

A Systematic Comparison of Machine Learning and NLP Techniques to Unveil Propaganda in Social Media

Deptii D. Chaudhari, Symbiosis Institute of Technology, Symbiosis International University (Deemed), India*

 <https://orcid.org/0000-0003-3632-3832>

Ambika V. Pawar, Symbiosis Institute of Technology, Symbiosis International University (Deemed), India

ABSTRACT

With the technological advancements and its reach, social media has become an essential part of our daily lives. Using social media platforms allows propagandist to spread the propaganda more effortlessly and faster than ever before. Machine learning and natural language processing applications to solve the problem of propaganda in social media has invited researcher attention in recent years. Several techniques and tools have been proposed to counter propagation of propaganda over social media. This work analyses the trends in research studies in the recent past which address this issue. The purpose is to conduct a comprehensive literature review of studies focusing on this area. The authors perform meta-analysis, categorization, and classification of several existing scholarly articles to increase the understanding of the state-of-the-art in the mentioned field.

KEYWORDS

Deep Learning, Machine Learning, Natural Language Processing, Propaganda, Propaganda Analysis, Social Bots, Social Media Analytics, Social Networks

1. INTRODUCTION

Along with technological advances, social media has become an integral part of our lives. As per a recent report published by J. Clement*, in February 2020 approximately 2 billion internet users are using social networks. The number of users is still expected to grow as mobile device usage and mobile social networks increasingly gain traction.

Once developed with the intent of communication and entertainment, social media has now become a breeding ground for a lot of misinformation, rumors, hatemongering, and propaganda. In recent times, all over the world, many incidents were evident that anti-social elements tactically use social media to manipulate public discontent, separate communities, and cause public violence. For many people, social media has become a primary source of their news and information. The

DOI: 10.4018/JITR.299384

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

factuality and sometimes intent behind this news or information remains questionable. Though the web provides a substantial amount of valuable information, it has also become the source of false claims and misinformation coming in many forms like erroneous quoting about politics, fake or made-up news about celebrities or politicians, factually wrong product reviews, etc. Along with fake news and claims, another emerging problem is the spread of digital propaganda.

The dictionary defines Propaganda as “information, ideas, opinions, or images that give one part of an argument, which are broadcast, published to influence people’s opinions” which is frequently referred to in political context. The word propaganda first came to existence in 1622 when Pope Gregory XV established the ‘Congregation for propagating the faith’ for advancing Catholic missionary activity. It was not until the beginning 19th century that this word is used in the context of dubious accuracy as a method of advancing a cause.

According to Jacques Ellul, 1965 modern Propaganda operates with many kinds of truth, half-truth, limited truth, and truth out of context. In recent times propaganda has been used by terrorist organizations for recruitment, by political parties during elections, and even by marketing agencies.

In 1937, social scientists, opinion leaders, historians, educators, and journalists founded the organization called the Institute for Propaganda Analysis (IPA). This organization was established to spread awareness among American citizens about political propaganda. IPA defined propaganda as “*Propaganda is an expression of opinion or action by individuals or groups deliberately designed to influence opinions or actions of other individuals or groups with reference to predetermined ends.*” In their work, the Institute of Propaganda Analysis also identified seven features indicating the use of propaganda as Name-Calling, Glittering Generality, Transfer, Testimonial, Plain Folks, Card Stacking, and Band Wagon.

Other than these seven techniques different authors have identified more propaganda techniques. Wikipedia has listed 69 of such techniques. Using several techniques propaganda is loaded with falsification and emotional appeals to spread misinformation. As propaganda can appear in the form of numerous techniques, it requires extensive and deeper analysis to be recognized. Propaganda is also very hard to detect without the complete landscape of information available on the topic. Identification of propaganda on social media poses an even bigger challenge due to the enormous and quick spread on social media. Also, digital propaganda is propagated using a combination of bots and human users which makes it further difficult to be discovered.

In recent years, new techniques, tools, and approaches are being proposed by researchers to address the problem of propaganda in social media. Several researchers have made efforts to identify propagandist techniques such as the use of social bots, emotional appeals, bias news, etc. The study presented in this paper aims to provide a systematic literature survey of the previous techniques by the researchers. This work provides a thorough analysis of new theories, features, research gaps, tools and techniques, and future directions of propaganda detection in social media.

1.1 Technical Challenges

The automatic identification of social bots and automated propaganda propagation on social media has emerged as a challenging problem for natural language processing and computational linguistic researchers. The tremendous amount of data generated by social media platforms and the dynamic nature of these platforms makes it hugely difficult to discover, analyze, and predict propaganda content. Most of the research work is focused on either specific event or time duration (Johnston & Weiss, 2018), (Kellner et al., 2019), (Baines & O’Shaughnessy, 2014), (Derbas et al., 2020), (Neyazi, 2020), (Baisa et al., 2019) for propaganda analysis. The enormous volume of data, it’s multi-media and multi-lingual nature creates challenges like discovering the trends and hidden patterns, connections between social bots, and techniques used for propaganda. Even though some research (Varol et al., 2017), (Agarwal et al., 2017) have explored the connections between social botnets the multi-media and multi-lingual aspect remains unexplored. The textual and multi-media nature of social media data results in highly unstructured and noisy content. The social media data is usually riddled with

misspellings, internet slangs, incorrect grammar, and multi-lingual contents. Even though some steppingstone research by (Barrón-Cedeño, Da San Martino, et al., 2019), (Rashkin et al., 2017), (Heidarysafa et al., 2019), (Araque & Iglesias, 2020), (Volkova et al., 2017), (Alhindi et al., 2019) is done in this direction sophisticated natural language processing tools and techniques are required to deal with these challenges.

