International Journal of Scientific Research & Engineering Trends
Volume 11, Issue3 , May-jun-2025, ISSN (Online): 2395-566X

Big Data analysis in social media

Ms. Muskan Shaikh, Dr.Jasbir Kaur, Mr.Suraj Kanal
Student of Master's in Computer Applications (MCA) at Guru Nanak Institute of Management Studies, Matunga, Mumbai, India, mca23.muskan.shaikh@gnims.com
Director GNIMS B-School, Head of Information Technology and HR at Guru Nanak Institute of Management Studies, Matunga, Mumbai, India, jasbir.it@gnims.com
Assistant Professor at Guru Nanak Institute of Management Studies, Matunga, Mumbai, India,

Abstract- This paper discusses the importance and advantages of big data analysis and application in social media marketing. With the popularity of social media platforms, big data analysis provides enterprises with opportunities to deeply understand user needs, optimize marketing strategies and improve marketing effects. This paper introduces the current situation of social media marketing, and expounds in detail the application of big data analysis in user portrait analysis, user behavior analysis and marketing effect evaluation. Through big data analysis, enterprises can formulate more accurate marketing strategies, improve marketing accuracy, optimize user experience and improve marketing efficiency. However, big data analysis also faces challenges such as data quality and privacy protection, which requires enterprises to pay attention to data security and compliance in the process of application.

Keywords- Social Media Marketing; Big Data Analysis; User Portrait; User Behavior.

I. INTRODUCTION

'Big data' is fast becoming an area of great importance for businesses in many areas, including education. In simple terms it refers to the combination of data from various sources and understanding patterns in the data which can be used for various purposes such as improving market intelligence and educational research. Businesses, large and small, are implementing (or planning to implement) big data strategies. Apart from market intelligence, it is being applied in diverse areas such as healthcare and other scientific research, complex manufacturing industries such as aviation and heavy machinery, improving public utilities and traffic management, oil and gas exploration, telecoms, retail, banking and insurance, defense and security. In this article we give an introduction to big data and some of its applications in various fields, including education. We also describe the use of big data for the monitoring of social media (for instance LinkedIn, Facebook and Twitter) for market growth and brand management. Some training courses in big data offered by various universities are mentioned in the article.

II. Background of the study

Social big data makes it easier to organize consumers according to their interests and tailor advertisements to them. Individuals are rated according on their location, interests, gender, and age. People are more likely to be interested in the same products if they visit the same accounts, watch videos, read articles on related subjects, and reside in the same area.

Nevertheless, the use of big data in marketing is sometimes the subject of scandals. For instance, Netflix was accused of racism in 2018 for displaying various movie and series posters to users according to their gender and ethnicity.

Media simultaneously measures audiences with the aid of Big Data analytics. In this instance, editorial policy may even be impacted by big data. It

For example, the Huffington Post uses a system that shows real-time statistics on visits, comments, and other user actions, and also prepares analytical report

III . Statement of the problem:

Social media platforms' explosive growth has produced a deluge of big data, which is defined by its great volume, velocity, and variety. Even while there is a great deal of promise for meaningful insights to be derived from this data, efficiently analyzing such complex and unstructured data is still quite difficult. When it comes to handling the dynamic, real-time nature of social media data, current analytical tools and frameworks frequently fall short.

Objective of the Study:
This study paper's main goal is to investigate the approaches, tools, and real-world applications of using Big Data analytics

in relation to social media platforms. Our particular objectives are to:

Discover the advantages and drawbacks of utilizing big data analytics to glean insights from social media data. Talk about the technologies, methods, and tools used to gather, store, process, and analyze vast amounts of social media data.

To illustrate how big data techniques are used in social media, look at various analytical methods such as sentiment analysis, trend detection, and user behavior modeling.

Evaluate actual case studies and applications where businesses have effectively used big data analytics for public opinion research, brand monitoring, social media marketing, and crisis management.

