

# Revolutionizing social media: The applications and implications of big data analytics

Qiyao Sun<sup>1,3</sup> and Lu Liu<sup>2</sup>

<sup>1</sup> Beijing Huijia Private School, No. 157, Changhuai Road, Changping, Beijing, P.R. China

<sup>2</sup> Beijing Huijia Private School, No. 157, Changhuai Road, Changping, Beijing, P.R. China

<sup>3</sup>jamsunqiyao@gmail.com

**Abstract.** The increasing deployment of online social channels has generated an unprecedented volume of data that can be analyzed to provide valuable insights into human behavior, preferences, and trends. The progression of extensive data analysis is clear. As a crucial technology for processing and making sense of this vast amount of data, enabling businesses and organizations to better understand their customers and improve their marketing strategies. Big data analytics have the potential to revolutionize social media by unlocking valuable insights and enhancing user experiences. However, the collection and analysis of user data raise ethical questions regarding consent and user control over their personal information. Moreover, algorithmic biases can lead to discriminatory outcomes and echo chambers, impacting society at large. It is crucial to address these issues and establish appropriate regulations and safeguards to mitigate potential harm. This paper explores the applications and implications of big data analytics in social media, highlighting its benefits and challenges and discussing its impact on society as a whole.

**Keywords:** human behavior, preferences, trends, privacy, data security.

## 1. Introduction

The advent of social media platforms such as Facebook, Twitter, Instagram, TikTok and LinkedIn has revolutionized communication and networking on a global scale. These platforms have become virtual arenas where billions of individuals share personal information, opinions, and experiences. Consequently, they generate an enormous volume of data, including text, images, videos, and user interactions. This abundance of data poses a significant challenge in terms of storage, processing, and analysis. However, big data analytics offers a powerful solution to extract valuable insights from this vast and complex dataset.

The applications of big data analytics in social media are manifold. One prominent area is marketing and advertising. By analyzing user preferences, demographics, and online behaviors, organizations can tailor their marketing strategies to target specific customer segments more effectively. This data-driven approach enables companies to optimize their advertising campaigns, enhance customer engagement, and improve overall business performance.

Moreover, big data analytics can facilitate sentiment analysis and opinion mining, providing a deeper understanding of public opinion and sentiment towards various products, services, or social issues. This information can be invaluable for policymakers, businesses, and researchers in making informed decisions and developing effective strategies.

Furthermore, big data analytics in social media has the potential to enhance customer service and support. By analyzing customer feedback, complaints, and inquiries on social media platforms, organizations can identify emerging trends, address issues promptly, and provide personalized and timely support.

However, while big data analytics presents numerous opportunities, it also raises concerns regarding privacy, security, and ethical considerations. The vast amount of personal information available on social media platforms raises questions about data protection, consent, and potential misuse. Addressing these challenges is crucial to ensure the responsible and ethical use of big data analytics in the social media realm.

This paper explores the applications and implications of big data analytics in social media, highlighting its benefits and challenges and discussing its impact on society as a whole. In particular, this paper aims to examine the key applications of big data analytics in social media. Moreover, this paper will critically analyze the ethical, legal, and social implications of big data analytics in social media, and evaluate the positive and negative impact of big data analytics on society.

## **2. Literature review**

The literature on the applications and implications of big data analytics in social media is extensive and covers a wide range of topics. This section provides a brief overview of some of the key findings from recent studies on this topic.

One area of research focuses on the use of big data analytics in social media for public health surveillance. Studies have shown that data analysis of social media can be used to track the source of transmissible infectious diseases and monitor public sentiment and behavior such as vaccination [1]. This information can be used to inform public health interventions and improve health outcomes.

Another area of research focuses on the use of big data analytics in social media for political campaigning and elections. Research has shown that social media data can be used to identify voter preferences and target campaign messages to specific demographic groups [2]. There are also concerns that these technologies could be used to manipulate public opinion or undermine democratic processes [3].