1.2 Contributions of This Work

This work provides an exhaustive and rigorous literature survey on the topic of propaganda identification in social media. Our work will provide insights to the researchers working to address the problem of propaganda in social media. We present an in-depth meta-analysis of techniques employed in the existing literature. This detailed analysis offers the intuitions to uncover the trends, research gaps, and future directions in the field of computational propaganda detection.

2. RESEARCH SCOPE AND FOCUS

This literature survey is focused on the crossroads of the following major disciplines:

1. Social Media Platforms
2. Natural Language Processing and Text Mining
3. Security Informatics

Although this literature review is limited to the use of social media for the spread of propaganda causing security threats and manipulation of public opinion. We focus our work on the use of web and social media contents with various modalities like text, images, videos. We also constrain our analysis of natural language processing, machine learning, and deep learning techniques. This literature survey also covers the research work related to issues close to propaganda such as radicalization, fake news, and news trustworthiness and misinformation on social media platforms.

3. FRAMEWORK FOR LITERATURE REVIEW

The general framework followed for the literature review in this work is illustrated in Figure 2.

A comprehensive list of keywords related to the topic was created. We used several keywords strings for example, 'propaganda and social media', 'propaganda analysis and social media or social bots or social intelligence' along with secondary keywords such as politics, terrorism, bias news, etc. An extensive search on various scholarly articles platforms such as Scopus, Google Scholar, IEEE, etc were conducted. The returned scholarly articles as a result of the search were studied. For details, the related articles and cited articles were also looked into. Further, the articles were categorized based on the topics into the use of social bots/botnets for propaganda spread, extremism propaganda, and political propaganda.

4. SURVEY OF RESEARCH PUBLICATIONS

The seminal work in the direction of propaganda identification was done by the Institute of Propaganda Analysis from 1937 to 1942. The IPA was established to provide guidelines to help public educated discussions on topics and inspire rational thinking. Propaganda has been widely studied in sociology, political science, journalism, and recently it has captured the attention of computational researchers due to its widespread in social media.

In recent times researchers have shown interest in studying bias and misinformation in news and social media by challenging the truth worthiness of news sources and social media as well as studying

Figure 1. Scope and Focus of Literature Review

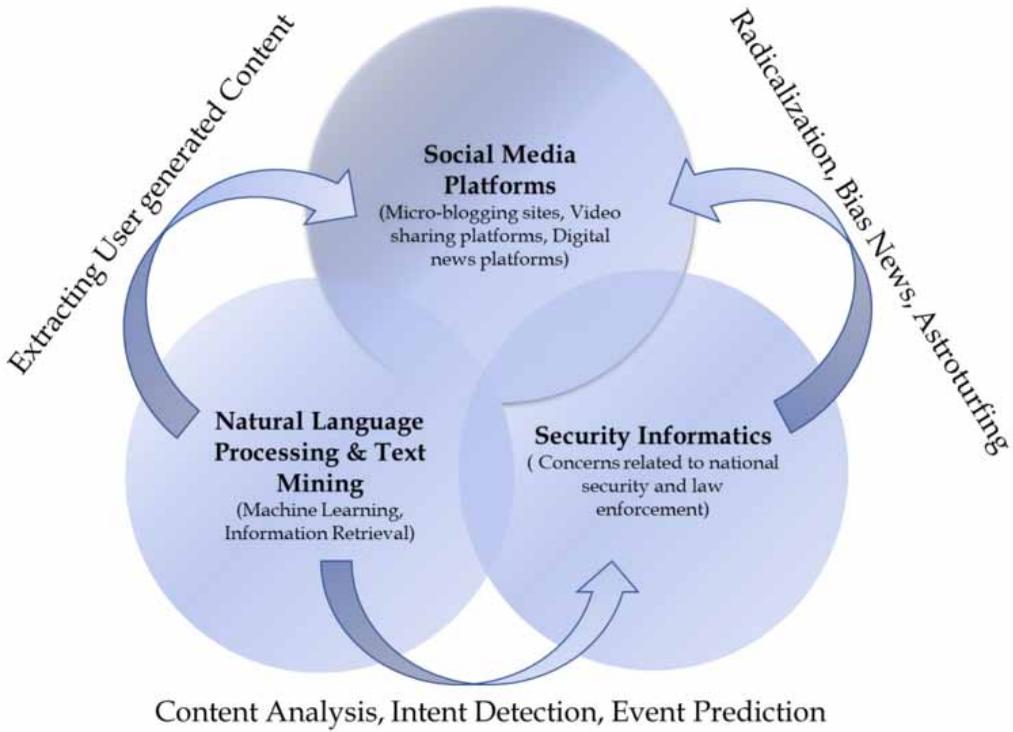
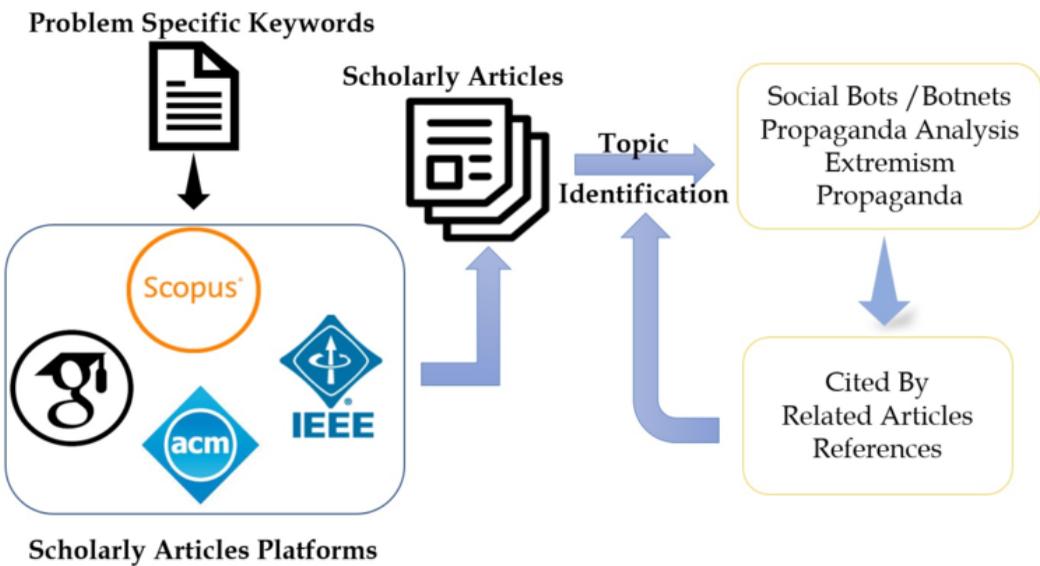


