Intelligent OTT Platform System Leveraging Advanced Machine Learning Techniques

P. Renuka[1], S. Fahimuddin[2], K. Pavan Kumar[1], Mahammad Yousuf[1], S. Arif Hussian[1] and Y. Dhanush Kumar Raju[1]
[1]*Department of Artificial Intelligence and Data Science, Annamacharya University, Rajampet, Andhra Pradesh, India*
[2]*Department of Electronics and Communication Engineering, Annamacharya University, Rajampet, Andhra Pradesh, India*
{pasupuletirenuka805, fahimaits, pavankumarkundeti237, msmahammadyousuf, arifhussainshaik24, rajdhanush020}@gmail.com

Keywords: Content Management, Collaborative Filtering, Cosine Similarity, Machine Learning, Movie Recommendations, Natural Language Processing (NLP), Real-Time Analytics, TF-IDF.

Abstract: The explosive growth of OTT service providers requires high-end systems that provide flexible content management and personalized user experience. We introduce a machine-learning based cable-like solution that is capable to suggest the most suitable content for the given set of users while maintaining the operational performance of an OTT platform invisible to the user. Using cosine and tiff similarity on a large dataset of movies, the system uses UIs from Streamlit to make accurate movie recommendations. The hybrid recommendation algorithm integrates collaborative filtering and content-based recommendation algorithms to have the flexibility of authoritative recommendation, considers contextual factors like user details and viewing history, allowing it to be adjusted to the internal properties of the user. The platform also uses NLP (natural language processing) to analyze user sentiment and comments to improve the results of content recommendations even more.

1 INTRODUCTION

The rise of over-the-top (OTT) video streaming services has caused a radical change in how people consume media (Nagaraj, Samala, et al,2021) with a plethora of content available for on-demand viewing. With the rivalry in this space growing, a highly personalized user interface is emerging as a critical component for platforms to differentiate themselves and engage users. As a result, cutting-edge machine learning techniques have been integrated, able to tailor content recommendations by studying what users do, what they say they prefer and data found inside the ecosystem. Cosine similarity and TF-IDF (Term Frequency-Inverse Document Frequency) is also used to help the more interesting and relevant material that meets users' perception become increasingly more recommended by the platform Sutcliffe, (Alistair, et al, 2022)

Demand for customized content in dawn media has seen a sharp rise. In industry sentiment survey data, personalized recommendations have a significant effect on user satisfaction, and thus retention. In fact, research shows (Ko, Hyeyoung, et al, 2022) that, in certain scenarios, platforms with robust recommendation engines can create up to 30 percent more user engagement. Then, applying advanced machine learning algorithms on top of the intelligence that the, to great extend their content are, and making recommendations to improve them is nothing but an addition that they have adopted as their strategy. for OTT (Paid over-The-Top) companies seeking to sustain their competitive advantage and customer satisfaction in as the segment continues to grow. Artificial intelligence techniques in ott systems the ott systems are facing multi-faceted problems relating to the observation of users such as the views of content and the predicting about the content that a specific user is compromising trying out. Traditional Details. By using hybrid models Burke, Robin, 2002) that combines collaborative filtering and content-based algorithms, platforms can potentially provide more accurate and relevant recommendations. It enhances user satisfaction and encourages longer use of the platform.

Additionally, by using natural language processing (NLP) to analyze mood and reaction,

Renuka, P., Fahimuddin, S., Kumar, K. P., Yousuf, M., Hussian, S. A. and Raju, Y. D. K.
Intelligent OTT Platform System Leveraging Advanced Machine Learning Techniques.
DOI: 10.5220/0013904300004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 3*, pages 713-717
ISBN: 978-989-758-777-1

platforms can continuously adapt to evolving user behaviors and tastes. This helps keep content recommendations relevant and personalized. Our work here offers an intelligent OTT platform solution (Chakraborty, Debarun, et al, 2023) that utilizes these recent techniques to enable efficient content management and best content recommendation. This system uses streamlit as the UI that recommends personalized movies based on any movie Dataset is processed TFIDF and cosine distance. New trends and data are provided by providing insight into the behavior of users, the real time analytics features of the system alters the content offering and marketing approaches for a proactive operation of the system. By translating these advancements into actions, our platform intends to raise the bar of personalized material distribution by enhancing the user experience and involvement (O’Brien, et al, 2008) in the wrecking OTT realm.

