DOI: 10.54254/2755-2721/2025.LD27259

A Construction Method for Personalized Movie Recommendation System Based on Embedding

Wenwei Zeng

South China Normal University, Shanwei, China
952133592@qq.com

Abstract. Personalized recommendation systems are essential for enhancing user experience and platform profitability. This study presents a deep learning-based movie recommendation model employing embedding techniques, trained on the MovieLens-25M dataset. Unlike traditional collaborative filtering methods, the proposed system maps users and movies into a low-dimensional latent space, capturing complex nonlinear interactions through a multi-layer perceptron. To address data sparsity and cold-start limitations, negative sampling and a leave-one-out evaluation strategy were implemented. Experimental results demonstrate strong performance, achieving a Hit Ratio @10 of 0.85, thereby validating the effectiveness of embeddings in capturing user preferences. While the model shows robust recommendation quality, it is constrained by the lack of multimodal data and temporal modeling. Future work will explore Transformer architectures, large language models, and multimodal integration to enhance personalization, interpretability, and adaptability in dynamic environments.

Keywords: Movie Recommendation System, MovieLens, Deep Learning, Embedding Techniques

1. Introduction

With the rapid advancement of internet technologies and the advent of the big data era, users are confronted daily with an overwhelming volume of diverse information [1]. Efficiently and accurately filtering this vast sea of data to identify content that aligns with users' genuine interests has become a critical challenge for major internet platforms. Against this backdrop, recommendation systems have emerged as a pivotal solution. As a key technology for mitigating information overload, recommendation systems not only enhance user experience and engagement but also enable enterprises to uncover latent demand, thereby boosting commercial profitability.

Recommendation systems are widely deployed across e-commerce, video streaming, news, and music platforms, functioning as the “second search engine” of the internet industry. Among these domains, movie recommendation stands out as particularly representative and competitive. With the exponential growth of the film industry, the number of available movies has surged dramatically, leaving users overwhelmed by choice. An intelligent and efficient movie recommendation system that accurately captures user preferences can simultaneously enhance user satisfaction and platform revenue, thus effectively resolving these challenges. Building a system capable of deeply understanding user interests and delivering highly accurate recommendations holds both practical significance and strong research value. However, as user demands become increasingly diverse and personalized, traditional recommendation techniques fall short of meeting contemporary expectations. Classical collaborative filtering (CF) methods, despite early success, exhibit inherent limitations. These methods rely heavily on user ratings or interaction data; when the user-item matrix is sparse and behavioral data insufficient, both accuracy and coverage of recommendations deteriorate markedly. Moreover, CF struggles to model users' deeper interest features and lacks semantic understanding of content, often reducing "personalization" to a superficial level. It fails to deliver truly individualized recommendations. Critically, CF methods are ineffective in addressing the cold-start problem. When new users or new movies are introduced, the absence of historical behavioral data leads to significantly degraded recommendation performance.

In recent years, the rise of deep learning has opened new avenues for recommendation system development. Among these, embedding techniques—central to deep learning—offer powerful representational capabilities and have introduced novel approaches for building efficient recommendation models. By mapping users and movies into a continuous low-dimensional vector space, embedding enables users with similar preferences and movies with similar features to cluster closely in this space. Such dense vector representations can capture complex, non-linear relationships between users and items, while neural networks automatically extract latent features, significantly improving model generalization and recommendation quality. Furthermore, embedding methods can incorporate unstructured data such as text, images, and social signals, enabling the construction of multimodal recommendation frameworks that further enhance personalization.

In light of this technological context, we propose and implement a personalized movie recommendation system based on embedding techniques. This system demonstrates excellent performance and recommendation quality, with a deeper understanding of user preferences and more accurate, diverse recommendation outcomes. Empirical results confirm its strong adaptability and practical value, offering an efficient and viable solution for the movie recommendation domain.

2. Previous works

With the continuous evolution of the technological landscape and shifting user demands, recommendation systems have undergone several developmental stages. From early rule-based approaches to today's intelligent models powered by deep learning, the core objective remains consistent: to deliver highly personalized and accurate content recommendations to users.

