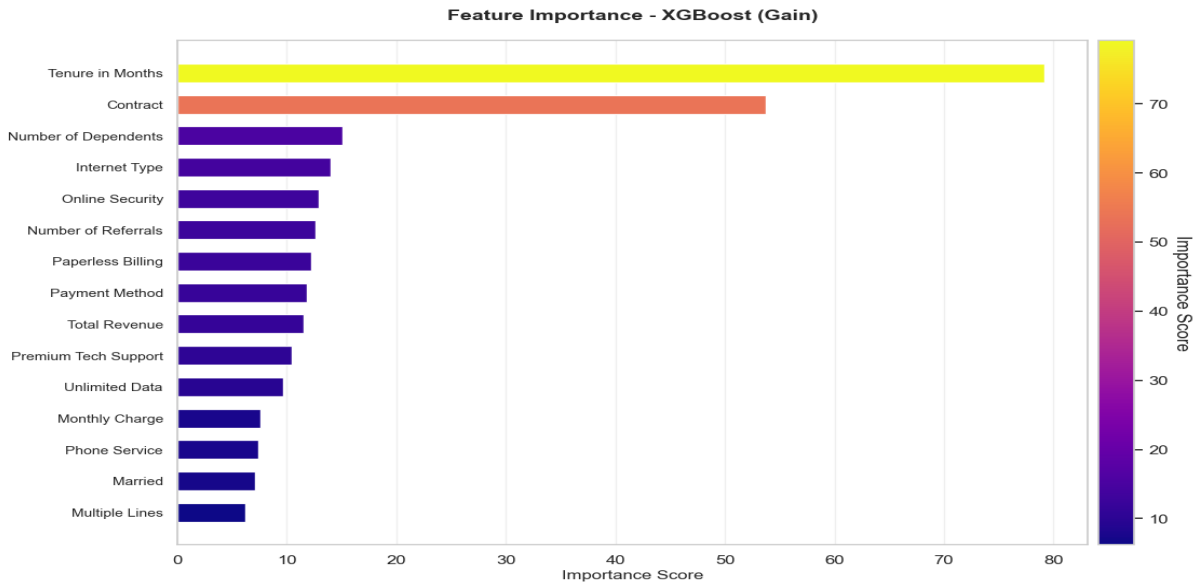
A machine learning model using Extreme Gradient Boosting (XGBoost) achieved a test accuracy of 87.92%, making it a reliable tool for predicting churn. This model can be effectively used to identify customers likely to churn, enabling proactive interventions and preventing revenue loss. It also provides actionable insights for budgeting, planning, and retention efforts.



ROC Curve



Cofusion Matrix



Feature Importance

Business Recommendations

To reduce churn and increase customer lifetime value (CLV), it is recommended to:

• Promote longer-term contracts (e.g., 12- or 24-month plans) to increase customer commitment and reduce churn likelihood.

• Improve fiber optic service experience, focusing on speed, reliability, and support.

• Communicate value clearly to high-paying customers — consider bundling, VIP support, or loyalty benefits.

• Implement customer segmentation strategies to focus on high-value, high-risk segments.

• Launch targeted retention campaigns, such as personalized offers or discounts for at-risk customers.

Future Work

To further improve model accuracy and prediction reliability:

• Explore ensemble learning techniques like stacking and tuned boosting.

• Experiment with advanced models such as LightGBM, CatBoost, or deep learning (ANNs).

• Conduct feature engineering and selection to enhance model interpretability and efficiency.

• Use cross-validation and AutoML tools for robust model development.

• Integrate external data sources (support logs, web activity, sentiment analysis) for deeper insights.

✅ **Actionable Focus:**

Use model predictions to prioritize retention strategies for high-risk customers and implement personalized interventions to maximize loyalty and minimize CLV loss.