International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 6 June 2023, pp: 653-658 www.ijaem.net ISSN: 2395-5252

A Survey on Identifying Radicalized Content on Social Media Applications

Hiral Jain[1], Prof. Shailendra Bhalla[2]
Department of CSE, SDBCT, Indore[1,2]

Date of Submission: 14-06-2023 Date of Acceptance: 24-06-2023

ABSTRACT: With advancements in technology and cyber warfare, terrorist and radical content is being used as a tools to spread violence and social unrest. It is important to note that any system used for radical content identification should be designed with care to ensure fairness, avoid bias, and respect privacy and ethical considerations. Regular monitoring and updating of the models are also crucial to adapt to emerging radical content patterns. Machine learning approaches have used thus far to identify potentially radical content. This paper presents a comprehensive review on filtering out potentially radical social media content based on machine learning approaches.

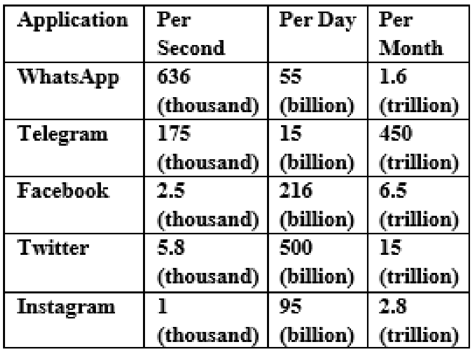
Keywords: Social Media, Radicalized content, machine learning, accuracy.

I. INRRODUCTION

The recent times have seen an extensive use of internet and social media. Different business groups and communities use social media for business and economic purposes. The world has become heavily dependent on technology today. The proliferation of social media is paramount. With the economies growing at rapid rate, these are using the digital infrastructure for better use. Social media has also evolved over the past few years. Not only individual users but several businesses are being run digitally. There is a varied and diverse content on social media that is being consumed world over. Any one can post content on social media by having an account. The positive aspects of social media are countless. But there are also some cons. Artificial Intelligence has also seen an increase in its usage over the recent few years. It can help in categorising and classifying the various types of content on social media platforms. Some key point's related to social media and digital content are as follows:-

- Social media use has increased massively in the recent few years.
- The social media is all about content. It is immensely content driven. Various kinds of content can shared on it.
- It is important to regularize the type of content it has as it is on the internet which is accessible to anyone and everyone.
- The social media and digitization has immensely helped large tech based companies to expand their business and economy.

Figure 1 depicts the typical message magnitude on different social media platforms [3].



Application	Per Second	Per Day	Per Month
WhatsApp	636 (thousand)	55 (billion)	1.6 (trillion)
Telegram	175 (thousand)	15 (billion)	450 (trillion)
Facebook	2.5 (thousand)	216 (billion)	6.5 (trillion)
Twitter	5.8 (thousand)	500 (billion)	15 (trillion)
Instagram	1 (thousand)	95 (billion)	2.8 (trillion)

Fig.1 Message content of Different Social Media Applications

The rampant use of social media has led to the increase in spread of all types of misinformation. The freedom of using the internet has also opened the ways towards using it in many wrong ways. The spread of radical content is one of the most harmful ways of creation of political turbulence [4]. The proliferation of the social network on the web has given rise to a number of possibilities to share and spread any type of content. The benefit of anonymity of dark web and freedom also gives the extremist groups the advantage of doing radicalized activities on the internet through social media engagement[5]. This helps these extremists and terrorist organizations to lure and influence common people into joining their organizations and promote radically motivated acts. This also leads to politically controversial activities and increases the participation and growth of such unlawful and terror acts [6].

Radicalized information that is spread through various online mediums can endanger the security of a nation. Henceforth, it is of enormous importance to check and prevent this kind of radicalization through various social media. For this purpose, correct and accurate identification of radical content is important and proper classification is necessary. As the social media contains a huge amount of information with a huge user base, hence it is important to use a artificial intelligence and machine learning (AI & ML) based methods for identifying and classifying online radical content [7].

II. LITERATURE REVIEW

This section presents the existing work in the domain.

In [1], Kapitonov et al. studied about the malicious messages that the terrorist communities used to share through the use of social media. The messages on propaganda are circulated on these instant messaging handles and the social network. The only method to combat such messages is to block the handles and such spread of messages. To carry this out, the researchers require processing significantly huge amounts of data. In this paper, the authors present a method based on artificial intelligence and machine learning that could automate the classification process of separating the radical content. The only limitation of this approach is that it was more based on the conventional statistical implementation.

