Systematics Review on the Application of Social Media Analytics for Detecting Radical and Extremist Group

Systematics Review on the Application of Social Media Analytics for Detecting Radical and Extremist Group

R T Adek[1], Bustami[2] and M Ula[3*]
[1,2]Informatics Department, Universitas Malikussaleh, Lhoksumawe
[3]Information System Department, Universitas Malikussaleh, Lhoksumawe

E-mail: munirulula@unimal.ac.id

Abstract. Recently, social media platforms such as Twitter, Tumblr, Facebook, YouTube, blogs and discussion forums are being mistreated by radical groups to promote their ideologies and encourage radicalization. Social medias also have been used to create online extremist community and recruit new followers. In this paper work, the authors conduct a schematics literatures review on all available techniques and perform a comprehensive analysis on the application of social media analytic for detecting radical group to understand the circumstances, trends and its gaps. Further, the author provides the characterization, classification and meta-analysis in order to achieve a better understanding of the literature on the extremist detection through intelligent social media. It is found that for over the last 10 years many researchers have been conduct deep investigation on the use of social media analytics on predicting and identifying online radicalization. Besides, data source, features, geolocation, language, machine learning techniques, and tools have been applied on the recent literatures to detect those cyber-extremist activists. This paper also highlighting the performance measurement methods that have been used by researcher for detecting extremist group and radical communities. The goal of this research is to provide an academic base for ongoing research in the developing machine learning algorithm for detecting extremist and radical contents in social medias.

1. Introduction

Over the last decade, social media has emerged into a dynamic form of interpersonal communication around the world. It facilitates users to constantly share information and continuously connect and convey opinion to the world through medias. For example - social networking (Facebook), micro-blogging (Twitter, Tumblr), image sharing (Imgur, Flickr) and video hosting and sharing (YouTube, Dailymotion, Vimeo) [1]. Likewise, the simplicity of navigation, low barriers to publication (users only need to have a valid account on the website) and anonymity (freedom to upload any content without revealing their real identity) have caused users to abuse this website in several ways by uploading offensively. and illegal data [1][2]. Further, popular social media websites, blogs, forums are often misused by divers of groups to promote online radicalization (also referred to as cyber-extremists, cyber-crime and cyber hate propaganda radicalization (also referred to as cyber-extremists, cyber-crime and cyber hate propaganda).

Previous research showed that extremist groups often delivered a hate speech, offensive comments and messages that focused on their mission. a large number of hate groups use popular social media websites to disseminate their ideology with extremist content for their readers [2] [3]. They communicate with other existing groups and form virtual communities on social media and

sharing a common agenda. They use social networks as a platform to hire new members by gradually reaching a worldwide audience that helps persuade others to commit violence and terrorism [4][5]. Researchers from various disciplines such as psychology, social science and computer science are constantly developing tools and knowledge on fighting technique and these problems online radicalization.

Hence, automatic detection and radicalizing content analysis is an important issue in the research domain of intelligent Security Informatic. Monitoring the presence of such content on social media and this issue in real time is predominant for security analysts who works on law enforcement agencies.

Since 2005, for more than 10 years, Information Security Intelligent (ISI) [5] [6] [7] has been proposed number of approaches, techniques, algorithms and tools to mine the data and provide solutions on emerging problems. The aim of the research in this journal is to carry out a systematic literature survey of the previous techniques as documented in scientific articles. Accordingly, the author's goal is to conduct a comprehensive analysis of those articles in order to get better understanding on research gap, techniques and future recommendation.

2. Research Method

This systematic literature review has been done by applying 6 stages process for a systematic review research method [8]. The process was started by research question formulation. The second stage was to determine the required characteristic for the study. The third stage was collecting relevant papers from the research databases. The fourth stage was selecting related research papers and cleansing the irrelevant papers. The fifth stage was synthesizing the pertinent information from the literatures and the last stage was tabulation and report writing [9].

To provide guideline of the systematics review, the following research question was addressed:

1. What data sources and features are used in detecting radical and extremist group?
2. What machine learning techniques are used in detecting radical and extremist group?
3. What performance evaluation method are used in detecting radical and extremist group?

