Multi Cross Domain Recommendation Using Item Embedding And Canonical Correlation Analysis

Masahiro Kazama
Recruit Technologies Co., Ltd.
Tokyo, Japan
masahiro_kazama@r.recruit.co.jp

István Varga
Recruit Technologies Co., Ltd.
Tokyo, Japan
vistvan@r.recruit.co.jp

ABSTRACT

In a multi-service environment it is crucial to be able to leverage user behavior from one or more domains to create personalized recommendations in the other domain. In our paper, we present a robust transfer learning approach that successfully captures user behavior across multiple domains. First, we vectorize users and items in each domain independently. Second, using a handful of common users across domain pairs, we project each domain vector space into a common vector space using canonical correlation analysis (CCA). Next, recommendations can be performed by recommending the items in any domains that are closest to the user's vector in the common space. We also experimented on what kind of domain combination works well.

KEYWORDS

Recommender Systems, Canonical Correlation Analysis, Transfer Learning, Item Embedding

ACM Reference format:
Masahiro Kazama and István Varga. 2017. Multi Cross Domain Recommendation Using Item Embedding And Canonical Correlation Analysis. *RecSys '17 Poster Proceedings, Como, Italy, August 27–31, 2017,* 2 pages.

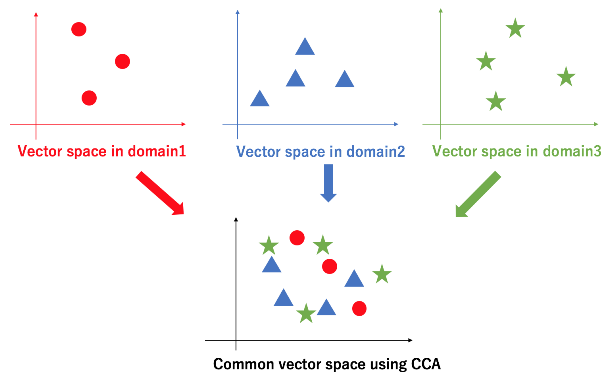

Figure 1: Overview of this paper: Users and items are vectorized in each domain independently and those vectors are mapped into a common space by canonical correlation analysis(CCA). An object in the figure denotes a user or an item.

1 INTRODUCTION

In recent years the ever-increasing ubiquity of e-commerce is allowing us to purchase virtually every product or service that we could desire. With this increasing growth, one can also observe a trend in the interconnection of e-commerce businesses by means of common IDs. With such common IDs, services are not only getting increased visibility, but consumers are also receiving personalized product recommendations in a new domain. In this paper we propose a simple and robust transfer learning method that facilitates cross domain recommendation that leverages canonical correlation analysis (CCA) to represent multiple domains in a single vector space. All users and items are represented as vectors in the common space; therefore, items from any domain can be recommended to users by calculating the similarity between the users' vector and the items' vector in the common space. Figure 1 shows the overview of our research.

Our contributions in this papers are as follows:

- We applied item embedding technique and CCA to multi cross domain recommendation in a simple and robust way.
- We experimented on what kind of domain combination works well.

RecSys '17 Poster Proceedings, Como, Italy
2017.

2 RELATED WORK

There are two main research areas related to our proposal.

CCA [2] is actively explored in the field of multimodal representation. Numerous studies have been conducted where using CCA, the embeddings of various types of data (e.g., image, text) by deep learning and word2vec are projected into a common vector space, where various other subsequent tasks can be performed [1, 4].

Another relevant area is recommendation systems, some studies proposed recommendation methods using CCA [5]. Our proposal retains its simplicity and robustness with three or more domains.

3 PROPOSED METHOD

Our approach consists of two steps. First, we calculate the vector representation of each domain using word2vec. Word2vec is a natural language processing method that generates semantic representations of words [3] , but it can also be used for rating data in the following way: by considering an item as a word and the sequence of items evaluated favorably by a user (i.e., rated with 4 or 5 stars) as a sentence, the resulting sentences can be fed into word2vec to achieve item embeddings. We used skip-gram model with hierarchical softmax. Next, we define the user vector as the average of item vectors favorably evaluated by the user. In this step, the vector representations of each domain are calculated independently, so it is easy to parallelize them.

