Day: ________ Date: ________

Bayesian Classifier

$$P(A|B) = \frac{P(A\cap B)}{P(B)} \quad — (1)$$

$$P(B|A) = \frac{P(B\cap A)}{P(A)} \quad — (2)$$

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

$$P(B|A) = \frac{P(A\cap B)}{P(A)} \quad — (3)$$

Divide (1) by (3)

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

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

Bayes theorem

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

Day:
Date:

hypothesis

$$P(H/X) = \frac{P(X/H)P(H)}{P(X)}$$

likelyhood (probability after evidence occur)

posterior probability

evidence

prior probability

$$P(H/X) = \frac{P(X/H)P(H)}{P(X)}$$

posterior probability

likelyhood

prior probability (already known)

hypothesis
1. Orange
2. Apple
3. cherry

and

$P(\text{orange} / \text{red} \cap \text{round})$ — ①

$P(\text{apple} / \text{red} \cap \text{round})$ — ②

$P(\text{cherry} / \text{red} \cap \text{round})$ — ③

$$P(\text{orange}/\text{red} \cap \text{round}) = \frac{P(\text{red} \cap \text{round} / \text{orange}) \times P(\text{orange})}{P(\text{red} \cap \text{round})}$$

maximize

$P(\text{orange}) = \frac{3}{10}$

$P(\text{red} \cap \text{round}) = \frac{5}{10}$

Scanned with CamScanner

Day: ________ Date: ________

MAP

maximize a posterior probability

Naive Bayesian Classifier

↓

(Simple Baysian Classifier)

P(red round | orange) is computationally expensive

we make an assumption

Store all the possible possibilities

12 possibilities 2

→ 4(red round → orange)

→ 4(red round apple)

→ 4(red round → dery)

10 attributes

2^{10} 3^{10} 10 2×2×2

10 attributes

5^{10} = 9.7M

Day: Date:

$P(red \cap round \mid orange)$

Class conditional Independence

$\cancel{6\times 2}$

$\cancel{3\times 2}$

D	C
$\frac{1}{3}$	$\frac{1}{1}$

$\frac{1}{12}$

Independent events

$P(D=3 \wedge C=\cancel{2})$

$= P(D=3) \wedge P(C=2)$

$= \frac{1}{6} \times \frac{1}{2}$

$= \frac{1}{12}$

$3\times 3\times 3\times 3$ 3^{8}

3

4^3 2^3

orange red

$2\times 2\times 2$

~~orange round~~ 4

3

2

Day: Independent eveil

Date:

$P(red \wedge round / orange) =$

12

$P(red / orange) \times$

3

$P(round / orange)$

3

10 attributes

every one is binary

$2^{10} = 1024 \times 3$

$2^{10} = 3076$

10 attributes

$2 + 2 + 2 + \dots 2$

red / orange
not red / orange

sweet / orange
bitter / orange

10 values

2×10

$= 20$ for each class

20×3

$= 60$

Day:

Date:

training

stored probabilities

testing

result on these stored probabilities

100 attributes

binary

$2^{100} \times 2$

$\rightarrow$ 2 + 2 ....

200

200×2

$= 400$

A	B	C	D	E	Class
T	T	T	T	T	
F	F	F	F	F	X
T	T	F	T	T	Y
					X
					X

Independ

$P(A=T \mid X)$

$P(A=F \mid X)$

$P(A=T \mid Y)$

$P(A=F \mid Y)$

$2^5 = 32$ for one class

$32 \times 2 = 64$

Independence $2 \times 5 = 10$

$10 \times 2 = 20$

Date:

A	B	C	Class
a_1	b_1	c_1	y
a_2	b_2	c_2	y
a_3			
a_1	b_1	c_3	
a_2	b_2	c_1	x

$= 3 \times 2 \times 3$ ~~3 × 2 × 3~~

$= 18$ for X.

$= 18$ for y.

36

for X $3+2+3 = 8$

for y $3+2+3 = 8$

16

Day:

Date:

buy computer dated

$3 \times 3 \times 2 \times 2$

$= 36$

$= 36 \times 2$

72

① $36 \times 2 = 72$

~~$3 + 3 + 2 + 2$~~

$= 10$

$= 10 \times 2$

$= 20$

X = {age = "< 30", income = medium, student = yes, credit rating = fair}

P (Buy computer = yes | X) =

$P(Y|H) = (P(<30|Yes) \times P(medium|Yes) \times P(Yes|Yes) \times P(fair|Yes)$

~~P (Buy computer = not X)~~

$\times P(Yes)$ ÷ (P(X))

