# CS 171: Intro to ML and DM

Christian Shelton

**UC** Riverside

Slide Set 2: Probabilities



#### Slides from CS 171

- From UC Riverside
  - CS 171: Introduction to Machine Learning and Data Mining
  - Professor Christian Shelton
- DO NOT REDISTRIBUTE
  - ► These slides contain copyrighted material (used with permission) from
    - ► Elements of Statistical Learning (Hastie, et al.)
    - ► Pattern Recognition and Machine Learning (Bishop)
    - An Introduction to Machine Learning (Kubat)
    - Machine Learning: A Probabilistic Perspective (Murphy)
  - ► For use only by enrolled students in the course

	$x_1$ (weekend?)	$x_2$ (weather)	$x_3(game?)$	$y(faster\ route)$
Γ	no	sunny	no	Rte-113 7
	no	sunny	yes	Rte-113
İ	no	cloudy	yes	Rte-75
İ	yes	sunny	no	Rte-113
İ	no	raining	no	Rte-113
-	no	raining	yes	Rte-75
	yes	cloudy	yes	Rte-75
	no	sunny	no	Rte-113
	yes	sunny	no	Rte-75
	no	cloudy	no	Rte-113

	$x_1$ (weekend?)	$x_2$ (weather)	$x_3(game?)$	$y(faster\ route)$
Γ	no	sunny	no	Rte-113 7
	no	sunny	yes	Rte-113
١	no	cloudy	yes	Rte-75
İ	yes	sunny	no	Rte-113
İ	no	raining	no	Rte-113
İ	no	raining	yes	Rte-75
	yes	cloudy	yes	Rte-75
	no	sunny	no	Rte-113
	yes	sunny	no	Rte-75
	no	cloudy	no	Rte-113

$x_1$ (weekend?)	$x_2$ (weather)	$x_3(game?)$	$y(faster\ route)$			
no	sunny	no	Rte-113 7			
no	sunny	yes	Rte-113			
no	cloudy	yes	Rte-75			
yes	sunny	no	Rte-113			
no	raining	no	Rte-113			
no	raining	yes	Rte-75			
yes	cloudy	yes	Rte-75			
no	sunny	no	Rte-113			
yes	sunny	no	Rte-75			
no	cloudy	no	Rte-113			
$x_2 = cloudy \to y = ????$						

	$x_1$ (weekend?)	$x_2$ (weather)	$x_3(game?)$	$y(faster\ route)$
Γ	no	sunny	no	Rte-113 ]
	no	sunny	yes	Rte-113
	no	cloudy	yes	Rte-75
İ	yes	sunny	no	Rte-113
İ	no	raining	no	Rte-113
İ	no	raining	yes	Rte-75
	yes	cloudy	yes	Rte-75
l	no	sunny	no	Rte-113
	yes	sunny	no	Rte-75
	no	cloudy	no	Rte-113
		0000		1

$$\begin{aligned} x_2 &= \mathsf{cloudy} \to y = ???? \\ x_2 &= \mathsf{raining} \to y = ???? \\ x_2 &= \mathsf{sunny} \to y = ???? \end{aligned}$$

	$x_1$ (weekend?)	$x_2$ (weather)	$x_3(game?)$	$y(faster\ route)$	
Γ	no	sunny	no	Rte-113 7	
	no	sunny	yes	Rte-113	
İ	no	cloudy	yes	Rte-75	
İ	yes	sunny	no	Rte-113	
İ	no	raining	no	Rte-113	
İ	no	raining	yes	Rte-75	
ļ	yes	cloudy	yes	Rte-75	
	no	sunny	no	Rte-113	
	yes	sunny	no	Rte-75	
L	no	cloudy	no	Rte-113	
$x_2 = cloudy \to y = Rte\text{-}75$					

counts	Rte-113	Rte-75	
cloudy	1	2	٦
raining	1	1	
sunny	4	1	

 $x_2 = \mathsf{raining} \to y = ????$  $x_2 = \mathsf{sunny} \to y = \mathsf{Rte-113}$ 

