

Article

Not peer-reviewed version

Application of Recommendation System Technology and Architecture in Video Streaming Platforms

[Qin Yang](#) *

Posted Date: 8 January 2025

doi: 10.20944/preprints202412.1839.v2

Keywords: Recommendation systems; video streaming; personalized recommendation; deep learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Application of Recommendation System Technology and Architecture in Video Streaming Platforms

Qin Yang

Georgia Institute of Technology, USA; yqin0709@gmail.com

Abstract: With the rapid development of video streaming platforms and the diversification of user content consumption habits, personalized recommendation systems have gradually become a critical technology for enhancing user experience, increasing user engagement, and optimizing platform profitability. This paper systematically studies the technical architecture and application scenarios of recommendation systems in video streaming platforms, focusing on the unique challenges in handling large volumes of video data, analyzing user behavior, and delivering personalized content recommendations. By examining the overall architecture of recommendation systems, data processing, feature engineering, model design, and optimization strategies, this paper summarizes the application effects of various commonly used recommendation algorithms on video streaming platforms and explores the value of recommendation systems in these platforms through practical case studies. Additionally, this paper discusses the future challenges and potential technological breakthroughs in recommendation systems, such as deep learning and multi-modal feature fusion, reinforcement learning and real-time recommendations, as well as privacy protection and fairness research. The research findings indicate that a well-designed and optimized recommendation system can effectively enhance the depth of user content exploration and the overall activity level of the platform, providing strong support for the sustainable development of video streaming platforms.

Keywords: recommendation systems; video streaming; personalized recommendation; deep learning

1. Introduction

In recent years, the rapid advancement of internet technology and the growing demand for digital content consumption have led to the emergence of video streaming platforms. These platforms have become primary channels for accessing entertainment, education, news, and more. Compared to traditional distribution methods, video streaming offers advantages such as abundant content, real-time updates, and personalized recommendations, effectively meeting users' diverse needs. However, as user numbers and content variety expand, delivering accurate and personalized recommendations has become a critical challenge for enhancing user experience and competitiveness[1]. Recommendation systems have been widely adopted by major video streaming platforms, providing tailored content based on users' historical behavior and preferences. These systems significantly improve platform engagement, extend user session durations, and increase conversion rates[2]. However, designing and implementing effective recommendation systems for video streaming presents unique challenges. Unlike e-commerce product recommendations, video content is time-sensitive, emotionally impactful, and complex[3]. User behavior during video consumption—such as watch duration, retention rate, and skip behavior—is more intricate than in traditional text or product contexts. Efficiently processing and analyzing this data to uncover user interests and form effective recommendation strategies is crucial[4]. Moreover, with vast content libraries and complex user data, these systems must offer real-time responses and dynamically adjust recommendations to evolving user preferences. Additional challenges include cold start issues, data

sparsity, long-tail content distribution, and ensuring privacy and fairness, all of which heighten the demands on algorithm design and system architecture[5].

Recent advancements in related fields have provided insights and methodologies that can be applied to video streaming recommendation systems. For example, Wenqing et al.[6] propose MAT, combining Mamba and Transformer models to improve time series forecasting, which can be adapted for dynamic content recommendation. In a similar vein, DiaoSu et al.[7] introduce a deep sequence model for predicting mechanical ventilator pressure, offering relevant techniques for understanding user behavior patterns in complex, time-sensitive systems. Tangtang et al.[8] explore hybrid approaches using ARIMA and LSTM models for electricity price forecasting, shedding light on model selection strategies that can be useful for handling the inherent complexities of video recommendation systems. Meanwhile, Yimeng et al.[9] investigate machine learning techniques like PCA combined with XGBoost to enhance classification accuracy, providing valuable insights into improving recommendation models' precision. Moreover, Liumin et al.[10] propose a DeepFM-based transfer learning model with an attention mechanism, which could benefit video content recommendations by incorporating transfer learning strategies. Further, Xinyu et al.[11] develop deep learning models to enhance prediction accuracy, a technique with promising applications in predicting user preferences. Finally, Haosen et al.[12] offer a Regional Prior Fusion framework that could inform image processing techniques relevant to video feature extraction.

