

Detection of Propaganda from News Articles using Deep Learning



SuthanthiraDevi P¹, Karthika S², Sowmya K³, Srinidhi S⁴, Pavithra S⁵

¹SSN College of Engineering, India, devisaran2004@gmail.com

²SSN College of Engineering, India, skarthika@ssn.edu.in

³SSN College of Engineering, India, sowmyak17108@it.ssn.edu.in

⁴SSN College of Engineering, India, srinidhis17112@it.ssn.edu.in

⁵SSN College of Engineering, India, pavithras17064@it.ssn.edu.in

ABSTRACT

In the Big data era, a news article provides the biggest amount of unstructured data to the social media users. Propaganda is considered as the information that is purposefully shaped to foster a predetermined agenda. Propaganda uses psychological and rhetorical techniques to reach its purpose. The techniques include the use of logical fallacies and appealing to the emotions of the audience. These techniques are intended to go unnoticed to achieve maximum effect. The main objective of this work is to identify the presence of propaganda in the given news articles by examining the text for as many as 14 propaganda patterns with the help of Convolution Neural Networks (CNN) with Glove word embedding. The use of CNN helps achieve high accuracy in the prediction rate for a given news article. The proposed system can also be used to predict the propaganda from social media. A comparative study is performed to prove the proposed system is better than the n-gram models.

Key words : Big Data, Propaganda, CNN, Glove, n-gram .

1. INTRODUCTION

Dealing with propaganda and extrapolated facts has been a concern for a long while. So much that attempts to identify propaganda were considered to be 'saving democracy' [10]. This intention led to the establishment of The Institute for Propaganda Analysis (IPA) which operated from 1932 to 1947. The IPA helped coin the seven major tricks used as a means of extrapolating information. Though the IPA was set up with good intentions, it had to be dissolved in 1947 as it allegedly created "more destructive skepticism than intelligent reflectiveness". It was also argued that IPA's approach was too simplistic because many messages fell in more than one category [11]. The propaganda of the 21st century is characterized by psychological operations for political means, i.e appealing to the public's emotions to take a particular side for political gain [15]. Also, the enhancement of technology has led to automation in the spreading of propagandistic content across social media platforms like

twitter[4]. Such content consists of conspiratorial, sensationalist, and extremist news article that is barely based on facts and research and can sway public opinion especially for political events such as elections [9].

Several techniques have been tried and tested over the years to appeal to the emotions of the public, and many have succeeded. From using German techniques of limiting ideas and repeating them, to playing on the fear and uncertainties of the victims, propaganda has come a long way[8]. There are a few famous techniques that need to be familiarized to identify and deal with propaganda. Introducing irrelevant material to the issue being discussed, so that everyone's attention is diverted away from the points made is a common strategy of misdirection and is also called Red Herring. Attempts can also be made to discredit the opponent by blaming them of hypocrisy and by not responding to the statement posed - a classic case of Whataboutism. People also are inclined to believe when they are given a single reason for a problem, even though there may be many more, due to lack of research and in-depth knowledge. Such Causal Oversimplification has come a long way in corrupting rhyme and reason in the victims. Obfuscation, deliberate vagueness in statements made tend to confuse understanding facts among the victims leading to multiple interpretations. Many times, untrue content is spread just because perpetrators claim that the content is from an expert on the subject. Such appeal to authority takes undue advantage of the victims' trust.

Labeling the object of the propaganda campaign as either something the target audience fears, hates, finds undesirable or loves, praise is a method that is commonly being used. The emotional side of the target audience is appealed to, by sandwiching baseless opinions in the emotionally loaded language of high praise or hate. Exaggeration or minimization of statements is an accessory to creating an opinionated audience in the world of social and visual media. The propaganda perpetrators may take undue advantage of the victims' sense of patriotism while promoting their products or opinions. Such techniques of flag-waving are commonly observed at times of political events or any social crisis. Forcing the victims to accept a profitable opinion is also common, as such opinions are advertised as the "best course of action" and are accepted on a large scale.

Gina Haspel, New CIA Director

President Donald Trump nominated Gina Haspel as the new director of the CIA¹, announcing the news on Twitter.

Mike Pompeo, the previous director, was nominated to run the State Department to replace the ousted Rex Tillerson.

* * * * *

IN MAY 2013, the Washington Post²s Greg Miller reported that the head of the CIA²s clandestine service was being shifted out of that position as a result of a management shake-up by then-Director John Brennan.

