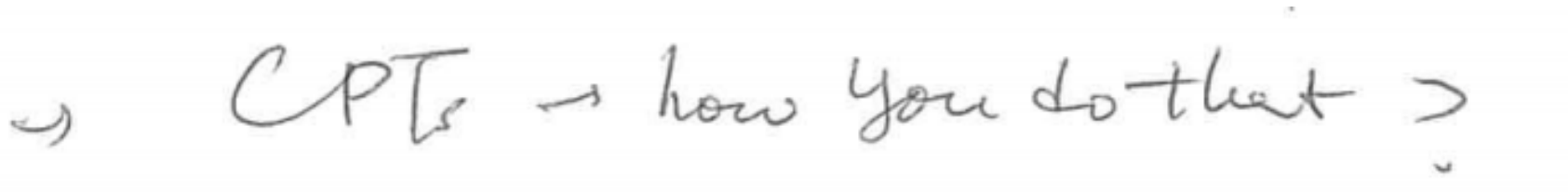


PHD - Progress update 10/11/2015

Fauzy Bin Che Yayah

Nov 10, 2015



The CPT (Conditional Probability Table) can be generated via Recursive Partitioning and Regression Trees

- Download sample Trouble Ticket using the Impala
- Sample original Trouble Ticket dataset below :-

```
suppressWarnings(suppressMessages(library(ggplot2)))
suppressWarnings(suppressMessages(library(xlsx)))
suppressWarnings(suppressMessages(library(sentiment)))
suppressWarnings(suppressMessages(library(wordcloud)))
suppressWarnings(suppressMessages(library(RODBC)))
suppressWarnings(suppressMessages(library(pander)))
suppressWarnings(suppressMessages(library(ggmap)))
suppressWarnings(suppressMessages(library(RColorBrewer)))
suppressWarnings(suppressMessages(library(Rstem)))
suppressWarnings(suppressMessages(library(tm)))
suppressWarnings(suppressMessages(library(NLP)))
suppressWarnings(suppressMessages(library(rmarkdown)))
suppressWarnings(suppressMessages(library(gridExtra)))
suppressWarnings(suppressMessages(library(Rmisc)))
suppressWarnings(suppressMessages(library(png)))
suppressWarnings(suppressMessages(library(grid)))
suppressWarnings(suppressMessages(library(igraph)))
suppressWarnings(suppressMessages(library(RPMG)))
suppressWarnings(suppressMessages(library(Hmisc)))
suppressWarnings(suppressMessages(library(rpart)))
suppressWarnings(suppressMessages(library(rpart)))
suppressWarnings(suppressMessages(library(rpart.plot)))
suppressWarnings(suppressMessages(library(RColorBrewer)))
suppressWarnings(suppressMessages(library(party)))
suppressWarnings(suppressMessages(library(partykit)))
suppressWarnings(suppressMessages(library(caret)))
suppressWarnings(suppressMessages(library(xtable)))
suppressWarnings(suppressMessages(library(knitr)))
suppressWarnings(suppressMessages(library(ascii)))
suppressWarnings(suppressMessages(library(pander)))
```

```
# Make big tree
b <- read.csv("cat.csv")
kable(head(b), digits = 4, longtable = TRUE , format = "pandoc" , padding = 0)
```

tt_num	ccat_before	ccat_after	symp_before	symp_after	cc_before	cc_after	rc_before	rc_after
1-10406822634	null	TM_CPE	Line Disconnect	Line Disconnect	null	TM_CPE_PG Hang	null	Change Wireless Chan
1-10407333624	null	Customer	Slow streaming	Slow streaming	null	Other_Fault	null	Advise Customer
1-10407889932	null	Customer	No Dial Tone	No Dial Tone	null	Customer_Behaviour	null	Advise Customer
1-10407906915	null	Customer	Line Disconnect	Line Disconnect	null	Customer_Not Reachable	null	Advise Customer
1-10407922815	null	Third Party Damage	Line Disconnect	Line Disconnect	null	Drop/Int. Fiber Damage by pest	null	3rd Party_Drop/Int.F
1-10407923040	null	Customer	No Dial Tone	No Dial Tone	null	Clear While Localising	null	Normalise_Power Adp

