# AN ENHANCING MOVIE RECOMMENDATION SYSTEM USING HYBRID MODELS

**R.Kirubahari[1], P.Kiruthika[2], S.Lakshita[3], J.M.Namritha Shree[4]**

*[1]Associate Professor, Dept. of Computer Science and Engineering, KLN College of Engineering, Tamil Nadu, India*
*[234]UG student, Dept. of Computer Science and Engineering, KLN College of Engineering, Tamil Nadu, India*

---

**Abstract -** The cold start problem in recommendation systems, particularly in the movie recommendation domain, arises when the system has insufficient user data, making it challenging to deliver personalized and accurate suggestions. This issue is especially significant when dealing with new users or items, which lack historical interaction data. To mitigate this problem, by integrating text analysis techniques, specifically utilizing text blob and sentiment analysis algorithms, to extract meaningful insights from the minimal user input. These algorithms help analyze initial interactions, such as user reviews or feedback, to generate relevant movie recommendations early in the user's journey, even with sparse data. Additionally, recognizing that user preferences can fluctuate depending on their emotional states, the system integrates emotion recognition techniques to refine its recommendations.

***Key Words***: Cold start problem, Recommendation systems, Sentiment analysis, User preferences, Recommendation accuracy.

## 1. INTRODUCTION

The cold start problem is a prevalent challenge in recommendation systems, particularly in the movie recommendation domain, where systems struggle to provide personalized and accurate suggestions for users who lack historical interaction data. This problem is especially significant when dealing with new users or items, as the absence of sufficient data hampers the system's ability to deliver relevant recommendations. This system addresses the challenge of noise and missing data in recommendation systems, which negatively impacts recommendation accuracy. It proposes collaborative denoising auto-encoders to improve Top-N recommendations by effectively filtering out noisy information and reconstructing user-item interactions[4]. To address this issue, by integrating the text analysis techniques, specifically leveraging Text Blob and sentiment analysis algorithms, to extract valuable insights from minimal user input. By analyzing initial interactions such as user reviews or feedback, these algorithms help generate relevant movie recommendations early in the user journey, even when data is sparse.

To further enhance personalization, proposed model incorporates geographical location data, allowing the system to recognize region-specific movie trends, preferences, and cultural nuances that influence viewing habits. This geographical data enables the system to align recommendations with local tastes and regional popularity, improving the relevance of suggestions for users across different locations.

Recognizing the dynamic nature of user preferences, which may fluctuate based on emotional states, the system also integrates emotion recognition techniques. By analyzing the user's current emotional state from textual inputs or interactions, the system can recommend movies that either align with or counterbalance their emotions. For instance, users exhibiting signs of stress or sadness may be presented with uplifting or calming movie options, creating a more personalized and emotionally attuned experience.

### 1.1 PROBLEM STATEMENT

The cold Start Problem in Recommendation Systems Many existing recommendation systems struggle with the cold start problem, where they are unable to generate accurate suggestions for new users or items with limited interaction data. Recommender Systems face challenges like scalability, diversity, accuracy, and data sparsity, impacting prediction quality. Tuning the hyperparameters of Restricted Boltzmann Machines (RBM) is difficult and affects performance. The Improved RBM with Bayesian Optimization (IRBM-BO) is proposed to optimize hyperparameters like learning rate, momentum, and light-cost. This method enhances prediction accuracy, verified on datasets like Movie lens and Netflix using MAE and RMSE metrics [5]. This system addresses the limitations of traditional movie recommendation systems, such as cold start, data sparsity, and lack of diversity. It proposes a hybrid method combining social similarity and item attributes to improve recommendation accuracy and relevance. This approach aims to enhance user satisfaction by providing diverse and personalized movie suggestions[1]. The proposed model aims to address this issue by utilizing hybrid models that incorporate text sentiment analysis and geospatial data, enabling the system to provide relevant recommendations even in scenarios where little user history is available.

