XGBoost Model

Saptarshi Maiti (A20447671)

12/4/2020

#Load library

```r
library(xgboost)
library(Matrix)
library(caret)
```

```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

#Load Data

```r
filepath <- '/Users/saptarshimaiti/Desktop/Statistical Learning/Project/Project/Data/'
df_train <- read.csv(file = paste0(filepath, "H1.csv"),sep=",",stringsAsFactors = FALSE, na.strings = "N
df_test <- read.csv(file = paste0(filepath, "H2.csv"),sep=",",stringsAsFactors = FALSE, na.strings = "NU
print(paste0('Number of training data set: ', nrow(df_train)))
```

```
## [1] "Number of training data set: 40060"
```

```r
print(paste0('Number of testing data set: ', nrow(df_test)))
```

```
## [1] "Number of testing data set: 79330"
```

#Data Transformation for training

```r
leadtime<-as.numeric(df_train$LeadTime)
country<-as.numeric(factor(df_train$Country))
marketsegment<-as.numeric(factor(df_train$MarketSegment))
deposittype<-as.numeric(factor(df_train$DepositType))
customertype<-as.numeric(factor(df_train$CustomerType))
rcps<-as.numeric(df_train$RequiredCarParkingSpaces)
week<-as.numeric(df_train$ArrivalDateWeekNumber)
IsCanceled<-as.numeric(factor(df_train$IsCanceled))

IsCanceled[IsCanceled == "1"] <- "0"
IsCanceled[IsCanceled == "2"] <- "1"

df_train <- data.frame(leadtime,country,marketsegment,deposittype,customertype,rcps,week,IsCanceled)
```

#Training and Validation Split

```r
trainIndex = createDataPartition(df_train$IsCanceled, p = .8, list = FALSE)
val <- df_train[-trainIndex, -length(df_train)]
IsCanceled_val <- df_train[-trainIndex, length(df_train)]
train <- df_train[trainIndex, -length(df_train)]
IsCanceled_train <- df_train[trainIndex, length(df_train)]
train <- as.matrix(train)
```

```r
val <- as.matrix(val)
train=as(train, "dgCMatrix")
#train
val=as(val, "dgCMatrix")
#val
IsCanceled_train <- as.numeric(factor(IsCanceled_train))
IsCanceled_train[IsCanceled_train == "1"] <- "0"
IsCanceled_train[IsCanceled_train == "2"] <- "1"

IsCanceled_val <- as.numeric(factor(IsCanceled_val))
IsCanceled_val[IsCanceled_val == "1"] <- "0"
IsCanceled_val[IsCanceled_val == "2"] <- "1"
```

Model Fitting XGBoost

```r
bst <- xgboost(data = train, label = IsCanceled_train, max.depth = 2, eta = 1, nthread = 2, nrounds = 60
```

```
## [1]   train-error:0.237043
## [2]   train-error:0.197760
## [3]   train-error:0.189397
## [4]   train-error:0.185404
## [5]   train-error:0.178820
## [6]   train-error:0.181441
## [7]   train-error:0.175669
## [8]   train-error:0.178227
## [9]   train-error:0.173765
## [10]  train-error:0.173765
## [11]  train-error:0.172642
## [12]  train-error:0.172392
## [13]  train-error:0.173048
## [14]  train-error:0.171831
## [15]  train-error:0.173453
## [16]  train-error:0.170614
## [17]  train-error:0.169210
## [18]  train-error:0.166901
## [19]  train-error:0.164561
## [20]  train-error:0.165372
## [21]  train-error:0.165153
## [22]  train-error:0.164561
## [23]  train-error:0.164873
## [24]  train-error:0.165091
## [25]  train-error:0.165372
## [26]  train-error:0.165185
## [27]  train-error:0.165434
## [28]  train-error:0.164498
## [29]  train-error:0.164311
## [30]  train-error:0.164186
## [31]  train-error:0.163718
## [32]  train-error:0.162595
## [33]  train-error:0.162907
## [34]  train-error:0.162813
## [35]  train-error:0.161752
```

