

ABSTRACT

The proliferation of misleading information in everyday access media outlets such as social media feeds, news blogs, and online newspapers make it challenging to identify trustworthy news sources, thus increasing the need for computational tools able to provide insights into the reliability of online content. Generally, it is misinformation and specifically it can take many forms. It can range from misinformation for profit, political gain, pranks, click bait or political misinformation.

The extensive spread of fake news has the potential for extremely negative impacts on individuals and society. Therefore, fake news detection has recently become an emerging research that is attracting tremendous attention. Fake news detection on social media presents unique characteristics and challenges that make existing detection algorithms from traditional news media ineffective or not applicable.

Combating the spread of false stories is important but difficult. The obvious solution – fact-checking is ineffective. This project uses the applications of NLP (Natural Language Processing) and ML (Machine Learning) techniques for the detection of 'fake news' that is, misleading news stories that comes from non-reputable sources. This helps in providing the people with accurate news so as to gain the confidence of them.

INTRODUCTION

Machine learning is the science of getting computers to act without being explicitly programmed. In the past decade, machine learning has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it.

Fake news detection has recently attracted a growing interest from the general public and researchers as the circulation of misinformation online increases, particularly in media outlets such as social media feeds, news blogs, and online newspapers. For instance, a recent report by the Jump shot Tech Blog found that Facebook referrals accounted for 50% of the total traffic to fake news sites and 20% total traffic to reputable websites. Since the majority of US adults - 62% - (but only 2% of Indian adults) gets news on social media, being able to identify fake content in online sources is a pressing need.

By fake we refer to intentional media manipulation by changing its content and context to bring changes that influence opinions, facts and representations of real world events. Fake can be achieved with a multitude of different manipulation techniques:

- Adding details into media by inserting regions or object from the same image or from another images and adapting them to fit into the entire media environment.
- Deletion of media details, by removing scene elements (regions) and replacing them by others.
- Incorrect classification, ordering and placement in presentations of media collections.
- Generating montages by combining separate images, also called composition and splicing in literature

In 2016, the prominence of disinformation within American political discourse was the subject of substantial attention, particularly following the election of President Trump. The term 'fake news' became common parlance for the issue, particularly to describe factually incorrect and misleading articles published mostly for the purpose of making money through page views. In this project, we seek to produce a model that can accurately predict the likelihood that a given article is fake news.

Facebook has been at the epicenter of much critique following media attention. They used a chrome plug called BS detector that identified bad sources using it's database but now they've dropped it and implemented their own feature for users to flag fake news on the site, however, it is clear from their public announcements that they are actively researching their ability to distinguish these articles in an automated way. Indeed, it is not an easy task. A given algorithm must be politically unbiased – since fake news exists on both ends of the spectrum – and also give equal balance to legitimate news sources on either end of the spectrum. In addition, the question of legitimacy is a difficult one. We need to determine what makes a news site 'legitimate' and a method to determine this in an objective manner.

The top benefit of our model would be reliability. With the help of this model people can accept the news they receive online with no doubts . And not just direct deception but also opinion biases and false critics would be easily identified, hence people would be clear of those as well.

The downside would be misuse of the model that is if the model could be used to promote favoured content if the wrong hands gain control over it. It can be used to promote fake rumours, sponsored content or basically any other source of information on the internet.

With the increasing use of the internet and online news and information the need for distinguishing reliable sources from the unreliable ones is expected to increase drastically and being confident in the news we receive will certainly be a requirement in the future.

PROBLEM STATEMENT

The project aims at building a model that can detect fake news in all its forms on its own after it has been trained with the appropriate dataset.

The aim is to achieve this using:

- Machine learning
- Natural language processing

With techniques like:

- Confidence levels: confidence levels are awarded to each source of news that indicates how reliable they are based on their previous articles.
- Linguistic cues: The goal in the linguistic approach is to look for instances of leakage of deceptive language or, so called “predictive deception cues” found in the content of a message.
- Stance identification: The stance of the author is detected and classified into categories such as agree, disagree and neutral with respect to the title and providing weights to each of these categories will help classify the article into real or fake.

The main objective of the model is to be able to distinguish between “real news” from “fake news” in all its forms such as clickbait, sponsored content, false rumours etc .

