**R-Code for Naïve Bayes**

**########## *Using the partially cleansed dataset from KNN* ###########################**

naive\_train<-new\_data2

**############ *Filtering out the missing values and N/A’s* ###############################**

naive\_train<-naive\_train[complete.cases(naive\_train),]

naive\_train<-na.omit(naive\_train)

**######## *Adding a data label for Departure Delay* ####################################**

naive\_train$Label\_Departure\_Delay<-ifelse(naive\_train$DepDelay>10,"Yes","No" )

**####### *Repeating the above steps for test dataset* ##################################**

naive\_test<-test\_subset2

naive\_test<-naive\_test[complete.cases(naive\_test),]

naive\_test<-na.omit(naive\_test)

str(naive\_test)

**################## *Adding Delay Label to testing data* ###########################**

naive\_test$Label\_Departure\_Delay<-ifelse(naive\_test$DepDelay>10,"Yes","No")

head(naive\_test)

**####### *Converting the training data variables into factor values* ##################**

naive\_train$Month<-as.factor(naive\_train$Month)

naive\_train$DayofMonth<-as.factor(naive\_train$DayofMonth)

naive\_train$DayOfWeek<-as.factor(naive\_train$DayOfWeek)

str(naive\_train)

**######### *Converting the testing data variables into factor values* ################**

naive\_test$DayOfWeek<-as.factor(naive\_test$DayOfWeek)

naive\_test$DayofMonth<-as.factor(naive\_test$DayofMonth)

naive\_test$Month<-as.factor(naive\_test$Month)

str(naive\_test)

**######### *Applying Naïve Bayes model to training dataset* #####################**

naive\_model <- naiveBayes(Label\_Departure\_Delay ~ ., data=naive\_train)

naive\_preds<-predict(naive\_model,naive\_test)

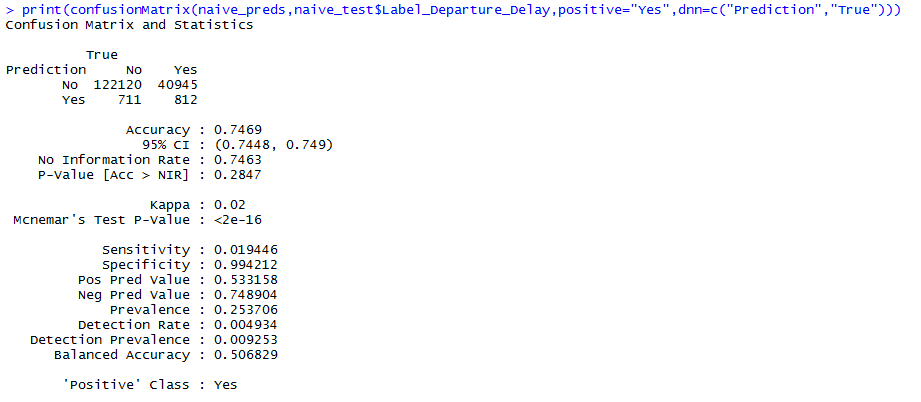
**############ *Install “caret” package, required to creating confusion matrix* #########**

Install.packages(“caret”)

Library(caret)

**####### *Creating confusion matrix for actual labels and prediction labels* ###########**

print(confusionMatrix(naive\_preds, naive\_test$Label\_Departure\_Delay,positive="Yes",dnn=c("Prediction","True")))



**Accuracy = 74.690743%**