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

A toy problem

x_1(weekend?)	x_2(weather)	x_3(game?)	y(faster route)
no	sunny	no	Rte-113
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

A toy problem

x_1(weekend?)	x_2(weather)	x_3(game?)	y(faster route)
no	sunny	no	Rte-113
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

A toy problem

x_1(weekend?)	x_2(weather)	x_3(game?)	y(faster route)
no	sunny	no	Rte-113
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

$x_2 = \text{cloudy} \rightarrow y = ????$

A toy problem

x_1(weekend?)	x_2(weather)	x_3(game?)	y(faster route)
no	sunny	no	Rte-113
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

$x_2 = \text{cloudy} \rightarrow y = ????$

$x_2 = \text{raining} \rightarrow y = ????$

$x_2 = \text{sunny} \rightarrow y = ????$

A toy problem

x_1(weekend?)	x_2(weather)	x_3(game?)	y(faster route)
no	sunny	no	Rte-113
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

$x_2 = \text{cloudy} \rightarrow y = \text{Rte-75}$

$x_2 = \text{raining} \rightarrow y = ????$

$x_2 = \text{sunny} \rightarrow y = \text{Rte-113}$

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

A toy problem

x_1(weekend?)	x_2(weather)	x_3(game?)	y(faster route)
no	sunny	no	Rte-113
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

$x_2 = \text{cloudy} \rightarrow y = \text{Rte-75}$

$x_2 = \text{raining} \rightarrow y = \text{????}$

$x_2 = \text{sunny} \rightarrow y = \text{Rte-113}$

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

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

Bayes' rule

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

Bayes' rule

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

Bayes' rule

$\hat{P}(x_2, y)$

	Rte-113	Rte-75
cloudy	0.1	0.2
raining	0.1	0.1
sunny	0.4	0.1

$\hat{P}(x_2)$

cloudy	0.3
raining	0.2
sunny	0.5

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

Bayes' rule

$\hat{P}(x_2,y)$

	Rte-113	Rte-75
cloudy	0.1	0.2
raining	0.1	0.1
sunny	0.4	0.1

$\hat{P}(x_2)$

cloudy	0.3
raining	0.2
sunny	0.5

$\hat{P}(y|x_2)$

	Rte-113	Rte-75
cloudy	0.33	0.67
raining	0.50	0.50
sunny	0.80	0.20

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

Bayes' rule

$\hat{P}(x_2,y)$

	Rte-113	Rte-75
cloudy	0.1	0.2
raining	0.1	0.1
sunny	0.4	0.1

$\hat{P}(x_2)$

cloudy	0.3
raining	0.2
sunny	0.5

$\hat{P}(y|x_2)$

	Rte-113	Rte-75
cloudy	0.33	0.67
raining	0.50	0.50
sunny	0.80	0.20

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

Bayes' rule

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

$\hat{P}(x_2)$	
cloudy	0.3
raining	0.2
sunny	0.5

$\hat{P}(y\|x_2)$	Rte-113	Rte-75
cloudy	0.33	0.67
raining	0.50	0.50
sunny	0.80	0.20

$\hat{P}(y)$	Rte-113	Rte-75
	0.6	0.4

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

$\hat{P}(x_2\|y)$	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

Bayes' rule

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

$\hat{P}(x_2)$	
cloudy	0.3
raining	0.2
sunny	0.5

$\hat{P}(y\|x_2)$	Rte-113	Rte-75
cloudy	0.33	0.67
raining	0.50	0.50
sunny	0.80	0.20

$\hat{P}(y)$	Rte-113	Rte-75
	0.6	0.4

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

$\hat{P}(x_2\|y)$	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$$\begin{aligned} P(A{=}a, B{=}b) &= P(A{=}a)P(B{=}b \mid A{=}a) \\ &= P(B{=}b)P(A{=}a \mid B{=}b) \end{aligned}$$

Bayes' rule

$\hat{P}(x_2,y)$

	Rte-113	Rte-75
cloudy	0.1	0.2
raining	0.1	0.1
sunny	0.4	0.1

$\hat{P}(x_2)$

cloudy	0.3
raining	0.2
sunny	0.5

$\hat{P}(y|x_2)$

	Rte-113	Rte-75
cloudy	0.33	0.67
raining	0.50	0.50
sunny	0.80	0.20

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

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

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$$\begin{aligned} P(A{=}a, B{=}b) &= P(A{=}a)P(B{=}b \mid A{=}a) \\ &= P(B{=}b)P(A{=}a \mid B{=}b) \end{aligned}$$

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

Bayes' rule

$\hat{P}(x_2,y)$

	Rte-113	Rte-75
cloudy	0.1	0.2
raining	0.1	0.1
sunny	0.4	0.1

$\hat{P}(x_2)$

cloudy	0.3
raining	0.2
sunny	0.5

$\hat{P}(y|x_2)$

	Rte-113	Rte-75
cloudy	0.33	0.67
raining	0.50	0.50
sunny	0.80	0.20

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

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

$$\begin{aligned} P(A{=}a, B{=}b) &= P(A{=}a)P(B{=}b \mid A{=}a) \\ &= P(B{=}b)P(A{=}a \mid B{=}b) \end{aligned}$$

