Legislative Prediction with Dual Uncertainty Minimization from Heterogeneous Information

Yu Cheng *, Ankit Agrawal *, Huan Liu ɨ,
Alok Choudhary *

ɨ CSE, Arizona State University
* EECS, Northwestern University

Talk Agenda

Motivation

Method

- Previous Methods and Comparisons
- Dual Uncertainty Minimization over Heterogeneous

Experiments

- Setting and Result
- Basic Analysis

Conclusion

Problem Scenario

House or Senate

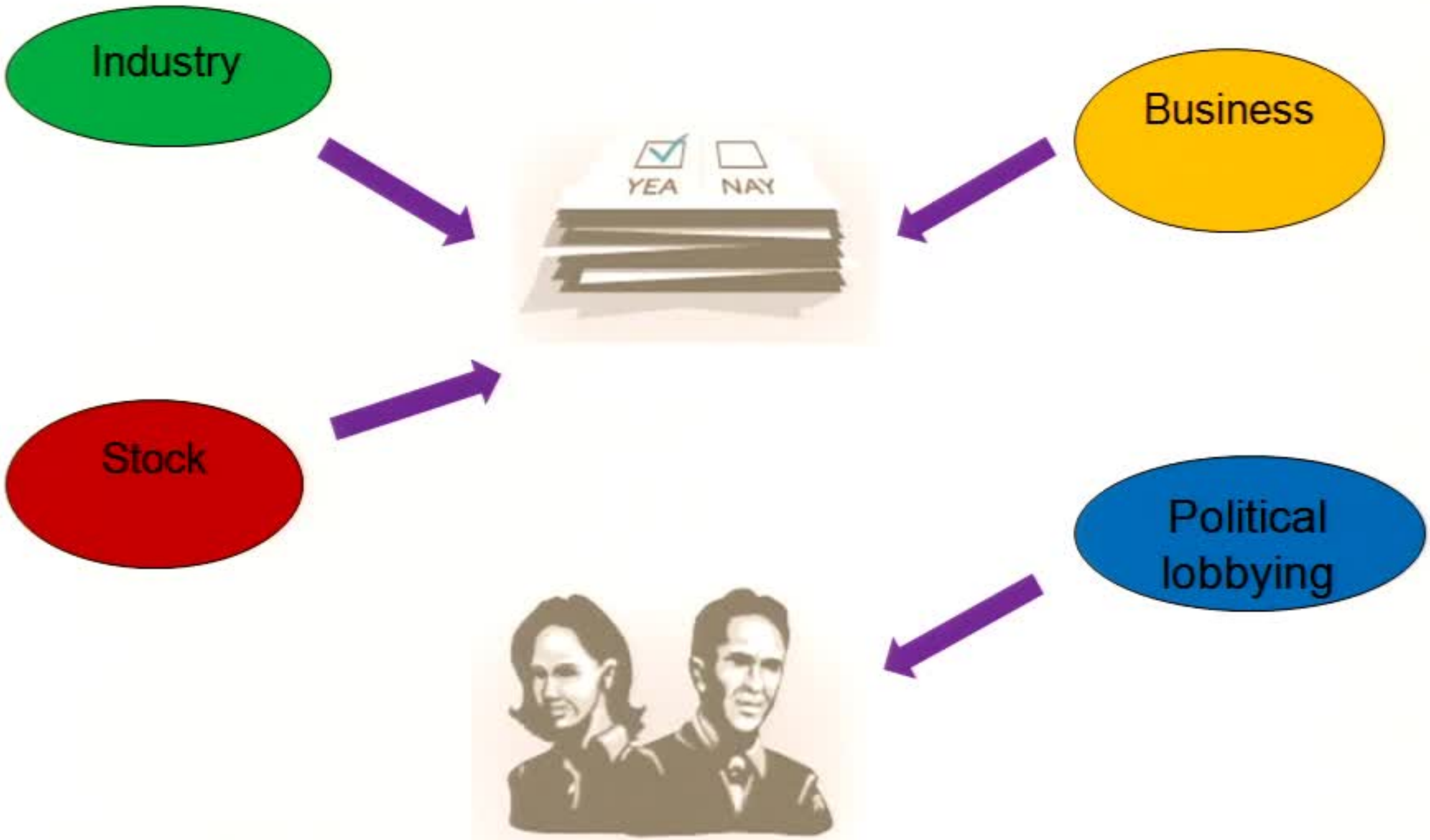
