WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

MACHINE LEARNING APPROACHES FOR SOCIAL MEDIA CONTENT ANALYSIS

Sigabadi Ghagan Kumar

CSE GMRIT Rajam,India

ABSTRACT :

The integration of AI and ML has greatly transformed the way businesses communicate with their customers on social media. This paper discusses the impact of AI on marketing and customer service on different social media networks. AI enables companies to establish more customized and individualized marketing messages through the analysis of user behavior. Additionally, it can enhance interactions between a business and its customers through chatbots and recommendation systems among other tools. With AI, businesses can conceive more intelligent advertisements, produce better content, and reach their audiences more effectively. This causes higher engagement with their brands. This paper also covers how AI is shaping social media, making it an integral part of contemporary marketing strategies. Businesses need to use AI effectively to stay competitive, connecting with target audiences and meeting the prescribed marketing objectives. Adaptation to such technological advancement can increase the existence and success of companies in the digital world.

Keywords— social media, posts, artificial intelligence, machine learning, social media analysis, tools

Introduction :

In the modern era of technology, social media channels have transformed into the central means of sending and receiving messages, sharing information, and creating content, thereby influencing the day-to-day activities of millions of users from all over the world. Machine learning (ML) is one of the more recent tools developed to help in dealing with such challenges, especially the processing of huge volumes of data and retrieval of useful information hidden in its depths. In the case of social networks, applying ML algorithms helps not only to analyze existing data, but also helps businesses and researchers to visualize trends, anticipate changes, and improve the overall experience of users like never before. The objectively analyzed data included the posts made on social media by ML systems from the ML Algorithms for Social Media Analysis book. With its capacity to spot trends, allow forecasts, and perform decision-making operations, ML is changing the landscape of analyzing data collected from social media. From improved marketing tactics and academic research, and armed with the most advanced ML techniques, there are effective means of meeting the challenge of big data and finding business value in it.

The social media sites have been incorporated into the lives of flesh and blood social network participants who are internet users from across the globe. Most people use social networks to retrieve information and to publish contents where they exhibit their abilities. For instance, one might find a visual of an individual's professionally developed video resume in the portfolio of that person. People share various types of posts on social media which include pictures, written information, emojis and videos. There are few restrictions as to what can be posted or shared which creates a very large volume of data. This data that encompasses every aspect of big data can further be exploited and put into various uses – commercial and educational in nature as well. Numerous firms operate on the basis of collecting and selling information as a service through APIs targeting other organizations. However, traditional techniques such as SQL cannot provide the solution to the analysis of this wide range of data. Advanced statistical analysis techniques with machine learning techniques are the only solution for that problem. This chapter presents a number algorithms directed to different machine learning approaches to social and big data networks.

The algorithms used in the analysis of social media content from machine learning perspectives are diverse and include various techniques starting from Natural Language Processing (NLP) like Sentiment analysis, and Topic Modelling to image processing techniques like Content Based Image Retrieval. These sampled methods are critical in dealing with deep learning to appreciate the available raw data and give room for its utilization in industries including but not limited to advertising, political science, risk communication, and global health. In this article, we work on the existing literature on machine learning techniques for social media content analysis focusing one by one the approaches, benefits, drawbacks, and future directions of each technique in this area of research.

