Scholar Publishing Group
Machine Learning Theory and Practice
https://doi.org/10.38007/ML.2025.050114
ISSN 2790-0983 Vol. 5, Issue 1: 133-140

Design of Data-Driven Social Network Platforms and Optimization of Big Data Analysis

Huijie Pan

Identity Department, PayPal Inc., San Jose, California, 95131, United States

Keywords: Big data analysis; Social networking; Platform design; Recommendation system

Abstract: In the context of the rapid expansion of social network service platforms, big data analysis technology plays a crucial role in enhancing user interaction and platform operation efficiency. This study deeply analyzes the design and architecture of a data-based social network service platform and at the same time integrates big data analysis technology to enhance the platform function. Through in-depth analysis of the interaction between social networks and big data, a scientific service platform architecture and data processing mechanism are formulated, and the technology application in key links such as recommendation algorithm, user behavior analysis, information promotion and public opinion monitoring is emphatically discussed. The research results contribute a solid theoretical basis for the intelligent upgrade of social network service platform.

1. Introduction

In the context of the widespread application of social media platforms, the generation of massive data provides a strong support for the growth of platforms. The current research focuses on how to efficiently mine and use these data to enhance the interactive experience of users and the service function of the platform. As an effective technical tool, big data analysis not only helps the platform achieve personalized content recommendation, but also improves the platform structure design and user interaction process. This study focuses on the design principles of data-driven social media platforms and the optimization path of big data analysis technology, discusses its specific applications in user behavior research, personalized recommendation, public opinion monitoring and other fields, and provides strategic suggestions for the sustainable development of the platform.

2. The relationship between big data analysis and social networks

On the social platform, there will be a huge amount of user information every day, such as interactive dynamics, content creation, message feedback, like interaction and information sharing. This information contains a lot of value points to be explored, the use of big data technology for in-depth analysis of this information, can reveal the user's personalized preferences, behavioral habits and other characteristics, help the social platform to more accurately grasp the needs of users. Through detailed analysis of this data, the platform can build a detailed personal portrait of the user, and then achieve customized content push, enhance user activity and loyalty. In addition, big data analysis can also guide the platform to carry out accurate advertising positioning, ensure that advertising is more efficient, and target potential consumer groups, so as to enhance the influence

and conversion efficiency of advertising.

In addition, big data analysis helps the platform to accurately optimize content layout and user interaction. Through in-depth exploration of social media information dissemination, the platform can gain insight into which content is more popular with users and which time points are more likely to get attention, so as to optimize the production and push of content. More critically, big data analysis technology provides social platforms with real-time monitoring of the dynamics of public opinion. Through the immediate processing of various comments, articles, discussions and other materials published by users, the platform can quickly detect possible public opinion crises, implement effective control and management, and prevent the spread of negative information. On the social network platform, big data analysis is not only a tool to improve user satisfaction, it has also become a key strategy to enhance the operational efficiency of the platform and ensure the robust operation and security of the platform.

3. Design and architecture of data-driven social network platform

3.1 Platform architecture design

When constructing the social communication network architecture with data as the core, we should not only focus on the operation efficiency, scalability and immediacy of response of the system, but also ensure that it can carry massive information and meet the diversified and complex needs of users. This type of architecture typically consists of multiple levels and functional modules, each with a specific function, which work closely together and interact frequently to ensure the smooth operation of the entire platform. In addition, the architecture needs to demonstrate strong adaptability and expansion potential to cope with the possible future diversification of demands and technological advances.

