Building Consumer Loyalty

- Making The Smarter Choice -

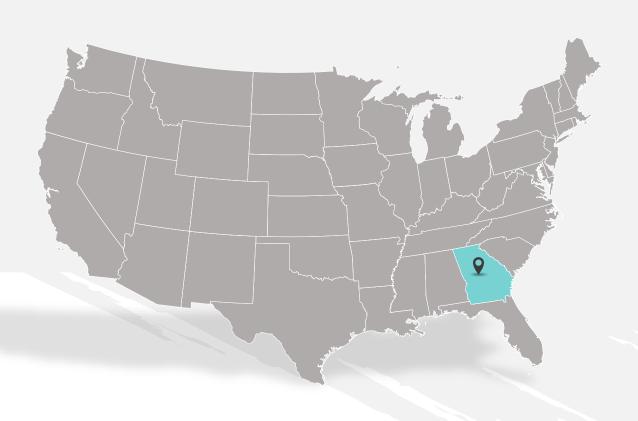BigML Software Project

Dhananjay Singh

Scenario- Introduction

Moby

Moby is a mid sized telecom provider based out of Atlanta started by UGA and GA TECH alumni

Majority of their customers are based out of GA and surrounding states

Scenario- Introduction

Moby takes pride in their 3 pillars to success – which have been instrumental to their ascend in the past 4 years

1. Quality infrastructure

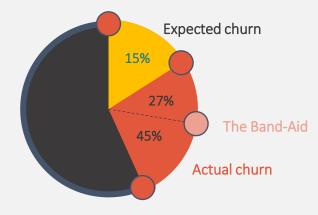2. Agility and Adaptability to technology

3. High customer acquisition rates

Scenario- Moby in troubled waters

Their third pillar is under threat

Moby started to see a heavy customer attrition by the end of their second year when they withdrew their introductory pricing which involved two year contracts at a heavily discounted prices

Moby was prepared for a 15% churn but the actual numbers were **staggering-** 45%

They re-introduced the promo last qtr. as a stop-gap arrangement but the **churn was still high at 27%**
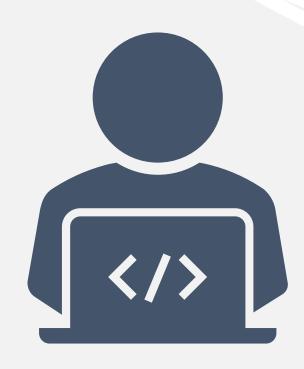
Expected churn

15%

27%

The Band-Aid

45%

Actual churn

Scenario- Moby in troubled waters

Responding to the problem

As a response, the management turned to their analytics team for some answers

They wanted to understand what is causing this churn

And create a **customer retention strategy**

The task at hand

Task allocated to the team is to answer these questions:

- What is driving the high customer attrition rate?
- How can we accurately predict the customer's churn?
- What actions should be taken to maintain customer loyalty?

About the dataset:

- Moby has provided us with a random sample of 7000 customers which represents their larger population of customers
- Total of 46 features include plan data, demographics, referrals, customer service requests etc.

Data preparation and modeling summary

- Understanding the data
- Dependent categorical feature: **Churn value (0 or 1)**
- Removing non deterministic features
- Final number of features used: 38
- 80:20 split used to create training and test dataset

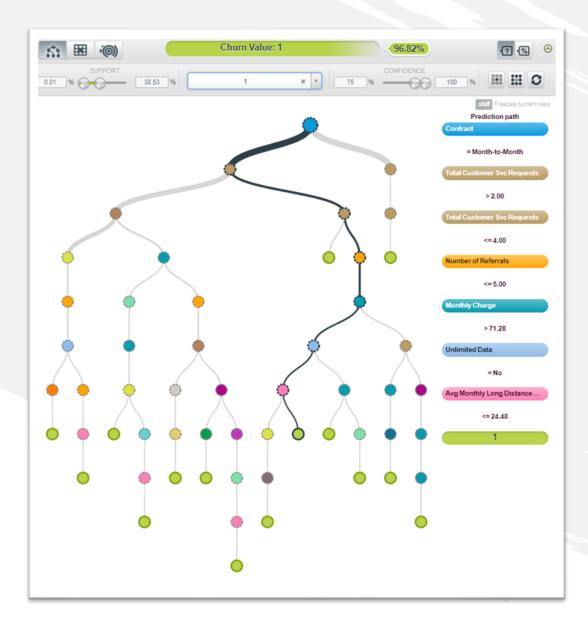- Creating and testing **ensemble** of 10 models to improve accuracy of the model

- Providing key insights and recommendations

Data wrangling

Take1: Modeling & testing

Take 2 : Remodel

Optimization

Implementation

- Using Big ML platform to create a decision treee **model** to predict customer churn value
- Ranking the **most important variables** for the exercise
- **Testing** the trained model against the test dataset

- Weighing the importance of **type 1 and type 2 errors**
- Adjusting probability threshold to **optimize cost**
- Introduce random use cases and predict the churn using the ensemble