Lack of Emotion and Location Context in Traditional Models Current movie recommendation systems often fail to consider the emotional and locational context of users. This model identifies the gap in traditional recommendation engines that solely rely on historical data, ignoring how a user's current emotional state or geographic location might influence their movie preferences. The proposed system integrates emotion detection and geospatial analysis to refine its suggestions based on real-time user contexts. Computational Complexity and Feedback Dependency Existing systems that attempt to combine multiple filtering techniques, like collaborative and content-based filtering, often face increased computational complexity and rely heavily on user feedback for accuracy. This model tackles the issue by designing a hybrid model that balances performance and feedback dependency while ensuring computational efficiency, minimizing.

### 1.2 OBJECTIVE

The primary goal of this model is to improve the accuracy of movie recommendations using a hybrid approach. This system addresses the limitations of existing tourism recommendation systems, such as inaccurate suggestions and lack of personalization. It proposes a hybrid filtering approach combined with association rule mining to enhance recommendation accuracy and relevance for users[3]. The proposed model involves integrating various techniques such as text sentiment analysis and geospatial data to tailor recommendations based on user preferences, emotional states, and geographical trends. By incorporating a more holistic analysis of user behavior, the system aims to deliver highly relevant suggestions that cater to both local and emotional contexts. This system addresses the item cold-start problem in recommendation systems, where new items lack sufficient data for accurate recommendations. It proposes a personalized feature selection method to improve the recommendation of new items by leveraging relevant user preferences[2]. The proposed model seeks to overcome the cold start problem, which typically occurs when there is insufficient initial user data to make accurate recommendations. Through the use of geospatial algorithms and sentiment analysis, the system can generate meaningful suggestions even with minimal user input. This enables the recommendation engine to function effectively for new users or when historical data is scarce. Another objective is to ensure that the recommendation system continuously adapts to users' changing environments and emotional states. By leveraging real-time data on user locations and emotions, the system offers personalized movie recommendations that evolve with the user, providing a dynamic and tailored viewing experience that maximizes user engagement.

## 2. PROCESS FLOW DIAGRAM

The flowchart represents the process flow of a hybrid movie recommendation system. It begins with collecting user input, such as preferences, emotional state, and location data. The next step involves data preprocessing, where the raw data is cleaned and prepared for analysis. The system then performs sentiment analysis to assess the emotional tone of the user's reviews, followed by geospatial analysis to evaluate location-based trends. These two analyses are combined in a hybrid model to create a personalized recommendation system. At the recommendation generation step, the system checks whether enough data is available. If not, it loops back to gather more input. If enough data is available, recommendations are generated. The system then collects user feedback to refine future recommendations, forming a feedback loop. The process ends once the recommendation generation and feedback are complete.

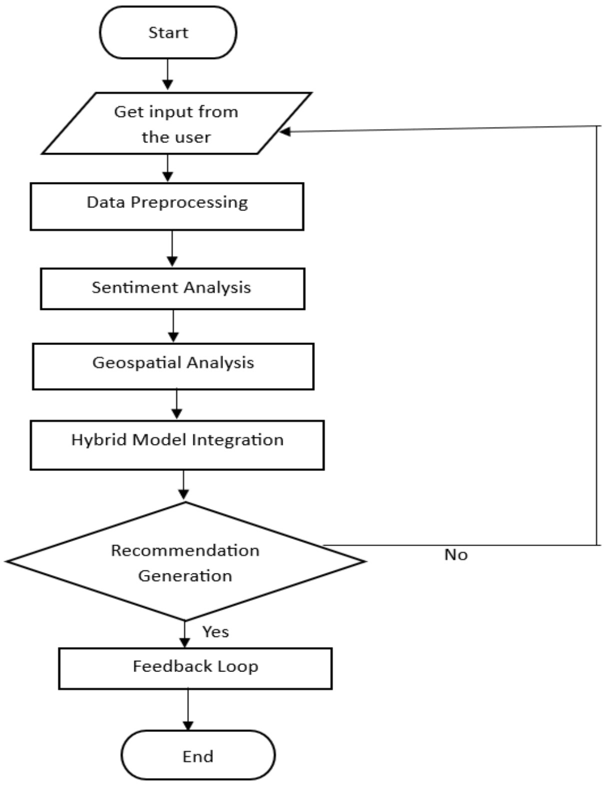


Figure 2.1: Flowchart of recommendation system

## 3. PROPOSED SYSTEM

The user selects the movie recommendation system, which integrates context-aware features such as user emotions and location to refine recommendations. By employing hybrid approaches like text blob sentiment analysis and geospatial algorithms, the system generates accurate recommendations for new users or for movies with limited interaction data. It collects and utilizes location data to tailor movie recommendations according to the user's current environment, such as home, work, or travel. The recommendation system continuously updates based on real-time changes in the user's location and emotional state, with a feedback loop to enhance system accuracy over time.

