IJARCCE
ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 10, October 2021
DOI: 10.17148/IJARCCE.2021.101016

A Comparative Study on the Different Approaches Used for a Recommendation System for Games

Justin Janz[1], Deepa V Jose[2]

Department of computer science, CHRIST(Deemed to be University), Bangalore, India[1]

Department of computer science, CHRIST(Deemed to be University), Bangalore, India[2]

Abstract: Recommender systems can be found in almost every domain in the current world. It is a multidisciplinary field, utilizing data mining and machine learning and some other similar techniques as per the domain. Be it a shopping site, media streaming platforms, while navigating with Google maps or even booking an appointment. In the current world of overloaded technology, users are bombarded with recommendations where ever one goes. Here the focus is on a game recommending system which suggests its users what game to buy next. The different approaches used for recommending games for a particular user is compared and contrasted. We see how the approaches have their own perks and losses. We take a look at the content-based filtering approaches for a game recommendation system and a collaborative filtering system. Also gives a closer look at a deep learning system to see if that bridges the gap between the content-based and collaborative approaches.

Keywords: Recommender systems, content-based filtering, collaborative filtering, hybrid methods,Deep learning, reinforcement learning.

I. INTRODUCTION

Recommender Systems are dynamic frameworks; they become familiar with a user profile and continues refreshing it as per the input. When the user collaborates with the framework, it learns his preferences and assembles a profile, and later it utilizes this profile to deliver the suggestions. Recommender systems are used for providing recommendations to users on the basis of their taste preferences and past usage history, or simply we can say we use it to predict what the user will like. Usually, these are found in the E-Commerce industry like in popular websites such as Amazon, Flipkart Myntra etc. These systems have now spread to many other domains such as in video streaming platforms like Netflix and YouTube, Music streaming apps such as Spotify and Amazon music and similar areas where some kind of e-commerce is used. It has very scarcely been used in the gaming industry even though we are getting many online platforms for games like Steam, Origin, Uplay and other video game distribution service by value. With the ever-increasing popularity and the number of games released finding the game that a user might like is harder; hence the requirement for the recommender system is a requirement.
Video games are one of the world's largest entertainment industry. Video gamers collectively spend 3 billion hours per week in front of their screens. It alone gives us an idea of how the popularity of games has increased. In the past, computer games have been considered as a distraction from more 'worthy' activities, such as homework or playing outside. Now studies have shown that video games are not only the source of entertainment but have shown to improve and develop, some beneficial real-life skills for people like developing their mind to think better and also increases their problem-solving ability.
Recommender system plays a very major role in personalizing the experience of people using the above-mentioned game distribution services. One of the most common techniques used for building a recommendation system is collaborative filtering (CF) where we find out a subset of users who have similar interests to that of a target user and make suggestions to the user[1]. The subset of user can range from family and friends to anybody who has similar interest as the user. Another approach is the content-based filtering recommender systems which give predictions based on the content rather than ratings by a user. It uses the item-to-item correlation for making recommendations. It has one of the major drawbacks as the filtering and recommendation is limited to the initial recommendation of the item. Then there is the demographic approach where the system assumes that all the users of same age group or gender, in a general demographic group, have the same preferences which itself is one of its major drawbacks. A hybrid recommendation system is a combination of two or more of the mentioned recommendation systems[2]. A higher form of a hybrid system is the context-based recommender system which takes into consideration the dynamic aspects of the environment that it is deployed in like the location, mood of the user, Season etc[3].

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 10, October 2021

DOI: 10.17148/IJARCCE.2021.101016

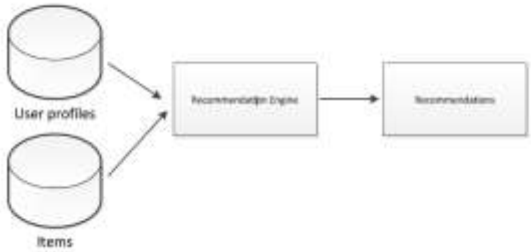

Fig 1: Overall Look of a recommender system

There are researches going on to improve the prediction of recommender systems in other platforms such as video streaming and e-commerce websites but not a lot of active work on the Video game distribution service platforms which is currently there. The integration of recommendation systems into video games is a relatively new area of research. While there are several approaches to the problem of developing recommendation systems, we will explore more in the lines of how to improve the prediction accuracy of video games distribution systems by introducing a new feature. This research work is done to perform a comparative analysis on the various strategies used to improve a recommender system using a collaborative filtering approach.

