**Proactive Attrition Management**

**Logistic Regression Business Case**

**1. Data cleaning including missing values, outliers and multi-collinierity. Describe your predictive churn model. How did you select variables to be included in the model?**

**Ans1 :**

* I have done complete exploratory data analysis and calculated percentiles , upper and lower cap , central tendencies and missing values for all continuous data. For categorical data I have calculated missing values, frequencies and it’s proportions.
* I have used mean to fill missing values in case of numerical and mode for categorical data. Outliers are treated for continuous variables using 1 and 99 percentile.
* Factor analysis and selection of variable has been done using various techniques. For categorical data I have used chi square test to compare with CHURNDEP while for numerical data STEP AIC is used on the model build with all numerical values.

**2. Demonstrate the predictive performance of the model.**

**Ans2 :** The concordance of the train data was 0.626361. The confusion matrix calculated for the model gives good accuracy of 60%

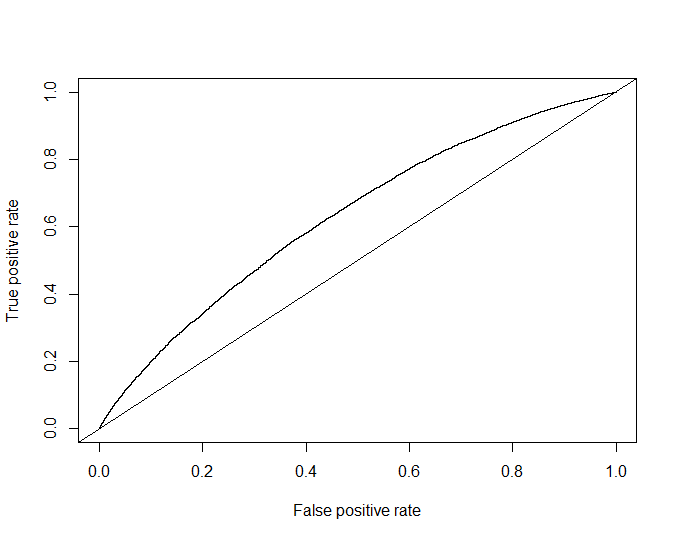
Confusion matrix:

0 1

FALSE 6849 4288

TRUE 7191 9672

**Auc Curve**



**3. What are the key factors that predict customer churn? Do these factors make sense?**

**Ans3:** Myfinal model is as below :

final\_model<- glm(CHURNDEP ~ REVENUE + MOU + RECCHRGE + OVERAGE + ROAM + CHANGEM + CHANGER + BLCKVCE + THREEWAY + OUTCALLS + INCALLS + PEAKVCE +

DROPBLK + MONTHS + PHONES + EQPDAYS + SETPRC + UNIQSUBS +

ACTVSUBS + AGE1 + CREDITAA + CREDITB + CREDITDE + REFURB +

WEBCAP + MARRYNO + MAILRES + RETCALLS + NEWCELLY + SETPRCM,

data = training,

family = binomial(logit))

Most of the variables form 3 sets , 1. Calling pattern of customer 2. Company and customer relation 3. Customer device and his/her properties. All three sets influence the relationship between the company and customer.

**4. What offers should be made to which customers to encourage them to remain with Cell2Cell? Assume that your objective is to generate net positive cash flow, i.e., generate additional customer revenues after subtracting out the cost of the incentive.**

**Ans4 :**

As from above answer we know the driving factors so following could be done by the company:

1. Understand the calling pattern of customer like recharge,revenue,calls,roaming etc. create segments of similar customers and provide a good plan to target segments. Eg: Giving discounts and offers on usage or data.
2. Improving on customer and company interaction. Addressing to the customer’s issues efficiently and effectively. Customer could have a single platform where his/her issues could be resolved or information could be provided.
3. With different types of technologies present in the market different customers will be using different types of devices will require different plans and services as per their use. Company could identify different types of platform and devices that will run it services and engineer the best quality output for these identified platforms.

**5 Assuming these actions were implemented, how would you determine whether they had worked?**

**Ans5 :** Data collection and observations could be done after the new techniques are implemented. By active marketing campaigns and advertisements company could promote and sell it’s new plan and see change in usage and patterns in service use . Company through feedback systems could have data on new customer service changes . Improvement in engineering could be tested technically and random samples should be collected after new deployments. This new data can be utilised and studied to understand changes.