week11

Jiayin Qu

4/1/2020

Library Imports

```
knitr::opts_chunk$set(echo = TRUE)

library(tidyverse)
library(stringr)
library(haven)
library(glmnet)
library(caret)
library(tictoc)
library(mice)
library(MLmetrics)
library(xgboost)
library(lattice)
```

Data Import

- import dataset
- use zap_empty to clean up blank cells (show them as NAs)
- isolate personality and health variables
- get rid of labels from labelled vectors
- turn variables into numeric
- combine personality and health data for the final dataset
- delete obs that does not have DV

```
gss <- read_sav("../data/GSS2006.sav") %>%
  mutate_all(zap_label) %>%
  mutate_all(zap_labels) %>%
  mutate_all(zap_formats) %>%
  mutate_all(zap_widths) %>%
  as_tibble(lapply(., as.character))
gss_per <- gss %>%
  select(starts_with("BIG5")) %>%
  as_tibble(lapply(., as.numeric))
gss_health <- gss %>%
  select(ends_with("HEALTH")) %>%
  lapply(., as.numeric)
final_gss <- cbind(gss_per, gss_health)
final_gss <- final_gss[!is.na(final_gss$HEALTH),]
```

Analysis

- run OLS model

```r
lm_model <- train(
  HEALTH ~ .^2,
  final_gss,
  method = "lm",
  preProcess = c("center", "scale", "zv", "medianImpute"),
  trControl=trainControl(method = "cv", number = 10, verboseIter = T),
  na.action = na.pass
)
```

```r
summary(lm_model)
```

```
##
## Call:
## lm(formula = .outcome ~ ., data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.66736 -0.79616 -0.03874  0.51031  2.39264
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2.026886   0.021261  95.335   <2e-16 ***
## BIG5A2           -0.047320   0.213133  -0.222   0.8243
## BIG5D2            0.424344   0.177119   2.396   0.0166 *
## BIG5B2           -0.371796   0.166539  -2.232   0.0256 *
## BIG5C2            0.367555   0.167765   2.191   0.0285 *
## BIG5D1            0.414713   0.192012   2.160   0.0309 *
## BIG5B1            0.386642   0.196042   1.972   0.0487 *
## BIG5C1            0.195280   0.133241   1.466   0.1428
## BIG5E2           -0.228552   0.143854  -1.589   0.1122
## BIG5E1           -0.053518   0.198589  -0.269   0.7876
## BIG5A1            0.011602   0.185951   0.062   0.9503
## `BIG5A2:BIG5D2` -0.029674   0.086844  -0.342   0.7326
## `BIG5A2:BIG5B2`  0.252848   0.099258   2.547   0.0109 *
## `BIG5A2:BIG5C2` -0.138423   0.126060  -1.098   0.2723
## `BIG5A2:BIG5D1` -0.102432   0.092356  -1.109   0.2675
## `BIG5A2:BIG5B1`  0.155744   0.086622   1.798   0.0723 .
## `BIG5A2:BIG5C1` -0.122111   0.094258  -1.296   0.1952
## `BIG5A2:BIG5E2`  0.002030   0.083681   0.024   0.9807
## `BIG5A2:BIG5E1`  0.001507   0.080562   0.019   0.9851
## `BIG5A2:BIG5A1`  0.036107   0.062283   0.580   0.5621
## `BIG5D2:BIG5B2`  0.126864   0.109067   1.163   0.2448
## `BIG5D2:BIG5C2` -0.308370   0.134877  -2.286   0.0223 *
## `BIG5D2:BIG5D1` -0.073239   0.067891  -1.079   0.2808
## `BIG5D2:BIG5B1` -0.149429   0.095037  -1.572   0.1160
## `BIG5D2:BIG5C1` -0.111813   0.093291  -1.199   0.2308
## `BIG5D2:BIG5E2` -0.074657   0.098129  -0.761   0.4468
## `BIG5D2:BIG5E1` -0.071406   0.088047  -0.811   0.4174
## `BIG5D2:BIG5A1` -0.002397   0.102764  -0.023   0.9814
## `BIG5B2:BIG5C2`  0.091332   0.127173   0.718   0.4727
## `BIG5B2:BIG5D1`  0.123507   0.087612   1.410   0.1587
## `BIG5B2:BIG5B1` -0.107777   0.079328  -1.359   0.1744
## `BIG5B2:BIG5C1`  0.058375   0.092373   0.632   0.5275
## `BIG5B2:BIG5E2`  0.121163   0.093174   1.300   0.1936
```

```
## `BIG5B2:BIG5E1`  0.089814   0.093477   0.961   0.3367
## `BIG5B2:BIG5A1` -0.071251   0.091408  -0.779   0.4357
## `BIG5C2:BIG5D1` -0.241272   0.116488  -2.071   0.0384 *
## `BIG5C2:BIG5B1` -0.066309   0.117628  -0.564   0.5730
## `BIG5C2:BIG5C1`  0.008763   0.096796   0.091   0.9279
## `BIG5C2:BIG5E2`  0.052759   0.117889   0.448   0.6545
## `BIG5C2:BIG5E1`  0.012169   0.115319   0.106   0.9160
## `BIG5C2:BIG5A1` -0.162902   0.116628  -1.397   0.1626
## `BIG5D1:BIG5B1` -0.193413   0.082391  -2.347   0.0190 *
## `BIG5D1:BIG5C1`  0.041746   0.099144   0.421   0.6737
## `BIG5D1:BIG5E2`  0.117734   0.084162   1.399   0.1619
## `BIG5D1:BIG5E1` -0.092908   0.080294  -1.157   0.2473
## `BIG5D1:BIG5A1` -0.017901   0.079527  -0.225   0.8219
## `BIG5B1:BIG5C1` -0.014718   0.086849  -0.169   0.8654
## `BIG5B1:BIG5E2` -0.135218   0.083219  -1.625   0.1043
## `BIG5B1:BIG5E1`  0.009690   0.076618   0.126   0.8994
## `BIG5B1:BIG5A1`  0.018780   0.075772   0.248   0.8043
## `BIG5C1:BIG5E2`  0.005952   0.084711   0.070   0.9440
## `BIG5C1:BIG5E1` -0.086103   0.091111  -0.945   0.3447
## `BIG5C1:BIG5A1` -0.001080   0.086158  -0.013   0.9900
## `BIG5E2:BIG5E1`  0.129881   0.070133   1.852   0.0641 .
## `BIG5E2:BIG5A1`  0.083143   0.081892   1.015   0.3100
## `BIG5E1:BIG5A1`  0.055571   0.073081   0.760   0.4471
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8276 on 3460 degrees of freedom
## Multiple R-squared:  0.03506,    Adjusted R-squared:  0.01973
## F-statistic: 2.286 on 55 and 3460 DF,  p-value: 2.694e-07
```

- dummy code DV and create folds

```
# turn Health into factor
final_per <- final_gss[,1:10]
final_health <- factor(final_gss$HEALTH, levels = c(1, 2, 3, 4), labels = c("Excellent", "Good", "Fair"
final_gss <- cbind(final_per, final_health)

# set 10 folds
gssControl <- trainControl(
  method = "repeatedcv",
  number = 10,
  summaryFunction = multiClassSummary,
  classProbs = TRUE,
  verboseIter = TRUE,
  savePredictions = TRUE
)
```

