Bayesian Networks

CISC 5800
Professor Daniel Leeds

Approaches to learning/classification

For classification, find highest probability class given features

- $P(x_1,\ldots,x_n|y=?)$

Approaches:

- Learn/use function(s) for probability
  - $P(\text{light}|Y=\text{eclipse})=N(\mu_{eclipse},\sigma_{eclipse})$
- Learn/use probability look-up table for each combination of features:

letter_1	$P(\text{letter}_1 \mid \text{word}=\text{“duck”})$
“a”	0.001
“b”	0.010
“c”	0.005
“d”	0.950

Joint probability over N features

Problem with learning table with N features:

- If all dependent, exponential number of model parameters

Burglar breaks in	Alarm goes off	Jill gets call	Zack gets call	P(A,J,Z|B)
Y	Y	Y	Y	0.3
Y	Y	Y	N	0.03
Y	Y	N	Y	0.03
Y	Y	N	N	0.06
		⋮		

Joint probability over N features

Naïve Bayes – all independent

- Linear number of model parameters

What if only **some** features are independent?

Burglar breaks in	Alarm goes off	Jill gets call	Zack gets call	P(A,J,Z|B)
Y	Y	Y	Y	0.3
Y	Y	Y	N	0.03
Y	Y	N	Y	0.03
Y	Y	N	N	0.06
		⋮		

Bayes nets: conditional independence

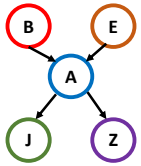
B – Burglar
E – Earthquake
A – Alarm goes off
J – Jill is called
Z – Zack is called

In Naïve Bayes: $P(x_1,x_2,x_3|y) = P(x_1|y)P(x_2|y)P(x_3|y)$

In Bayes nets, some variables depend on other variables:

Alarm depends on **B**urglar and **E**arthquake

Jill and **Z**ack calls each depend only on **A**larm



• P(B, E, A, J, Z) = P(B) P(E) P(A|B,E) P(J|A) P(Z|A)

5

Bayes nets: conditional independence

B – Burglar
E – Earthquake
A – Alarm goes off
J – Jill is called
Z – Zack is called

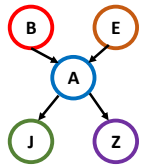
In Bayes nets, some variables depend on other variables:

• P(B, E, A, J, Z) = P(B) P(E) P(A|B,E) P(J|A) P(Z|A)

In general for Bayes nets:

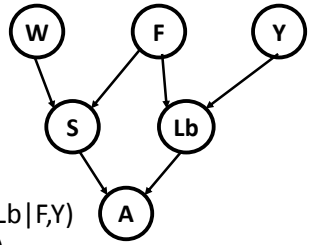
• $P(x_1,\ldots,x_n) = \prod_i P(x_i|Pa(x_i))$

• $Pa(x_i)$ are the "parents" of x_i – the variables x_i is conditioned on



6

Another Example: Health probabilities



P(W,F,Y,S,Lb,A) =
P(W)P(F)P(Y)
P(S|W,F)P(Lb|F,Y)
P(A|S,Lb)

F – Flu
S – Stress
Y – Age (years)
Lb – Weight
W – Weather
A – Activity

7

Probability review

Conditional Probabilities:

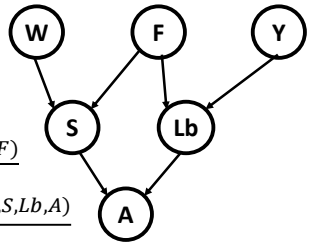
• $P(A|B) = \frac{P(A,B)}{P(B)}$

Marginal Probability

• $P(A) = \sum_{b \in B} P(A, B = b)$

8

Health probabilities, find P(S,Lb,A | F)



F – Flu
S – Stress
Y – Age (years)
Lb – Weight
W – Weather
A – Activity

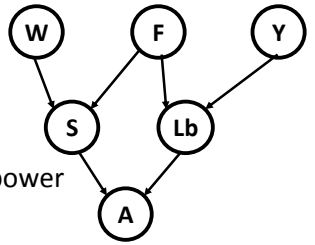
$$P(S,Lb,A|F) = \frac{P(S,Lb,A,F)}{P(F)}$$

$$= \frac{\sum_{w \in W} \sum_{y \in Y} P(W,F,Y,S,Lb,A)}{P(F)}$$

$$= \frac{\sum_{w \in W} \sum_{y \in Y} \boldsymbol{P(F)P(W)P(Y)P(S|W,F)P(Lb|F,Y)P(A|S,Lb)}}{\boldsymbol{P(F)}}$$

9

Health probabilities, find P(S,Lb,A | F)



Moving variables out of irrelevant summation loops saves computation power

$$P(S,Lb,A|F)$$

$$= \frac{\sum_{w \in W} \sum_{y \in Y} \boldsymbol{P(F)P(W)P(Y)P(S|W,F)P(Lb|F,Y)P(A|S,Lb)}}{\boldsymbol{P(F)}}$$

$$= \frac{\cancel{P(F)} \sum_{w \in W} P(W)P(S|W,F) \sum_{y \in Y} P(Y)P(Lb|F,Y)P(A|S,Lb)}{\cancel{P(F)}}$$

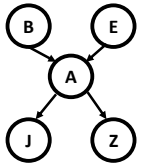
10

Conditional independence

If two variables x_i and x_j share same "parent," the x_i and x_j are independent given that parent

J and Z are independent given A: "J ⊥ Z|A"

B – Burglar
E – Earthquake
A – Alarm goes off
J – Jill is called
Z – Zack is called



11

Example evaluation of Bayes nets

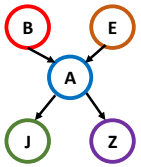
Use joint probabilities to find more probable class-variable value

Compute P(E=yes|A,J,Z), P(E=no|A,J,Z)

B – Burglar
E – Earthquake
A – Alarm goes off
J – Jill is called
Z – Zack is called

$$\boldsymbol{P(E|A,J,Z)} = \frac{\boldsymbol{P(E,A,J,Z)}}{\boldsymbol{P(A,J,Z)}} = \frac{\sum_B \boldsymbol{P(E,B,A,J,Z)}}{\sum_E \sum_B \boldsymbol{P(E,B,A,J,Z)}}$$

$$= \frac{\sum_B \boldsymbol{P(E)P(B)P(A|E,B)P(J|A)P(Z|A)}}{\sum_E \sum_B \boldsymbol{P(E)P(B)P(A|E,B)P(J|A)P(Z|A)}}$$



12