2.1. Systematic Review Process

The data collection process in this systematic review is as shown in Figure 1. The research papers for literature review were retrieved from many credible online databases dan google scholar. This research utilize Google Scholar is because it provides more articles coverage and index and also offer a strong mechanism to explore related works, ideas and authors [9] [10]. Besides, Google Scholar also produces data on how often and how recently journals have been cited which is also a useful metric (assessing impact) for conducting author literature surveys. Variation of keywords used in search engines and online databases for data collection are “Extremist Detection, Radical Detection, social media, machine learning, 'online community radicalization detection”. After the research articles were collected. Title, Keywords and Abstracts were analysed. A meta-analysis process was conducted on the article and determines the relevance of the articles to the author's topic or focus area. The author conducts a one-class classification (Rule Based Classifier) on each article and checks whether it meets the specified scope and focus.

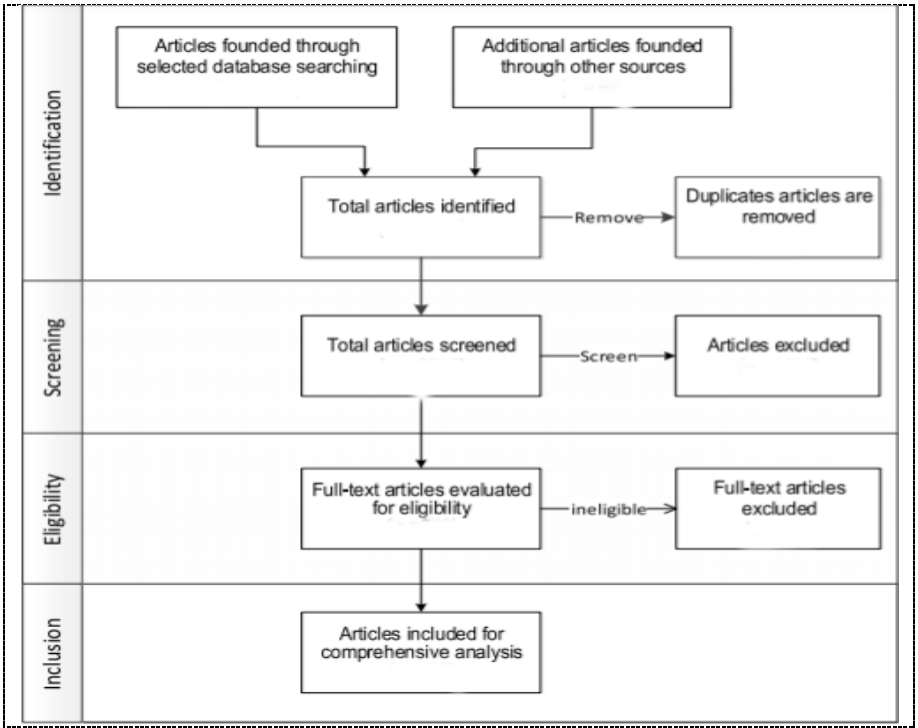

Figure 1. Systematic review process

3. Systematic Review Result

3.1. Data Source

From 2010 to 2020, there are more than 40 publications on online extremism detection. This review has identified many publications using multiple data sources and experimental datasets over the last decade. Figure 2 reveals that there is some work on all popular social networking platforms, micro-blogging websites [8] [9], video sharing [10] [11] [12] [13] and image hosting websites [14]. Other researchers have used blogs [15], forums [16], web documents, news articles [17] and websites [18] for hate and extremism detection. The figure shows that most of the research publications using social media platform as their data source. The most famous social media website is twitter. Because twitter offer easy data scrapping using API for developer. Other articles have utilizing blogs, discussion forums, news articles and websites as data sources. However, YouTube, also gradually become popular data source for researcher for detecting extremist and radical contents. Moreover, the preceding research indicate that YouTube is the most prevalent platform used for hate promotion [19].

Figure 2 shows that there are 28% of research papers have used Twitter, 24% has used multi data source, 19% has used Facebook, and 21% YouTube as a data source. The pie chart above shows that Twitter, Facebook, YouTube and multiple sources have been being used as data source by many researchers for detecting extremist and radical communities for the last decade [11] [12] [20] [21]. Twitter as a famous micro-blogging website is very popular platforms for radical group and community. It has been extensively used by the extremist to deliver their messages and propaganda to their communities. However, many researchers have used multiple platforms as their data sources for detecting extremist and radical communities. These multiple sources are including Facebook, Instagram, Twitter and others. Other data sources like news, blog, and forum have gradually received less attention from researchers in the last decade [22] [23] [24].