In [2], Lopez et al. proposed a unique approach and framework for automated monitoring of radical information in the occurring in the social network Twitter. It mainly focussed on two main aspects; detection of the majority of the users who had radical content propagation agenda and then supervision of the interaction of the radical content between users that were involved in such things. The authors also did a case study. The main problem with this approach was that the model couldn’t specify the exact demarcating boundary between the content. Many of the datasets overlapped with each other and higher precision was needed.

In [3], Bobashev et al. presented and researched on a case of 1995 Paris metro Attack. Where, the incidents of the classes were traced to find the messages of radical content belonging to them. This led to the origination of the group that did such activities. Dynamic visualization and analysis were performed for the tracing of the formation of the terrorist group. This was the subtle use of the Natural Language processing. The major limitation here analysed is that the NLP used was in its very initial stage and better NLP modelling and methods could fetch better results.

In [4], Sun et al. put forth that method of prediction of the terrorist attacks by groups. They studied that though it is a very important issue around intelligence and security analytics, it is also a hard task to perform it. Conventional schemes and approaches are not enough for classifying such complex datasets. So the there is a need to build a robust and very advanced system to classifying such information. The amount of data that is there on the digital platforms is enormous. Feature extraction is a very important part of machine learning. Better the features extracted, better is the neural network trained. This is one of the areas which could have been improved to increase the accuracy and precision.

In [5], Tundis et al. worked on an approach to state that using a high end computer associated approach could yield viable results. With every positive aspect, there are also chances of using the social media in a malicious way. In some of the latest research works in this context, it has been found that several radical content are being spread using the medium of social networks. Many illegal community groups who choose to remain anonymous are misusing the platform to spread radical content and misinformation. This approach used text analysis methods by considering multiple language aspects. But one of the major issues that surfaced was that it was hard to classify based on one token of word taken at a time. Better multi token word analysis could be evaluated for the same purpose.

In [6], Zevairi et al. studied the traits and researched on the Islamic terrorist activities and involvement. They specifically studied their characteristics and motives that differentiated from other sets of radicals. Supervised machine learning approach was used to implement this that was a combination of 4. Using the different combination of learning algorithms helped but the understated use of pre processing posed as a major problem. Pre-processing of data is a major necessity for the machine learning methods which helps in improved classification of data.

In [7], Johnston et al. explored about the Sunni extremists groups that were involved in the jihadist radical content propaganda. Dark net is usually a well targeted place of engagement in such illegal and radical activities. Invoking terrorist activities on the cyberspace and spreading communal posts has seen a recent increase. Terrorism on cyberspace is equally dangerous and can inflict harm. This method combines the neural network with deep learning approach. While the results were good and the neural network worked well, but the use of a single neural network could not process the multitude of training datasets. Hybrid or ensemble approach could be used to help with this approach for better outcomes and accuracy.

.In [8], Ishitaki et al. studied about the Tor application that was being utilized for propagating radical content and information. The spread was attributed to the anonymity of the users that helped in such rapid and enormous circulation of the messages. This was the tool for rolling out malicious content on the digital landscape. Conventional schemes and approaches are not enough for classifying such complex datasets. So the there is a need to build a robust and very advanced system to classifying such information. The amount of data that is there on the digital platforms is enormous. It's very important to classify such information to stop it from being broadcasted. They used the deep learning on tor web server. The system could yield better accuracy if probabilistic approach was used for such complex set of data.

In [9], Lourentzou et al. studied on the use of the deep neural networks based on geographical locations. They studied the use on the geographical prediction and tried to figure out the radical content based on locations. Without any specific and clear boundary that can demarcate the radical and non radical data, the classification of such content is also an uphill task. The warfare of recent times is just not confined to the battleground alone. The cyberspace has also become a well targeted medium. Invoking terrorist activities on the cyberspace and spreading communal posts has seen a recent increase. Deep Neural method has been implemented but there need more variety of data sets for proper training and better performance.

In [10], Lara-Cabrera et al. provided insight into the works of some extremists groups that use the medium of social media to propagate malicious content the social network. Terrorism on cyberspace is equally dangerous and can inflict harm. This method combines the neural network with deep learning approach. While the results were good and the neural network worked well, but the use of a single neural network could not process the multitude of training datasets. The indicators of radical content were researched upon. Better pre-processing mechanisms could be enforced for better classification of the data sets used for the analysis of the proposed system.

