**Random forest in R**

> data("iris")

> attach(iris)

> str(iris)

Data Partition

> set.seed(123)

> ind<-sample(2,nrow(iris),replace=T,prob = c(0.8,0.2))

> train<-iris[ind==1,]

> test<-iris[ind==2,]

Random Forest

> library(randomForest)

> set.seed(222)

> rf<-randomForest(Species~.,data = train)

> print(rf)

> attributes(rf)

Prediction and Confusion Matrix on Training Data

> rf$confusion

> library(caret)

> p1<-predict(rf,train)

> head(p1)

> head(train$Species)

> library(caret)

> install.packages("e1071")

> confusionMatrix(p1,train$Species)

> p2<-predict(rf,test)

Prediction with test data

> confusionMatrix(p2,test$Species)

Error rate of Random Forest

> plot(rf)

> t<-tuneRF(train[,-5],train[,5],stepFactor = 1,plot = TRUE,ntreeTry = 350,trace = TRUE,improve = 0.05)

> t<-tuneRF(train[,-5],train[,5],stepFactor = 0.5,plot = TRUE,ntreeTry = 350,trace = TRUE,improve = 0.05)

> rf<-randomForest(Species~.,data = train,ntree=300,mtry=2,importance=TRUE,proximity=TRUE)

> print(rf)

No of nodes in the trees

> hist(treesize(rf),main="no. of nodes for the trees",col = "red")

Variable Importance

> varImpPlot(rf)

> varImpPlot(rf,sort=T,n.var = 3,main="top 3")

> importance(rf)

Variable Used no of times

> varUsed(rf)

Partial Dependence plot

> partialPlot(rf,train,Petal.Length,"setosa")

Extract Single Tree

> getTree(rf,1,labelVar = TRUE)