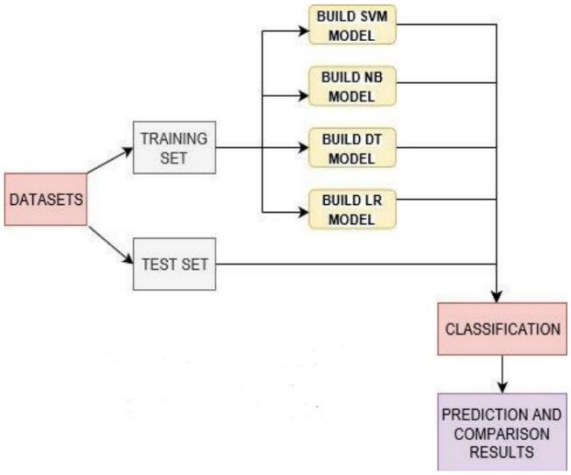


Fig.1.Architecture of Methodology

This review paper focuses on the different machine learning techniques used in analyzing content on social media, starting from supervised and unsupervised learning, all the way up to deep learning and natural language processing techniques. While doing so, the paper also considers the applications of these approaches in order to elaborate on their advantages and disadvantages, and more importantly the role that ML plays in the comprehension and analysis of social media content. According to L'heureux et al., machine learning is identified as a process of self-learning that takes place as a result of processing large amounts of data. Traditional methods known as 'symbolic AI' relied on logical algorithmic instructions designed to compute the solution within a predefined class known as the target for every possible instance or set of command inputs. Machine learning strategies on the other hand are quite different in this respect since they utilize large sets of data to 'learn' and therefore do not need to define the mathematical relations or temporal sequences linking inputs and outputs. The interactions taken place over such network has also helped millions of users across the world on the rise of internet communication. Social Network Analysis, in fact, contributes to the users' mapping and aggregation of their ideas and preferences (which are the nodes), diagrams showing the density of users' communications (the links), and providing interpretation to online activities. Network analysis or SNA has found diverse usefulness in many sectors including business and research.

Literature survey :

Machine learning tools are revolutionizing content analysis in social media by providing efficient and accurate insights from vast unstructured sources, as identified through the study of Anashkin and Malyshenko [1]. The study led on to show just how easy machine learning may make content analysis, going as far as giving real-time monitoring and data-driven decisions. In the same manner, through Arasu, Seelan, and Thamaraiselvan's study, machine learning techniques can actually benefit social media marketing strategies better in terms of targeting a particular audience [2].

The recent trends of neural networks have been incorporated into social media portals to make them more functional with new horizons in one-to-one content delivery and user engagement, reveals research by Taherdoost [3]. Another paper further moves along this trend by presenting artificial intelligence and machine learning as ways to increase the relevance of content and the interaction rates on social media, as discussed by Uygun and Gujrati [4]. This is further supported by Sharma and Kumar through data mining and analytics for superior social media monitoring with machine learning techniques [5].

Machine learning applications extend to social media analysis beyond merely marketing to how to handle other subjects like cyberbullying. Bansal et al. could realize algorithms that would help find harmful contents, thus making the online environment safer [6]. In dealing with problematic social media use, Savci, Tekin, and Elhai apply predictive approaches in machine learning to understand behavioral tendencies related to excessive use on social media [7]. Furthermore, Fan et al. outline the application of machine learning for archiving data from social media and processing it into accessibly friendly cultural research [8].

The application of machine learning has also been explored in the area of detecting hate speech online. For instance, Marshan et al. made a comparison between traditional and deep learning methods to pour light upon the effectiveness of these methods in identifying and categorizing hate comments by severity [9]. Expanding on this, Dangsawang and Nuchitprasitchai apply machine learning in detecting customs fraud from social media, therefore indicating greater applications of unstructured data analysis [10].

Bayesian deployments, Zohaib and Khanam presents how practical approaches can be enacted for the use of machine learning on nuanced social media analysis focusing on probabilistic models in order to better generate insights [11]. Emotional recognition in social media conversation is advanced by Chowanda, Sutoyo, and Tanachutiwat, emphasizing that machine learning might be able to acquire human emotions clearly through some textual data [12].

Methodology :

Machine learning approaches for content analysis in social media involve extracting insights from user-generated data. Techniques like graph theory, neural networks, natural language processing, and random forests help identify trends, analyze sentiments, detect influencers, and classify topics. These methods enable understanding user behavior, monitoring trends, and improving content moderation effectively.

