

```
1 /* CIND 119 Final Project */
2
3 /* Import Data */
4
5 PROC import DATAFILE= '/home/u63311549/TMU-CIND119/TMU-CIND119-Final/churn.csv'
6     DBMS= CSV replace
7     OUT= churn;
8     GETNAMES= YES;
9
10 /* Examine data */
11 PROC print CONTENTS= 'Examine Data' data= churn(obs = 10);
12 title 'Show the first 10 rows for brief look';
13 RUN;
14 title; /*delete the title*/
15
16 /* Renaming columns -if needed */
17
18 /* rename time1=temp1 time2=temp2 time3=temp3; */
19
20 ods graphics on;
21
22 /* Decision Tree - All 20 vars + GINI + without pruning */
23
24 PROC hpsplit data= churn;
25 /* A SAS name literal is a name token that is expressed as a string within quotation marks, followed by the upper- or lowercase letter n. */
26 class 'churn?'n "Int'l Plan"n 'VMail Plan'n
27 /* The following are to be removed WITH Variable selection*/
28 state 'Area Code'n Phone;
29 model 'churn?'n = 'Account Length'n "Int'l Plan"n 'VMail Plan'n 'VMail Message'n
30 'Day Mins'n 'Day Calls'n 'Eve Mins'n 'Eve Calls'n 'Night Mins'n 'Night Calls'n
31 'Intl Mins'n 'Intl Calls'n 'CustServ Calls'n
32 /* The following are to be removed WITH Variable selection */
33 state 'Area Code'n Phone
34 'Day Charge'n 'Eve Charge'n 'Night Charge'n 'Intl Charge'n;
35 grow GINI;
36 partition fraction(validate=0.3 seed=123);
37 prune OFF;
38
39 RUN;
40
41 /* Decision Tree - 13 vars + GINI + without pruning */
42
43 PROC hpsplit data= churn;
44 /* A SAS name literal is a name token that is expressed as a string within quotation marks, followed by the upper- or lowercase letter n. */
45 class 'churn?'n "Int'l Plan"n 'VMail Plan'n ;
46 model 'churn?'n = 'Account Length'n "Int'l Plan"n 'VMail Plan'n 'VMail Message'n
47 'Day Mins'n 'Day Calls'n 'Eve Mins'n 'Eve Calls'n 'Night Mins'n 'Night Calls'n
48 'Intl Mins'n 'Intl Calls'n 'CustServ Calls'n ;
49 grow GINI;
50 partition fraction(validate=0.3 seed=123);
51 prune OFF;
52 /* where is 'Vmail Message' ? */
53 RUN;
54
55 /* Decision Tree - 13 vars + GINI + with costcomplexity pruning */
56
57 PROC hpsplit data= churn;
58 /* A SAS name literal is a name token that is expressed as a string within quotation marks, followed by the upper- or lowercase letter n. */
59 class 'churn?'n "Int'l Plan"n 'VMail Plan'n ;
60 model 'churn?'n = 'Account Length'n "Int'l Plan"n 'VMail Plan'n 'VMail Message'n
61 'Day Mins'n 'Day Calls'n 'Eve Mins'n 'Eve Calls'n 'Night Mins'n 'Night Calls'n
62 'Intl Mins'n 'Intl Calls'n 'CustServ Calls'n ;
63 grow GINI;
64 partition fraction(validate=0.3 seed=123);
65 prune costcomplexity;
66 /* where is 'Vmail Message' ? */
67 RUN;
68
69
70
71
72 /* Naive Bayes - All Vars except Phone Number; without prescreening and variable selection */
73 /* There can be more than one input statement, one for NOMinal vars and the other for INTerval vars */
74 PROC hpbnet data= churn structure=Naive maxparents=1 prescreening= 0 varselect= 0;
75 target 'churn?'n;
76 input "Int'l Plan"n 'VMail Plan'n
77 /* The following are to be removed WITH Variable selection */
78 state 'Area Code'n ;
79 /* There can be more than one input statement, one for NOMinal vars and the other for INTerval vars */
80 input 'Account Length'n 'VMail Message'n
81 'Day Mins'n 'Day Calls'n 'Eve Mins'n 'Eve Calls'n 'Night Mins'n 'Night Calls'n
82 'Intl Mins'n 'Intl Calls'n 'CustServ Calls'n
83 /* The following are to be removed WITH Variable selection */
84 'Day Charge'n 'Eve Charge'n 'Night Charge'n 'Intl Charge'n/level=INT;
85 output network=network fit=fit PRED=prediction PARAMETER=parameter
86 VARSELECT=varselect VARINFO=varinfo VARLEVEL=varlevel;
87 partition fraction(validate=0.3 seed=123);
88
89 RUN;
90
91 proc print data=network noobs label;
92 RUN;
93
94 proc print data=fit noobs label;
```


