Assignment2-Script.R

omkar

2021-04-27

#install.packages('mlbench')  
#install.packages('rpart')  
#install.packages('rpart.plot')  
#install.packages('caret')  
#install.packages('lattice')  
#install.packages('e1071')  
#install.packages('doParallel')  
#install.packages('foreach')  
#install.packages('iterators')  
#install.packages('parallel')  
  
# Omkar Jagdale osj170000 CS4375.0W2  
  
library(mlbench)

## Warning: package 'mlbench' was built under R version 4.0.4

library(rpart)

## Warning: package 'rpart' was built under R version 4.0.4

library(rpart.plot)

## Warning: package 'rpart.plot' was built under R version 4.0.4

library(ggplot2)  
library(lattice)

## Warning: package 'lattice' was built under R version 4.0.4

library(e1071)

## Warning: package 'e1071' was built under R version 4.0.4

library(caret)

## Warning: package 'caret' was built under R version 4.0.4

library(parallel)  
library(iterators)

## Warning: package 'iterators' was built under R version 4.0.4

library(foreach)

## Warning: package 'foreach' was built under R version 4.0.4

library(doParallel)

## Warning: package 'doParallel' was built under R version 4.0.4

library(dplyr)

## Warning: package 'dplyr' was built under R version 4.0.4

##   
## 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

mydata <- read.csv("C:\\Users\\omkar\\Downloads\\bank-additional.csv", sep=";", header = TRUE)  
summary(mydata)

## age job marital education   
## Min. :18.00 Length:4119 Length:4119 Length:4119   
## 1st Qu.:32.00 Class :character Class :character Class :character   
## Median :38.00 Mode :character Mode :character Mode :character   
## Mean :40.11   
## 3rd Qu.:47.00   
## Max. :88.00   
## default housing loan contact   
## Length:4119 Length:4119 Length:4119 Length:4119   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## month day\_of\_week duration campaign   
## Length:4119 Length:4119 Min. : 0.0 Min. : 1.000   
## Class :character Class :character 1st Qu.: 103.0 1st Qu.: 1.000   
## Mode :character Mode :character Median : 181.0 Median : 2.000   
## Mean : 256.8 Mean : 2.537   
## 3rd Qu.: 317.0 3rd Qu.: 3.000   
## Max. :3643.0 Max. :35.000   
## pdays previous poutcome emp.var.rate   
## Min. : 0.0 Min. :0.0000 Length:4119 Min. :-3.40000   
## 1st Qu.:999.0 1st Qu.:0.0000 Class :character 1st Qu.:-1.80000   
## Median :999.0 Median :0.0000 Mode :character Median : 1.10000   
## Mean :960.4 Mean :0.1903 Mean : 0.08497   
## 3rd Qu.:999.0 3rd Qu.:0.0000 3rd Qu.: 1.40000   
## Max. :999.0 Max. :6.0000 Max. : 1.40000   
## cons.price.idx cons.conf.idx euribor3m nr.employed   
## Min. :92.20 Min. :-50.8 Min. :0.635 Min. :4964   
## 1st Qu.:93.08 1st Qu.:-42.7 1st Qu.:1.334 1st Qu.:5099   
## Median :93.75 Median :-41.8 Median :4.857 Median :5191   
## Mean :93.58 Mean :-40.5 Mean :3.621 Mean :5166   
## 3rd Qu.:93.99 3rd Qu.:-36.4 3rd Qu.:4.961 3rd Qu.:5228   
## Max. :94.77 Max. :-26.9 Max. :5.045 Max. :5228   
## y   
## Length:4119   
## Class :character   
## Mode :character   
##   
##   
##

mydataNA <- na.omit(mydata)  
improvedMyData <- unique(mydataNA[complete.cases(mydataNA),])  
  
for(i in c(2:10, 15))  
 improvedMyData[[i]] <- as.factor(improvedMyData[[i]])  
  
summary(improvedMyData)

## age job marital education   
## Min. :18.00 admin. :1012 divorced: 446 university.degree :1264   
## 1st Qu.:32.00 blue-collar: 884 married :2509 high.school : 921   
## Median :38.00 technician : 691 single :1153 basic.9y : 574   
## Mean :40.11 services : 393 unknown : 11 professional.course: 535   
## 3rd Qu.:47.00 management : 324 basic.4y : 429   
## Max. :88.00 retired : 166 basic.6y : 228   
## (Other) : 649 (Other) : 168   
## default housing loan contact month   
## no :3315 no :1839 no :3349 cellular :2652 may :1378   
## unknown: 803 unknown: 105 unknown: 105 telephone:1467 jul : 711   
## yes : 1 yes :2175 yes : 665 aug : 636   
## jun : 530   
## nov : 446   
## apr : 215   
## (Other): 203   
## day\_of\_week duration campaign pdays previous   
## fri:768 Min. : 0.0 Min. : 1.000 Min. : 0.0 Min. :0.0000   
## mon:855 1st Qu.: 103.0 1st Qu.: 1.000 1st Qu.:999.0 1st Qu.:0.0000   
## thu:860 Median : 181.0 Median : 2.000 Median :999.0 Median :0.0000   
## tue:841 Mean : 256.8 Mean : 2.537 Mean :960.4 Mean :0.1903   
## wed:795 3rd Qu.: 317.0 3rd Qu.: 3.000 3rd Qu.:999.0 3rd Qu.:0.0000   
## Max. :3643.0 Max. :35.000 Max. :999.0 Max. :6.0000   
##   
## poutcome emp.var.rate cons.price.idx cons.conf.idx   
## failure : 454 Min. :-3.40000 Min. :92.20 Min. :-50.8   
## nonexistent:3523 1st Qu.:-1.80000 1st Qu.:93.08 1st Qu.:-42.7   
## success : 142 Median : 1.10000 Median :93.75 Median :-41.8   
## Mean : 0.08497 Mean :93.58 Mean :-40.5   
## 3rd Qu.: 1.40000 3rd Qu.:93.99 3rd Qu.:-36.4   
## Max. : 1.40000 Max. :94.77 Max. :-26.9   
##   
## euribor3m nr.employed y   
## Min. :0.635 Min. :4964 Length:4119   
## 1st Qu.:1.334 1st Qu.:5099 Class :character   
## Median :4.857 Median :5191 Mode :character   
## Mean :3.621 Mean :5166   
## 3rd Qu.:4.961 3rd Qu.:5228   
## Max. :5.045 Max. :5228   
##

