🖥 **pkonsomu2020** / **dsc-phase_3_project** (Public)

End of phase 3 project in Moringa School.

⭐ **0** stars   ⑂ **0** forks

☆ Star   👁 Unwatch ⌄

| ⟨⟩ Code | ⊙ Issues | ↕ Pull requests | ▶ Actions | 🗔 Projects | 📖 Wiki | ⋯ |

ᛘ main ⌄   ⋯

Peter Onsomu ⋯	1 minute ago 🕘
📁 .ipynb_checkpoints	1 minute ago
📄 Presentation.pdf	1 minute ago
📄 README.md	1 minute ago
📄 customerchurnbg.jpg	1 minute ago
📄 customerchurndata.csv	4 days ago
📄 student - Jupyter Notebook.pdf	1 minute ago
📄 student.ipynb	1 minute ago

☰ README.md   ✏

Phase_3 Project Prediction

Overview

Churn is a measure of how many customers stop using a product or service due to some certain reasons making them not to use them. It is calculated by dividing the number of customers who have stopped using the product or service by the total number of customers who have ever used the product or service. The churn rate is usually expressed as a percentage. A high churn rate can indicate that there is a problem with the product or service. It can also be a sign that the company is not doing a good job of retaining customers.

In this project, we will try to determine the churn rate for future customers and find a way of minimizing it. This project basically, focuses on predicting customer churn for SyriaTel, a telecommunications company. The main goal is to develop a classifier that can identify customers who are likely to stop doing business with SyriaTel in the near future. By anticipating churn, SyriaTel can take proactive measures to retain customers, reduce financial losses, and improve customer satisfaction.

Business and Data Understanding

Stakeholder audience

The primary audience for this project is SyriaTel's telecom business itself. The stakeholders include executives, marketing teams, and customer relationship management teams who are interested in reducing churn and optimizing customer retention strategies. The insights gained from the analysis will help the stakeholders understand the patterns and factors influencing churn and enable them to make data-driven decisions to mitigate customer churn.

Dataset Choice

The dataset used for this project contains historical customer data from SyriaTel, including various features such as customer demographics, usage patterns, account information, and customer churn status. The dataset provides valuable insights into customer behavior and characteristics that can be leveraged to predict churn. By analyzing this dataset, we aim to identify predictable patterns and develop a robust churn prediction model. The dataset of our project has **3,333** number of rows and **21** number of columns.

Modeling

In the modeling phase, we employ several classification algorithms to predict customer churn. The following models were evaluated:

- Baseline Model: Logistic Regression
- Decision Tree Classifier
- K-Nearest Neighbors (KNN) Model
- Random Forest Classifier

These models were chosen due to their suitability for binary classification tasks and their ability to handle different types of data. Each model was trained using a subset of the dataset and evaluated based on their performance metrics.

Evaluation

The models were evaluated using various metrics to assess their performance and effectiveness in predicting customer churn. These metrics include accuracy, precision, recall, and F1-score. Additionally, confusion matrices were used to visualize the model's performance in terms of true positives, true negatives, false positives, and false negatives.

Conclusion

Based on the evaluation, the random forest classifier emerged as the most effective model for predicting customer churn in SyriaTel. It consistently achieved the highest accuracy and precision, indicating its ability to accurately identify potential churners. By implementing the random forest model, SyriaTel can gain insights into customer behavior and preferences, enabling them to take proactive measures to retain customers and reduce financial losses.

Further steps could involve fine-tuning the selected model, conducting feature engineering, and regularly monitoring and retraining the model to adapt to changing customer behavior. The insights gained from this analysis can guide targeted marketing campaigns and personalized retention strategies to optimize customer retention efforts.

By leveraging the predictive power of machine learning, SyriaTel can make data-driven decisions and enhance their customer retention strategies, ultimately improving customer satisfaction and business profitability.

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

● **Jupyter Notebook** 100.0%