- run 10-fold elastic net regression

```
model_glmnet <- train(
  x = final_per,
  y = final_health,
  metric = "ROC",
  method = "glmnet",
  trControl = gssControl,
  tuneGrid = expand.grid(
```

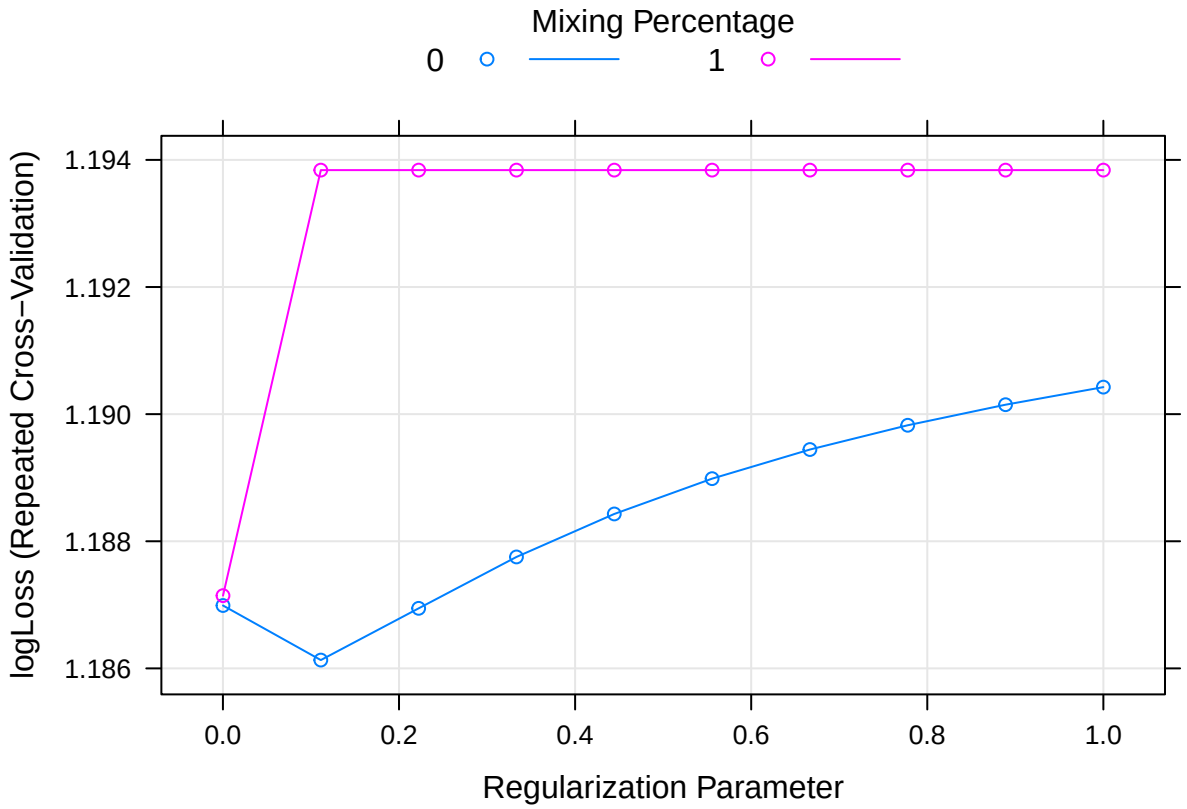
```
    alpha = 0:1,
    lambda = seq(0.0001, 1, length = 10)),
  preProcess = c("medianImpute","zv","center", "scale", "pca"),
  na.action = na.pass,
  tuneLength = 10
)
```

```
model_glmnet
```

```
## glmnet
##
## 3516 samples
##   10 predictor
##    4 classes: 'Excellent', 'Good', 'Fair', 'Poor'
##
## Pre-processing: median imputation (10), centered (10), scaled
##  (10), principal component signal extraction (10)
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 3165, 3165, 3164, 3164, 3164, 3165, ...
## Resampling results across tuning parameters:
##
##   alpha  lambda  logLoss   AUC        prAUC      Accuracy   Kappa
##   0      0.0001  1.186990  0.5377741  0.2880897  0.4707127  0.021463180
##   0      0.1112  1.186131  0.5394073  0.2881784  0.4695698  0.012061941
##   0      0.2223  1.186945  0.5389051  0.2876172  0.4687159  0.006331791
##   0      0.3334  1.187753  0.5380218  0.2872855  0.4670089  0.000928998
##   0      0.4445  1.188430  0.5381388  0.2878190  0.4667248  0.000000000
##   0      0.5556  1.188985  0.5380471  0.2877870  0.4667248  0.000000000
##   0      0.6667  1.189443  0.5381164  0.2883613  0.4667248  0.000000000
##   0      0.7778  1.189825  0.5388997  0.2885927  0.4667248  0.000000000
##   0      0.8889  1.190148  0.5390717  0.2883215  0.4667248  0.000000000
##   0      1.0000  1.190425  0.5392496  0.2884205  0.4667248  0.000000000
##   1      0.0001  1.187145  0.5390343  0.2881575  0.4707127  0.021856252
##   1      0.1112  1.193839  0.5000000  0.0000000  0.4667248  0.000000000
##   1      0.2223  1.193839  0.5000000  0.0000000  0.4667248  0.000000000
##   1      0.3334  1.193839  0.5000000  0.0000000  0.4667248  0.000000000
##   1      0.4445  1.193839  0.5000000  0.0000000  0.4667248  0.000000000
##   1      0.5556  1.193839  0.5000000  0.0000000  0.4667248  0.000000000
##   1      0.6667  1.193839  0.5000000  0.0000000  0.4667248  0.000000000
##   1      0.7778  1.193839  0.5000000  0.0000000  0.4667248  0.000000000
##   1      0.8889  1.193839  0.5000000  0.0000000  0.4667248  0.000000000
##   1      1.0000  1.193839  0.5000000  0.0000000  0.4667248  0.000000000
##   Mean_F1  Mean_Sensitivity  Mean_Specificity  Mean_Pos_Pred_Value
##   NaN      0.2592886         0.7544847         NaN
##   NaN      0.2549485         0.7525544         NaN
##   NaN      0.2524148         0.7513861         NaN
##   NaN      0.2503578         0.7502021         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2594992         0.7545545         NaN
```

```
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   NaN      0.2500000         0.7500000         NaN
##   Mean_Neg_Pred_Value  Mean_Precision  Mean_Recall  Mean_Detection_Rate
##   0.7824402            NaN             0.2592886    0.1176782
##   0.8001029            NaN             0.2549485    0.1173925
##   0.8324419            NaN             0.2524148    0.1171790
##   0.8042938            NaN             0.2503578    0.1167522
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##   0.7817933            NaN             0.2594992    0.1176782
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##         NaN            NaN             0.2500000    0.1166812
##   Mean_Balanced_Accuracy
##   0.5068866
##   0.5037514
##   0.5019005
##   0.5002799
##   0.5000000
##   0.5000000
##   0.5000000
##   0.5000000
##   0.5000000
##   0.5000000
##   0.5070268
##   0.5000000
##   0.5000000
##   0.5000000
##   0.5000000
##   0.5000000
##   0.5000000
##   0.5000000
##   0.5000000
##   0.5000000
##
## logLoss was used to select the optimal model using the smallest value.
## The final values used for the model were alpha = 0 and lambda = 0.1112.
```

```
plot(model_glmnet)
```



The tuning parameters worked best for the elastic net model were alpha = 0 and lambda = 0.1112. Alpha = 0, meaning that the "best" model is a full ridge model.

