Supplementary Material for Estimating Accuracy from Unlabeled Data: A Probabilistic Logic Approach

Emmanouil A. Platanios
Carnegie Mellon University
Pittsburgh, PA
e.a.platanios@cs.cmu.edu

Hoifung Poon
Microsoft Research
Redmond, WA
hoifung@microsoft.com

Tom M. Mitchell
Carnegie Mellon University
Pittsburgh, PA
tom.mitchell@cs.cmu.edu

Eric Horvitz
Microsoft Research
Redmond, WA
horvitz@microsoft.com

1 Grounding Algorithm

Algorithm 1: Grounding algorithm.

Input: $\hat{f}_j^d(X)$, for $d = 1, \dots, D$, and $j = 1, \dots, N^d$, only for observed values, set of all pairwise mutual-exclusion constraints $\mathit{ME} = \{d_1^i, d_2^i\}_{i=1}^M$, and set of all subsumption constraints $\mathit{SUB} = \{d_1^i, d_2^i\}_{i=1}^S$.

Create empty sets G_p and G_l
foreach *observed* $\hat{f}_j^d(X)$ **do**
  Add $\hat{f}_j^d(X)$, e_j^d, and $f^d(X)$ to G_p
  Add $\hat{f}_j^d(X) \land \neg e_j^d \rightarrow f^d(X)$ and $\neg \hat{f}_j^d(X) \land \neg e_j^d \rightarrow \neg f^d(X)$ to G_l
  Add $\hat{f}_j^d(X) \land e_j^d \rightarrow \neg f^d(X)$ and $\neg \hat{f}_j^d(X) \land e_j^d \rightarrow f^d(X)$ to G_l
  Add $\hat{f}_j^d(X) \rightarrow f^d(X)$ and $\neg \hat{f}_j^d(X) \rightarrow \neg f^d(X)$ to G_l
  foreach *pair* (d_1, d_2) *in ME* **do**
    if $d_1 = d$ **then**
      Add $f^{d_2}(X)$ to G_p
      Add $\mathrm{ME}(d_1, d_2) \land \hat{f}_j^{d_1}(X) \land f^{d_2}(X) \rightarrow e_j^{d_1}$ to G_l
    else if $d_2 = d$ **then**
      Add $f^{d_1}(X)$ to G_p
      Add $\mathrm{ME}(d_2, d_1) \land \hat{f}_j^{d_2}(X) \land f^{d_1}(X) \rightarrow e_j^{d_2}$ to G_l
  foreach *pair* (d_1, d_2) *in SUB* **do**
    if $d_1 = d$ **then**
      Add $f^{d_2}(X)$ to G_p
      Add $\mathrm{SUB}(d_1, d_2) \land \neg \hat{f}_j^{d_1}(X) \land f^{d_2}(X) \rightarrow e_j^{d_1}$ to G_l

Output: Set of ground predicates G_p and set of ground rules G_l.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

2 PSL Consensus ADMM Inference Algorithm

Algorithm 2: PSL consensus ADMM inference algorithm.

Input: Observed ground predicate values $\mathbf{X}$, objective terms $\boldsymbol{\ell}$, $\boldsymbol{p}$, rule weights $\boldsymbol{\lambda}$, parameter ρ, and mapping from variable copies' indices to consensus variables' indices $\mathcal{G}$.

Randomly initialize all $\mathbf{Y}$ (consensus variables) and α_j (Lagrange multipliers) for $j = 1, \ldots, k$, and then randomly initialize the variable copies $\mathbf{y}_j$ for $j = 1, \ldots, k$, corresponding to each subproblem

while *not converged* **do**

  for $i = 1, \ldots, k$ **do**

    $\alpha_j \leftarrow \alpha_j + \rho(\mathbf{y}_j - \mathbf{Y}_{\mathcal{G}(j,:)})$

    $\mathbf{y}_j \leftarrow \arg\min_{\mathbf{y}_j} \big[\lambda_j [\max\{\ell_j(\mathbf{X}, \mathbf{y}_j)\}]^{p_j}$
    $\quad + \frac{\rho}{2} \|\mathbf{y}_j - \mathbf{Y}_{\mathcal{G}(j,:)} + \frac{1}{\rho}\alpha_j\|_2^2\big]$

  for $i = 1, \ldots, \mathit{length}(\boldsymbol{Y})$ **do**

    $\mathbf{Y}_i \leftarrow \frac{\sum_{\mathcal{G}(j,d)=i} \left([\mathbf{y}_j]_d + \frac{1}{\rho}[\alpha_j]_d\right)}{\sum_{\mathcal{G}(j,d)=i} \mathbf{1}}$

    Project $\mathbf{Y}_i$ on the interval $[0, 1]$

Output: Inferred ground predicate values $\mathbf{Y}$.