Table.1 Summary of Literature Review

S.No.	Authors	Approach and findings
1.	Kapitoanov et al	The approach uses the machine learning based approach using the Naïve Bayes Classifier for identifying radicalized content for twitter data.
2.	Sadiq et al.	The approach used the ensemble of Bi-LSTM and CNN (CNN-Bi-LSTM) based approach for filtering extremist Arabic texts.
3.	Asif et al.	The approach presented the linear support vector classifier (SVC) model for filtering extremist content from Facebook datasets
4.	Fraiwan et al.	The approach proposed the SVM-OAA (one against all) and SVM-AAA (all against all) methods for classifying radical content using semantic lexicons and emotion features.
5.	Owoeye et al.	The approach proposed the use of support vector machine (SVM) as the technique to identify and classify potential radical content
6.	Rosewelt et al.	This approach proposed a convolutional neural network (CNN) based approach for semantic analysis of social medial tweets
7.	Lansley et al.	The approach proposed fuzzy logic and decision trees based machine learning approaches for identifying radical content.
8.	Hitest et al.	The approachproposed a word to vector based random forests (RF) algorithm for sentiment

		analysis
9.	Padmaja et al.	The approach proposed the use of adaptive neuro fuzzy inference systems (ANFIS) and genetic algorithm for sentiment classification of twitter data
10.	Katarya et al.	Genetic algorithm used for sentiment analysis as positive, negative and neutral classes.

RESEARCH FRAMEWORK

The proposed work aims at filtering out radical content and classify testing samples as potentially radical or un-radical. This however is challenging owing to the fact that there exists no clear demarcation among positive, negative and neutral data which is mined.

Fig.2 Sentiment Analysis

Figure 2 illustrates the concept of sentiment analysis from mined data.With the evolution of social networks, it has definitely changed the digital landscape and how the digital content is consumed. With every positive aspect, there are also chances of using the social media in a malicious way. In some of the latest research works in this context, it has been found that several radical content are being spread using the medium of social networks. Many illegal community groups who choose to remain anonymous are misusing the platform to spread radical content and misinformation. With the ease of social media sharing this allows anyone to post anything at liberty and this can be very risky sometimes.

The mathematical model of the artificial neural network is depicted in figure 3.

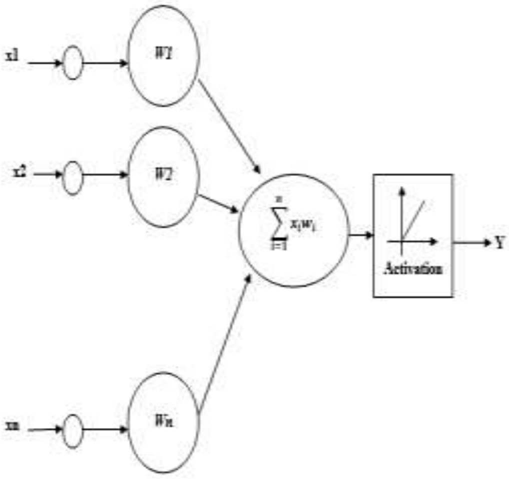


Fig.3 Mathematical model of ANN

The output of the neural network can be related to the inputs as:

$$\mathrm{output} = f[\textstyle\sum_{i=1}^{n} x_i w_i + \theta] \qquad (1)$$

Here,

Output vector correcsponds to the output vector of the network.

The inputs and weights are represented by x and w respectively.

The additional term of bias termed as θ is added.

III. CONCLUSION

Form the previous discussions, it can be concluded that with advancements in technology and cyber warfare, terrorist and radical content is being used as a tools to spread common vendetta which can cause violence and social unrest. Due to the complexity and the enormity of the data size, it is humanly infeasible to analyze the data manually or even statistically using dictionary based learning. Thus machine learning based approaches are indispensable for the detection and classification of the radical and possible terrorist activity conducive content. This paper presents the contemporary approaches for identification of radical content.

REFERENCES

[1]. K Chetioui, B Bah, AO Alami, A Bahnasse, "Overview of Social Engineering Attacks on Social Networks",

Procedia Computer Science, Elsevier 2022, vol. 198, pp.656-661.

[2]. A Pimentel, KF Steinmetz, “Enacting social engineering: the emotional experience of information security deception” Crime, Law and Social Change, Springer, 2022, vol. 77, pp.341–361.

