

DOI: <http://dx.doi.org/10.21123/bsj.2021.18.1.0199>

## Detecting Textual Propaganda Using Machine Learning Techniques

*Akib Mohi Ud Din Khanday*\*

*Qamar Rayees Khan*

*Syed Tanzeel Rabani*

Department of Computer Sciences, Baba Ghulam Shah Badshah University -185234, Rajouri, J&K, India.

\*Corresponding author: [akibkhanday@bgsbu.ac.in](mailto:akibkhanday@bgsbu.ac.in), [rayees.dcs@gmail.com](mailto:rayees.dcs@gmail.com), [syedtanzeel@bgsbu.ac.in](mailto:syedtanzeel@bgsbu.ac.in)

\*ORCID ID: <https://orcid.org/0000-0001-6804-4905>, <https://orcid.org/0000-0002-7628-173X>, <https://orcid.org/0000-0001-7596-5964>

Received 13/5/2020, Accepted 8/10/2020, Published Online First 6/12/2020, Published 1/3/2021



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

### Abstract:

Social Networking has dominated the whole world by providing a platform of information dissemination. Usually people share information without knowing its truthfulness. Nowadays Social Networks are used for gaining influence in many fields like in elections, advertisements etc. It is not surprising that social media has become a weapon for manipulating sentiments by spreading disinformation. Propaganda is one of the systematic and deliberate attempts used for influencing people for the political, religious gains. In this research paper, efforts were made to classify **Propagandist text from Non-Propagandist text using supervised machine learning algorithms**. Data was collected from the news sources from July 2018-August 2018. After annotating the text, feature engineering is performed using techniques like term frequency/inverse document frequency (TF/IDF) and Bag of words (BOW). The relevant features are supplied to support vector machine (SVM) and Multinomial Naïve Bayesian (MNB) classifiers. The fine tuning of SVM is being done by taking kernel Linear, Poly and RBF. SVM showed better results than MNB by having precision of 70%, recall of 76.5%, F1 Score of 69.5% and overall Accuracy of 69.2%.

**Key words:** Bag of Words, Disinformation, Propaganda, Social Networks, Term Frequency.

### Introduction:

In this era of technology, computer science plays an essential role in providing solutions to almost all the emerging fields. With the advent of the internet in the 1970s, there has been a revolutionary improvement in computer science and has found its place in multidisciplinary subjects like remote sensing, technical diagnosis, traffic control systems, criminology, medical imaging, image processing, data mining, and automatic surveillance. Due to these applications, there is tremendous growth in the number of hardware and software products in the market(1). Nowadays data analytics is the major research area which discovers the patterns from large data sets, it is integrated with various fields like bioinformatics, Natural Language Processing (NLP), Machine Learning etc.(2). In Data mining important information is mined from the text, image and video etc. The data mining can perform the descriptive or predictive task, for characterizing data there are descriptive tasks and predictive tasks make the projection of the future on the basis of old data(3). Clustering, Correlation, Pattern Finding are some of the tasks in data mining(4). Analyzing Online Social Networks is a

challenging process, due to its enormous usage, variety, large volume, veracity, and real-time data. Online Social Networks (OSN) are communicating via computer-mediated tools which assist the formation and distribution of information, concepts, business welfares and new practices of communication via virtual communities and links (4). OSNs are gradually used for purposes like advertising, brand construction, political campaigning, and recommender systems, customer response, sentiment analysis(6,7) etc. Social network saturation worldwide is increasing day by day, in 2017, about 71 per cent of internet consumers were using social network and the statistics are estimated to raise in 2019 and is assessed that around 2.77 billion social media users will be in 2019, rising from 2.46 billion in 2017<sup>1</sup> the statistics is shown in Fig.1. The modern age is the age of technology, the advancement in information and communication technology (ICT) changed the classical process of sending and receiving the

<sup>1</sup> <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>

information. Everybody in this universe wants to be updated irrespective of their behavioural differences. The adversarial use of social media is common for spreading deceptive/misleading information which is social, economic, and political

intimidation (8). Social media has been effectively used by fear mongers for triggering mass hysteria and panic(9). There are three types of attacks that took place in cyber network operations – physical, syntactic and semantic attacks.

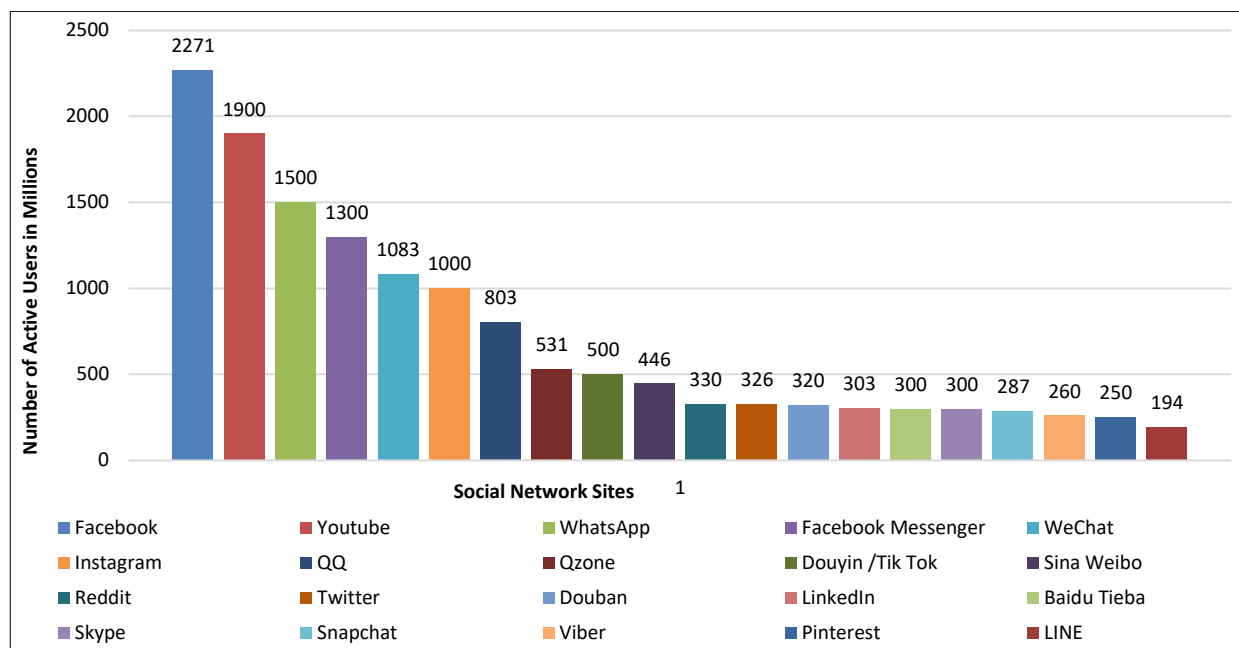


Figure 1. Number of Social Networks World Wide as of January 2019(2)

### Physical attacks

Physical attacks are attacks that affect the hardware of the system and occur at the physical level. By removing the hardware, the system disappears as well, but physical attacks as such need no elaboration here due to our diversity in the research problem.

