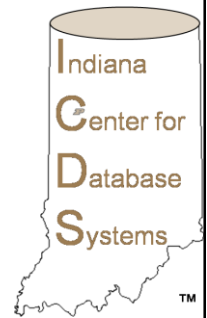


CS57300: Data Mining

Bayesian Networks

Prof. Chris Clifton

14 April 2022



Bayes Classifier

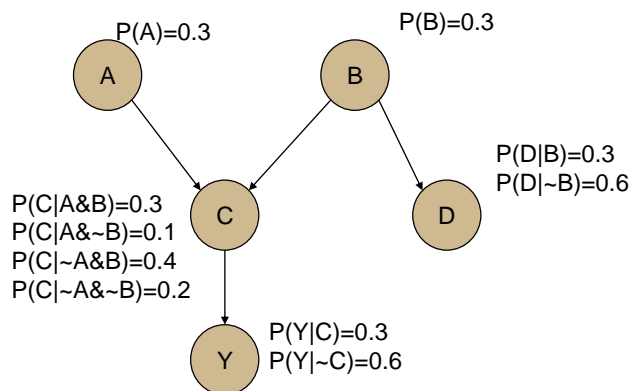
- Bayes rule: $P(Y|X) = \frac{P(X|Y) P(Y)}{P(X)}$
- Bayes theorem: $y_{MAP} = \operatorname{argmax}_{y \in Y} P(y|x) = \operatorname{argmax}_{y \in Y} \frac{P(x|y) P(y)}{P(x)}$
 - But how can we estimate these probabilities unless we've seen the exact case before?

Simplifications

- Naïve Bayes: Conditional independence
 - Only need to have seen enough cases of each attribute to estimate it's probability
- Bayes Network: Assume dependencies known
 - A bayesian network is a graphical model for probabilistic relationships among a set of variables

3

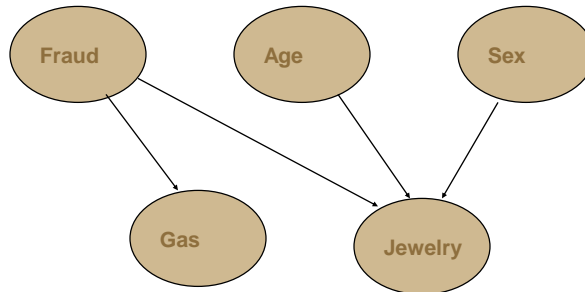
Bayes Net



4

An Example

Detecting Credit - Card Fraud



5

Learning a Bayes Net

Bayesian Inference:

Learn the probabilities

- Like Naïve Bayes:
Estimation of each of the probabilities
 - Look at the data
 - MAP
- *More to estimate, but not as many as full bayes*

Learn the network

- Try all possible and choose the best?
 - Too many possible
- Idea: Penalize complex models

6

Scoring a Structure (Friedman & Yakhini, UAI'96)

- Score = $-\frac{N_{params}}{2} \log R + \sum_{j=1}^m \sum_{k=1}^{parent\ values} \sum_{v=1}^{arity\ of\ X_j} P(V_k) P(X_j = v | V_k) \log P(X_j = v | V_k)$
- Still too complex to search exhaustively
 - Simulated annealing: start with simple structure, make small changes and use ones that work well
 - Random restarts in case we start with a “bad idea”