The use of big data analytics in social media also raises concerns about bias and discrimination. Studies have shown that algorithms used in social media platforms can perpetuate existing biases and discrimination, particularly with respect to race and gender [4].

Recent literature has also explored the potential for big data analytics in social media to support mental health interventions. The use of social media data to identify individuals at risk of depression [5]. Researchers analyzed data from Twitter and found that linguistic features such as first-person pronouns or words of negative emotion could be used to predict whether a speaker was depressed. This information can be used to develop targeted interventions and support for individuals at risk of mental health issues.

A paper explored the use of big data analytics in social media marketing [6]. The researchers analyzed data from WeChat, a popular social media platform in China, and found that personalized recommendations based on user behavior and interests significantly improved the effectiveness of marketing campaigns. This suggests that big data analytics can help marketers better target their audiences and improve ROI.

Overall, the literature suggests that big data analytics in social media has the potential to revolutionize public health, political campaigning, mental health interventions, and other domains. However, careful consideration must be given to the ethical and social implications of its use, in order to minimize potential harms and maximize its potential benefits.

### 3. Applications of big data analytics in social media

The application of big data analysis in social media mainly has the following directions:

**Targeted Advertising:** Social media platforms are known to generate vast amounts of user behavior data, which can be used to target advertisements to specific groups of people, and then by analyzing user data such as past purchases, browsing history, and social media activity, advertising merchants can analyze the effect of advertisements, and can also create more matching advertisements to find more suitable customers. A study performed found that personalized recommendations based on user behavior and interests significantly improved the effectiveness of marketing campaigns [6].

**Brand Reputation Management:** Social media can be a valuable source of information for companies looking to manage their brand reputation. By analyzing social media data, companies can identify negative sentiment towards their brand and take proactive steps to address the underlying issues. A research paper analyzed social media data to identify factors contributing to negative sentiment towards a hotel brand [7]. The researchers found that issues with cleanliness and customer service were major drivers of negative sentiment, which allowed the hotel to take steps to address these issues.

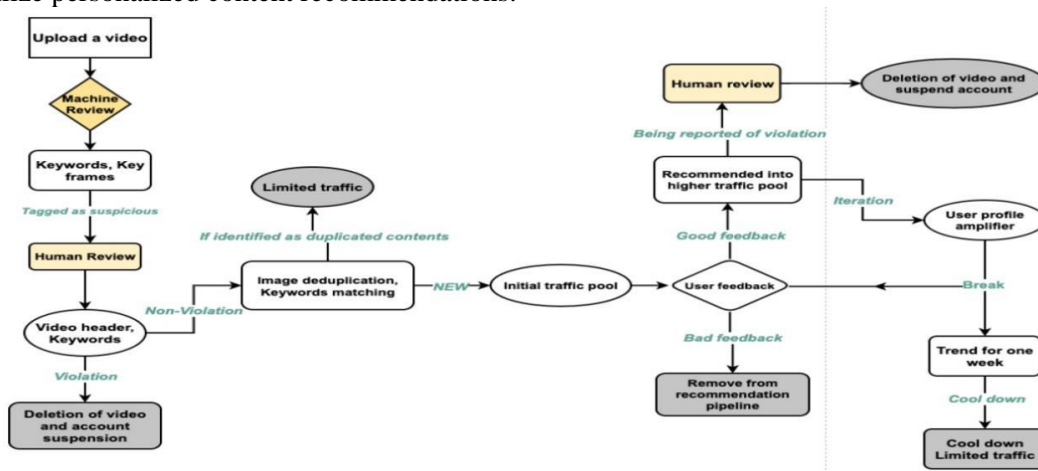
**Customer service:** Social media can also be a tool to improve after-sales service ratings. By monitoring social media channels, companies can quickly respond to customer issues and deal with their issues and complaints, thereby increasing customer satisfaction and loyalty. A research analyzed social media data to identify common issues customers had with a popular ride-sharing app [8]. The company was able to use this information to improve their customer service processes and reduce the number of complaints.

