Statistical Methods in AI and ML

Nicholas Ruozzi
University of Texas at Dallas

The Course

One of the **most exciting** advances in AI/ML in the last decade

Judea Pearl won the Turing award for his work on Bayesian networks!

(among other achievements)

Prob. Graphical Models

Exploit **locality** and structural features of a given model in order to gain insight about **global properties**

The Course

- What this course is:
  - Probabilistic graphical models
  - Topics:
    - representing data
    - exact and approximate statistical inference
    - model learning
    - variational methods in ML

UTD

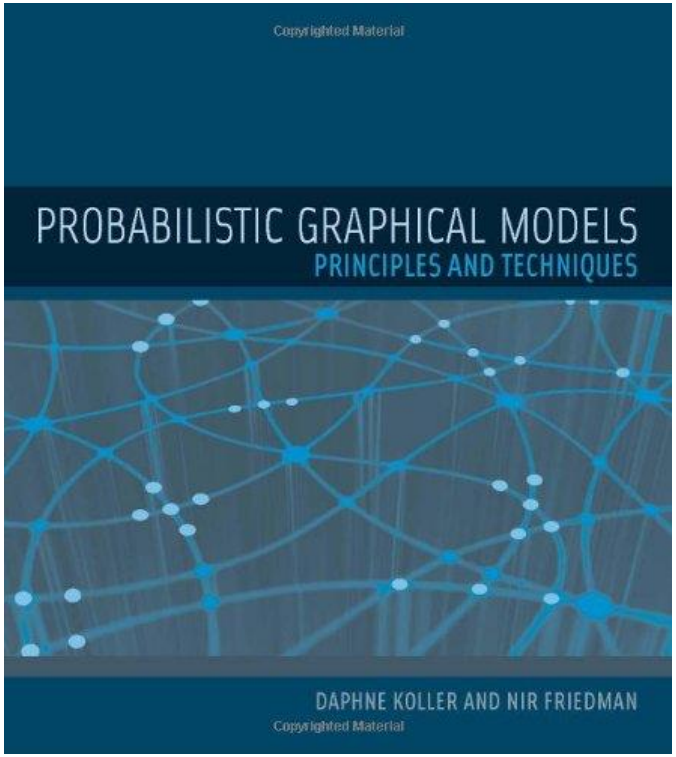
Prerequisites

- CS 5343: Algorithm Analysis and Data Structures
- CS 3341: Probability and Statistics in Computer Science and Software Engineering
- Basically, comfort with probability and algorithms (machine learning is helpful, but not required)

Textbook

Required readings will be posted online before each lecture

Check the course website for additional resources and papers



Grading

- 4-5 problem sets (70%)
  - See collaboration policy on the web
- Final project (25%)
- Class participation & extra credit (5%)

-subject to change-

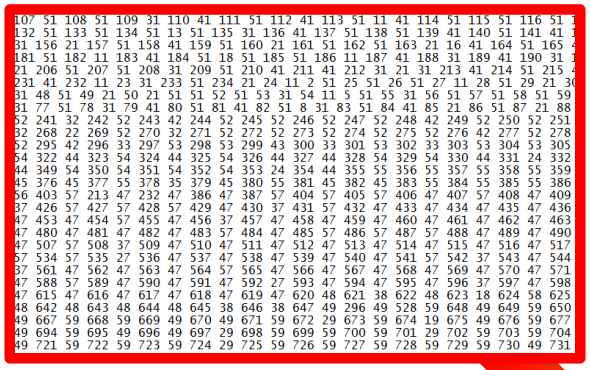
Course Info.

- Instructor: Nicholas Ruozzi
  - Office: ECSS 3.409
  - Office hours: Tues. 11am - 12pm and by appointment
- TA: Prasanna Kothalkar
  - Office hours and location TBD
- Course website: http://www.nrouzzi.me/cs6347/

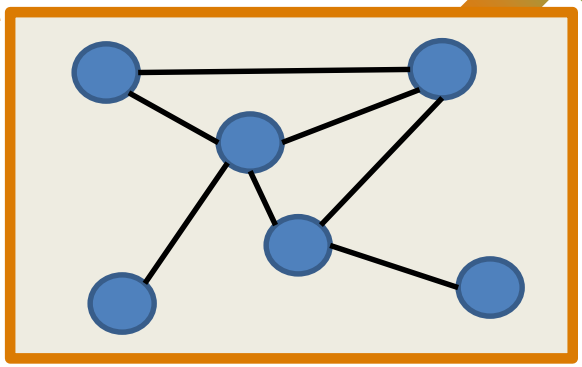
Main Ideas

- Model the world (or at least the problem) as a collection of random variables related through some joint probability distribution
  - Compactly represent the distribution
  - Undirected graphical models
  - Directed graphical models
- Learn the distribution from observed data
  - Maximum likelihood, SVMs, etc.
- Make predictions (statistical inference)

Inference and Learning



Collect Data

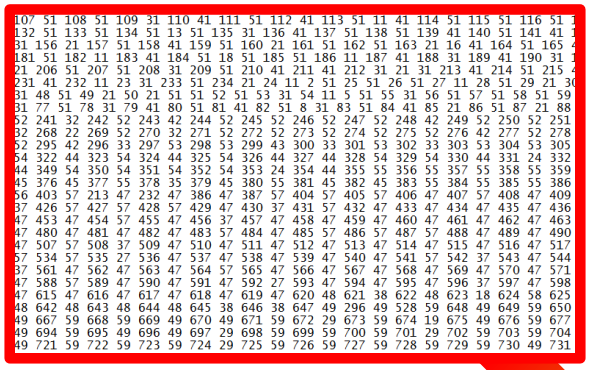


"Learn" a model that represents the observed data

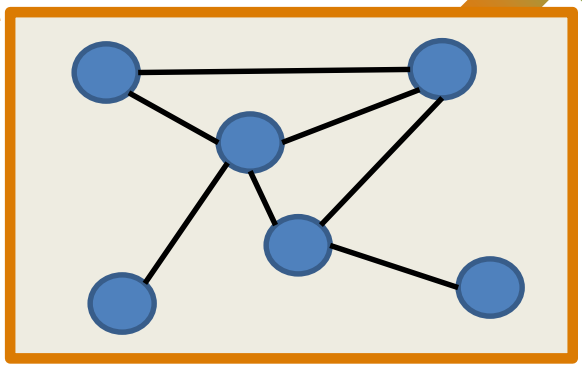
$$Z(\theta) = \sum_{x} p(x;\theta)$$

Use the model to do inference / make predictions

Inference and Learning



Data sets can be large



Data must be compactly modeled

$$Z(\theta) = \sum_{x} p(x;\theta)$$

Inference needs to be fast

Applications

- Computer vision
- Natural language processing
- Robotics
- Computational biology
- Computational neuroscience
- Text translation
- Text-to-speech
- Many more…

Graphical Models

- A graphical model is a graph together with "local interactions"
- The graph and interactions model a global optimization or learning problem
- The study of graphical models is concerned with how to exploit local structure to solve these problems either exactly or approximately

Optimization Problems on Trees

Time for a motivating example!

(chalk board)