Figure 2. Framework for Literature Review Followed



credibility, influence, and bias. A lot of work is also been done in the direction of the identification of true and fake news.

Propaganda analysis even though close to fake news takes a different stance as it is not entirely fake or untrue. The use of several propaganda techniques makes it even more difficult to uncover. This literature survey focuses on four major aspects of propaganda as follows.

4.1 Propaganda Spread Using Bots

With the advent of social media, the use of digital propaganda has seen a rise. Digital propaganda implies the use of computer algorithms and automation for rallying and influencing public opinions. According to (Abokhodair et al., 2015), social bots can unsettle or impact public opinion by many means like spam hashtags, scam Twitter users, and astroturfing. In this study, the authors focused on Syrian social bots and identified the distinguishing features of botnets from normal users. (Agarwal et al., 2017) focused their work on two events in the Russian region (2014 Crimean Water Crisis and the 2015 Dragoon Ride Exercise) to study how botnets were highly used to spread the propaganda. The authors uncovered that the behavior of these coordinated and automated social bots is becoming increasingly sophisticated. In the work of (Neyazi, 2020) the Twitter posts on the Uri attack, Surgical Strike, and Indian television viewership data were studied to understand the use of political bots in the Indian political landscape. The author after using statistical analysis of automated tweets concludes that the use of social bots has fuelled the polarization of politics in India and at times it infiltrates even the mainstream media. The author also highlights the immediate need for computational techniques to uncover political bots. Instead of focusing on the use of bots, (Wagner et al., 2012) focused on the identification of susceptible users and their level of susceptibility to falling prey to social bots. Their findings suggest that active users are also more vulnerable to social bots attacks.

4.2 Terrorist Propaganda Analysis

In the recent past, it is widely evident that terrorist organizations use social media for radicalization and recruitment. The terrorist propaganda is spread through several mechanisms using soft propaganda, hard propaganda, and symbolic propaganda. According to (Hashemi & Hall, 2019) violent extremist organizations (VEOs) pose a dangerous threat to national security and citizens through the spread of their extremist and radical propaganda spread. The authors focused on the detection and intent-based classification of online visual propaganda by Islamic State (ISIS). Using Convolutional Neural Networks authors have shown the generalization accuracy of 97.02% and an F1 score of 97.89% for the detection of visual propaganda. In the work of (Ferrara, 2017) good practices for data collection, validation, and analysis to study online radicalization were discussed. In the research works of (Jun et al., 2018) (Johnston & Weiss, 2018) and (Heidarysafa et al., 2019) the online magazines and websites by terrorist organizations were studied to understand how the emotional appeals and specific targets like women and Sunni sect are exploited for propaganda spread. These works have also shown that natural language processing techniques like topic modeling and word features can help uncover the use of propaganda proliferation. In the work of (Baines & O'Shaughnessy, 2014) the propaganda and persuasion techniques employed by terrorist organizations over 10 years were studied. The author emphasizes the need for more assessment of propaganda to recognize the countermeasures.

4.3 Political Propaganda in Social Media

After the allegations of Russian meddling and influencing the 2016 U.S. presidential elections, the discussions about the use of social media for political influence and propaganda has gained more attention. The authors (Kellner et al., 2019) studied the German Federal Elections for evidence of social bots used for political opinion swinging and automated propaganda. The authors concluded that the political landscape is greatly influenced by the propaganda on social media. The findings of this work are similar to (Neyazi, 2020) which also indicates the use of social bots in Indian political terrain.

4.4 Propaganda Through Biased News

Social media due to its dynamic nature sees faster dissemination of information and news as compared to traditional news media. This has also given rise to fake news, misinformation, rumors, and hoax. Many efforts are reported to tackle the misinformation and fact-checking organizations like FactCheck, Snopes, and Politifact have emerged.

TSHP-17 (trusted, satire, hoax, and propaganda 2017 corpus) (Rashkin et al., 2017) and Hyperpartisan News Dataset from SemEval-2019 (Saleh et al., 2019) are the prominent datasets used for the analysis of news articles. Some studies (Popat et al., 2019), (Wang et al., 2018), (Qazvinian et al., 2011), (Baly et al., 2020), (Kwon et al., 2013) have worked in the direction of rumor detection and fact-checking whereas (Saleh et al., 2019), (Barrón-Cedeño, Jaradat, et al., 2019), (Rashkin et al., 2017), (da San Martino et al., 2020), (Baisa et al., 2019) have worked to uncover the political propaganda in news articles.

