Journal of Technology Informatics and Engineering
Vol.1, No.1 April 2022
e-ISSN: 2961-8215; p-ISSN: 2961-9068, Hal 30-49

HYBRID MODEL MACHINE LEARNING FOR DETECTING HOAXES

Budi Hartono[1], Munifah[2], Sindhu Rakasiwi[3]
Universita Sains dan Teknologi Komputer

Abstract

Unlimited availability of content provided by users on social media and websites facilitates aggregation around a broad range of people's interests, worldviews, and common narratives. However, over time, the internet, which is a source of information, has become a source of hoaxes. Since the public is commonly flooded with information, they occasionally find it difficult to distinguish misinformation disseminated on net platforms from true information. They may also rely massively on information providers or platform social media to collect information, but these providers usually do not verify their sources.

The purpose of this research is to propose the use of machine learning techniques to establish hybrid models for detecting hoaxes. The research methodology used here is a feature extraction experiment, in which a series of features will be analyzed and grouped in an experiment to detect hoax news and hoax, especially in the political sphere by considering five modalities.

The outcome of this research indicates that the relation between publisher Prejudice and the attitude of hyper-biased news sources makes them more possible than other sources to spread illusive articles, besides that the correlation between political Prejudice and news credibility is also very strong. This shows that the experiment using a hybrid model to detect hoaxes works. well. To achieve even better results in future research, it is highly recommended to analyze user-based features in terms of attitudes, topics, or credibility.

Keywords*: Hybrid Model, Social Media, Machine Learning, Hoax.*

Received Maret 05, 2022; Revised Maret 20, 2022; April 13, 2022

INTRODUCTION

Detecting hoax is a crucial effort. Not only does it help ensure that users accept reliable information, but it also helps to manage a trusted news ecosystem. The growth of social media and websites on a large scale can provide and disseminate various information without checking facts first, this makes various information posted on websites and social media cannot be fully trusted. Internet access, which is increasing rapidly until 2022, allows anyone to share various information easily and various information that is not necessarily true can be easily accessed by anyone, anytime, and anywhere, making information seekers swallow misinformation without them knowing it. “Misinformation can take many forms such as vandalism” (Abhishek & Philip, 2016), “rumors” (Zhang et al, 2003), Spam (Liu et al, 2017), clickbait, fake websites, fake product reviews (Borella & Rossinelli, 2017), “Hoax” (Shu et al, 2017), etc.

Hoaxes are low-quality messages designed to spreading obfuscation and delude readers. Hoaxes and rumors spread more easily, which can trigger an information crisis. For individuals and communities in various circles, the hoax has a very detrimental effect (Kai et al., 2017) because the general public will be misled by hoaxes and mistrustful information (Zisiadis et al, 2011) so they become confused and clash news. wrong. Besides that, hoaxes can also substitute for how people react to real news. Hoax spreading without borders will throw the trust of the entire news ecosystem into chaos, so from here, finding fake news on social media is very important. Hoax is very difficult to detect, usually, the hoax is deliberately reported to delude information seekers. To find valid and correct information, it is necessary to dig up additional information from various perspectives and be careful.

A study aimed on tweet on social media disclosed that significantly more hoax is spreading on Twitter than real news (valid news), in a higher and quick way, with the user being responsible for spreading it, having fewer followers on average, followed by approximately limited people, and was significantly inactive in tweets. In addition, personal attitude contributes more to the increase of hoaxes than actual news, especially if the news matches their pre-existing attitudes and beliefs that make them believe easily and deliberately believe the news because it fits their thinking. Additionally, Twitter's bot functionality is critical to hoaxes and the increase of hoaxes, with the significant spread of fake news in tweets being driven by human activity.

The volume of illusive news on social media has matured in popularity in modern years. Shearer & Gottfried, (2017) explain that “In 2017, PRC found that 67% of adults (18+) get their news from social media, which is a 5% increase since 2016”. Craig, (2016) found that “An analysis of news leading up to the 2016 election conducted by BuzzFeed, found that there was more engagement with illusive leading news stories than actual news”. This makes disclosure more reachable and across the board than before. Nonetheless, dissemination of information also contributes to the spread of illusive messages and facilitates the advancement of different methods of judging the legitimacy of messages. Such methods were developed by analyzing grammatical personality similar components and grammatical information that determine readability (Home & Adali, 2017).

These methods generally emulate those commonly believed to be highest efficient. Read the message and rate its accuracy. However, with the proliferation of ambiguous news, it is absurd, if not useless, that anyone will spend much time reading numerous magazines and authorities. Moreover, in another study by Gabielkov et al, (2016:179) construct “the evidence that the number of news shares is an inaccurate measure of true readership”. As a result, people are westering in information on socmed, often shared without users reading or considering the validity of the content, and that information can go viral.

The impact of the diffusion of hoaxes is enormous, affecting the news media ecosystem, causing corruption of political information, influencing social media marketing, and also corrupting individual opinions. Therefore, in this research, a machine learning technique to develop a hybrid model is used in detecting hoaxes. All available data and news including titles, content, related images, social networks of users spreading hoaxes, and Source Prejudice are used to detect illusive news. The analysis used in this study highlights the correlation between publishers' political Prejudice and their credibility. Analysis of the collected information shows that nonpartisan news authorities are more likely to viral ambiguous news than other authorities.

