R version 3.3.3 (2017-03-06) -- "Another Canoe"

Copyright (C) 2017 The R Foundation for Statistical Computing

Platform: x86\_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.

You are welcome to redistribute it under certain conditions.

Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.

Type 'contributors()' for more information and

'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or

'help.start()' for an HTML browser interface to help.

Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

Error in as.environment(namespaceName) :

no item called "[[.data.frame()" on the search list

> #TRUE: server settings / FALSE: local settings

> isServerRun = FALSE

> model.name <- "LDA - Feature Selection"

> model.formula <-as.formula("factor(churn)~.")

>

> print(sprintf("model: %s formula: %s", model.name, deparse(model.formula)))

[1] "model: LDA - Feature Selection formula: factor(churn) ~ ."

>

> #########################################

> ############## Drive Config #############

> #########################################

> if(isServerRun){

+ setwd('/host/dsm1/fmare001/stats/svm/deliverables')

+ }else{

+ #setwd('C:/Users/audrey.ekuban/dev/goldsmiths/mlsdm/assignment3')

+ #setwd('C:/Users/john/dev/goldsmiths/mlsdm/assignment3')

+ setwd('C:/Users/Fred/Desktop/Studies/MSc-DataScience/Statistical Learning/Assignments/Assignment3/deliverables')

+ }

>

> #########################################

> ########### Load Dependencies ###########

> #########################################

> source("init\_data.r")

Loading required package: caret

Loading required package: lattice

Loading required package: ggplot2

Loading required package: DMwR

Loading required package: grid

> source("exploratory\_functions.r")

Loading required package: dplyr

Attaching package: ‘dplyr’

The following objects are masked from ‘package:stats’:

filter, lag

The following objects are masked from ‘package:base’:

intersect, setdiff, setequal, union

Loading required package: reshape2

> source("pre\_processing\_functions.r")

> source("feat\_selection.r")

randomForest 4.6-12

Type rfNews() to see new features/changes/bug fixes.

Attaching package: ‘randomForest’

The following object is masked from ‘package:dplyr’:

combine

The following object is masked from ‘package:ggplot2’:

margin

> require(randomForest)

>

> forceReloadPreCanned1 = TRUE

> if (forceReloadPreCanned1) {

+ #List the pre-processing functions

+ model.preProcessingFunctions <- c(

+ convert\_to\_factors,

+ drop\_na\_cols,

+ remove\_correlated\_predictors,

+ convert\_NAs\_to\_level,

+ remove\_linear\_dependencies,

+ bin\_negative\_levels\_churn,

+ keep\_top\_10\_levels,

+ impute\_data

+ )

+ # Need to do these up-front, otherwise we might end up with mis-matched

+ # levels between training and test folds in the cross-validation loop.

+ # Shouldn't introduce any bias since nothing is being imputed.

+ model.data <- apply\_pre\_processing(train, model.preProcessingFunctions)

+ write.csv(model.data, file = "train.bin.neg.top10.churn.csv", row.names = FALSE)

+ } else {

+ #reload from file

+ model.data <- read.csv("train.bin.neg.top10.churn.csv", stringsAsFactors = FALSE)

+ model.data <- convert\_to\_factors(model.data)

+ }

[1] "Number of removed linearly dependent col(s): 0"

>

>

> random\_forest\_feature\_selection(model.formula, model.data,500)

