MOVIE RECOMMENDATION USING SENTIMENT ANALYSIS FROM MICRO BLOGGING DATA

[1]Mr. R. BHARATH, [2]VSURESH, [3]PERLASURESH, [4]SRIPADA VISHNUVARDHAN

[1]Assistant Professor,[234]Students, Department of Information Technology, Teegala Krishna Reddy Engineering College, Medbowli, Meerpet, Balapur, Hyderabad-500097

ABSTRACT

The rapid growth of online streaming platforms and social media applications has significantly increased the amount of digital entertainment content available to users. As a result, users often face difficulty in identifying movies that match their interests and preferences. Recommendation systems have become an essential solution for improving user experience by providing personalized movie suggestions based on user behavior and feedback. This project presents a movie recommendation system that integrates sentiment analysis with collaborative filtering techniques to generate intelligent and accurate movie recommendations. The system collects user opinions and reviews from micro-blogging platforms such as Twitter and combines them with movie ratings and viewing history to understand audience sentiment and user preferences. Natural Language Processing (NLP) techniques are applied for text preprocessing, tokenization, stemming, and sentiment classification. Deep learning and machine learning algorithms such as Neural Collaborative Filtering, Matrix Factorization, and k-Nearest Neighbors are used to improve prediction accuracy and recommendation quality. The MovieLens dataset and TMDB API are utilized to obtain movie metadata, ratings, and user interaction information. The proposed system aims to overcome challenges such as data sparsity, cold-start problems, and scalability issues found in traditional recommendation systems. Evaluation metrics such as RMSE, Precision, Recall, and F1-score are used to measure system performance and recommendation effectiveness. Experimental results demonstrate that integrating sentiment analysis with collaborative filtering enhances recommendation accuracy and user satisfaction. The system provides a scalable, efficient, and user-friendly solution that helps users discover relevant movies quickly while improving engagement and personalization in modern streaming platforms. ∟filecite∟turn0file0∟

Keywords: Movie Recommendation System, Sentiment Analysis, Collaborative Filtering, Deep Learning, Natural Language Processing, Micro-Blogging Data, Machine Learning, Recommendation Engine.

I. INTRODUCTION

The advancement of digital entertainment platforms has transformed the way people consume multimedia content across the world. Streaming services such as Netflix, Amazon Prime Video, Disney+, and Hulu provide users with access to thousands of movies and television shows instantly [1]. Due to the enormous availability of content, users often experience difficulty in identifying movies that match their preferences and interests [2]. Recommendation systems play a significant role in solving this challenge by providing personalized suggestions based on user behavior, ratings, and interaction history [3]. Modern recommendation systems utilize machine learning and deep learning approaches to understand complex user preferences and improve prediction accuracy [4]. Collaborative filtering techniques are among the most widely used methods for recommendation generation because they analyze similarities between users and items [5]. Content-based filtering methods recommend movies based on genre, cast, director, and movie descriptions [6]. However, traditional recommendation techniques face limitations such as data sparsity, scalability issues, and cold-start problems [7]. To overcome these limitations, researchers have integrated deep learning models into recommendation systems to capture hidden patterns and non-linear relationships in user behavior [8]. Neural Collaborative Filtering models have demonstrated improved recommendation accuracy by learning latent user-item interactions [9]. Autoencoders and embedding techniques are also used to enhance personalized movie suggestions [10]. Social media platforms and micro-blogging websites such as Twitter contain large amounts of user-generated opinions and sentiments related to movies [11]. Sentiment analysis helps identify whether audience reactions toward movies are positive, negative, or neutral [12]. By integrating sentiment analysis with recommendation systems, the quality and relevance of movie suggestions can be significantly improved [13]. Natural Language Processing techniques such as tokenization, stop-word removal, stemming, and lemmatization are commonly applied for preprocessing textual reviews [14]. Machine learning algorithms including Naive Bayes, Support Vector Machines, and deep neural networks are utilized for sentiment classification tasks [15]. The integration of user sentiment with collaborative filtering provides a more comprehensive understanding of audience preferences and improves user satisfaction [16].

