IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, April 2025
Impact Factor: 7.67

News Application with Personalized News Recommendation and OCR

Danish Shaikh, Tauseeb Siddiqui, Sufiyan Shaikh, Abuzar Gore
Students, Department of IT
MH Saboo Siddik College of Engineering, Mumbai, India

Abstract: *With the increasing shift towards digital news consumption, users are exposed to an overwhelming amount of content from various sources. News Recommender Systems (NRS) play a crucial role in addressing information overload by providing users with relevant and reliable news articles tailored to their interests. The internet has become the dominant medium for news consumption, offering real-time updates and free accessibility. According to a 2018 Pew Research Center survey, approximately 93% of U.S. adults consume news through digital platforms, including online newspapers, social media, and mobile applications. Despite advancements in technology, research suggests that the criteria for determining news relevance in digital media remain similar to those in traditional print journalism. This paper explores the most effective algorithm for a news recommendation system. It focuses on content-based and collaborative filtering approaches, which not only analyze real-time news but also build user interest profiles based on browsing behavior. These profiles enable personalized recommendations that align with individual preferences. The discussion is divided into two parts: the first section provides an overview of recommendation techniques, while the second presents a literature review. We evaluated multiple recommendation models, including the RBM model with an accuracy of 0.8238, the DKN model with 0.4652, the DMF model with 0.4603, and the BPR model achieving 0.9501 accuracy. This study highlights the importance of recommender systems in enhancing the online news consumption experience*

Keywords: Content recommendation, Collaborative filtering, Content filtering, Recommender systems, Personalization, News Recommendation System

I. INTRODUCTION

With the advancement of interactive communication technologies, the internet has become a primary source of news due to its unrestricted access, instant updates, and availability at any time and place. The shift from traditional print media to digital platforms has transformed how people consume news. According to a 2018 Pew Research Center survey, approximately 93% of adults in the United States access news through digital mediums such as mobile apps, social media, and online newspapers. Despite these technological advancements, studies indicate that the criteria used to determine the relevance and significance of news in digital media remain similar to those used in traditional print journalism. This growing reliance on online news sources necessitates the use of intelligent technologies, such as recommender systems, to provide users with personalized news updates based on their specific information needs.

As internet-based services continue to evolve, the amount of accessible information has expanded exponentially. Users can now interact with digital content by rating, reviewing, and commenting on news articles, leading to an influx of data that often results in information overload. Due to the overwhelming volume of news available online, there has been an increasing demand for efficient filtering techniques. This has fueled the development of recommender systems, which serve as intelligent filtering tools that deliver highly relevant content to users based on their preferences and past interactions.

The primary objective of a news recommendation system is to minimize the effort users must exert in searching for relevant information. By analyzing historical user behavior, recommender systems can personalize the news consumption experience. For instance, the Tapestry system enables users to classify emails according to their interests, allowing them to prioritize which messages to read. Similarly, content-based recommendation models focus on

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, April 2025
Impact Factor: 7.67

analyzing real-time news, constructing user interest profiles based on browsing behavior, and recommending articles that align with those interests but have not yet been accessed. The accuracy of such systems depends on developing precise user interest models and effectively measuring the similarity between different news articles to ensure relevant recommendations.

II. LITERATURE REVIEW

Recommender systems play a crucial role in filtering information based on user preferences. Raza et al. (2021) emphasize that these systems must account for both user and item attributes, along with their interactions. Content-based filtering (CBF) algorithms function by comparing user and item profiles within a shared attribute space, while collaborative filtering (CF) remains independent of content, relying on user behavior, such as ratings and interaction histories. Unlike CF, CBF dynamically updates user profiles, accommodating shifts in interests through continuous exposure to new content.

Roy et al. (2022) explore model-based approaches in recommender systems, introducing a neuro-fuzzy technique to identify relationships between user groups and item relevance. K-Means clustering proves effective for large datasets, while decision trees facilitate easy interpretation. Additionally, Bayesian Networks, grounded in conditional probability and Bayes' theorem, aid classification tasks. The authors highlight the CNN technique's capability to extract key dataset features without supervision, offering a computationally efficient and accurate recommendation mechanism.

Mitova et al. (2022) adopt a scoping study methodology to synthesize emerging research in recommender systems. Their study systematically classifies and organizes relevant literature, defining key criteria that structure the news recommendation system (NRS) domain. By grouping studies into overarching research perspectives, they provide a thematic synthesis that evaluates the depth, scope, and characteristics of NRS research.