```
## [36] train-error:0.161315
## [37] train-error:0.161472
## [38] train-error:0.161784
## [39] train-error:0.161908
## [40] train-error:0.161503
## [41] train-error:0.161253
## [42] train-error:0.160910
## [43] train-error:0.161659
## [44] train-error:0.162408
## [45] train-error:0.162064
## [46] train-error:0.161097
## [47] train-error:0.160036
## [48] train-error:0.159880
## [49] train-error:0.158882
## [50] train-error:0.159100
## [51] train-error:0.158570
## [52] train-error:0.159225
## [53] train-error:0.158726
## [54] train-error:0.158476
## [55] train-error:0.157134
## [56] train-error:0.157197
## [57] train-error:0.156167
## [58] train-error:0.155262
## [59] train-error:0.154825
## [60] train-error:0.155169
## [61] train-error:0.155137
## [62] train-error:0.155075
## [63] train-error:0.154607
## [64] train-error:0.154170
## [65] train-error:0.153453
## [66] train-error:0.153952
## [67] train-error:0.153889
## [68] train-error:0.153921
## [69] train-error:0.153609
## [70] train-error:0.153796
## [71] train-error:0.153827
## [72] train-error:0.153952
## [73] train-error:0.153640
## [74] train-error:0.153515
## [75] train-error:0.153390
## [76] train-error:0.153671
## [77] train-error:0.153234
## [78] train-error:0.152579
## [79] train-error:0.153203
## [80] train-error:0.152579
## [81] train-error:0.151768
## [82] train-error:0.152360
## [83] train-error:0.152173
## [84] train-error:0.152048
## [85] train-error:0.152828
## [86] train-error:0.152142
## [87] train-error:0.152610
## [88] train-error:0.151736
## [89] train-error:0.151955
```

```
## [90] train-error:0.151768
## [91] train-error:0.152485
## [92] train-error:0.152080
## [93] train-error:0.151768
## [94] train-error:0.151393
## [95] train-error:0.151487
## [96] train-error:0.151019
## [97] train-error:0.150894
## [98] train-error:0.150832
## [99] train-error:0.150894
## [100]     train-error:0.150426
## [101]     train-error:0.149989
## [102]     train-error:0.149989
## [103]     train-error:0.150114
## [104]     train-error:0.150239
## [105]     train-error:0.150676
## [106]     train-error:0.150800
## [107]     train-error:0.150800
## [108]     train-error:0.149147
## [109]     train-error:0.150395
## [110]     train-error:0.150083
## [111]     train-error:0.149552
## [112]     train-error:0.150332
## [113]     train-error:0.148959
## [114]     train-error:0.148772
## [115]     train-error:0.148741
## [116]     train-error:0.148554
## [117]     train-error:0.148460
## [118]     train-error:0.148491
## [119]     train-error:0.148367
## [120]     train-error:0.148585
## [121]     train-error:0.147930
## [122]     train-error:0.147618
## [123]     train-error:0.147992
## [124]     train-error:0.148211
## [125]     train-error:0.148273
## [126]     train-error:0.148460
## [127]     train-error:0.148273
## [128]     train-error:0.147805
## [129]     train-error:0.148055
## [130]     train-error:0.148398
## [131]     train-error:0.148055
## [132]     train-error:0.148772
## [133]     train-error:0.147680
## [134]     train-error:0.147992
## [135]     train-error:0.148086
## [136]     train-error:0.148429
## [137]     train-error:0.147867
## [138]     train-error:0.148055
## [139]     train-error:0.148429
## [140]     train-error:0.148554
## [141]     train-error:0.148179
## [142]     train-error:0.148086
## [143]     train-error:0.147774
```

