# 8

Pawan Pandey Roll no:21

2024-05-31

1

```
library(car)
```

```
## Warning: package 'car' was built under R version 4.3.3
```

```
## Loading required package: carData
```

```
## Warning: package 'carData' was built under R version 4.3.3
```

```
set.seed(21)
# Data partition :
#Do random sampling to divide the cases into two independent samples
ind <- sample(2, nrow(Arrests), replace = T, prob = c(0.8, 0.2))
#Data partition
train.data <- Arrests[ind==1,]
test.data <- Arrests[ind==2,]
```

2.

```
# Logistic Regression
model <- glm(released ~ colour+age+sex+employed+citizen , data = train.data, family = binomial)
```

```
# Naive Bayes
library(e1071)
model.nb <- naiveBayes(  released ~ colour+age+sex+employed+citizen ,data = train.data)
```

3.

```
# Predictions from logistic model
predictions <- predict(model, newdata = test.data, type = "response")
#prediction to binary
# Convert predictions to binary values (0/1)
predictions_binary <- as.numeric(ifelse(predictions > 0.5, 1, 0))
#Get the confusion matrix
conf_mat <- table(factor(predictions_binary, levels = c(0, 1)), test.data$released)
conf_mat
```

```
##
##      No Yes
##   0   0   0
##   1 225 872
```

```
y_pred <- predict(model.nb,newdata = test.data)
cm <- table(test.data$released,y_pred)
cm
```

```
##      y_pred
##        No Yes
##   No   12 213
##   Yes  21 851
```

From above we can see that navie bayes has less number on errors so ,navie bayes is the best classification.