### Syntactic attacks

Syntactic attacks occur due to the technologies, this attack occurs at the syntactic level which contains instructions for designers and users which they give to a machine and contains protocols through which machines interact with one another. Hacking is one type of syntactic attacks whose main aim is to steal data. A hacker can alter everything regarding a system arbitrarily. Since our work is regarding text analytics there is no role of ours in this type of attacks.

### Semantic Attacks

In semantic attacks human-computer interfaces are attacked. The effects of semantic attacks are not visible like other types of attacks (physical, syntactic). Semantic attacks are the most dangerous attacks which change the information content or the meaning of information(10). Semantic attacks are divided into many categories viz Overt attack (include Phishing, Spam etc.) and Covert attack. Our focus would be on Covert attacks i.e. misinformation, disinformation and

propaganda (11). This domain is considered as an emerging field of research and is being critically studied by researchers. It has been suggested by many authors to use the social computing properties of the users on online social networks for determining the credibility of the information. Social networks are being used enormously for spreading fabricated information. The spread of information becomes easier for users with the increase in usage of OSN's. Information can be truthful or dubious, the doubtful information can be shared deliberately or un-deliberately. The doubtful information can be categorized as misinformation and disinformation. Misinformation is that information where the user does not know the truthfulness of information that is being spread whereas disinformation is that information in which the user deliberately shares/spread false/true information(12). Disinformation usually occurs in politics, health, finance, technology etc. Orchestrated Astroturf is used for manipulating political conversations, even during election times(13)(14). The compromised accounts are being used for spreading disinformation and these accounts may also be used for spreading propaganda. Propaganda comes under the type of disinformation which is defined as the systematic and deliberate process to shape opinions, influence thoughts, and direct behaviour of a person for

achieving the desired intention of a propagandist. Propaganda identification in online social networks is the need of an hour and is mainly used for gaining the people's faith in some person or some community or party. Propaganda plays a major role in politics, specifically political propaganda has gained much interest of the researchers of the world. In the USA presidential election 2016, political propaganda has a major role in the winning of Donald Trump(15). Radical propaganda can be shared by posting messages, related to violence, religion, sectors and events(16). Due to the semantical nature of propagandistic text it is difficult to differentiate text into the binary class (propaganda and non-propaganda). Till now propaganda is being analysed manually. Machine learning has shown the promising results in every field of life, whether it is biological science or arts and humanities(17). The Prevalence of manipulation tactics on social media in 65 countries is shown in Fig. 2 .A framework is being used which will classify text into binary class using machine learning algorithms. This system will somehow help researchers to perform more work in this area. Hybrid feature engineering is used by combining TF/IDF and Bag of Words. The paper is organized under various sections starting from the literature survey from the already existing literature which is related to the proposed work followed by the proposed methodology using data extraction, pre-processing & feature extraction and classification algorithms. In corresponding sections, discussion on the findings and the experimental results and future scope of the study in this area of research are provided.

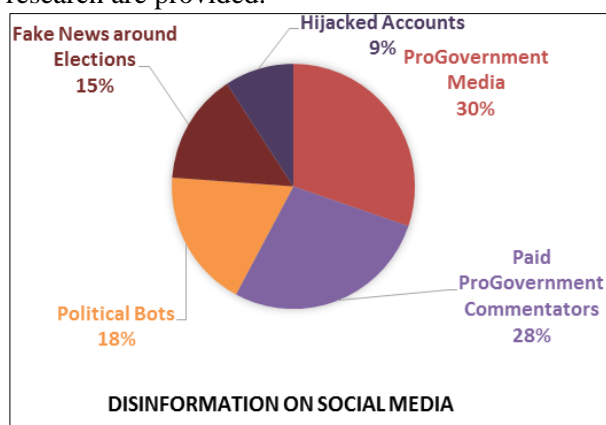


Figure 2. Prevalence of manipulation tactics on social media in 65 countries (2).

### Motivation

The modern age is the age of the internet especially social media and with the increasing use of online social network, there are more chances of exploitation of this platform. Social networks are used to share the information, but the information

shared or posted can be misinformation or disinformation. On social networks propaganda is being shared enormously and is usually used for political influences. By detecting the propaganda within time limits a lot of false information can be stopped which is being shared by hatemongers. Thus, it can help in many activities like in law enforcement, cyber-crime analysis etc.

The significant contribution of this work are as follows:

- A novel dataset of 6k records extracted/collected is built by annotating manually with the knowledge of 16 different techniques of propaganda.
- Hybrid approach of feature engineering is done by combining two techniques of feature extraction.
- To our knowledge this is a first work where accuracy of human annotated text for classification of propagandist text and non-propagandist text is significantly replicated by machine classification.

### Related Work:

The ever-growing attractiveness and beauty of using social networks directly or indirectly affects our daily life, by causing us to trust on other people's ideas & suggestions when making small or big decisions, from buying low-cost minor products and voting online in elections for the formation of the new government. It is not surprising that social media has become a weapon for manipulating sentiments by spreading disinformation as per the trend. False content and propaganda are extensively used on social media and must be identified and countered. H. Gao *et.al* (18) studied the characteristics of the malicious accounts wherein they focused on the messages that contain URL's in text form. With the increase in social media users, the events related to politics, government schemes are mostly discussed and the abuse in social media has become common. J. Ratkiewicz *et.al*(19) proposed a method for detecting and tracking political abuse in social networks. At the time of the election, the structure of the network shows some deflection from the normal structure. A. Halu *et.al*(20) studied the structure of social networks during the election campaign. Model for opinion dynamics was proposed that showed wide region parties survive by gathering a finite fraction of the votes at election time. N. Ramakrishnan *et.al* (21) analyzed and mined data regarding civil unrest in online social networks. They performed data extraction related to the specific event, identified the contributing factors, event evolution analysis. J. S. Liu *et.al* (22) identified the inner circles of

government political power holders underneath formal work relations and observed how the selected political groups form and change over time. They performed three main processes for this purpose network construction, community identification, and community evolution tracking. S. Y. Bhat *et.al* (23) devised a novel approach for detecting the stealthy accounts in online social networks. The various steps that were included are community detection at the node level, features, classification, and finding of stealthy sybils. F. M. F. Wong, *et.al* (24) computed political leaning from tweets, retweets, and re-tweeters. The systematic step is calculating Twitter users' behaviour. With the usage of social media, various events occur concurrently. F. Zarrinkalam *et.al* (25) identified events in the social network. Three models detecting unspecified events in social networks are used, for topic modelling LDA, for document clustering TF/IDF, and for feature clustering wavelet analysis is used. S. Lightfoot (26) studied the effect of social bots on politics (political propaganda through social bots). According to them, social bots have an only negative effect on politics. They found that social bots played a vital role in spreading fake news and accounts that continuously spread misinformation are significantly more likely to be bots. People with radical thoughts use social media only for spreading hate and fear. M. Ashcroft *et.al*(27) proposed a semantic graph-based approach for radicalization detection in social media. The pro-ISIS users tend to discuss more about religion, historical events and ethnicity while anti-ISIS users focus more on

politics, geographical locations and interventions against ISIS. They also shared the dubious information in the social network. K. K. Kumar *et.al* (28) detected misinformation on online social networks using cognitive psychology. The cognitive process consists consistency of message, the coherency of a message, the credibility of the source and general acceptability of message. The collected data sets used keywords Syria, Egypt etc.by using Twitter API. G. Mazzoleni *et.al* (29) identified the socially mediated type of populist communication profoundly affected by the specific nature of social media. They measured the degree of populist communicative ideology on the basis of appeal to the people, attacking the elite and by ostracizing the others.