Draw attention to upcoming developments and issues, such as data privacy, moral dilemmas, and the use of AI and big data in social media analysis.

IV . Significance of the Study:

Big data analytics has grown more and more important in recent years for understanding social media trends, behaviors, and patterns. The purpose of this study is to examine different methods, tools, and approaches for evaluating vast amounts of data produced by social media sites. Organizations, scholars, and marketers can learn a great deal about consumer preferences, brand perception, public opinion, and new trends by doing this. Strategic decision-making, targeted marketing, and realtime monitoring are made possible by the capacity to evaluate social media big data. It enables companies to manage their online reputation, maximize interaction tactics, and react swiftly to consumer input. Furthermore, advanced analytical techniques like sentiment analysis, trend detection, and user behavior prediction might reveal trends that conventional analysis may not identify. Applying big data analytics to social media also has the important benefit of automating data-driven insights, which can increase accuracy while saving time and money. This study highlights the significance of developing scalable, safe, and effective analytical models in order to fully utilize social media data, given that data is still growing at an unprecedented rate.

V . Literature Review:

Big data analytics is an essential tool for businesses looking to get insights from online user interactions because of the enormous amounts of unstructured data generated by the exponential expansion of social media platforms. Using a variety of academic sources, this review paper attempts to provide an overview of big data analytics' methods, uses, difficulties, and potential future directions in relation to social media.

VI . Overview of Big Data Analysis in Social Media:

The term "big data" describes datasets that are too big or complicated for conventional data processing technologies to handle. Social media, which creates content constantly on sites like Facebook, Instagram, LinkedIn, and Twitter, makes an important contribution to big data. Businesses and government agencies alike benefit greatly from realtime monitoring, sentiment analysis, and trend forecasting made possible by the integration of big data with social media analytics.

VII . Social Media Analytics Tools:

These tools track and evaluate the performance of social media content. Some examples include:

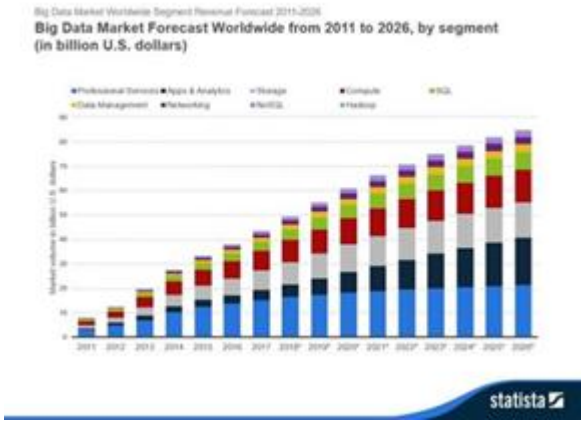
- **Sprout Social:** An application that facilitates social media management, including influencer marketing management, performance monitoring, and post scheduling and publication. Several social networks, such as Facebook, Instagram, and Twitter, are supported.
- **Hootsuite**: A tool for content planning, expanding reach, and monitoring performance on all of the main social media networks.
- **Keyhole:** An application that offers real-time social media research for Facebook, YouTube, LinkedIn, Instagram, TikTok, and X. It can watch rival behavior, identify popular hashtags as well, and keep an eye on what fans are saying about your company.
- **Mention:** A tool that lets you keep an eye on, listen to, and evaluate your audience interactions and postings.
- **To put it briefly**: A tool that assists brands in understanding their online performance and contrasting it with that of other companies**.**

VIII . Future of Big Data Analytics:

Emerging technologies like retrieval-augmented generation (RAG) and generative artificial intelligence (GenAI) have the potential to revolutionize large data analytics. GenAI automates content generation and creates synthetic datasets, pushing the limits of conventional data analysis. This invention creates new opportunities for data visualization and predictive analytics.

By incorporating real-time data retrieval, RAG improves AI models and guarantees precise and contextually relevant insights. Because RAG is dynamic, integrating it into data systems requires sophisticated knowledge of data pipeline architecture.