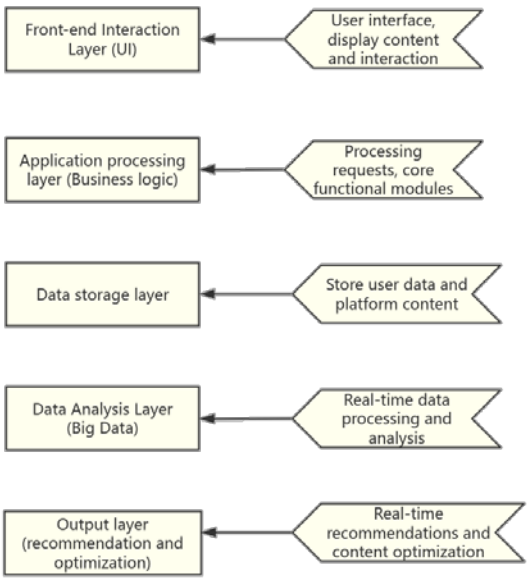

Figure 1. Platform architecture diagram

The platform architecture consists of front-end interaction layer, application processing layer, data storage layer, data analysis layer and other important modules (see Figure 1). The front-end interaction layer is mainly used to process various interactive behaviors of users, including reading, publishing and social interaction of information, and is presented to users through mobile phones

and computers. This layer must provide an efficient human-computer interaction interface and a fast response mechanism to ensure the user experience. The application processing layer is the backbone of the whole system, which undertakes the core tasks of handling user instructions, social interaction logic, personalized recommendation algorithm, content distribution and message notification. Relying on a refined algorithm model, the recommendation system at this layer creates personalized content push by analyzing user behavior data, and continuously improves user experience through back-end algorithms. In the data storage layer, its functions cover relational databases (mainly used to store user information and basic data) and distributed data storage systems (dealing with the storage and management of massive data). This layer is designed to ensure efficient data storage and quick retrieval. The final data analysis layer focuses on real-time processing and analysis of user behavior and social activity data. Leveraging big data processing technologies, such as Hadoop and Spark, for deep data mining and analytics to assist the platform in accurate content recommendation, optimization, and prediction of user behavior.

3.2 Data collection and processing mechanism

Data collection is the core of platform design. Major social network platforms adopt a variety of methods to gather users' behavioral information on the platform (including but not limited to clicking, interactive comments, approval and information dissemination), and collect the original data of uploaded content information (such as articles, pictures, videos, etc.). The collected information needs to be initially screened and purified, and then properly stored in the database for subsequent analysis and mining. Assuming that the platform needs to process user click behavior data, the data collection formula can be expressed as:

$$D_{click} = \{U_i, C_j, T_k\} \tag{1}$$

Among them, D_{click} is click data, U_i is user ID, C_j is content ID, and T_k is click time stamp.Through this structured data storage, the platform can connect each user to the content, laying the foundation for subsequent recommendations and behavioral analysis.

The data processing mechanism covers data preprocessing, feature extraction, data storage and real-time processing. In the process of data cleaning and preprocessing, the main task is to remove impurities and supplement incomplete data items. Then, the system will use feature extraction technology to model the user's behavior information, so as to provide necessary data support for the subsequent operation of personalized recommendation algorithm.

3.3 Recommendation system design in social networks

Recommendation systems in social networks generate personalized content recommendations based on users' historical behavior data. Common recommendation algorithms include collaborative filtering, content-based recommendation and hybrid recommendation. In order to improve the accuracy of recommendations, a combination of algorithms is often used.

A common collaborative filtering algorithm is user-based collaborative filtering algorithm, which recommends the content liked by other similar users by calculating the similarity between users. Assuming that user U scores content C as $r(U,C)$, user U is predicted score can be expressed as:

$$\hat{r}(U,C) = \frac{\sum_{V \in N_U} \omega(U,V) \cdot r(V,C)}{\sum_{V \in N_U} |\omega(U,V)|} \tag{2}$$

Where, N_U is the set of users similar to user U, $\omega(U,V)$ is the similarity between user U and user V, and $\hat{r}(U,C)$ is the predicted score of user U on content C. By constantly adjusting the algorithm, the recommendation system can provide users with more accurate content recommendation, thus improving the user's activity and platform stickiness.

3.4 Interactive interface and user experience optimization

The interaction hub between users and social network platforms is the interactive interface, the quality of which is directly related to user satisfaction. In the process of data-based social platform design, the improvement of interactive interface includes not only the optimization of visual presentation, but also the dynamic optimization of interface layout and function based on user behavior data.

