Naïve Bayes – Spam Detection

CS385 – Machine Learning - Classification

Outline

- Probability – a quick review
- Spam Detection
  - Maximum Likelihood Estimate
  - Maximum A Posterior
  - (Overfitting and MLE) vs. (Laplace Smoothing and MAP)
- Naïve Bayes – Conclusion
- Gaussian Naïve Bayes

Probability - Review

CS385 – Machine Learning

Independent Event

P(H H H H)+
P(H H H T)+
P(H H T H)+
P(H T H H)+
P(T H H H)

- Coin flip
  - Head P(H)=0.5
  - Tail P(T)=0.5

} normal coin

- P({X1 X2 X3 X4} >=3 Heads)=?

  - Head P(H)=0.25
  - Tail P(T)=0.75

} magic coin
loaded coin

- complementary event

  P(A)=p P(~A)=1-p

- Independent event
  - X and Y are independent
    P(X)P(Y)=P(XY)

  - P(H)=0.5 P(H H H)=?

P(H).P(H).P(H)= 0.5×0.5×0.5
= ...

  - Xi: result of the i-th flipping, P(H)=0.5
  - P(X1=X2=X3=X4)=?

P(H H H H)
P(T T T T) +

Dependence

- First flip a fair coin,
  - P(X1=H)=0.5
- If X1=H, then flip a loaded coin
  - P(X2=H|X1=H) = 0.9
- If X1=T, then flip another loaded coin,
  - P(X2=T|X1=T) = 0.8

P(X2=H)=?

$P(X_2=H)$

$= \underbrace{P(X_2=H \mid X_1=H)}_{0.9} \cdot \underbrace{P(X_1=H)}_{0.5}$

$+ \underbrace{P(X_2=H \mid X_1=T)}_{1-0.8} \cdot \underbrace{P(X_1=T)}_{1-0.5}$

- Total probability

$$P(Y)=\sum_i P(Y \mid X=i)P(X=i)$$

- P(~X|Y)=1-P(X|Y)
- P(X|~Y) ?= 1-P(X|Y)

$!=$

Cancer Test

- $P(C)=0.01$
- $P(\sim C)=?$

- $P(+|C)=0.9$
- $P(-|C)=?$

- $P(+|\sim C)=0.2$
- $P(-|\sim C)=0.8$

- $P(+,C)=?$ $P(+|C)\cdot P(C)$
- $P(-,C)=?$ $P(-|C)\cdot P(C)$

- $P(+,\sim C)=?$
- $P(-,\sim C)=?$

- $P(C|+)=?$

$P(C)$ 0.01 $\uparrow$ $+$

$\downarrow$ $=$

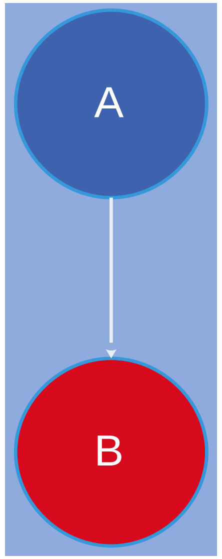
Bayes Rule

posterior likelihood Prior

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

$$P(c|+) = \frac{① P(+|c)\cdot P(c)}{① P(+|c)\cdot P(c) + P(+|\sim c)\cdot P(\sim c)}$$

Hidden Var. C



observable Var. Test $\{ {+ \atop -}$

- A: Not observable
- B: Observable
- Diagnose Reasoning:
  - P(A|B)
  - P(A|~B)
  - How many parameters?

Spam Detection

CS385 – Machine Learning

Bag of Words

- Hello I will say hello!

- Dictionary
  - Hello 2
  - I 1
  - Will 1
  - Say 1

Spam Detection – Training Data

Spam

- Offer is secret
- Click secret link
- Secret sports link

Ham

- Play sports today
- Went play sports
- Secret sports event
- Sport is today
- Sport costs money

1 Size of vocabulary? 12

2 P(Spam) = ? 3/8

Maximum Likelihood Estimate (1)

P(SSSHHHHH) = P(S) P(S) · P(S) · P(H) ... P(H)

- SSSHHHHH
  - P(S)=Π
  - I.ID I: Independant ID: Identical distribution

Y
- 11100000 y_i

$$P(y_i) = \begin{cases} \Pi & \text{if } y_i = S \\ 1 - \Pi & \text{if } y_i = H \end{cases} \qquad P(y_i) = \Pi^{y_i} \cdot (1 - \Pi)^{1-y_i}$$

Π 1-Π

p(11100000)

$$P(data) = \Pi^{count(y_i=1)} \cdot (1 - \Pi)^{count(y_i=0)} = \Pi^3(1 - \Pi)^5$$

$$\log P(data) = 3 \log \Pi + 5 \log(1 - \Pi)$$

Maximum Likelihood Estimate (2)

$$P(data) = \Pi^{count(y_i=1)} \cdot (1-\Pi)^{count(y_i=0)} = \Pi^3(1-\Pi)^5$$

$$\log P(data) = 3\log\Pi + 5\log(1-\Pi)$$

$$\arg\max_\pi \log P(data)$$

$$\frac{\delta \log P(data)}{\delta\pi} = 0 = \frac{3}{\pi} - \frac{5}{1-\pi}$$

$$\pi = \frac{3}{8}$$

$$p(\pi \mid data) = \frac{p(data \mid \pi) \cdot p(\pi)}{p(data)}$$

MLE Solutions for conditional probability

$P(spam \mid \text{"secret"})$ $p.(spam)$

$= \dfrac{① \; P(\text{"secret"} \mid spam)}{① + P(\text{"secret"} \mid Ham) \cdot P(Ham)}$