Converting the dataset into factor which is suitable for Regression Trees. Choose only 3 independent variables. (Assume this variables is listed after the features selection process)

```
for(i in names(b)){

  num <- as.numeric(as.factor(b[,i]))
  b <- cbind(b,num)
  names(b)[names(b)=="num"] <- paste(names(b[i]),"_factor",sep = "")
  #print(paste(names(b[i]),"1",sep = ""))
}
b <- b[9:18]
b$tt_num_factor <- NULL
b$cc_before <- NULL
b$cc_before_factor <- NULL
b$ccat_before <- NULL
b$rc_before_factor <- NULL
b$rc_before_factor <- NULL
b$ccat_before_factor <- NULL
b$rc_after_factor <- NULL
```

```
b$symp_before_factor <- NULL
kable(head(b), digits = 4, longtable = TRUE , format = "pandoc" , padding = 0)
```

rc_after	ccat_after_factor	symp_after_factor	cc_after_factor
Change Wireless Channel Mode	32	68	200
Advise Customer	11	110	175
Advise Customer	11	91	49
Advise Customer	11	68	54
3rd Party_Drop/Int.FiberRestrd	31	68	76
Normalise_Power AdptrConnct.	11	91	30

Using the RPart and applied to the sample dataset of the Trouble Ticket for 21k+ sample records with 3 independent variables below :-

```
names(b)
```

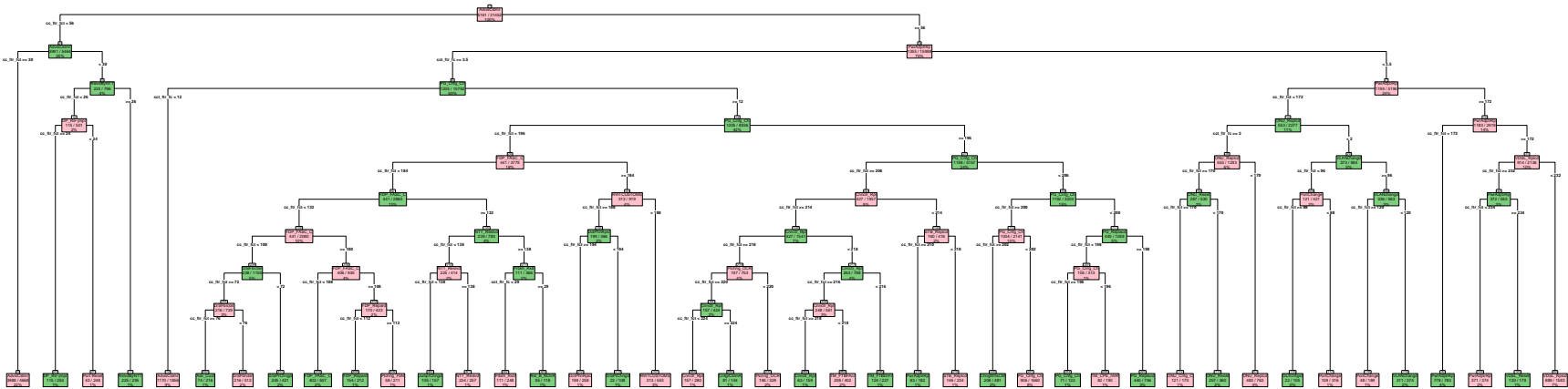
```
## [1] "rc_after"          "ccat_after_factor" "symp_after_factor"
## [4] "cc_after_factor"
```