In terms of improving the user experience, personalized interface adjustment is one of the key links. By analyzing user behavior records, the platform can gain insight into user preferences, and then automatically personalize content display or adjust interface elements. For example, if a user frequently interacts with video content, the platform will prioritize video content and increase the visibility of the video player. A common formula for optimizing user behavior is:

$$UX_{opt} = \sum_{i=1}^{n} (\omega_i \cdot f_i) \tag{3}$$

Where, UX_{opt} indicates the degree of user experience optimization, ω_i indicates the weight of each optimization indicator, and f_i indicates the value of the corresponding optimization indicator (such as loading speed and interface friendliness). Through weighted summation, the platform can comprehensively evaluate the optimization effect of user experience and make targeted adjustments. The platform can further optimize the interactive interface design by monitoring user feedback and behavior in real time, so as to continuously improve the user experience and enhance user stickiness and satisfaction.

4. Application of big data analysis and optimization in social network platforms

4.1 User behavior analysis and platform optimization

The interactive footprint of users on social media constitutes the platform's valuable resource library, which includes users' clicks, browsing, interactive comments and content sharing. Through the in-depth mining of these data, the platform can accurately capture the actual needs of users, dig out potential usage habits, and then promote the continuous upgrading of platform functions and service quality.

By analyzing user behavior data, social platforms can grasp users' preferences and optimize the presentation of content, advertising positioning strategy and content recommendation mechanism in real time, so as to enhance users' sense of belonging and activeness. For example, the platform can customize personalized content or friend recommendations for users based on their browsing and interaction logs to optimize the user's personalized experience. At the same time, the analysis of user behavior data can also help the platform distinguish between active users and users who are likely to lose, so as to implement effective user retention strategies in a timely manner.

Table 1: Key indicators of user behavior analysis and platform optimization

index	Description	Data source	Analytical application
User activity (DAU/MAU)	Measure the daily/monthly active users of the platform, reflecting the level of activity of the platform	User logins, posts, comments, likes, shares and other behavioral data	Optimize content push and social interaction modules
User stickiness	A measure of user dependence on the platform, usually expressed in terms of retention or activity	The login frequency and duration of the user	Identify lost users and optimize retention strategies
User interest preference	Analyze users' interests and types of content through their interactive behavior	Content data that users view, click, like, and share	Recommend precise content and social circles
User social behavior	Users' social interactions on the platform, including comments, sharing and interaction	User comments, shares, private messages, likes	Enhance social interaction and enhance platform activity

Through the analysis of the key indicators in Table 1 above, the platform can deeply understand the behavior pattern of users in the process of use, so as to accurately optimize the content display, advertisement delivery and recommendation system, and improve the user experience of the platform. At the same time, detailed analysis of behavioral data helps the platform identify potential loss users and take timely intervention measures to improve user retention and engagement.

4.2 Content recommendation and precision marketing optimization

The content recommendation system of the social media platform plays an extremely important role. It displays the information associated with the user according to the user's preference, interaction behavior and interpersonal circle, aiming to improve the user's interaction frequency and platform use time. Relying on big data analysis technology, the content recommendation mechanism can analyze the user's online browsing trajectory, interactive information and social network structure, and continue to improve the accuracy of the recommendation algorithm to achieve highly customized content distribution.

Precision marketing strategy relies on in-depth data analysis, detailed mining for users' personalized information, and provides more accurate advertising targeting services for businesses. In the social network space, the precision of advertising delivery has a decisive impact on the revenue growth of the platform. Through multidimensional data analysis of users' behavior patterns, interest points, social networks, etc., the platform can help advertisers implement accurate advertising push, thereby improving the click-through rate and transaction rate of advertisements, and reducing advertising cost.

As shown in Table 2, through the intelligent recommendation of content and the precise adjustment of marketing strategies, the platform can present users with more appropriate information and services to their needs, and at the same time create more effective promotion plans for advertisers. Relying on the accurate analysis of big data, this optimization path significantly enhances the interactive experience of users, and significantly improves the commercial efficiency and income level of the platform.

