Predicting Customer Churn for SyriaTel

This project focuses on a binary classification problem aimed at enhancing customer retention for SyriaTel. By leveraging machine learning, we identify customers most likely to churn, supporting proactive retention strategies and protecting vital revenue streams.

The Business Challenge: Understanding Customer Churn

Customer churn is a significant challenge for any telecommunications company, leading to substantial lost revenue and increased customer acquisition costs.

Retaining existing customers is consistently more cost-effective than acquiring new ones. SyriaTel needs a robust method to pinpoint at-risk customers before they decide to leave.

Project Objective: Empowering Retention

Predict Churn Risk

Accurately identify which customers are likely to churn in the near future.

Enable Targeted Strategies

Provide insights to deploy effective and personalized retention campaigns.

Support Data-Driven Decisions

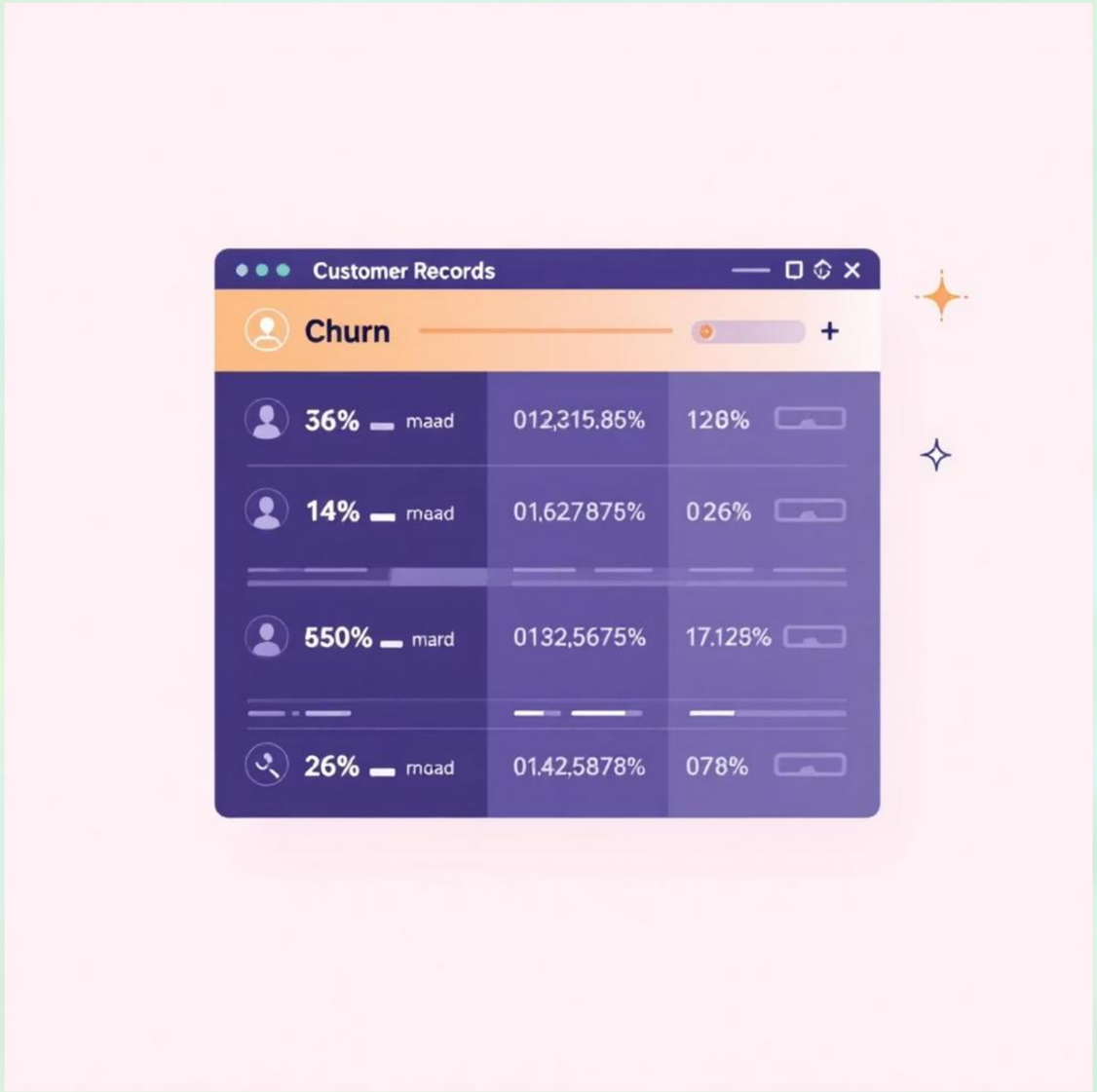
Integrate predictive analytics into customer retention management processes.

Comprehensive Data Overview

Our dataset comprises **3,333 individual customer records**, providing a rich foundation for analysis.