2. Overview of Recommendation System Technology

2.1. Basic Concepts of Recommendation Systems

Recommendation systems originated from information retrieval and filtering technologies, with the primary goal of solving the problem of how to filter out personalized content from massive datasets that matches users' preferences. The basic concepts of recommendation systems include four core elements: users, items, features, and models. "Users" refer to the individuals using the system, typically represented as viewers or platform users in the context of video streaming. "Items" are the content units that the recommendation system aims to suggest to users, such as movies, TV series, and short videos[13]. "Features" are the attributes or behavior data of both users and items, such as users' watch time, number of likes and comments, content tags, etc[14]. "Models" are algorithms and rules that analyze the relationship between user and item features to learn and predict the items that a user might be interested in. The working principle of a recommendation system generally involves several steps. The first is the data collection and processing stage, where the system gathers, cleans, and pre-processes user behavior data and item data, transforming them into feature data suitable for algorithm learning[15]. The next step is the model training stage, where the system uses historical data to train the recommendation model, learning the potential relationships between users and items to predict content that users might be interested in in the future. The final step is the generation and application of recommendation strategies, where the system dynamically adjusts the recommendation list based on the user's current status (e.g., the video content being watched, search keywords, etc.) and presents the recommendation results in a sorted or grouped format[9]. In practical applications, recommendation systems need to consider several factors, such as how to handle the cold start problem (when there is limited data about new users or items), data sparsity (when interaction data between users and items is sparse), long-tail effects (how to guide users to less popular or niche content), and privacy protection (how to achieve accurate recommendations without compromising user privacy)[16]. Effectively addressing these issues is key to the success of recommendation systems in video streaming platforms. A deep understanding of the basic concepts and working principles of recommendation systems provides a solid theoretical foundation for subsequent system architecture design and algorithm optimization.

2.2. Characteristics of Recommendation Systems in Video Streaming Platforms

The application of recommendation systems in video streaming platforms presents unique challenges compared to other content platforms due to content diversity, user behavior complexity, real-time performance requirements, and issues like cold start and data sparsity[17]. Video content encompasses multiple modalities, including visuals, audio, subtitles, and emotional elements, necessitating consideration of visual, semantic, and emotional features during data processing and feature extraction[18]. This multi-modal fusion relies on deep learning, natural language processing, and computer vision technologies to extract high-dimensional features from videos, which vary widely in length, subject matter, and style. User behavior data on video streaming platforms is also more intricate[19]. Explicit actions like clicks and likes are supplemented by implicit behaviors such as watch duration and skip frequency, reflecting user interests and habits. Modeling these implicit actions requires complex feature extraction and integration strategies, as users' viewing habits are often cyclical and dynamic, necessitating real-time updates to profiles and recommendation strategies based on current states. Real-time performance is essential, as users expect quick access to relevant content. Recommendation systems must rapidly process data, extract features, and generate predictions in a short timeframe. They should dynamically adapt to real-time feedback, requiring architectures that support caching, indexing, and parallel computing for responsiveness and scalability. Cold start and data sparsity are significant challenges as well. Cold start occurs when there is insufficient historical data for new users or items, hindering accurate recommendations. Initial user or content profiles can be established through content analysis and contextual recommendations[20]. Data sparsity, involving low interaction frequency with an extensive video library, complicates model effectiveness. Solutions include cross-domain recommendations and matrix completion strategies to enrich data. Lastly, managing long-tail content is vital for attracting niche audiences and increasing session duration. Recommendation strategies should balance the distribution of popular and long-tail content, guiding users to explore diverse options. In conclusion, recommendation systems for video streaming platforms must innovate in algorithms, architecture, and data processing to enhance effectiveness and competitiveness[21].