Neural Networks

Neural networks are the backbone of a machine learning model, and are particularly effective in dealing with complexity in social media content. Essentially, these networks mimic the human brain, using layers of interconnected nodes called neurons to process and analyze data. In analyzing the content shared on social media, neural networks are particularly efficient at detecting patterns in unstructured data, such as in text, images, and videos.
In fact, with text content containing sequential data, the neural networks, particularly Recurrent Neural Networks and more advanced versions such as Long Short-Term Memory networks, can adequately analyze the given text to identify sentiment, classify topics, or detect trends in posts or comments. For visual content, Convolutional Neural Networks are used to analyze the image or video, for example, to identify specific objects, faces, or scenes, so that within the visual the trends can be identified or harmful content can be detected.Modern neural network models therefore usually incorporate transformers, BERT and GPT being some of the well-known ones, for better natural language understanding. These models are convenient in handling the informal, divergent language style on social media, using slang words, emojis, as well as multilingual content. Neural networks also facilitate content personalization through predictions of user preferences and presentation thereof.
They are computationally expensive and require a lot of training data. They also tend to be hard to interpret with their "black-box" nature. However, they are very effective in extracting valuable information from this explosion of diversity happening on social media with better content moderation, trend analysis, and audience engagement.

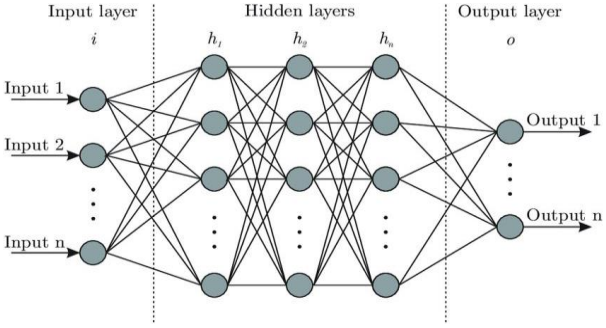


Fig.2.methodology of neural networks

This is what defines neural networks-the ability to adapt and generalize, which are especially well-suited for social platforms whose nature is dynamic and constantly evolving. Neural networks go much beyond text analysis and image analysis-the interaction patterns between the users allow it to detect bots, fake accounts, or coordinated misinformation campaigns. It plays a more critical role in maintaining the integrity of these virtual spaces.
Advanced applications include the analysis of different modes, where neural networks collect inputs from both text and images, as well as video data at the same time. A post having both an image and its caption can be analyzed holistically to determine the sentiment or underlying message. This is invaluable in terms of understanding complex social media content that combines different kinds of data.
Neural networks also form the basis for real-time applications like flagging offensive or objectionable content at the time of posting. Since the diversity of data on which models can be trained makes them highly robust against the diversity of global users in varied accents, languages, and contexts.
Despite the challenges they have—interpretability, high computational demands, and the possibility of bias from training data-neural networks are revolutionizing how social media content is analyzed. Scalability and efficiency enable the platforms and organizations to map insights into action, contribute toward more improved user experience, and generally address important social media challenges effectively.

Natural Language Processing(NLP)

Natural Language Processing, NLP-stands for that sub-field of Artificial Intelligence, which is oriented toward the interaction between computers and human language. Regarding social media, NLP is likely to play a very important role in the analysis of and in making sense of most of the unprecedented volumes of unstructured text data that arise from user activity for instance, describing tweets, Facebook posts, Instagram image captions, comments on reviews, or even in private messages. Techniques of NLP help decode this extraordinarily rich, often illiterate language and make it possible to analyze the content in such ways as sentiment analysis, topic modeling, and trending among so much else.

Understanding Social Media Text

Social media content is informal, carrying slang words, abbreviations, emojis, hashtags, and even nonstandard grammar. Therefore it's more complicated in its interpretation for more conventional language processing systems. Those are usually developed for more formal or structured text. NLP techniques have evolved to easily address this kind of informality, thus making them highly efficient in analyzing social media content. Such models could allow NLP to pay attention to the detail of social media language while still optimally capturing the nuances of language and its underlying structures, ultimately guiding users toward deeper insights about their behavior, sentiment, or trends.

