SyriaTel Customer Churn Prediction

Using Predictive Analytics to Protect Revenue and Maximize Customer Value

Prepared by:

Morvine Otieno

Overview

- **Goal:**
  Help SyriaTel identify customers who are likely to stop using the service so the company can act early and reduce lost revenue.

- **Why this matters:**
- Losing customers directly impacts revenue
- Retaining customers is cheaper than acquiring new ones
- Early intervention increases the chance of keeping customers

Business and Data Understanding

- **Business Problem**
- SyriaTel needs a way to **spot at-risk customers before they leave,** rather than reacting after churn has already happened.

- **Key Business Question:**
- Which customers are likely to churn soon so SyriaTel can intervene early?

What Data We Used

- We analyzed historical customer records that included:
- Call usage (day, evening, night, international)
- Customer service interactions
- Service plans (international and voicemail)
- Account details
- Each record shows whether the customer **eventually stayed or left.**

Why Classification Is Useful

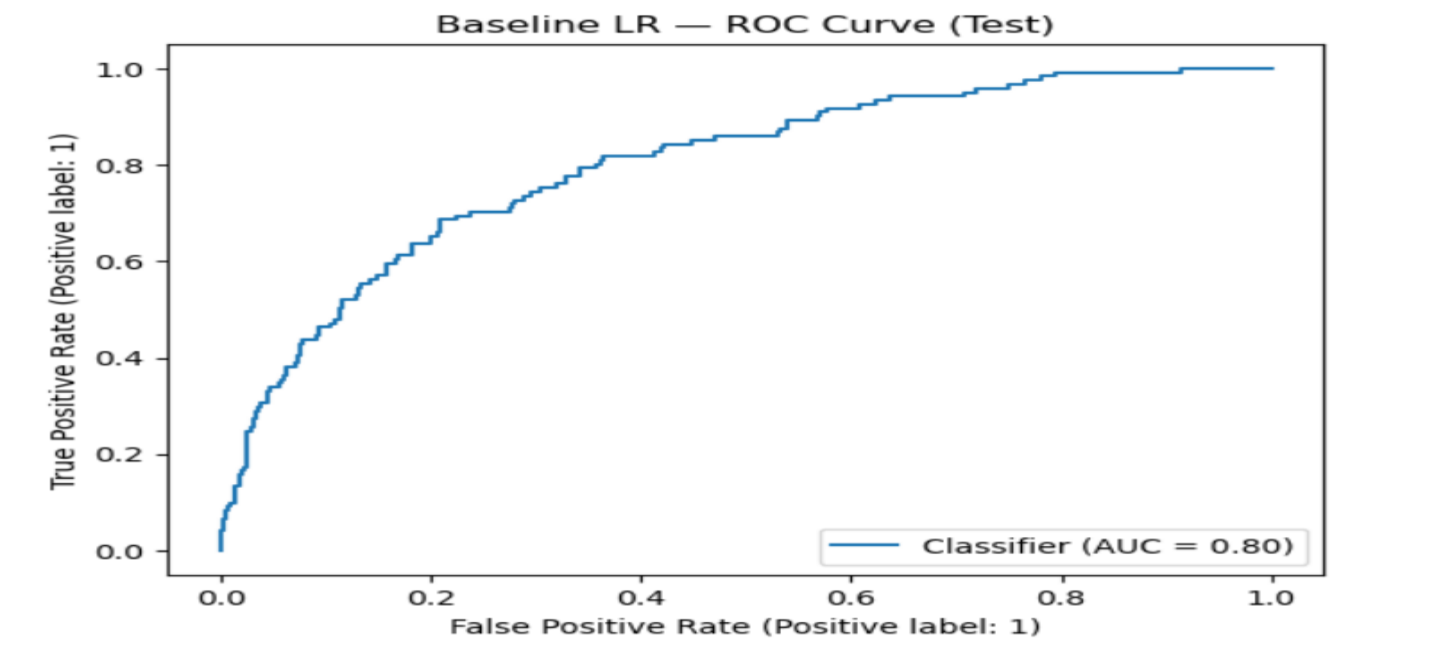
- This is a **yes/no decision:**
- Will the customer churn?
- Or will they stay?
- Classification allows us to **group customers into high-risk and low-risk,** making it easier for business teams to decide **who to contact first.**

Modeling

- **How We Approached the Problem**
- We tested several approaches to predict churn and compared how well they identified customers who actually left.
- Our focus was **not perfection**, but **catching as many future churners as possible.**
- **What "Good Performance" Means**
- For SyriaTel, the most costly mistake is **missing a customer who will churn.**
- So we focused on:
- **Identifying as many future churners as possible**
- Accepting that a small number of contacted customers may not actually leave
- This supports proactive retention.

