Bayesian Classifier

Bayesian Classifier

- Statistical Classifier
- Predict Class Membership Probabilities
- Based on Bayes Theorem
- High Accuracy and Speed in large Databases
- Prior Probability: Probability of X, P(X)
- Posterior Probability: Probability of X when consider some condition Y, P(X|Y)
- A Simple Bayesian Classifier Naive Bayesian Classifier
 - Class Conditional Independence

MLE vs MAP

Maximum Likelihood Estimation

$$egin{aligned} heta_{MLE} &= rg\max_{ heta} f(X_1, X_2, \ldots, X_n \,|\, heta) \ &= rg\max_{ heta} \sum_i \log f(X_i \,|\, heta) \end{aligned}$$

Maximum A Priori

$$egin{aligned} heta_{MAP} &= rg \max_{ heta} f(heta \, | \, X_1, X_2, \ldots, X_n) \ &= rg \max_{ heta} \left(\log g(heta) + \sum_i \log f(X_i \, | \, heta)
ight) \end{aligned}$$

Conditional Independence

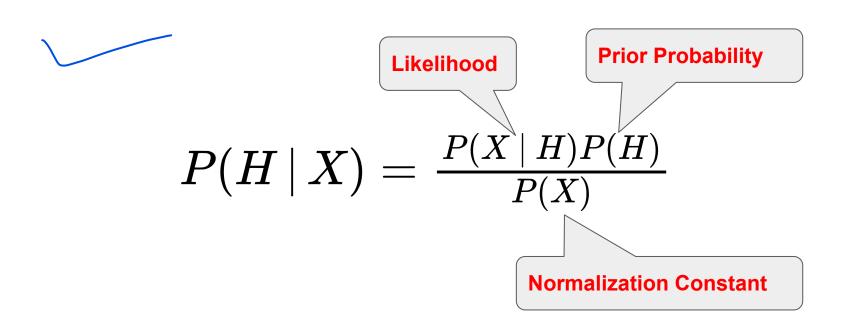
A and B are independent if,

$$P(A\cap B)=P(A) imes P(B) \ orall_{a,b}: P(A=a\cap B=b)=P(A=b) imes P(B=b)$$

A and B are conditionally independent given C if,

$$P(A,B\,|\,C) = P(A\,|\,C) imes P(B\,|\,C) \ orall_{a,b,c}: P(A=a\cap B=b\,|\,C=c) = P(A=b\,|\,C=c) imes P(B=b\,|\,C=c)$$

Bayes Theorem



Naive Bayes Classifier

- Attributes are conditionally independent
- ullet Consider *n*-dimensional attribute vector, $X=(X_1,X_2,\ldots,X_n)$
- ullet Consider m classes, $C=(C_1,C_2,\ldots,C_m)$
- Naive Bayes predicts that a tuple belongs to some class C_i for given condition X if and only if,

$$P(C_i \mid X) > P(C_j \mid X) \ for \ 1 \leq j \leq m, \ j \neq i$$

ullet Goal: maximize $P(C_i \,|\, X)$

Naive Bayes Equations

Likelihood:
$$P(X \mid C_i) = \prod_{k=1}^n P(X_k \mid C_i)$$
 $= P(X_1 \mid C_i) imes P(X_2 \mid C_i) imes \dots imes P(X_n \mid C_i)$

Categorical Attribute:

$$P(X_k \,|\, C_i) = rac{|D_{X_k,C_i}|}{D_{C_i}}$$

Continuous-valued Attribute: Gaussian Distribution

$$egin{aligned} P(X_k \,|\, C_i) &= g(X_k, \mu_{C_i}, \sigma_{C_i}) \ g(X, \mu, \sigma) &= rac{1}{\sqrt{2\pi}\sigma} e^{-rac{(X-\mu)^2}{2\sigma^2}} \end{aligned}$$

Sample Data: D

iD	age	income	student	creditRating	buyComputer
1	youth	high	yes	fair	no
2	youth	low	no	excellent	no
3	middleaged	medium	no	fair	yes
4	middleaged	medium	no	fair	yes
5	senior	medium	yes	fair	yes
6	senior	low	no	excellent	no
7	midleaged	high	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	high	yes	fair	yes
11	youth	high	yes	excellent	yes
12	middleaged	medium	no	excellent	yes
13	middleaged	low	yes	fair	yes
14	senior	high	yes	excellent	no
ID	age	income	student	creditRating	buyComputer
1	youth	low	no	fair	???

