Enhancing Collaborative Variational Autoencoder with Tag and Citation Information for Scientific Article Recommendation

Ho, Xanh
University of Science VNUHCM

Aizawa, Akiko
National Institute of Informatics

KYUSHU UNIVERSITY

Enhancing Collaborative Variational Autoencoder with Tag and Citation Information for Scientific Article Recommendation

Xanh Ho[1] and Akiko Aizawa[2]

[1] University of Science VNUHCM, Vietnam
[2] National Institute of Informatics, Japan

Abstract. Hybrid methods such as collaborative deep learning (CDL) and collaborative variational autoencoder (CVAE) have become state-of-the-art methods in recommender systems for scientific articles. However, they typically use only information from titles and abstracts of articles, and ignore potentially useful information in the tags and citations. Therefore, they may miss articles that contain vastly different content from other articles, although those articles present the same topic. We addressed this problem by developing the CiT-CVAE model that considers tag and citation information when providing recommendations. Our experimental results indicate that the proposed model achieves consistent improvement compared with CDL and CVAE.

Keywords: Recommender systems, Scientific article recommendation, Hybrid methods, Citation information, Tag information

1 Introduction

Scientific articles are being published in increasing numbers annually; however, researchers generally lack the time required to process all of them or be unaware of the ones attracting potential interest owing to different terminologies being used. A recommender system that can effectively match the interests of a user with the attributes of the articles would solve this problem. Recommender systems can be divided into three groups: content based, collaborative filtering (CF) based, and hybrid methods [3]. Content-based methods [7] utilize the description of an item, e.g., scientific article and user preferences to generate a recommendation. Meanwhile, CF-based methods [13] use solely the history of user-item interactions to provide a suggestion. Although CF-based methods yield better results than content-based methods, they perform ineffectively in practice owing to the sparse user-item interaction matrix. Moreover, items with no previous record are never recommended. Hybrid methods [1, 19] have been developed to overcome such limitations by combining auxiliary information (content information or user profile) and collaborative information. Depending on the interaction between the auxiliary and collaborative information, hybrid methods can be categorized as loosely and tightly coupled [23].

This paper focuses on tightly coupled methods applied to recommender systems for scientific articles. The first exploration of this type was collaborative topic regression (CTR) [19]. CTR is a probabilistic model that combines traditional collaborative filtering, e.g., probabilistic matrix factorization (PMF) [14] and probabilistic topic modeling, e.g., latent Dirichlet allocation (LDA) [2]. Although CTR produces promising results and provides an interpretable latent structure for users and items, the learned latent representation does not suffice when the auxiliary information is extremely sparse. Wang et al. [23] addressed this issue by collaborative deep learning (CDL) using stacked denoising autoencoder (SDAE) [18] with PMF. The model learns (1) embedding vectors from the auxiliary information, and (2) the latent vectors of the users and items from the collaborative information. CDL with denoising autoencoders (DAEs) contains no Bayesian nature and the denoising scheme of the DAEs is not from a probabilistic perspective [9]. Therefore, the incorporation of SDAE and PMF is not sufficiently effective. Li and She [9] thus proposed the collaborative variational autoencoder (CVAE) that uses a variational autoencoder (VAE) [6] instead of the SDAE. Throughout the experiments, the authors demonstrated that the CVAE significantly outperformed CTR and CDL.

All of the abovementioned models are limited by the information about the title and abstract of one article, and skip the citation graph and tag set that are also important information in the scientific domain. In practice, when researchers read an article j, they can obtain relevant articles by following a list of citations of the article. That is, if user i is interested in article j, and j cites another article k, we can infer that the probability that user i is interested in article k is high. Therefore, the use of a citation graph is important in recommending scientific articles. Similarly, tag information can be used to infer the transference of a user's interest from one article to another.

To explore the effect of using tags and citation graphs in recommender systems, we introduce the CiT-CVAE model, in which we incorporate both types of information. We first use the tag information as the text content of an article, and subsequently embed the tag into the content matrix. Hence, we enrich the context matrix as a large number of overlapping words exist between the list of tags and list of words extracted from the titles and abstracts. The content matrix is subsequently fed into the conventional CVAE model to learn a rating matrix for each user. Next, a user-cite matrix is produced by integrating the citation graph with the implicit feedback. We finally revise the CVAE rating matrix by combining it with the user-cite matrix. We applied our model to the CiteULike dataset [19], of which the tag information and the citation graph were collected from CiteULike[3] and Google Scholar[4], respectively [20]. The experimental results indicate that our CiT-CVAE model achieves better performance than the CDL and CVAE models.