Among traditional recommendation techniques, collaborative filtering (CF) has been one of the most widely adopted [2]. This method primarily relies on historical user–item interaction data. User-based CF infers a target user's preferences by identifying similar users. It offers advantages such as simplicity, intuitive logic, and independence from item content. However, in the presence of sparse user behavior data, similarity calculations between users become unreliable, leading to reduced recommendation accuracy. Moreover, as the user base expands, computing pairwise similarities incurs high time and space costs. In contrast, item-based CF calculates similarities between items, recommending movies similar to those the user has previously engaged with. This approach typically exhibits greater data stability, as item attributes are relatively constant compared to user behavior, allowing for reusable computations. Nonetheless, when the number of items grows large, computing the item similarity matrix remains computationally expensive. Additionally, item-based CF still struggles with cold-start issues, especially for newly released movies. To address these limitations, hybrid recommendation systems have emerged. These methods integrate user-based and item-based CF or further combine content-based strategies with collaborative filtering, aiming to balance accuracy, diversity, and coverage. For instance, some hybrid approaches prioritize collaborative filtering for users with rich behavioral history while shifting to content features when user data is sparse, thereby mitigating cold-start problems. Although hybrid models generally improve overall recommendation performance, their structures are complex, tuning is labor-intensive, and the systems are sensitive to feature engineering and parameter configuration. Beyond CF, content-based recommendation represents another critical branch of traditional methods [3]. This approach matches static features of movies—such as genre, director, cast, and plot descriptions —with those from a user's liked items to generate recommendations. Since it does not rely on other users' data, it partially alleviates data sparsity and cold-start issues. Moreover, its interpretability is stronger, as the reasons for recommendations can be explicitly traced. However, content-based methods depend heavily on manually extracted features and often struggle to uncover latent user interests. As a result, they tend to produce the so-called "information cocoon" effect, limiting diversity and exploration in recommendations. Several improvements on similarity-based methods have also been proposed within traditional frameworks. Techniques using cosine similarity, Pearson correlation, or Jaccard index aim to refine the measurement of user or item relationships [4]. While these methods are grounded in neighborhood-based CF, their varied similarity metrics seek to enhance both accuracy and personalization. Despite their foundational role in early recommender systems, the limitations of these traditional approaches have become increasingly evident. As user demands grow in complexity and scale, such methods are no longer sufficient for robust interest modeling.

With breakthroughs in deep learning, neural network-based models have gradually been introduced into recommender systems to extract high-order features and capture latent relations automatically. Convolutional neural networks (CNNs) can effectively leverage visual information such as movie posters and stills to improve recommendation quality [5]. Recurrent neural networks (RNNs) are well-suited to model temporal dependencies in user behavior sequences, capturing dynamic interest shifts more accurately. Compared with traditional methods, these deep models handle more complex nonlinear patterns and yield significantly improved recommendations. However, they also require large-scale data and high computational resources, and often suffer from poor interpretability.

Amid this trend, embedding techniques—core to deep learning—have begun to play a prominent role in recommendation systems. By mapping users and movies into dense, low-dimensional vector spaces, embeddings efficiently capture latent features and nonlinear relationships between users and items. Compared to conventional collaborative filtering, embedding-based approaches offer greater robustness under sparse data and cold-start conditions. Furthermore, they can integrate multimodal information such as text and images, enabling more precise and personalized recommendations.

