# Investigation on the Self-Improving Algorithm of TikTok Based on Extensive User Interactions

Xiaoxing Chen[a]
*Department of Computer and Science Engineering, The Ohio State University, Ohio, U.S.A.*

Keywords: Tiktok, Social Media, Self-Growing, Tag Analyze, Time-Dependent Content Recommendation.

Abstract: The ubiquity of short video apps in contemporary society is epitomized by the widespread adoption of TikTok on mobile devices. The platform's escalating user rates and engagement duration are indicative of its growing influence. This paper investigates the TikTok algorithm's ability to process extensive data sets to curate and recommend user-preferred content. Conducted through a series of surveys and analytical studies across various age demographics within university populations, this research emphasizes the pivotal role of metadata tags and the platform's autonomous enhancement algorithms. By harnessing advanced machine learning and artificial intelligence technologies—such as Graph Neural Networks (GNN), Reinforcement Learning (RL), Temporal Convolutional Networks (TCN), Natural Language Processing (NLP), Generative Adversarial Networks (GANs), and Attention Mechanisms—TikTok effectively tailors its algorithmic learning to user interactions. This strategic integration allows for the progressive refinement of user recommendations, fostering personalized content delivery while ensuring privacy, and enhancing the overall quality of content and user engagement. The study's findings reveal that these technological integrations enable TikTok to more accurately discern user preferences, thus facilitating the delivery of more engaging and relevant content. Ultimately, these improvements have substantial implications for the enrichment of user experience on the platform.



## 1 INTRODUCTION

TikTok was started in September 2016 and expanded to the United States in 2017 before spreading to 150 countries. TikTok has experienced a meteoric rise in its user base. By February 2024 TikTok had amassed 4.7 billion downloads. The platform's short video format resonates with users by capturing moments. Hashtags play a role in driving trends and content recommendations on TikTok, which evolve over time. Continuous research and analysis are crucial for understanding these dynamics. Through an examination of label usage trends during periods I have personally observed shifts in content suggestions, on TikTok. Reflecting on my experience of encountering food related videos on the app prompted me to delve into TikTok intricate data algorithms. In today's world it's common to see people of all ages. Whether young people, to adults, or even children. Spending an amount of time on social media platforms like video apps. According to data from Datareportal it is projected that by 2023 individuals will allocate about 35.8% of their time to media activities (DataReportal, 2024). Surprisingly this surpasses the time spent on sleeping or working for some individuals. Handling data requires robust algorithms for support. An algorithm can be described as a procedure embedded within a computer system that aids in problem solving and calculations (Amjady, 2001). The algorithms utilized in media platforms are advanced and efficient(Ma, 2014). Platforms like TikTok utilize user data such as information, engagement levels, location details, viewing habits and other relevant data to tailor content for users with precision(Ma, 2014).

## 2 DTAT SURVEY

This paper finds that more than 110 college students who have been using social media for at least 5 years, random school professors and ordinary passers-by to

[a] https://orcid.org/0009-0000-2464-8117



Chen, X.
Investigation on the Self-Improving Algorithm of TikTok Based on Extensive User Interactions.
DOI: 10.5220/0012925100004508

In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024)*, pages 227-233
ISBN: 978-989-758-713-9

conduct a questionnaire survey to get their opinions on what TikTok's big data analysis is based on. In the questionnaire, the time spent on social media is included, do you authorize the location of the social media you use, do you think the videos TikTok sends you are your favorite and so on. For anyone who has been on social media for less than a year (and very few), this paper chose to ignore their responses. They also ignored extreme responses, such as 99 years of social media use. The questions included in the questionnaire include age range, gender, daily use time, main use period, main use purpose, algorithm accuracy, algorithm basis, understanding of algorithm working principle, data and privacy processing, and an open question about how Tik Tok will improve its algorithm or user experience in the future.