II. RELATED WORK

Research on recommender systems started in the ’90s and increased with Netflix Prize competition. Recommender systems are a specific type of information filtering systems that rank the existing items based on users’ preferences[4]. Recommender systems are profitable for both customers and service providers. Typically, recommender systems use customers’ action history to learn users’ personality and preferences and then it gives feedback to recommend relevant items back to the user according to that. So far, most of the researches have focused on explicit feedbacks of users, such as like and dislike and have used only discrete prediction metrics. Unfortunately, many users do not participate in rating items while they implicitly show their opinion. In order to produce a list of interesting items to a user as a suggestion or a recommendation, Recommendation systems should first predict that a specific item is worth recommending for a specific user. This prediction can be obtained by following multiples strategies, which classifies a recommender system. Methods can be divided into three main groups:

- content-based filtering
- collaborative filtering
- Hybrid methods

Deep learning and reinforcement learning methods have been used too. Content-based filtering methods use the items’ features and users’ history of actions to learn users’ preferences[5]. Recommender systems using this Content-based filtering recommendation strategy, analyse a set of documents (features) of the items rated by the user and then create the user profile based on the features of the objects rated by that user. It basically tries to match up the features of the item against the user profile. The best possible matches are included in the recommendations. The recommendation process is performed in three steps, each of which is handled by a separate component: content analyser, profile learner and the filtering component [6].
Content Analyzer: Pre-processes the data to be used in the succeeding step.
Profile learner: Constructs the user profile based on the content analysers output.
Filtering component: Produces the recommendations based on the profiles generated.

Collaborative filtering methods use a user-item-rate matrix to predict the rating of items for each user. Collaborative filtering methods can be divided into two main groups: model-based and memory-based methods. In the model-based methods, the recommendation is based on a model that will be learned in a learning process [7]. Memory-based methods use the KNN algorithm to find similar users/items. Originally designed as basic memory-based methods, collaborative filtering has evolved into model-based methods that commonly involve machine learning techniques, such as matrix factorization, probabilistic models [8], and deep neural networks. Feature dependency and train data limitation are some of the drawbacks for the content-based methods. Cold start problem and data sparsity are drawbacks for the collaborative filtering methods. The hybrid method is the combination of content-based and collaborative filtering methods that can address some of these problems[9].
Cross-Domain Recommender Systems with Side Information: In this category, it is assumed that some side information about the entities is available. Collective matrix factorization (CMF) is designed for scenarios where a user-item rating matrix and an item-attribute matrix for the same group of items are available[10].

IJARCCE
ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 10, October 2021
DOI: 10.17148/IJARCCE.2021.101016

Cross-Domain Recommender Systems with Non-overlapping Entities: This category covers the methods that handle two domains with non-overlapping entities and transfer knowledge at the group level. Users and items are clustered into groups, and knowledge is shared through group-level rating patterns[1].

Text-reviews of the games were used for games recommendation to concerned users [9]. Pair of words were used containing adjectives and context words; the phenomenon is known as information-theoretic co-clustering, thus lessening the dimensionality of vectors. Jose P. Zagal et al. [10] emphasized on the usage of the term gameplay in user-submitted game reviews on various websites. They filtered the adjectives that altered gameplay using Natural Language Processing (NLP).

Hybrid Recommender systems are a blend of the above-mentioned techniques. A hybrid system combining techniques A and B tries to use the advantages of A to fix the disadvantages of B. For instance, Collaborative Filtering methods suffer from new-item problems, i.e., they cannot recommend items that have no ratings[11]. This does not limit content-based approaches since the prediction for new items is based on their description (features) that are typically easily available. So, combining these two should theoretically produce a better result.

