Kavinyasheilla / **dsc-phase3-project**

Code Issues Pull requests Actions Projects Wiki Security Insights Set

0 stars 0 forks 0 watching Branches Activity
Tags

Public repository

1 Branch 0 Tags Go to file t Go to file Add file Code

Kavinyasheilla Create README.md e81a615 · yesterday

data	Read me	yesterday
Non technical presentation....	Non technical presentation	yesterday
README.md	Create README.md	yesterday
README.md.md	Read me	yesterday
index.ipynb.ipynb	Renamed project notebook	2 days ago
~$Non technical presentati...	Non technical presentation	yesterday

README

SyriaTel Customer Churn Prediction

Sheilla Kavinya Muli

Classifying "at-risk" customers to reduce attrition and increase revenue.

Python 3.8+ Library Scikit-Learn Status Complete

Project Overview

SyriaTel, a telecommunications company, is facing a costly problem: **Customer Churn**. It is significantly more expensive to acquire new customers than to retain existing ones. This project analyzes customer behavior to identify patterns leading to attrition and builds a predictive classifier to flag "at-risk" customers before they leave.

Business Problem

The stakeholders (Marketing & Sales Teams) need a proactive way to minimize revenue loss.

- **The Goal:** Build a binary classification model to predict whether a customer will "soon" stop doing business with SyriaTel.
- **The Strategy:** We optimized for **Recall**.
  - **Why?** A "False Negative" (missing a customer who leaves) is much more costly than a "False Positive" (offering a discount to a happy customer). We want to catch as many churners as possible.

Key Findings (EDA)

Through Exploratory Data Analysis, we identified two major drivers of churn:

1. **The "Rage Quit" Threshold (Customer Service Calls):**
   - Customers who make **more than 3 calls** to customer service have a churn rate of over **45%**, compared to <11% for those with fewer calls.
   - *Insight:* 4 calls is the "tipping point" where patience runs out.
2. **International Plan Issues:**
   - Customers with an International Plan churn at a rate of **42%**, significantly higher than those without.
   - *Insight:* This suggests dissatisfaction with pricing or call quality for international travelers.

Modeling Process

We followed an iterative approach, building three separate models:

1. **Baseline Model (Logistic Regression):**
   - *Result:* High Accuracy (85%) but extremely low Recall (~0.11).
   - *Verdict:* Failed to identify actual churners.
2. **Model 2 (Decision Tree - Balanced):**
   - *Result:* Balanced Recall (~0.57) and Precision.
   - *Verdict:* **Selected as Final Model.** It effectively captures non-linear relationships (like the 3-call threshold) without over-flagging every customer.
3. **Model 3 (Tuned Decision Tree):**
   - *Result:* Perfect Recall (0.99) but poor Precision (0.15).
   - *Verdict:* Rejected. This model predicted "Churn" for almost everyone, which would make retention campaigns too expensive.

Recommendations

Based on the modeling results, we recommend:

1. **Implement a "3-Call" Alert:** Automatically flag any customer who calls customer service for a 3rd time. Agents should be empowered to offer immediate resolution or perks to prevent the 4th call.
2. **Revamp International Plans:** Conduct a price/quality audit of the international plan, as 42% of these users are leaving.
3. **Targeted Retention:** Use the **Balanced Decision Tree** model to score customers monthly. Focus retention budget on the top 20% of high-risk users identified by the model.

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

- **Jupyter Notebook** 100.0%