Support Vector Model

Author: Rajat Jain

Last Updated: 2018-05-08

Contents

Setting Training Control Params

Using 10-fold Cross Validation with 10 repeatitions.

```
ctrl <- trainControl(method="repeatedcv",
                     number=10,
                     repeats=10,
                     classProbs=TRUE,
                     savePredictions=TRUE,
                     allowParallel=TRUE)
set.seed(123)
```

Training - SVM Model

Caret is Awesome!

```
fit <- train(class ~ lr_cc_usage + lr_cl_usage + storage_usage + ps_usage + stock_usage,
             data=usage.data, method = "svmLinear", trControl = ctrl)
```

Summarize trained model.

```
## Support Vector Machines with Linear Kernel
##
## 2650 samples
##    5 predictor
##    2 classes: 'OTHER', 'PHOTOGRAPHER'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 10 times)
## Summary of sample sizes: 2385, 2385, 2385, 2385, 2385, 2385, ...
## Resampling results:
##
##   Accuracy   Kappa
##   0.5787547  0.05659445
##
## Tuning parameter 'C' was held constant at a value of 1
```

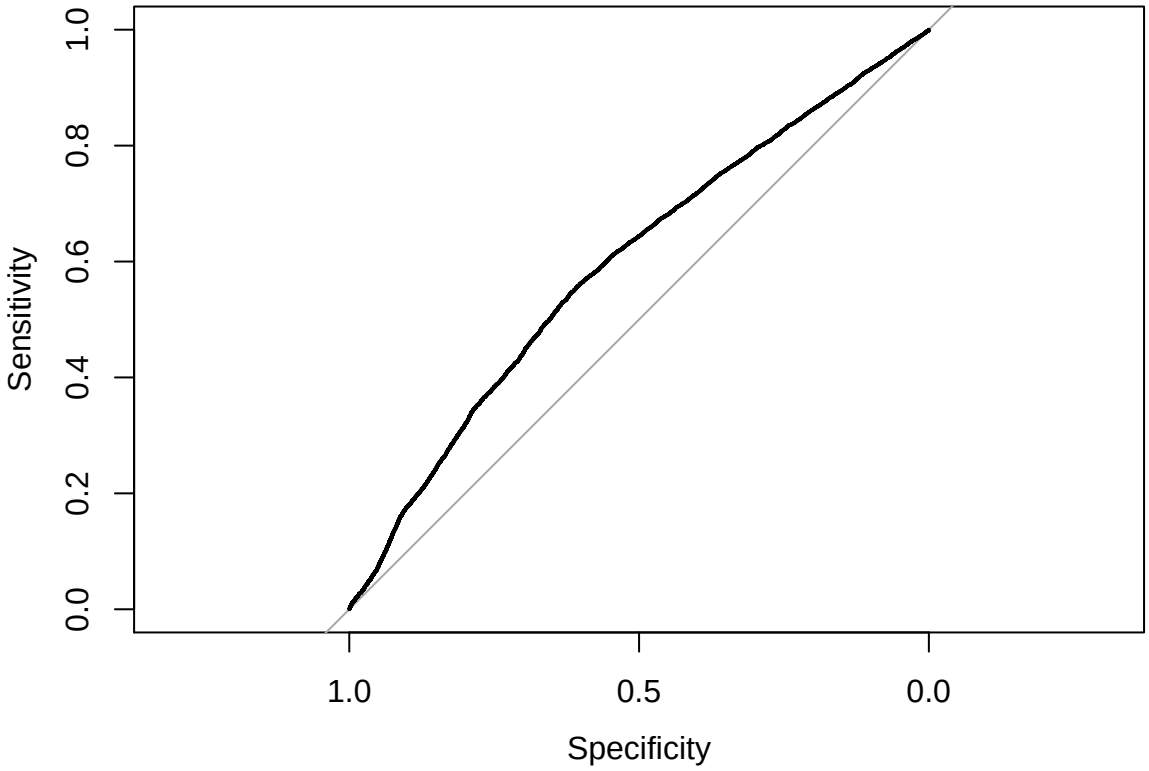
Performance

Based on the measure defined in the FPS, we will use classification accuracy as our performance measure.

Confusion Matrix

```
## Cross-Validated (10 fold, repeated 10 times) Confusion Matrix
##
## (entries are percentual average cell counts across resamples)
##
##                Reference
## Prediction      OTHER PHOTOGRAPHER
##   OTHER          52.5         38.0
##   PHOTOGRAPHER    4.1          5.4
##
##  Accuracy (average) : 0.5788
```

ROC Curve




```
##
## Call:
## plot.roc.default(x = fit$pred$obs, predictor = fit$pred$PHOTOGRAPHER)
##
## Data: fit$pred$PHOTOGRAPHER in 15000 controls (fit$pred$obs OTHER) < 11500 cases (fit$pred$obs PHOTO
## Area under the curve: 0.5948
```

Accuracy

- Kohen's Kappa: 0.06
- Observed Accuracy : 57.88%
- Desired accuracy : 70%
- Performance is Not Satisfactory.