

Assignment 3

Akhil Seliveri

2023-10-16

Problem Statement

The file `accidentsFull.csv` contains information on 42,183 actual automobile accidents in 2001 in the United States that involved one of three levels of injury: NO INJURY, INJURY, or FATALITY. For each accident, additional information is recorded, such as day of week, weather conditions, and road type. A firm might be interested in developing a system for quickly classifying the severity of an accident based on initial reports and associated data in the system (some of which rely on GPS-assisted reporting).

Our goal here is to predict whether an accident just reported will involve an injury ($\text{MAX_SEV_IR} = 1$ or 2) or will not ($\text{MAX_SEV_IR} = 0$). For this purpose, create a dummy variable called INJURY that takes the value “yes” if $\text{MAX_SEV_IR} = 1$ or 2 , and otherwise “no.”

1. Using the information in this dataset, if an accident has just been reported and no further information is available, what should the prediction be? (INJURY = Yes or No?) Why?
2. Select the first 24 records in the dataset and look only at the response (INJURY) and the two predictors WEATHER_R and TRAF_CON_R. Create a pivot table that examines INJURY as a function of the two predictors for these 12 records. Use all three variables in the pivot table as rows/columns. 1. Compute the exact Bayes conditional probabilities of an injury (INJURY = Yes) given the six possible combinations of the predictors. 2. Classify the 24 accidents using these probabilities and a cutoff of 0.5. 3. Compute manually the naive Bayes conditional probability of an injury given WEATHER_R = 1 and TRAF_CON_R = 1. 4. Run a naive Bayes classifier on the 24 records and two predictors. Check the model output to obtain probabilities and classifications for all 24 records. Compare this to the exact Bayes classification. Are the resulting classifications equivalent? Is the ranking (= ordering) of observations equivalent?
3. Let us now return to the entire dataset. Partition the data into training (60%) and validation (40%). 1. Run a naive Bayes classifier on the complete training set with the relevant predictors (and INJURY as the response). Note that all predictors are categorical. Show the confusion matrix. 2. What is the overall error of the validation set?

Summary

The code begins by exploring a dataset named ‘accident’ to understand the distribution of injuries (INJURY) using a table.

A subset of the dataset containing the first 24 records with relevant columns (INJURY, WEATHER_R, TRAF_CON_R) is created and examined.

Pivot tables are created to examine the relationship between injury and weather, traffic conditions for the first 24 records.

Conditional probabilities of injury (Yes/No) based on different combinations of weather and traffic conditions are calculated manually.

Each of the 24 records is classified into “Yes” or “No” for injury using calculated Bayes probabilities and a cutoff of 0.5.

A Naive Bayes classifier is created and used to predict injury for the first 24 records. Predicted probabilities and classifications are compared with manual Bayes probabilities.

The dataset is split into training (60%) and validation (40%) sets. A Naive Bayes classifier is trained using the training set and then used to predict injuries in the validation set. A confusion matrix is generated to evaluate the model's performance on the validation set, calculating the overall error rate.

Data Input and Cleaning

Load the required libraries and read the input file

```
library(e1071)
library(caret)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: lattice
```

```
#Loading the dataset from desktop using read.csv()
accident <- read.csv("/Users/akhilchintu/Downloads/accidentsFull.csv")
#Creating dummy variable "INJURY" with "yes" or "no" based on "MAX_SEV_IR" variable
accident$INJURY = ifelse(accident$MAX_SEV_IR>0,"yes","no")

# Converting all the variables to factor
for (i in c(1:dim(accident)[2])){
  accident[,i] <- as.factor(accident[,i])
}
head(accident,n=24)
```

##	HOUR_I_R	ALCHL_I	ALIGN_I	STRATUM_R	WRK_ZONE	WKDY_I_R	INT_HWY	LGTCN_I_R
## 1	0	2	2	1	0	1	0	3
## 2	1	2	1	0	0	1	1	3
## 3	1	2	1	0	0	1	0	3
## 4	1	2	1	1	0	0	0	3
## 5	1	1	1	0	0	1	0	3
## 6	1	2	1	1	0	1	0	3
## 7	1	2	1	0	0	1	1	3
## 8	1	2	1	1	0	1	0	3
## 9	1	2	1	1	0	1	0	3
## 10	0	2	1	0	0	0	0	3
## 11	1	2	1	0	0	1	0	3
## 12	1	2	1	1	0	1	0	3
## 13	1	2	1	1	0	1	0	3
## 14	1	2	2	0	0	1	0	3
## 15	1	2	2	1	0	1	0	3
## 16	1	2	2	1	0	1	0	3
## 17	1	2	1	1	0	1	0	3
## 18	1	2	1	1	0	0	0	3
## 19	1	2	1	1	0	1	0	3
## 20	1	2	1	0	0	1	0	3
## 21	1	2	1	1	0	1	0	3
## 22	1	2	2	0	0	1	0	3
## 23	1	2	1	0	0	1	0	3
## 24	1	2	1	1	0	1	9	3
##	MANCOL_I_R	PED_ACC_R	RELJCT_I_R	REL_RWY_R	PROFIL_I_R	SPD_LIM	SUR_COND	
## 1	0	0	1	0	1	40	4	
## 2	2	0	1	1	1	70	4	
## 3	2	0	1	1	1	35	4	
## 4	2	0	1	1	1	35	4	
## 5	2	0	0	1	1	25	4	
## 6	0	0	1	0	1	70	4	
## 7	0	0	0	0	1	70	4	
## 8	0	0	0	0	1	35	4	
## 9	0	0	1	0	1	30	4	
## 10	0	0	1	0	1	25	4	
## 11	0	0	0	0	1	55	4	
## 12	2	0	0	1	1	40	4	
## 13	1	0	0	1	1	40	4	
## 14	0	0	0	0	1	25	4	
## 15	0	0	0	0	1	35	4	
## 16	0	0	0	0	1	45	4	
## 17	0	0	0	0	1	20	4	
## 18	0	0	0	0	1	50	4	
## 19	0	0	0	0	1	55	4	
## 20	0	0	1	1	1	55	4	
## 21	0	0	1	0	0	45	4	
## 22	0	0	1	0	0	65	4	
## 23	0	0	0	0	0	65	4	
## 24	2	0	1	1	0	55	4	
##	TRAF_CON_R	TRAF_WAY	VEH_INVL	WEATHER_R	INJURY_CRASH	NO_INJ_I	PRPTYDMG_CRASH	
## 1	0	3	1	1	1	1	0	

