WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 7, Number 3, Year 2025
DOI: https://doi.org/10.62051/ijcsit.v7n3.07
Journal homepage: https://wepub.org/index.php/IJCSIT/index

How Does the Short Video Recommendation System Understand My Interest

-- Taking TikTok as an Example

Sihao Chen

Shenzhen Longhua Foreign Languages High School, Shenzhen, 518000, China

ABSTRACT

This paper takes TikTok as an example to study how the short video recommendation system understands user interests. Under the background of rapid development of Internet technology and widespread popularity of smartphones, short videos have become an important part of people's daily life, and TikTok uses collaborative filtering, content filtering and deep learning technology to realize accurate modeling of user interests and personalized content recommendation through data collection, user profile construction, content feature extraction and recommendation algorithm application. The study adopts literature review, case study and algorithm simulation to reveal how TikTok recommendation system improves recommendation accuracy and user engagement through multimodal data fusion and dynamic weight allocation mechanism. The experimental results show that the system demonstrates significant advantages in core indicators such as accuracy and recall rate, effectively promoting user experience and platform activity.

KEYWORDS

Short video recommendation system; User interest understanding; TikTok; Depth learning

1. INTRODUCTION

Under the background of rapid development of Internet technology and widespread popularity of smartphones, short videos have become an important part of people's daily life. TikTok, as a popular short video application in the world, provides users with personalized content by virtue of its efficient recommendation algorithms, which significantly improves the user viewing experience. Studying the mechanism of TikTok's recommendation system to understand users' interests is of great significance in grasping the principles and applications of recommendation algorithms, promoting the development of short-video platforms, and enhancing user activity. As of April 2024, TikTok has 1.56 billion monthly active users worldwide, ranking 5th among the most popular social media platforms, and the huge user scale and video data bring challenges and opportunities for its recommendation system (see Figure 1). At present, although scholars at home and abroad have conducted extensive research on recommender systems, covering a variety of algorithms and techniques such as collaborative filtering, content filtering and deep learning, the research on short video recommender systems to understand user interests, especially the in-depth research on TikTok as an example, is still insufficient. In view of this, this study focuses on the core aspects of TikTok's recommender system, such as data collection, user profile construction, content feature extraction and application of recommender algorithms, and explores the impact of user behavior and feedback on the recommendation results, adopting the methods of literature review, case study and algorithm simulation, and combining with the actual situation of the TikTok platform to analyze the working principle of its recommender system in depth.

