Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD29735

Comparision of Recommenders System

Muye Chen

Northwestern University, Evanston, USA
muyechen2026@u.northwestern.edu

Abstract. Today, video games have already become a mainstream way of entertainment. There are a lot of video games with different types online. In this case, a good recommender system that can help customers find what they want is important for the platform to attract users. This is because sometimes users may not have heard of the game that they actually like. In this case, showing the game that users like on the homepage is important, so the user will purchase the game and the platform can make revenue. This paper compared the performance of two widely used recommender systems: Collaborative filtering and a neural network (GCN). This data is from Kaggle, which is statistics from Steam, since it's the most popular platform for people to download PC video games. So its data may be more representative and can avoid some bias caused by the data. According to this research, if the training data is small, which means most of the user relations in the test dataset are not included in the training dataset, then Collaborative filtering may have a better performance than the neural network.

Keywords: Collaborative filtering, neural network, Steam, GCN, recommender system

1. Introduction

This project is about using collaborative filtering and neural networks to make a recommender system and compare their performance on Steam video games. This project uses the dataset from Kaggle, which includes more than 40 million comments and is larger than 2 GB. Because of the limitations of the hardware equipment, especially the space of memory, this project is not based on all the data but sampling from the dataset. All the training and testing are based on the sample chosen from the dataset. So the reduction of the amount of data may especially affect the performance of the model. For the result, the precision of the collaborative filtering method is about 0.017 for the top 20 items. And the precision of the neural network method is about 0.015.

Today, more and more people own personal computers, and game technology is more and more perfect; computer games have become the most popular way of entertainment and leisure in people's daily lives, especially for most young people. "In 2024, the revenue from the worldwide gaming market was estimated at almost 455 billion U.S. dollars [1]." However, with the continuous publishing every year, the number of games available for purchase on the market is extremely high. Therefore, players are undoubtedly unable to browse through all the games in sequence and find the most suitable ones. So enabling players to easily see the games they are interested in on the home page and download them can definitely enhance the competitiveness of the game download

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD29735

platform. Therefore, a reliable and effective recommendation system, as a tool to suggest items of interest to users from a vast array of different types, is indispensable [2,3]. Collaborative filtering and neural networks are today's most common and reliable recommendation algorithms.

Collaborative filtering appears earlier than neural networks, but neural networks have shown superior performance in different fields in recent years. In this case, this study will attempt to compare the performance of these two different recommendation algorithms for the same data set. The recommendation system is not only applicable to game platforms, but also has used on other platforms such as online shopping platforms and short-video platforms. However, this research will focus on the recommendation system for games based on the Steam platform. The reason for choosing the Steam platform is that it is very famous and holds a monopolistic position in the field of game platforms, so its data is representative of the field of video games. It is the first choice for most people to purchase and download games. It's also because it includes most of the video games, which can avoid the deviation caused by the insufficient amount of data.

2. Methodology

This experiment mainly compared the performance of collaborative filtering and neural networks.

Collaborative filtering: Collaborative filtering is a widely used recommendation algorithm in recommendation systems and an effective way to suggest products [4,5]. It can make recommendations based on user similarity and item similarity. "User-based collaborative filtering methods were introduced in 1992, while item-based collaborative filtering methods emerged in 1999 [6]." This research employed user-based recommendation, which evaluates items using others' opinions [7]. This way assumes that users with similar behaviors will make similar choices [8]. "For instance, if both the first and second users have previously made analogous item selections, and the first user subsequently acquires a new product, such as a smartphone, which the second user has not yet encountered or acquired, the role of CF is to propose this previously undiscovered item to the second user [9]." This recommendation algorithm will select the most similar users based on the similarity between the target user and other users and determine the recommendation degree of the target user according to the recommendation degree of these similar users to the target item. To achieve this goal, it is necessary to calculate the cosine similarity first and convert it into a data frame for the subsequent operations. "The similarity between two documents can be measured by treating each document as a vector of word frequencies and computing the cosine of the angle formed by the frequency vectors. This formalism can be adopted in collaborative filtering, which uses users or items instead of documents and ratings instead of word frequencies. Formally, if R is the m × n user-item matrix, then the similarity between two items, i and j, is defined as the cosine of the n dimensional vectors corresponding to the ith and jth column of matrix R. Vector cosine similarity between items i and j is given by

$$w_{i,j} = \cos\left(\overrightarrow{i,j}\right) = \frac{\overrightarrow{i} \cdot \overrightarrow{j}}{\left\|\overrightarrow{i}\right\| * \left\|\overrightarrow{j}\right\|}$$

, where "•" denotes the dot-product of the two vectors [10]." If the user does not exist within the dataframe of similarity, that is to say, the user is not included in the training set, the algorithm will recommend using the average score of this game or the average score of this user for other games,

DOI: 10.54254/2755-2721/2025.LD29735

depending on which score exists in the training set. If none of the above data exists, the expected rating of the user for the game will be 0. This experiment adopts the user-based top-N recommendation method; this way, we will first sort some users who are most similar to the user we will predict. Then, the users' ratings for all of the items will be used to generate the predicted score for all of the items for the user we are trying to predict. Then we will pick the top n items, for which we have the highest predicted score, and compare these n items to the real n items that the user really prefers.