## 2	0	3	2	2	0	0	1
## 3	1	2	2	2	0	0	1
## 4	1	2	2	1	0	0	1
## 5	0	2	3	1	0	0	1
## 6	0	2	1	2	1	1	0
## 7	0	2	1	2	0	0	1
## 8	0	1	1	1	1	1	0
## 9	0	1	1	2	0	0	1
## 10	0	1	1	2	0	0	1
## 11	0	1	1	2	0	0	1
## 12	2	1	2	1	0	0	1
## 13	0	1	4	1	1	2	0
## 14	0	1	1	1	0	0	1
## 15	0	1	1	1	1	1	0
## 16	0	1	1	1	1	1	0
## 17	0	1	1	2	0	0	1
## 18	0	1	1	2	0	0	1
## 19	0	1	1	2	0	0	1
## 20	0	1	1	2	0	0	1
## 21	0	3	1	1	1	1	0
## 22	0	3	1	1	0	0	1
## 23	2	2	1	2	1	2	0
## 24	0	2	2	2	1	1	0

##	FATALITIES	MAX_SEV_IR	INJURY
## 1	0	1	yes
## 2	0	0	no
## 3	0	0	no
## 4	0	0	no
## 5	0	0	no
## 6	0	1	yes
## 7	0	0	no
## 8	0	1	yes
## 9	0	0	no
## 10	0	0	no
## 11	0	0	no
## 12	0	0	no
## 13	0	1	yes
## 14	0	0	no
## 15	0	1	yes
## 16	0	1	yes
## 17	0	0	no
## 18	0	0	no
## 19	0	0	no
## 20	0	0	no
## 21	0	1	yes
## 22	0	0	no
## 23	0	1	yes
## 24	0	1	yes

Questions

1. Using the information in this dataset, if an accident has just been reported and no further information is available, what should the prediction be? (INJURY = Yes or No?) Why?

```
#Calculating for number of YES and No values to understand the dataset for INJURY prediction
table(accident$INJURY)
```

```
##
##      no      yes
## 20721 21462
```

If there's no specific accident-related information, we can rely on the dataset's prevailing trend to predict whether an accident caused an injury (INJURY = Yes or No). Analysis of the dataset reveals that "Yes" (INJURY = Yes) is the more common outcome, occurring 21,462 times compared to "No" (INJURY = No) which occurred 20,721 times. Therefore, based on this statistical insight, the likely prediction for INJURY is "Yes." This interpretation assumes that the dataset accurately represents the broader population of accidents and that the observed distribution of injuries provides a reliable estimate for accidents lacking additional details.

2. Select the first 24 records in the dataset and look only at the response (INJURY) and the two predictors WEATHER_R and TRAF_CON_R. Create a pivot table that examines INJURY as a function of the two predictors for these 12 records. Use all three variables in the pivot table as rows/columns. 1. Compute the exact Bayes conditional probabilities of an injury (INJURY = Yes) given the six possible combinations of the predictors. 2. Classify the 24 accidents using these probabilities and a cutoff of 0.5. 3. Compute manually the naive Bayes conditional probability of an injury given WEATHER_R = 1 and TRAF_CON_R = 1. 4. Run a naive Bayes classifier on the 24 records and two predictors. Check the model output to obtain probabilities and classifications for all 24 records. Compare this to the exact Bayes classification. Are the resulting classifications equivalent? Is the ranking (= ordering) of observations equivalent?

```
#Creating dataframe with only first 24 records from the accidents dataset with 3 variables
accident24 <- accident[1:24,c("INJURY","WEATHER_R","TRAF_CON_R")]
head(accident24)
```

```
##      INJURY WEATHER_R TRAF_CON_R
## 1      yes          1           0
## 2      no          2           0
## 3      no          2           1
## 4      no          1           1
## 5      no          1           0
## 6      yes          2           0
```

```
#Creating a pivot table that examines injury as function of 2 predictors
p.1 <- ftable(accident24)
#Pivot table only for conditions
p.2 <- ftable(accident24[,-1])
p.1
```

```
##                TRAF_CON_R 0 1 2
## INJURY WEATHER_R
## no      1                3 1 1
##          2                9 1 0
## yes     1                6 0 0
##          2                2 0 1
```

```
p.2
```

```
##                TRAF_CON_R 0 1 2
## WEATHER_R
## 1                9 1 1
## 2               11 1 1
```

