Learning Latent Groups with Hinge-loss Markov Random Fields

Stephen H. Bach, Bert Huang, and Lise Getoor
University of Maryland, College Park

Introduction

- Can we build **tractable** models that learn **latent mixed-memberships** in **rich data**?
- We achieve this goal using a graphical model that encodes **complex, yet interpretable dependencies** among
  - group membership
  - language usage
  - social interactions
- We demonstrate our approach by learning indicators of **latent political preferences**

Hinge-loss Markov Random Fields

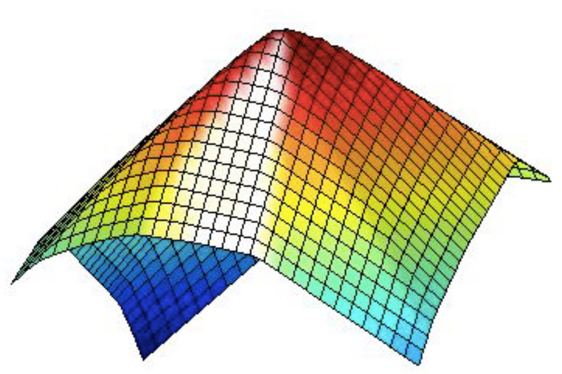
Hinge-loss Markov random fields are undirected graphical models analogous to discrete MRFs.

- Variables are **continuous valued** in [0,1]
- Potentials are hinge-loss functions

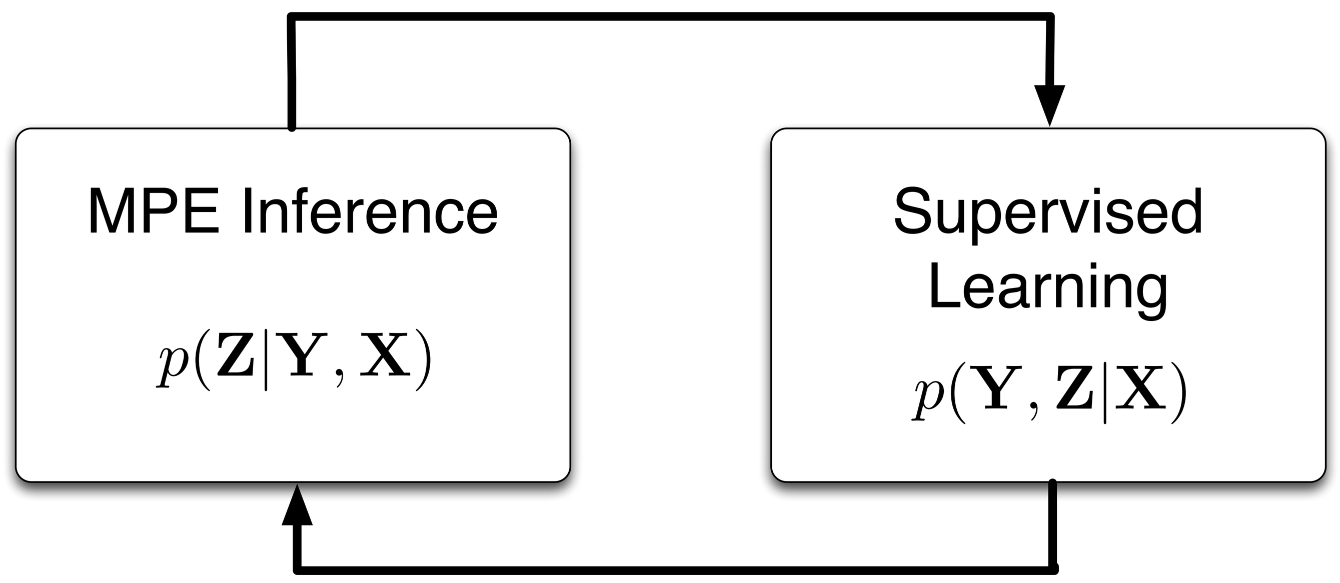
$$p(\mathbf{Y}|\mathbf{X}) = \frac{1}{Z}\exp\left[-\sum_{j=1}^{m} w_j \max\left\{\ell_j(\mathbf{Y},\mathbf{X}), 0\right\}^{p_j}\right]$$