LITERATURE REVIEWS

To detect illusive news, many studies consider social media engagement, including news content (title, text, picture), social networks among users, and built-in features including built-in features such as "*Share, comment, and Discuss forums*" and the mixed way that contemplates both. A survey by Shu, et al. (2017:18), precisely defines hoaxes and provides a complete review of the approach for a perceptive hoax on soc-media. The paper characterizes hoaxes by comparing numerous assumptions and properties in traditional news media and social media. Existing algorithms for detecting hoaxes in traditional media depend solely on news content. This method is not effective in the casing of soc-media and therefore exploiting this problem with the additional information of social context helps to detect hoaxes more efficiently.

Features based on news content include metadata-based features such as grammatical aspects extracted from news text, news sources (authors and editors), titles, and visual features extracted from news-related images and videos. Potthast et al. (2017:76) used an article writing style to identify highly biased news stories from neutral ones using a technique called unmasking. A data set consisting of 1.6K materials from the Buzzfeed dataset was used in this research. Aspect equally "n-grams", "stop-words", "parts-of-speech", & "readability" are treated in this research. While there is greater accuracy in determining mainstream versus nonpartisan (accuracy 0.75 percent based on characteristic features and 0.71 percent for point), this study is finite in disentangling among actual and hoax (accuracy 0.55 percent for characteristic and 0.52 percent for the point).

Horne and Adali (2017) consider news bodies and titles to determine news validity. They included three data sets: a data set constructed by "Buzzfeed" dominant up to the 20106 appointments, a set created by researchers containing fake, actual, and innuendo authority, and a 3rd data set of far-reaching actual and sarcastic articles from past research. Based on textual features extracted from content and titles, they found that fake and actual news content differed greatly as they were able to obtain an accuracy of 0,71 when seeing the number of nouns, grammatical repetitions, number of words, and number of citations. Furthermore, this research found that the fake titles contained various types of words namely slang words and overly specific words than the original news article titles which resulted in an accuracy of 0.78.

Picture in newspapers also plays a character in discriminating against illusive news Antol et al, (2015); Gupta et al, (2012); Jin et al, (2017); Tian et al, (2013). Fake pictures are used in news material to foment emotional feedback from soc-media users. Pict is the most consideration-rapacious variety of news contented. Most people are interested in reading hoaxes by looking at a catchy title and supporting pict, so it is very important to consist of pictures investigation in the hoax disclosure approach. Jin et al, (2017:59) “only used visual and statistical features extracted from news images for microblogging news verification and obtained an accuracy of 0.83 on the image data set collected at general news events”. Features based on social context take into account the profiles and characteristics of users who create and share messages. Total of followers or following, the total of posts, and the credibility of users. Also the average of all users associated with a particular story. User opinions and reactions to social media posts (which may be hoaxes), multiple society organizations, co-event organizations, or networks shaped fixed on the total of posts a user has written, related stories, or The dissemination of information between distribution networks whose ends represent paths between users.

Kim et al, (2017:12) suggested an approach to not only catch the hoax but further hamper the advancement of a hoax by creating users’ hoax emblem and using positive 3 parts to fact-check disclosure contented. To do this, they establish a wired conclusion that works when users spread hoaxes, preventing them from spreading. Jin et al. (2016:598) establish an approach to detect hoaxes by finding the pros and cons from the user's point of view and building a credibility transmission network using that relationship. Social media users tend to organize with a consentaneous community and then take in and stake messages that further their enthusiasm and mentality, creating an echo chamber effect. Therefore, creating a different variety of networks, equally attitude, co-occurrence, friendship, and propagation networks, and extracting these network-based traits can help infer network patterns and identify hoaxes.

An attitude network has nodes and edges, where nodes represent news-related posts and edges represent viewpoint similarity weights. The co-event organization is based on user engagement by counting the number of posts the user has written on the news. A friendship network represents a network pattern of followers and followers of users who post relevant messages. A dissemination organization pursues the direction of information dissemination among users. Characterize diffusion and friendship networks using network metrics such as degree coefficient and clustering. Finally, the hybrid method combines the previous two approaches. Ruchansky et al. (2017:797) used the user's transient behavior and their responses, as well as the content of the message body to detect hoaxes. They propose a CSI (Capture, Score, and Integrate) model for classifying news articles. Shu et al. (2016:22) Exploit hoax content and relationships between administrator, disclosure, and users to disclose hoaxes.

Respecting clickbait disclosure categorically, Chakraborty et al. (2016:541) built a customized automated blocker for clickbait titles adopting an easy set of aspects that employ the structure of sentences, patterns’ word, N-gram aspect, and clickbait language. Their browser extension 'Stop-Clickbait' alerts users to potentially clickable titles. Potthast et al, (2016:810) “used Twitter datasets to identify messages on social media that lead to clickbait”. They collect tweets from the different builders and build aspects based on teaser messages, linked web pages, and meta information. Anand et al. (2017) used three variants of the two-way RNN model (LSTM, GPU, and standard RNN) to detect clickbait titles. They use two particular enclose approaches equally distributed embedding and aspect-matched embedding. Chen et al, (2015) “investigated a hybrid approach to clickbait detection using text-based and non-text-based clickbait cues. While textual cues use text-based semantic and syntactic analysis, non-textual cues deal with image analysis and user behavior”.

DATASET

Data Collection

As shown in Table 1, several constitutional disclosure records have been used to detect hoaxes of news.

Table 1. Possible data set for illusive disclosure detection.

