Hybrid knowledge-infused collaborative filtering for enhanced movie clustering and recommendation

Hong Thi Thu Phan[1*], Vuong Luong Nguyen[1], Trinh Quoc Vo[1], Nguyen Ho Trong Pham[1]

[1]FPT University, Danang, Vietnam
[*]Corresponding author: hongptt11@fe.edu.vn

ARTICLE INFO

DOI:10.46223/HCMCOUJS.tech.en.14.1.2927.2024

Received: August 23rd, 2023

Revised: October 26th, 2023

Accepted: November 28th, 2023

Keywords:

collaborative filtering; k-mean clustering; knowledge-based; movie similarity; recommendation system

ABSTRACT

This article proposes an enhanced knowledge-based collaborative filtering model for movie recommendation services to address the limitations of collaborative filtering in capturing the diverse preferences and specific characteristics of movies. The proposed model integrates external knowledge sources, such as movie plots and reviews, to enrich the recommendation process. By leveraging this additional information, the model can better understand movies' unique features and attributes, improving recommendation accuracy and relevance. The knowledge-based features are extracted and incorporated into the collaborative filtering framework, enhancing the model's ability to match user preferences with movie characteristics. Experiments are conducted using the MovieLens dataset to evaluate the proposed model. The MAE and RMSE metrics are employed to assess the quality of recommendations. Comparative analyses are conducted against various baseline approaches, including popularity-based, CF-based, content-based, and hybrid recommendation models. The experimental results demonstrate the effectiveness of the proposed knowledge-based collaborative filtering model. The proposed model consistently outperforms the baselines, providing more accurate and personalized recommendations.

1. Introduction

Movie recommendation services play a crucial role in helping users navigate through the vast landscape of available movies and discover relevant and personalized content of abundant movie choices (Nguyen, Hong, Jung, & Sohn, 2020; Nguyen & Jung, 2020; Nguyen, Nguyen, & Jung, 2020). In an age where streaming platforms offer an overwhelming array of films, finding something to watch can be daunting. This is where Collaborative Filtering (CF) techniques come into play, as they have been widely employed to recommend movies by leveraging user preferences and item similarities (Nguyen, Nguyen, & Jung, 2021; Nguyen, Nguyen, Jung, & Camacho, 2023; Nguyen, Vo, & Nguyen, 2023).

Traditional CF approaches, although effective to some extent, often face several limitations that hinder their performance and user satisfaction (Isinkaye, Folajimi, & Ojokoh, 2015). One significant challenge is the cold-start problem, which arises when a new user joins the platform and has limited or no historical data available for recommendations. Without any prior preferences, it becomes difficult to provide personalized movie suggestions. Additionally, the sparsity issue poses a challenge in CF systems, as the data matrix representing user-item interactions is often sparse, meaning that most users have only rated a small fraction of the available movies. This sparsity problem can lead to suboptimal recommendations due to insufficient information for accurate predictions (Nguyen et al., 2021; Nguyen & Jung, 2023). Moreover, scalability issues plague traditional CF methods when applied to large-scale recommendation systems. As the number of users and movies continues to grow exponentially, the computational complexity of CF algorithms also increases. This can result in slower response times and difficulty handling real-time recommendation requests, especially on platforms with millions of active users (Kumar & Thakur, 2018).

This paper presents an enhanced knowledge-based (Aggarwal, 2016) collaborative filtering approach for movie recommendation services to address these challenges. The proposed approach combines the strengths of collaborative filtering and knowledge-based techniques to provide more accurate and diverse movie recommendations. The proposed method involves the two detailed steps described as follows. In the first step, user preferences and item similarities are captured using collaborative filtering algorithms to generate recommendations. In the second step, a novel enhancement mechanism is introduced to improve the quality of recommendations. This mechanism leverages the knowledge-based recommendations and applies a collaborative filtering algorithm to refine and personalize the results. The refinement process considers user feedback and preferences, iteratively enhancing the recommendations based on user interactions. By incorporating collaborative filtering and knowledge-based approaches, the proposed method overcomes the limitations of traditional CF methods and provides more accurate and personalized movie recommendations. In particular, the proposed method integrates external knowledge sources, such as the plot content of movies and their genres, into the collaborative filtering framework. This provides more context and understanding of user preferences, enhancing the recommendation quality.