Features

- **Easily constructed and interpreted** by mapping between logical rules and hinge-loss functions
- **Linear constraints** enable mixture models
- Log-concave density function admits **fast, exact MPE inference** using the alternating direction method of multipliers (ADMM)



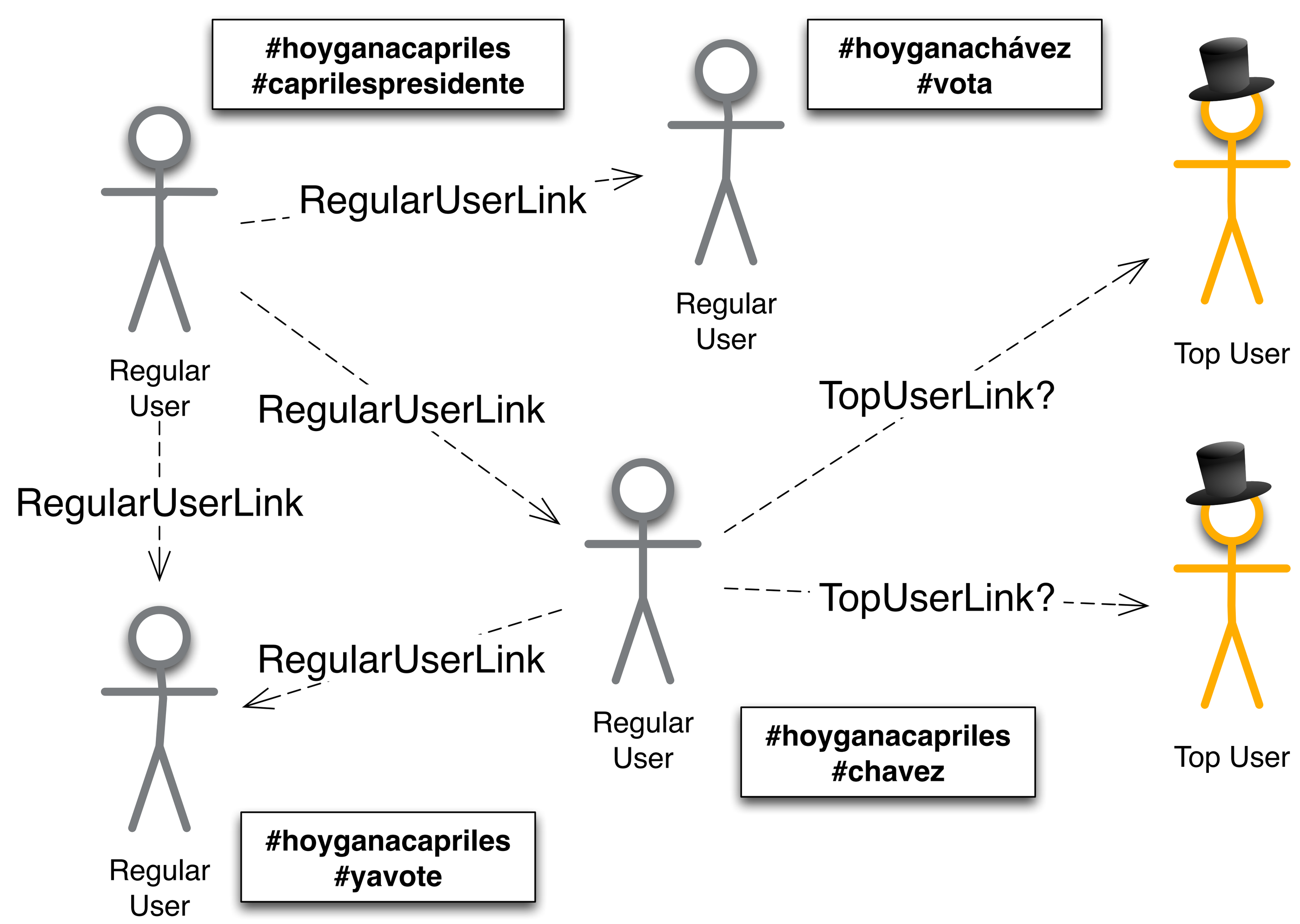
Latent-Variable Learning



Hard EM excels because of tractable inference and learning with expressive continuous variables