Dataset	Size	Text	Images
BuzzFeedNews	1,627	✓	
Horne and Adali DS1	71	✓	
Craig Silverman	225	✓	
Gupta et al	480	✓	
FakeNewsNet	384	✓	✓

The BuzzFeedNews database of election-related news is announced on FB by 9 post outlets. The dataset identified 356 left-leaning news articles and right-leaning (545) news materials, of which were mostly accurate, with mixed accurate and false 212, and false 87. A study by Horne and Adali (2017:9) used two datasets in the work by Horne and Adali (2017:3). While the early record, DS-1, contains 36 true traditions and 35 hoaxes, Craig Silverman's third column lists him as having 75 true stories, fantasies, and caricatures, i.e. 75 stories in each category is included. The main drawback of both datasets is that labels are assigned based on the credibility of the news source rather than fact-checking. However, news sources vary in reliability and may contain factual and fantastical information. FakeNewsNet Wang et al. (2017:426) is the only highest development data set far-reaching message outside the modalities of news content and political domains. In terms of relevance and importance, this dataset is used to analyze this study. As Table 1 shows, label collection requires time-consuming fact-checking, which generally limits the use of large-scale benchmarks for hoax detection. In a study by Shu et al. (2017:36), another dataset was used for a related task but was not suitable for analysis in this study as it did not contain appropriate news articles. For example, LIAR (William, 2017:422) contains a short statement labeled "human", whereas CREDBANK by Tanushree & Gilbert (2015:258) contains a log of events, where each event is a collection of tweets. Finally, the MediaEval Verification Multimedia Usage Benchmark Dataset by Symeon et al. (2015:11) Uses that include images or tweets, not news articles.

Fake-News-Net Data Collection

The FakeNewsNet dataset consists of news content details, disclosure news, and community commitment by Wang et al, (2017:422). Basic authenticity stamps are quiet from expert journalists such as “Buzzfeed” and the fact-checking website “Politifact”. The data set is split between two networks, Buzzfeed and Politifact, with news content collected from his web links on Facebook. The record contains all available downloaded images associated with the news in this record. Publisher Prejudice is obtained from the data set described in the next section. For this study, we combined disclosure from “Politifact” and “Buzzfeed” to make a bigger data set available. After clearing the data set that was missing message titles, there were a total of 384 messages, 175 fictitious and 209 factual.

Figure 1. Amount of publishers per division in the Media Prejudice/Fact-Check data set.

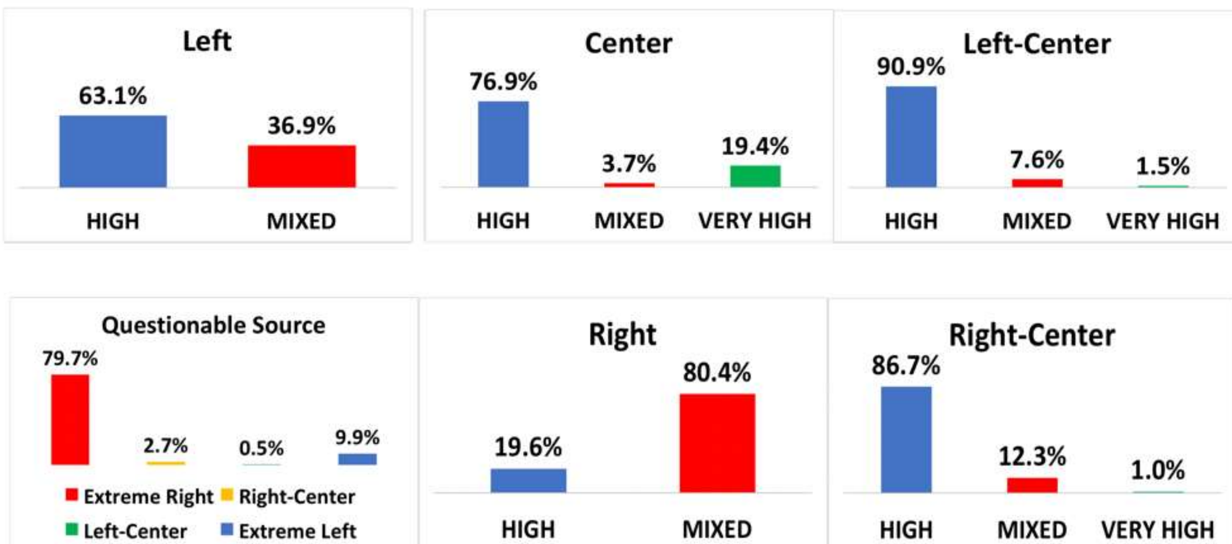

Figure 2. creator integrity per Prejudice and distribution of Prejudice within the source in question in the MediaBias/FactCheck data set.

media prejudice/FactCheck Dataset

To take advantage of biased information from news sources, several websites were investigated to develop the civil about media prejudice and misleading disclosure manners. This web consists of an all-inclusive checklist of disclosure sources, their biases, and the reliability of their findings in fact-based reporting. Here, the publisher's political Prejudice is definite using 7 degrees of prejudice: extreme-right, right, center-right, neutral, center-left, left, and extreme-left. The factual reporting scores of all news sources are aggregated into five categories: Left Prejudice (moderate to strong Prejudice towards liberal causes), Mid-Left (mild to moderate liberal prejudice), Slight (minimal prejudice), Mid-Right (slight to moderate conservative prejudice), and Right Prejudice (moderate to very strong Prejudice towards conservative causes). This publisher's credibility score falls into 3 categories: highest, High and Mixed (meaning the source is not always the correct source or other prejudiced/mixed authority). Publisher Prejudice is also collected in the Questionable Sources category which contains publishers who are highly biased mainly by disinformation and/or print illusive disclosure. The total of publishers in all classes studied is served in table fig.1 and there are 1,7K publishers.

METHOD FOR EXTRACTING FEATURES

Textual Features

To analyze the text content, using the following feature groups are used, and those features will be calculated for the content of the news body and the title.

