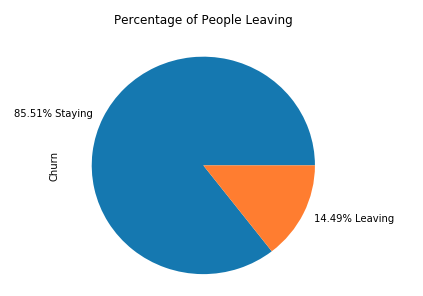
Pratik Patel

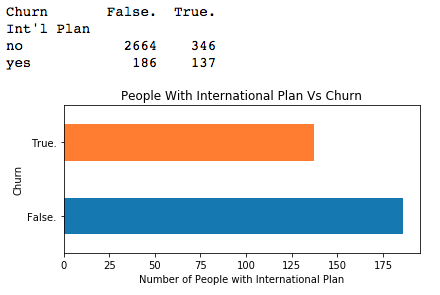
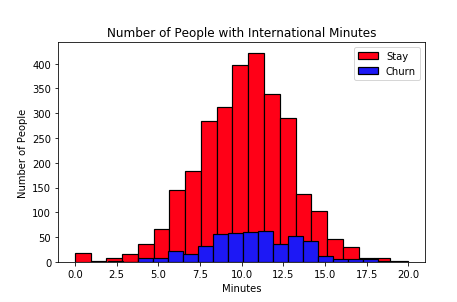
Professor Vucetic

CIS 3715

20 March 2018

Churn EDA

With Alexander Graham Bell developing the first telephone in 1854, the future of a vastly profitable business was not out of sight for Bell. Bell’s creation later allowed the creation of what we know as of today as telecom company. With nearly every person having their own communication device and there being so many different telecommunication companies for people to choose from, it is important for companies to need to know the churn rate. The churn rate is the percentage of subscribers of a telecom company who will end their subscription with the company and most likely change companies. With the data, we got from a telecom company we can see that for this company about 14.49% of the subscribers’ churn which may not seem a lot but any amount of churn rate is bad for company as that is some revenue being lessened.

Rather than just knowing the churn rate, companies will see what user’s churn to see if there is an underlying reason of why they are churning or to predict if new customers will churn or not so they can begin to take action to keep them. An understanding of the data set of what each attribute is and other aspects needs to be known before it can be used to predict if new customers will churn. Each entry in the data set is a subscriber telling information about user itself as well as what plan they have through a simple yes and no as well as the number of calls, number of mins, and the charge for the day, evening, night, and international calls. Using the data a few insights we can draw from how the other attributes related to churn. One of these insights is that from the cross table and bar graph we can see that about 28% of the people who churn have an international plan and of all people who do have an international plan 42% churn showing how international plan could relate to churn. This is further supported by the following histogram graph showing how the churn users have a more skewed distribution to the right showing how they are more likely to make longer calls compared to the almost perfect normal distribution of the people who do not. Overall, just looking at a few attributes is not enough to predict if a new customer will churn and therefore we must use supervised learning to look at all attributes to accurately make predications.