## 4. SYSTEM ARCHITECTURE

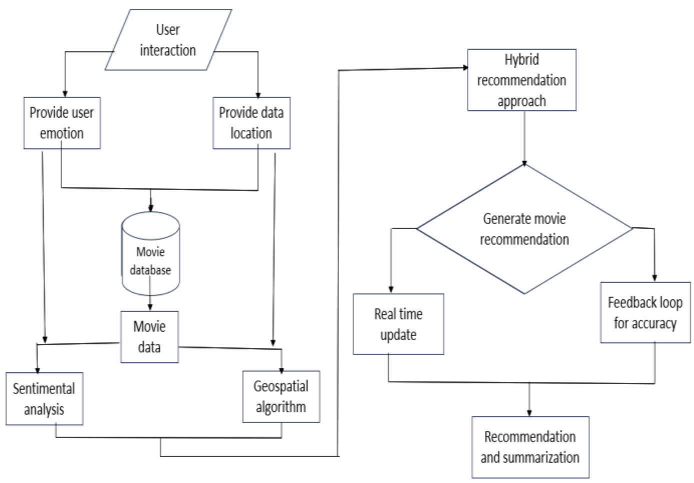


Figure 4.1: Architecture Diagram

## 5. MODULES

### 5.1 Data Collection

To collect demographic information and user preferences along with interaction data such as watch history, watchlists, and ratings. This data helps in building comprehensive user profiles that reflect individual preferences. Track detailed viewing history, including movies watched, timestamps, and duration. Engagement metrics like time spent on a movie, rewatches, and trailer views provide insights into user preferences and interests.

Gather data on users' social connections and their viewing habits, Collect feedback on recommended movies to refine the algorithms and enhance the accuracy of future suggestions. Utilize film attributes (genres, cast, and ratings) and integrate external data, such as critics' reviews and trending topics, to enrich the recommendation engine's understanding of movie quality and popularity. This combination aids in delivering more relevant recommendations.

### 5.2 Preprocessing

This involves removing duplicates, handling missing values, and correcting inconsistencies in the dataset, Normalization ensures that different scales of data are standardized, enabling more accurate comparisons across different data sources. Extracting relevant features from the raw data is crucial for enhancing the recommendation process. This may include identifying key attributes such as genre, cast, director, and release year from movie metadata, as well as user-specific features like demographic information and viewing habits. Selecting the most informative features helps improve the performance of the hybrid model. User profiles are created from historical viewing data, ratings, and interactions, while item profiles include attributes like genres and metadata. This helps the hybrid model understand user preferences and movie similarities. Data Transformation and Encoding is the process of converting raw data into a numerical format suitable for analysis and make predictions.

### 5.3 Sentiment Analysis

Text Blob can analyze user-generated movie reviews to extract sentiment scores, determining whether the sentiments expressed are positive, negative, or neutral. This sentiment analysis helps gauge overall viewer satisfaction and can be integrated into user profiles to better understand individual preferences and Emotional responses to films. By applying Text Blob to movie descriptions, synopses, and reviews, sentiment scores can be derived and added as additional features in item profiles. These sentiment features can help the hybrid model assess the overall reception of a movie, thereby influencing recommendations based on both qualitative (sentiment) and quantitative (ratings) data Incorporating sentiment analysis allows the hybrid model to offer more accurate recommendations by considering not only user ratings but also the underlying sentiments in reviews.

### 5.4 Geospatial Clustering

Geospatial clustering allows for the identification of location-based viewing trends, enabling recommendations tailored to users' geographical contexts. By clustering users based on their geographic locations, the model can analyze regional movie popularity and preferences. This information can enhance recommendations by highlighting films that are trending or well-received in a user's area, making the suggestions more timely and appealing. Geospatial clustering can capture cultural differences that influence movie preferences. By understanding the cultural context of specific regions, the recommendation system can suggest films that resonate with local values, traditions, and social norms, thereby improving user satisfaction. As users travel or relocate, geospatial clustering enables the recommendation system to adapt dynamically to their new environments.

