install.packages("pROC")

install.packages("mlbench")

install.packages("caret")

library(caret)

library(pROC)

library(mlbench)

Type 'citation("pROC")' for a citation.

Attaching package: ‘pROC’

The following objects are masked from ‘package:stats’:

cov, smooth, var

zoo <- read.csv(file.choose()) #read the file zoo.csv

str(zoo)

data.frame': 101 obs. of 18 variables:

$ animal.name: Factor w/ 100 levels "aardvark","antelope",..: 1 2 3 4 5 6 7 8 9 10 ...

$ hair : int 1 1 0 1 1 1 1 0 0 1 ...

$ feathers : int 0 0 0 0 0 0 0 0 0 0 ...

$ eggs : int 0 0 1 0 0 0 0 1 1 0 ...

$ milk : int 1 1 0 1 1 1 1 0 0 1 ...

$ airborne : int 0 0 0 0 0 0 0 0 0 0 ...

$ aquatic : int 0 0 1 0 0 0 0 1 1 0 ...

$ predator : int 1 0 1 1 1 0 0 0 1 0 ...

$ toothed : int 1 1 1 1 1 1 1 1 1 1 ...

$ backbone : int 1 1 1 1 1 1 1 1 1 1 ...

$ breathes : int 1 1 0 1 1 1 1 0 0 1 ...

$ venomous : int 0 0 0 0 0 0 0 0 0 0 ...

$ fins : int 0 0 1 0 0 0 0 1 1 0 ...

$ legs : int 4 4 0 4 4 4 4 0 0 4 ...

$ tail : int 0 1 1 0 1 1 1 1 1 0 ...

$ domestic : int 0 0 0 0 0 0 1 1 0 1 ...

$ catsize : int 1 1 0 1 1 1 1 0 0 0 ...

$ type : int 1 1 4 1 1 1 1 4 4 1 ...

zoo1 <- zoo[,2:18]

str(zoo1)

zoo1$hair <- as.factor(zoo1$hair)

zoo1$feathers <- as.factor(zoo1$feathers)

zoo1$eggs <- as.factor(zoo1$eggs)

zoo1$milk <- as.factor(zoo1$milk)

zoo1$airborne <- as.factor(zoo1$airborne)

zoo1$aquatic <- as.factor(zoo1$aquatic)

zoo1$predator <- as.factor(zoo1$predator)

zoo1$toothed <- as.factor(zoo1$toothed)

zoo1$backbone <- as.factor(zoo1$backbone)

zoo1$breathes <- as.factor(zoo1$breathes)

zoo1$venomous <- as.factor(zoo1$venomous)

zoo1$fins <- as.factor(zoo1$fins)

zoo1$legs <- as.factor(zoo1$legs)

zoo1$tail <- as.factor(zoo1$tail)

zoo1$domestic <- as.factor(zoo1$domestic)

zoo1$catsize <- as.factor(zoo1$catsize)

zoo1$type <- as.factor(zoo1$type)

# Data partition

set.seed(123)

ind <- sample(2,nrow(zoo1), replace = T, prob = c(0.7,0.3))

train <- zoo1[ind==1,]

test <- zoo1[ind==2,]

# KNN Model

trcontrol <- trainControl(method = "repeatedcv", number = 10,repeats = 3

# classprobs are needed when u want to select ROC for optimal K Value

)

set.seed(222)

fit <- train(type ~., data = train, method = 'knn', tuneLength = 20,

trControl = trcontrol, preProc = c("center","scale"))

fit

plot(fit)

varImp(fit)

pred <- predict(fit, newdata = test )

confusionMatrix(pred, test$type)