
Fundamentos de Aprendizaje Automático.

Práctica 1

Grupo1461
Mario Valdemaro García Roque
Manuel Reyes Sánchez

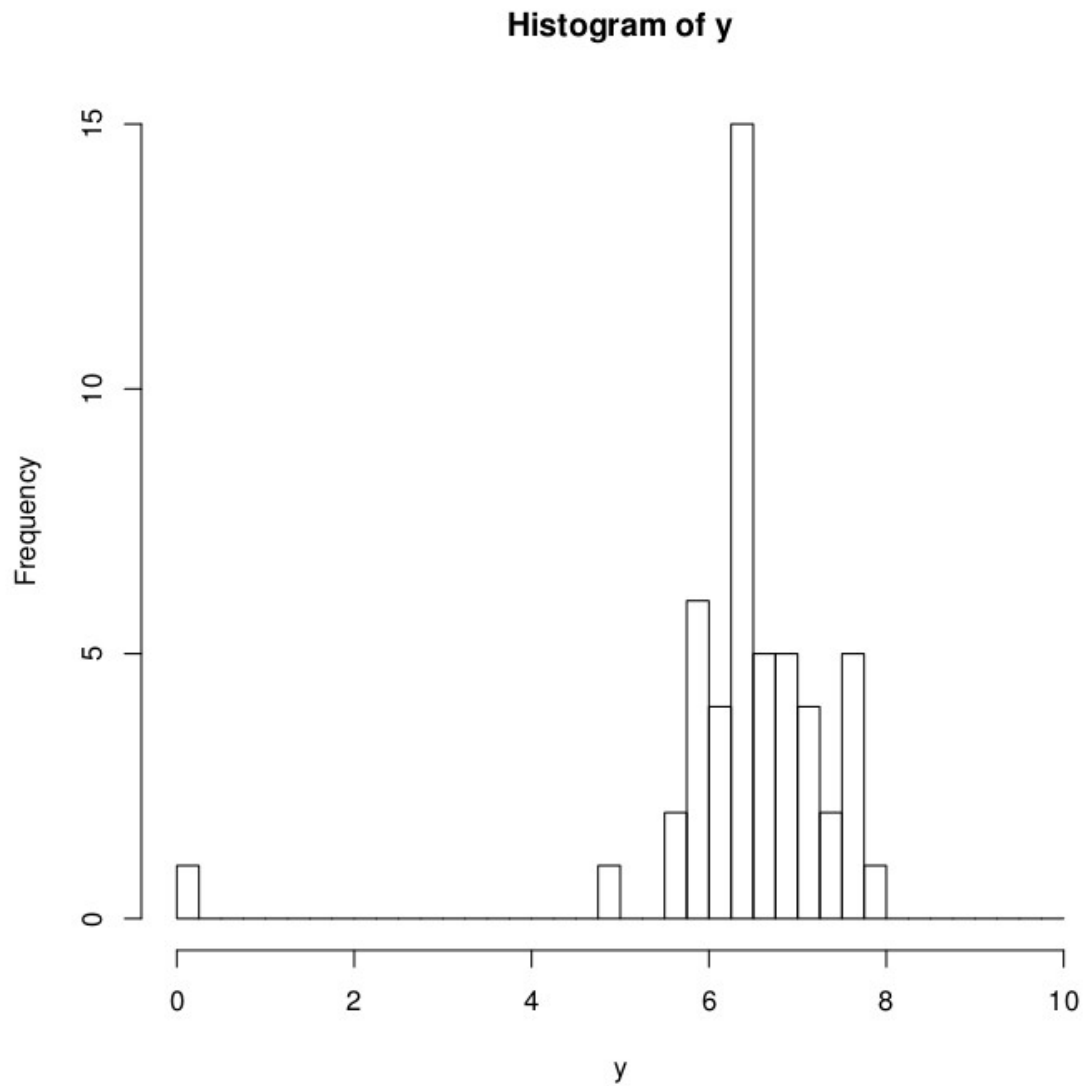
Apartados 1 y 2

Validación cruzada 5:

	A Priori	Naive-Bayes	Naive-Bayes Laplace	Naive-Bayes dist. Normal	Naive-Bayes Laplace y dist. normal
Tic-tac-toe	34,66%	30,89%	29,53%	30,79%	30,05%
crx	44,99%	37,1%	30,14%	40,58%	39,42%
Iris	78,0%	15,3%	8,0%	8,0%	6,0%

