

Anime Recommendation using Item Based Collaborative Filtering and Content Based Filtering

Aidan Eka Prasajo
Data Science Major
Binus University
Jakarta, Indonesia
aidan.prasajo@binus.ac.id

Narendra Nusantara Handradika
Data Science Major
Binus University
Jakarta, Indonesia
narendra.handradika@binus.ac.id

Stavanus Febrian
Data Science Major
Binus University
Jakarta, Indonesia
stevanus.febrian@binus.ac.id

Abstract—This paper proposes a two-anime recommendation system that consist of collaborative filtering and content-based filtering methods. Item-based collaborative filtering identifies preference similarities between users to provide recommendations, meanwhile content-based filtering builds a profile for each item and compares them with the user's preference to recommend suitable items. These methods complement each other, with collaborative filtering uncovering unexpected preferences and content-based filtering tackling the cold start problem. Moreover, the system's recommendations are evaluated by comparing them with user preferences to assess their effectiveness. By comparing item-based collaborative filtering and content-based filtering approaches, we can analyze which system model can provide accurate and personalized anime recommendations, improving user satisfaction.

(Abstract)

Keywords—recommendation system, collaborative filtering, item-based collaborative filtering, anime

I. INTRODUCTION (HEADING 1)

Anime is a widely recognized animation style from Japan, typically created through hand-drawn or computer-generated techniques. It encompasses a diverse range of genres including romance, action, fantasy, horror, and comedy, among others. Animes are frequently based on literary works such as manga or light novels, although they can also be original productions. These factors contribute to the immense popularity of Anime, making it a prominent and influential part of Japan's entertainment industry.

Anime recommendation is an interesting theme to explore due to its massive popularity worldwide. There are many reasons why someone might like an anime, encompassing complex plot, good character building, and beautiful visuals to name a few. However, as the years go by, anime continues to increase in amount, making it difficult for newcomers or long-time fans to seek new watching material. [1-2].

Anime theme recommendations are important, especially for viewers who want to find anime that suits their preferences [3]. This topic will discuss various factors that can influence the choice of anime recommendations, namely genres and similarities. Analysis of these two factors was carried out using item-based recommendations and user-based recommendations.

In a previous research, collaborative filtering using user-based methods became the algorithm employed in an anime recommender system. The algorithm was then tested on a dataset consisting of 100,000 rows out of 1,000,000 rows, comprising details of 73,516 users and 12,294 anime. The system worked by solely considering the user's watch history. Therefore, to differentiate our attempt at an anime

recommender system, we decided upon using a hybrid approach, by not only utilizing user-based but also item-based collaborative filtering.

Item-based collaborative filtering and content-based filtering will be the two methods chosen as a recommendation system, item-based collaborative filtering can be used to find similarities between users and items, so that they can recommend items that have not been seen by users based on the activities and preferences of the users who have been seen. similar. Meanwhile, content-based filtering focuses on the characteristics of the item's own content by identifying the relevant attributes of the item, such as genre, synopsis, etc.

In this study, we will introduce and compare the method of item-based collaborative filtering and content-based filtering in the context of anime recommendation systems [10]. We will implement both methods and evaluate their performance in providing relevant and personalized recommendations to users. By understanding the strengths and weaknesses of each method, we can gain better insight into the optimal approach to improve the quality of the anime recommendation system.

II. DATASET

This dataset is taken from Kaggle, which has its own database on MyAnimeList. This dataset consists of several files, including "animelist.csv" which lists all the anime registered by the user with their respective scores, viewing status, and number of episodes watched. There is also a "rating_complete.csv" file which is a subset of "animelist.csv" which only considers fully watched and scored anime. There is also an "anime.csv" file which contains general information about each anime such as genre, statistics, studios, etc.

TABLE I
Data of rating_df

user_id	anime_id	rating
0	67	9
0	6702	7
0	242	10
0	4898	0

This dataset contains information on 17,562 anime and the preferences of 325,772 different users. This dataset includes a per-user list of anime, including abandoned, finished watching, about to watch, currently watching, and postponed Anime. Apart from that, this dataset also includes the ratings

given by users for anime that they have watched in full, as well as information about anime such as genre, statistics, studios, and so on.

