







SyriaTel-Customer-Churn_Prediction-Analysis Public

  1 Branch  0 Tags  

 Go to file  Go to file + Add file About Code ... 

 **LeoKRono** Update README.md 666d4e2 · 2 minutes ago 

 data	Initial commit	yesterday
 .gitignore	Initial commit	4 days ago
 README.md	Update README.md	2 minutes ago
 index.ipynb	Final commit	44 minutes ago
 notebook.pdf	Add files via upload	26 minutes ago
 presentation.pdf	Add files via upload	41 minutes ago

This project aims to analyse and predict customer churn at SyriaTel (a telecommunications company in Syria) by using Machine Learning algorithms in order to improve their customer retention and reduce financial losses.

 Readme

 Activity

 0 stars

 0 watching

 0 forks

Releases

No releases published

[Create a new release](#)

Packages

 README

SyriaTel Customer Churn Prediction Analysis

This project focuses on predicting customer churn for SyriaTel, leveraging machine learning techniques. The goal is to identify customers likely to churn and understand the key factors driving this behavior.

Project Overview

- **Objective:** To develop a predictive model that accurately identifies customers likely to churn.
- **Business Problem:** Syriatel faces a high churn rate and aims to reduce it by understanding and predicting customer churn. Long-term customer relationships are more effective than acquiring new customers.
- **Data Source:** The dataset is sourced from Kaggle and contains customer information, usage patterns, and churn status.

Data Analysis

- **Dataset Size:** The dataset contains 3,333 records with 21 columns (4 categorical and 17 numerical).
- **Class Imbalance:** The dataset has a class imbalance, with churned customers representing a smaller portion of the data.
- **Feature Correlation:** Low correlation between most features, but a perfect positive correlation between total charge and total minutes at different times. Total day minutes, total day charge, and customer service calls have a weak positive correlation with churn.

Data Preparation

- **Multicollinearity:** Columns showing total charge at different times are dropped to address multicollinearity.

No packages published
[Publish your first package](#)

Languages

- Jupyter Notebook 100.0%

- **Train-Test Split:** Data is split into training and testing sets.
- **Categorical Variables:** Dummy variables are created for categorical features.
- **SMOTE:** SMOTE is used to handle class imbalance problems by oversampling the minority class with replacement.

Modeling and Evaluation

- **Models Used:** Logistic Regression, Decision Tree, and Random Forest algorithms. Hyperparameter tuning is applied to Decision Tree and Random Forest.
- **Evaluation Metric:** Recall score is used to evaluate model performance.
- **Best Model:** Decision Tree with tuned hyperparameters resulted in the lowest number of false negatives and the highest recall score. XGBClassifier also showed promising results.
- **Feature Importance:** The most important features for predicting customer churn are total day minutes, total evening minutes, customer service calls, and total international minutes.

Key Findings and Recommendations

- **High Churn Areas:** Offer discounts or promotional offers to customers in area codes 415 and 510, as these areas have a higher churn rate.
- **Customer Service:** Improve customer service quality and reduce the number of customer service calls.
- **Pricing Structure:** Evaluate the pricing structure for day, evening, night, and international charges.
- **Retention Strategies:** Focus on customer retention strategies in states with higher churn rates, such as Texas, New Jersey, Maryland, Miami, and New York.

- **Voicemail Plans:** Enhance the value proposition of the voicemail plan to increase adoption among customers.
- **Call Charges:** The company should look into the call charge rates in comparison to the competitors, and consider if they should lower the charges of calls per minute to prevent customers from churning.
- **Customer Service Follow-Up:** Syriatel should ensure effective customer service to meet customer expectations and analyze customer interactions, following up on both positive and negative feedback.

Repository Structure

├── README.md <- The README for reviewers of this project
├── index.ipynb <- Narrative documentation of analysis in Jupyter Notebook
├── presentation.pdf <- PDF version of project presentation
└── telecom.csv <- Dataset used in the analysis

Next Steps

- Increase the training data size to reduce overfitting and improve model performance.
- Undertake further feature engineering to boost the recall score.

For More Information

- Review the full analysis in the [Jupyter Notebook](#).
- View the project [presentation](#).