TELCO CUSTOMERS CHURN

Shuchita Mishra

ABSTRACT

Customer churn analysis helps us understand the journey a customer goes through on the service and the key areas where they fall off, which in turn helps us understand where we can improve our strategies and increase brand loyalty. This project proposes to predict the customer churn for a telecom company using predictive modeling.

INTRODUCTION

Customer churn is the rate at which a business's customers leave the company (Fig1). For a telecom company like "Interconnect", this makes a significant business impact directly affecting the profits. Focusing on long-term relationships and observing customer behavior is more cost-effective than the marketing campaigns to attract new customers [1].

It costs 25 times more to acquire a new customer as compared to retaining customers. This alone shows why tracking customer churn is so important. Also, the more we learn about customer behavior, the better understanding we'll gain of the future expected revenue. For this reason, large telecommunications corporations are seeking to develop models to predict which customers are more likely to change and take actions accordingly.

We would like to be able to predict the churn of clients for this company.

In this project, I aim to analyze the dataset [2] and create ML classifier models to perform "Uplift Modeling" by targeting potential customers with the intention of reducing marketing costs while preserving the profit margins.

FUTURE WORK

This prediction is used to identify the factors which influence churning of clients so that the company can offer incentives for them to stay. I will build a logistic regression classifier which will be used as a baseline model. Other models such as Naive Bayes, KNN, Decision Trees, Random Forest and Support Vector Machine will be implemented along with cross validation to identify the best hyperparameter values. Then the performance of these models will be evaluated based on the AUC score generated by the ROC curve with the help of the confusion matrix.

Logistic regression uses the sigmoid function and a predefined threshold value (set to 0.5 by default) to classify the data points to one of the existing labels/classes. In this case, I will use the Logistic regression algorithm on the dataset features $(x_1, x_2 \ldots x_d)$ to classify them as Churn (Yes) or Not Churn (No). The equation for Logistic regression cost function using Lasso regularization for feature selection is as follows:

$$\Sigma_{i=1\ to\ n}(y_i - \beta_0 - \Sigma_{j=1\ to\ d}\beta_{ij}x_{ij})^2 + \lambda\Sigma_{j=1\ to\ n}|\beta j| \quad \ldots (1)$$

Naive Bayes uses Bayes rule with the assumption of independent and identical features:

$$P(y_i \mid x_1, x_2, \ldots, x_n) = P(x_1, x_2, \ldots, x_n \mid y_i) * P(y_i) \div P(x_i, x_2, \ldots, x_n) \quad \ldots (2)$$

FIGURES

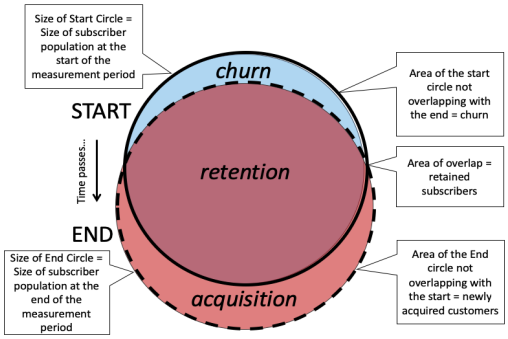

Figure 1: Churn Illustration

This diagram explains how the Churn is calculated. It is basically the area of the acquisition circle overlapping with the churn circle, that's the number of retained customers.

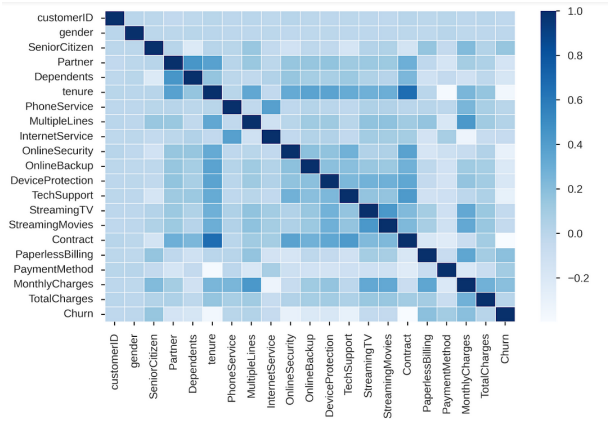

Figure 2: Correlation Heatmap

This plot shows the correlation of the features in the dataset. Correlation is how we gauge the importance of each feature and decide to include in our modes.

REFERENCES

[1] *What is customer churn prediction and why is it important?* (2019, February 1). Avaus. https://www.avaus.com/blog/predicting-customer-churn/

[2] Dataset: https://www.kaggle.com/blastchar/telco-customer-churn