3. Architecture Design of Recommendation Systems for Video Streaming Platforms

3.1. Overall Architecture of Recommendation Systems

In the context of media convergence, the cultural tourism industry is adopting digital and interactive marketing strategies to adapt to changing consumer behaviors and stand out in a competitive market[22]. Digital strategies leverage technologies like big data, artificial intelligence (AI), and blockchain to achieve precision and personalization, while interactive strategies focus on engaging two-way communication between brands and consumers, enhancing audience loyalty and participation. Combining these two strategies not only expands brand influence but also improves the overall tourist experience, promoting the sustainable growth of the cultural tourism industry[23]. Firstly, the application of digital technologies significantly enhances marketing precision and management efficiency[24]. Cultural tourism enterprises can use big data to analyze visitor demographics, interests, and consumption habits, leading to more accurate audience targeting and personalized recommendations for tourism products or services. For instance, analyzing search records and online reviews can help predict visitor interest in specific attractions and send them relevant promotions. Moreover, digital technology can track visitor behavior within scenic spots, providing insights into tour paths and preferences, enabling the optimization of site layouts, product offerings, and service processes to improve visitor satisfaction[25]. Secondly, AI introduces new possibilities for intelligent marketing. AI-based recommendation systems can suggest tailored travel routes, attractions, and accommodations to visitors based on their preferences[26]. AI-powered chatbots can provide real-time assistance during tours, offering information on scenic spots,

navigation guidance, and weather updates, improving service quality. AI can also analyze social media comments and trending topics to forecast market trends and potential risks, providing valuable data support for decision-making. Thirdly, interactive marketing strategies can effectively boost brand loyalty and audience engagement. Today's consumers are not just passive receivers of information but active content creators through platforms like social media and short videos. Cultural tourism enterprises can launch interactive activities, such as travel-related discussions, challenges, and online campaigns, encouraging visitors to share their experiences and generate user-generated content (UGC). This enhances the relationship between visitors and the brand while attracting more potential customers, expanding brand influence and awareness. Establishing online communities or fan groups further promotes long-term interaction, fostering a sense of belonging and loyalty among visitors[27]. Lastly, the combination of digital and interactive marketing strategies can greatly enhance brand communication and competitiveness. For example, by analyzing visitor behavior with big data, enterprises can place targeted ads on various platforms, while interactive campaigns increase engagement and conversion rates. Integrating online interactions with offline experiences, such as hosting online events to attract visitors to offline scenic spots and then encouraging post-visit sharing, can create a seamless marketing loop. This model boosts visitor experiences, strengthens brand recognition, and promotes repeat visits. In summary, digital and interactive marketing strategies are crucial for cultural tourism enterprises seeking to innovate and thrive under media convergence. Digital technologies like big data and AI enable precise and personalized marketing, while interactive strategies strengthen connections with consumers. The combined approach not only differentiates cultural tourism enterprises in a complex market but also supports their sustainable development.

3.2. Data Processing and Feature Engineering in Video Streaming Recommendation Systems

Data processing and feature engineering are critical in constructing recommendation systems, particularly for video streaming platforms due to the complexity and variety of data. Effective strategies in these stages are necessary to accurately capture user behavior and video content characteristics, providing quality inputs for model training and generating recommendation strategies[28]. The main steps in data processing and feature engineering for video streaming platforms include data collection, preprocessing, multi-modal feature extraction, and feature selection. Firstly, data collection and preprocessing form the foundation. Video streaming platforms gather data from multiple sources, including user behavior, video content, and platform context. User behavior data, such as clicks, views, skips, comments, and likes, indicate user preferences[29]. Video content data, like titles, tags, descriptions, and subtitles, provide information on the video's theme and style. Platform context, such as device type and location, offers background data for user profiling. During data collection, completeness and accuracy must be ensured, while preprocessing addresses missing values, noise reduction, and data normalization to facilitate smooth feature engineering. Next, multi-modal feature extraction and fusion are crucial for video streaming recommendation systems, as they handle diverse content formats like text, image, audio, and video. Specific strategies are designed for each modality. For example, convolutional neural networks (CNN) extract visual features from images, word embedding models (e.g., Word2Vec) extract semantic features from text, and neural networks analyze audio characteristics. Effective fusion of these features, using methods such as concatenation or attention mechanisms, integrates diverse data into a unified feature space, enabling a comprehensive analysis of content and user preferences. After feature extraction, feature selection and interaction modeling further refine the data. High-dimensionality in video streaming data often includes redundant features, making feature selection essential to identify the most informative attributes. Techniques like variance analysis and L1 regularization help filter out less relevant features. In the feature interaction stage, models like factorization machines or deep and cross networks capture complex relationships between features. For instance, interactions between watch time, skip behavior, and content tags can reveal user preferences for specific video types. Additionally, incorporating temporal and contextual features