-1 1 MeanDecreaseAccuracy MeanDecreaseGini

V118 -2.6199382 100.4283703 24.9644096 425.383188

V224 -31.5345499 39.8470293 -10.3868338 190.869553

V165 35.6367976 78.2659105 68.4706058 180.331619

V84 18.6072804 30.1157981 32.1025012 163.825983

V155 21.1626288 -9.5775248 15.5688147 131.304265

V14 14.6157892 -0.4714576 14.6966475 127.521662

V116 22.3089917 21.0958957 36.5342824 116.058750

V149 43.3975999 -10.3134515 43.6123411 113.586061

V111 2.5029160 -0.6058037 1.8342359 109.424305

V156 2.6920071 0.9857594 2.8346678 107.417386

V70 6.2130942 2.8186126 7.1760059 103.982547

V192 15.5830785 -4.6396510 14.5083321 96.703571

V101 9.8095439 0.9288836 9.8038904 92.496492

V222 16.1566593 -7.9156901 14.7081646 92.361413

V161 22.3279203 -10.8889591 19.3483698 90.044591

V191 14.2061960 -1.1740478 13.8760361 88.053563

V141 13.5883309 -3.4314036 12.8318729 87.650779

V42 18.4316874 -7.4079607 18.1635202 85.706907

V55 23.0215395 -9.8091699 20.0110784 85.067424

V91 14.1242172 -3.0756211 13.2051111 79.108926

V134 16.4857750 -10.3650481 12.9391754 78.634142

V166 19.0718457 -8.1107503 17.9103954 72.775683

V151 20.4716359 -9.0942109 19.2472350 71.445441

V37 12.0160042 -7.2970581 11.1541426 70.511844

V119 15.1798294 -3.1695815 12.8134276 68.973829

V189 23.0257103 -10.1926138 19.9532008 68.855275

V124 15.9741111 -3.9682819 14.7068488 68.446782

V121 10.5895596 -3.6379619 8.9431944 67.553537

V1 11.6622197 -2.9981391 10.8736242 66.986688

V81 14.3636074 -7.0442230 14.5109790 64.678335

V96 18.2584174 -9.8533894 17.3085657 63.615316

V202 15.9256077 -6.0559496 13.8620182 62.479325

V75 18.5275859 -6.4884738 16.3082605 61.780022

V197 16.0526309 -8.5737438 13.3328282 60.826783

V187 15.7570417 12.4216914 19.3918823 58.940110

V90 15.4613235 10.8007187 19.0136859 58.548376

V93 17.5317876 10.6980281 20.4236418 57.983448

V206 17.3714158 -5.0987751 16.1652980 57.649808

V71 14.6046587 -4.8316845 12.6748578 47.971994

V15 15.0589103 -4.3678443 13.7830331 46.594490

V150 12.1554412 -2.1842956 12.4962430 45.107210

V21 11.5330043 -5.4137898 9.3397560 44.253849

V67 17.0051877 -10.4996185 14.9408775 44.187246

V143 -6.1839512 7.2968563 -1.4463186 41.669496

V190 20.4480496 -13.7014614 19.7605024 35.940938

V58 5.6977156 -3.8450442 3.3340589 31.170821

V29 20.1401819 -6.5893192 19.7678224 29.580360

V107 5.9104077 1.3417254 6.7956819 29.516523

V115 13.9089446 -6.2310459 13.5335665 28.154219

V204 16.9550796 -7.4204158 16.5957844 27.373223

V212 14.7911173 -4.7462976 15.3681611 26.951612

V138 8.2003876 -0.7964282 7.4063008 26.621891

V114 4.5232749 -0.8315718 3.6983800 25.509016

V95 8.3010071 -2.8239620 7.9637961 25.499419

V112 13.4490191 -2.7739510 14.0088717 24.666796

V79 1.2483943 -1.2393679 0.5147195 24.364353

upselling -14.9311827 27.9443974 14.9867492 20.562757

V128 17.4435724 -10.9879192 16.0864924 20.552931

V154 11.1435301 -4.8913185 11.0333890 20.163867

V30 5.8462771 -1.4812895 5.1832209 17.251225

V171 2.4250734 -1.4933854 1.8696536 16.439751

V168 11.3318173 -3.5904151 11.4511929 12.575271

V230 7.8854753 -6.2794925 7.5013899 11.378421

V196 2.6471682 -1.9690063 1.3628226 11.302171

V11 8.9483746 -3.8961530 8.0384216 8.741601

V125 4.5453722 -0.1210787 4.1532009 7.328994

V113 7.7550632 -4.6636178 7.6265013 5.381543

V94 8.6729872 -4.5272405 8.2223165 4.867223

appetency 8.0294117 9.9756779 10.0782101 4.519600

V219 8.0102242 -6.7770857 7.7261976 3.246380

V73 0.8632167 1.3407685 1.3069686 3.142204