2(1). Compute the exact Bayes conditional probabilities of an injury (INJURY = Yes) given the six possible combinations of the predictors.

```
#Calculating six possible outcome for Injury = YES
p1 = p.1[3,1] / p.2[1,1] # Injury = YES, Weather=1, Traffic=0
p2 = p.1[4,1] / p.2[2,1] # Injury = YES, Weather=2, Traffic=0
p3 = p.1[3,2] / p.2[1,2] # Injury = YES, Weather=1, Traffic=1
p4 = p.1[4,2] / p.2[2,2] # Injury = YES, Weather=2, Traffic=1
p5 = p.1[3,3] / p.2[1,3] # Injury = YES, Weather=1, Traffic=2
p6 = p.1[4,3] / p.2[2,3] # Injury = YES, Weather=2, Traffic=2

#Calculating six possible outcome for Injury = NO
n1 = p.1[1,1] / p.2[1,1] # Injury = NO, Weather=1, Traffic=0
n2 = p.1[2,1] / p.2[2,1] # Injury = NO, Weather=2, Traffic=0
n3 = p.1[1,2] / p.2[1,2] # Injury = NO, Weather=1, Traffic=1
n4 = p.1[2,2] / p.2[2,2] # Injury = NO, Weather=2, Traffic=1
n5 = p.1[1,3] / p.2[1,3] # Injury = NO, Weather=1, Traffic=2
n6 = p.1[2,3] / p.2[2,3] # Injury = NO, Weather=2, Traffic=2
print(c(p1,p2,p3,p4,p5,p6))
```

```
## [1] 0.6666667 0.1818182 0.0000000 0.0000000 0.0000000 1.0000000
```

```
print(c(n1,n2,n3,n4,n5,n6))
```

```
## [1] 0.3333333 0.8181818 1.0000000 1.0000000 1.0000000 0.0000000
```

2(2). Classify the 24 accidents using these probabilities and a cutoff of 0.5.

```
prob.inj <- rep(0,24)

for (i in 1:24) {
  print(c(accident24$WEATHER_R[i],accident24$TRAF_CON_R[i]))
  if (accident24$WEATHER_R[i] == "1") {
    if (accident24$TRAF_CON_R[i]=="0"){
      prob.inj[i] = p1
    }
    else if (accident24$TRAF_CON_R[i]=="1") {
      prob.inj[i] = p3
    }
    else if (accident24$TRAF_CON_R[i]=="2") {
      prob.inj[i] = p5
    }
  }
  else {
    if (accident24$TRAF_CON_R[i]=="0"){
      prob.inj[i] = p2
    }
    else if (accident24$TRAF_CON_R[i]=="1") {
      prob.inj[i] = p4
    }
    else if (accident24$TRAF_CON_R[i]=="2") {
      prob.inj[i] = p6
    }
  }
}
```



```
## [1] 1 0
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
## [1] 2 1
## Levels: 1 2 0
## [1] 1 1
## Levels: 1 2 0
## [1] 1 0
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
## [1] 1 0
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
## [1] 1 2
## Levels: 1 2 0
## [1] 1 0
## Levels: 1 2 0
## [1] 1 0
## Levels: 1 2 0
## [1] 1 0
## Levels: 1 2 0
## [1] 1 0
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
## [1] 1 0
## Levels: 1 2 0
## [1] 1 0
## Levels: 1 2 0
## [1] 2 2
## Levels: 1 2 0
## [1] 2 0
## Levels: 1 2 0
```

```
accident24$prob.inj <- prob.inj
accident24
```

##	INJURY	WEATHER_R	TRAF_CON_R	prob.inj
## 1	yes	1	0	0.6666667
## 2	no	2	0	0.1818182
## 3	no	2	1	0.0000000
## 4	no	1	1	0.0000000
## 5	no	1	0	0.6666667
## 6	yes	2	0	0.1818182
## 7	no	2	0	0.1818182
## 8	yes	1	0	0.6666667
## 9	no	2	0	0.1818182
## 10	no	2	0	0.1818182
## 11	no	2	0	0.1818182
## 12	no	1	2	0.0000000
## 13	yes	1	0	0.6666667
## 14	no	1	0	0.6666667
## 15	yes	1	0	0.6666667
## 16	yes	1	0	0.6666667
## 17	no	2	0	0.1818182
## 18	no	2	0	0.1818182
## 19	no	2	0	0.1818182
## 20	no	2	0	0.1818182
## 21	yes	1	0	0.6666667
## 22	no	1	0	0.6666667
## 23	yes	2	2	1.0000000
## 24	yes	2	0	0.1818182

```
accident24$pred.prob <- ifelse(accident24$prob.inj>0.5, "yes", "no")
accident24
```