### Proposed Methodology:

An elaborated analytical and experimental methodology is used based on the extensive literature survey and the corresponding discussion with the experts of the field on the problem. Based on the various factors that influence the performance of the model, real-time data was collected from the twitter and the data is submitted to the model and the final results are generated. The overall methodology comprises of five main steps:

- (i) Corpus/Data collection (ii) Data annotation (iii) Pre-processing and feature extraction (iv) Machine classification (v) Evaluation and validation.

The complete picture of the proposed methodology that is used in our work is depicted below in Fig.3.

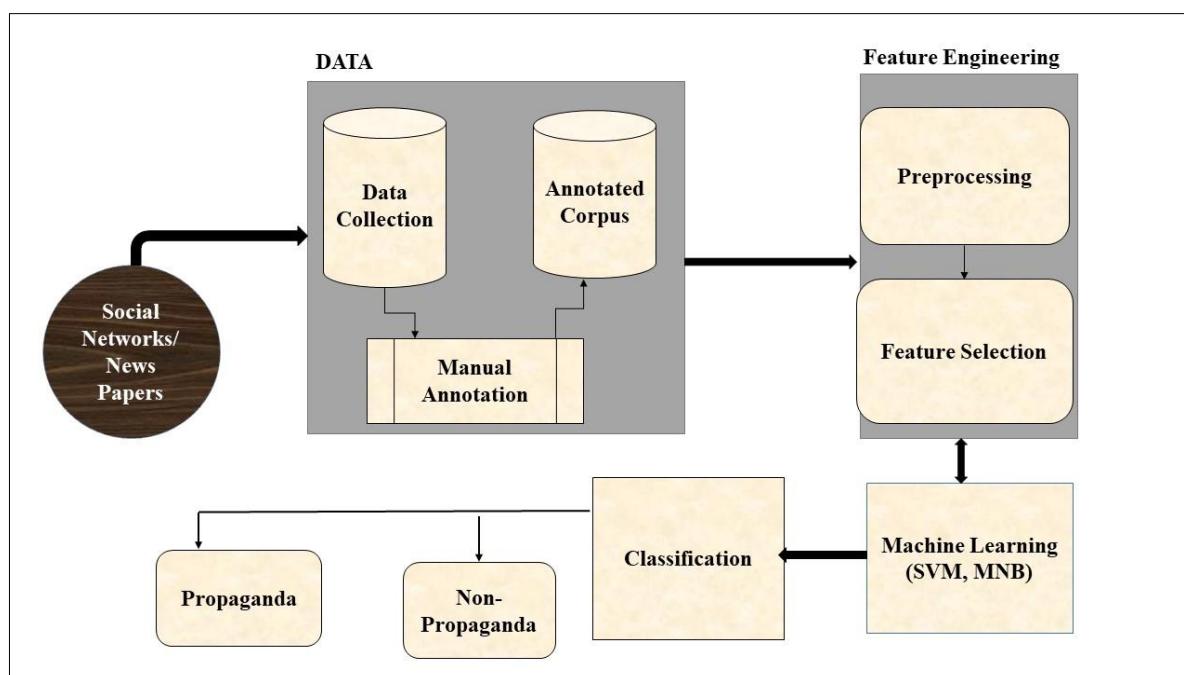


Figure 3. Proposed Methodology



## Data

Data collection is the first phase of the proposed methodology, in this phase, Twitter API is used for the extraction of data from Twitter. Various steps are to be followed before accessing twitter API these are:

1. Apply for a developer account.
2. Make an application related to your work.
3. Get the credentials like Access Token, Secret Token etc.
4. Use these credentials for data extraction with the help of some programming language.

The data is extracted using various hashtags like #propaganda, #Hoaxes, #Falseflag etc. R Language is used for extraction of data from Twitter.

Data is also extracted from online news lets and break them into sentences such that the length of each news sentence does not exceed 300 characters. Figure 4 gives the length distribution of News/Tweets. After the data is extracted, it needs to be classified for performing the classification supervised data is needed. For this task, various annotators were used for labelling the data.

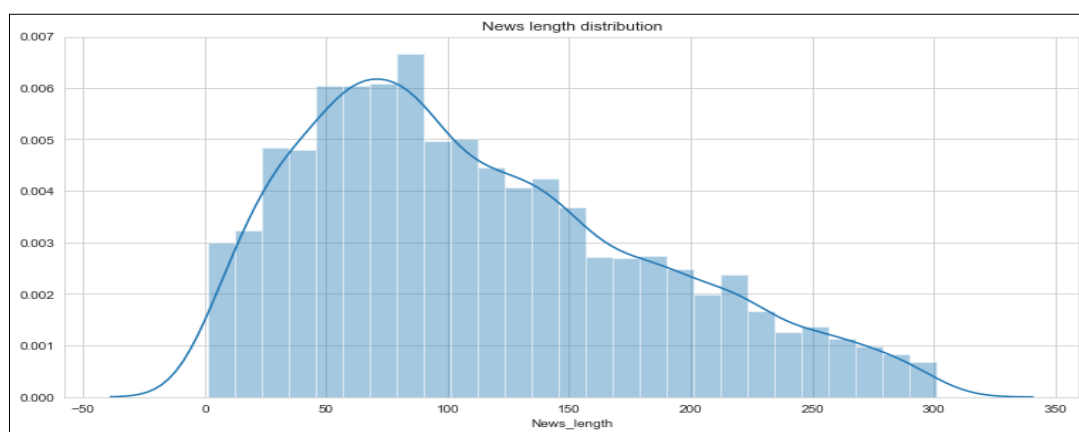


Figure 4. News Length Distribution

## Annotation

In order to classify the posts whether they are propagandist posts or not, various experts of the field were consulted for their expert input and suggestions. Two leading journalists and two computer science professionals were given the task for annotating the data. The annotators were asked to classify the data on the basis of semantics and context. Binary classification is being used in this work. The annotators checked the relevance of the post, whether it is propaganda or non-propaganda by taking different techniques in consideration like a bandwagon, glittering, card stacking, repetition, transfer and name-calling etc.

## Corpus Collection

The experiments were performed on twitter, a social networking site that is mostly used by celebrities, politicians etc. for expressing their views regarding the latest issues that are occurring around the world. Various hashtags have been used for extraction of data from twitter like #Propaganda,

#Hoaxes, #FalseFlag etc. After extracting the 20K Number of tweets an annotation was performed and the tweets were labelled into two categories (0, 1). 0 means non-propaganda and 1 means Propaganda with the help of various researchers that are working on similar type of problems. Figure 5 shows the visual representation of labelled classes with their corresponding length in characters. A supervised dataset has been developed in which there are 6K number of tweets/News. For removing biasness equal number of labelled tweets and news are used that is if

$T$  = total number of labelled tweets.

