

---

---

# Customer Churn Model

— Adriana Ramirez Franco —  
July 25, 2024

---

---

# Summary

- Develop a **binary classifier** to predict whether a customer will soon stop doing business with SyriaTel.
- Identify **predictable patterns** in customer behavior to reduce losses from customer churn.

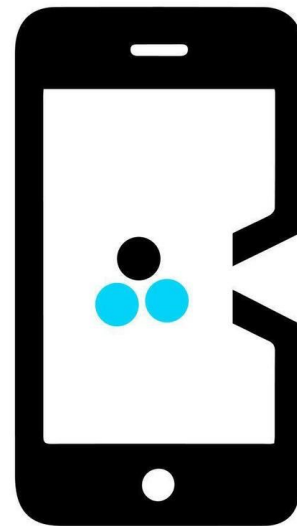


# Outline

- Business Problem
- Data & Methods
- Features of Importance
- Algorithms and their results
- Conclusions

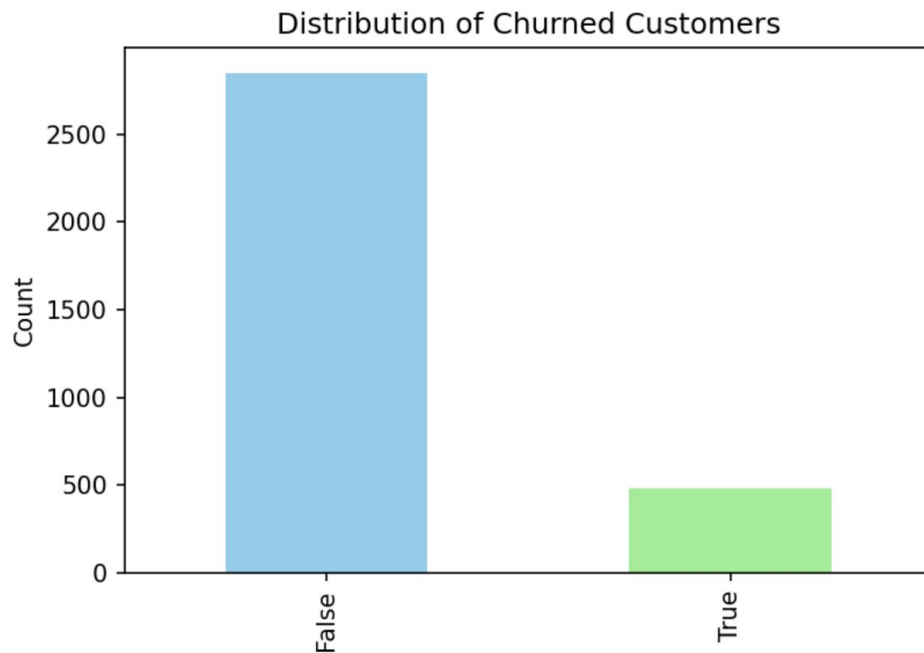
# Business Problem

- SyriaTel aims to **identify factors** contributing to customer churn and **increase retention**.
- **Predict** at-risk customers and **enabling** proactive retention strategies.
- **Achieve** a high recall score of 0.8, **identify** churn factors, and **demonstrate** the model's **effectiveness** in reducing revenue loss.



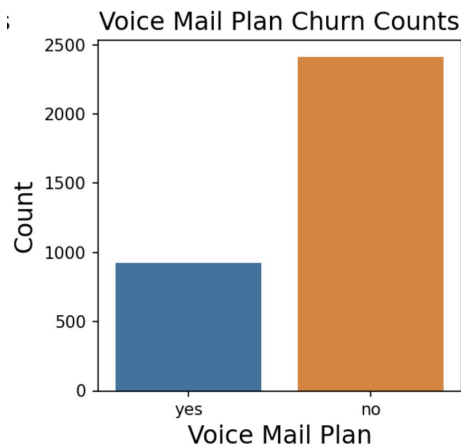
# Data & Methods

- Utilized data from **Kaggle** 'Churn in Telecom's dataset'.
- **No missing or duplicate** features found.
- Target column '**churns**' showed data imbalance.



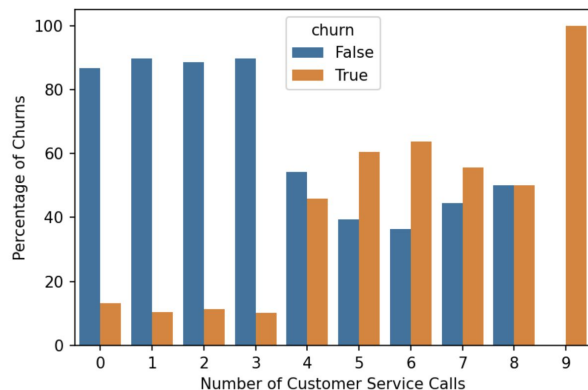
# Features of importance

## 1. Voicemail Plan



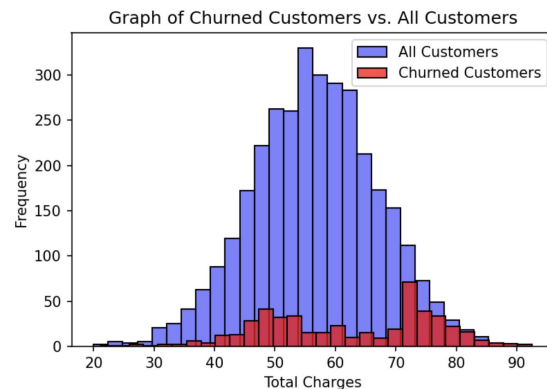
**28%** (922) customers churned.