)
٦
İ
İ
İ
l

counts	Rte-113	Rte-75
cloudy	1	$2  \rceil$
raining	1	1
sunny	4	1
$\hat{P}(x_2,y)$	Rte-113	Rte-75
cloudy	$\begin{bmatrix} 0.1 \end{bmatrix}$	0.2
raining	0.1	0.1
sunny	0.4	0.1

$$x_2 = \mathsf{cloudy} \to y = \mathsf{Rte-75}$$
  
 $x_2 = \mathsf{raining} \to y = ????$   
 $x_2 = \mathsf{sunny} \to y = \mathsf{Rte-113}$ 

$\hat{P}(x_2,y)$	Rte-113	Rte-75	
cloudy	0.1	0.2	1
raining	0.1	0.1	1
sunny	0.4	0.1	

$\hat{P}(x_2,y)$	Rte-113	Rte-75	$\hat{P}(x_2)$	
cloudy	$\begin{bmatrix} 0.1 \end{bmatrix}$	0.2	cloudy	0.3
raining	0.1	0.1	raining	0.2
sunny	0.4	0.1	sunny	0.5

$$\begin{array}{c|cccc} \hat{P}(x_2,y) & \text{Rte-113} & \text{Rte-75} & & \hat{P}(x_2) \\ \text{cloudy} & \begin{bmatrix} 0.1 & 0.2 \\ 0.1 & 0.1 \\ \end{bmatrix} & \text{cloudy} & \begin{bmatrix} 0.3 \\ 0.2 \\ \text{sunny} \end{bmatrix} \\ \text{sunny} & \begin{bmatrix} 0.4 & 0.1 \\ \end{bmatrix} & \text{sunny} & \begin{bmatrix} 0.5 \\ 0.5 \\ \end{bmatrix}$$

$$\begin{array}{cccc} \hat{P}(y) & & \text{Rte-113} & \text{Rte-75} \\ & \left[ & 0.6 & 0.4 & \right] \end{array}$$

$\hat{P}(x_2,y)$	Rte-113	Rte-75	$\hat{P}(x_2)$	$\hat{P}(y x_2)$	Rte
cloudy	0.1	0.2	cloudy $\begin{bmatrix} 0.3 \end{bmatrix}$	cloudy	0
raining	0.1	0.1	raining 0.2	raining	0
sunny	0.4	0.1	sunny $\left[\begin{array}{c}0.5\end{array}\right]$	sunny	0

$$\begin{array}{ccccc} \hat{P}(y|x_2) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.33 & 0.67 \\ 0.50 & 0.50 \\ 0.80 & 0.20 \\ \end{bmatrix} \\ \end{array}$$

$\hat{P}(x_2,y)$	Rte-113	Rte-75	$\hat{P}(x_2)$
cloudy	$\begin{bmatrix} 0.1 \end{bmatrix}$	0.2	cloudy
raining	0.1	0.1	raining
sunny	0.4	0.1	sunny

$$\begin{array}{ccccc} \hat{P}(x_2) & & \hat{P}(y|x_2) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.3 \\ 0.2 \\ \text{sunny} \end{bmatrix} & & \text{cloudy} & \begin{bmatrix} 0.33 & 0.67 \\ 0.50 & 0.50 \\ \text{sunny} \end{bmatrix} & & \text{cloudy} & \begin{bmatrix} 0.50 & 0.50 \\ 0.80 & 0.20 \end{bmatrix} \end{array}$$

$$\begin{array}{cccc} \hat{P}(y) & \text{Rte-113} & \text{Rte-75} \\ & \left[ & 0.6 & 0.4 & \right] \end{array}$$

$$\begin{array}{cccc} \hat{P}(x_2|y) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \\ \end{bmatrix} \\ \text{sunny} & \begin{bmatrix} 0.67 & 0.25 \\ 0.67 & 0.25 \\ \end{bmatrix}$$

$$\begin{array}{cccc} \hat{P}(y) & \text{Rte-113} & \text{Rte-75} \\ & \left[ & 0.6 & 0.4 & \right] \end{array}$$