```
## [144]    train-error:0.147743
## [145]    train-error:0.147680
## [146]    train-error:0.147462
## [147]    train-error:0.146588
## [148]    train-error:0.147181
## [149]    train-error:0.146931
## [150]    train-error:0.146900
## [151]    train-error:0.147025
## [152]    train-error:0.147025
## [153]    train-error:0.146869
## [154]    train-error:0.146307
## [155]    train-error:0.146151
## [156]    train-error:0.146494
## [157]    train-error:0.146089
## [158]    train-error:0.145839
## [159]    train-error:0.145402
## [160]    train-error:0.145496
## [161]    train-error:0.145777
## [162]    train-error:0.145558
## [163]    train-error:0.145558
## [164]    train-error:0.145496
## [165]    train-error:0.145371
## [166]    train-error:0.145402
## [167]    train-error:0.145028
## [168]    train-error:0.144903
## [169]    train-error:0.145059
## [170]    train-error:0.144872
## [171]    train-error:0.144841
## [172]    train-error:0.145434
## [173]    train-error:0.145278
## [174]    train-error:0.145184
## [175]    train-error:0.145122
## [176]    train-error:0.144716
## [177]    train-error:0.144497
## [178]    train-error:0.144997
## [179]    train-error:0.144435
## [180]    train-error:0.145028
## [181]    train-error:0.145122
## [182]    train-error:0.144903
## [183]    train-error:0.144903
## [184]    train-error:0.144747
## [185]    train-error:0.144685
## [186]    train-error:0.144591
## [187]    train-error:0.144373
## [188]    train-error:0.144123
## [189]    train-error:0.143873
## [190]    train-error:0.144279
## [191]    train-error:0.144092
## [192]    train-error:0.144435
## [193]    train-error:0.144341
## [194]    train-error:0.144248
## [195]    train-error:0.144217
## [196]    train-error:0.144154
## [197]    train-error:0.144061
```

```
## [198]    train-error:0.144061
## [199]    train-error:0.144279
## [200]    train-error:0.144217
## [201]    train-error:0.144279
## [202]    train-error:0.144123
## [203]    train-error:0.143842
## [204]    train-error:0.143624
## [205]    train-error:0.143842
## [206]    train-error:0.143593
## [207]    train-error:0.143561
## [208]    train-error:0.143624
## [209]    train-error:0.143717
## [210]    train-error:0.143405
## [211]    train-error:0.143780
## [212]    train-error:0.143873
## [213]    train-error:0.143655
## [214]    train-error:0.143468
## [215]    train-error:0.143218
## [216]    train-error:0.143312
## [217]    train-error:0.143530
## [218]    train-error:0.143468
## [219]    train-error:0.143499
## [220]    train-error:0.143437
## [221]    train-error:0.143218
## [222]    train-error:0.143031
## [223]    train-error:0.143062
## [224]    train-error:0.143062
## [225]    train-error:0.143187
## [226]    train-error:0.143062
## [227]    train-error:0.143249
## [228]    train-error:0.143437
## [229]    train-error:0.143312
## [230]    train-error:0.142906
## [231]    train-error:0.143125
## [232]    train-error:0.143093
## [233]    train-error:0.143156
## [234]    train-error:0.142750
## [235]    train-error:0.142750
## [236]    train-error:0.142688
## [237]    train-error:0.143374
## [238]    train-error:0.143156
## [239]    train-error:0.142345
## [240]    train-error:0.142750
## [241]    train-error:0.142750
## [242]    train-error:0.142657
## [243]    train-error:0.143312
## [244]    train-error:0.142875
## [245]    train-error:0.142750
## [246]    train-error:0.143062
## [247]    train-error:0.143405
## [248]    train-error:0.143187
## [249]    train-error:0.143312
## [250]    train-error:0.142657
## [251]    train-error:0.142937
```

```
## [252]    train-error:0.141689
## [253]    train-error:0.142033
## [254]    train-error:0.142064
## [255]    train-error:0.141783
## [256]    train-error:0.141658
## [257]    train-error:0.142438
## [258]    train-error:0.142438
## [259]    train-error:0.141533
## [260]    train-error:0.141564
## [261]    train-error:0.141440
## [262]    train-error:0.141408
## [263]    train-error:0.141720
## [264]    train-error:0.142095
## [265]    train-error:0.141627
## [266]    train-error:0.141970
## [267]    train-error:0.141939
## [268]    train-error:0.141502
## [269]    train-error:0.141502
## [270]    train-error:0.140878
## [271]    train-error:0.140847
## [272]    train-error:0.140940
## [273]    train-error:0.141221
## [274]    train-error:0.140940
## [275]    train-error:0.141471
## [276]    train-error:0.141252
## [277]    train-error:0.141315
## [278]    train-error:0.141346
## [279]    train-error:0.141970
## [280]    train-error:0.141658
## [281]    train-error:0.141471
## [282]    train-error:0.141096
## [283]    train-error:0.141252
## [284]    train-error:0.141658
## [285]    train-error:0.141128
## [286]    train-error:0.141003
## [287]    train-error:0.141128
## [288]    train-error:0.141221
## [289]    train-error:0.141252
## [290]    train-error:0.141252
## [291]    train-error:0.141346
## [292]    train-error:0.141221
## [293]    train-error:0.141128
## [294]    train-error:0.141221
## [295]    train-error:0.141128
## [296]    train-error:0.141034
## [297]    train-error:0.140909
## [298]    train-error:0.140940
## [299]    train-error:0.140784
## [300]    train-error:0.140816
## [301]    train-error:0.140722
## [302]    train-error:0.140691
## [303]    train-error:0.140847
## [304]    train-error:0.140410
## [305]    train-error:0.139973
```