enhances recommendation accuracy. User behavior often shows temporal patterns, such as varying viewing habits between weekdays and weekends. Temporal models like RNN or LSTM can capture these patterns, allowing the system to dynamically adjust recommendations based on time. Contextual features like geographic location or device type further refine recommendations, ensuring content suitability (e.g., shorter videos for mobile users). In summary, data processing and feature engineering in video streaming recommendation systems involve complex processes, focusing on extracting and integrating multi-modal data while employing effective feature selection and interaction modeling. Properly designed strategies enable video streaming platforms to generate high-quality inputs for recommendation models, resulting in more accurate and personalized recommendations.

4. Application Scenarios of Recommendation Systems in Video Streaming Platforms

Recommendation systems are essential for video streaming platforms, serving various purposes such as personalized recommendations, content discovery, long-tail content suggestions, social interaction enhancements, and advertisement optimization. Firstly, personalized recommendations are a core application that significantly enhances user experience and engagement. By analyzing users' historical behavior, interests, and real-time states, these systems create customized content lists that entice users to click and watch. They dynamically adjust recommendation strategies to boost overall platform activity. Effective implementation relies on algorithms like collaborative filtering and deep learning, as well as real-time data processing systems to continuously update user profiles and accurately predict changes in interests. Secondly, content discovery and long-tail content recommendations guide users toward exploring niche or less popular content, increasing content diversity and traffic utilization. The recommendation system employs content feature analysis, random exploration strategies, and community-based suggestions to effectively distribute new and niche content. This enhances users' exploration depth and prevents overemphasis on popular content. Social interaction recommendations integrate users' social networks with content suggestions. By using strategies such as friend recommendations and highlighting trending content, these systems foster social engagement and interactivity. This approach encourages users to participate in discussions and interact with friends on the platform, reinforcing social attributes and user loyalty. In advertisement optimization, recommendation systems utilize user profiling and click-through rate (CTR) predictions to intelligently select ad content and timing based on users' interests, viewing history, and current contexts. This improves ad click-through and conversion rates while maintaining relevance between ads and video content. By considering ad duration and display formats, these systems minimize the negative impact of ads on the viewing experience, balancing platform ad revenue with user satisfaction. In summary, recommendation systems significantly enhance various application scenarios on video streaming platforms. They improve content distribution efficiency and user satisfaction while increasing the platform's commercial value and user engagement through intelligent recommendation strategies. As recommendation technology continues to evolve, these systems are expected to offer even more sophisticated and precise services in complex scenarios, further benefiting video streaming platforms.

5. The Frontiers of Recommendation System Technology in Video Streaming

The frontiers of recommendation system technology in video streaming focus on deep learning advancements, multi-modal feature fusion, reinforcement learning, privacy protection, and fairness research. Continuous progress in deep learning has led to the widespread application of models such as deep neural networks (DNN), convolutional neural networks (CNN), recurrent neural networks (RNN), and transformer-based models (e.g., BERT, GPT) in video streaming recommendations. These models automatically extract complex, non-linear features from user behavior and video content, enhancing the understanding of user interests and improving recommendation accuracy. For

instance, long short-term memory networks (LSTM) effectively model users' sequential behaviors, capturing dynamic changes in preferences over time. Multi-modal feature fusion is also a key research area. Video content often comprises various modalities, including text, images, audio, and video, which complement one another in expressing content attributes and user preferences. Researchers are exploring methods to fuse features from these different modalities and employing attention mechanisms to weight them, thereby enriching the system's understanding of content features. Reinforcement learning has made significant strides as well. Unlike traditional recommendation systems that rely on static modeling with historical data, reinforcement learning dynamically adjusts strategies through real-time user interactions, allowing adaptation to changing preferences. For example, exploration-exploitation strategies based on the multi-armed bandit model optimize recommendations in dynamic environments. With increasing concerns over data privacy and fairness, recommendation system technology is being applied to address these issues. Federated learning facilitates decentralized model training, protecting user privacy while enhancing performance. Differential privacy mechanisms safeguard sensitive data during training and recommendation. Additionally, fairness research focuses on designing de-biasing algorithms and metrics to prevent bias based on attributes like gender or race, enhancing credibility. In conclusion, advancements in recommendation system technology for video streaming span model optimization to data protection, offering robust solutions for dynamic environments. These developments will continue to improve content distribution, user experience, and commercial value in the future.

