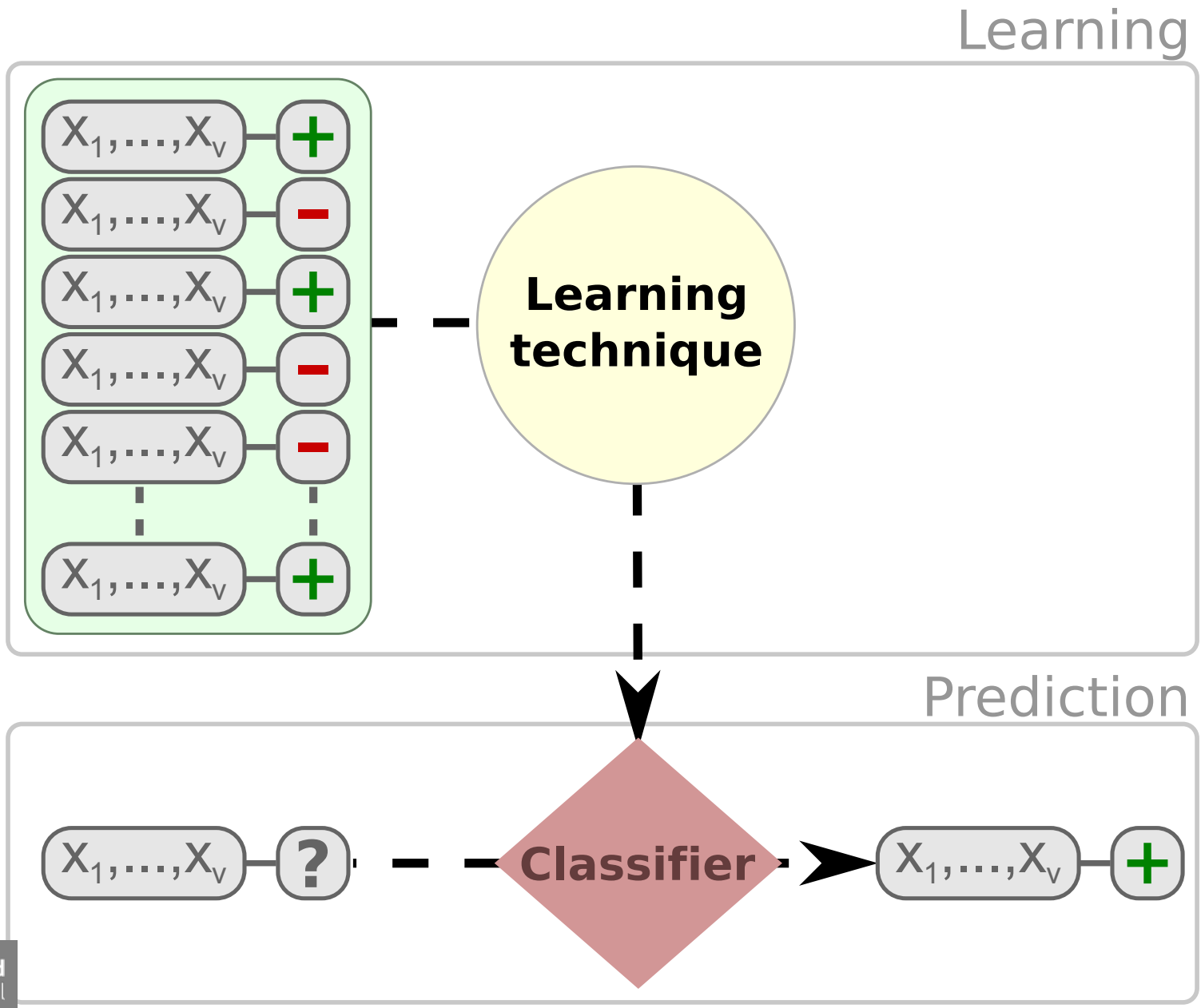
Aprendizaje no supervisado

VC08: Naive Bayes

Félix José Fuentes Hurtado
felixjose.fuentes@campusviu.es

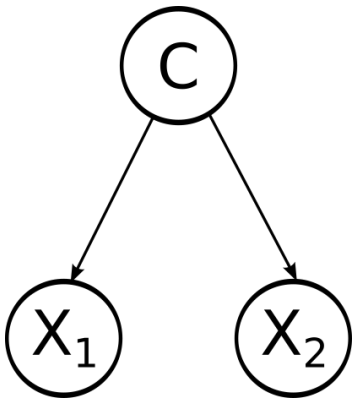
Universidad Internacional de Valencia

Clasificación Supervisada



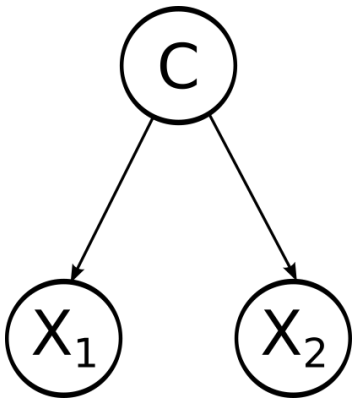
Naive Bayes

$$p(X_1, X_2, C) = p(C|X_1, X_2) \times p(X_1, X_2)$$
$$p(C|X_1, X_2) = \frac{p(X_1, X_2, C)}{p(X_1, X_2)}$$



Naive Bayes

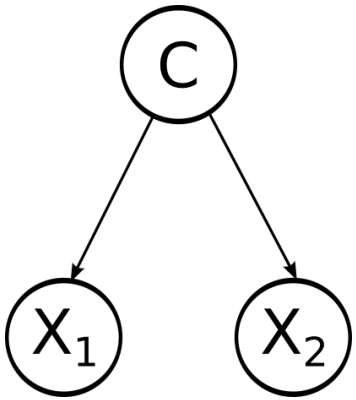
$$p(X_1, X_2, C) = p(C|X_1, X_2) \times p(X_1, X_2)$$
$$p(C|X_1, X_2) = \frac{p(X_1, X_2, C)}{p(X_1, X_2)}$$



Regla de la cadena

$$p(X_1, X_2, C) = p(X_1|X_2, C) \times p(X_2, C)$$
$$p(X_1, X_2, C) = p(X_1|X_2, C) \times p(X_2|C) \times p(C)$$

Naive Bayes



$$p(X_1, X_2, C) = p(C|X_1, X_2) \times p(X_1, X_2)$$