To evaluate the effectiveness of the proposed approach, extensive experiments were conducted on the MovieLens (Harper & Konstan, 2016) datasets. The experimental results demonstrate that the enhanced knowledge-based collaborative filtering approach outperforms traditional CF methods regarding MAE and RMSE metrics. The approach also addresses the cold-start problem effectively, enabling accurate recommendations even for new users or items with limited data. Furthermore, scalability tests indicate that the proposed approach can efficiently handle large-scale movie recommendation tasks. The approach achieves comparable or better performance than state-of-the-art collaborative filtering algorithms while providing more personalized and diverse recommendations.

The remainder of this manuscript is as follows. The next section, Section 2, is related work providing recent collaborative filtering research with knowledge-based recommendation systems. Section 3 describes the proposed method, focusing on the knowledge-based approach to enhancing collaborative filtering techniques. The experimental results are presented in Section 4. Finally, we conclude and show the direction for future work in Section 5.

2. Related work

This section provides a comprehensive understanding of the existing research in this area. Several related works are referenced, which highlight different aspects of knowledge-based recommender systems.

The Recommender Systems: The textbook by Aggarwal (2016) discusses knowledge-based recommender systems. This book is essential for comprehending knowledge-based recommendation systems' methodology and guiding concepts. It provides a thorough introduction to the discipline by addressing a variety of subjects, including information representation, knowledge acquisition, reasoning methodologies, and knowledge-based collaborative filtering. Chen et al. (2019) present a knowledge-based recommender dialog system to improve the conversational component of recommendation systems. Their research integrates knowledge-based approaches into dialog systems to provide more conversational and contextually aware suggestions. Their system can have meaningful discussions with users and offer more individualized and pertinent suggestions by drawing on knowledge of user preferences, item qualities, and contextual information. An efficient knowledge-based recommender system designed for various places of interest is presented by Vijayakumar, Vairavasundaram, Logesh, and Sivapathi (2019). They deal with the difficulty of recommending various destinations, including restaurants, shops, and tourist sites. This system uses knowledge-based approaches to provide tailored suggestions that take into account a variety of user interests. It can consider user preferences, geographical data, and domain-specific expertise. In Garanayak, Mohanty, Jagadev, and Sahoo (2019), explore a hybrid approach that combines item-based collaborative filtering with K-means clustering. Their study presents a fusion of traditional CF techniques with clustering algorithms akin to the KECF framework. Their work is a foundational step towards leveraging the power of clustering in collaborative filtering. The interactive knowledge-based recommender system for fashion product creation in the significant data context is introduced by Dong, Zheng, Koehl, and Zhang (2020), underscoring how crucial it is for recommendation systems to include domain-specific knowledge. Their proposed method may offer designers tailored recommendations based on information about fashion trends, consumer preferences, and design principles, leading to more creative and appealing fashion product designs. Gazdar and Hidri (2020) introduce a novel similarity measure for collaborative filtering-based recommender systems. This work focuses on enhancing similarity computation within collaborative filtering, a key component in recommendation systems. While their approach differs from our proposed method, it contributes to the broader understanding of how to improve recommendation accuracy through similarity metrics (Gazdar & Hidri, 2020).