Wang et al. (2023) introduce a personalized news recommendation system designed to mitigate information overload. Built on the Django framework with frontend optimization via Bootstrap and jQuery, the system utilizes Flask, HTML, and JavaScript for backend communication. Their study proposes an enhanced collaborative filtering algorithm, incorporating a temporal attenuation factor to refine recommendations. Experimental results indicate superior recall, accuracy, and F1-score compared to existing models. The authors elaborate on the system's architecture, covering frontend-backend integration and framework operations.

Tran et al. (2022) present CupMar, a deep learning-based personalized news recommender model. This model integrates contextual user profiles with multi-aspect article representations. The research discusses related works, CupMar's architectural design, and the technical specifications of its News Encoder and User-Profile Encoder. The study also details model experimentation and evaluation, with its source code made publicly available for further research.

Ali et al. (2021) explore collaborative filtering-based recommendation systems, analyzing various methodologies, including content-based and hybrid approaches. Their study identifies the strengths and limitations of collaborative filtering while addressing the challenges associated with its implementation.

Andreeaa et al. (2022) provide a comprehensive examination of key recommender models, covering content-based, collaborative filtering, and hybrid approaches. They discuss critical challenges such as data sparsity, cold-start problems, personalization, diversity, and privacy concerns. Their research suggests leveraging external data sources like knowledge graphs and ontologies to improve recommendation accuracy and mitigate common issues.

Shaoa et al. (2020) emphasize the significance of recommendation systems in alleviating information overload. They highlight how these systems serve as advanced filtering mechanisms, offering personalized content delivery and commercial value. Their study also addresses the interpretability challenges of recommendation models and calls for further research into designing more effective reward functions for optimizing user-agent interactions.

Joris et al. (2021) investigate audience perceptions of news selection mechanisms. Conducting a survey in Belgium, they examine how users perceive algorithm-driven news recommendations versus human editorial selection. Findings suggest skepticism towards algorithmic transparency and biases, while human editors are perceived as more reliable. The study underscores the importance of media transparency in algorithmic curation to maintain public trust.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, April 2025
Impact Factor: 7.67

Wang et al. (2020) review the application of graph learning in recommender systems, highlighting improvements in precision, reliability, and interpretability. They categorize common issues in graph-based recommendation models and analyze state-of-the-art methodologies. Their study explores two major types of knowledge-based graph recommender systems: Side Information-based KGRS (SKGRS) and Common Knowledge-based KGRS (CKGRS), which enhance recommendation accuracy through implicit relationship identification.

Milano et al. (2020) examine ethical concerns surrounding recommender systems, identifying key issues such as privacy, fairness, transparency, manipulation, and societal impact. They propose a taxonomy of ethical implications, distinguishing direct, indirect, and systemic effects. Their study also discusses the interdisciplinary nature of recommender ethics, with contributions from fields such as philosophy, anthropology, and technology studies.

Felfernig et al. (2021) explore recommender system applications across various domains, including software, financial services, movies, and books. Their study focuses on feature modeling in configuration systems, highlighting how machine learning aids in feature selection and recommendation refinement. The research compiles methodologies, application scenarios, and unresolved questions for future development.

Bangari et al. (2021) analyze the role of reinforcement learning in news recommendation systems. They compare multiple reinforcement learning algorithms, including Twin Delayed DDPG (TD3), Deep Deterministic Policy Gradient (DDPG), and Deep Q-learning Network (DQN), concluding that TD3 is the most effective. Their study highlights the benefits of reinforcement learning, particularly in enhancing long-term user engagement and adapting to dynamic user interactions.

Heitz et al. (2022) investigate the impact of news recommendations on democratic engagement. Their study tests various recommendation strategies using a news aggregation app featuring diverse political sources. Results indicate that exposure to diverse recommendations increases political awareness and engagement, reinforcing democratic values.

Yu et al. (2023) introduce FedNRM, a privacy-preserving personalized news recommendation model. Utilizing secure multi-party computation and pre-trained language models, the system enhances user data security while optimizing recommendation relevance. Their study presents experimental validations demonstrating the effectiveness of FedNRM in real-world applications.

Gao et al. (2020) propose an improved news recommendation algorithm

Comparison of Different Machine Learning Algorithms

The Deep Knowledge-aware Network (DKN) model, which uses semantic and knowledge-based information from news data has potential, however, it struggles with overfitting when working with scant data. The Dual Matrix Factorization (DMF) and Bayesian Personalized Ranking (BPR) models are also discussed in the study, with a focus on DMF's scaling problems as the user-item matrix expands and BPR's limitations in providing reliable recommendations for new users or items with scant interaction data.