$$p(C|X_1, X_2) = \frac{p(X_1, X_2, C)}{p(X_1, X_2)}$$

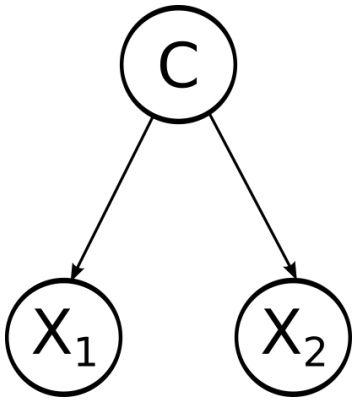
Regla de la cadena

$$p(X_1, X_2, C) = p(X_1|X_2, C) \times p(X_2, C)$$
$$p(X_1, X_2, C) = p(X_1|X_2, C) \times p(X_2|C) \times p(C)$$

Asunción de independencia X's—C del Naive Bayes $p(X_1, X_2, C) = p(X_1|C) \times p(X_2|C) \times p(C)$

$$p(C|X_1, X_2) = \frac{p(X_1|C) \times p(X_2|C) \times p(C)}{p(X_1, X_2)}$$

Naive Bayes



$$p(X_1, X_2, C) = p(C|X_1, X_2) \times p(X_1, X_2)$$
$$p(C|X_1, X_2) = \frac{p(X_1, X_2, C)}{p(X_1, X_2)}$$

Regla de la cadena

$$p(X_1, X_2, C) = p(X_1|X_2, C) \times p(X_2, C)$$
$$p(X_1, X_2, C) = p(X_1|X_2, C) \times p(X_2|C) \times p(C)$$

Asunción de independencia X's—C del Naive Bayes $p(X_1, X_2, C) = p(X_1|C) \times p(X_2|C) \times p(C)$