5. DESCRIPTION AND CATEGORIZATION OF ARTICLES BASED UPON META-ANALYSIS

This section of the literature review elaborates on the categorization of previous literature work. After a critical analysis of each research paper, we listed the components to demonstrate the statistics. Further, these components were divided into sub-components based on the properties as described in Table 1. The aspects considered are as follows:

1. Modality
2. Data Type

Table 1. List of components and their properties used for meta-analysis of research articles

Component	Label	Property	Description
Modality	M1	News Websites	News Websites on the Internet
	M2	Online Magazines	Privately and publicly accessible magazines on the web
	M3	Social Media	Online social networking websites like Facebook, Twitter, YouTube, etc.
Data Type	D1	Text	Documents comprising of only textual content
	D2	Text + Multimedia	Documents containing both text and multimedia
Evaluation Parameters	E1	Precision	Evaluation measure for the exactness of the technique
	E2	Recall	Evaluation measure for completeness of the technique
	E3	F1 Measure	Evaluation measure of a test's accuracy [$F1 = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$]
	E4	Accuracy	Evaluation measure for the correctness of the technique
Language	L1	English	Documents with English matter
	L2	Arabic	Documents with Arabic matter
	L3	Multilingual	Documents with multiple languages
	L4	Other	Documents with matter other than English and Arabic
Type of Propaganda addressed	P1	Political	Propaganda for political purpose or astroturfing
	P2	Terrorist / Radicalization	Terrorism or radicalization propaganda
	P3	News / Media Bias	Propaganda spread through the use of bias/fake news

Table 2. Summary of existing literature

Paper	Year	Modality			Data Type		Evaluation				Language				Propaganda Type		
		M1	M2	M3	D1	D2	E1	E2	E3	E4	L1	L2	L3	L4	P1	P2	P3
(Araque & Iglesias, 2020)	2020		✓		✓				✓		✓					✓	
(Stukal et al., 2019)	2019			✓	✓		✓	✓					✓	✓			
(Barrón-Cedeño, Jaradat, et al., 2019)	2019	✓			✓				✓	✓	✓						✓
(da San Martino et al., 2020)	2020	✓			✓		✓	✓	✓		✓						✓
(Gavrilenko et al., 2020)	2020	✓			✓					✓	✓						✓
(Gupta et al., 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Saleh et al., 2019)	2019	✓			✓		✓	✓	✓	✓	✓						✓
(aggarwal & Sadana, 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Alhindi et al., 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Hua, 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Heidarysafa et al., 2019)	2019		✓		✓						✓	✓				✓	
(Ferreira Cruz et al., 2019)	2019	✓			✓		✓	✓	✓	✓	✓						✓
(Baisa et al., 2019)	2019		✓		✓								✓				
(Li et al., 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Mapes et al., 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Al-Omari et al., 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Yoosuf & Yang, 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Hou & Chen, 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Tayyar Madabushi et al., 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Vlad et al., 2019)	2019	✓			✓		✓	✓	✓		✓						✓
(Baisa et al., 2019)	2019	✓			✓				✓	✓				✓	✓		
(Nizzoli et al., 2019)	2019			✓	✓				✓		✓					✓	
(Hashemi & Hall, 2019)	2019			✓		✓	✓	✓	✓		✓	✓				✓	
(Kellner et al., 2019)	2019			✓	✓								✓	✓			
(Beğenilmiş & Uskudarli, 2018)	2018			✓	✓				✓	✓	✓					✓	
(Johnston & Weiss, 2018)	2017	✓			✓		✓	✓	✓	✓			✓			✓	
(Recalde et al., 2019)	2017			✓	✓						✓					✓	
(Agarwal et al., 2017)	2017			✓	✓						✓					✓	

3. Evaluation Parameters
4. Language
5. Type of propaganda addressed

5.1 Machine Learning and Deep Learning Techniques Used

Figure 3 and Figure 4 illustrate the machine learning and deep learning techniques used in the existing research. A range of supervised machine learning algorithms such as Support vector machines, Logistic regression, Random forest, etc. have been used for propaganda detection by the authors. (Kellner et al., 2019) have demonstrated the use of an ensemble of machine learning techniques for the detection of automated propaganda in German Federal elections. Although most authors have

