Journal of
Mobile Computing, Communications &
Mobile Networks

ISSN: 2349-901X
Volume 11, Issue 1, 2024
DOI (Journal): 10.37591/JoMCCMN

https://journals.stmjournals.com/jomccmn/

Review JoMCCMN

Role and Importance of Machine Learning in Social Media

Vaishali Jha[1,*], Kapil Sharma[2], Ashok Kumar Shrivastava[2], Samta Jain Goyal[2]

Abstract

The widespread adoption of social media platforms has transformed the way individuals interact and communicate. Going beyond personal connections, social media has evolved into a potent tool for sharing information, shaping ideas, and fostering participation across various industries. Machine learning is pivotal in enhancing social media's impact. Social media generates vast data daily, and machine learning is essential for extracting insights. Sentiment analysis, a machine learning application, identifies emotions in social media content, aiding brand tracking and customer insights. Machine learning accurately predicts user behaviour on social platforms, enabling tailored content and personalized recommendations, improving user experiences. It is crucial in social media marketing to optimize ad campaigns by targeting specific demographics, enhancing engagement, and boosting ROI. In addressing social issues, machine learning detects hate speech, false information, and cyberbullying, creating safer online environments and supporting fact-checking efforts. In governance, it tracks public sentiment and informs policymaking. Challenges include privacy, biases, and ethics. As social media evolves, machine learning must adapt to new data sources.

Keywords: Machine learning, social media, demographics, proliferation, sentiment analysis

INTRODUCTION

Through the introduction of social media, communication has undergone a revolution and become an integral element of contemporary life. In the current digital environment, where vast amounts of data are constantly flowing, this study examines the crucial role played by machine learning. User experience improvement, sentiment analysis, content personalization, and other areas are just a few of the many ways that machine learning is changing social media. To better understand machine learning's crucial importance in the social media space, this study's main goal is to provide examples. Its uses, advantages, difficulties, and moral issues are all examined. This study expects to offer guidance for practitioners, decision-makers, and researchers interested in maximizing the potential of machine learning in social media by concentrating on this dynamic intersection. This study sheds light on how these two domains coalesce to impact society, marketing, governance, and information integrity. The findings can guide responsible and innovative use of technology in the digital age as social media's influence grows. The study is divided into sections that explore various facets of social media machine learning. The study starts by reviewing the literature to lay the groundwork for our ideas. The parts that follow examine sentiment analysis, predictive

***Author for Correspondence**
Vaishali Jha
E-mail: vaishalijha281532@gmail.com

[1]Student, Department of Computer Science & Engineering, Amity School of Engineering & Technology, Amity University, Gwalior, Madhya Pradesh, India
[2]Associate Professor, Department of Computer Science & Engineering, Amity School of Engineering & Technology, Amity University, Gwalior, Madhya Pradesh, India

Received Date: February 14, 2024
Accepted Date: February 16, 2024
Published Date: April 03, 2024

Citation: Vaishali Jha, Kapil Sharma, Ashok Kumar Shrivastava, Samta Jain Goyal. Role and Importance of Machine Learning in Social Media. Journal of Mobile Computing, Communications & Mobile Networks. 2024; 11(1): 23–30p.

analytics, and marketing, dealing with social issues, applications in government, difficulties, potential trends, and ramifications and then wrap off with a summary of the most important conclusions and suggestions.

LITERATURE REVIEW

Through the use of machine learning techniques, any machine may be taught how to understand a system that has a set of input/dependent variables and the desired output. Indents are required for each paragraph. Decision tree-based, logistic, and linear regression-based, and neural network-based machine learning techniques will be categorized under these headings. Three different learning modes can be used by machine learning. These three learning methods are reinforcement, supervised, and unsupervised [1]. Machine learning algorithms come in many different varieties, including Support Vector Machines (SVM) [2], K-Nearest Neighbour (K-NN) [3], and Generalized Instance Set (GIS). A machine learning method based on statistical learning theory is support vector machine. Due to the so-called "curse of dimension", high dimensional feature spaces pose the most difficult problems for other machine learning algorithms.

