Interest Inference via Structure-Constrained Multi-Source Multi-Task Learning

NExT Research Centre

Xuemeng SONG, Liqiang NIE, Luming ZHANG, Maofu LIU, Tat-Seng CHUA

{sxmustc, nieliqiang, zglumg}@gmail.com, liumaofu@wust.edu.cn, chuats@comp.nus.edu.sg

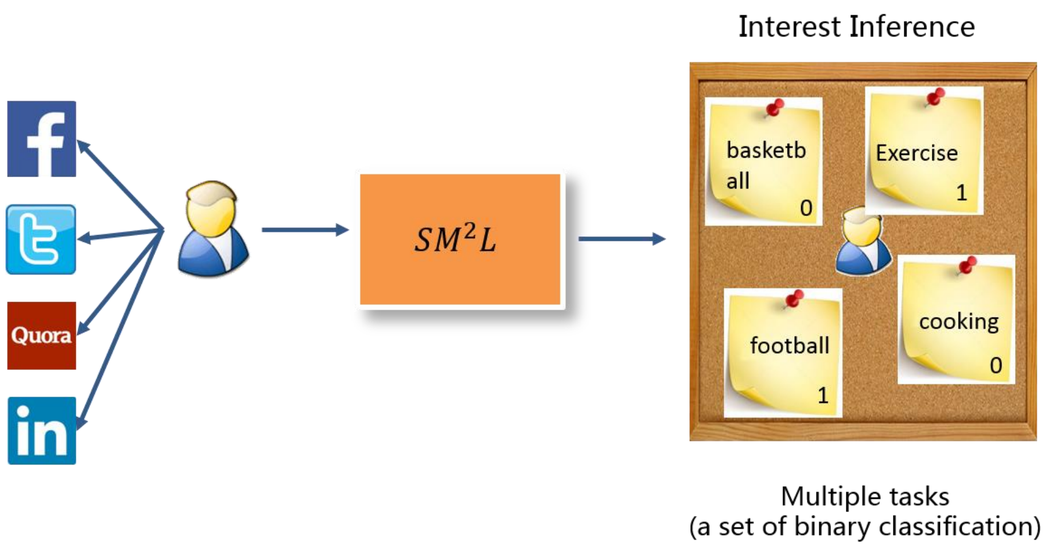
❖ Background

- User interest inference is the basis for many applications, such as adaptive E-learning and personalized service.
- Multi-platform use is on the rise: 52% of online adults now use two or more social media sites, a significant increase from 2013, when it stood at 42% of internet users.
- Fusing cues from multiple sources can potentially boost the performance of user interest inference by a large margin.

❖ Objective

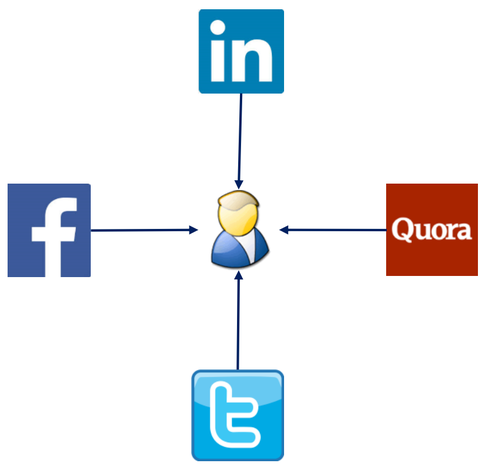
- We aim to propose a structure-constrained multi-source multi-task learning (SM2L) scheme to infer users' interests from multiple social networks.
- This scheme jointly regularizes two important aspects: source consistency and tree-guided task relatedness modeling.

❖ Architecture

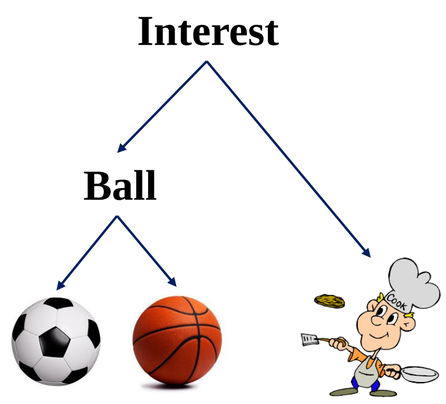


❖ Prior Knowledge

- Source consistency: The rationale is that interests reflected by different social networks for the same person should be similar, and hence the disagreement among the prediction results should be penalized.
- Tree-guided task relatedness modeling: We believe that interests are usually not independent but correlated in a non-uniform way.