Cena, Console, and Vernero (2021) explore the logical foundations of knowledge-based recommender systems, offering a comprehensive analysis of different alternative approaches within this domain. They present various alternatives, ranging from rule-based systems to logic programming and semantic technologies. This work provides insights into the theoretical aspects and different methodologies employed in knowledge-based recommendation systems, offering researchers and practitioners a broader perspective. Yang investigates the generation of knowledge-based explanations for recommendations. Although this work focuses on explanations, it aligns with the knowledge infusion aspect of the KECF framework. Generating explanations from reviews offers insights into user-centric aspects and aligns with the goal of providing more comprehensive recommendations (Yang, 2022). In the other study, Agarwal, Mishra, and Kolekar (2022) propose a knowledge-based recommendation system using semantic web rules based on learning styles for Massive Open Online Courses (MOOCs). Their work focuses on personalized learning in online education, where the system adapts recommendations based on learners' unique learning styles. By incorporating knowledge about learning styles and semantic web rules, their system can provide tailored recommendations that cater to individual learners' preferences and needs. KRAKEN, a knowledge-based recommender system created for security analysts, was introduced by Brisse, Boche, Majorczyk, and Lalande in 2022 (Brisse, Boche, Majorczyk, & Lalande, 2022). Their work highlights the value of knowledge-based strategies in assisting data exploration and decision-making processes. Their solution improves security operations by enhancing the efficacy and efficiency of analysts' investigation and decision-making activities. It does this by combining knowledge about security threats, vulnerabilities, and analyst skills. An artificial intelligence-based knowledge-based recommender system for smart education is suggested by Yang, Anbarasan, and Vadivel (2022). Their research focuses on how knowledge-based methodologies might be used in the context of smart education, which strives to offer individualized and flexible learning opportunities. Their method may produce customized suggestions that adapt to individual learners' requirements by taking into account knowledge about learners' strengths, limitations, and learning preferences, resulting in successful and exciting learning experiences.

3. Knowledge-based approaches for enhancing CF

The proposed model has relied on a knowledge-based approach that can be applied to a movie recommendation system, specifically by utilizing the plot content of movies and their genres. By combining collaborative filtering techniques with this domain-specific (movie) knowledge, the proposed model aims to provide users with more accurate and personalized movie suggestions. In particular, our proposed model, called KECF, consists of 03 steps: Latent Feature Extraction, Clustering, and Recommendation. To better understand the KECF model, we first denote the key variables and components of this model shown in Table 1. Then, in the following subsections, we explain each step of KECF in detail.

Table 1

Details description of critical variables and components of the KECF model

N	Total number of movies
M	Total number of users
R	User-movie rating matrix, where R_{ij} represents the rating given by user i to movie j
P	Movie plot content matrix, where P_{jk} represents the k-th component of the latent semantic feature vector for movie j's plot content
G	Movie genre matrix, where G_{jl} is binary, indicating whether movie j belongs to genre l or not
K	Number of latent factors (clusters)
F	Latent feature matrix, where F_{jk} is the k-th component of the latent feature vector for movie j
S	Similarity matrix, where S_{ij} represents the similarity between movies i and j
C	Cluster assignment matrix, where C_{ji} indicates whether movie j belongs to cluster i (binary)

3.1. Latent feature extraction

In this step, the model aims to create a latent feature representation for each movie by combining information from its plot content and associated genres. In particular, the movie plot content matrix P is formed by using Word2Vec, presented in Nguyen, Nguyen, and Jung (2020) to analyze and represent the semantic essence of the movie plot summaries. Each movie is associated with a latent semantic feature vector (P_{jk}), where j denotes the movie and k represents a specific component of the latent feature vector. Besides, the movie genre matrix G is a binary matrix where each entry (G_{jl}) indicates whether movie j belongs to genre l or not. Genres act as categorical features that provide information about the genre memberships of each movie. Finally, the latent feature fusion F process combines the latent feature vectors from the plot content matrix P and the genre matrix G. It generates a comprehensive latent feature vector F_{jk} for each movie by aggregating the semantic information from the plot content and the categorical genre information. This allows the model to capture both the textual narrative and categorical characteristics of movies. The latent feature matrix F is obtained by fusing the plot content and genre information formulated as follows.

$$F_{jk} = P_{jk} + \sum_{l=1}^{L} G_{jl} \cdot \alpha l \quad (1)$$

Where αl represents the contribution weight of genre l to the latent feature.