The SVM algorithm proved to be particularly successful in addressing these issues. When compared to other teaching techniques, K-NN provides clear target functions as examples of training. A family of learning algorithms known as Memory-Based Learning or sample-based as KNN and Rocchio is used in this method. Only samples for their training are stored by the algorithms, or the training example is changed to a different object before storage. An algorithm like this one is Generalized Instance Set (GIS) [4]. To reduce some of the noise in the training set, GIS replaces positive and negative training documents judiciously within the training document pool. It automatically chooses examples that are good, representative, and generic. [5] Because there is an increase in harmful content on social networks, automated methods must be used to discover and delete it.

Creating a supervised machine learning classification model to track the spread of harmful content in online social networks (ONSs) was the aim of this project. Multisource characteristics have been used to identify social network messages that contain malicious Uniform Resource Locators (URLs). These URLs may lead users to websites that include malicious software, drive-by downloads, phishing, spam, and scams. The dataset was labelled using Virus Total, while the data collection phase utilized the Twitter streaming application programming interface (API). Using a random forest classification algorithm, data from several sources were combined. Using a random forest classification algorithm, data from several sources were combined. And feature selection had a recall value of 0.89. But with more investigation, strategies for feature selection, and parameter modification, we were able to improve the classifier performance to 0.92 in recall [4].

The natural language expressions of people can be easily conveyed through brief yet impactful text messages on social networks, which represent a burgeoning and demanding industry. The polarity of texts, that is, whether they are positive, negative, or neutral, can be inferred from social settings in important ways [6]. If users who posted bullying-related tweets could subsequently want to erase those posts, Xu *et al*. investigated regret behaviours in those messages [7]. Up to 60.7% cross-validation accuracy was reported. The feature sets employed by Dadvar *et al.* are content-based, cyberbullying-based, and user-based [8]. The best recall was achieved with user-based and pronoun-profanity window feature sets (55.0% recall, 77.0% precision, and 64.0% F1 measure). Dinakar *et al.* deconstructed the detection of cyberbullying into the identification of sensitive topics, which is likely to lead to talks about bullying including topics like sexuality, racism, IQ, and profanity. [9]. With SVM, the accuracy for the topic of sexuality is 79%. Nahar *et al*. used user ranking and probabilistic features to obtain 99% accuracy [10]. Using a variety of features, such as content, sentiment, and contextual ones, Yin *et al*. demonstrated 59.5% recall, 35.2% precision, and 44.4% F1 measure. [11]. However, these techniques use supervised learning and involve giving a learner direct access to the entire input feature space. These methods cannot manage the noisy and unbalanced data [12].

Journal of Mobile Computing, Communications & Mobile Networks
Volume 11, Issue 1
ISSN: 2349-901X

SENTIMENT ANALYSIS IN SOCIAL MEDIA

On social media, a variety of information is posted and shared in the form of text, videos, photographs, and audio. Social media is full of raw, unprocessed data, and technological advancements, particularly in machine learning and artificial intelligence, have made it possible to process and turn that data into meaningful information that may help most commercial organizations [13]. The method of utilizing machine learning and natural language processing to identify the emotional tone or sentiment expressed in text data is known as sentiment analysis, commonly referred to as opinion mining. This in social media refers to classifying user-generated content as favourable, unfavourable, or neutral, enabling the evaluation of public attitude towards subjects, goods, or companies [14]. Three levels: sentence level, document level, and feature level, are used to categorize sentiment analysis. The objective is to categorize the viewpoint into positive and negative sentiment based on the language, document, or feature [15]. Machine learning and lexicon-based approaches have been established as the two basic methodologies of sentiment analysis. While lexicon-based approaches work by counting the positive and negative words that are associated to the data, machine learning approaches use algorithms to extract and detect sentiment from data. A fresh, precise model for sentiment analysis has been created by academics. But creating a model when the majority of it is designed for the English language is a hurdle. However, a recent study demonstrates that sentiment analysis model design exists in various languages such as Korean [16], Thailand [17], Arabic [18], Malay [19], Portuguese [20] and Chinese [21].

Methods Used in Sentimental Analysis

All the papers showed how to implement sentiment analysis using a lexicon-based method, a machine learning method, or a combination of both. The findings indicate that, while performing sentiment analysis, 10 publications employ machine learning, seven papers use lexicon-based methods, and seven papers combine both methods. Unsupervised learning is a term used to describe the lexicon-based approach. The lexicon technique solely relies on the dictionary and does not need any training data. When doing sentiment analysis, most of the studies adopted the TF-IDR and Sentiwordnet methods. This method is based on the frequency with which the phrases appear in the text data along with other positive or negative words from polarity lexicons that have already been constructed, such as Sentiwordnet [22, 23].