Baseline Model (Interpretable)

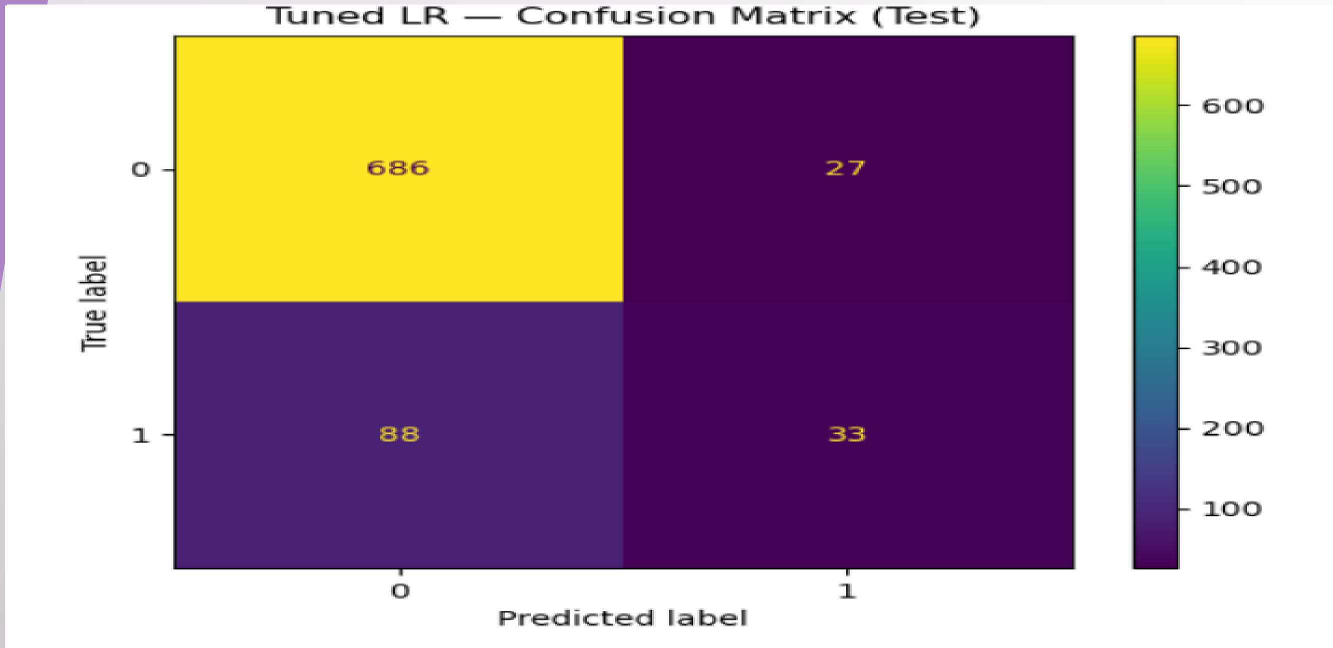
- Baseline: Logistic Regression



- The AUC (Area Under the Curve) is 0.80: This means the model is generally good at ranking risk. If we randomly selected one customer who churned and one who didn't, the model correctly identifies the churner as having the higher risk score 80% of the time.