```
## [306]    train-error:0.140348
## [307]    train-error:0.140160
## [308]    train-error:0.140160
## [309]    train-error:0.140285
## [310]    train-error:0.140192
## [311]    train-error:0.140160
## [312]    train-error:0.140160
## [313]    train-error:0.140566
## [314]    train-error:0.141003
## [315]    train-error:0.141284
## [316]    train-error:0.140878
## [317]    train-error:0.140660
## [318]    train-error:0.140410
## [319]    train-error:0.140691
## [320]    train-error:0.140691
## [321]    train-error:0.140628
## [322]    train-error:0.140285
## [323]    train-error:0.139942
## [324]    train-error:0.140036
## [325]    train-error:0.140036
## [326]    train-error:0.140036
## [327]    train-error:0.140192
## [328]    train-error:0.140129
## [329]    train-error:0.139786
## [330]    train-error:0.139724
## [331]    train-error:0.139817
## [332]    train-error:0.140098
## [333]    train-error:0.140067
## [334]    train-error:0.139755
## [335]    train-error:0.139786
## [336]    train-error:0.139692
## [337]    train-error:0.139817
## [338]    train-error:0.140285
## [339]    train-error:0.139630
## [340]    train-error:0.139661
## [341]    train-error:0.139692
## [342]    train-error:0.139817
## [343]    train-error:0.139692
## [344]    train-error:0.139786
## [345]    train-error:0.139880
## [346]    train-error:0.139880
## [347]    train-error:0.139724
## [348]    train-error:0.139661
## [349]    train-error:0.139661
## [350]    train-error:0.139661
## [351]    train-error:0.139724
## [352]    train-error:0.139474
## [353]    train-error:0.139318
## [354]    train-error:0.139412
## [355]    train-error:0.139630
## [356]    train-error:0.139568
## [357]    train-error:0.139318
## [358]    train-error:0.139817
## [359]    train-error:0.139474
```

```
## [360]    train-error:0.139630
## [361]    train-error:0.139412
## [362]    train-error:0.139256
## [363]    train-error:0.138850
## [364]    train-error:0.138694
## [365]    train-error:0.138507
## [366]    train-error:0.139037
## [367]    train-error:0.139037
## [368]    train-error:0.139256
## [369]    train-error:0.139162
## [370]    train-error:0.138819
## [371]    train-error:0.138538
## [372]    train-error:0.138413
## [373]    train-error:0.138943
## [374]    train-error:0.138694
## [375]    train-error:0.139162
## [376]    train-error:0.139037
## [377]    train-error:0.139037
## [378]    train-error:0.139162
## [379]    train-error:0.139100
## [380]    train-error:0.138819
## [381]    train-error:0.139037
## [382]    train-error:0.138663
## [383]    train-error:0.138631
## [384]    train-error:0.139037
## [385]    train-error:0.138881
## [386]    train-error:0.138694
## [387]    train-error:0.138444
## [388]    train-error:0.138413
## [389]    train-error:0.138226
## [390]    train-error:0.138101
## [391]    train-error:0.138007
## [392]    train-error:0.137976
## [393]    train-error:0.137883
## [394]    train-error:0.138351
## [395]    train-error:0.138382
## [396]    train-error:0.138070
## [397]    train-error:0.137883
## [398]    train-error:0.137758
## [399]    train-error:0.137914
## [400]    train-error:0.138070
## [401]    train-error:0.138257
## [402]    train-error:0.138163
## [403]    train-error:0.137727
## [404]    train-error:0.138163
## [405]    train-error:0.137851
## [406]    train-error:0.137945
## [407]    train-error:0.137945
## [408]    train-error:0.138007
## [409]    train-error:0.137945
## [410]    train-error:0.137571
## [411]    train-error:0.137477
## [412]    train-error:0.137321
## [413]    train-error:0.137477
```