As Miller documented, this official ‘ whom the paper did not name because she was a covert agent at the time ‘ was centrally involved in the worst abuses of the CIA²s Bush-era torture regime.

As Miller put it, **she was directly involved in its controversial interrogation program and had an extensive role in torturing detainees².**

Even more troubling, she had run a secret prison in Thailand ‘ part of the CIA²s network of ‘black sites’ ‘ where **two detainees were subjected to waterboarding and other harsh techniques².** The Senate Intelligence Committee²s report on torture also detailed the central role she played in the particularly gruesome torture of detainee Abu Zubaydah.

Beyond all that, **she played a vital role in the destruction of interrogation videotapes that showed the torture of detainees both at the black site she ran and other secret agency locations.** The concealment of those interrogation tapes, which violated multiple court orders as well as the demands of the 9/11 commission and the advice of White House lawyers, was condemned as ‘obstruction’ by commission chairs Lee Hamilton and Thomas Keane.

Figure 1: Sample news article with propaganda technique

This tactic called bandwagoning is commonly used in commercial opinions[13]. These are just a few of the many techniques that should be known to identify and combat propaganda. To produce an automated model to detect propaganda in news articles by training the model in different types of propaganda techniques and to obtain the optimum values for parameters to achieve maximum accuracy. The news article containing some of the propaganda techniques namely 1. Name Calling 2. Loaded Language 3. Whataboutism, is shown in the figure 1. The corpus based data was used to identify the language features [5]. The main contribution of this paper is to identify the propaganda in the news article. In this paper, Section 2 discusses the problem statement. Section 3 presents the related work for identifying propaganda. Section 4 designs a propaganda detection system to identify the possibility of propaganda. Section 5 discusses the results of the model and its effectiveness of the model. Section 6 concludes the proposed model.

2. PROBLEM STATEMENT

Propaganda makes use of logical fallacies to falsely affect the public opinion for personal gain. It appeals to and exploits the emotions of the target audience to gain a profitable opinion or action. The boom of social media is

making it all the simpler to create and spread propaganda, especially in news articles. Thus, to detect such content, on

a large scale, there is a need to automate the detection of propaganda, especially in news articles.

A Convolutional Neural Network (CNN) is proposed to address the problem of propaganda in news articles. The problem is mathematically identified as Article = {Propaganda, Non-propaganda} for a news article. The proposed propaganda detection system is exposed in Figure 1. The system is trained by using the newspaper3k library news article dataset. The preprocessed data is split into training (0.80) and testing(0.20) sets. The word embedding GloVe method is used to extract meaningful features from the news article. The important features are fed into the CNN system to learn and identify the propaganda in the data set.

3. RELATED WORKS

Propaganda became mainstream mainly after the first World War. Germans are said to have used propagandistic measures to boast of a ‘‘miracle weapon’’ to both raise the morale of their people and scare away their enemies[8]. Careful analysis of multiple cases of propaganda revealed that considering the psychic conditionings of the target audience and utilizing public media are major factors in spreading propaganda. At first, the IPA brought to light seven basic techniques that are being used to generate

propaganda[10]. Research later on added on the techniques of using different modes, such as any interest of the target audience[1]. Analysis of propaganda was first formulated as a 10 step process, from identifying the propagandist to evaluating the effect of the news on the target audience[4]. But it turned out to be tedious with the large amount of news being spread through various social media handles. Thus there came a need to automate the analysis of content in search of propaganda .A deep content analysis was used to identify the sentiment level text features [14].

To combat propagandistic content, there has been a lot of interest in automating propaganda detection, especially as a part of Natural Language Processing. Advancement of technology-enabled the use of deep learning techniques such as neural networks in the domain of Natural Language Processing. Propopy - a system to unmask propaganda and cluster data based on their propaganda likelihood using various NLP techniques, was developed[2]. Propaganda and disinformation identification was also performed using recent state-of-the-art deep learning models such as ELMO, BERT and RoBERTa to compare and ultimately improve the efficiency of propaganda detection[7].

Convolutional Neural Networks (CNN), mainly known for classification tasks in Computer vision have been eventually shown to be effective for sentence classification as well. A single layer CNN was found to produce remarkable results in sentence classification with an accuracy of greater than 85% [6].