Model Iteration

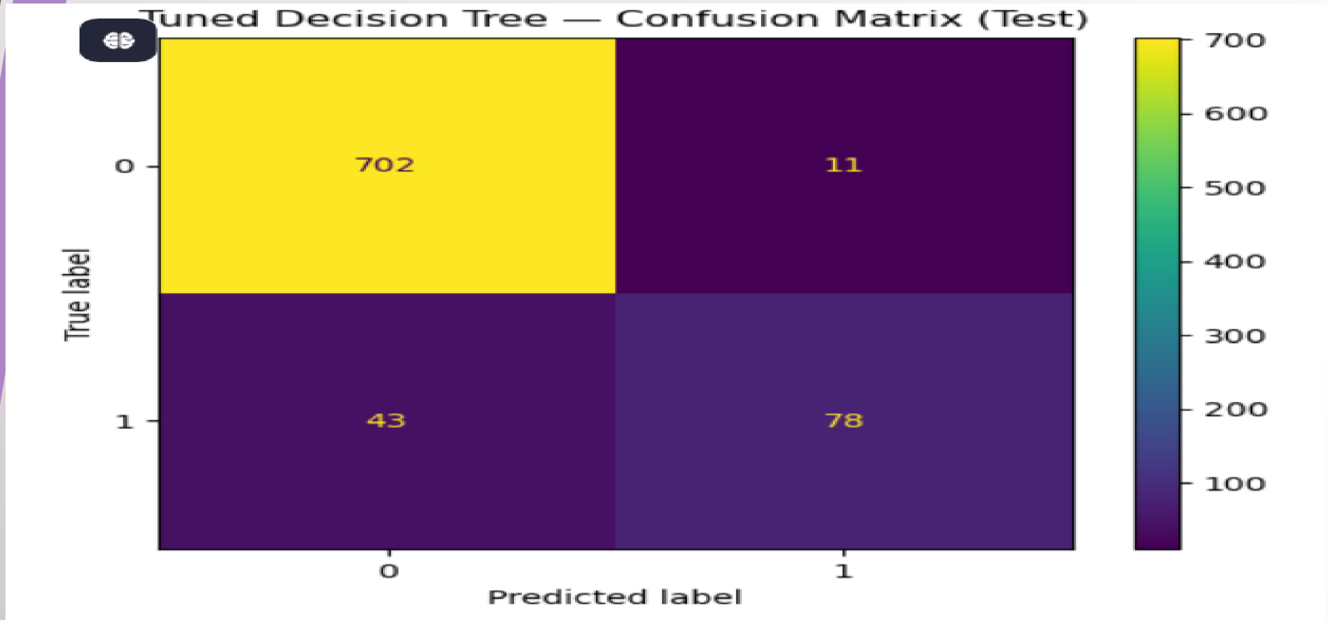
- Tuned Logistic Regression



- This Confusion Matrix shows we missed 88 actual churners (False Negatives) while only successfully identifying 33. This low catch rate of 27.3% (Recall) is what drove our decision to reject this model and move to the superior Decision Tree, which achieved a $64.5% Recall.

Nonparametric Model

- Decision Tree (Tuned)



- The Decision Tree successfully caught 64.5% of all customers who actually left. This is the goal of the project and represents a significant improvement over the 27.3% catch rate of the Logistic Regression model.

Evaluation

	model	recall	precision	f1	roc_auc
2	Tuned Decision Tree	0.644628	0.876404	0.742857	0.852457
1	Tuned Logistic Regression	0.272727	0.550000	0.364641	0.792195
0	Baseline Logistic Regression	0.264463	0.592593	0.365714	0.798813

- Model Comparison (Business View)

We compare test performance across models, prioritizing **churn recall** while monitoring precision and F1.

Why the Final Model Was Chosen

The final model:

- Identified **nearly two-thirds of customers who churned**
- Performed **more than twice as well** as the baseline approach
- Also showed strong accuracy, meaning outreach efforts are efficient
- This makes it the best option for business use.

Key Drivers of Churn (Plain Language)

- Customers are more likely to leave when they:
- Have an **international plan**
- Make **multiple customer service calls**
- Do **not** have a voicemail plan

These are clear signals of dissatisfaction or unmet needs.

Recommendations

1. Target High-Risk Customers First

- Use the model weekly to rank customers by churn risk
- Focus retention efforts on the **top 10-15% most at-risk customers**
- This ensures efficient use of retention budgets.

2. Act on Key Risk Signals

- **International plan customers:**
  Offer better international packages or discounts
- **Frequent customer service callers:**
  Assign priority support and proactive follow-ups
- **No voicemail plan customers:**
  Offer bundled features to increase engagement

3. Use the Model as a Support Tool

- The model should **guide decisions**, not replace human judgment
- Retention teams should combine model insights with customer context

Next Steps

Short-Term Improvements

- Introduce more advanced models to further improve accuracy
- Fine-tune how aggressively customers are contacted

Business Optimization Opportunity

- Measure the cost of:
  - Losing a customer
  - Offering retention incentives
- Adjust outreach thresholds to **maximize profit**, not just predictions

Project Links

- **GitHub Repository** (Technical Project):

https://github.com/Morvine-otieno/phase-3-project/tree/main

Thank You

💬 Questions?

📌 Contact Information

Name: Morvine Otieno

LinkedIn: www.linkedin.com/in/morvine-otieno-14837a180

Email: morvineawuor93@gmail.com