Table 2: Key applications of content recommendation and precision marketing optimization

Application field	Description	Data source	Optimization objective
Content recommendation system	Recommend content of interest to users based on their historical behavior and social relationships	Users click, browse, share, comment and other behavioral data	Increase user engagement and increase time on the platform
Targeted advertising	Provide advertisers with accurate advertising recommendations based on user profiles and behavioral data	Users' interests, behavior, geographic location, social data	Increase AD conversion rate and optimize AD revenue
Social circle recommendations	Recommend potential friends or groups based on the user's social connections and similarity of interests	The user's friend relationship, interaction history, interest tag data	Increase user social stickiness and optimize user network effect
Content diversity and personalized	Recommend diverse content according to users' preferences to avoid excessive uniformity of information	User browsing history, behavior feedback data	Improve user experience and avoid recommendation fatigue

4.3 Sentiment analysis and public opinion monitoring in social networks

The content of social networking platforms is full of emotional expressions from users, including comments, posts, hashtags, etc. With the rapid expansion of user groups, the generation and dissemination of public opinion is also accelerating, especially the rapid spread of negative emotions, which may have adverse effects on the brand reputation of social media and the interaction between users. In view of this, accurate analysis of emotions and implementation of public opinion monitoring are particularly critical in the operation of social media.

Through natural language processing technology, sentiment analysis can judge the emotional tendency of the content generated by users, and distinguish the positive, negative or neutral emotions contained in the text. This analysis process often requires the automation of large volumes of user-generated content (UGC) so that the platform can spot potentially risky issues, adverse emotional reactions, and early signs of crisis. The key to sentiment analysis is to assign an emotion value to each piece of information through an algorithm to provide decision support to the platform. A common sentiment analysis formula is:

$$S = \sum_{i=1}^{n} \omega_i \cdot t_i \tag{4}$$

Where S represents the emotion score of the text, n is the number of words in the text, and ω_i is the emotion weight of word t_i, usually obtained by pre-trained emotion dictionaries or machine learning models. The algorithm deduces the emotional color of the whole text by summarizing the emotional value of the words in the text. A positive score indicates that the text conveys positive emotions, while a negative score indicates that the text has negative emotions. A score approaching zero means that the emotion presented in the text is neutral.

With the help of sentiment analysis technology, the platform can instantly track the trend of public opinion on social networks, quickly capture and respond to user mood fluctuations, effectively deal with possible crisis situations, and prevent further deterioration of public opinion. For example, once the detected negative sentiment indicator exceeds the established limit, the platform will activate the emergency response procedure, issue an official statement or adjust the information distribution strategy to maintain the brand image and good interaction among users.

4.4 Platform performance optimization and big data application

The smooth operation of social media platforms depends on the continuous improvement of their performance, especially in the context of surging user visits and massive data interaction, system reliability and rapid response become the core elements. In this context, data analytics plays an indispensable role in improving platform performance. Through real-time monitoring and in-depth analysis of the platform's running status, such as the monitoring of user operation frequency, server workload, database feedback time and other indicators, the platform can quickly discover and deal with possible performance obstacles and implement corresponding optimization strategies.

Through the use of distributed computing and storage technologies, big data significantly enhances the efficiency of processing and resource allocation, which is critical to optimizing platform operations. For example, the platform can rely on real-time traffic data, independently achieve load balancing, flexibly adjust the allocation of computing resources, and ensure that users' requests at different points in time can be quickly fed back. At the same time, through real-time monitoring and analysis of data streams and logs, the platform can detect possible hidden faults or performance deterioration signs in time, so as to give early warning and take debugging measures to prevent large-scale system failures or service interruptions.

5. Conclusion

With the rapid progress of big data technology, the operation strategy and service improvement of social media have entered a period of innovation. Through in-depth mining of user behavior patterns, customized information push, accurate marketing methods, and the use of emotion tracking and public opinion monitoring, major platforms can more accurately grasp user needs and enhance user loyalty and activity. In addition, the platform uses massive data analysis to continuously improve the efficiency of the system, ensuring smooth operation in traffic peaks and changing environments. With the continuous expansion of data scale and continuous upgrading of technology, data-centric service optimization strategies will play an increasingly critical role in the field of social media, promoting the in-depth development of platform intelligence and personalization.

