Simple Regularization for Aligning Embedding Spaces for Cross-brand Recommendation

Ioannis Partalas[1]

[1] *Expedia Group, Geneva, Switzerland*

Abstract

In online platforms it is often the case to have multiple brands under the same group which may target different customer profiles, or have different domains. For example, in the hospitality domain, on-line travel platforms may have multiple brands which have either different traveler profiles or are more relevant in a local context.

In this context, learning embeddings for hotels that can be leveraged in recommendation tasks in multiple brands requires to have a common embedding that can be induced using alignment approaches. At the same time, one needs to ensure that this common embedding space does not degrade the performance in any of the brands.

In this work, we build upon the hotel2vec model and propose a simple regularization approach for aligning hotel embeddings of different brands via domain adaptation. We also explore alignment methods previously used in cross-lingual embeddings to align spaces of different languages. We present results on the task of next-hotel prediction using click sessions from two brands. The results show that the proposed approach can align the two embedding spaces while achieving good performance in both brands. Additionally, with respect to single-brand training we observed that the proposed approach can significantly reduce training time and improve the predictive performance.

Keywords

Cross-brand recommendations, Embeddings, Hotel embeddings

1. Introduction

Online platforms often have multiple brands, for the same line of business, under the same group which may target different customer profiles. As an example, in the hospitality and retail domains the on-line platforms may have different brands, that can either have different profiles of customers or be more relevant locally. A main task in retail platforms, is to recommend products to customers which requires to learn an embedding space that captures their salient attributes. Hence, enable similarity comparisons that can be leveraged from recommendation systems.

In the recent years approaches that learn product embeddings from the interactions of the customers with the on-line platform have been proposed [1], [2], [3] as well as approaches tailored to the hospitality domain [4],[5]. These approaches leverage the seminal word2vec model [6] in order to learn the embedding space by treating the clicked items as tokens in a sentence. While it is common to learn such embeddings in a single domain/brand, in the context of electronic commerce we would like to be able to leverage such embeddings across different domains/brands. As mentioned previously, one can learn hotel embeddings on a specific brand and leverage them in another brand in order to bootstrap or improve the hotel embeddings in the latter case. Subsequently these hotel representations can be used in tasks like personalized recommendations.

To do so, one would need to align the embedding spaces of the different brands and use this aligned space to capture the intent of the users while searching on the on-line platform. In a very recent work, Bianchi et al. [7] study the alignment of product embeddings to enable zero-shot learning in a cross-shop scenario. Their setting is more general as in our case the multi-brands belong to the same domain, and usually we dispose of a partial overlap of the inventories across brand domains.

Workshop on Recommenders in Tourism (RecTour 2025), September 22, 2025, co-located with the 19th ACM Conference on Recommender Systems, Prague, Czech Republic.

ipartalas@expediagroup.com (I. Partalas)
https://ioannispartalas.github.io/ (I. Partalas)

In this work we propose to align embeddings from different brands using a simple regularization approach from domain adaptation. We build upon the hotel2vec model [4] for learning hotel embeddings and extend it to accommodate alignment of embedding spaces. Our approach, can also be thought of as a transfer learning one as we are able to bootstrap the learning procedure of hotel2vec in a target brand using the hotel2vec embeddings from a source brand.

The main contributions of this paper are the following: i) we propose a simple yet effective domain adaptation approach that adds a regularization term in the loss function of hotel2vec, ii) we present empirical results on the task of next-hotel prediction for two brands and iii) we also implement an alignment method borrowed from the task of alignment of cross-lingual embeddings. We show empirically, that to perform such a transpose to another domain one should be careful of the particularities of it.

2. Related Work

In recommendation systems learning embeddings for users and items that capture the semantics is a critical part. In the domain of electronic commerce, approaches as prod2vec have been proposed in order to learn representations of products [1], [8], [5] which leverage the skip-gram model [6]. In [3] the authors propose to learn embeddings for YouTube videos by combining multiple features, which are used for candidate generation and ranking. Other approaches have also been proposed that include metadata [2] or try to capture different aspects of the product from different sources (clickstream data, text and images) [9].