CNN is a class of deep learning networks that employs a mathematical operation called convolution on the input. In these models, the convolution operation is applied to at least one of the input layers instead of matrix multiplication [12]. Convolution operation can be thought of as a sliding window function applied to a matrix to simplify large data matrices and find patterns. These functions (or filters) are of a predetermined size and train themselves to perform better. The weights of a neural network can be adjusted to an optimum value using optimizers like gradient descent, Adam, etc. Optimizers aim to reduce the training cost of the dataset to achieve a low-cost model for classification. Though CNN's are usually attributed to computer vision, they can also be applied in the NLP domain. In this case, the input sentences are preprocessed with the help of embedding tools like word2vec and GloVe to create a character/word matrix to be given as the input [3]. CNN's are known to process a huge amount of data at a quick rate and have efficient representation. The Bag of Words model has been a standard approach for the NLP domain and has led to pretty good results.

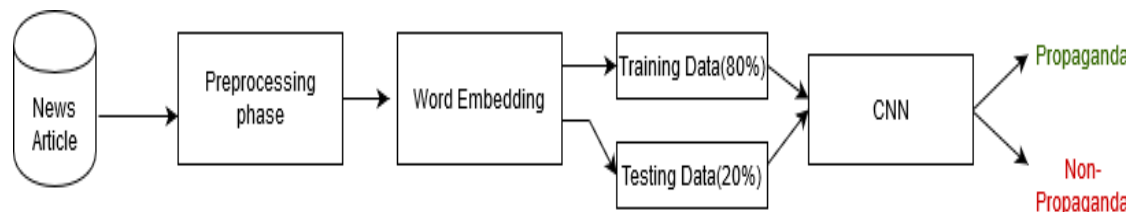


Figure 2: The architecture of the proposed system

4. METHODOLOGY

The proposed architecture of the propaganda detection system is shown in the figure 2. This system contains three major sections as follows

1. Data preprocessing
2. Feature Extraction
3. CNN Layer with Glove embedding

4.1 Dataset

A dataset containing news articles is required to train and test the neural network for classification. The dataset comprises a corpus of 200 articles in which fragments

Containing one out of 10 propaganda techniques have been annotated. This dataset is obtained from the proceedings of the SemEval2020 International Workshop.

The data provided is divided into train articles and test articles and is given in plain text format. Each article has been retrieved with the newspaper3k library and sentence splitting has been performed automatically with NLTK sentence splitter.

4.2 Propaganda Detection system

The data needs to be preprocessed to make it uniform and suitable for sending into the CNN layer as input. The preprocessing steps are shown in figure 3.

Article to Sentence Conversion: The news articles are taken one by one and are converted into sentences by a tokenizer. These individual sentences are stored in a CSV file.

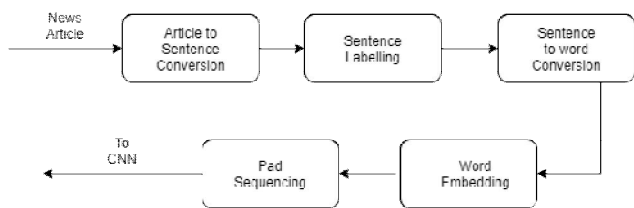


Figure 3: Data Preprocessing

Sentence Labeling: The sentences in the CSV file are manually labeled to be either propaganda or non-propaganda. This CSV file acts as the dataset for training the CNN Layer.

Word Embedding: Word Embedding is a learned representation for text where words with similar meanings are clustered together. In this model, GloVe(Global Vector) word embedding is used. This process breaks the sentences down to words and groups them into vectors based on their similarity. The similarity is ascertained from a pre-existing corpus and the input dataset is tested.

Pad Sequencing: The number of words per sentence is uneven in most of the data corpora, but it is necessary to give a uniform length input to the CNN layer for training and testing. Thus the input sentence is padded with zeros if it has a shorter length.