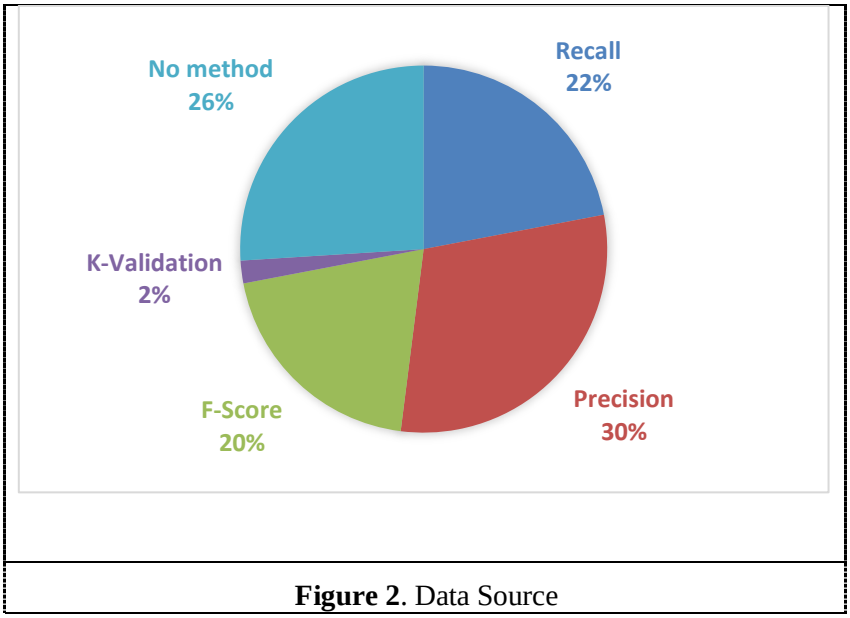

Figure 2. Data Source

3.2. Data Features

This review classifies research articles based on three discriminatory features as follows:

1. Text;
   Text defines as the contextual metadata of a post, tweet content, video title, hashtags in the video description, etc.
2. Link
   Link defines as the connection between two user account. It could be a "repost blog by" on Tumblr, "followers" on Twitter, "liked videos" on YouTube, etc.
3. Demographic information
   This information is a metadata based on post statistics and user profiles. For example, number of comments on videos, number of favourites on tweets, notes on Tumblr posts etc.

The systematics review results show that; 16 papers using text as their data source, 24 papers use links as their data, and 28 research papers use demographic information as their data. Furthermore, this review records that 25% of research papers use all three features in their analysis, while only 30% of these papers use text and demographic information to detect the extremist group and radical communities. Only 5% of the reviewed papers, that only use link as their data for detecting radical contents and extremist groups.

3.3. Type of Contents

In this review, type of contents in research articles are classified into 3 sub-categories namely 'content', 'user profile' and 'community' based on the type of content analyzed to provide solutions to fight and combat online extremism. Results of literature analysis discovers that out of 36 articles, in 35 papers, the authors conducted experiments on contextual metadata while in 18 papers, the authors extracted user profile information. Likewise, in 16 articles, they mine the profiles of linked users and their communities. Content analysis is an important feature for identifying extremist content.

3.4. Language

Social medias contain multi-lingual text in metadata, therefore, in this review, authors categorize research papers into three language classification. The review results reveal that among the 36 research articles, 14 articles are successfully detect the extremist group and radical content posted in Arabic. Furthermore, among these 14 articles, the methodology applied in 7 articles is able to detect

other non-English texts as well. The review also exposed that all the machine learning technique used in the past research were able to detect extremist and radical contents posted in English.

3.5. Geolocation

Several research papers on detecting extremist group and radical communities have been focused on the metadata determined and originates from a specific region and country. Therefore, in this review, authors have classified 36 research papers and articles into three categories which are Domestic, International, and Others. The review result shows that there are 20 research papers and articles that focus on detecting extremist and radical group taking place worldwide and other 5 articles focused on US domestic region and 9 research papers focus on middle east and Asian region.

3.6. Machine Learning Techniques

The graph in Figure 3 shows that Support Vector Machine (SVM) and Naïve Bayes (NB) are among the famous machine learning techniques used for detecting extremist and radical communities in last decade [25] [26]. Other machine learning techniques also used by researchers are KNN text classification, Rule Based Classifier, Decision Tree, Clustering, Exploratory Data Analysis (EDA), Topical Crawler are the most widely used techniques for online radicalization detection on social media websites [27] [28] [29]. Twitter, Facebook, Instagram and YouTube are users generated content websites. Those websites generated terabyte multimedia data in a month. Therefore, text classification techniques like EDA and KBF algorithms are very well and commonly used to identify extremist content on social media [21] [28].

