Evaluate testing data (regression) - Lasso

EVE W.

2019-11-16

Contents


```
## user input
project_home <- "~/EVE/examples"
project_name <- "lasso_regression_noSplitCV"
```

0. Load Data

```
## 300 of samples were used
```

```
## 100 of full features
```

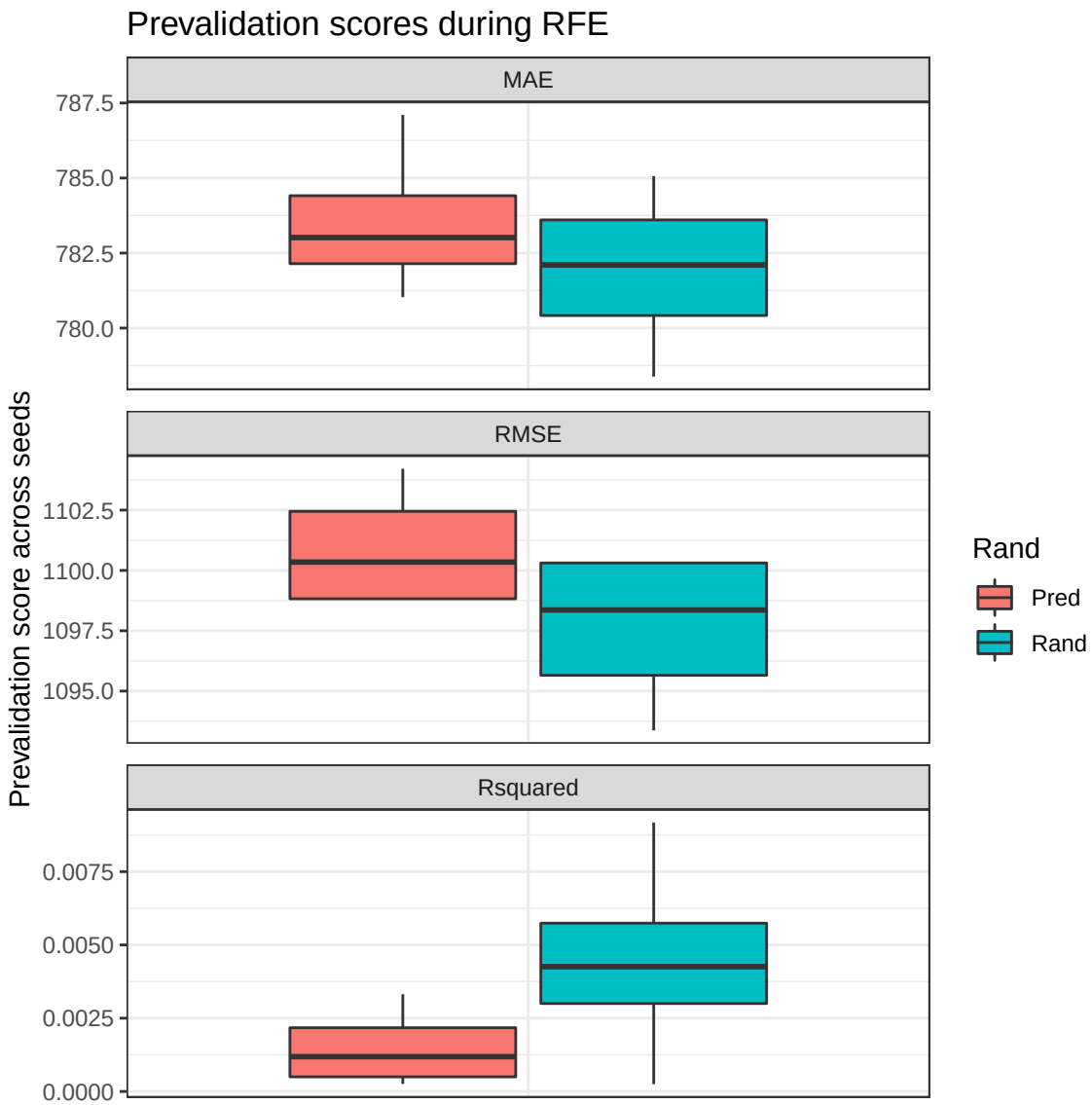
```
## 4 runs, each run contains 3 CVs.
```

```
## os_time :
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0   182.8   480.0   889.4  1221.2  7125.0
```

run with lasso.r.