$X_1$  = number of tweets labelled as propaganda.

$X_2$  = number of tweets labelled as non-propaganda, then

$$T = \sum_{i=1}^2 X_i. \text{ And } X_1 = X_2$$

After annotation, 60% of the tweets/news was found to be about politics and are propagandist in nature.

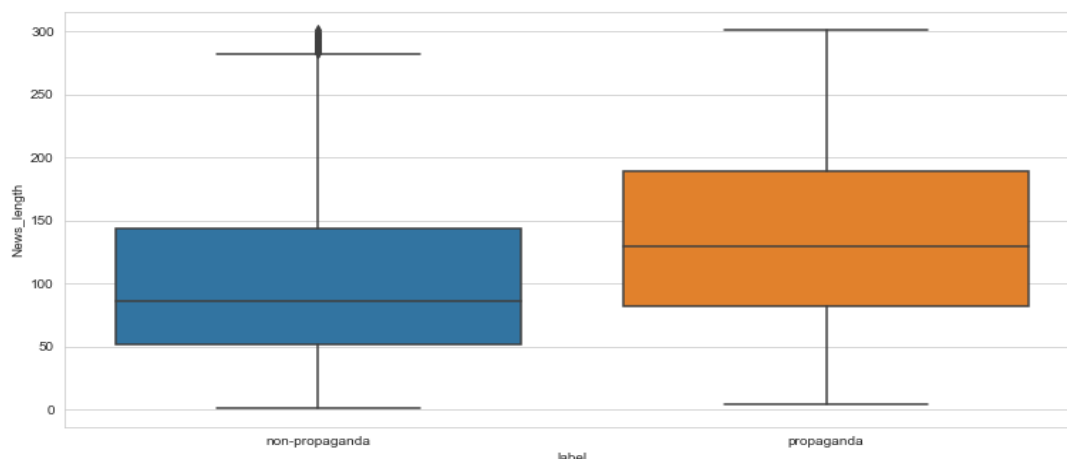


Figure 5. News/Tweet Length of Propaganda and Non-Propaganda Class.

## Feature Engineering:

### Data Pre-processing

Given the collected data, pre-processing and feature design is concerned with the source text. Unicode was used for the conversion of input raw string such that variation is reduced by generating unicode characters into ASCII. The data was pre-processed by removing various uneven text like non-english

text, hashtags, URLs, stop words etc. The next step in data pre-processing is data tokenization in which text is being tokenized by splitting off punctuation from words and part of speech tagging is being done to each token by assigning them morphosyntactic category. In the last step, lemmatization is being performed by leading to the base form of each token, as shown in Fig.6.

	A	B	C	D	E	F	G
1	Text	label	News_lengt	To Lower	Punctuation	Lemmitization	StopWord Removal
2	Trump Pardons Hammond!	non-propaganda	23	trump pardons hammonds!	trump pardons hammonds	trump pardon hammonds	trump pardon hammonds
3	Now, this is good news!	non-propaganda	23	now, this is good news!	now this is good news	now this be good news	good news
4	On Tuesday, President Trump Oregon cattle ranchers, Dwight	non-propaganda	120	on tuesday, president trump oregon cattle ranchers,	on tuesday president trump oregon cal	on tuesday president trump oregon cattle ran	tuesday president trump oregon cattle ranchers dwig
5	A Statement from the White House read as follows:	non-propaganda	49	a statement from the white house read as follows:	a statement from the white house read a	statement from the white house read as fol	statement white house read follow
6	Today, President Donald J. Trump signed Executive Grants	non-propaganda	145	today, president donald j. trump signed executive gr	today president donald j trump signed	today president donald j trump sign executiv	today president donald j trump sign executive grant
7	The Hammonds are multi-generation cattle ranchers in Ore	non-propaganda	165	the hammonds are multi-generation cattle ranchers	the hammonds are multi-generation ci	the hammonds are multi-generation cattle ra	hammonds multi-generation cattle ranchers oregon
8	The evidence at trial regarding the Hammond's case res	non-propaganda	144	the evidence at trial regarding the hammond's case	the evidence at trial regarding the han	evidence trial regard hammonds's case res	evidence trial regard hammonds's case res
9	take our poll - story continues below Will Brett Kavanaugh	non-propaganda	93	take our poll - story continues below will brett kavanaugh	take our poll - story continues below t	take our poll - story continue below will brett	take poll - story continue brett kavanaugh confir
10	Will Brett Kavanaugh be confirmed to the Supreme Court?	non-propaganda	55	will brett kavanaugh be confirmed to the supreme co	will brett kavanaugh be confirmed to t	will brett kavanaugh be confirm to the supre	brett kavanaugh confirm supreme court
11	Will Brett Kavanaugh be confirmed to the Supreme Court?	non-propaganda	55	will brett kavanaugh be confirmed to the supreme co	will brett kavanaugh be confirmed to t	will brett kavanaugh be confirm to the supre	brett kavanaugh confirm supreme court
12	* Yes, he will be confirmed.	non-propaganda	28	* yes, he will be confirmed.	* yes he will be confirmed	* yes he will be confirm	* yes confirm
13	No, he will not be confirmed.	non-propaganda	29	no, he will not be confirmed.	no he will not be confirmed	no he will not be confirm	confirm
14	At the Hammond's original sentencing, the judge n	non-propaganda	265	at the hammond's original sentencing, the judge n	at the hammond's original si	at the hammond's original sentenc	hammonds's original sentence judge note
15	As a result, the judge imposed significantly lesser sentences.	non-propaganda	62	as a result, the judge imposed significantly lesser sen	as a result the judge imposed signific	as a result the judge impose significantly less	result judge impose significantly lesser sentence
16	The previous administration, however, filed an overzealous	propaganda	136	the previous administration, however, filed an over	the previous administration however f	the previous administration however file an	previous administration however file overzealous a
17	This was unjust.	non-propaganda	16	this was unjust.	this was unjust	this be unjust	unjust
18	Dwight Hammond is now 76 years old and has served approx	non-propaganda	86	dwight hammond is now 76 years old and has served dwight	dwight hammond is now 76 years old ;	dwight hammond be now 76 years old and has	dwight hammond 76 years old serve approximate
19	Steven Hammond is 49 and has served approximately four y	non-propaganda	71	steven hammond is 49 and has served approximately	steven hammond is 49 and has served	steven hammond be 49 and have serve appr	steven hammond 49 serve approximately four yea
20	They have also paid \$400,000 to the United States to settle	non-propaganda	81	they have also paid \$400,000 to the united states to	they have also paid \$400000 to the u	they have also pay \$400000 to the unite sta	also pay \$400000 unite state settle relate civil s
21	The Hammonds are devoted family men, respected contrib	propaganda	200	the hammonds are devoted family men, respected c	the hammonds are devoted family me	the hammonds be devote family men respect	hammonds devote family men respect contributors
22	Justice is overdue for Dwight and Steven Hammond, both of	non-propaganda	124	justice is overdue for dwight and steven hammond, t	justice is overdue for dwight and steve	justice be overdue for dwight and steven han	justice overdue dwight steven hammond entirel
23	Well, it took long enough, but thank you President Trump.	non-propaganda	57	well, it took long enough, but thank you president t	well it took long enough but thank you	well it take long enough but thank you presi	well take long enough thank president trump
24	You did the right thing in this matter.	non-propaganda	39	you did the right thing in this matter.	you did the right thing in this matter	you do the right thing in this matter	right thing matter
25	And for all those who took the time to keep this story alive	non-propaganda	142	and for all those who took the time to keep this sto	and for all those who took the time to	and for all those who take the time to keep	take time keep story alive urge people petitio
26	It should be noted that the protests that took place in Ore	non-propaganda	137	it should be noted that the protests that took place	it should be noted that the protests th	it should be note that the protest that take	note protest take place oregon couple years ag
27	As a result, Robert "LaVoy" Finicum was killed by Oregon	non-propaganda	70	as a result, robert lavoy finicum was killed by oregon	as a result robert lavoy finicum was ki	as a result robert lavoy finicum be kill by ore	result robert lavoy finicum kill oregon state polic
28	Those who led the protest were all acquitted of all charges	propaganda	120	those who led the protest were all acquitted of all ch	those who led the protest were all acqui	those who lead the protest be all acquit of all	lead protest acquit charge reporter pete santilli
29	No doubt, Finicum would have been found not guilty as well,	non-propaganda	99	no doubt, finicum would have been found not guilty	no doubt finicum would have been fou	no doubt finicum would have be find not guil	doubt finicum would find guilty well tyrants wo
30	Today is a day to celebrate a wrong that has not been fully	non-propaganda	121	today is a day to celebrate a wrong that has not been	today is a day to celebrate a wrong thi	today is a day to celebrate a wrong that hav	today, day celebrate wrong fully make right de
31	Article posted with permission from The Washington Stand	non-propaganda	59	article posted with permission from the washington	article posted with permission from t	article post with permission from the wash	article post permission washington standard

