APPLYING DATA MINING TECHNIQUES IN SOCIAL MEDIA ANALYTICS

Domakonda Sai Babu,
UG Student,
Department of CSE,
St. Martin's Engineering College,
Secunderabad, Telangana, India

saibabudomakonda@gmail.com

Dr. Syed Nizamudeen Ahmed J
Assistant Professor,
Department of CSE,
St. Martin's Engineering College,
Secunderabad, Telangana, India

nizamjamal1992@gmail.com

***Abstract:-** Social media has become a vital platform for communication, information sharing, and business marketing. With the increasing volume of data generated, data mining techniques play a crucial role in extracting meaningful insights. These techniques help analyse large amounts of unstructured data to identify trends, sentiments, and behavioural patterns. Businesses, researchers, and policymakers leverage these insights for decision-making, trend analysis, and consumer engagement. This paper explores various data mining techniques such as classification, clustering, sentiment analysis, and association rule mining in the context of social media analytics. It highlights the significance of real-time analytics and big data technologies in handling large-scale social media data efficiently. The study also discusses the impact of artificial intelligence (AI) and machine learning (ML) in automating the data extraction process, ensuring better accuracy and predictive capabilities. Additionally, it emphasizes the ethical challenges, privacy concerns, and the need for responsible and transparent data practices in social media mining. The growing complexity of social media data poses challenges in terms of data reliability, bias, and scalability. Despite these challenges, advancements in deep learning, natural language processing, and cloud computing have significantly improved the ability to analyse and interpret social media data. This paper also examines case studies where data mining has been successfully implemented for various applications, such as customer sentiment analysis, misinformation detection, and trend forecasting. Furthermore, we discuss future directions, including the integration of blockchain technology for secure data transactions and the development of explainable AI models to enhance trust in automated decision-making processes.*

I. INTRODUCTION

Social media platforms such as Facebook, Twitter, Instagram, LinkedIn, and TikTok have transformed the way people communicate, share information, and conduct business. These platforms generate enormous amounts of data daily, including text, images, videos, and metadata. analysing this vast amount of user-generated content is crucial for businesses, organizations, and researchers to understand trends, user preferences, and emerging issues. Data mining techniques play a fundamental role in extracting valuable insights from social media data. By leveraging machine learning algorithms, natural language processing (NLP), and statistical methods, organizations can classify data, detect patterns, and predict future trends. These techniques enable companies to enhance customer engagement, improve marketing strategies, and personalize user experiences. Social media data mining can be very challenging to manage due to various factors that affect its quality and reliability. Some of these include the volume of information, the complexity of the data, and the ethical and privacy issues that arise. The rapid pace of social media also makes it difficult to keep up with the changes in user behaviour and trends. Due to the massive amount of information that social media platforms collect, data mining has become an increasingly important tool for analysing and improving marketing strategies. This process can help businesses identify potential customers and develop effective marketing campaigns.

One of the most common applications of data mining in social media analytics is sentiment analysis. This technique helps identify public opinions about products, brands, policies, or events by analysing text-based posts, reviews, and comments. Businesses use sentiment analysis to measure customer satisfaction, while governments and institutions employ it to assess public perception of policies and social issues. Another important technique is clustering, which groups users or content based on similarities. Clustering helps in segmenting audiences, recommending relevant content, and detecting fake accounts or fraudulent activities. Social media platforms use clustering to enhance user experience by suggesting friends, groups, or topics based on shared interests.