```
## [414]    train-error:0.137633
## [415]    train-error:0.137571
## [416]    train-error:0.137571
## [417]    train-error:0.137383
## [418]    train-error:0.137415
## [419]    train-error:0.137446
## [420]    train-error:0.137820
## [421]    train-error:0.137945
## [422]    train-error:0.137851
## [423]    train-error:0.137883
## [424]    train-error:0.137789
## [425]    train-error:0.137820
## [426]    train-error:0.137789
## [427]    train-error:0.137789
## [428]    train-error:0.137758
## [429]    train-error:0.137695
## [430]    train-error:0.137727
## [431]    train-error:0.137820
## [432]    train-error:0.137321
## [433]    train-error:0.137571
## [434]    train-error:0.137664
## [435]    train-error:0.136978
## [436]    train-error:0.136978
## [437]    train-error:0.136915
## [438]    train-error:0.137290
## [439]    train-error:0.137415
## [440]    train-error:0.137383
## [441]    train-error:0.137695
## [442]    train-error:0.137851
## [443]    train-error:0.137914
## [444]    train-error:0.137945
## [445]    train-error:0.137945
## [446]    train-error:0.137945
## [447]    train-error:0.137758
## [448]    train-error:0.137539
## [449]    train-error:0.138195
## [450]    train-error:0.138132
## [451]    train-error:0.138101
## [452]    train-error:0.138195
## [453]    train-error:0.138226
## [454]    train-error:0.138007
## [455]    train-error:0.138101
## [456]    train-error:0.138351
## [457]    train-error:0.138132
## [458]    train-error:0.138195
## [459]    train-error:0.137883
## [460]    train-error:0.137820
## [461]    train-error:0.137976
## [462]    train-error:0.138101
## [463]    train-error:0.138226
## [464]    train-error:0.138070
## [465]    train-error:0.138070
## [466]    train-error:0.138070
## [467]    train-error:0.137945
```

```
## [468]    train-error:0.137727
## [469]    train-error:0.137914
## [470]    train-error:0.137914
## [471]    train-error:0.138007
## [472]    train-error:0.138101
## [473]    train-error:0.137976
## [474]    train-error:0.138101
## [475]    train-error:0.138132
## [476]    train-error:0.138101
## [477]    train-error:0.138163
## [478]    train-error:0.138319
## [479]    train-error:0.138070
## [480]    train-error:0.138132
## [481]    train-error:0.137820
## [482]    train-error:0.137851
## [483]    train-error:0.137914
## [484]    train-error:0.137976
## [485]    train-error:0.137664
## [486]    train-error:0.137727
## [487]    train-error:0.137383
## [488]    train-error:0.137446
## [489]    train-error:0.137789
## [490]    train-error:0.137571
## [491]    train-error:0.137571
## [492]    train-error:0.137789
## [493]    train-error:0.137820
## [494]    train-error:0.138288
## [495]    train-error:0.138163
## [496]    train-error:0.138070
## [497]    train-error:0.138226
## [498]    train-error:0.138195
## [499]    train-error:0.138319
## [500]    train-error:0.137976
## [501]    train-error:0.138007
## [502]    train-error:0.138101
## [503]    train-error:0.138132
## [504]    train-error:0.138070
## [505]    train-error:0.137945
## [506]    train-error:0.137571
## [507]    train-error:0.137383
## [508]    train-error:0.137321
## [509]    train-error:0.137352
## [510]    train-error:0.137290
## [511]    train-error:0.137383
## [512]    train-error:0.137290
## [513]    train-error:0.137290
## [514]    train-error:0.137477
## [515]    train-error:0.137477
## [516]    train-error:0.137259
## [517]    train-error:0.137290
## [518]    train-error:0.137259
## [519]    train-error:0.137321
## [520]    train-error:0.137477
## [521]    train-error:0.137446
```

