✯ Naive Bayes

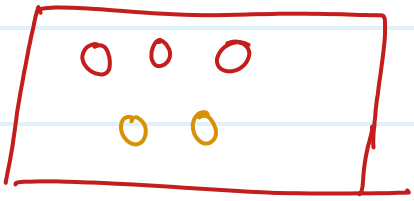
It works on probability based approch, and It is used to solve classification problem.

Independent event

coin toss (H) = $\frac{1}{2}$

Dice = $\frac{1}{6}$

Dependent probability



$Pr(R) = \frac{3}{5}$

After remove one red malbel

$Pr(Y) = \frac{2}{4}$

conditional probability

R = Red
G = Green

$$Pr(R \text{ and } G) = P(R) \times P(G/R)$$

formula

$$Pr(A \text{ and } B) = P(A) \times P(B/A)$$

$$Pr(A \text{ and } B) = Pr(B \text{ and } A)$$

$$P(A) \times P(B/A) = Pr(B) \times P(A/B)$$

formula

$$P(B/A) = \frac{Pr(B) \times P(A/B)}{P(A)}$$

Baye's Theorem

Example

$\overbrace{x_1\ x_2\ x_3\ \cdots\ x_n}^{A}$ $\quad B$

y

$$P(Y/(x_1 x_2 x_3 \cdots x_n)) = \frac{P(Y) \times P[(x_1 x_2 x_3 \cdots x_n)/Y]}{P(x_1 x_2 x_3 \cdots x_n)}$$

Y = Yes / No

Binary class

$$P[Yes/(x_1, x_2 x_3)] = \frac{P(Yes) \times P(x_1/Yes) \times P(x_2/Yes) \times P(x_3/Yes)}{P(x_1)\ P(x_2)\ P(x_3)}$$

constant

$$P[No/(x_1, x_2 x_3)] = \frac{P(No) \times P(x_1/No) \times P(x_2/No) \times P(x_3/No)}{P(x_1)\ P(x_2)\ P(x_3)}$$

constant

Example - Dataset

outlook, x_1	Temp. x_2	humidity x_3	wind x_4	play tennis y

★ outlook

	yes	no	P(y)	P(N)
sunny	2	3	2/9	3/5
overcast	3	0	3/9	0/5
Rain	4	2	4/9	2/5
	9	5		

$\frac{9}{14}$ $\frac{5}{14}$

Temp

	Yes	No	P(Y)	P(N)
Hot	2	2	2/9	2/5
mild	4	2	4/9	2/5
cold	3	1	3/9	1/5
	9	5		

$$P(yes) = \frac{9}{14} \qquad P(No) = \frac{5}{14}$$

For new dataset

Test (sunny, hot) = play ??

$$p(Y/sunny, hot) = p(Y) \times p(sunny/y) \times p(hot/y)$$

$$p(N/sunny, hot) = p(N) \times p(sunny/No) \times p(hot/N)$$

$$p(Y/sunny, hot) = \frac{9}{14} \times \frac{2}{9} \times \frac{2}{9} = 0.031$$

$$p(N/sunny, hot) = \frac{5}{14} \times \frac{3}{5} \times \frac{2}{5} = 0.085$$

$$p(Y/sunny, hot) = \frac{0.031}{0.031 + 0.085}$$

$$= 0.27 = 27\%$$

$$P(N/Sunny, hot) = \frac{0.085}{0.085 + 0.031}$$

$$= 0.73 \Rightarrow 73\%$$

// Naive bayes's classifier used for NLP (natural language processing)

NLTK - Natural Language Toolkit

RE - Regular expression.

He is eating.

- - —

He

He