Term of Frequency Inverse in Doc Frequency (Tef-Idf)

In this research, Tef-Idf former to produce text, where any word is characterized by a score to express the importance of the word in a corpus based on the frequency level of the word that appears in the document and also how many other documents contain Tef-Idf. Text preprocessing is done by applying Stemming and punctuation and removing stopwords. The basic Tef-Idf scoring function is provided in the following Eq.

$$tef - idf_{v,di,D} = \left(\frac{C_{wd}}{|d_i|}\right) . \log\left(\frac{|D|}{1 + |d \in D : w \in d|}\right)$$

The 1st phrase serves as the syllable frequency *(tef)* of the word *v*, which is the ratio of the number of occurrences of the word *(C_{wd})* to the total number of words in the document *($|d_i|$)*. The second term is reverse document frequency *(idf)* which increases more descriptive words and reduces the impact of frequently used words such as articles, and pronouns. Idf is calculated by taking the logarithm of the total number of documents in the corpus *($|D|$)* divided by the number of documents with words offset by 1 to avoid 0 denominators *($1+|d \in D{:}w \in d|$)*. Words that appear in almost all documents will have an idf close to 0 and words that come out only in selected documents will have a larger idf value, increasing their Tef-Idf weight.

LI-WC

LI-WC is a translucent text investigation tool that calculates the total of words of text in intellectually substantial classes. The LI-WC 93 measurement is used to analyze cognitive, affective, and semantic development in texts. To explore the differences amongst accurate and illusive news writing styles, LI-WC features are divided into four categories: “Linguistic”, “Punctuation”, “Psychological”, and “Summary” Pennebaker et al, (2015:2).

Linguistic features refer to features that represent the functionality of the text such as the average number of words per sentence and the misspelling rate. Accordingly, the words total function, as well as negation under this category, are selected. *The punctuation feature is* used to overstate or sensationalize a story which can be analyzed through the types of punctuation used in news such as (.), (,), (:), (;), (?), (!), (x), (“…”), (‘’), {}, and another interval. *Psychological features* target emotional, social, and cognitive processes. Affective processes (positive and negative emotions), social processes, cognitive processes, perceptual processes, biological processes, time orientation, relativity, personal attention, and informal language (swear words, non-fluency) can be used to examine the emotional part of news stories. *The summary feature* determines the frequency of words that reflect the author's thoughts, perspectives, and honesty. It consists of investigative reasoning, Impact, accuracy, passionate tone, Word per-decision words is more than 6 letters and Dictionary words under these categories.

Legibility

Readability measures how easily readers can read and understand the text. Text complexity is measured using attributes such as word length, sentence length, and the number of syllables. Popular readability measures used in the search: "*Flesh Reading Ease", "Flesh Kincaid Grade Level", "Coleman Liau Index", "Gunning Fog Index", "Simple Measure of Gobbledygook Index"* (SMOG), "*Automatic Readability Index"* (ARI), "*Lycee International Xavier Index"* (LIX), and "*the Dale-Chall Score" (DL)*. A higher score of Flesch's readability indicates that the text is easier to read and a lower score indicates it is crucial to learn. The Coleman Liau index relies on word characters to measure text comprehension. The "Dale-Chall" readability test uses a list of words fourth graders know to resolve the crisis level of a text.

Image Features

To analyze news-related images, a sophisticated deep learning-based technique is used to extract features from images. In the research by Cozens et al, (2005: 337) "Promoting a positive image and maintaining the built environment regularly ensures that the physical environment continues to function effectively and transmits positive signals to all users". Efficient fast management, and up-to-date maintenance are the main components of this site feature. The concept of 'image' with its emotional basis determines the unique appeal of an attraction. This influences the nature of visitors attracted to the site and the desired behavior. Sometimes stakeholder engagement can be influenced by the consequence of allowable visitor behavior by Ekblom (2011a:17). Although conservation of material affects the performance of the site, management, which includes maintenance, affects the opportunities for crime to occur. The broken study by Wilson & Kelling, (1982:294) emphasizes the role of management, while other work links inadequate management practices to crime drivers Wortley & Mazerolle, (2012:3). Opportunities to create a positive image and to ensure prompt repair and maintenance depend on the nature of the attraction property. Vulnerable property features are considered easy targets and easily damaged (Fyall et al., 2008).

NeuralTalk2

NeuralTalk2 is an efficient drawing text model, coded in Torch running on GPU. It is similar to the original NeuralTalk, but this model's implementation is bundled, uses Torch, runs on the GPU, and supports CNN enhancements. All of this together results in a sizable increase in practice acceleration for the LM (~100x). The "NeuralTalk2" approach is planted on a new compound of Convolutional Neural Networks over image regions, bidirectional Recurrent Neural Networks over sentences, and structured goals that align the two modalities through multimodal embedding. Then they designed a Multimodal Recurrent Neural Network architecture that used inferred alignments to learn to generate new descriptions of image regions. Models can be trained using loadcaffe using VGGNet. But in this study, a pre-training model checkpoint was used to extract captions describing image content from news-related images by using NeuralTalk2 (Vinyals et al, 2017), a pre-training recurrent neural network that summarizes image content in a single sentence description. After that, the Tef-Idf to represent the text of the text is calculated and considered as an additional feature in the analysis.