Latent-Group Model

Domain



Rules

$$w_{h,g} : \textsc{UsedHashtag}(U,h) \rightarrow \textsc{InGroup}(U,g) \qquad \forall h \in H, \forall g \in \mathcal{G}$$

$$w_{\text{social}} : \textsc{RegularUserLink}(U_1,U_3) \wedge \textsc{RegularUserLink}(U_2,U_3) \wedge U_1 \neq U_2 \wedge \textsc{InGroup}(U_1,G) \rightarrow \textsc{InGroup}(U_2,G).$$

$$w_{g,t} : \textsc{InGroup}(U,g) \rightarrow \textsc{TopUserLink}(U,t) \qquad \forall g \in \mathcal{G}, \forall t \in T$$

Learn more: http://psl.umiacs.umd.edu

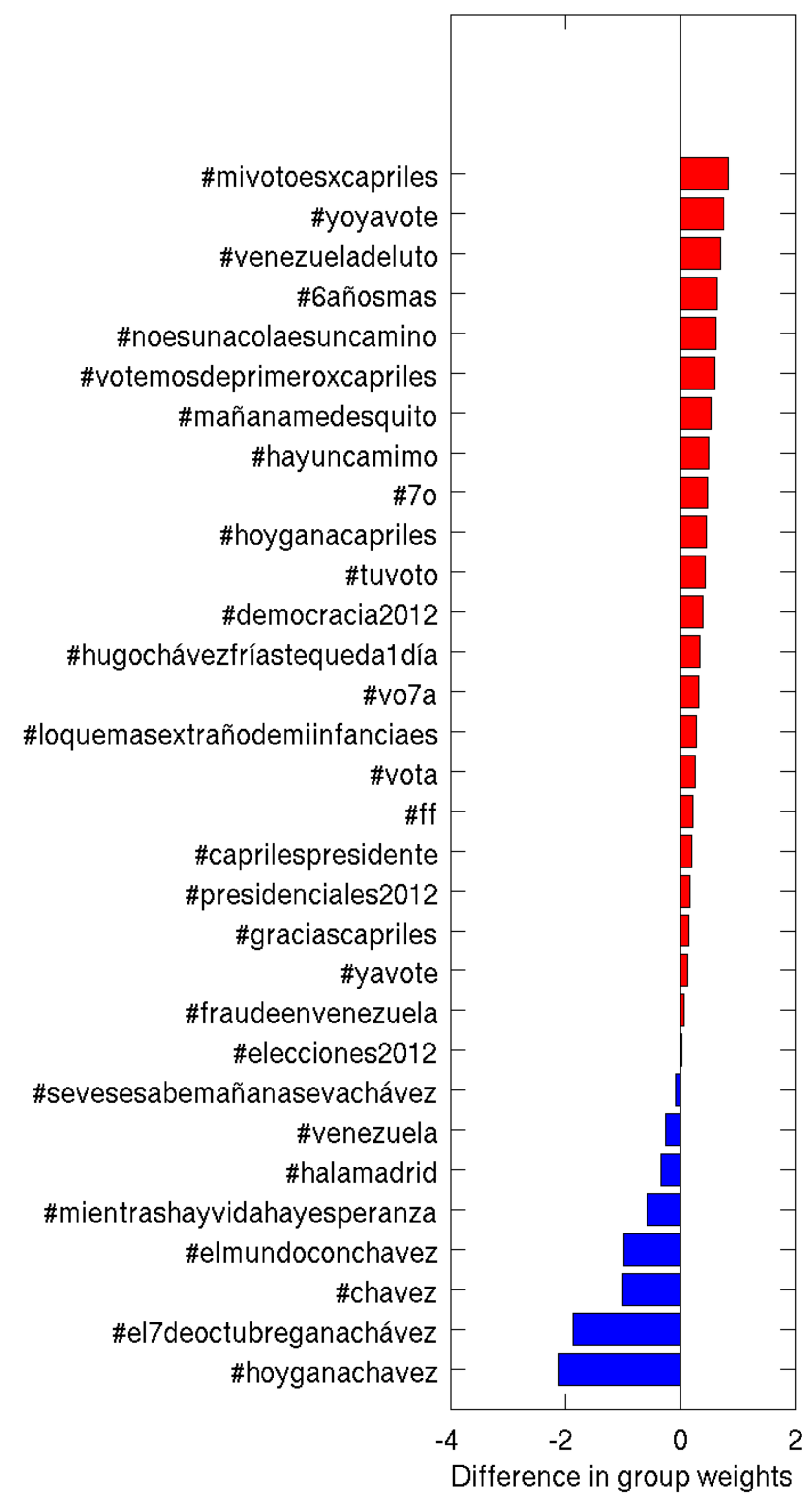
Case Study: 2012 Venezuelan Presidential Election

