

Unveiling Anime Preferences: A Data-driven Analysis using MyAnimeList API

Ashwin Pillai, Advin Manhar

Amity University Chhattisgarh, Raipur, Chhattisgarh, India

ARTICLE INFO

Article History:

Accepted: 03 June 2023

Published: 16 June 2023

Publication Issue

Volume 9, Issue 3

May-June-2023

Page Number

488-495

ABSTRACT

This research paper presents a data-driven analysis of anime preferences using the MyAnimeList (MAL) API. The objective of this study is to uncover valuable insights into the preferences of anime viewers by analyzing data obtained through MAL API. The paper begins with a comprehensive background that includes a detailed literature survey to understand anime recommendation systems.

The methodology employed in this study involves utilizing the MAL API to collect and analyze data related to anime. The API provides access to a wide range of anime related information including ratings, ranks, genres etc. The collected data is then preprocessed to ensure its quality and suitability for analysis, involving data cleaning and transformation.

Various data analysis techniques are applied to the preprocessed data. These techniques help identify patterns and trends in different fields such as genres, studios, media type etc. The findings from the data analysis provide valuable insights into the preferences of anime viewers and help to identify features that may be crucial while building the model. Then using cosine similarity model, content-based anime recommendation system is built after feature engineering.

The results of this research highlight the importance of leveraging data-driven approaches in understanding and catering to the diverse preferences of anime viewers. By EDA uncovering patterns and trends in anime preferences, anime recommendation systems can be enhanced to provide more relevant recommendations to viewers. The implications of this research extend to the anime industry, where a deeper understanding of viewer preferences can inform content production and distribution strategies.

In conclusion, this research paper presents a data-driven analysis and content-based recommendation system for anime using the MAL API and data science. The methodology employed involves data collection, preprocessing, Exploratory Data Analysis(EDA), Feature engineering and Model building. The findings contribute to enhancing anime recommendation systems and understanding the diverse preferences of anime viewers.

Keywords : MyAnimeList API, Exploratory Data Analysis(EDA), Content Based Recommendation System, Cosine Similarity, Anime Recommendation System.

I. INTRODUCTION

Background

Anime is hand-drawn and computer-generated animation originating from Japan. Outside of Japan and in English, anime refers specifically to animation produced in Japan. However, in Japan and in Japanese, anime describes all animated works, regardless of style or origin. Animation produced outside of Japan with similar style to Japanese animation is commonly referred to as anime-influenced animation.

Anime has gained immense popularity worldwide, attracting a diverse community of viewers with varied preferences. With the exponential growth of anime content available, it has become more and more difficult for viewers to discover new shows that align with their interests. This creates a need for effective anime recommendation systems that can cater to individual preferences and enhance the overall viewing experience.

In recent years, data-driven approaches have revolutionized recommendation systems across various domains. By analyzing large datasets, valuable insights can be extracted to understand user preferences and provide personalized recommendations. In the context of anime, leveraging data-driven analysis can significantly contribute to improving the accuracy and relevance of recommendations.

II. OVERVIEW

MyAnimeList, also known as MAL, is an anime and manga social networking and social cataloging application website run by volunteers. The site provides its users with a list-like system to organize and score anime and manga. It facilitates finding users who share similar tastes and provides a large database on anime and manga. Their large subset of data related to anime and manga can be accessed through their API.

This research paper focuses on analyzing anime preferences using the MAL API, MAL is a widely used platform that provides comprehensive information

about anime, including ratings, genres, and studios. MAL is the most popular website in the anime community to get information or discuss about anime related stuff. This API provides access to a vast collection of anime-related information. The objective of this study is to gain a deeper understanding of viewer preferences by analyzing data obtained through the MAL API.

The research methodology involves collecting data from the MAL API, which offers a rich source of information on anime series, movies, and related metadata. The collected data is then preprocessed to ensure its quality and suitability for analysis. This preprocessing step involves cleaning the data, handling missing values, and transforming it into a structured format suitable for analysis.

Exploratory Data Analysis (EDA) will be conducted to uncover patterns and trends in anime preferences, such as popular genres, highly rated shows, and type of media content produced. Additionally, feature engineering techniques will be applied to identify key attributes that influence viewer preferences.

Based on the insights gained from the data analysis, a content-based recommendation system will be built using cosine similarity. This system will leverage the identified features and provide personalized recommendations to anime viewers based on their similarities with other shows or also have an option of selecting a genre and get recommendations based on their ranking.