Figure 6. Pre-processed Dataset.  
Fig 6 Labeled Classes

## Feature Selection

After performing pre-processing, the set of features have been defined for identifying propaganda. Features selected were TF-IDF, sentence length, readability grade level, emotion, LIWC, emphatic and semantic features for identifying propaganda. Term Frequency/ Inverse Document Frequency, numeric measure for scoring a word/term in a tweet/news based on how often it appears in the tweet/news and a given collection of tweets. If in a tweet/news a term is frequently appearing then the word/term should be important and the high score should be assigned to that term/word. Term frequency is the amount of words/terms appearing in each news/tweet. The

following Equation is calculating the TF/IDF in context to our corpus.

$$TFIDF(t, w, D) = TF(t, w) * IDF(t, D)$$

$$IDF(t, D) = \log \frac{|D|}{1 + |\{w \in D: t \in w\}|}$$

Where  $t$  is the term as a feature,  $w$  denotes each tweet in the corpus and  $D$  is the total number of tweets/news in the dataset (Document space). Bag of Words features: Consists of words and lemma uni, bi and trigrams. Bigrams and trigram words were included such that more information can be extracted from the text. The features extracted are shown in Fig.7.

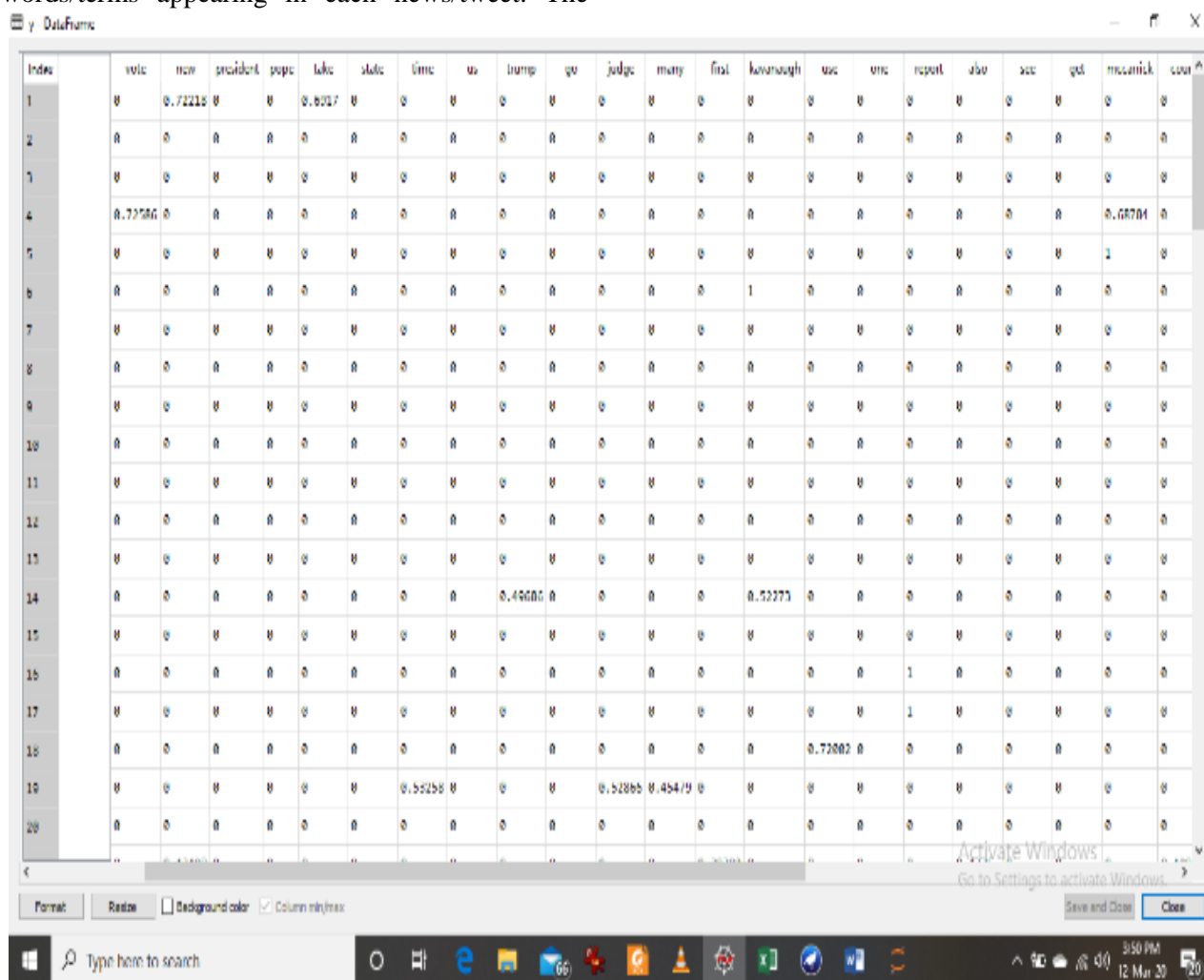


Figure 7. Features and their respective weights.