The most pervasive use of NLP in social media is sentiment analysis-that is, whether a comment, tweet, or post has a positive, negative, or neutral disposition. The sentiment is of utmost importance to businesses, marketers, and researchers as a verdict of public opinion when the platform users frequently declare their emotions, opinions, and reactions. NLP models can identify sentiment, even amidst slangs, sarcasms, and mixed sentiments, by analyzing words, phrases, and context. More advanced models in sentiment analysis pay attention not only to the context, but also to what is being attacked-such as a product, celebrity, political figure, etc-in this case of a tweet-end.

Another very important application of NLP in content analysis on social media is topic modeling. Topic modeling algorithms, including Latent Dirichlet Allocation, automatically choose the greatest themes or topics that dominate a collection of social media posts. This allows for tracking which subjects are popular, seeing emerging trends, and knowing what users are speaking about. For instance, a social media analytic tool could use NLP on exactly the same batch of tweets to extract topics including "climate change," "technology," or "sports," which in turn will help organizations or researchers monitor what is being discussed and if these topics give traction over time.

Another method of NLP is the named entity recognition that identifies and classifies entities of interest to social media-related content such as the names of persons, locations, organizations, products, or any events. Social media NER helps pull priceless insights about brand mentions that are usually made by users mentioning celebrities or places. Thus, the tracking of brand mentions across different platforms will have insight into public perception, potential marketing opportunities, or signals of a possible public relations crisis.

Using NLP to automatically identify and classify these entities may enable organizations to systematically analyze large volumes of content.

Text Classification is the task of applying predefined labels to social media posts based on content. This can run the gamut from positive/neutral/negative, through topic-specific labeling, and into picking out things like spam, hate speech, or marketing-type material. It could all be scaled out using training on these machine learning models, trained upon labeled datasets, to then allow companies the capability to filter out unwanted or damaging content, among other abusive language, or detect marketing campaigns. Because hashtags are one of the most crucial elements of social media communication, one fundamental application of NLP is the analysis of the relationships between the hashtags used and, therefore, monitoring trending topics and prediction of the virality of certain discussions that might occur.

For instance, an NLP system, by analyzing the usage of certain hashtags in group terms, can identify new emerging trends or conversations before they are at their peak popularity. This way, businesses, media companies, and researchers can respond more quickly. Although NLP has proven pretty effective in analyzing social media content, there remain plenty of challenges ahead. The main challenge arises from the informality and diversity of language used on the social media platform. Many users write with informal styles, utilize abbreviations, and use emoticons or slang, features that make the traditional NLP models extremely challenging. In addition, sarcasm and irony are widely used in social media; these are difficult for machines to correctly deduce the meaning of the language.

For instance: Any such tweet that says "Love getting stuck in traffic for hours, great experience! " could be interpreted as a positive sentence by a basic sentiment analysis model, though it is intended to be negative.

The second one is the multilingual context. For instance, Twitter and Facebook are used by people from basically all over the world, and a post can be written in a number of languages, dialects, or even mixed with other languages. Models have to cope with the linguistic diversity in this type of content and switch between languages and cultural contexts to analyze it.Context is also important in the interpretation of social media text. Users may allude to specific events, celebrities, or trends that are premised on understanding their background deeper cultural or social context. NLP models should thus be able to discern such context in the posts for them to capture the meaning being conveyed.

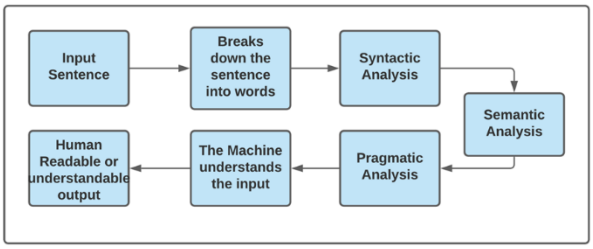
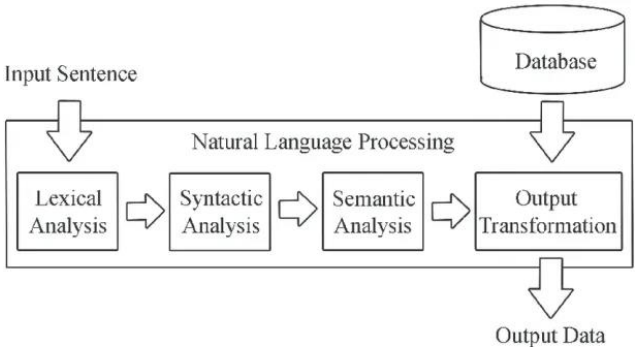


