COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

IMPROVING CUSTOMER RETENTION FOR TELECOM COMPANIES

IEOR 4650-Business Analytics Project
Fall 2017

Hrishikesh Jadhav-Lara Jalwan-Shikhar Mittal-Rohan Pimprikar- Chandana Rao

Introduction

In today's world, gaining a new customer is not enough anymore. In order to succeed, a company must be able to retain its customers by creating a relationship between the customers and the brand. Increasing retention requires careful consideration of all aspects of the customer experience and the company's strategy.

Taking the case of a telecommunication company, we will study the data set related to its customers in order to predict the customer behavior and increase retention. Companies make profit on customers when they try to sell to their existing customers rather than losing them to another telecom since it is cheaper to retain customers. It increases ROI, loyalty and lifetime value of customers.

The goal of this project is to find the best plan for each category in order to decrease the cost of losing customer, and increase the retention.

Data Set and Methodology

This report was made through the analysis of two public datasets. In both datasets, we are looking at whether customer churns or not. In the first dataset, customer demographics and their monthly payment and services they use are provided. Through this data set we will show how should customers be grouped based on their characteristics and what best plan to apply to each group based on their churn rate. This data set contains information for 7043 customers.

The second dataset provides information about the usage of the plan, the length of calls: night, international, customer services and so on. Through this data set we were able to visualize the factors in a way that affects the churn rate the most. This data set contains information on 3333 different customers.

From previous research, we found that the best method to predict churn rate is by the analysis made using logistic regression and classification trees. Both datasets were randomly divided into train and test in order to avoid over fitting. The models were then trained on train data and its predictions tested on test data.

Visualization & Results

This graph shows that the customers who have a short tenure are more likely to leave. The churn rate of people who have a month-to-month contract is 42.7% vs 11.3% for the customers with a one-year contract. (Fig. 1). Another major difference we notice in the services provided was the Tech Support. Most of the people who are churning are not using tech support. One of the reasons could be that this service is not free, and offering it as free to these customers could help retain them.