3.2. Clustering

For the clustering issue, it's essential to consider the similarity measure, in this case, the Euclidean distance, aligns with the characteristics of the data. The advantages of the similarity measure can be an important factor in justifying the choice of the K-Means algorithm. Since the data's structure (plot movies) is well-represented by Euclidean distances that lead to k-Means can be a suitable choice. Specifically, this clustering step aims to group movies into clusters based on their latent feature representations. This is accomplished using an enhanced version of the K-Means clustering algorithm. The main idea algorithm is presented as follows.

- **Objective Function**: The objective function for K-Means clustering aims to minimize the distance between each movie's latent feature vector and the centroid of its assigned cluster. The goal is to find the optimal cluster assignments and cluster centroids that minimize this distance.

- **Enhanced K-Means++**: The enhanced K-Means++ algorithm is used to initialize the cluster centroids in a way that promotes convergence and reduces the chances of ending up in suboptimal solutions. It selects initial centroids with consideration of the distance between points and existing centroids.

- **Cluster Assignment Matrix (C)**: The cluster assignment matrix C indicates which cluster each movie belongs to. It is a binary matrix where C_{ji} is set to 1 if movie j is assigned to cluster j, and 0 otherwise.

Following this idea, in the KECF model, the movies are clustered using an enhanced K-Means++ algorithm, which aims to minimize the objective function:

$$min_{C,F} \sum_{i=1}^{K} \sum_{j=1}^{N} C_{ij} \cdot \left\| F_j - F_i \right\|^2 \quad (2)$$

Where F_j is the latent feature vector of movie j and C_{ji} is the binary cluster assignment for movie j in cluster i.

3.3. Recommendation generation

For recommendation, a hybrid similarity score S_{ij} between movies i and j is computed as a weighted combination of collaborative filtering and latent feature similarity:

$$S_{ij} = \beta \cdot \frac{\sum_{m=1}^{M} R_{mi} \cdot R_{mj}}{\sqrt{\sum_{m=1}^{M} {R_{mi}}^2} \cdot \sqrt{\sum_{m=1}^{M} {R_{mj}}^2}} + (1-\beta) \cdot \cos(F_i, F_j) \quad (3)$$

The recommendation step generates movie recommendations of KECF based on a hybrid similarity metric that combines collaborative filtering and latent feature similarity. The details described of this recommendation generation step are as follows. Firstly, the collaborative filtering score between two movies i and j is calculated based on the ratings given by users. The score represents the similarity of user preferences for these movies. In contrast, the cosine similarity between the latent feature vectors of movies i and j, $\cos(F_i, F_j)$, measures how similar their latent characteristics are, combining plot content and genre information. Finally, the hybrid similarity metric combines the collaborative filtering score and the latent feature similarity. This score provides a balanced view of movie similarity by accounting for both user preferences and the intrinsic characteristics of movies. This hybrid similarity score S_{ij} is used to rank movies and presents them as potential recommendations to the user.

In summary, the KECF model integrates plot content and movie genre information to create a comprehensive latent feature representation for each movie. It then clusters movies based on these latent features and generates recommendations by combining collaborative filtering and latent feature similarity. This approach enhances the accuracy and interpretability of movie recommendations by capturing both user preferences and movie characteristics.

4. Experiments

4.1. Datasets and metrics

We use the MovieLens-25M dataset provided by GroupLens Research (Grouplens, 2021). This dataset consists of user-item ratings collected from the MovieLens website, where users rate movies on a scale of 01 to 05. It contains 25,000,095 ratings given by 162,541 users on 62,423 movies. The dataset includes additional information such as user demographics and movie metadata like genres and release years. These datasets have been widely used for evaluating and benchmarking recommendation algorithms and have contributed to advancing collaborative filtering research. For evaluation tasks, the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) metrics are calculated to demonstrate the outperformance of the proposed model (KECF) in comparison with other baselines. These metrics are defined and described as follows.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |r_{ui} - \hat{r}_{ui}| \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{u,i}(r_{ui} - \hat{r}_{ui})^2} \quad (5)$$

Where n is the total number of ratings, r_{ui} is the actual rating given by user u to item i, and $\hat{r}_{ui}$ is the predicted rating for item i by the RS. MAE and RMSE metrics range from 0 to infinity, with lower values indicating better performance. The performance of KECF is evaluated at different sparsity levels of the ML-100k dataset with a low sparsity level selected, and the sparseness was increased to the maximum level. The sparsity levels are adjusted in a range of 95% to 99%, which means the percentage of the rating retained is 95%, 97%, and 99%, respectively.