6. Conclusion

Economic transformation is a crucial strategy for driving sustainable development in nations and regions, and it is also a necessary path for improving economic growth quality and social well-being. Under the combined influence of globalization and technological revolution, traditional economic models are facing increasingly severe challenges, including excessive resource consumption, environmental pollution, a single industrial structure, and sluggish economic growth. Therefore, governments and regions are actively implementing measures to optimize industrial structures, enhance technological innovation capabilities, accelerate digital transformation, and promote green development to establish new economic structures characterized by high added value, low energy consumption, and high efficiency. Economic transformation not only requires the introduction of advanced manufacturing, emerging technology industries, and modern services at the industrial level but also the creation of a favorable policy environment and innovation atmosphere at the macro level to provide institutional support and resource backing for economic development. At the same time, the successful implementation of economic transformation depends on the dual support of technology and talent. By promoting digital transformation in enterprises and improving the skills of the workforce, economies can effectively enhance their competitiveness and resilience. In this process, the emergence of new economic forms such as the digital economy, green economy, and sharing economy serves as both a driving force and an important component of economic transformation. In the future, with the further advancement of global technological progress and regional economic cooperation, countries and regions should place greater emphasis on the quality and efficiency of economic growth, achieving coordinated development of the economy, society, and environment based on innovation-driven and green development. Economic transformation is a long-term and complex systemic project that requires the collaboration of governments, enterprises, and society through scientific decision-making and flexible responses to inject new vitality into the economic system and bring about a fairer and more sustainable prosperity for society. In summary, a successful economic transformation will create a new landscape of high-quality development for nations and regions and make a positive contribution to global economic stability and social progress.