---

[3] http://www.citeulike.org/

[4] https://scholar.google.com/

2   Related Work

2.1   Hybrid methods for Recommender Systems

In addition to CTR, CDL, CVAE models (Section 1), other autoencoder-based models have been applied to recommender systems [8,21,22,24]. Wang et al. [21] proposed a relational stacked denoising autoencoder (RSDAE) model that combines a stacked denoising autoencoder and relational learning in a principle manner for tag recommendation. Another variant of CDL is deep learning collaborative filtering [8], which tightly couples matrix factorization with a deep learning algorithm, namely marginalized denoising autoencoders. Unlike these models, [22] proposed collaborative recurrent autoencoder (CRAE) that combines the order-aware generation of sequences (content information) and the rating information. As the target of this study is to leverage the importance of citation and tag information in recommender systems for scientific articles, we selected CDL and CVAE—state-of-the-art models for the task, to compare with our model.

2.2   Citation and Tag for Recommender Systems

Citation and tag are important information that have been used widely in recommender systems [4, 5, 10–12, 15–17, 25, 26]. Sugiyama and Kan [15] exploited the explicit citation network of publications as a source of knowledge to improve recommendation accuracy. [16,17] proposed a new method of using the explicit citation network. They used collaborative filtering to identify potential citation papers that facilitate in representing target papers. In addition, they investigated which sections of the papers can be utilized to represent papers effectively. In [10, 12], the authors leveraged heterogeneous bibliographic networks to provide a recommendation. Meanwhile, previous works [5, 11, 25, 26] incorporated tag information in their recommender systems. Specifically, [26] demonstrated that tag and time information are important when predicting a user's preferences by building a resource-recommendation model that combined the two types of information. In [5], the authors proposed a hybrid item recommendation system that combined tag expansion and implicit trust relationship simultaneously. They used tag expansion to generate relevant items and extracted an implicit trust relationship on the tagging information to estimate user similarity.

3   Notation and Problem Formulation

Given a list of M users and a list of N articles, the $M \times N$ rating matrix $\mathcal{R} = (r_{ij}) \in \{0,1\}^{M \times N}$ is defined as follows:

$$r_{ij} = \begin{cases} 1, \text{ if user } i \text{ has article } j \text{ in his or her personal library} \\ 0, \text{otherwise.} \end{cases} \tag{1}$$

The case $r_{ij} = 0$ can be interpreted in two aspects: (1) user i is not interested in article j, or (2) user i does not know about article j. The rating matrix $\mathcal{R}$

is divided into the training set and test set. The details of how to split the rating matrix $\mathcal{R}$ is described in Section 5. The entire collection of N articles is represented by a content matrix $\mathcal{X} = (x_{ij}) \in \mathbb{R}^{N \times S}$, where S is the size of the vocabulary. Let $\mathbb{X} = \{x_1, \ldots, x_S\}$ be a set of all vocabularies corresponding to S columns in matrix $\mathcal{X}$. A word x_j is represented by column $\mathcal{X}_j$ for $j = 1, \ldots, S$. The content matrix $\mathcal{X}$ is computed based on the title and the abstract of an article [9,19,23]. The recommendation task is defined as given the content matrix $\mathcal{X}$, and the training part of the rating matrix $\mathcal{R}$, predicts the missing ratings (i.e., the testing part) in $\mathcal{R}$.

Assuming that there are Q tags assigned to all articles, we denote the tag matrix as $\mathcal{T} = (t_{ij}) \in \mathbb{R}^{N \times Q}$. Let $\mathbb{T} = \{t_1, \ldots, t_Q\}$ be a set of all tags corresponding to Q columns in matrix $\mathcal{T}$. A tag t_j is represented by column $\mathcal{T}_j$ for $j = 1, \ldots, Q$. Similarly, the citation matrix is a binary matrix $\mathcal{A} = (a_{ij}) \in \{0, 1\}^{N \times N}$, where $a_{ij} = 1$ implies that article i cites article j, or article i is cited by article j. In this work, both the tag matrix $\mathcal{T}$ and citation matrix $\mathcal{A}$ were used as the input of our model. Similar to previous models, our model predicts the missing ratings in $\mathcal{R}$. More details about our model are presented in Section 4.