Conditional Probability Table generated below :-

	var	n	wt	dev	yval	complexity	ncompete	nsurrogate	yval2	percentage
1	cc_after_factor	21452	21452	16291	25	0.0492	2	1	25.0000	24.0583628566101
2	cc_after_factor	5464	5464	1473	25	0.0136	2	1	25.0000	73.0417276720351
4		4668	4668	680	25	0.0000	0	0	25.0000	85.4327335047129
5	cc_after_factor	796	796	571	194	0.0071	2	2	194.0000	28.2663316582915
10	cc_after_factor	501	501	386	83	0.0026	1	2	83.0000	22.9540918163673
20		253	253	138	83	0.0000	0	0	83.0000	45.4545454545455
21		248	248	205	180	0.0000	0	0	180.0000	17.3387096774194
11		295	295	70	194	0.0000	0	0	194.0000	76.271186440678
3	ccat_after_factor	15988	15988	14633	181	0.0492	2	1	181.0000	8.47510632974731
6	ccat_after_factor	10792	10792	9587	173	0.0492	2	1	173.0000	11.1656782802076
12		1856	1856	686	25	0.0000	0	0	25.0000	63.0387931034483
13	cc_after_factor	8936	8936	7731	173	0.0266	2	2	173.0000	13.4847806624888
26	cc_after_factor	3779	3779	3338	93	0.0192	2	2	93.0000	11.669753903149
52	cc_after_factor	2860	2860	2419	93	0.0147	2	2	93.0000	15.4195804195804
104	cc_after_factor	2080	2080	1639	93	0.0125	2	0	93.0000	21.2019230769231
208	cc_after_factor	1150	1150	912	72	0.0112	2	1	72.0000	20.695652173913
416	cc_after_factor	729	729	513	72	0.0045	1	2	72.0000	29.6296296296296
832		216	216	142	24	0.0000	0	0	24.0000	34.2592592592593
833		513	513	297	72	0.0000	0	0	72.0000	42.1052631578947
417		421	421	216	74	0.0000	0	0	74.0000	48.6935866983373
209	cc_after_factor	930	930	524	93	0.0102	0	1	93.0000	43.6559139784946
418		507	507	105	93	0.0000	0	0	93.0000	79.2899408284024
419	cc_after_factor	423	423	253	94	0.0026	0	2	94.0000	40.1891252955083
838		212	212	58	94	0.0000	0	0	94.0000	72.6415094339623
839		211	211	153	166	0.0000	0	0	166.0000	27.4881516587678
105	cc_after_factor	780	780	541	137	0.0062	2	2	137.0000	30.6410256410256
210	cc_after_factor	414	414	189	137	0.0062	1	2	137.0000	54.3478260869565
420		157	157	52	116	0.0000	0	0	116.0000	66.8789808917197
421		257	257	33	137	0.0000	0	0	137.0000	87.15953307393
211	ccat_after_factor	366	366	255	183	0.0058	2	1	183.0000	30.327868852459
422		248	248	137	183	0.0000	0	0	183.0000	44.758064516129
423		118	118	23	199	0.0000	0	0	199.0000	80.5084745762712
53	cc_after_factor	919	919	606	207	0.0122	2	2	207.0000	34.0587595212187
106	cc_after_factor	366	366	167	76	0.0014	1	2	76.0000	54.3715846994536
212		258	258	59	76	0.0000	0	0	76.0000	77.1317829457364
213		108	108	86	74	0.0000	0	0	74.0000	20.3703703703704
107		553	553	240	207	0.0000	0	0	207.0000	56.6003616636528
27	cc_after_factor	5157	5157	3959	173	0.0258	2	2	173.0000	23.2305604033353
54	cc_after_factor	1957	1957	1530	58	0.0098	2	1	58.0000	21.8191108840061
108	cc_after_factor	1541	1541	1114	58	0.0061	2	2	58.0000	27.709279688514
216	cc_after_factor	753	753	556	167	0.0061	1	1	167.0000	26.1620185922975
432	cc_after_factor	424	424	267	58	0.0050	1	2	58.0000	37.0283018867925
864		280	280	123	58	0.0000	0	0	58.0000	56.0714285714286
865		144	144	63	55	0.0000	0	0	55.0000	56.25
433		329	329	133	167	0.0000	0	0	167.0000	59.5744680851064
217	cc_after_factor	788	788	525	58	0.0061	0	0	58.0000	33.3756345177665
434	cc_after_factor	561	561	313	58	0.0014	0	0	58.0000	44.2067736185383
868		159	159	96	58	0.0000	0	0	58.0000	39.622641509434
869		402	402	194	232	0.0000	0	0	232.0000	51.7412935323383
435		227	227	103	231	0.0000	0	0	231.0000	54.625550660793
109	cc_after_factor	416	416	256	216	0.0042	1	1	216.0000	38.4615384615385
218		182	182	99	181	0.0000	0	0	181.0000	45.6043956043956
219		234	234	88	216	0.0000	0	0	216.0000	62.3931623931624
55	cc_after_factor	3200	3200	2008	173	0.0155	1	1	173.0000	37.25
110	cc_after_factor	2141	2141	1137	173	0.0068	1	0	173.0000	46.8939747781411
220		481	481	273	46	0.0000	0	0	46.0000	43.2432432432432
221		1660	1660	754	173	0.0000	0	0	173.0000	54.578313253012
111	cc_after_factor	1059	1059	619	174	0.0064	1	1	174.0000	41.5486307837583
222	cc_after_factor	313	313	208	173	0.0029	0	1	173.0000	33.5463258785943
444		123	123	52	173	0.0000	0	0	173.0000	57.7235772357724
445		190	190	108	222	0.0000	0	0	222.0000	43.1578947368421
223		746	746	306	174	0.0000	0	0	174.0000	58.9812332439678