Figure 3. Machine Learning Techniques applied by researchers in active research

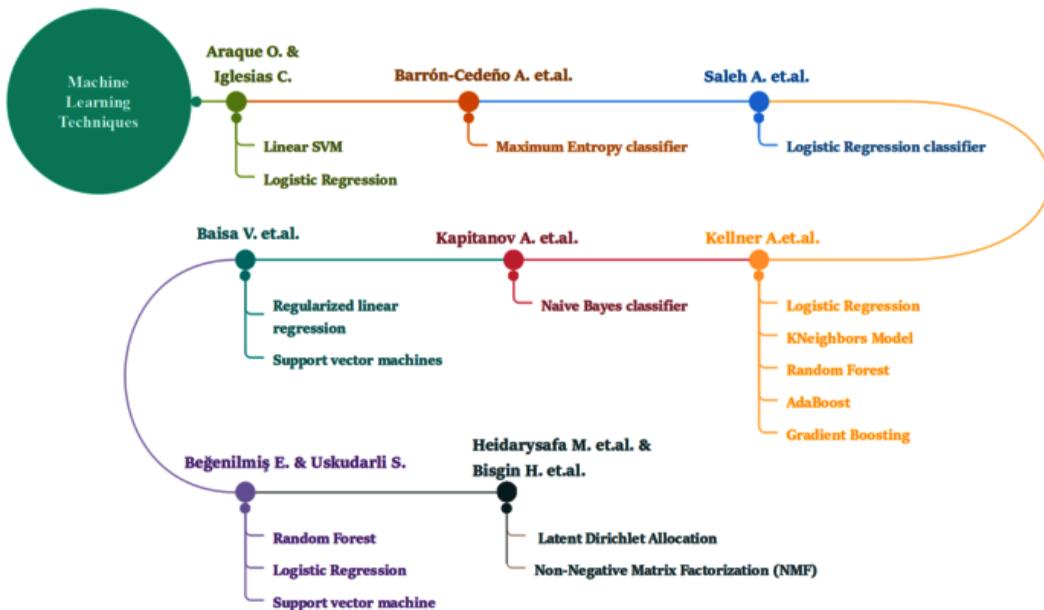


Figure 4. Deep Learning Techniques applied by researchers in active research



resorted to supervised techniques, authors (Heidarysafa et al., 2019) and (Bisgin et al., 2019). et.al. have employed the unsupervised technique of topic modeling to uncover the terrorist propaganda elements used to influence the emotions of readers of extremist magazines.

The recent developments in deep learning have shown impressive results for language processing. Researchers have demonstrated the effectiveness of deep learning models for the task of propaganda detection. (Al-Omari et al., 2019) and (Ferreira Cruz et al., 2019) have used ensemble deep learning techniques whereas other authors have relied on conventional deep learning techniques such as Convolutional Neural Network (CNN) and Long-Short-Term-Memory (LSTM).

5.2 Features Used in Natural Language Processing Techniques

In combination with machine learning and deep learning techniques as explained above, several authors have also used textual and language features. These features are used for training the classifiers and play a crucial role in the classification process. We list important features utilized by the researchers in Table 3.

BERT (Bidirectional Encoder Representations from Transformers) is a transformer-based NLP model. BERT has shown state-of-the-art results for NLP tasks outperforming other language models like word2vec and Glove. At 2nd Workshop on NLP for Internet Freedom (NLP4IF): Censorship, Disinformation, and Propaganda, several authors proposed the use of BERT for sentence-level and fragment-level propaganda detection in news articles. We have listed some of the variations of BERT-based models utilized by the authors. Among the two tasks, the sentence-level task has been proved to be easier than the fragment level. The collection of work under the NLP4IF workshop also proved the overall effectiveness of BERT-based models for discovering propaganda contents.

5.3 Data Sources

We analyzed around 40 research articles to comprehend the use of various social media platforms explored in the existing literature. We observed that News articles and news websites are the most explored media followed by micro-blogging site Twitter. It also emphasizes the fact that news articles are being widely used for propaganda spread through disinformation, bias news, and fake news. Due to its wide reach and easy access through APIs, Twitter also seems to be a choice of data source for many researchers.

Stukal et al., (2019); Gavrilenko et al., (2020); Kellner et al., (2019) studied Twitter bots in a political context whereas Nizzoli et al., (2019) and Hashemi & Hall, (2019) focused on the extremist

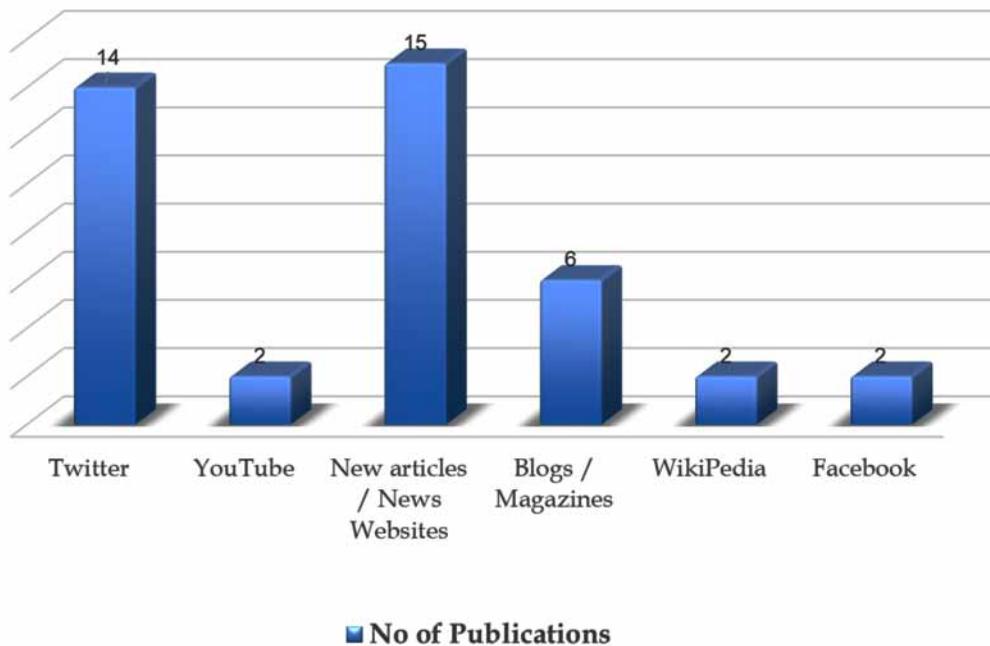
Table 3. Features used by authors in the existing literature