This paper analyzes the data gathered from a survey questionnaire as shown in Fig.1. Focusing on the sections previously introduced in Data collection let's delve into each of them individually. Firstly, regarding age range is the first part of the questionnaire. The majority of participants (70%) fall within the 18-24 age bracket aligning with college student demographics. Secondly in terms of gender distribution is next. It's fairly balanced with 45% male, 45% female, 10% genders and those preferring not to disclose. Thirdly, let's can look at usage time. The highest percentage falls within the range of 30 minutes to 2 hours indicating usage patterns among college students. Fourthly, it is the social media usage time. Most TikTok activity occurs during the night (30%) followed by afternoon and morning slots. Fifthly, let's analyze the primary usage purposes which after collect. Entertainment (40%) and educational information seeking (30%) are highlighted as reasons for using TikTok—a platform that serves both informative and entertaining roles effectively. Sixth, algorithm accuracy assessment is the part people care so much. A large portion finds the algorithm accurate" (40%) or "average" (30%) suggesting high user satisfaction, with the recommended content. Seventhly, the algorithms foundation lies in user engagement (35%). Time spent viewing content (30%) which are considered the influences on algorithmic suggestions. Additionally, when it comes to comprehending how algorithms operate the majority of individuals possess a " understanding" (35%) or a "moderate understanding" (25%) suggesting a restricted awareness of algorithm functioning. Lastly, data and privacy handling is the thing that have to analyze. TikTok's handling of user data and privacy is mostly "average" (40%), indicating that there is room for improvement.

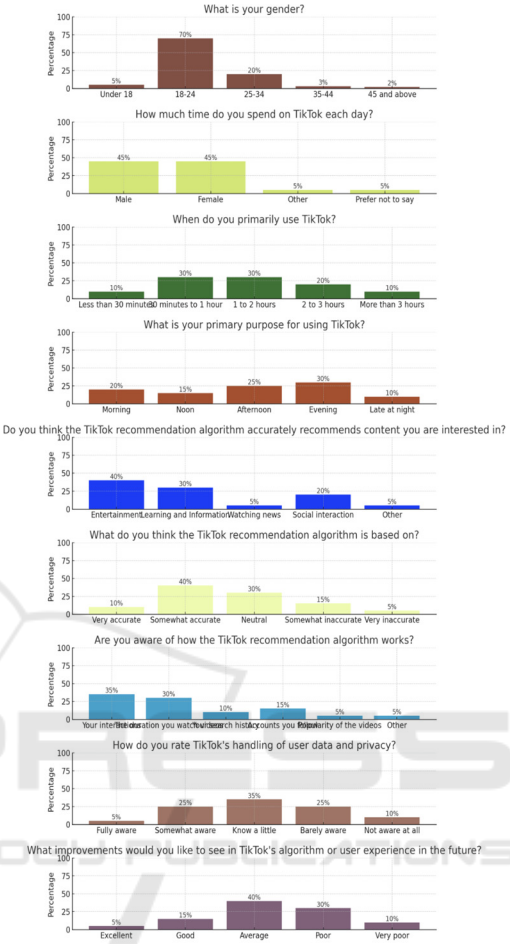


Figure 1: Data Chart of questionnaire survey.

This paper also collects and analyzes a question at the end of the questionnaire. This paper divided them into 10 broad categories and analyzed them briefly. Let us expand it one by one. The first thing that need to know is More personalized content recommendations. Students may want algorithms to better understand their interests and needs and provide more relevant video content. The second one is Increased privacy protection which is the thing that people care so much. With the increasing emphasis on personal privacy, many students may expect platforms to provide stronger privacy Settings and transparency. The third one is Reduce advertising and promotional content. Students may want to encounter fewer ads while scrolling through videos to provide a smoother user experience. The fourth one is Optimize the interface and interaction design which is intuitive display in front of the user. Some users may mention

that they want a more intuitive and user-friendly interface, including improved search and video classification. The fifth one is Provide more education and knowledge content. As students, they may want to see more education related content, such as professional knowledge, study skills, etc. The sixth one is Add content creation tools and features which can attract more target users. Expect more innovative video editing tools and creation features to encourage creative expression. The seventh one is Strengthen community and interactive features. Students may want better comment and interactive features to promote community building. The eighth one is to Improve fairness and diversity of algorithms. Algorithms are required to fairly promote creators from different backgrounds and cultures. The ninth one is to Improve content moderation systems. Avoid blocking or restricting harmless content by mistake and remove harmful content quickly. The last one is to support multilingual and cross-cultural communication because the world today is very diverse. Multilingual users may request improved translation and subtitling features to better enjoy internationa11 content.