- run SVM model

```
model_SVM <- train(
  x = final_per,
  y = final_health,
  metric = "ROC",
  method = "svmLinear",
  trControl = gssControl,
  preProcess = c("medianImpute", "zv","center", "scale", "pca")
)
```

```
model_SVM
```

```
## Support Vector Machines with Linear Kernel
##
## 3516 samples
##   10 predictor
##    4 classes: 'Excellent', 'Good', 'Fair', 'Poor'
##
## Pre-processing: median imputation (10), centered (10), scaled
##  (10), principal component signal extraction (10)
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 3164, 3165, 3165, 3164, 3165, 3165, ...
## Resampling results:
##
```

```
##   logLoss   AUC        prAUC      Accuracy   Kappa         Mean_F1
##   1.194138  0.5299337  0.2732627  0.4670096  0.0006310196  NaN
##   Mean_Sensitivity  Mean_Specificity  Mean_Pos_Pred_Value
##   0.2502551         0.750133          NaN
##   Mean_Neg_Pred_Value  Mean_Precision  Mean_Recall  Mean_Detection_Rate
##   0.8669912            NaN             0.2502551    0.1167524
##   Mean_Balanced_Accuracy
##   0.500194
##
## Tuning parameter 'C' was held constant at a value of 1
```

- run Extreme Gradient Boosted model

```
model_xgb <- train(
  x = final_per,
  y = final_health,
  metric = "ROC",
  method = "xgbTree",
  trControl = gssControl,
  preProcess = c("medianImpute", "zv","center", "scale", "pca")
)
```

```
model_xgb
```

```
## eXtreme Gradient Boosting
##
## 3516 samples
##   10 predictor
##    4 classes: 'Excellent', 'Good', 'Fair', 'Poor'
##
## Pre-processing: median imputation (10), centered (10), scaled
##  (10), principal component signal extraction (10)
## Resampling: Cross-Validated (10 fold, repeated 1 times)
## Summary of sample sizes: 3165, 3164, 3163, 3163, 3165, 3166, ...
## Resampling results across tuning parameters:
##
##   eta  max_depth  colsample_bytree  subsample  nrounds  logLoss
##   0.3  1          0.6               0.50        50      1.195526
##   0.3  1          0.6               0.50       100      1.202145
##   0.3  1          0.6               0.50       150      1.206923
##   0.3  1          0.6               0.75        50      1.192120
##   0.3  1          0.6               0.75       100      1.199391
##   0.3  1          0.6               0.75       150      1.206151
##   0.3  1          0.6               1.00        50      1.192694
##   0.3  1          0.6               1.00       100      1.197641
##   0.3  1          0.6               1.00       150      1.202611
##   0.3  1          0.8               0.50        50      1.194805
##   0.3  1          0.8               0.50       100      1.201466
##   0.3  1          0.8               0.50       150      1.208882
##   0.3  1          0.8               0.75        50      1.194342
##   0.3  1          0.8               0.75       100      1.201539
##   0.3  1          0.8               0.75       150      1.205657
##   0.3  1          0.8               1.00        50      1.192577
##   0.3  1          0.8               1.00       100      1.197591
##   0.3  1          0.8               1.00       150      1.202295
##   0.3  2          0.6               0.50        50      1.207071
```

```
##   0.3  2          0.6              0.50        100      1.226799
##   0.3  2          0.6              0.50        150      1.246584
##   0.3  2          0.6              0.75         50      1.205014
##   0.3  2          0.6              0.75        100      1.222265
##   0.3  2          0.6              0.75        150      1.238947
##   0.3  2          0.6              1.00         50      1.204186
##   0.3  2          0.6              1.00        100      1.224284
##   0.3  2          0.6              1.00        150      1.241163
##   0.3  2          0.8              0.50         50      1.207944
##   0.3  2          0.8              0.50        100      1.227147
##   0.3  2          0.8              0.50        150      1.246667
##   0.3  2          0.8              0.75         50      1.203822
##   0.3  2          0.8              0.75        100      1.225407
##   0.3  2          0.8              0.75        150      1.244263
##   0.3  2          0.8              1.00         50      1.202631
##   0.3  2          0.8              1.00        100      1.218727
##   0.3  2          0.8              1.00        150      1.235737
##   0.3  3          0.6              0.50         50      1.217347
##   0.3  3          0.6              0.50        100      1.253700
##   0.3  3          0.6              0.50        150      1.292504
##   0.3  3          0.6              0.75         50      1.214060
##   0.3  3          0.6              0.75        100      1.245111
##   0.3  3          0.6              0.75        150      1.276578
##   0.3  3          0.6              1.00         50      1.213481
##   0.3  3          0.6              1.00        100      1.240271
##   0.3  3          0.6              1.00        150      1.271575
##   0.3  3          0.8              0.50         50      1.232391
##   0.3  3          0.8              0.50        100      1.267399
##   0.3  3          0.8              0.50        150      1.303180
##   0.3  3          0.8              0.75         50      1.218752
##   0.3  3          0.8              0.75        100      1.256615
##   0.3  3          0.8              0.75        150      1.288810
##   0.3  3          0.8              1.00         50      1.217346
##   0.3  3          0.8              1.00        100      1.247858
##   0.3  3          0.8              1.00        150      1.282220
##   0.4  1          0.6              0.50         50      1.198944
##   0.4  1          0.6              0.50        100      1.208623
##   0.4  1          0.6              0.50        150      1.210407
##   0.4  1          0.6              0.75         50      1.195744
##   0.4  1          0.6              0.75        100      1.205775
##   0.4  1          0.6              0.75        150      1.211340
##   0.4  1          0.6              1.00         50      1.194381
##   0.4  1          0.6              1.00        100      1.201618
##   0.4  1          0.6              1.00        150      1.206574
##   0.4  1          0.8              0.50         50      1.199768
##   0.4  1          0.8              0.50        100      1.207813
##   0.4  1          0.8              0.50        150      1.213787
##   0.4  1          0.8              0.75         50      1.195353
##   0.4  1          0.8              0.75        100      1.205485
##   0.4  1          0.8              0.75        150      1.211541
##   0.4  1          0.8              1.00         50      1.194812
##   0.4  1          0.8              1.00        100      1.201620
##   0.4  1          0.8              1.00        150      1.207057
##   0.4  2          0.6              0.50         50      1.216964
```

```
##   0.4  2          0.6               0.50       100      1.244310
##   0.4  2          0.6               0.50       150      1.269209
##   0.4  2          0.6               0.75        50      1.211972
##   0.4  2          0.6               0.75       100      1.237467
##   0.4  2          0.6               0.75       150      1.261256
##   0.4  2          0.6               1.00        50      1.211675
##   0.4  2          0.6               1.00       100      1.235652
##   0.4  2          0.6               1.00       150      1.257692
##   0.4  2          0.8               0.50        50      1.219767
##   0.4  2          0.8               0.50       100      1.252701
##   0.4  2          0.8               0.50       150      1.274785
##   0.4  2          0.8               0.75        50      1.213487
##   0.4  2          0.8               0.75       100      1.240907
##   0.4  2          0.8               0.75       150      1.265618
##   0.4  2          0.8               1.00        50      1.211946
##   0.4  2          0.8               1.00       100      1.235217
##   0.4  2          0.8               1.00       150      1.258622
##   0.4  3          0.6               0.50        50      1.244748
##   0.4  3          0.6               0.50       100      1.298550
##   0.4  3          0.6               0.50       150      1.349013
##   0.4  3          0.6               0.75        50      1.236337
##   0.4  3          0.6               0.75       100      1.281672
##   0.4  3          0.6               0.75       150      1.320908
##   0.4  3          0.6               1.00        50      1.231253
##   0.4  3          0.6               1.00       100      1.272194
##   0.4  3          0.6               1.00       150      1.310125
##   0.4  3          