## 2. Customer Service Calls



Most customers churn when they get from **4 to 9** calls.

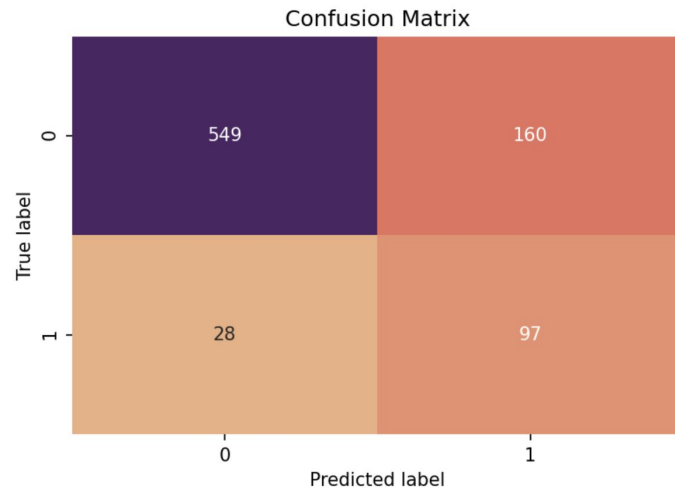
## 3. Total Charges.



Customer churn happens when the charge is **greater than 60\$**.

# 1st Algorithm: Logistic Regression

- The logistic regression model has a **macro average score of 0.78**.
- The model can identify 78% of the actual **positive instances correctly**.

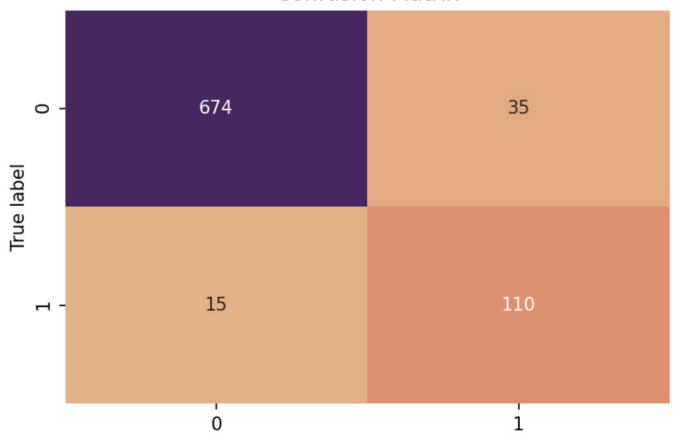


	precision	recall	f1-score	support
0.0	0.95	0.77	0.85	709
1.0	0.38	0.78	0.51	125
accuracy			0.77	834
macro avg	0.66	0.78	0.68	834
weighted avg	0.87	0.77	0.80	834

## 2nd Algorithm: Decision Tree Classifier

- The decision tree model has a macro **average recall score of 0.92**.
- The model can identify around 92% of the actual **positive instances correctly**.

Confusion Matrix



True label	Predicted label					
	0	1	precision	recall	f1-score	support
0	674	35	0.98	0.95	0.96	709
1	15	110	0.76	0.88	0.81	125
accuracy					0.94	834
macro avg			0.87	0.92	0.89	834
weighted avg			0.95	0.94	0.94	834

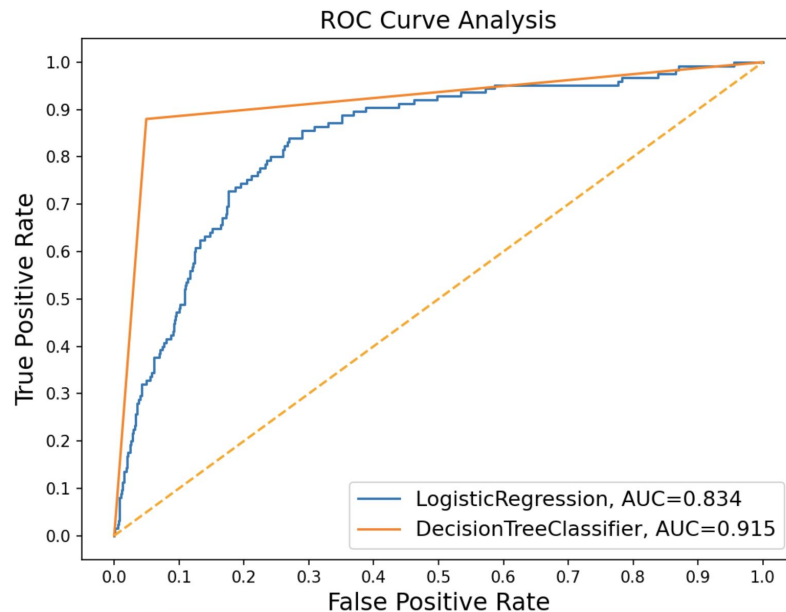


# Conclusions

- The Decision Tree Classifier is the **best prediction model**.

## Recommendations:

- **Reduce** the number of customer service calls.
- Offer adjusted phone plans with **discounts** on voicemail services.
- Provide **additional discounts** to retain and attract customers.



# Thank you!

Email: [aramirezfr20@gmail.com](mailto:aramirezfr20@gmail.com)

Github: [@aramirezfr](https://github.com/aramirezfr)