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

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

Bayes' rule

$\hat{P}(x_2,y)$

	Rte-113	Rte-75
cloudy	0.1	0.2
raining	0.1	0.1
sunny	0.4	0.1

$\hat{P}(x_2)$

cloudy	0.3
raining	0.2
sunny	0.5

$\hat{P}(y|x_2)$

	Rte-113	Rte-75
cloudy	0.33	0.67
raining	0.50	0.50
sunny	0.80	0.20

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

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

$$\begin{aligned} P(A{=}a, B{=}b) &= P(A{=}a)P(B{=}b \mid A{=}a) \\ &= P(B{=}b)P(A{=}a \mid B{=}b) \end{aligned}$$

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

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

Bayes' rule

$\hat{P}(x_2,y)$

	Rte-113	Rte-75
cloudy	0.1	0.2
raining	0.1	0.1
sunny	0.4	0.1

$\hat{P}(x_2)$

cloudy	0.3
raining	0.2
sunny	0.5

$\hat{P}(y|x_2)$

	Rte-113	Rte-75
cloudy	0.33	0.67
raining	0.50	0.50
sunny	0.80	0.20

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

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

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

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

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

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

Example: $\hat{P}(\text{Rte-113} \mid \text{sunny}) = \frac{\hat{P}(\text{sunny}|\text{Rte-113})\hat{P}(\text{Rte-113})}{\hat{P}(\text{sunny})} = \frac{0.67 \times 0.6}{0.5} = 0.8$

Bayes-optimal classifier

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

Bayes-optimal classifier

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

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

Bayes-optimal classifier

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

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)
no	sunny	no	Rte-113
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

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

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)
no	sunny	no	Rte-113
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

$\hat{P}(x,y)$	Rte-113	Rte-75
(no,cloudy,no)	0.1	0.0
(no,cloudy,yes)	0.0	0.1
(no,raining,no)	0.1	0.0
(no,raining,yes)	0.0	0.1
(no,sunny,no)	0.2	0.0
(no,sunny,yes)	0.1	0.0
(yes,cloudy,no)	0.0	0.0
(yes,cloudy,yes)	0.0	0.1
(yes,raining,no)	0.0	0.0
(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 $x = (x_1, x_2, \ldots, x_n)$, predict $y = \arg\max_c P(y = c \mid x)$

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

Naïve Bayes

Assume: $P(x_1, x_2, \ldots, x_n \mid y) = \prod_{i=1}^{n} P(x_i \mid y)$ [X_i is independent of X_j, given Y]

Naïve Bayes

Assume: $P(x_1, x_2, \ldots, x_n \mid y) = \prod_{i=1}^{n} P(x_i \mid y)$ [X_i is independent of X_j, given Y]

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

Naïve Bayes

Assume: $P(x_1, x_2, \ldots, x_n \mid y) = \prod_{i=1}^{n} P(x_i \mid y)$ [X_i is independent of X_j, given Y]

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

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

Naïve Bayes

Assume: $P(x_1, x_2, \ldots, x_n \mid y) = \prod_{i=1}^{n} P(x_i \mid y)$ [X_i is independent of X_j, given Y]

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

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

Naïve Bayes rule:
Given $x = (x_1, x_2, \ldots, x_n)$,
Pick c that maximizes $\hat{P}(y{=}c)\prod_{i=1}^{n}\hat{P}(x_i \mid y{=}c)$

Naïve Bayes Example

x_1(weekend?)	x_2(weather)	x_3(game?)	y(faster route)
no	sunny	no	Rte-113
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

Naïve Bayes Example

x_1(weekend?)	x_2(weather)	x_3(game?)	y(faster route)
no	sunny	no	Rte-113
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