Legislator

Pass or Not Pass

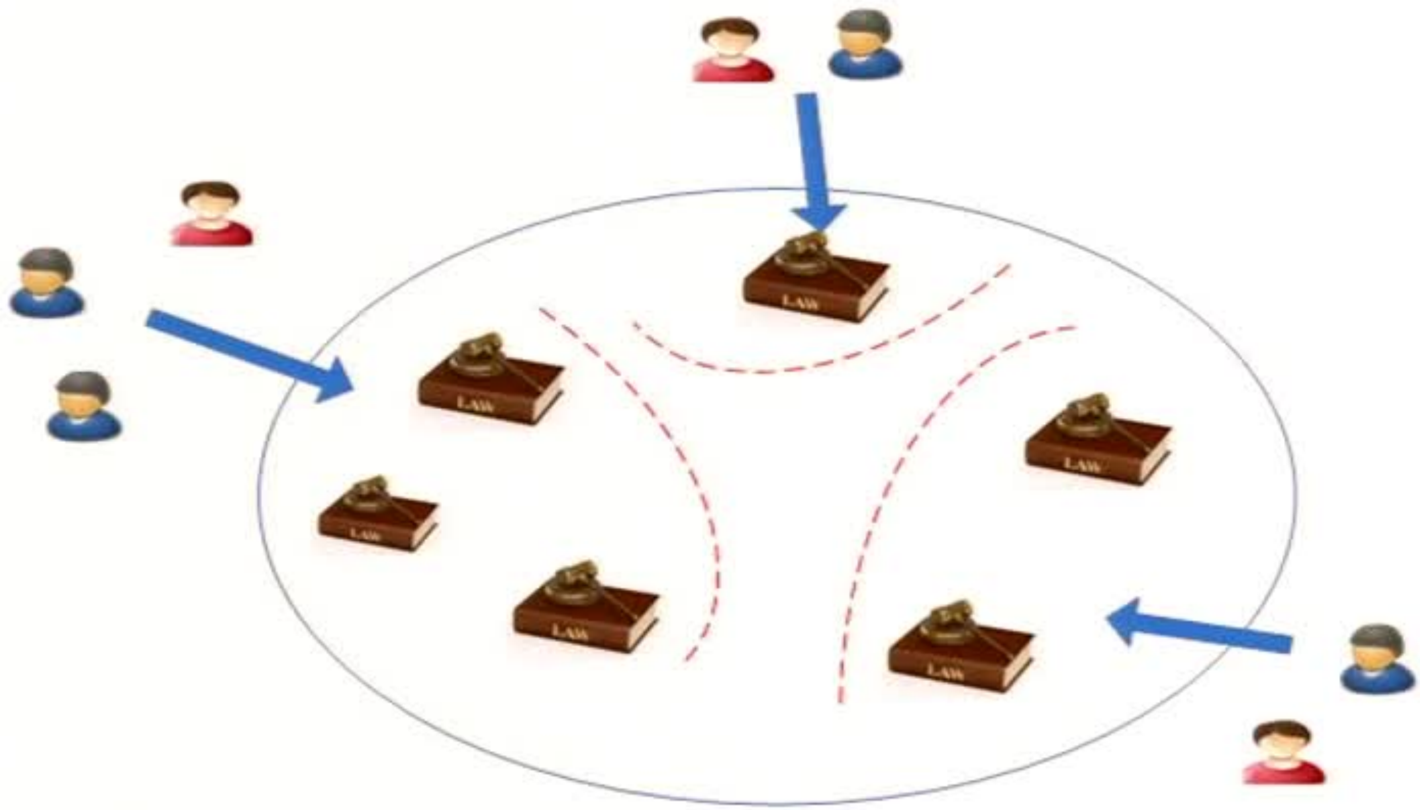
Legislative

Problem Importance

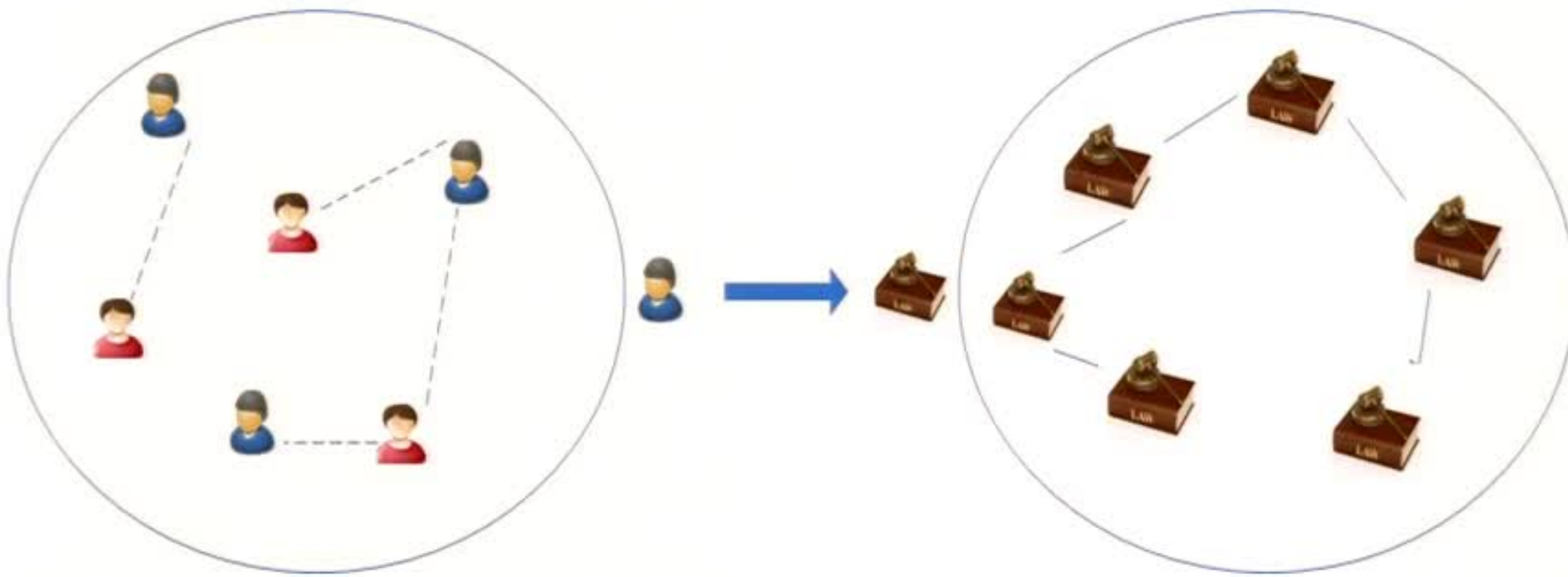
Potential Applications



Previous Method #1: Ideal Point Model



Previous Method #2: Heterogeneous Mixture