The phrase frequency-inverse document frequency technique is used to calculate the TF-IDR method, which converts words into numbers [24]. The strategies rely on lexical resources, and the quality of those resources has a significant impact on how well the whole strategy works. A text's polarity can be determined by looking at the polarity of the words that make up the text. This technique is not intended to address every facet of language because natural languages are so complicated, especially when it comes to slang, irony, and negation [25]. Using sentimental language is insufficient. There are various issues, such as the fact that some words have different meanings depending on the context, that some phrases with sentimental words may not actually be expressing anything, and that many statements without sentimental terms can also suggest an opinion [26]. The lexicon-based approach does, however, have some advantages of its own, such as simple counting of positive and negative terms, adaptability to other languages, and speed of analysis.

Implement counting of positive and negative words, adaptability to other languages, and quickness of analysis. The supervise learning category includes machine learning, which needs training data to be processed. The SVM and Nave Bayes models are the two machine learning techniques that are most frequently utilized. These are the most widely used machine learning models, while there are more. Applying Nave Bayes to a text corpus with good structure yields positive results [27]. When using a support vector machine, a low shape dataset performs well. Despite this, machine learning methods struggle on Facebook, where posts are frequently misspelled and are of arbitrary length. To improve these methods, a substantial amount of training data must be collected [26, 28]. Additionally, using a complex machine learning model to analyse data takes a lot of time, especially if training is necessary.

A smaller training dataset speeds up the procedure, but the classification accuracy suffers [29]. Therefore, combining the two approaches is advised to improve the results because they will enhance one another and produce a better end result than employing one way alone. Combining methods can be useful in identifying phenomena [29].

Applications of Sentiment Analysis

Sentiment analysis is used in a variety of social media contexts. Businesses use it to assess client feedback, modify their marketing plans, and enhance their product offers. Sentiment analysis in politics offers insights into trends in popular opinion, assisting parties in customizing their campaigns. It is used by media companies to track audience responses and modify material as necessary. Sentiment analysis is used in a variety of fields, including business and marketing, politics, health, and social policy. Sentiment analysis has a wide range of applications that can help with decision-making across many different domains. Sentiment analysis can be used to examine reactions to global events like current conflicts, sporting events, and natural disasters [30].

Several notable case studies showcase the efficacy of sentiment analysis. For instance, during product launches, companies analyse social media sentiment to assess initial reactions and make real-time adjustments. Political campaigns use sentiment analysis to gauge public sentiment towards candidates, guiding campaign strategies. These examples highlight how sentiment analysis can provide actionable insights.

Predictive Analytics: Concepts and Methods

Utilizing statistical algorithms and historical data, predictive analytics forecasts future results. This refers to using user-generated data to forecast user behaviour, trends, and content engagement in the context of social media. To help with this process, machine learning methods like collaborative filtering and decision trees are frequently used [31].

Predicting User Behaviour on Social Media

Social media predicts user behaviour with the help of predictive analytics. Predictive models can determine what material a user will likely engage with in the future by examining historical activity, such as likes, shares, and clicks. This functionality is used to curate news feeds, recommend friends, and propose content, improving the user experience [31].

Personalized Recommendations

Personalization is a cornerstone of predictive analytics in social media. Algorithms examine user data to provide tailored content recommendations, including articles, products, and connections. Popular platforms like Netflix and LinkedIn employ these techniques to enhance user engagement, resulting in longer user sessions and increased user satisfaction [31].

Case Studies and Success Stories

Predictive analytics in social media is effective, as demonstrated by several well-known case studies. Predictive modelling is used by Netflix's recommendation system, for example, to make recommendations for movies and television episodes, greatly boosting user retention and viewing hours. Like this, Amazon's product recommendation engine has significantly increased sales by making tailored product recommendations. These success tales demonstrate the revolutionary impact that predictive analytics can have on user engagement and business outcomes in the social media space [31].

SOCIAL MEDIA MARKETING AND MACHINE LEARNING

Targeted Advertising with Machine Learning

In social media marketing, machine learning provides accurate targeting. To ensure that ads are seen by the correct people, algorithms examine user data to determine optimal demographics and interests. This maximizes advertising ROI while also improving ad relevancy [32].

