

Knowledge Transformation for Cross-Domain Sentiment Classification

Tao Li *

School of Computer Science
Florida International University
taoli@cs.fiu.edu

Chris Ding

CSE Department
University of Texas at Arlington
chqding@uta.edu

Vikas Sindhwani

IBM T.J. Watson Research
Center
vsindhw@us.ibm.com

Yi Zhang

School of Computer Science
Florida International University
yzhan004@cs.fiu.edu

ABSTRACT

With the explosion of user-generated web2.0 content in the form of blogs, wikis and discussion forums, the Internet has rapidly become a massive dynamic repository of public opinion on an unbounded range of topics. A key enabler of opinion extraction and summarization is sentiment classification: the task of automatically identifying whether a given piece of text expresses positive or negative opinion towards a topic of interest. Building high-quality sentiment classifiers using standard text categorization methods is challenging due to the lack of labeled data in a target domain. In this paper, we consider the problem of cross-domain sentiment analysis: can one, for instance, download *rated* movie reviews from rottentomatoes.com or IMBD discussion forums, learn linguistic expressions and sentiment-laden terms that *generally characterize* opinionated reviews and then successfully transfer this knowledge to the target domain, thereby building high-quality sentiment models without manual effort? We outline a novel sentiment transfer mechanism based on constrained non-negative matrix tri-factorizations of term-document matrices in the source and target domains. We report some preliminary results with this approach.

Categories and Subject Descriptors: I.2 [Artificial Intelligence]: Learning

General Terms: Algorithms, Experimentation

Keywords: Sentiment analysis, Transfer learning, Nonnegative matrix factorization

1. INTRODUCTION

Suppose that we download movie reviews from an online review site and compile a large term-document matrix, X_1 representing m terms and n documents. Assume further that the sentiment associated with at least some of these documents is known from the explicit ratings given by movie enthusiasts that visit these sites. From such a dataset, we can hope to identify general linguistic indicators of positive

*The work is partially supported by NSF grants IIS-0546280 and DMS-0844513.

and negative opinions, e.g., sentiment-laden terms such as “great” and “awful” respectively. The end goal of this exercise would be to apply this knowledge to gauge sentiment around documents in a new *target* domain X_2 , which for example may be a collection of blogs posts talking about products and services of keen interest to a company. In this poster, we propose a mechanism to transform document label information in one domain to another *via words*, using a principled matrix factorization framework.

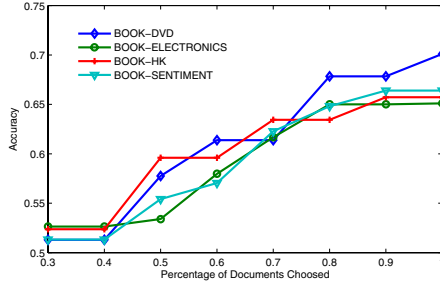
Transfer learning and domain adaptation (see e.g., [1]) are quite natural tools for sentiment analysis applications, but most efforts in these directions have been fairly recent (see e.g., [2] and references therein). A in-depth survey of sentiment analysis literature [5] shows emphasis on standard text classification or lexical rule based approaches mainly geared towards single-domain applications. In this regard, our main objective in this poster is to report progress on a new contribution to cross-domain sentiment analysis.

2. NON-NEGATIVE FACTORIZATIONS

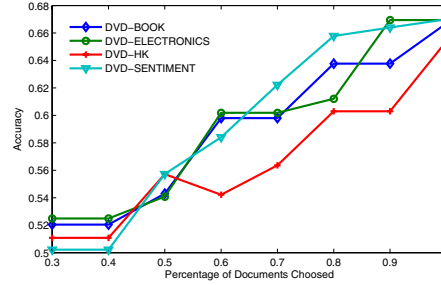
We first recall nonnegative matrix factorization models [3] to give a background for our methodology. We build on tri-factorization based co-clustering models closely related to probabilistic Latent Semantic Indexing (pLSI) and probabilistic Topic models. Given a $m \times n$ term-document matrix X from a domain, we construct a tri-factorization, $X \approx FSG^T$, where G is an $n \times k$ nonnegative matrix whose i^{th} row represents the posterior probability that the i^{th} document belongs to one of k classes or topics. Similarly, F is an $m \times k$ matrix specifying posterior class-probabilities for words. The $k \times k$ core factor S gives a condensed view of X – in the pLSI context, it may be understood as class probability distribution. Under orthogonality constraints on F, G and non-negativity constraints on F, S, G the factorization above can be computed using simple iterative update rules. See [3] for applications to clustering. In the sentiment analysis case, we set $k = 2$ with the intention of interpreting the first and second columns of F, G as positive and negative sentiment probability respectively. In order to enable transfer learning, we also need to add further constraints that express prior knowledge. These are discussed next.

3. KNOWLEDGE TRANSFER MECHANISM

Let X_1 and X_2 be term-document matrices in the source



(a) Source Domain: BOOK



(b) Source Domain: DVD

and target domains respectively. We will shortly consider constrained factorizations $X_1 \approx F_1 S_1 G_1$ and $X_2 \approx F_2 S_2 G_2$ respectively. We assume that the number of terms is the same: if the vocabularies differ, we simply pad zero columns and re-express the matrices under the same unified vocabulary so that the column indices in both matrices correspond to the same word. Moreover, let V represent a $m \times m$ diagonal matrix with $V_{ii} = 1$ if i is a shared word, i.e., it occurs in both domains, or in other words the associated column is non-zero for both X_1 and X_2 .