III. METHODOLOGY

Assessing a recommender system is no simple task. To achieve this feat of comparing different collaborative filtering approaches to recommendation systems, some variety of methods were used. Different aspects about a model were tested and can be generalized into how good the system recommended games, the performance of the system in terms of recommending users. First, we used different methods available on the steam 200k dataset the below models.

Alternate Least Squares: The general idea of matrix factorization is that we are given a large user-item matrix, R, where each cell represents the user's implicit feedback on the item. This could be whether the user interacted with the item, how many times the user interacted with the item, or how long the user interacted with the item for, depending on the dataset. In our case, it was how much time the user spent playing the game. This matrix is then factored out into a user factor X and an item factor Y[12].

Bayesian Personalized Ranking: BPR is different from most collaborative filtering models in that, rather than finding relationships between users and items. If a user has interacted with item 1 and not item 2, we say that the user prefers item 1. If the user has interacted with both or neither, we cannot say whether the user prefers one over the other. A Bayesian formulation is then used in order to optimize the parameter vector[13].

Logistic Matrix Factorization: As another matrix factorization model, the setup for LMF is very similar to that of ALS: we factor out a matrix R into user factor X and item factor Y. The big difference here is that we are trying to calculate the probability that user u interacts with item i[14].

IV. EXPERIMENTAL SETUP

A game data set containing 345667 games on over 50 platforms, including mobiles, was used to test the different methods. RAWG is the largest video game database in the world with 300,000+ titles, 2M screenshots, and 425,000 user ratings[11]. It can be considered the IMDb for games. Each row contains information about one game. There are several features, namely Game Title, Description, platforms, genres, rating, price, hours played etc. It contained instances that provided how long a Steam user played a specific game. We used this data to train and 2 compare each of the models mentioned above.

V. RESULTS AND DISCUSSIONS

There were two metrics that we tracked to compare the four models. The first consisted of having each of the models recommend five games per user. The recommendation was considered successful if the game that the user played in the test set was one of the five games that the model recommended for that user. The percent of users that had their game successfully predicted in the test set were then calculated. The second metric was to measure the diversity of the games that each of the models were recommending. We measured this by taking the number of unique games that the model recommended to all of the users and divided it by the number of unique games in the training dataset. Each model had parameters that were fine-tuned to ensure that the most accurate results were being acquired.

All of the models performed better than the baseline, most of them by a very large margin. The ALS model was able to accurately predict a game that a user played from five recommended games for 44.23 % of the users. It was also able to recommend 89.5% of the games that were in the training set. The BPR model was able to accurately predict a game that a user played from the five recommended games for 45.34% of the users. It was also able to recommend 95.3% of the games in the training set. Again, this was much better than the baseline, but this model also performed slightly better than ALS for accuracy and significantly better in percentage of games recommended. The LMF model accurately predicted a game that a user played from the five recommended games for only 28.44% of the users and only

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 10, October 2021

DOI: 10.17148/IJARCCE.2021.101016

recommended 45.68% of the games in the training set. This model did not perform significantly better than the baseline, and it performed significantly worse than the other models. The results are mentioned Table 1.

Table 1: Evaluation Results

Method	Evaluation 1	Evaluation 2
Alternate Least Squares	44.23%	89.5%
Bayesian Personalized Ranking	45.34%	95.3%
Logistic Matrix Factorization	28.44%	45.68%

CONCLUSIONS

After running tests with all of the models, we determined that the Bayesian Personalized Ranking model was the most accurate followed closely by and Alternating Least Squares, with Logistic Matrix Factorization far behind. In terms of percentage of games in the training set recommended by the model, BPR performed the best by a wide margin, leading ALS by almost close to 6%, with LMF again far behind.