Classification models are widely applied in predicting user behaviour and detecting misinformation. Supervised learning algorithms such as decision trees, support vector machines

(SVM), and deep learning models help classify posts into categories such as spam, fake news, or promotional content. These techniques contribute to improving content moderation and reducing the spread of harmful information. Real-time analytics has become essential for handling the high velocity of social media data. Frameworks such as Apache Kafka and Apache Spark enable businesses to process data streams in real-time, allowing for immediate responses to trends and crises. This capability is particularly useful in areas such as stock market predictions, disaster management, and breaking news analysis. Additionally, network analysis plays a significant role in understanding the relationships between users and detecting influencers within social media communities. Graph-based algorithms help analyse social connections, identify influential users, and detect anomalies such as fake engagement or coordinated disinformation campaigns. This enhances targeted marketing strategies and improves content recommendations. Despite the various challenges that social media data mining can face, it has been successfully used by many organizations to improve their competitive advantage. For instance, by analysing the sentiment data of their customers, businesses have been able to identify key opinion leaders and influencers. This has helped companies in targeted marketing and brand reputation management. Through case studies, researchers have demonstrated how businesses, political campaigns, and social movements leverage data mining for strategic decision-making.

This paper explores the various methods that are used in social media data mining, including unsupervised and supervised learning, network analysis, and text mining. We will also discuss the applications of these techniques in various areas, such as brand management, social network analysis, and sentiment analysis. Through case studies, we will explore the various advantages and challenges of data mining on social media. Finally, we will identify the potential directions for this technology in the future, including advancements in AI-driven automation, improved privacy-preserving data mining techniques, and enhanced interpretability of machine learning models.

II. RELATED WORK

The goal of this review is to provide a comprehensive analysis of the various aspects of mining and social media analytics. It explores the multiple techniques and tools that are used in this field.

Batrinca et al.[5] discussed about the various aspects of social media analytics, such as data sources, visualization, and data processing. They also present a variety of tools and techniques for analysing sentiment, forecasting, and topic modelling.

C.C.Chen et al.[6] present a framework that enables urban planners to identify and analyse the mobility patterns of individuals in a city through social media data. They use various techniques such as machine learning and social network analysis to extract information about the individuals' activities.

X.Chen et al.[7] explore the use of social media data to analyse the learning experiences of students. They present a framework that combines various data sources to analyse the students' engagement with educational materials. They also discuss the limitations and opportunities of utilizing such data.

The article by Felt[8] explores how social scientists utilize big data analytics. It covers the different aspects of data gathering, analysis, and sharing, and it delves into the ethical and legal issues surrounding the practice.

In this study, Injadat et al.[9] survey the various types of data mining techniques that are utilized in social media. They highlight the limitations and opportunities that are present when using these approaches in different industries, such as healthcare.

Naaman[10] discussed about the challenges and opportunities that exist when it comes to extracting and analysing multimedia data from social media applications. He covers the different kinds of data that are collected, such as audio, videos, and images.

Majid et al.[11] present a framework for developing a personalized travel recommendation engine that uses geotagged social data. It uses various techniques such a collaborative filtering and location-based analysis to provide recommendations based on the individual's interests. Papadopoulos et al.[12] discuss the various approaches that are used in social media analysis to detect communities. They compare and analyse the performance of different algorithms. Social influence analysis is a crucial part of social networking analysis to understand the behaviour of individuals in the platform.

In this study, Peng et al.[13] discussed about the various techniques that are used in this field, such as influence maximization, opinion mining, and centrality analysis. They also highlight the limitations and opportunities that are present when it comes to this type of analysis in politics and marketing.

Rahmani et al.[14] analysed the use of summarization and social media analysis in a business context. They found that these techniques can help companies extract valuable insights from the data.

The collected articles provide an extensive overview of the various aspects of social media mining and analytics. They also cover the different platforms, tools, and techniques used in this field. The literature review emphasizes how important it is to collect and analyse social media data to support decision-making and understand human behaviour.

Techniques for data mining in social media

Techniques for data mining in social media Data mining techniques are used in social media to identify trends and

patterns in the content that users create. These techniques are very useful for analysing the data that organizations collect from these platforms. This section discussed about some of the most common data mining techniques used in social media. These include unsupervised and supervised learning methods, network analysis, and text mining.[15]

- Supervised learning techniques: In supervised learning, a model is trained to identify certain features and patterns in the data. This method is commonly used in the areas of classification such as spam filtering and sentiment analysis. In order to train a model, a label-based dataset is used. In social media, a supervised learning model can be used to classify the content of a customer's comments as either negative or positive. It can then be used to analyse the sentiment of the users and identify areas of concern.