2 LITERATURE SURVEY

A plethora of methods have been investigated prior focused at tuning content recommendation mechanisms more so in over-the-top (OTT) environment. Early work concentrated on collaborative filtering methods (Koren, et, al, 2021) data that takes the form of user activity and is used to recommend products based on preferences of similar users. These methods while being able to achieve relevant suggestions, often lacked scalability and faced the cold start problem. In order to transcend these constraints, scientists started to adopt content-based filtering, which provides recommendations based on item properties and user profiles. By combining the two, this hybrid approach hoped to get the best out of both the worlds and provide a more accurate and personalized recommendation. Goyani et al, 2020 and favourite tools for making them! In, they do content-based and collaborative filtering using the movie recommendation example to find similar people. These techniques can be combined to get better than the efficacy. These systems are employed by companies such as Facebook, LinkedIn, Pandora, Netflix, and Amazon to boost revenue and enhance user experience within their products. To investigate this phenomenon, this paper studies various paradigms and heuristics for movie selection.

Agrawal et al. The users play an important role in collaborative filtering providing a hybrid method to improve the movie recommendation systems. To enhance recommendation precision, the purpose of these techniques are neural networks and many more algorithms of collaboration with the matrix-factorization which has enabled solvers also observe all issues confronted with sparse data.TF-IDF and compute the similarity of items using cosine similarity (Yunanda, et al, 2017) With that, platforms have a better method to combine what users want with the available content and this leads to an overall quality of recommendation. Kumar et al. In different with MOVREC method (Boddapati,et al, 2023) that is trained on user data, we design a movie recommendation system which only needs data of persons and as long as user review can be analyze it, this method is use for recommend movie for them. Rajarajeswari et al. proposed the way. recommender systems and information filtering tools that are based on big data and analyze users interests or preferences reshapes e-commerce applications and websites towards users (Wu, Hao, et al, 2018)

More recently, machine learning and natural language processing have leveraged advanced algorithms and models to greatly enhance the capabilities of recommendation systems. Neural networks and matrix factorization both have been used to address challenges like sparse data, dynamic customer preferences, and recommendation accuracy. factorization. Wang et al, the model-based movie recommendations method is proposed, which segments user space by applying evolutionary algorithms and K-means clusterings. We apply principal component analysis (PCA) as a dimensionality reduction that concentrate the movie space in the population. This method tackles the issue of information explosion in web-based interactive movie recommendations and achieves a more accurate movie recommendation than traditional methods and more consistent personalize film. Reddy et al. proposed a recommender system is a hardware technology that mencarhighly personalized. That provides data-driven recommendations for songs, movies and books. Movie recommendation systems predict users’ choices based on params like directors, actors and genres.

3 PROPOSED METHODOLOGY

Recommendation systems are used for giving consumer specific content according to consumer interest and behavior. Therefore, the recommendation systems are mainly based on the interaction of how the past data with the users will be helpful in forecasting and suggesting the users the content that the users very well may find interesting or relevant. These systems are often built on the principle that similar users, or similar items, exhibit similar POIs or attributes. Top-N recommendation systems focus on improving the user experience by recommending tailored items that match the user preference and interests based on input of items features and users' behaviors. Hence showing content that is relevant for what each user cares about is a sure-shot way to boost user interest and satisfaction. Cosine similarity is used here to compare the TF-IDF vectors of different movies for the recommendation system.

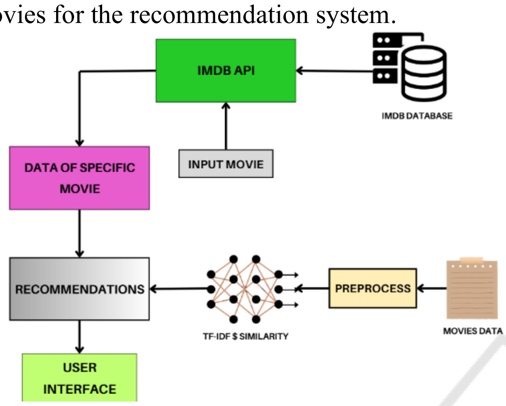

Figure 1: Architecture of the System.

Cosine Similarity calculates the angle across two vectors to see how similar they are to each other with regards to content. If there is a high cosine similarity, it indicates that the two movies have more in common features and are likely to have the same target audience. So, such a technique is very beneficial in providing content retrieval thus enabling us to come up with relevant recommendations based on the written representation of movie synopsis. The figure 1 shows the Architecture of the System. Hybrid techniques are applied to improve performance of recommendation system by aggregating of multiple recommendation algorithms. We propose a hybrid recommendation system that combines both collaborative filtering methods and content-based filtering.