The findings of this research hold significant implications for the anime industry and the development of anime recommendation systems. Moreover, the research contributes to the ongoing efforts to enhance the accuracy and effectiveness of anime recommendation systems, ultimately improving the overall anime viewing experience for users.

III. Literature Survey

"Analyzing User Requests for Anime Recommendations": The paper analyzes user questions

from a Korean Q&A website to understand how people seek anime recommendations. The findings highlight the importance of robust metadata and allowing users to specify known anime and series titles for better recommendations.

"Analyzing anime users' online forum queries for recommendation using content analysis": This study uses content analysis of recommendation request threads from an online forum to identify important anime information features. The findings can enhance anime-related databases and information systems for better organization, browsing, retrieval, and recommendation of anime.

"AniReco: Japanese Anime Recommendation System": The paper proposes AniReco, an animation work recommendation system that reflects users' potential preferences. The system is capable of recommending animation works and their related contents based on users' preferences.

"ANIME RECOMMENDATION SYSTEM": This paper aims to develop a recommendation system for newcomers to the anime world. It utilizes KNN and SVD algorithms to offer suggestions and recommendations based on user preferences.

"Collaborative Recommendation System in Users of Anime Films": The study focuses on recommending anime films based on ratings of previously watched films. It uses collaborative filtering techniques and an ALS method to match user histories and recommend anime based on similarities.

"Explainable recommendation based on knowledge graph and multi-objective optimization": The paper proposes an explainable recommendation framework based on a knowledge graph and multi-objective optimization. The framework aims to optimize precision, diversity, and explainability in recommendations.

"Diversified Interactive Recommendation with Implicit Feedback": The paper introduces a novel diversified recommendation model, DC2B, for interactive recommendation with users' implicit feedback. The

model employs determinantal point process to promote diversity in recommendation results.

"A Deep Learning Recommender System for Anime": The research proposes a deep learning-based recommender system for personalized and accurate anime recommendations. The model utilizes a collaborative filtering approach and is evaluated using metrics such as Mean Squared Error and Mean Absolute Error.

"AniReco: Japanese Anime Recommendation System": This paper introduces AniReco, an animation work recommendation system capable of recommending animation works and related contents while reflecting users' preferences.

"Facet Analysis of Anime Genres: The Challenges of Defining Genre Information for Popular Cultural Objects": The study conducts a facet analysis of anime genre terms used in English-language databases and websites. The identified terms can improve genre definitions and search experiences in different organizational systems.

"Using Posters to Recommend Anime and Mangas in a Cold-Start Scenario": The research addresses the item cold-start problem in anime and manga recommendations. It utilizes a deep learning technique to extract tag information from posters and proposes a collaborative filtering model, BALSE, for improved recommendations.

"Recommendation System For Anime Using Machine Learning Algorithms": The paper focuses on developing a recommendation system for anime using machine learning algorithms. Content-based filtering, collaborative-based filtering, and popularity-based filtering techniques are applied to provide top matching recommendations.

"Improving Recommender Systems Performance with Cross-domain Scenario: Anime and Manga Domain Studies": The study explores cross-domain recommendation techniques in the anime and manga domain. The results show that cross-domain recommenders outperform single-domain

recommenders in manga recommendations when anime domains are used as auxiliary domains.

"Mangaki: an Anime/Manga Recommender System with Fast Preference Elicitation": The paper presents Mangaki, a manga and anime recommender system that implements a preference elicitation process based on welcome decks.

"A Study on Accuracy, Miscalibration, and Popularity Bias in Recommendations": The research investigates the accuracy, miscalibration, and popularity bias in recommendations. It explores the impact of different recommendation models and algorithms on user experiences and suggests strategies to improve recommendation accuracy.

Theory:

Cosine similarity:

Cosine similarity is a measure commonly used to quantify the similarity between two vectors in a multi-dimensional space. It is particularly effective in scenarios where the magnitude of the vectors is less important than their orientations and relative angles.

The cosine similarity ($\cos\theta$) is calculated using the following formula:

$$\cos\theta = \frac{(A \cdot B)}{||A|| * ||B||}$$

Where:

- A and B are the two vectors being compared.
- $A \cdot B$ represents the dot product of vectors A and B.
- $||A||$ and $||B||$ denote the magnitudes (or Euclidean norms) of vectors A and B, respectively.