- Each record represents a unique SyriaTel customer.
- Includes detailed usage patterns, such as call duration and data consumption.
- Information on subscribed service plans and billing details.
- Customer service interaction history.

The crucial **target variable** clearly indicates whether a customer has churned or remained active.



Methodology: A Machine Learning Approach

01

Initial Model Selection

Began with simple, interpretable machine learning classification models to establish a baseline.

02

Iterative Performance Improvement

Systematically enhanced model performance through rigorous hyperparameter tuning and feature engineering.

03

Comparative Model Evaluation

Compared multiple models against each other to identify the most robust and effective solution.

Models Under Evaluation

We rigorously evaluated several classification models, both in their baseline configurations and after fine-tuning, to ensure comprehensive assessment.

Logistic Regression

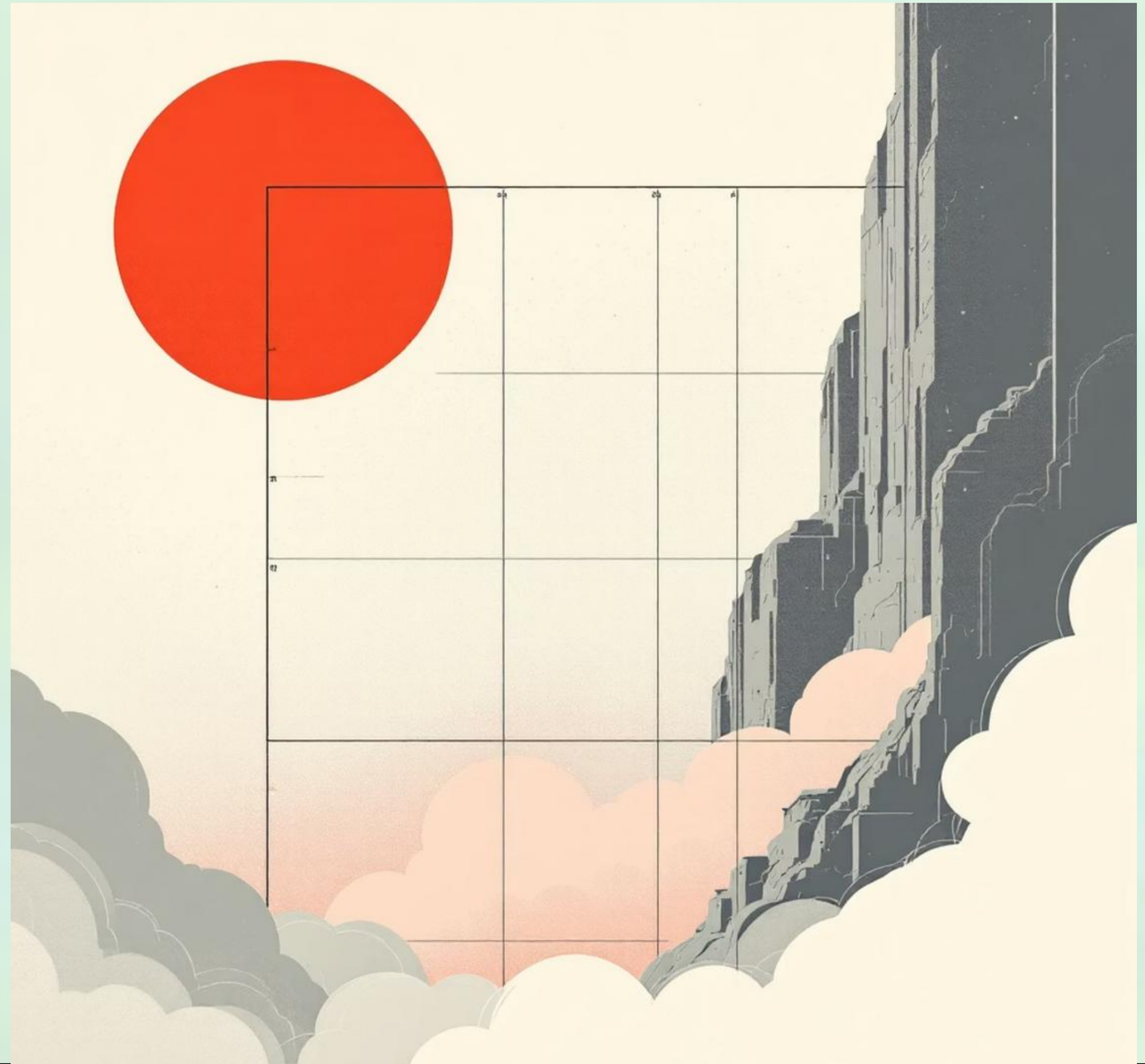
- Baseline performance
- Tuned for optimal parameters

Decision Tree

- Baseline performance
- Tuned for optimal parameters

All models were rigorously evaluated on unseen test data to ensure generalizability and prevent overfitting, focusing on metrics relevant to business objectives.