0.8               0.50        50      1.254979
##   0.4  3          0.8               0.50       100      1.303927
##   0.4  3          0.8               0.50       150      1.350997
##   0.4  3          0.8               0.75        50      1.240009
##   0.4  3          0.8               0.75       100      1.284385
##   0.4  3          0.8               0.75       150      1.330493
##   0.4  3          0.8               1.00        50      1.229501
##   0.4  3          0.8               1.00       100      1.272682
##   0.4  3          0.8               1.00       150      1.316826
##   AUC        prAUC      Accuracy   Kappa        Mean_F1
##   0.5295856  0.2756149  0.4618765  0.016309865        NaN
##   0.5276458  0.2764456  0.4598975  0.019832858  0.2439888
##   0.5264783  0.2747347  0.4579072  0.020708299        NaN
##   0.5322883  0.2784832  0.4607425  0.009466261        NaN
##   0.5315253  0.2784033  0.4590307  0.015196789        NaN
##   0.5296181  0.2758359  0.4561865  0.015515868        NaN
##   0.5215655  0.2725867  0.4675519  0.017847233        NaN
##   0.5230598  0.2730158  0.4598821  0.012277511        NaN
##   0.5235756  0.2736889  0.4601654  0.016540155        NaN
##   0.5290210  0.2794827  0.4635851  0.019107932        NaN
##   0.5285399  0.2779332  0.4607287  0.023927640        NaN
##   0.5239311  0.2733254  0.4590194  0.024496801        NaN
##   0.5255563  0.2774185  0.4650047  0.020263018        NaN
##   0.5270281  0.2757473  0.4564698  0.012478854        NaN
##   0.5290470  0.2760380  0.4587571  0.019654850        NaN
##   0.5201054  0.2726888  0.4664163  0.015765733        NaN
##   0.5239707  0.2731502  0.4598805  0.012336775        NaN
##   0.5232530  0.2736947  0.4604511  0.017048446        NaN
```

```
##   0.5293417  0.2767054  0.4522140  0.015108229  0.2435261
##   0.5262441  0.2739446  0.4536385  0.028268645        NaN
##   0.5261214  0.2732037  0.4488049  0.024849393  0.2282619
##   0.5247067  0.2751721  0.4578854  0.020647146        NaN
##   0.5248303  0.2742377  0.4547668  0.025418197  0.2281368
##   0.5233170  0.2706289  0.4476572  0.017343072  0.2250605
##   0.5216145  0.2722724  0.4590169  0.017155864        NaN
##   0.5193274  0.2692180  0.4550590  0.020610911  0.2153192
##   0.5196107  0.2678144  0.4502318  0.019136251  0.2119551
##   0.5296732  0.2784487  0.4582051  0.025143170  0.2392120
##   0.5289796  0.2744352  0.4561978  0.033654080  0.2390308
##   0.5270200  0.2753581  0.4527935  0.033676451  0.2355696
##   0.5289867  0.2761953  0.4559259  0.016738969  0.2235002
##   0.5190564  0.2697686  0.4547806  0.025409180  0.2354883
##   0.5193073  0.2691092  0.4513803  0.026531506  0.2467597
##   0.5270068  0.2728550  0.4598741  0.019753391        NaN
##   0.5280686  0.2741191  0.4559186  0.022864637  0.2151037
##   0.5232458  0.2721800  0.4544949  0.026732530  0.2372874
##   0.5340162  0.2824125  0.4533463  0.026874141  0.2234923
##   0.5316536  0.2798455  0.4490906  0.028670331  0.2366308
##   0.5218892  0.2717835  0.4462343  0.029085387        NaN
##   0.5309020  0.2797700  0.4507895  0.015739636  0.2273042
##   0.5301705  0.2774978  0.4516410  0.029193370  0.2621117
##   0.5291290  0.2766975  0.4482359  0.028221930  0.2419706
##   0.5249378  0.2740894  0.4587732  0.024494766        NaN
##   0.5215149  0.2722752  0.4505103  0.020366962        NaN
##   0.5209641  0.2706739  0.4493698  0.025585456  0.2564799
##   0.5192464  0.2688652  0.4408333  0.005022908        NaN
##   0.5257462  0.2719493  0.4385630  0.012343637        NaN
##   0.5260836  0.2740004  0.4419761  0.021609064        NaN
##   0.5318020  0.2783801  0.4522156  0.021351624        NaN
##   0.5222103  0.2715691  0.4468025  0.023031861  0.2524660
##   0.5276631  0.2747136  0.4473739  0.028058302  0.2526703
##   0.5256069  0.2722648  0.4539121  0.019351278  0.2432774
##   0.5200237  0.2705975  0.4490865  0.021033060  0.2692431
##   0.5170187  0.2677314  0.4479397  0.025244619  0.2492802
##   0.5270091  0.2743248  0.4570372  0.012247624        NaN
##   0.5277808  0.2750749  0.4553399  0.020091158  0.2351466
##   0.5344583  0.2775265  0.4590258  0.028594751  0.2289400
##   0.5262089  0.2776150  0.4635875  0.020047509        NaN
##   0.5274187  0.2750669  0.4593131  0.021798371        NaN
##   0.5282653  0.2757241  0.4564754  0.020175788        NaN
##   0.5223462  0.2731662  0.4647215  0.016502553        NaN
##   0.5261820  0.2744649  0.4624374  0.020380628        NaN
##   0.5245237  0.2736439  0.4575997  0.014685444        NaN
##   0.5254858  0.2755657  0.4621606  0.022513489        NaN
##   0.5280750  0.2762550  0.4550356  0.019066464  0.2314518
##   0.5260907  0.2745053  0.4547701  0.021378326  0.2326605
##   0.5319475  0.2776375  0.4630169  0.018714369        NaN
##   0.5272570  0.2773072  0.4607344  0.026504697        NaN
##   0.5278416  0.2752002  0.4598797  0.027756618        NaN
##   0.5217312  0.2727678  0.4650007  0.017061872        NaN
##   0.5232694  0.2733144  0.4615883  0.019117397        NaN
##   0.5233517  0.2729606  0.4576070  0.015436959        NaN
```

```
##   0.5263087  0.2739387  0.4576102  0.027375289  0.2520079
##   0.5251368  0.2706802  0.4513521  0.029609109  0.2629058
##   0.5222801  0.2684707  0.4513618  0.034065162  0.2618652
##   0.5310832  0.2765782  0.4522165  0.015217187  0.2234305
##   0.5239602  0.2720817  0.4448308  0.014660654  0.2349332
##   0.5248747  0.2715930  0.4459615  0.023063663  0.2494724
##   0.5206884  0.2716297  0.4584431  0.021578137  0.2341656
##   0.5186745  0.2699510  0.4476499  0.013483446  0.2215525
##   0.5192497  0.2698544  0.4459462  0.017405376  0.2410727
##   0.5286790  0.2770622  0.4539137  0.024026995  0.2449550
##   0.5259584  0.2722859  0.4479405  0.025919748  0.2450027
##   0.5189490  0.2681944  0.4436661  0.019840536  0.2487994
##   0.5274040  0.2730546  0.4567692  0.023400592  0.2311108
##   0.5219769  0.2695784  0.4490914  0.023376469  0.2406643
##   0.5195916  0.2696349  0.4448325  0.019113532  0.2543537
##   0.5190738  0.2706568  0.4618741  0.028427319  0.2371547
##   0.5183520  0.2700950  0.4522043  0.022918834  0.2421541
##   0.5223276  0.2717440  0.4490874  0.024249997  0.2427682
##   0.5202409  0.2707243  0.4527838  0.033852762  0.2552239
##   0.5199062  0.2714090  0.4422796  0.023832880  0.2427317
##   0.5199186  0.2699042  0.4388665  0.018227915        NaN
##   0.5220407  0.2704507  0.4471190  0.016538698  0.2508553
##   0.5162869  0.2669558  0.4414145  0.014954165  0.2364319
##   0.5215302  0.2702223  0.4374283  0.013538180  0.2424934
##   0.5196341  0.2675355  0.4513585  0.021397825  0.2397638
##   0.5162514  0.2645758  0.4442433  0.017170090        NaN
##   0.5192435  0.2659409  0.4428309  0.019014652        NaN
##   0.5173419  0.2682035  0.4448324  0.019741688  0.2512581
##   0.5212357  0.2683443  0.4371652  0.012037830  0.2645977
##   0.5228781  0.2701366  0.4422877  0.024140957  0.2707473
##   0.5252080  0.2719098  0.4462327  0.015459824  0.2299835
##   0.5218457  0.2702111  0.4496580  0.031034896  0.2333469
##   0.5207536  0.2690242  0.4388381  0.016917815  0.2238713
##   0.5261310  0.2727589  0.4527838  0.023509842  0.2389075
##   0.5194012  0.2688432  0.4470922  0.023759486  0.2623557
##   0.5188315  0.2670788  0.4411149  0.017466001  0.2232180
##   Mean_Sensitivity  Mean_Specificity  Mean_Pos_Pred_Value
##   0.2572038         0.7533861         0.3107187
##   0.2595736         0.7540732         0.3088453
##   0.2590761         0.7543319         0.2929840
##   0.2544343         0.7518444         0.2795927
##   0.2569945         0.7531221         0.3005518
##   0.2564103         0.7532903         0.2852319
##   0.2575730         0.7537454         0.3656716
##   0.2555242         0.7524835         0.3306430
##   0.2573209         0.7534061         0.3320172
##   0.2580124         0.7540667         0.2924405
##   0.2603937         0.7550511         0.2897077
##   0.2604935         0.7552452         0.3020787
##   0.2588759         0.7542329         0.3364871
##   0.2557390         0.7524904         0.2963607
##   0.2585129         0.7541427         0.2949715
##   0.2566526         0.7533182         0.3537109
##   0.2556242         0.7524607         0.3012842
```

```
