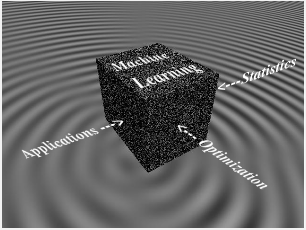
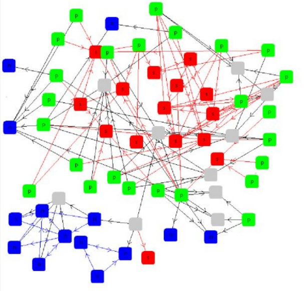
# Algorithms for Graph Label Prediction: Settings

18 November 2008



# General



1. Universum = set of $n < \infty$ objects $\mathcal{V} = \{v_i\}_{i=1}^n$
2. Edges

$$\mathcal{E} = \{a_{ij} = a_{ji} \in \{0,1\}\}_{v_i, v_j \in \mathcal{V}}$$

... or weighted edges

$$\mathcal{E} = \{a_{ij} = a_{ji} \geq 0\}_{v_i, v_j \in \mathcal{V}}$$

3. Organized in a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
4. Each Object has a true label $\{y_i \in \{-1,1\}\}_{i=1}^n$, or $y : \mathcal{V} \to \{-1,1\}$.
5. Subset $S \subset \mathcal{V}$ of Labels is revealed
6. Predict the remaining labels $T = \mathcal{V}/S$
7. Hypothesis $q \in \mathcal{H}$ where

$$\mathcal{H} \subset \left\{ \mathcal{V} \to \{-1,1\} \right\}$$

1 Inference

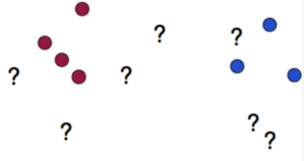
2 Online Learning

3 Hypothesis Spaces

4 Conclusions

# I. Inference

# Transductive Inference




1. $S$ sampled randomly without replacement from $\mathcal{V}$
2. Only $y(S)$ given
3. Risk of a labeling $q : \mathcal{V} \to \{-1, 1\}$

$$\begin{aligned} \mathcal{R}(q) &= P\left(q(V) \neq y(V)\right) \\ &= E\left[I(q(V) \neq y(V))\right] \\ &= \frac{1}{n}\sum_{i=1}^{n} I(q(v_i) \neq y(v_i)) \end{aligned}$$

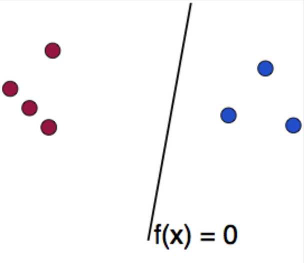
where $V$ random vertex from $\mathcal{V}$

4. Empirical Risk given $S$

$$\mathcal{R}_S(q) = \frac{1}{|S|}\sum_{i \in S} I(q(v_i) \neq y(v_i))$$

5. Risk on Test Set

$$\mathcal{R}_T(q) = \frac{1}{|T|}\sum_{i \in T} I(q(v_i) \neq y(v_i))$$




1. $S$ sampled iid from $P$ from $\mathcal{V}$
2. Only $\mathcal{G}|S$ and $y(S)$ given.
3. Risk of a labeling $q : \mathcal{V} \to \{-1, 1\}$

$$\begin{aligned} \mathcal{R}(q) &= P\left(q(V) \neq y(V)\right) \\ &= E\left[I(q(V) \neq y(V))\right] \end{aligned}$$

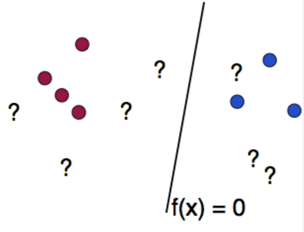
where $V$ random vertex from $\mathcal{V}$

4. Empirical Risk given $S$

$$\mathcal{R}_S(q) = \frac{1}{|S|} \sum_{i \in S} I(q(v_i) \neq y(v_i))$$

5. Risk on Test Set

$$\mathcal{R}_T(q) = \frac{1}{|T|} \sum_{i \in T} I(q(v_i) \neq y(v_i)) \approx \mathcal{R}(q)$$




1. $S$ sampled iid from $P$ from $\mathcal{V}$, $n \to \infty$
2. $S_s$ random subsample from $S$, and $S_t = S/S_s$
3. Only $\mathcal{G}|S$ and $y(S_s)$ given.
4. Risk of a labeling $q : \mathcal{V} \to \{-1, 1\}$

$$\begin{aligned} \mathcal{R}(q) &= P\left(q(V) \neq y(V)\right) \\ &= E\left[I(q(V) \neq y(V))\right] \end{aligned}$$

   where $V$ random vertex from $\mathcal{V}$

5. Empirical Risk given $S$

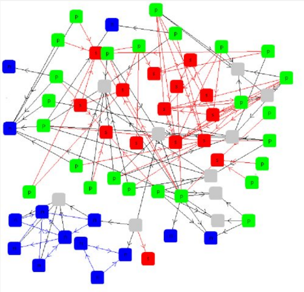
$$\mathcal{R}_S(q) = \frac{1}{|S|} \sum_{i \in S} I(q(v_i) \neq y(v_i))$$