**Public Opinion and Sentiment Analysis:** Social media can also be used to gain insights into public opinion and sentiment. Kuchler and colleagues used social media data to track the spread of COVID-19 and to identify trends and patterns related to public attitudes and behaviors related to the pandemic [9].

At the same time, for big data analysis in social media, these points are worthy of attention.

#### 3.1. Identify the key applications

There are many social media platforms emerging, such as TikTok, a short video platform that has grown to be a more popular social media application in the mainstream. TikTok has been downloaded 3.5 billion times worldwide on the Android and IOS. TikTok's core recommendation algorithm and video classification technology uses AI deep learning to analyze and understand multimodal information such as video images, audio, and text, determine video topics and user interests, and realize personalized content recommendations.



**Figure 1.** Recommendation algorithm workflow.

TikTok never reveals its core algorithm to the public. Catherine Wang draw up the above TikTok's the recommendation work flow based on the fragmented information posted via the company, and trails discovered by geeks using the reverse engineer techniques [11].

One of the main algorithms for TikTok videos is dual review, machine review and human review before posting a successful video as we can see from the above workflow. the Duo-audit model is based on computer vision that recognizes the user's video image and keywords to perform initial machine review. For content marked as suspicious by the Duo-Audit model, technicians will manually review it, and if violations are determined, this video will be deleted and blocked. When the content passes the double audit, it will be put into the initial traffic pool. After the video is published and gets many likes and favorites, it will get more exposure from the TikTok. We can see the right-side workflow , after passing through the initial traffic pool, videos can receive thousands of views, and this data is in turn collected and analyzed, again with metrics such as: views, full views, comments, likes, follows, etc. The engine then measures the weight score of the content based on this initial video data and your account score.

User Profile Amplifier : Higher quality content will be pushed and exposed to a wide range of specific user classification groups (e.g. lifestyle, entertainment, sports, etc.)[11].

If a content passes the machine and human review, iterations, and amplifiers, the creator's account will receive a disproportionate amount of exposure, user interaction, and followers. Usually, this high exposure, high traffic skew will last for about a week. After this time frame, this content and account will cool down, and even subsequent videos will hardly become hot trending.

Entertainment user group scores						
User	Complete views 1	Likes 2	Comments 3	Reposts 4	Shares 5	Scores
1	5*1	5*2	3*3	2*4	1*5	37
2	2*1	1*2	0*3	0*4	0*5	4
3	4*1	4*2	2*3	2*4	1*5	31
4	1*1	1*2	0*3	0*4	0*5	3
Lifestyle user group scores						
User	Complete views 1	Likes 2	Comments 3	Reposts 4	Shares 5	Scores
1	2*1	1*2	0*3	0*4	0*5	4
2	4*1	4*2	2*3	2*4	2*5	36
3	1*1	1*2	0*3	0*4	0*5	3
4	5*1	5*2	3*3	2*4	2*5	42

**Figure 2.** User group feedback scores.

Suppose there are 12 TikTok users, and these users watch videos with 4 kinds of preferences: Entertainment, Lifestyle, Sports and Fashion, how to determine the preferences of each user? For example there are 20 videos,5 each of Entertainment, Lifestyle, Music, Sports and Fashion, and the type of each video is known. These 20 videos are randomly placed to these 12 users, and then rated according to their feedback. Assume that a complete view is 1 point, a like is 2 points, a comment is 3 points, a repost is 4 points, and a share is 5 points. In this way, after placing 20 videos, the user's feedback on the video will be available and then the user's preference will be known. For example, User1 watched all of entertainment videos, and liked 5 videos, commented 3 videos, reposted 2 videos, and shared one video, then, this user scored 37 points which means user 1 got higher scores in the group of entertainment and likes entertainment videos more, so this user can be rated as a entertainment video lover. Therefore 12 users can be grouped, a user can also have multiple preferences, that is, one user can belong to multiple groups (as shown in Figure 3).