The “animelist.csv” is defined as “rating df”, and originally contains 90,000,000 rows. And filtered to only users that have rated at least 400 shows and left with 71,418,114 rows. And we split for test dataset to 100,000 rows and the remaining rows used for training dataset which has 71,318,113 rows.

III. METHODOLOGY

A. Building Model

The embedding layers are used to convert the user and anime indexes into 128 dimensional vectors. Where it represents each user and anime. After that, the two are connected to the dot product layer to calculate the similarity between the user vector and the anime, the results will be used as input for the dense layer to predict ratings. The model is using the Adam optimizer with mean-squared-error as the loss-function with a total of 13,977,734 trainable parameters.

B. Item-based filtering

Item-based collaborative filtering is a powerful recommendation technique that leverages embedding layers and dense layers to generate accurate predictions. It takes into account two crucial inputs, namely user preferences and anime attributes. These inputs are then analyzed, allowing the model to deduce relationships between users and their preferred animes. This process made it possible to predict output ratings that will be used for providing personalized recommendations to users based on their interests and tastes [4].

C. Content-based filtering

Content-based filtering is a technique that recommends items based on their similarity to previously interacted items by the user. The similarity of the target item is determined by comparing its features with those of other items. This process is facilitated by employing a library known as Sentence Transformer for the content-based filtering model. The utilization of Sentence Transformer allows for the model to effectively capture the semantic meaning behind an anime's description and in turn provide the user with more accurate recommendations.

IV. RESULT AND DISCUSSION

In this research, we apply two models, namely item-based collaborative filtering and content-based filtering which uses transformer based to assist the embedding process which is applied in the recommendation system in anime. Item-based collaborative filtering can identify anime that have similarities in user preferences. The existence of rating data allows us to implement a system that can recommend similar anime based on user preferences.

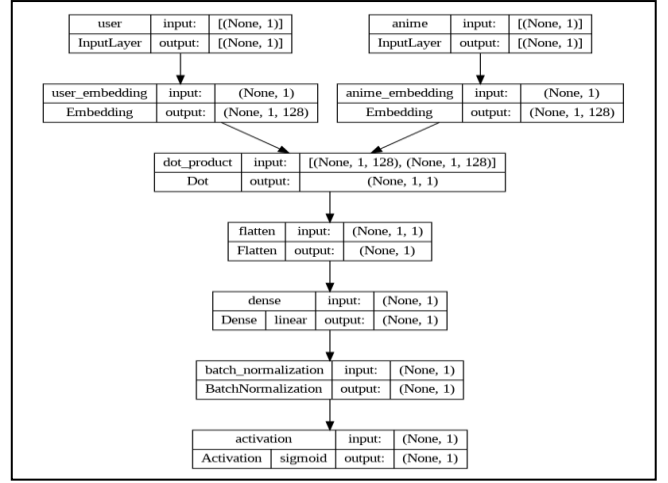


Fig. 1. Graphical representation of collaborative filtering model

Content-based filtering uses anime content information, such as a description or synopsis to provide recommendations. We use transformers based to produce embedding representation of anime content. The system can provide anime recommendations that have similarities in description or theme by utilizing the cosine similarity method. Here we compare item-based filtering and content-based filtering to show a list of anime with highest similarity to the prompted anime.

TABLE II.
Item-based Recommendation Result

Item Based Filtering	
Anime name	Similarity
Dragon Ball	0.974
Dragon Ball GT	0.882
Dragon Ball Z Kai	0.741
Dragon Ball Z Special 1:Bardock, The Father of Goku	0.709
Dragon Ball Z Special 2:The History of Trunks	0.695

TABLE III.
Content-based Recommendation Result

Content Based Filtering	
Anime name	Similarity
Dragon Ball Z Movie 14: Kami to Kami	0.705
Dragon Ball Kai	0.692
Rokushin Gattai GodMars: Juunana-sai no Densetsu	0.603
Dragon Ball Z Movie 12: Fūkkatsu no Fusion!! Gokū to Vegeta	0.584
Hagure Yuusha no Aesthetica	0.572

This research shows the effectiveness of item-based collaborative filtering and content-based filtering in the development/implementation of anime recommendation systems. The following is a discussion regarding the two methods. Some of the advantages of item-based collaborative filtering are that it does not require anime content information, so it can provide more personal recommendations without the need to use descriptions or other information.