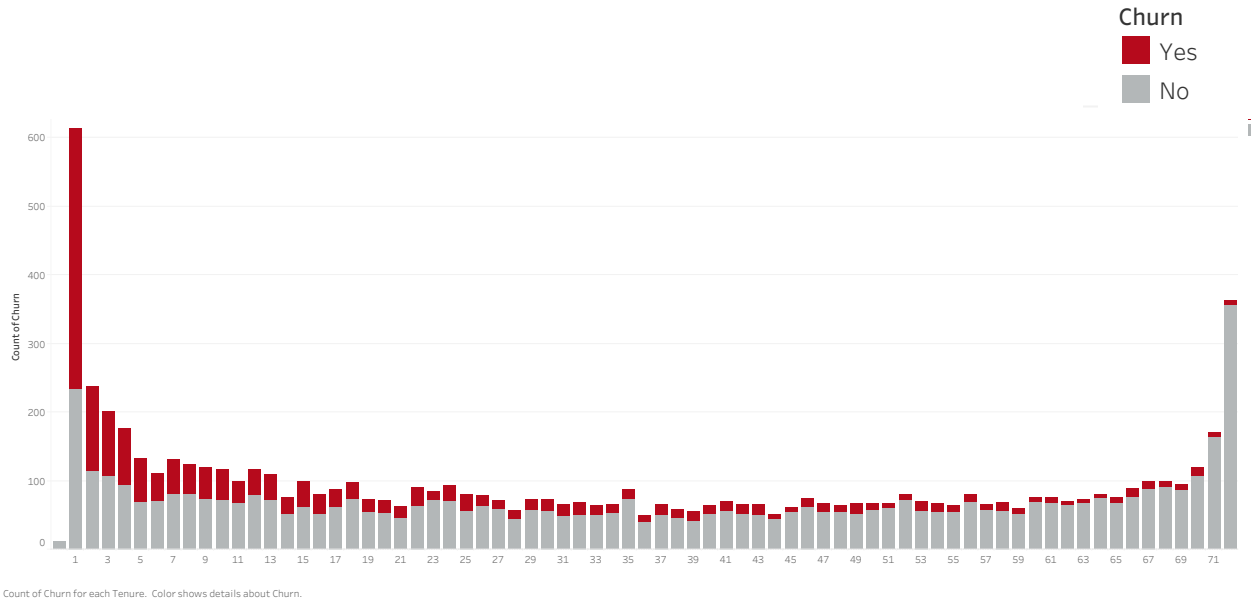

Figure 1 Churn Rate vs Tenure

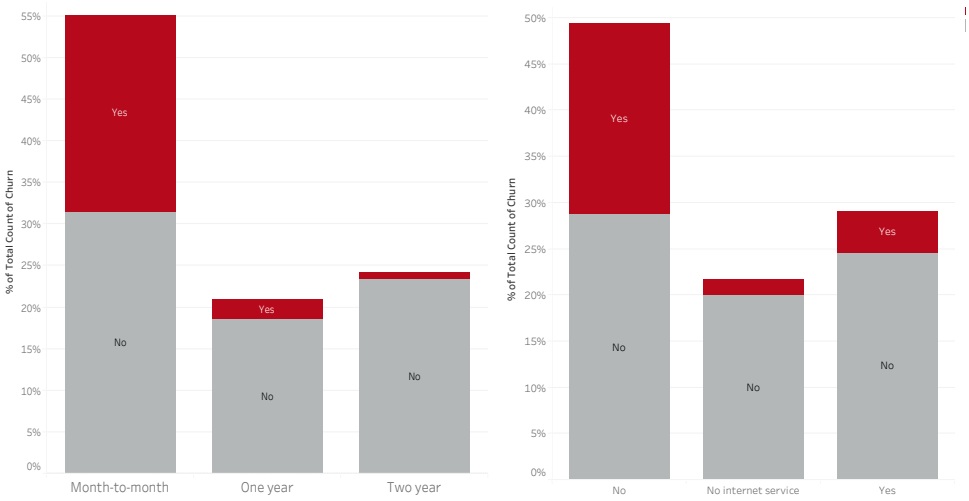

Figure 2 Churn Rate vs Contract Type (left) Churn Rate vs Tech Support (right)

Exploring the second data set, we can see that most of the people do not use an international plan. As seen in Figure 3 the churn rate for users who does not use international plan is of 11.3% and those that use them are more likely to churn (Churn rate = 42.4%)

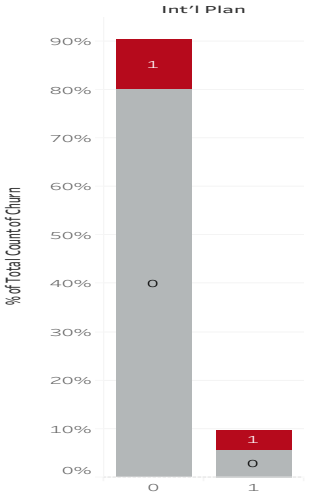

Figure 3 Churn Rate for International Plan Users

In Figure 4 we can see that the churn rates in some states is much higher than in others. We explore this by looking into network towers and customer service calls as being key issues.