Real-time insights, automated decision-making, data governance, cloud scalability, data variety management, democratization, no-code solutions, microservices, data markets, and the data mesh idea are characteristics of the future of big data analytics. Regardless of size or budget, adopting these trends will enable firms to fully utilize their data. For those who adjust and use big data analytics in creative ways, the future is bright.



IX . Applications Big Data Analysis in Social Media:

Big data analytics have significantly altered how institutions and companies use social media to glean insightful information. Big data tools and techniques facilitate better informed, real-time decision-making because of the massive volume and velocity of data generated across platforms such as Facebook, Instagram, LinkedIn, and Twitter.

- Monitoring public opinion and sentiment analysis Sentiment analysis, which processes user-generated content to comprehend public sentiments, attitudes, and opinions, is one of the most popular applications. Businesses use it extensively to evaluate consumer happiness, and governments use it to determine how the public reacts to events or policy.

- Analysis of Consumer Behavior and Tailored Marketing Personalized marketing techniques are made possible by big data tools, which assist firms in understanding consumer preferences and behaviors.

- Monitoring Events in Real Time and Crisis Management Monitoring catastrophes like pandemics, natural disasters, and political upheaval in real time is made possible via social media analytics. Organizations may react to crises swiftly and efficiently by examining trending themes, hashtags, and tracked data.

- Brand Tracking and Competitive Analysis Analytics platforms are used by brands to track competition activity, customer feedback, and reputation. Businesses can enhance their goods, services, and communication tactics by using insights from user evaluations, mentions, and online discussions.

X . Literature Review Methodologies Used for Big Data Research:

Research questions

This SLR was carried out according with the protocol outlined by Kitchenham and Charters. Therefore, we established the research questions (RQ) that will guide the entire review process in the initial phase.
Finding a balance between the scope and depth of our examination is essential when we formulate the research questions that will direct our SLR. The following three research questions have been selected after much deliberation in order to preserve a targeted and significant scope for our review:

- RQ1: Which fields of applications for big data analytics are most prevalent, and how have they changed over time?

- RQ2: what are the major challenges and limitations that researchers have encountered in Big Data analysis, and how have they been addressed?
- RQ3: what are the emerging research trends and directions in Big Data that will likely shape the field in the next 5 to 10 years? Search strategy SLR begins by looking for relevant studies related to our research questions. To do this, we find appropriate search terms using the method outlined by Kitchenham and Charters, which suggests to consider three aspects: Population (P), Interventions (I), and Outcomes (O). We identified the following relevant search terms for each aspect in our review:
- Population: Big Data, real-time data analytics, large datasets.
- Intervention: methodologies, techniques, domains, architectures, solutions.
- Results: SLR, Systematic Literature Review, emerging technologies, research trends, and future directions. The following is how the search string was put together: P stands for population terms, I for intervention terms, and O for outcome terms. These are all related by the boolean operators AND and OR. (P1 OR P2... OR Pn) AND (I1 OR I2... OR In) AND (O1 OR O2... OR On). The exemplar form of a search string could look like this:

AND ("methodologies," "techniques,"
"domains," "architectures," or
"solutions") AND ("research trends," "future directions,"
"emerging technologies," "challenges," "SLR," or "Systematic Literature Review" AND "big data," "real-time data analytics,"
or
"large datasets."

Since we need to find and study primary studies related to our research questions, the selection of appropriate digital libraries/search engines to search for the articles needed is essential. For this reason, it has been decided to use the following state-of-the-art sources:

- Scopus: a multidisciplinary database that covers a broad range of research fields.
- IEEE Xplore: an invaluable resource for technology and engineering-related SLR.
- ACM Digital Library: a comprehensive collection of relevant articles, conference papers, and journals focused on computer science and information technology.
- SpringerLink: an extensive collection of academic articles in the fields that align closely with our research interests.
- Google Scholar: a freely accessible web search engine that indexes scholarly literature across various disciplines.