Aligning embedding spaces is a topic that has been extensively studied in the NLP domain and specifically in the context of cross-lingual embeddings. A simple and efficient approach to align two embedding spaces of different languages is to learn a linear projection [10]. In [11] the authors employ an iterative approach starting from seed mappings of words (source language to target language) and do the linear projection by imposing an orthogonality constraint in the projection matrix. A survey on cross-lingual word embedding can be found in [12]. In this work we study the effectiveness of such alignment approaches in the context of hotel embeddings.

Domain adaptation is also a topic that has attracted interest in the recommendation systems space [13], [14] where the goal is to transfer knowledge from a source to a target task. In our setting we employ a domain adaptation approach in order to align different embedding spaces and we follow a straightforward regularization approach [15]. Such approaches have been included in a single framework for domain adaptation [16]. Other works propose adversarial approaches for domain adaptation that learn in the same time the alignment and an embedding space that is invariant to the domain [17], [18], [19]. More recently, graph and Transformer based methods have been proposed to address cross-domain recommendations [20, 21, 22] as well as LLM-based models [23]. In [24] the authors propose to use a proximal operator to learn shared latent factors through a Matrix Factorization approach.

The approach most similar to our work is that presented in [7] where the authors propose to align product embeddings to enable cross-shop recommendations. Specifically, they propose different approaches either based on content (images and text) or using the clicked products in the different shops as supervised signals in methods for alignment of cross-lingual embeddings. Our use case has a simpler setting as we dispose a partial overlap among the products of the different embeddings spaces we wish to align.

3. Alignment through Regularized Domain Adaptation

In order to learn hotel representations, we follow the hotel2vec model [4], which implements a neural model and is trained with the skip-gram model and negative sampling. The model learns different embeddings for click (V_c), hotel properties (V_a) and geographical information (V_g) which are fused to learn an enriched representation. Specifically, the embeddings are calculated as follows: $V_c = f(I_c; W_c), V_a = f(I_a; W_a), V_g = f(I_g; W_g)$ where $f(x; W) = \text{ReLU}(\frac{xW}{\|xW\|_2})$ and I_c, I_a and I_g refer to the input features for the click, amenity and geographical embedding. The final hotel2vec embedding is calculated as a projection of the concatenated embeddings: $V_h = \mathrm{ReLU}([V_c, V_a, V_g]W_e)$.

Let $\mathbf{V_{h_i}}$ be the representation of the hotel h_i as calculated by the above equation where $h_i \in H$. Then hotel2vec model minimizes the following loss function:

$$J(\theta) = -\sum_{(h_t,h_c)\in D^+} \log\sigma\left(V_{h_t}V_{h_c}\right) + \sum_{h_i\in N_c} \log\sigma\left(-V_{h_t}V_{h_i}\right) \tag{1}$$

where D^+ are the skip-gram pairs of clicked hotels in the session that are generated using a fixed length window. N_c are the negative samples that are sampled from the same market of clicked hotels in the session, as a traveler searches for a specific destination. Finally, $\sigma()$ is the sigmoid function.

In this work, we propose to extend the hotel2vec model in order to accommodate embedding spaces alignment in the case of multi-brand representations.

We do so by employing a domain adaptation scheme. Denote $\mathscr{D}^{\mathrm{S}}$ the source domain where we already learned a hotel2vec model by minimizing the loss function in Equation 1, $J_s(\theta)$. Then in the target domain $\mathscr{D}^{\mathrm{T}}$ we learn hotel2vec representations by minimizing the following regularized function:

$$J_t(\theta) = J(\theta) + \lambda\|V_{h_t}^{\mathrm{T}} - V_{h_t}^{\mathrm{S}}\|_2^2$$

where $V_{h_t}^{\mathrm{S}}$ refers to the corresponding embedding of hotel h_t in the source domain. Note that the embeddings of the source domain are fixed and not re-trained. Also, λ is a parameter that controls how much knowledge we would like to transfer from the source domain. As in our case we want to align the embedding spaces, we would like to constraint the model to be as close as possible in the source domain, hence set λ equal to 1.0. In the experimental section we experiment with this parameter to understand the impact in the downstream task. Note that we define the strength of regularization globally but in a more fine version it could be defined per hotel. We leave this for future work.

The regularization framework is evaluated under the assumption of partially overlapping property inventories across brand domains, where approximate correspondences between entities are inferred for the purposes of this paper.