Fig.3.flowchart of nlp

Recent advancements in NLP have made it an even more effective tool for content analysis on social media. Transformer-based models, such as BERT and GPT, revolutionized NLP, where previous models were unable to understand the nuances of language. Such models are mainly pre-trained on massive amounts of text with the ability to fine-tune these models for specific tasks on social media, like sentiment analysis or trend detection. They can even take care of complex language, understand context, and even make their way through informal language and slang.

NLP offers some really powerful tools to analyze and make sense of user-generated content at scale in the context of social media: from sentiment analysis and topic modeling to trend detection and important entities. Businesses, researchers, and platforms can gain deeper insights into user behavior and public opinion and understand emerging trends in more detail. Challenges such as informalities in language and sarcasm aside, improvement in techniques for NLP has constantly enhanced accuracy and adaptability, and it's soon going to be an indispensable tool for social media content analysis.



NLP is a transformed analytical tool for analyzing social media text, giving insights into user sentiment, trends, and behavior. Still, challenges abound; people use informal language, and sarcasm, among others. However, improvements in NLP models improve accuracy and scalability and are the only way through which businesses, researchers, and platforms will find a way to effectively engage with social media data.

Random Forest

Random Forest is considered to be one of the most popular ensemble techniques in machine learning for classification and regression tasks. In general, using this technique to carry out analysis on social media data has been proved to be very efficient in handling big amounts of complex data consisting of user interactions, text data, images, and much more. For instance, it is particularly useful for tasks in sentiment analysis, topic classification, prediction of user behavior, and content moderation. It can combine multiple decision trees, improving the accuracy of predictions, reducing overfitting, and handling noisy or unstructured data that is really common on social media platforms.

How Random Forest Works:

Random Forest is an ensemble method that constructs a collection or "forest" of decision trees during the training phase. This collection is built by constructing each tree with a random subset of the data and a random subset of features. This ensures that all the individual trees are different, and therefore, the model is more robust and it can capture any and all aspects of the data. At test time, the Random Forest algorithm takes the majority vote for classification tasks or averages over all the trees for regression tasks. This aggregation of multiple trees helps to reduce errors and overfitting, which is a common problem in decision trees when they are used individually.

The key components of a Random Forest algorithm include:

Bootstrap Aggregating, or Bagging: In this process, it samples random subsets of the training data with replacement and trains each tree on different subsets. This helps not overtrain the model on any particular feature set or data points, hence increasing generalization.
Feature Randomization: At every node during the training of each decision tree, only a random subset of the features is considered for splitting. This further diversified the trees and made the model less likely to overfit on particular features.

