

# Evaluate testing data (survival) - xgboost

*Andrew Chang*

*2018-12-04*

## Contents

0. Load Data . . . . .	1
1. Scores . . . . .	1
2. Important Features . . . . .	4
3. Hyper-parameters . . . . .	8

Label: os\_time

```
## user input
project_home <- "~/EVE/tests"
project_name <- "xgboost_survival_outCV_test"
## end of user input
```

## 0. Load Data

300 of samples were used

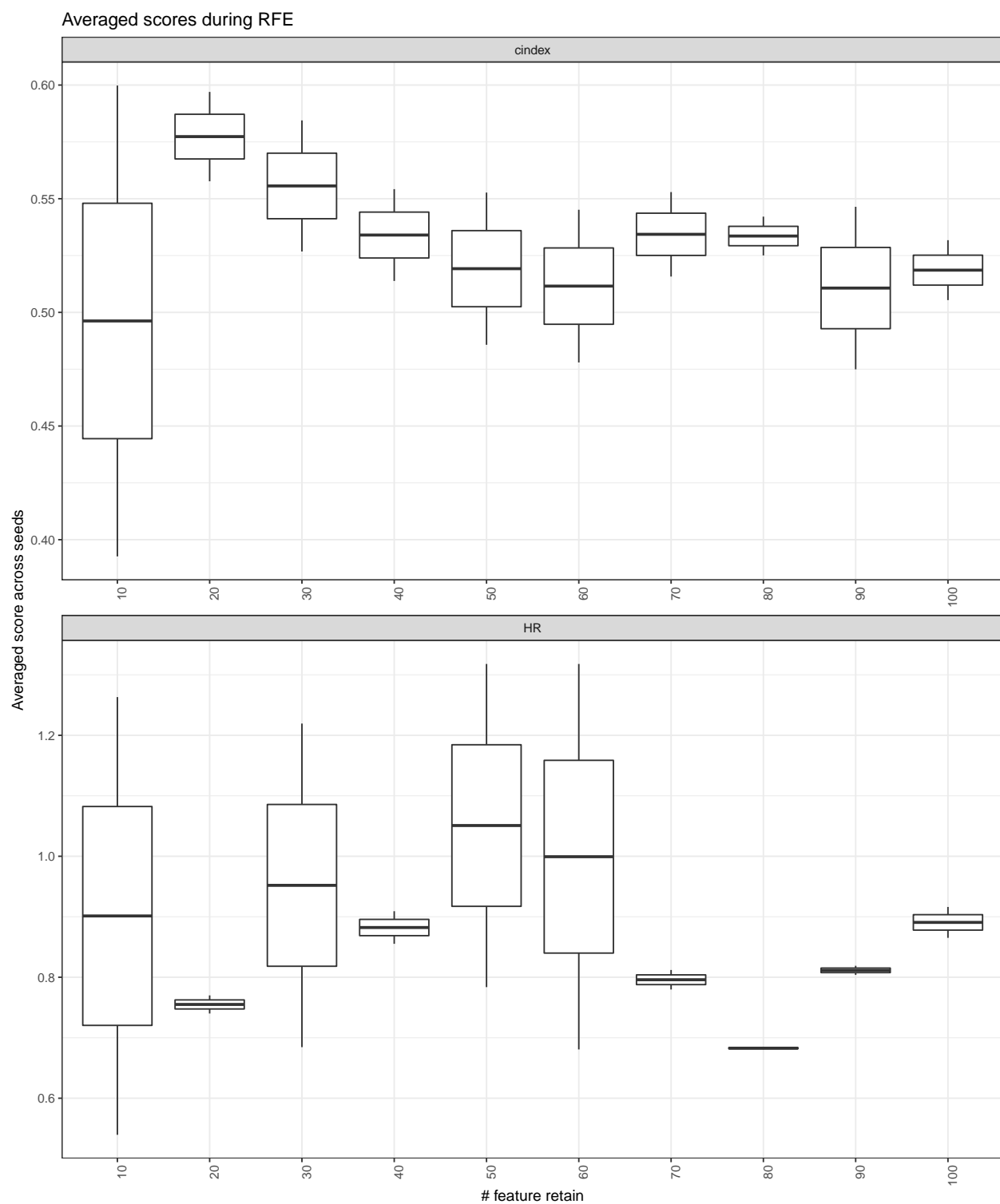
100 of full features

2 runs, each run contains 3 CVs.