### 5.5 Hybrid Model

Geospatial data allows the recommendation system to consider the user's location when suggesting movies. For instance, the system can recommend films that are popular in the user's area or those that feature local landmarks, enhancing the relevance of the suggestions based on regional preferences and cultural context. Utilizing text blob sentiment analysis enables the system to evaluate the emotional tone of user reviews and comments about movies By analyzing sentiment scores, the model can prioritize movies with positive sentiments from users similar to the target user, thereby improving the accuracy of recommendations based on overall viewer satisfaction. The hybrid model can merge insights from both geospatial data and sentiment analysis to generate nuanced recommendations. By continuously analyzing incoming geospatial and sentiment data, the hybrid model can provide real-time updates to recommendations

## 6. SAMPLE OUTPUT

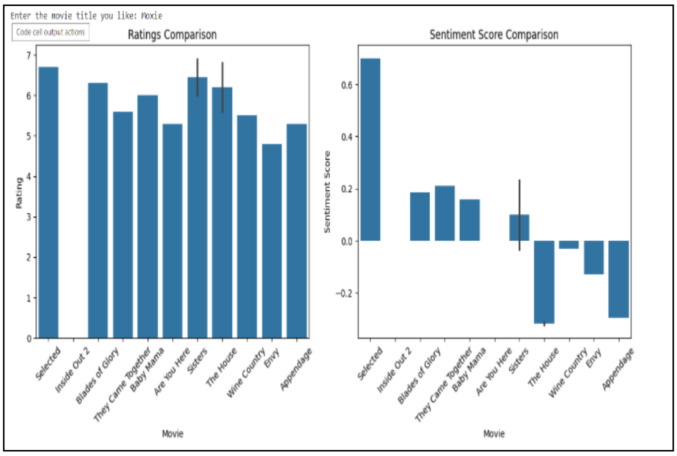


Figure 6.1 : Rating Comparison and Sentiment Score Comparison

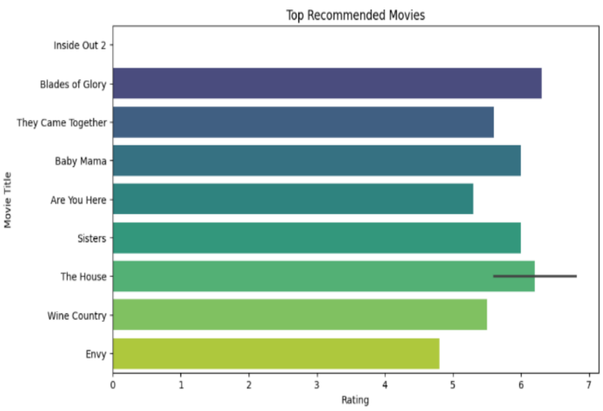


Figure 6.2 : Proposed model-Top Recommended Movies

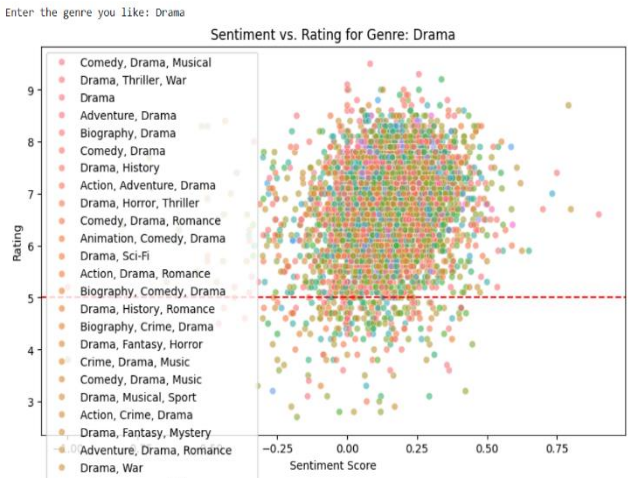


Figure 6.3 : Sentiment Vs Rating for Genre

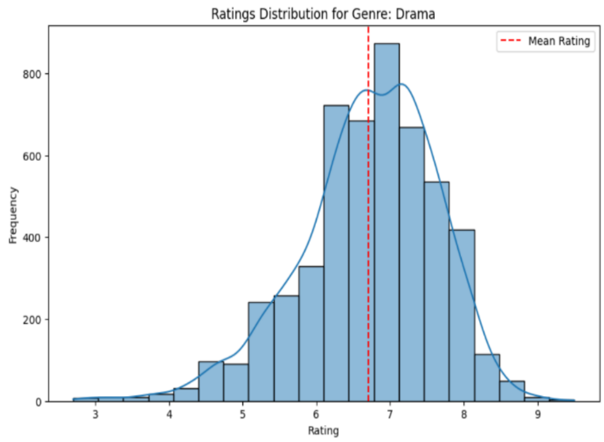


Figure 6.4 : Rating Distribution for Genre

## 7. CONCLUSIONS

The proposed hybrid movie recommendation system built on hybrid models combining sentiment analysis and geospatial data offers a significant improvement over traditional recommendation systems. By addressing the cold-start problem, it ensures that users are presented with relevant recommendations even in the absence of extensive historical data. This system has the capability to personalize recommendations based on both emotional and locational factors, making the recommendations more contextually appropriate and user-centric. The proposed system not only solves current challenges faced by recommendation engines but also opens the door for future advancements. As user data continues to grow, the model can continuously refine its algorithms, delivering more precise recommendations and improving user satisfaction. Its integration of cutting-edge AI techniques, like NLP and geospatial clustering, sets it apart from existing models, offering a more holistic approach to content recommendations. Overall, the hybrid movie recommendation system provides a flexible and scalable solution to enhance user engagement on streaming platforms. With its adaptive feedback loop and real-time updates, it promises to remain relevant and effective in the rapidly evolving landscape of digital media consumption.