## Classification

The literature of text classification, it was found that Support Vector Machine (SVM) and Multinomial Naïve Bayes (MNB) are frequently used algorithms for performing classification task. SVM, supervised machine learning approach used for classification tasks as well as for regression problems whereas Multinomial Naïve Bayes (MNB) uses a classical Bayes algorithm for text

classification (30). Here support vector machine (SVM) is used for classification purpose, as our data size is not too large. In our work, data is plotted in the  $n$ -dimensional space (where  $n$  is number of relevant features, in our work 40 relevant features are selected), and the value of each feature is the value of a particular coordinate. By giving  $N$  data points for training set:  $(y_k, x_k)_1^n$

The aim for Support vector machine classifier is to construct classifier which can be formed by Equation 1:

$$y(x) = \text{sign}[\sum_{k=1}^n \alpha_k y_k \psi(x, x_k) + b] \quad (1)$$

Where:  $\alpha_k$  = positive real constant.  
 $b$  = real constant.

$$\psi(x, x_k) = \begin{cases} x_k^T x: & \text{Linear SVM} \\ (x_k^T x + 1)^d: & \text{Polynomial SVM Degree } d \\ \exp(-||x - x_k||_2^2 / \sigma^2): & \text{RBF SVM} \end{cases}$$

Where  $k, \sigma$  are constants.

The classifier is built by assuming the following:

$$\omega^T \varphi(x_k) + b \geq 1, \text{ if } y_k = +1$$

$$\omega^T \varphi(x_k) + b \leq -1, \text{ if } y_k = -1$$

Which is equivalent to Equation 2:

$$y_k[\omega^T \varphi(x_k) + b] \leq -1, \text{ if } y_k = -1, \quad k = 1, \dots, n \quad (2)$$

Where  $\varphi(\cdot)$  = Nonlinear function that maps input spaces into higher dimensional space. The hyperplane is formed by which classification is performed. The hyperplane differentiates the two classes (propaganda and non-propaganda) for this a new variable  $\xi_k$  is introduced. Equation 3 is the equation for hyperplane:

$$y_k[\omega^T \varphi(x_k) + b] \geq 1 - \xi_k, \quad k = 1, \dots, n \quad (3)$$

$$\xi_k \geq 0, \quad k = 1, \dots, n$$

Whereas in Multinomial Naïve Bayesian (MNB) posterior probability is calculated by:

$$P(c) = (N - c)/N$$

Where  $N - c$  = Number of Tweets/News in the Class,

$N$  = Total Number of Tweets/News.

$$P(U|C) = \text{Count}(U|C) + \left( \frac{1}{\text{Count}(C)} \right) + |\text{Vocab}|$$

Where  $\text{Count}(U|C)$  = Count of words in the Class,

$\text{Count}(C)$  = Count of all words in the Class,

$\text{Vocab}$  = Words used in Sentence.

By tuning SVM and MNB properly, it has been observed that they show more accuracy than other complex machine learning algorithms.

### Experimental Results and Discussions:

In our experiment, linear SVM and Multinomial Naïve Bayesian classification algorithms were used. After fine-tuning of SVM, the results showed better accuracy and performance on the given data set. The dataset was divided in the ratio of 70:30 in which 70 per cent of labelled data was used for training and 30 per cent was used for testing purpose. The dataset was not balanced as more textual data was regarding non-propaganda as compared to propaganda data. For performing statistical derivations like mean and standard deviation, statistical equations are used. The statistical results are shown in Fig.8. The statistical analysis is being performed such that much knowledge about the data can be gained which is used for classification purpose.

1	Feature	Standard Deviation	Mean	Minimum	Maximum
2	cardinal	0.029656995	0.14257	0	1
3	ford	0.018193369	0.1146	0	1
4	continuo	0.013808926	0.09728	0	1
5	came	0.024063467	0.13523	0	1
6	ray	0.016012491	0.10581	0	1
7	vate	0.01594771	0.10996	0	1
8	new	0.015949827	0.10744	0	1
9	president	0.017218344	0.11264	0	1
10	pape	0.020469633	0.13311	0	1
11	take	0.019260909	0.1179	0	1
12	state	0.015109876	0.10569	0	1
13	time	0.015201134	0.10522	0	1
14	ur	0.022776423	0.13501	0	1
15	trump	0.023046996	0.12755	0	1
16	qa	0.016218233	0.10721	0	1
17	judge	0.016106411	0.10936	0	1
18	many	0.028394832	0.13741	0	1
19	first	0.024561689	0.13314	0	1
20	kavanaugh	0.019231869	0.12065	0	1
21	wre	0.026997508	0.13874	0	1
22	one	0.015090731	0.10773	0	1
23	report	0.018210036	0.12399	0	1
24	ala	0.015907013	0.1068	0	1
25	ree	0.020471401	0.1201	0	1
26	get	0.026714662	0.14038	0	1
27	mccarrick	0.020174947	0.12202	0	1
28	court	0.016533577	0.10887	0	1
29	know	0.027349941	0.13791	0	1
30	people	0.016872763	0.11161	0	1
31	may	0.054810034	0.19477	0	1
32	would	0.016422881	0.11489	0	1
33	email	0.031358579	0.14968	0	1
34	like	0.019312089	0.11526	0	1
35	even	0.026195422	0.13305	0	1
36	make	0.024135267	0.13266	0	1
37	church	0.043605471	0.17499	0	1
38	claim	0.033382729	0.15819	0	1
39	arrange	0.017434389	0.11173	0	1
40	story	0.013268811	0.09848	0	1
41	guardian	0.026272882	0.13926	0	1
42					

Figure 8. Standard Deviation, Mean, Min and Max



Python language was used for implementing the above work using Spyder IDE. The input is given in the form of a table by having different values for 40 features that have been selected, as shown in Fig.7. Various libraries are being used for performing the given task. Some of the used libraries are Pandas, Sckitlearn, and NLTK etc. SVM is Fine Tuned as:

Kernal = Linear, Poly, Rbf., Degree = 3, 4, 5, Cost parameters = 0.001, 0.01, and 0.1, Gamma = 1, 10,100, Probability = True. Shown in Fig. 9. MNB as: Alpha=1.0, Class prior=None, Fit prior=True.