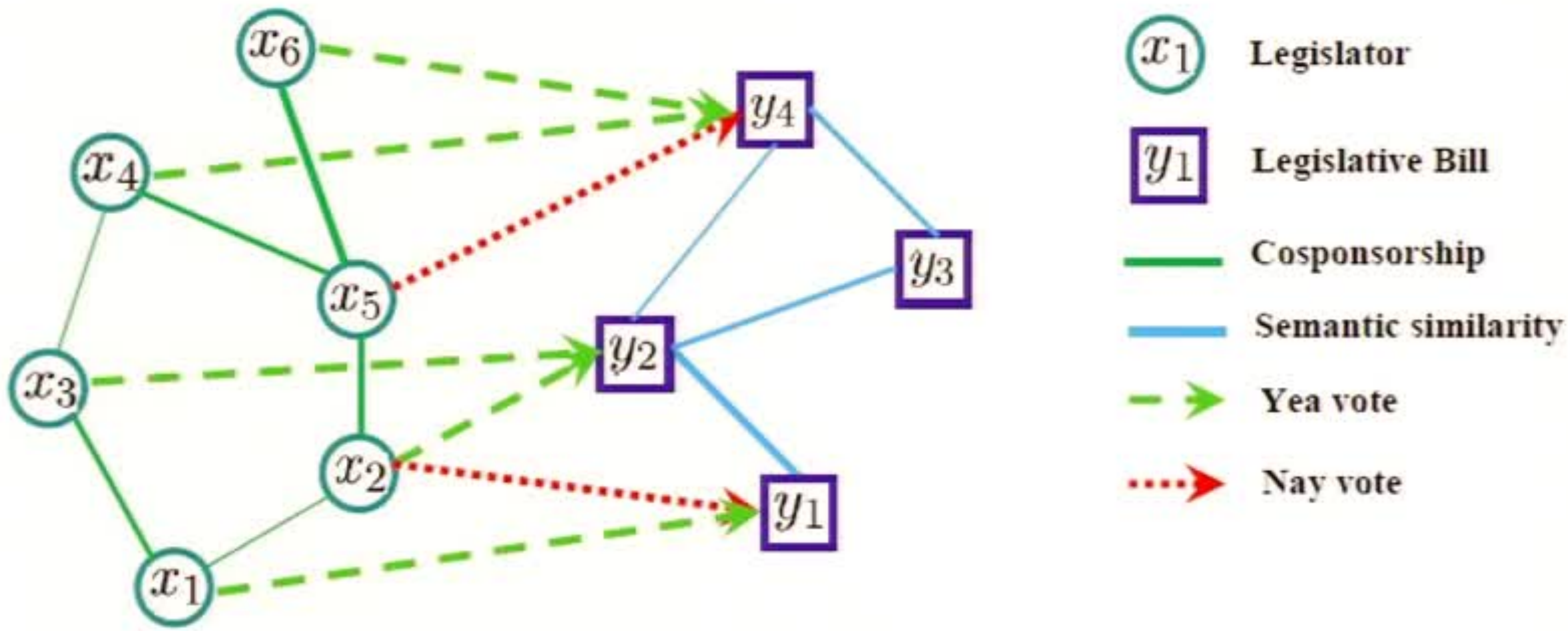


co-sponsorship

semantic similarity

Previous Method #3: Random Walk Over Heterogeneous

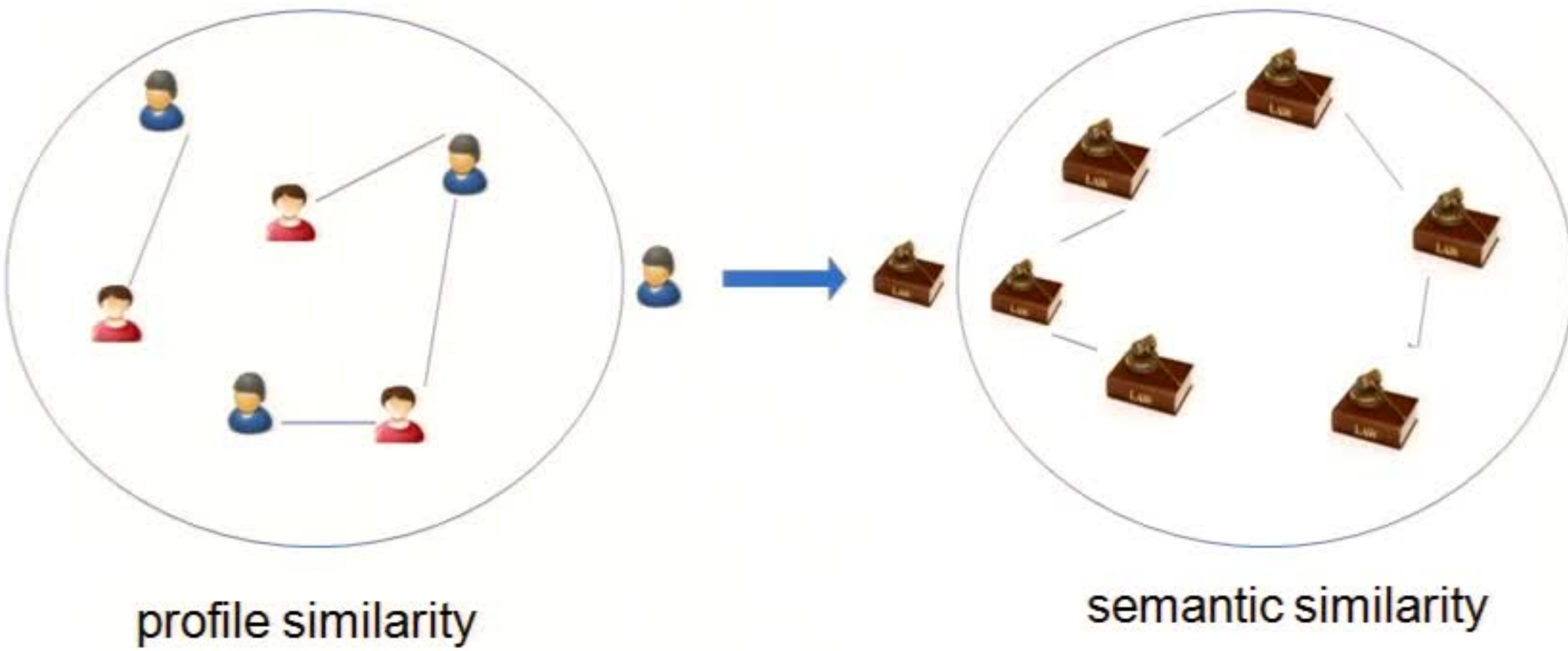
Q: new legislative? missing votes? scalability?



