CS412 office hour

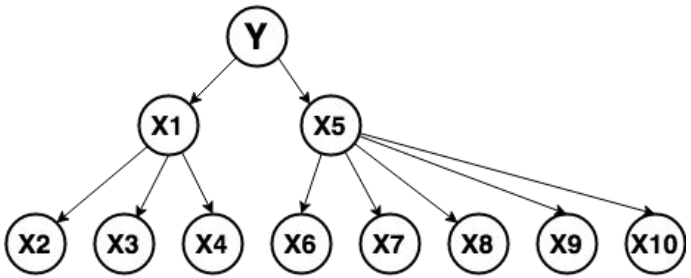
Apr 12, 2019

Today’s Office Hour

- Counting parameters
- QA

#parameters in a Bayesian network

Example: Suppose our data has 3 different classes and 10 attributes, each with 2 possible values. How many parameters will our model have if we use a Bayesian network with three layers as pictured below?

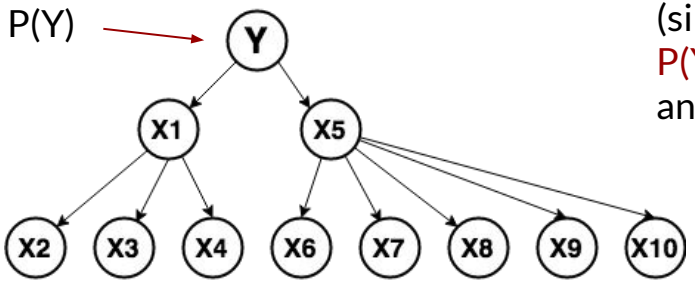


#parameters in a Bayesian network

Example: Suppose our data has 3 different classes and 10 attributes, each with 2 possible values. How many parameters will our model have if we use a Bayesian network with three layers as pictured below?

Y=a1	Y=a2	Y=a3
p1	p2	p3

3 parameters



OR 2 free parameters
(since the probabilities adds up to 1)
P(Y=a1), P(Y=a2)
and P(Y=a3) can be derived from them

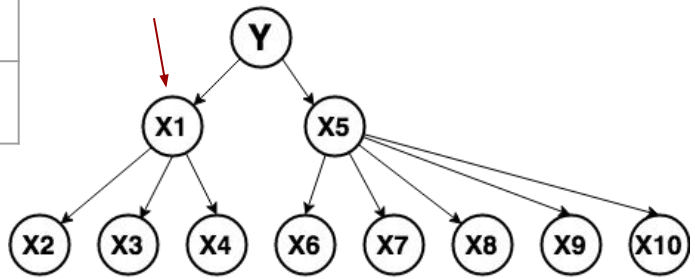
#parameters in a Bayesian network

Example: Suppose our data has 3 different classes and 10 attributes, each with 2 possible values. How many parameters will our model have if we use a Bayesian network with three layers as pictured below?

	Y=a1	Y=a2	Y=a3
X=b1			
X=b2			

6 parameters

P(X1 | Y)



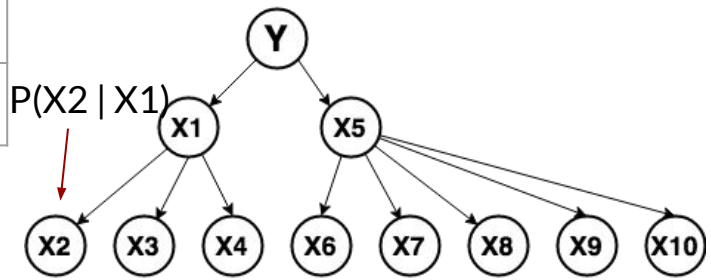
OR 3 free parameters
P(X1=b1| Y=a1)
P(X1=b1| Y=a2)
P(X1=b1| Y=a3)
and P(X1=b2 | Y=a1), P(X1=b2 | Y=a2), P(X1=b3 | Y=a3) can be derived from them

#parameters in a Bayesian network

Example: Suppose our data has 3 different classes and 10 attributes, each with 2 possible values. How many parameters will our model have if we use a Bayesian network with three layers as pictured below?

	X1=b1	X1=b2
X2=b1		
X2=b2		

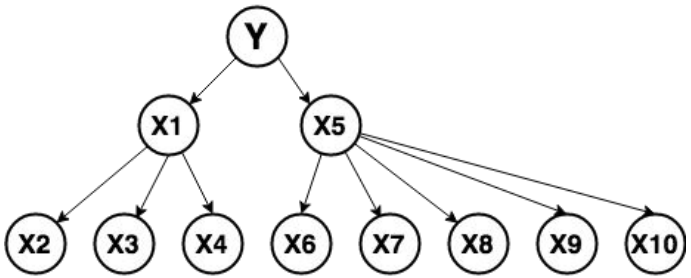
4 parameters



OR 2 free parameters
P(X2=x1| X1=x1)
P(X2=x1| X1=x2)
and P(X2=b2 | X1=b1), P(X2=b2 | X1=b2) can be derived from them

#parameters in a Bayesian network

Example: Suppose our data has 3 different classes and 10 attributes, each with 2 possible values. How many parameters will our model have if we use a Bayesian network with three layers as pictured below?



3 OR 2

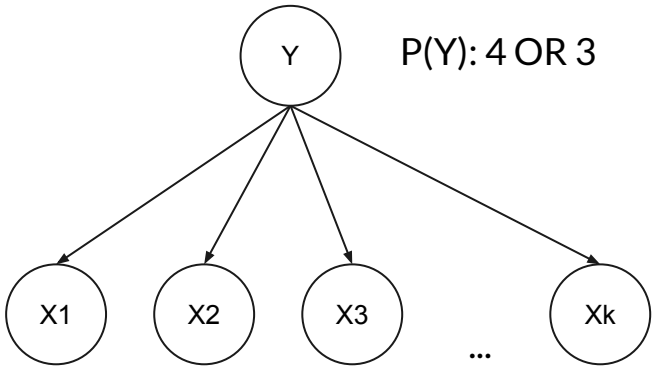
2*6 OR 2*3

8*4 OR 8*2

Total 47 OR 24

#parameters of naive Bayes

- A special case of Bayesian networks



P(Y): 4 OR 3

#classes (Y): 4
#value_for_each_dim: 3
#feature_dim: 10

P(X1 | Y): 3*4

Counting all: 10*3*4 + 4 = 124
Counting free prams: 10*2*4 + 3 = 83

Free: 2*4