3. Dataset and preprocessing

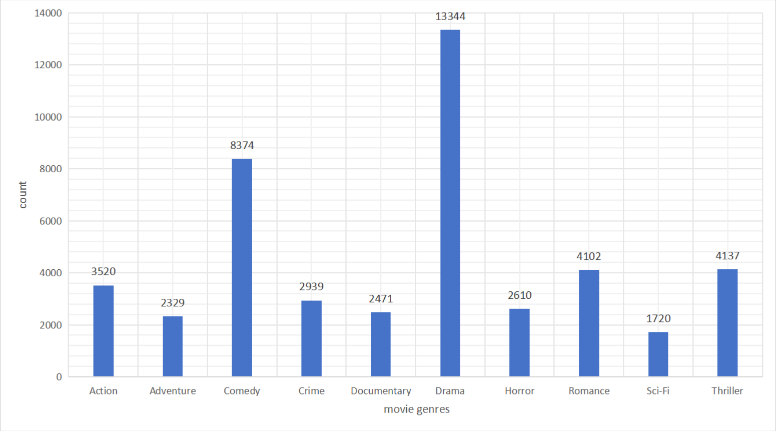

Figure 1. Distribution of movie genres in the dataset

This study employs the MovieLens-25M dataset, publicly released by GroupLens Research [6]. The dataset contains 25,000,095 five-star rating records, encompassing the interactions of 162,541 users with 62,423 movies. Each record includes a user ID, movie ID, a rating score ranging from 0.5 to 5.0 in increments of 0.5, and a timestamp. In addition, the dataset provides movie titles, genre labels, and external identifiers linked to IMDB and TMDB, facilitating the enrichment of metadata. Figure 1 presents the distribution of the top 10 most frequent genres within the MovieLens-25M movie catalog. Figure 2 illustrates the distribution of user rating scores for the movies in the dataset.

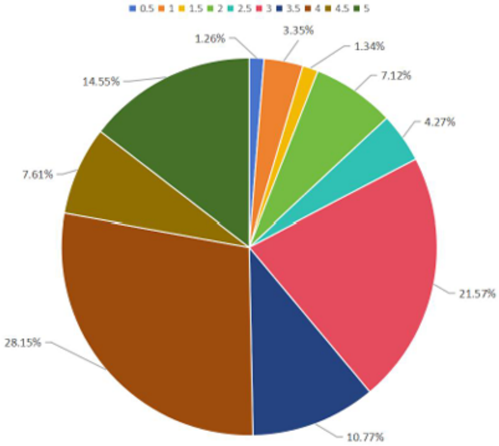

Figure 2. Distribution of rating scores in the dataset

During data preprocessing, a random subset of users was selected from the original rating data to maintain manageable data scale and to remove unnecessary redundant information. To construct the training and test sets, each user's rating history was sorted chronologically by timestamp. Following the leave-one-out strategy, the most recent rating record of each user was placed into the test set, while the remaining interactions formed the training set. This setting simulates a realistic recommendation scenario for predicting future behavior. In the training set, all user–movie interactions were treated as signals of positive feedback. Given that recommendation systems often suffer from sparse positive feedback, negative samples were introduced to better capture users' true preference distribution. Specifically, for each positive user–movie pair in the training set, multiple negative samples were randomly selected from the pool of movies the user had not interacted with. These negatives were combined with the positives at a fixed ratio (e.g., 4:1), resulting in a balanced training dataset comprising both positive and negative samples.

4. Model and results

In this study, a personalized movie recommendation model was developed based on a deep learning framework. The model aims to capture latent associative features between users and movies through embedding techniques, and to efficiently model nonlinear relationships using neural networks. Unlike traditional collaborative filtering methods, this approach does not directly rely on the sparse user–movie rating matrix. Instead, users and movies are mapped into low-dimensional dense vector representations (embeddings), and their similarity is computed within a continuous vector space [7]. This design not only alleviates the issue of data sparsity but also enables a more natural capture of deep user interest features in the embedding space.

Table 1. Architecture of the proposed recommendation model

	Name	Type	Params
0	User input	Embedding	1 M
1	Movie input	Embedding	1 M
2	Neural Network	Linear	1 K
3	Neural Network	Linear	2 K
4	Output	Linear	33

As illustrated in Table 1, the overall architecture consists of two embedding layers and three fully connected layers. User and movie IDs are first independently transformed into 8-dimensional dense vectors via separate embedding layers. This representation maps discrete user and movie identifiers into a continuous latent feature space, providing an efficient input for subsequent feature interaction and pattern learning. The resulting user and movie embeddings are then concatenated and passed into a multi-layer perceptron (MLP). Two nonlinear fully connected layers, comprising 1K and 2K parameters respectively, progressively extract higher-order feature interactions. Nonlinear activation functions, such as ReLU, are applied to facilitate the learning of complex mappings between user interests and movie attributes. The final layer is a linear output layer that yields a score representing the predicted match or preference between a user and a movie.