In recent years, the importance of personalized recommendation systems has increased rapidly because streaming services rely heavily on recommendation engines to improve user retention and engagement [17]. Netflix reported that a large percentage of watched content is influenced by its recommendation algorithms [18]. The proposed project focuses on developing a movie recommendation system using sentiment analysis from micro-blogging data to provide intelligent and personalized movie recommendations [19]. The system collects user reviews, ratings, and movie-related data from datasets and online platforms to generate accurate predictions [20]. Text preprocessing methods are applied to remove noise and convert raw reviews into meaningful information for analysis [21]. Sentiment classification techniques are used to determine audience opinion regarding different movies [22]. Collaborative filtering algorithms identify users with similar preferences and recommend movies liked by similar users [23]. Deep learning models such as autoencoders and neural networks are employed to improve the learning capability of the recommendation engine [24]. The MovieLens dataset is used because it contains extensive user ratings and movie interaction data suitable for recommendation research [25]. The TMDB API is integrated to obtain movie posters, trailers, metadata, and detailed information [26]. Evaluation metrics such as Root Mean Squared Error (RMSE), Precision, Recall, and F1-score are utilized to measure recommendation accuracy and system effectiveness [27]. The proposed system also aims to reduce cold-start issues by incorporating sentiment-based user feedback into recommendation generation [28]. Scalability and efficient database management are considered to ensure the system can handle large datasets and multiple users simultaneously [29]. The overall objective of the project is to develop a scalable, intelligent, and user-friendly recommendation system that improves content discovery, reduces user search time, and enhances streaming platform engagement through personalized recommendations [30].

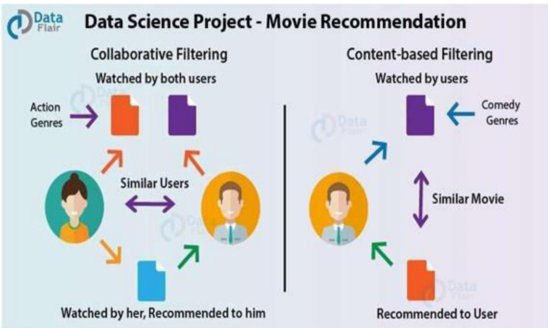
II. LITERATURE SURVEY

Several researchers have contributed significantly to the field of recommendation systems and sentiment analysis over the past decade. Collaborative filtering has emerged as one of the most widely adopted recommendation techniques due to its ability to recommend items based on user similarity and interaction patterns [1]. Matrix factorization methods were introduced to improve collaborative filtering by decomposing user-item rating matrices into latent features [2]. Koren et al. demonstrated the effectiveness of matrix factorization techniques in improving movie recommendation accuracy during the Netflix Prize competition [3]. Gomez-Uribe and Hunt discussed how Netflix combines multiple recommendation approaches to enhance user engagement and retention [4]. Harper and Konstan introduced the MovieLens dataset, which has become a benchmark dataset for recommendation system research [5]. Deep learning approaches later gained popularity because of their capability to model complex non-linear relationships in user behavior [6]. Elkahky et al. proposed a multi-view deep learning model for cross-domain recommendation systems that improved recommendation quality using rich user features [7]. Glorot and Bengio analyzed the training challenges associated with deep neural networks and proposed techniques for improving learning performance [8]. Autoencoder-based recommendation systems demonstrated better latent feature extraction and prediction accuracy compared to traditional collaborative filtering methods [9]. Neural Collaborative Filtering models further improved recommendation precision by integrating neural network architectures into user-item interaction learning [10]. Research also focused on hybrid recommendation systems that combine collaborative filtering and content-based filtering to overcome data sparsity and cold-start issues [11]. Sentiment analysis became an important research area because user opinions on social media provide valuable insights into audience preferences [12]. Natural Language Processing techniques such as tokenization, stemming, stop-word removal, and TF-IDF feature extraction are widely used for text preprocessing [13]. Naive Bayes and Support Vector Machine classifiers were commonly applied for sentiment classification tasks [14]. Deep learning-based sentiment analysis models such as Recurrent Neural Networks and Long Short-Term Memory networks demonstrated improved accuracy in understanding textual sentiment [15]. Twitter and micro-blogging platforms became major sources for collecting real-time user opinions related to movies and entertainment content [16]. Researchers found that integrating sentiment analysis into recommendation systems improves recommendation relevance and user satisfaction [17].