# 3 TIK TOK ALGORITHMS

## 3.1 The Role of Tag in Massive Algorithms

Based on the above data analysis, this paper found that people's opinions about TikTok's algorithm are mainly based on the definition of labels. In fact, tags must occupy the core foundation of TikTok's algorithm(Klug, 2021). Users generally believe that cumulative tags, such as hashtags, are one of the factors that influence algorithm recommendations. The study empirically confirmed that video interaction, release time, and the use of tags are the key factors that improve the probability of video recommendation. TikTok assigns descriptive tags to new videos based on computer vision analysis, tag mentions, post descriptions, sound effects, and embedded text. These tags are then used to align videos with groups of users whose interests match, optimizing the recommendation process(Klug, 2021). According to Fig.2, the data suggests that the best times to post on TikTok are Tuesday, Wednesday and Thursday from 2pm to 5pm (Sprout Social, 2024). In this way, uploaded videos have a better chance to be seen by more people, because it happens that more than 50 percent of people in the questionnaire survey use short video apps such as TikTok in the afternoon and evening.

Use 15-point type for the title, aligned to the center, line space exactly at 17-point with a bold font style and initial letters are capitalized. No formulas or special characters of any form or language are allowed in the title.

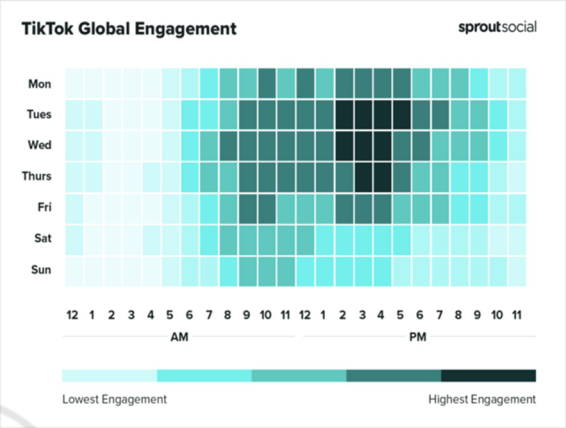


Figure 2: TikTok Global Engagement chart from SproutSocial.

## 3.2 Time-Dependent Content Recommendation

Understanding the Complexities of TikTok's Time-Dependent Content Recommendation Strategy. In today's digital age, social media platforms have become an integral part of people's lives, with TikTok leading the way in short-form video content. Behind the scenes, TikTok's algorithm works tirelessly to provide users with personalized and engaging content. However, the intricacies of its time-dependent content recommendation strategy go beyond what meets the eye. To effectively curate content based on users' preferences, TikTok's algorithm incorporates various disciplines such as data science, sociology, marketing, and computer science. In a recent study, seven key aspects were identified that significantly impact the platform's time-dependent content recommendations.

User Behavior and Interaction on TikTok involve a complex algorithm which takes into account users' past activities and viewing habits. By analyzing your history on the platform, including the types of content you engage with and when, the algorithm adjusts the push content accordingly. This explains why you might find yourself scrolling through mouth watering food videos during dinner time.

Time-Sensitive Content Strategy is content creators on TikTok understand the significance of

timing. They recognize the topics that are likely to resonate with users during specific time periods. For instance, pushing food-related content during mealtimes takes advantage of users' daily activity patterns and their interest in culinary delights.
Cultural Habits and Algorithm Design is the thing that must pay attention on. Food cultures vary across regions, and TikTok acknowledges this diversity. Algorithms are designed to reflect cultural habits, recommending specific types of content at specific times. This ensures that users are more likely to explore and engage with content that aligns with their cultural preferences.