Ad Campaign Optimization

Ad campaigns are continuously modified in real-time by machine learning algorithms, which optimize elements including ad placement, bidding tactics, and content. This dynamic tuning increases the effectiveness of advertisements, resulting in higher click-through and conversion rates [32].

Measuring ROI and Performance

Machine learning aids in the measurement of social media marketing ROI. Advanced analytics and attribution models track user journeys from ad exposure to conversion, providing valuable insights into campaign effectiveness and areas for improvement [32].

Ethical Considerations in Social Media Marketing

Privacy problems with user data, ad transparency, and algorithmic prejudice are just a few ethical challenges that social media marketers need to be aware of. It is critical to strike a balance between efficient marketing and ethical data utilization to uphold credibility and adhere to moral obligations [32].

ADDRESSING SOCIAL ISSUES ON SOCIAL MEDIA

Detecting Hate Speech and Cyberbullying

Hate speech and cyberbullying can be instantly detected and flagged by machine learning algorithms. They do text and user interaction analyses to assist social media companies in taking quick action to safeguard users and promote a safer environment online [33].

Combating Misinformation and Fake News

Misinformation and fake news stories are identified and categorized using machine learning models. By comparing material with reliable sources, fact-checking algorithms help to stop the spread of incorrect information on social media [34].

Promoting Information Integrity

To promote information integrity, social media platforms employ machine learning to monitor content for authenticity and reliability. This includes identifying deep fakes, manipulated media, and content that violates platform policies [33].

Ensuring Online Safety and Inclusivity

By detecting and halting harassment, discrimination, and harmful content, machine learning helps to ensure online safety. Additionally, by recognizing and eliminating prejudice in algorithmic decisions and content suggestions, it contributes to the creation of a more inclusive and equal online space on social media platforms [34].

CHALLENGES AND ETHICAL CONSIDERATION

Privacy Concerns and Data Security

There are significant privacy concerns raised by the widespread collecting and analysis of user data. To avoid breaches, abuse, and unauthorized access to sensitive information, protecting personal information and data security is essential [35].

Algorithmic Bias and Fairness

Biases existing in training data might be inherited by machine learning algorithms, producing discriminating results. The hard task of ensuring fairness in algorithmic decision-making calls for ongoing monitoring, bias reduction, and transparency [35].

Ethical Use of Machine Learning in Social Media

The range of ethical considerations is wide, ranging from transparent algorithmic procedures and responsible data processing to combating misinformation and upholding user rights. Adherence to ethical principles and rules is necessary for achieving ethical use, which promotes trust and accountability in the creation and application of machine learning in social media [35, 36].

Emerging Technologies in Social Media

Emerging technologies like augmented reality (AR), virtual reality (VR), and 5G connections are poised to transform user experiences as social media continues to change. With the integration of various technologies, machine learning will be essential for providing immersive and individualized experiences [37].

The Future of Machine Learning in Social Media

Future applications of machine learning in social media promise even more precise sentiment analysis, content personalisation, and user behaviour forecasting. Advanced natural language processing, reinforcement learning, and deep learning are poised to take the lead [37].

Potential Societal Impact

Social media has a significant impact on society in terms of how we use information, interact with others, and even make decisions. While it creates chances for good change, it also raises issues with privacy, bias, and false information. To create a harmonious society, it will be crucial to strike a balance between innovation and responsible use [38].

Dimensionality Reduction in Social Media Dataset

Dimensionality reduction techniques are increasingly applied in the realm of social media analysis to handle the vast amount of data generated by users. Social media platforms generate massive datasets with high-dimensional features such as user profiles, posts, comments, likes, shares, and other interactions. Dimensionality reduction methods can help in Visualization, Feature Extraction, Noise Reduction, Clustering and Classification, Anomaly Detection etc. [39, 40].

CONCLUSION

This study has examined how important machine learning is in the world of social media. It has shed light on how user experiences and societal dynamics are changing because of the usage of sentiment analysis, predictive analytics, marketing, and social issue resolution. The promise of machine learning to improve information integrity, enable data-driven decision-making, and establish safer online environments is highlighted by key findings. To promote openness, fairness, and user data protection, recommendations include the creation of strong ethical frameworks, ongoing bias monitoring, and the implementation of industry standards. Social media and machine learning can significantly influence the digital environment as they continue to converge. Stakeholders, from people and enterprises to governments and technology providers, share responsibility for navigating the difficulties and maximizing the advantages. The ethical integration of machine learning will have an impact on society as a whole and influence social media's digital future.

