Jeff Nguyen
Jn6298
05NOV2018
SDS358

Regression Project Proposal

At Charter Communications, actionable insights are extracted from available data to improve the customer experience. Through exploratory data analysis, pain points can be identified allowing a Line of Business areas to target and improve.

Due to Charter policies company data cannot be used for this project, however analogous data is fortuitously available on Kaggle. Logistic regression will be used to determine: "How do the factors of Total Charges, Streaming Movies, Tech Support, Phone Service, Tenure, Gender, Multiple Lines, and Internet Service determine whether or not a customer churns?"

We want to see if Churn (If a customer drops services) can be predicted by:

- Total Charges (Quantitative, Continuous)
- StreamingMovies (Categorical, Dichotomous)
- TechSupport (Categorical, Dichotomous)
- PhoneService (Categorical, Dichotomous)
- Tenure (Quantitative, Discrete)
- Gender (Categorical, Dichotomous)
- MultipleLines (Categorical, Ordinal)
- Internet Service (Categorical, Ordinal)

By using logistic regression, the relationship between variables can be described allowing for greater insight as to how individual explanatory variables relate to Churn. This insight can possibly allow for a Line of Business to create strategies to retain customers that intend to cancel services and improve existing services to ensure happy customers.

Bulleted Summary can be found on Page 2

Jeff Nguyen
Jn6298
05NOV2018
SDS358

Regression Project Proposal

Research Question

"How do the factors of Total Charges, Streaming Movies, Tech Support, Phone Service, Tenure effect, Gender, Multiple Lines, and Internet Service determine whether or not a customer churns?"

Technique you'll use to answer the question

- Logistic Regression

Outcome Variable, Variables of Interest, Nuisance Variables

Outcome:

- Churn (Categorical, Dichotomous)

Variables of Interest:

- Total Charges (Quantitative, Continuous)
- StreamingMovies (Categorical, Dichotomous)
- TechSupport (Categorical, Dichotomous)
- PhoneService (Categorical, Dichotomous)
- Tenure (Quantitative, Discrete)
- Gender (Categorical, Dichotomous)
- MultipleLines (Categorical, Ordinal)
- Internet Service (Categorical, Ordinal)

Where you'll be getting data

Kaggle - "Telecom Churn Dataset (IBM Watson Analytics)"

https://www.kaggle.com/zagarsuren/telecom-churn-dataset-ibm-watson-analytics#WA_Fn-UseC_-Telco-Customer-Churn.xlsx