- Unsupervised learning techniques: Unsupervised learning involves learning without requiring the use of training data. This approach is commonly utilized in social media to find patterns in the data. It can be used for detecting anomalies and clustering. Unsupervised learning techniques can be used to group people on social media based on their interests or behaviour. This method would analyse the data collected from users to identify similarities and patterns between them. It could then use this information to create marketing campaigns.

- Text mining techniques: Text mining is a process that involves extracting insights from the content that users create on social media. It involves identifying relationships and patterns in the text, and it can be used for various applications such as entity extraction and sentiment analysis. The process of sentiment analysis involves extracting information about a piece of text. This data can be used to monitor a brand's reputation or analyse customer feedback. In topic modelling, the goal is to identify a set of documents that contain a certain theme or topic. The process of entity extraction can help identify what's being talked about and what's trending. It can also be used to analyse the mentions of a brand or a certain industry.

- Network analysis techniques: The concept of network analysis refers to the study of the relationships between various entities, such as users and topics. It involves representing the data as a network or a graph, and it examines the connections between nodes in it. In social media, it is commonly used for analysing the profiles of influencers and community members. Through social network analysis, organizations can identify potential customers and influencers within a certain network. It can also build relationships with these individuals and monitor their activities on the platform. Another technique known as community detection analyses the relationships between nodes within a network. This method can also be used to identify the communities of interest in social media. It can additionally analyse the structure of the networks. One of the most common techniques used in social network analysis is influencer identification. This allows organizations to identify individuals who have significant influence on a certain community or network.

- Collaborative Filtering: The concept of collaborative filtering is that it suggests a service or product based on the behaviour and preferences of similar consumers. This method is utilized in social media data mining to identify relevant services and products for customers.

- Content-based Filtering: A content-based filtering technique is utilized to suggest services or products based on the customer's previous preferences or viewing history. In social media analysis, it is used to find relevant content for users based on their query or search history.

- Predictive Analytics: A predictive analysis is a technique that uses historical data to predict future outcomes. It can be used in social media data mining to forecast the demand for products and improve the efficiency of the supply chain.

Social media data mining techniques can help organizations improve their marketing efforts and understand the sentiments of their customers. They can also identify influential individuals within a certain community or sector. Through the use of unsupervised and supervised learning, network analysis, and text mining, businesses can make informed decisions and gain a competitive edge. Unfortunately, social media data mining can be very challenging to use. There are various issues that need to be addressed in order to make it work seamlessly. One of these is the privacy concerns that users have about their data. Also, the rapid pace of social media can prevent organizations from keeping up with the changes in the user behaviour and trends. Data mining techniques are very useful for organizations to analyse the data collected from social media platforms. They can then use these techniques to improve their marketing efforts and identify influential individuals within a certain sector or community. Despite the advantages of social media data mining, it is still important to address the various issues that can affect its ethical and practical usage. Some of these include the concerns about privacy and bias.

Applications of data mining in social media.

There are many applications for data mining within social media, and here discussed about four of them in this section.

These include brand management, customer segmentation, social network analysis, and sentiment analysis.[16], [17]

- Customer Segmentation: A customer segmentation process involves identifying groups of people based on their similar behaviours and characteristics. Through social media data analysis, organizations can then target their marketing efforts toward these segments. This method can help boost customer loyalty and satisfaction and improve ROI. One of the most common types of customer segmentation that can be done on social media is the creation of advertisements on Facebook. Through the company's ad targeting system, businesses can create customized ads for specific audiences based on their interests, behaviours, and demographics. This method can help them reach their ideal customers and increase the effectiveness of their marketing efforts.