#A  
  
  
accuracy <- function(truth, prediction) {  
 tbl <- table(truth, prediction)  
 sum(diag(tbl))/sum(tbl)  
}

index <- 1:nrow(improvedMyData)  
index <- sample(index)  
  
fold <- rep(1:10, each=nrow(improvedMyData)/10)[1:nrow(improvedMyData)]  
  
folds <- split(index, fold) ### create list with indices for each fold  
  
summary(index)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1 1030 2060 2060 3090 4119

testingError <- vector(mode="numeric")  
trainingError <- vector(mode="numeric")  
# Create a tree for each fold and run it on the first fold  
# i.e. the testing set  
############################################  
# 6 in Assignment 4  
for(i in 1:length(folds)) {  
 tree <- rpart(y ~., data=improvedMyData[-folds[[i]],], control=rpart.control(minsplit=2, cp=0.128497202))  
 accs <- accuracy(improvedMyData[folds[[i]],]$y, predict(tree, improvedMyData[folds[[i]],], type="class"))  
 testingError[i] <- 1 - accs  
 trainingError[i] <- 1 - accuracy(improvedMyData[-folds[[i]],]$y, predict(tree, improvedMyData[-folds[[i]],], type="class"))  
 }  
testingError

## [1] 0.10705596 0.09489051 0.09245742 0.10462287 0.11192214 0.12652068  
## [7] 0.09732360 0.11435523 0.11678832 0.13138686

trainingError

## [1] 0.1097627 0.1111111 0.1113808 0.1100324 0.1092233 0.1076052 0.1108414  
## [8] 0.1089536 0.1086839 0.1070658

mean(accs)

## [1] 0.8686131

mean(testingError)

## [1] 0.1097324

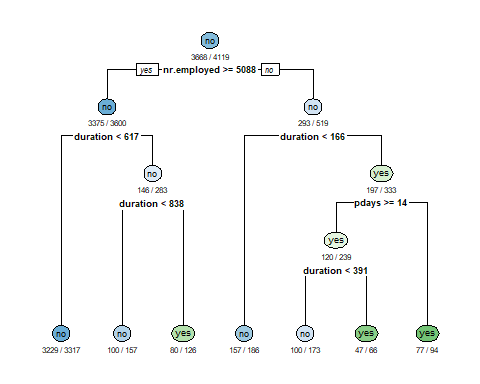
# cp=0.003197442  
# Average accuracy = 0.9148418  
# Average testing error = 0.09902676  
  
# cp=0.006705747  
# Average accuracy = 0.9051095  
# Average testing error = 0.0946472  
  
# cp=0.036903810  
# Average accuracy = 0.9148418  
# Average testing error = 0.09440389  
  
# cp=0.064481748  
# Average accuracy = 0.8734793  
# Average testing error = 0.1060827  
  
# cp=0.128497202  
# Average accuracy = 0.8734793  
# Average testing error = 1097324  
  
# cp = 0.003197442 and cp = 0.036903810 have the highest accuracy  
# of 0.9148418. However, cp = 0.036903810 has a lower testing error  
# of 0.09440389 as compared to cp = 0.003197442's 0.009902676.  
  
# cp = 0.036903810 is the choice  
######################################################################  
#7  
  
fit <- train(y ~ ., data = improvedMyData , method = "rpart",  
 control=rpart.control(minsplit=2),  
 trControl = trainControl(method = "cv", number = 10),  
 tuneLength=5)  
  
fit

## CART   
##   
## 4119 samples  
## 20 predictor  
## 2 classes: 'no', 'yes'   
##   
## No pre-processing  
## Resampling: Cross-Validated (10 fold)   
## Summary of sample sizes: 3707, 3706, 3707, 3707, 3707, 3708, ...   
## Resampling results across tuning parameters:  
##   
## cp Accuracy Kappa   
## 0.009977827 0.9145453 0.4902842  
## 0.011086475 0.9155161 0.4985705  
## 0.029933481 0.9111454 0.5239550  
## 0.037694013 0.9041037 0.4533486  
## 0.067627494 0.8939124 0.1904793  
##   
## Accuracy was used to select the optimal model using the largest value.  
## The final value used for the model was cp = 0.01108647.

fit$finalModel

## n= 4119   
##   
## node), split, n, loss, yval, (yprob)  
## \* denotes terminal node  
##   
## 1) root 4119 451 no (0.8905074 0.1094926)   
## 2) nr.employed>=5087.65 3600 225 no (0.9375000 0.0625000)   
## 4) duration< 616.5 3317 88 no (0.9734700 0.0265300) \*  
## 5) duration>=616.5 283 137 no (0.5159011 0.4840989)   
## 10) duration< 837.5 157 57 no (0.6369427 0.3630573) \*  
## 11) duration>=837.5 126 46 yes (0.3650794 0.6349206) \*  
## 3) nr.employed< 5087.65 519 226 no (0.5645472 0.4354528)   
## 6) duration< 165.5 186 29 no (0.8440860 0.1559140) \*  
## 7) duration>=165.5 333 136 yes (0.4084084 0.5915916)   
## 14) pdays>=13.5 239 119 yes (0.4979079 0.5020921)   
## 28) duration< 390.5 173 73 no (0.5780347 0.4219653) \*  
## 29) duration>=390.5 66 19 yes (0.2878788 0.7121212) \*  
## 15) pdays< 13.5 94 17 yes (0.1808511 0.8191489) \*

