Reducing Customer Churn:
A Data-Driven Approach to Improving Customer Service

1 What is Churn?

In the customer management, *customer churn* refers to a decision made by the customer about ending the business relationship. It is also referred as loss of clients or customers. Customer loyalty and customer churn always add up to 100

It is very important to predict the users likely to churn from business rela-tionship and the factors affecting the customer decisions. This analysis shows how logistic regression model, support vector machines and the random forest model can be used to identify the customer churn in the telecom dataset.

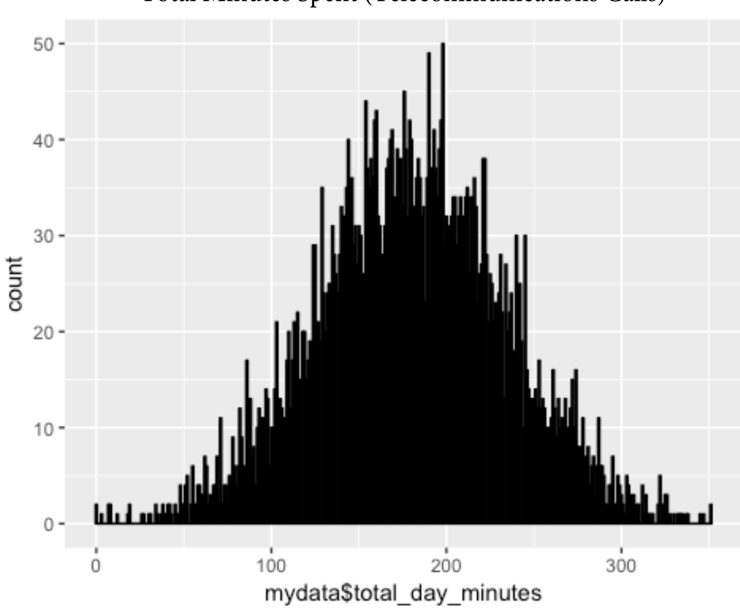
2 The Data

The "churn" data set was developed to predict telecom cus-tomer churn based on information about their account. I used the same analysis during my internship at SitterFriends and was able to reduce customer churn in the company using its data. From deriving insight using the appropriate model (as shown below), I was able to reduce customer churn by over 20% and reduce costs associated with these inefficiencies.

	account_length <int>	international_plan <dbl>	voice_mail_plan <dbl>	number_vmail_messages <int>	total_day_minutes <dbl>	total_day_calls <int>	total_day_charge <dbl>	total_eve_minutes <dbl>	total_eve_calls <int>
1	101	0	0	0	70.9	123	12.05	211.9	73
2	137	0	0	0	223.6	86	38.01	244.8	139
3	103	0	1	29	294.7	95	50.10	237.3	105
4	99	0	0	0	216.8	123	36.86	126.4	88
5	108	0	0	0	197.4	78	33.56	124.0	101

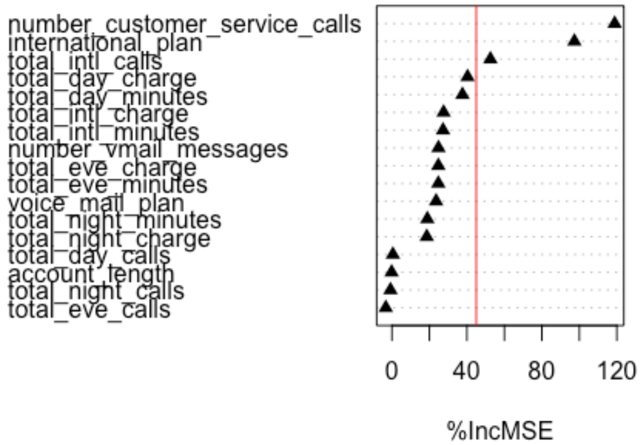
1-5 of 6 rows | 1-10 of 18 columns Previous 1 2 Next

Total Minutes Spent (Telecommunications Calls)



The figure to the left shows all data points of calls spent with a customer. Each point refers to how long it took to solve the customers needs (e.g. some customers require multiple calls, but this graph gives the total time needed to solve a customer issue).