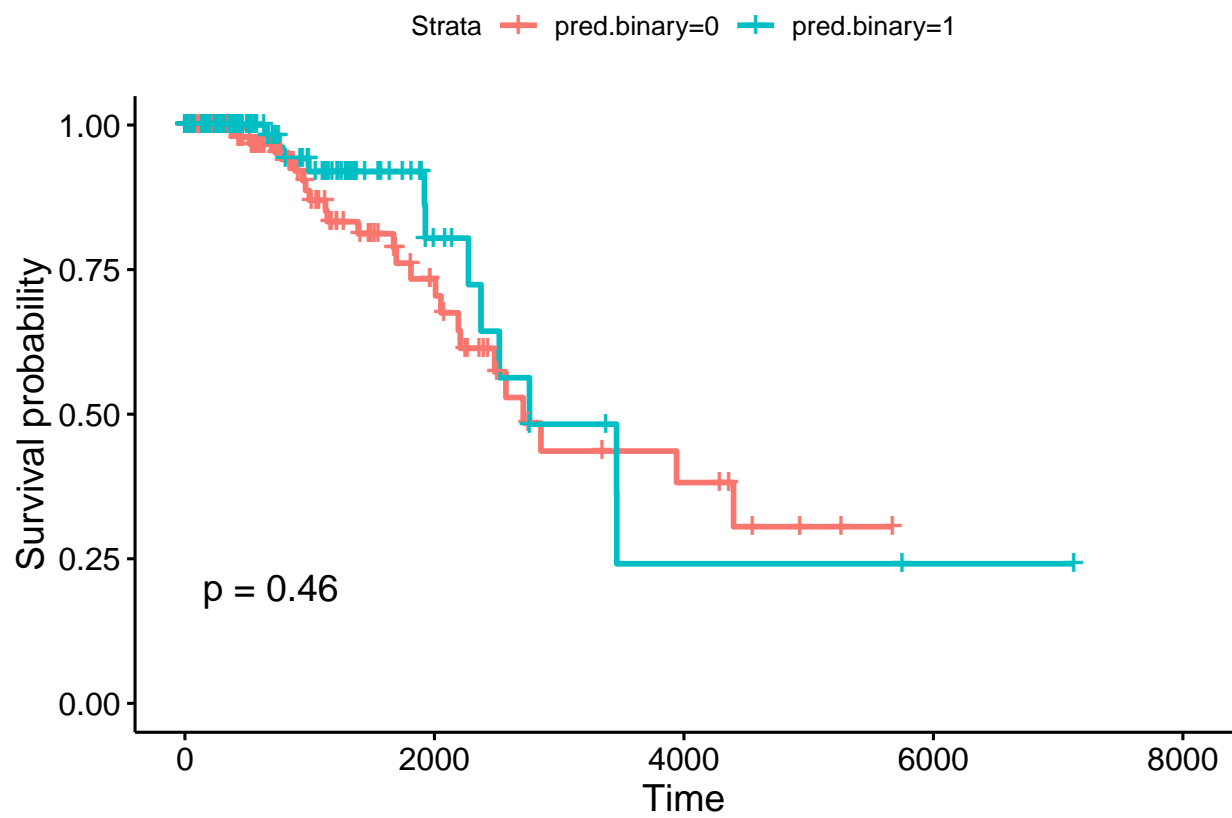
## 1. Scores

**average = T:** scores are based on pre-validation (combine predictions from all CVs per job and then calculate a single score per job).