4 Model

In this section, we first introduce CVAE model [9], the state-of-the-art model in recommending scientific articles. We subsequently describe our proposed model CiT-CVAE in which we incorporate tags into the content information and combine the citation with the CVAE model.

4.1 Collaborative Variational Autoencoder

CVAE is a Bayesian probabilistic model that unifies collaborative filtering and content information through a stochastic deep learning model and a graphical model [9]. While PMF infers latent vectors of users and items from rating information, VAE learns latent representations from content data.

With VAE and PMF as key components, the generative process of the CVAE is defined as follows:

1. For each user i, a user latent variable $u_i \in \mathbb{R}^K$ is drawn from $u_i \sim \mathcal{N}(0, \lambda_u^{-1} I_K)$, where K is the dimension of the shared latent low space between users and items.
2. For each item j, a collaborative latent variable $v_j^\dagger$ is drawn from $v_j^\dagger \sim \mathcal{N}(0, \lambda_v^{-1} I_K)$.
3. The content of item x_j is generated from its latent variable z_j through a generation neural network parameterized by θ $x_j \sim p_\theta(x_j|z_j)$.
4. The latent variable of the content z_j is a unit normal distribution $z_j \sim \mathcal{N}(0, I_K)$.
5. The item latent variable v_j is composed by combining the collaborative latent variable with the content latent variable $v_j = v_j^\dagger + z_j$.

By defining maximum a posteriori estimates and maximizing the corresponding function, the CVAE can learn u_i, v_j, and the parameters of the model using a stochastic gradient descent algorithm. Once the training process is completed, the user latent matrix $\mathcal{U} \in \mathbb{R}^{M \times K}$ and item latent matrix $\mathcal{V} \in \mathbb{R}^{N \times K}$ are obtained.

4.2 Proposed Model — CiT-CVAE

CiT-CVAE is a model that incorporates tag into the content information and combines the citation graph with the CVAE, as shown in Figure 1.
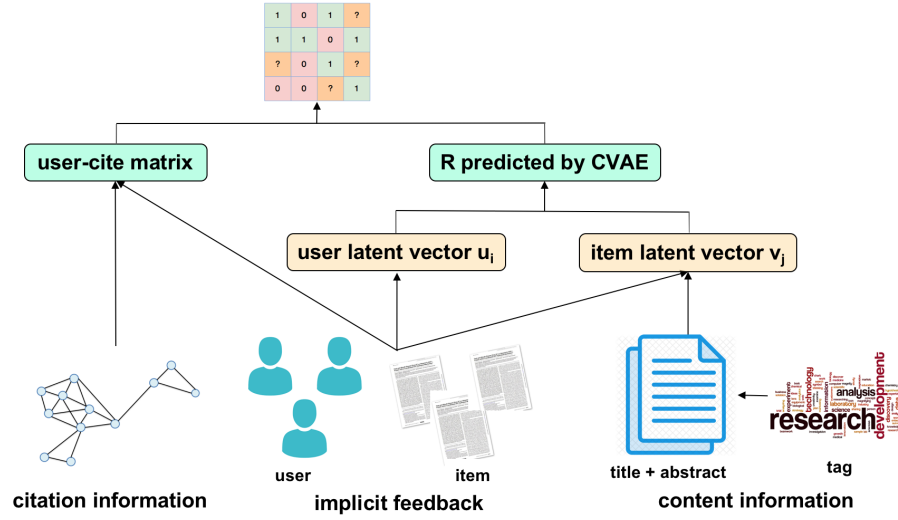

Fig. 1. CiT-CVAE model

Let $\mathbb{C} = \mathbb{X} \cup \mathbb{T}$, $\mathbb{K} = \mathbb{T} \setminus \mathbb{X}$, and $k = |\mathbb{K}|$. Assume that $\delta_{ij} = 1$ if $x_i = t_j$ and $\delta_{ij} = 0$; otherwise, for any $x_i \in \mathbb{X}$ and $t_j \in \mathbb{T}$. Our model contains three steps. First, a new content matrix $\mathcal{X}_{\text{new}}$ is generated from the $N \times S$ content matrix $\mathcal{X}$ (using only the title and abstract) and the $N \times Q$ tag matrix $\mathcal{T}$ (using tag) in Section 3 as follows:

$$\mathcal{X}_{\text{new}} = \left[\mathcal{X}_1 + \sum_{j=1}^{Q} \delta_{1j}\mathcal{T}_j \ldots \mathcal{X}_S + \sum_{j=1}^{Q} \delta_{Sj}\mathcal{T}_j \; \mathcal{T}_{j_1} \ldots \mathcal{T}_{j_k} \right] \qquad (2)$$

where $\{t_{j_1}, \ldots, t_{j_k}\} = \mathbb{K}$. It is noteworthy that the size of $\mathcal{X}_{\text{new}}$ is $N \times |\mathbb{C}|$.

Next, a user-cite matrix is produced by integrating the citation graph with the implicit feedback (rating matrix) as follows. Let $\mathcal{R}_{\text{train}}$ be a part of the rating matrix $\mathcal{R}$ used for training. Denote a row vector of $\mathcal{R}_{\text{train}}$ as $r_{i,*}(\in \mathbb{R}^{1 \times N})$. $r_{i,*}$ represents the interest of user i. For example, if user i is interested in item $id = 41$, subsequently, $r_{i,41} = 1$. Denote a column of the citation matrix $\mathcal{A}$ by

$a_{*,j}(\in \mathbb{R}^{N \times 1})$. This vector contains information about a list of cited articles of item j. We define the user-cite matrix $\hat{\mathcal{R}}$ as follows:

$$\hat{\mathcal{R}}_{ij} = \begin{cases} 1, \Leftrightarrow \sum_{k=1}^{N} r_{ik}a_{kj} \geq 1 \\ 0, \Leftrightarrow \sum_{k=1}^{N} r_{ik}a_{kj} = 0. \end{cases} \tag{3}$$

With this definition, the user-cite matrix can capture both citation information and interest of users. Let $\mathcal{R}_{\text{CVAE}}$ be the output of the CVAE model, $\mathcal{R}_{\text{CVAE}} = \mathcal{U} \times \mathcal{V}^{T}$ (correspond "*R predicted by CVAE*" in Figure 1). Finally, we calculate the final predicted rating matrix by a linear combination of the user-cite matrix $\hat{\mathcal{R}}$ and $\mathcal{R}_{\text{CVAE}}$.

5 Experiments

We evaluated our proposed model quantitatively and qualitatively to demonstrate the importance of the citation graph and tag information, and compared the results with those of two models, CDL [23] and CVAE [9].

5.1 Experimental Setup

We used CiteULike-a—a publicly accessible dataset[5] in our experiments. The statistics for this dataset are summarized in Table 1.

Table 1. Statistics of CiteULike-a.

#Users	#Articles	#Interactions	#Tags	#Citations
5,551	16,980	204,986	46,391	44,709

We compared the performance of the following three models:

- **CDL:** Collaborative deep learning [23] combines probabilistic stacked denoising autoencoder and probabilistic matrix factorization to form a unified probabilistic graphical model.
- **CVAE:** Collaborative variational autoencoder [9] is a Bayesian probabilistic model that unifies the collaborative filtering and content information through a stochastic deep learning model and graphical model.
- **CiT-CVAE:** This is our proposed model as described in Section 4.

We used the same parameters for CDL and CVAE as described in their corresponding papers.

---

[5] http://www.citeulike.org/faq/data.adp

5.2 Evaluation Scheme

We define P_{all} as the total number of articles that a user rated in the rating matrix. For our experiments, we randomly selected P $(P < P_{all})$ articles from the rating information to form the training set, and used the remaining articles $(P_{all} - P)$ as the testing set. Three different values of P, $P = 1$, $P = 5$, and $P = 9$ were used. Unlike the previous work [9, 19, 23], we did not use $P = 10$ because the number of users interested in more than 10 articles is small for the dataset. The case of $P = 5$ was considered as neither sparse nor dense. For each value of P, we performed five runs with the training P articles and reported the average performance.

Similar to the previous work [9, 19, 23], we used recall at the top-M recommended articles to evaluate the performance of each system:

$$recall@M = \frac{\text{number of articles that a user likes in top M}}{\text{total number of articles that the user likes}}. \qquad (4)$$

The final result was the average recall of all users in the system.