There is more that can be done further, one aspect of the data cleaning that could be looked into is the issue of overlapping game names. For example, there is one listing for Call of Duty: Black Ops, and one for Call of Duty: Black Ops - Multiplayer. In the future purchase data should also be factored into the training set. The original data set had a behaviour column of either Purchase or Play, where Purchase indicated that the user had bought the game but had never played it. These rows were taken out and only considered games that users had played, but there may be some way to factor the Purchase behaviour into our calculations.

REFERENCES

[1] A. Pujahari and V. Padmanabhan, "Group Recommender Systems: Combining User-User and Item-Item Collaborative Filtering Techniques," *Proc. - 2015 14th Int. Conf. Inf. Technol. ICIT 2015*, pp. 148–152, 2016, doi: 10.1109/ICIT.2015.36.
[2] S. Sharma, A. Sharma, Y. Sharma, and M. Bhatia, "Recommender system using hybrid approach," *Proceeding - IEEE Int. Conf. Comput. Commun. Autom. ICCCA 2016*, pp. 219–223, 2017, doi: 10.1109/CCAA.2016.7813722.
[3] S. Jain, A. Grover, P. S. Thakur, and S. K. Choudhary, "Trends, problems and solutions of recommender system," *Int. Conf. Comput. Commun. Autom. ICCCA 2015*, pp. 955–958, 2015, doi: 10.1109/CCAA.2015.7148534.
[4] T. Ruotsalo, *Methods and Applications for Ontology-Based Recommender Systems*. 2010.
[5] P. Wei, S. Xia, R. Chen, J. Qian, C. Li, and X. Jiang, "A Deep-Reinforcement-Learning-Based Recommender System for Occupant-Driven Energy Optimization in Commercial Buildings," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6402–6413, 2020, doi: 10.1109/JIOT.2020.2974848.
[6] P. Lops, M. De Gemmis, and G. Semeraro, *Recommender Systems Handbook*. 2011.
[7] P. Bertens, A. Guitart, P. P. Chen, and A. Perianez, "A Machine-Learning Item Recommendation System for Video Games," *IEEE Conf. Comput. Intell. Games, CIG*, vol. 2018-Augus, 2018, doi: 10.1109/CIG.2018.8490456.
[8] C. Yu, Q. Tang, Z. Liu, B. Dong, and Z. Wei, "A Recommender System for Ordering Platform Based on an Improved Collaborative Filtering Algorithm," *ICALIP 2018 - 6th Int. Conf. Audio, Lang. Image Process.*, pp. 298–302, 2018, doi: 10.1109/ICALIP.2018.8455852.
[9] G. Zhang and X. Zhou, "Similarity Based Matrix Factorization for Recommender Systems," *Proc. - 2017 10th Int. Symp. Comput. Intell. Des. Isc. 2017*, vol. 1 2018-Jan, pp. 7–11, 2017, doi: 10.1109/ISCID.2017.22.
[10] A. Zigomitros, A. Papageorgiou, and C. Patsakis, "A practical k-anonymous recommender system," *IISA 2016 - 7th Int. Conf. Information, Intell. Syst. Appl.*, 2016, doi: 10.1109/IISA.2016.7785379.
[11] A. S. B. A. Kamal, S. Saaidin, and M. Kassim, "Recommender System: Rating predictions of Steam Games Based on Genre and Topic Modelling," *2020 IEEE Int. Conf. Autom. Control Intell. Syst. I2CACIS 2020 - Proc.*, no. June, pp. 212–218, 2020, doi: 10.1109/I2CACIS49202.2020.9140194.
[12] Y. Hu, C. Volinsky, and Y. Koren, "Collaborative filtering for implicit feedback datasets," *Proc. - IEEE Int. Conf. Data Mining, ICDM*, pp. 263–272, 2008, doi: 10.1109/ICDM.2008.22.
[13] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "BPR: Bayesian personalized ranking from implicit feedback," *Proc. 25th Conf. Uncertain. Artif. Intell. UAI 2009*, pp. 452–461, 2009.
[14] C. Johnson, "Logistic matrix factorization for implicit feedback data," *Adv. Neural Inf. Process. Syst.*, vol. 27, pp. 1–9, 2014.