**average = F:** report all the scores from each CV across entire jobs.

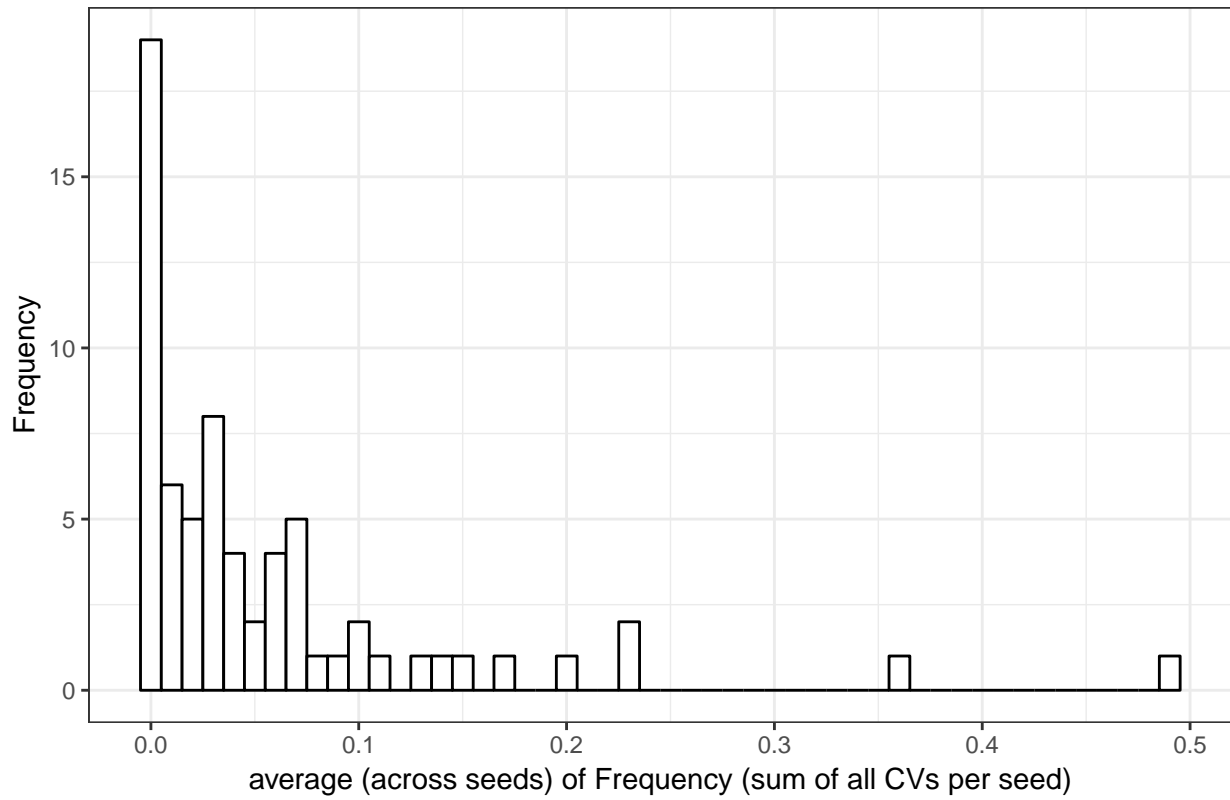


size	metrics	med