##   0.2576730         0.7534946         0.3354955
##   0.2587641         0.7527895         0.2909931
##   0.2627564         0.7558637         0.2898571
##   0.2610171         0.7553902         0.2695969
##   0.2595754         0.7542757         0.2658225
##   0.2626726         0.7553553         0.2942958
##   0.2590801         0.7535671         0.2831520
##   0.2582201         0.7533982         0.3018147
##   0.2603836         0.7542530         0.2826788
##   0.2597350         0.7539687         0.2843905
##   0.2622729         0.7552236         0.3099605
##   0.2686060         0.7569255         0.3174707
##   0.2681144         0.7571648         0.3079610
##   0.2599797         0.7532695         0.3469993
##   0.2635117         0.7554793         0.3073325
##   0.2665719         0.7552723         0.3192291
##   0.2592051         0.7540333         0.3002264
##   0.2616639         0.7547122         0.3066102
##   0.2632399         0.7555865         0.3025696
##   0.2648761         0.7554632         0.3337933
##   0.2664875         0.7560290         0.3324314
##   0.2632557         0.7562662         0.2684184
##   0.2585279         0.7530909         0.2974124
##   0.2647868         0.7560974         0.2997962
##   0.2641786         0.7560598         0.2905900
##   0.2609646         0.7551957         0.2862567
##   0.2594165         0.7542496         0.2596700
##   0.2640828         0.7552784         0.2933943
##   0.2536910         0.7506714         0.2548908
##   0.2554576         0.7525022         0.2530519
##   0.2614408         0.7543683         0.2658334
##   0.2602255         0.7543902         0.2830966
##   0.2638184         0.7545046         0.3078345
##   0.2647350         0.7558864         0.2918075
##   0.2608775         0.7538343         0.3107813
##   0.2605120         0.7544868         0.2891659
##   0.2644089         0.7553195         0.2843250
##   0.2549128         0.7525724         0.2455365
##   0.2608910         0.7540980         0.3106538
##   0.2647930         0.7559325         0.3086608
##   0.2584261         0.7542580         0.3125129
##   0.2593053         0.7546122         0.2989211
##   0.2585116         0.7542725         0.2869246
##   0.2572893         0.7534217         0.3603506
##   0.2589415         0.7542762         0.3220185
##   0.2568548         0.7529090         0.3040417
##   0.2599183         0.7546455         0.2809251
##   0.2594658         0.7538524         0.3003829
##   0.2608373         0.7543041         0.3029723
##   0.2578241         0.7539800         0.2562764
##   0.2613282         0.7556557         0.2889373
##   0.2621795         0.7559120         0.3034245
##   0.2577462         0.7534723         0.3485442
##   0.2580831         0.7540845         0.3784407
```

```
##   0.2567593         0.7531483         0.2962737
##   0.2648919         0.7555775         0.3355970
##   0.2657486         0.7562003         0.3065258
##   0.2683216         0.7571081         0.3078741
##   0.2598984         0.7527732         0.3083904
##   0.2591119         0.7529442         0.2940238
##   0.2638321         0.7546378         0.2962337
##   0.2611586         0.7543340         0.3232678
##   0.2587079         0.7525710         0.3032626
##   0.2590528         0.7535455         0.2794864
##   0.2623564         0.7550918         0.2985028
##   0.2644296         0.7554732         0.3282662
##   0.2604480         0.7541239         0.2851579
##   0.2620959         0.7547734         0.2986872
##   0.2630195         0.7548305         0.2960793
##   0.2612683         0.7538281         0.2865743
##   0.2645323         0.7558027         0.3331389
##   0.2619186         0.7546536         0.3001067
##   0.2620724         0.7550180         0.2888287
##   0.2663479         0.7572161         0.3117491
##   0.2625606         0.7547411         0.2896286
##   0.2581002         0.7536315         0.2620437
##   0.2607519         0.7531498         0.3303437
##   0.2624038         0.7527007         0.3113463
##   0.2605450         0.7524424         0.2960268
##   0.2628020         0.7544416         0.3366276
##   0.2578622         0.7535865         0.2592389
##   0.2582497         0.7541012         0.2604036
##   0.2612349         0.7538924         0.2962538
##   0.2569269         0.7522647         0.2799789
##   0.2623364         0.7550065         0.2920382
##   0.2600654         0.7528779         0.3024474
##   0.2652163         0.7565924         0.2882581
##   0.2599788         0.7535416         0.2659733
##   0.2621741         0.7548381         0.2856854
##   0.2621798         0.7550100         0.2792387
##   0.2588154         0.7536808         0.2604732
##   Mean_Neg_Pred_Value  Mean_Precision  Mean_Recall  Mean_Detection_Rate
##   0.7628276            0.3107187       0.2572038    0.1154691
##   0.7590943            0.3088453       0.2595736    0.1149744
##   0.7590729            0.2929840       0.2590761    0.1144768
##   0.7504180            0.2795927       0.2544343    0.1151856
##   0.7555277            0.3005518       0.2569945    0.1147577
##   0.7567052            0.2852319       0.2564103    0.1140466
##   0.7703048            0.3656716       0.2575730    0.1168880
##   0.7542237            0.3306430       0.2555242    0.1149705
##   0.7566491            0.3320172       0.2573209    0.1150414
##   0.7618725            0.2924405       0.2580124    0.1158963
##   0.7639588            0.2897077       0.2603937    0.1151822
##   0.7636301            0.3020787       0.2604935    0.1147548
##   0.7650066            0.3364871       0.2588759    0.1162512
##   0.7534760            0.2963607       0.2557390    0.1141174
##   0.7588987            0.2949715       0.2585129    0.1146893
##   0.7696528            0.3537109       0.2566526    0.1166041
```

```
##   0.7530781          0.3012842      0.2556242   0.1149701
##   0.7566380          0.3354955      0.2576730   0.1151128
##   0.7490787          0.2909931      0.2587641   0.1130535
##   0.7590659          0.2898571      0.2627564   0.1134096
##   0.7606145          0.2695969      0.2610171   0.1122012
##   0.7597316          0.2658225      0.2595754   0.1144714
##   0.7600195          0.2942958      0.2626726   0.1136917
##   0.7547979          0.2831520      0.2590801   0.1119143
##   0.7566955          0.3018147      0.2582201   0.1147542
##   0.7580924          0.2826788      0.2603836   0.1137648
##   0.7566707          0.2843905      0.2597350   0.1125580
##   0.7622250          0.3099605      0.2622729   0.1145513
##   0.7622425          0.3174707      0.2686060   0.1140494
##   0.7629514          0.3079610      0.2681144   0.1131984
##   0.7533698          0.3469993      0.2599797   0.1139815
##   0.7628419          0.3073325      0.2635117   0.1136951
##   0.7578632          0.3192291      0.2665719   0.1128451
##   0.7579597          0.3002264      0.2592051   0.1149685
##   0.7573488          0.3066102      0.2616639   0.1139796
##   0.7605031          0.3025696      0.2632399   0.1136237
##   0.7579043          0.3337933      0.2648761   0.1133366
##   0.7597035          0.3324314      0.2664875   0.1122727
##   0.7599648          0.2684184      0.2632557   0.1115586
##   0.7533186          0.2974124      0.2585279   0.1126974
##   0.7596806          0.2997962      0.2647868   0.1129102
##   0.7608205          0.2905900      0.2641786   0.1120590
##   0.7635684          0.2862567      0.2609646   0.1146933
##   0.7580647          0.2596700      0.2594165   0.1126276
##   0.7588019          0.2933943      0.2640828   0.1123425
##   0.7472827          0.2548908      0.2536910   0.1102083
##   0.7521516          0.2530519      0.2554576   0.1096407
##   0.7554540          0.2658334      0.2614408   0.1104940
##   0.7586372          0.2830966      0.2602255   0.1130539
##   0.7555674          0.3078345      0.2638184   0.1117006
##   0.7597094          0.2918075      0.2647350   0.1118435
##   0.7560262          0.3107813      0.2608775   0.1134780
##   0.7596417          0.2891659      0.2605120   0.1122716
##   0.7601044          0.2843250      0.2644089   0.1119849
##   0.7554735          0.2455365      0.2549128   0.1142593
##   0.7567928          0.3106538      0.2608910   0.1138350
##   0.7621136          0.3086608      0.2647930   0.1147565
##   0.7646615          0.3125129      0.2584261   0.1158969
##   0.7610314          0.2989211      0.2593053   0.1148283
##   0.7599710          0.2869246      0.2585116   0.1141189
##   0.7620020          0.3603506      0.2572893   0.1161804
##   0.7615843          0.3220185      0.2589415   0.1156093
##   0.7544778          0.3040417      0.2568548   0.1143999
##   0.7627464          0.2809251      0.2599183   0.1155401
##   0.7570889          0.3003829      0.2594658   0.1137589
##   0.7572232          0.3029723      0.2608373   0.1136925
##   0.7632905          0.2562764      0.2578241   0.1157542
##   0.7636759          0.2889373      0.2613282   0.1151836
##   0.7622399          0.3034245      0.2621795   0.1149699
##   0.7597053          0.3485442      0.2577462   0.1162502
```