OBJECTIVES

The main objective is to detect the fake news ,which is a classic text classification problem with a straight forward proposition .It is needed to build a model that can differentiate between “Real” and “Fake” news.

Functions:

- **Rumor Classification:** The model should be able to detect whether the rumors about an ambiguous situation within an article are true or not.
- **Click bait detection :** The model should be able to detect if the heading of the article matches with its content to make sure it's not merely eye- catching headline to make the reader enter the link.
- **Spammer and bot detection :** The model should be able to detect spammers on social media, which aims to capture malicious users that coordinate among themselves to launch various attacks, such as spreading ads, delivering viruses, and phishing.
- **Truth Detection:** The model should be able to detect true facts from multiple conflicting sources.

LITERATURE REVIEW

Title and Author	Methodology	Advantages	Limitations
<p>“Fake news Detection using Naïve Bayes classifier in Indonesian Language” by Ingrid Yamur Risca Pratiwi, Rosa Andrie Asmara, Faisal Rahutomo</p> <p>2017 International Conference on Information and Communication Technology and system</p>	<p>First the html tags in the article are removed.</p> <p>This is followed by case folding, tokenizing, splitting of sentences into words.</p> <p>Stop word removal: words that occur more than usual and verbs, prepositions, dverbs are removed.</p> <p>Document frequency is the number of document in which a term occurs.</p> <p>The article is manually tagged by reviewers. Each dataset is scored and average score is calculated and updated in the training dataset.</p>		<p>Naïve Bayes is an independent feature algorithm and SVM takes a huge amount of time to train a large dataset.</p>
<p>“Fake news detection in social media” by Kelly Stahl CA95382</p>	<p>A combination of Naïve Bayes Classifier, Support Vector Machine algorithms and semantic analysis is used to detect fake news using a three part method subdivided into supervised learning techniques, and natural language processing methods.</p>	<p>Using Naïve Bayes in conjunction with SVM and semantic analysis drastically improves the efficiency of detecting fake news.</p> <p>Semantic analysis supplements for the independent feature concept of Naïve Bayes.</p>	<p>Naïve Bayes overestimates the probabilities of dependencies. SVM takes more time as the dataset size increases.</p> <p>Associating descriptors with extracted attributes is hard using semantic analysis.</p>

Title and Author	Methodology	Advantages	Limitations
<p>Kai Shu , Amy Sliva, Suhang Wang , Jiliang Tang and Huan Liu</p> <p>“Fake News Detection on Social Media: A Data Mining Perspective”</p> <p>3 Sep 2017</p>	<p>A narrow definition of fake news is news articles that are <u>intentionally false</u> and could mislead readers .</p> <ul style="list-style-type: none"> • User-based: characteristics of the users which post messages • Post-based: information from the posts to infer the veracity of news 	<p>Along with fake news detection, source verification and author credibility check is done as a part of method to find fake news.</p>	<p>From a dataset perspective, there is no existing benchmark dataset that includes resources to extract all relevant features. A neutral stance of any tweets can't prompt the veracity of the news.</p>
<p>“Media-Rich Fake News Detection: A Survey” by Shivam B. Parikh ; Pradeep K. Atrey, in IEEE conference at miami, USA.</p> <p>10-12 April 2018</p>	<p>Contaminated news stories tend to promote interactivity and encouragement that actually attracts users.</p> <p>Choice of vocabulary plays an important role in convincing readers to believe in the story</p> <p>i) Lexical and Semantic Levels of Analysis</p> <p>ii) Syntactic and Pragmatic Levels of Analysis</p>	<p>Instead of verification of source/author, it aims at breaking down the article and identifying the fake news.</p> <p>This method's result is dependent on the article itself and usually is very accurate.</p>	<p>Multi-modal data set: This opens up an opportunity for researchers to create a multi-modal data-set that covers all the fake news data types.</p> <p>Source Verification Author Credibility check</p>
<p>“Detecting fake news at its source”</p> <p>Adam Conner-Simons CSAIL conference at MIT university, USA.</p> <p>Date of Conference: October 4, 2018,</p>	<p>More general approach classifying articles from unknown sources as generally agreeing or disagreeing with sources of known credibility.</p> <p>Program search the database for thousands of articles related to the keywords.</p> <p>It determines the stance of the article to the input claim.</p> <p>If lots of reputable sources all agree with your claim, then it's probably true.</p>	<p>Verifies articles from unknown sources.</p> <p>Source verification and author credibility check is done.</p>	<p>This method works fine only for articles from multiple reputed sources.</p> <p>Even though author credibility check is done, it doesn't contribute to the verification process.</p>

