Experiment-2.2

Student Name: Aayush Gurung **UID:20BCS5323**
Branch: CSE **Section/Group:DM_607(A)**
Semester: 6 **Date of Performance:06-04-2023**
Subject Name: DM LAB **Subject Code: 20CST-376**

1. Aim:

To perform the classification using Bayesian classification algorithm using R.

2. Code:

```
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)
```

3. Output:

```
Naive Bayes Classifier for Discrete Predictors

Call:
naiveBayes.default(x = X, y = Y, laplace = laplace)

A-priori probabilities:
Y
    setosa versicolor  virginica
 0.3333333  0.3333333  0.3333333

Conditional probabilities:
            Sepal.Length
Y                [,1]      [,2]
  setosa     4.973333 0.3084257
  versicolor 5.966667 0.4929386
  virginica  6.520000 0.6764002

            Sepal.Width
Y                [,1]      [,2]
  setosa     3.426667 0.3561609
  versicolor 2.776667 0.2712466
  virginica  2.976667 0.3607304

            Petal.Length
Y                [,1]      [,2]
  setosa     1.453333 0.1775957
  versicolor 4.243333 0.4328600
  virginica  5.496667 0.5505379

            Petal.Width
Y                 [,1]       [,2]
  setosa     0.2333333 0.09222661
  versicolor 1.3233333 0.19419743
  virginica  1.9900000 0.27586853
```

```
> cm
            y_pred
             setosa versicolor virginica
  setosa         20          0         0
  versicolor      0         19         1
  virginica       0          1        19
>
>
>
> # Model Evaluation
>
> confusionMatrix(cm)
Confusion Matrix and Statistics

            y_pred
             setosa versicolor virginica
  setosa         20          0         0
  versicolor      0         19         1
  virginica       0          1        19

Overall Statistics

               Accuracy : 0.9667
                 95% CI : (0.8847, 0.9959)
    No Information Rate : 0.3333
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.95
```

```
 Mcnemar's Test P-Value : NA

Statistics by Class:

                     Class: setosa Class: versicolor
Sensitivity                 1.0000            0.9500
Specificity                 1.0000            0.9750
Pos Pred Value              1.0000            0.9500
Neg Pred Value              1.0000            0.9750
Prevalence                  0.3333            0.3333
Detection Rate              0.3333            0.3167
Detection Prevalence        0.3333            0.3333
Balanced Accuracy           1.0000            0.9625
                     Class: virginica
Sensitivity                    0.9500
Specificity                    0.9750
Pos Pred Value                 0.9500
Neg Pred Value                 0.9750
Prevalence                     0.3333
Detection Rate                 0.3167
Detection Prevalence           0.3333
Balanced Accuracy              0.9625
> 
```