Boosting Course Recommendation Explainability: A Knowledge Entity Aware Model Using Deep Learning

Tianyuan YANG[a*], Baofeng REN[a], Boxuan MA[b], Tianjia HE[a], Chenghao GU[a] & Shin'ichi KONOMI[b]
[a]*Graduate School and Faculty of ISEE, Kyushu University, Japan*
[b]*Faculty of Arts and Science, Kyushu University, Japan*
*yang.tianyuan.791@s.kyushu-u.ac.jp

Abstract: Course recommender systems can assist students in identifying suitable or appealing courses by leveraging user interaction data. However, a prevalent issue with existing course recommender systems is their tendency to prioritize accuracy over explainability. To address this limitation, we propose a novel Knowledge Entity-Aware Model for course recommendation called KEAM, which supports explicit user profile generation based on detailed information from a knowledge graph to enhance comprehension of the students. Specifically, we exploit the information within knowledge graphs using neural networks. Then, KEAM captures students' preferences and creates profiles for explainable recommendations. Comprehensive experiments are conducted on two datasets to verify the effectiveness and explainability of KEAM.

Keywords: Course Recommendations, Knowledge Graphs, Explainability

1. Introduction

Massive Open Online Courses (MOOCs) have garnered considerable attention as an alternative educational pattern to university, which remains a challenge for students to select the appropriate ones. Course recommender systems demonstrate their potential to assist students and effectively alleviate the problem of information overload (Ma et al., 2021).

The key factor of generating effective recommendations relies on the accurate creation of the modeling of user preferences or the user's profile. Many works (Jing, 2017; Morsomme et al., 2019; Zhang et al., 2019) have been proposed to construct user profiles that reflect their preferences based on historical data. While these methods excel in modeling user preferences, they do have some limitations. One significant drawback is the limited emphasis on the explainability of the recommendations. Due to the 'black-box' nature of most deep learning models, the representations of user preferences often remain inscrutable. Recent research has emphasized the importance of explainable recommender systems (Alkan et al., 2019; Pu et al., 2012; Ma et al., 2021). One approach to address the issue is to provide a descriptive summary of the user's profiles, allowing users to review and adjust their profiles as needed. Given the importance of course selection, we consider explicitly presenting the user's profile as an opportunity to enhance users' understanding of the recommendation results.

To address the challenges mentioned above, we present a novel deep learning-based recommender system called the Knowledge Entity-Aware Model (KEAM) for explainable course recommendations. KEAM explicitly constructs user profiles leveraging knowledge and interaction graphs and provides personalized recommendations with enhanced explainability. Specifically, we first propose an autoencoder architecture that integrates course features from knowledge graphs into its hidden layers to generate comprehensive user profiles. After training, we derive recommendations tailored to the obtained user profiles. To validate the efficacy and explainability of KEAM, extensive experiments are conducted on two real-world datasets. Ablation studies are also performed to investigate the impact of knowledge graphs.