	var	n	wt	dev	yval	complexity	ncompete	nsurrogate	yval2	percentage
7	cc_after_factor	5196	5196	3997	181	0.0330	2	2	181.0000	23.0754426481909
14	ccat_after_factor	2277	2277	1724	156	0.0229	2	2	156.0000	24.286341677646
28	cc_after_factor	1293	1293	740	156	0.0133	1	1	156.0000	42.76875483372
56	cc_after_factor	530	530	243	157	0.0056	1	0	157.0000	54.1509433962264
112		170	170	49	152	0.0000	0	0	152.0000	71.1764705882353
113		360	360	103	157	0.0000	0	0	157.0000	71.3888888888889
57		763	763	280	156	0.0000	0	0	156.0000	63.302752293578
29	cc_after_factor	984	984	611	249	0.0052	1	0	249.0000	37.9065040650407
58	cc_after_factor	421	421	300	179	0.0006	0	0	179.0000	28.7410926365796
116		105	105	83	147	0.0000	0	0	147.0000	20.952380952381
117		316	316	207	179	0.0000	0	0	179.0000	34.4936708860759
59	cc_after_factor	563	563	227	249	0.0014	1	0	249.0000	59.6802841918295
118		189	189	141	179	0.0000	0	0	179.0000	25.3968253968254
119		374	374	63	249	0.0000	0	0	249.0000	83.1550802139037
15	cc_after_factor	2919	2919	1736	181	0.0312	1	1	181.0000	40.5275779376499
30		783	783	5	181	0.0000	0	0	181.0000	99.3614303959131
31	cc_after_factor	2136	2136	1222	247	0.0213	1	0	247.0000	42.7902621722846
62	cc_after_factor	553	553	181	181	0.0079	0	1	181.0000	67.2694394213381
124		374	374	3	181	0.0000	0	0	181.0000	99.1978609625668
125		179	179	49	248	0.0000	0	0	248.0000	72.6256983240224
63		1583	1583	694	247	0.0000	0	0	247.0000	56.1591914087176

SSS



```

gx <- as.data.frame(table(predict(tree.1, data_test, type = "class")))
kable(gx, digits=2 , format = "html" , padding = 0)