20	cindex	0.577
80	HR	0.683

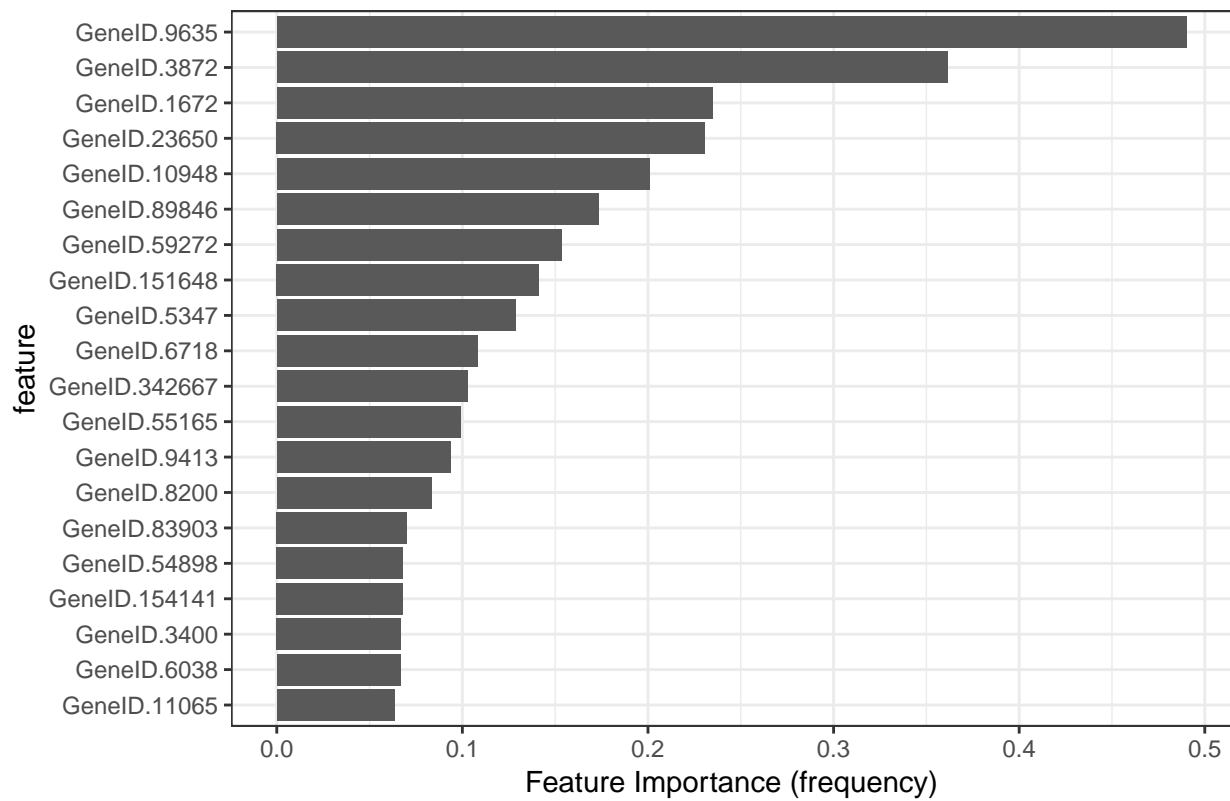


## 2. Important Features

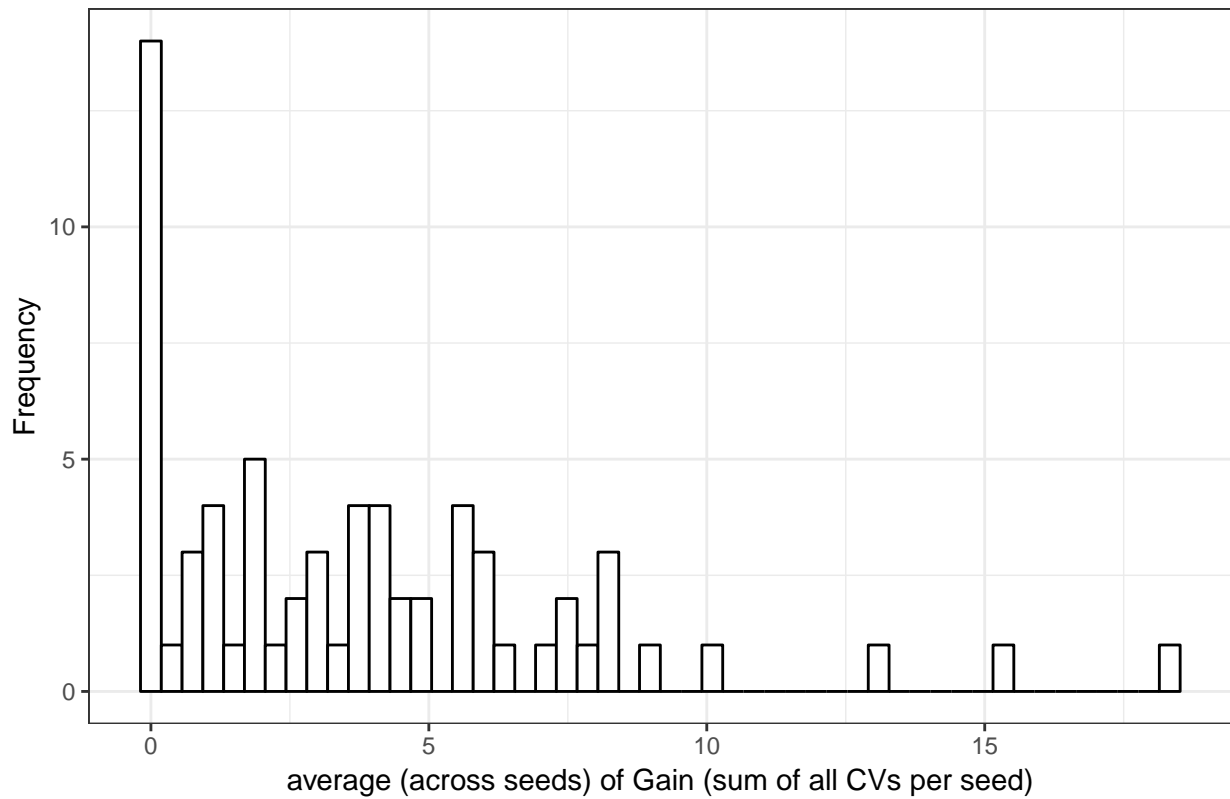
