

RoseKyalo /
Phase-III-Project

<> CodeIssuesPull requestsActionsProjectsWikiSecurityInsightsSet

MIT license

0 stars 0 forks 1 watching Activity

Public repository

master

BranchesTags

RoseKyalo Changes ...

11 minutes ago5

View code

README.md

SyriaTel Customer Churn Analysis



Image credit: [istockphoto](#)

Table of Contents

1. [Overview](#)

2. [Business Objectives](#)

3. [Data Understanding](#)

4. [Methodology](#)

[Exploratory Data Analysis \(EDA\)](#)

[Modeling](#)

[Final Model](#)

5. [Recommendations](#)

6. [Getting Started](#)

[Prerequisites](#)

[Installation](#)

[Usage](#)

7. [Contributing](#)

8. [License](#)

9. [Acknowledgments](#)

Overview

The SyriaTel Customer Churn Analysis project is aimed at conducting a comprehensive examination of customer data. Its primary goals are to better understand and reduce customer churn rates while also increasing revenue. This project plays a pivotal role in enhancing SyriaTel's competitiveness in the telecommunications industry. Its primary focus lies in predicting customer churn and pinpointing the factors contributing to it.

Business Objectives

i). Reduce Churn Rate

SyriaTel's primary objective is to minimize the number of customers churning. By identifying the factors that contribute to customer churning, they aim to take proactive measures to improve customer retention. Predictive analytics will play a crucial role in identifying potential churners, allowing SyriaTel to intervene in a timely manner.

ii). Revenue Enhancement

SyriaTel is committed to boosting its revenue. By comprehensively understanding the customer churn dynamics, they intend to implement strategic measures that will lead to increased revenue.

iii). Price Plan Optimization

SyriaTel seeks to optimize its pricing plans based on data-driven insights. Understanding the relationship between pricing, customer satisfaction, and churn is critical for achieving this objective.

iv). Network Quality Assessment


Identifying network issues is essential to maintain service quality. SyriaTel is dedicated to addressing these issues to improve the overall customer experience and reduce churn caused by network-related problems.

Data Understanding

The project is based on SyriaTel's customer data, which includes information on customer service calls, call minutes, charges, and customer churn status. The dataset provides valuable insights into customer behavior and churn patterns.

Methodology

Exploratory Data Analysis (EDA)

During the exploratory data analysis phase, we identified that customer service calls, total day minutes, and total day charges were the critical features affecting customer churn. It was also observed that an increase in total charges corresponded to higher churn rates. The correlation matrix revealed a strong positive relationship between "total day minutes" and churn, indicating network quality issues during daytime usage. 

Modeling

In the modeling phase, we employed advanced machine learning techniques, including the XGBoost model. This model was chosen for its ability to identify at-risk customers and enable targeted retention strategies. Data preprocessing, feature engineering and model tuning were performed to improve model performance. Our model evaluation focused on two key metrics: accuracy and F1 score. These metrics provide a comprehensive assessment of our models' performance.

The table below table concisely summarizes the models, their best hyperparameters, accuracy, and F1 scores for both churned and non-churned customers.

Model	Best Hyperparameters	Accuracy	F1-Score (Churned)	F1-Score (Non - Churned)
Logistic Regression	{'C': 100, 'penalty': 'l2', 'solver': 'liblinear'}	0.62	0.35	0.73
Decision Tree	{'Criterion': 'gini', 'Max Depth': 5, 'Random State': 42}	0.70	0.39	0.80
Bagging Classifier	{'estimator__max_depth': 15, 'max_features': 1.0, 'max_samples': 1.0, 'n_estimators': 30}	0.79	0.44	0.87
Random Forest	{'max_depth': 15, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 30}	0.78	0.43	0.86
XGBoost	{'learning_rate': 0.1, 'max_depth': 7, 'n_estimators': 200}	0.83	0.47	0.90

Final Model [↗](#)

The final model, an XGBoost-based machine learning model, exhibited strong performance in identifying potential churners. It will be used to predict customers at risk of churning, allowing SyriaTel to implement retention strategies effectively.

Recommendations [↗](#)

To achieve the business objectives:

- SyriaTel should improve customer service quality to reduce churn rates.
- Optimize pricing strategies to strike a balance between retaining customers and ensuring profitability.
- Address network quality issues, particularly during the day, to provide better service and reduce churn.
- Implement and continuously monitor the XGBoost model to identify at-risk customers and implement targeted retention strategies.

Getting Started [↗](#)

Follow these instructions to get a copy of the project up and running on your local machine for development and testing purposes.

Prerequisites [↗](#)

Make sure you have the following software installed on your machine:

- [Python](#)
- [Jupyter Notebook](#)

Installation [↗](#)

1. Clone the repository:

```
git clone https://github.com/your-username/syria-tel-churn-analysis.git
```



2. Change to the project directory:

```
cd syria-tel-churn-analysis
```



3. Install the required Python libraries:

```
pip install -r requirements.txt
```



Usage [↗](#)

The project code is organized into Jupyter notebooks and Python scripts. You can explore the analysis and model building by opening the notebooks in Jupyter. If you're keen to examine the final model, it's securely stored as "train_model.pkl" for your convenience.

1. To execute a notebook, run:

```
jupyter notebook notebook_name.ipynb
```



Follow the instructions within the notebooks to replicate the analysis and run the models.

Folder Structure [↗](#)

Phase III Project/



```
|
|├── data/
|│   ├── syria_tel_data.csv
|│   └── ...
|├── notebooks/
|│   ├── SyriaTel analysis
|│   └── ...
|├── non-technical presentation/
|│   ├── powerpoint
|│   └── ...
|
```

```
|— README.md
|— requirements.txt
|— train_model.pkl
|
```

Contributing [↗](#)

If you wish to contribute to this project, please open an issue or create a pull request. We welcome any improvements or suggestions.

License [↗](#)

This project is licensed under the MIT License - see the LICENSE file for details.

Acknowledgments [↗](#)

This project was completed as part of a data analysis course. We would like to acknowledge the creators of the open-source libraries and tools used in this project.

Releases

No releases published

[Create a new release](#)

Packages

No packages published

[Publish your first package](#)

Languages

● Jupyter Notebook 100.0%