Title and Author	Methodology	Advantages	Limitations
<p>“Fake News Detection” by Manisha Gahirwal, Sanjana Moghe, Tanvi Kulkarni, Devansh Khakhar, Jayesh Bhatia in the International Journal of advance research, Ideas and innovations in technology, volume 4, issue 1, January 2018</p>	<p>The url of the article is used to extract from it which is preprocessed using processes like tokenization and generation of word cloud.</p> <p>The outputs from the previous stage are used to detect the stance of the article i.e. the mental or emotional position adopted by the author such as agree, disagree and neutral with respect to the article.</p> <p>Giving each of these categories weights gives us the final conclusion of whether or not the article is fake or real.</p> <p>It also checks the similarity of the article with top google results.</p>	<p>The weighted categorization of the stance of the author is a very good criteria for detecting deception</p> <p>Since the article is mapped against top google results its authenticity is most likely to be found</p>	<p>The model is very inefficient in the pre-processing stage .</p> <p>No credibility checks are done for the author or the source.</p>
<p>“Linguistics based approach towards fake news detection” by A. A. Ajonye and Z. S. Isa June 2017</p>	<p>In the linguistic approach, the content of deceptive article is extracted and analyzed to associate language patterns with deception.</p> <p>The leakages of deception made by an author could be frequencies in patterns of pronoun, conjunction, and even word usage that are influenced by negative feelings or emotions.</p> <p>Some measures put in place to identify and flag these “leakages” are; data representation, deep syntax, semantic analysis, rhetorical structure and discourse analysis and training of classifiers.</p>	<p>Linguistics deception leakages are involuntary and therefore inevitable in a deceptive article so this method can always be used to detect a fake article</p>	<p>The method relies on language patterns and text alone to arrive at conclusions so the contextual information in the article which might indicate sarcasm or humour is completely disregarded and might lead to false positives during classification.</p>

Title and Author	Methodology	Advantages	Limitations
<p>“A Hybrid approach towards fake news detection on social media” by E. M. Okoro¹, B.A. Abara² April 2018</p>	<p>It combines the linguistics approach along with network analysis.</p> <p>This method relies on the perusal of existing body of collective human knowledge to assess the likelihood of new statements to be false .</p> <p>The method goes beyond the analysis of the questionable content itself to collect and compare a wide range of similar and related statement from various sources such as metatags and social network behaviour to ascertain the likelihood of the content being false.</p>	<p>Since this method combines current network information with the linguistics cues it is less likely to generate false fakes by misinterpreting satires and sarcastic comments.</p>	<p>Literature does not point to any standardized fake news database that</p> <p>(a) stores all social media news,</p> <p>(b) distinguish news into reliable source and non-reliable source,</p> <p>(c) distinguish URL based on</p> <p>(i) known for jokes and</p> <p>(ii) known for real news</p>
	<ul style="list-style-type: none"> • Data Representation: Representation of texts in the “bag of words” approach, which regards each word as a single, equally significant unit. In the bag of words approach, individual words or “n-grams” frequencies are aggregated and analyzed to reveal cues of deception. 	<p>Simplicity of representation.</p>	<p>This method relies exclusively on language, the method relies on isolated n-grams, often divorced from useful context information.</p>

