

Statistical Relational Learning: A Tutorial

Lise Getoor

University of Maryland, College Park

ECML/PKDD 2007 Tutorial

● ● ● acknowledgements

- This tutorial is a synthesis of ideas of many individuals who have participated in various SRL events, workshops and classes:
- Hendrik Blockeel, Mark Craven, James Cussens, Bruce D'Ambrosio, Luc De Raedt, Tom Dietterich, Pedro Domingos, Saso Dzeroski, Peter Flach, Rob Holte, Manfred Jaeger, David Jensen, Kristian Kersting, Daphne Koller, Heikki Mannila, Andrew McCallum, Tom Mitchell, Ray Mooney, Stephen Muggleton, Kevin Murphy, Jen Neville, David Page, Avi Pfeffer, Claudia Perlich, David Poole, Foster Provost, Dan Roth, Stuart Russell, Taisuke Sato, Jude Shavlik, Ben Taskar, Lyle Ungar and many others...

● ● ● Roadmap

- SRL: What is it?
- SRL Tasks & Challenges
- 4 SRL Approaches
- Applications and Future directions

● ● ● Why SRL?

- Traditional statistical machine learning approaches assume:
 - A random sample of homogeneous objects from single relation
- Traditional ILP/relational learning approaches assume:
 - No noise or uncertainty in data
- Real world data sets:
 - Multi-relational, heterogeneous and semi-structured
 - Noisy and uncertain
- Statistical Relational Learning:
 - newly emerging research area at the intersection of research in social network and link analysis, hypertext and web mining, graph mining, relational learning and inductive logic programming
- Sample Domains:
 - web data, bibliographic data, epidemiological data, communication data, customer networks, collaborative filtering, trust networks, biological data, natural language, vision

● ● ● What is SRL?

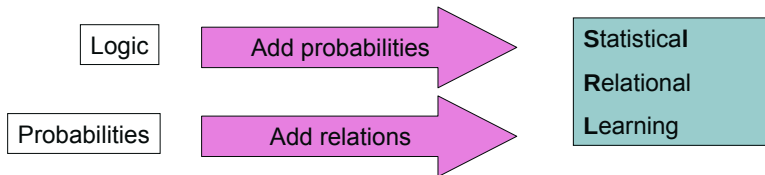
- Three views...

● ● ● View 1: Alphabet Soup



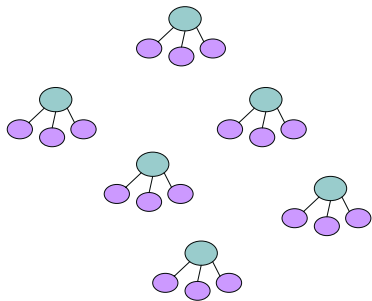
● ● ● View 2: Representation Soup

- Hierarchical Bayesian Model + Relational Representation

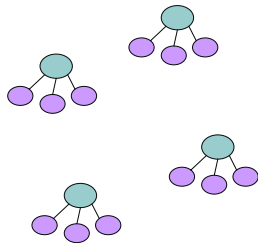


● ● ● View 3: Data Soup

Training Data

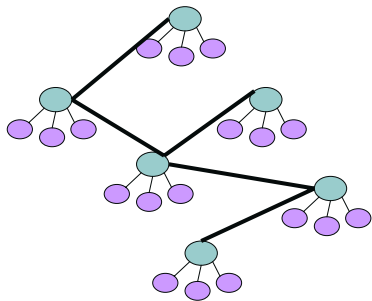


Test Data

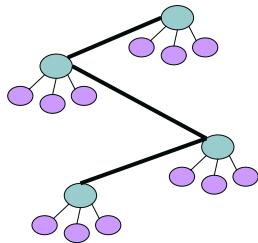


View 3: Data Soup

Training Data



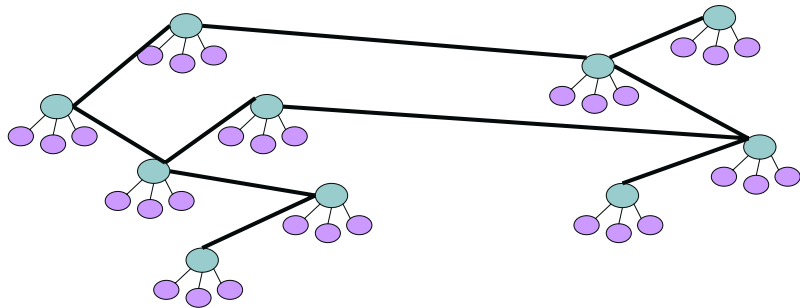
Test Data



View 3: Data Soup

Training Data

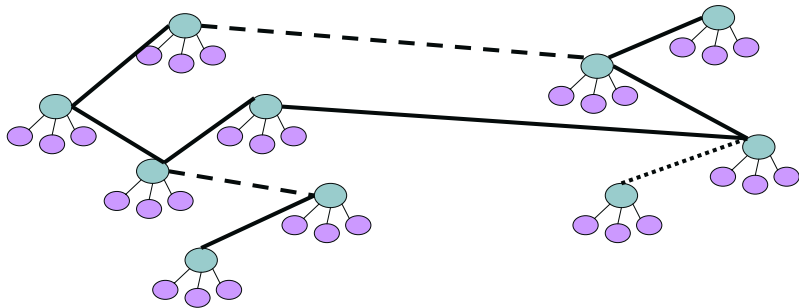
Test Data



View 3: Data Soup

Training Data

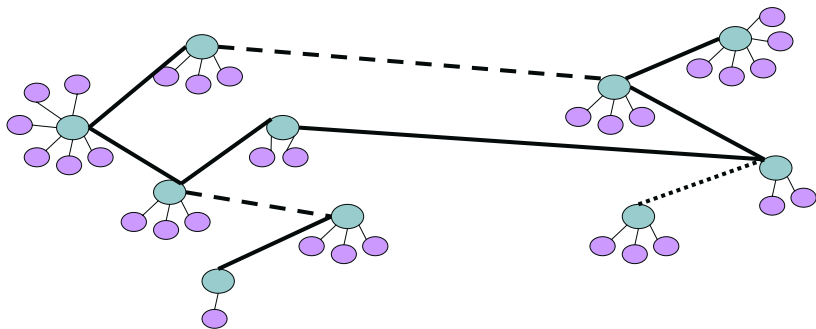
Test Data



View 3: Data Soup

Training Data

Test Data



View 3: Data Soup

Training Data

Test Data