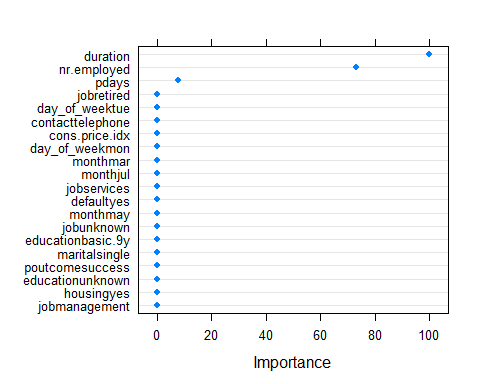
# The three most important attributes are  
# duration, nr.employed, and pdays  
  
rpart.plot(fit$finalModel, extra = 2, under = TRUE, varlen=0, faclen = 0)



varImp(fit, compete = FALSE)

## rpart variable importance  
##   
## only 20 most important variables shown (out of 53)  
##   
## Overall  
## duration 100.000  
## nr.employed 72.996  
## pdays 7.846  
## jobservices 0.000  
## educationhigh.school 0.000  
## housingyes 0.000  
## age 0.000  
## euribor3m 0.000  
## campaign 0.000  
## cons.conf.idx 0.000  
## educationprofessional.course 0.000  
## `jobblue-collar` 0.000  
## poutcomesuccess 0.000  
## monthmar 0.000  
## defaultunknown 0.000  
## loanunknown 0.000  
## day\_of\_weekwed 0.000  
## previous 0.000  
## educationbasic.9y 0.000  
## jobretired 0.000

dotPlot(varImp(fit, compete = FALSE))



#######################################################################  
#8  
# The 5 attributes used are:  
# duration, nr.employed, pdays, month, emp.var.rate  
subsetMyData <- improvedMyData[,c(9, 11, 13, 16, 20, 21)]  
subsetMyData[[1]] <- as.numeric(subsetMyData[[1]])  
myData\_scaled <- cbind(as.data.frame(scale(subsetMyData[,-6])), type = subsetMyData[,6])  
  
#######################################################################  
#9  
trainMyData <- createFolds(myData\_scaled$type, k=10)  
trainMyData