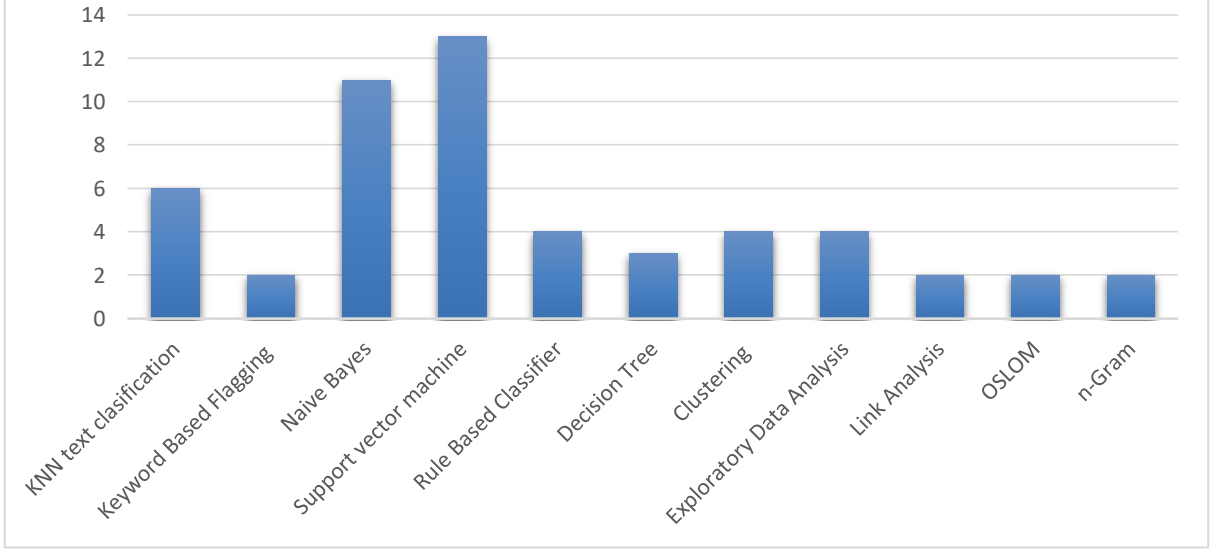

Figure 3. Machine learning techniques used by researchers

Topical crawler and link analysis is a technique used to enter social media websites through navigational links and identify similar users and find the hidden communities [11] [15] [22]. The crawler topical is a recursive process that adds and removes nodes after each iteration. Starting from the seed node, traverse in the graph navigating through multiple links and returning all relevant nodes to the given topic. Breadth First Search, Depth First Search and Best First Search are techniques to select neighbours and navigate them through external links. These links and neighbours are varied for different social networking websites. For example, if the user (u) posts a tweet (t) then the neighbours can become your users' followers, or the user likes and tweet or re-blogs (in the case of Tumblr) that

post. Likewise, on YouTube, if your user posts v video then the link can redirect back to the channel of the user who subscribed or posted a comment on video v. N-gram is another technique for classifying textual data in detecting radical contents based on some discriminatory features [4].

Other machine learning technique for detecting extremist is OSLOM (Order Statistics Local Optimization Method). It is one of the clustering methods designed for detecting graphs and social media networks. OSLOM has been applied for detecting extremist and radical group network [23] [24].