	Semantic analysis: signals of truthfulness have also been analyzed and achieved by characterizing the degree of compatibility between a personal experience (e.g., a hotel review) as compared to a content “profile” derived from a collection of analogous data	Prediction of falsehood is approximately 91% accurate with this method.	This method has so far been restricted to the domain of application.
“From Clickbait to Fake News Detection: An Approach based on Detecting the Stance of Headlines to Articles”. Peter Bourgonje, Julian Moreno Schneider, Georg Rehm, DFKI GmbH, Proceedings of the 2017 EMNLP Workshop on Natural Language Processing meets Journalism, pages 84–89 Copenhagen, Denmark, September 7, 2017.	The first step of deciding whether a headline/article pair is related or not is done based on n-gram matching (of lemmatised n-grams), using the CoreNLP Lemmatiser . To improve and to further classify the related pairs into “agree”, “disagree” or “discuss”, they used Mallet’s Logistic Regression classifier implementation (McCallum, 2002) trained on headlines only ,using the three classes. This resulted in a weighted score of 79.82. •	Accuracy of 61% is achieved with the first step itself.	Context information of the article is not accounted. This system lacks in sophisticated techniques to capture word relatedness in a knowledge-rich way.

LIMITATION OF EXISTING SYSTEM

A major issue of Naïve Bayes classifier is that it uses the “naïve” notion that all features are unrelated. In most cases, this assumption of independence is outrageously false. Suppose Naïve Bayes classifier is scanning an article and comes across “Barack,” in many cases the same article will also have “Obama” contained in it. Even though these two features are clearly dependent, the method will still calculate the probabilities “as if they were independent,” which does end up overestimating “the probability that an article belongs to a certain class” (Fan, 2017). Since Naïve Bayes classifier overestimates the probabilities of dependencies, it gives the impression that it would not work well for text classification.

The disadvantages of using the SVM approach are that it has difficulty with large datasets since “the training time with SVMs can be high” and it is “less effective on noisier [meaningless] datasets with overlapping classes” (Brambrick). In addition, the SVM method will not “directly provide probability estimates” (Ray et al., 2017).

Comparing profiles and the “description of the writer’s personal experience” there are potentially two limitations with the semantic analysis method (Conroy, Rubin, & Chen). In order to even “determine alignment between attributes and descriptors,” there needs to be a great amount of excavated content for profiles in the first place (Conroy, Rubin, & Chen). In addition, there also exists the challenge of being able to accurately associate “descriptors with extracted attributes” (Conroy, Rubin, & Chen)

RESEARCH GAP AND CHALLENGES

PROPOSED SYSTEM AND METHODOLOGY

This system aims to design an artificial intelligence based fake news detection model with a confidence of 75% or more. The model for this system is trained using an existing dataset which includes news published from 2017 onwards. This model we are building does not delete a article which is termed as fake it rather provides a confidence score which is indicative of the legitimacy of the given article. It is left to the user to decide whether to trust the source or not. By implementing this policy, this system does not give rise to biased news which can also be termed as another feature of our system.

The goal of the “Fake News Challenge” is to explore how artificial intelligence technologies, particularly machine learning and natural language processing, might be leveraged to combat the fake news problem. We believe that these AI technologies hold promise for significantly automating parts of the procedure human fact checkers use today to determine if a story is real or a hoax.

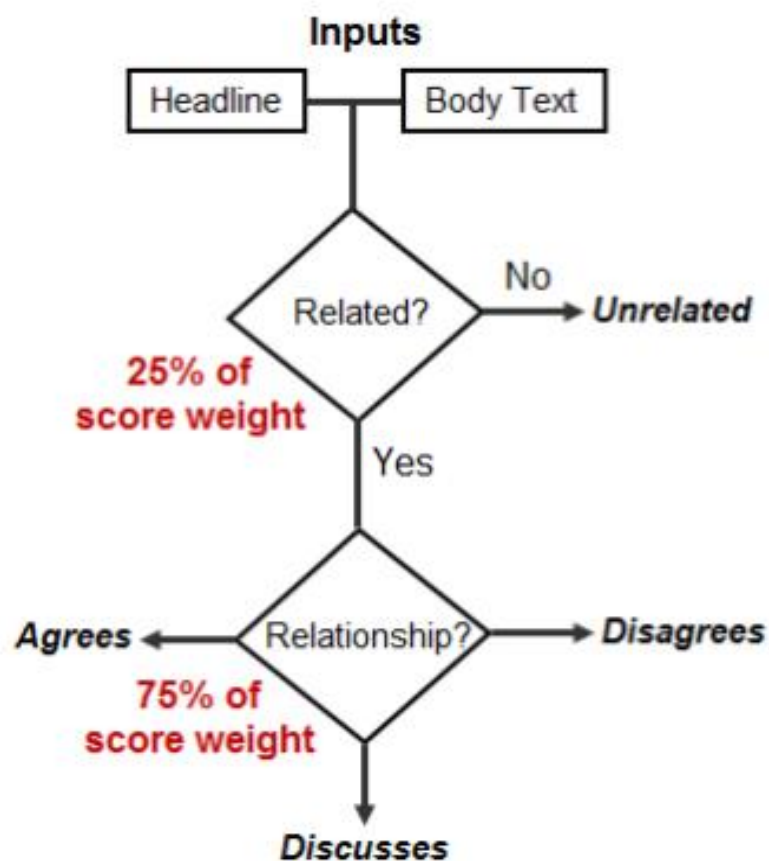
Assessing the veracity of a news story is a complex and cumbersome task, even for trained experts . Fortunately, the process can be broken down into 2 steps or stages.