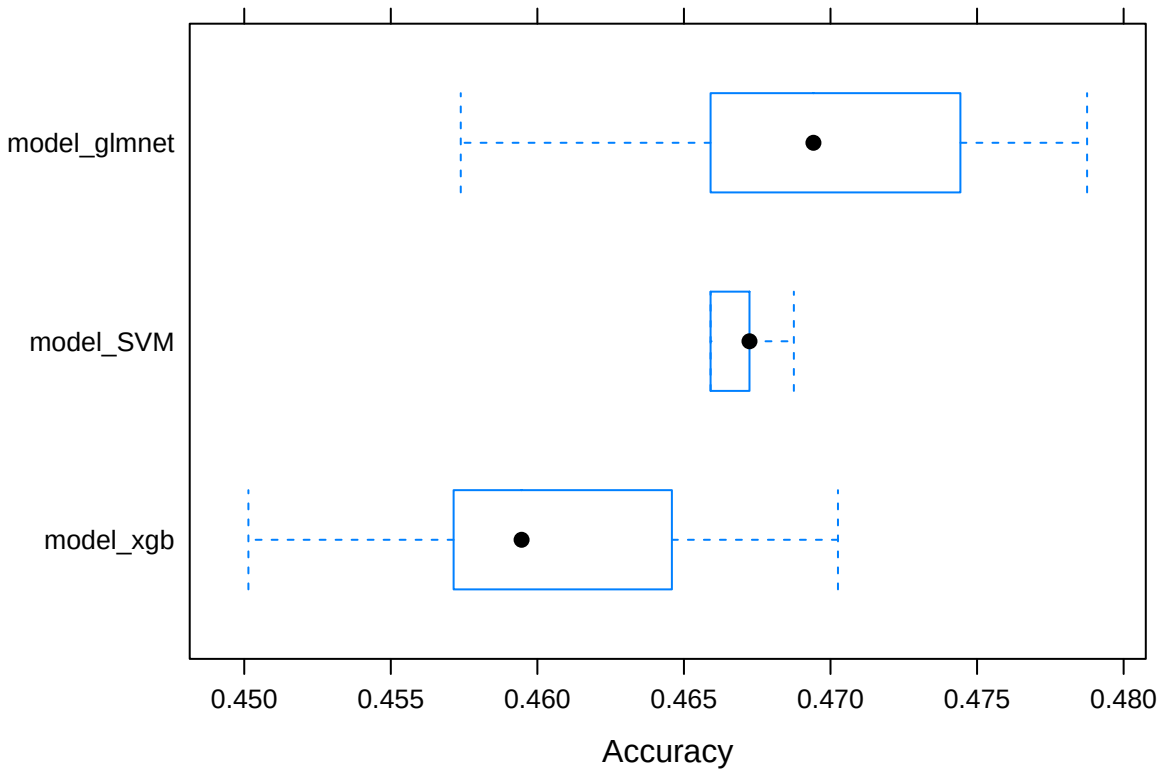
```
##   0.7615209           0.3784407       0.2580831    0.1153971
##   0.7559885           0.2962737       0.2567593    0.1144017
##   0.7607251           0.3355970       0.2648919    0.1144025
##   0.7598052           0.3065258       0.2657486    0.1128380
##   0.7605615           0.3078741       0.2683216    0.1128404
##   0.7508873           0.3083904       0.2598984    0.1130541
##   0.7534782           0.2940238       0.2591119    0.1112077
##   0.7553626           0.2962337       0.2638321    0.1114904
##   0.7579537           0.3232678       0.2611586    0.1146108
##   0.7527822           0.3032626       0.2587079    0.1119125
##   0.7548528           0.2794864       0.2590528    0.1114865
##   0.7611328           0.2985028       0.2623564    0.1134784
##   0.7593646           0.3282662       0.2644296    0.1119851
##   0.7562516           0.2851579       0.2604480    0.1109165
##   0.7587062           0.2986872       0.2620959    0.1141923
##   0.7583556           0.2960793       0.2630195    0.1122728
##   0.7547839           0.2865743       0.2612683    0.1112081
##   0.7620396           0.3331389       0.2645323    0.1154685
##   0.7575223           0.3001067       0.2619186    0.1130511
##   0.7578488           0.2888287       0.2620724    0.1122718
##   0.7635194           0.3117491       0.2663479    0.1131960
##   0.7555856           0.2896286       0.2625606    0.1105699
##   0.7538076           0.2620437       0.2581002    0.1097166
##   0.7523932           0.3303437       0.2607519    0.1117797
##   0.7502834           0.3113463       0.2624038    0.1103536
##   0.7510511           0.2960268       0.2605450    0.1093571
##   0.7583295           0.3366276       0.2628020    0.1128396
##   0.7554909           0.2592389       0.2578622    0.1110608
##   0.7572231           0.2604036       0.2582497    0.1107077
##   0.7547561           0.2962538       0.2612349    0.1112081
##   0.7517921           0.2799789       0.2569269    0.1092913
##   0.7584291           0.2920382       0.2623364    0.1105719
##   0.7518652           0.3024474       0.2600654    0.1115582
##   0.7608236           0.2882581       0.2652163    0.1124145
##   0.7554043           0.2659733       0.2599788    0.1097095
##   0.7590469           0.2856854       0.2621741    0.1131959
##   0.7596569           0.2792387       0.2621798    0.1117731
##   0.7559553           0.2604732       0.2588154    0.1102787
##   Mean_Balanced_Accuracy
##   0.5052949
##   0.5068234
##   0.5067040
##   0.5031394
##   0.5050583
##   0.5048503
##   0.5056592
##   0.5040039
##   0.5053635
##   0.5060396
##   0.5077224
##   0.5078694
##   0.5065544
##   0.5041147
##   0.5063278
```

```
##   0.5049854
##   0.5040424
##   0.5055838
##   0.5057768
##   0.5093101
##   0.5082037
##   0.5069256
##   0.5090140
##   0.5063236
##   0.5058092
##   0.5073183
##   0.5068518
##   0.5087482
##   0.5127657
##   0.5126396
##   0.5066246
##   0.5094955
##   0.5109221
##   0.5066192
##   0.5081881
##   0.5094132
##   0.5101696
##   0.5112583
##   0.5097609
##   0.5058094
##   0.5104421
##   0.5101192
##   0.5080801
##   0.5068330
##   0.5096806
##   0.5021812
##   0.5039799
##   0.5079045
##   0.5073079
##   0.5091615
##   0.5103107
##   0.5073559
##   0.5074994
##   0.5098642
##   0.5037426
##   0.5074945
##   0.5103627
##   0.5063421
##   0.5069588
##   0.5063920
##   0.5053555
##   0.5066088
##   0.5048819
##   0.5072819
##   0.5066591
##   0.5075707
##   0.5059020
##   0.5084919
##   0.5090458
```