##	INJURY	WEATHER_R	TRAF_CON_R	prob.inj	pred.prob
## 1	yes	1	0	0.6666667	yes
## 2	no	2	0	0.1818182	no
## 3	no	2	1	0.0000000	no
## 4	no	1	1	0.0000000	no
## 5	no	1	0	0.6666667	yes
## 6	yes	2	0	0.1818182	no
## 7	no	2	0	0.1818182	no
## 8	yes	1	0	0.6666667	yes
## 9	no	2	0	0.1818182	no
## 10	no	2	0	0.1818182	no
## 11	no	2	0	0.1818182	no
## 12	no	1	2	0.0000000	no
## 13	yes	1	0	0.6666667	yes
## 14	no	1	0	0.6666667	yes
## 15	yes	1	0	0.6666667	yes
## 16	yes	1	0	0.6666667	yes
## 17	no	2	0	0.1818182	no
## 18	no	2	0	0.1818182	no
## 19	no	2	0	0.1818182	no
## 20	no	2	0	0.1818182	no
## 21	yes	1	0	0.6666667	yes
## 22	no	1	0	0.6666667	yes
## 23	yes	2	2	1.0000000	yes
## 24	yes	2	0	0.1818182	no

2(3). Compute manually the naive Bayes conditional probability of an injury given WEATHER_R = 1 and TRAF_CON_R = 1.

```

#Probability of weather=1 given injury=YES
p.w.y = (p.1[3,1]+p.1[3,2]+p.1[3,3])/(p.1[3,1]+p.1[3,2]+p.1[3,3]+p.1[4,1]+p.1[4,2]+p.1[4,3])
#Probability of traffic=1 given injury=YES
p.t.y = (p.1[3,2]+p.1[4,2])/(p.1[3,1]+p.1[3,2]+p.1[3,3]+p.1[4,1]+p.1[4,2]+p.1[4,3])
#Probability of Injury=YES
p.y = (p.1[3,1]+p.1[3,2]+p.1[3,3]+p.1[4,1]+p.1[4,2]+p.1[4,3])/24
#Probability of weather=1 given injury=NO
p.w.n = (p.1[1,1]+p.1[1,2]+p.1[1,3])/(p.1[1,1]+p.1[1,2]+p.1[1,3]+p.1[2,1]+p.1[2,2]+p.1[2,3])
#Probability of traffic=1 given injury=NO
p.t.n = (p.1[1,2]+p.1[2,2])/(p.1[1,1]+p.1[1,2]+p.1[1,3]+p.1[2,1]+p.1[2,2]+p.1[2,3])
#Probability of Injury=NO
p.n = (p.1[1,1]+p.1[1,2]+p.1[1,3]+p.1[2,1]+p.1[2,2]+p.1[2,3])/24

#Result for probability of an injury given WEATHER_R = 1 and TRAF_CON_R = 1
res = (p.w.y*p.t.y*p.y)/((p.w.y*p.t.y*p.y)+(p.w.n*p.t.n*p.n))
res

```

```
## [1] 0
```

From the above result the naive bayes conditional probability of an injury given WEATHER_R = 1 and TRAF_CON_R = 1 is 0. A conditional probability of 0 in this context means that, according to the Naive Bayes model, there is no probability of an injury occurring when both WEATHER_R = 1 and TRAF_CON_R = 1 are observed.

2(4). Run a naive Bayes classifier on the 24 records and two predictors. Check the model output to obtain probabilities and classifications for all 24 records. Compare this to the exact Bayes classification. Are the resulting classifications equivalent? Is the ranking (= ordering) of observations equivalent?

```
#Creating a naive bayes model for accident24 dataframe
naiveb.model <- naiveBayes(INJURY ~ TRAF_CON_R + WEATHER_R,
                           data = accident24)
#Making predictions using naive bayes model for the same dataframe acc24
naiveb.predict <- predict(naiveb.model, newdata = accident24, type = "raw")
#Extracting the predicted probabilities for the "Yes" class (INJURY = Yes) and adding them to the original dataset accident24,
accident24$naivebpred.prob <- naiveb.predict[,2]
accident24
```

##	INJURY	WEATHER_R	TRAF_CON_R	prob.inj	pred.prob	naivebpred.prob
## 1	yes	1	0	0.6666667	yes	0.571428571
## 2	no	2	0	0.1818182	no	0.250000000
## 3	no	2	1	0.0000000	no	0.002244949
## 4	no	1	1	0.0000000	no	0.008919722
## 5	no	1	0	0.6666667	yes	0.571428571
## 6	yes	2	0	0.1818182	no	0.250000000
## 7	no	2	0	0.1818182	no	0.250000000
## 8	yes	1	0	0.6666667	yes	0.571428571
## 9	no	2	0	0.1818182	no	0.250000000
## 10	no	2	0	0.1818182	no	0.250000000
## 11	no	2	0	0.1818182	no	0.250000000
## 12	no	1	2	0.0000000	no	0.666666667
## 13	yes	1	0	0.6666667	yes	0.571428571
## 14	no	1	0	0.6666667	yes	0.571428571
## 15	yes	1	0	0.6666667	yes	0.571428571
## 16	yes	1	0	0.6666667	yes	0.571428571
## 17	no	2	0	0.1818182	no	0.250000000
## 18	no	2	0	0.1818182	no	0.250000000
## 19	no	2	0	0.1818182	no	0.250000000
## 20	no	2	0	0.1818182	no	0.250000000
## 21	yes	1	0	0.6666667	yes	0.571428571
## 22	no	1	0	0.6666667	yes	0.571428571
## 23	yes	2	2	1.0000000	yes	0.333333333
## 24	yes	2	0	0.1818182	no	0.250000000