```
## [522]	train-error:0.137009
## [523]	train-error:0.137415
## [524]	train-error:0.137103
## [525]	train-error:0.137134
## [526]	train-error:0.137165
## [527]	train-error:0.136791
## [528]	train-error:0.136947
## [529]	train-error:0.136822
## [530]	train-error:0.136791
## [531]	train-error:0.136947
## [532]	train-error:0.136947
## [533]	train-error:0.137009
## [534]	train-error:0.136978
## [535]	train-error:0.137040
## [536]	train-error:0.137134
## [537]	train-error:0.136915
## [538]	train-error:0.137477
## [539]	train-error:0.137227
## [540]	train-error:0.137415
## [541]	train-error:0.137415
## [542]	train-error:0.137383
## [543]	train-error:0.137290
## [544]	train-error:0.137352
## [545]	train-error:0.137352
## [546]	train-error:0.137383
## [547]	train-error:0.137383
## [548]	train-error:0.137352
## [549]	train-error:0.137071
## [550]	train-error:0.136947
## [551]	train-error:0.136603
## [552]	train-error:0.136728
## [553]	train-error:0.136541
## [554]	train-error:0.136947
## [555]	train-error:0.137071
## [556]	train-error:0.137134
## [557]	train-error:0.137103
## [558]	train-error:0.136759
## [559]	train-error:0.136759
## [560]	train-error:0.136728
## [561]	train-error:0.136541
## [562]	train-error:0.136541
## [563]	train-error:0.136385
## [564]	train-error:0.136697
## [565]	train-error:0.136447
## [566]	train-error:0.136572
## [567]	train-error:0.136385
## [568]	train-error:0.136635
## [569]	train-error:0.136416
## [570]	train-error:0.136510
## [571]	train-error:0.136385
## [572]	train-error:0.136572
## [573]	train-error:0.136479
## [574]	train-error:0.136479
## [575]	train-error:0.136541
```

```
## [576]     train-error:0.136447
## [577]     train-error:0.136635
## [578]     train-error:0.136572
## [579]     train-error:0.136479
## [580]     train-error:0.136572
## [581]     train-error:0.136635
## [582]     train-error:0.136447
## [583]     train-error:0.136385
## [584]     train-error:0.136198
## [585]     train-error:0.136104
## [586]     train-error:0.136291
## [587]     train-error:0.136354
## [588]     train-error:0.136510
## [589]     train-error:0.136510
## [590]     train-error:0.136479
## [591]     train-error:0.136135
## [592]     train-error:0.136135
## [593]     train-error:0.136323
## [594]     train-error:0.136291
## [595]     train-error:0.136229
## [596]     train-error:0.136042
## [597]     train-error:0.136260
## [598]     train-error:0.135917
## [599]     train-error:0.136260
## [600]     train-error:0.136260
```

```r
summary(bst)
```

```
##                  Length Class              Mode
## handle                1 xgb.Booster.handle externalptr
## raw              244154 -none-             raw
## niter                 1 -none-             numeric
## evaluation_log        2 data.table         list
## call                 17 -none-             call
## params                5 -none-             list
## callbacks             2 -none-             list
## feature_names         7 -none-             character
## nfeatures             1 -none-             numeric
```

Prediction Validation

```r
pred_val <- predict(bst, val)
prediction_val <- as.numeric(pred_val > 0.5)
```

Validation Missclassification Rate

```r
val_missclassification <-sum(prediction_val!=IsCanceled_val)/length(prediction_val)
val_missclassification
```

