Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD24513

Interaction Behavior Prediction Method for Movie Recommendation Systems Based on Neural Collaborative Filtering

Shilin Wang

Wuhan Institute of Technology, Wuhan, China
2107767932@qq.com

Abstract. With the rapid advancement of internet technologies and the explosive growth of information, movie recommendation systems have become increasingly important in alleviating information overload and enhancing user experience. Traditional collaborative filtering and content-based recommendation methods face challenges in real-world applications, such as the cold start problem, data sparsity, and limited capacity to model nonlinear interaction relationships. This study proposes a movie interaction behavior prediction method based on Neural Collaborative Filtering, which leverages the strength of deep learning in modeling high-order nonlinear interactions. Using the publicly available MovieLens dataset, we constructed a binary classification task for user-movie interactions. Dense vector representations were learned through embedding layers, and multi-layer perceptrons were employed to model deep feature interactions. Dropout and Batch Normalization mechanisms were introduced to enhance model robustness. Using Hit Ratio@10 as the evaluation metric, experiments demonstrated that the proposed model achieved excellent performance in predicting user preferences, reaching a hit ratio of 0.74, significantly outperforming traditional recommendation methods while exhibiting strong stability.

Keywords: Movie Recommendation, Deep Learning, Neural Collaborative Filtering, Embedding

1. Introduction

In today's digital era, the rapid development of internet technologies and the explosive increase in data have introduced massive volumes of information in domains such as news, music, and entertainment into people's daily lives. While such information abundance greatly facilitates access and communication, it also causes serious information overload [1]. Faced with excessive and irrelevant content, users often experience anxiety and confusion during information filtering and decision-making processes, which negatively impacts both mental health and work efficiency. Traditional information retrieval and filtering technologies can alleviate information overload to some extent, but they often fail to meet users' dynamic and uncertain personalized needs, and are prone to triggering the Matthew effect. In this context, recommendation systems have emerged as an

DOI: 10.54254/2755-2721/2025.LD24513

intelligent bridge between users and information, rapidly becoming a key technology in the digital economy.

In the film industry, recommendation systems are particularly significant. As the number of film and television works continues to grow, audiences are often overwhelmed by excessive choices. Movie recommendation systems analyze user viewing behaviors, interest preferences, and rating data to provide intelligent personalized content recommendations. These systems not only enhance the user's viewing experience and increase engagement and loyalty, but also help content producers accurately identify market demand, enabling targeted marketing and box office revenue growth, thus promoting the healthy development of the film industry [2].

In recent years, recommendation systems have been widely applied in e-commerce and content distribution globally. For example, Amazon introduced item-based collaborative filtering as early as 1998, allowing for precise product recommendations based on users' browsing and purchase histories, dynamically updating recommendation lists according to new behaviors. Domestic platforms such as Taobao, Meituan, Douban, and Tencent Video also utilize user ratings, browsing records, and geolocation data to build efficient and precise recommendation frameworks, achieving notable success in enhancing user experience and platform performance.

Traditional recommendation methods include item-based collaborative filtering [3], user-based collaborative filtering [4], content-based recommendation, and hybrid approaches [5]. These methods have achieved certain success but also face limitations. Item-based collaborative filtering, which recommends items based on their similarity, features strong real-time performance and interpretability but struggles with cold start and limited diversity. User-based collaborative filtering identifies users with similar interests to generate recommendations, dynamically capturing preference changes, but still suffers from data sparsity and cold start issues. Content-based methods rely on item attributes (e.g., genre, cast, director) to recommend similar items. While effective during cold starts and interpretable, they often lack recommendation diversity and struggle with complex interest modeling. Hybrid systems combine multiple strategies to overcome the shortcomings of individual approaches, improving accuracy and robustness, but increasing system complexity, computational cost, and reducing interpretability.

Although traditional methods remain valuable in specific applications, their limitations in shallow feature extraction and linear modeling are increasingly evident as user needs become more complex and diverse. A technological breakthrough is needed to further enhance recommendation quality. Recently, the rise of deep learning [6] has revitalized recommendation systems through embedding techniques based on deep neural networks. These models can automatically learn dense low-dimensional representations for users and items, uncover complex patterns behind user behavior, and capture nonlinear, high-order interactions, thereby greatly improving personalization and accuracy. Additionally, deep learning allows for integration of multimodal data (text, images, sequences), enabling modeling and prediction of dynamic user interests—capabilities far beyond those of traditional methods.

Based on this context, this paper conducts in-depth research on a movie recommendation approach based on deep neural network embeddings, aiming to improve recommendation performance through advanced representation learning and interest modeling techniques. This work provides theoretical support and practical insights for enhancing user experience and driving intelligent development in the film industry.