Journal of Technology Inforamatics and Engineering
Vol.1, No.1 April 2022
e-ISSN: 2961-8215; p-ISSN: 2961-9068, Hal 30-49

Source Prejudice

Several studies in journalism have theorized about the correlation between a publisher's political Prejudice and the correctness of the disclosure content it circulates. To approve this expectation, the communication amongst a news source's political Prejudice and its integrity was examined by identifying information about 1,7K publishers in the MediaBias/FactCheck data set. Figure 2 shows the distribution of credibility scores per category of political Prejudice (from left to right) and the distribution of Prejudice in the sources in question. The plots show that when a news source is moderately to heavily biased (both conservative and liberal), then that source is more likely to publish illusive stories than other news sources that are more moderate and expressed as center-left, center-right, or neutral. Also, Extreme-right (or very conservative) is the main Prejudice among the sources in question. Thus, news Source Prejudice was used as another feature in the experiment.

Social Network Features

The social network feature provides useful information about a user's social network, how quickly and at what level hoaxes spread through these networks. This research uses the relationship among “publisher & news”, “news & users”, and “users & users” to get the truth about each. Then, the news credibility score is calculated by modeling the problem as an MRF where the absurd BP algorithm from Chau, et al. (2011) is used to make inferences. In general, the MRF method treats each node as an aimless fickle and this research problem has three types of publisher, news, and user nodes. The aimless variables for each node are represented as pi ∈ {0.1}, nj ∈ {0.1}, and uk ∈ {0.1} where 0 is not credible and 1 is credible and the result is the marginal probability p(pi), p (nj), and p(uk) quantify the belief that node i belongs to class pi, node j belongs to nj, and node k belongs to uk. The initial probability of each node can be determined and represented by the operation of ∅ (pi), ∅ (nj), and ∅ (uk) which can be obtained from the data set. Given the bias, ∅ (pi) provides information about whether or not the publisher is credible. For example, for Right bias, the prior probability that there is a possibility of the publisher not being credible is higher as shown in Figure 2. For Left bias, even though the probability of not being credible is lower compared to right Prejudice but the chance of being not credible is higher compared to other biased publishers. Questionable sources are mostly from non-credible sources and the maximum percentage of publishers are extremely right-biased.

With a lack of prior knowledge and information about the news and based on ground truth, an assumption is made that the news can be 50% likely to be credible and 50% not credible. A new study from MIT proposes that human nature is responsible for the rapid spread of hoaxes rather than actual credible news. The research work analyzed more than 100,000 stories on Twitter for how many total tweets were posted and reposted, the time to reach engagement magnitude, and verified the account from which it was created. They have proven that users who spread hoaxes have far minor followers, follow minor people, and are less active on Twitter. This study is used in this study work to deduce the prior possibility of users on the net by calculating the degrees of entry, and degrees of exit. The features that represent the number of followers/followers are extracted using a graph mining library, SNAP, based on the network between users. This individual-level feature is used to assume the prospect and accuracy of all users who spreads disclosure on soc-net.

The function $\psi\psi$ ij is a hyperparameter that determines the conditional probabilities for each node and the credit score can be measured for edges using the edge potential function. The table below shows the choice of affinity matrix $\psi\psi$, for $\varepsilon\varepsilon > 0$, this choice of $\psi\psi$ assumes a correlation between nodes. Table 2 shows that if the publisher is not credible then it is likely to publish a hoax and low probability to publish actual news. Likewise, if the news is not credible, the user likely spreads the news is also not credible and has a small possibility of growing the good disclosure in Tables 3 and 4.

$\psi(p_i, nj)$	0	1
0	$1-\varepsilon$	ε
1	ε	$1-\varepsilon$

Table 2. Potential edge function between publisher nodes and news

$\psi(n_j, u_k)$	0	1
0	$1-\varepsilon$	ε
1	ε	$1-\varepsilon$

Table 3. Boundary possible functionality between disclosure nodes and users

$\psi(u_k, u_k)$	0	1
0	$1-\varepsilon$	ε
1	ε	$1-\varepsilon$

Table 4. Function possible edge between node and users

EXPERIMENT RESULTS

In this study, each feature group was used as input to the RandomForest classifier with 5-pleat cross-authorization to calculate the achievement of this feature in classifying factual vs. factual stories. illusive. Results serve in Table 7 correspond to the Area Under the ROC curve (Ar-Un-ROC), Measure of Fi1, and regular attention. Class weighting is used to address the class imbalance. Experiments also included classification using a continuous S-VM separator with L2 adjustment (5-pleat cross-authorization) and the results are reported in the Appendix. The results of the two classifiers were compared and it was found that the RandomForest separator performed better with these features.

News Body Contents

The first modality analyzed is the body of the news. Here, the Tef-Idf feature achieves the best results (0.888 Ar-Un-ROC, Fi1 size 0.811, and average precision 0.781). Furthermore, LI-WC features are the second-best feature group. Among them, psycho-linguistic features are the most important feature group, achieving comparable performance. After LI-WC, readability features did not appear to separate illusive news from factual news in this data set. Illusive news has a higher frequency of psychology-related words equally particular concern, relativity, social and biological action. The language used is more tentative words that give rise to uncertainty, is more informal, and contains more swear words. In contrast, accurate disclosure is extra crucial to understand and has words associated with higher risk, less anger, and more words of sadness. There are more parentheses in factual stories which are used to indicate additional content that provides more evidence about the story.

News Title

Among all the features considered for analyzing titles, it is seen that all LI-WC features combined perform best by all measures (for example, Ar-Un-ROC 0.791) and after that Tef-Idf with Ar-Un-ROC 0.733 performs better. Regarding punctuation, illusive titles tend to use periods, exclamation points, and semi-columns (which may indicate that they pack a lot of sentences in the title). According to the level of readability, factual titles are more complex to understand and show a higher "Flesch-Kincaid" score than illusive news which has more tentative words that raise uncertainty, is more informal, and has extra affirmed words as seen in the contents of the news. Overall, the analysis shows that factual political titles are written more professionally than illusive titles.

