RWKarimi / **Phase_3_Project**

<> **Code**   ⊙ Issues   ⅋ Pull requests   ▷ Actions   ▦ Projects   📖 Wiki   ⚠ Security   ∿ Insights   ⚙ Settings

☆ **0** stars   ⅋ **0** forks   👁 **0** watching   ⅋ **Branches**   ∿ Activity
                                              🏷 **Tags**

🌐 **Public repository**

⅋ m... ▾   ⅋ **2 Branches**   ⬦ **0 Tags**   ⅋   🏷   Go to file   t   Go to file   +   Add file ▾   Code   ···

🟢 **RWKarimi** Merge pull request #2 from RWKarimi/ft-Jupyter-Notebook •••   14171ce · 13 minutes ago   🕐

🗋 .gitignore	ch-converting-files-to-pdf	17 minutes ago
🗋 README.md	ch-fix-grammar-mistakes-in-notebook	34 minutes ago
🗋 SyriaTel_Churn_Prediction.ipynb	ch-converting-files-to-pdf	17 minutes ago
🗋 SyriaTel_Churn_Prediction.pdf	ch-converting-files-to-pdf	17 minutes ago
🗋 Syriatel.pdf	ch-converting-files-to-pdf	17 minutes ago
🗋 Telecom's data.csv	Initial commit	3 days ago
🗋 archive.zip	Initial commit	3 days ago

📖 **README**   ✏ ☰

SyriaTel Customer Churn Prediction

🎯 SyriaTel Churn Prediction

Open in `Colab`   Python `3.8+`   Framework `scikit-learn`

Overview

This project focuses on predicting customer churn for **SyriaTel**, a leading telecommunications operator in Syria. By leveraging machine learning techniques on historical usage and billing data, we aim to identify customers who are at high risk of leaving, enabling targeted retention strategies and improving overall business performance.

Business and Data Understanding

Business Context

SyriaTel faces intense competition in the telecom sector and aims to reduce customer attrition, which directly impacts revenue and market share. Identifying likely churners before they leave allows the company to deploy proactive retention offers and personalized engagement.

Stakeholder Audience

- **Senior Management:** To understand projected churn rates and ROI on retention campaigns.
- **Marketing & Customer Care Teams:** To design targeted promotions for at-risk customers.
- **Data Science Team:** To validate model performance and integrate into a production pipeline.

Dataset Choice

The dataset includes customer demographics, service usage metrics (daytime, evening, nighttime and international call minutes and charges), customer service interactions, and contract details. It provides both behavioral and transactional features necessary for robust churn prediction.

Modeling

We experimented with two supervised classification models:

1. **Logistic Regression**

   - Feature selection based on correlation and domain knowledge.
   - Standardization of numerical variables.
   - Use of SMOTE to address class imbalance.

2. **Decision Tree Classifier**

   - Default Gini impurity criterion.
   - Pruning parameters tuned via cross-validation to prevent overfitting.

Each model was trained on a stratified train-test split and evaluated on the holdout set.

Evaluation

Key performance metrics for the models (Churn = 1 class):

Metric	Logistic Regression	Decision Tree
Accuracy	0.81	0.92
Precision	0.39	0.76
Recall	0.57	0.62
F1-Score	0.47	0.68
AUC-ROC	0.76	0.85

Best Model: Decision Trees demonstrated the highest balance of precision and recall, making it the preferred choice for deployment.

Conclusion

- **Recommended Approach:** Deploy the decision tree model for real-time churn scoring.

- **Business Impact:** Targeted retention campaigns could reduce churn by an estimated 10–15%, translating to significant revenue savings.

- **Next Steps:**

   i. Integrate the model into SyriaTel's CRM system for live scoring.

ii. Monitor model drift and recalibrate quarterly.

iii. Explore ensemble methods or more complex algorithms (e.g., Random Forest, Gradient Boosting).

Presentation Slides

> *Presentation link Syriatel presentation.pdf*

Running the Notebook

Local Environment

1. Clone the repository:

```
git clone https://github.com/your-username/SyriaTel-Churn-Prediction.git
cd SyriaTel-Churn-Prediction
```

2. Create a virtual environment and install dependencies:

```
python3 -m venv venv
source venv/bin/activate  # On Windows: venv\Scripts\activate
pip install -r requirements.txt
```

3. Launch Jupyter Notebook:

```
jupyter notebook
```

Google Colab

1. Upload `SyriaTel_Churn_Prediction.ipynb` to Google Drive.

2. Open the notebook in Colab.

3. Install required packages by running at the top cell:

```
!pip install -r https://raw.githubusercontent.com/your-username/SyriaTel-Churn-Prediction/main/requirements.txt
```

4. Ensure data files are accessible (upload to Colab or mount Drive).

5. Run all cells to reproduce the analysis.

For detailed methodology, refer to the Jupyter Notebook. For any questions or feedback, please contact the data science team.

Made with ❤️ using scikit-learn, pandas, and Matplotlib in Jupyter Notebook

Author

Ryan Karimi

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

● **Jupyter Notebook** 100.0%