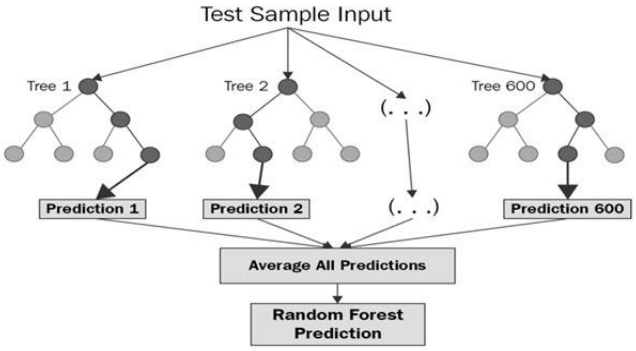


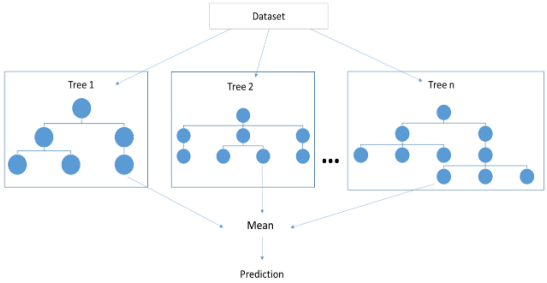
Fig.5.methodology of seq2seq

Application - Sentiment analysis: One of the most popular applications of random forests is in content analysis over social media. Social networking has produced millions of user-generated contents that contain opinions, emotions, and even responses to various events or topics. Such content can be categorized into sentiment types, for example, positive, negative, or neutral, using the random forest. It learns from labeled data and then can generalize to different types of content. It could be on a tweet, a Facebook post, or an Instagram comment. The advantages of using multiple decision trees increase the robustness of this model against informal language, slang, and emojis used on social media and in general improve the accuracy of the sentiment detection.
Topic Classification Another very important application of Random Forest is topic classification. Topic classification is a process in which posts are classified into predefined topics or themes. For example, posts related to politics, technology, sports, or entertainment can be classified into respective categories. Random Forest works very effectively because it can handle a high-dimensional feature space and capture complex patterns in the data. It handles text-based features such as words, phrases, and hashtags but can also think about even more context-specific features like the past behavior of a user, their geographic location, or history of how they have engaged with other content. Because Random Forest handles many combinations of features, it is well-suited for tackling the complex and large-scale data associated with social media
User Behavior Prediction Random Forest can also be used to predict user behavior in social media platforms. The model analyzes past user interactions, which include likes, comments, shares, and follows, to predict future behavior-say, predicting what post a user might engage with or what kind of content they could be interested in creating. The predictive power has great value in content recommendation systems, in marketing strategy, and so on. For example, service such as YouTube or Facebook can apply Random Forest to predict what messages a user will likely interact with; so this would personalize the feed of the user and thus improve the experience.
Content Moderation The other critical application of Random Forest in social media is content moderation. With millions of content posts appearing each day, social media requires automated systems that can flag problematic material, for example, hate speech, bullying, or graphic violence. Text-based

content like this may be classified as "acceptable" versus "harmful" by a Random Forest on features such as specific words, phrases, or patterns of interaction. It can also handle images and videos, though for those applications, Random Forest may be employed in combination with other machine learning techniques, such as CNNs. Utilizing the aggregated decision trees, Random Forest models could be able to make highly accurate predictions about the contents violating the guidelines of the community, and therefore such content may be removed from such media sites more effectively.

Detection of Emerging Trends: The identity of emergent trends in social media could also be an important application of Random Forest. Trends, hashtags, or newly emerged topics often breakout from social media streams first. Random Forest can be applied for trending tracking with categorizing trends by analyzing temporal patterns of content and how users engage as well as when particular hashtags are used. For example, a Random Forest model could determine when a new hashtag starts trending through the analysis of Twitter post streams, helping brands or news organizations stay up to date. Further, the algorithm can predict how long any given trend might run with the insights gathered through historical data, which again provides some great insights for marketers and businesses.

Spam Detection Spam detection is also an issue with social media platforms wherein there is common flooding of the system through spammy or harmful content, mainly through fake accounts, bots, and malicious users. Such accounts or posts would be picked out by Random Forest based on characteristics such as posting frequency, patterns of text, and history of engagement. The training of the algorithm by the labeled data mix of the legitimate and spammy posts would ensure that Random Forest classifies new content as either a non-spam or a spam to increase the quality of feeds within social mediaIn Conclusion The powerful and flexible machine learning approach, Random Forest, is truly great at content analysis for social media. It can handle diverse data types; its ability to improve accuracy in predicting and to generate insightful data about users' behavior and trends and sentiment make the overall tool a good one for a business on any social media site.