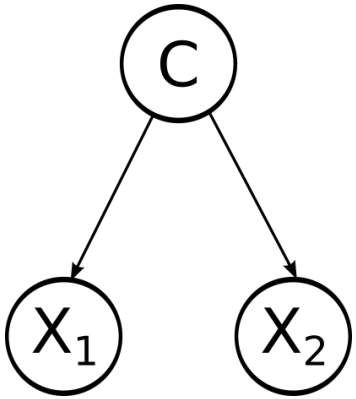
$$p(C|X_1, X_2) = \frac{p(X_1|C) \times p(X_2|C) \times p(C)}{p(X_1, X_2)}$$

$\text{argmáx}_c\, p(C|X_1, X_2)$

$$p(C|X_1, X_2) \propto p(X_1|C) \times p(X_2|C) \times p(C)$$

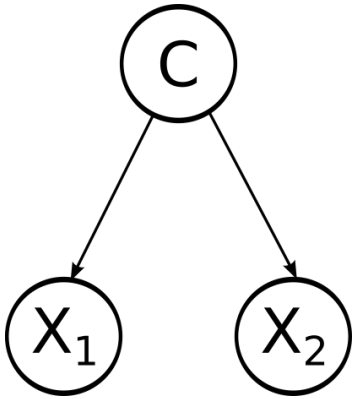
$\text{argmáx}_c\, p(X_1|C) \times p(X_2|C) \times p(C)$

Naive Bayes



$$\theta_C = p(C = 1)$$
$$\theta_{X_1|C=0} = p(X_1 = 1|C = 0)$$
$$\theta_{X_1|C=1} = p(X_1 = 1|C = 1)$$
$$\theta_{X_2|C=0} = p(X_2 = 1|C = 0)$$
$$\theta_{X_2|C=1} = p(X_2 = 1|C = 1)$$

Naive Bayes



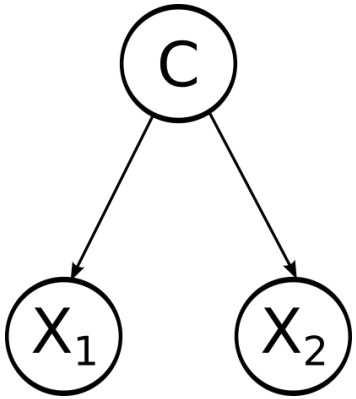
$$\theta_C = p(C = 1)$$
$$\theta_{X_1|C=0} = p(X_1 = 1|C = 0)$$
$$\theta_{X_1|C=1} = p(X_1 = 1|C = 1)$$
$$\theta_{X_2|C=0} = p(X_2 = 1|C = 0)$$
$$\theta_{X_2|C=1} = p(X_2 = 1|C = 1)$$

C	$p(C)$
0	0,4
1	0,6

X_1	C	$p(X_1\|C)$
0	0	0,50
1	0	0,50
0	1	0,33
1	1	0,66

X_2	C	$p(X_2\|C)$
0	0	0,25
1	0	0,75
0	1	0,66
1	1	0,33

Naive Bayes



$$\theta_C = p(C = 1)$$
$$\theta_{X_1|C=0} = p(X_1 = 1|C = 0)$$
$$\theta_{X_1|C=1} = p(X_1 = 1|C = 1)$$
$$\theta_{X_2|C=0} = p(X_2 = 1|C = 0)$$
$$\theta_{X_2|C=1} = p(X_2 = 1|C = 1)$$