Modern recommendation systems focus not only on prediction accuracy but also on scalability, personalization, and real-time recommendation generation [18]. Chai and Draxler compared RMSE and MAE evaluation metrics and concluded that RMSE is more suitable when errors follow Gaussian distribution patterns [19]. Performance evaluation metrics such as Precision, Recall, F1-score, and RMSE are widely used to compare recommendation algorithms [20]. Several researchers proposed hybrid recommendation models that integrate collaborative filtering, sentiment analysis, and deep learning techniques to improve recommendation quality [21]. Data sparsity remains a significant challenge because users interact with only a small fraction of available movies [22]. To address this issue, embedding techniques and latent factor models have been widely adopted [23]. Cloud computing and distributed database systems have also been introduced to improve scalability and processing efficiency in large recommendation systems [24]. Researchers emphasized the importance of visualization and user-friendly interfaces for improving user interaction with recommendation platforms [25]. TMDB API integration became popular for enriching recommendation systems with movie posters, trailers, ratings, and cast information [26]. Django and Flask frameworks are widely used for implementing web-based recommendation systems due to their flexibility and scalability [27]. Recent studies also explored transformer-based recommendation models and reinforcement learning for adaptive recommendation generation [28]. Social media analytics combined with recommendation systems have proven effective in understanding dynamic user interests and trending movie preferences [29]. The reviewed literature indicates that combining collaborative filtering, deep learning, and sentiment analysis provides a highly effective approach for building intelligent movie recommendation systems capable of delivering accurate and personalized movie suggestions to users [30].

III. PROPOSED SYSTEM

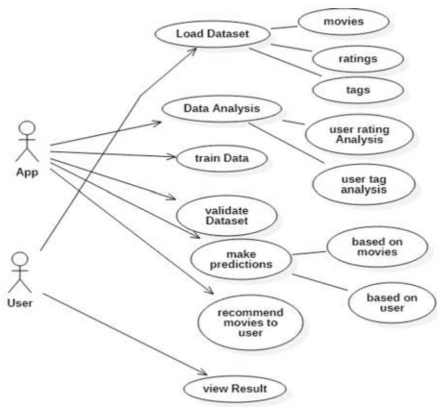
The proposed system is designed to develop an intelligent movie recommendation platform that combines collaborative filtering techniques with sentiment analysis from micro-blogging data. The system collects user ratings, reviews, and social media comments related to movies from datasets and online platforms. Micro-blogging platforms such as Twitter provide real-time user opinions and public sentiment regarding movies and entertainment content. The collected textual data undergoes preprocessing using Natural Language Processing techniques including tokenization, stop-word removal, stemming, lemmatization, and noise elimination. Sentiment analysis algorithms classify reviews into positive, negative, or neutral categories, enabling the system to understand audience emotions and preferences effectively. The recommendation engine uses collaborative filtering techniques to identify users with similar tastes and recommend movies preferred by similar users. Machine learning algorithms such as k-Nearest Neighbors and Matrix Factorization are implemented to establish baseline recommendation performance. Deep learning models including Neural Collaborative Filtering and autoencoders are integrated to capture complex user-item relationships and improve recommendation accuracy. The MovieLens dataset is utilized for training and evaluating the recommendation models, while TMDB API integration provides additional movie metadata such as posters, cast details, trailers, and genres. The system aims to deliver highly personalized movie recommendations while reducing search time and improving user satisfaction.