Day: Date:

$$\frac{P(X/H)\,P(H)}{P(X)}$$

$$\frac{P(<30/Yes) \times P(medium/Yes) \times P(Yes/Yes) \times P(Fair/Yes) \times P(Yes)}{\cancel{P(X)}}$$

No need same in both

P (By. comp = No / X) =

P (< 30/No) x P (Yes / No) x . . . .

$$\frac{(medium \cap Yes)}{P(Yes)}$$

(Yes) $\Rightarrow \frac{2}{9} \times \frac{4}{9} \times \frac{6}{9} \times \frac{6}{9} = 0.64$

(No) $\Rightarrow \frac{3}{5} \times \frac{2}{5} \times \frac{1}{5} \times \frac{2}{5} = 0.35$

map → which one is greater will win

= 0

Day: Date:

Normalize

$$\frac{0.028}{0.028+0.007} = 0.8$$

$$\frac{0.007}{0.028+0.007} = 0.2$$

Yes = 1 × 2 × 3 × Yes

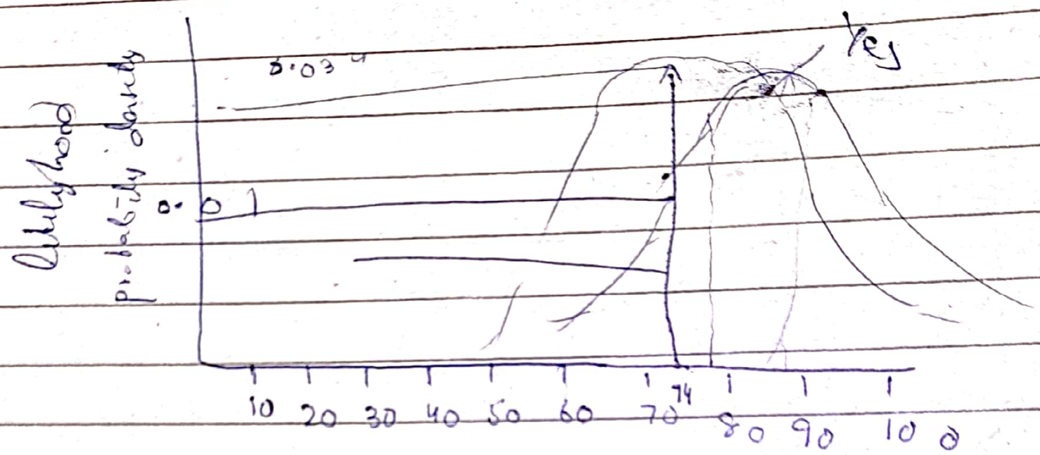
No = 1 × 2 × 3 × N

$P(Yes) \times P(X_s\ No$

P(H|X) P(H)

Day: Date:

Gaussian



yes 86 96 80 65 70 80 70 90 85

Mean std dev

79.1 10.2

No

86

74

74 → peak → greater

Gaussien for every column mean and median must be same

Day: Date:

λ x

Smoothing $\dot{x}$

laplace

$$\frac{c\,(x,y)+1}{c(cj)+1\,(3)}$$

Day: Date:

Multinomial Naïve Bayes Classifier

- fit → find columns
- transform → set values on columns

When to use

~~Features~~ documents have frequency counts

$$P(+ \mid GD) = P(good \mid +) * P(nice \mid +) * P(+)$$

$$\frac{4}{14} \times \frac{2}{14} \times \frac{3}{5}$$

$$P(- \mid GD) = P(good \mid -) * P(nice \mid -) * P(-)$$

$$\frac{2}{12} \times \frac{1}{12} \times \frac{2}{5}$$

Day:

Date:

→ Normalization

↓

training

↓

mean

↓

fit_transform

↓

on the basis of training count

test data

↓

new

↓

just adding into the columns

transform

↓

not fit again because its new

↓

not known about the word frequency.

transform

usi sanche me data ke srf wo words jo training me already use horae ho. isliye fit ni krte but transform krte k sef wali words

Bernoulli: NB
(0 - 1)

$\frac{d}{2d}$

$P(+|TD) =$
$P(nice|+) *$
$P(good|+) * P(+)$

$P(-|TD) =$
$P(nice|-) *$
$P(good|-) * P(-)$

Bernoulli:

$P(+|TD) = P(good|+) *$
$* (1 - P(liked|+)) * (1 - P(bad|+))$
$* (1 - P(unpleasant|+)) *$
$P(nice|+) * (1 - P(pleased|+)) *$
$(1 - P(rude|+))$.

$\cancel{P(-|TD)} =$

$P(-|TD) =$