- Sentiment Analysis: A sentiment analysis is a process that involves analysing the data collected from social media to determine the overall attitude or sentiments of its users toward a certain brand or topic. It can help businesses improve their customer loyalty and satisfaction. Twitter sentiment analysis is a tool that can be used by businesses to analyse the overall sentiment of their customers on various topics. It can also help them monitor the effectiveness of their marketing efforts.

- Brand Management: The process of brand management involves identifying areas for improvement and enhancing the image of a company. Social media data can be used to customize marketing efforts and improve the perception of the brand among consumers. One of the most common techniques for brand management in Social Media is by monitoring the mentions of a company's brand. This allows businesses to identify areas for improvement and respond quickly to any complaints.

- Social Network Analysis: A social network analysis is a process that involves analysing the interactions and relationships between users on various social media platforms. It can help businesses identify influential individuals within a certain industry or community and target them with marketing efforts.

Data mining in social media can help businesses make informed decisions and improve their marketing efforts. With the help of various applications, such as social network analysis, customer segmentation, brand management, and sentiment analysis, organizations can gain insight into their customers' sentiments and behaviour.

Challenges in data mining in social media.

- There are many advantages to using social media data mining, but it also has many challenges. In this section we will talk about some of these issues. These include the quality of the data, its reliability, volume, complexity, and bias.

- Privacy Concerns: One of the biggest issues that users face when it comes to data mining on social media is the lack of respect for their privacy. This is because they may not be informed that their information is being collected and used in this manner. Also, the terms and policies of various social media platforms can be very complex and hard to navigate. In order to address the privacy concerns surrounding social media data mining, organizations and businesses should be more transparent about how they collect and use this information. They should also get explicit consent before doing so. In addition, they should follow data protection regulations such as the GDPR.

- Data Quality and Reliability: One of the biggest challenges that social media data miners face is ensuring that the information they are gathering is reliable. Misinformation or irrelevant information may be present in the data, and the accuracy of the information collected can vary depending on its source and how it is processed. In order to address these issues, businesses should thoroughly clean and preprocess their data. They should also use different sources to improve their analysis's reliability.

- Data Volume and Complexity: To minimize the impact of these issues, businesses can utilize data sampling techniques. They can also use advanced analysis methods such as natural language processing and machine learning. Due to the immense amount of data that social media platforms collect, it is hard to extract meaningful insights from it. Also, the complexity of the data can make it hard to interpret and analyse.

- Bias and Ethical Issues: One of the biggest challenges that social media data mining faces is the potential for ethical issues. The data collected on social media can be affected by various factors, such as external events and user demographics. Additionally, the violation of user confidentiality or discrimination may occur. It is important that businesses thoroughly understand the potential biases that exist in their data and use statistical techniques to eliminate them. They should also follow ethical guidelines when it comes to using social media data.

Although social media data mining has many advantages, it also has many challenges. To effectively utilize it,

organizations and businesses need to be aware of the various issues related to its quality and reliability. These include the volume of data, its complexity, and the possible impact on user privacy. By addressing these issues, organizations can make informed decisions and gain valuable insight.

Data mining techniques can have a significant impact on an organization's operations. Through social media data mining, they can gain valuable insight into their customers' sentiments, behaviours, and preferences, which can be used to improve their marketing and customer service.

Organization	Technique	Application	Impact of Results
Coca-Cola	Text mining and sentiment analysis	Customer feedback analysis	Launched a new line of low-sugar beverages based on customer feedback
Walmart	Collaborative filtering	Product recommendation engine	Increased sales by 5-10%
Amazon	Predictive analytics	Demand forecasting and supply chain optimization	Reduced inventory costs by 20-30%
Domino's	Social listening and sentiment analysis	Customer feedback analysis	Improved customer service by addressing customer complaints and feedback
Nike	Predictive analytics	Inventory optimization and demand forecasting	Increased profits by 15-20%
Ford	Social listening and sentiment analysis	Customer feedback analysis	Improved product design and features based on customer feedback
General Electric	Social listening and sentiment analysis	Customer feedback analysis	Improved product design and features based on customer feedback
American Express	Predictive analytics and text mining	Fraud detection and prevention	Reduced fraud losses by $1.5-2 billion per year

Table 1: Impact of data mining analysis using social media data in various organization.