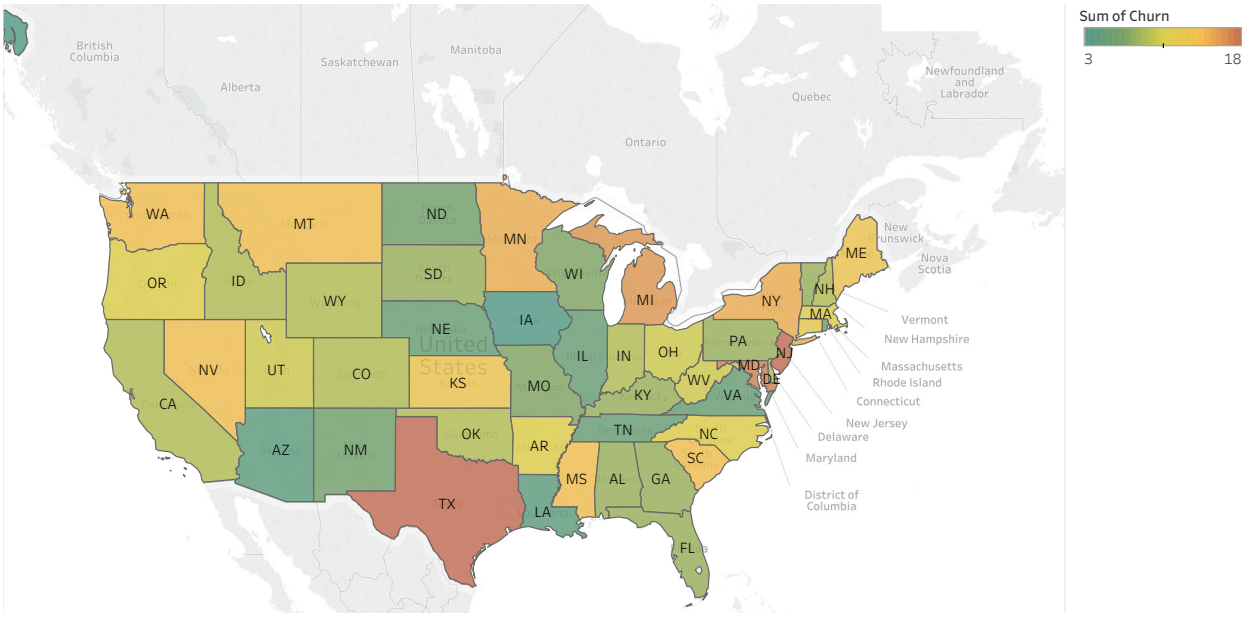

Figure 4 Churn Rate per State

Logistic regression results

First data set

The result shows that the coefficients that affect the churn rate the most are: tenure, contract type, electronic payment option and the total charges. The churn rate for the total sample is 26.53%. Looking into the customers grouped in categories, we find that people with a month-to-month contract have a much higher churn rate: 42.75% vs people with a one or two-year contract. Also, senior citizens have a churn rate of 71.5%. Hence, we looked for customized plans or promotions for these customers in order to decrease the churn rate in these categories.

Second data set

The significant predictors out of the logistic regression from the second dataset are the customer service calls, international plan, international calls and Vmail message. The MSE of our model on the test data is 0.11 which means that our model is able to predict the churn rate well.

Classification tree results

First data set

After the pruning, our tree shows that contract, tenure and internet service is the deciding factor for churn rate. The AUC for this model equals 73.54% and the MSE is 0.104 with the ROC curve predicting a significantly better prediction than logistic regression.

Second data set

With the second data set, we find that the decision tree gives us the following variables that are significant: day minutes,customer service calls, international calls, evening minutes and

international minutes. The AUC of this model is 71.54% with the ROC curve predicting a significantly better than logistic regression.

Cost Analysis

For the first data set, we use cost analysis to decide which method performs better and reduces the cost to the company. This is a better method of prediction for it is important to reduce the costs incurred by a telecommunication company so as to increase its profit margin from the existing customer base.

Deciding on assigning costs

Based on a few assumptions we assign a cost of $0 to true negatives, that is the people who don't leave and are predicted to not do so correctly by our model. The false negatives are the customers that are predicted to not leave when in reality they do. This means we lose the customer and have to acquire another customer to replace the one that left, thus, adding in more cost at the loss of revenue. The advertising, sales and administrative costs add up to give a higher figure in the range of hundreds. Here, we assume it to be $500. And for the customers that are identified to churn, we offer a $100 incentive. This is the cost associated with true positive. On the other hand, the false positive customers see $100 being wasted on them for they are actually staying but the model falsely predicts them to leave.