Sample Data: Problem Definition



ID	age	income	student	creditRating	buyComputer
1	youth	low	no	fair	???

Here, X = (age = youth, income = low, student = no, creditRating = fair)

Find,
$$= \max(P(C_{yes} \,|\, X), P(C_{no} \,|\, X))$$

$$= \max \left(rac{P(X\,|\,C_{yes})P(C_{yes})}{P(X)}, rac{P(X\,|\,C_{no})P(C_{no})}{P(X)}
ight)$$

$$= \max(P(X | C_{yes}) P(C_{yes}), P(X | C_{no}) P(C_{no}))$$

ID	age	income	student	creditRating	buyComputer
1	youth	low	no	fair	???

```
Necessary Statistics, P(buyComputer = yes) P(age = youth \mid buyComputer = yes) P(income = low \mid buyComputer = yes) P(student = no \mid buyComputer = yes) P(creditRating = fair \mid buyComputer = yes)
```

Necessary Statistics,

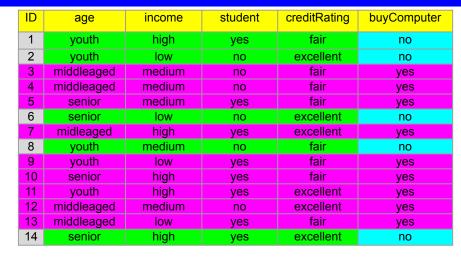
$$P(buyComputer = yes) = \frac{9}{14}$$

 $P(age = youth \mid buyComputer = yes)$

$$P(income = low \,|\, buyComputer = yes)$$

$$P(student = no \mid buyComputer = yes)$$

$$P(creditRating = fair | buyComputer = yes)$$



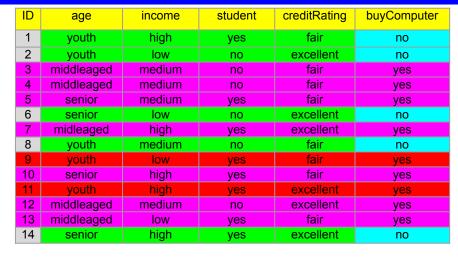
$$P(buyComputer = yes) = \frac{9}{14}$$

$$P(age = youth \mid buyComputer = yes) = \frac{2}{9}$$

$$P(income = low \,|\, buyComputer = yes)$$

$$P(student = no \mid buyComputer = yes)$$

$$P(creditRating = fair \, | \, buyComputer = yes)$$



age

youth

youth

middleaged

middleaged

senior

senior

midleaged

youth

youth

senior

vouth

middleaged

senior

Necessary Statistics,

$$P(buyComputer = yes) = \frac{9}{14}$$

$$P(ouyComputer = yes) = rac{3}{14}$$
 $P(age = youth \mid buyComputer = yes) = rac{2}{9}$

$$P(income = low \,|\, buyComputer = yes)$$

$$P(student = no \mid buyComputer = yes)$$

$$P(creditRating = fair \, | \, buyComputer = yes)$$

buyComputer

no

no

ves

ves

ves

no

ves

student

ves

no

no

no

ves

no

ves

no

ves

ves

ves

no

ves

ves

income

high

low

medium

medium

medium

low

high

medium

low

high

high

medium

low

high

creditRating

fair

excellent

fair

fair

fair

excellent

excellent

fair

fair

fair

fair

age

youth

youth

middleaged

middleaged

senior

senior

midleaged

youth

youth

senior

vouth

middleaged

middleaged

senior

11

14

student

yes

no

ves

no

ves

no

yes

ves

yes

no

ves

yes

income

high

low

medium

medium

medium

low

high

medium

low

high

high

medium

low

high

creditRating

fair

excellent

fair

fair

fair

excellent

excellent

fair

fair

fair

excellent

excellent

fair

excellent

buyComputer

no

no

ves

yes

no

yes

no

yes

yes

yes

ves

yes

no

$$P(buyComputer = yes) \quad \boxed{= rac{9}{14}}$$

$=\frac{1}{14}$	
$P(age = youth buyComputer = yes) \boxed{= rac{2}{9}}$	
$P(income = low buyComputer = yes) \hspace{0.5cm} = \hspace{0.5cm}$	$\frac{2}{9}$
P(student = no buyComputer = yes)	$=\frac{3}{9}$