Spam

- Offer is secret
- Click secret link
- Secret sports link

Ham

- Play sports today
- Went play sports
- Secret sports event
- Sport is today
- Sport costs money

$= \dfrac{1/3 \times 3/8}{1/3 \times 3/8 + 1/15 \cdot 5/8}$

MLE

1 P(“Secret” |Spam) =1/3

$3/9 = 1/3$

2 P(“Secret” |Ham) = 1/15

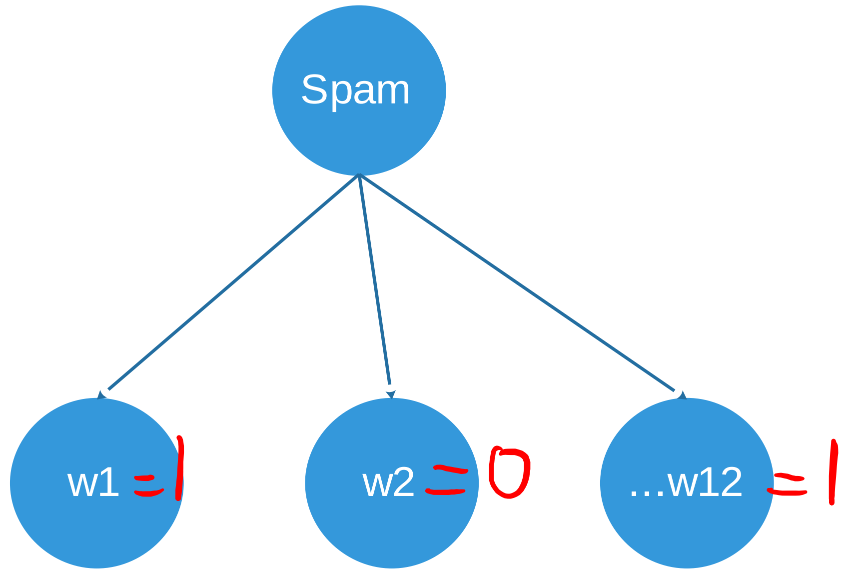
1

Spam Detection

(1, 0, 2, 4 ... 0, 2)

12

- How do we get the probability for each node?
- To learn from the training data by using Maximum Likelihood



Naïve Bayes

MLE Solutions for detection

Spam

- Offer is secret
- Click secret link
- Secret sports link

Ham

- Play sports today
- Went play sports
- Secret sports event
- Sport is today
- Sport costs money

1 P(Spam| "Sports") =1/6

2 P(Spam| "Secret is secret") =25/26

Spam

- Offer is secret
- Click secret link
- Secret sports link

Ham

- Play sports today
- Went play sports
- Secret sports event
- Sport is today
- Sport costs money

P(Ham| " — — ") = 1/26

P(Spam| "Secret is secret") =25/26

$$① \frac{1/3 \cdot 1/9 \cdot 1/3 \cdot 3/8}{1/3 \cdot 1/9 \cdot 1/3 \cdot 3/8 + 1/15 \cdot 1/15 \cdot 1/15 \cdot 5/8}$$

P(Spam)=3/8

P("Secret" |Spam) =1/3

P("Secret" |Ham) = 1/15

P("is" |Spam) =

P("is" |Ham) =

P("Secret is secret|spam)

conditional independent

Multiple classes

symptom	occupation	diagnosis
Sneeze	Nurse	Cold
Sneeze	Farmer	Allergy
Headache	Construction worker	Concussion
Headache	Construction worker	Cold
Sneeze	Lecturer	Cold
Headache	Lecturer	Concussion

(Sneeze, Construction worker)

Diagnosis ?

(Sneeze, Construction worker) Diagnosis ?

$$P(\text{cold} \mid \text{sneeze}, cw)$$

$$= \frac{① \; P(\text{sneeze}, cw \mid \text{cold}) \cdot P(\text{cold})}{① + ② \; P(\text{sneeze}, cw \mid \text{Allergy}) \cdot P(\text{Allergy}) + ③ \; P(\text{sneeze}, cw \mid \text{con}) \cdot P(\text{con})}$$

① $P(\text{sneeze} \mid \text{cold})\, P(cw \mid \text{cold}) \cdot P(\text{cold}) = 2/3 \cdot 1/3 \cdot 1/2$

② $P(\text{sneeze} \mid \text{Allergy})\, P(cw \mid \text{Allergy}) \cdot P(\text{Allergy}) = 1/1 \cdot 0/1 \cdot 1/6 = 0$

③ $P(\text{sneeze} \mid \text{con})\, P(cw \mid \text{con}) \cdot P(\text{con}) = 0 \cdot 1/2 \cdot 1/3 = 0$

$$\frac{①}{① + ② + ③} = 1$$