The data is to be given as input to the Neural Network Layer. The CNN Layer is inserted to take a large amount of preprocessed data and train itself to detect the presence of propaganda in a given sentence. The propaganda detector system is created as a Convolutional neural network with 2 layers having the first layer as the convolutional layer. The activations for CNN are ReLU (Rectified Linear Unit) and sigmoid functions respectively. These functions when set, activate the specific neuron. The ReLU function is given by

$$f(x) = \max(0, x) \tag{1}$$

The sigmoid function is used for the output layer as the predicted output is determined by the probability of the outcome being ‘propaganda article’, i.e, the output layer value should lie between 0 and 1. The sigmoid function is given by

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

If the probability is greater than 0.5 for any sentence in an article, the article is predicted by the model to have propagandistic content, else, the sentence is considered clean. The CNN layer is further optimized using Adam Optimizer. Optimizers update the weight parameters to minimize the loss function. Adam is an adaptive learning-rate optimizer i.e it computes individual learning rates intuitively for each parameter. Adam uses estimations of the first and second moments of the gradient to adapt the learning rate for each weight of the neural

network. Adam optimizer combines the advantages and gives better results than the other optimizers taken individually.

5. RESULTS AND DISCUSSION

The news article or a sentence is given as input to the propaganda detector system. The news article contains any one propaganda technique it is marked as propaganda. The output is binary, i.e either 0 or 1, 1 stands for True, i.e there is propaganda present in the given input. 0 stands for False, i.e the input is clean of propagandistic content.

5.1 Preprocessing

The news article contains noisy, dirty, and inconsistent data. This noisy data may cause incorrect results. The news article contains emoji, whitespaces, and special characters and symbols. The preprocessing rules namely Tokenization, stemming, lemmatization, and case conversions, whitespace removal are applied to clean the article. The following table 1 shows the original article and the cleaned article.

Table 1: Preprocessed Tweet

Original Tweet	Preprocessed Tweet
#John McCain (R-Ariz.)said @Americans need answers about her past involvement in state sponsored torture programs.	John mccain rariz said americans need answers past involvement state sponsored toure programs

5.2 Word Embedding

Word embedding is a technique of converting text corpus into a list of integers. The output of the data preprocessing is fed into the embedding layer and converts words into the indices.

Input: A 2013 Washington Post report on the new CIA head called her role in the torture programs of that era extensive and noted that at one point she ran a CIA black site in Thailand where torture occurred.

Output: [2, 254, 255, 256, 257, 22, 1, 79, 37, 121, 122, 38, 258, 6, 1, 39, 123, 3, 11, 259, 260, 124, 11, 10, 80, 261, 262, 263, 2, 37, 264, 265, 266, 55, 39, 267]

5.3 Machine Learning Models

The proposed news article dataset is fed into the machine learning algorithms namely Logistic Regression, Support Vector Machine, and Voting Classifier. These algorithms map the input vector space into the corresponding output vector space. A comparative analysis is seen in table 2. It has been observed that CNN with an embedding system

provides better accuracy as compared to the traditional machine learning algorithms. The proposed system utilizes the pre-trained word embedding techniques to generate feature vectors.

Table 2.Comparative analysis of accuracy

Model	Accuracy
Logistic Regression	57
Support Vector Machine	61
Voting Classifier	64
Proposed CNN	88.2

5.4 CNN Model

The annotated dataset is split into the training (75%) and testing set (25%). The classifier is used to classify the propaganda and non-propaganda data with good accuracy. It can be observed that the proposed CNN with Glove word embedding can produce an accurate and efficient result to classify the news article contains propaganda or not. This experiment is performed in the newspaper3k library. From the accuracy perspective, the proposed model achieves high accuracy as compared to the machine learning models. Figure 4 shows the training and testing accuracy of the CNN model.

In figure 4 x-axis labeled as epoch to represents the number of epoch in the system, whereas the y-axis represents the accuracy of the proposed model. It has been analyzed that training accuracy is higher than the testing accuracy. This is inferred that the loss is reduced gradually in all epoch.

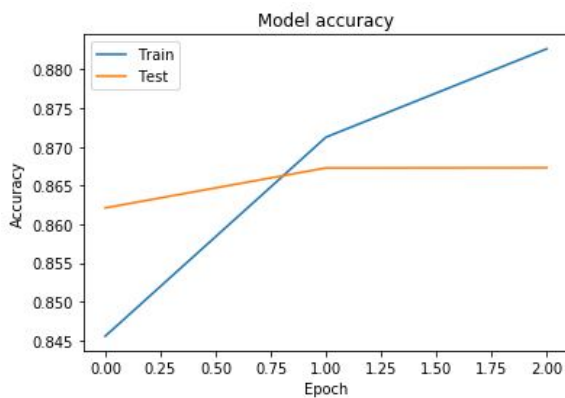


Figure 4: Training accuracy vs. validation accuracy of the propaganda detection model

6. CONCLUSION

The importance of detecting propaganda is quite evident considering the effect this unchecked content has on society. This research work aims to introduce a propaganda detecting system using a 2 layer Convolution Neural Network with GloVe pre-trained embedding. The text from news articles is preprocessed and converted to pre-trained vector and fed into the CNN layer. The neural network trains itself over multiple iterations to predict efficiently. This model has processed a dataset of over 200 news articles and achieved efficiency of 86.7 % in testing. A comparative

study is performed to prove the proposed system is better than the n-gram models.

