

SmartFly: Train model and validate via cross-validation

Cindy Lamm

January 12, 2015

Load prepared data from the previous step "Prepare Data For Modeling"

```
rm(list=ls()) #clear memory
load("../02_prepare_data_for_modeling/prepared_data.Rdata")
```

Model a flight being delayed as dependent of the taxi time in:

[illegible]

[illegible]

[illegible]

[illegible]


```

##
## 3687183 samples
##      20 predictor
##      2 classes: 'delayed', 'on_time'
##
## No pre-processing
## Resampling: Cross-Validated (3 fold)
##
## Summary of sample sizes: 2458122, 2458122, 2458122
##
## Resampling results across tuning parameters:
##
##   alpha      lambda  Accuracy   Kappa  Accuracy SD
##   0.0500000  0.1      0.7138059  NaN    0
##   0.0500000  0.2      0.7138059  NaN    0
##   0.0500000  0.3      0.7138059  NaN    0
##   0.0500000  0.4      0.7138059  NaN    0
##   0.0500000  0.5      0.7138059  NaN    0
##   0.1178571  0.1      0.7138059  NaN    0
##   0.1178571  0.2      0.7138059  NaN    0
##   0.1178571  0.3      0.7138059  NaN    0
##   0.1178571  0.4      0.7138059  NaN    0
##   0.1178571  0.5      0.7138059  NaN    0
##   0.1857143  0.1      0.7138059  NaN    0
##   0.1857143  0.2      0.7138059  NaN    0
##   0.1857143  0.3      0.7138059  NaN    0
##   0.1857143  0.4      0.7138059  NaN    0
##   0.1857143  0.5      0.7138059  NaN    0
##   0.2535714  0.1      0.7138059  NaN    0
##   0.2535714  0.2      0.7138059  NaN    0
##   0.2535714  0.3      0.7138059  NaN    0
##   0.2535714  0.4      0.7138059  NaN    0
##   0.2535714  0.5      0.7138059  NaN    0
##   0.3214286  0.1      0.7138059  NaN    0
##   0.3214286  0.2      0.7138059  NaN    0
##   0.3214286  0.3      0.7138059  NaN    0
##   0.3214286  0.4      0.7138059  NaN    0
##   0.3214286  0.5      0.7138059  NaN    0
##   0.3892857  0.1      0.7138059  NaN    0
##   0.3892857  0.2      0.7138059  NaN    0
##   0.3892857  0.3      0.7138059  NaN    0
##   0.3892857  0.4      0.7138059  NaN    0
##   0.3892857  0.5      0.7138059  NaN    0
##   0.4571429  0.1      0.7138059  NaN    0
##   0.4571429  0.2      0.7138059  NaN    0
##   0.4571429  0.3      0.7138059  NaN    0
##   0.4571429  0.4      0.7138059  NaN    0
##   0.4571429  0.5      0.7138059  NaN    0
##   0.5250000  0.1      0.7138059  NaN    0
##   0.5250000  0.2      0.7138059  NaN    0
##   0.5250000  0.3      0.7138059  NaN    0
##   0.5250000  0.4      0.7138059  NaN    0
##   0.5250000  0.5      0.7138059  NaN    0
##   0.5928571  0.1      0.7138059  NaN    0

```

```
## 0.5928571 0.2 0.7138059 NaN 0
## 0.5928571 0.3 0.7138059 NaN 0
## 0.5928571 0.4 0.7138059 NaN 0
## 0.5928571 0.5 0.7138059 NaN 0
## 0.6607143 0.1 0.7138059 NaN 0
## 0.6607143 0.2 0.7138059 NaN 0
## 0.6607143 0.3 0.7138059 NaN 0
## 0.6607143 0.4 0.7138059 NaN 0
## 0.6607143 0.5 0.7138059 NaN 0
## 0.7285714 0.1 0.7138059 NaN 0
## 0.7285714 0.2 0.7138059 NaN 0
## 0.7285714 0.3 0.7138059 NaN 0
## 0.7285714 0.4 0.7138059 NaN 0
## 0.7285714 0.5 0.7138059 NaN 0
## 0.7964286 0.1 0.7138059 NaN 0
## 0.7964286 0.2 0.7138059 NaN 0
## 0.7964286 0.3 0.7138059 NaN 0
## 0.7964286 0.4 0.7138059 NaN 0
## 0.7964286 0.5 0.7138059 NaN 0
## 0.8642857 0.1 0.7138059 NaN 0
## 0.8642857 0.2 0.7138059 NaN 0
## 0.8642857 0.3 0.7138059 NaN 0
## 0.8642857 0.4 0.7138059 NaN 0
## 0.8642857 0.5 0.7138059 NaN 0
## 0.9321429 0.1 0.7138059 NaN 0
## 0.9321429 0.2 0.7138059 NaN 0
## 0.9321429 0.3 0.7138059 NaN 0
## 0.9321429 0.4 0.7138059 NaN 0
## 0.9321429 0.5 0.7138059 NaN 0
## 1.0000000 0.1 0.7138059 NaN 0
## 1.0000000 0.2 0.7138059 NaN 0
## 1.0000000 0.3 0.7138059 NaN 0
## 1.0000000 0.4 0.7138059 NaN 0
## 1.0000000 0.5 0.7138059 NaN 0
##
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were alpha = 0.05 and lambda = 0.5.
```

Save the model result:

```
save(glmnetFit1, file="glmnetFit1.Rdata")
```