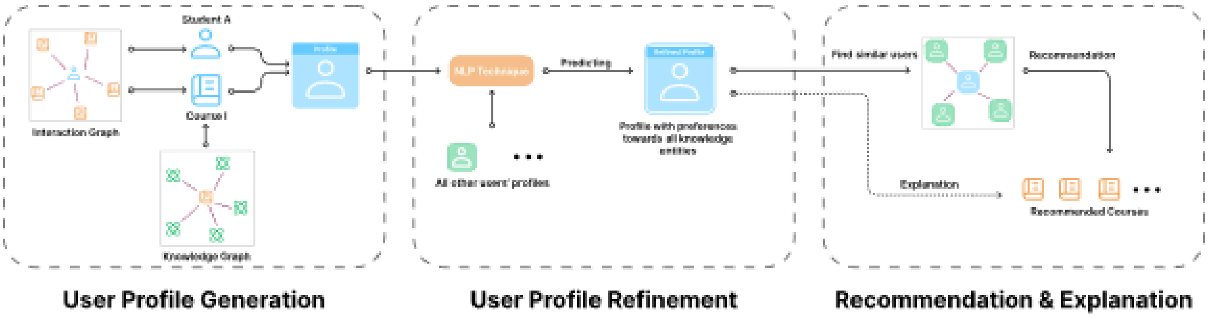

Figure 1. The architectural overview of the proposed KEAM model

2. Related Work

Historically, course recommender systems focused on content-based or collaborative filtering (CF) methods (Naren, 2020; Wager et al., 2023). Scholars Walk (Polyzou et al., 2019) employs a randomized wandering approach to acknowledge sequential course relationships. Deep learning is also widely used in course recommendations (Gong et al., 2020). Pardos et al. (2019) predict course probability distributions for recommendations. Gao et al. (2022) integrate a deep convolutional neural network with negative sequence mining. Knowledge graphs (Ye et al., 2021) are utilized to enrich user and item low-dimensional entity representations. The concept of meta-path representations (Ma et al., 2019) is proposed to provide a detailed description between user-item pairs. These studies (Yang et al., 2023; Jiang et al., 2023) explore how meta-graphs facilitate the representations of user/item. However, the 'black box' nature of deep learning models raises interpretability concerns (Ma et al., 2024). Thus, our primary objective is to delve into the potential of explainable course recommendation systems. Our approach explicitly constructs user profiles, which can enhance the explainability.

3. The Proposed KEAM Model

In this section, we introduce our model KEAM. Figure 1 shows an overview of KEAM. We shorten this section since we introduced our model detailed in this study (Yang et al., 2024).

3.1 User Profile Generation

Our model is designed as an autoencoder architecture (Bellini et al., 2018), which is shown in Figure 2. We inhibit the feed-forward and back-propagation steps for those neurons that are not connected in the KGs by employing a matrix M, where rows m and columns n represent courses c_j and knowledge entity k_{jx} in knowledge graph $K(j)$, respectively.

$$M_{m,n} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

where $M_{m,n}$ represents the adjacency matrix of the KG, $a_{i,j}$ is 1 if course i is connected to feature j, otherwise is 0. Hidden layer (h) and output layer (o) can be obtained as follows:

$$h = g\big(X \times (W_1 \circ M)\big), \qquad o = g\big(h \times (W_2 \circ M^T)\big),$$

where X is the input embedding of user u_i, W_1 and W_2 represent the weight matrix between courses and knowledge entities. After training, we extract the weights of the hidden layer and use them to generate a user profile $P(u_i)$ for user u_i, as follows:

$$P(u_i) = \{\langle f_{i1}, w_{i1}\rangle, \dots, \langle f_{im}, w_{im}\rangle\},$$

where f_i means the knowledge entity associated with the hidden neurons for each course and w_i represents its corresponding weight. Finally, we generate a profile of each user.

3.2 User Profile Refinement

Since users interacted with only a small subset of available courses, their profiles tend to be limited in scope. We introduce a refinement strategy leveraging collaborative information. As shown in Figure 3, we consider each knowledge entity and its corresponding weight in various user profiles as a distinct ‘word’ entity and treat each user profile as a ‘sentence’, which are fed into the Word2Vec model. After training, it gains the ability to predict word similarities. We leverage the model to generate similarity scores for all words used to refine profiles.

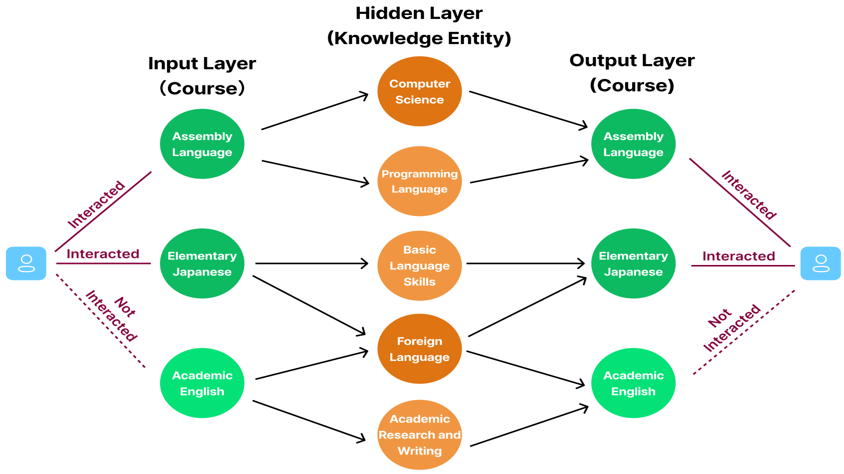

Figure 2. User Profile Generation in the KEAM

3.3 Recommendation and Explanation

Then, we employ the collaborative filtering technique to generate recommendations and generate a Top-N recommendation list for each user and present these recommendations alongside the user's individual profile, which includes the knowledge entities associated with each suggested course. This approach provides a clear rationale for each recommendation, ensuring it aligns with the user's interests and preferences.