4. Experimental Setting

We evaluate the proposed approach in the next item prediction task where we want to predict the next hotel clicked in the session based on the previous hotel clicked. We collect click sessions over one year of searches for two brands, namely **Brand A** and **Brand B**. Each dataset has millions of user interaction sessions and hundreds of thousands of unique properties distributed across distinct brand domains. We randomly split the sessions into training, validation, and test with a ration of 8:1:1.

We use a system with 64GB RAM, 8 CPU cores, and a Tesla V100 GPU. We use Python 3 as the programming language and the Tensorflow [25] library for the neural network architecture and gradient calculations. For hotel2vec we follow the experimentation methodology that authors employed in [4] for tuning the hyperparameters of the model. The model in both brands is trained with $L2$-regularization.

We compare the regularization approach with the linear projection alignment approach presented in [10]. In that case given the embeddings that we export from the trained models in brand A and brand B domains, denoted $\mathscr{V}^{\mathrm{S}}$ and $\mathscr{V}^{\mathrm{T}}$ respectively, we solve the following optimization problem:

$$\min_W \|\mathscr{V}^{\mathrm{S}}W - \mathscr{V}^{\mathrm{T}}\|_2^2$$

The alignment is learned only on the common hotel embeddings for the two sets of vectors in order to avoid injecting too much noise.

We compare the different approaches in terms of Hits@k and Mean Reciprocal Rank at k (MRR@k). Both metrics are calculated over the ranking induced by cosine similarity of the embedding vectors. Note that when we have a cross-brand setting this corresponds to a zero-shot learning framework.

Table 1
% uplift of Hits@k and MRR@k (abbreviated H@k and M@K respectively) for the different approaches for embeddings alignment.

	Brand A				Brand B			
	H@100	H@10	M@10	M@100	H@100	H@10	M@10	M@100
LP	-29.17	-24.48	-23.16	-25.05	-37.74	-33.91	-21.87	-24.61
A→B								
$\lambda = 1.0$	0.14	0.94	0.76	0.77	0.00	-0.54	-1.00	-0.86
$\lambda = 0.5$	-0.06	0.69	0.51	0.39	-0.48	-1.58	-1.63	-1.41
$\lambda = 0.1$	2.58	3.20	3.05	3.08	-1.53	-3.09	-2.18	-2.04
$\lambda = 0.01$	**7.02**	**6.97**	**7.63**	**7.71**	-5.44	-9.90	-6.99	-6.66
B→A								
$\lambda = 1.0$	**11.66**	36.95	-0.49	-0.37	0.54	0.65	-0.39	-0.25
$\lambda = 0.5$	11.53	37.01	-0.24	-0.37	0.80	1.18	0.00	0.17
$\lambda = 0.1$	11.60	**37.91**	**0.73**	**0.37**	2.74	5.03	2.86	2.72
$\lambda = 0.01$	8.90	35.09	-1.70	-2.20	**7.79**	**13.99**	**8.48**	**8.15**

5. Results

Table 1 presents the percentage uplift in terms of Hits@k and MRR@k for $k \in \{10, 100\}$ to predict next-clicked hotel in the test sets for both brands. The uplift is measured against the corresponding single-domain trained models with no further alignment. The first line presents the corresponding metrics for the linear projection alignment approach (LP). For the proposed approach that employs domain adaptation ($h2vec_{DA}$) we present the results when training the model with regularization for both pairs, that is when brand B is used in the regularization while training brand A (B→A) and vice versa.

Concerning the alignment approaches presented in the table, we notice that the LP method has a big drop in performance in both brands. As we mentioned earlier, even though we dispose of a partial overlap of the inventories between brands, the method fails to keep the similarities in the projected space. This shows that extra care should be taken when aligning embedding spaces, for example, requiring the projection matrix to be orthogonal.

Regarding the proposed approach we can observe that is able to achieve good performance in both brands. In fact, it achieves better performance than the in-domain model in brand A with an increase of more than 10% for Hits@100 and similar performance in brand B with stronger regularization strength. In the case of a weaker regularization strength ($\lambda = 0.01$), we can observe that the metrics improve. Smaller values of the parameter can balance between a full transfer of knowledge and in-brand learning. For the improvements, we hypothesize that allowing for weaker regularization we are able to allow more in-domain knowledge to flow and learn a better similarity space. Interestingly, when we use brand B as the source in the regularization we can improve the Hits@100 in brand A by 10%.