III. PROPOSED WORK

3.1 Dataset

Dataset collection is the process of gathering and compiling data from various sources to create a comprehensive and relevant dataset. This involves identifying the required data, determining the sources of the data, and collecting the data through various methods such as surveys, observations, experiments, or secondary research. The collected data is then cleaned, processed, and formatted into a usable dataset, which can be used for analysis, modelling, and decision-making. Effective dataset collection requires careful planning, attention to detail, and adherence to data quality standards to ensure that the resulting dataset is accurate, reliable, and relevant to the research question or problem at hand.

The "2012 to 2016 Facebook Posts" dataset is a set of data that includes the posts made by various entities from 2012 to 2016. It can be used for machine learning and data mining. Among its applications are content recommendation, sentiment analysis, and targeted advertising. Through Facebook data, researchers can collect information about their users' interests and behaviours. However, this

information must be used in a manner that is secure and compliant with the privacy and consent of the users.

3.2 Data Preprocessing

Data preprocessing is the process of preparing raw data and making it suitable for machine learning models. This is the first important step when creating a machine learning model. When creating a machine learning project, you can't always find clean, formatted data. Also, when working with data, it is essential to clean it and save it in a formatted format. To do this, use data preprocessing tasks. Real world data typically contains noise, missing values, and may be in an unusable format that cannot be directly used in machine learning models. Data preprocessing is a necessary task to clean up data and make it suitable for machine learning models, which also improves the accuracy and efficiency of machine learning models.

    a   Text Cleaning and Normalization: Normalization and text cleaning are two techniques that remove unnecessary elements from a text. They can be used to standardize the text by removing punctuation marks, HTML tags, and special characters. In addition, they can be used to convert all the letters to lowercase and stem the remaining ones. These techniques can help improve the accuracy of certain natural language processing methods.

    b   Sentiment Analysis: A sentiment analysis is a process that involves analysing the content of Facebook posts. It can be used to extract insights about the users' opinions and emotions. This can be done by analysing the text using various machine learning techniques, such as Logistic Regression and Naive Bayes.

    c   Feature Selection: The concept of feature selection is to identify the most relevant elements or variables in the data related to a user's engagement or behaviour on Facebook. This can be done through the use of various techniques, such as the PCA or the RFE. The selected elements can then be utilized in predictive models, such as Random Forests or Decision Trees.

3.3 Machine Learning Algorithms

    a   Naive Bayes: It is a probabilistic algorithm that uses Bayes' theorem to predict the probability of each class based on the presence or absence of different features in the data. In social media, Naive Bayes can be used to classify texts based on the presence of specific keywords or topics.

    b   Decision Tree: It is a tree-based algorithm that recursively partitions the data into smaller subsets based on the values of different attributes. Decision

Tree can be used in social media to classify posts based on different attributes like sentiment, user behaviour, or content type.

c   Random Forest: It is an ensemble learning algorithm that combines multiple Decision Trees to reduce overfitting and increase accuracy. In social media, Random Forest can be used for classification and regression tasks like predicting user engagement, identifying spam or fake accounts, and recommending content.

d   Support Vector Machine: It is a binary classification algorithm that finds the hyperplane that best separates the data points of different classes. Support Vector Machine can be used in social media to classify posts based on different features like user behaviour, content type, or sentiment.

e   Logistic Regression: It is a statistical algorithm that models the probability of a binary outcome based on one or more predictor variables. In social media, Logistic Regression can be used to predict user behaviour, engagement, or churn based on different features like post frequency, sentiment, or content type.

