

Submission 12

Shannon Rush

July 21, 2014

1 caret RF with full training set

Load and Source

```
source("../helpers/predictions.R")
source("../helpers/aws.R")
```

```
library(caret)
```

```
## Loading required package: lattice
## Loading required package: ggplot2
```

```
library(doMC)
```

```
## Loading required package: foreach
## Loading required package: iterators
## Loading required package: parallel
```

```
train <- read.csv("../data/original/training.csv")
test <- read.csv("../data/original/test.csv")
```

AWS

```
PutObject(key="train.RData", body="/Users/shannonmrush/git/HiggsBoson/data/train.RData")
```

Random Forests

Completed in AWS

```
predictors <- train[,setdiff(names(train),c("EventId","Label","Weight"))]
set.seed(111)
registerDoMC(cores=6)
rf1.fit <- train(x=predictors, y=train$Label, method="rf",trControl=trainControl(verboseIter=T))
```

```
download.file("https://s3.amazonaws.com/cloudier/rf1fit.RData", destfile="RData/rf1fit.RData",
              method="curl")
load("RData/rf1fit.RData")
rf1.fit
```

```
pred <- predict(rf1.fit, test, type="prob")
pred.final <- PrepPrediction(pred, test)
WriteSubmission(pred.final, 12)
```

Results

2.88717