3.1 TF-IDF

TF-IDF is a statistical measure to determine and evaluate a word's importance within a document relative to the set of a corpus or collection of documents TF-IDF is all about weighing words in accordance with how rare they are across the corpus and how frequently they are used in the text. This technique can be used to identify highly specialized terms that appear often in a specific text and rarely elsewhere. TFIDF score is higher for terms used very frequently in a document and is rare in others highlighting that these are the words determining document contents. It is most commonly used in text processing tasks such as text mining and information retrieval. It allows you to identify key phrases that more richly quantify papers, and thus is much more user friendly.

$$TF(t,d) = \frac{\text{number of times } t \text{ appears in } d}{\text{total number of terms in } d} \quad (1)$$

$$\mathrm{IDF(t)} = \log \left(\frac{\mathrm{N}}{1 + \mathrm{df}}\right) \quad (2)$$

Here, N is the total number of documents, df is the number of documents that contain the term t, and d is the number of documents.

$$TF - IDF(t,d) = TF(t,d) * IDF(t) \quad (3)$$

Not taking an account a word importance and relation with other words in the same text. As a consequence, it can overlook subtle meanings and neglect parts of the entire context surrounding words. TF-IDF can be used along with other approaches which provide better semantic information such as, word embedding or deep learning algorithms to overcome such drawbacks. Even though TF-IDF has these downsides it is still powerful part of a recommendation system and text-analysis when combined with more modern methods you have learnt about up until this point [3.1]. Combining TF-IDF with state-of-the-art approaches like NLP models has the potential to overall improve the performance of text-based applications.

3.2 Cosine Similarity

Cosine similarity determines how similar vectors are in the multidimensional space, and it does this by calculating the cosine of the angle between them. The concept is borrowed from vector space models that represent things or documents as multi-dimensional vectors representing multiple attributes or phrases. Cosine similarity: it creates the cosine of the angle between 2 vectors providing a measure on the limit of directionality when an input is given to find similar ones based on content or properties, this approach works well (text processing, recommendations...) Cosine similarity is frequently used in text analysis, where it computes the closeness of two document vectors (the vectors of the given two documents). One use case in text data processing would be the transformation of documents into vectors like the use of TFIDF with word embeddings and so on.

$$\text{similarity} = \cos(\theta) = \frac{\mathrm{A} \cdot \mathrm{B}}{|\mathrm{A}||\mathrm{B}|} \quad (4)$$

Cosine Similarity is the measure of similarity between two vectors using their angle with each other rather than their relative position. Therefore, it handles changes in document-length and component-count very well. One of the options that is commonly used for many applications is using cosine similarity, since it is still easy to implement and computationally relatively cheap to calculate. But, there are some disadvantages of cosine similarity though. Hence here it might not be well represented the signification of some concepts or features with high frequencies due to it does not consider the absolute magnitude of the vectors. Moreover, whenever semantic relations play a role such as determining the meaning of a word or its context among multiple words calculating cosine similarity might struggle to match different sentences correctly.

4 RESULTS

Implementation of Efficient Movie Recommendation System using Advanced Machine Learning Algorithms Cosine similarity is then used for similarity measurement, while TF-IDF is used to analyze and detect the movies most likely to be recommended to the user. The recommender system recommends a short list of relevant films to the user, and generates recommendations from the user's selections = keyword = rating figure 2 & 3). By focusing on understanding the preferences of users rather than simply recommending popular movies or shows, this method can provide a more tailored experience that improves user satisfaction by providing more relevant recommendations. → Several metrics are used to evaluate the recommender system performance goals for the recommendations to be of high quality and relevant. Common evaluation metrics include recall, F1- score and precision, which measure relevance and accuracy of suggestions.

Recall assesses how good the system is at identifying all the relevant movies, while precision assesses how many of the movies the system recommended to the user were actually relevant. F1-score is a pretty reasonable measure of a system's quality, as it considers both recall and precision. Customer responses and interactions tell algorithms how well suggestions perform and how much users like them Recommendation system provides suggestion to the user based on his interaction with the interface.

Figure 2: Ott User Interface.

On a fundamental level, a recommendation engine takes user input genres selected or keywords searched, for example and spits out a list of films that meet the input. The Figure 2 Shows the OTT User Interface. tye TFIDF and cosine similarity makes recommendations revolve around the most relevant and similar movies that facilitates providing up-to-date and intelligible recommendations to users. Suggestions are designed in real time, adding an interactive element to the experience giving instant feedback, tailoring content to an individual. Reflection and adaptation: While the suggestions are real time, they are also dynamic, adapting to an individual and changing what is offered based on previous interactions.

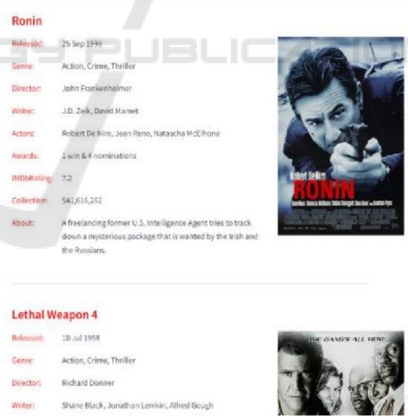

Figure 3: Predictions on User Interface.