$$P(creditRating = fair \, | \, buyComputer = yes)$$

Necessary Statistics,

$$P(buyComputer = yes) \quad \boxed{= rac{9}{14}}$$

·	13	middleaged	low	yes
$P(buyComputer = yes) = \frac{9}{14}$	14	senior	high	yes
$f'(uageomparer = ges) = \frac{1}{14}$				
P(age = youth buyComputer = yes)		$\frac{2}{2}$		
		9		
P(income = low buyComputer = yes	()	$=\frac{2}{9}$		
	,	9		
P(student = no buyComputer = yes	3)		$=\frac{3}{9}$	
, , , , , , , , , , , , , , , , , , , ,	,		9	
P(creditRating = fair buyCompute	er	= ues	=	$=\frac{4}{0}$
		<i>J)</i>		9

creditRating

fair

excellent

fair

fair

fair

excellent

excellent

fair

fair

fair

excellent

excellent

fair

excellent

income

high

low

medium

medium

medium

low

high

medium

low

high

high

medium

age

youth

youth

middleaged

middleaged

senior

senior

midleaged

youth

youth

senior

youth

middleaged

2

student

yes

no

no

no

yes

no

ves

no

ves

ves

yes

no

buyComputer

no

no

ves

ves

yes

no

yes

no

yes

ves

yes

yes

yes

no

$$P(buyComputer = yes) = \frac{9}{14}$$
 $P(age = youth \mid buyComputer = yes) = \frac{2}{9}$
 $P(income = low \mid buyComputer = yes) = \frac{2}{9}$
 $P(student = no \mid buyComputer = yes) = \frac{3}{9}$
 $P(creditRating = fair \mid buyComputer = yes) = \frac{4}{9}$

$$P(X | buyComputer = yes) imes P(buyComputer = yes) = rac{2}{9} imes rac{2}{9} imes rac{3}{9} imes rac{4}{9} imes rac{9}{14} = rac{8}{1701}$$

ID	age	income	student	creditRating	buyComputer
1	youth	low	no	fair	???

 $egin{aligned} P(buyComputer = no) \ P(age = youth \,|\, buyComputer = no) \ P(income = low \,|\, buyComputer = no) \ P(student = no \,|\, buyComputer = no) \end{aligned}$

P(creditRating = fair | buyComputer = no)

Necessary Statistics,

$$P(buyComputer = no) = \frac{5}{14}$$

 $P(age = youth \mid buyCom\overline{puter} = no)$

 $P(income = low \mid buyComputer = no)$

P(student = no | buyComputer = no)

P(creditRating = fair | buyComputer = no)

ID	age	income	student	creditRating	buyComputer
1	youth	high	yes	fair	no
2	youth	low	no	excellent	no
3	middleaged	medium	no	fair	yes
4	middleaged	medium	no	fair	yes
5	senior	medium	yes	fair	yes
6	senior	low	no	excellent	no
7	midleaged	high	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	high	yes	fair	yes
11	youth	high	yes	excellent	yes
12	middleaged	medium	no	excellent	yes
13	middleaged	low	yes	fair	yes
14	senior	high	yes	excellent	no

Necessary Statistics,

$$P(buyComputer = no) = \frac{5}{14}$$

$$P(age = youth | buyComputer = no)$$

 $P(income = low \mid buyComputer = no)$

$$P(student = no | buyComputer = no)$$

$$P(creditRating = fair | buyComputer = no)$$

age	income	student	creditRating	buyComputer
youth	high	yes	fair	no
youth	low	no	excellent	no
middleaged	medium	no	fair	yes
middleaged	medium	no	fair	yes
senior	medium	yes	fair	yes
senior	low	no	excellent	no
midleaged	high	yes	excellent	yes
youth	medium	no	fair	no
youth	low	yes	fair	yes
senior	high	yes	fair	yes
youth	high	yes	excellent	yes
middleaged	medium	no	excellent	yes
middleaged	low	yes	fair	yes
senior	high	yes	excellent	no
	youth youth middleaged middleaged senior senior midleaged youth youth senior youth middleaged middleaged	youth high youth low middleaged medium middleaged medium senior medium senior low midleaged high youth medium youth low senior high youth high middleaged medium middleaged low	youth high yes youth low no middleaged medium no middleaged medium no senior medium yes senior low no midleaged high yes youth medium no youth low yes senior high yes youth high yes middleaged medium no middleaged low yes	youth high yes fair youth low no excellent middleaged medium no fair senior medium yes fair senior low no excellent midleaged high yes excellent youth medium no fair youth low yes fair youth low yes fair senior high yes fair senior high yes excellent middleaged medium no excellent middleaged medium no excellent middleaged low yes fair