$$P(A=a) = \sum_{b} P(A=a, B=b)$$

$$\begin{array}{cccc} \hat{P}(x_2|y) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \\ \end{array} \\ \text{sunny} & \begin{bmatrix} 0.67 & 0.25 \\ 0.67 & 0.25 \\ \end{bmatrix}$$

$$P(A=a) = \sum_{b} P(A=a, B=b)$$

$$\begin{array}{cccc} \hat{P}(x_2|y) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \\ \end{array} \\ \text{sunny} & \begin{bmatrix} 0.67 & 0.25 \\ 0.67 & 0.25 \\ \end{bmatrix}$$

$$P(A=a, B=b) = P(A=a)P(B=b \mid A=a)$$
  
=  $P(B=b)P(A=a \mid B=b)$ 

$$\begin{array}{c|cccc} \hat{P}(x_2) & & \hat{P}(y|x_2) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.3 \\ 0.2 \\ \text{sunny} \end{bmatrix} & \text{cloudy} & \begin{bmatrix} 0.33 & 0.67 \\ 0.50 & 0.50 \\ \text{sunny} \end{bmatrix} & \begin{bmatrix} 0.50 & 0.50 \\ 0.80 & 0.20 \end{bmatrix} \end{array}$$

$$\hat{P}(y)$$
 Rte-113 Rte-75  $\left[ \begin{array}{cc} 0.6 & 0.4 \end{array} \right]$ 

$$P(A=a) = \sum_{b} P(A=a, B=b)$$

$$\begin{array}{cccc} \hat{P}(x_2|y) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \\ \end{array} \\ \text{sunny} & \begin{bmatrix} 0.67 & 0.25 \\ 0.67 & 0.25 \\ \end{bmatrix}$$

$$P(A=a, B=b) = P(A=a)P(B=b \mid A=a)$$
  
=  $P(B=b)P(A=a \mid B=b)$   
 $P(A=a \mid B=b) = \frac{P(A=a, B=b)}{P(B=b)}$ 

$$\hat{P}(y)$$
 Rte-113 Rte-75  $\begin{bmatrix} 0.6 & 0.4 \end{bmatrix}$ 

0.6 0.4 ] 
$$P(A=a) = \sum_b P(A=a, B=b)$$

$$\begin{array}{cccc} \hat{P}(x_2|y) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \\ \end{array} \\ \text{sunny} & \begin{bmatrix} 0.67 & 0.25 \\ 0.67 & 0.25 \\ \end{bmatrix}$$

$$P(A=a, B=b) = P(A=a)P(B=b \mid A=a)$$
  
=  $P(B=b)P(A=a \mid B=b)$   
 $P(A=a \mid B=b) = \frac{P(A=a, B=b)}{P(B=b)}$ 

$$\hat{P}(y)$$
 Rte-113 Rte-75  $\begin{bmatrix} 0.6 & 0.4 \end{bmatrix}$ 

0.6 0.4 ] 
$$P(A=a) = \sum_b P(A=a, B=b)$$

$$P(A=a, B=b) = P(A=a)P(B=b \mid A=a)$$
  
=  $P(B=b)P(A=a \mid B=b)$ 

$$P(A=a \mid B=b) = \frac{P(A=a,B=b)}{P(B=b)}$$

$$P(A=a \mid B=b) = \frac{P(B=b|A=a)P(A=a)}{P(B=b)}$$
 (Bayes' Theorem)

### Bayes-optimal classifier

$$(x_1 = no, x_2 = sunny, x_3 = no) \rightarrow y = ????$$

$$(x_1 = \text{no}, x_2 = \text{sunny}, x_3 = \text{no}) \to y = ????$$
 Bayes-optimal classifier: Given  $x = (x_1, x_2, \dots, x_n)$ , predict  $y = \arg\max_c P(y = c \mid x)$ 

```
(x_1 = \text{no}, x_2 = \text{sunny}, x_3 = \text{no}) \to y = ????
Bayes-optimal classifier: Given x = (x_1, x_2, \dots, x_n), predict y = \arg\max_c P(y = c \mid x)
Using estimated probabilities: Given x = (x_1, x_2, \dots, x_n), predict y = \arg\max_c \hat{P}(y = c \mid x)
```