Overall, TikTok's recommendation algorithm mechanism is the process of rating the videos posted by the creators. This process is based on all the data of TikTok users watching videos to classify them. In addition to metric data about the user's video viewing, personal information, Internet history, and age will be collected to create a more accurate and personalized user profile so that the user stays in their system for as long as possible.

### 3.2. Analyze the ethical, legal, and social implications

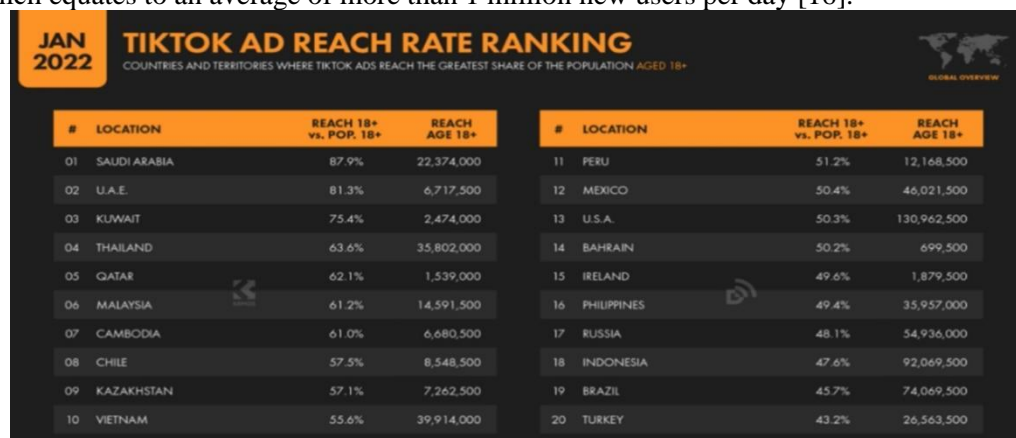
The utilization of data analytics on social media platforms at a large scale can potentially collect and analyze personal information, which raises concerns about privacy [12]. Users often have limited

knowledge about the types of data collected, how it is used, or who has access to it [13]. This lack of transparency leaves individuals unsure about how their personal information is being utilized, compromising their privacy and control over their own data. Big data analytics may enable surveillance of users on social media platforms, which raises concerns about government surveillance and infringement on individual rights [14]. Social media platforms are attractive targets for hackers due to the vast amount of user data they possess. Data breaches can lead to significant legal consequences, including financial penalties and reputational damage for the companies involved.

Algorithmic bias is another privacy concern. Big data analytics heavily relies on algorithms to analyze and interpret vast amounts of information [15]. However, these algorithms are not immune to biases, and if they are trained on biased or incomplete datasets, they may perpetuate and amplify existing social, racial, or gender biases. This can lead to discriminatory practices and violate users' privacy by influencing the content they see or the decisions made about them based on their data.

### 3.3. Explore future directions and trends of big data analytics in social media

According to DIGITAL 2022: GLOBAL OVERVIEW REPORT, There were 4.62 billion social media users worldwide in January 2022, a figure equivalent to 58.4% of the world's total population. Over the past 12 months, global social media users have grown by more than 10%, adding 424 million new users, which equates to an average of more than 1 million new users per day [16].



**Figure 3.** TikTok ad reach rate ranking.

TikTok's ad audience also continues to grow at an impressive rate. The latest data released by Bytedance Advertising Resources shows that TikTok's ad reach has increased by 60 million users in the last 90 days (+7.3%) and is approaching 885 million worldwide by early 2022. However, it is important to note that users under 18 are not counted in the data due to national data security requirements, but obviously there are users under 18 in the factual situation, so marketers will still cover users 13 and older when planning.