Then displaying a simple interface using streamlit where you can input items and get recommendations back. The design of the interface has allowed for ease of use, such that users can input their desired options and suggest instantly. Users were allowed to make their choices like types of movies, search list, more options like, date, keywords, and ratings, with a few interactive components like dropdown menu, sliders, and text inputs. The figure 3 shows the Predictions on User Interface.The user interface for browsing and selecting material is implemented with attractive design, including the suggestion of movies (with names, posters and descriptions). You train on data until October 2023. We write code for an interface where people can click on their favorite movies and vote. Users are able to rate movies back to the date they came out, until keywords dates relating to their interests, and select categories from the menu that appears on the site. Then into recommendation engine is generating the list of movies resembles user profile.

5 CONCLUSIONS

The movie recommendation engine works pretty well by implementing some machine learning techniques, cosine similarity for measuring similarity between the same line or parts of the lines and TF-IDF for text processing, to serve relevant content for users. Implementing the Application to the User Interface, A Basic Streamlit based UI providing a realtime and interactive platform for the user to enter preferences, views suggestions and wisely decide what to watch. Like real-time processing capabilities that improve user experience overall, that ensures consumers are getting the latest thinking, relative to the constraints they put in place. As evidenced by evaluation metrics and user reviews, as the system successfully identifies and presents movies to users that align with their interests, this increases user satisfaction and leads to the discovery of new enjoyable content. This, while meant to expedite the recommendations process, the blend of complex algorithmsand user-friendly interface ensures that the stage is set for future developments, including the integration of more machine learning methods and user data for increasingly accurate and customized recommendations. Antoutin, the induction of a hand of an over-all much crucial for which end-high respect system assstirs the tool(s) of content personalization, that offers Artists Another upstream of insight data, as well as, enhancing a parasitism watching experience, simply high-via the changing consumer needs.

REFERENCES

Agrawal, Shreya, and Pooja Jain. "An improved approach for movie recommendation system." 2017 international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC). IEEE, 2017.

Boddapati, Mohan Sai Dinesh, et al. "Creating a Protected Virtual Learning Space: A Comprehensive Strategy for Security and User Experience in Online Education." International Conference on Cognitive Computing and Cyber Physical Systems. Cham: Springer Nature Switzerland, 2023.

Burke, Robin. "Hybrid recommender systems: Survey and experiments." User modeling and user-adapted interaction 12 (2002): 331-370.

Chakraborty, Debarun, et al. "Watching is valuable: Consumer views–Content consumption on OTT platforms." Journal of Retailing and Consumer Services 70 (2023): 103148.

Goyani, Mahesh, and Neha Chaurasiya. "A review of movie recommendation system: Limitations, Survey and Challenges." ELCVIA: electronic letters on computer vision and image analysis 19.3 (2020): 0018-37.

Ko, Hyeyoung, et al. "A survey of recommendation systems: recommendation models, techniques, and application fields." Electronics 11.1 (2022): 141.

Koren, Yehuda, Steffen Rendle, and Robert Bell. "Advances in collaborative filtering." Recommender systems handbook (2021): 91-142.

Nagaraj, Samala, Soumya Singh, and Venkat Reddy Yasa. "Factors affecting consumers' willingness to subscribe to overthe-top (OTT) video streaming services in India." Technology in Society 65 (2021): 101534.

O'Brien, Heather L., and Elaine G. Toms. "What is user engagement? A conceptual framework for defining user engagement with technology." Journal of the American society for Information Science and Technology 59.6 (2008): 938-955.

Rajarajeswari, S., et al. "Movie recommendation system." Emerging Research in Computing, Information, Communication and Applications: ERCICA 2018, Volume 1. Springer Singapore, 2019.

Reddy, S.R. S., et al. "Content-based movie recommendation system using genre correlation." Smart Intelligent Computing and Applications: Proceedings of the Second International Conference on SCI 2018, Volume 2. Springer Singapore, 2019.

Sutcliffe, Alistair. Designing for user engagment: Aesthetic and attractive user interfaces. Springer Nature, 2022.

Wang, Zan, et al. "An improved collaborative movie recommendation system using computational intelligence." Journal of Visual Languages & Computing 25.6 (2014): 667675.

Wu, Hao, et al. "Dual-regularized matrix factorization with deep neural networks for recommender systems." KnowledgeBased Systems 145 (2018): 46-58.

Yunanda, Gisela, Dade Nurjanah, and Selly Meliana. "Recommendation system from microsoft news data using TFIDF and cosine similarity methods." Building of informatics, technology and science (BITS) 4.1 (2022): 277-284.