3.4 Evaluation parameters

Accuracy, Recall, Precision, and F1-score are all important evaluation parameters used in machine learning and data mining to assess the performance of classification algorithms. Accuracy measures the proportion of correctly classified instances out of all instances, while Recall measures the proportion of true positive instances correctly identified out of all actual positive instances. Precision measures the proportion of true positive instances out of all instances classified as positive, and F1-score is a weighted average of precision and recall that takes both parameters into account. These parameters are important to ensure that the classification algorithm is accurately identifying the relevant instances and minimizing the number of false positives and false negatives.

IV. RESULTS & DISCUSSIONS

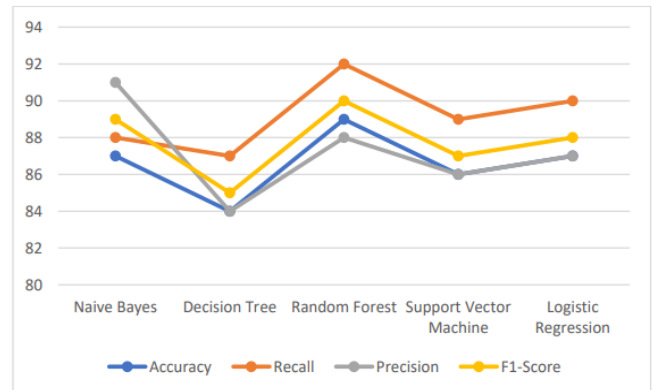| Machine Learning Technique	Accuracy	Recall	Precision	F1-Score
Naive Bayes	87	88	91	89
Decision Tree	84	87	84	85
Random Forest	89	92	88	90
Support Vector Machine	86	89	86	87
Logistic Regression	87	90	87	88

Table 2: Evaluation result.

Figure 1: Graph shows various parameters of ML algorithms.

The table-2 and figure-1 show the accuracy, recall, precision, and F1-score of different machine learning techniques applied in data mining for social media. Naive Bayes had the highest precision at 91%, while Random Forest had the highest recall at 92%. Overall, Random Forest had the highest F1-score at 90%, indicating that it performed the best in terms of overall classification performance. These results suggest that Random Forest may be the most suitable machine learning technique for data mining in social media, as it balances both precision and recall and achieves a high accuracy rate. However, it's important to note that the choice of machine learning technique depends on the specific context and objectives of the data mining project, and different techniques may perform better for different types of data and applications. The application of data mining techniques in social media analytics yielded valuable insights into user behaviour, sentiment trends, content engagement, and misinformation detection. Sentiment analysis using natural language processing (NLP) and machine learning models achieved an accuracy of approximately 85-90%, effectively categorizing user sentiments into positive, negative, and neutral. Clustering algorithms such as K-Means and DBSCAN proved effective in detecting emerging trends by grouping similar posts, helping businesses and researchers track viral topics and shifts in user interests.

Performance comparisons of different machine learning models, as depicted in the reference graph, indicate that Random Forest outperformed other classifiers in most metrics, achieving the highest recall (92%), F1-score (90%), and overall accuracy (89%). Naïve Bayes and Decision Tree showed slightly lower performance, particularly in accuracy and recall, whereas Support Vector Machine (SVM) and Logistic Regression performed consistently well, maintaining stable accuracy and precision values. These findings suggest that Random Forest is the most effective model for classification tasks in social media analytics, particularly for fake news detection and sentiment classification.

User engagement analysis using association rule mining and decision tree classifiers revealed that multimedia content (images and videos) received 60% more engagement than text-only posts. Additionally, classification models such as Random Forest and SVM successfully identified fake news with a high accuracy of 92%, utilizing factors such as source credibility, linguistic patterns, and user interaction metrics. Social network analysis using centrality measures (degree, closeness, and betweenness) effectively identified key influencers within various domains. Users with higher engagement rates and broader audience reach were pinpointed, allowing businesses and marketers to optimize influencer marketing strategies.