Author	Features
(Araque & Iglesias, 2020)	EmoFeat- Emotion Word Features
(Stukal et al., 2019)	Textual Feature - Unigrams, Bigrams, hashtags, mentions, and links
(Gupta et al., 2019)	Linguistic, Layout and Topical features
(Aggarwal & Sadana, 2019)	Context-sensitive features produced by ELMo
(Alhindi et al., 2019)	Linguistic Inquiry and Word Count (LIWC) features, punctuation features
(Beğenilmiş & Uskudarli, 2019)	User Features, Temporal Features,
(Ferreira Cruz et al., 2019)	Linguistic Style and Complexity features
(Al-Omari et al., 2019)	TweetToEmbeddings, TweetToInputLexicon, TweetToLexicon, TweetToSentiStrength
(Kellner et al., 2019)	Metadata-based, Text-based, Time-based, and User-based features

Table 4. BERT-based models used in existing literature for propaganda detection

Author	Variation of BERT - Models used
(Alhindi et al., 2019)	BERT, BiLSTM-CRF
(Mapes et al., 2019)	BERT-based attention transformer model
(Fadel et al., 2019)	BERT (Cased/ Uncased), Universal Sentence Encoder (USE) Model
(Hou & Chen, 2019)	Context-dependent BERT model
(Tayyar Madabushi et al., 2019)	Cost-Sensitive BERT
(Vlad et al., 2019)	BERT, BiLSTM Capsule

Figure 5. Social Media Platforms used as Data Sources



propaganda on Twitter. The online magazine by the terrorist organizations’ ISIS, Al-Qaeda such as Dabiq and Rumiya were the data source for the studies conducted by Araque & Iglesias, (2020); Bisgin et al., (2019); Heidarysafa et al., (2019), etc.

Even though Facebook is the most popular social networking platform, it is observed that very few researchers have explored Facebook data for their studies. It is also observed that most of the research articles have addressed textual contents on social media avoiding the multi-media contents on the web.

6. CRITICAL ANALYSIS OF EXISTING LITERATURE

In this section, we discuss the existing literature critically based on the type of propaganda addressed, techniques used, trends seen in the field of research, and limitations observed.

6.1 Type of Propaganda Addressed

In this work, we focus on three major categories of propaganda spread political, terrorism, and propaganda through disinformation. Our analysis reveals that after the disclosure of social bots uses in the 2016 US presidential elections, the issue of propaganda through social bots gained attention. The political propaganda studies throw light on the use of social bots in the manipulation of public opinion during elections. The radicalization and terrorism propaganda spread by extremist organizations pose a great deal of threat on social media. Due to its open access, social media has become a breeding ground for terrorist, radical, and extremist propaganda. Researchers have proposed solutions to detect such propaganda, but the onus falls on the social media platforms organizations to mitigate these propaganda contents. The propaganda has also infiltrated mainstream media and comes in the form of bias news and disinformation. Recently, the studies in this direction have gained traction and have shown promising results to unveil the various propaganda techniques used in news.

6.2 Techniques

Researchers highly use machine learning and deep learning techniques for the analysis and detection of propaganda. To recognize the propaganda contents natural language processing and several language models have been explored in the existing literature. A major line of work also addresses the use of social bots or botnets for propaganda proliferation. To expose the use of social botnets researchers have demonstrated the effective usage of network analysis techniques.

7. CLOSING REMARKS AND CONCLUSION

Applying social media intelligence for analysis and prediction of propaganda has attracted the interest of many researchers in the recent past. It is observed that the emergence of fake news has invited the attention towards the use of news for propaganda spread more prominently during the previous three years. Another leading area where propaganda is studied widely is terrorist propaganda followed by political propaganda.

This literature review also reveals that social media platforms like Twitter, blogs, YouTube are the breeding grounds for propaganda and misinformation. Even though being the most popular social networking platform, the use of Facebook has not been studied in this context. The strict access regulations by Facebook for the access of third-party APIs do not provide wide access to data to the researchers which might be attributed to this lack of work in this context. The video-sharing platform YouTube has been analyzed for visual propaganda by few studies yet needs more attention. By far the micro-blogging site Twitter has been leading the pack as the source of social media data for propaganda detection studies. The literature survey also unveils that due to its features like retweets and follower-followed relations Twitter has been instrumental in propaganda facilitation and dissemination.

This literature review shows that several machine learning, information retrieval techniques are used for the analysis of propaganda. It is observed that Graph modeling techniques and Network analysis methods have been used to illustrate the diffusion of propaganda through social networks. Among the several techniques Naïve Bayes, Support Vector Machine, Decision Tree, KNN (K Nearest Neighbour), and Link analysis remain the most explored techniques. To examine the efficacy of the results accuracy and F1 score remains the most used evaluation metrics.

Even though propaganda spread is a global phenomenon, more than 90% of studies can be attributed to the English language while Arabic remains the focus in second place. The literature survey also uncovers the urgent need for studies in the multi-lingual text. Also, it is observed that most research works focus on a few specific events, country, or region for the propaganda analysis where there is a need for a generic framework which will go beyond these specificities.

FUNDING AGENCY

The publisher has waived the Open Access Processing fee for this article.