with 20 features based on Frequency



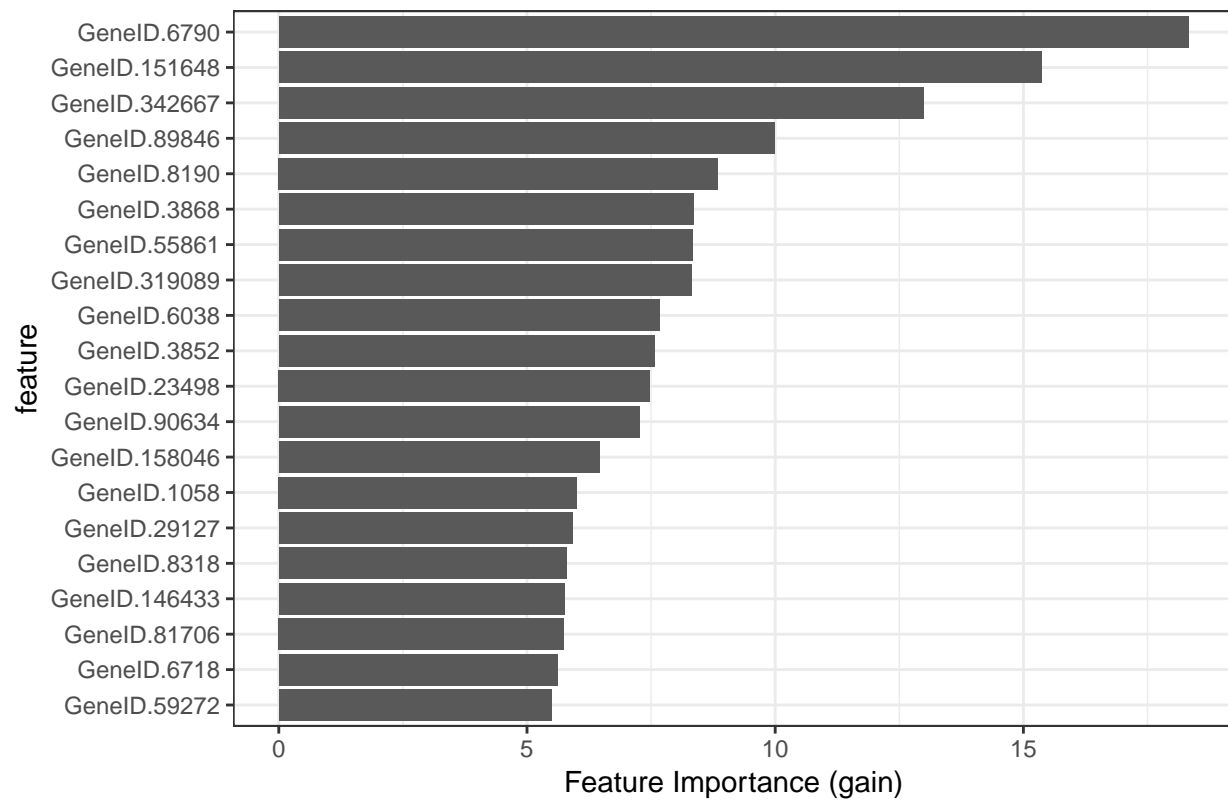
Top 20 features at 20 feature set based on Frequency