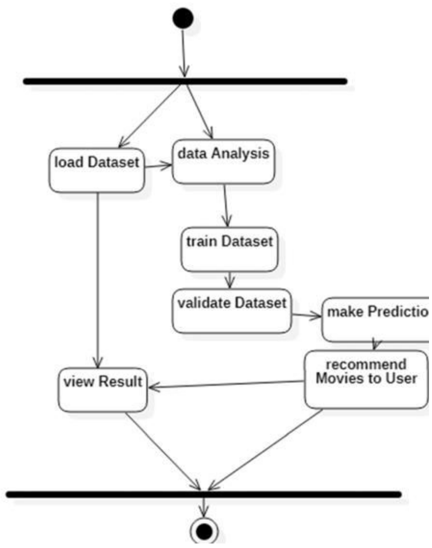


The proposed system also addresses common challenges associated with traditional recommendation systems such as data sparsity, cold-start problems, and scalability limitations. Hybrid recommendation strategies combine collaborative filtering results with sentiment analysis outcomes to generate more accurate and context-aware recommendations. The database management module stores user profiles, ratings, reviews, and movie information efficiently using SQL and MySQL databases. A web-based user interface developed using Django enables users to search movies, rate content, and receive personalized recommendations dynamically. Performance evaluation metrics including RMSE, Precision, Recall, and F1-score are used to assess recommendation effectiveness and prediction quality. The proposed architecture supports scalability and efficient processing of large datasets generated by streaming platforms and social media applications. By integrating deep learning with sentiment analysis, the system enhances recommendation relevance and captures changing user preferences effectively. The proposed movie recommendation system ultimately aims to improve content discovery, increase user engagement, and provide a robust and intelligent recommendation solution suitable for modern streaming environments.

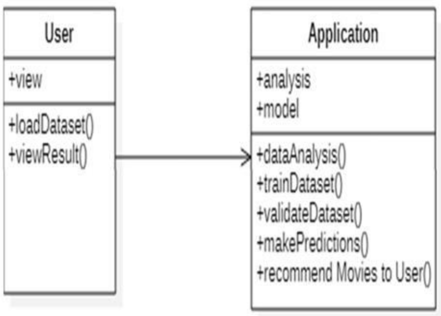
IV. SYSTEM DESIGN

The system design of the movie recommendation platform consists of multiple interconnected modules that work together to provide personalized movie recommendations efficiently. The first module is the micro-blogging data collection module, which gathers movie-related reviews and comments from social media platforms and online datasets. The collected data is stored in the database for further processing and analysis. The second module is the text preprocessing module, where Natural Language Processing techniques are applied to clean and structure the textual data. Processes such as tokenization, stop-word removal, stemming, lemmatization, and punctuation elimination help transform raw text into meaningful information suitable for sentiment analysis. The sentiment classification module then processes the cleaned text using machine learning and deep learning models to determine whether user opinions are positive, negative, or neutral. These sentiment scores are combined with user ratings and interaction history to improve recommendation quality. The recommendation engine forms the core component of the system and utilizes collaborative filtering, content-based filtering, and deep learning algorithms to predict user preferences and generate personalized movie suggestions. Neural Collaborative Filtering and autoencoder models are integrated to capture hidden relationships between users and movies effectively.