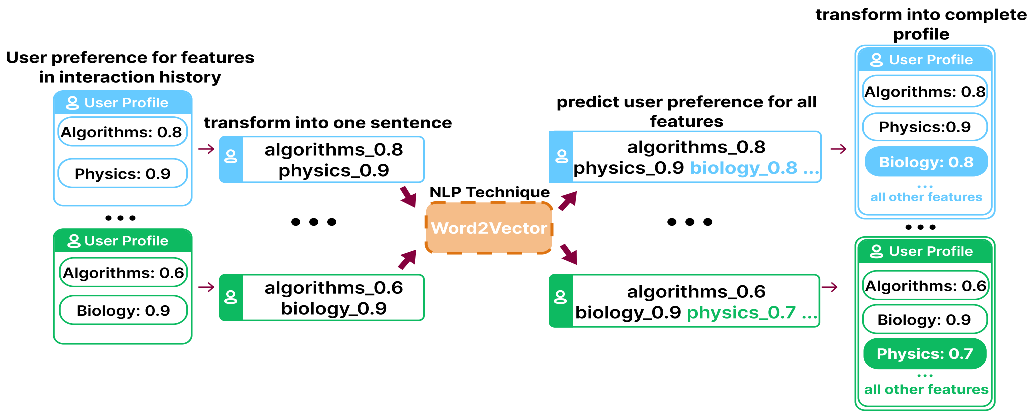

Figure 3. User Profile Refinement in the KEAM

4. Evaluation and Results

In this section, we introduce our model KEAM. Figure 1 shows an overview of KEAM. We shorten this section since we introduced our model detailed in this study (Yang et al., 2024).

4.1 Evaluation

In this work, our experiments use two datasets (Zhang et al., 2019; Yu et al., 2020) collected from the XuetangX[1] MOOC platform. Student information has been anonymized. The dataset statistics are introduced in Table 1. To estimate the effectiveness of our model, we rely on three metrics: Recall@K, HR@K and NDCG@K of Top-K recommendations. KEAM is compared with different baselines following our previous work (Yang et al., 2024).

4.2 Performance Comparison

We test the performance of our model with baselines, which is detailed in this study (Yang et al., 2024). We summarize the result that KEAM outperforms all baseline models on both two datasets. Since our model incorporates knowledge entities and integrates knowledge graph information. It verifies the effectiveness of KEAM in handling complex data structures.

Table 1. *Statistics of the two datasets*

Dataset	XuetangX	MOOCCube
Users	82,535	198,950
Courses	1,302	678
Interactions	458,454	677,019
Interactions Density	0.4266%	0.5019%

4.3 Model Study

4.3.1 Influence of Similar User Group Size

To explore the influence of the size of similar user groups, we evaluated different group sizes. As shown in Figure 4, we observed that KEAM's performance initially increases with the group size but then starts to decline as the size continues to grow. The possible reason is that small group size may not adequately capture the essential information. On the other hand, an excessively large group size can lead to increased model complexity and noise, which may not yield better outcomes, especially when dealing with sparse data.

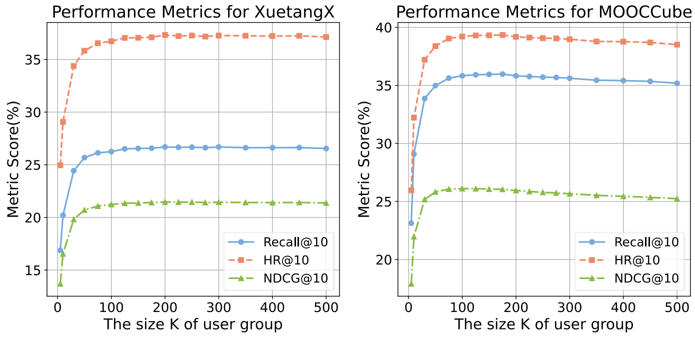

Figure 4. The impact of similar user group size on KEAM

4.3.2 Effectiveness of Knowledge Graphs

To analyze the contributions of various knowledge graphs, we conducted an experiment where we compared the KEAM leveraging different kinds of knowledge graphs for two datasets.