The resulting value ranges between -1 and 1. A similarity of 1 indicates perfect alignment, meaning that the two vectors are in the same direction. A similarity of -1 indicates complete opposition, meaning that the two vectors are in opposite directions. A similarity of 0 suggests orthogonality, meaning that the two vectors are perpendicular to each other.

In the context of a content-based recommendation system for anime, cosine similarity is employed to assess

the similarity between feature vectors representing different anime titles. Using cosine similarity, the recommendation system can offer personalized recommendations based on the underlying content and attributes of the anime, rather than relying solely on user ratings or collaborative filtering techniques. By comparing the feature vectors, the system identifies anime titles that share similar characteristics and recommends them to users.

In summary, cosine similarity is a fundamental concept within the content-based recommendation system. It enables the measurement of similarity between anime titles based on their feature vectors using the cosine similarity formula. By leveraging cosine similarity, the system provides personalized recommendations that align with user preferences and aid in the discovery of new anime titles with similar attributes.

IV. Methodology

The methodology employed in this research paper combines data collection, preprocessing, exploratory data analysis (EDA), feature engineering, and model building to analyze anime preferences and develop a content-based recommendation system. This section provides a detailed overview of the steps involved in conducting the study.

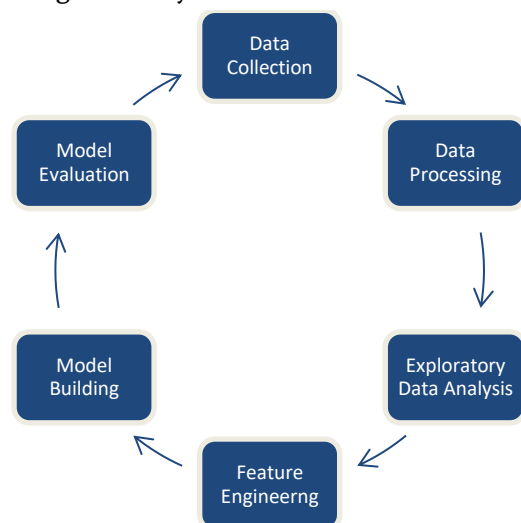


Fig.1 – Different Stages in The Process of Making Recommendation System

1. **Data Collection:** The primary source of data for this research is the MAL API. Through API requests, a substantial dataset is gathered, encompassing a wide range of anime titles and associated metadata. First, we had to set up credentials and other parameters for accessing the API. Then with the help of the requests library we made API calls and stored the responses.
2. **Data Preprocessing:** To ensure the quality and suitability of the data for analysis, a thorough preprocessing stage is conducted. This includes data cleaning, handling missing values, and transforming the json data into pandas dataframe and stored it in a csv file. Cleaning involves removing duplicate entries, correcting inconsistencies, and addressing any data integrity issues. Missing values are handled through a combination of appropriate imputation and removal techniques, ensuring minimal impact on subsequent analysis.
3. **Exploratory Data Analysis (EDA):** EDA is performed to gain insights into the collected dataset and identify patterns, trends, and distributions related to anime preferences. Various statistical techniques, visualizations, and summary statistics are utilized to explore the data comprehensively. EDA helps uncover relationships between variables, identify popular genres, media type, and reveal any notable trends.
4. **Feature Engineering:** Feature engineering plays a crucial role in identifying relevant attributes that influence anime preferences. This step involves selecting, transforming, and creating new features from the dataset. Features such as genres, studios, rank and similarity metrics are derived to enhance the recommendation system's performance and accuracy.
5. **Model Building:** Based on the identified features, a content-based recommendation system is developed using cosine similarity. This model utilizes the extracted features to measure the similarity between

different anime titles and recommend shows that align with a user's preferences. Cosine similarity provides a measure of similarity by comparing the vector representations of anime titles based on their feature values.

6. **Model Evaluation:** The developed recommendation system is evaluated based on the quality of recommendations provided. The system's performance is assessed by surveying the recommended anime titles with user preferences and analyzing the alignment between the two. The evaluation process helps gauge the effectiveness and reliability of the recommendation system.

By following this methodology, a comprehensive analysis of anime preferences is conducted, and a content-based recommendation system is developed to provide personalized anime recommendations to users based on their preferences and similarities with other shows.