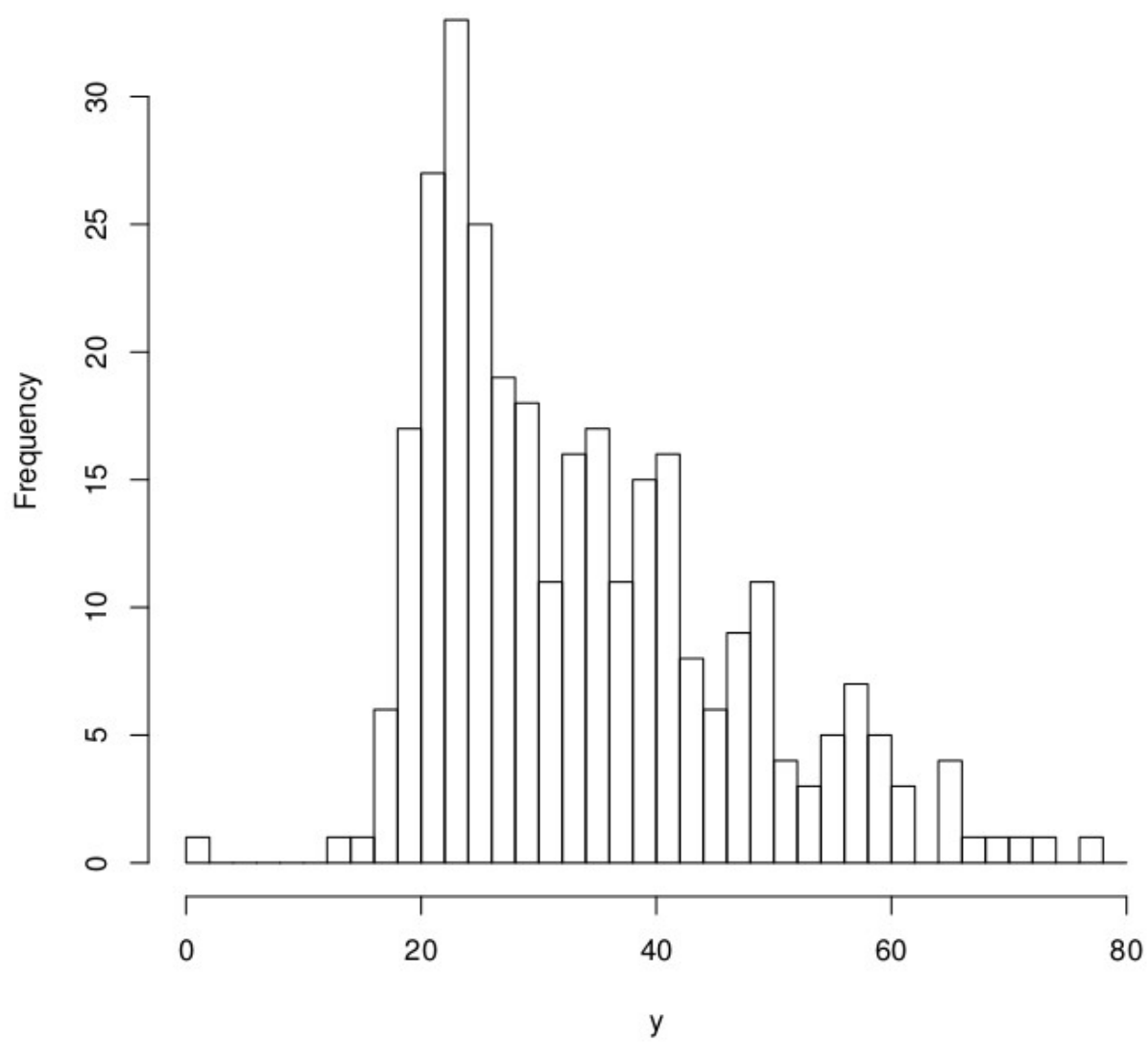
Los datos de las **tres en raya** se ve como en su mayoría son iguales independientemente del clasificador, esto se debe a que todos los atributos de estos datos son nominales y al ser nominales usar clasificadores como normalización no mejora el resultado, ya que para estos datos no se hace. Otro tema en el que tampoco mejora es en Laplace. Esto se debe a que usar Laplace y no usarlo en estos datos nos va dar igual y que la mayoría de los casos están contemplados y es muy improbable que se de un caso en que no tengamos información.

Como podemos observar en la tabla de clasificadores los mejores valores los obtenemos en **Iris**, que según vamos utilizando distintos clasificadores va mejorando, tanto usando Laplace mejora, ya que habrá casos para los que no tengamos información al tener una columna de atributos muchos posibles valores. Y por otro lado vemos como al normalizar también mejora esto se debe a que los valores de las tablas de atributos continuos siguen una distribución normal, por tanto clasificarlos según este modelo nos va a dar muy buenas predicciones. Como muestra damos una gráfica del atributo 1 para una determinada clase:



Los datos de crx son los peores y se observa como empeoran mucho para la normal a diferencia de los datos de lirs. Esto se debe a que los atributos que son continuos de de estos datos no siguen una distribución normal sino algo mas parecido a una chi cuadrado u otras funciones, como muestra:

Histogram of y



Validación cruzada 10:

	A Priori	Naive-Bayes	Naive-Bayes Laplace	Naive-Bayes dist. Normal	Naive-Bayes Laplace y dist. normal
Tic-tac-toe	34,65%	29,95%	30,26%	29,75%	30,16%
crx	44,99%	39,71%	30,14%	40,58%	39,85%
Iris	78,6%	13,99%	8,66%	7,99%	7,33%

Como se ve la diferencia entre una y otra tabla no es apreciable y si se repite la prueba varias veces dará que usar un particionado es mejor que otro arbitrariamente, lo que no nos da una conclusión rotunda sobre si es mejor usar un cantidad de particionado que otro.

Apartado 3

$P(\text{Class}=\text{negative}) = 0,35$
 $P(\text{Class}=\text{positive}) = 0,65$
 $P(\text{MLeftSq}=b|\text{Class}=\text{positive}) = 0,274760$
 $P(\text{TRightSq}=x|\text{Class}=\text{negative}) = 0,370482$

Para este conjunto de datos no cambia el resultado al aplicar Laplace, pues hay ejemplo de todas las opciones (x, o, b) de los atributos para las clases.

Apartado 4

$P(\text{Class}=-) = 0,55$
 $P(\text{Class}=+) = 0,44$
 $P(A7=bb|\text{Class}=+) = 0,081433$
 $P(A4=u|\text{Class}=-) = 0,686684$
Atributo A2, Media=33.500814332247565, Varianza=169.84281529777505
Atributo A14, Media=161.2084690553746, Varianza=26091.083576483557
Atributo A15, Media=2038.85993485342, Varianza=5.8480869677450076E7