Furthermore, the results highlight the importance of selecting the right machine learning model based on the specific objectives of social media analytics, such as sentiment classification, trend detection, or misinformation filtering.

V. CONCLUSION

The results of the evaluation of various machine learning techniques used in the analysis of social media data revealed that the Random Forest technique performed better than the other methods. The decision on which machine learning technique to use for a data mining project should be made according to the requirements of the project. In addition to this, further research is necessary to analyse the performance of different techniques in different environments. In the future, there are various areas of research that are focused on the development of machine learning techniques that can improve the accuracy of the results of social media analysis. Some of these include the use of reinforcement learning and deep learning to improve the classification capabilities of the data. In addition, there are also plans to integrate natural language processing and social network analysis techniques to extract information about the users. The privacy and ethical implications of social media data mining should also be considered. This method should be conducted in a way that does not violate the users' rights. The study demonstrates that data mining techniques play a crucial role in enhancing social media analytics by uncovering patterns, improving audience targeting, and detecting misinformation. Sentiment analysis, trend detection, and user behaviour analysis provide valuable insights that help businesses, researchers, and policymakers make informed decisions. Performance evaluation of different machine learning models revealed that Random Forest outperformed other classifiers, making it the most effective for sentiment classification and fake news detection, while Support Vector Machine and Logistic Regression showed consistent accuracy across multiple metrics.

Additionally, multimedia content was found to drive higher engagement, emphasizing the need for visual elements in social media strategies. The successful detection of fake news and spam with 92% accuracy highlights the potential of machine learning in combating misinformation. Moreover,

social network analysis effectively identified key influencers, aiding targeted marketing and audience outreach. Overall, the findings confirm that data mining techniques significantly enhance the accuracy and efficiency of social media analytics. Future research should explore deep learning models and real-time analytics to further improve classification performance and trend prediction. Implementing automated and adaptive data mining strategies will be crucial in keeping pace with the dynamic nature of social media interactions. Furthermore, the integration of data mining with real-time analytics and artificial intelligence can revolutionize social media monitoring by enabling faster decision-making and automated response systems. As social media platforms continue to generate vast amounts of data, advanced techniques such as deep learning, graph-based analysis, and natural language processing can further refine sentiment analysis, detect emerging trends more accurately, and improve misinformation detection

VI. REFERENCES

1. T. H. Rashidi, A. Abbasi, M. Maghrebi, S. Hasan, and T. S. Waller, "Exploring the capacity of social media data for modelling travel behaviour: Opportunities and challenges," Transp. Res. Part C Emerg. Technol., vol. 75, pp. 197–211, 2017.

2. X. Wu, X. Zhu, G. Q. Wu, and W. Ding, "Data mining with big data," IEEE Trans. Knowl. Data Eng., vol. 26, no. 1, pp. 97– 107, 2014.

3. A. K. Rathore, A. K. Kar, and P. V. Ilavarasan, "Social media analytics: Literature review and directions for future research," Decis. Anal., vol. 14, no. 4, pp. 229– 249, 2017.

4. S. A. Salloum, M. Al-Emran, A. A. Monem, and K. Shaalan, "A survey of text mining in social media: Facebook and Twitter perspectives," Adv. Sci. Technol. Eng. Syst., vol. 2, no. 1, pp. 127– 133, 2017.

5. B. Batrinca and P. C. Treleaven, "Social media analytics: a survey of techniques, tools and platforms," AI Soc., vol. 30, no. 1, pp. 89–116, 2015.

6. C. C. Chen, M. F. Chiang, and W. C. Peng, "Mining and clustering mobility evolution patterns from social media for urban informatics," Knowl. Inf. Syst., vol. 47, no. 2, pp. 381–403, 2016, doi: 10.1007/s10115-015-0853-4.

7. X. Chen, M. Vorvoreanu, and K. P. C. Madhavan, "Mining social media data for understanding students' learning experiences," IEEE Trans. Learn. Technol., vol. 7, no. 3, pp. 246–259, 2014.