## $Fold01  
## [1] 6 12 13 38 39 44 76 81 106 108 110 111 117 125 135  
## [16] 147 164 172 207 215 218 224 228 231 233 238 248 249 254 268  
## [31] 291 295 303 309 318 320 323 363 367 383 387 392 415 428 436  
## [46] 441 452 454 456 459 481 483 497 509 515 541 548 552 557 561  
## [61] 585 597 619 639 649 656 666 678 696 709 730 744 753 784 798  
## [76] 823 824 829 856 868 876 897 904 914 920 937 938 960 964 984  
## [91] 987 996 997 1008 1014 1030 1053 1054 1061 1063 1092 1106 1107 1120 1137  
## [106] 1144 1153 1161 1168 1207 1213 1215 1244 1263 1279 1285 1294 1318 1319 1327  
## [121] 1332 1350 1361 1367 1373 1384 1389 1391 1396 1398 1402 1434 1468 1471 1485  
## [136] 1508 1526 1540 1544 1555 1568 1594 1596 1605 1610 1634 1637 1646 1661 1664  
## [151] 1674 1678 1681 1688 1706 1710 1713 1725 1741 1753 1756 1772 1784 1819 1820  
## [166] 1831 1832 1833 1841 1863 1876 1877 1898 1907 1924 1933 1942 1946 1953 1961  
## [181] 1986 2026 2028 2031 2032 2049 2058 2062 2098 2134 2139 2150 2162 2164 2165  
## [196] 2192 2211 2217 2220 2259 2294 2303 2314 2321 2323 2326 2341 2343 2352 2355  
## [211] 2359 2361 2364 2365 2367 2377 2407 2410 2412 2417 2418 2422 2434 2450 2461  
## [226] 2486 2490 2494 2497 2501 2506 2509 2520 2521 2524 2533 2541 2549 2556 2559  
## [241] 2561 2579 2589 2605 2606 2609 2619 2635 2653 2655 2659 2667 2690 2694 2710  
## [256] 2734 2744 2752 2800 2810 2829 2841 2846 2864 2889 2892 2904 2911 2915 2922  
## [271] 2924 2930 2936 2975 2981 2985 2988 3001 3002 3005 3007 3012 3014 3030 3040  
## [286] 3041 3042 3061 3068 3072 3077 3081 3082 3084 3088 3092 3117 3118 3150 3152  
## [301] 3153 3166 3169 3176 3177 3197 3201 3209 3210 3214 3221 3235 3242 3244 3252  
## [316] 3255 3256 3258 3285 3299 3310 3321 3334 3337 3346 3357 3361 3373 3374 3400  
## [331] 3401 3406 3412 3419 3429 3440 3445 3459 3471 3477 3485 3489 3496 3499 3508  
## [346] 3527 3528 3544 3551 3557 3562 3569 3574 3589 3597 3602 3619 3621 3628 3631  
## [361] 3634 3654 3667 3694 3696 3697 3701 3702 3706 3708 3718 3753 3762 3764 3769  
## [376] 3774 3799 3804 3808 3815 3831 3835 3848 3864 3874 3888 3892 3894 3895 3920  
## [391] 3925 3946 3949 3952 3959 3971 3977 3979 3988 3990 3992 3996 4007 4009 4024  
## [406] 4038 4066 4067 4095 4104 4106 4109 4116  
##   
## $Fold02  
## [1] 9 21 22 26 45 46 49 63 100 148 159 183 221 242 250  
## [16] 261 266 267 276 286 287 294 300 304 311 312 317 330 332 333  
## [31] 337 340 343 362 368 371 372 382 386 438 443 461 467 480 490  
## [46] 492 493 494 499 523 536 537 545 547 549 559 565 569 571 593  
## [61] 634 647 655 668 684 691 693 694 711 765 781 786 792 805 809  
## [76] 831 839 840 849 853 887 913 916 924 973 986 993 1002 1003 1010  
## [91] 1011 1015 1038 1041 1049 1058 1071 1087 1088 1089 1104 1122 1136 1145 1157  
## [106] 1158 1160 1183 1195 1202 1216 1227 1233 1236 1246 1247 1256 1258 1262 1268  
## [121] 1275 1278 1281 1297 1339 1348 1371 1372 1374 1380 1387 1410 1423 1428 1430  
## [136] 1442 1446 1451 1466 1472 1480 1484 1487 1494 1501 1505 1517 1528 1549 1550  
## [151] 1557 1562 1572 1574 1577 1580 1581 1586 1609 1626 1663 1668 1670 1685 1699  
## [166] 1701 1702 1703 1708 1715 1723 1728 1732 1757 1769 1774 1798 1815 1824 1835  
## [181] 1846 1856 1870 1880 1882 1887 1889 1909 1915 1917 1932 1944 1945 1966 1969  
## [196] 1978 1994 2008 2017 2024 2033 2038 2039 2045 2048 2051 2064 2081 2082 2091  
## [211] 2094 2103 2106 2110 2122 2126 2142 2156 2168 2191 2200 2210 2240 2241 2262  
## [226] 2273 2275 2302 2304 2306 2324 2328 2349 2368 2369 2380 2381 2385 2395 2397  
## [241] 2405 2408 2411 2425 2430 2459 2462 2471 2477 2518 2543 2550 2567 2570 2587  
## [256] 2595 2599 2608 2610 2618 2624 2627 2642 2671 2682 2686 2689 2693 2707 2717  
## [271] 2723 2729 2770 2774 2782 2784 2786 2788 2804 2809 2813 2815 2819 2824 2830  
## [286] 2833 2835 2839 2884 2891 2910 2932 2969 2971 2996 3000 3011 3017 3026 3027  
## [301] 3046 3059 3070 3099 3107 3108 3112 3113 3128 3133 3137 3163 3190 3228 3230  
## [316] 3232 3246 3253 3277 3291 3302 3311 3312 3316 3345 3348 3356 3359 3367 3383  
## [331] 3384 3389 3392 3393 3404 3415 3423 3425 3427 3430 3431 3433 3449 3452 3453  
## [346] 3460 3472 3480 3483 3493 3511 3516 3526 3536 3552 3558 3565 3577 3580 3611  
## [361] 3614 3617 3620 3633 3641 3660 3661 3663 3672 3681 3692 3698 3703 3705 3713  
## [376] 3714 3728 3760 3775 3778 3780 3809 3829 3832 3833 3844 3850 3851 3852 3865  
## [391] 3873 3890 3899 3912 3921 3928 3933 3936 3944 3951 3963 3967 3970 3975 3986  
## [406] 4026 4032 4036 4051 4061 4069 4092  
##   
## $Fold03  
## [1] 1 3 5 8 10 23 30 31 41 42 59 74 84 85 86  
## [16] 89 99 101 116 122 124 144 145 153 160 163 166 168 174 178  
## [31] 182 187 209 234 251 255 256 