[3]. A.I. Kapitanov, I. I. Kapitanova, V. M. Troyanovskiy, V. F. Shangin and N. O. Krylikov, "Approach to automatic identification of terrorist and radical content in social networks messages," 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), 2018, pp. 1517-1520

[4]. M. Nouh, J. R. C. Nurse and M. Goldsmith, "Understanding the Radical Mind: Identifying Signals to Detect Extremist Content on Twitter," 2019 IEEE International Conference on Intelligence and Security Informatics (ISI), 2019, pp. 98-103.

[5]. S Davis, B Arrigo, “The Dark Web and anonymizing technologies: Legal pitfalls, ethical prospects, and policy directions from radical criminology”, Crime, Law and Social Change, Springer 2021, vol. 76, pp.367–386.

[6]. M. Ashcroft, A. Fisher, L. Kaati, E. Omer and N. Prucha, "Detecting Jihadist Messages on Twitter," 2015 European Intelligence and Security Informatics Conference, 2015, pp. 161-164.

[7]. S Mussiraliyeva, M Bolatbek, B Omarov, Detection of extremist ideation on social media using machine learning techniques”, Computational Collective Intelligence. ICCCI 2020. Lecture Notes in Computer Science. Vol. 12496, pp.743–752.

[8]. A.I. Kapitanov, I. I. Kapitanova, V. M. Troyanovskiy, V. F. Shangin and N. O. Krylikov, :A. I. Kapitanov, I. I. Kapitanova, V. M. Troyanovskiy, V. F. Shangin and N. O. Krylikov, "Approach to automatic identification of terrorist and radical content in social networks messages”, IEEE Access, 2021, pp. 1517-1520.

[9]. S Sadiq, A Mehmood, S Ullah, M Ahmad, “Aggression detection through deep neural model on twitter”, Future Generation Computer Systems, Elsevier 2021, vol.114, pp. 120-129.

[10]. M. Asif, A. Ishtiaq, H. Ahmad, H. Aljuaid, and J. Shah, ‘‘Sentiment analysis of extremism in social media from textual information,’’ Telematics Information., Elsevier 2020, vol. 48, 101345.

[11]. M. Fraiwan, ‘‘Identification of markers and artificial intelligence-basedclassification of radical Twitter data,’’ Applied Computing and Information, Emerald Publication, 2020, vol. 16, no.1.

[12]. K. O. Owoeye and G. R. S. Weir, "Classification of Extremist Text on the Web using Sentiment Analysis Approach," 2019 International Conference on Computational Science and Computational Intelligence (CSCI), 2019, pp. 1570-1575.

[13]. A Rosewelt, A Renjit, “Semantic analysis-based relevant data retrieval model using feature selection, summarization and CNN”, Soft Computing, Springer 2020, vol.24, pp.16983–17000.

[14]. M Lansley, F Mouton, S Kapetanakis, “SEADer++: social engineering attack detection in online environments using machine learning”, Journal of Information and Telecommunication, Taylor and Francis, 2020, vol.4, no.3, pp.346-362.

[15]. M. Hitesh, V. Vaibhav, Y. J. A. Kalki, S. H. Kamtam and S. Kumari, "Real-Time Sentiment Analysis of 2019 Election Tweets using Word2vec and Random Forest Model," 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), 2019, pp. 146-151.

[16]. K. Padmaja and N. P. Hegde, "Twitter sentiment analysis using adaptive neuro-fuzzy inference system with genetic algorithm," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 2019, pp. 498-503.

[17]. R. Katarya and A. Yadav, "A comparative study of genetic algorithm in sentiment analysis," 2018 2nd International Conference on Inventive Systems and Control (ICISC), 2018, pp. 136-14.

[18]. CN Kamath, SS Bukhari, A Dengel, “Comparative study between traditional machine learning and deep learning approaches for text classification”, DocEng '18: Proceedings of the ACM Symposium on Document Engineering, ACM 2018, Article No.14, pp.1-11.

[19]. M. Ebrahimi, A. H. Yazdavar and A. Sheth, "Challenges of Sentiment Analysis for Dynamic Events," in IEEE Intelligent Systems, vol. 32, no. 5, pp. 70-75.

[20]. R. M. Bütler, C. Häger, H. D. Pfister, G. Liga and A. Alvarado, "Model-Based Machine Learning for Joint Digital Backpropagation and PMD Compensation," in Journal of Lightwave Technology, vol. 39, no. 4, pp. 949-959.

[21]. X Yuan, L Xie, M Abouelenien, "A regularized ensemble framework of deep learning for cancer detection from multi-class, imbalanced training data", Pattern Recognition, Elsevier 2018, vol. 77, pp.160-172.