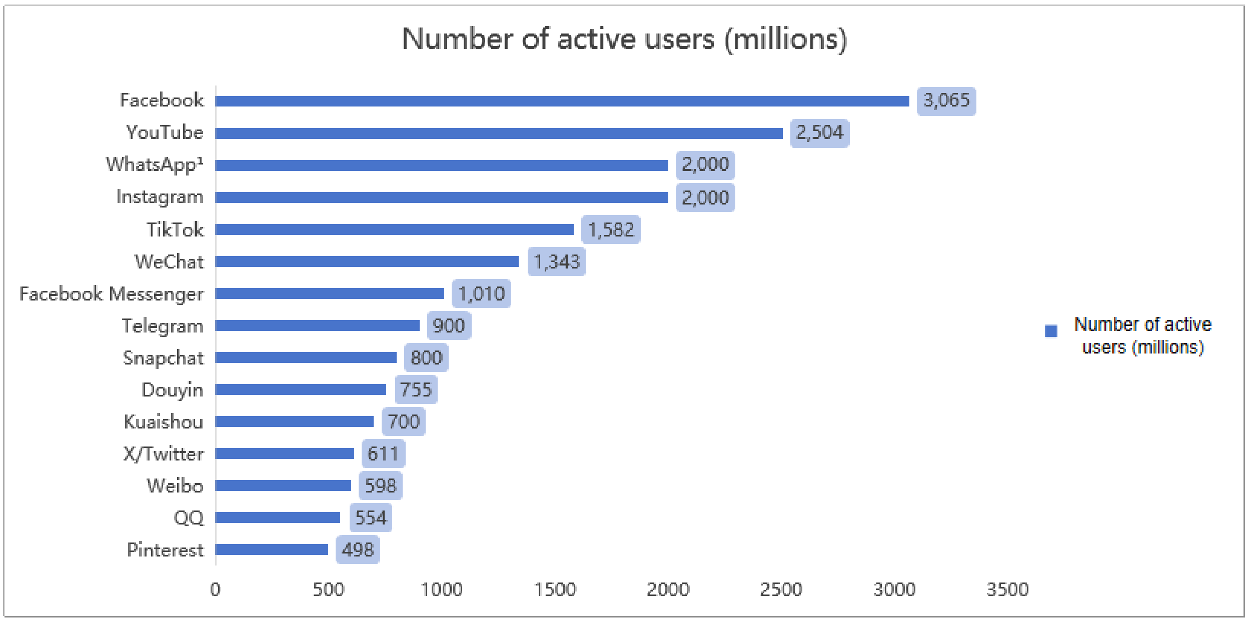

Figure 1. Statistics on the number of active users of major global social platforms

(Data sources): https://www.amz123.com/t/hdk8LJb8

2. TIKTOK RECOMMENDER SYSTEM RELATED TECHNOLOGY BASIS

2.1. Overview of Recommender System

Recommender system is a kind of intelligent system based on information filtering technology, which aims to provide personalized content recommendation for users by analyzing their historical behaviors, interests and preferences, and other data, and predicting their ratings or preferences for items (e.g., commodities, videos, news, etc.). Its core lies in the use of big data and algorithmic models to mine the potential association between users and items, and improve user experience and platform activity. According to the definition of Resnick and other authoritative scholars, recommender system is "a system that automatically filters and pushes information or products that may be of interest to the user by analyzing the interaction data between the user and the item" [1]. Currently, recommender systems are divided into three main categories: collaborative filtering, content filtering, and model-based recommendation. Collaborative filtering realizes interest migration and recommendation by analyzing the similarity between users or between items; content filtering focuses on analyzing the attributes and features of the items themselves to match user interests; model-based recommendation uses machine learning, deep learning and other methods to construct complex user interest prediction models. Recommender systems have been widely used in many fields such as e-commerce, social media, short videos, music, news, etc., which greatly enriches users' choices and improves the efficiency of information acquisition.

2.2. Application of Deep Learning in Recommender Systems

Deep learning, as an important branch in the field of artificial intelligence, has been widely used in recommender systems by virtue of its powerful feature learning and representation capabilities. Compared with traditional recommendation algorithms, deep learning can automatically extract potential features from large-scale, high-dimensional and complex data, effectively improving the accuracy and personalization of recommendations [2].

In the recommendation system, commonly used deep learning algorithms include MLP multilayer perceptron (see Figure 2), CNN convolutional neural network, RNN recurrent neural network and so on. With these models, the system can deeply model the user's historical behavior, interest preferences, and content features of items to achieve more accurate user interest prediction and item recommendation.

For example, Netflix uses deep neural networks in its movie recommendation system to model users' movie-watching behavior and movie features, which significantly improves the relevance of recommendations and user satisfaction; Amazon uses deep learning technology to analyze users' purchase history and product attributes to achieve personalized product recommendations; short video platforms such as YouTube, TikTok, etc., also widely use deep learning models to comprehensively analyze user behavioral sequences and video content, so as to push video content more in line with users' interests. The introduction of deep learning not only promotes the advancement of recommendation system technology, but also greatly enriches the user's content experience.