Measuring Model Performance: Focus on Recall



Our primary performance metric was **Recall**.

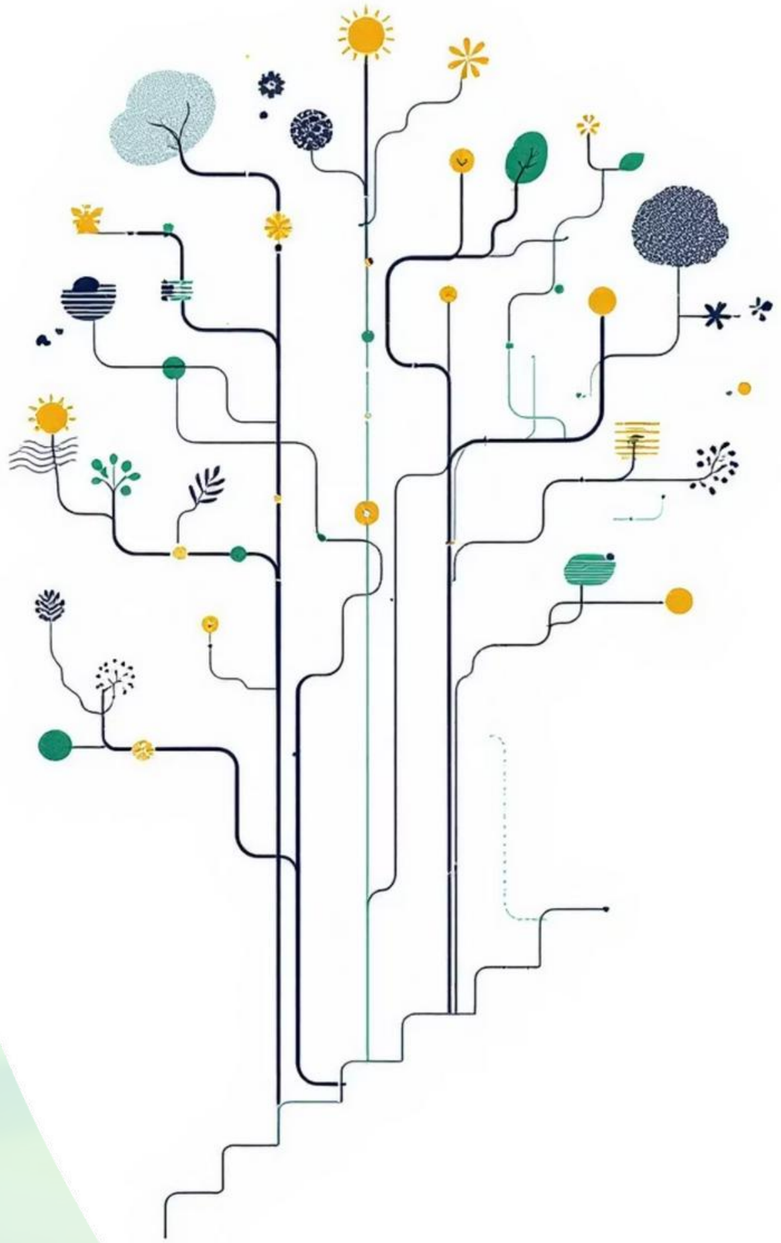
- **Recall** measures the model's ability to correctly identify all positive instances (i.e., customers who actually churn).
- In a churn prediction scenario, missing a churning customer (a false negative) is significantly more costly than incorrectly contacting a non-churner (a false positive).

Secondary metrics were also considered to ensure a balanced and reliable assessment of model performance.

The Champion Model: Tuned Decision Tree

After thorough evaluation and tuning, the **Tuned Decision Tree** emerged as the superior model.

- It achieved the **highest Recall score**, effectively minimizing the number of undetected churners.
- This model demonstrates a robust ability to identify customers most critically at risk of churn.
- It strikes an optimal balance between predictive accuracy and tangible business impact.



Impact for SyriaTel: Proactive Retention

Early Identification

High-risk customers can be identified well in advance of their potential departure.

Targeted Interventions

Retention efforts can be precisely tailored and deployed to the most vulnerable customer segments.

Optimized Resource Allocation

Resources can be strategically focused where they yield the highest return on investment.

Reduced Revenue Loss

Minimize financial impact by preventing unexpected customer churn.

Strategic Recommendations for Implementation

To fully leverage the predictive power of this model, we recommend the following strategic actions:

- **Flag High-Risk Customers:** Systematically identify and tag customers with a high churn probability.
- **Prioritize Retention Campaigns:** Design and deploy specific campaigns (e.g., personalized offers, dedicated support) for these flagged customers.

- **Integrate Predictions:** Embed the churn prediction model directly into existing customer relationship management (CRM) systems and operational workflows.