

“**Experiment 1.6”**

**Student Name: SUMIT KUMAR UID: 20BCS8226**

**Branch: CSE Section/Group: DM-720 A**

**Subject Name: Data Mining Lab Subject Code: 20CSP-376**

**Aim:**

To perform classification using Bayesian classification algorithm using R.

**Theory:**

Naive Bayes algorithm is based on Bayes theorem. Bayes theorem gives the conditional probability of an event A given another event B has occurred.

where,

P(A|B) = Conditional probability of A given B. P(B|A) = Conditional probability of B given A.

P(A) = Probability of event A.

P(B) = Probability of event B.

For many predictors, we can formulate the posterior probability as follows:

P(A|B) = P(B1|A) \* P(B2|A) \* P(B3|A) \* P(B4|A) …

**Code:**

# Loading data data(iris)

#

Structure str(iris)

# Installing Packages



install.packages("e1071")

install.packages("caTools") install.packages("caret")

# Loading package library(e1071) library(caTools) library(caret)

# Splitting data into train # and test data split <- sample.split(iris, SplitRatio = 0.7) train\_cl <- subset(iris, split == "TRUE") test\_cl <- subset(iris, split == "FALSE")

# Feature Scaling

train\_scale <- scale(train\_cl[, 1:4]) test\_scale <- scale(test\_cl[, 1:4])

# Fitting Naive Bayes Model # to training dataset set.seed(120) # Setting Seed classifier\_cl <- naiveBayes(Species ~ ., data = train\_cl) classifier\_cl

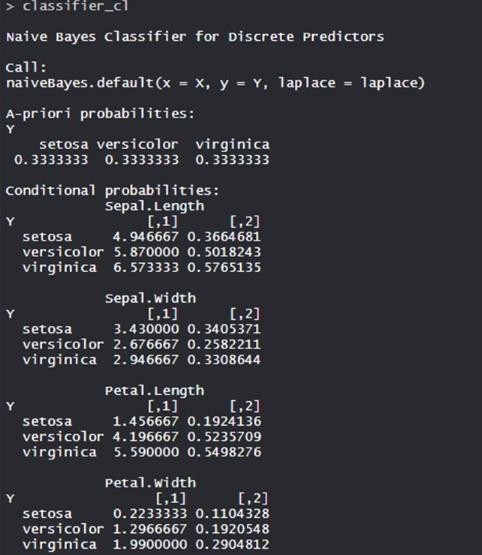
# Predicting on test data' y\_pred <- predict(classifier\_cl, newdata = test\_cl) #



Confusion Matrix cm <- table (test\_cl$Species , y\_pred) cm

# Model Evaluation confusionMatrix(cm)

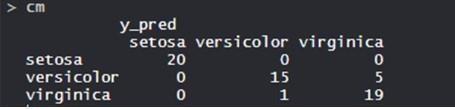
**OUTPUT:**





The Conditional probability for each feature or variable is created by model separately. The apriori probabilities are also calculated which indicates the distribution of our data.

1. Confusion Matrix:



So, 20 Setosa are correctly classified as Setosa. Out of 16 Versicolor, 15 Versicolor are correctly classified as Versicolor, and 1 are classified as virginica. Out of 24 virginica, 19 virginica are correctly classified as virginica and 5 are classified as Versicolor.

Model Evaluation