5.3 Quantitative Comparison

Figure 2 shows the results of the three models with the three settings of P, $P = 1$, $P = 5$ and $P = 9$. As shown, CiT-CVAE achieved better performance than the state-of-the-art methods—CVAE and CDL. Specifically, CiT-CVAE outperformed CDL by a margin of 3.2%, 5.0%, and 5.7% (in average) in the three settings of $P = 1$, $P = 5$, and $P = 9$, respectively. Compared with CVAE, the relative improvements over CVAE were 2.5%, 2.9%, and 1.2%. It is noteworthy that when the rating matrix is sparse ($P = 1$ and $P = 5$), all approaches must overcome the cold-start issue by depending heavily on the content of the items, citations, and tags. Hence, CiT-CVAE naturally obtained better results than CVAE because it incorporates more information. In contrast, the weights on the rating and content information become more balanced when $P = 9$; thus, not much change is found in making use of citations and tags. As the rating matrix is often extremely sparse in practice, using other information of items is necessary.

To investigate the important of tag and citation individually, we conducted two experiments that used only the tag, namely T-CVAE, and only the citation, namely Ci-CVAE, on the setting of $P = 5$. The results in Figure 3 indicate that (1) both Ci-CVAE and T-CVAE are better than CVAE, and (2) CiT-CVAE achieves the best performance. As mentioned above, when the rating matrix is sparse ($P = 5$), most of the recommendations are based on the content of the items, citations, and tags. Therefore, CiT-CVAE, Ci-CVAE, and T-CVAE performed slightly better than CVAE. Incorporating the tag information (T-CVAE) produced better results than the citation (Ci-CVAE) because a tag is less sparse than a citation (1/3 of items do not contain citation information, while only 1/5 do not contain tag information). This experiment reveals that tag and citation information are highly useful. Although such information is sparse, CiT-CVAE can still leverage it to improve the recommendation results.
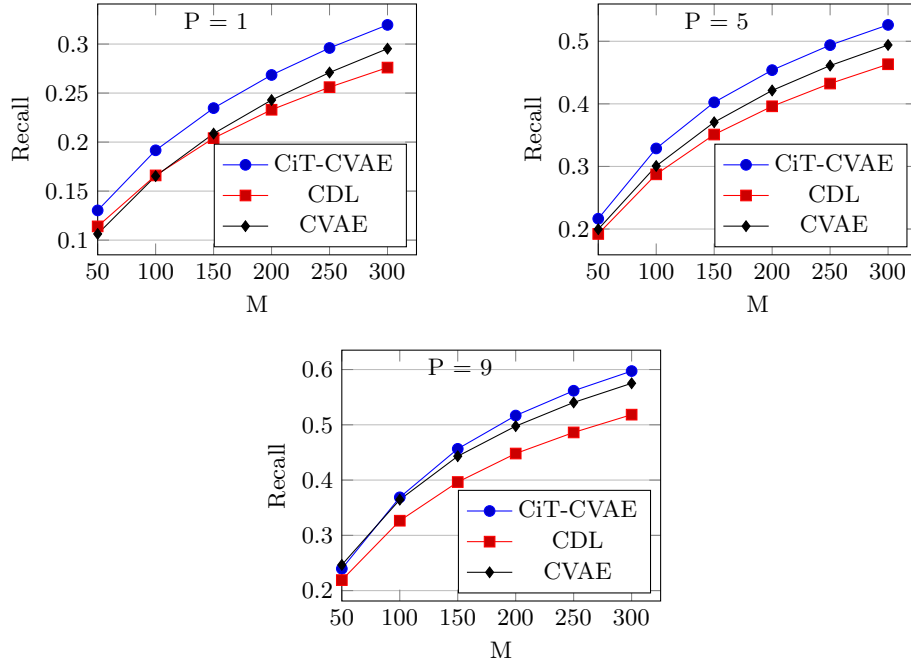
Fig. 2. Performance of CiT-CVAE, CDL, CVAE based on Recall@M in the three settings P = 1, P = 5 and P = 9.

5.4 Qualitative Comparison

Similar to [9, 23], to gain a better insight of CiT-CVAE, we use the results of CVAE and CiT-CVAE in the setting of *P = 1* to recommend articles to users. Table 2 shows the top-10 items recommended to the same user.