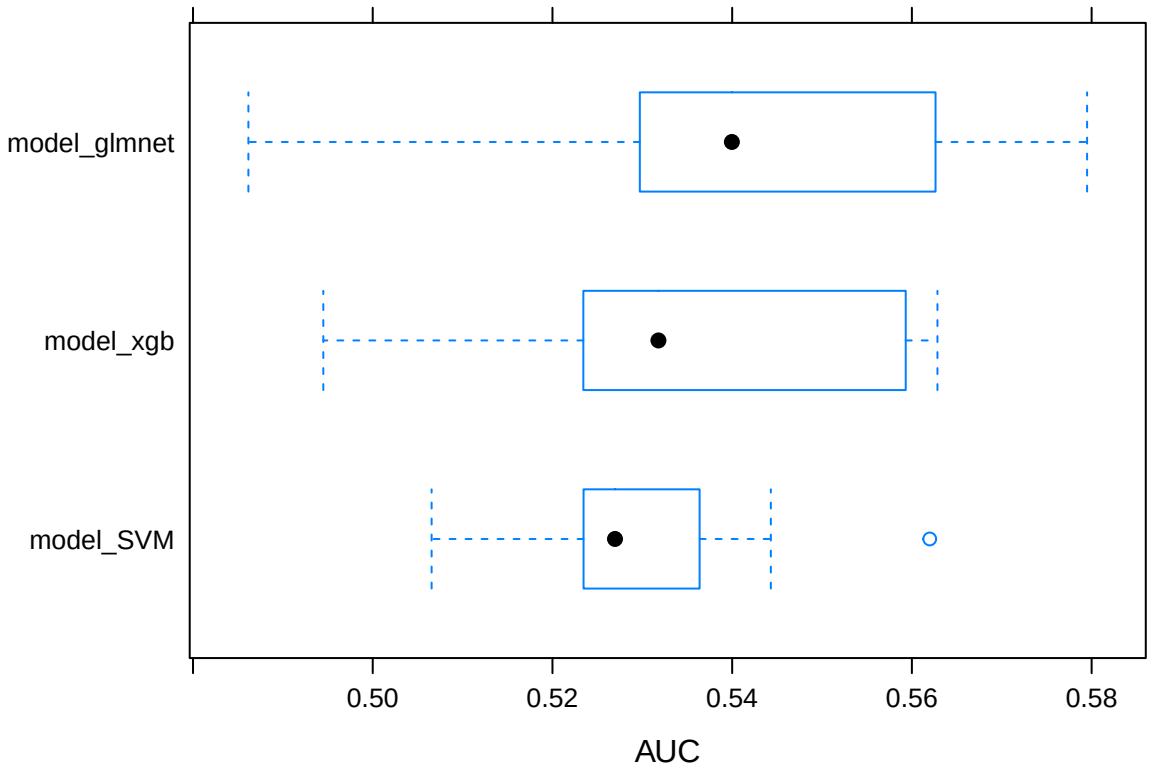
```
##   0.5056092
##   0.5060838
##   0.5049538
##   0.5102347
##   0.5109744
##   0.5127148
##   0.5063358
##   0.5060280
##   0.5092349
##   0.5077463
##   0.5056395
##   0.5062992
##   0.5087241
##   0.5099514
##   0.5072860
##   0.5084346
##   0.5089250
##   0.5075482
##   0.5101675
##   0.5082861
##   0.5085452
##   0.5117820
##   0.5086509
##   0.5058659
##   0.5069509
##   0.5075523
##   0.5064937
##   0.5086218
##   0.5057244
##   0.5061754
##   0.5075637
##   0.5045958
##   0.5086715
##   0.5064717
##   0.5109043
##   0.5067602
##   0.5085061
##   0.5085949
##   0.5062481
##
## Tuning parameter 'gamma' was held constant at a value of 0
##
## Tuning parameter 'min_child_weight' was held constant at a value of 1
## logLoss was used to select the optimal model using the smallest value.
## The final values used for the model were nrounds = 50, max_depth = 1,
##  eta = 0.3, gamma = 0, colsample_bytree = 0.6, min_child_weight = 1
##  and subsample = 0.75.
```