### Bayes-optimal classifier

	$x_1$ (weekend?)	$x_2(weather)$	$x_3(game?)$	$y(faster\ route)$
Γ	no	sunny	no	Rte-113 7
	no	sunny	yes	Rte-113
١	no	cloudy	yes	Rte-75
İ	yes	sunny	no	Rte-113
İ	no	raining	no	Rte-113
İ	no	raining	yes	Rte-75
	yes	cloudy	yes	Rte-75
	no	sunny	no	Rte-113
	yes	sunny	no	Rte-75
L	no	cloudy	no	Rte-113

```
(x_1=\operatorname{no},x_2=\operatorname{sunny},x_3=\operatorname{no}) 	o y=???? Bayes-optimal classifier: Given x=(x_1,x_2,\ldots,x_n), predict y=\arg\max_c P(y=c\mid x) Using estimated probabilities: Given x=(x_1,x_2,\ldots,x_n), predict y=\arg\max_c \hat{P}(y=c\mid x)
```

#### Bayes-optimal classifier

	$x_1$ (weekend?)	$x_2$ (weather)	$x_3(game?)$	$y(faster\ route)$	$\hat{P}(x,y)$	Rte-113	Rte-75
ſ	no	sunny	no	Rte-113	(no,cloudy,no)	0.1	0.0
	no	sunny	yes	Rte-113	(no,cloudy,yes)	0.0	0.1
İ	no	cloudy	yes	Rte-75	(no,raining,no)	0.1	0.0
İ	yes	sunny	no	Rte-113	(no,raining,yes)	0.0	0.1
İ	no	raining	no	Rte-113	(no,sunny,no)	0.2	0.0
İ	no	raining	yes	Rte-75	(no,sunny,yes)	0.1	0.0
	yes	cloudy	yes	Rte-75	(yes,cloudy,no)	0.0	0.0
	no	sunny	no	Rte-113	(yes,cloudy,yes)	0.0	0.1
	yes	sunny	no	Rte-75	(yes,raining,no)	0.0	0.0
	no	cloudy	no	Rte-113	(yes,raining,yes)	0.0	0.0
				•	(yes,sunny,no)	0.1	0.1
					(yes,sunny,yes)	0.0	0.0

 $(x_1 = \text{no}, x_2 = \text{sunny}, x_3 = \text{no}) \rightarrow y = ????$ Bayes-optimal classifier: Given  $x = (x_1, x_2, \dots, x_n)$ , predict  $y = \arg n$ 

Bayes-optimal classifier: Given  $x=(x_1,x_2,\ldots,x_n)$ , predict  $y=\arg\max_c P(y=c\mid x)$ 

Using estimated probabilities: Given  $x = (x_1, x_2, \dots, x_n)$ , predict  $y = \arg\max_c \hat{P}(y = c \mid x)$ 

Assume:  $P(x_1, x_2, \dots, x_n \mid y) = \prod_{i=1}^n P(x_i \mid y)$ 

 $[X_i \text{ is independent of } X_j, \text{ given } Y]$ 

Assume:  $P(x_1, x_2, \dots, x_n \mid y) = \prod_{i=1}^n P(x_i \mid y)$ 

 $[X_i \text{ is independent of } X_j, \text{ given } Y]$ 

Then use  $\hat{P}(x_i \mid y)$  instead of  $P(x_i \mid y)$ .