269 270 302 321 339 344 346 365  
## [46] 377 397 401 410 421 435 449 470 496 504 506 508 512 513 517  
## [61] 521 531 538 579 583 584 591 595 608 610 611 621 628 641 654  
## [76] 672 683 697 701 720 723 736 739 751 758 760 773 774 776 782  
## [91] 793 802 810 821 830 834 835 842 850 863 869 872 880 902 907  
## [106] 915 922 926 933 936 946 952 959 989 995 1001 1005 1007 1022 1032  
## [121] 1042 1043 1064 1067 1068 1103 1117 1123 1130 1138 1149 1159 1166 1176 1179  
## [136] 1185 1190 1197 1218 1223 1245 1248 1255 1295 1303 1308 1323 1333 1356 1357  
## [151] 1359 1362 1364 1366 1369 1370 1388 1418 1422 1437 1461 1467 1478 1488 1491  
## [166] 1499 1502 1507 1513 1521 1523 1538 1541 1543 1547 1553 1571 1623 1655 1666  
## [181] 1667 1694 1698 1707 1721 1722 1739 1746 1773 1780 1788 1792 1808 1825 1838  
## [196] 1839 1843 1852 1862 1864 1867 1869 1873 1878 1892 1893 1894 1897 1902 1904  
## [211] 1913 1923 1930 1931 1939 1943 1948 1959 1987 1989 1991 2003 2015 2052 2053  
## [226] 2063 2090 2095 2108 2113 2114 2119 2141 2172 2184 2185 2186 2188 2227 2234  
## [241] 2254 2267 2270 2272 2284 2298 2301 2311 2333 2335 2342 2345 2370 2371 2372  
## [256] 2392 2393 2398 2403 2414 2416 2426 2441 2442 2443 2445 2505 2519 2538 2539  
## [271] 2542 2565 2568 2574 2578 2581 2615 2631 2633 2649 2657 2672 2677 2692 2697  
## [286] 2705 2721 2737 2739 2746 2753 2805 2811 2820 2826 2828 2838 2844 2856 2862  
## [301] 2893 2928 2949 2952 2978 2982 2993 2995 3019 3023 3034 3038 3053 3080 3087  
## [316] 3091 3098 3125 3135 3143 3155 3172 3184 3193 3216 3217 3248 3280 3281 3284  
## [331] 3289 3297 3300 3309 3313 3322 3323 3347 3351 3353 3370 3377 3395 3396 3399  
## [346] 3413 3414 3434 3455 3466 3509 3519 3521 3560 3568 3572 3575 3587 3615 3616  
## [361] 3623 3624 3629 3638 3642 3655 3658 3664 3669 3693 3711 3712 3722 3730 3744  
## [376] 3750 3751 3757 3759 3765 3772 3792 3797 3802 3806 3821 3828 3837 3838 3876  
## [391] 3891 3893 3904 3910 3914 3934 3942 3953 3954 3955 3958 3987 3997 4018 4052  
## [406] 4082 4087 4096 4108 4110 4117 4118  
##   
## $Fold04  
## [1] 16 24 34 52 57 58 67 90 96 119 132 146 152 157 167  
## [16] 180 181 197 199 208 211 217 247 257 274 284 310 327 328 338  
## [31] 351 357 374 378 391 396 407 411 433 437 439 442 447 460 472  
## [46] 498 516 528 530 534 540 542 544 556 568 578 581 589 600 622  
## [61] 643 644 652 657 658 663 665 677 689 726 737 745 783 873 875  
## [76] 881 885 895 909 911 929 983 1000 1016 1020 1025 1031 1046 1065 1066  
## [91] 1070 1073 1102 1116 1143 1147 1150 1155 1156 1164 1169 1173 1180 1187 1204  
## [106] 1209 1214 1225 1231 1232 1235 1237 1241 1270 1284 1287 1289 1302 1304 1320  
## [121] 1331 1338 1377 1381 1408 1432 1435 1438 1443 1445 1447 1465 1469 1476 1486  
## [136] 1496 1503 1504 1506 1510 1511 1512 1515 1524 1532 1536 1542 1588 1590 1595  
## [151] 1613 1619 1620 1624 1628 1629 1636 1649 1659 1669 1682 1684 1686 1696 1717  
## [166] 1724 1730 1734 1740 1743 1778 1781 1783 1791 1794 1796 1805 1810 1817 1829  
## [181] 1861 1868 1874 1905 1908 1914 1928 1938 1940 1954 1962 1967 1990 1992 1996  
## [196] 1998 2007 2014 2018 2020 2022 2036 2040 2061 2071 2074 2111 2121 2133 2157  
## [211] 2167 2170 2179 2183 2187 2189 2208 2213 2216 2218 2222 2230 2232 2261 2263  
## [226] 2276 2281 2283 2292 2297 2307 2318 2332 2360 2362 2363 2389 2390 2394 2421  
## [241] 2431 2482 2492 2493 2495 2503 2507 2508 2514 2540 2555 2563 2588 2593 2602  
## [256] 2616 2630 2643 2658 2687 2703 2704 2712 2713 2718 2742 2749 2777 2778 2780  
## [271] 2792 2831 2843 2854 2861 2863 2866 2875 2881 2902 2907 2919 2929 2933 2935  
## [286] 2940 2945 2946 2947 2954 2958 2973 2998 3009 3025 3035 3057 3062 3064 3067  
## [301] 3085 3094 3106 3111 3122 3132 3151 3157 3164 3185 3186 3192 3198 3204 3233  
## [316] 3240 3251 3263 3264 3267 3272 3288 3304 3324 3325 3330 3335 3380 3381 3409  
## [331] 3432 3438 3444 3446 3484 3497 3498 3503 3514 3518 3523 3529 3533 3539 3566  
## [346] 3578 3583 3586 3603 3612 3625 3637 3643 3653 3656 3668 3682 3685 3699 3700  
## [361] 3719 3724 3725 3729 3739 3752 3755 3782 3796 3810 3814 3823 3834 3836 3839  
## [376] 3855 3858 3872 3877 3885 3917 3927 3929 3931 3932 3940 3956 3968 3984 3995  
## [391] 4003 4006 4025 4028 4039 4040 4041 4048 4056 4058 4068 4075 4078 4079 4081  
## [406] 4089 4093 4094 4097 4100 4103 4107  
##   
## $Fold05  
## [1] 15 20 25 29 32 55 70 77 91 103 109 115 120 134 139  
## [16] 150 156 169 170 173 185 195 202 212 223 225 239 245 260 262  
## [31] 278 289 313 326 373 379 381 384 385 398 399 409 427 445 462  
## [46] 469 478 505 507 511 514 518 520 524 526 529 532 553 554 560  
## [61] 577 586 594 598 603 605 631 648 661 686 703 721 735 738 743  
## [76] 746 749 761 