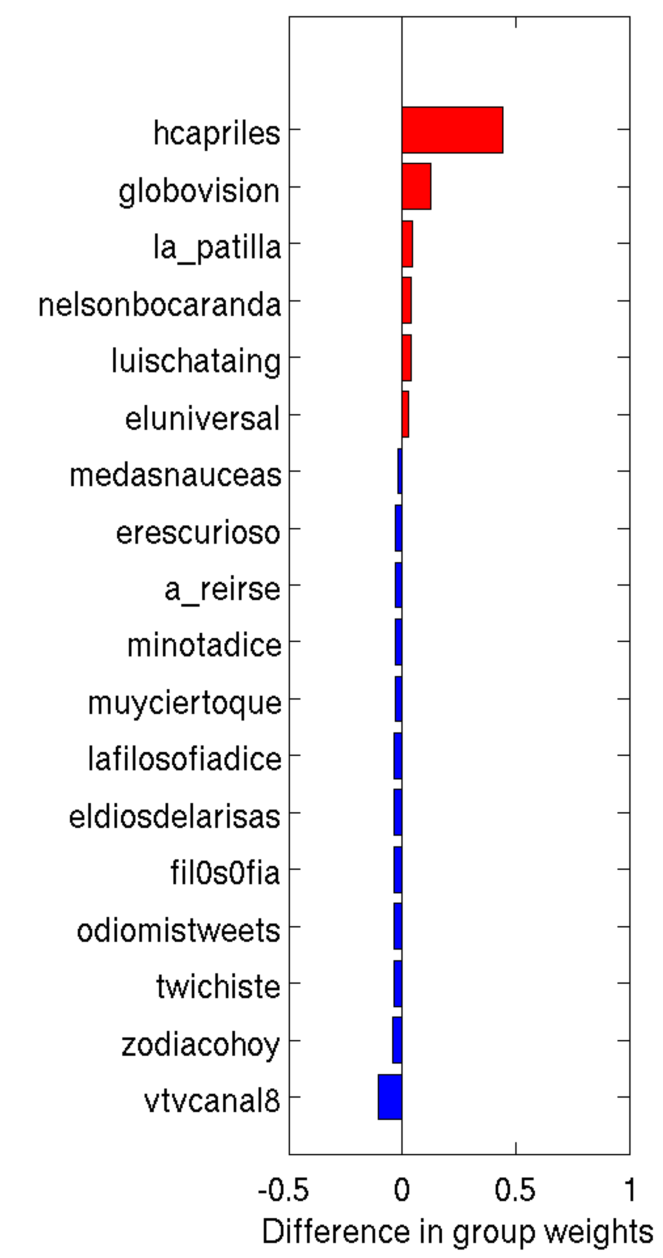
Candidates

Hugo Chávez
Incumbent
Won with 55% of vote

Henrique Capriles
Challenger

Learned Groups





Highlights

- Pro-Chávez and pro-Capriles hashtags separated
- Anti-Chávez hashtags identified as pro-Capriles
- Superficially pro-Chávez hashtag "#6añosmas" is correctly identified as pro-Capriles
- Capriles's account associated with pro-Capriles group
- Independent and state-owned media accounts separated

This work was supported by NSF grant CCF0937094 and IARPA via DoI/NBC contract number D12PC00337. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the U.S. Government.