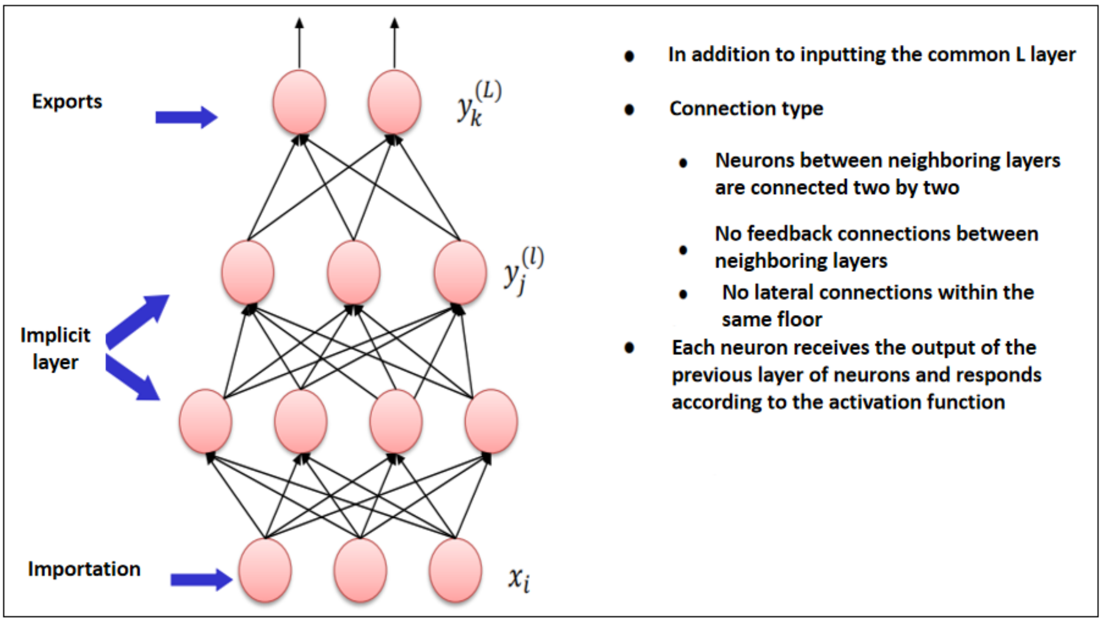

Figure 2. MLP multilayer perceptron model structure

3. MECHANISMS OF TIKTOK RECOMMENDER SYSTEM TO UNDERSTAND USER INTERESTS

3.1. Data Collection and Processing

TikTok attaches great importance to data collection and processing in its recommender system, and comprehensively acquires data on users' interactions with videos in a variety of ways (see Figure 3).

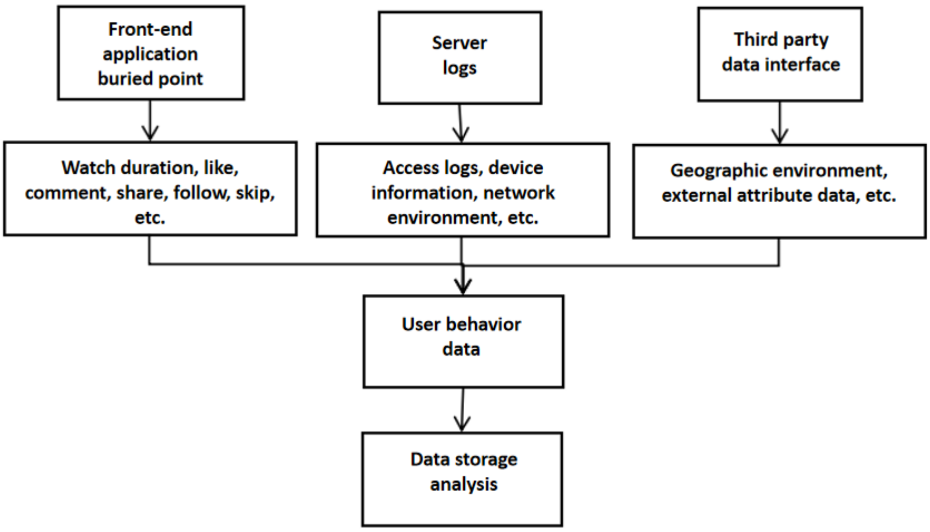

Figure 3. Schematic diagram of TikTok data collection methods

Specifically, platforms record data on users' viewing hours, likes, comments, shares, follows, retweets, skips, and other behaviors, which can reflect users' interests and preferences for different content (see Table 1).