Our Proposed Method

Q: new legislative? missing votes? scalability?

A: profile similarity; semi-supervised; inductive



Overall Comparisons

	Information Leveraged			Algorithm Characteristic			
	Bill	Legislator	Roll Call	NB	NL	MS	IN
Ideal Point Model [1]	○	×	○	○	×	×	×
Ideal Point Topic Model [5]	○	×	○	○	×	×	○
Time-Evolving Voting [27]	○	×	○	○	×	×	○
Random Walks on a Heterogeneous Graph [29]	○	○	○	○	×	×	×
k-NN bills [29]	○	▽	○	○	×	×	×
k-NN legislators [29]	×	▽	○	×	×	×	×
Dual Uncertainty Minimization from Heterogenous Information	○	○	○	○	○	○	○

NB: new bill prediction NL: new legislator prediction MS: missing vote IN: inductive
○:Yes ×:NO ▽:Partial

Problem Solver

A good framework to estimate α β[G. Niu]:

Squared-loss mutual information (v over x):

$$\mathrm{SMI}(\mathbf{x}, \mathrm{v}) = \mathrm{p}(\mathbf{x})\mathrm{p}(\mathrm{v})(\frac{\mathrm{p}(\mathbf{x}, \mathrm{v})}{\mathrm{p}(\mathbf{x})\mathrm{p}(\mathrm{v})} - 1)^2 \mathbf{x}\,\mathrm{v}$$

Kernel version:

$$\mathrm{SMI}(\mathbf{x}, \mathrm{v}) = \mathrm{Const.} + \frac{\mathrm{c}}{2\mathrm{n}} \sum_{\mathrm{v} \in \mathbf{V}} \alpha_{\mathrm{v}}^{\mathrm{T}} \mathbf{K}^2 \alpha_{\mathrm{v}}$$

Add empirical loss:

$$\min_{\alpha} \mathcal{L}(v, \hat{v}) - \gamma \widehat{\mathrm{SMI}}$$

empirical loss: labeled data

minimize uncertainty on all data

G. Niu et al. Squared-loss mutual information regularization: A novel information-theoretic approach to semi-supervised learning. ICML 2013.

Problem Formulation

Legislators: $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^{m} \in \mathbf{R}^{d_l \times m}$ Votes: $\mathbf{V} \in R^{m \times n}$

Bills: $\mathbf{Y} \in \mathbf{R}^{d_b \times n}$

$$v_{ij} = \begin{cases} 1 & : \ \mathbf{x}_i \ \ votes\ yea\ on \ \ \mathbf{y}_j \\ -1 & : \ \mathbf{x}_i \ \ votes\ nay\ on \ \ \mathbf{y}_j \\ 0 & : \ otherwise. \end{cases}$$

missing votes

Given a training set, estimate:

$$\hat{v} = \arg\max_{v \in \mathcal{V}} p(v|\mathbf{x}, \mathbf{y})$$

Approximate with a combined kernel model:

$$p(v|\mathbf{x}, \mathbf{y}) \approx q(v|\mathbf{x}, \mathbf{y}; \alpha, \beta) = \sum_{i=1}^{n} \alpha_i k_b(\mathbf{y}, \mathbf{y_i}) + \sum_{j=1}^{m} \beta_j k_l(\mathbf{x}, \mathbf{x_j}).$$

Assumptions: 1) similar legislators have similar voting behavior; 2) one legislators will have the same votes on similar bills

Problem Solver

A good framework to estimate α β[G. Niu]:

Squared-loss mutual information (v over x):

$$\mathrm{SMI}(\mathbf{x}, \mathrm{v}) = \mathrm{p}(\mathbf{x})\mathrm{p}(\mathrm{v})\left(\frac{\mathrm{p}(\mathbf{x}, \mathrm{v})}{\mathrm{p}(\mathbf{x})\mathrm{p}(\mathrm{v})} - 1\right)^2 \mathbf{x}\,\mathrm{v}$$

Kernel version:

$$\mathrm{SMI}(\mathbf{x}, \mathrm{v}) = \mathrm{Const.} + \frac{\mathrm{c}}{2\mathrm{n}} \sum_{\mathrm{v} \in \mathbf{V}} \alpha_{\mathrm{v}}^{\mathrm{T}} \mathbf{K}^2 \alpha_{\mathrm{v}}$$

Add empirical loss:

$$\min_{\alpha} \mathcal{L}(v, \hat{v}) - \gamma \widehat{\mathrm{SMI}}$$

empirical loss: labeled data

minimize uncertainty on all data

G. Niu et al. Squared-loss mutual information regularization: A novel information-theoretic approach to semi-supervised learning. ICML 2013.

Problem Solver

Empirical loss:

$$\mathcal{L}(v,\hat{v}) = \int_{\mathcal{Y}} \sum_{v\in\mathbf{V}} (p(v|\mathbf{y}) - q(v|\mathbf{y},\alpha))^2 p(\mathbf{y})$$

SMI (v over x,y)

$$\widehat{\mathrm{SMI}} := \mathrm{SMI}(\mathbf{x},\mathrm{v}) + \mathrm{SMI}(\mathbf{y},\mathrm{v})$$
$$= \mathrm{Const.} + \frac{1}{\mathrm{n}} \sum_{\mathrm{v}\in\mathbf{V}} \alpha_{\mathrm{v}}^{\mathrm{T}} \mathbf{K}_{\mathrm{b}}^{2} \alpha_{\mathrm{v}} + \frac{1}{\mathrm{m}} \sum_{\mathrm{v}\in\mathbf{V}} \beta_{\mathrm{v}}^{\mathrm{T}} \mathbf{K}_{\mathrm{l}}^{2} \beta_{\mathrm{v}}$$

Put together:

$$\min_{\alpha,\beta} \mathcal{L}(v,\hat{v}) - \gamma\widehat{\mathrm{SMI}}$$

Out-of-sample Predication

New bills,
New Legislators,
New bill & legislators: approximate α β from others

linear neighborhood reconstruction [Roweis]:

$$W^* = \arg\min_{W} \quad ||\mathbf{x} - \sum_{\mathbf{x}_j \in \mathcal{N}(\mathbf{x})} W_{ij}\mathbf{x}_j||$$

$$\sum_j W_{ij} = 1, W_{ij} \geq 0$$

$$\alpha = \sum_{\mathbf{x}_j \in \mathcal{N}(\mathbf{x})} W^*_{ij}\alpha_{(j)}, \quad \beta = \sum_{\mathbf{x}_j \in \mathcal{N}(\mathbf{x})} W^*_{ij}\beta_{(j)}$$

S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *SCIENCE*, 290:2323–2326, 2000

Performance Evaluation – Setting

- Source: https://www.govtrack.us/ Wikipedia Pages
  110-111: 1585 bills, 631 unique legislators, 638,955 votes
  112-113: 695 bills, 628 legislators, 289,067 valid votes

- Evaluation: 1) accuracy on random missing voting; 2) accuracy on sequential voting

- Compared approaches:
  - Yes
  - IPTM
  - RWHG
  - DUMHI

Prediction Results

- 110-111 sessions

Method	A-Accuray	G-Sim
Yes	0.8548	0.8698
IPTM	0.887	0.87
RWHG	0.911	0.9036
DUMHI	0.9315	0.9223