As a second step, we project each vector space into a common space using CCA. Let us illustrate this projection with three domains. Note that it can be easily applied to $N(> 3)$ domains. Let the vector of user and item in domain $i(=1,2,3)$ be x_i and y_i. Let C_{ij} be the covariance matrix of x_i and x_j. Note that it is not required to have users who are active in all domains. CCA is calculated when there are some common users in each domain pair. By solving the following eigenvalue equation, the transformation vector $w = (w_1^T, w_2^T, w_3^T)^T$ can be obtained.

$$\begin{pmatrix} 0 & C_{12} & C_{13} \\ C_{21} & 0 & C_{23} \\ C_{31} & C_{32} & 0 \end{pmatrix} w = \lambda \begin{pmatrix} C_{11} & 0 & 0 \\ 0 & C_{22} & 0 \\ 0 & 0 & C_{33} \end{pmatrix} w$$

Using the transformation vector w, both items and users from each domain can be projected to the common vector space. As a result, personalized recommendations can be performed by simply recommending the items closest to the user's vector in the common vector space. Even if the user is active in only one domain, we can recommend the other domain's items. We employed cosine similarity for a similarity measure.

4 EXPERIMENTS

We attempted to investigate the domain characteristics that can be used to improve recommendation performance. As a baseline, two domains are projected into one common vector space using CCA with the usage of 80% of the common users. 20% of common users were held out for testing purposes: based on user actions on one domain, we predict the items on the other domain, comparing the recommended items with the actual actions.

Next we added a new third domain and calculated the common vector space with these three domains using CCA. Our hypothesis is that if the new domain correlates highly with at least one of the already existing domains, the new domain will enrich the common vector space, thus improving performance. However, if the new domain is less similar to the already existing ones, this will mostly introduce noise, meaning performance will either drop or not improve significantly.

For the experiments we used the yelp rating dataset[1] in which users rated various items. After a basic sanity check (i.e., a removal of multiple categories), we conducted experiments using the top 5 categories: Restaurants, Shopping, Food, Beauty&Spas, and Health&Medical.

Table 1 shows the item prediction performance in the Food category, based on user behavior from the Shopping category when the additional categories (i.e., Beauty&Spas, Health&Medical, and Restaurants) were added. Baseline performance (recall @ 50 = 12.8%) did not increase with the addition of Beauty&Spas and even decreased by 1.5 points with the addition of Health&Medical (11.3%). However, with the addition of Restaurants, we achieved an improvement of 2.0 points (14.8%), showing that information from this new domain strengthened the relationship between our original two categories, Food and Shopping.

Table 2 shows the similarity between our original two domains and the additional domains. Similarity is calculated using the average of the top 5 canonical correlations calculated using CCA. We can observe that the original domains have a relatively low correlation with Beauty&Spas and Health&Medical. On the other hand, Restaurants displays a high correlation with both Food and Shopping.

[1] https://www.yelp.com/dataset_challenge

Table 1: Result of recall@50: Baseline uses Shopping and Food only. We experimented 5 times and shows the average and standard diviation.

Model	Recall@50(%)
Baseline	12.8 ± 0.4
Add Restaurants	**14.8 ± 0.7**
Add Beauty&Spas	12.9 ± 0.7
Add Health&Medical	11.3 ± 0.6

Table 2: Correlation between each domain: correlation is the average of top 5 canonical correlations between the users' vectors in one domain and the other domain.

	Restaurants	Beauty&Spas	Health&Medical
Shopping	**0.92**	0.74	0.76
Food	**0.96**	0.80	0.69

As a result, we can assume that recommendation performance can be increased by adding a highly correlated domain to an already participating domain, but there is likely to be an adverse effect if a new domain that does not correlate well is introduced.

5 CONCLUSIONS

We proposed a simple and robust recommendation method that works with multiple domains. We experimented what kinds of domain combinations increase recommendation performance. As a result, we found that if we add a domain that is highly correlated (e.g., based on the top N canonical correlations calculated using CCA) with an already added domain, the recommendation performance increases.

REFERENCES

[1] Ruka Funaki and Hideki Nakayama. 2015. Image-Mediated Learning for Zero-Shot Cross-Lingual Document Retrieval. In *Proceedings of EMNLP 2015*. 585–590. http://aclweb.org/anthology/D/D15/D15-1070.pdf
[2] Jon R Kettenring. 1971. Canonical analysis of several sets of variables. *Biometrika* 58, 3 (1971), 433–451.
[3] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of NIPS 2013*. 3111–3119.
[4] Nikhil Rasiwasia, Jose Costa Pereira, Emanuele Coviello, Gabriel Doyle, Gert RG Lanckriet, Roger Levy, and Nuno Vasconcelos. 2010. A new approach to cross-modal multimedia retrieval. In *Proceedings of ACMMM 2010*. ACM, 251–260.
[5] Shaghayegh Sahebi and Peter Brusilovsky. 2015. It Takes Two to Tango: An Exploration of Domain Pairs for Cross-Domain Collaborative Filtering. In *Proceedings of RecSys 2015*. 131–138. https://doi.org/10.1145/2792838.2800188