Assume:  $P(x_1, x_2, \dots, x_n \mid y) = \prod_{i=1}^n P(x_i \mid y)$ 

 $[X_i \text{ is independent of } X_j, \text{ given } Y]$ 

Then use  $\hat{P}(x_i \mid y)$  instead of  $P(x_i \mid y)$ .

$$\hat{P}(y \mid x) = \frac{\hat{P}(x \mid y)\hat{P}(y)}{\hat{P}(x)}$$
$$= \frac{\hat{P}(y)\prod_{i=1}^{n}\hat{P}(x_i \mid y)}{\hat{P}(x)}$$

Assume:  $P(x_1, x_2, \dots, x_n \mid y) = \prod_{i=1}^n P(x_i \mid y)$ 

 $[X_i \text{ is independent of } X_j, \text{ given } Y]$ 

Then use  $\hat{P}(x_i \mid y)$  instead of  $P(x_i \mid y)$ .

$$\hat{P}(y \mid x) = \frac{\hat{P}(x \mid y)\hat{P}(y)}{\hat{P}(x)}$$
$$= \frac{\hat{P}(y)\prod_{i=1}^{n} \hat{P}(x_i \mid y)}{\hat{P}(x)}$$

Naïve Bayes rule:

Given 
$$x = (x_1, x_2, \dots, x_n)$$
,

Pick c that maximizes  $\hat{P}(y=c)\prod_{i=1}^{n}\hat{P}(x_i\mid y=c)$ 

	$x_1$ (weekend?)	$x_2$ (weather)	$x_3(game?)$	$y(faster\ route)$
Γ	no	sunny	no	Rte-113 7
	no	sunny	yes	Rte-113
İ	no	cloudy	yes	Rte-75
İ	yes	sunny	no	Rte-113
İ	no	raining	no	Rte-113
	no	raining	yes	Rte-75
	yes	cloudy	yes	Rte-75
	no	sunny	no	Rte-113
l	yes	sunny	no	Rte-75
L	no	cloudy	no	Rte-113

	$x_1$ (weekend?)	$x_2(weather)$	$x_3(game?)$	$y(faster\ route)$
Γ	no	sunny	no	Rte-113 7
	no	sunny	yes	Rte-113
İ	no	cloudy	yes	Rte-75
İ	yes	sunny	no	Rte-113
İ	no	raining	no	Rte-113
İ	no	raining	yes	Rte-75
	yes	cloudy	yes	Rte-75
	no	sunny	no	Rte-113
	yes	sunny	no	Rte-75
L	no	cloudy	no	Rte-113