REFERENCES

- Abokhodair, N., Yoo, D., & McDonald, D. W. (2015). Dissecting a social Botnet: Growth, content and influence in Twitter. *CSCW 2015 - Proceedings of the 2015 ACM International Conference on Computer-Supported Cooperative Work and Social Computing*, 839–851. doi:10.1145/2675133.2675208
- Agarwal, N., Al-Khateeb, S., Galeano, R., & Goolsby, R. (2017). Examining the use of botnets and their evolution in propaganda dissemination. *Defence Strategic Communications*, 2(1), 87–112. doi:10.30966/2018.riga.2.4
- Aggarwal, K., & Sadana, A. (2019). *NSIT@NLP4IF-2019: Propaganda Detection from News Articles using Transfer Learning*. 10.18653/v1/D19-5021
- Al-Omari, H., Abdullah, M., AlTiti, O., & Shaikh, S. (2019). *JUSTDeep at NLP4IF 2019 Task 1: Propaganda Detection using Ensemble Deep Learning Models*. 10.18653/v1/D19-5016
- Alhindi, T., Pfeiffer, J., & Muresan, S. (2019). *Fine-Tuned Neural Models for Propaganda Detection at the Sentence and Fragment levels*. 10.18653/v1/D19-5013
- Araque, O., & Iglesias, C. A. (2020). An approach for radicalization detection based on emotion signals and semantic similarity. *IEEE Access: Practical Innovations, Open Solutions*, 8, 17877–17891. doi:10.1109/ACCESS.2020.2967219
- Baines, P. R., & O'Shaughnessy, N. J. (2014). Al-Qaeda messaging evolution and positioning, 1998–2008: Propaganda analysis revisited. *Public Relations Inquiry*, 3(2), 163–191. doi:10.1177/2046147X14536723
- Baisa, V., Herman, O., & Horák, A. (2019). Benchmark dataset for propaganda detection in Czech newspaper texts. *International Conference Recent Advances in Natural Language Processing, RANLP*, 77–83. doi:10.26615/978-954-452-056-4_010
- Baly, R., Karadzhov, G., Alexandrov, D., Glass, J., & Nakov, P. (2020). Predicting factuality of reporting and bias of news media sources. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018*, 3528–3539. doi:10.18653/v1/D18-1389
- Barrón-Cedeño, A., Da San Martino, G., Jaradat, I., & Nakov, P. (2019). Proppy: A System to Unmask Propaganda in Online News. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33, 9847–9848. doi:10.1609/aaai.v33i01.33019847
- Barrón-Cedeño, A., Jaradat, I., Da San Martino, G., & Nakov, P. (2019). Proppy: Organizing the news based on their propagandistic content. *Information Processing & Management*, 56(5), 1849–1864. doi:10.1016/j.ipm.2019.03.005
- Beğenilmiş, E., & Uskudarli, S. (2018). Organized behavior classification of tweet sets using supervised learning methods. *ACM International Conference Proceeding Series*. doi:10.1145/3227609.3227665
- da San Martino, G., Yu, S., Barrón-Cedeño, A., Petrov, R., & Nakov, P. (2020). Fine-grained analysis of propaganda in news articles. *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, 4, 5636–5646.
- Derbas, N., Dusserre, E., Padró, M., & Segond, F. (2020). Eventfully Safapp: Hybrid approach to event detection for social media mining. *Journal of Ambient Intelligence and Humanized Computing*, 11(1), 87–95. doi:10.1007/s12652-018-1078-7
- Fadel, A., Tuffaha, I., & Al-Ayyoub, M. (2019). *Pretrained Ensemble Learning for Fine-Grained Propaganda Detection*. Advance online publication. doi:10.18653/v1/D19-5020
- Ferrara, E. (2017). Contagion dynamics of extremist propaganda in social networks. *Information Sciences*, 418–419, 1–12. doi:10.1016/j.ins.2017.07.030
- Ferreira Cruz, A., Rocha, G., & Lopes Cardoso, H. (2019). *On Sentence Representations for Propaganda Detection: From Handcrafted Features to Word Embeddings*. 10.18653/v1/D19-5015
- Gavrilenko, O., Oliinyk, Y., & Khanko, H. (2020). Analysis of Propaganda Elements Detecting Algorithms in Text Data. In *Advances in Intelligent Systems and Computing* (Vol. 938). Springer International Publishing. doi:10.1007/978-3-030-16621-2_41

