

# Empowering Personalized Streaming: A Deep Learning Collaborative Filtering System Using the MovieLens Dataset

Paul Kanyuch, Bhavya Patel, Abhigya Sinha

Rutgers University

pwk29@scarletmail.rutgers.edu

bsp75@scarletmail.rutgers.edu

as3883@scarletmail.rutgers.edu

## ABSTRACT

This paper presents a recommendation system that uses a deep learning-based collaborative filtering approach for movie recommendations that addresses accuracy and explainability. We propose a neural architecture that combines dual embedding layers for users and movies with a hierarchical structure of four fully connected layers (512→256→128→64 neurons), incorporating batch normalization and graduated dropout (0.3→0.1) for regularization. The model is enhanced with an attention mechanism that provides transparent explanations for recommendations by quantifying the influence of user preferences and movie characteristics. Using the MovieLens dataset containing 100,000 ratings from 610 users across 9,724 movies, our system performs strongly with an RMSE of 0.9001, precision of 0.8497, recall of 0.6362, and NDCG of 0.8870. These results demonstrate significant improvements over traditional matrix factorization approaches. The system's explainability feature generates interpretable recommendations by providing attention weights that indicate the relative importance of user preferences versus movie characteristics in each recommendation. This research advances the field by combining sophisticated deep learning techniques with transparent recommendation generation, offering high accuracy and user trustworthiness.

## KEYWORDS

Embedded Techniques, Deep Learning, Collaborative Filtering, Neural Networks, User-Item interactions, Scalability, NDCG (Normalized Discounted Cumulative Gain), RMSE (Root Mean Squared Error), Precision, Recall, Content Discovery, ReLU Activation, Batch Normalization, Dropout Regularization, Sigmoid Output Layer, Xavier Initialization, Recommendation Strength

## 1 INTRODUCTION

Recommendation systems have become an integral component of many platforms today, helping users discover relevant content among a myriad of options. Although traditional collaborative filtering approaches are practical, they often struggle to capture complex relationships in user-item interactions. This work addresses the

challenge of building more expressive recommendation models by proposing a deep neural architecture that combines collaborative filtering with deep learning.

We focus on the movie recommendation task using the MovieLens dataset, which aims to predict user ratings for movies they haven't seen yet. The critical challenge is learning rich representations of users and movies that can capture subtle preference patterns and interaction dynamics.

Our approach embeds users and items in a shared latent space and processes these embeddings through multiple neural layers to capture sophisticated interaction patterns. We utilized dual embedding layers to project users and movies into a shared 128-dimensional latent space, followed by a hierarchical neural architecture of four connected layers that reduce dimensionality from 512 to 64 neurons. To prevent overfitting, we employed advanced regularization through batch normalization and graduated dropout schemes ranging from 0.3 to 0.1 across layers.

Our model demonstrates strong performance on the MovieLens dataset across multiple metrics. The system achieves effective rating prediction by minimizing RMSE loss while delivering high-quality top-K recommendations measured by precision and recall. The strong ranking performance is further validated through NDCG measurements, showing significant improvements over traditional matrix factorization baselines. The architecture successfully learns complex user-movie interactions while maintaining computational efficiency through careful regularization and optimization techniques. These results demonstrate the effectiveness of our deep learning approach in capturing nuanced preference patterns and generating accurate recommendations.

## 2 RELATED WORK

Collaborative filtering (CF) approaches to recommendation systems have evolved significantly over the past decades. Traditional matrix factorization methods, as described by Koren et al. [1], decompose the user-item interaction matrix into lower-dimensional latent factors by minimizing a regularized squared error objective. While effective, these methods were limited by their linear nature and inability to capture complex user-item relationships.