Because the training part of the rating matrix is often sparse, many recommendations have to rely solely on the content. However, in practice, when researchers are investigating on a specific topic, they would like to discover more information about the topic through surveys, evaluations, and the existing models and systems. If a recommendation model uses only the content information, it cannot cover sufficient articles that researchers are interested in. Such situation occurs with the CVAE model, i.e., the CVAE model cannot recommend articles related to evaluations or surveys of the topic of interest. With a relatively simple extension, our proposed CiT-CVAE model can alleviate the situation by providing suggestions based on both citation and content information. As shown in Table 2, CiT-CVAE can recommend articles about the survey (Article 1) and evaluation (Articles 3 and 6), which are more informative than those by CVAE.

In summary, tags and citations play the supportive role for item information, and the experimental results indicate that the more integrated the information, the better the model. We can apply a similar approach to other domains. For
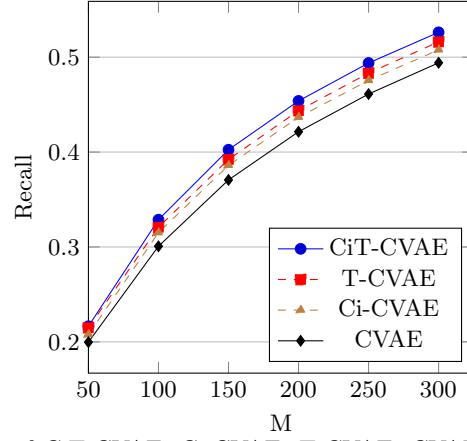
Fig. 3. Performance of CiT-CVAE, Ci-CVAE, T-CVAE, CVAE based on Recall@M when $P = 5$.

Table 2. Qualitative comparison between CiT-CVAE and CVAE.

user I (CiT-CVAE)	in user's lib?
1. Toward the next generation of Recommender Systems: A survey of the state-of-the-art and possible extensions	**yes**
2. Collaborative filtering with privacy	no
3. Evaluating collaborative filtering Recommender Systems	**yes**
4. GroupLens: An open architecture for collaborative filtering of Netnews	**yes**
5. Eigentaste: A constant time collaborative filtering algorithm	**yes**
6. Methods and metrics for cold-start recommendations	**yes**
7. Slope one predictors for online rating-based collaborative filtering	no
8. Learning collaborative information filters	**yes**
9. Recommendation as Classification: Using social and content-based information in recommendation	**yes**
10. Empirical analysis of predictive algorithms for collaborative filtering	**yes**
user I (CVAE)	in user's lib?
1. Amazon.com Recommendations: Item-to-Item collaborative filtering	**yes**
2. Solving the apparent diversity-accuracy dilemma of Recommender Systems	no
3. Collaborative filtering by personality diagnosis: A hybrid memory- and model-based approach	no
4. Taxonomy-driven computation of product recommendations	**yes**
5. Unifying collaborative and content-based filtering	no
6. Collaborative filtering with privacy	no
7. Unified relevance models for rating prediction in collaborative filtering	no
8. Collaborative filtering with temporal dynamics	no
9. Trust in Recommender Systems	no
10. Using mixture models for collaborative filtering	no

instance, with movies, instead of using only plots and titles to represent the content, we can use other information of the movies such as the tag, director, and actor.

6 Conclusion

Unlike previous recommendation models, our proposed CiT-CVAE model for scientific articles considered information from tags and citations. Quantitative and qualitative evaluations demonstrated that CiT-CVAE performed better than both collaborative deep learning and collaborative variational autoencoder. These results also verified that using information from tags and citations improved the performance of recommender systems. Our future work would be to evaluate the effect of using other information such as author, topic, and full text on the performance of recommender systems.

Acknowledgments

References

1. Agarwal, D., Chen, B.C.: Regression-based latent factor models. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 19–28. KDD '09, ACM, New York, NY, USA (2009)
2. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (Mar 2003)
3. Bobadilla, J., Ortega, F., Hernando, A., GutiéRrez, A.: Recommender systems survey. Know.-Based Syst. **46**, 109–132 (Jul 2013)
4. Deng, S., Huang, L., Xu, G., Wu, X., Wu, Z.: On deep learning for trust-aware recommendations in social networks **28**, 1–14 (02 2016)
5. Kim, H., Kim, H.J.: A framework for tag-aware recommender systems. Expert Syst. Appl. **41**(8), 4000–4009 (Jun 2014)
6. Kingma, D.P., Welling, M.: Auto-encoding variational bayes. In Proceedings of the 2nd International Conference on Learning Representations. (2013)
7. Lang, K.: Newsweeder: Learning to filter netnews. In: Proceedings of the 12th International Machine Learning Conference (ML95) (1995)
8. Li, S., Kawale, J., Fu, Y.: Deep collaborative filtering via marginalized denoising auto-encoder. In: Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. pp. 811–820. CIKM '15, ACM, New York, NY, USA (2015)
9. Li, X., She, J.: Collaborative variational autoencoder for recommender systems. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 305–314. KDD '17, ACM, New York, NY, USA (2017)