Let us use Caret

```
naiveb2 <- train(INJURY ~ TRAF_CON_R + WEATHER_R,
                 data = accident24, method = "nb")
```

```
## Warning: model fit failed for Resample01: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1
```

```
## Warning: model fit failed for Resample02: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample03: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample04: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample05: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample06: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample07: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample08: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample09: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1
```

```
## Warning: model fit failed for Resample10: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1
```

```
## Warning: model fit failed for Resample11: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :  
## Zero variances for at least one class in variables: TRAF_CON_R1
```

```
## Warning: model fit failed for Resample12: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1
```

```
## Warning: model fit failed for Resample13: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2, WEATH
ER_R2
```

```
## Warning: model fit failed for Resample14: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1
```

```
## Warning: model fit failed for Resample15: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1
```

```
## Warning: model fit failed for Resample16: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample17: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1
```

```
## Warning: model fit failed for Resample18: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1
```

```
## Warning: model fit failed for Resample19: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample20: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample21: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample22: usekernel=FALSE, fL=0, adjust=1 Error in Na
iveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample23: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample24: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning: model fit failed for Resample25: usekernel=FALSE, fL=0, adjust=1 Error in NaiveBayes.default(x, y, usekernel = FALSE, fL = param$fL, ...) :
## Zero variances for at least one class in variables: TRAF_CON_R1, TRAF_CON_R2
```

```
## Warning in nominalTrainWorkflow(x = x, y = y, wts = weights, info = trainInfo,
## : There were missing values in resampled performance measures.
```

```
## Warning in train.default(x, y, weights = w, ...): missing values found in
## aggregated results
```

```
predict(naiveb2, newdata = accident24[,c("INJURY", "WEATHER_R", "TRAF_CON_R")])
```

```
## [1] no no no no no no no no no no no no no no no no no no no no no no no no
## Levels: no yes
```

```
predict(naiveb2, newdata = accident24[,c("INJURY", "WEATHER_R", "TRAF_CON_R")],
        type = "raw")
```

```
## [1] no no no no no no no no no no no no no no no no no no no no no no no no
## Levels: no yes
```

3. Let us now return to the entire dataset. Partition the data into training (60%) and validation (40%).

```
#Important to ensure that we get the same sample if we rerun the code
```

```
set.seed(1)
```

```
train.df<- sample(row.names(accident),0.6*dim(accident)[1])
valid.df <- setdiff(row.names(accident),train.df)
train.accident <- accident[train.df,]
valid.accident <- accident[valid.df,]
print(paste("The size of training data is:",nrow(train.accident)))
```

```
## [1] "The size of training data is: 25309"
```

```
print(paste("The size of validation data is:",nrow(valid.accident)))
```

```
## [1] "The size of validation data is: 16874"
```