News Source Prejudice

Source Prejudice is a strong diviner of disclosure integrity as it achieves an Ar-Un-ROC of 0.884, an average precision of 0.917, and a FiI1 measure of 0.854. This result further confirms the interactions among Source Prejudice and the integrity of the disclosure it disseminates. It should be noted that the publisher message is autonomous of disclosure stamp as the departed is quiet from "MediaBias/FactCheck", which the latter is from "Buzzfeed" and "Politifact".

News Image

Images related to the news were used to determine disclosure potency and it was found that the Tef-Idf feature of picture captions from NeuralTalk2 performed superior and was equal to disclosure titles (0.743 of Fi1 size, 0.600 Ar-Un-ROC, and 0.725 mean precision). Through manual analysis of the images included in the data set used, the trend of images used in illusive news and real news became clear. One such trend is that genuine news articles include more images that focus on a character speaking, whereas illusive disclosure articles consist of more pics of the community with only their facial expressions. Furthermore, the images in actual news portray a more positive impression than illusive news. A final note from a manual check of the data set is that the illusive news image has most likely been shot by combining the two images and that the image is of lower quality than the image from the original news data set.

News Social Network

Social network features are used to calculate a credibility score for each story based on how users on the social network share the news. To get the useful features of the social network, the initial probability for news is set as 50% credible and 50% non-credible due to a lack of prior knowledge of the news.

The prior probabilities for each user are derived based on the degrees in, degrees out, and combining the two degrees. Inbound and outbound rates against the percentage of users in each category are plotted. All users are plotted, users who spread the news more than 6, 7, 8, 9, and 10, and the threshold to get the initial probability for each user are set based on the histogram plot. Figure 3 shows the indegree, outdegree, and percentage distribution of users with users spreading the word greater or more than 8. Two different prior probabilities for publishers were also tried, one with Source Prejudice and another with the previous default as 50% for both credible and not credible.

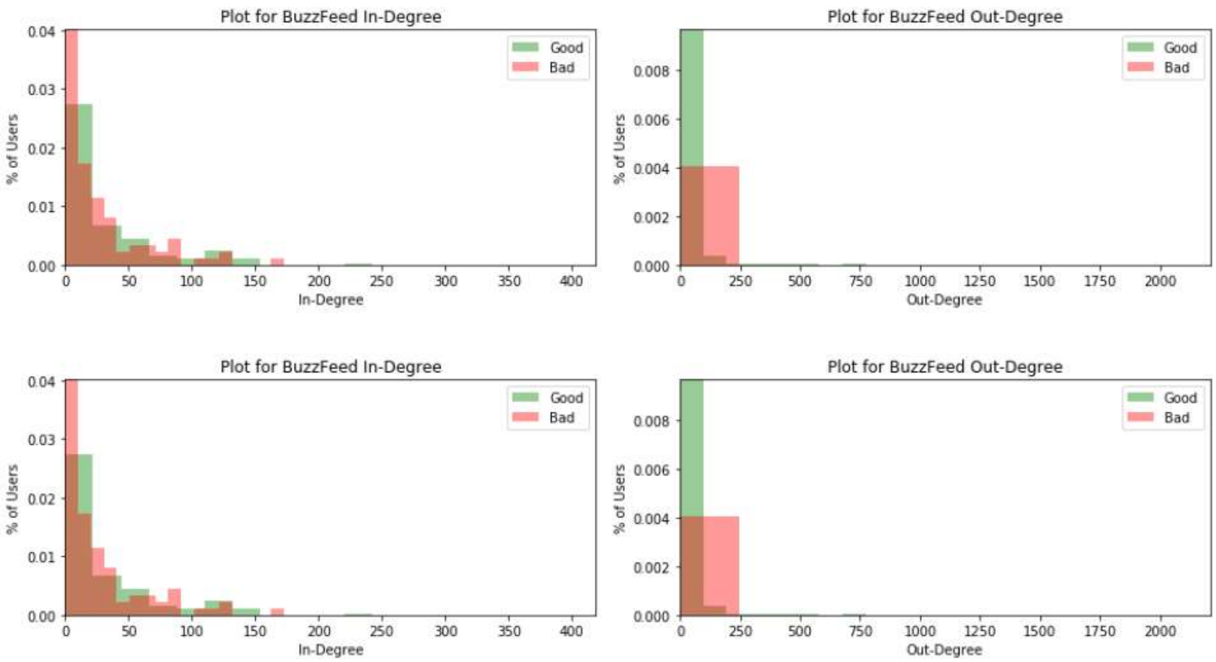

Figure 3. In-degree & Out-degree and percentage distribution of users

The next step is to use sparse matrix confidence propagation, an altered form of crazy assumption breeding that encodes a graph structure with a sparse matrix to infer a credibility score for each node. In this case, this is focused on the combination Markov aimless plane that approaches the posterior margin of each node. Various possibilities are tried and tested equally connecting Prejudice by changing the prior probabilities to calculate the credibility score using crazy belief propagation. Experiments were also conducted with all users and also with those users who posted a total number of stories greater than 6,7,8,9, and 10, to observe whether users only create noise by sharing only one or two stories. The metrics used to calculate the credit score are Ar-Un-ROC and the mean precision and results for users sharing more than 6 and 7 stories are shown in Tables 5 and 6 for the BuzzFeed and PolitiFact data sets. From this experiment, it is evident that the integrity scores for the user with different news counts are almost the same with small differences, and the indegree and outdegree features work for the “BuzzFeed” dataset and the “in-degree” features give better results for “Politifact”.