Big data analysis is widely used in social media marketing, different from traditional forms of advertising and marketing, mainly in the form of multi-dimensional data to analyze and identify the effect of the state of brand operation or promotion, as a support, to summarize the subsequent optimization strategy plan; The company's social media operation can rely on various data (in the case of TikTok, for example, the number of active followers, the number of highly interactive followers, the amount of videos watched, etc.) as a basis for evaluation, and analyze what kind of content users like to see the most by publishing different kinds of content. In this way, social media platforms can quantitatively and precisely deliver advertisements.

Big data analytics and machine learning form artificial intelligence is changing various industries at an accelerated pace, and the streaming media field is probably one of the fastest changing. With the problem of algorithms related to neural networks solved, artificial intelligence technology has been developed rapidly in recent years, and the penetration of AI technology in the field of streaming media has led to a new breakthrough in this technology [17]. Pictures, long videos, short videos, live

streaming, AR and other various media forms occupy the Internet circle, and in today's media content and forms are very rich, how to identify and analyze these contents and feed them through AI is the focus of all tech giants at present, among which image recognition, voice semantic recognition, simultaneous interpretation, subtitle recognition and other application scenarios are further explored, all need to collect a large amount of data to perform analysis.

Social media platforms can better serve user needs and further improve user experience through automatic data processing and personalized recommendations. In addition, social media advertising realizations and commercialization have also been improved by the continuous development of technology. However, these technologies also bring some hidden dangers. For example, automatically pushed content may enhance the filtering of information by users. This information filtering may affect that user's vision and comprehensive access to information [18]. Therefore, future research needs to think carefully about how this technology can be applied to better protect user needs and interests.

Search engines apply artificial intelligence as well as big data to improve the accuracy and speed of search results. For example, in the process of intelligent recommendation, the search engine platform can use the user's historical search records, geographic location, purchase behavior and other information to provide customized recommendation services for users. Meanwhile, the development of technologies such as semantic understanding and natural language processing can also provide more accurate and intelligent search services. These applications provide a better search experience for people [19]. However, the recommendation algorithms of search engine platforms are not perfect. Recommendation algorithms may lead to information polarization problems, further driving users to filter information on their own. Therefore, further research should be more careful in using these technologies and protecting our personal information from misuse.

In conclusion, the use of big data analytics and AI in social media and search engines has obvious effects on improving people's experience and increasing efficiency. However, these technologies can also have far-reaching effects on our society and humanity. We should think carefully about the potential problems with these technologies and take appropriate measures to take full advantage of their benefits while avoiding the negative aspects of their impact.

#### **4. Implications of big data analytics in social media**

Social media has evolved into an essential component of contemporary society, with a vast number of users actively engaging with various platforms on a daily basis. With the vast amount of data generated from these interactions, The significance of big data analytics in comprehending and utilizing the valuable insights derived from social media has grown exponentially. While big data analytics has many potential benefits in social media, it also has several implications that need to be considered. In this essay, we will explore the implications of big data analytics in social media.

One of the major implications of big data analytics in social media is the potential impact on user privacy. Social networking platforms amass extensive volumes of data from their users, including behavioral data such as age, gender, browsing history, search queries, and more. While these data can be used to improve user experience and personalize content, it can also lead to privacy breaches, identity theft, and other forms of cybercrime.

Another implication of big data analysis in social media is the potential for bias and discrimination. Are the algorithms or data used in big data analysis percentage fair and fair? In fact, If the data used to train these algorithms exhibits bias, the subsequent outcomes may also be biased. The consequences of such biases could suddenly lead to discrimination against certain groups of users and could exacerbate existing social inequalities. Therefore, it is imperative to ensure that the data used in big data analysis is diverse and representative of all users.

Despite these implications, big data analytics also has several positive impacts on society. For example, it can help identify and track the spread of infectious diseases, which can lead to faster and more effective responses. It can also be used to monitor and analyze social media conversations related to mental health, which can lead to better understanding and treatment of mental health

conditions. Additionally, big data analytics can help identify and address online harassment and hate speech, which can make social media a safer and more inclusive space. As the number of social media content creation platforms and the diversity of content make data analysis increasingly complex, AI can help content creators analyze and predict data more accurately to better develop subsequent content plans and marketing strategies.