770 775 780 789 816 820 836 847 860 874 882 886  
## [91] 888 901 912 921 942 948 951 962 966 972 980 981 991 998 1009  
## [106] 1019 1029 1035 1039 1040 1056 1072 1074 1076 1085 1109 1141 1177 1181 1191  
## [121] 1192 1211 1224 1229 1238 1243 1265 1272 1280 1292 1300 1316 1317 1326 1334  
## [136] 1337 1341 1351 1360 1378 1383 1390 1400 1401 1406 1407 1413 1417 1420 1424  
## [151] 1425 1427 1453 1460 1463 1470 1482 1493 1514 1519 1520 1537 1561 1564 1566  
## [166] 1569 1578 1603 1614 1618 1621 1638 1639 1652 1704 1705 1726 1735 1736 1744  
## [181] 1752 1759 1771 1776 1790 1802 1804 1828 1830 1847 1859 1911 1912 1918 1926  
## [196] 1934 1950 1951 1974 1993 2001 2006 2009 2035 2066 2070 2075 2089 2124 2131  
## [211] 2144 2146 2159 2194 2206 2209 2221 2233 2239 2243 2251 2252 2257 2258 2260  
## [226] 2266 2269 2282 2293 2300 2316 2317 2329 2330 2334 2340 2348 2356 2374 2384  
## [241] 2401 2409 2423 2432 2438 2451 2460 2466 2480 2483 2502 2512 2534 2535 2552  
## [256] 2572 2575 2617 2632 2654 2661 2666 2683 2685 2699 2714 2720 2732 2740 2747  
## [271] 2755 2771 2772 2776 2787 2789 2794 2797 2812 2814 2823 2848 2851 2860 2878  
## [286] 2886 2890 2906 2914 2918 2920 2934 2950 2962 2967 2970 2994 3006 3008 3024  
## [301] 3036 3045 3052 3054 3063 3095 3104 3109 3116 3167 3168 3175 3187 3188 3194  
## [316] 3196 3199 3222 3224 3225 3238 3254 3265 3269 3270 3276 3279 3293 3308 3315  
## [331] 3319 3328 3336 3338 3350 3388 3390 3398 3405 3418 3424 3428 3441 3442 3443  
## [346] 3458 3467 3468 3470 3473 3478 3492 3502 3504 3505 3524 3550 3553 3556 3561  
## [361] 3581 3588 3590 3596 3601 3607 3635 3645 3649 3659 3665 3666 3671 3674 3679  
## [376] 3686 3717 3721 3723 3737 3746 3748 3763 3790 3816 3820 3840 3847 3849 3861  
## [391] 3866 3878 3879 3884 3896 3915 3916 3919 3922 3938 3950 3993 3998 4016 4029  
## [406] 4033 4034 4037 4085 4090 4112 4113  
##   
## $Fold06  
## [1] 4 14 19 50 69 73 75 93 95 113 129 133 155 161 171  
## [16] 191 200 204 205 210 216 232 241 244 253 259 272 273 280 283  
## [31] 335 336 341 342 349 354 355 366 370 380 389 400 417 419 432  
## [46] 434 457 466 473 476 484 488 501 502 522 572 573 582 588 596  
## [61] 606 626 650 667 671 681 682 685 690 699 700 707 708 714 718  
## [76] 719 725 731 733 734 741 754 755 767 778 785 797 800 812 817  
## [91] 819 838 841 848 858 878 884 903 917 928 941 945 950 969 970  
## [106] 982 985 990 1026 1028 1034 1044 1048 1051 1060 1069 1075 1082 1083 1095  
## [121] 1099 1110 1111 1112 1114 1119 1132 1146 1162 1170 1178 1198 1199 1226 1228  
## [136] 1239 1249 1254 1269 1271 1282 1283 1286 1301 1312 1330 1335 1349 1354 1355  
## [151] 1368 1382 1403 1404 1419 1440 1449 1450 1464 1477 1489 1522 1525 1530 1534  
## [166] 1554 1556 1560 1567 1570 1575 1582 1583 1591 1608 1633 1645 1653 1658 1673  
## [181] 1687 1689 1718 1742 1747 1754 1761 1763 1764 1765 1770 1775 1777 1793 1800  
## [196] 1806 1809 1849 1850 1855 1857 1858 1881 1886 1890 1895 1919 1941 1947 1949  
## [211] 1952 1956 1957 1958 1964 1970 1997 2000 2010 2041 2042 2046 2060 2077 2078  
## [226] 2080 2096 2101 2104 2125 2132 2138 2151 2161 2163 2169 2173 2177 2193 2196  
## [241] 2199 2201 2203 2214 2229 2253 2271 2277 2290 2309 2315 2331 2379 2386 2436  
## [256] 2444 2453 2455 2463 2467 2474 2485 2513 2523 2536 2544 2545 2546 2548 2553  
## [271] 2557 2583 2598 2611 2626 2646 2648 2674 2675 2754 2761 2768 2791 2793 2818  
## [286] 2825 2840 2845 2847 2857 2859 2865 2868 2883 2899 2901 2903 2923 2926 2941  
## [301] 2944 2956 2959 2965 2966 2976 2980 2984 2989 2990 2992 3003 3004 3015 3021  
## [316] 3037 3049 3055 3076 3079 3089 3119 3129 3165 3181 3202 3208 3227 3243 3245  
## [331] 3249 3250 3261 3268 3286 3306 3307 3314 3317 3327 3342 3352 3354 3366 3382  
## [346] 3394 3402 3420 3426 3435 3462 3463 3507 3512 3513 3534 3545 3547 3563 3570  
## [361] 3573 3594 3599 3600 3606 3608 3609 3636 3644 3657 3662 3676 3678 3684 3688  
## [376] 3731 3740 3756 3768 3781 3783 3791 3800 3818 3819 3822 3827 3842 3845 3881  
## [391] 3886 3907 3908 3913 3924 3937 3945 3981 3989 4008 4013 4014 4019 4035 4042  
## [406] 4045 4050 4057 4060 4074 4086  
##   
## $Fold07  
## [1] 7 17 28 53 64 65 80 92 97 105 121 127 131 138 142  
## [16] 151 165 186 189 213 214 237 252 279 299 305 306 307 308 322  
## [31] 324 325 331 347 348 356 358 361 364 408 412 413 418 422 423  
## [46] 440 463 479 482 487 495 500 503 567 570 574 590 612 617 623  
## [61] 624 625 629 630 636 646 653 662 669 670 706 710 712 713 717  
## [76] 724 750 756 762 769 795 799 826 827 828 843 851 859 862 865  
## [91] 890 892 898 900 934 947 961 967 974 976 1006 1012 1023 1027 1036  
## [106] 1052 1055 1084 1090 1096 1097 1100 1113 1126 1140 1148 1188 1194 1219 1222  