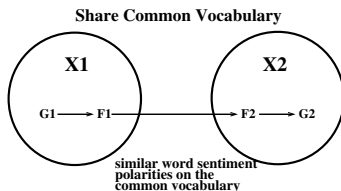
In domain X_1 , we have some labeled documents. The partial labels on documents can be described using G_1^0 where $(G_1^0)_{i1} = 1$ if the document expresses positive sentiment, and $(G_1^0)_{i2} = 1$ for negative sentiment. Note that one may also use soft sentiment polarities though our experiments are conducted with hard assignments. We first transfer this document label knowledge to words by learning F in a 2-way clustering,

$$\min_{F, G, S} \|X_1 - FSG^T\|^2 + \beta \text{Tr} \left[(G - G_1^0)^T C_1 (G - G_1^0) \right] \quad (1)$$

where the notation $\text{Tr}(A)$ means trace of the matrix A . Here, $\beta > 0$ is a parameter which determines the extent to which we enforce $G_1 \approx G_1^0$, C_1 is a $n \times n$ diagonal matrix whose entry $(C_1)_{ii} = 1$ if the category of the i -th document is known (i.e., specified by the i -th row of G_1^0) and $(C_1)_{ii} = 0$ otherwise. Note that if $C_1 = I$, then we know the class orientation of all the documents and thus have a full specification of G_1^0 . Let the solution be F_1, S_1 , which contains the knowledge to be transferred to X_2 . This transfer is achieved by next solving the following 2-way clustering:

$$\min_{F, G, S} \|X_2 - FSG^T\|^2 + \alpha \text{Tr} \left[(F - F_1)^T V (F - F_1) \right] \quad (2)$$

The key part here is the second term which enforces the word sentiment polarity on X_2 to be approximately close to F_1 which is learnt from X_1 ; the extent of this approximation is determined by parameter $\alpha > 0$. The constraint only applies to the common terms in X_1, X_2 as enforced by the diagonal matrix V . The solution for Eq.(2) gives (F_2, S_2, G_2) . This mechanism therefore transfers the document sentiment G_1 of domain X_1 to the document sentiment G_2 in domain X_2 , via word sentiment polarities in F_1 , schematically shown below.



The optimization in Eq.(1) can be solved using the following update rules:

$$S_{ik} \leftarrow S_{ik} \frac{(F^T XG)_{ik}}{(FF^T SGTG)_{ik}}, F_{ik} \leftarrow F_{ik} \frac{(XGS^T)_{ik}}{(FF^T XGS^T)_{ik}}$$

$$G_{jk} \leftarrow G_{jk} \frac{(X^T FS + \beta C_1 G_1^0)_{jk}}{(GG^T X^T FS + \beta GG^T C_1 G_1^0)_{jk}} \quad (3)$$

The optimization in Eq.(2) can be solved using the following update rules (updates for S are same as in Eq.(3)):

$$G_{jk} \leftarrow G_{jk} \frac{(X^T FS)_{jk}}{(GG^T X^T FS)_{jk}}, F_{ik} \leftarrow F_{ik} \frac{(XGS^T + \alpha V F_1)_{ik}}{(FF^T XGS^T + \alpha V F_1)_{ik}}$$

The algorithm consists of an iterative procedure using the above three rules until convergence. The correctness and convergence of the updating rules can be rigorously proved using the standard auxiliary function approach [4].

4. EXPERIMENTS

The Figure at the top of this page presents preliminary experiments on cross domain sentiment analysis on Amazon.com reviews for 4 product types: Houseware-Kitchen (HK), Books, DVDs, and Electronics [2]. Results are shown for two source domains Books and DVDs, and four target domains. As can be seen, as supervision is increased in the source domain in the form of labeled data, performance gains effectively show up in the target domains also. We also note the effect of task relatedness: for example, DVD transfers more effectively than Books to Electronics but less effectively to Houseware-Kitchen. For each source domain, the top 10 highest entropy words as per F_1 are as follows: (1) HK – *easy return love poor perfect excellent waste disappointed broke clean*, (2) Electronics: *return excellent price terrible waste perfect highly easy poor returned*, (3) DVD: *waste worst bad boring horrible love wonderful enjoy excellent ridiculous*, (4) BOOKS: *boring disappointing waste excellent bad wonderful poor poorly love easy*. Our initial results are promising but more empirical studies need to be done before we can accurately evaluate the potential of our method.

5. REFERENCES

- [1] S.Ben-David, J. Blitzer, K. Crammer, and F. Pereira. Analysis of representations for domain adaptation. In *NIPS*, 2007.
- [2] J. Blitzer, M. Dredze, and F. Pereira. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In *ACL*, pages 440–447, 2007.
- [3] C. Ding, T. Li, W. Peng, and H. Park. Orthogonal nonnegative matrix tri-factorizations for clustering. In *KDD*, pages 126–135, 2006.
- [4] D. Lee and H. Seung. Algorithms for non-negative matrix factorization. In *NIPS*, 2001.
- [5] B. Pang and L. Lee. *Opinion mining and sentiment analysis*. Foundations and Trends in Information Retrieval, 2(1):1–135.