

Naives Bayes Classification

Amin Fesharaki

11/14/2021

```
fram_train <- read.csv(file = "/Users/datascience/Desktop/ADS 502 Data Sets/Website Data Sets/framingham")
fram_test <- read.csv(file = "/Users/datascience/Desktop/ADS 502 Data Sets/Website Data Sets/framingham")
```

31. Run the Naïve Bayes classifier to classify persons as living or dead based on sex and education.

```
nb01 <- naiveBayes(formula = Death ~ Sex + Educ, data = fram_train)
nb01
```

```
##
## Naive Bayes Classifier for Discrete Predictors
##
## Call:
## naiveBayes.default(x = X, y = Y, laplace = laplace)
##
## A-priori probabilities:
## Y
##      0      1
## 0.45 0.55
##
## Conditional probabilities:
##      Sex
## Y      [,1]      [,2]
## 0 1.591111 0.4921759
## 1 1.440000 0.4968388
##
##      Educ
## Y      [,1]      [,2]
## 0 2.011111 0.9954735
## 1 1.798182 0.9886406
```

32. Evaluate the Naïve Bayes model on the framingham_nb_test dataset. Display the results in a contingency table. Edit the row and column names of the table to make the table more readable. Include a total row and column.

```
ypred <- predict(object = nb01, newdata = fram_test)

c.pred <- table(fram_test$Death, ypred)
rownames(c.pred) <- c("Actual: Alive", "Actual: Dead")
colnames(c.pred) <- c("Predicted: Alive", "Predicted: Dead")
addmargins(A = c.pred, FUN = list(Total=sum), quiet = TRUE)
```

```
##
##      ypred
##      Predicted: Alive Predicted: Dead Total
## Actual: Alive      167      283    450
## Actual: Dead      122      428    550
```

```
##      Total                289                711   1000
```

33. According to your table in the previous exercise, find the following values for the Naïve Bayes model:
(A) Accuracy, (B) Error Rate

```
# Assigning General Form of Table to matrix values
TN <- c.pred[1,1]
FN <- c.pred[2,1]
FP <- c.pred[1,2]
TP <- c.pred[2,2]

# Accuracy
Acc <- (TN + TP) / (TN + FN + FP + TP)
# Error Rate
Error <- 1 - Acc

cat ("MODEL 2", "\nAccuracy = ", Acc, "\nError =", Error)
```

```
## MODEL 2
## Accuracy = 0.595
## Error = 0.405
```

34. According to your contingency table, find the following values for the Naïve Bayesmodel: (A) How often it correctly classifies dead person (B) How often it correctly classifies living persons
- (a) The model correctly classifies a dead person 77% (428/550) of the time.
 - (b) The model correctly classifies an alive person 37% (167/450) of the time.