C	$p(C)$
0	0,4
1	0,6

$= \theta_C$

X_1	C	$p(X_1\|C)$
0	0	0,50
1	0	0,50
0	1	0,33
1	1	0,66

$= \theta_{X_1|C=0}$

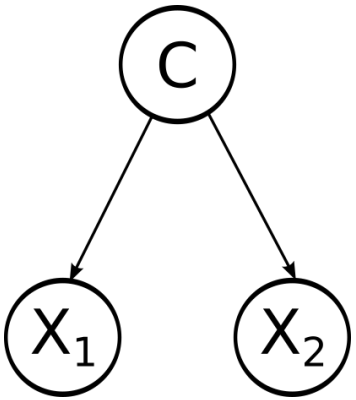
$= \theta_{X_1|C=1}$

X_2	C	$p(X_2\|C)$
0	0	0,25
1	0	0,75
0	1	0,66
1	1	0,33

$= \theta_{X_2|C=0}$

$= \theta_{X_2|C=1}$

Naive Bayes



C	$p(C)$
0	0,4
1	0,6

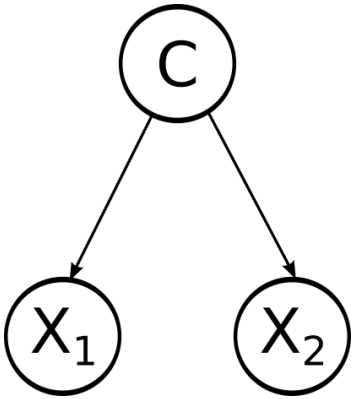
X_1	C	$p(X_1\|C)$
0	0	0,50
1	0	0,50
0	1	0,33
1	1	0,66

X_2	C	$p(X_2\|C)$
0	0	0,25
1	0	0,75
0	1	0,66
1	1	0,33

$$\hat{c} = f(\mathbf{x}) = \underset{c}{\text{argmáx}}\, p(c) \prod_{i=1}^{2} p(x_i|c)$$

X_1	X_2	C
0	0	?
1	0	?

Naive Bayes



C	$p(C)$
0	0,4
1	0,6

X_1	C	$p(X_1\|C)$
0	0	0,50
1	0	0,50
0	1	0,33
1	1	0,66

X_2	C	$p(X_2\|C)$
0	0	0,25
1	0	0,75
0	1	0,66
1	1	0,33

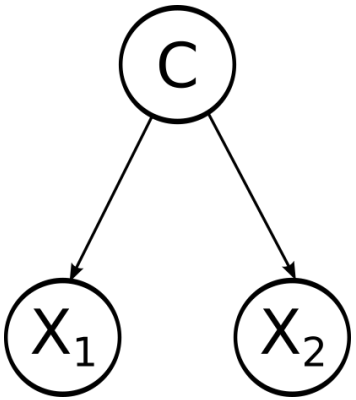
$$\hat{c} = f(\mathbf{x}) = \underset{c}{\text{argmáx}}\, p(c) \prod_{i=1}^{2} p(x_i|c)$$

X_1	X_2	C
0	0	?
1	0	?

C=0 $\quad (1-\theta_C) \times (1-\theta_{X_1|C=0}) \times (1-\theta_{X_2|C=0})$

C=1 $\quad \theta_C \times (1-\theta_{X_1|C=1}) \times (1-\theta_{X_2|C=1})$

Naive Bayes



C	$p(C)$
0	0,4
1	0,6

X_1	C	$p(X_1\|C)$
0	0	0,50
1	0	0,50
0	1	0,33
1	1	0,66

X_2	C	$p(X_2\|C)$
0	0	0,25
1	0	0,75
0	1	0,66
1	1	0,33

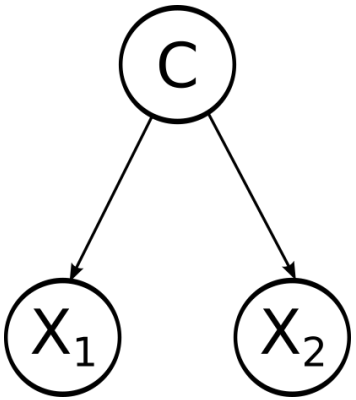
$$\hat{c} = f(\mathbf{x}) = \underset{c}{\text{argmáx}}\, p(c) \prod_{i=1}^{2} p(x_i|c)$$

X_1	X_2	C
0	0	?
1	0	?