Insightful visualizations:

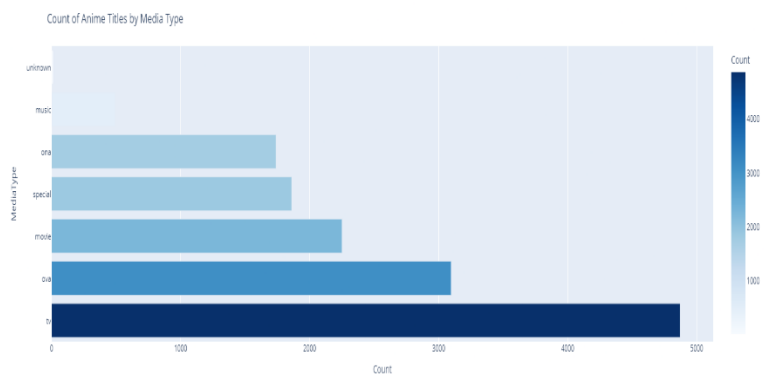


Fig.2 – Horizontal Bar Graph of Count of Anime Titles by Media Type

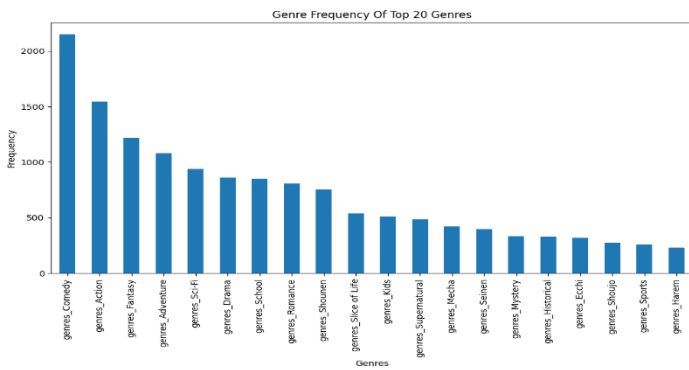


Fig.3 – Bar Graph of Top 20 Anime Genres

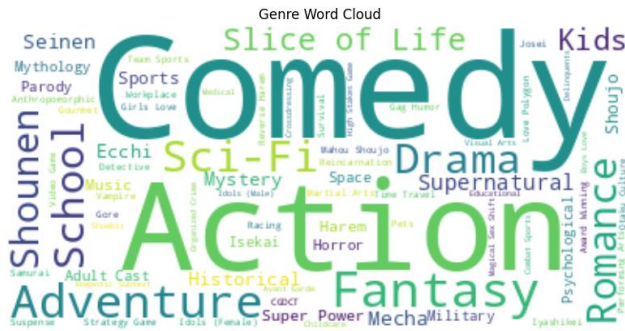


Fig.4 – Word cloud Of Genres Present in The Anime Dataset

Media Type	Count
0 tv	4871
1 ova	3096
2 movie	2250
3 special	1860
4 ona	1738
5 music	493
6 unknown	13

Fig.5 – Table Showing Count of Anime Titles Present in Each Media Type

	id	num_episodes	num_users	popularity	rank
count	14321.000000	14321.000000	1.432100e+04	14321.000000	14321.000000
mean	23275.142937	12.336988	3.259079e+04	8177.606801	9479.574471
std	17932.337522	38.337222	1.217818e+05	5598.263150	7330.167920
min	1.000000	0.000000	0.000000e+00	1.000000	1.000000
25%	4820.000000	1.000000	3.190000e+02	3606.000000	3624.000000
50%	22763.000000	3.000000	2.119000e+03	7353.000000	7444.000000
75%	38648.000000	12.000000	1.320900e+04	11733.000000	14172.000000
max	55571.000000	1818.000000	2.655870e+06	24597.000000	24606.000000

Fig.6 – Summary of Numeric Data Columns

V. Result:

The analysis conducted using the MAL API revealed several key findings regarding anime preferences:

1. More than 30% of anime is sourced through manga.
2. The mean number of episodes in an anime is 12.33.
3. The most popular genres among users are Comedy, Action, and Fantasy.
4. The majority of anime falls under the TV media type, followed by OVA and movies.
5. A list of top-rated anime series based on user ratings was identified, providing valuable recommendations for anime enthusiasts.
6. A combination of features, including title, genres, studios, and ranking, was found to contribute to the best recommendations.

VI. Discussion:

The high percentage of anime sourced through manga suggests a strong connection between the two mediums, indicating potential opportunities for cross-promotion and content adaptation.

The mean number of episodes in an anime can serve as a reference for both viewers and content creators, providing an understanding of the typical length of anime series. This information can guide viewers in

selecting anime based on their time availability and preferences, while also aiding creators in structuring their storytelling and pacing.