Deepjyoti Roy and Mala Dutta provide a thorough analysis of recommender systems in another publication titled "A Systematic Review and Research Perspective on Recommender Systems." They discuss various recommendation algorithms, such as Convolutional Neural Networks (CNN), which is thought to be extremely effective for real-time product recommendations but has a high time complexity, and K-nearest neighbours (KNN), which is described as a scalable model with good precision but unsuitable for real-time applications. For those interested in learning more about different recommendation methods in the subject of recommender systems, this paper is an invaluable resource.

Diversity and Area Under Curve(AUC) Score:

Paper Name	Author(s)	Summary	Algorithm	AUC Score	Key Challenges

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, April 2025
Impact Factor: 7.67

A Regularized Model for Trade-off Between Accuracy and Diversity in a News Recommender System	Shaina Raza & Chen Ding	Uses **DKN** to fuse semantic and knowledge-based features for click prediction.	DKN	0.4652	Prone to overfitting when working with limited data.
A Regularized Model for Trade-off Between Accuracy and Diversity in a News Recommender System	Shaina Raza & Chen Ding	Collaborative Filtering (CF) method that identifies similarities among users and items for ranking recommendations.	DMF	0.4603	Becomes computationally expensive as the user-item matrix grows.
A Regularized Model for Trade-off Between Accuracy and Diversity in a News Recommender System	Shaina Raza & Chen Ding	Bayesian Personalized Ranking (BPR) for pairwise learning to enhance recommendations.	BPR	0.9501	Struggles to generate accurate recommendations for new users or items with limited interaction history.
A Regularized Model for Trade-off Between Accuracy and Diversity in a News Recommender System	Shaina Raza & Chen Ding	Restricted Boltzmann Machine (RBM)-based generative neural CF for improved recommendation accuracy.	RBM	0.8238	Computationally expensive to train, particularly on large datasets.
CupMar: A Deep Learning Model for Personalized News Recommendation	Dai Hoang Tran, Quan Z. Sheng, Wei Emma Zhang, Nguyen H. Tran, Nguyen Lu Dang Khoa	Factorization Machines (FM) for recommendation tasks with improved feature engineering.	FM	0.5661	Suffers from the "cold start" problem, especially when historical data is limited.
CupMar: A Deep Learning Model for Personalized News Recommendation	Dai Hoang Tran, Quan Z. Sheng, Wei Emma Zhang, Nguyen H. Tran, Nguyen Lu Dang Khoa	Neural Recommendation Model (NRMS) designed to capture complex user-item interactions for enhanced recommendations.	NRMS	0.6491	Struggles with sparse data where users interact with only a small subset of available items.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, April 2025
Impact Factor: 7.67

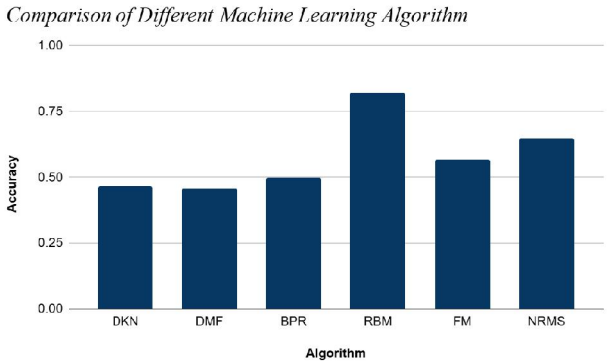


Figure 1 Caption: Various Machine Learning algorithms with accuracy comparison. Figure 1 Alt Text [85 words]: The image compares the accuracy of six algorithms: DKN, DMF, BPR, RBM, FM, and NRMS. The y-axis represents accuracy, ranging from 0.00 to 1.00, and the x-axis lists the algorithms. The bars show that BPR has the highest accuracy, 0.9501, while DKN and DMF have the lowest 0.4652 and 0.4603. RBM and FM have 0.8238 and 0.5661 and NRMS is 0.6491. The graph is used to visually represent the performance of these algorithms in terms of accuracy.