6. Risk on Test Set

$$\mathcal{R}_T(q) = \frac{1}{|T|} \sum_{i \in T} I(q(v_i) \neq y(v_i)) \approx \mathcal{R}(q)$$

# II. Online Learning

# Online + Transductive

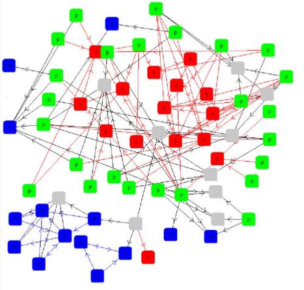


Iterate $t = 1, \ldots \quad (S^0 = \{\})$

1. Adversary asks $'y(v_t)?'$
2. $S^t = S^{t-1} \cup v_t$
3. Prediction $q^t$ based on $\mathcal{G}$ and $y(S^{t-1})$
4. Nature reveals $y(v_t)$
5. Regret

$$\sum_{i=1}^{t} P\Big(q^i(V) \neq y(V)\Big) - \min_{q \in \mathcal{H}} \sum_{i=1}^{t} P\Big(q(V) \neq y(V)\Big)$$

# Online + Supervised



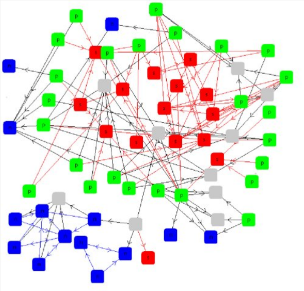
Iterate $t = 1, \ldots \quad (S^0 = \{\})$

1. Adversary asks $'y(v_t)?'$
2. $S^t = S^{t-1} \cup v_t$
3. Prediction $q^t$ based on $\mathcal{G}|S^t$ and $y(S^{t-1})$
4. Nature reveals $y(v_t)$
5. Regret

$$\sum_{i=1}^{t} P\Big(q^i(V) \neq y(V)\Big) - \min_{q \in \mathcal{H}} \sum_{i=1}^{t} P\Big(q(V) \neq y(V)\Big)$$

$\rightarrow \mathcal{G}$ only revealed incrementally

# Online + Semi-Supervised

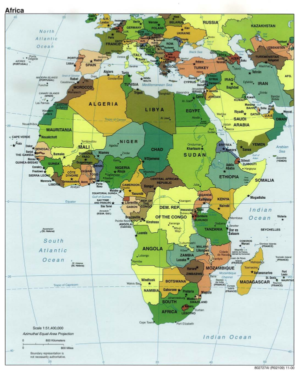


?Transductive Learning Phase + Extension Operator

# III. Hypothesis Spaces

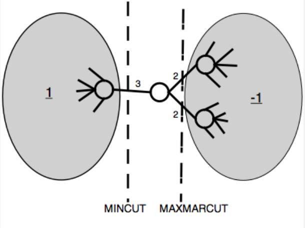
# What is a good Hypothesis Space

For learning Concept of 'Continent'

# Plausible Hypothesis Spaces

$$\mathcal{H} \subset \left\{ q : \mathcal{V} \to \{-1, 1\} \right\}$$



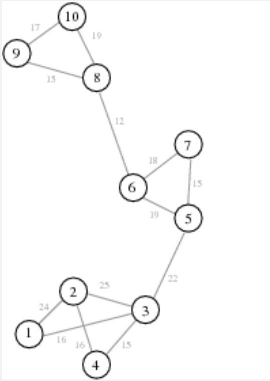

1. Counterpart to margin?
2. Graph Cut

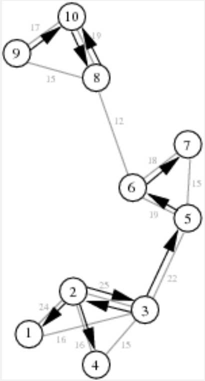
$$\text{cut}(q) = \sum_{q(v_i) \neq q(v_j)} a_{ij} = \frac{1}{4} q^T L q$$

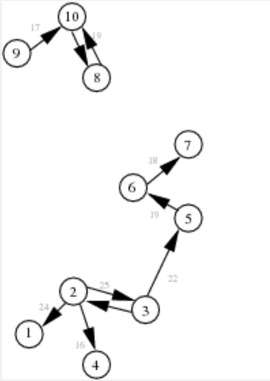
3. Bounded Graph Cut

$$\mathcal{H}_B = \left\{ q : \mathcal{V} \to \{-1, 1\} : \ \text{cut}(q) \leq B \right\}$$

Local Consistency with 1NN



# IV. Conclusions

# Open Problems

1. Form $\mathcal{G}$
2. Suitable $\mathcal{H}$ and regularization?
3. Model Selection?
4. Tuning Graph or Algorithm?
5. Extension Operator?
6. Noise Sources/ Robustness w.r.t. $\mathcal{G}$