The project on propaganda analysis has a good future scope. It can be expanded to include the classification of the test article by the type of propaganda techniques used. It will enhance the specificity of the model, thus increasing the scope of use. Propaganda detection can be used in social media and other online platforms where the risk of fake news is high to avoid turning the public opinion for personal gain. Thus, the automation of news policing will help in retaining original opinions in society, thereby supporting the spirit of democracy.

REFERENCES

- [1] Alfred McClung Lee, **The Analysis of Propaganda: A Clinical Summary**, American Journal of Sociology 51, no. 2 (Sep., 1945): 126-135. <https://doi.org/10.1086/219744>
- [2] Barrón-Cedeno, Alberto, Giovanni Da San Martino, Israa Jaradat, and Preslav Nakov. **Proppy: A system to unmask propaganda in online news**, In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 9847-9848. 2019. <https://doi.org/10.1609/aaai.v33i01.33019847>
- [3] Denny Britz, **Understanding Convolutional Neural Networks for NLP**
- [4] <http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/>
- [5] Noorli Khamis , Nurul Farahin Musa **Corpus-based Data for Determining Specialised Language Features** Volume 9, No.1,pp.36-41 January – February 2020 <https://doi.org/10.30534/ijtcse/2020/07912020>
- [6] James Combs and Dan Nimmo (1993), **The New Propaganda: The Dictatorship of Palaver in Contemporary Politics**. New York: Longman.
- [7] Kalchbrenner, Nal, Edward Grefenstette, and Phil Blunsom. **A convolutional neural network for modelling sentences**, arXiv preprint arXiv:1404.2188 (2014).
- [8] Kartik Aggarwal, Anubhav Sadana. NSIT@ NLP4IF-2019: **Propaganda Detection from News Articles using Transfer Learning**. <https://www.aclweb.org/anthology/D19-5021.pdf> <https://doi.org/10.18653/v1/D19-5021>
- [9] Klaus Krippendorf. The Content Analysis Reader. Chapter 1.5: **Propaganda Analysis**, Acase Study from World War II
- [10] Philip N. Howard, Gillian Bolsover, Bence Kollanyi, Samantha Bradshaw, Lisa-Maria Neudert. **Junk News and Bots during the U.S. Election: What Were Michigan Voters Sharing Over Twitter?** Data Memo 2017.1. Oxford, UK: Project on Computational Propaganda. www.politicalbots.org.
- [11] Wikipedia contributors. (2018, April 19). **Institute for Propaganda Analysis**. In

- Wikipedia, The Free Encyclopedia. Retrieved 21:30, May 7, 2020, from https://en.wikipedia.org/w/index.php?title=Institute_for_Propaganda_Analysis&oldid=83
- [12] Wikipedia contributors. (2020, January 25). **Overview of 21st-century propaganda.** In Wikipedia, The Free Encyclopedia. Retrieved 21:55, May 7, 2020, from https://en.wikipedia.org/w/index.php?title=Overview_of_21stcentury_propaganda&oldid=937461420
- [13] Wikipedia contributors, **Convolutional neural network**, Wikipedia, The Free Encyclopedia <https://propaganda.qcri.org/annotations/definitions.html>
- [14] Maganti shymala, Nalini N J, **A Deep Analysis on Aspect based Sentiment Text Classification Approaches**, International Journal of Advanced Trends in Computer Science and Engineering Volume 8, No.5, pp.1795-1801 September - October 2019 <https://doi.org/10.30534/ijatcse/2019/01852019>
- [15] Garth.S.Jowett, Victoria O'Donnell, **Propaganda and Persuasion**, Chapter 6: How to analyze propaganda https://www.sagepub.com/sites/default/files/upm-inaries/95393_Chapter_6_How_to_Analyze_Propaganda.pdf