- compare models

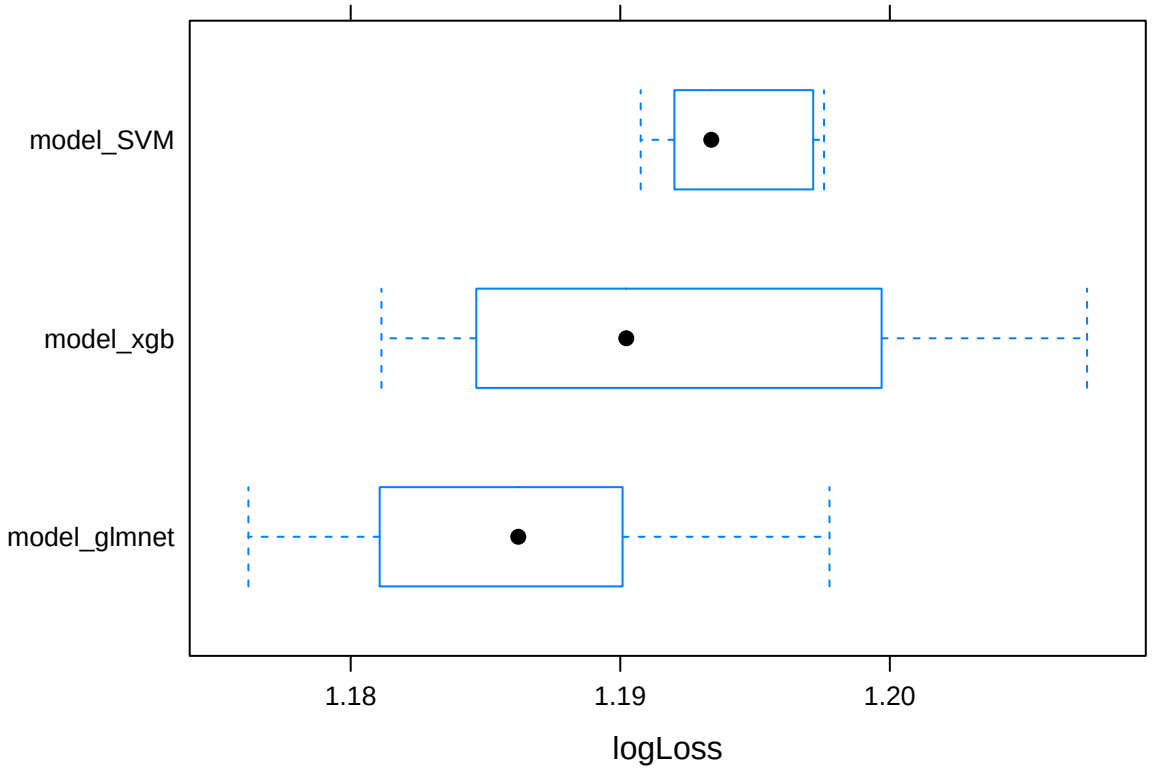
```
model_list <- list(model_glmnet = model_glmnet, model_SVM = model_SVM, model_xgb = model_xgb)
resamples <- resamples(model_list)
bwplot(resamples, metric = "Accuracy")
```




```
bwplot(resamples, metric = "AUC")
```




```
bwplot(resamples, metric = "logLoss")
```



It appears that none of the model will be able to predict as AUCs are around 0.5 and logloss is large as well. Therefore, although model glmnet seems to perform the best based on log lost and accuracy among the three models, the difference is small and none of the model seems to significantly outperform the others.

I would prefer to choose OLS model. Because: - although OLS model does not seem to explain much of the variance (adjusted R^2 = 2%), none of the three machine learning model (glmnet, SVM, xgb) seems to predict HEALTH based on personality variables better than chance (AUC ~ 0.5 and logLoss > 1). - in this case, N is orders of magnitude larger than k (3000+ observations with 10 predictors), therefore, machine learning shouldn't get much better prediction.