$ \begin{bmatrix} 0.6 & 0.4 \end{bmatrix} $ Rte-113 Rte-75 no $ \begin{bmatrix} 0.83 & 0.50 \\ 0.17 & 0.50 \end{bmatrix} $ Rte-113 Rte-75 cloudy $ \begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \end{bmatrix} $ Rte-113 Rte-75 no $ \begin{bmatrix} 0.83 & 0.25 \\ 0.17 & 0.75 \end{bmatrix} $	$\hat{P}(y)$	Rte-113	Rte-75	i
$\begin{array}{c} \text{no} \\ \text{yes} \end{array} \left[ \begin{array}{ccc} 0.83 & 0.50 \\ 0.17 & 0.50 \end{array} \right] \\ \hat{P}(x_2 y) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & 0.17 & 0.50 \\ 0.17 & 0.25 \\ \text{sunny} & 0.67 & 0.25 \end{array} \right] \\ \hat{P}(x_3 y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \left[ \begin{array}{ccc} 0.83 & 0.25 \end{array} \right] \end{array}$	[	0.6	0.4	]
$\begin{array}{c} \text{no} \\ \text{yes} \end{array} \left[ \begin{array}{ccc} 0.83 & 0.50 \\ 0.17 & 0.50 \end{array} \right] \\ \hat{P}(x_2 y) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & 0.17 & 0.50 \\ 0.17 & 0.25 \\ \text{sunny} & 0.67 & 0.25 \end{array} \right] \\ \hat{P}(x_3 y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \left[ \begin{array}{ccc} 0.83 & 0.25 \end{array} \right] \end{array}$				
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$\hat{P}(x_1 y)$	Rte-11	3 Rte	-75
$ \hat{P}(x_2 y)  \text{Rte-113}  \text{Rte-75} \\ \text{cloudy}  \begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \end{bmatrix} \\ \hat{P}(x_3 y)  \text{Rte-113}  \text{Rte-75} \\ \text{no}  \begin{bmatrix} 0.83 & 0.25 \end{bmatrix} $	no	[0.83]	0.	50 ]
cloudy raining sunny $\begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \end{bmatrix}$ $\hat{P}(x_3 y)$ Rte-113 Rte-75 no $\begin{bmatrix} 0.83 & 0.25 \end{bmatrix}$	yes	0.17	0.	50
cloudy raining sunny $\begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \end{bmatrix}$ $\hat{P}(x_3 y)$ Rte-113 Rte-75 no $\begin{bmatrix} 0.83 & 0.25 \end{bmatrix}$		_		_
$ \begin{array}{c cccc} {\rm raining} & 0.17 & 0.25 \\ {\rm sunny} & 0.67 & 0.25 \\ \end{array} \\ \hat{P}(x_3 y) & {\rm Rte-113} & {\rm Rte-75} \\ {\rm no} & \left[ \begin{array}{cccc} 0.83 & 0.25 \end{array} \right] $	$\hat{P}(x_2 y)$	Rte-11	l3 Rt	e-75
sunny $\begin{bmatrix} 0.67 & 0.25 \end{bmatrix}$ $\hat{P}(x_3 y)$ Rte-113 Rte-75 no $\begin{bmatrix} 0.83 & 0.25 \end{bmatrix}$	cloudy	[0.17]	7 0	.50
$\hat{P}(x_3 y)$ Rte-113 Rte-75 no $\left[ egin{array}{ccc} 0.83 & 0.25 \end{array}  ight]$	raining	0.17	7 0	.25
no $\begin{bmatrix} 0.83 & 0.25 \end{bmatrix}$	sunny	0.67	7 0	.25
no $\begin{bmatrix} 0.83 & 0.25 \end{bmatrix}$				
.	$\hat{P}(x_3 y)$	Rte-11	3 Rte	-75
yes $0.17  0.75$	no	[0.83]	0.	25
	yes	[0.17]	0.	75

Input: 
$$x = (no, cloudy, no)$$

Input: 
$$x = (no, cloudy, no)$$

Possibility: 
$$y = \text{Rte-}113$$

$$\hat{P}(y)$$
 Rte-113 Rte-75  $\left[ egin{array}{ccc} 0.6 & 0.4 \end{array} 
ight]$ 

$$\begin{array}{ccc} \hat{P}(x_1|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \left[ \begin{array}{ccc} 0.83 & 0.50 \\ 0.17 & 0.50 \end{array} \right] \end{array}$$

$$\begin{array}{cccc} \hat{P}(x_2|y) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \\ \end{array} \\ \text{sunny} & \begin{bmatrix} 0.67 & 0.25 \\ 0.67 & 0.25 \\ \end{bmatrix}$$

$$\begin{array}{cccc} \hat{P}(x_3|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.25 \\ 0.17 & 0.75 \\ \end{array}$$

Input: 
$$x = (no, cloudy, no)$$

Possibility: y = Rte-113

$$\hat{P}(y = \mathsf{Rte-113}) \hat{P}(x_1 = \mathsf{no} \mid y = \mathsf{Rte-113}) \hat{P}(x_2 = \mathsf{cloudy} \mid y = \mathsf{Rte-113}) \hat{P}(x_3 = \mathsf{no} \mid y = \mathsf{Rte-113})$$

$$\hat{P}(y)$$
 Rte-113 Rte-75  $\begin{bmatrix} 0.6 & 0.4 \end{bmatrix}$ 

$$\begin{array}{ccccc} \hat{P}(x_1|y) & & {\rm Rte\text{-}}113 & {\rm Rte\text{-}}75 \\ & {\rm no} & \left[ & 0.83 & 0.50 \\ & {\rm yes} & \left[ & 0.17 & 0.50 \right. \end{array} \right]$$

$$\begin{array}{cccc} \hat{P}(x_2|y) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \\ \end{array} \\ \text{sunny} & \begin{bmatrix} 0.67 & 0.25 \\ 0.67 & 0.25 \\ \end{bmatrix}$$

$$\begin{array}{cccc} \hat{P}(x_3|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.25 \\ 0.17 & 0.75 \\ \end{array}$$