The credit score increases when using publisher bias, but to blend with other methods this research uses the credibility score with a compound of total news greater than 8 and both degrees for Buzzfeed which gives better results (average precision 0.722, Ar-Un-ROC 0.636), and indegree for "Politifact" (0.513 mean precision, 0.560 Ar-Un-ROC) to avoid repetition. Overall with all modalities, the results achieved are 0.880 Ar-Un-ROC, 0.858 Measures of Fi1, and 0.779 average precision, which indicates that the utilization of social networking features will increase in detecting illusive news.

Features	AUROC > 8	Avg.Precision > 8	AUROC > 9	Avg.Precision > 9
Indegree	0.584	0.604	0.603	0.638
Outdegree	0.415	0.503	0.397	0.493
Both degree	0.636	0.722	0.648	0.738
All Default	0.500	0.506	0.500	0.506

Table 5. Integrity Average for aspect with more than 6 and 7 news spread by users "BuzzFeed"

Features	ArUn-ROC > 8	Avg.Precision > 8	ArUn-ROC > 9	Avg.Precision > 9
Indegree	0.560	0.513	0.558	0.511
Outdegree	0.441	0.446	0.442	0.448
Both degree	0.440	0.446	0.441	0.448
All Default	0.500	0.500	0.500	0.500

Table 6. Integrity Scores for aspects with the number of disclosure growth by users higher than 6 and 7 "PolitiFact"

Features	Fi1	ArUn-ROC	Avg.Precision
News Content			
TF-IDF	0.811	0.888	0.781
Readability	0.642	0.682	0.629
Punctuation (LIWC)	0.704	0.766	0.636
Linguistic (LIWC)	0.719	0.787	0.650
Psychological (LIWC)	0.711	0.799	0.646
Summary (LIWC)	0.673	0.725	0.604
All LIWC	0.761	0.836	0.691
All News Content	0.848	0.874	0.771
News Headline			
TF-IDF	0.663	0.733	0.644
Readability	0.539	0.560	0.565
Punctuation (LIWC)	0.644	0.727	0.644
Linguistic (LIWC)	0.660	0.725	0.605
Psychological (LIWC)	0.635	0.676	0.574
Summary (LIWC)	0.657	0.705	0.600
All LIWC	0.722	0.791	0.654
All Headline	0.845	0.816	0.752
Image			
NeuralTalk2	0.743	0.600	0.725
Bias	0.854	0.884	0.917
Social Network	0.738	0.627	0.731
All	0.858	0.880	0.779

Table 7. RandomForest Classification Results with multi-modal features.

Refer to Table 8, Title, bias, image features, and social features combined to see if this further improves the detection of illusive news. The results show that certain feature sets are effective for categorizing political news articles as factual or not. Feature Prejudice plays an important role in detecting illusive news and the 2nd highest crucial aspect is the title. On the other hand, Shearer and Gottfried (2017) show that titles are more informative than content (78% vs. 71% accuracy). The results show that instead of "reading" news articles for validity, considering news metadata such as titles, biases, social networks, and images can achieve comparable or even higher performance (0.90 Ar-Un-ROC versus 0.88). Therefore, viewing news snippets with the characteristics of the title in mind, checking for publisher Prejudice and title keywords, and paying close attention to related images provide an efficient tool for detecting illusive news. If this signal can be thought of by humans, it is hoped that this research can avert the community from growing non-accurate disclosure massively over online soc-media.

Table 8: Measurements of Fi1, Ar-Un-ROC, and mean accuracy results with a compound of prejudice, title, pict, and social features.

Features	Fi1	ArUn-ROC	Avg.Precision
Title + Content + Prejudice + Pict + Social	0.858	0.880	0.779
Title + Prejudice + Pict + Social	0.865	0.901	0.786
Title + Content + Prejudice + Pict	0.860	0.879	0.777

Table 9. Linear SVM Classifier results with multi-modal features

Features	Fi1	ArUn-ROC	Avg.Precision
Title + Content + Prejudice + Pict + Social	0.814	0.797	0.802
Title + Prejudice + Pict + Social	0.817	0.809	0.796
Title + Content + Prejudice + Pict	0.824	0.803	0.805
Title + Content + Prejudice + Social	0.814	0.802	0.802
Title + Content + Pict + Social	0.789	0.749	0.788
Content + Prejudice + Pict + Social	0.833	0.826	0.821
Content + Prejudice + Pict	0.846	0.825	0.827
Title + Content+ Social	0.780	0.720	0.815
Title + Prejudice + Social	0.835	0.821	0.814
Prejudice + Pict + Social	0.835	0.873	0.876

Table 10. Measurements of Fi1, Ar-Un-ROC, and mean accuracy results with a compound of prejudice, title, pict, and social features.

Features	Fi1	ArUn-ROC	Avg.Precision
News Content			
TF-IDF	0.818	0.875	0.791
Readability	0.585	0.639	0.596
Punctuation (LIWC)	0.671	0.708	0.606
Linguistic (LIWC)	0.684	0.729	0.620
Psychological (LIWC)	0.695	0.735	0.632
Summary (LIWC)	0.637	0.678	0.581
All LIWC	0.729	0.780	0.667
All News Content	0.812	0.773	0.794
News Headline			
TF-IDF	0.672	0.730	0.654
Readability	0.573	0.593	0.591
Punctuation (LIWC)	0.640	0.742	0.653
Linguistic (LIWC)	0.608	0.640	0.568
Psychological (LIWC)	0.607	0.628	0.573
Summary (LIWC)	0.551	0.555	0.529
All LIWC	0.675	0.720	0.626
All Headline	0.785	0.704	0.772
Image			
NeuralTalk2	0.721	0.670	0.761
Bias	0.843	0.878	0.890
Social Network	0.444	0.538	0.690
All	0.814	0.797	0.802

CONCLUSIONS AND RECOMMENDATIONS

In this study, the proportionate effect of various news manner as “body”, “titles”, “Source Prejudice”, “visual content”, and “social networks” is analyzed in discriminating illusive political news. Especially, Source Prejudice has not been analyzed before, and the detention of this research shows that the relation between news plausibility and political Prejudice is very strong. In addition, for the validity assessment problem here it is proven that the analysis of news bodies is unnecessary because it is time-consuming for the users. Comparable results can be obtained by considering different methods such as title features, authority prejudice, and visual content.