```

Var1

Freq

*PMW/RFC completed

0

*TNB Power Restore

0

3rd Party_D/S cable replaced

0

3rd Party_D/S cable restored

0

3rd Party_D/S Pair Changed

0

3rd Party_D/S Pair restored

0

3rd Party_Drop/Int.FiberReplce

0

3rd Party_Drop/Int.FiberRestrd

0

3rd Party_FDC_Replaced

0

3rd Party_FDP_PatchNormalised

0

3rd Party_FOC_D/S CableReplce

0

3rd Party_FOC_E/S CableRestord

0

3rd Party_FOC_E/S Switch Stdby

0
3rd Party_FOC_Joint Respliced
0
3rd Party_FOC_PatchNormalised
0
3rd Party_FOC_Tie CableReplce
0
3rd Party_FOC_Tie CableRestord
0
3rd Party_Jumper replaced
0
3rd Party_Jumper restored
0
3rd Party_Tie Copper Problem
0
ACPDB reset
0
ADSL MODEM CHANGED
0
ADSL MODEM RECONFIGURED
0
Adv_Cust
64
Advise Customer
1944
Advise TMUC
0
AdviseTMUCtoProvide full info
0
Alarm_Disabled
0
ATA Replaced
0
ATA&Dect Phone_Replaced
0
ATA&PG_Replaced
0
ATA_Replaced
0
Auto-Reset bypass
0
Auto-Reset replaced
0
BRAS_*Resolved By ANOC
0
BRAS_Reset Card
0
Building power restored
0
Cable Repair Completed
0
Cable_Reconnected
0
CCP_CPE Replaced
0
CCS BTU Configured
0
CCS CPE Configured
0

Change cabling/connector
0
Change port
0
Change Rectifier
0
Change Wireless Channel Mode
135
Change/Reset Card
0
Change__Cable
0
Change__ONU
0
Change__STB
0
CKC Restored
0
Clear While Localising
0
Clear while localizing
0
ConfigurationError
0
Configure Customer Account
39
Configure__BRAS
0
Configure__STB
0
Connector__Replaced
130
Control card replaced
0
Control card reset
0
CUST CONFIRM CCT UP&SATISFIED
0
CUST TO CHECK STATUS
0
Customer Abroad
0
Customer CPE Own Configuration
0
Customer__Reconnect CPE
0
Customer__ResetModem
0
Customer__ResetONU
0
Customer__Reuse Supplied TM CPE
0
CWL CANNOT GET CUSTOMER
0
D/S Cable Replaced
0
D/S Cable Restored
0
D/S Fiber cable replaced

150
D/S Fiber cable restored
0
D/S Pair Changed
156
D/S Pair Restored
0
Dect Phone Replaced
79
Dect Phone_Replaced
0
Dispatch activity to NGN NOC
0
Dispatch activity to NOC1 HBC
0
Dispatch activity to RNO
0
Dispatch activity to TVNOC
0
Dispatch_Activity_to_NOC1_HBC
0
DP_Re-jumper
71
E/S Fiber cable replaced
0
E/S Fiber cable restored
0
E/S pair replaced
0
E/S stdby fiber cable restored
0
FDC_Change Port
0
FDP_Change
0
FDP_Change Connector
0
FDP_Change Port
0
FDP_FASC_Connector_Cleaned
0
FDP_FASC_Connector_Replaced
141
FDP_Repaired
67
FDP_Replaced
0
FDP_Splitter Replace
0
Firmware & Software upgrading
0
Firmware_Update
0
FOC_D/S cable replaced
0
FOC_D/S cable restored
0
FOC_D/S core changed
0

FOC_D/S core restored
0
FOC_E/S cable replaced
0
FOC_E/S cable restored
0
FOC_E/S core changed
0
FOC_E/S core restored
0
FOC_OFM Pull n Resplice Cable
0
FOC_Tie cable replaced
0
FOC_Tie cable restored
0
Follow-up on new TR
0
HDSL MODEM RESET
0
HomePlug_Replaced
0
Incorrect Info
0
Install Lightning Arrestor
0
Int.__Wiring Replaced/Repaired
0
Jumper Changed
44
Jumper_changed
0
Lock port done
0
ME_OFM_Splicing fiber
0
Modify_Account
0
Modify_EPG
0
MSAN fan replaced
0
MSAN IGMP setting reconfigured
0
MSAN mcast setting reconfig
0
MSAN PON card replaced
0
MSAN PON card reset
0
MSAN POTS card replaced
0
MSAN POTS card reset
0
MSAN reconfigured
0
Network_Link_Restored
0
NOD_IMS Modify Routing