In terms of training speed, the models that use the proposed regularization scheme can converge much faster by reducing training time by 50%, which is substantial when dealing with large-scale datasets.

6. Conclusions

We presented in this work a simple yet effective regularization approach for aligning embedding spaces in a multi-brand scenario. For example, in the hospitality/retail domains the on-line platforms have multiple brands that operate in the same domain. The idea is to add a regularizer in the objective function of the model that learns the embeddings in order to force them to be as close as possible to the embeddings of the source brand, hence performing domain adaptation. This kind of approaches has also been explored in the past in Natural Language Processing tasks [15].

We evaluated the proposed approach in the next-hotel prediction task for two brands. We measured performance in terms of hits@k and MRR@k metrics. We also, compared with linear projection alignment borrowed by the cross-lingual approaches for aligning embeddings of different languages [10].

The results showed that the proposed approach can align the spaces of the multiple brands achieving good performance in both brands. Indeed the results showed that we can outperfom the single-domain models. We also observed that an approach like the linear projection without taking into account some particularities of the domain leads to worse performance.

7. Future Work

For future work we would like to add a more adaptive regularization parameter that can be defined per hotel rather than being global. In that way we may wish to transfer knowledge when we are certain that a pair of hotels in the source and target brands should have the same embedding. We would also like to explore the use of multiple source brands in order to align in the same time the embedding spaces. Multi-task approaches can be leveraged to align the different embedding spaces [16]. Also, other alignment approaches could be explored [26]. Finally, we would like to explore adversarial cross-domain adaptation for aligning the embedding spaces [18]. In this case, we want to leverage the similar features across the brands while also learning specific embeddings for each brand.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] M. Grbovic, V. Radosavljevic, N. Djuric, N. Bhamidipati, J. Savla, V. Bhagwan, D. Sharp, E-commerce in your inbox: Product recommendations at scale, in: Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining, 2015, pp. 1809–1818.

[2] F. Vasile, E. Smirnova, A. Conneau, Meta-prod2vec: Product embeddings using side-information for recommendation, in: Proceedings of the 10th ACM Conference on Recommender Systems, ACM, 2016, pp. 225–232.

[3] P. Covington, J. Adams, E. Sargin, Deep neural networks for youtube recommendations, in: Proceedings of the 10th ACM Conference on Recommender Systems, RecSys '16, ACM, New York, NY, USA, 2016, pp. 191–198. URL: http://doi.acm.org/10.1145/2959100.2959190. doi:10.1145/2959100.2959190.

[4] I. Partalas, A. Morvan, A. Sadeghian, S. Minaee, X. Li, B. Cowan, D. Z. Wang, Hotel2vec: Learning hotel embeddings from user click sessions with side information., in: Proceedings of the Workshop on Recommenders in Tourism co-located with the 15th ACM Conference on Recommender Systems (RecSys 2021), 2021, pp. 69–84.

[5] C. Grbovic, Real-time personalization using embeddings for search ranking at airbnb, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD, 2018, pp. 311–320.

[6] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, in: Advances in neural information processing systems, 2013, pp. 3111–3119.

[7] F. Bianchi, J. Tagliabue, B. Yu, L. Bigon, C. Greco, Fantastic embeddings and how to align them: Zero-shot inference in a multi-shop scenario, in: Proceedings of the SIGIR 2020 eCom workshop, July 2020, Virtual Event, 2020.