C=0 $\quad 0,4 \times 0,5 \times 0,25$

C=1 $\quad 0,6 \times 0,33 \times 0,66$

Naive Bayes



C	$p(C)$
0	0,4
1	0,6

X_1	C	$p(X_1\|C)$
0	0	0,50
1	0	0,50
0	1	0,33
1	1	0,66

X_2	C	$p(X_2\|C)$
0	0	0,25
1	0	0,75
0	1	0,66
1	1	0,33

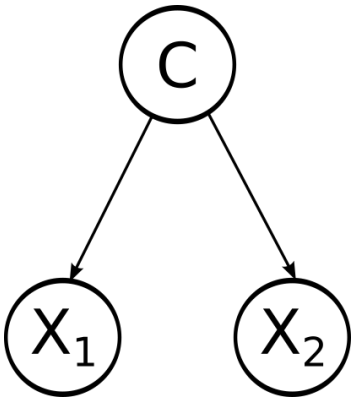
$$\hat{c} = f(\mathbf{x}) = \underset{c}{\text{argmáx}}\, p(c) \prod_{i=1}^{2} p(x_i|c)$$

X_1	X_2	C
0	0	?
1	0	?

C=0 $\quad (1-\theta_C) \times \theta_{X_1|C=0} \times (1-\theta_{X_2|C=0})$

C=1 $\quad \theta_C \times \theta_{X_1|C=1} \times (1-\theta_{X_2|C=1})$

Naive Bayes



C	$p(C)$
0	0,4
1	0,6

X_1	C	$p(X_1\|C)$
0	0	0,50
1	0	0,50
0	1	0,33
1	1	0,66

X_2	C	$p(X_2\|C)$
0	0	0,25
1	0	0,75
0	1	0,66
1	1	0,33

$$\hat{c} = f(\mathbf{x}) = \underset{c}{\text{argmáx}}\, p(c) \prod_{i=1}^{2} p(x_i|c)$$

X_1	X_2	C
0	0	?
1	0	?

C=0 $\quad 0,4 \times 0,5 \times 0,25$

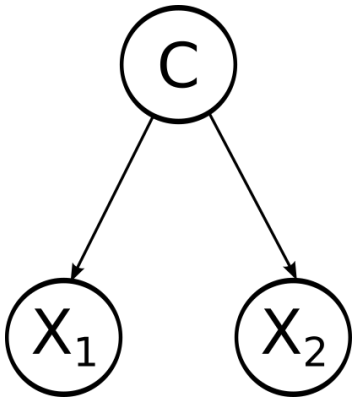
C=1 $\quad 0,6 \times 0,66 \times 0,66$

Naive Bayes

Aprendizaje del modelo

Estimando los parámetros (tablas de probabilidades):

X_1	X_2	C
1	0	0
0	0	1
0	1	0
0	1	0
1	0	1
0	0	1
1	1	0
1	1	1
0	1	0
1	1	0



$p(C)$

$p(X_1|C)$

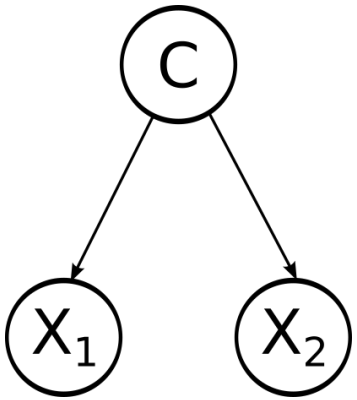
$p(X_2|C)$

Naive Bayes

Aprendizaje del modelo

Estimando los parámetros (tablas de probabilidades):

X_1	X_2	C
1	0	0
0	0	1
0	1	0
0	1	0
1	0	1
0	0	1
1	1	0
1	1	1
0	1	0
1	1	0



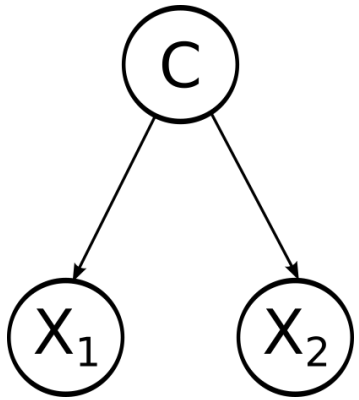
C	$p(C)$
0	
1	

Naive Bayes

Aprendizaje del modelo

Estimando los parámetros (tablas de probabilidades):