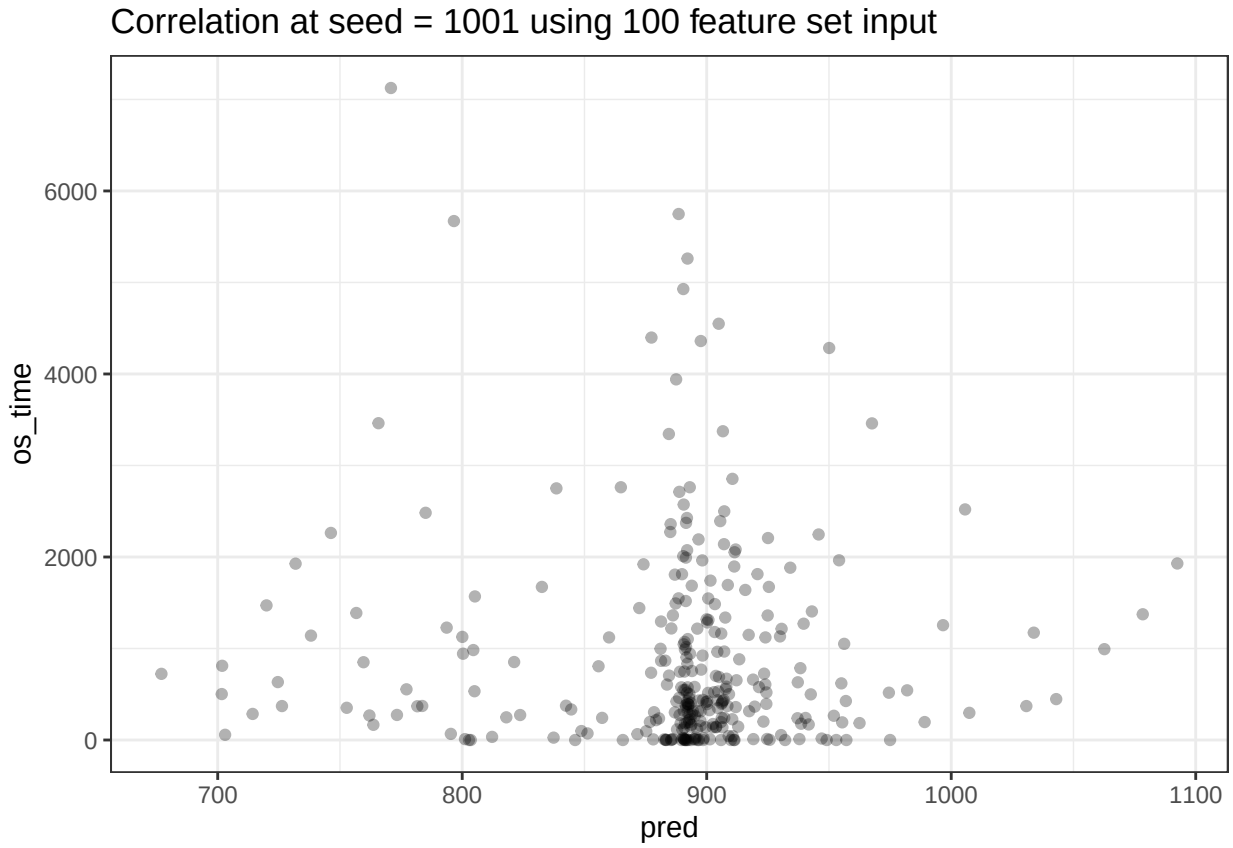
1. Scores



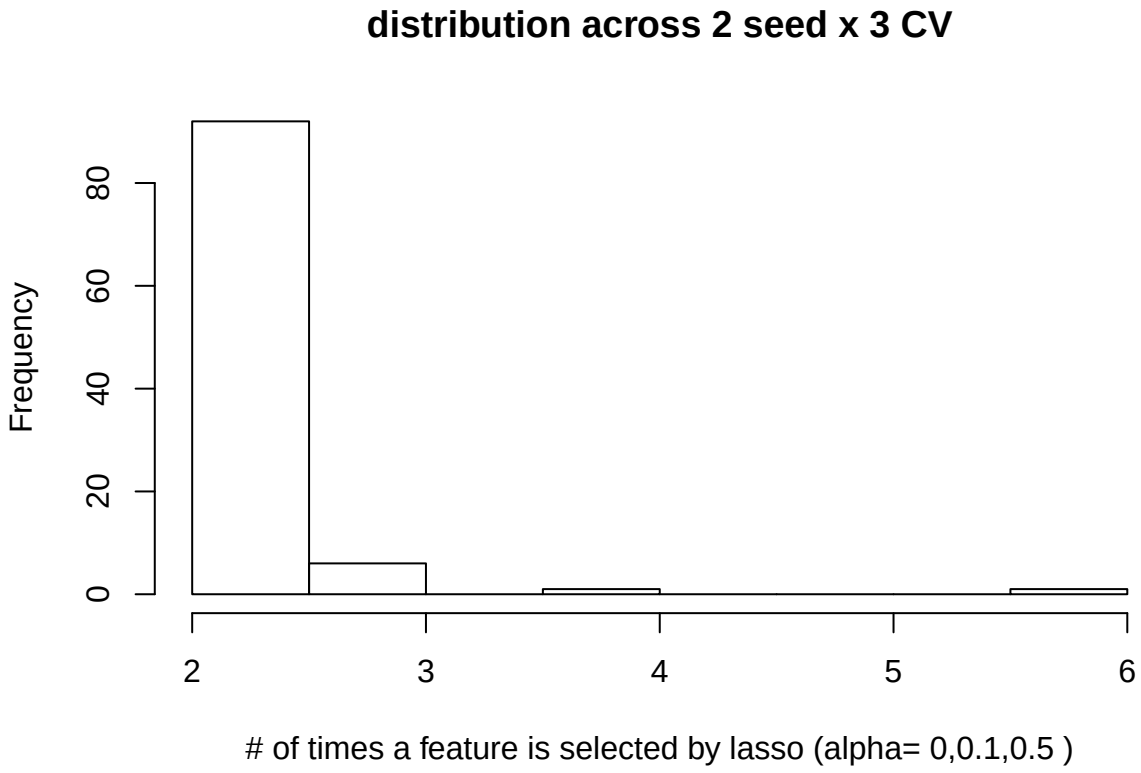
'Pred' compares the actual CV prediction with observed value. 'Rand' compares permuted CV prediction with observed to mimic random prediction.