The popularity of Comedy, Action, and Fantasy genres highlights the audience's interest in these themes. This finding can inform content creators and distributors in curating and promoting anime titles that align with these popular genres.

Understanding the distribution of anime across different media types (TV, OVA, movies) helps stakeholders in the anime industry to optimize their production and distribution strategies. This information can guide decisions on resource allocation, marketing efforts, and platform selection for reaching the target audience effectively.

The identification of top-rated anime series based on user ratings provides a valuable resource for anime enthusiasts seeking high-quality content recommendations. These highly regarded titles can attract a larger viewership and contribute to the overall growth and reputation of the anime industry.

By combining various features such as title, genres, studios, and ranking, the recommendation system can provide more accurate and relevant suggestions to viewers. This holistic approach takes into account multiple aspects of an anime's characteristics, enhancing the user experience by offering personalized recommendations that align with their preferences.

VII. Conclusion & Future Scope

Conclusion:

In conclusion, this research paper utilized the MyAnimeList API and data-driven analysis to gain insights into anime preferences. The findings shed light on various aspects of anime, including the influence of manga as a source, the average episode count, popular genres, media types, and top-rated series. The implementation of a content-based recommendation system demonstrated the importance of considering multiple features for providing accurate and relevant anime suggestions to viewers.

Limitations:

However, it is important to acknowledge the limitations of this study. Firstly, the analysis was conducted solely based on data obtained from the MyAnimeList API, which may not represent the entire anime-viewing population. The results might be skewed towards the preferences and behaviors of MyAnimeList users. Additionally, the analysis primarily focused on quantitative data, and factors such as qualitative user feedback and subjective opinions were not extensively considered.

Future Scope

To further enhance our understanding of anime preferences, future research should consider integrating data from multiple anime platforms and incorporating user reviews, ratings, and sentiment analysis. Additionally, exploring collaborative filtering techniques can improve the accuracy and personalized nature of the recommendation system.

Furthermore, conducting surveys or interviews with anime viewers can provide valuable qualitative insights into their preferences and help in uncovering underlying factors driving their choices. Incorporating demographic and contextual information can also enable more targeted and tailored recommendations. Investigating the impact of cultural and regional differences on anime preferences can offer valuable insights for content creators and distributors to tailor their offerings for specific markets.

VIII. REFERENCES

- [1]. Jin Ha Lee, Yuna Shim, and Jacob Jett. 2015. Analyzing User Requests for Anime Recommendations. In Proceedings of the 15th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '15). Association for Computing