```
## [1] 0.1465485
```

Validation Confusion Matrix

```
confusionMatrix(as.factor(prediction_val), as.factor(IsCanceled_val), positive="1")
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##          0 5332  719
##          1  455 1505
##
##                Accuracy : 0.8535
##                  95% CI : (0.8455, 0.8611)
##     No Information Rate : 0.7224
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.6208
##
##  Mcnemar's Test P-Value : 1.644e-14
##
##             Sensitivity : 0.6767
##             Specificity : 0.9214
##          Pos Pred Value : 0.7679
##          Neg Pred Value : 0.8812
##              Prevalence : 0.2776
##          Detection Rate : 0.1879
##    Detection Prevalence : 0.2447
##       Balanced Accuracy : 0.7990
##
##        'Positive' Class : 1
##
```
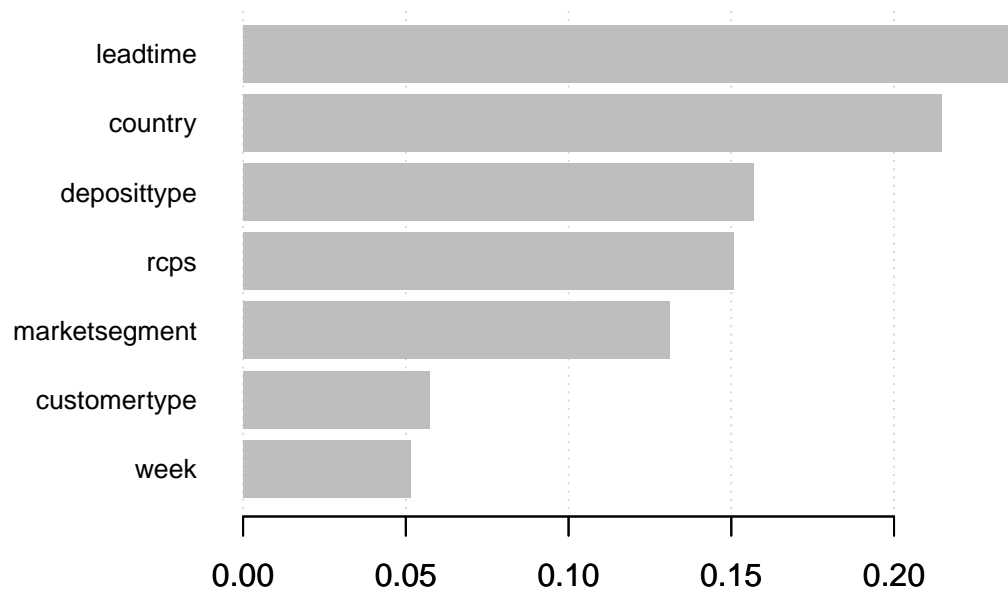
Importance Matrix

```
importance_matrix <- xgb.importance(model = bst)
print(importance_matrix)
```

```
##          Feature       Gain       Cover  Frequency
## 1:      leadtime 0.23830459 0.475959045 0.435050392
## 2:       country 0.21449303 0.194075491 0.187569989
## 3:   deposittype 0.15682525 0.017277025 0.013437850
## 4:          rcps 0.15056102 0.008122998 0.005039194
## 5: marketsegment 0.13099125 0.060116688 0.066629339
## 6:  customertype 0.05742093 0.039261305 0.045352744
## 7:          week 0.05140392 0.205187447 0.246920493
```

```
xgb.plot.importance(importance_matrix = importance_matrix)
```

Testing

Data Transformation

```
leadtime<-as.numeric(df_test$LeadTime)
country<-as.numeric(factor(df_test$Country))
marketsegment<-as.numeric(factor(df_test$MarketSegment))
deposittype<-as.numeric(factor(df_test$DepositType))
customertype<-as.numeric(factor(df_test$CustomerType))
rcps<-as.numeric(df_test$RequiredCarParkingSpaces)
week<-as.numeric(df_test$ArrivalDateWeekNumber)
IsCanceled_test<-as.numeric(factor(df_test$IsCanceled))

IsCanceled_test[IsCanceled_test == "1"] <- "0"
IsCanceled_test[IsCanceled_test == "2"] <- "1"

df_test <- data.frame(leadtime,country,marketsegment,deposittype,customertype,rcps,week)

df_test_mtx <- as.matrix(df_test)
test <- as(df_test_mtx, "dgCMatrix")
```

#Prediction for test dataset

```
pred <- predict(bst, test)
prediction <- as.numeric(pred > 0.5)
```

#Missclassification Rate for test dataset

```
pred_missclassification <-sum(prediction!=IsCanceled_test)/length(prediction)
pred_missclassification
```

```
## [1] 0.3245304
```

Test Confusion Matrix

```r
confusionMatrix(as.factor(prediction), as.factor(IsCanceled_test), positive="1")
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction     0     1
##          0 35458 14975
##          1 10770 18127
##
##                Accuracy : 0.6755
##                  95% CI : (0.6722, 0.6787)
##     No Information Rate : 0.5827
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.3204
##
##  Mcnemar's Test P-Value : < 2.2e-16
##
##             Sensitivity : 0.5476
##             Specificity : 0.7670
##          Pos Pred Value : 0.6273
##          Neg Pred Value : 0.7031
##              Prevalence : 0.4173
##          Detection Rate : 0.2285
##    Detection Prevalence : 0.3643
##       Balanced Accuracy : 0.6573
##
##        'Positive' Class : 1
##
```