Discussion Panel: Open Problems

Objective function

- What are we trying to optimize? And is that actually related to what we want?
  - Build PDF; take level set => MLE/MAP
  - One-class kernel approach => structural risk
  - Find compact covering set => compactness, connectedness
  - User preference function => utility
  - MIL?
  - Active learning?

- Modeling what you "expect" to see (fits in "std" ML)
- Of the residual: what do we actually like?

Validation

- Data sets?
- How do we know that the models we produce actually work well?
  - Holdout data/cross-validation -- hard to get labeled data (what do labels mean?)
  - Expert evaluation -- hard to compare methods

Background knowledge

- Bayesian prior, vs...
- Model (hypothesis) space, vs...
- User preference function, vs...
- ???

Nonstationarity (Concept drift)

- Deepak's method for tracking time series
- Maloof
- Lot of work in COLT
- ???

Feature selection/ construction

- What is the right set of features to look at...
- ... to detect something that you don't know about yet?

Adversarial learning

- Attacker doesn't like us -- trying to break our learner
  - Strong theoretical bounds
  - Tease apart attack from normal
  - Game theory
- Subversion/training attacks

Data types

- Time series
- Images/video
- Text distribution
- Networks -- are we detecting interesting things in...
  - Data related to the network, or...
  - ... interesting structure in the network?

Sensor fusion

- Merging (radically) different data sources
- Which data sources to look at (feature selection)