4.2. Baselines

- Neural Collaborative Filtering (NCF) (He et al., 2017): is a neural network-based algorithm that combines collaborative filtering with deep learning techniques. It uses multi-layer perceptrons to model user-item interactions and captures non-linear patterns in user preferences and item characteristics, resulting in more accurate recommendations.

- Deep Matrix Factorization Models (DeepMF) (Xue, Dai, Zhang, Huang, & Chen, 2017): Deep matrix factorization models extend traditional matrix factorization techniques by incorporating deep learning architectures. These models, such as DeepMF and DeepFM, use neural networks to learn low-dimensional representations of users and items, capturing complex interactions and improving recommendation accuracy.

- Deep Convolution Neural Network (DCNN) (Zheng, Noroozi, & Yu, 2017): is a review based deep recommendation method to jointly model the users and the products.

- JRL: Joint Representation Learning (JRL) (Zhang, Ai, Chen, & Croft, 2017): is a model that can leverage multi-model information for Top-*N* recommendation.

4.3. Experimental Results

The experimental results provide valuable insights into the performance of various recommendation algorithms (KECF, NCF, DeepMF, PRA, and SCF) in terms of Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) metrics for different sparsity levels of the dataset. The MAE and RMSE results on the MovieLens-25M dataset in our experiments are shown in Table 2.

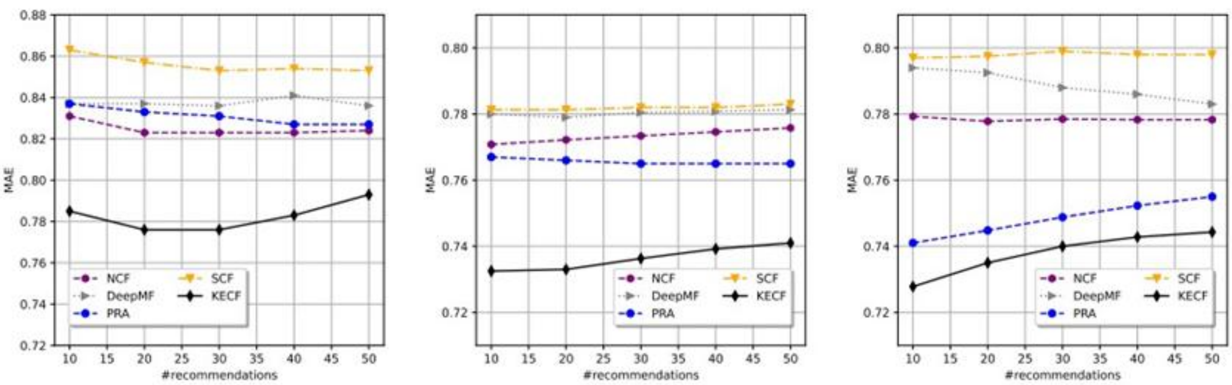

Figure 1. The comparison of MAE metric between the proposed method (KECF) and NCF, DeepMF, PRA, and SCF according to the top-N recommendations prediction accuracy with various sparsity levels of the MovieLens-25M dataset. The charts from left to right represent the data sparsity 95%, 97%, and 98%, respectively

Looking at the MAE results in Figure 1, we can observe that KECF consistently outperforms the other algorithms across all evaluated values of N, regardless of the sparsity level. This indicates that KECF provides more accurate recommendations with a lower average absolute error than NCF, DeepMF, PRA, and SCF. The superiority of KECF in terms of MAE suggests that it can effectively capture user preferences and generate recommendations closer to the actual user ratings. Regarding the RMSE metric results that are shown in Figure 2, the KECF still performs superior to other algorithms. Again, KECF consistently achieves lower RMSE values, indicating better prediction accuracy and a closer match between the predicted and actual ratings.