MOOCCube dataset provides three different types of information for constructing knowledge graphs: Field, Subject, and Concept. Field refers to broad categories and coarse granularity information, such as *"Computer Science and Technology"* for course *Data Structure*. Subject pertains to more detailed subject-level and medium granularity knowledge within those categories, such as *"Theoretical Computer Science"*. Concept includes specific,

[1] http://www.xuetangx.com

fine-grained knowledge concept with a course, such as *"Decision Making"* and *"Sorting"*. However, the XuetangX dataset only offers a broad categorization of courses at the field level. To better investigate the influence of different knowledge graphs, we enrich the knowledge graph to enhance its richness and utility with ChatGPT and expert knowledge. About extension of the knowledge graph using expert knowledge, we first employ NLP techniques for generating detailed categorization, followed by judgments of two experts based on descriptions. Therefore, we construct three different types of knowledge graphs in the XuetangX dataset: Field, Subject by GPT, Subject.

The results of the experiment in the XuetangX and MOOCCube are depicted in the top and bottom parts of Figure 5, respectively. We observe the following conclusions: (1) There is an enhancement in the performance with increasing granularity of knowledge entities on both datasets. It can be attributed to the richer and higher quality information content of more granular knowledge entities, facilitating more accurate user profile generation. (2) In the XuetangX dataset, the KEAM-Subject by GPT and the KEAM-Subject models both outperform the KEAM-Field. However, the KEAM-Subject model achieves better results, possibly because the knowledge entities generated by GPT may not be accurate as knowledge entities manually extracted by human experts. (3) In the MOOCCube dataset, the performance enhancement observed from Field to Subject is less than that from Subject to Concept. We argue that the discrepancy is due to the knowledge entities KEAM-Concept utilizing being more concept-oriented. Given that the interests of users may align more closely with specific concepts within a course rather than the broader categories.

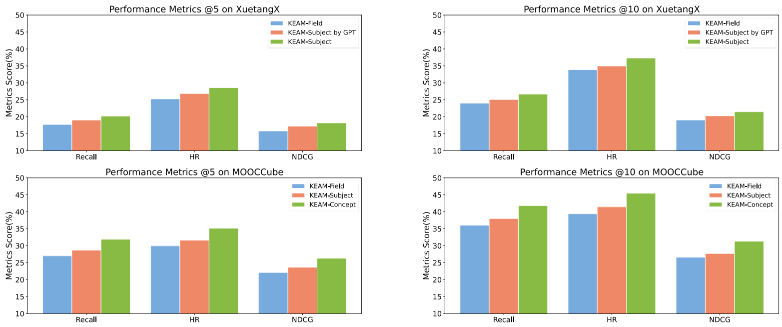

Figure 5. The effect of different knowledge graphs in KEAM

4.4 Explainability

During the recommendation, we explicitly show users their profiles about the knowledge entities that have their preference weights. Additionally, we provide textual explanations of the recommendations to facilitate user understanding. As shown in Figure 6, the left part is the recommendation list, and the right part shows the user profile which means the user preference on different knowledge entities. In this case, it becomes evident that the student shows an interest in knowledge entities related to 'nursing' while exhibiting lesser interest in history and automobiles. The result demonstrates that our model can provide effective recommendations and generate a user profile to provide reasonable explanations.

Figure 6. Case study

5. Conclusion

In this paper, we introduced a novel Knowledge Entity-Aware Model for explainable course recommendations, which can explicitly build students' profiles by incorporating knowledge graphs. We highlighted the significance of knowledge graphs, which can help generate more accurate and explainable recommendations. Through extensive experiments, we validated the effectiveness and explainability of KEAM. However, one significant constraint is that it treats all knowledge entities uniformly, disregarding their hierarchical relationships. Additionally, we intend to conduct a comprehensive user study involving real users.

Acknowledgements

This work was supported by JST SPRING, Grant Number JPMJSP2136, JSPS KAKENHI Grant Numbers JP20H00622 and JP24K20903.

References

Alkan, O., Daly, E. M., Botea, A., Valente, A. N., & Pedemonte, P. (2019, March). Where can my career take me? harnessing dialogue for interactive career goal recommendations. In Proceedings of the 24th International Conference on Intelligent User Interfaces (pp. 603-613).

Bellini, V., Schiavone, A., Di Noia, T., Ragone, A., & Di Sciascio, E. (2018, October). Knowledge-aware autoencoders for explainable recommender systems. In Proceedings of the 3rd workshop on deep learning for recommender systems (pp. 24-31).

