Churn Analysis in Telecom Industry

Business Overview

Churn analysis is a critical process for telecom companies as it helps them identify customers who are likely to leave the service provider and take their business elsewhere. It involves analyzing customer behavior and usage patterns to identify factors that lead to churn and developing strategies to retain customers.

Churn analysis is crucial for the telecom industry as it is directly related to the profitability and sustainability of a telecom company. A high churn rate can lead to a decline in revenue and market share, while a low churn rate can help companies retain their customers and gain a competitive edge. It is vital for the telecom industry to analyze the churn rate, understand the reasons for it, and take the necessary steps to retain customers.

Why Data Science?

Data science plays a crucial role in churn analysis in telecom. Data scientists use a range of techniques, including machine learning and predictive analytics, to analyze customer data and develop models that predict which customers are at risk of churning. These models can then be used to develop targeted retention strategies and improve overall customer satisfaction.

Key Benefits of Churn Analysis

- Improved customer retention: Churn analysis helps telecom companies identify customers who are at risk of leaving and develop strategies to retain them, which can help improve customer retention rates

- Increased revenue: By retaining more customers, telecom companies can increase their revenue and profitability

- Improved customer satisfaction: Telecom companies can improve their overall customer satisfaction by analyzing customer data and identifying factors that lead to churn

- Competitive advantage: Telecom companies can gain a competitive advantage over their rivals by reducing churn rates and improving customer retention

Overall, churn analysis is a critical process for telecom companies that helps them understand customer behavior and develop strategies to retain their customers. By leveraging data science and analytics, telecom companies can gain valuable insights into customer behavior and improve their overall business performance.

Aim

1. To build a predictive model for churn in the Telecom industry
2. To build a feedback loop to help maintain the quality of results

Data Description

The telecom company from the US provided the data with 98230 customers over 73 unique features. The features are related to customer demographics, personal information, and usage.

Tech stack

- Language - Python, SQL
- Cloud - AWS
- Libraries
    - Pandas: For Data Analysis and Manipulation
    - Pyodbc : For connecting with aws cloud to fetch the data

- Numpy: For performing mathematical operations over data
- matplotlib, seaborn: For Data visualization
- scikit-learn,xgboost : For model building
- deepchecks: for continuous model monitoring

Approach

In this project, we will tackle the problem of churn faced by telecom companies which have a direct impact on their revenue and profitability.

We will start by exploring the data helping us to create model-ready data. We will train various models such as logistic regression, random forest classifier, xgboost classifier etc.

Finally, after the model is built we will check for data drift and model drift to monitor the model in the future. We will also create a feedback loop for better model performance.

Project Takeaways

1. How to load data from AWS SQL using pyodbc and pandas
2. Missing Value Imputation
3. Outlier detection and Imputation
4. Exploratory Data Analysis
5. Categorical Feature Encoding
6. Importance of Feature Scaling
7. Building Base Model with logistic regression
8. Training XGBoost Model
9. Hyperparameter Tuning
10. What is Data Drift?
11. What is Model Drift?
12. How to check drift using deepchecks?
13. Feedback loop for model retraining