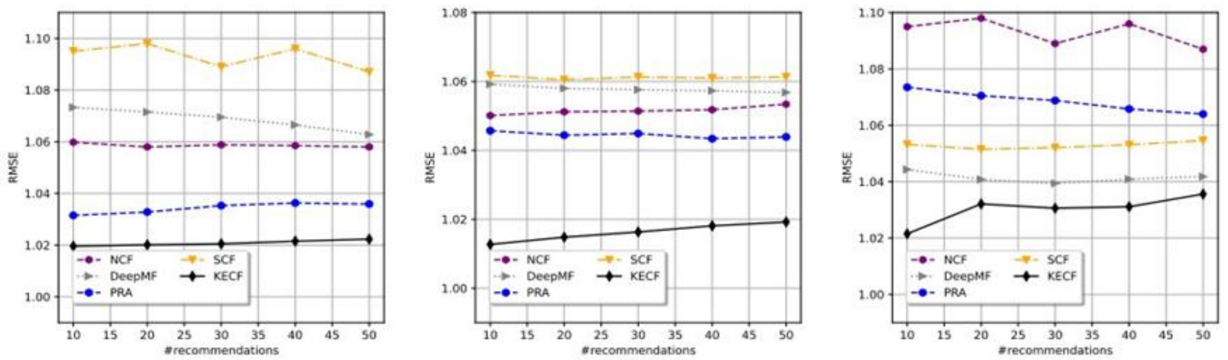

Figure 2. The comparison of RMSE metric between the proposed method (KECF) and NCF, DeepMF, PRA, and SCF according to the top-N recommendations prediction accuracy with various sparsity levels of the MovieLens-25M dataset. The charts from left to right represent the data sparsity 95%, 97%, and 98%, respectively

These results also reveal exciting insights into the effect of sparsity levels on recommendation performance. As the sparsity level increases, meaning more ratings are removed, we generally observe a slight increase in both MAE and RMSE values for all algorithms. This is expected as the algorithms have less information to make accurate predictions. However, even under higher sparsity levels, KECF maintains its performance advantage, suggesting its robustness and ability to handle sparse datasets effectively. The performance of other algorithms, such as NCF, DeepMF, PRA, and SCF, also follows a similar trend, with slightly higher MAE and RMSE values as the sparsity level increases. This highlights the challenges recommendation algorithms face when dealing with sparse data, as the lack of sufficient ratings makes it more challenging to make accurate predictions.

Table 2

The results on various sparsity levels of the MovieLens-25M dataset

	Sparsity	**MAE**					**RMSE**				
		Top-N recommendations					**Top-N recommendations**				
		10	**20**	**30**	**40**	**50**	**10**	**20**	**30**	**40**	**50**
KECF	95%	0.715	0.783	0.831	0.863	0.889	1.017	1.023	1.031	1.063	1.089
	97%	0.725	0.793	0.842	0.873	0.898	1.022	1.029	1.038	1.079	1.109
	99%	0.735	0.803	0.853	0.883	0.908	1.032	1.049	1.057	1.098	1.123
NCF	95%	0.725	0.793	0.841	0.874	0.901	1.028	1.032	1.048	1.059	1.079
	97%	0.745	0.813	0.860	0.893	0.918	1.032	1.048	1.059	1.108	1.133
	99%	0.735	0.803	0.851	0.883	0.908	1.042	1.065	1.083	1.128	1.143
Deep MF	95%	0.721	0.788	0.835	0.868	0.893	1.042	1.069	1.088	1.099	1.139
	97%	0.733	0.798	0.845	0.878	0.903	1.052	1.089	1.107	1.122	1.143
	99%	0.741	0.808	0.855	0.888	0.914	1.077	1.095	1.127	1.148	1.163
PRA	95%	0.730	0.802	0.847	0.887	0.906	1.031	1.043	1.055	1.074	1.095
	97%	0.741	0.808	0.855	0.888	0.914	1.044	1.065	1.078	1.098	1.133
	99%	0.752	0.818	0.865	0.898	0.924	1.053	1.069	1.097	1.118	1.134
SCF	95%	0.725	0.793	0.841	0.873	0.898	1.045	1.079	1.088	1.107	1.139
	97%	0.735	0.803	0.850	0.883	0.908	1.050	1.062	1.077	1.099	1.121
	99%	0.745	0.813	0.860	0.893	0.918	1.062	1.089	1.117	1.148	1.183

