#install.packages("caret")

library("caret")

## Importing the dataset

health <- read.csv("/Users/lakshmi/Downloads/healthdataset")

## Reassigning the dependent variable into numeric

health2 <- health;

health2[health2$RiskLevel == "high risk","RiskLevel"] <- 2

health2[health2$RiskLevel == "low risk","RiskLevel"] <- 0

health2[health2$RiskLevel == "mid risk","RiskLevel"] <- 1

head(health2)

str(health2)

#health2$RiskLevel<- as.numeric(as.character(health2$RiskLevel))

health2$RiskLevel<- as.factor(health2$RiskLevel)

## Preprocessing the column name Age

colnames(health2)[1]<-("Age")

str(health2)

head(health2)

health2 <- health2[ -c(3) ]

head(health2)

library(caret)

set.seed(100)

trctrl <- trainControl(method = "cv", number = 10)

nb\_fit <- train(factor(RiskLevel) ~., data = health2, method= "nb", trControl=trctrl)

nb\_fit

set.seed(100)

trctrl <- trainControl(method = "cv", number = 10)

nb\_fit <- train(factor(RiskLevel) ~., data = health2, method= "knn", trControl=trctrl)

nb\_fit

#install.packages("ISLR")

#library(ISLR)

#library(C50)

set.seed(100)

trctrl <- trainControl(method = "cv", number = 10)

nb\_fit <- train(factor(RiskLevel) ~., data = health2, method= "rpart", trControl=trctrl)

nb\_fit

Result :

Naive Bayes

1014 samples

6 predictor

3 classes: '0', '1', '2'

No pre-processing

Resampling: Cross-Validated (10 fold)

Summary of sample sizes: 912, 911, 913, 913, 914, 913, ...

Resampling results across tuning parameters:

usekernel Accuracy Kappa

FALSE 0.5995096 0.3716426

TRUE 0.6774970 0.4985567

Tuning parameter 'fL' was held constant at a value of 0

Tuning parameter 'adjust' was

held constant at a value of 1

Accuracy was used to select the optimal model using the largest value.

The final values used for the model were fL = 0, usekernel = TRUE and adjust = 1.

k-Nearest Neighbors

1014 samples

6 predictor

3 classes: '0', '1', '2'

No pre-processing

Resampling: Cross-Validated (10 fold)

Summary of sample sizes: 912, 911, 913, 913, 914, 913, ...

Resampling results across tuning parameters:

k Accuracy Kappa

5 0.6814488 0.5141716

7 0.6618204 0.4839999

9 0.6282046 0.4303108

1014 samples

6 predictor

3 classes: '0', '1', '2'

No pre-processing

Resampling: Cross-Validated (10 fold)

Summary of sample sizes: 912, 911, 913, 913, 914, 913, ...

Resampling results across tuning parameters:

cp Accuracy Kappa

0.03947368 0.6252429 0.4245073

0.06578947 0.5995772 0.3696273

0.31250000 0.5040374 0.1931959

Accuracy was used to select the optimal model using the largest value.

The final value used for the model was cp = 0.03947368.

>