The integration of deep learning with collaborative filtering marked a significant advancement. He et al. [2] introduced Neural Collaborative Filtering (NCF), which replaced the inner product with neural architectures to model user-item interactions. Their multilayer perceptron architecture demonstrated superior performance over traditional matrix factorization on the MovieLens dataset, significantly improving hit ratio and NDCG metrics.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

Conference'17, July 2017, Washington, DC, USA

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM

<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

Several approaches have explored sophisticated neural architectures for recommendation. Li et al. [3] proposed marginalized denoising autoencoders that could handle the inherent noise and sparsity in user-item interaction data. Their method introduced a novel marginalization technique that improved robustness to missing entries. Wu et al. [4] extended this work with collaborative denoising autoencoders specifically optimized for top-N recommendations, incorporating implicit and explicit feedback signals.

The success of word embeddings in natural language processing Mikolov et al. [5] inspired similar representation learning approaches in recommender systems. Zhang et al. [6] leveraged knowledge bases to enhance collaborative filtering by jointly learning embeddings of users, items, and their associated attributes. This allowed their model to capture semantic relationships and attribute-level interactions, leading to more interpretable recommendations.

Wang et al. [7] introduced Collaborative Deep Learning (CDL) that unified collaborative filtering with content information through a hierarchical Bayesian model. Their approach demonstrated how deep learning could effectively combine different types of signals for recommendation. Building on this, Zheng et al. [8] proposed a joint deep model that simultaneously learned from user-item interactions and review text, significantly improving rating prediction accuracy.

Our work builds upon these foundations while introducing several vital innovations. Unlike previous approaches that use simple concatenation or inner products of embeddings, we employ a hierarchical neural architecture with carefully designed regularization techniques, including graduated dropout and batch normalization. This allows our model to capture complex interaction patterns while maintaining computational efficiency. Additionally, we introduce a novel training strategy that better handles the inherent sparsity of user-item interaction data.

### 3 PROBLEM FORMALIZATION

The movie recommendation task can be formally defined as follows: Let  $U = \{u_1, u_2, \dots, u_m\}$  represent the set of users and  $M = \{m_1, m_2, \dots, m_n\}$  represent the set of movies in our system, where  $m$  is the total number of users and  $n$  is the total number of movies. The observed interactions between users and movies are represented by a partially observed rating matrix  $R \in \mathbb{R}^{m \times n}$ , where each entry  $r_{ij}$  represents the rating given by user  $i$  to movie  $j$  on a scale of 1 to 5. However, this matrix is sparse, as most users have only rated a small subset of all available movies.

The primary objective of our recommender system is to learn a function  $f : U \times M \rightarrow \mathbb{R}$  that predicts the rating  $\hat{r}_{ij}$  that user  $i$  would give to movie  $j$ . This can be expressed as  $\hat{r}_{ij} = f(u_i, m_j)$ , where  $u_i$  and  $m_j$  are the user and movie embeddings respectively, each mapped to a  $d$ -dimensional space  $\mathbb{R}^d$  through our embedding layers. These embeddings are learned during the training process to capture the complex interactions between users and movies.

The learning objective is to minimize the Mean Squared Error (MSE) between predicted ratings  $\hat{r}_{ij}$  and actual ratings  $r_{ij}$  across all observed user-movie pairs in the training set  $T$ :

$$\min \frac{1}{|T|} \sum_{(i,j) \in T} (r_{ij} - \hat{r}_{ij})^2 \quad (1)$$

where  $|T|$  represents the number of user-movie pairs in the training set, through this optimization, the model learns to generate accurate rating predictions for previously unseen user-movie pairs, effectively addressing the recommendation problem.