3 Insights



To the right shows significant variables in minimizing customer churn. Using SVM, I was able to derive what factors on a customer service call are important in determining/ predicting customer churn. Here, we see that the number of customer service calls, the international call plan the company buys and the total number of international calls are highly significant in explaining customer churn. That is saying these are the most important factors in determining customer churn. Intuitively, this makes sense. A customer who has to receive many customer service calls to resolve an issue would likely become frustrated and leave their business with the company.

```
> plot.new()
> varImpPlot(rf, type = 1, pch = 17, col = 1, cex = 1.0, main =
  "")

> abline(v= 45, col= "red")
```

Business Insights: Changing the Way we Interact with Customers

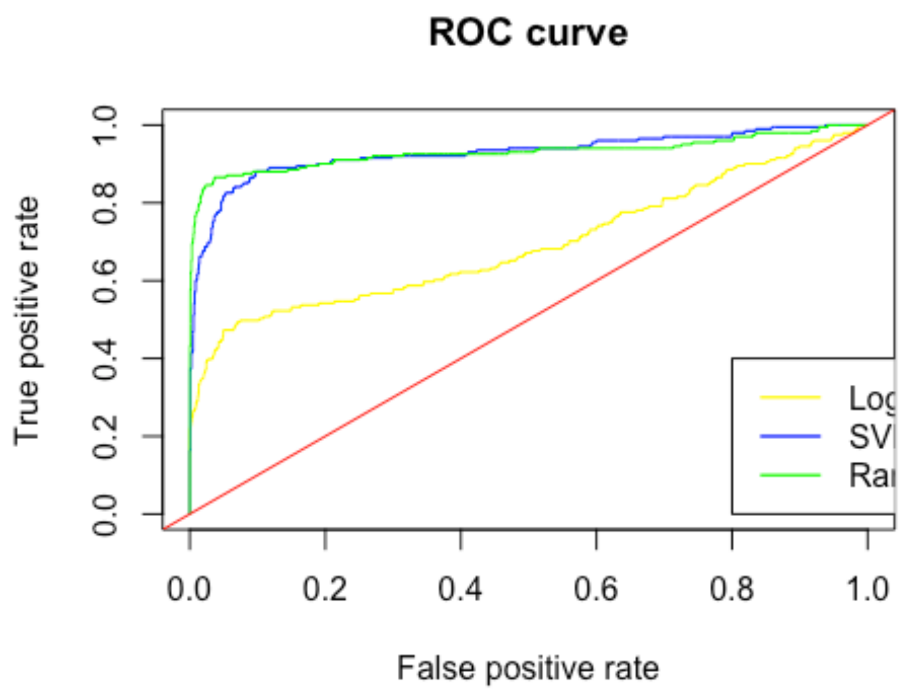
Rule 9: (82/1, lift 6.9)
    account_length > 69
    voice_mail_plan <= 0
    total_day_minutes > 277.7
    total_eve_minutes > 152.7
    number_customer_service_calls <= 3
    -> class 1 [0.976]

Rule 10: (114/3, lift 6.8)
    voice_mail_plan <= 0
    total_day_minutes > 245.1
    total_eve_minutes > 201
    total_night_charge > 8.54
    number_customer_service_calls <= 3
    -> class 1 [0.966]

Rule 11: (130/5, lift 6.8)
    international_plan <= 0
    voice_mail_plan <= 0
    total_day_minutes > 263.4
    total_eve_minutes > 184.9
    -> class 1 [0.955]

There are many rules here, but let's look at rule 9. It is saying if the total minutes in the day exceed 277.7 and total evening minutes exceed 152.7, the customer is likely to leave the company. This can be used to derive business insight.

4 Conclusion



My final model- the random forest model - is about 92.9% accurate. I used this model using my company data to reduce customer churn by over 20% and therefore reduce costs associated with these inefficiencies. This analysis can be used by any company that uses customer service practices to interact with its customer base and provide assistance with any issues. Small changes to how companies serve their customers can mean a great difference in customer retention and satisfaction.