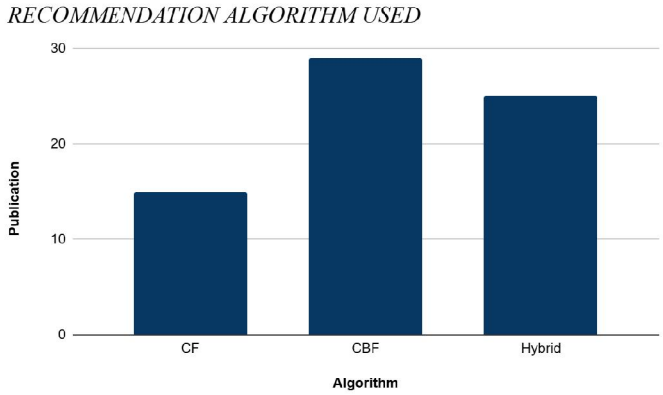


Figure 2 Caption: Algorithms used in NRS[1]

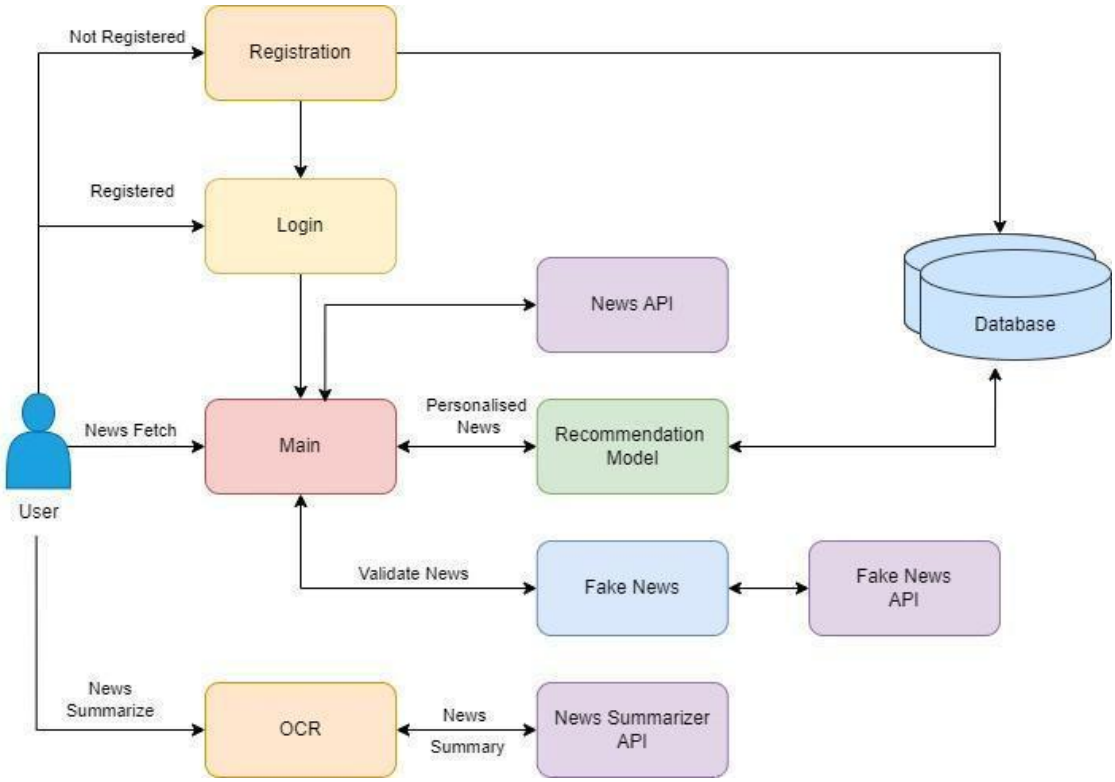
Figure 2 Alt Text [60 words]: The image shows the number of publications for each of the three recommendation algorithm types: Collaborative Filtering, Content-Based Filtering, and Hybrid. The y-axis is marked "PUBLICATIONS" with a range from 0 to 40, and the x-axis lists the algorithms. The blue bars indicate CF has around 15 publications, CBF has the highest with around 29, and Hybrid has about 25

III. PROPOSED SYSTEM

This research introduces an enhanced approach to recommendation systems by leveraging the Bayesian Personalized Ranking (BPR) algorithm. Rooted in Bayesian principles, the model emphasizes user-specific customization while efficiently processing implicit feedback data. By integrating matrix factorization with probabilistic inference, it ensures optimized computations and more accurate recommendations.

To enhance performance, the system incorporates automated hyperparameter tuning and evaluates results using multiple performance metrics. A comparative analysis against state-of-the-art algorithms highlights its effectiveness. Additionally, the model's scalability makes it well-suited for large-scale applications. In summary, the proposed BPR-based approach represents a significant advancement in recommendation system technology, with broad potential across various industries.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, April 2025
Impact Factor: 7.67



IV. CONCLUSION

As digital news consumption continues to evolve, this study delves into the dynamic landscape of news recommender systems (NRS). By analyzing various algorithms such as RBM, DKN, DMF, and BPR, the research aims to identify the most effective approach for delivering personalized news recommendations while addressing the challenge of information overload.