Gao, M., Luo, Y., & Hu, X. (2022). Online course recommendation using deep convolutional neural network with negative sequence mining. Wireless Communications and Mobile Computing, 2022.

Ehara, Y. (2023). Measuring Similarity between Manual Course Concepts and ChatGPT-generated Course Concepts. Proceedings of the 16th International Conference on Educational Data Mining, 474--476. https://doi.org/10.5281/zenodo.8115758

Gong, J., Wang, S., Wang, J., Feng, W., Peng, H., Tang, J., & Yu, P. S. (2020, July). Attentional graph convolutional networks for knowledge concept recommendation in moocs in a heterogeneous view. In Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval (pp. 79-88).

Jiang, L., Liu, K., Wang, Y., Wang, D., Wang, P., Fu, Y., & Yin, M. (2023). Reinforced explainable knowledge concept recommendation in MOOCs. ACM Transactions on Intelligent Systems and Technology, 14(3), 1-20.

Jing, X., & Tang, J. (2017). Guess you like: course recommendation in MOOCs. In Proceedings of the international conference on web intelligence (pp. 783-789).

Ma, B., Lu, M., Taniguchi, Y., & Konomi, S. I. (2021). CourseQ: the impact of visual and interactive course recommendation in university environments. Research and practice in technology enhanced learning, 16, 1-24.

Ma, B., Lu, M., Taniguchi, Y., & Shin'ichi Konomi. (2021, April). Exploration and Explanation: An Interactive Course Recommendation System for University Environments. In IUI Workshops.

Ma, B., Lu, M., Taniguchi, Y., & Konomi, S. I. (2021). Investigating course choice motivations in university environments. Smart Learning Environments, 8, 1-18.

Ma, W., Zhang, M., Cao, Y., Jin, W., Wang, C., Liu, Y., ... & Ren, X. (2019, May). Jointly learning explainable rules for recommendation with knowledge graph. In The world wide web conference (pp. 1210-1221).

Morsomme, R., & Alferez, S. V. (2019). Content-Based Course Recommender System for Liberal Arts Education. International educational data mining society.

Morsy, S., & Karypis, G. (2019). Will this course increase or decrease your gpa? towards grade-aware course recommendation. arXiv preprint arXiv:1904.11798.

Naren, J., Banu, M. Z., & Lohavani, S. (2020). Recommendation system for students' course selection. In Smart Systems and IoT: Innovations in Computing: Proceeding of SSIC 2019 (pp. 825-834). Springer Singapore.

Pardos, Z. A., Fan, Z., & Jiang, W. (2019). Connectionist recommendation in the wild: on the utility and scrutability of neural networks for personalized course guidance. User modeling and user-adapted interaction, 29, 487-525.

Pardos, Z. A., & Jiang, W. (2019). Combating the filter bubble: Designing for serendipity in a university course recommendation system. arXiv preprint arXiv:1907.01591.

Polyzou, A., Nikolakopoulos, A. N., & Karypis, G. (2019). Scholars Walk: A Markov Chain Framework for Course Recommendation. International Educational Data Mining Society.

Wagner, K., Merceron, A., Sauer, P., & Pinkwart, N. (2023). Can the Paths of Successful Students Help Other Students with Their Course Enrollments?. International Educational Data Mining Society.

Yang, T., Ren, B., Ma, B., Khan, M. A. Z., He, T., & Konomi, S. I. (2024). Making Course Recommendation Explainable: A Knowledge Entity-Aware Model using Deep Learning. In Proceedings of the 17th International Conference on Educational Data Mining (pp. 658-663).

Yang, Y., Zhang, C., Song, X., Dong, Z., Zhu, H., & Li, W. (2023). Contextualized knowledge graph embedding for explainable talent training course recommendation. ACM Transactions on Information Systems, 42(2), 1-27.

Ye, Q., Hsieh, C. Y., Yang, Z., Kang, Y., Chen, J., Cao, D., ... & Hou, T. (2021). A unified drug–target interaction prediction framework based on knowledge graph and recommendation system. Nature communications, 12(1), 6775.

Zhang, J., Hao, B., Chen, B., Li, C., Chen, H., & Sun, J. (2019, July). Hierarchical reinforcement learning for course recommendation in MOOCs. In Proceedings of the AAAI conference on artificial intelligence (Vol. 33, No. 01, pp. 435-442).