We aim to ensure a comprehensive and focused literature search by utilizing these sources, thereby facilitating a thorough and methodical research.

XI . Inclusion/Exclusion criteria

In this stage of the SLR, we need to make an accurate selection of the studies extracted. To do this, we must define some rigorous inclusion/exclusion criteria, to decide which studies are going to be useful for our purpose. To achieve this, studies were excluded based on the following criteria:

- Studies published before the 15-year time frame
- Studies in languages other than English
- Exclude non-academic sources, including blogs, news articles, marketing materials, and reports from non-academic organizations
- Studies that are only marginally related to Big Data or the specific topics within our research questions.

In conclusion, all those studies that are not cut off by the exclusion criteria above are to be considered as included. They are called “Primary Studies” (PS).

XII . Study quality assessment

Kitchenham and Charters stresses the necessity of assessing the quality of primary studies to reduce bias and enhance the validity of the evaluation process. In our research, we employ a study quality assessment to make sure that we have only the most relevant results for our research.
To achieve this, we formulated a five question study quality questionnaire, which serves as the foundation for assessing the quality of the primary studies:

- QA1: has the primary study established a well-defined research objective?
- QA2: did the primary study comprehensively describe its research methods and data sources?
- QA3: has the technique or approach undergone a trustworthy validation?
- QA4: has the primary study effectively identified and discussed the significant challenges and limitations encountered in Big Data analysis?
- QA5: are the findings, research trends, and directions clearly presented and directly connected to the study’s objectives or goals?

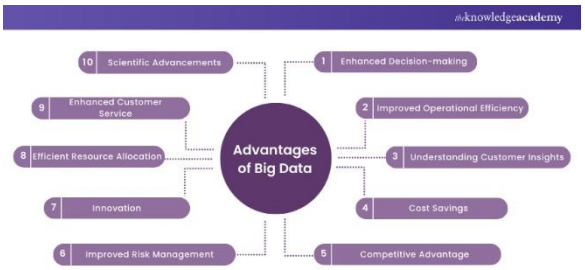
Hence, we applied the formulated questionnaire to the included PSs to assess their quality. The output of this SLR stage will be discussed in Section 4.

International Journal of Scientific Research & Engineering Trends
Volume 11, Issue3 , May-jun-2025, ISSN (Online): 2395-566X

XIII . Data extraction

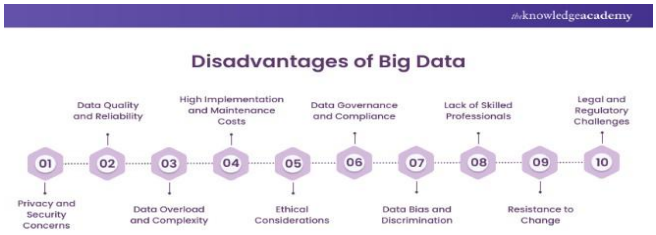
The data extraction process entails gathering relevant information from the chosen primary studies to address the research questions. To facilitate this process, we have created a dedicated data extraction form, as shown in Table1. As suggested in Kitchenham and Charters, we used the test-retest process to check the consistency and accuracy of the extracted data with respect to the original sources. After finishing the data extraction for all the selected studies, we randomly selected 3 primary studies and performed a second extraction of the data. No inconsistencies were detected .