Neural network: The neural network used in this study consists of the Input layer, an Embedding layer, a flatten layer, a Concatenate layer, a Dense layer, and a dropout layer. Early stop was used during compilation, with a batch size of 8192 and epochs of 64. The optimizer of the neural network is “Adam” and the loss is “mean squared error”. The Dense layer uses “relu” as the activation function and uses L2 regularization. This study used GCN as the neural network. “Generally speaking, graph convolutional network models are a type of neural network architectures that can leverage the graph structure and aggregate node information from the neighborhoods in a convolutional fashion. Graph convolutional networks have a great expressive power to learn the graph representations and have achieved a superior performance in a wide range of tasks and applications [11].” GCN is a subclass of GNN, the GNN is a new type of neural network that gets features from the data, is non-Euclidean space, and does convolution using a scalable kernel on this data [12]. In recent years, GNN has been the advanced way to solve recommendation problems with its ability to discover high-order information and deal with structured data [13]. Compare to the traditional method, “GNN can naturally and explicitly encode the crucial collaborative signal (i.e., topological structure) to improve the user and item representations [14].”

3. Experiment and results

After sampling the number of users by 0.0001, we used 8537 users’ data, among which 20% was the test dataset and 80% was the training dataset. In this case, the precision of the top 20 in the final collaborative filtering is about 0.017, and that of the neural network is approximately 0.015. The RMSE for the collaborative filtering is about 0.38 and 0.3 for the neural network. The accuracy rate of the results of this experiment is relatively low. This might be due to the small sample size, which causes many users in the test set to not be in the training set, which makes it impossible to calculate the similarity. Therefore, using methods such as reducing batch size to reduce memory usage to allow the training of larger datasets may improve the accuracy of the recommendation system. Also, there are more type of recommender systems appear. For example, in the Deep Learning area, beyond GNNs, there are also RNNs and Transformers, which prompt the development of recommender systems [15]. In this case, we can compare the performance of more models in the future.

References

[1] Clement, Jessica. “Topic: Video Game Industry.” Statista, 6 Nov. 2024, www.statista.com/topics/868/video-games/#topicOverview.
[2] Burke, Robin, Alexander Felfernig, and Mehmet H. Göker. "Recommender systems: An overview." Ai Magazine 32.3 (2011): 13-18.
[3] Ekstrand, Michael D., John T. Riedl, and Joseph A. Konstan. "Collaborative filtering recommender systems." Foundations and Trends® in Human–Computer Interaction 4.2 (2011): 81-173.
[4] Al-Bashiri, Hael, et al. "Collaborative filtering recommender system: Overview and challenges." Advanced Science Letters 23.9 (2017): 9045-9049.

DOI: 10.54254/2755-2721/2025.LD29735

[5] Zhou, Hongde, Fei Xiong, and Hongshu Chen. "A comprehensive survey of recommender systems based on deep learning." Applied Sciences 13.20 (2023): 11378.
[6] Aljunid, Mohammed Fadhel, et al. "A collaborative filtering recommender systems: Survey." Neurocomputing 617 (2025): 128718.
[7] Schafer, J. Ben, et al. "Collaborative filtering recommender systems." The adaptive web: methods and strategies of web personalization. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. 291-324.
[8] Li, Yang, et al. "Recent developments in recommender systems: A survey." IEEE Computational Intelligence Magazine 19.2 (2024): 78-95.
[9] Ibrahim, Osman Ali Sadek, et al. "Revisiting recommender systems: an investigative survey." Neural Computing and Applications 37.4 (2025): 2145-2173.
[10] Su, Xiaoyuan, and Taghi M. Khoshgoftaar. "A survey of collaborative filtering techniques." Advances in artificial intelligence 2009.1 (2009): 421425.
[11] Zhang, Si, et al. "Graph convolutional networks: a comprehensive review." Computational Social Networks 6.1 (2019): 1-23.
[12] Liang, Fan, et al. "Survey of graph neural networks and applications." Wireless Communications and Mobile Computing 2022.1 (2022): 9261537.
[13] Gao, Chen, et al. "Graph neural networks for recommender system." Proceedings of the fifteenth ACM international conference on web search and data mining. 2022.
[14] Wu, Shiwen, et al. "Graph neural networks in recommender systems: a survey." ACM Computing Surveys 55.5 (2022): 1-37.
[15] Ren, Xubin, et al. "A comprehensive survey on self-supervised learning for recommendation." ACM Computing Surveys (2024).