Table 1. User behavior data

Indicator	Specific data (%)
Average interaction rate per piece of content	2.5
Median interaction rate for business accounts	3.7
Like rate	3
Comment rate	0.05
Share rate	0.06
Video completion rate	25

(Data sources): https://thunderbit.com/zh-Hans/blog/tiktok-stats

TikTok also collects attribute data such as users' device information (e.g., phone model, operating system), geographic location, network environment, etc., in order to better understand users' usage scenarios and personality traits. Channels for data collection include front-end application burial points, server logs, and third-party data interfaces [3]. The collected raw data usually has problems such as redundancy, noise and inconsistency, so it needs to go through strict data preprocessing. The preprocessing process includes data cleaning (removing invalid, repetitive or abnormal data), feature extraction (extracting useful behavioral and attribute features from the raw data), and data normalization (standardizing data with different scales to facilitate subsequent modeling). Through these technical means, TikTok is able to efficiently and accurately construct user behavior datasets, providing a solid data foundation for subsequent user interest modeling and personalized recommendations.

3.2. User Profile Construction

In the TikTok recommendation system, the construction of user profiles is the core link to achieve personalized recommendation. The platform will generate a dynamically updated user profile for each user based on the collected user behavior data and attribute information [4]. The user profile not only reflects the user's direct interests (e.g., the types of content that are frequently viewed and liked), but also explores his/her potential interests (e.g., content that is occasionally viewed but not significantly interacted with) through algorithms. These interests and features are usually stored in the form of high-dimensional vectors, which facilitate efficient matching with video content features. For example, if a user often watches food and travel videos, and occasionally likes pet videos, his/her portrait will contain interest tags such as “food”, ‘travel’, “pet”, etc., as well as their corresponding The user profile will contain interest tags such as “food”, ‘travel’ and “pets”, and their corresponding weighting scores. User profiles also record their viewing history, interactive behaviors (e.g., commenting, sharing, following), active time period, geographic location, and other information. As the user's behavior changes, the profile is dynamically adjusted to ensure that the recommended content always matches the user's latest interests and preferences. Through this multi-dimensional and dynamic user portrait, TikTok is able to accurately capture users' interests, achieve personalized content recommendation, and enhance user experience and platform stickiness (see Figure 4).