Benefits of Big Data



- Enhanced Decision-making: Big Data provides organisations with access to a vast amount of information from various sources, enabling them to make data-driven decisions. By analyzing trends, patterns, and correlations within the data, businesses can seek valuable insights that guide their strategies.
- Improved Operational Efficiency: Big Data analytics helps organisations optimise their operations by identifying inefficiencies, bottlenecks, and areas for improvement. By streamlining processes and workflows, businesses can enhance productivity, reduce costs, and deliver better results.
- Personalisation and Customer Insights: Big Data enables organisations to understand their customers at a deeper level. By analysing customer data, businesses can identify preferences, behaviours, and needs, allowing them to personalise products, services, and marketing campaigns to enhance the customer experience.
- Cost Savings: Big Data analytics can uncover cost-saving opportunities for organisations. By analysing data related to operations, supply chains, and resource allocation, businesses can identify areas where costs can be minimised, such as optimising inventory management or reducing energy

Limitations of Big Data



- Privacy and Security Concerns: The accumulation and examination of extensive data sets raise privacy and security concerns. Organisations must implement robust data protection measures to secure sensitive information and comply with relevant regulations to ensure individuals' privacy is maintained.
- Data Quality and Reliability: Big Data poses challenges related to data quality and reliability. The volume, as well as a variety of data sources can lead to issues such as incomplete or inconsistent data, which may affect the accuracy and validity of analysis.
- Data Overload and Complexity: Dealing with massive volumes of data can be overwhelming. Managing and processing large datasets requires advanced infrastructure, tools, and expertise. The complexity of Big Data can make it challenging to extract meaningful insights efficiently.
- High Implementation and Maintenance Costs: Implementing Big Data solutions can be costly for organisations. Investments in infrastructure, technology, and skilled personnel are required, which may pose challenges, especially for small and mediumsized businesses.

XIV . Discussion:

Data is changing our world – and fast. There is no denying this fact. What we buy, what we eat, how we communicate, how we are governed, how we live are all affected by the use of data. However, it should be noted that using data in day-to-day life is not a new concept. Ancient civilizations designed their calendars by predicting planetary movements based on data from prior recordings. More recently the advancement in digital and telecommunication technologies has led to an explosion of the amount of data available. The world has never been so interconnected. Each person who uses the internet, the telephone, or credit cards leaves a trail of information which can be used by organizations to predict their behavior and adapt accordingly. The same is true of anyone who pays a utility bill, files a tax return or is registered with government in some way (electoral registration office, health services, etc.). Big data is also being used in government initiatives as well as in all areas of research

including health, economics, manufacturing, defense and security and education.

XV . Conclusion:

In conclusion, big data analysis plays a pivotal role in enhancing social media marketing strategies by providing businesses with detailed insights into user behavior, preferences, and engagement patterns. The ability to understand and anticipate customer needs enables companies to create more personalized and effective marketing campaigns. While the benefits of big data, such as improved targeting, optimized user experience, and better marketing efficiency, are clear, challenges related to data quality and privacy protection must be carefully managed. As technological advancements, including the rise of Generative AI and real-time data retrieval systems, continue to reshape the landscape, businesses must adapt to these changes to stay competitive. The future of big data in social media marketing is promising, offering unprecedented opportunities for growth, but it requires strategic planning, ethical data handling, and continuous innovation to fully harness its potential.

References:

1. "Big Data and Social Media Analytics" by Vikas Dhawan and Nadir Zanini
2. "Impact of Big Data and Social Media on society" by Savita Kumari
3. "Big Data in Social Media Environment" by Dr.S. Matilda
4. "Big Data Analytics in Social Media: A Triple T" by Md. Safiur Rehman and Hassan Reza
5. Big Data and Social Media to Improve the Quality of Higher Education" by Dr. Savita Kumari
6. Incorporating Big Data Tools for Social Media Analytics in a Business Analytics Course" by Amir H. Zadeh, Hamed M. Zolbanin and Ramesh Sharda
7. Social Media Big Data: The Good, The Bad, and the Ugly (un) truths" by Dinesh Gunasekeran
8. Social Media Big Data: The Good The Bad The Ugly (Un)truths " by Dinesh Visva Gunasekeran
9. "Big Data Analytics: Approaches and Applications in Online Marketing" by Shikha Gupta
10. "Impact of Social Media Networks Big Data Analysis for High-Level Bisiness" by
11. T. Veeramani, P. Srinuvasarao, B Rama Krishna R. ThilagaVathy