The literature review provides valuable insights into the diverse methodologies employed in recommender systems, including content-based filtering, collaborative filtering, and hybrid models, each with its strengths and limitations. Additionally, the study highlights key challenges faced by NRS, such as data sparsity, cold-start problems, personalization, diversity, and privacy concerns.

In conclusion, this research underscores the pivotal role of recommender systems in adapting to the rapidly shifting landscape of digital news consumption. By offering users relevant and reliable content, NRS serve as essential tools for navigating the overwhelming influx of online news. As advancements in this field continue, future developments are expected to introduce even more sophisticated and user-centric solutions for enhancing the news consumption experience.

V. DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

[1] Raza, S., & Ding, C. (2022). A comprehensive review of advancements, challenges, and future directions in news recommender systems. Artificial Intelligence Review, 1-52.

[2] Roy, D., & Dutta, M. (2022). Analyzing the evolution of recommender systems: A systematic literature review and future research directions. Journal of Big Data, 9(1), 59.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, April 2025
Impact Factor: 7.67

[3] Han, K. (2020). Enhancing personalized news recommendations through improved collaborative filtering techniques. Complexity, 2020, 1-12.
[4] Mitova, E., Blassnig, S., Strikovic, E., Urman, A., Hannak, A., de Vreese, C. H., & Esser, F. (2023). A structured analysis of research in news recommender systems. Annals of the International Communication Association, 47(1), 84-113.
[5] Wang, X., & Liu, C. (2023). Optimizing user preferences in news recommendations with enhanced collaborative filtering. Mobile Information Systems, 2023.
[6] Tran, D. H., Sheng, Q. Z., Zhang, W. E., Tran, N. H., & Khoa, N. L. D. (2023). CupMar: A deep learning-based contextual model for personalized news recommendations. World Wide Web, 26(2), 713-732.
[7] Ali, S. I. M., & Majeed, S. S. (2021). Evaluating the performance of collaborative filtering in modern recommendation systems. MJPS, 8(1).
[8] Iana, A., Alam, M., & Paulheim, H. (2022). A knowledge-driven perspective on news recommendation algorithms: A systematic survey. Semantic Web (Preprint), 1-62.
[9] Shao, B., Li, X., & Bian, G. (2021). Research trends in recommendation systems: Insights from a knowledge graph perspective. Expert Systems with Applications, 165, 113764.
[10] Joris, G., Grove, F. D., Van Damme, K., & De Marez, L. (2021). Understanding audience perceptions of algorithm-driven news selection. Digital Journalism, 9(5), 589-618.
[11] Wang, S., Hu, L., Wang, Y., He, X., Sheng, Q. Z., Orgun, M., & Yu, P. S. (2020). A review of graph learning techniques applied to recommendation systems. arXiv preprint arXiv:2004.11718.
[12] Milano, S., Taddeo, M., & Floridi, L. (2020). Ethical dilemmas in recommender systems and their societal impact. AI & Society, 35, 957-967.
[13] Felfernig, A., Le, V. M., Popescu, A., Uta, M., Tran, T. N. T., & Atas, M. (2021, February). Integrating machine learning and feature modeling for intelligent recommender systems. Proceedings of the 15th International Working Conference on Variability Modelling of Software-Intensive Systems, 1-8.
[14] Bangari, S., Nayak, S., Patel, L., & Rashmi, K. T. (2021, March). Examining the role of reinforcement learning in news recommendation systems and its associated challenges. Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 260-265. IEEE.
[15] Heitz, L., Lischka, J. A., Birrer, A., Paudel, B., Tolmeijer, S., Laugwitz, L., & Bernstein, A. (2022). Exploring the democratic benefits of diversity in news recommendations: A user study. Digital Journalism, 10(10), 1710-1730.
[16] Yu, S., Jie, Z., Wu, G., Zhang, H., & Shen, S. (2023). FedNRM: A privacy-preserving personalized news recommendation model using federated learning. Intelligent Automation & Soft Computing, 37(2).
[17] Gao, Y., Zhao, H., Zhou, Q., Qiu, M., & Liu, M. (2020, October). Improving text-based news recommendation through enhanced similarity algorithms. Proceedings of the 3rd International Conference on Smart Blockchain (SmartBlock), 132-136. IEEE.
[18] Raza, S., & Ding, C. (2020). Balancing accuracy and diversity in news recommendations using a regularized model. Proceedings of the 2020 IEEE International Conference on Big Data (Big Data). IEEE.