DOI: 10.54254/2755-2721/2025.LD24513

2. Previous works

In recent years, recommendation system technologies have evolved rapidly, resulting in various mainstream approaches. Traditional methods primarily include user-based collaborative filtering, item-based collaborative filtering, content-based methods, and hybrid systems.

User-based collaborative filtering recommends items to a target user by identifying other users with similar interests . While this method is intuitive and easy to implement, it suffers from high computational complexity in large-scale systems and is sensitive to cold start and sparsity issues. Item-based collaborative filtering recommends items similar to those the user has previously rated. This method offers better stability and suits systems with extensive historical user behavior, but still faces challenges in sparsity and dynamic preference modeling. Content-based methods utilize item features (e.g., genre, cast, director) to model user preferences, recommending items with similar characteristics. Though effective with limited user data and interpretable, they often produce narrow recommendations lacking diversity. Hybrid methods combine multiple strategies, such as collaborative filtering and content-based recommendations, to mitigate the shortcomings of individual methods. They improve accuracy and robustness but at the cost of increased system complexity and computational demands.

In recent years, deep learning–based recommendation models have attracted significant attention. Compared to traditional methods, deep models offer stronger nonlinear modeling capabilities and can uncover complex latent feature relationships from user-item interaction data, thereby significantly improving recommendation performance. Multi-layer perceptrons (MLPs) stack multiple fully connected layers to model high-order nonlinear feature interactions, effectively capturing diverse user preferences. Convolutional Neural Networks (CNNs) extract local patterns in user behavior, useful for analyzing short-term interest shifts. Recurrent Neural Networks (RNNs) and their variants (e.g., LSTM, GRU) capture temporal dependencies in user behavior sequences, suitable for sequential modeling.

Among deep learning–based methods, Neural Collaborative Filtering (NCF) replaces traditional matrix factorization with neural networks, demonstrating strong performance in interaction prediction tasks. NCF introduces nonlinear transformation layers to model complex user-item relationships, surpassing the linear inner-product limitations and capturing richer interaction signals. Moreover, NCF supports integration of additional features (e.g., user attributes, contextual data), making it highly scalable. Across various evaluation metrics, NCF often outperforms traditional methods in accuracy and recall, especially suitable for fine-grained prediction of user behavior in movie recommendation scenarios. However, it has higher training overhead and limited interpretability, which warrants further investigation.

DOI: 10.54254/2755-2721/2025.LD24513

3. Dataset and preprocessing

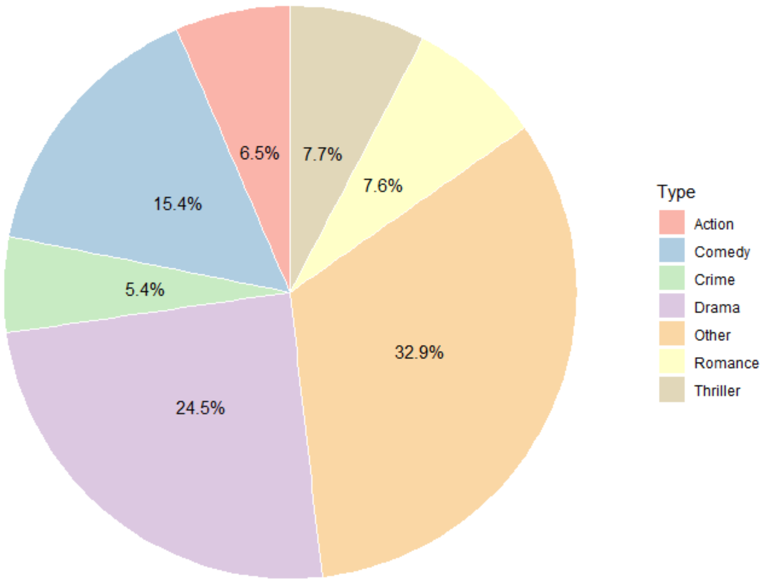

Figure 1: Movie genre distribution

The dataset used in this study is the widely recognized open-source MovieLens dataset, released by the GroupLens research group [7]. It contains extensive user rating information and movie metadata, and is commonly used in recommendation research. The dataset includes 27,278 movies across 20 genres (Figure 1), with 138,493 users contributing a total of 20,000,263 rating records (Figure 2).

During preprocessing, the raw rating data were loaded and the timestamp fields normalized to standard datetime format to enable time-series ordering and behavior modeling. Given the large dataset size, to reduce computational cost and improve training efficiency, 30% of users were randomly sampled, and all their rating records were retained. The final subset contained 41,547 users and 6,027,314 ratings, preserving behavioral diversity while controlling resource consumption.