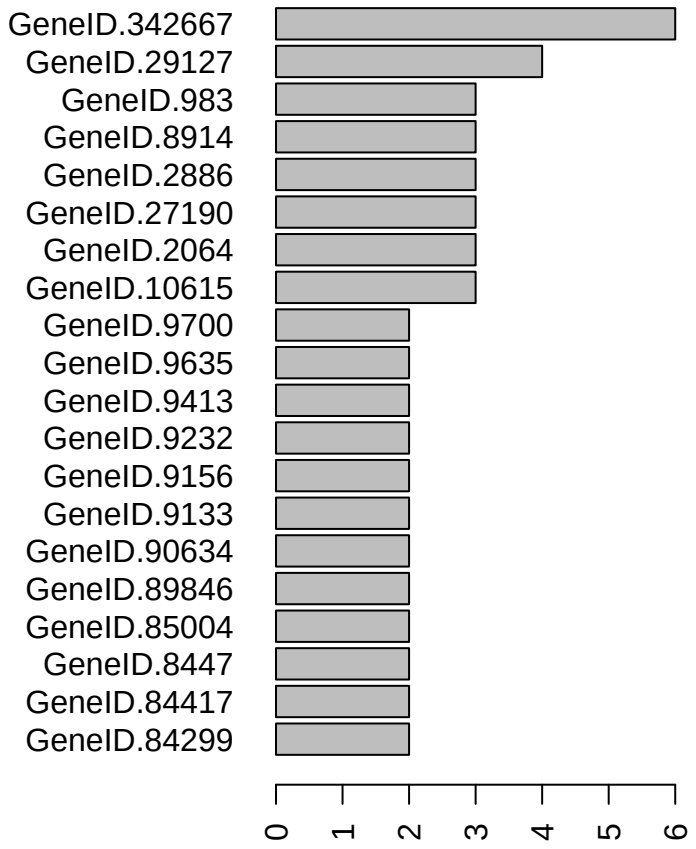
correlation

```
##
##
## Table: Averaged pearson correlation across seeds
##
##  cor.avg   cor.sdt
## --------  --------
##   -0.035   0.01867
```