Item-based collaborative filtering can provide recommendations that focus on similarities in anime content, so users can find anime with themes or genres that are similar to their favorite or preferred anime. By using content-based collaborative filtering assisted by transformers based for the embedding process, the recommendation results based on similarity are still lower than item-based collaborative filtering.

Transformers-based is important in the content-based filtering model. Because it can generate an embedding representation, this utilization can understand and produce important features of the description or synopsis of the anime, thereby increasing the accuracy and relevance of recommendations.

V. CONCLUSION AND FUTURE WORKS

After exploring and implementing item-based collaborative filtering and content-based filtering methods for anime recommendation systems. The results of the research show that using the item-based collaborative filtering method can provide better recommendation results through its similarity value compared to using the content-based filtering method which uses transformers-based. Which makes sense because content-based only uses metadata from anime, so the recommendations given are not very good. [5-9].

Although this research has succeeded in implementing collaborative filtering and content-based filtering in the anime recommendation system, there are still several areas that can be explored and improved. such as using hybrid approaches, incorporating user context, evaluation metrics,

etc. with further exploration, it is expected to create a recommendation system that is more precise and effective in meeting user needs.

REFERENCES

- [1] T. Liu, S. Liu, and I. Rahman, "International anime tourists' experiences: A netnography of popular Japanese anime tourism destinations," *Asia Pacific Journal of Tourism Research*, vol. 27, no. 2, pp. 135–156, 2022. doi:10.1080/10941665.2021.1998163
- [2] M. Pellitteri, "The European experience with Japanese animation, and what it can reveal about the transnational appeal of anime," *Asian Journal of Communication*, vol. 31, no. 1, pp. 21–42, 2020. doi:10.1080/01292986.2020.1862263
- [3] B. Soni, D. Thakuria, N. Nath, N. Das, and B. Boro, "Rikonet: A novel anime recommendation engine," *Multimedia Tools and Applications*, 2023. doi:10.1007/s11042-023-14710-9
- [4] A. S. Girsang, B. Al Faruq, H. R. Herlianto, and S. Simbolon, "Collaborative recommendation system in users of anime films," *Journal of Physics: Conference Series*, vol. 1566, no. 1, p. 012057, 2020. doi:10.1088/1742-6596/1566/1/012057
- [5] C. S. Murty, G. P. Saradhi Varma, and Ch. Satyanarayana, "Content-based collaborative filtering with hierarchical agglomerative clustering using user/ item based ratings," *Journal of Interconnection Networks*, vol. 22, no. Supp01, 2022. doi:10.1142/s0219265921410267
- [6] F. Xue et al., "Deep item-based collaborative filtering for top-N recommendation," *ACM Transactions on Information Systems*, vol. 37, no. 3, pp. 1–25, 2019. doi:10.1145/3314578
- [7] A. Da'u and N. Salim, "Recommendation system based on Deep Learning Methods: A systematic review and New Directions," *Artificial Intelligence Review*, vol. 53, no. 4, pp. 2709–2748, 2019. doi:10.1007/s10462-019-09744-1
- [8] S. Bag, S. K. Kumar, and M. K. Tiwari, "An efficient recommendation generation using relevant Jaccard similarity," *Information Sciences*, vol. 483, pp. 53–64, 2019. doi:10.1016/j.ins.2019.01.023
- [9] Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, and R. Kashef, "Recommendation systems: Algorithms, challenges, metrics, and business opportunities," *Applied Sciences*, vol. 10, no. 21, p. 7748, 2020. doi:10.3390/app10217748
- [10] U. Javed et al., "A review of content-based and context-based recommendation systems," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 16, no. 03, p. 274, 2021. doi:10.3991/ijet.v16i03.18851
- [11] Y. Tian, B. Zheng, Y. Wang, Y. Zhang, and Q. Wu, "College Library personalized recommendation system based on hybrid recommendation algorithm," *Procedia CIRP*, vol. 83, pp. 490–494, 2019. doi:10.1016/j.procir.2019.04.126