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

$\hat{P}(x_1|y)$

	Rte-113	Rte-75
no	0.83	0.50
yes	0.17	0.50

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$\hat{P}(x_3|y)$

	Rte-113	Rte-75
no	0.83	0.25
yes	0.17	0.75

Naïve Bayes Example

Input: $x = (\text{no}, \text{cloudy}, \text{no})$

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

$\hat{P}(x_1|y)$

	Rte-113	Rte-75
no	0.83	0.50
yes	0.17	0.50

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$\hat{P}(x_3|y)$

	Rte-113	Rte-75
no	0.83	0.25
yes	0.17	0.75

Naïve Bayes Example

Input: $x = (\text{no}, \text{cloudy}, \text{no})$

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

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

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

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

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

Naïve Bayes Example

Input: $x = (\text{no}, \text{cloudy}, \text{no})$

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

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

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

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

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

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

Naïve Bayes Example

Input: $x = (\text{no}, \text{cloudy}, \text{no})$

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

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

$= 0.6 \times 0.83 \times 0.17 \times 0.83 = 0.0694$

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

$\hat{P}(x_1|y)$

	Rte-113	Rte-75
no	0.83	0.50
yes	0.17	0.50

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$\hat{P}(x_3|y)$

	Rte-113	Rte-75
no	0.83	0.25
yes	0.17	0.75

Naïve Bayes Example

Input: $x = (\text{no}, \text{cloudy}, \text{no})$

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

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

$= 0.6 \times 0.83 \times 0.17 \times 0.83 = 0.0694$

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

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

$\hat{P}(x_1|y)$

	Rte-113	Rte-75
no	0.83	0.50
yes	0.17	0.50

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$\hat{P}(x_3|y)$

	Rte-113	Rte-75
no	0.83	0.25
yes	0.17	0.75

Naïve Bayes Example

Input: $x = (\text{no}, \text{cloudy}, \text{no})$

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

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

$= 0.6 \times 0.83 \times 0.17 \times 0.83 = 0.0694$

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

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

$\hat{P}(y)$	Rte-113	Rte-75
	0.6	0.4

$\hat{P}(x_1 \mid y)$	Rte-113	Rte-75
no	0.83	0.50
yes	0.17	0.50

$\hat{P}(x_2 \mid y)$	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$\hat{P}(x_3 \mid y)$	Rte-113	Rte-75
no	0.83	0.25
yes	0.17	0.75

Naïve Bayes Example

Input: $x = (\text{no}, \text{cloudy}, \text{no})$

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

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

$= 0.6 \times 0.83 \times 0.17 \times 0.83 = 0.0694$

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

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

$= 0.4 \times 0.50 \times 0.50 \times 0.25 = 0.025$

$\hat{P}(y)$	Rte-113	Rte-75
	0.6	0.4

$\hat{P}(x_1 \mid y)$	Rte-113	Rte-75
no	0.83	0.50
yes	0.17	0.50

$\hat{P}(x_2 \mid y)$	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$\hat{P}(x_3 \mid y)$	Rte-113	Rte-75
no	0.83	0.25
yes	0.17	0.75

Naïve Bayes Example

Input: $x = (\text{no}, \text{cloudy}, \text{no})$

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

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

$= 0.6 \times 0.83 \times 0.17 \times 0.83 = 0.0694$

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

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

$= 0.4 \times 0.50 \times 0.50 \times 0.25 = 0.025$

Therefore, the rule reports Rte-113

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

$\hat{P}(x_1|y)$

	Rte-113	Rte-75
no	0.83	0.50
yes	0.17	0.50

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$\hat{P}(x_3|y)$

	Rte-113	Rte-75
no	0.83	0.25
yes	0.17	0.75

Naïve Bayes Example II

Input: $x = (\text{yes}, \text{raining}, \text{no})$

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

$\hat{P}(y{=}\text{Rte-113})\hat{P}(x_1{=}\text{yes} \mid y{=}\text{Rte-113})\hat{P}(x_2{=}\text{raining} \mid y{=}\text{Rte-113})\hat{P}(x_3{=}\text{no} \mid y{=}\text{Rte-113})$

$= 0.6 \times 0.17 \times 0.17 \times 0.83 = 0.0139$

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

$\hat{P}(y{=}\text{Rte-75})\hat{P}(x_1{=}\text{yes} \mid y{=}\text{Rte-75})\hat{P}(x_2{=}\text{raining} \mid y{=}\text{Rte-75})\hat{P}(x_3{=}\text{no} \mid y{=}\text{Rte-75})$

$= 0.4 \times 0.50 \times 0.25 \times 0.25 = 0.0125$

Therefore, the rule reports Rte-113

$\hat{P}(y)$

Rte-113	Rte-75
0.6	0.4

$\hat{P}(x_1|y)$

	Rte-113	Rte-75
no	0.83	0.50
yes	0.17	0.50

$\hat{P}(x_2|y)$

	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$\hat{P}(x_3|y)$

	Rte-113	Rte-75
no	0.83	0.25
yes	0.17	0.75

Naïve Bayes Example III

Input: $x = (\text{yes}, \text{sunny}, \text{yes})$

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

$\hat{P}(y{=}\text{Rte-113})\hat{P}(x_1{=}\text{yes} \mid y{=}\text{Rte-113})\hat{P}(x_2{=}\text{sunny} \mid y{=}\text{Rte-113})\hat{P}(x_3{=}\text{yes} \mid y{=}\text{Rte-113})$

$= 0.6 \times 0.17 \times 0.67 \times 0.17 = 0.0111$

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

$\hat{P}(y{=}\text{Rte-75})\hat{P}(x_1{=}\text{yes} \mid y{=}\text{Rte-75})\hat{P}(x_2{=}\text{sunny} \mid y{=}\text{Rte-75})\hat{P}(x_3{=}\text{yes} \mid y{=}\text{Rte-75})$

$= 0.4 \times 0.50 \times 0.25 \times 0.75 = 0.0375$

Therefore, the rule reports Rte-75

$\hat{P}(y)$	Rte-113	Rte-75
	0.6	0.4

$\hat{P}(x_1 \mid y)$	Rte-113	Rte-75
no	0.83	0.50
yes	0.17	0.50

$\hat{P}(x_2 \mid y)$	Rte-113	Rte-75
cloudy	0.17	0.50
raining	0.17	0.25
sunny	0.67	0.25

$\hat{P}(x_3 \mid y)$	Rte-113	Rte-75
no	0.83	0.25
yes	0.17	0.75