HW06_RF

Omkar Sinha

11/11/2021

```
## Load the data




rm(list=ls())
file<-file.choose()


bc_RF<-  read.csv(file,na.strings = "?",colClasses=c("Sample"="character",
                                          "F1"="factor","F2"="factor","F3"="factor",
                                          "F4"="factor","F5"="factor","F6"="character",
                                          "F7"="factor","F8"="factor","F9"="factor",
                                          "Class"="factor"))
summary(bc_RF$F6)
```

```
##    Length     Class      Mode
##       699 character character
```

```
bc_RF[is.na(bc_RF$F6),"F6"]<- "M"

bc<-data.frame(bc_RF[,1:6],F6=as.factor(bc_RF$F6),bc_RF[,8:11])



index<-sort(sample(nrow(bc),round(.30*nrow(bc))))
training<-bc [-index,]
test<-bc[index,]


##installing required package
library(randomForest)
```

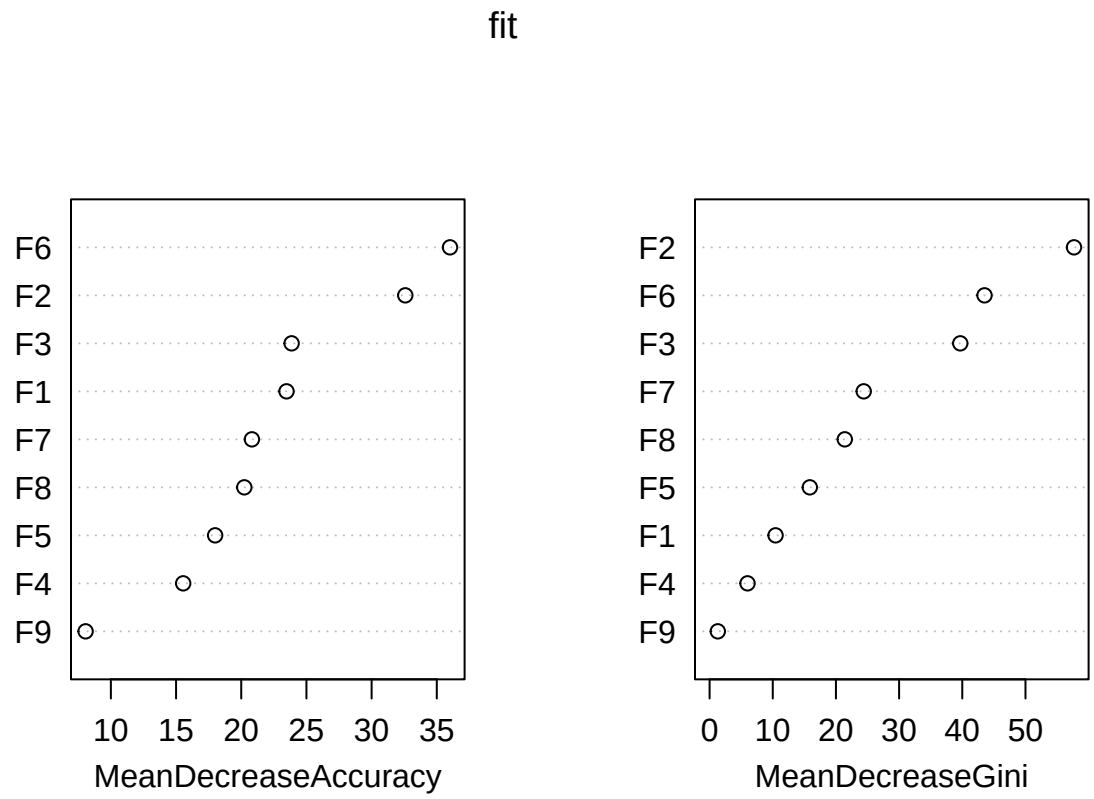
```
## randomForest 4.6-14
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
fit <- randomForest( Class~., data=training[,-1], importance=TRUE, ntree=1000)
importance(fit)
```

```
##            2         4 MeanDecreaseAccuracy MeanDecreaseGini
## F1 19.947747 17.953128            23.472372        10.444093
## F2 24.135407 22.568170            32.582881        57.685158
## F3 13.889390 18.558881            23.862805        39.681465
## F4 12.719085 10.356133            15.546928         6.005639
## F5 16.833059  8.086317            17.996258        15.866297
## F6 25.419332 31.639872            36.014410        43.518595
## F7 15.634724 16.658068            20.818281        24.399016
## F8 18.446309 11.551158            20.234546        21.399848
## F9  7.614888  3.158479             8.059551         1.301521
```

```
varImpPlot(fit)
```




```
dev.off()
```

```
## null device
##           1
```

```
Prediction <- predict(fit, test[,-1])
table(actual=test$Class,Prediction)
```

```
##       Prediction
## actual   2   4
##      2 134   4
##      4   1  71
```

```
wrong<- (test$Class!=Prediction )
error_rate<-sum(wrong)/length(wrong)
error_rate
```

```
## [1] 0.02380952
```