## [121] 1250 1257 1259 1266 1306 1311 1325 1344 1353 1376 1385 1393 1395 1411 1416  
## [136] 1421 1429 1436 1454 1456 1457 1459 1479 1490 1495 1565 1573 1593 1601 1630  
## [151] 1632 1635 1640 1641 1651 1656 1657 1660 1662 1677 1679 1680 1695 1714 1733  
## [166] 1738 1789 1801 1822 1827 1837 1840 1848 1853 1860 1871 1903 1921 1937 1960  
## [181] 1968 1972 1975 2011 2013 2021 2023 2037 2043 2044 2056 2057 2059 2079 2084  
## [196] 2088 2100 2102 2105 2107 2112 2123 2130 2145 2149 2154 2155 2166 2175 2176  
## [211] 2182 2205 2212 2215 2219 2223 2225 2226 2237 2249 2264 2268 2280 2286 2289  
## [226] 2305 2308 2322 2336 2347 2353 2357 2358 2366 2391 2399 2400 2446 2447 2454  
## [241] 2457 2473 2476 2487 2491 2500 2510 2564 2566 2571 2577 2580 2586 2591 2594  
## [256] 2600 2604 2622 2637 2638 2639 2647 2668 2679 2688 2701 2702 2706 2711 2715  
## [271] 2724 2741 2745 2760 2762 2763 2769 2773 2775 2796 2799 2801 2817 2836 2850  
## [286] 2855 2867 2870 2871 2872 2873 2874 2877 2879 2898 2905 2912 2937 2948 2986  
## [301] 2991 3018 3031 3039 3044 3050 3056 3058 3060 3069 3071 3078 3083 3103 3114  
## [316] 3121 3124 3138 3156 3160 3161 3170 3171 3178 3182 3183 3189 3195 3211 3215  
## [331] 3219 3226 3234 3237 3266 3274 3290 3294 3331 3341 3358 3369 3386 3387 3456  
## [346] 3457 3465 3469 3475 3506 3520 3530 3548 3564 3567 3593 3618 3627 3630 3670  
## [361] 3675 3689 3715 3716 3726 3732 3736 3749 3761 3770 3771 3785 3817 3824 3841  
## [376] 3846 3862 3868 3869 3870 3882 3897 3903 3918 3930 3941 3957 3965 3966 3969  
## [391] 3972 3980 3982 3999 4004 4005 4011 4015 4023 4030 4044 4046 4049 4073 4077  
## [406] 4083 4101 4102 4105 4111 4115 4119  
##   
## $Fold08  
## [1] 18 27 33 36 43 48 71 79 82 87 102 104 112 123 126  
## [16] 128 137 154 162 175 176 177 184 192 196 198 201 263 275 285  
## [31] 290 292 301 319 329 350 353 359 388 390 402 404 406 414 416  
## [46] 426 429 444 446 450 451 453 464 475 477 485 491 510 525 533  
## [61] 546 558 564 618 620 627 635 637 640 659 660 664 674 676 688  
## [76] 692 704 715 727 748 752 757 772 787 788 794 803 806 807 808  
## [91] 818 833 845 854 857 861 864 867 870 891 893 894 896 905 908  
## [106] 910 919 925 932 949 955 968 975 977 1004 1021 1033 1037 1047 1050  
## [121] 1062 1077 1079 1081 1086 1124 1127 1134 1139 1151 1154 1189 1193 1208 1220  
## [136] 1221 1230 1234 1242 1251 1261 1267 1273 1274 1276 1299 1305 1313 1321 1336  
## [151] 1340 1345 1415 1439 1448 1452 1462 1474 1500 1509 1527 1529 1535 1545 1551  
## [166] 1579 1585 1625 1631 1642 1644 1647 1648 1650 1654 1665 1683 1700 1709 1719  
## [181] 1727 1729 1737 1750 1751 1755 1762 1786 1795 1797 1807 1811 1814 1818 1834  
## [196] 1836 1842 1845 1872 1879 1885 1891 1896 1899 1925 1935 1965 1971 1979 1983  
## [211] 1984 2016 2025 2029 2047 2050 2054 2065 2083 2087 2097 2099 2116 2128 2135  
## [226] 2136 2143 2153 2160 2174 2190 2202 2224 2244 2245 2246 2247 2250 2256 2274  
## [241] 2279 2285 2291 2337 2338 2354 2373 2375 2382 2388 2406 2419 2424 2448 2456  
## [256] 2464 2465 2470 2478 2481 2484 2515 2517 2551 2554 2562 2584 2590 2596 2601  
## [271] 2603 2612 2629 2636 2662 2665 2676 2678 2681 2684 2691 2709 2719 2722 2731  
## [286] 2733 2756 2759 2766 2779 2803 2806 2822 2834 2837 2842 2849 2852 2887 2888  
## [301] 2896 2900 2909 2927 2938 2939 2942 2953 2972 3016 3020 3029 3047 3048 3075  
## [316] 3093 3097 3100 3105 3115 3123 3134 3136 3141 3144 3146 3149 3158 3180 3203  
## [331] 3207 3247 3260 3275 3278 3296 3301 3326 3332 3333 3343 3364 3372 3385 3391  
## [346] 3410 3417 3436 3451 3482 3501 3522 3537 3541 3542 3543 3546 3559 3571 3592  
## [361] 3595 3598 3604 3605 3622 3639 3648 3652 3677 3680 3687 3695 3709 3710 3720  
## [376] 3734 3747 3767 3789 3793 3794 3801 3803 3805 3811 3812 3813 3825 3830 3856  
## [391] 3867 3871 3880 3902 3909 3926 3943 3948 3961 3973 3985 4002 4010 4017 4027  
## [406] 4043 4053 4065 4070 4080 4084  
##   
## $Fold09  
## [1] 35 56 60 66 68 78 83 88 107 118 141 149 158 194 203  
## [16] 219 222 226 229 235 240 271 282 298 314 315 352 360 394 403  
## [31] 420 425 448 455 458 465 471 474 519 527 543 550 555 580 587  
## [46] 592 599 607 609 614 632 633 642 675 680 687 695 702 722 729  
## [61] 742 763 764 768 771 777 813 814 815 822 825 832 837 846 852  
## [76] 855 877 889 899 918 923 930 939 940 954 957 958 971 978 979  
## [91] 992 999 1017 1018 1024 1057 1078 1091 1115 1125 1129 1131 1133 1135 1165  
## [106] 1171 1175 1196 1200 1201 1203 1205 1206 1210 1212 1217 1253 1260 1264 1290  
## [121] 1291 1307 1314 1315 1322 1328 1329 1342 1343 1352 1358 1379 1386 1392 1394  
## [136] 1397 1405 1426 1433 1441 1475 1481 1483 1492 1497 1516 1539 1548 1558 1559  
## [151] 1563 1576 1584 1587 1592 