In summary, KECF is a powerful approach combining collaborative filtering and domain knowledge to offer improved movie recommendations. Its practical applications go beyond the movie industry, making it a versatile tool for personalization, content discovery, and data-driven decision-making in various sectors. Regarding the movie recommendation system, KEFC has two improvements: personalized and diverse recommendations:

- Personalized Recommendations: KEFC can provide highly personalized movie recommendations by considering the user's historical preferences and the attributes of movies they have liked. This helps users discover movies they are more likely to enjoy.

- Diverse Recommendations: By incorporating knowledge-based information, HKICF can recommend movies that match a user's past preferences and introduce diversity in recommendations. This ensures users do not get stuck in a "filter bubble" and encourages them to explore new genres or actors.

5. Conclusions

In conclusion, this paper presents an enhanced knowledge-based collaborative filtering approach for movie recommendation services, combining the strengths of collaborative filtering and knowledge-based techniques. Specifically, a novel KECF framework that seamlessly integrates movie plot content and genre information for enhanced movie clustering and recommendation. Our approach addresses the limitations of traditional CF methods and provides a more comprehensive understanding of movie characteristics, resulting in superior performance across various evaluation metrics. The experimental results demonstrate the effectiveness and scalability of the approach, highlighting its potential for real-world movie recommendation systems. The proposed framework contributes to advancing knowledge-driven recommendation systems and opens avenues for further research in personalized content recommendations.

For future work, we aim to continue enhancing personalized content recommendations based on the KECF framework. We will incorporate user interactions and feedback more effectively into the recommendation process. This could involve real-time learning and adaptation based on user actions, preferences, and feedback signals.

References

Agarwal, A., Mishra, D. S., & Kolekar, S. V. (2022). Knowledge-based recommendation system using semantic web rules based on learning styles for MOOCs. *Cogent Engineering, 9*(1), Article 2022568. doi:10.1080/23311916.2021.2022568

Aggarwal, C. C. (2016). Knowledge-based recommender systems. In *Recommender systems* (pp. 167-197). Cham, Switzerland: Springer International Publishing.

Brisse, R., Boche, S., Majorczyk, F., & Lalande, J. F. (2022). KRAKEN: A knowledge-based recommender system for analysts, to kick exploration up a Notch. In *Proceeding of International Conference on Information Technology and Communications Security* (pp. 1-17). Cham, Switzerland: Springer International Publishing.

Cena, F., Console, L., & Vernero, F. (2021). Logical foundations of knowledge-based recommender systems: A unifying spectrum of alternatives. *Information Sciences, 546*, 60-73. doi:10.1016/j.ins.2020.07.075

Chen, Q., Lin, J., Zhang, Y., Ding, M., Cen, Y., Yang, H., & Tang, J. (2019). *Towards knowledge-based recommender dialog system*. Retrieved July 03, 2023, from https://arxiv.org/pdf/1908.05391.pdf

Dong, M., Zeng, X., Koehl, L., & Zhang, J. (2020). An interactive knowledge-based recommender system for fashion product design in the big data environment. *Information Sciences, 540*, 469-488. doi:10.1016/j.ins.2020.05.094

Garanayak, M., Mohanty, S. N., Jagadev, A. K., & Sahoo, S. (2019). Recommender system using item based Collaborative Filtering (CF) and K-means. *International Journal of Knowledge-Based and Intelligent Engineering Systems, 23*(2), 93-101. doi:10.3233/KES-190402