Finding the optimal threshold

Most of the cost to the company will be added by the customers in the false negative category due to the cost running in hundreds. Hence, we minimize a cost function that is as follows:

Equation 1

$500 * FN(C) + $0 * TN(C) + $100 * FP(C) + $100 * TP(C)

where FN(C) signifies a false negative percentage of the cutoff C and so on for the rest of the variables.

We get the predictions for both the models and calculated the cost incurred by the company based on a threshold value. This cost analysis is done comparing both models with a baseline model wherein nobody churns. The cost associated with the baseline model is $62000 while the cost reduction according to classification tree is $12400 while for logistic regression it is $15400.

This clearly adjudges classification tree to be the better method to predict churn rate. Using this model, the company should take further actions to increase customer retention.

Proposal for Customer Retention

Best plan for senior citizens

We propose a discount coupon system in place to help retain them. The major challenge then lies in determining an optimum discount rate that keeps the cost to the company at a lower value. From the decision tree, we can see that tenure is an important classifier. Senior citizens with a tenure greater than 3.5 and less than 55.5 tend to churn. Hence, we will use this threshold to find an optimum discount rate for this category. Below we can see the confusion matrix, with the

above tenure threshold:

	PREDICT	
TRUTH	FALSE	TRUE
NO	147	231
YES	19	248

This gives us a True Positive Rate: 92.8%, False Positive Rate: 61.1% and Total Error Rate: 38.76%. We try a cost analysis with this threshold, and check whether any promotion can reduce the cost of losing customers.

Finding the optimal promotion:

In this case we simulated different discount rate and compared them with the cost of no promotion at all. The analysis shows that the cost of no promotion is:

Equation 2

$0*TN(C) - (average monthly charge) * TP(C) = - $21270

We then calculated the cost of offering a promotion based on Equation 1 . In this case, the 100$ incentive cost will be replaced with the discount cost. We will simulate different discount offered, and below will be the cost used:
FN cost: 200$ + average monthly charges
TN cost: $0
FP cost: $discount *average monthly charges
TP cost: $ (average monthly charges - discount *average monthly charges) * effectiveness

Equation 3

$FN * FN(C) + $0 * TN(C) + $FP* FP(C) + $TP * TP(C)

The plot below shows the cost for the different discount rates. It is clear that with only a 10% discount we would have a much lower cost: **$9673.**

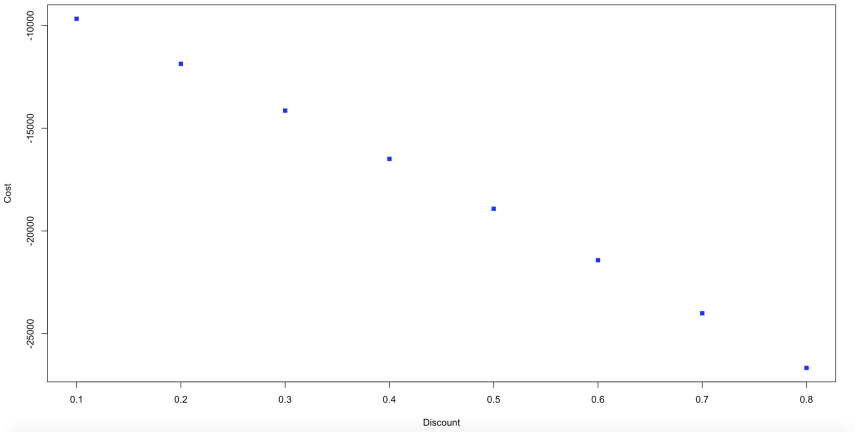

Figure 5 Cost vs Discount Rate

Best plan for month-to-month customers

Similar to a discount plan for senior citizens, we also suggest offering discount coupons to the category of customers that have a month to month contract since these are the ones most likely to increase the churn rate. Hence, we specifically focus on the customers with this contract and build a logistic model with the most significant variables that affect the churn rate, namely, Seni -or Citizen, Tenure, Multiple Lines, Paperless Billing, Payment Method and Total Charges.
By building a classification tree, we find that customers with less than 2.5 tenure are the most likely to leave. Hence, we offer discounts to this category. We followed the same logic with senior citizens.