3(1).Run a naive Bayes classifier on the complete training set with the relevant predictors (and INJURY as the response). Note that all predictors are categorical. Show the confusion matrix.

```
nb.accident <- naiveBayes(INJURY ~ TRAF_CON_R + WEATHER_R, data = train.accident)
nb.predictaccident <- predict(nb.accident, newdata= valid.accident)
nb.predictaccident
```



```

##      [1] yes no  no  yes yes yes no  no  yes no  no  no  no  no  yes no  no  no
##      [19] no  no  yes no  no  no  no  no  no  no  no  no  yes no  no  yes yes yes
##      [37] no  no  no  yes yes no  yes yes yes no  no  yes no  no  no  no  yes no
##      [55] no  no  yes no  yes yes yes no  yes yes no  no  yes yes yes no  no  no
##      [73] no  no  no  no  no  no  no  yes no  yes no  no  yes no  no  no  no  no
##      [91] no  no  no  no  no  no  yes no  no  no  yes no  yes no  no  no  no  no
##     [109] no  yes yes yes yes no  no  no  no  no  yes no  no  no  no  no  yes no
##     [127] no  yes yes no  no  no  no  no  no  no  no  no  no  yes yes yes no  no
##     [145] no  yes yes yes yes yes yes no  no  no  no  no  no  no  yes yes yes no
##     [163] no  no  no  yes yes yes no  no  no  no  yes yes yes yes yes yes yes
##     [181] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [199] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [217] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [235] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [253] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [271] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [289] yes yes yes yes yes yes yes yes yes yes yes yes no  yes yes yes yes yes
##     [307] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [325] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [343] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [361] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [379] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [397] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [415] yes yes yes no  yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [433] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [451] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [469] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [487] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [505] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [523] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [541] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [559] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [577] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [595] yes no  no  yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [613] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
##     [631] yes yes yes yes no  no  no  no  yes yes no  yes yes no  no  yes no  no
##     [649] no  no  yes no  yes no  yes no  no  yes no  no  no  no  yes no  yes yes no
##     [667] yes no  no  no  yes no  no  yes yes yes yes no  yes no  no  no  yes no
##     [685] no  no  no  no  no  yes yes yes no  no  no  no  no  no  no  no  yes no
##     [703] no  yes yes no  no  no  no  no  yes no  no  no  no  yes no  no  no  no
##     [721] yes no  no  no  no  no  no  no  no  no  no  no  yes yes no  no  yes no  yes
##     [739] no  no  no  yes no  no  yes yes no  no  no  no  no  no  no  no  no  yes
##     [757] no  no  yes no  no  no  no  no  no  no  no  no  no  yes no  no  no  yes
##     [775] yes yes yes no  no  no  no  no  yes no  no  no  no  no  no  no  yes no
##     [793] no  no  no  yes no  no  yes yes yes no  no  no  no  yes yes yes no  no
##     [811] no  no  no  no  yes no  no  no  yes no  no  no  no  no  no  no  no  no
##     [829] yes yes no  no  no  yes yes no  no  no  no  no  no  no  no  no  no
##     [847] no  no  yes no  yes no  no  yes no  yes no  no  no  no  no  no  yes yes
##     [865] yes no  yes no  yes no  no  no  no  no  yes no  no  no  no  no  no  no
##     [883] no  no  no  no  no  no  no  yes yes no  yes no  no  no  no  no  no  no
##     [901] no  no  yes no  yes yes yes no  no  no  no  no  no  no  no  no  no
##     [919] no  yes yes no  no  no  no  no  no  no  no  no  no  no  no  no  no

```

##	[937]	no	no	no	no	no	yes	no	no	no	yes	yes	yes	yes	yes	yes	no	no	no
##	[955]	no	no	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no
##	[973]	no	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[991]	no	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	yes	no	no	no	no
##	[1009]	no	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no
##	[1027]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	yes	yes
##	[1045]	yes	yes	yes	no	no	no	no	no	no	no	no	yes	yes	yes	no	no	no	no
##	[1063]	no	yes	yes	yes	no	no	yes	no	yes	no	no	no	yes	yes	yes	yes	yes	no
##	[1081]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	no
##	[1099]	yes	no	no	no	no	yes	yes	no	no	no	yes	yes	yes	yes	yes	no	no	no
##	[1117]	no	no	no	no	no	no	no	no	yes	no	no	yes	yes	yes	yes	no	no	no
##	[1135]	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	yes	yes
##	[1153]	yes	yes	no	no	yes	yes	yes	no	no	yes	no	no	yes	yes	yes	yes	yes	no
##	[1171]	no	no	no	no	no	no	no	no	no	no	no	yes	no	no	no	no	no	no
##	[1189]	no	no	no	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1207]	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no
##	[1225]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[1243]	no	no	no	no	no	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1261]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1279]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1297]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1315]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1333]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1351]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1369]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1387]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	yes	yes	yes	yes
##	[1405]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1423]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1441]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1459]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1477]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1495]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1513]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1531]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1549]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1567]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1585]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1603]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1621]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1639]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	yes
##	[1657]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1675]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1693]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1711]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1729]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1747]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1765]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1783]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1801]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1819]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1837]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[1855]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

file:///Users/akhilchintu/Downloads/assignment-3-fml.html

```
## [2809] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [2827] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [2845] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [2863] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [2881] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [2899] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [2917] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [2935] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [2953] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [2971] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [2989] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3007] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3025] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3043] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3061] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3079] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3097] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3115] yes yes yes no no yes no no no no no no no no no no no yes yes yes
## [3133] yes no no no no yes no yes no no yes yes yes yes yes yes yes yes yes
## [3151] yes yes no no no no no no no no no no no yes no no no yes no
## [3169] no no no yes yes yes yes yes no no no no no no no no no no
## [3187] no no yes no no no yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3205] yes no no no no no no no no no no no no no no no no no no
## [3223] no yes yes no no no no no no no yes no no no no no no no yes
## [3241] no no no yes no no no no no no no yes no no yes no yes no
## [3259] yes yes no no no no no no no no yes yes yes no yes yes no no
## [3277] yes yes no no yes no no no yes no no no yes no no no no no
## [3295] no no no no yes yes yes yes no no no no no no no no no no
## [3313] no no no no no no no no yes yes yes no yes yes no no no no
## [3331] no no yes yes yes no no no yes yes yes yes yes yes yes yes yes yes
## [3349] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes no no
## [3367] no no no no no no no no no no no no no no no no no no no
## [3385] no no no no no no no no no no no no no no yes yes no yes
## [3403] yes no yes yes no yes no no no yes yes no no no no no no no
## [3421] no yes no no no yes no no no no no no no no no yes no yes
## [3439] yes no yes yes no no no no no no no no no yes no no yes
## [3457] no no no no yes yes yes no no no no no no yes yes yes yes
## [3475] yes yes yes yes yes yes yes yes no no no no no no no no no
## [3493] no no no no no no no no no no no no no no no no yes yes
## [3511] yes no no yes yes no no yes no no no no yes no no no yes no
## [3529] no yes yes no no yes yes yes yes no no no yes yes yes no no no
## [3547] no no yes yes yes yes yes yes no no no yes yes yes yes yes yes
## [3565] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3583] yes yes yes yes yes yes yes yes yes yes no no no no no no no
## [3601] no no no no no no no no no no no no no no no no no no
## [3619] no no no no no no no no no no no no no no no no no no
## [3637] no no no no no no no no no no no no no no no no no no
## [3655] no no no no yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3673] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3691] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3709] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [3727] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
```

