## Predicting Customer Churn: Decision Tree Model in SAS

/\*Create your own library in SAS like here it is libref and mention the path \*/

libname libref “/home/aro1260/deep”;

/\*Importing churn\_dataset in your assigned library \*/

PROC IMPORT DATAFILE= “/home/aroragaurav1260/data/churn\_dataset.csv”

DBMS=CSV Replace

OUT=libref.churn;

GETNAMES=YES;

RUN;

/\*To check the contents of the data \*/

PROC CONTENTS DATA=libref.churn;

RUN;

/\*Descriptive statistics of the data\*/

proc means data = libref.churn;

var Term Monthly\_Charges;

run;

/ Applying Proc freq to see the frequency of the data \*/

Proc freq data = libref.churn;

tables Internet\_service Churn Churn \* Internet\_service;

run;

/\* Applying proc univariate to get more detailed summary of the data \*/

proc univariate data = libref.churn;

var Monthly\_Charges;

histogram Monthly\_Charges/normal;

run;

/\* Growing Decision tree on Full data\*/

**Program2:**

ods graphics on;

proc hpsplit data=libref.churn cvmodelfit seed=123 maxdepth=3;

class Churn Sex Marital\_Status Phone\_service International\_plan Voice\_mail\_plan Multiple\_line Internet\_service Technical\_support Streaming\_Videos Agreement\_period;

model Churn (event='Yes') = Sex Marital\_Status

Term Phone\_service International\_plan Voice\_mail\_plan Multiple\_line Internet\_service Technical\_support Streaming\_Videos Agreement\_period Monthly\_Charges Total\_Charges;

grow entropy;

prune costcomplexity(leaves=6);

run;

/\* Splitting dataset into training and testing dataset in 70:30 and growing deciison tree \*/

**Program2.1:**

ods graphics on;

proc hpsplit data=libref.churn

plots=zoomedtree(nodes=('0') depth=3);

class Churn Sex Marital\_Status Phone\_service International\_plan Voice\_mail\_plan Multiple\_line Internet\_service Technical\_support Streaming\_Videos Agreement\_period;

model Churn (event='Yes') = Sex Marital\_Status

Term Phone\_service International\_plan Voice\_mail\_plan Multiple\_line Internet\_service Technical\_support Streaming\_Videos Agreement\_period Monthly\_Charges Total\_Charges;

grow entropy;

prune costcomplexity(leaves=6);

partition fraction(validate=0.3 seed=123);

code file= "/home/aroragaurav1260/data/scorefile.sas";

rules file="/home/aroragaurav1260/data/noderules.txt";

run;

/\*scoring the data to predict the probability of customer churn\*/

data libref.finalscore;

set libref.churn;

%include "/home/aro1260/data/scorefile.sas";

run;