3.7. Performance Evaluation Method

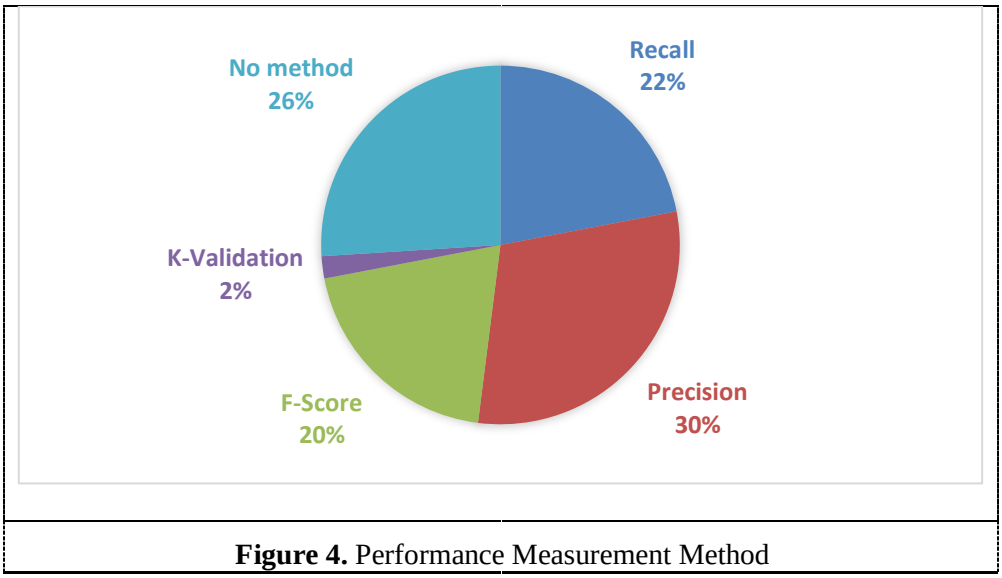

Figure 4. Performance Measurement Method

In this review we also analyse the performance measurement method used by many researchers. This section purposes are to provide answer for the last research question; What performance evaluation method are used in detecting radical and extremist group? The result of this systematics review shows that Precision, Retrieval, K-cross validation are commonly used to measure the performance of the machine learning technique used. In this work, author classify performance evaluation method into 5 sub-categories which are F-score, Precision, Retrieval, K-cross validation. F-score is the average harmonic of Precision and Recall while Accuracy calculates the "correctness" of the proposed method. *Figure 4* shows the distribution of the number of articles using various measures to examine the performance proposed over the last decade. Figure 4 also reveals that there is only one article using K-cross validation for evaluation in 7 articles no evaluation method is mentioned. The authors also consider that although 50% of the articles analysed only textual data to identify extremist content, only 30% of the articles used precision, while only 20% of the articles used recall as a measure for evaluation.

4. Conclusion

Applying intelligent in social media to identify online radicalization is an area that has attracted the attention of several researchers over the past 10 years. Characterization and meta-analysis conducted on existing studies on online radicalization detection revealed that identifying the presence of metadata based on extreme contextual content is the most commonly used feature. However, demographic and activity information from user profiles and links between two users are discriminatory features for finding extremists. The author observes that many existing techniques are capable of mining multi-lingual texts such as Arabic and capturing relevant information. The author also observed that precision is the most commonly used evaluation method to test the effectiveness of

the results in community detection and identification of extremism content, and Social Network analysis.