[8] O. Barkan, N. Koenigstein, Item2vec: neural item embedding for collaborative filtering, in: 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP), IEEE, 2016, pp. 1–6.
[9] L. Singh, S. Singh, S. Arora, S. Borar, One embedding to do them all, arXiv preprint arXiv:1906.12120 (2019).
[10] T. Mikolov, Q. V. Le, I. Sutskever, Exploiting similarities among languages for machine translation, CoRR abs/1309.4168 (2013). URL: http://arxiv.org/abs/1309.4168. arXiv:1309.4168.
[11] M. Artetxe, G. Labaka, E. Agirre, Learning bilingual word embeddings with (almost) no bilingual data, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Vancouver, Canada, 2017, pp. 451–462. URL: https://www.aclweb.org/anthology/P17-1042. doi:10.18653/v1/P17-1042.
[12] S. Ruder, A survey of cross-lingual embedding models, CoRR abs/1706.04902 (2017). URL: http://arxiv.org/abs/1706.04902. arXiv:1706.04902.
[13] F. Yuan, L. Yao, B. Benatallah, Darec: Deep domain adaptation for cross-domain recommendation via transferring rating patterns, in: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19, International Joint Conferences on Artificial Intelligence Organization, 2019, pp. 4227–4233. URL: https://doi.org/10.24963/ijcai.2019/587. doi:10.24963/ijcai.2019/587.
[14] H. Kanagawa, H. Kobayashi, N. Shimizu, Y. Tagami, T. Suzuki, Cross-domain recommendation via deep domain adaptation, in: L. Azzopardi, B. Stein, N. Fuhr, P. Mayr, C. Hauff, D. Hiemstra (Eds.), Advances in Information Retrieval, Springer International Publishing, Cham, 2019, pp. 20–29.
[15] H. Daumé III, Frustratingly easy domain adaptation, in: Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Association for Computational Linguistics, Prague, Czech Republic, 2007, pp. 256–263. URL: https://www.aclweb.org/anthology/P07-1033.
[16] W. Lu, H. L. Chieu, J. Löfgren, A general regularization framework for domain adaptation, in: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Austin, Texas, 2016, pp. 950–954. URL: https://www.aclweb.org/anthology/D16-1095. doi:10.18653/v1/D16-1095.
[17] S. Motiian, Q. Jones, S. Iranmanesh, G. Doretto, Few-shot adversarial domain adaptation, in: I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), Advances in Neural Information Processing Systems, volume 30, Curran Associates, Inc., 2017.
[18] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, Domain-adversarial training of neural networks, J. Mach. Learn. Res. 17 (2016) 2096–2030.
[19] X. Chen, Y. Sun, B. Athiwaratkun, C. Cardie, K. Weinberger, Adversarial deep averaging networks for cross-lingual sentiment classification, Transactions of the Association for Computational Linguistics 6 (2018) 557–570.
[20] Z. Xu, W. Pan, Z. Ming, A multi-view graph contrastive learning framework for cross-domain sequential recommendation, in: Proceedings of the 17th ACM Conference on Recommender Systems, RecSys '23, Association for Computing Machinery, 2023, p. 491–501.
[21] G. Lin, C. Gao, Y. Zheng, J. Chang, Y. Niu, Y. Song, K. Gai, Z. Li, D. Jin, Y. Li, M. Wang, Mixed attention network for cross-domain sequential recommendation, in: Proceedings of WSDM conference, WSDM '24, New York, NY, USA, 2024, p. 405–413.
[22] H. Ma, R. Xie, L. Meng, X. Chen, X. Zhang, L. Lin, J. Zhou, Triple sequence learning for cross-domain recommendation, ACM Trans. Inf. Syst. 42 (2024).
[23] A. Petruzzelli, C. Musto, L. Laraspata, I. Rinaldi, M. de Gemmis, P. Lops, G. Semeraro, Instructing and prompting large language models for explainable cross-domain recommendations, in: Proceedings of the 18th ACM Conference on Recommender Systems, RecSys '24, Association for Computing Machinery, 2024, p. 298–308.
[24] A. Samra, E. Frolov, A. Vasilev, A. Grigorevskiy, A. Vakhrushev, Cross-domain latent factors sharing via implicit matrix factorization, in: Proceedings of the 18th ACM Conference on Recommender Systems, RecSys '24, Association for Computing Machinery, New York, NY, USA, 2024, p. 309–317. doi:10.1145/3640457.3688143.

[25] M. Abadi, A. Agarwal, P. Barham, et. al., TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL: https://www.tensorflow.org/, software available from tensorflow.org.
[26] F. Bianchi, V. D. Carlo, P. Nicoli, M. Palmonari, Compass-aligned distributional embeddings for studying semantic differences across corpora, CoRR abs/2004.06519 (2020). URL: https://arxiv.org/abs/2004.06519. arXiv:2004.06519.