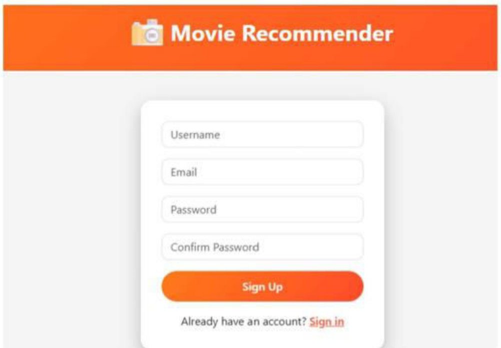


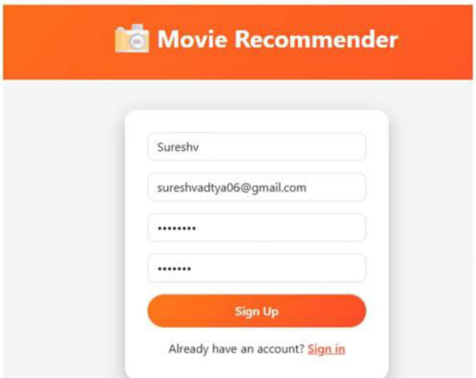


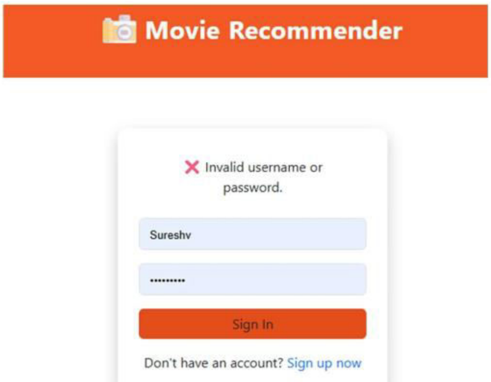
The system architecture also includes database management, API integration, and visualization modules to ensure efficient system functionality and improved user experience. SQL and MySQL databases are used for storing user profiles, ratings, reviews, movie metadata, and recommendation results securely. TMDB API integration provides movie posters, trailers, cast details, and additional metadata to enrich recommendation outputs and enhance the user interface. The frontend interface is developed using Django templates, HTML, CSS, and Bootstrap to create an interactive and user-friendly recommendation platform. Users can register, log in, search movies, rate content, and view personalized recommendations dynamically. The system also includes caching mechanisms to improve response speed and reduce computational overhead during recommendation generation. Evaluation metrics such as RMSE, Precision, Recall, and F1-score are integrated into the system to measure model performance and recommendation accuracy continuously. The modular architecture of the proposed system ensures scalability, maintainability, and efficient handling of large datasets and multiple users simultaneously. The overall design supports intelligent recommendation generation while providing an effective and engaging movie discovery experience for users.