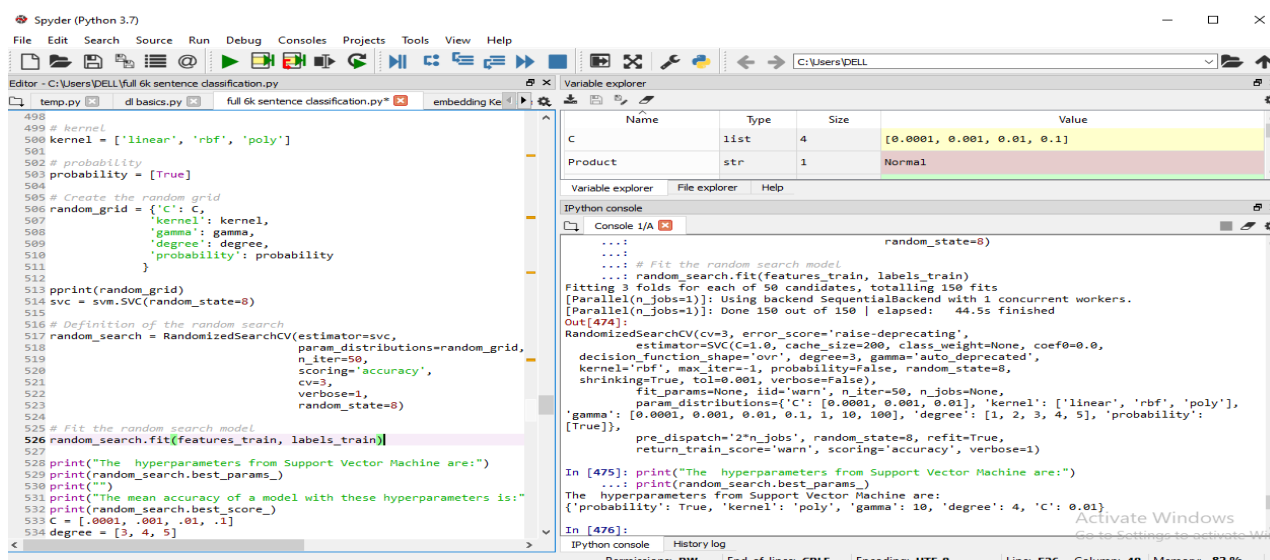


Figure 9. Screenshot of Spyder IDE showing hyper parameters of SVM.

Table 1. Below shows the classification results generated using Support Vector Machine (SVM) and Table 2. Shows the results generated using Multinomial Naïve Bayes (MNB).

Table 1. Classification results generated using SVM.

	Precision	Recall	F1-Score	Support
Non	0.69	0.99	0.81	502
Propaganda	0.71	0.54	0.58	300
Micro Avg.	0.69	0.69	0.69	882
Macro Avg.	0.70	0.765	0.695	882
Weighted Avg.	0.70	0.69	0.59	882
SVM Accuracy	69.2% Testing Accuracy			

Table 2. Classification results generated using MNB.

	Precision	Recall	F1-Score	Support
Non	0.68	0.80	0.78	502
Propaganda	0.71	0.60	0.58	300
Micro Avg.	0.68	0.68	0.67	882
Macro Avg.	0.70	0.70	0.68	882
Weighted Avg.	0.70	0.68	0.59	882
MNB Accuracy	68% Testing Accuracy			

Comparing the results of the algorithms, it was found that Support Vector Machine is better for performing text classification for our problem. The

comparative analysis of these algorithms is shown in Table 3 and also graphically in Fig.10.

Table 3. Comparison between SVM and MNB

	Label	SVM	MNB
Precision	Propaganda	0.71	0.71
	Non-Propaganda	0.69	0.68
Recall	Propaganda	0.54	0.60
	Non-Propaganda	0.99	0.80
F1 Score	Propaganda	0.58	0.58
	Non-Propaganda	0.81	0.78
Accuracy		69.2%	68%

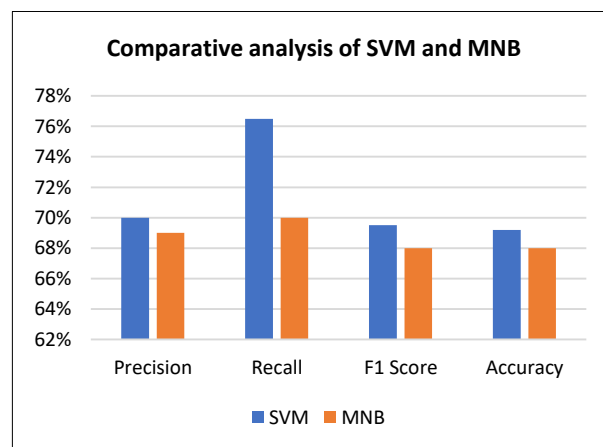


Figure 10. Comparative analysis

After performing analysis it was revealed that propagandistic text have greater length than non-propagandistic text. The majority of data used in our research was related to politics. More and more data that span many fields need to be collected for better analysis of propaganda. More human effort is needed to perform the labelling of the text into various categories. As the data increases, manual annotation becomes tough. Automatic manual annotation program is needed that will learn from the semantics of supplied text. More feature engineering is needed for achieving the better accuracy. Long Short Term Memory (LSTM) may be used in future for achieving the better results.

### Conclusion:

As data sharing has become comfortable with the increasing use of the latest social networking platforms, dubious information is being posted tremendously. In this paper, filthy information, a propagandist text which can be true or false were considered. Propaganda is widely used for gaining influence among common persons for mostly at times of elections. The Support Vector Machine and Multinomial Naïve Bayesian Algorithms are chosen for performing classification. Various features are being selected for performing this task. The results have shown that SVM has 69% accuracy with F1-Score 0.81 for Non-Propagandist class & 0.58 for Propaganda class respectively. In contrast, MNB showed 68% accuracy with F1-Score 0.78 for Non-Propagandist class & 0.58 for Propaganda class respectively. It was also revealed that propagandistic sentences are having larger length then non-propagandistic. Results also showed that SVM outperforms MNB by having a recall of 0.99 and 0.54 with precision 0.69 and 0.71 respectively of two classes. In future other Machine Learning algorithms can be used for achieving better accuracy, with good F1-Score, precision and recall. Ensemble and deep learning techniques may also be used. More feature extraction methods would be used to extract relevant features.

### Authors' declaration:

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are ours. Besides, the Figures and images, which are not ours, have been given the permission for re-publication attached with the manuscript.
- Ethical Clearance: The project was approved by the local ethical committee in Baba Ghulam Shah Badshah University.