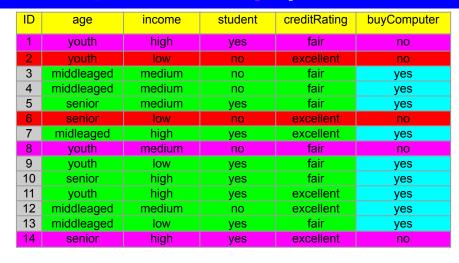
$$P(buyComputer = no) = \frac{5}{14}$$

$$P(age = youth | buyComputer = no)$$

$$P(income = low \,|\, buyComputer = no)$$

$$P(student = no \,|\, buyComputer = no)$$

$$P(creditRating = fair | buyComputer = no)$$



Necessary Statistics,

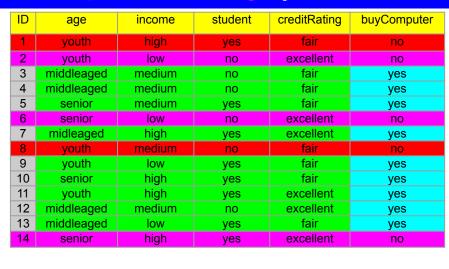
$$P(buyComputer = no) = \frac{5}{14}$$

P(age = youth | buyComputer = no)

 $P(income = low \,|\, buyComputer = no)$

P(student = no | buyComputer = no)

P(creditRating = fair | buyComputer = no)



$$P(buyComputer = no)$$
 $= \frac{5}{14}$
 $P(age = youth \mid buyComputer = no)$ $= \frac{3}{5}$
 $P(income = low \mid buyComputer = no)$ $= \frac{2}{5}$
 $P(student = no \mid buyComputer = no)$ $= \frac{3}{5}$
 $P(creditRating = fair \mid buyComputer = no)$ $= \frac{2}{5}$

$$P(X | buyComputer = no) imes P(buyComputer = no) = rac{3}{5} imes rac{2}{5} imes rac{3}{5} imes rac{3}{5} imes rac{5}{14} = rac{18}{875}$$

Sample Data: Prediction

$$P(X | buyComputer = yes) \times P(buyComputer = yes) < P(X | buyComputer = no) \times P(buyComputer = no)$$

$$buyComputer = no$$

Zero Probability

$$egin{aligned} &= P(age = middleaged \, | \, buyComputer = no) \ &= rac{0}{5} \end{aligned}$$

ID	age	income	student	creditRating	buyComputer
1	youth	high	yes	fair	no
2	youth	low	no	excellent	no
3	middleaged	medium	no	fair	yes
4	middleaged	medium	no	fair	yes
5	senior	medium	yes	fair	yes
6	senior	low	no	excellent	no
7	midleaged	high	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	high	yes	fair	yes
11	youth	high	yes	excellent	yes
12	middleaged	medium	no	excellent	yes
13	middleaged	low	yes	fair	yes
14	senior	high	yes	excellent	no

$$X = (age = middleaged, income = low, student = no, creditRating = fair)$$

Laplacian Correction

Add an extra count for each different value.

For attribute age,

- → count(youth | no) = 4 (Actually 3)
- → count(middleaged | no) = 1 (Actually 0)
- → count(senior | no) = 3 (Actually 2)
- \rightarrow count(no) = 8 (Actually 5)

=P(age=middleaged% =aged) +aged	$ \mathit{buyComputer} = no)$
1	
$=\frac{-}{8}$	

ID	age	income	student	creditRating	buyComputer
1	youth	high	yes	fair	no
2	youth	low	no	excellent	no
3	middleaged	medium	no	fair	yes
4	middleaged	medium	no	fair	yes
5	senior	medium	yes	fair	yes
6	senior	low	no	excellent	no
7	midleaged	high	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	high	yes	fair	yes
11	youth	high	yes	excellent	yes
12	middleaged	medium	no	excellent	yes
13	middleaged	low	yes	fair	yes
14	senior	high	yes	excellent	no

$$X = (age = middleaged, income = low, student = no, creditRating = fair)$$

Thank You

References

- Chapter 6, Data Mining Concepts and Techniques, Jiawei Han and Micheline Kamber, 2nd Edition.
- Chapter 4, Machine Learning in Action, Peter Harrington
- https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/
- https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c
- https://www.geeksforgeeks.org/naive-bayes-classifiers/
- https://www.cs.cmu.edu/~10601b/slides/NBayes.pdf
- https://web.stanford.edu/class/archive/cs/cs109/cs109.1176/lectures/22-NaiveBayes.pdf