

Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias de la Computación

Máquinas de aprendizaje

Reporte: Problema 2 con Naive Bayes



BUAP

Docente: Abraham Sánchez López

Alumno

Taisen Romero Bañuelos

Matrícula

202055209

Naive Bayes y el partido nazi

Bien, nuestro objetivo es determinar la probabilidad de que los profesores sean miembros del partido nazi según su religión, cohorte, residencia y género, obviamente, usando Naïve Bayes.

Los datos se cargaron en una única columna, por lo tanto, hay que separar los datos en diferentes columnas.

```
> head(data)
R...C...Re..G...M...Co
1 1 1 1 1 1 66
2 1 1 1 2 1 1
3 2 1 1 1 1 22
4 2 1 1 2 1 0
5 3 1 1 1 1 11
6 3 1 1 2 1 3
> summary(data)
R...C...Re..G...M...Co
Length:120
Class :character
Mode :character
> summary(data$R...C...Re..G...M...Co)
Length Class Mode
120 character character
> #Búsqueda de datos nulos
> colSums(is.na(data))
R...C...Re..G...M...Co
0
> |
```

De paso convertimos los datos en factores excepto la columna Co, Co será un integer.

```
> summary(data)
R      C      Re      G      M      Co
1:40  1:24  1:60  1:60  0:60  Min.   : 0.0
2:40  2:24  2:60  2:60  1:60  1st Qu.: 20.0
3:40  3:24                      Median : 71.5
      4:24                      Mean   :113.0
      5:24                      3rd Qu.:186.8
                        Max.   :501.0
> str(data)
'data.frame': 120 obs. of 6 variables:
 $ R : Factor w/ 3 levels "1","2","3": 1 1 2 2 3 3 1 1 2 2 ...
 $ C : Factor w/ 5 levels "1","2","3","4",...: 1 1 1 1 1 1 2 2 2 2 ...
 $ Re: Factor w/ 2 levels "1","2": 1 1 1 1 1 1 1 1 1 1 ...
 $ G : Factor w/ 2 levels "1","2": 1 2 1 2 1 2 1 2 1 2 ...
 $ M : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 2 ...
 $ Co: int 66 1 22 0 11 3 153 14 62 9 ...
> |
```

Haciendo el análisis exploratorio de datos obtenemos los siguiente