References

[1] Ahmad S, Asghar M Z, Alotaibi F M, and Awan I, 2019. Detection and classification of social media-based extremist afliations using sentiment analysis techniques, Human Centric Computing Information Science, (2019) 9:24, https://doi.org/10.1186/s13673-019-0185-6.
[2] Drus Z, Khalid H 2019, Sentiment Analysis in Social Media and Its Application: Systematic Literature Review. Procedia, The Fifth Information Systems International Conference 2019
[3] Sharif O, Hoque M M, Kayes A. S. M., Nowrozy R, and Sarker H I, 2020, Detecting Suspicious Texts Using Machine Learning Techniques. MDPI - Applied sciences Article 2020.
[4] Garg U and Kaur S, 2020. Systematic Review on Machine Learning Approaches for Sentiment Analysis. International Journal of Advance Science and Technology Vol. 29, No. 10S, (2020), pp. 2760-2764
[5] Arunachalam, R. dan Sarkar, S. 2013., 'The new eye of government: Citizen sentiment analysis in social media'. IJCNLP 2013 Workshop on Natural Language Processing for Social Media (SocialNLP) hal. 23–28.
[6] BBC. 2016. MIT dan JAT: Dua Kelompok Teror Indonesia Terkait ISIS. Diunduh dari www.bbc.com/indonesia/berita_indonesia/2016/01/160115_indonesia_explainer_kelompok_teror, accessed on 17 May 2019.
[7] Manning D C, Surdeanu M, Bauer J, Finkel J, Bethard S. J, and McClosky D. 2014. The Stanford CoreNLP natural language processing toolkit. In Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pages 55–60, Baltimore, Maryland, June 2014. Association for Computational Linguistics
[8] Christian D F, Kembro J, and Andreas. 2017 A New Paradigm for Systematic Literature Reviews in Supply Chain Management. Journal of Supply Chain Management Wieland 53 (4): 67-85
[9] Liu H, Christiansen T, Baumgartner Jr W A, and Verspoor K. 2012 Biolemmatizer: a lemmatization tool for morphological processing of biomedical text. J. Biomedical Semantics, 3(3):17, 2012.
[10] Hashemi, Mahmood. S, Jingsha He. 2017. LA-Based Approach for IoT Security. Journal of Robotics, Networking and Artificial Life, Vol. 3, No. 4 (March 2017), 240-248.
[11] IHS Jane's Defence Weekly. (2012). Annual Defence Report: Middle East and Africa, 7 December 2012.
[12] Liu, B. (2012). Sentiment Analysis And Opinion Mining: Morgan dan Claypool Publisher.
[13] Pak, A. dan Paroubek, P. (2010). 'Twitter as a corpus for sentiment analysis and opinion mining'.Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10) hal. 1320–1326.
[15] Selvam, B. dan Abiram, S. (2013). A survey on opinion mining framework.
[16] Agarwal S, Sureka A (2015) Topical Crawler for Uncovering Hidden Communities of Extremist Micro-Bloggers on Tumblr, 5th Workshop on Making Sense of Microposts, Microposts 2015, May 18th, 2015, Florence, Italy.
[17] Agarwal S, and Sureka A, Using KNN and SVM Based One-Class Classifier for Detecting Online Radicalization on Twitter, ICDCIT 2015, Springer International Publishing Switzerland 2015, LNCS 8956, pp. 431–442, 2015.
[18] Agarwal A, and Sureka A, 2015 Topic-Specific YouTube Crawling to Detect Online Radicalization DNIS 2015, Springer International Publishing Switzerland 2015, LNCS 8999, pp. 133–151, 2015.

[19] Agarwal A, and Sureka A, 2014, A Focused Crawler for Mining Hate and Extremism Promoting Videos on YouTube, HT'14, ACM, September 1–4, 2014, Santiago, Chile. 978-1-4503-2954- 5/14/09.
[20] Scanlon J R, and Gerber M S, 2014 Automatic detection of cyber-recruitment by violent extremists, Scanlon and Gerber Security Informatics 2014, 2014 Scanlon and Gerber; licensee Springer.
[21] Chen F, Daniel B. Neill, 2014 Non-Parametric Scan Statistics for Event Detection and Forecasting in Heterogeneous Social Media Graphs, KDD'14, 2014 ACM, August 24–27, 2014, New York, NY, USA.Copyright 978-1-4503-2956-9/14/08
[22] Jiejun Xu, Tsai-Ching Lu, Compton R, and Allen D, 2014 Civil Unrest Prediction: A Tumblr-Based Exploration, Springer International Publishing Switzerland 2014, SBP 2014, LNCS 8393, pp. 403–411, 2014.
[23] D. O'Callaghan, D. Greene, M. Conway, J. Carthy, and P. Cunningham. Uncovering the wider structure of extreme right communities spanning popular online networks. arXiv, preprint arXiv:1302.1726, 2013.
[24] Wadhwa P. and Bhatia M, 2013 Tracking on-line radicalization using investigative data mining. In Communications (NCC), 2013 National Conference on, pages 1{5. IEEE, 2013.
[25] I.-H. Ting, H.-M. Chi, J.-S. Wu, and S.-L. Wang. (2013) An approach for hate groups detection in facebook. The 3rd International Workshop on Intelligent Data Analysis and Management, Springer Proceedings in Complexity, pages 101{106. Springer Netherlands, 2013.
[26] Goodwin M, 2013 The Roots of Extremism: The English defence League and the Counter-Jihad Challenge. Chatham House.
[27] Patil G., Manwade K., and Landge M. P. 2013. A novel approach for recognized and overcrowding of terrorist websites. International Journal of Engineering Trends and Technology, Volume4Issue3-2013, ISSN: 2231-5381, pp 463-470
[28] Deshwal A.and Sharma S., 2016 Twitter sentiment analysis using various classification algorithms, 2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO).
[29] Ferrara E, Wang WQ, Varol O, Flammini A, Galstyan A (2016) Predicting online extremism, content adopters, and interaction reciprocity. International conference on social informatics. Springer, New York, pp 22–39