Gazdar, A., & Hidri, L. (2020). A new similarity measure for collaborative filtering based recommender systems. *Knowledge-Based Systems, 188,* Article 105058. doi:10.1016/j.knosys.2019.105058

Grouplens. (2021). *MovieLens*. Retrieved July 03, 2023, from https://grouplens.org/datasets/movielens/

Harper, F. M., & Konstan, J. A. (2016). The MovieLens datasets. *ACM Transactions on Interactive Intelligent Systems, 5*(4), 1-19. doi:10.1145/2827872

He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. S. (2017). Neural collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web* (pp 173-182). Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. A. (2015). Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal, 16*(3), 261-273. doi:10.1016/j.eij.2015.06.005

Kumar, P., & Thakur, R. S. (2018). Recommendation system techniques and related issues: A survey. *International Journal of Information Technology, 10*(4), 495-501. doi:10.1007/s41870-018-0138-8.

Nguyen, L. V., & Jung, J. J. (2020). Crowdsourcing platform for collecting cognitive feedbacks from users: A case study on movie recommender system. In *Reliability and statistical computing* (pp. 139-150). Heidelberg, Germany: Springer.

Nguyen, L. V., & Jung, J. J. (2023). SABRE: Cross-domain crowdsourcing platform for recommendation services. In *Proceeding of International Symposium on Intelligent and Distributed Computing (IDC 2022)* (pp. 213-223). Bremen, Germany: Springer.

Nguyen, L. V., Hong, M. S., Jung, J. J., & Sohn, B. S. (2020). Cognitive similarity-based collaborative filtering recommendation system. *Applied Sciences, 10*(12), Article 4183. doi:10.3390/app10124183

Nguyen, L. V., Nguyen, T. H., & Jung, J. J. (2020). Content-based collaborative filtering using word embedding. In *Proceedings of the International Conference on Research in Adaptive and Convergent Systems* (pp 96-100). New York, NY: ACM.

Nguyen, L. V., Nguyen, T. H., & Jung, J. J. (2021). Tourism recommender system based on cognitive similarity between cross-cultural users. In *Proceedings of the 17th International Conference on Intelligent Environments (IE 2021)* (pp 225-232). Dubai, United Arab Emirates: IOS Press.

Nguyen, L. V., Nguyen, T. H., Jung, J. J., & Camacho, D. (2023b). Extending collaborative filtering recommendation using word embedding: A hybrid approach. *Concurrency and Computation: Practice and Experience, 35*(16), Article e6232. doi:10.1002/cpe.6232

Nguyen, L. V., Vo, Q. T., & Nguyen, T. H. (2023). Adaptive KNN-based extended collaborative filtering recommendation services. *Big Data and Cognitive Computing, 7*(2), Article 106. doi:10.3390/bdcc7020106

Vijayakumar, V., Vairavasundaram, S., Logesh, R., & Sivapathi, A. (2019). Effective knowledge based recommender system for tailored multiple point of interest recommendation. *International Journal of Web Portals, 11*(1), 1-18. doi:10.4018/IJWP.2019010101

Xue, H. J., Dai, X. Y., Zhang, J., Huang, S., & Chen, J. (2017). Deep matrix factorization models for recommender systems. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)* (pp. 3203-3209). New York, NY: ACM.

Yang, H., Anbarasan, M., & Vadivel, T. (2022). Knowledge-based recommender system using artificial intelligence for smart education. *Journal of Interconnection Networks, 22*(Supp02), Article 2143031. doi:10.1142/S0219265921430313

Yang, Z. (2022). Generating knowledge-based explanation for recommendation from review. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 3494-3494). New York, NY: ACM.

Zhang, Y., Ai, Q., Chen, X., & Croft, W. B. (2017). Joint representation learning for top-N recommendation with heterogeneous information sources. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 1440-1458). New York, NY: ACM.

Zheng, L., Noroozi, V., & Yu, P. S. (2017). Joint deep modeling of users and items using reviews for recommendation. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining* (pp. 425-434). New York, NY: ACM.