- Machinery, New York, NY, USA, 269–270. <https://doi.org/10.1145/2756406.2756969>
- [2]. Cho, H., Schmalz, M.L., Keating, S.A. and Lee, J.H. (2018), "Analyzing anime users' online forum queries for recommendation using content analysis", *Journal of Documentation*, Vol. 74 No. 5, pp. 918-935. <https://doi.org/10.1108/JD-08-2017-0122>
- [3]. Syoichiro Ota, Hayaki Kawata, Masahumi Muta, Soh Masuko, Jun'ichi Hoshino. AniReco: Japanese Anime Recommendation System. 16th International Conference on Entertainment Computing (ICEC), Sep 2017, Tsukuba City, Japan. pp.400-403, [ff10.1007/978-3-319-66715-7_49ff](https://doi.org/10.1007/978-3-319-66715-7_49ff). [ffhal-01771283f](https://doi.org/10.1007/978-3-319-66715-7_49ff)
- [4]. Shriram, G., Prithika, U.G., & Dhivya, S. (2022). ANIME RECOMMENDATION SYSTEM.
- [5]. A S Girsang, B Al Faruq, H R Herlianto and S Simbolon Collaborative Recommendation System in Users of Anime Films 2020 J. Phys.: Conf. Ser. 1566 012057
- [6]. Xie, L., Hu, Z., Cai, X. et al. Explainable recommendation based on knowledge graph and multi-objective optimization. *Complex Intell. Syst.* 7, 1241–1252 (2021). <https://doi.org/10.1007/s40747-021-00315-y>
- [7]. Liu, Y., Xiao, Y., Wu, Q., Miao, C., Zhang, J., Zhao, B., & Tang, H. (2020). Diversified Interactive Recommendation with Implicit Feedback. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(04), 4932-4939. <https://doi.org/10.1609/aaai.v34i04.5931>
- [8]. Mutteppagol, Vidyashree Mahaling (2021) A Deep Learning Recommender System for Anime (2021) Masters thesis, Dublin, National College of Ireland. <https://norma.ncirl.ie/id/eprint/5201>
- [9]. Ota, S., Kawata, H., Muta, M., Masuko, S., Hoshino, J. (2017). AniReco: Japanese Anime Recommendation System. In: Munekata, N., Kunita, I., Hoshino, J. (eds) *Entertainment Computing – ICEC 2017*. ICEC 2017. Lecture Notes in Computer Science(), vol 10507. Springer, Cham. https://doi.org/10.1007/978-3-319-66715-7_49
- [10]. Cho, H., Disher, T., Lee, W.C., Keating, S., & Lee, J. (2018). Facet Analysis of Anime Genres: The Challenges of Defining Genre Information for Popular Cultural Objects. *KO KNOWLEDGE ORGANIZATION*, 45(6), 484–499.
- [11]. J. -J. Vie et al., "Using Posters to Recommend Anime and Mangas in a Cold-Start Scenario," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 2017, pp. 21-26, doi: 10.1109/ICDAR.2017.287.
- [12]. Jena, Abhipsa and Jaiswal, Arunima and Lal, Dakshita and Rao, Soumya and Ayubi, Afshan and Sachdeva, Nitin, Recommendation System For Anime Using Machine Learning Algorithms (May 27, 2022). *Proceedings of the International Conference on Innovative Computing & Communication (ICICC) 2022*, Available at SSRN: <https://ssrn.com/abstract=4121831> or <http://dx.doi.org/10.2139/ssrn.4121831>
- [13]. R. B. Bahaweres and A. R. Almujaiddi, "Improving Recommender Systems Performance with Cross-domain Scenario: Anime and Manga Domain Studies," 2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Jakarta, Indonesia, 2022, pp. 361-365, doi: 10.23919/EECSI56542.2022.9946560.
- [14]. Jill-Jénn Vie, Camille Laily, Solène Pichereau "Mangaki: an Anime/Manga Recommender System with Fast Preference Elicitation" (2015) https://jill-jenn.net/_static/works/mangaki-recsys2015.pdf
- [15]. Dominik Kowald, Gregor Mayr, Markus Schedl, & Elisabeth Lex. (2023). A Study on Accuracy, Miscalibration, and Popularity Bias in Recommendations. *builtin.com*
- [16]. Michael Sun, Tai Vu, Robert Yang. DeepAniNet: Deep NLP-based Representations for a Generalizable Anime Recommender System.

Stanford CS224N Custom Project.
https://web.stanford.edu/class/cs224n/reports/final_reports/report050.pdf

Cite this article as :

Ashwin Pillai, Advin Manhar, "Unveiling Anime Preferences: A Data-driven Analysis using MyAnimeList API", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 3, pp.488-495, May-June-2023.

Available at doi :

<https://doi.org/10.32628/CSEIT23903115>

Journal URL : <https://ijsrcseit.com/CSEIT23903115>

- [17]. N. S. Patankar, Y. S. Deshmukh, R. D. Chintamani, K. Vengatesan and N. L. Shelake, "An Efficient Machine Learning System using Sentiment Analysis for Movie Recommendations," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022, pp. 1284-1291, doi: 10.1109/ICOSEC54921.2022.9951906.
- [18]. Daryl Bagley. Combining Content Information with an Item-Based Collaborative Filter 2017 The Alpha Chi Journal of undergraduate scholarship, vol 2, issue 2, 2017 DOI: 10.21081/AX0126
- [19]. Garanayak, M., Nayak, S. K., Sangeetha K., Choudhury, T., & Shitharth S. (2022). Content and Popularity-Based Music Recommendation System. International Journal of Information System Modeling and Design (IJISMD), 13(7), 1-14. <http://doi.org/10.4018/ijismd.315027>
- [20]. scholar.google.com
- [21]. [semanticscholar.org](https://www.semanticscholar.org)
- [22]. [iopscience.iop.org](https://www.iopscience.iop.org)
- [23]. ieeexplore.ieee.org
- [24]. en.wikipedia.org
- [25]. [kaggle.com](https://www.kaggle.com)
- [26]. github.com
- [27]. pypi.org
- [28]. python-mal.readthedocs.io
- [29]. myanimelist.net
- [30]. gitlab.com
- [31]. [pythonspot.co](https://pythonspot.com)
- [32]. en.wikipedia.org
- [33]. [youtube.com](https://www.youtube.com)