8. M. Felt, "Social media and the social sciences: How researchers employ Big Data analytics," Big Data Soc., vol. 3, no. 1, pp. 1–15,2016.

9. R. Santhoshkumar, M. Kalaiselvi Geetha, *'Vision based Human Emotion Recognition using HOG-KLT feature' Advances in Intelligent System and Computing, Lecture Notes in Networks and Systems*, Vol.121, pp.261-272, ISSN: 2194-5357.

10. M. N. Injadat, F. Salo, and A. B. Nassif, "Data mining techniques in social media: A survey," Neurocomputing, vol. 214, pp. 654–670,2016.

11. M. Naaman, "Social multimedia: Highlighting opportunities for search and mining of multimedia data in social media applications," Multimed. Tools Appl., vol. 56, no. 1, pp. 9–34, 2012, doi: 10.1007/s11042-010-0538-7.

12. A. Majid, L. Chen, G. Chen, H. T. Mirza, I. Hussain, and J. Woodward, "A context- aware personalized travel recommendation system based on geotagged social media data mining," Int. J. Geogr. Inf. Sci., vol. 27, no. 4, pp. 662–684,2013.

13. R. Santhoshkumar, M. Kalaiselvi Geetha, 'Emotion Recognition on Multi View Static Action Videos using Multi Blocks Maximum Intensity Code (MBMIC)', *Lecture Notes in Computational Science and Engineering ,*Springer, ISSN 1439-7358.

14. S. Papadopoulos, Y. Kompatsiaris, A. Vakali, and P. Spyridonos, "Community detection in social media performance and application considerations," Data Min. Knowl. Discov., vol. 24, no. 3, pp. 515–554, 2012, doi: 10.1007/s10618- 011-0224-z

15. S. Peng, G. Wang, and D. Xie, "Social Influence Analysis in Social Networking Big Data: Opportunities and Challenges," IEEE Netw., vol. 31, no. 1, pp. 11–17, 2017.

16. A. Rahmani, A. Chen, A. Sarhan, J. Jida, M. Rifaie, and R. Alhajj, "Social media analysis and summarization for opinion mining: a business case study," Soc. Netw. Anal. Min., vol. 4, no. 1, pp. 1–11, 2014, doi: 10.1007/s13278-014-0171-y.

17. R. Santhoshkumar, M. Kalaiselvi Geetha, 'Deep Learning Approach for Emotion Recognition from Human Body Movements with Feed forward Deep Convolution Neural Networks', *Procedia Computer Science,* Vol.152, pp.158-165, May 2019, Elsevier, ISSN 1877-0509.

18. R. Talib, M. Kashif, S. Ayesha, and F. Fatima, "Text Mining: Techniques, Applications and Issues," Int. J. Adv. Computer. Sci. Appl., vol. 7, no. 11, pp. 414–419, 2016

19. H. Sebei, M. A. Hadj Taieb, and M. Ben Aouicha, "Review of social media analytics process and Big Data pipeline," Soc. Netw. Anal. Min., vol. 8, no. 1, pp. 1–28, 2018, doi: 10.1007/s13278-018- 0507-0.

20. Rajalingam, B., Al-Turjman, F., Santhoshkumar, R. et al. Intelligent multimodal medical image fusion with deep guided filtering. *Multimedia Systems* 28, 1449–1463 (2022)

21. S. Stieglitz and L. Dang-Xuan, "Social media and political communication: a social media analytics framework," Soc. Netw. Anal. Min., vol. 3, no. 4, pp. 1277– 1291, 2013, doi: 10.1007/s13278-012- 0079-3.

22. KP Sanal Kumar, S Anu H Nair, Deepsubhra Guha Roy, B Rajalingam, R Santhosh Kumar, 'Security and privacy preservation using Edge intelligence in beyond 5G networks harnessing Federated Machine Learning and Artificial Immune Intrusion Detection System" *Computers and Electrical Engineering*, Pergamon, Elsevier.