Personalized Recommendations and Group Behavior on TikTok's algorithm not only considers individual user behavior but also analyzes large-scale data to identify broader patterns of user behavior. By understanding that most people crave food content in the evening, the algorithm pushes relevant videos during this time, catering to the collective interests of the user community.

Real-Time Trends and Content Recommendations can be expanded in the algorithm captures real-time trends and adjusts the recommended content accordingly. If a particular food topic suddenly gains popularity in the evening, the algorithm quickly increases the feed of videos related to that topic, ensuring users stay up to date with the latest trends.

Social Media Marketing Strategies is merchants and brands strategically deliver food-related marketing content during peak meal times. TikTok's algorithm prioritizes this content to attract potential customers, benefiting both businesses and users seeking relevant recommendations.

Multimodal Data Analysis on TikTok's algorithms go beyond text analysis. They incorporate image recognition, analyzing food images, time stamps, and geolocation data to provide more precise and personalized content recommendations. This level of analysis ensures that users receive content tailored to their individual tastes and preferences.

## 3.3 The Self-Growing of Tik Tok

The self-growing nature of TikTok's algorithm is another fascinating aspect. It constantly optimizes itself through user interactions, utilizing the principles of big data and machine learning. While the specific algorithms used by TikTok remain a closely guarded secret, it is likely that advanced methods from the fields of artificial intelligence, machine learning, data mining, and user behavior analysis contribute to the platform's content optimization and user engagement strategies.

Although the exact details of TikTok's algorithm may remain undisclosed, the platform undoubtedly relies on cutting-edge technology to provide a seamless and captivating user experience. As users continue to engage with the platform, TikTok's algorithm evolves and adapts, ensuring that content recommendations remain fresh, relevant, and highly addictive. The Evolution of TikTok Algorithms: Unleashing the Power of AI.

TikTok, the popular social media platform, has captivated millions of users worldwide with its short-form videos and endless entertainment. Behind the scenes, the TikTok algorithm works tirelessly to curate personalized content streams for users, keeping them engaged and coming back for more. The secret to TikTok's success lies in its cutting-edge use of artificial intelligence (AI) and machine learning techniques. In this article, author will explore some of the key AI techniques that power the self-growing TikTok algorithm.

Graph Neural Networks (GNNs) are revolutionizing the way social networks analyze user interactions and content. With their ability to simulate complex relationships and interactions, GNNs enable TikTok to understand user preferences by analyzing the global structure of user interactions. This allows the algorithm to predict which content is most likely to be relevant and engaging, enhancing the user experience(Wu et al, 2021).

Reinforcement Learning (RL) on TikTok leverages reinforcement learning algorithms to dynamically adjust content recommendations based on user feedback in real-time. By analyzing metrics such as likes, shares, and viewing time, the algorithm learns to optimize content suggestions and maximize user engagement. Through trial and error, the RL algorithm continuously fine-tunes its recommendations, ensuring a personalized and addictive user experience(Sutton and Barto, 2018).

Federated Learning is the part that TikTok use. Privacy concerns have become a major issue in today's digital landscape. TikTok addresses this by employing federated learning, a decentralized approach that enables machine learning models to be trained on user devices without compromising privacy. With federated learning, TikTok can personalize content recommendations without relying on centralized user data, preserving privacy while delivering tailored recommendations(Konečný et al, 2016).

Temporal Convolutional Networks (TCNs) can Predicting user behavior over time is crucial for

TikTok's algorithm. TCNs excel at sequence prediction problems, making them ideal for analyzing past user activity and predicting future interactions. By utilizing TCNs, TikTok improves the timing of content recommendations, ensuring that users are presented with the right content at the right moment(Bai et al, 2018).

Zero-Sample Learning is a way to enhance content discovery and diversify recommendations, TikTok employs zero-sample learning. This technique allows the algorithm to recognize objects or patterns that it has not encountered during training. By recommending new or niche content without relying solely on historical interaction data, TikTok expands the horizons of its content discovery, surprising and delighting users with fresh and exciting recommendations(Xian et al, 2017).