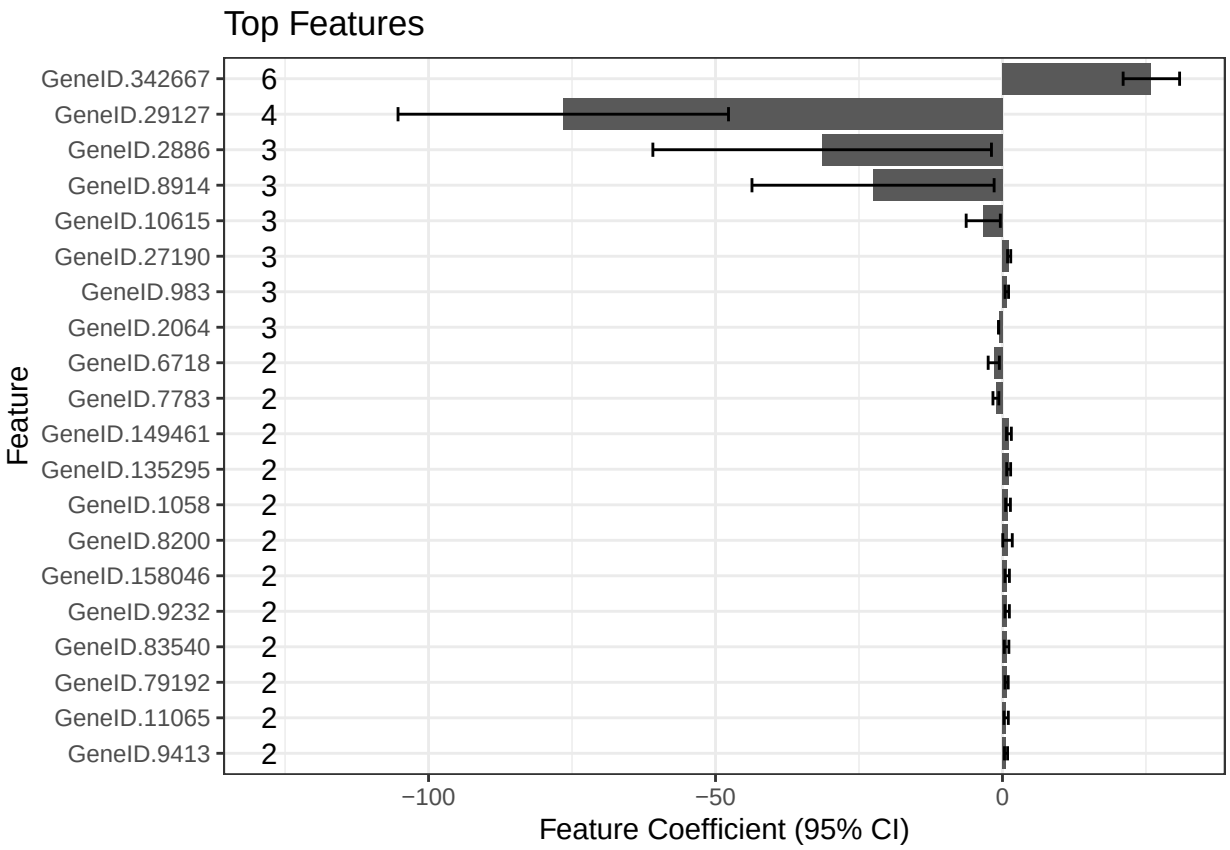
2. Important Features

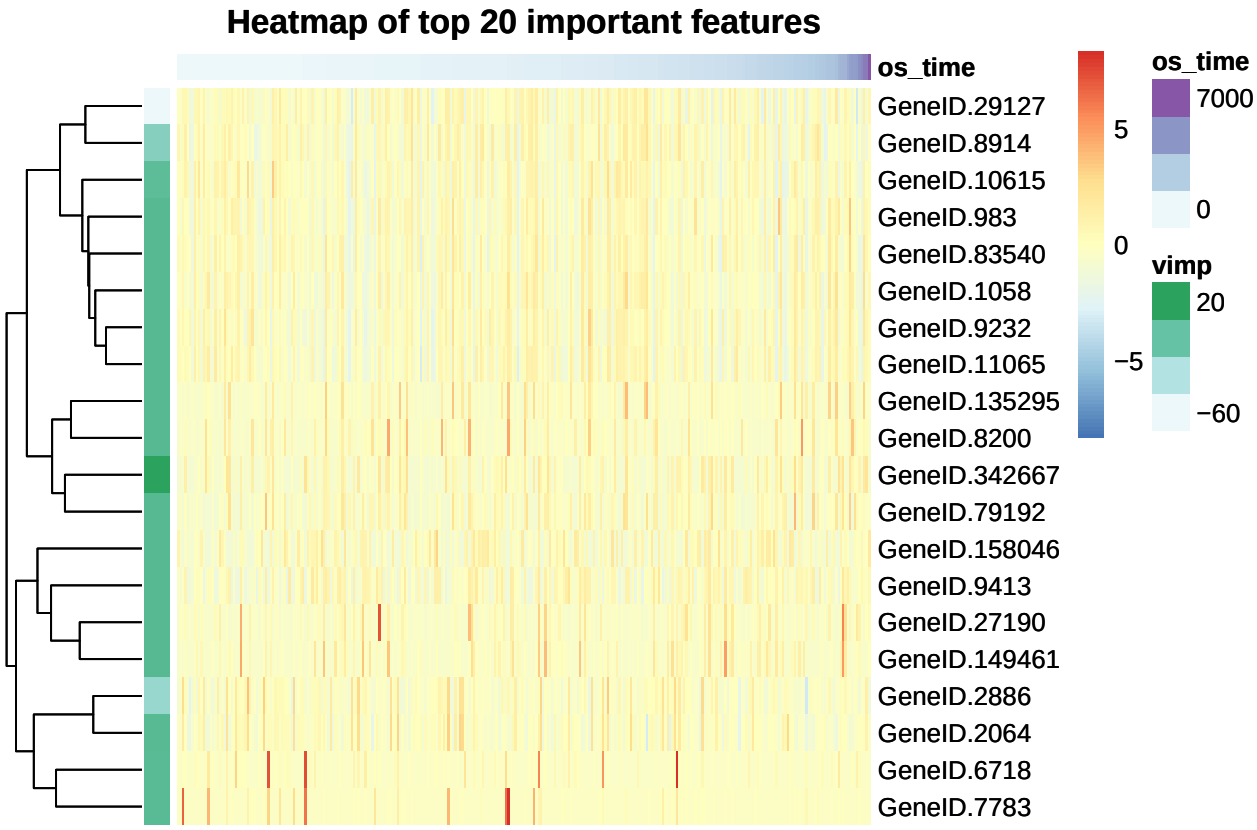





```
## (currently only Lasso has this graph)[1] "there are 100 unique features used from the 100 feature se
## [1] "summary of number of features used in each run under 2 seeds and 3 CVs"

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2.00    2.50    4.00   35.33   76.00  100.00
```


Top Features
GeneID.342667
GeneID.29127
GeneID.2886
GeneID.8914
GeneID.10615
GeneID.27190
GeneID.983
GeneID.2064
GeneID.6718
GeneID.7783
GeneID.149461
GeneID.135295
GeneID.1058
GeneID.8200
GeneID.158046
GeneID.9232
GeneID.83540
GeneID.79192
GeneID.11065
GeneID.9413
6
4
3
3
3
3
3
3
2
2
2
2
2
2
2
2
2
2
2
2
Feature
−100
−50
0
Feature Coefficient (95% CI)


Heatmap of top 20 important features
os_time
GeneID.29127
GeneID.8914
GeneID.10615
GeneID.983
GeneID.83540
GeneID.1058
GeneID.9232
GeneID.11065
GeneID.135295
GeneID.8200
GeneID.342667
GeneID.79192
GeneID.158046
GeneID.9413
GeneID.27190
GeneID.149461
GeneID.2886
GeneID.2064
GeneID.6718
GeneID.7783
5
0
−5
os_time
7000
0
vimp
20
−60