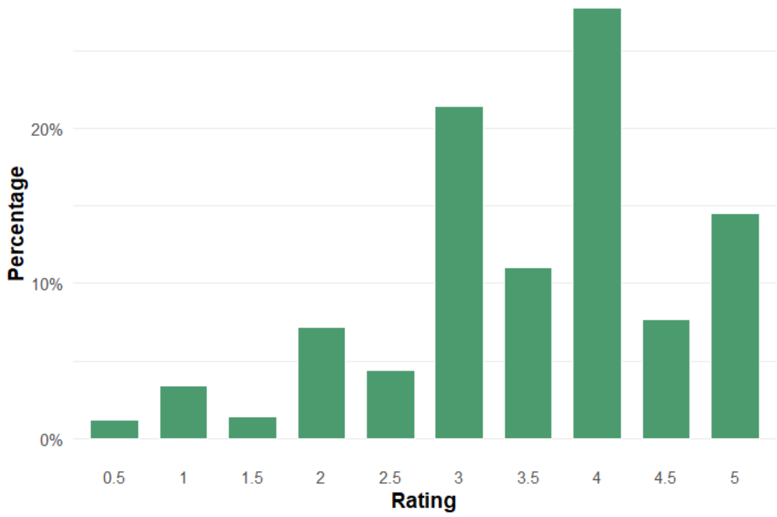

Figure 2: Rating distribution

DOI: 10.54254/2755-2721/2025.LD24513

To align with the real-world goal of predicting future behavior, we adopted a time-based train-test split. For each user, their ratings were sorted chronologically, with the latest rating used as the test instance and the rest as training data. This approach better reflects real-world user behavior prediction compared to random splits and avoids data leakage, ensuring rigorous evaluation. Only core features—user ID, movie ID, and rating—were retained, and redundant fields removed to reduce dimensionality and noise. Since many practical recommendation systems lack explicit ratings and rely on observed interactions, ratings were binarized into implicit feedback: 1 for interaction, 0 otherwise. This simplifies the modeling task and aligns with behaviors such as clicks and views.

To enhance the model's discrimination ability, negative sampling was applied. For each positive interaction, four negative samples (unrated movies) were randomly selected, forming user-movie non-interaction pairs. This expanded the training set and mitigated sparsity issues. All samples were formatted with user ID, movie ID, and interaction label for supervised training. Finally, the data were organized into batch-friendly structures suitable for deep learning, allowing efficient loading and iterative optimization during training to maximize performance while minimizing resource usage.

4. Model and results

Table 1: Model

	Layer Name	Type	#Params
0	user_embedding	Embedding	1 M
1	item_embedding	Embedding	1 M
2	fc1	Linear	2 K
3	fc2	Linear	8 K
4	fc3	Linear	2 K
5	fc4	Linear	528
6	output	Linear	17
7	dropout	Dropout	0
8	batch_norm1	BatchNorm1d	256
9	batch_norm2	BatchNorm1d	128
10	batch_norm3	BatchNorm1d	64

This study implemented an NCF-based movie recommendation model using the PyTorch Lightning framework to improve training efficiency and modularity. The architecture (Table 1) includes embedding layers for users and movies, multiple fully connected layers, Dropout, and Batch Normalization. A Sigmoid activation is used for binary classification of user-movie interaction. Given user and movie IDs, their dense embeddings are retrieved and concatenated, then passed through four fully connected layers with ReLU activation, batch normalization, and dropout, enhancing nonlinearity, stability, and generalization. The final Sigmoid output compresses predictions into the [0,1] range, representing interaction probability. The model was trained using Binary Cross Entropy Loss, optimized with Adam (learning rate = 0.001), batch size = 512, for a maximum of 5 epochs. Logging and checkpointing were disabled to reduce overhead. Training was accelerated using a single 2080Ti GPU.

DOI: 10.54254/2755-2721/2025.LD24513

For evaluation, the widely-used Hit Ratio@10 metric was adopted, which measures whether the ground truth item appears in the top-10 recommended list. For each test interaction, 99 unseen negative items were sampled, forming 100 candidates with the true item. The top-10 predictions were evaluated, and a “hit” was recorded if the true item appeared. Overall hit ratio was computed across all test samples, offering practical and meaningful assessment of real-world recommendation performance. The NCF model achieved a hit ratio of 0.74 on the MovieLens dataset, meaning it successfully recommended the correct movie 74% of the time on average. This indicates high predictive accuracy and robustness. Compared to traditional methods like similarity-based collaborative filtering or matrix factorization, NCF’s neural network architecture more effectively models nonlinear user-item interactions, overcoming issues of sparsity and cold starts. Moreover, the model demonstrated stability across user types. Active users with rich histories achieved higher hit ratios due to more informative embeddings, while even less active users received reasonable recommendations—highlighting the model’s generalization capability and cold start resilience.