Equation 4

$0*TN(C) - (average monthly charge) * TP(C) = - $194702.

We first find how much it costs us to give these customers no discount promotion at all which c omes up to be about $194702, based on the above formula. Below is the classification table wit h a division based on tenure.

	PREDICT	
TRUTH	**FALSE**	**TRUE**
NO	**252**	**1413**
YES	**372**	**869**

True positive rate = 70.02%. False positive rate = 84.86%.

Finding the optimal solution:

We experiment with different discount rates and find that giving a discount rate of 10% is sufficient to decrease the cost to $114435.4

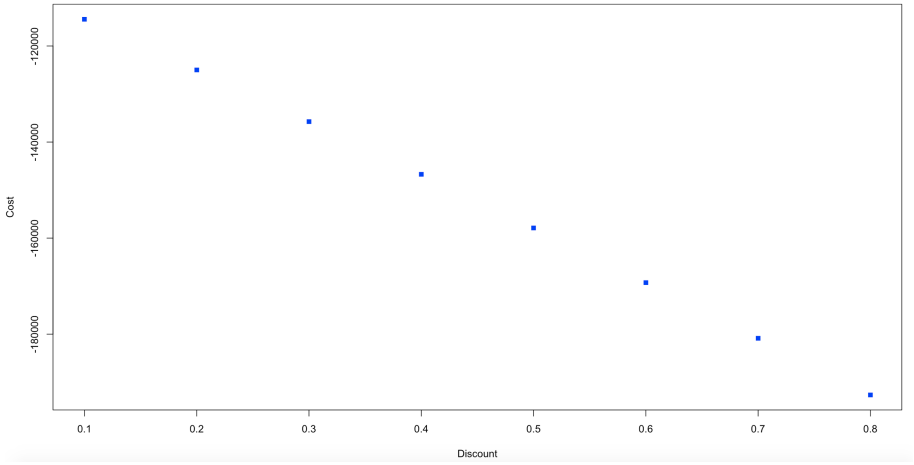

Figure 6 Cost vs Discount Rate

CONCLUSIONS

The annual churn rate for this company is 26.53%, quite higher than the normal average of 10% for other companies in the telecommunication industry. Customer acquisition has a very low rate and hence it is important to decrease the churn rate. Two different data sets are used to analyze and reveal the following insights:

- High churn rate among customers with international plans. This might point to reduced quality of the plan which is adding in to customer dissatisfaction. The company should focus on identifying issues like network problems and pricing of the plan.
- High churn rate among senior citizens and month to month customers. It is likely that higher bills are driving some of these customers to look for cheaper options. For these two issues, we have suggested a discount offer to be given to them at the least cost to the company, which amounts to about 10%. Specifically, in the case of senior citizens, we have observed that they are more prone to just using a single phone line and internet service. These being their major focus, the company should focus on price reductions and better service offers with these products.
- Customers with shorter tenure are more likely to leave the company in addition to customers with large number of customer service calls and tech support. This adds to the churn rate exponentially. The company should focus on increasing their quality of service in these two sectors and offer incentives to upgrade their month to month plan to yearly contracts.
- The most promising churn model to fit this dataset is a CRT decision tree model – decided on the basis of cost analysis. The CRT model gives the least cost to the company of $12400 with a higher accuracy of prediction than logistic regression. Using this model may reduce churn rate by up to 50%, leading telecom revenues to increase by millions of dollars.

However, we should be aware that customer behavior changes over time and therefore the model should be validated periodically to account for the changes. Additionally, these insights are based on correlating behaviors and as data points, correlation is not always causation.