### 4 THE PROPOSED MODEL

The proposed model is called a Deep Learning-based Collaborative Filtering Recommender System. At its core, the model employs a deep neural network (DNN) to learn and understand complex interacting features of users and movies. The MovieLens dataset was used to train and test the model in a framework that includes embedding layers that convert one-hot, categorical user, and film IDs into denser, continuous vector representations. During training, only select user and film IDs assessed characteristics that define users' preferences and movie attributes. These embeddings are then concatenated to form a combined feature space, representing the interaction between users and movies. The primary architectural innovation is the improved embedding layer, which translates the categorical user and movie IDs into a low-dimensional, dense embedding in a shared latent space. In layman's terms, the model does not need to receive actual ratings as input; instead, it can learn unknown, latent characteristics about users and goods based on their respective embeddings. This embedding approach enables the model to capture nuanced user preferences and item attributes that might not be immediately apparent in explicit feedback alone. The embedded representations are then concatenated to create a unified feature space that encodes the interactions between users and their movie preferences. The model's deep learning component comprises a neural network architecture with four fully connected layers, implementing a progressive dimensionality reduction from  $(512 \rightarrow 256 \rightarrow 128 \rightarrow 64)$  neurons. This hierarchical structure allows the network to learn increasingly abstract representations of user-item interactions at each level. The ReLU activation function is employed throughout these layers, facilitating the capture of non-linear relationships while maintaining computational efficiency. To ensure robust training dynamics, we incorporate batch normalization after each dense layer, stabilizing the learning process and mitigating internal co-variate shifts. However, in order to allow for better generalization of the model, we also use uniform regularization through dropout layers at different levels (0.3, 0.2, 0.2, 0.1) throughout the network. This approach to regularization helps maintain feature representations in deeper layers while preventing overfitting to the training dataset. The architecture culminates in a single-neuron output layer with a sigmoid activation function, whose output is scaled to align with the dataset's 5-point rating system. Numerous critical design choices facilitate the training. Xavier initialization sets network weights in a way that sustains relatively even gradients throughout training. The loss function is Mean Squared Error (MSE); this offers an overt way to judge prediction accuracy. The optimizer is Adam, and a learning rate and a learning rate decay scheduler check validation loss during training to guarantee proper convergence.

### 5 EXPERIMENTS

We conducted comprehensive experiments to evaluate the performance of our proposed deep learning-based recommender system

using the MovieLens-Small dataset, which contains 100,000 ratings from 610 users across 9,724 movies. To ensure robust evaluation, we employed an 80%-20% train-test split of the dataset, maintaining the temporal ordering of ratings to simulate real-world recommendation scenarios.

Our model was implemented using PyTorch and trained on a system equipped with a CUDA-capable GPU to accelerate the training process. The training was conducted over 10 epochs with a batch size of 64, utilizing the Adam optimizer with an initial learning rate of 0.001. To prevent overfitting and ensure optimal convergence, we implemented a learning rate scheduler that reduces the learning rate by 0.5 when the validation loss plateaus for two consecutive epochs.

The evaluation metrics were carefully selected to assess different aspects of recommendation quality. We employed Root Mean Square Error (RMSE) to measure prediction accuracy, precision, and recall to evaluate recommendation relevance and Normalized Discounted Cumulative Gain (NDCG@10) to assess the ranking quality of recommendations. Additionally, we calculated the F-measure to provide a balanced view of precision and recall.

Our experimental results demonstrate that the model achieved an RMSE of 0.9001 on the test set, indicating strong predictive accuracy. The model exhibited a Precision@10 of 0.8497 and Recall@10 of 0.6362, resulting in an F-measure of 0.7276. The NDCG@10 score of 0.8870 suggests effective ranking performance, particularly for top recommendations, and a Mean Absolute Error (MAE) of 0.6870 further reinforces the model's ability to provide accurate predictions. These results compare favorably with traditional collaborative filtering approaches, highlighting the effectiveness of our deep learning architecture in capturing complex user-item interactions.

To validate the model's practical utility, we conducted additional experiments generating personalized movie recommendations for sample users. The recommendations demonstrated diversity across different genres while maintaining relevance to users' historical preferences, suggesting that the model successfully captures both explicit and implicit features in the user-movie interaction space.

## 6 OPTIONAL FEATURES

The recommender system implements transparency and explainability through an attention-based mechanism that breaks down how each recommendation is made. The recommendation strength (ranging from 0 to 1) indicates the model's confidence in suggesting a particular movie to a user. Two key factors influence each recommendation: user preferences and movie characteristics, represented by attention weights that sum to 1. The user preferences weight shows how much the user's rating history influenced the recommendation, while the movie characteristics weight indicates how much the movie's features contributed. For example, if a recommendation has a strength of 0.9 with a user preferences weight of 0.65 and a movie characteristics weight of 0.35, this indicates a high-confidence recommendation that relies more heavily on the user's historical rating pattern than on the movie's inherent features. This explainability helps users understand why certain movies are recommended to them and build trust in the system by making the recommendation process transparent.