Take 1: Modeling

According to one of the branches, we can have a confidence level of 96.82% for a churn when the highlighted features are considered
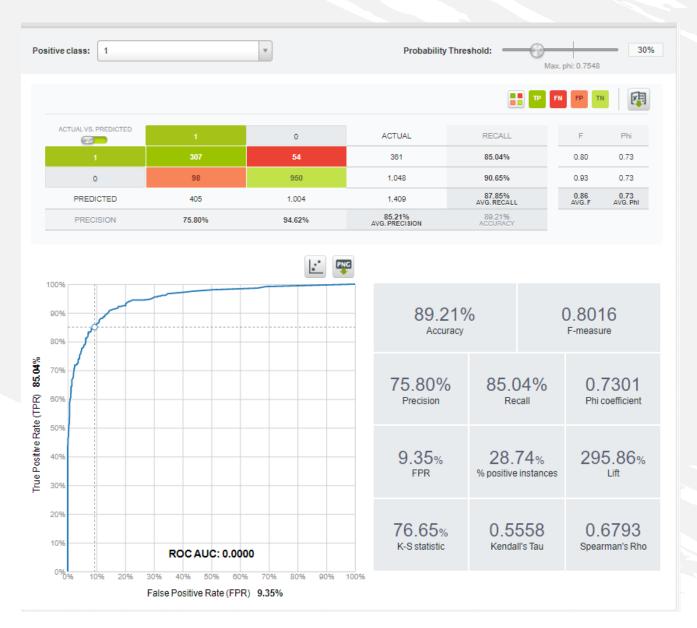
Take 1 – Model Evaluation and error type costs

ACTUAL VS. PREDICTED	1	0	ACTUAL	RECALL
1	286	75	361	79.22%
0	104	944	1,048	90.08%
PREDICTED	390	1,019	1,409	84.65% AVG. RECALL
PRECISION	73.33%	92.64%	82.99% AVG. PRECISION	87.30% ACCURACY

- After testing the model against the 20% test data set: we found the model to be fairly accurate but there was still scope for improvement in the Precision / recall (based on the cost of a type 1 or a type 2 error)
- For customer churn to be true, a **false negative will be more costly,** thus **minimizing Type 2 errors** is a priority: **recall has to be higher**
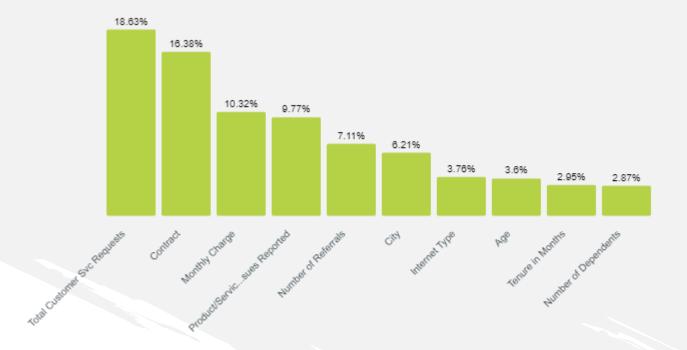
87.30% Accuracy	0.7616 F-measure

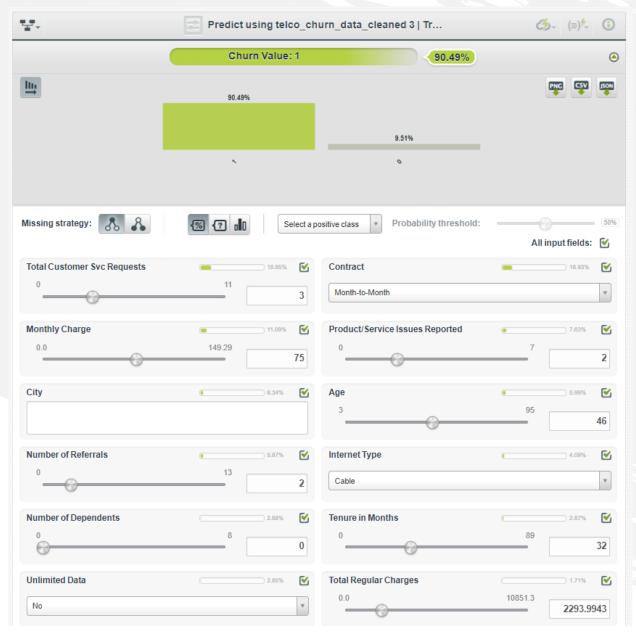| 73.33% Precision	79.22% Recall	0.6762 Phi coefficient
9.92% FPR	27.68% % positive instances	286.22% Lift
71.68% K-S statistic	0.5363 Kendall's Tau	0.6436 Spearman's Rho

Take 2: Ensemble an tweaking

- Testing the new ensemble against the test data, we can see the accuracy improved slightly
- By adjusting the probability threshold to 30% we find a significant improvement in the model recall which, in turn, reduces the more costly type 2 error
- False negative cases came down from 75 to 54

Field Importance

- Top 5 fields contributed to almost 50% of the impact on churn value prediction
- These fields include:
  - Total Customer service requests
  - Contract type
  - Monthly charge to the customer
  - Product service requests
  - Number of referrals a customer has done so far

A Test Case

By only changing the top 5 fields as shown and keeping the rest at data average, we can successfully predict a churn with a 90.49% probability

Recommendations to Management

On the basis of the prediction models built, we were able to figure out the four most important recommendations

1 Instead of focusing on just their pricing strategy, the company must look deeply into the "Complains" (>3 per qtr.) raised by customers i.e. their service requests via telephone, app or other channels & resolve them promptly

2 Based on the aggregate of the complains raised, Moby must resolve the issues pro-actively, in the order of their frequency

3 Customers who have enrolled to their monthly plan should be targeted first and incentivized to move to annual plans

4 Customer churn probability drastically increases for customers who have monthly charge > $75 and have higher than 3 complaints a quarter. Premium services such as dedicated customer service reps and other bundled perks for high- tier customers could be employed

How can we improve the model

- Adding more observations to the sample set & improving representation can further make the model more accurate
- Adding a field such as " Type/Category of complaint/ Service request" will help us further drill down to the more specific causes of churn

Thank You