During training, observed user–movie interactions are treated as positive samples. Negative samples are generated by randomly sampling movies that the user has not watched, forming a training dataset with a 1:4 ratio of positive to negative instances. The model is optimized using a cross-entropy loss function. Parameters in both the embedding and fully connected layers are updated via gradient descent to maximize the model's ability to distinguish positive from negative samples. To evaluate recommendation performance, Hit Ratio @10 is employed as the primary metric. This measures whether the set of top 10 recommended movies includes the item the user actually interacted with. Experimental results show that the proposed model achieves a Hit Ratio @10 of 0.85 on the MovieLens dataset, indicating that the model successfully captures true user preferences in most top-10 recommendation scenarios, and demonstrating strong recommendation quality and practical value.

5. Conclusion

This study focuses on personalized movie recommendation and proposes a deep learning recommendation model based on embedding techniques using the MovieLens dataset. The model maps users and movies into low-dimensional dense vectors and employs a multi-layer perceptron (MLP) to capture their latent interaction patterns, enabling effective characterization of user preference profiles. Experimental results demonstrate that the model achieves a Hit Ratio @10 of 0.85, validating the effectiveness and feasibility of embedding techniques in the task of movie recommendation.

Despite the promising results, the study presents several limitations. First, the experiments were conducted solely on the MovieLens dataset, which, although widely adopted as a benchmark in recommender system research, remains limited in both scale and diversity compared to real-world commercial platforms involving tens or even hundreds of millions of user interactions. This scale constraint prevents comprehensive evaluation of the model's robustness and generalization in large-scale, complex environments, and limits its assessment on long-tail or cold-start users. Second, the model primarily relies on users’ historical rating data for feature construction, without leveraging multimodal information such as movie descriptions, user reviews, or movie posters. These unstructured data sources often contain rich semantic and emotional signals, which are crucial for a more comprehensive representation of user interests and movie attributes. The lack of such auxiliary signals weakens the model’s ability to uncover deep user preferences and to learn nuanced content features, potentially affecting the diversity and interpretability of the recommendations. Additionally, the current model adopts a relatively simple architecture, where feature interactions are modeled through embeddings and an MLP. Although this captures certain nonlinear relationships, it falls short in modeling complex behavioral sequences and context dependencies. The model does not consider the temporal dynamics of user interests, nor does it explore more advanced structures such as attention mechanisms or graph neural networks, which may limit its performance in highly personalized recommendation tasks.

Future work can be extended in several directions. First, introducing the Transformer architecture is a promising direction. Transformers have demonstrated strong capabilities in sequential modeling by capturing global dependencies and fine-grained patterns in behavior sequences through multi-head self-attention mechanisms. Compared to traditional RNN or LSTM architectures, Transformers offer superior parallel computation and sequence representation capacity, enabling a more accurate understanding of the temporal dynamics of user interests [8]. Incorporating Transformers into recommender systems is expected to enhance context modeling of user history and produce more personalized recommendations. Second, combining large language models (LLMs) with multimodal information could significantly enhance model performance [9]. Movie data include not only structured ratings, but also rich textual descriptions, reviews, and visual data such as posters and stills. LLMs can deeply understand semantic information in reviews and plot summaries, enabling precise semantic modeling of user interests and movie characteristics. Meanwhile, multimodal learning—integrating text, image, and structured data—can compensate for the limitations of single-modality models, offering a more comprehensive depiction of preferences and item features, and improving both diversity and interpretability of recommendations. Third, exploring contrastive learning [10] and transfer learning [11] offers effective strategies for enhancing model generalization. Contrastive learning builds similar and dissimilar sample pairs to learn highly discriminative representations, particularly beneficial under data sparsity or weak supervision. Transfer learning enables knowledge from other domains or tasks to be adapted to movie recommendation, reducing reliance on large labeled datasets and improving performance under cold-start and cross-domain settings. The combination of these two techniques allows recommendation models to retain high accuracy and stability in low-resource or dynamic environments. Lastly, studying more efficient negative sampling strategies and dynamic user modeling is also of great significance. Current negative sampling relies on random selection, which may not accurately represent the boundaries of user preferences. Future approaches may adopt hard negative sampling or adversarial generation methods to create more informative and challenging negatives, thereby improving the model's discriminative power. In parallel, dynamic user modeling could better capture temporal variations in interest patterns, allowing for more personalized and timely recommendation outputs.