- A helpful first step towards identifying fake news is to understand what other news organizations are saying about the topic. We believe automating this process, called **Stance Detection**, and could serve as a useful building block in an AI-assisted fact-checking pipeline. So stage #1 of the **Fake News detection** focuses on the task of Stance Detection. Stance Detection involves estimating the relative perspective (or stance) of two pieces of text relative to a topic, claim or issue. For this model we have chosen the task of estimating the stance of a body text from a news article relative to a headline. Specifically, the body text may agree, disagree, discuss or be unrelated to the headline.

- Second step involves fact verification. A news article is compared with similar articles from other reputable publishers. Scoring is given to each of the articles based on following evaluation technique.

- **Stage 1:** Classify headline and body text as *related* or *unrelated* 25% score weighting.
- **Stage 2:** Classify related pairs as *agrees*, *disagrees*, or *discusses* 75% score weighting.

Flowchart



SYSTEM REQUIREMENT SPECIFICATION

Software requirements

- Operating System: Windows/Linux
- Languages: Python
- Libraries: Scikit Learn ,TensorFlow ,numpy , BeautifulSoup.
- Dataset: Kaggle(kaggle.com/mrisdal/fake-news)

Hardware requirements

- Processor: Any processor > 2Ghz
- Internal Memory: 2 GB and above
- Memory: More than 5GB

References

- “Fake news Detection using naives bayes classifier in Indonesian Language” by Ingrid Yamur Risca Pratiwi, Rosa Andrie Asmara, Faisal Rahutomo at the 2017 International Conference on Information and Communication Technology and system
- “Fake news detection in social media” by Kelly Stahl(CA95382)
- Kai Shu , Amy Sliva, Suhang Wang , Jiliang Tang and Huan Liu
- “Fake News Detection on Social Media: A Data Mining Perspective” 3 Sep 2017
- “Media-Rich Fake News Detection: A Survey” by Shivam B. Parikh ; Pradeep K. Atrey, in IEEE conference at miami, USA. 10-12 April 2018
- “Detecting fake news at its source” Adam Conner-Simons | CSAIL conference at MIT university, USA. Date of Conference: October 4, 2018
- “Fake News Detection” by Manisha Gahirwal, Sanjana Moghe, Tanvi Kulkarni, Devansh Khakhar, Jayesh Bhatia in International Journal of advance research, Ideas and innovations in technology, volume 4, issue 1, January 2018
- “From Clickbait to Fake News Detection: An Approach based on Detecting the Stance of Headlines to Articles” , Peter Bourgonje, Julian Moreno Schneider, Georg Rehm, Proceedings of the 2017 EMNLP Workshop on Natural Language Processing meets Journalism, pages 84–89 Copenhagen, Denmark, DOI:10.18653/v1/W17-4215, September 7, 2017.
- “Automatic Deception Detection: Methods for Finding Fake News”, Niall J. Conroy, Victoria L. Rubin, and Yimin Chen ,Proceedings of the 78th ASIS&T Annual Meeting :Information Science with Impact: Research in and for community. Article N.82 ,ISBN :0-87715-547-X, November 06 - 10, 2017.

