**Practical No.5: Implementing Classification in R (Naïve Bayes)**

|  |
| --- |
| **Naïve Bayes with R** |
| Naive Bayes in R can be implemented using different packages. Here, we are using the ‘e1071’  package to implement Naive Bayes on a weather dataset. The e1071 package contains the Naive  Bayes function. It allows numeric and factor variables to be used in the Naive Bayes model. If the  package ‘e1071’ is not installed, then first install the package using the following command.  **> install.packages(“e1071”) ##Installation of package “e1071”**  After installation, load the package using the ‘***library***’ command given as follows.  **> library(“e1071”) ##Load the package “e1071”** |
| Then, the ‘**Weather**’ dataset may be employed to illustrate how the Naive Bayes classification  can be performed in R. After loading, convert the dataset into a data frame  **> library(“e1071”) ##Load the package “e1071”**  **> data<-read.csv(“F:\\dataset\\weather.csv”,header=TRUE) ##Load the .csv file of dataset**  **> weather \_ df=as.data.frame(data) #Conversion of dataset into data frame**  The following command displays the instances of dataset as shown in Fig. 6.42.  **> weather \_ df #To display instances of weather dataset** |



|  |
| --- |
| > Naive \_ Bayes \_ Model=naiveBayes(play ~., data=weather \_ df)  > Naive \_ Bayes \_ Model |

After applying Naive Bayes, a confusion matrix can be created which helps in summarizing the

results of testing the algorithm for further inspection using the following commands. The confusion

matrix is shown in Figure

|  |
| --- |
| > NB \_ Predictions=predict(Naive \_ Bayes \_ Model,weather \_ df)  > table(NB \_ Predictions,weather \_ df$play, dnn=c(“Prediction”,”Actual”)) |



For the weather dataset, we have a total of 14 instances. Out of which, there are 9 instances that

specify that play will be held in the given weather conditions and 5 instances that specify that play

will not be held. Here, the confusion matrix tells us that Naive Bayes properly classified 9 instances

for play to be held and 4 instances for play not to be held. However, it classified 1 instance of play

into the ‘Yes’ class that actually belongs to the ‘No’ play class.

The Naive Bayes function includes the Laplace parameter. Whatever positive integer this is set

to will be added into for every class. In this example, we set the Laplace parameter to 1 as given in

the following commands.

|  |
| --- |
| > nb \_ laplace1 <- naiveBayes(play~., data=weather \_ df, laplace=1)  > laplace1 \_ pred <- predict(nb \_ laplace1, weather \_ df, type=”class”)  > table(laplace1 \_ pred, weather \_ df$play,dnn=c(“Prediction”,”Actual”)) |