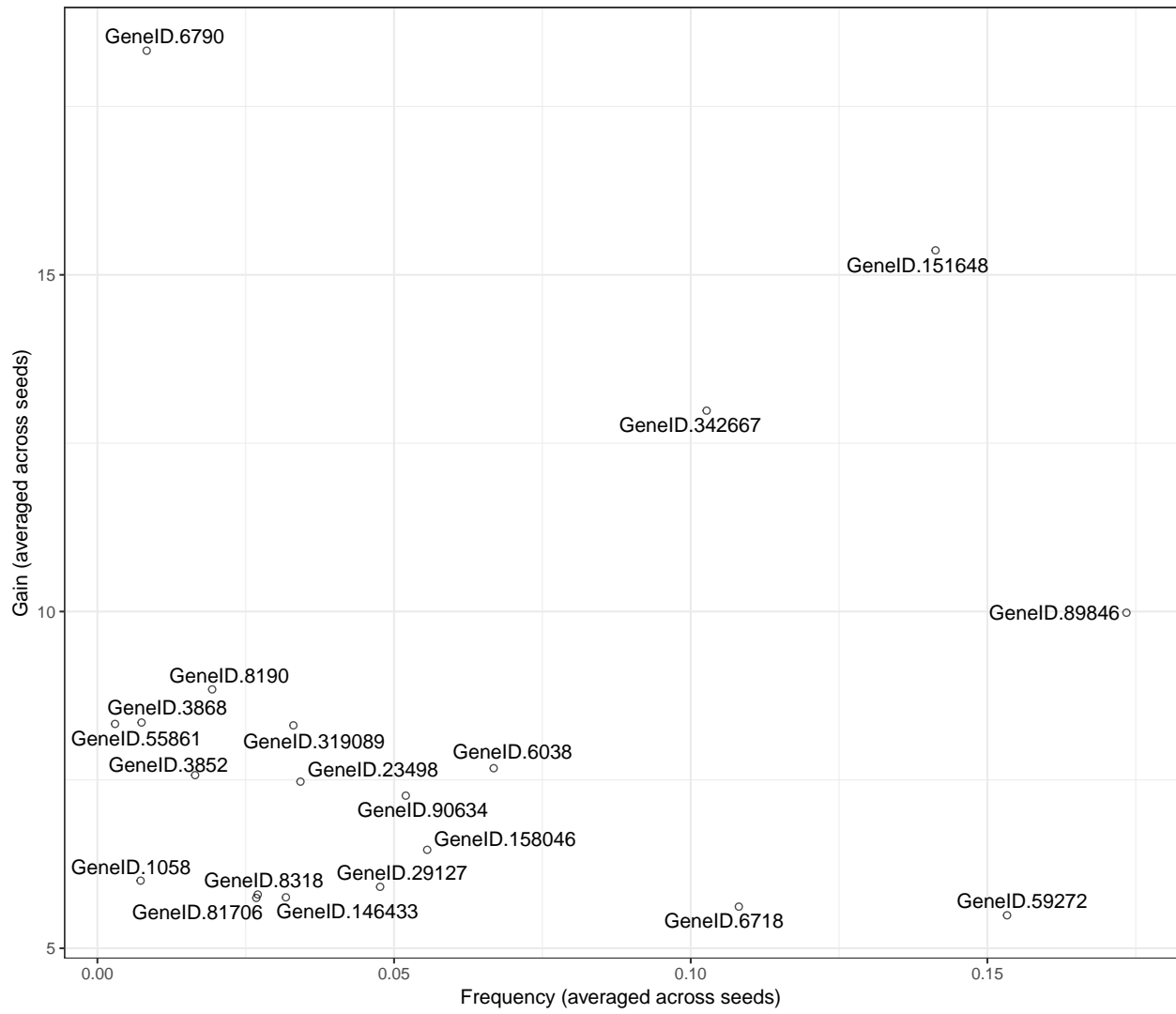
with 20 features based on Gain



Top 20 features at 20 feature set based on Gain



Top 20 features at 20 feature set



### 3. Hyper-parameters

