Supervised learning

① Naive Bayes

↓ classification

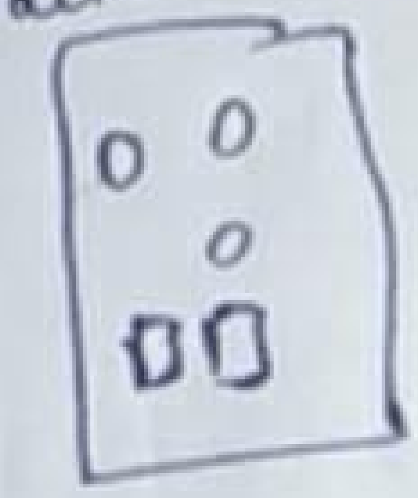
Works on completely conditional probability.

① Condition probability
② independent event → toss 2 coins independent each other ($\frac{1}{2}$ $\frac{1}{2}$ = 0.5, 0.5)
③ dependent event.

↓

Prob keep increase & independent variables

In Bag have 5 marbles (1 event) → pick one black marbles $\boxed{\frac{2}{5}}$ → next event (next one black) $\boxed{\frac{1}{4}}$



Condition prob:

A, B → 2 events.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

↓ given

I want findout what probability marbles take continously?

Event (A) Black marbles : 2/5 (P(B)

Event (B) On more Black marbles

$P(B|A)$ → already A event performed.

$: 1/4$

$P(B) = 2/5$

$P(B|A) = 1/4$

$2/5 * 1/4 = 2/20 = 1/10$

↓ What is prob Both event A & B.

How check match or not?

$$P(B|A) = \frac{P(A \cap B)}{P(B)}$$

$= 1/10 / 2/5$

$= \frac{1}{10_2} \cdot \frac{5}{2}$

$= 1/4$

$P(B|A)$ Try to find out prob of event and this event already performed. (taken in place)

dependent. ←

Baye's theorem:

try take this condition prob and deriv

Bayes theorm

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

$$P(B|A) = \frac{P(B\cap A)}{P(A)}$$

$P(A\cap B) = P(B\cap A)$

① take this dinometer put left ~~right~~ side

$$P(A\cap B) = P(A|B) * P(B)$$

$$P(B\cap A) = P(B|A) * P(A)$$

$$P(A|B) * P(B) = \underbrace{P(B|A)}_{\text{likelihood}} * P(A)$$

$$\underbrace{P(A|B)}_{\text{posterior prob}} = \frac{P(B|A) * \overbrace{P(A)}^{\text{prior prob.}}}{\underbrace{P(B)}_{\text{marginal prob}}}$$

(Based on Baye theorm)

(Bayes theorm)

Scanned by TapScanner

(2) Naive Bayes classifier

Basic of classification

Bayes theorem

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

actually derived by conditional prob.

convert this equation based on our

dataset → $x = \{x_1, x_2, \overset{B}{x_3 \ldots x_n} \mid \overset{A}{y}\}$ o/p

i/p

Convert equation:

$$P(y|x_1, x_2 \ldots x_n) = \frac{P(x_1|y)\, P(x_2|y) \ldots P(x_n|y) * P(y)}{P(x_1)\, P(x_2) \ldots P(x_n)}$$

find / compute

already given

(It is multiplication prob based on i/p features) so

$$\frac{P(y) \cdot \prod_{i=1}^{n} P(x_i|y)}{P(x_1)\, P(x_2) \ldots P(x_n)}$$

constant same for every record

$$\boxed{P(y|x_1, x_2 \cdots x_n) \propto P(y) \prod_{i=1}^{n} P(x_i|y)}$$

we need to find (take) argmax what ever computation to found (doing) with respect to P(y) * P(all off the feature)

argmax → highest prob

ex; yes : 0.7 ✓

no : 0.4.

$$y = \text{argmax}_y \, P(y) \prod_{i=1}^{n} P(x_i|y)$$

Example: I have some feature outlook, temp binary whether play or not

Outlook

	yes	NO	P(y)	P(N)
sunny	2	3	2/9	3/5
overcast	4	0	4/9	0/5
Rainy	3	2	3/9	2/5
total	9	5	100%	100%

temperature

hot	2	2	2/9	2/5
cool	4	2	4/9	2/5
mild	3	1	3/9	1/5
	9	5	100%	100%

play

Yes	9	9/14
NO	5	5/14
total	14	100%

Scanned by TapScanner

problem: Example

Today is {Sunny, HOT} check whether

today play or not ?