Input: 
$$x = (no, cloudy, no)$$

Possibility: y = Rte-113

$$\hat{P}(y)$$
 Rte-113 Rte-75  $\begin{bmatrix} 0.6 & 0.4 \end{bmatrix}$ 

$$\begin{array}{cccc} \hat{P}(x_1|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.50 \\ 0.17 & 0.50 \end{array} \\ \end{array}$$

$$\begin{array}{cccc} \hat{P}(x_3|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.25 \\ 0.17 & 0.75 \\ \end{array}$$

Input: x = (no, cloudy, no)

Possibility: y = Rte-113

$$\begin{array}{l} \hat{P}(y=\text{Rte-113})\hat{P}(x_1=\text{no}\mid y=\text{Rte-113})\hat{P}(x_2=\text{cloudy}\mid y=\text{Rte-113})\hat{P}(x_3=\text{no}\mid y=\text{Rte-113})\\ =0.6\times0.83\times0.17\times0.83=0.0694 \end{array}$$

Possibility: y = Rte-75

$$\begin{array}{cccc} \hat{P}(x_1|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.50 \\ 0.17 & 0.50 \end{array} \\ \end{array}$$

$$\begin{array}{cccc} \hat{P}(x_2|y) & \text{Rte-113} & \text{Rte-75} \\ \text{cloudy} & \begin{bmatrix} 0.17 & 0.50 \\ 0.17 & 0.25 \\ 0.67 & 0.25 \\ \end{array} \\ \text{sunny} & \begin{bmatrix} 0.67 & 0.25 \\ 0.67 & 0.25 \\ \end{bmatrix}$$

$$\begin{array}{cccc} \hat{P}(x_3|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.25 \\ 0.17 & 0.75 \\ \end{array}$$

Input: 
$$x = (no, cloudy, no)$$

Possibility: y = Rte-113

$$\begin{array}{l} \hat{P}(y=\text{Rte-113})\hat{P}(x_1=\text{no}\mid y=\text{Rte-113})\hat{P}(x_2=\text{cloudy}\mid y=\text{Rte-113})\hat{P}(x_3=\text{no}\mid y=\text{Rte-113})\\ =0.6\times0.83\times0.17\times0.83=0.0694 \end{array}$$

Possibility: y = Rte-75

$$\hat{P}(y = \mathsf{Rte-75}) \hat{P}(x_1 = \mathsf{no} \mid y = \mathsf{Rte-75}) \hat{P}(x_2 = \mathsf{cloudy} \mid y = \mathsf{Rte-75}) \hat{P}(x_3 = \mathsf{no} \mid y = \mathsf{Rte-75})$$

$$\begin{array}{cccc} \hat{P}(x_1|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.50 \\ 0.17 & 0.50 \end{array} \\ \end{array}$$

$$\begin{array}{cccc} \hat{P}(x_3|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.25 \\ 0.17 & 0.75 \\ \end{array}$$

Input: 
$$x = (no, cloudy, no)$$

Possibility: y = Rte-113

$$\begin{array}{l} \hat{P}(y=\!\!\!\text{Rte-113})\hat{P}(x_1=\!\!\!\!\text{no}\mid y=\!\!\!\!\text{Rte-113})\hat{P}(x_2=\!\!\!\!\text{cloudy}\mid y=\!\!\!\!\!\text{Rte-113})\hat{P}(x_3=\!\!\!\!\!\text{no}\mid y=\!\!\!\!\!\!\text{Rte-113})\\ = 0.6\times0.83\times0.17\times0.83 = 0.0694 \end{array}$$

