

- Introduction
- Features
- <u>Installation</u>
- Usage
- Exploratory Data Analysis
- Modeling and Evaluation
- Feature Importance
- Recommendations
- Contributing
- License
- Contact

# Introduction

SyriaTel, a telecommunications company, is experiencing customer churn, which impacts its revenue and growth. This project aims to predict whether a customer will soon stop using SyriaTel's services by analyzing various features such as customer demographics, usage patterns, and service details. The goal is to provide actionable insights and recommendations to SyriaTel's management to implement targeted retention strategies and improve customer satisfaction.

# **Features**

- Predicts customer churn with high accuracy
- Analyzes customer data including usage patterns and service details
- Provides insights into factors contributing to churn
- Visualizes data to understand churn patterns

#### Installation

To install and set up the project, follow these steps:

```
ďЪ
# Clone the repository
git clone https://github.com/shuruebale/syriatel-customer-churn.git
cd syriatel-customer-churn
# Install required packages
pip install -r requirements.txt
To use the project, run the following commands:
# Prepare the data
python prepare_data.py
# Perform exploratory data analysis
python exploratory data analysis.py
# Train the model and make predictions
python train_model.py
## Exploratory Data Analysis
To perform exploratory data analysis, use the `exploratory_data_analysis.py` script. This script will generate visualizations to
- **Distribution of Total Day Minutes by Churn**:
  - Churned customers tend to have lower total day minutes.
  - Non-churned customers have higher total day minutes.
- **Churn Rates Across Different States**:
  - States with higher proportions of churned customers indicate higher churn rates.
  - Regional patterns and outliers can be observed, providing insights into geographic factors influencing churn.
## Modeling and Evaluation
To train and evaluate the model, use the `train_model.py` script. This script trains both logistic regression and random forest
- **Logistic Regression Model**:
  - Decent accuracy but struggles with identifying churned customers.
- **Random Forest Model**:
  - Outperforms logistic regression in terms of accuracy.
 - Better performance in identifying churned customers.
The random forest model is recommended for predicting customer churn due to its higher accuracy and better performance in ident:
## Feature Importance
The top contributing factors for predicting customer churn identified by the random forest model are:
1. **Total Day Charge**: 13.03%
2. **Total Day Minutes**: 11.82%
3. **Customer Service Calls**: 10.47%
4. **International Plan**: 6.55%
5. **Total Evening Minutes**: 5.71%
6. **Total Evening Charge**: 5.61%
7. **Total International Minutes**: 4.25%
8. **Total International Charge**: 4.18%
```

```
9. **Total International Calls**: 4.01%
10. **Total Night Charge**: 3.88%
## Recommendations
Based on the analysis, the following recommendations are made:
1. **Optimize Pricing Plans**:
   - Focus on optimizing pricing plans and service quality during peak hours to retain customers.
2. **Improve Customer Service**:
   - Prioritize improving customer service experiences and resolving issues promptly.
3. **Tailor International Plans**:
   - Better serve customers with international calling needs by tailoring offerings and communication strategies.
4. **Enhance Evening Packages**:
   - Offer competitive pricing and attractive packages for evening usage.
5. **Analyze International Usage**:
   - Tailor international calling plans to improve satisfaction among international callers.
## Contributing
Contributions are welcome! Please open an issue or submit a pull request for any feature requests or bug fixes.
## License
This project is licensed under the MIT License.
## Contact
For any questions or inquiries, please contact [shuruebale]
(mailto:shuruxebale@gmail.com).
```



### Releases

No releases published Create a new release

# **Packages**

No packages published

Publish your first package

# Languages

Jupyter Notebook 100.0%