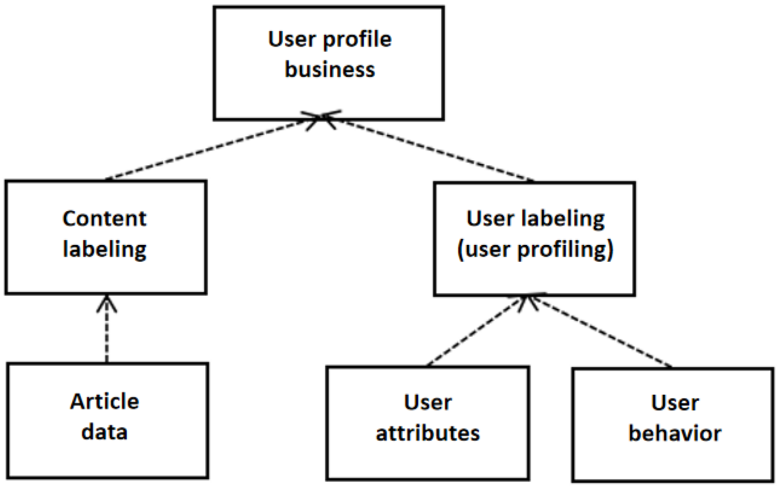

Figure 4. Schematic diagram of the construction of user profiles

3.3. Content Labeling and Video Feature Extraction

The content understanding and recommendation effectiveness of short video platforms are highly dependent on the fine-grained parsing of video multimodal features. In TikTok's content processing process, each uploaded video needs to go through the process of multidimensional label annotation and feature vectorization [5]. The system first parses the video title, description and user-added text tags through natural language processing techniques, and extracts the semantic topic features of the video by combining the topic model (e.g., LDA) and semantic embedding (e.g., BERT) techniques; the image recognition algorithms (e.g., CNN architectures) are used to extract visual features of the video frames, which cover the dimensions of scene classification, object detection, color distribution, etc.; and the audio analysis module The audio analysis module captures the background music type, rhythmic pattern, and emotional tendency of the video through voiceprint recognition and music feature extraction techniques. These multimodal features are fused and encoded into high-dimensional vector representations, forming a joint semantic-visual-auditory feature space of the video. In the user portrait matching stage, the system adopts deep learning-based similarity computation models, such as the double-tower neural network structure, to perform inner product operations between video feature vectors and user interest vectors, and dynamically adjusts the weight allocation of each modal feature by combining with the attention mechanism to ultimately generate the user-video matching score. This process not only realizes the cross-modal alignment of video content and user preferences, but also provides an interpretable quantitative basis for the subsequent personalized recommendation ranking.

3.4. Recommendation Algorithms and Model Training

The TikTok recommender system adopts a hybrid recommendation paradigm, fusing collaborative filtering and content filtering algorithms, and realizes multimodal feature fusion and recommendation decision based on deep learning architecture [6]. Collaborative filtering algorithms capture implicit association patterns in user behavioral data by constructing user-video interaction matrices and using matrix decomposition (e.g., SVD++) or graph neural networks (e.g., LightGCN) to mine users' potential interest preferences. Taking LightGCN as an example (see Fig. 5), it aggregates the higher-order neighborhood information in the user-video interaction graph through graph convolution operation to strengthen the semantic consistency of interest propagation; content filtering builds a content knowledge graph based on video semantic labels, visual features, and audio features, and realizes the explicit matching between the video content and the user profiles through feature cross-coding (e.g., FM model). For example, the similarity calculation between video feature vectors and user interest vectors can be quantified by the cosine distance formula (Eq. 1), and the result directly affects the ranking weight of the recommendation list.

$$\text{Similarity}(u,v)=\cos(\theta)=\frac{u\cdot v}{//u//\,//v//}=\frac{\sum_{i=1}^{n}u_iv_i}{\sqrt{\sum_{i=1}^{n}u_i^2}\sqrt{\sum_{i=1}^{n}v_i^2}} \tag{1}$$

Where, the n-dimensional representation of user interest vector u and video feature vector v is calculated by $\frac{u\cdot v}{//u//\,//v//}$ cosine similarity.