In conclusion, the implications of big data analytics in social media are complex and multifaceted. While there are concerns about user privacy, bias and discrimination, and political influence, there are also many potential benefits for society, including improved health outcomes and increased safety on social networking platforms. It is important to ensure that the development and use of big data analytics in social media is guided by ethical principles and a commitment to transparency and accountability.

## References

- [1] Luo, C., Chen, A., Cui, B., & Liao, W. (2021). Exploring public perceptions of the COVID-19 vaccine online from a cultural perspective: Semantic network analysis of two social media platforms in the United States and China. *Telematics and Informatics*, 65, 101712.
- [2] Weismueller, J., Harrigan, P., Coussemont, K., & Tessitore, T. (2022). What makes people share political content on social media? The role of emotion, authority and ideology. *Computers in Human Behavior*, 129, 107150.
- [3] Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European journal of communication*, 33(2), 122-139.
- [4] Noble, S. U. (2018). Algorithms of oppression. In *Algorithms of oppression*. New York University Press.
- [5] Ortega-Mendoza, R. M., Hernández-Farías, D. I., Montes-y-Gómez, M., & Villaseñor-Pineda, L. (2022). Revealing traces of depression through personal statements analysis in social media. *Artificial Intelligence in Medicine*, 123, 102202.
- [6] Teng, T., Li, H., Fang, Y., & Shen, L. (2022). Understanding the differential effectiveness of marketer versus user-generated advertisements in closed social networking sites: An empirical study of WeChat. *Internet Research*.
- [7] Yadav, M. L., & Roychoudhury, B. (2019). Effect of trip mode on opinion about hotel aspects: A social media analysis approach. *International Journal of Hospitality Management*, 80, 155-165.
- [8] Islam, T., Kundu, A., Lima, R. J., Hena, M. H., Sharif, O., Rahman, A., & Hasan, M. Z. (2023). 5 Review Analysis of Ride-Sharing Applications Using Machine Learning Approaches. *Computational Statistical Methodologies and Modeling for Artificial Intelligence*.
- [9] Kuchler, T., Russel, D., & Stroebel, J. (2022). JUE Insight: The geographic spread of COVID-19 correlates with the structure of social networks as measured by Facebook. *Journal of Urban Economics*, 127, 103314.
- [10] Wang, C. (2020). Why TikTok made its user so obsessive? The AI Algorithm that got you hooked. *Towards Data Science*.
- [11] Appel, G., Grewal, L., Hadi, R., & Stephen, A. T. (2020). The future of social media in marketing. *Journal of the Academy of Marketing science*, 48(1), 79-95.
- [12] Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information & Management*, 57(2), 103169.
- [13] Engin, Z., & Treleaven, P. (2019). Algorithmic government: Automating public services and supporting civil servants in using data science technologies. *The Computer Journal*, 62(3), 448-460.

- [14] Di Minin, E., Fink, C., Hausmann, A., Kremer, J., & Kulkarni, R. (2021). How to address data privacy concerns when using social media data in conservation science. *Conservation Biology*, 35(2), 437-446.
- [15] Simon Kemp. (2022). DIGITAL 2022: GLOBAL OVERVIEW REPORT. <https://datareportal.com/reports/digital-2022-global-overview-report>.
- [16] Regan, P. M., & Jesse, J. (2019). Ethical challenges of edtech, big data and personalized learning: Twenty-first century student sorting and tracking. *Ethics and Information Technology*, 21, 167-179.
- [17] Ofli, F., Alam, F., & Imran, M. (2020). Analysis of social media data using multimodal deep learning for disaster response. *arXiv preprint arXiv:2004.11838*.
- [18] Sarker, I. H., Hoque, M. M., Uddin, M. K., & Alsanoosy, T. (2021). Mobile data science and intelligent apps: concepts, AI-based modeling and research directions. *Mobile Networks and Applications*, 26, 285-303.