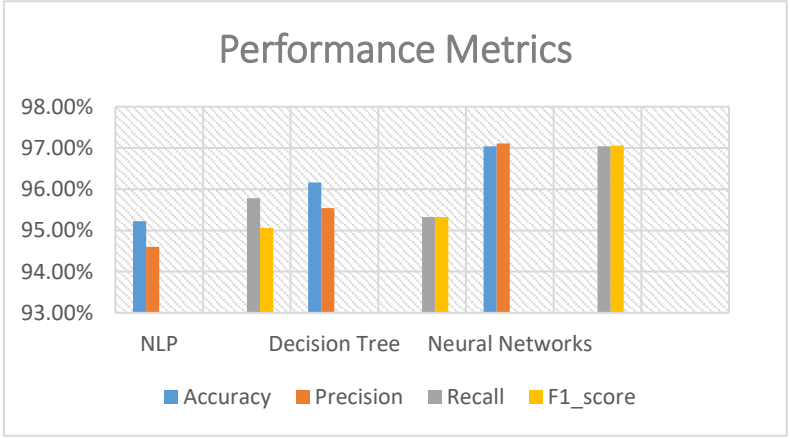
Results and discussion :

Neural Networks could be applied perfectly to process high-dimensional and complex data like images, texts, and videos, really delivering high accuracy regarding tasks like sentiment analysis and image recognition. They are powerful in learning subtle patterns without doing laborious feature engineering; however, they need big datasets along with heavy computative resources and risk overfitting if regularization is not applied properly, especially to small datasets.NLP excels in text analysis to identify sentiment and emotion, among several content moderation possibilities. Quite advanced models like BERT and GPT draw on context and tone quite well. However, it is not-so-great with less formal language, slang, and multi-language text and requires hundreds of thousands of computations to run large models. And for classification and prediction in structured data like predicting metrics of engagement, Random Forest stands quite tall. It combines several decision trees for robust predictions and handles features well, avoiding overfitting; however, does not too well on unstructured data like raw text or images. Further, it can fall behind some more modern neural networks for complex pattern recognition.

Model	***Accuracy***	***Precision***	***Recall***	***F1_score***
NLP	95.22%	94.60%	95.78%	95.06%
Decision Tree	96.16%	95.54%	95.32%	95.32%
Neural Networks	97.04%	97.11%	97.04%	97.05%

comparison among methodologies used

fig.6.graphical representation of various performance metrics



Conclusion :

The study had Neural Networks, Graph Theory, Natural Language Processing (NLP), and Random Forest regarding social media content analysis. All of them have unique strengths; Neural Networks, especially deep learning models, are great at identifying very complex patterns, mostly suited to image and text classification. Again, NLP, especially Transformer-based architectures, makes it suitable for sentiment analysis and emotion detection. Graph Theory would be helpful to understand structures of networks. Identify communities and influential users. Random Forest is excellent for structured data classification, but it fails to classify unstructured content in the case of social media. For content analysis tasks on social media content Neural Networks and NLP models generally give the best outcome and maximum accuracy. Neural networks are very accurate for deep learning models in complex patterns of multimedia data and NLP models, especially with transformer-based architectures, in parsing and interpreting the subtleties of social media text.

REFERENCES :

Reference 1: Arasu, B. Senthil, B. Jonath Backia Seelan, and N. Thamaraiselvan. "A machine learning-based approach to enhancing social media marketing." *Computers & Electrical Engineering* 86 (2020): 106723.

Reference 2 : Taherdoost, Hamed. "Enhancing social media platforms with machine learning algorithms and neural networks." *Algorithms* 16.6 (2023): 271.

Reference 3 : Uygun, H. A. Y. R. İ., and Rashmi Gujrati. "Role of artificial intelligence & machine learning in social media." *International Journal of Mechanical Engineering* 7, no. 5 (2022): 494-498.

Reference 4 : Sharma, Anu, and Vivek Kumar. "Machine Learning Prospects: Insights for Social Media Data Mining and Analytics." *International Journal of Innovative Research in Computer Science & Technology* 11.3 (2023): 12-19.