0
Normalise_CPE Connection
0
Normalise_Power AdptrConnct.
0
NSP_Activate/Deactive Signallng
0
NSP_Restart Signalling
0
NTT in Progress
0
NTT_Resolved
78
ODF Patch_Conntr. Repatched
0
ODF Patch_Opt TieCableReplaced
0
ODF Patch_Patch Cord Replaced
0
ODF Patching_Fiber Normalised
0
ODF repaired
0
OLT fan replaced
0
OLT IGMP setting reconfigured
0
OLT mcast setting reconfigured
0
OLT PON card reset
0
OLT reconfigured
30
ONU&DectPhone_Replaced
0
ONU&PG_Replaced
0
ONU&STB_Replaced
0
ONU_Change port
0
ONU_Config_Changed
50
ONU_Hang
0
ONU_Missing Config
0
ONU_Reconfigure password
0
ONU_Replaced
224
ONU_Reset
107
Others
0
Others_Prob_Action by RNO
0
Packaged_Added
0

Patch cord at cabinet replaced
0
Patch_Opt Connectors Cleaned
0
Patch_Opt Tie cable Repair
0
Patch_Opt Tie cable Replaced
0
Patching_Connector Repatched
0
Patching_Fiber Normalised
54
Patching_Opt Connct Replaced
94
Patching_Opt. Connct Repatched
0
Patching_Patch Cord Replaced
0
Patching_Replace Patch Cord
0
PG&Dect Phone_Replaced
0
PG&STB_Replaced
0
PG_Config_Changed
544
PG_Replaced
218
PG_Reset
0
PLAN MTCE BY CUST/VENDOR
0
Planned Maintenance Completed
0
PMW in Progress
0
Port Change
145
Port Reset
66
Power Adaptor Replaced
399
Premise Under Renovation
0
Problem_Restored
59
Re-activate
0
re-configure router
0
Reboot_OLT
0
Reboot_STB
0
Reconfig Sync
0
Replaced
0
Reset

0
Reset_OLT_Port
0
Reset_Password
0
Resolved By Account Service
0
Resolved By NTT
85
Resolved_By_ANOC
0
Resolved_By_Billing
0
Resolved_By_HBC
0
Resolved_By_NOC
0
Resolved_By_NOVA
33
Resolved_By_RNO
0
Resolved_By_Vendor
0
Restored by Others
0
RESTORED WHILE COORDINATING
0
Restored while localising
0
RETIGHTEN CONNECTOR BY CUST
0
Room aircond replaced
0
RWT CUSTOMER SATISFIED
166
SLT Phone replaced
0
STB&Dect Phone_Replaced
0
STB_Cable_Replaced
0
STB_Config_Changed
0
STB_Firmware_Downgraded
0
STB_Firmware_Upgraded
0
STB_MAC_Add_Registered
0
STB_Reboot
0
STB_Replaced
71
STB_Reset
0
Surge Arrestor Replacement
0
Tested Normal
0

TM TEST OK
0
TM_CPE_IW_RJ11 Repaired
0
TM_CPE_IW_RJ11 Replaced
61
TM_CPE_IW_RJ11 Replaced OTC
0
TM_CPE_IW_RJ45 Repaired
0
TM_CPE_IW_RJ45 Replaced
0
TM_CPE_IW_RJ45 Replaced OTC
0
TM_CPE_PG Config Problem
0
TM_CPE_PG Hang
0
TM_CPE_Voice Change Number
0
TM_FTB Drop Fiber repaired
0
TM_FTB Drop Fiber replaced
61
TM_FTB FASC connector replaced
122
TM_FTB Replaced
0
TMCPE Replaced
0
TNB Power Supply Restored
0
TNB/SESB/SESCO power restored
0
Trunk/uplink card replaced
0
Trunk/uplink card reset
0
Unbind_MAC_Add
0
User_Account
0
User_ID ReActivate
0
VDSL_Config_Changed
0
VDSL_Jumper Reconnect
0
VDSL_Modem&DectPhone_Replaced
0
VDSL_Modem&PG_Replaced
0
VDSL_Modem&STB_Replaced
0
VDSL_Replaced
470
VDSL_Reset
53
VLAN changed
109
VSAT ACTION BY CUSTOMER
0

Next update

Prediction via SVM