Possibility: y = Rte-75

$$\hat{P}(y)$$
 Rte-113 Rte-75  $\begin{bmatrix} 0.6 & 0.4 \end{bmatrix}$ 

$$\begin{array}{cccc} \hat{P}(x_1|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.50 \\ 0.17 & 0.50 \end{array} \\ \end{array}$$

$$\begin{array}{cccc} \hat{P}(x_3|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.25 \\ 0.17 & 0.75 \\ \end{array}$$

Input: 
$$x = (no, cloudy, no)$$

Possibility: 
$$y = \text{Rte-}113$$

$$\begin{split} &\hat{P}(y = \text{Rte-113}) \hat{P}(x_1 = \text{no} \mid y = \text{Rte-113}) \hat{P}(x_2 = \text{cloudy} \mid y = \text{Rte-113}) \hat{P}(x_3 = \text{no} \mid y = \text{Rte-113}) \\ &= 0.6 \times 0.83 \times 0.17 \times 0.83 = 0.0694 \end{split}$$

Possibility: 
$$y = \text{Rte-75}$$

$$\begin{array}{l} \hat{P}(y=\text{Rte-75})\hat{P}(x_1=\text{no} \mid y=\text{Rte-75})\hat{P}(x_2=\text{cloudy} \mid y=\text{Rte-75})\hat{P}(x_3=\text{no} \mid y=\text{Rte-75}) \\ = 0.4 \times 0.50 \times 0.50 \times 0.25 = 0.025 \end{array}$$

#### Therefore, the rule reports Rte-113

$$\begin{array}{cccc} \hat{P}(x_1|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \left[ \begin{array}{ccc} 0.83 & 0.50 \\ 0.17 & 0.50 \end{array} \right] \end{array}$$

$$\begin{array}{cccc} \hat{P}(x_3|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.25 \\ 0.17 & 0.75 \\ \end{array}$$

Input: 
$$x = (yes, raining, no)$$

Possibility: y = Rte-113

Possibility: y = Rte-75

$$\begin{array}{l} \hat{P}(y=\!\!\!\text{Rte-75}) \hat{P}(x_1=\!\!\!\text{yes} \mid y=\!\!\!\text{Rte-75}) \hat{P}(x_2=\!\!\!\!\text{raining} \mid y=\!\!\!\!\text{Rte-75}) \hat{P}(x_3=\!\!\!\!\text{no} \mid y=\!\!\!\!\!\text{Rte-75}) \\ = 0.4 \times 0.50 \times 0.25 \times 0.25 = 0.0125 \end{array}$$

Therefore, the rule reports Rte-113

$$\begin{array}{cccc} \hat{P}(x_1|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \left[ \begin{array}{ccc} 0.83 & 0.50 \\ 0.17 & 0.50 \end{array} \right. \\ \end{array}$$

$$\begin{array}{ccccc} \hat{P}(x_3|y) & & {\rm Rte\text{-}}113 & {\rm Rte\text{-}}75 \\ & {\rm no} & \left[ & 0.83 & 0.25 \\ & {\rm yes} & \left[ & 0.17 & 0.75 \end{array} \right. \end{array}$$

Input: x = (yes, sunny, yes)

Possibility: y = Rte-113

Possibility: y = Rte-75

$$\begin{array}{l} \hat{P}(y=\text{Rte-75})\hat{P}(x_1=\text{yes} \mid y=\text{Rte-75})\hat{P}(x_2=\text{sunny} \mid y=\text{Rte-75})\hat{P}(x_3=\text{yes} \mid y=\text{Rte-75}) \\ = 0.4 \times 0.50 \times 0.25 \times 0.75 = 0.0375 \end{array}$$

Therefore, the rule reports Rte-75

$$\begin{array}{cccc} \hat{P}(x_1|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \left[ \begin{array}{ccc} 0.83 & 0.50 \\ 0.17 & 0.50 \end{array} \right. \end{array}$$

$$\begin{array}{cccc} \hat{P}(x_3|y) & \text{Rte-113} & \text{Rte-75} \\ \text{no} & \begin{bmatrix} 0.83 & 0.25 \\ 0.17 & 0.75 \\ \end{array}$$