- 112-113 sessions

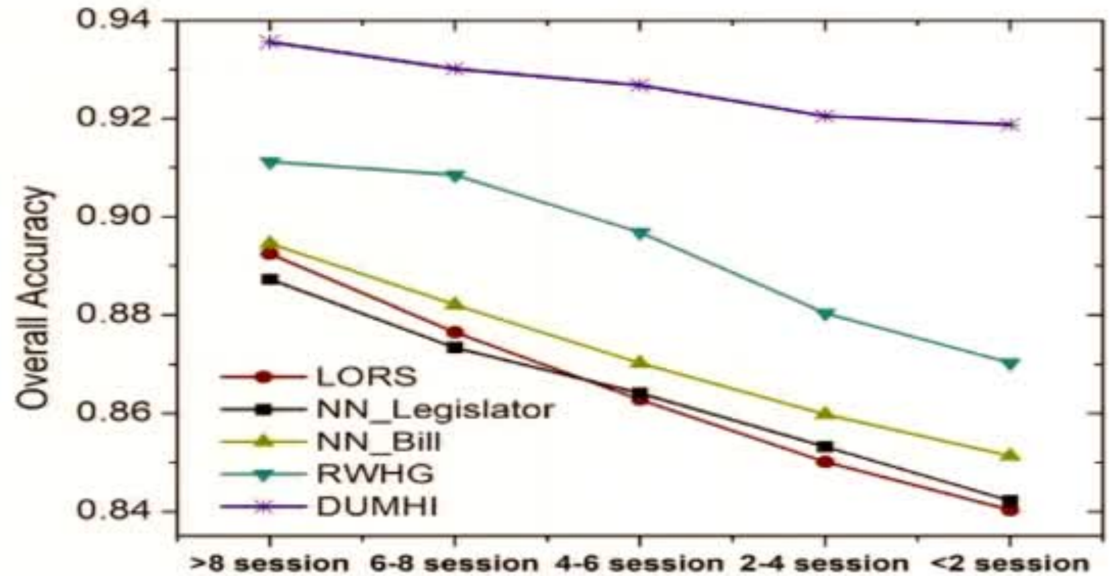
Method	A-Accuray	G-Sim
Yes	0.8219	0.8326
IPTM	0.8689	0.8712
RWHG	0.8973	0.8859
DUMHI	0.9204	0.9086

FiscalNote

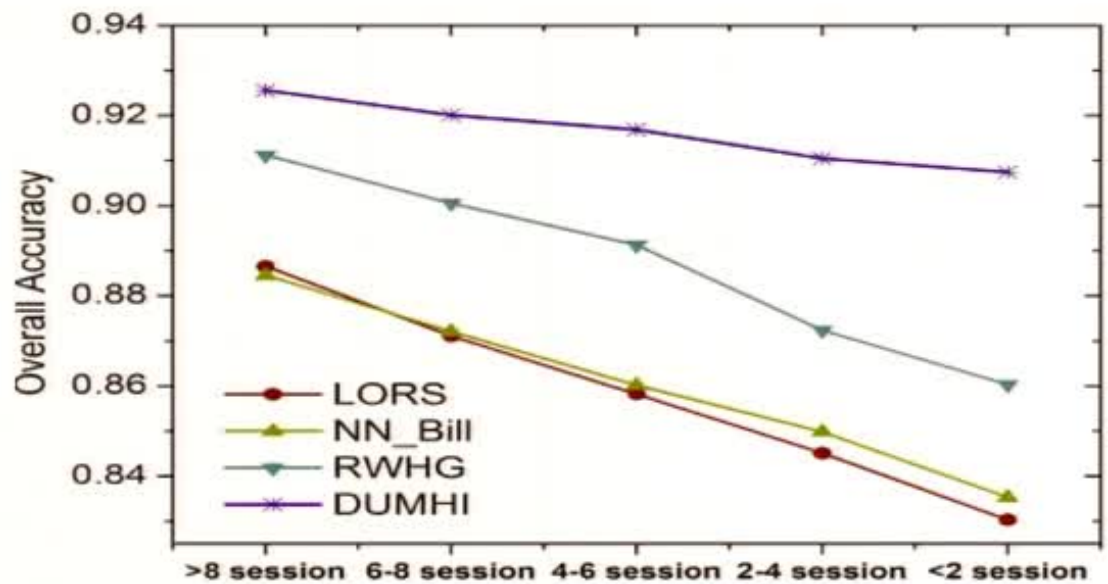
Reported accuracy: 93%, but no method provided

Predication vs. Service Time

- 110-111



- 112-113



Factor Feature Analysis

Profile Factor	Information Gain
Leadership	0.1307
Ideology	0.2658
Legislator Type	0.0456
Religion	0.1877
Years of Searvice	0.0872
Partisanship	0.2025
Gender	0.0263