X_1	X_2	C
1	0	0
0	0	1
0	1	0
0	1	0
1	0	1
0	0	1
1	1	0
1	1	1
0	1	0
1	1	0



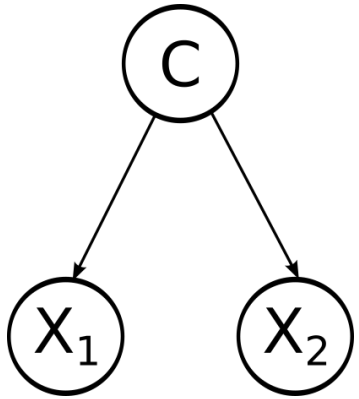
C	$p(C)$
0	$6/10 = 0{,}6$
1	$4/10 = 0{,}4$

Naive Bayes

Aprendizaje del modelo

Estimando los parámetros (tablas de probabilidades):

X_1	X_2	C
1	0	0
0	0	1
0	1	0
0	1	0
1	0	1
0	0	1
1	1	0
1	1	1
0	1	0
1	1	0



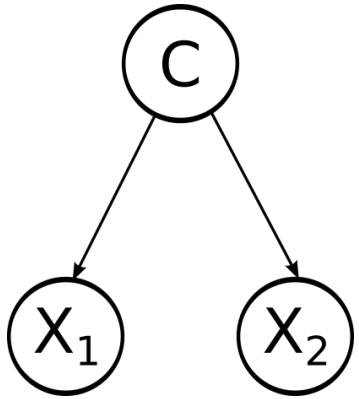
X_1	C	$p(X_1 \mid C)$
0	0	
1	0	
0	1	
1	1	

Naive Bayes

Aprendizaje del modelo

Estimando los parámetros (tablas de probabilidades):

X_1	X_2	C
1	0	0
0	0	1
0	1	0
0	1	0
1	0	1
0	0	1
1	1	0
1	1	1
0	1	0
1	1	0



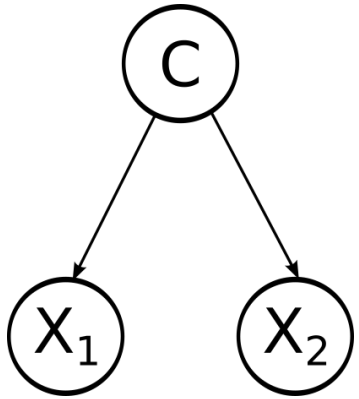
X_1	C	$p(X_1 \mid C)$
0	0	$3/6 = 0{,}50$
1	0	$3/6 = 0{,}50$
0	1	$2/4 = 0{,}50$
1	1	$2/4 = 0{,}50$

Naive Bayes

Aprendizaje del modelo

Estimando los parámetros (tablas de probabilidades):

X_1	X_2	C
1	0	0
0	0	1
0	1	0
0	1	0
1	0	1
0	0	1
1	1	0
1	1	1
0	1	0
1	1	0



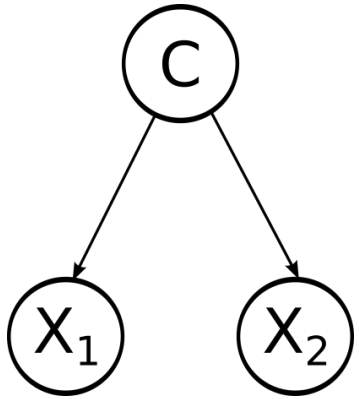
X_2	C	$p(X_2\|C)$
0	0	
1	0	
0	1	
1	1	

Naive Bayes

Aprendizaje del modelo

Estimando los parámetros (tablas de probabilidades):

X_1	X_2	C
1	0	0
0	0	1
0	1	0
0	1	0
1	0	1
0	0	1
1	1	0
1	1	1
0	1	0
1	1	0



X_2	C	$p(X_2\|C)$
0	0	$1/6 = 0{,}17$
1	0	$5/6 = 0{,}83$
0	1	$3/4 = 0{,}75$
1	1	$1/4 = 0{,}25$

Aprendizaje no supervisado

VC08: Naive Bayes

Félix José Fuentes Hurtado
felixjose.fuentes@campusviu.es

Universidad Internacional de Valencia