```
library(caret)

# Test all descriptors
swp = t(sapply(colnames(datr),function(nm){
  c(
    h=wilcox.test(datr[which(cat==3),nm],datr[which(cat==1),nm])$p.value,
    b=wilcox.test(datr[which(cat==2),nm],datr[which(cat==1),nm])$p.value
  )
}))

# Keep relevant descriptors
pt = 0.0125 #p-value threshold
hbp = rownames(swp)[which(swp[,'b']>pt & swp[,'h']<pt)]

# Create binary category
cat2 = cat;cat2[which(cat2==2)]=1

# Supersampling for unbalanced data
sm = SMOTE(cat ~ ., data  = data.frame(cat=as.factor(cat2),datr[,hbp]))

# Create random forest model
set.seed(123)
mod = train(cat~.,data=sm, method='rf',classwt=c(1,5))$finalModel

# Predict original dataset to get confusion matrix
predv = predict(mod,datr[,hbp]);conf=confusionMatrix(data = predv, reference
= as.factor(cat2))
```