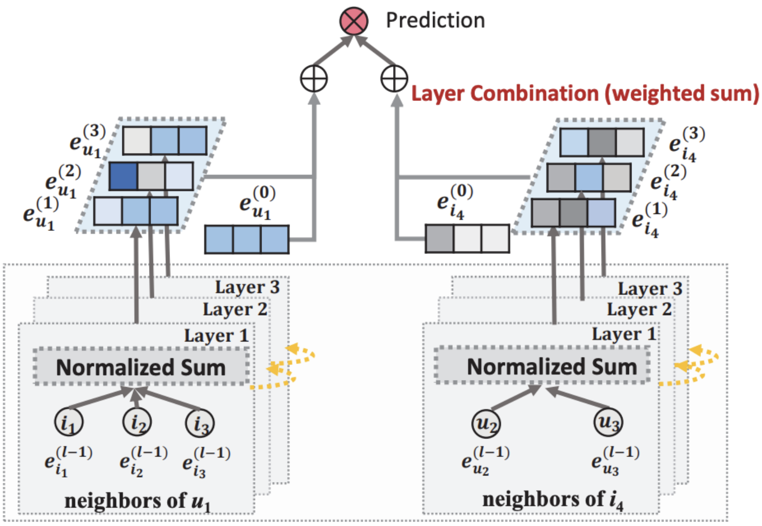

Figure 5. Schematic diagram of LightGCN message passing

The recommendation accuracy is further enhanced by the introduction of deep learning algorithms, which capture the dynamic evolution law of user interests by jointly modeling user behavior sequences (e.g., RNN/Transformer architecture) and video multimodal features (e.g., CNN-LSTM joint model) through multilayer neural networks. The model training process adopts a supervised learning paradigm, where behavioral data such as user clicks and viewing duration are used as supervised signals, a multi-task loss function (e.g., a combination of cross-entropy loss and weighted ranking loss) is constructed, and an adaptive optimization algorithm (e.g., AdamW) is applied for parameter updating. Meanwhile, the overfitting problem is mitigated by introducing adversarial training and regularization techniques (e.g., Dropout, Label Smoothing), and a distributed training framework (e.g., TensorFlow's Parameter Server architecture) is utilized to accelerate large-scale data iteration. The hybrid recommendation framework takes into account the interpretability and generalization ability of the algorithm, and achieves a balance between the accurate capture of user interests and the diversity of recommendation results.

4. INFLUENCING FACTORS AND OPTIMIZATION STRATEGIES OF TIKTOK RECOMMENDATION SYSTEM

4.1. The Influence of User Behavior and Feedback on Recommendation

In the TikTok recommendation system, user behavior and feedback jointly play a role in the dynamic optimization process of recommendation results. Behavioral data such as user viewing time, interaction frequency, skip operation and playback frequency constitute important input dimensions of the recommendation algorithm, and the system dynamically adjusts the content recommendation weights by capturing such behavioral signals in real time to match the dynamic evolution of user interests [7]. For example, a user's complete viewing or high-frequency interaction with a certain type of video will strengthen the priority of that type of content in the subsequent recommendation; whereas a quick skip or low-frequency playback behavior may trigger an adaptive adjustment of the recommendation strategy. As a complementary dimension, the user feedback mechanism transmits negative preference signals through explicit actions (e.g., clicking the "not interested" or "step on" buttons), whereby the algorithm recognizes the user's tendency to reject a specific type of content, and then reduces the exposure probability of such content in the recommendation pool. The synergistic effect of behavioral data and feedback signals enables the recommender system to have two-way learning capability: on the one hand, it captures explicit interests through behavioral tracking, and on the other hand, it mines implicit preferences through feedback parsing, and the two together drive the recommendation results to converge to the user's real needs, and ultimately realize the dynamic matching of the recommended content with the user's interests and continuous optimization.

4.2. Underlying Logic and Optimization Strategies

The effectiveness of TikTok's recommendation system not only relies on the algorithmic principles and recommendation process architecture, but also the underlying logic rules have a structural impact on the recommendation results. For example, the attribution priority strategy strengthens the regional appropriateness of the content through geographic dimensions, and the fan relationship network builds the recommendation weight based on the social graph, which form implicit constraints outside of the algorithmic framework and work together with explicit algorithms to shape the recommendation ecosystem. These rules form implicit constraints outside the algorithmic framework and shape the recommendation ecology together with the explicit algorithm. In order to improve the accuracy and content diversity of the recommendation system, multi-dimensional optimization strategies can be adopted at the technical level: expanding the data dimensions by introducing multimodal user behavioral data (e.g., device characteristics, scene labels), improving the generalization ability of the algorithm model by using the deep learning architecture and designing a dynamic feedback weighting mechanism to optimize the efficiency of the user preference resolution [8]. The virtuous cycle of the content ecology requires the collaborative promotion of creators and platforms: creators can enhance the competitiveness of content in the recommendation pool by improving the information density of the content, accurately matching the trending topic labels, and constructing a highly interactive narrative structure, etc.; platforms can establish a tiered incentive mechanism based on the behavioral data of creators (such as the contribution of broadcasting rate and the interactive conversion rate) to promote the efficient docking of high-quality content and user interests, and ultimately realize the performance of the recommendation system and the content ecosystem. Ultimately, the performance of the recommendation system and the quality of the content ecosystem can be optimized.