Okay, first look dataset

whenthe O.L Sunny has 2 time yes 3 time No

So the Prob yes 2/9 → total no of Yes value.

Prob No 3/5 → total no of No value

Next HOT in temp has 2 times yes 2 time No.

Apply Naive Bayes:

$$P(Y/\text{Today}) = \frac{P(\text{sunny}|\text{yes}) * P(\text{HOT}|\text{yes}) * P(\text{yes})}{P(\text{Today})}$$

$$= 2/9 * 2/9 * 9/14$$

$$= \boxed{0.031}$$

Similarly No

$$P(N|\text{Today}) = \frac{P(\text{sunny}|\text{NO}) * P(\text{Hot}|\text{NO}) * P(N)}{P(\text{Today})}$$

$P(\text{No}|\text{Today}) = 3/5 * 2/5 * 5/14$

$= \boxed{0.08571}$

How did determine whether Y or N?

You can take values if I want to calculate P(y) with respect to condition

I have to output that, output normalized

$$\frac{0.031}{0.031 + 0.0857}$$

$P(y)$ $\boxed{\simeq 0.27}$

likely

P(N) * I have subtract $1 - P(y)$

$= 1 - 0.27$

$\boxed{= 0.73}$

argmax So today not play.

Scanned by TapScanner

Summary:

We have dataset outlook, temp i/p features and one dependent feature (target is today play or not based on outlook, temp features 14 records actually say yes or No.

based on the particular prob calculated.

But we show today is sunny, not apply naive bayes based on bayes theorem. calculating finally normalize which is maximum that is our output. Here No is maximum.

No is higher argument than yes.

Scanned by TapScanner

③ Apply naive Bayes classifier on text data.

How does it actually work on text data

I taken Restaurant Review

text → good
→ bad.

⇒ bow nlp concepts → Remove som (preprocessing)

good - 1
bad - 0

After preprocessing we get vector like this:

f_1	f_2	f_3	f_4	o/p
The	food	delicious	bad	

$$P(y = yes \mid \underset{(x_1 \cdots x_n)}{sent}) = P(y = yes \mid (x_1, x_2, x_n))$$

$$\alpha \simeq P(y) * \prod_{i=1}^{n} P(x_i \mid y = yes) *$$

$$P(x_2 \mid y = yes) \cdots$$

$$P(x_n \mid y = yes)$$

Sen1: The food is delicious

sent2: The food is bad.

sent Food is Bad

f_1	f_2	f_3	f_4	o/p
The	good	delicious	bad	
1	1	1	0	1
1	1	0	1	0
0	1	0	1	0
0	1	1	0	1
0	0	0	1	0

$$\boxed{p(y=yes) * p(x_1|y=yes) * p(x_2|y=yes) * \cdots p(x_n|y=yes)}$$

$$= 2/5 * 1/2 * 2/2 * 2/2 * 0/3$$

$$= 2/5 * 1/2 * 1/2$$

$$= 1/5 * 1/2 = 1/10 = 0.01$$

Scanned by TapScanner

Similarly Compute No.

$$\boxed{P(y=NO) * P(x_1|y=NO) * \dots P(x_n|y=NO)}$$

$= 3/5 * 1/2 * 2/4 * 0/2 * 2/3$

$\simeq$ ~~0.03~~ $3/5 * 1/4 = 3/20 \simeq 0.15$

Normalized

$$= \frac{0.01}{0.01+0.15}$$

$$P(yes) = \frac{0.01}{0.25} = ~~0.06~~\ 0.4 = 40\%.$$

$$P(NO) = 1-0.4$$

$$= 0.6 = 60\%.$$

max of $P(y) = 60\%$.

so $\boxed{No}$.

Scanned by TapScanner