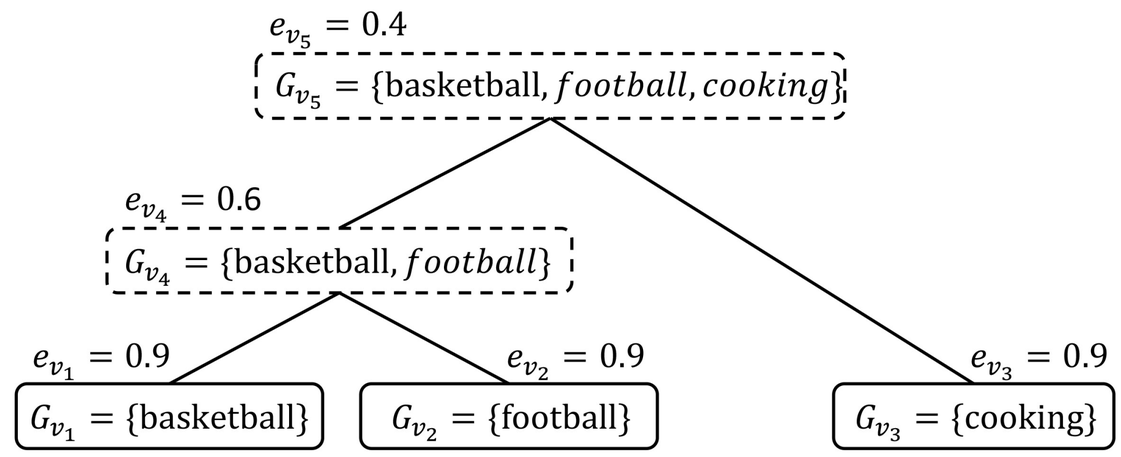


(1) Source consistency



(2) Tree-guided task relatedness modeling

❖ Tree Structure



❖ Problem Formulation

$$\mathbf{\Gamma} = \underbrace{\frac{1}{2N}\sum_{t=1}^{T}\left\|\mathbf{y}_t - \sum_{s=1}^{S}\frac{1}{S}\mathbf{X}_s\mathbf{w}_{st}\right\|^2}_{\text{Loss function}} + \underbrace{\frac{\lambda}{2}\sum_{s=1}^{S}\sum_{d=1}^{D_s}\sum_{v\in\mathcal{V}} e_v\left\|\mathbf{w}^d_{sG_v}\right\|}_{\text{Tree-guided group lasso}} + \underbrace{\frac{\mu}{2N}\sum_{t=1}^{T}\sum_{s=1}^{S}\sum_{s'\neq s}\left\|\mathbf{X}_s\mathbf{w}_{st} - \mathbf{X}_{s'}\mathbf{w}_{s't}\right\|^2}_{\text{Source consistency}}$$

❖ Data Sources

❖ Experiment Results

Table1: Performance comparison among various models

Approaches	$P@K$ (%)	$S@K$ (%)
SVM	8.69	54.69
RLS	24.32	73.86
regMVMT	24.69	74.54
SM²L-eu	25.50	73.80
SM²L-iu	24.56	74.11
SM²L-e	**25.72**	**74.57**
SM²L-i	**26.50**	**74.85**

Table2: Contribution of individual social network and their combinations

Social network combinations	$P@K$ (%)	$S@K$ (%)
Twitter	24.75	73.05
Facebook	19.59	69.74
Quora	20.97	68.19
Twitter+Facebook	25.51	74.98
Twitter+Quora	24.89	74.41
Facebook+Quora	22.52	71.80
Twitter+Facebook+Quora	**26.50**	**74.85**

Acknowledgement:
This research is supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.

Email: nieliqiang@gmail.com
Web : next.comp.nus.edu.sg