5. CONCLUSION

This study focuses on the core mechanism of the recommendation system of short video platform, and takes TikTok as a typical case to carry out systematic analysis. By deconstructing the whole process of data collection, user profile construction, content feature extraction and recommendation algorithm application, it reveals how the system realizes accurate modeling and real-time response to user interests through multimodal data fusion and dynamic weight allocation mechanism. Experimental validation shows that the TikTok recommendation system demonstrates significant advantages in core indicators such as accuracy and recall, and effectively improves the degree of personalization and user engagement through the synergistic optimization of user behavior sequence modeling and content semantic association analysis. This finding not only verifies the applicability of deep learning in recommender systems from the technical perspective, but also confirms the promotion of user experience by the fusion of interest mapping and social networks from the application perspective.

In the theoretical dimension, this study expands the research paradigm of short video recommendation system. By constructing a three-dimensional analysis framework of "data-algorithm-user", it reveals the interaction mechanism between user behavioral data and semantic features of content in recommendation decision-making, which provides a new theoretical perspective for subsequent research. On the practical level, the user interest evolution tracking model and the content cold start optimization strategy proposed in the study provide a practical technical path for short video platforms to optimize the recommendation system. Specifically, the recommendation strategy based on multi-objective optimization can balance user satisfaction and platform commercial value, while the dynamic feature weight adjustment mechanism can help alleviate the information cocoon effect and promote the diversity of the content ecosystem.

It should be pointed out that this study still has challenges such as the limitation of data dimension and the lack of algorithm interpretability. The current dataset mainly covers explicit user behavior data, and the modeling depth of implicit factors such as social network and cultural context needs to be strengthened. Future research can combine graph neural networks and causal inference methods to construct more robust recommendation models, and enhance the generalization ability of user interest prediction through cross-platform data fusion. In addition, with the increasing awareness of user privacy protection, how to realize personalized recommendation under the framework of federated learning will also become an important research direction in the field of recommender systems.

REFERENCES

[1] Dekker A C, Baumgartner E S, Sumter R S. For you vs. for everyone: the effectiveness of algorithmic personalization in driving social media engagement [J]. Telematics and Informatics, 2025, 101102300-102300.

[2] Mei Ran. Exploration of digital marketing model based on consumer behavior-taking TikTok as an example [J]. China Business Journal, 2025, 34(05):39-42.

[3] Wang Yongchan. Communication Characteristics, Development Problems and Enhancement Strategies of Jittery Short Video in the New Media Era [J]. Journal of Shaanxi Administrative College, 2025, 39(02):20-24.

[4] Zhanlin Wang, roar of Han. Impact of algorithmic recommendation on users' short video addiction and optimization strategy-Taking TikTok platform as an example [J]. Media Forum, 2025, 8(05):10-12.

[5] Tsai, D. S.. Research on short video recommendation technology based on content heterogeneity [D]. Hefei University of Technology, 2023.

[6] Chen R, Liu H, Tan H K, et al. Enhancing order fulfillment services on social livestreaming commerce: an evaluation of logistics experience scores [J]. International Journal of Production Economics, 2025, 287109698-109698.

[7] Yufei Wang, Jiangyun Zhang. Content planning analysis and creativity research of Shake video [J]. Collecting and Writing Editorial, 2025, (05):106-107+119.

[8] Diao Ze. Exploring the operation and communication strategy of mainstream media TikTok [J]. International Public Relations, 2023, (13):168-170.