## 8. FUTURE ENHANCEMENT

In the future enhancement of the proposed model, the movie recommendation system could incorporate advanced machine learning techniques, such as deep learning models, to improve recommendation accuracy. By integrating reinforcement learning, the system could evolve based on user interactions over time, continuously learning and adapting to preferences. Expanding the emotional recognition component could also involve facial recognition and voice tone analysis, allowing the system to assess emotional states in real-time for more personalized recommendations. Another possible enhancement would be the inclusion of social media integration. By analysing data from platforms like Twitter, Instagram, or Facebook, the system could detect trending movies, allowing it to stay current and recommend popular films based on real-time user sentiments.. Furthermore, the system could be expanded to include additional content types like TV shows, web series, or documentaries. By integrating data from different content formats, the recommendation engine could suggest relevant content across multiple categories, enhancing user engagement and satisfaction.

## REFERENCES

[1] C. Yang, X. Chen, L. Liu, T. Liu, and S. Geng, ''A hybrid movie recommendation method based on social similarity and item attributes,'' in Proc. 9th Int. Conf. Adv. Swarm Intell., Shanghai, China, 2018.

[2] Y.-F. Chen, X. Zhao, J.-Y. Liu, B. Ge, and W.-M. Zhang, "Item coldstart recommendation with personalized feature selection," J. Comput. Sci. Technol., vol. 35, no. 5, pp. 1217–1230, 2020.

[3] M. Gandhi and G. Sonali, "An enhanced approach for tourism recommendation system using hybrid filtering and association rule mining," Asian J. Convergence Technol., vol. 36, no. 1, pp. 1–4, 2019.

[4] Y.Wu, C. DuBois, A. X. Zheng, and M. Ester, ''Collaborative denoising auto-encoders for Top-N recommender systems,'' in Proc. 9th ACM Int. Conf. Web Search Data Mining (WSDM), 2016, pp. 153–162.

[5] R. Kirubahari and S. Miruna Joe Amali(2024). "An Improved Restricted Boltzmann Machine Using Bayesian Optimization for Recommender Systems." Systems, Vol.15(3),pp.1099-1111 (https://doi.org/10.1007/s12530-023-09520-1).