Natural Language Processing (NLP) and Transformer Models are powerful tools. Beyond video analysis, TikTok harnesses advanced NLP techniques to understand reviews, descriptions, and labels associated with content. By utilizing transformer models, renowned for their effectiveness in processing linguistic data, TikTok gains a deeper understanding of the text environment surrounding videos. This enhances the recommendation engine and ensures that content suggestions align with user preferences(Vaswani et al, 2017).

Generative Adversarial Networks (GANs) is typically associated with content creation, can also play a role in TikTok's content planning process. By generating synthetic data, GANs can train more robust recommendation models while enhancing user-generated content in innovative ways. This opens up new avenues for creativity and expands TikTok's content ecosystem(Goodfellow et al, 2014).

Attention Mechanisms is integrating attention mechanisms into deep learning models empowers TikTok's algorithm to prioritize crucial parts of the data. Whether it's a particular frame in a video or a paragraph in a user's interaction history, attention mechanisms help determine what is most important for making accurate predictions. By focusing on the most relevant user interactions, TikTok refines the personalization of content streams, ensuring an engaging and tailored user experience(Bahdanau et al, 2014).

Through the application of these AI techniques, TikTok's algorithms achieve self-growth and optimization. The continuous evolution of the algorithm, driven by cutting-edge technologies, enables TikTok to deliver personalized, addictive, and captivating content to its users. As TikTok continues to innovate and explore new frontiers in AI, users can expect even more exciting experiences and discoveries on this ever-evolving platform. A Comprehensive Look at TikTok's Recommendation Algorithm. TikTok, the popular social media platform, has gained immense popularity due to its highly accurate and engaging content recommendations. Behind the scenes, numerous technologies work together seamlessly to create a personalized experience for each user. In this article, author will explore how these technologies collaborate to solve problems and enhance user satisfaction.

The first step in TikTok's recommendation algorithm is to collect and understand data. Graph Neural Networks (GNN) play a crucial role in this process. By simulating complex relationships between users and content, GNN helps to understand the interaction structure within social networks(Wu et al, 2021). This understanding allows the algorithm to predict with greater accuracy which content will be most relevant to a particular user's preferences.

Natural Language Processing (NLP) and transformation models are also vital components of TikTok's algorithm. NLP is used to parse and understand user-generated text such as comments and tags, which provide context for the relevance of conten(Vaswani et al, 2017). Transformation models like BERT or GPT dive deeper into these texts, providing richer data input to the recommendation system.

Once the data is collected and understood, the algorithm focuses on real-time optimization and adaptation. Reinforcement Learning (RL) plays a significant role in this stage(Sutton and Barto, 2018). By adjusting recommendations in real-time based on user feedback, such as viewing time, likes, and shares, the algorithm increases user engagement. It learns which content types are most appealing to users and adjusts its recommendation strategy accordingly. Temporal Convolutional Networks (TCN) are also instrumental in this process(Bai et al, 2018). By analyzing user behavior over time, TCN predicts future interactions and optimizes the timing of content recommendations.

Personalization and privacy protection are crucial aspects of TikTok's algorithm. Federated Learning allows the algorithm to make personalized recommendations without direct access to user data. By training models locally on user devices, TikTok enhances privacy protections while still providing personalized content recommendations. Diversity of content and user engagement are essential considerations for TikTok's recommendation algorithm. Zero-shot learning allows the algorithm to

recommend new or niche content that does not appear in the training data, thereby increasing the diversity of content discovery. Additionally, Generative Adversarial Networks (GANs) can generate synthetic data to train recommendation systems or enhance user-generated content in novel ways, improving the diversity and appeal of content. To refine recommendations further, attention mechanisms come into play. These mechanisms determine the most important parts of the data and enable models to predict users' interests more accurately. This, in turn, enables more personalized content recommendations.

By combining all these technologies, TikTok's algorithms continually grow and optimize themselves in different dimensions. The algorithm learns from user behavior, feedback, and interaction patterns, constantly adapting and improving its recommendation mechanism with new data inputs. This dynamic process aims to enhance user satisfaction and engagement.