References

[1] *Tian X, Zhu R, Xu G. Construction of Big Data Analysis Platform for College Students' Sports Training Driven by Wireless Communication Network. International Journal of Emerging Technologies in Learning, 2024, 19(1).*

[2] *Dong R.A Study on the Application of Social Network Analysis Based Approach in the Optimization of Cultural Communication Paths and Efficiency. Applied Mathematics and Nonlinear Sciences, 2024, 9(1).*

[3] *Manupati V K, Putnik G, Jacob S M, et al. A social network and simulation-based optimization approach for manufacturing resilience in a distributed manufacturing system. Procedia CIRP, 2024, 130:1221-1226.*

[4] *Alghamdi A. A Hybrid Method for Customer Segmentation in Saudi Arabia Restaurants Using Clustering, Neural Networks and Optimization Learning Techniques. Arabian Journal for Science & Engineering (Springer Science & Business Media B.V.), 2023, 48(2).*

[5] *W. Han, "Using Spark Streaming Technology to Drive the Real-Time Construction and Improvement of the Credit Rating System for Financial Customers," 2025 International Conference on Intelligent Communication Networks and Computational Techniques (ICICNCT), Bidar, India, 2025, pp. 1-6, doi: 10.1109/ICICNCT66124.2025.11232932.*

[6] J. Huang, "Research on Multi-Model Fusion Machine Learning Demand Intelligent Forecasting System in Cloud Computing Environment," 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), Hassan, India, 2025, pp. 1-7, doi: 10.1109/IACIS65746.2025.11210946.

[7] Ding, J. (2025). Research On CODP Localization Decision Model Of Automotive Supply Chain Based On Delayed Manufacturing Strategy. arXiv preprint arXiv:2511.05899.

[8] X. Liu, "Research on User Preference Modeling and Dynamic Evolution Based on Multimodal Sequence Data," 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), Hassan, India, 2025, pp. 1-7, doi: 10.1109/IACIS65746.2025.11211273.

[9] D. Shen, "Complex Pattern Recognition and Clinical Application of Artificial Intelligence in Medical Imaging Diagnosis," 2025 International Conference on Intelligent Communication Networks and Computational Techniques (ICICNCT), Bidar, India, 2025, pp. 1-8, doi: 10.1109/ICICNCT66124.2025.11232821.

[10] Q. Hu, "Research on Dynamic Identification and Prediction Model of Tax Fraud Based on Deep Learning," 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), Hassan, India, 2025, pp. 1-6, doi: 10.1109/IACIS65746.2025.11211426.

[11] F. Liu, "Transformer XL Long Range Dependency Modeling and Dynamic Growth Prediction Algorithm for E-Commerce User Behavior Sequence," 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), Hassan, India, 2025, pp. 1-6, doi: 10.1109/IACIS65746.2025.11211467.

[12] Zhu, P. (2025). The Role and Mechanism of Deep Statistical Machine Learning In Biological Target Screening and Immune Microenvironment Regulation of Asthma. arXiv preprint arXiv:2511.05904.

[13] Liu, B. (2025). Design and Implementation of Data Acquisition and Analysis System for Programming Debugging Process Based On VS Code Plug-In. arXiv preprint arXiv:2511. 05825.

[14] Zou, Y. (2025). Design and Implementation of a Cloud Computing Security Assessment Model Based on Hierarchical Analysis and Fuzzy Comprehensive Evaluation. arXiv preprint arXiv:2511.05049.

[15] Su H, Luo W, Mehdad Y, et al. Llm-friendly knowledge representation for customer support[C]//Proceedings of the 31st International Conference on Computational Linguistics: Industry Track. 2025: 496-504.

[16] Wu Y. Software Engineering Practice of Microservice Architecture in Full Stack Development: From Architecture Design to Performance Optimization. 2025.

[17] Wu Y. Optimization of Generative AI Intelligent Interaction System Based on Adversarial Attack Defense and Content Controllable Generation. 2025.

[18] Sun J. Quantile Regression Study on the Impact of Investor Sentiment on Financial Credit from the Perspective of Behavioral Finance. 2025.

[19] Wang Y. Application of Data Completion and Full Lifecycle Cost Optimization Integrating Artificial Intelligence in Supply Chain. 2025.

[20] Mingjie Chen. (2025). Exploration of the Application of the LINDDUN Model in Privacy Protection for Electric Vehicle Users. Engineering Advances, 5(4), 160-165.