References

1. Gao D, Shenoy R, Yi S, et al. Synaptic resistor circuits based on Al oxide and Ti silicide for concurrent learning and signal processing in artificial intelligence systems[J]. *Advanced Materials*, 2023, 35(15): 2210484.
2. Zhao, Q., Hao, Y., & Li, X. (2024). Stock price prediction based on hybrid CNN-LSTM model. *Applied and Computational Engineering*, 104, 110-115.
3. Mo K, Chu L, Zhang X, et al. DRAL: Deep reinforcement adaptive learning for multi-UAVs navigation in unknown indoor environment[J]. *arXiv preprint arXiv:2409.03930*, 2024.
4. Tang X, Wang Z, Cai X, et al. Research on heterogeneous computation resource allocation based on data-driven method[C]//2024 6th International Conference on Data-driven Optimization of Complex Systems (DOCS). IEEE, 2024: 916-919.
5. Yan, Hao, et al. "Research on image generation optimization based deep learning." *Proceedings of the International Conference on Machine Learning, Pattern Recognition and Automation Engineering*. 2024.
6. Zhang W, Huang J, Wang R, et al. Integration of Mamba and Transformer--MAT for Long-Short Range Time Series Forecasting with Application to Weather Dynamics[J]. *arXiv preprint arXiv:2409.08530*, 2024.
7. Diao S, Wei C, Wang J, et al. Ventilator pressure prediction using recurrent neural network[J]. *arXiv preprint arXiv:2410.06552*, 2024.
8. Xu Q, Wang T, Cai X. Energy market price forecasting and financial technology risk management based on generative AI[J]. 2024.
9. Wu X, Sun Y, Liu X. Multi-class classification of breast cancer gene expression using PCA and XGBoost[J]. 2024.
10. Min, Liu, et al. "Financial Prediction Using DeepFM: Loan Repayment with Attention and Hybrid Loss." 2024 5th International Conference on Machine Learning and Computer Application (ICMLCA). IEEE, 2024.
11. Shi X, Tao Y, Lin S C. Deep Neural Network-Based Prediction of B-Cell Epitopes for SARS-CoV and SARS-CoV-2: Enhancing Vaccine Design through Machine Learning[J]. *arXiv preprint arXiv:2412.00109*, 2024.
12. Wang, H., Zhang, G., Zhao, Y., Lai, F., Cui, W., Xue, J., Wang, Q., Zhang, H., & Lin, Y. (2024). RPF-ELD: Regional Prior Fusion Using Early and Late Distillation for Breast Cancer Recognition in Ultrasound Images. *Preprints*. <https://doi.org/10.20944/preprints202411.1419.v1>
13. Zhao R, Hao Y, Li X. Business Analysis: User Attitude Evaluation and Prediction Based on Hotel User Reviews and Text Mining[J]. *arXiv preprint arXiv:2412.16744*, 2024.
14. Guo H, Zhang Y, Chen L, et al. Research on vehicle detection based on improved YOLOv8 network[J]. *arXiv preprint arXiv:2501.00300*, 2024.
15. Zhao Y, Hu B, Wang S. Prediction of brent crude oil price based on lstm model under the background of low-carbon transition[J]. *arXiv preprint arXiv:2409.12376*, 2024.
16. Wang Z, Chen Y, Wang F, et al. Improved Unet model for brain tumor image segmentation based on ASPP-coordinate attention mechanism[J]. *arXiv preprint arXiv:2409.08588*, 2024.
17. Yang H, Cheng Z, Zhang Z, et al. Analysis of Financial Risk Behavior Prediction Using Deep Learning and Big Data Algorithms[J]. *arXiv preprint arXiv:2410.19394*, 2024.
18. Huang B, Lu Q, Huang S, et al. Multi-modal clothing recommendation model based on large model and VAE enhancement[J]. *arXiv preprint arXiv:2410.02219*, 2024.
19. Zhang J, Zhang W, Tan C, et al. YOLO-PPA based efficient traffic sign detection for cruise control in autonomous driving[J]. *arXiv preprint arXiv:2409.03320*, 2024.
20. Li X, Cao H, Zhang Z, et al. Artistic Neural Style Transfer Algorithms with Activation Smoothing[J]. *arXiv preprint arXiv:2411.08014*, 2024.
21. Wu Z. Large Language Model Based Semantic Parsing for Intelligent Database Query Engine[J]. *Journal of Computer and Communications*, 2024, 12(10): 1-13.
22. Huang S, Yang H, Yao Y, et al. Deep adaptive interest network: personalized recommendation with context-aware learning[J]. *arXiv preprint arXiv:2409.02425*, 2024.
23. Li X, Wang X, Qi Z, et al. DTSGAN: Learning Dynamic Textures via Spatiotemporal Generative Adversarial Network[J]. *Academic Journal of Computing & Information Science*, 7(10): 31-40.

24. Yang H, Sui M, Liu S, et al. Research on Key Technologies for Cross-Cloud Federated Training of Large Language Models[J]. arXiv preprint arXiv:2410.19130, 2024.
25. Ma D, Yang Y, Tian Q, et al. Comparative analysis of x-ray image classification of pneumonia based on deep learning algorithm[J]. Theoretical and Natural Science, 2024, 56: 52-59.
26. Cheng Y, Yang Q, Wang L, et al. Research on Credit Risk Early Warning Model of Commercial Banks Based on Neural Network Algorithm[J]. arXiv preprint arXiv:2405.10762, 2024.
27. Xiang A, Zhang J, Yang Q, et al. Research on splicing image detection algorithms based on natural image statistical characteristics[J]. arXiv preprint arXiv:2404.16296, 2024.
28. Wu Z, Chen J, Tan L, et al. A lightweight GAN-based image fusion algorithm for visible and infrared images[C]//2024 4th International Conference on Computer Science and Blockchain (CCSB). IEEE, 2024: 466-470.
29. Yang H, Yun L, Cao J, et al. Optimization and Scalability of Collaborative Filtering Algorithms in Large Language Models[J]. arXiv preprint arXiv:2412.18715, 2024.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.