### References:

1. Khanday AMUD, Amin A, Manzoor I, Bashir R. Face Recognition Techniques: A Critical Review. STM Journals [Internet]. 2018;5(2):24–30. Available from: [https://www.researchgate.net/publication/330872439\\_Face\\_Recognition\\_Techniques\\_A\\_Critical\\_Review](https://www.researchgate.net/publication/330872439_Face_Recognition_Techniques_A_Critical_Review)
2. O'Neil C, Schutt R. Doing data science: Straight talk from the frontline. " O'Reilly Media, Inc."; 2013 Oct 9.
3. J Han J, Kambel M. Data Mining: Concepts and Techniques, Mor Third. ELSEVIER; 2012. 14–15 p.
4. Ravindran D. Fog computing resource optimization: a review on current scenarios and resource management. Baghdad Sci J. 2019;16(2):419–27.
5. Kietzmann JH, Hermkens K, McCarthy IP, Silvestre BS. Social media? Get serious! Understanding the functional building blocks of social media. Bus Horiz. 2011;54(3):241–51.
6. AL-Jumaili AS. A Hybrid Method of Linguistic and Statistical Features for Arabic Sentiment Analysis. Baghdad Sci J. 2020;17(1(Suppl.)):0385.
7. Verma P, Khanday AMUD, Rabani ST, Mir MH, Jamwal S. Twitter sentiment analysis on Indian government project using R. Int J Recent Technol Eng. 2019;8(3):8338–41.
8. World Economic Forum. The Global Risks Report 2017 12th Edition. Glob Compet Risks Team [Internet]. 2017;103. Available from: [http://www3.weforum.org/docs/GRR/WEF\\_GRR16.pdf](http://www3.weforum.org/docs/GRR/WEF_GRR16.pdf)
9. Gupta A, Lamba H, Kumaraguru P. \$1.00 per rt# bostonmarathon# prayforboston# Analyzing fake content on twitter. In: 2013 APWG eCrime researchers summit 2013 Sep 17 (pp. 1-12). IEEE.
10. Arts THE, Policy C, Justice C, Security N, Safety P, Security H. Rand\_Mg877 [Internet]. Available from: [papers2://publication/uuid/EEFC2DA5-FA25-4EA1-A3F2-674BEC499C03](https://papers2://publication/uuid/EEFC2DA5-FA25-4EA1-A3F2-674BEC499C03)
11. Cybenko G, Giani A, Thompson P. Cognitive hacking: A battle for the mind. Computer (Long Beach Calif). 2002;35(8):50–6.
12. Kumar KPK, Srivastava A, Geethakumari G. A psychometric analysis of information propagation in online social networks using latent trait theory. Computing. 2016;98(6):583–607.
13. Howard PN, Kollanyi B. Bots, # StrongerIn, and # Brexit: computational propaganda during the UK-EU referendum. Available at SSRN 2798311. 2016 Jun 20.
14. Varol O, Ferrara E, Menczer F, Flammini A. Early detection of promoted campaigns on social media. EPJ Data Sci [Internet]. 2017;6(1). Available from: <http://dx.doi.org/10.1140/epjds/s13688-017-0111-y>
15. Introduction C, Data M, Various CI, Spring A, Street OW, Presidential S, et al. First Monday , Volume 21 , Number 11 - 7 November 2016 Electronic copy available at: <https://ssrn.com/abstract=2982233>. 2017;21(11):1–14.
16. Badawy A, Ferrara E. The rise of jihadist propaganda on social networks. Journal of Computational Social Science. 2018 Sep 1;1(2):453-70.

17. Khanday AM, Rabani ST, Khan QR, Rouf N, Din MM. Machine learning based approaches for detecting COVID-19 using clinical text data. International Journal of Information Technology. 2020 Sep;12(3):731-9. Available from: <https://doi.org/10.1007/s41870-020-00495-9>
18. Gao H, Hu J, Zhao BY, Barbara S, Barbara UCS, Chen Y. Detecting and Characterizing Social Spam Campaigns Categories and Subject Descriptors. Proc Internet Meas Conf (IMC. 2010;
19. Ratkiewicz J, Conover MD, Meiss M, Gonçalves B, Flammini A, Menczer FM. Detecting and tracking political abuse in social media. InFifth international AAAI conference on weblogs and social media. 2011 Jul 5;297-304.
20. Halu A, Zhao K, Baronchelli A, Bianconi G. Connect and win: The role of social networks in political elections. Epl. 2013;102(1).
21. Ramakrishnan N, Tech V, Chen F, Arredondo J, Mares D, Summers K. Analyzing Civil Unrest through Social Media. Computer . 2013;80-4.
22. Liu JS, Ning KC, Chuang WC. Discovering and characterizing political elite cliques with evolutionary community detection. Soc Netw Anal Min. 2013;3(3):761-83.
23. Bhat SY, Abulaish M. Communities A gainst Deception in Online Social Networks 1 The Platform 2 The Mischief. IEEE. 2014;2014(2):8-16.
24. Wong FMF, Tan CW, Sen S, Chiang M. Quantifying political leaning from tweets, retweets, and retweeters. IEEE Trans Knowl Data Eng. 2016;28(8):2158-72.
25. Zarrinkalam F, Bagheri E. Event identification in social networks. Encycl with Semant Comput Robot Intell. 2017;01(01):1630002. [Internet] Available from: <http://www.worldscientific.com/doi/abs/10.1142/S2425038416300020>
26. Lightfoot S. Political Propaganda Spread Through Social Bots. Media, Culture, & Global Politics. 2018;(March):0-22.
27. Ashcroft M, Fisher A, Kaati L, Omer E, Prucha N. Detecting Jihadist Messages on Twitter. Proc - 2015 Eur Intell Secur Informatics Conf EISIC. 2015. 2016;161-4.
28. Kumar KK, Geethakumari G. Detecting misinformation in online social networks using cognitive psychology. Human-centric Comput Inf Sci. 2014;4(1):1-22.
29. Mazzoleni G, Bracciale R. Socially Mediated Populism: The Communicative Strategies of Political Leaders on Facebook. Ssrn [Internet]. 2018. Available from: <http://dx.doi.org/10.1057/s41599-018-0104-x>
30. Khanday AM, Khan QR, Rabani ST. Identifying propaganda from online social networks during COVID-19 using machine learning techniques. International Journal of Information Technology. 2020 Oct 29:1-8.

## الكشف عن الدعاية النصية باستخدام تقنيات التعلم الآلي

سيد تنزيل رباني

قمر ريس خان

أكيب محي الدين خندي

قسم علوم الكمبيوتر ، جامعة بابا غلام شاه بادشاه -185234 ، راجوري ، جي أند كيه ، الهند

### الخلاصة:

سيطرت الشبكات الاجتماعية على العالم بأسره من خلال توفير منصة لنشر المعلومات. عادة ما يشارك الناس المعلومات دون معرفة صدقها. في الوقت الحاضر ، تُستخدم الشبكات الاجتماعية لاكتساب النفوذ في العديد من المجالات مثل الانتخابات والإعلانات وما إلى ذلك ، وليس من المستغرب أن تصبح وسائل التواصل الاجتماعي سلاحًا للتلاعب بالمشاعر من خلال نشر معلومات مُضللة. الدعاية هي إحدى المحاولات المنهجية والمتعمدة التي تستخدم للتأثير على الناس لتحقيق مكاسب سياسية ودينية. في هذه الورقة البحثية ، تم بذل جهود لتصنيف النص الدعائي من النص غير الدعائي باستخدام خوارزميات التعلم الآلي الخاضعة للإشراف. تم جمع البيانات من مصادر الأخبار في الفترة من يوليو 2018 إلى أغسطس 2018. بعد إضافة التعليقات التوضيحية على النص ، يتم تنفيذ هندسة الميزات باستخدام تقنيات مثل مصطلح تردد / تردد الوثيقة العكسي (TF / IDF) وحقيبة الكلمات (BOW). يتم توفير الميزات ذات الصلة لدعم المصنفات المتجهة (SVM) و Multinomial Naïve Bayesian (MNB). يتم إجراء ضبط دقيق لـ SVM عن طريق أخذ kernel Linear و Poly و RBF. أظهر SVM نتائج أفضل من MNB من خلال دقة 70٪ واسترجاع 76.5٪ ودرجة F1 69.5٪ ودقة كلية 69.2٪.

الكلمات المفتاحية: الشبكات الاجتماعية ، المعلومات المضللة ، الدعاية ، التردد ، حقيبة الكلمات