ECE 6504: Advanced Topics in Machine Learning
Probabilistic Graphical Models and Large-Scale Learning

Topics:
- Bayes Nets: Representation/Semantics
  - v-structures
  - Probabilistic influence, Active Trails

Readings: Barber 3.3; KF 3.3.1-3.3.2

Dhruv Batra
Virginia Tech
Plan for today

- Notation Clarification

- Errata #1: Number of parameters in disease network

- Errata #2: Car start v-structure example

- Bayesian Networks
  - Probabilistic influence & active trails
  - d-separation
  - General (conditional) independence assumptions in a BN
A general Bayes net

- Set of random variables
- Directed acyclic graph
  - Encodes independence assumptions
- CPTs
  - Conditional Probability Tables
- Joint distribution:

\[ P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid \text{Pa}_{X_i}) \]
Factorized distributions

• Given
  – Random vars $X_1, \ldots, X_n$
  – $P$ distribution over vars
  – BN structure $G$ over same vars

• $P$ factorizes according to $G$ if

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P\left(X_i \mid \text{Pa} X_i\right)$$
How many parameters in a BN?

- Discrete variables $X_1, \ldots, X_n$

- Graph
  - Defines parents of $X_i$, $\text{Pa}_{X_i}$

- CPTs – $P(X_i| \text{Pa}_{X_i})$
Independencies in Problem

World, Data, reality:

True distribution $P$ contains independence assertions

BN:

Graph $G$ encodes local independence assumptions
Bayes Nets

• BN encode (conditional) independence assumptions.
  – $I(G) = \{X \text{ indep of } Y \text{ given } Z\}$

• Which ones?
• And how can we easily read them?
Local Structures

• What’s the smallest Bayes Net?
Local Structures

Indirect causal effect:

\[ X \rightarrow Z \rightarrow Y \]

Indirect evidential effect:

\[ X \leftarrow Z \leftarrow Y \]

Common cause:

\[ X \rightarrow Z \rightarrow Y \]

Common effect:

\[ X \rightarrow Z \rightarrow Y \]
Car starts BN

- 18 binary attributes
- Inference
  - \( P(\text{BatteryAge}|\text{Starts}=f) \)
- \(2^{18}\) terms, why so fast?
Bayes Ball Rules

• Flow of information
  – on board
Active trails formalized

• Let variables $O \subseteq \{X_1, \ldots, X_n\}$ be observed

• A path $X_1 \rightarrow X_2 \rightarrow \cdots \rightarrow X_k$ is an active trail if for each consecutive triplet:
  
  - $X_{i-1} \rightarrow X_i \rightarrow X_{i+1}$, and $X_i$ is **not observed** ($X_i \notin O$)
  
  - $X_{i-1} \leftarrow X_i \leftarrow X_{i+1}$, and $X_i$ is **not observed** ($X_i \notin O$)
  
  - $X_{i-1} \leftarrow X_i \rightarrow X_{i+1}$, and $X_i$ is **not observed** ($X_i \notin O$)
  
  - $X_{i-1} \rightarrow X_i \leftarrow X_{i+1}$, and $X_i$ is **observed** ($X_i \in O$), or one of its descendents **is observed**
• **Theorem**: Variables $X_i$ and $X_j$ are independent given $Z$ if
  - no active trail between $X_i$ and $X_j$ when variables $Z \subseteq \{X_1, \ldots, X_n\}$ are observed
Name That Model

Naïve Bayes:

\[ p(y, x) = p(y) \prod_{j=1}^{D} p(x_j | y) \]
Name That Model

Tree-Augmented Naïve Bayes (TAN)
Name That Model

$Y_1 = \{a, \ldots, z\}$ → $Y_2 = \{a, \ldots, z\}$ → $Y_3 = \{a, \ldots, z\}$ → $Y_4 = \{a, \ldots, z\}$ → $Y_5 = \{a, \ldots, z\}$

$X_1 = \hat{b}$ → $X_2 = \hat{c}$ → $X_3 = \hat{d}$ → $X_4 = \hat{e}$ → $X_5 = \hat{f}$

*Hidden Markov Model (HMM)*