In summary, this study demonstrates the effectiveness of embedding techniques in personalized movie recommendation and offers a practical framework for building efficient recommender systems. With the ongoing advances in deep learning and multimodal modeling, future recommender systems are expected to exhibit stronger capabilities in user understanding and intelligence, delivering more accurate and richer recommendation experiences.

References

[1] A. Da’u and N. Salim, “Recommendation system based on deep learning methods: A systematic review and new directions, ” Artif. Intell. Rev., vol. 53, no. 4, pp. 2709–2748, Apr. 2020, doi: 10.1007/s10462-019-09744-1.

[2] M. Srifi, A. Oussous, A. Ait Lahcen, and S. Mouline, “Recommender systems based on collaborative filtering using review texts—a survey, ” Information, vol. 11, no. 6, p. 317, June 2020, doi: 10.3390/info11060317.

[3] S. A. Thorat, G. Ashwini, and M. Seema, “Survey on collaborative and content-based recommendation systems, ” in 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), Jan. 2023, pp. 1541–1548. doi: 10.1109/ICSSIT55814.2023.10061072.

[4] S. C. Mana and T. Sasipraba, “Research on cosine similarity and pearson correlation based recommendation models, ” J. Phys. Conf. Ser., vol. 1770, no. 1, p. 012014, Mar. 2021, doi: 10.1088/1742-6596/1770/1/012014.

[5] H. Daneshvar and R. Ravanmehr, “A social hybrid recommendation system using LSTM and CNN, ” Concurr. Comput. Pract. Exp., vol. 34, no. 18, p. e7015, 2022, doi: 10.1002/cpe.7015.

[6] F. M. Harper and J. A. Konstan, “The MovieLens datasets: History and context, ” ACM Trans Interact Intell Syst, vol. 5, no. 4, p. 19: 1-19: 19, Dec. 2015, doi: 10.1145/2827872.

[7] S. Li et al., “Embedding compression in recommender systems: A survey, ” ACM Comput Surv, vol. 56, no. 5, p. 130: 1-130: 21, Jan. 2024, doi: 10.1145/3637841.

[8] H. I. Pohan, H. L. H. S. Warnars, B. Soewito, and F. L. Gaol, “Recommender system using transformer model: A systematic literature review, ” in 2022 1st International Conference on Information System & Information Technology (ICISIT), July 2022, pp. 376–381. doi: 10.1109/ICISIT54091.2022.9873070.

[9] Z. Zhao et al., “Recommender systems in the era of large language models (LLMs), ” IEEE Trans. Knowl. Data Eng., vol. 36, no. 11, pp. 6889–6907, Nov. 2024, doi: 10.1109/TKDE.2024.3392335.

[10] Z. Liu, Y. Ma, Y. Ouyang, and Z. Xiong, “Contrastive learning for recommender system, ” Jan. 05, 2021, arXiv: arXiv: 2101.01317. doi: 10.48550/arXiv.2101.01317.

[11] J. Fu et al., “Exploring adapter-based transfer learning for recommender systems: Empirical studies and practical insights, ” in Proceedings of the 17th ACM International Conference on Web Search and Data Mining, in WSDM ’24. New York, NY, USA: Association for Computing Machinery, Mar. 2024, pp. 208–217. doi: 10.1145/3616855.3635805.