1598 1599 1600 1602 1607 1616 1617 1627 1671 1693  
## [166] 1697 1745 1748 1758 1760 1768 1779 1782 1803 1812 1816 1823 1851 1854 1866  
## [181] 1888 1900 1901 1910 1920 1922 1927 1929 1973 1977 1982 1988 2002 2004 2005  
## [196] 2019 2030 2034 2055 2067 2069 2072 2093 2115 2117 2129 2152 2171 2180 2181  
## [211] 2195 2197 2198 2207 2228 2235 2236 2238 2242 2248 2255 2287 2288 2295 2296  
## [226] 2299 2312 2325 2339 2344 2376 2413 2415 2428 2437 2439 2449 2468 2475 2479  
## [241] 2489 2498 2511 2522 2525 2528 2530 2531 2532 2560 2569 2573 2576 2582 2585  
## [256] 2613 2620 2621 2623 2625 2640 2644 2651 2656 2660 2664 2669 2673 2700 2726  
## [271] 2728 2730 2743 2750 2757 2764 2765 2781 2783 2785 2795 2798 2807 2821 2827  
## [286] 2869 2880 2894 2897 2916 2931 2955 2957 2961 2964 2968 2983 2987 2999 3010  
## [301] 3022 3032 3033 3066 3073 3074 3090 3096 3101 3102 3127 3130 3131 3139 3145  
## [316] 3147 3148 3159 3162 3173 3179 3191 3200 3205 3206 3212 3218 3223 3231 3257  
## [331] 3259 3262 3273 3292 3295 3303 3305 3318 3349 3355 3363 3371 3376 3397 3403  
## [346] 3437 3447 3464 3474 3476 3479 3481 3487 3488 3491 3500 3525 3531 3532 3549  
## [361] 3555 3576 3579 3582 3584 3585 3591 3613 3626 3647 3650 3683 3704 3707 3727  
## [376] 3738 3741 3742 3754 3766 3773 3777 3784 3787 3788 3807 3826 3854 3860 3875  
## [391] 3887 3901 3906 3911 3923 3939 3947 3960 3964 3974 3976 3978 4001 4012 4020  
## [406] 4063 4064 4071 4072 4088 4098 4114  
##   
## $Fold10  
## [1] 2 11 37 40 47 51 54 61 62 72 94 98 114 130 136  
## [16] 140 143 179 188 190 193 206 220 227 230 236 243 246 258 264  
## [31] 265 277 281 288 293 296 297 316 334 345 369 375 376 393 395  
## [46] 405 424 430 431 468 486 489 535 539 551 562 563 566 575 576  
## [61] 601 602 604 613 615 616 638 645 651 673 679 698 705 716 728  
## [76] 732 740 747 759 766 779 790 791 796 801 804 811 844 866 871  
## [91] 879 883 906 927 931 935 943 944 953 956 963 965 988 994 1013  
## [106] 1045 1059 1080 1093 1094 1098 1101 1105 1108 1118 1121 1128 1142 1152 1163  
## [121] 1167 1172 1174 1182 1184 1186 1240 1252 1277 1288 1293 1296 1298 1309 1310  
## [136] 1324 1346 1347 1363 1365 1375 1399 1409 1412 1414 1431 1444 1455 1458 1473  
## [151] 1498 1518 1531 1533 1546 1552 1589 1597 1604 1606 1611 1612 1615 1622 1643  
## [166] 1672 1675 1676 1690 1691 1692 1711 1712 1716 1720 1731 1749 1766 1767 1785  
## [181] 1787 1799 1813 1821 1826 1844 1865 1875 1883 1884 1906 1916 1936 1955 1963  
## [196] 1976 1980 1981 1985 1995 1999 2012 2027 2068 2073 2076 2085 2086 2092 2109  
## [211] 2118 2120 2127 2137 2140 2147 2148 2158 2178 2204 2231 2265 2278 2310 2313  
## [226] 2319 2320 2327 2346 2350 2351 2378 2383 2387 2396 2402 2404 2420 2427 2429  
## [241] 2433 2435 2440 2452 2458 2469 2472 2488 2496 2499 2504 2516 2526 2527 2529  
## [256] 2537 2547 2558 2592 2597 2607 2614 2628 2634 2641 2645 2650 2652 2663 2670  
## [271] 2680 2695 2696 2698 2708 2716 2725 2727 2735 2736 2738 2748 2751 2758 2767  
## [286] 2790 2802 2808 2816 2832 2853 2858 2876 2882 2885 2895 2908 2913 2917 2921  
## [301] 2925 2943 2951 2960 2963 2974 2977 2979 2997 3013 3028 3043 3051 3065 3086  
## [316] 3110 3120 3126 3140 3142 3154 3174 3213 3220 3229 3236 3239 3241 3271 3282  
## [331] 3283 3287 3298 3320 3329 3339 3340 3344 3360 3362 3365 3368 3375 3378 3379  
## [346] 3407 3408 3411 3416 3421 3422 3439 3448 3450 3454 3461 3486 3490 3494 3495  
## [361] 3510 3515 3517 3535 3538 3540 3554 3610 3632 3640 3646 3651 3673 3690 3691  
## [376] 3733 3735 3743 3745 3758 3776 3779 3786 3795 3798 3843 3853 3857 3859 3863  
## [391] 3883 3889 3898 3900 3905 3935 3962 3983 3991 3994 4000 4021 4022 4031 4047  
## [406] 4054 4055 4059 4062 4076 4091 4099

knnFit <- train(type ~ ., method = "knn", data = myData\_scaled,  
 tuneLength = 5, tuneGrid = data.frame(k=1:10),  
 trControl = trainControl(  
 method = "cv", indexOut = trainMyData))  
knnFit

## k-Nearest Neighbors   
##   
## 4119 samples  
## 5 predictor  
## 2 classes: 'no', 'yes'   
##   
## No pre-processing  
## Resampling: Cross-Validated (10 fold)   
## Summary of sample sizes: 3707, 3707, 3708, 3707, 3707, 3707, ...   
## Resampling results across tuning parameters:  
##   
## k Accuracy Kappa   
## 1 0.9844636 0.9209402  
## 2 0.9337295 0.6532539  
## 3 0.9378534 0.6618288  
## 4 0.9291125 0.6080917  
## 5 0.9332346 0.6231811  
## 6 0.9252249 0.5723725  
## 7 0.9278960 0.5864044  
## 8 0.9261993 0.5750890  
## 9 0.9240143 0.5687691  
## 10 0.9228001 0.5612401  
##   
## Accuracy was used to select the optimal model using the largest value.  
## The final value used for the model was k = 1.

# the final value used for k is 1 as it has the highest accuracy:  
# 0.9837331