Reference 5 : Arasu, B. Senthil, B. Jonath Backia Seelan, and N. Thamaraiselvan. "A machine learning-based approach to enhancing social media marketing." *Computers & Electrical Engineering* 86 (2020): 106723.

Reference 6 : Bansal, A., Baliyan, A., Yadav, A., Kamlesh, A. and Baranwal, H.K., 2022. Cyberbullying Detection on Social Networks Using Machine Learning Approaches. *International Research Journal of Engineering and Technology (IRJET) Volume*, *9*.

Reference 7 : Savci, Mustafa, Ahmet Tekin, and Jon D. Elhai. "Prediction of problematic social media use (PSU) using machine learning approaches." *Current Psychology* 41.5 (2022): 2755-2764.

Reference 8 : Fan, Lizhou, Zhanyuan Yin, Huizi Yu, and Anne J. Gilliland. "Using machine learning to enhance archival processing of social media archives." *Journal on Computing and Cultural Heritage (JOCCH)* 15, no. 3 (2022): 1-23.

Reference 9 : Marshan, Alaa, Farah Nasreen Mohamed Nizar, Athina Ioannou, and Konstantina Spanaki. "Comparing Machine Learning and Deep Learning Techniques for Text Analytics: Detecting the Severity of Hate Comments Online." *Information Systems Frontiers* (2023): 1-19.

Reference 10 : Dangsawang, Bundidth, and Siranee Nuchitprasitchai. "A machine learning approach for detecting customs fraud through unstructured data analysis in social media." *Decision Analytics Journal* 10 (2024): 100408.

Reference 11 : Zohaib, Muhammad, and Maria Khanam. "Practical Bayesian Implementation of Machine Learning Techniques for Social Media Analysis."

Reference 12 : Chowanda, Andry, Rhio Sutoyo, and Sansiri Tanachutiwat. "Exploring text-based emotions recognition machine learning techniques on social media conversation." *Procedia Computer Science* 179 (2021): 821-828.

Reference 13 : Gannarapu, S., Dawoud, A., Ali, R.S. and Alwan, A., 2020, November. Bot detection using machine learning algorithms on social media platforms. In *2020 5th international conference on innovative technologies in intelligent systems and industrial applications (CITISIA)* (pp. 1-8). IEEE.

Reference 14 : Akhter, M. P., Jiangbin, Z., Naqvi, I. R., AbdelMajeed, M., & Zia, T. (2022). Abusive language detection from social media comments using conventional machine learning and deep learning approaches. *Multimedia Systems*, *28*(6), 1925-1940.

Reference 15 : Poecze, F., Ebster, C., & Strauss, C. (2018). Social media metrics and sentiment analysis to evaluate the effectiveness of social media posts. *Procedia computer science*, *130*, 660-666.

Reference 16 : Kokab, S. T., Asghar, S., & Naz, S. (2022). Transformer-based deep learning models for the sentiment analysis of social media data. *Array*, *14*, 100157.

Reference 17 : Shin, D., He, S., Lee, G. M., Whinston, A. B., Cetintas, S., & Lee, K. C. (2020). *Enhancing social media analysis with visual data analytics: A deep learning approach* (pp. 1459-1492). Amsterdam, The Netherlands: SSRN.

Reference 18 : Vermeer, S. A., Araujo, T., Bernritter, S. F., & van Noort, G. (2019). Seeing the wood for the trees: How machine learning can help firms in identifying relevant electronic word-of-mouth in social media. *International Journal of Research in Marketing*, *36*(3), 492-508.

Reference 19 : Sharma, A., Sharma, M. K., & Dwivedi, R. K. (2023). Exploratory data analysis and deception detection in news articles on social media using machine learning classifiers. *Ain Shams Engineering Journal*, *14*(10), 102166.

Reference 20 : Boyd-Graber, J., Glasgow, K., & Zajac, J. S. (2013). Spoiler alert: Machine learning approaches to detect social media posts with revelatory information. *Proceedings of the American Society for Information Science and Technology*, *50*(1), 1-9.