- Gupta, P., Saxena, K., Yaseen, U., Runkler, T., & Schütze, H. (2019). *Neural Architectures for Fine-Grained Propaganda Detection in News*. 10.18653/v1/D19-5012
- Hashemi, M., & Hall, M. (2019). Detecting and classifying online dark visual propaganda. *Image and Vision Computing*, 89, 95–105. doi:10.1016/j.imavis.2019.06.001
- Heidarysafa, M., Kowsari, K., Odukoya, T., Potter, P., Barnes, L. E., & Brown, D. E. (2019). *Women in ISIS Propaganda: A Natural Language Processing Analysis of Topics and Emotions in a Comparison with Mainstream Religious Group*. <https://arxiv.org/abs/1912.03804>
- Hou, W., & Chen, Y. (2019). *Sentence-Level Propaganda Detection Using BERT with Context-Dependent Input Pairs*. Academic Press.
- Hua, Y. (2019). *Understanding BERT performance in propaganda analysis*. 10.18653/v1/D19-5019
- Johnston, A. H., & Weiss, G. M. (2018). Identifying Sunni extremist propaganda with deep learning. *2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings*, 1–6. doi:10.1109/SSCI.2017.8280944
- Jun, S. P., Yoo, H. S., & Choi, S. (2018). Ten years of research change using Google Trends: From the perspective of big data utilization and applications. *Technological Forecasting and Social Change*, 130(November), 69–87. 10.1016/j.techfore.2017.11.009
- Kapitanov, A. I., Kapitanova, I. I., Troyanovskiy, V. M., Shangin, V. F., & Krylikov, N. O. (2018). Approach to automatic identification of terrorist and radical content in social networks messages. *Proceedings of the 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2018*, 1517–1520. doi:10.1109/ElConRus.2018.8317386
- Kellner, A., Rangosch, L., Wressnegger, C., & Rieck, K. (2019). *Political Elections Under (Social) Fire? Analysis and Detection of Propaganda on Twitter*. <https://arxiv.org/abs/1912.04143>
- Kwon, S., Cha, M., Jung, K., Chen, W., & Wang, Y. (2013). Prominent features of rumor propagation in online social media. *Proceedings - IEEE International Conference on Data Mining, ICDM*, 1103–1108. doi:10.1109/ICDM.2013.61
- Li, J., Ye, Z., & Xiao, L. (2019). *Detection of Propaganda Using Logistic Regression*. 10.18653/v1/D19-5017
- Mapes, N., White, A., Medury, R., & Dua, S. (2019). *Divisive Language and Propaganda Detection using Multi-head Attention Transformers with Deep Learning BERT-based Language Models for Binary Classification*. 10.18653/v1/D19-5014
- Neyazi, T. A. (2020). Digital propaganda, political bots and polarized politics in India. *Asian Journal of Communication*, 30(1), 39–57. doi:10.1080/01292986.2019.1699938
- Nizzoli, L., Avvenuti, M., Cresci, S., & Tesconi, M. (2019). Extremist propaganda tweet classification with deep learning in realistic scenarios. *WebSci 2019 - Proceedings of the 11th ACM Conference on Web Science*, 203–204. doi:10.1145/3292522.3326050
- Popat, K., Mukherjee, S., Strötgen, J., & Weikum, G. (2019). Where the truth lies: Explaining the credibility of emerging claims on the web and social media. *26th International World Wide Web Conference 2017, WWW 2017 Companion*, 1003–1012. doi:10.1145/3041021.3055133
- Qazvinian, V., Rosengren, E., Radev, D. R., & Mei, Q. (2011). Rumor has it Identifying Misinformation in Microblogs. *Conference on Empirical Methods in Natural Language Processing*, 1589–1599.
- Rashkin, H., Choi, E., Jang, J. Y., Volkova, S., & Choi, Y. (2017). Truth of varying shades: Analyzing language in fake news and political fact-checking. *EMNLP 2017 - Conference on Empirical Methods in Natural Language Processing, Proceedings*, 2931–2937. doi:10.18653/v1/D17-1317
- Recalde, L., Mendieta, J., Boratto, L., Terán, L., Vaca, C., & Baquerizo, G. (2019). Who You Should Not Follow: Extracting Word Embeddings from Tweets to Identify Groups of Interest and Hijackers in Demonstrations. *IEEE Transactions on Emerging Topics in Computing*, 7(2), 206–217. doi:10.1109/TETC.2017.2669404
- Saleh, A., Baly, R., Barrón-Cedeño, A., Da San Martino, G., Mohtarami, M., Nakov, P., & Glass, J. (2019). *Team QCRI-MIT at SemEval-2019 Task 4: Propaganda Analysis Meets Hyperpartisan News Detection*. 10.18653/v1/S19-2182

- Stukal, D., Sanovich, S., Tucker, J. A., & Bonneau, R. (2019). For Whom the Bot Tolls: A Neural Networks Approach to Measuring Political Orientation of Twitter Bots in Russia. *SAGE Open*, 9(2). Advance online publication. doi:10.1177/2158244019827715
- Tayyar Madabushi, H., Kochkina, E., & Castelle, M. (2019). *Cost-Sensitive BERT for Generalisable Sentence Classification on Imbalanced Data*. 10.18653/v1/D19-5018
- Varol, O., Ferrara, E., Menczer, F., & Flammini, A. (2017). Early detection of promoted campaigns on social media. *EPJ Data Science*, 6(1), 13. Advance online publication. doi:10.1140/epjds/s13688-017-0111-y
- Vlad, G.-A., Tanase, M.-A., Onose, C., & Cercel, D.-C. (2019). *Sentence-Level Propaganda Detection in News Articles with Transfer Learning and BERT-BiLSTM-Capsule Model*. 10.18653/v1/D19-5022
- Volkova, S., Shaffer, K., Jang, J. Y., & Hodas, N. (2017). Separating facts from fiction: Linguistic models to classify suspicious and trusted news posts on Twitter. *ACL 2017 - 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)*, 2, 647–653. doi:10.18653/v1/P17-2102
- Wagner, C., Mitter, S., Körner, C., & Strohmaier, M. (2012). When social bots attack: Modeling susceptibility of users in online social networks. *CEUR Workshop Proceedings*, 838, 41–48.
- Wang, L., Wang, Y., De Melo, G., & Weikum, G. (2018). Five shades of untruth: Finer-grained classification of fake news. *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018*, 593–594. doi:10.1109/ASONAM.2018.8508256
- Yoosuf, S., & Yang, Y. (2019). *Fine-Grained Propaganda Detection with Fine-Tuned BERT*. Advance online publication. doi:10.18653/v1/D19-5011

Deptii Chaudhari is currently a research scholar with Symbiosis Institute of Technology, Symbiosis International (Deemed University). Her primary research interest includes Natural Language Processing, Machine Learning & Deep Learning, and Information Retrieval.

Ambika Pawar has extensive experience in Security & Privacy, Algorithms, Cloud Computing, Machine Learning & Deep Learning and publications in reputed journals and conferences.