file:///Users/akhilchintu/Downloads/assignment-3-fml.html

```
## [4681] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4699] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4717] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4735] yes no yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4753] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4771] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4789] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4807] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4825] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4843] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4861] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4879] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4897] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4915] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4933] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4951] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4969] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [4987] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [5005] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [5023] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [5041] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [5059] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [5077] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [5095] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [5113] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [5131] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes no no no
## [5149] no no yes yes yes no yes yes yes no yes no no yes no no no no no
## [5167] no yes no yes no yes no yes yes no no no yes yes no no no no yes
## [5185] yes yes yes yes yes yes yes yes yes yes yes yes no no no no no no no
## [5203] no no no no no no yes no no yes yes no yes no no no no yes yes
## [5221] yes yes yes no yes no no yes no no no no yes yes no no yes no
## [5239] no no yes no no no yes yes yes yes no no no no no no no no
## [5257] no no no no yes no yes no no no yes yes no no no yes yes yes
## [5275] no yes yes yes yes yes yes yes yes yes yes yes no no no no no no
## [5293] no no yes yes yes yes yes yes no yes yes yes yes yes yes yes yes
## [5311] no no no no no no no no no no yes yes no no no no no no yes
## [5329] no no no yes no no no no no no yes yes no no no no no no no
## [5347] no no no no no no no no no no no yes no no no no no no no
## [5365] no no yes no no no no no yes no no no no no no no yes no
## [5383] no no no yes no no no no no no no no yes yes no yes no
## [5401] yes no no no no no no no yes no no yes no yes no no no no
## [5419] no no no no no no yes no yes yes yes no no no no no no no
## [5437] no no no no no no no no no no no no no no no yes no no
## [5455] no no no no yes no no no no no no no no no no yes no no
## [5473] no no no yes no no no yes yes no no no no no no no no no
## [5491] no no no no no yes yes yes no no no no no yes yes yes no
## [5509] no no no no no no no no no no no no no no no no no yes
## [5527] yes yes no no no no yes yes no no no no no no no no yes
## [5545] no no yes no no no no no no no no no no yes yes no no
## [5563] no no no no no no no yes no no no no no no no no yes
## [5581] yes yes yes no no no no no no yes yes yes no no no no no
## [5599] no no no no no no yes yes no no no no no no no no no
```

```

## [5617] no no no yes yes no no no no no no no no no yes yes yes no
## [5635] no no no no no no no no no no no no no no no yes yes yes yes
## [5653] yes yes no no no no no no no no no no no no no no no no no
## [5671] no no no no no no no no no no no yes yes yes no no no no yes
## [5689] yes no yes yes no no no yes no yes yes no no no no yes yes no
## [5707] no no no no yes no no no no no yes no yes no yes yes yes no
## [5725] no no no no no no no no no no yes yes no no no no no no
## [5743] no no no yes yes yes no no no yes no no yes no no no no yes
## [5761] no yes yes yes yes yes yes yes yes yes yes yes yes yes no no no
## [5779] no no no no no no no no no no no no no no no no no no no
## [5797] no no no no no no no yes yes yes yes no no no no yes yes yes
## [5815] yes yes yes yes yes yes yes yes yes yes yes yes no no no no no
## [5833] no no no no no no no no no no no no no no no no no no no
## [5851] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [5869] yes yes yes yes yes yes yes yes yes yes yes yes yes no no no no
## [5887] no no no no no no no no no no no no no no no no no no no
## [5905] no no no no no no no no no no no no no no no no no no no
## [5923] no no no no no no no no no no no no no no no no no no no
## [5941] no no no yes no yes no yes no no yes yes no yes yes yes no no
## [5959] no yes no no no yes no no no no yes yes yes yes no no no yes
## [5977] yes no no no no no yes yes no no no yes yes yes yes yes no no
## [5995] no no yes no no no no yes no yes yes yes yes yes yes no no no
## [6013] no no no no no no no no no yes yes no no no no no no no no
## [6031] no no no yes yes yes yes no no no no no yes yes no yes no no
## [6049] no no no yes yes yes no no no no no no no no yes yes no no
## [6067] no no no no no no no no no yes no no no no no no no no no
## [6085] no no no yes yes yes yes yes no no no no no no yes yes yes yes
## [6103] yes yes yes no no no no no no no no no no no no no no no no
## [6121] no no yes yes yes yes yes yes yes yes yes yes no no no no no no
## [6139] no no no no no no no no no no no no no no no no no no no no
## [6157] no no no yes yes yes yes yes no no no no no no no no no no no
## [6175] no no no no yes yes yes yes yes yes no no no no no no no no no
## [6193] no yes yes yes yes yes yes yes yes yes yes yes no no no no no no
## [6211] no no no no no no no no no no no no no no no no no no no no
## [6229] no no yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [6247] yes no no no no no no no no no no no no no no no no no no no
## [6265] no no no no no no no no no no no no no no no no no yes yes
## [6283] yes yes yes yes yes yes no no no no no no no no no no no no no
## [6301] no no no no no no no no yes yes yes yes yes yes no no no no no
## [6319] no no yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [6337] no no no no no no no no no no no no no no no no no no no no
## [6355] no no no no no no no no yes no yes yes yes yes yes yes yes yes
## [6373] yes yes yes yes no no no no no no no no no no no no no no no
## [6391] no no no no no no no no no no no no no no no no no no no no
## [6409] no no no no no no no no no no yes yes yes yes yes yes yes yes
## [6427] yes yes yes yes no no no no no no no no no no no no no no no
## [6445] no no no no no no no no no no no no no no no no no no no no
## [6463] no no no no no no no no no no no no no no no no no no no no
## [6481] no no no no no no yes yes yes yes yes no no yes yes yes yes yes
## [6499] yes yes yes yes yes no no no no no no no no no no no no no no
## [6517] no yes no yes yes yes yes no no no no no no no no no no no no
## [6535] no no no no yes yes yes no no no no no no no no no no no no