REFERENCES

1. Du KL, Swamy MN. Neural networks and statistical learning. London: Springer Science & Business Media; 2013 Dec 9.
2. Dubey GP, Gupta N, Bhujade RK. A novel approach to intrusion detection system using rough set theory and incremental SVM. Int J Soft Comput Eng (IJSCE). 2011 Mar; 1(1): 1448.
3. Raghuwanshi A, Dubey GP. Enhanced PCA-KNN Classifier Method for Effective Diagnosis of a Parkinson’s Disease. Int J Emerg Technol Adv Eng (IIJETAE). 2020; 10(1): 67–72.
4. Munawar Z, Siswoyo B, Herman NS. Machine learning approach for analysis of social media. ADRI Int Journal Information Technol. 2017 Jan 13; 1(1): 5–8.
5. Jang–Jaccard J, Nepal S. A survey of emerging threats in cybersecurity. J Comput Syst Sci. 2014 Aug 1; 80(5): 973–93.
6. Pozzi FA, Fersini E, Messina E, Liu B. Sentiment analysis in social networks. Massachusetts, United States; 2016 Oct 6.

7. Xu JM, Burchfiel B, Zhu X, Bellmore A. An examination of regret in bullying tweets. In Proceedings of the 2013 conference of the North American chapter of the Association for Computational Linguistics: Human language technologies. 2013 Jun; 697–702.
8. Dadvar M, Trieschnigg D, Ordelman R, De Jong F. Improving cyberbullying detection with user context. In Advances in Information Retrieval: 35th European Conference on IR Research, ECIR 2013, Moscow, Russia, March 24–27, 2013. Proceedings 35. Berlin Heidelberg: Springer; 2013; 693–696.
9. Dinakar K, Reichart R, Lieberman H. Modeling the detection of textual cyberbullying. In Proceedings of the International AAAI Conference on Web and Social Media. 2011; 5(3): 11–17.
10. Nahar V, Unankard S, Li X, Pang C. Sentiment analysis for effective detection of cyber bullying. In Web Technologies and Applications: 14th Asia-Pacific Web Conference, APWeb 2012, Kunming, China, April 11–13, 2012. Proceedings 14. Berlin, Heidelberg: Springer; 2012; 767–774.
11. Yin D, Xue Z, Hong L, Davison BD, Kontostathis A, Edwards L. Detection of harassment on web 2.0. Proceedings of the Content Analysis in the WEB. 2009 Apr 21; 2(0): 1–7.
12. Nahar V, Al-Maskari S, Li X, Pang C. Semi-supervised learning for cyberbullying detection in social networks. In Databases Theory and Applications: 25th Australasian Database Conference, ADC 2014, Brisbane, QLD, Australia, July 14-16, 2014. Proceedings 25, Springer International Publishing; 2014; 160–171.
13. Sivarajah U, Kamal MM, Irani Z, Weerakkody V. Critical analysis of Big Data challenges and analytical methods. J Bus Res. 2017 Jan 1; 70: 263–86.
14. Gursoy UT, Bulut D, Yigit C. Social media mining and sentiment analysis for brand management. Global Journal of Emerging Trends in e–Business, Marketing and Consumer Psychology (GJETeMCP). 2017; 3(1): 497–551.
15. Sharma A, Tahiliani K, Dubey GP. Reactive-Optimized Sentence Detection in Kubernetes Using OpenNLP and Native GraalVM Image with Framework Metric Comparison. In 2023 IEEE 4th International Conference for Emerging Technology (INCET). 2023 May 26; 1–9.
16. Song M, Park H, Shin KS. Attention-based long short-term memory network using sentiment lexicon embedding for aspect-level sentiment analysis in Korean. Inf Process Manag. 2019 May 1; 56(3): 637–53.
17. Sanguansat P. Paragraph2vec-based sentiment analysis on social media for business in Thailand. In 2016 IEEE 8th International Conference on Knowledge and Smart Technology (KST). 2016 Feb 3; 175–178.
18. Itani M, Roast C, Al-Khayatt S. Developing resources for sentiment analysis of informal Arabic text in social media. Procedia Comput Sci. 2017 Jan 1; 117: 129–36.
19. Chekima K, Alfred R. Sentiment analysis of Malay social media text. In Computational Science and Technology: 4th ICCST 2017, Kuala Lumpur, Malaysia, 29–30 November, 2017. Singapore: Springer; 2018; 205–219.
20. Cirqueira D, Pinheiro MF, Jacob A, Lobato F, Santana Á. A literature review in preprocessing for sentiment analysis for Brazilian Portuguese social media. In 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI). 2018 Dec 3; 746–749.
21. Zhang S, Wei Z, Wang Y, Liao T. Sentiment analysis of Chinese micro-blog text based on extended sentiment dictionary. Future Gener Comput Syst. 2018 Apr 1; 81: 395–403.
22. Hamouda A, Rohaim M. Reviews classification using sentiwordnet lexicon. In World congress on computer science and information technology. 2011 Jan; 23: 104–105.
23. El Rahman SA, AlOtaibi FA, AlShehri WA. Sentiment analysis of twitter data. In 2019 IEEE international conference on computer and information sciences (ICCIS). 2019 Apr 3; 1–4.
24. Das B, Chakraborty S. An improved text sentiment classification model using TF-IDF and next word negation. arXiv preprint arXiv:1806.06407. 2018 Jun 17.
25. Khan MT, Durrani M, Ali A, Inayat I, Khalid S, Khan KH. Sentiment analysis and the complex natural language. Complex Adapt Syst Model. 2016 Dec; 4(1): 1–9.
26. Akter S, Aziz MT. Sentiment analysis on facebook group using lexicon based approach. In 2016 IEEE 3rd international conference on electrical engineering and information communication technology (ICEEICT). 2016 Sep 22; 1–4.