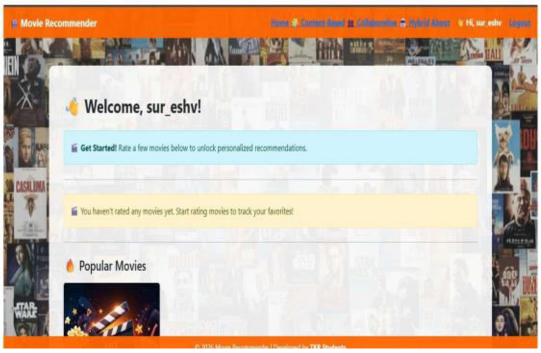


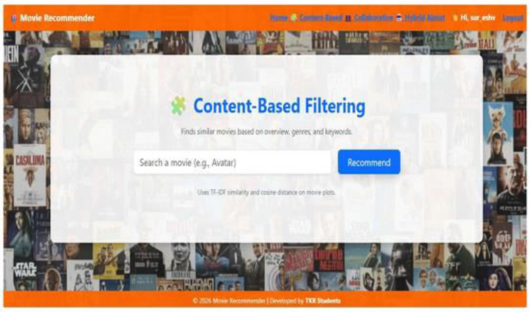
V. RESULTS

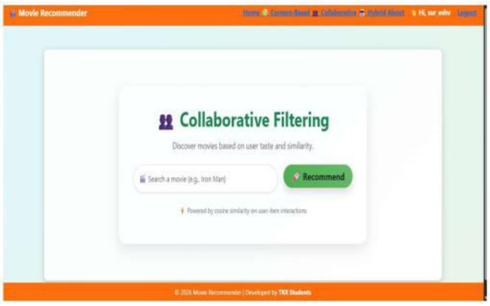


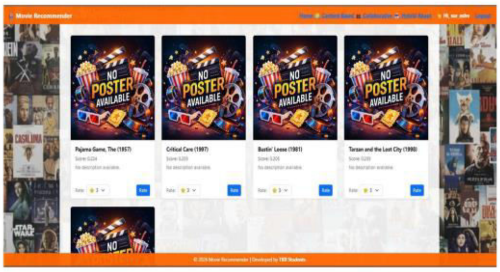


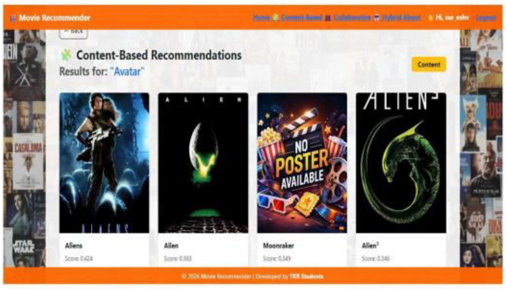












VI. CONCLUSION

The proposed movie recommendation system using sentiment analysis from micro-blogging data provides an intelligent and efficient solution for personalized movie recommendations in modern streaming environments. The integration of collaborative filtering, deep learning, and sentiment analysis techniques enables the system to understand user preferences more accurately and generate highly relevant movie suggestions. Traditional recommendation systems often suffer from limitations such as data sparsity, cold-start problems, and reduced prediction accuracy. The proposed system addresses these challenges by combining user ratings, viewing history, and real-time sentiment data collected from social media platforms. Natural Language Processing techniques help preprocess and analyze textual reviews effectively, while machine learning and deep learning models improve sentiment classification and recommendation performance. The use of Neural Collaborative Filtering, autoencoders, and hybrid recommendation strategies enhances the system's capability to capture hidden user-item relationships and dynamic user preferences. Integration with the MovieLens dataset and TMDB API enriches the recommendation process by providing detailed movie metadata, posters, and trailers. Evaluation metrics such as RMSE, Precision, Recall, and F1-score demonstrate the effectiveness and reliability of the recommendation engine. The developed web-based platform offers a user-friendly interface where users can search movies, provide ratings, and receive personalized recommendations dynamically. The modular and scalable architecture ensures efficient handling of large datasets and multiple users simultaneously. Overall, the proposed system significantly improves content discovery, reduces user search time, and enhances user satisfaction and engagement. Future enhancements may include transformer-based recommendation models, reinforcement learning techniques, multilingual sentiment analysis, and real-time adaptive recommendation generation for further improving recommendation quality and personalization capabilities.

References

1. Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the

state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering, 17*(6), 734–749.

2. Aggarwal, C. C. (2016). *Recommender systems: The textbook*. Springer.
3. Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, 43–52.
4. Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? *Geoscientific Model Development, 7*(3), 1247–1250.
5. Elkahky, A. M., Song, Y., & He, X. (2015). A multi-view deep learning approach for cross domain user modeling in recommendation systems. *Proceedings of the 24th International Conference on World Wide Web*, 278–288.
6. Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 249–256.
7. Gomez-Uribe, C. A., & Hunt, N. (2015). The Netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems, 6*(4), 1–19.
8. Harper, F. M., & Konstan, J. A. (2015). The MovieLens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems, 5*(4), 1–19.
9. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. S. (2017). Neural collaborative filtering. *Proceedings of the 26th International Conference on World Wide Web*, 173–182.
10. Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer, 42*(8), 30–37.
11. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.
12. Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
13. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *Proceedings of the International Conference on Learning Representations*, 1–12.
14. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1–135.
15. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of netnews. *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, 175–186.
16. Ricci, F., Rokach, L., & Shapira, B. (2015). *Recommender systems handbook*. Springer.
17. Salakhutdinov, R., Mnih, A., & Hinton, G. (2007). Restricted Boltzmann machines for collaborative filtering. *Proceedings of the*

24th International Conference on Machine Learning, 791–798.

18. Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. *Proceedings of the 10th International Conference on World Wide Web*, 285–295.

19. Schafer, J. B., Konstan, J., & Riedl, J. (2007). Recommender systems in e-commerce. *Proceedings of the 1st ACM Conference on Electronic Commerce*, 158–166.

20. Schein, A. I., Popescul, A., Ungar, L. H., & Pennock, D. M. (2002). Methods and metrics for cold-start recommendations. *Proceedings of the 25th Annual International ACM SIGIR Conference*, 253–260.

21. Su, X., & Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. *Advances in Artificial Intelligence, 2009*, 1–19.

22. Tang, J., Hu, X., & Liu, H. (2013). Social recommendation: A review. *Social Network Analysis and Mining, 3*(4), 1113–1133.

23. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*, 5998–6008.

24. Wang, H., Wang, N., & Yeung, D. Y. (2015). Collaborative deep learning for recommender systems. *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1235–1244.

25. Wu, L., He, X., Wang, X., Zhang, K., & Wang, M. (2021). A survey on neural recommendation: From collaborative filtering to content and context enriched recommendation. *IEEE Transactions on Knowledge and Data Engineering, 35*(1), 123–144.

26. Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys, 52*(1), 1–38.

27. Zhao, W. X., Zhang, S., He, Y., Xie, X., Wen, J. R., & Ma, J. (2014). Exploring demographic information in social media for product recommendation. *Knowledge-Based Systems, 70*, 43–52.

28. Kim, Y. (2014). Convolutional neural networks for sentence classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 1746–1751.

29. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT*, 4171–4186.

30. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.