5. Discussion and conclusion

This study developed a deep learning–based movie recommendation system using Neural Collaborative Filtering and validated it on the MovieLens dataset. Experimental results showed strong performance in Top-N recommendation tasks, achieving a Hit Ratio@10 of 0.74, demonstrating excellent user interest prediction. By incorporating embeddings, MLP architectures, Dropout, and Batch Normalization, the model achieved superior generalization and robustness. These findings confirm the potential of deep learning in modeling complex nonlinear interactions in recommendation systems, providing an effective approach for personalized and accurate movie recommendations.

Despite these achievements, several limitations remain. First, the study used only the MovieLens dataset, which, while representative, features well-structured data and focused user behavior that may not fully reflect the diversity of real-world scenarios. Second, the model still depends on historical interaction data, and cannot completely resolve the cold start problem. Additionally, the deep learning model operates as a "black box," lacking interpretability and potentially exhibiting biases toward active users or popular items.

Future research can explore several promising directions to further enhance model expressiveness, generalization, diversity, and personalization. Recommendation systems are evolving from collaborative filtering and shallow models to more intelligent deep learning architectures. Emerging AI technologies—such as large language models (LLMs) [8], graph neural networks (GNNs) [9], multimodal fusion, and transfer learning [10]—offer new possibilities. LLMs, trained via autoregressive or autoencoding strategies, have excelled in text understanding and generation. Integrating them into recommendation systems can improve understanding of unstructured data such as user reviews and movie descriptions. They can also support conversational recommendation via natural language interaction, improving interpretability and engagement. GNNs specialize in modeling graph-structured data, capturing complex relationships among users, items, genres, actors, and directors. By aggregating information across multi-typed heterogeneous graphs, they provide rich and structured embeddings. Multimodal techniques combine various data types—text, video, images, audio—into a unified learning framework. For instance, integrating visual and textual modalities helps capture movie styles and themes, aligning better with user aesthetic preferences. Finally, transfer learning allows knowledge transfer across domains or tasks, especially valuable in new-user, new-item, or cross-platform scenarios. Using pretrained models or fine-tuning

DOI: 10.54254/2755-2721/2025.LD24513

can reduce data requirements and development time. Combined with domain adaptation, self-supervised learning, or meta-learning, this can further improve cross-domain recommendation.

In summary, this study presents an NCF-based interaction prediction model for movie recommendation and verifies its effectiveness using a public dataset. Results suggest that deep learning significantly enhances modeling of user preferences. Future work will explore richer features and advanced architectures to further improve recommendation accuracy and generalization for real-world personalized applications.

References

[1] X. Zhang, Ding , Xiaoyan, and L. and Ma, “The influences of information overload and social overload on intention to switch in social media, ” Behaviour & Information Technology, vol. 41, no. 2, pp. 228–241, Jan. 2022, doi: 10.1080/0144929X.2020.1800820.
[2] H. Yuan and A. A. Hernandez, “User Cold Start Problem in Recommendation Systems: A Systematic Review, ” IEEE Access, vol. 11, pp. 136958–136977, 2023, doi: 10.1109/ACCESS.2023.3338705.
[3] M. Gao, Z. Wu, and F. Jiang, “Userrank for item-based collaborative filtering recommendation, ” Information Processing Letters, vol. 111, no. 9, pp. 440–446, Apr. 2011, doi: 10.1016/j.ipl.2011.02.003.
[4] L. H. Son, “HU-FCF: A hybrid user-based fuzzy collaborative filtering method in Recommender Systems, ” Expert Systems with Applications, vol. 41, no. 15, pp. 6861–6870, Nov. 2014, doi: 10.1016/j.eswa.2014.05.001.
[5] Z. Huang, C. Yu, J. Ni, H. Liu, C. Zeng, and Y. Tang, “An Efficient Hybrid Recommendation Model With Deep Neural Networks, ” IEEE Access, vol. 7, pp. 137900–137912, 2019, doi: 10.1109/ACCESS.2019.2929789.
[6] H. Zhou, F. Xiong, and H. Chen, “A Comprehensive Survey of Recommender Systems Based on Deep Learning, ” Applied Sciences, vol. 13, no. 20, Art. no. 20, Jan. 2023, doi: 10.3390/app132011378.
[7] 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.
[8] ChangYupeng et al., “A Survey on Evaluation of Large Language Models, ” ACM Transactions on Intelligent Systems and Technology, Mar. 2024, doi: 10.1145/3641289.
[9] G. Corso, H. Stark, S. Jegelka, T. Jaakkola, and R. Barzilay, “Graph neural networks, ” Nat Rev Methods Primers, vol. 4, no. 1, pp. 1–13, Mar. 2024, doi: 10.1038/s43586-024-00294-7.
[10] M. Gholizade, H. Soltanizadeh, M. Rahmanimanesh, and S. S. Sana, “A review of recent advances and strategies in transfer learning, ” Int J Syst Assur Eng Manag, vol. 16, no. 3, pp. 1123–1162, Mar. 2025, doi: 10.1007/s13198-024-02684-2.