27. Hassan AU, Hussain J, Hussain M, Sadiq M, Lee S. Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression. In 2017 IEEE international conference on information and communication technology convergence (ICTC). 2017 Oct 18; 138–140.
28. Mahtab SA, Islam N, Rahaman MM. Sentiment analysis on Bangladesh cricket with support vector machine. In 2018 IEEE international conference on Bangla speech and language processing (ICBSLP). 2018 Sep 21; 1–4.
29. Dhaoui C, Webster CM, Tan LP. Social media sentiment analysis: lexicon versus machine learning. J Consum Mark. 2017 Sep 11; 34(6): 480–8.
30. Peng H, Cambria E, Hussain A. A review of sentiment analysis research in Chinese language. Cognit Comput. 2017 Aug; 9(4): 423–35.
31. Johnson LA. Predictive Analytics in Social Media Marketing: A Case Study of Facebook Advertising. Marketing Science Quarterly. 2020; 25(2): 189–207.
32. GAO Report. (2021). Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities. Government Accountability Office. [Online]. https://www.gao.gov/assets/gao-21-519sp.pdf Accessed on 12-Dec-2023.
33. Febriana T, Budiarto A. Twitter dataset for hate speech and cyberbullying detection in Indonesian language. In 2019 IEEE International Conference on Information Management and Technology (ICIMTech). 2019 Aug 19; 1: 379–382.
34. Dame Adjin-Tettey T. Combating fake news, disinformation, and misinformation: Experimental evidence for media literacy education. Cogent Arts Humanit. 2022 Dec 31; 9(1): 2037229.
35. Kaddoura S, Al Husseiny F. The rising trend of Metaverse in education: challenges, opportunities, and ethical considerations. Peer J Comput Sci. 2023 Feb 13; 9: e1252.
36. Dubey GP, Bhujade RK. Investigating the Impact of Feature Reduction Through Information Gain and Correlation on the Performance of Error Back Propagation Based IDS. Int J Electr Electron Res (IJEER). 2021; 9(3): 27–34.
37. Chen H, Chiang RH, Storey VC. Business intelligence and analytics: From big data to big impact. MIS Q. 2012 Dec 1; 36(4): 1165–88.
38. Himthani P, Dubey GP, Sharma BM, Taneja A. Big data privacy and challenges for machine learning. In 2020 IEEE 4th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC). 2020 Oct 7; 707–713.
39. Dubey GP, Bhujade RK, Himthani P. Significance of Dimensionality Reduction in Intrusion Detection Dataset. In International Conference on Intelligent Vision and Computing. Cham: Springer International Publishing; 2021 Oct 3; 518–529.
40. Hashmi SK, Dubey GP, Chourasia V, Yadav A. Generating an Optimal Feature Subset for Performance Enrichment of Intrusion Detection System using GINI. Int J Appl Eng Technol. 2023 Jan; 5(1): 60–70.