TikTok's recommendation algorithm utilizes a wide range of technologies to provide users with a highly personalized and engaging experience. From data collection and understanding to real-time optimization, personalization, and diversity of content, each component plays a crucial role in ensuring that users are presented with content that aligns with their preferences and interests. As TikTok continues to evolve, its algorithm will undoubtedly become even more sophisticated, enhancing user satisfaction and cementing its position as a leading social media platform.

## 4 CONCLUSIONS

In the realm of personalized content recommendation systems, TikTok has established itself as a true leader. Leveraging cutting-edge machine learning and artificial intelligence technologies, TikTok has revolutionized the way users discover and engage with content. Through extensive analysis and the implementation of various advanced techniques, such as graph neural networks, reinforcement learning, federated learning, temporal convolutional networks, zero-sample learning, natural language processing, generative adversarial networks, and attention mechanisms, TikTok has elevated the user experience to new heights. Not only does this optimize user satisfaction, but it also increases overall engagement with the app. The combination of these innovative technologies allows TikTok to gain a deep understanding of user preferences, ensuring that the content served is not only relevant but also highly engaging. This commitment to personalization has led to a significant diversification and interactivity of content, all while prioritizing user privacy. As time progresses and technology advances, TikTok's algorithmic self-growth mechanism ensures that it remains at the forefront of the short video platform competition. By providing users with a personalized and highly participatory platform, TikTok has solidified its position as a pioneer in the industry. This achievement not only showcases TikTok's exceptional ability to handle vast amounts of data and optimize recommendation algorithms, but it also emphasizes its pivotal role in driving innovation and progress in the digital media landscape. Finally, TikTok's personalized content recommendation system has truly revolutionized the way users interact with short videos. By utilizing state-of-the-art technologies, TikTok continues to push boundaries and provide users with a platform that is both tailored to their preferences and highly engaging. Its commitment to privacy, innovation, and progress cements its position as a leader in the ever-evolving digital media landscape.

## REFERENCES

Fangfang., 2011, Research on power load forecasting based on Improved BP neural network. Harbin Institute of Technology.

DataReportal., 2024, The time we spend on social media. Retrieved from https://datareportal.com/reports/digital-2024-deep-dive-the-time-we-spend-on-social-media

Amjady, N., 2001, Short-term hourly load forecasting using time series modeling with peak load estimation capability. IEEE Transactions on Power Systems, 16(4), 798-805.

Ma, K., 2014, Short term distributed load forecasting method based on big data. Changsha: Hunan University.

Klug, D. et al., 2021, "An Empirical Investigation of Personalization Factors on TikTok." arXiv preprint arXiv:2201.12271.

Sprout Social., 2024, TikTok algorithm. Retrieved [March 29, 2024], from https://sproutsocial.com/insights /tiktok-algorithm/

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S., 2021, "A Comprehensive Survey on Graph Neural Networks." IEEE Transactions on Neural Networks and Learning Systems, 32(1), 4-24.

Sutton, R. S., & Barto, A. G., 2018, "Reinforcement Learning: An Introduction." Second Edition, The MIT Press.

Konečný, J., McMahan, H. B., Yu, F., Richtárik, P., Suresh, A. T., & Bacon, D., 2016, "Federated Learning: Strategies for Improving Communication Efficiency." arXiv preprint arXiv:1610.05492.

Bai, S., Kolter, J. Z., & Koltun, V., 2018, "An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling." arXiv preprint arXiv:1803.01271.

Xian, Y., Lampert, C. H., Schiele, B., & Akata, Z., 2017, "Zero-Shot Learning - A Comprehensive Evaluation of the Good, the Bad and the Ugly." arXiv preprint arXiv:1707.00600.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I., 2017, "Attention Is All You Need." In Advances in Neural Information Processing Systems (pp. 5998-6008).

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y., 2014, "Generative Adversarial Nets." In Advances in Neural Information Processing Systems (pp. 2672-2680).

Bahdanau, D., Cho, K., & Bengio, Y., 2014, "Neural Machine Translation by Jointly Learning to Align and Translate." arXiv preprint arXiv:1409.0473.