10. Liu, X., Yu, Y., Guo, C., Sun, Y.: Meta-path-based ranking with pseudo relevance feedback on heterogeneous graph for citation recommendation. In: Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management. pp. 121–130. CIKM '14, ACM, New York, NY, USA (2014)
11. Ma, T., Zhou, J., Tang, M., Tian, Y., Al-Dhelaan, A., Al-Rodhaan, M., Lee, S.: Social network and tag sources based augmenting collaborative recommender system **E98.D**, 902–910 (04 2015)
12. Ren, X., Liu, J., Yu, X., Khandelwal, U., Gu, Q., Wang, L., Han, J.: Cluscite: Effective citation recommendation by information network-based clustering. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 821–830. KDD '14, ACM, New York, NY, USA (2014)
13. Rendle, S., Freudenthaler, C., Gantner, Z., Schmidt-Thieme, L.: Bpr: Bayesian personalized ranking from implicit feedback. In: Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence. pp. 452–461. UAI '09, AUAI Press, Arlington, Virginia, United States (2009)
14. Salakhutdinov, R., Mnih, A.: Probabilistic matrix factorization. In: Proceedings of the 20th International Conference on Neural Information Processing Systems. pp. 1257–1264. NIPS'07, Curran Associates Inc., USA (2007)
15. Sugiyama, K., Kan, M.Y.: Scholarly paper recommendation via user's recent research interests. In: Proceedings of the 10th Annual Joint Conference on Digital Libraries. pp. 29–38. JCDL '10, ACM, New York, NY, USA (2010)
16. Sugiyama, K., Kan, M.Y.: Exploiting potential citation papers in scholarly paper recommendation. In: Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries. pp. 153–162. JCDL '13, ACM, New York, NY, USA (2013)
17. Sugiyama, K., Kan, M.Y.: A comprehensive evaluation of scholarly paper recommendation using potential citation papers. Int. J. Digit. Libr. **16**(2), 91–109 (Jun 2015)
18. Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.A.: Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. J. Mach. Learn. Res. **11**, 3371–3408 (Dec 2010)
19. Wang, C., Blei, D.M.: Collaborative topic modeling for recommending scientific articles. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 448–456. KDD '11, ACM, New York, NY, USA (2011)
20. Wang, H., Chen, B., Li, W.J.: Collaborative topic regression with social regularization for tag recommendation. In: Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence. pp. 2719–2725. IJCAI '13, AAAI Press (2013)
21. Wang, H., Shi, X., Yeung, D.Y.: Relational stacked denoising autoencoder for tag recommendation. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. pp. 3052–3058. AAAI'15, AAAI Press (2015)
22. Wang, H., Shi, X., Yeung, D.Y.: Collaborative recurrent autoencoder: Recommend while learning to fill in the blanks. In: Proceedings of the 30th International Conference on Neural Information Processing Systems. pp. 415–423. NIPS'16, Curran Associates Inc., USA (2016)
23. Wang, H., Wang, N., Yeung, D.Y.: Collaborative deep learning for recommender systems. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 1235–1244. KDD '15, ACM, New York, NY, USA (2015)

24. Wu, Y., DuBois, C., Zheng, A.X., Ester, M.: Collaborative denoising auto-encoders for top-n recommender systems. In: Proceedings of the Ninth ACM International Conference on Web Search and Data Mining. pp. 153–162. WSDM '16, ACM, New York, NY, USA (2016)
25. Zhang, Z.K., Zhou, T., Zhang, Y.C.: Tag-aware recommender systems: A state-of-the-art survey. J. Comput. Sci. Technol. **26**(5), 767–777 (Sep 2011)
26. Zheng, N., Li, Q.: A recommender system based on tag and time information for social tagging systems. Expert Syst. Appl. **38**(4), 4575–4587 (Apr 2011)