

Data Churn Prediction

Instructions

1. Conduct a preliminary analysis of the data and share some key insights as well as statistical matrices. Please use SQL and comment on the code.
2. Look at how we can use insights and analytics to drive down data churn. Use the dataset to prepare a 15 mins presentation to the leadership on your findings.
3. Construct a model predicting data droppers (subscribers that will stop using data) using the dataset shared.
 - a. The purpose of this model is to predict if a customer will use data in the coming week.
 - b. The model scoring will set Monday as day 1 of the predicted window and Sunday as the 1st day of the data input window.
 - c. Evaluate at least 3 different models.
 - d. Compare the model's performance and suggest which model is recommended.
 - e. Complete this using any modeling tool of your choice including commentary in code.
4. Based on the data shared and the work done above, segment the results from your chosen model and recommend what targeted campaigns can be conducted and what is the expected business impact.

Notes

- Data consumption is the life blood in the telecoms industry.
- There exists scenarios where subscribers reduce and stop using data, this is referred to as data churn.
- The dataset contains data from a mobile telecoms operator Yellow mobile.

- The data are for 6 weeks, and a data dictionary is attached.
- In the Dictionary, "week xx" represents the week number ranging from 01 to 05, where 05 is the most recent week and 01 is the last week of the input window. i.e a variable with week xx will occur 5 times