```

##	[6553]	no	no	no	no	yes	yes	yes	no	no	no	no	no	no	yes	no	no	no	no
##	[6571]	no	no	yes	no	no	no	no	no	no	no	no	no	no	yes	yes	no	no	yes
##	[6589]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
##	[6607]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[6625]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[6643]	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no
##	[6661]	no	no	no	no	no	no	no	no	no	yes	no	no	yes	no	no	no	no	yes
##	[6679]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no
##	[6697]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[6715]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[6733]	no	yes	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[6751]	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[6769]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[6787]	no	no	yes	yes	yes	no	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[6805]	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	yes	yes	yes	yes	no
##	[6823]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	yes	yes	yes
##	[6841]	yes	no	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[6859]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[6877]	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no
##	[6895]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[6913]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[6931]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[6949]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[6967]	no	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no	yes	yes
##	[6985]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7003]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no
##	[7021]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[7039]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[7057]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[7075]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[7093]	no	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7111]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7129]	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[7147]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[7165]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[7183]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[7201]	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
##	[7219]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7237]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7255]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7273]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7291]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7309]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7327]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7345]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7363]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7381]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7399]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7417]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7435]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7453]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
##	[7471]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

file:///Users/akhilchintu/Downloads/assignment-3-fml.html

file:///Users/akhilchintu/Downloads/assignment-3-fml.html

[illegible]

file:///Users/akhilchintu/Downloads/assignment-3-fml.html

29/35

file:///Users/akhilchintu/Downloads/assignment-3-fml.html

file:///Users/akhilchintu/Downloads/assignment-3-fml.html

file:///Users/akhilchintu/Downloads/assignment-3-fml.html

file:///Users/akhilchintu/Downloads/assignment-3-fml.html

file:///Users/akhilchintu/Downloads/assignment-3-fml.html

```
## [16849] yes yes yes yes yes yes yes yes yes yes yes yes yes yes yes
## [16867] yes no  yes yes yes yes yes yes
## Levels: no yes
```

```
valid.accident$INJURY <- as.factor(valid.accident$INJURY)
confusion.matrix <- confusionMatrix(nb.predictaccident,valid.accident$INJURY)
confusion.matrix
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction   no  yes
##           no 1285 1118
##           yes 6934 7537
##
##           Accuracy : 0.5228
##           95% CI : (0.5152, 0.5304)
##           No Information Rate : 0.5129
##           P-Value [Acc > NIR] : 0.005162
##
##           Kappa : 0.0277
##
##           Mcnemar's Test P-Value : < 2.2e-16
##
##           Sensitivity : 0.15635
##           Specificity : 0.87083
##           Pos Pred Value : 0.53475
##           Neg Pred Value : 0.52083
##           Prevalence : 0.48708
##           Detection Rate : 0.07615
##           Detection Prevalence : 0.14241
##           Balanced Accuracy : 0.51359
##
##           'Positive' Class : no
##
```

3(2).What is the overall error of the validation set?

```
#Calculating overall erroe
overall.error <- (confusion.matrix$table[1,2]+confusion.matrix$table[2,1])/sum(confusion.matrix$table)
overall.error
```

```
## [1] 0.4771838
```

In this specific case, the overall error rate is approximately 0.4771838, which means that the model's predictions are incorrect for about 47.72% of the cases in the validation dataset. In other words, it has a 47.72% error rate, indicating that the model's accuracy is approximately 52.28%.