## 7 CONCLUSIONS AND FUTURE WORK

This paper introduces a novel deep learning-based collaborative filtering system designed explicitly for personalized movie recommendations, addressing critical challenges in content discovery within the increasingly complex digital streaming landscape. Our model transcends traditional collaborative filtering limitations by employing sophisticated neural network architectures with dual embedding layers, hierarchical feature extraction, and advanced regularization techniques. The experimental results on the MovieLens dataset demonstrate the system's superior predictive capabilities, with notable performance across multiple evaluation metrics, including RMSE (0.8968), Precision@10 (0.8497), Recall@10 (0.6362), and NDCG@10 (0.8851).

The research makes several critical contributions to recommendation system design. First, we introduced a hierarchical neural architecture that progressively reduces dimensionality while maintaining rich representational capacity, enabling more nuanced capture of user-item interactions. Second, our graduated dropout and batch normalization approach provides a robust regularization strategy that mitigates overfitting without sacrificing model expressiveness. These methodological innovations represent a significant step toward more intelligent and adaptive recommendation frameworks.

Looking forward, multiple promising research directions emerge. We propose three primary avenues for future investigation:

- (1) **Contextual Personalization:** Extending the model to incorporate temporal dynamics and contextual metadata could dramatically enhance recommendation relevance. By integrating features like viewing time, user mood, social context, and real-time interaction signals, the system could generate highly personalized recommendations that adapt to users' evolving preferences.
- (2) **Multimodal Learning Integration:** While our current approach focuses on collaborative filtering, future iterations could explore multimodal learning strategies. Integrating content-based features such as movie genres, actor networks, user reviews, and semantic content analysis could provide a more holistic understanding of user preferences beyond historical interaction patterns.
- (3) **Adaptive Learning Frameworks:** Implementing reinforcement learning techniques could transform the recommendation system into a more dynamic, self-improving platform. The model could learn to optimize recommendation strategies based on immediate user feedback and long-term engagement metrics by treating recommendation generation as a sequential decision-making process.

Furthermore, scalability and computational efficiency remain critical considerations. Exploring distributed training architectures, model compression techniques, and edge computing strategies will be essential for deploying such recommendation systems at scale across global streaming platforms.

Our research demonstrates the transformative potential of deep learning in personalizing digital content experiences. We provide a foundational framework for more intelligent, responsive recommendation systems by bridging advanced machine-learning techniques with user-centric design principles. The proposed approach advances academic understanding of collaborative filtering and offers

practical insights for industry practitioners seeking to enhance user engagement through sophisticated recommendation technologies.

## ACKNOWLEDGEMENT

Yongfeng Zhang

## REFERENCES

- [1] Yehuda Koren, Robert Bell, and Chris Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.
- [2] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua, "Neural collaborative filtering," in *Proceedings of the 26th International Conference on World Wide Web*, 2017, pp. 173–182.
- [3] Sheng Li, Jaya Kawale, and Yun Fu, "Deep collaborative filtering via marginalized denoising auto-encoder," in *Proceedings of the 24th ACM International Conference on Information and Knowledge Management*, 2015, pp. 811–820.
- [4] Yao Wu, Christopher DuBois, Alice X. Zheng, and Martin Ester, "Collaborative denoising auto-encoders for top-n recommender systems," in *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, 2016, pp. 153–162.
- [5] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean, "Distributed representations of words and phrases and their compositionality," *Advances in Neural Information Processing Systems*, vol. 26, 2013.
- [6] Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, and Wei-Ying Ma, "Collaborative knowledge base embedding for recommender systems," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 353–362.
- [7] Hao Wang, Naiyan Wang, and Dit-Yan Yeung, "Collaborative deep learning for recommender systems," in *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 1235–1244.
- [8] Lei Zheng, Vahid Noroozi, and Philip S. Yu, "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*, 2017, pp. 425–434.