Random Forest

5500 samples

16 predictor

2 classes: '0', '1'

Pre-processing: scaled (16), centered (16)

Resampling: Cross-Validated (5 fold, repeated 5 times)

Summary of sample sizes: 4399, 4400, 4401, 4401, 4399, 4401, ...

Resampling results across tuning parameters:

mtry Accuracy Kappa

1 0.7700746 0.5398622

2 0.7705833 0.5409212

3 0.7726203 0.5450163

4 0.7721106 0.5440065

5 0.7726556 0.5451162

6 0.7718915 0.5435830

Accuracy was used to select the optimal model using the largest value.

The final value used for the model was mtry = 5. Acc=0.509

Accuracy was used to select the optimal model using the largest value.

The final values used for the model were n.trees = 110, interaction.depth =

5, shrinkage = 0.1 and n.minobsinnode = 20. Acc= 0.50920.

Boosting

Call:

summary.resamples(object = boosting\_results)

Models: nn, gbm, rf

Number of resamples: 25

Accuracy

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

nn 0.7106 0.7182 0.7364 0.7311 0.7386 0.7534 0

gbm 0.7648 0.7818 0.7861 0.7856 0.7920 0.8048 0

rf 0.7539 0.7636 0.7727 0.7727 0.7809 0.7909 0

Kappa

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

nn 0.4202 0.4384 0.4729 0.4619 0.4775 0.5061 0

gbm 0.5293 0.5631 0.5723 0.5710 0.5842 0.6093 0

rf 0.5074 0.5272 0.5456 0.5451 0.5614 0.5816 0

ENSEMBLE

Call:

summary.resamples(object = results)

Models: gbm, mlp, rf

Number of resamples: 25

ROC

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

gbm 0.8495 0.8599 0.8633 0.8635 0.8676 0.8785 0

mlp 0.7569 0.7703 0.7794 0.7794 0.7894 0.7979 0

rf 0.8349 0.8511 0.8569 0.8556 0.8603 0.8681 0

Sens

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

gbm 0.7307 0.7457 0.7556 0.7603 0.7704 0.8109 0

mlp 0.5802 0.6825 0.7026 0.7034 0.7414 0.7681 0

rf 0.7038 0.7333 0.7481 0.7468 0.7613 0.7857 0

Spec

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

gbm 0.7464 0.7749 0.7915 0.7890 0.8055 0.8232 0

mlp 0.6443 0.6847 0.7268 0.7244 0.7494 0.8171 0

rf 0.7421 0.7803 0.7877 0.7863 0.8030 0.8217 0

> modelCor(results)

gbm mlp rf

gbm 1.0000000 0.4185210 0.8852719

mlp 0.4185210 1.0000000 0.4213178

rf 0.8852719 0.4213178 1.0000000

ENSEMBLE II

Call:

summary.resamples(object = results)

Models: gbm, mlp

Number of resamples: 25

ROC

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

gbm 0.8495 0.8591 0.8616 0.8630 0.8674 0.8793 0

mlp 0.7447 0.7693 0.7762 0.7754 0.7799 0.7984 0

Sens

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

gbm 0.7281 0.7408 0.7568 0.7596 0.7721 0.8109 0

mlp 0.5917 0.6774 0.6953 0.6929 0.7214 0.7802 0

Spec

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

gbm 0.7416 0.7749 0.7880 0.7897 0.8072 0.8183 0

mlp 0.6220 0.6858 0.7198 0.7225 0.7500 0.8190 0

gbm mlp

gbm 1.0000000 0.4817746

mlp 0.4817746 1.0000000

The following models were ensembled: gbm, mlp

They were weighted:

2.511 -4.732 -0.3269

The resulting ROC is: 0.5

The fit for each individual model on the ROC is:

method ROC ROCSD

gbm 0.8629674 0.007049532

mlp 0.7753859 0.011113011

Average predictors

All three average

predictions

targets 1 2

1 468 71

2 182 378

Average of nn,rf

predictions

targets 1 2

1 399 140

2 133 427

Average of nn,gbm

predictions

targets 1 2

1 431 108

2 150 410

Average of rf,gbm

predictions

targets 1 2

1 418 121

2 117 443

Based on these we choose gbm+rf – acc=0.42