One of the qualifications of this research is the size of the data set considered, but no other currently available data set contains all the information on the four modalities considered. As such, collecting larger datasets would be of great help to refine the investigation as a future study. Furthermore, by extracting pulse from the images of news, one may achieve better achievement in analyzing illusive news, because guide examination of images in the data set shows that images associated with illusive news are more emotional than factual news. A cross-domain efficiency test of alternative news modalities was tested because so far it has only been investigated for content news content (Gupta et al, 2012). From social networks, user credibility and network-based features such as network diffusion are analyzed and estimated.

Suggestions for future research

- For future research, it is highly recommended to analyze user-based features similarly to user figure, user assumption, and also the post based on features representing users' social responses in attitude, topic, etc. as this will help achieve better results.
- It is also recommended to find out impressive features and early detection figure of the hoax on social media and websites because hoax commonly develops quickly on socmed as well as on free websites.

REFERENCE

Abhijnan Chakraborty, Bhargavi Paranjape, Sourya Kakarla, and Niloy Ganguly. Stop click baits: Detecting and preventing click baits in online news media. In 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2016, San Francisco, CA, USA, August 18-21, 2016, pages 9–16, 2016.

Abhishek Bhati, Philip Pearce. (2016). Tourist attractions in Bangkok and Singapore; linking vandalism and setting characteristics. http://creativecommons.org/licenses/by/4.0/

Ankesh Anand, Tanmoy Chakraborty, and Noseong Park. We use neural networks to detect clickbait: You won't believe what happened next! In Advances in Information Retrieval - 39th European Conference on IR Research, ECIR 2017, Aberdeen, UK, April 8-13, 2017, Proceedings, pages 541–547, 2017.

Benjamin D Horne and Sibel Adali. 2017. This is just in: fake news packs a lot in title, uses simpler, repetitive content in the text body, more similar to satire than real news. arXiv preprint arXiv:1703.09398 (2017).32

Borella, CA, & Rossinelli, D. (2017). Fake news, Immigration, and Opinion Polarization. http://essuir.sumdu.edu.ua/handle/123456789/66314

Chau, et al. (2011). "Polonium: Tera-scale graph mining and inference for malware detection." In Proceedings of the 2011 SIAM International Conference on Data Mining, pp. 131-142. Society for Industrial and Applied Mathematics, 2011.

Cozens, P., Saville, G., & Hillier, D. (2005). Crime prevention through environmental design (CPTED): A review and modern bibliography. Property Management, 23(5), 328e356. http://dx.doi.org/10.1108/02637470510631483 .

Craig Silverman. 2016. This analysis shows how viral fake election news stories outperformed real news on Facebook. BuzzFeed News 16 (2016).

D. Zisiadis, S. Kopsidas, A. Varalis, and L. Tassiulas, “Mailbook: A social network against spamming,” in Proceedings of the 2011 International Conference for Internet Technology and Secured Transactions, ICITST 2011, pp. 245–249, are December 2011.

Ekblom, P. (2011a). Deconstructing CPTED and reconstructing it for practice, knowledge management, and research. European Journal on Criminal Policy and Research, 17(1), 7e28. http://dx.doi.org/10.1007/s10610-010-9132-9 .

Fyall, A., Garrod, B., Leask, A., & Wanhill, S. (2008). Conclusion. In A. Fyall, B. Garrod, A. Leask, & S. Wanhill (Eds.), Managing visitor attractions (2nd ed., pp. 347e353). Oxford: Butterworth-Heinemann.

James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of LIWC 2015. Technical Report.

Jin, et al. (2017). "Novel visual and statistical image features for microblog news verification." IEEE transactions on multimedia 19, no. 3 (2017): 598-608.

Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter 19, 1 (2017), 22–36.

Kai Shu, Suhang Wang, and Huan Liu. 2017. Exploiting Tri-Relationship for Fake News Detection. arXiv preprint arXiv:1712.07709 (2017).

Maksym Gabielkov, Arthi Ramachandran, Augustin Chaintreau, and Arnaud Legout. 2016. Social clicks: What and who gets read on Twitter? ACM SIGMET- RICS Performance Evaluation Review 44, 1 (2016), 179–192.

Manish Gupta, Peixiang Zhao, and Jiawei Han. 2012. Evaluating event credibility on Twitter. In Proceedings of the 2012 SIAM International Conference on Data Mining. SIAM, 153–164.

Martin Potthast, Sebastian Ko¨psel, Benno Stein, and Matthias Hagen. Clickbait detection. In Advances in Information Retrieval - 38th European Conference on IR Research, ECIR 2016, Padua, Italy, March 20-23, 2016. Proceedings, pages 810–817, 2016.

Natali Ruchansky, Sungyong Seo, and Yan Liu. CSI: A hybrid deep model for fake news detection. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM 2017, Singapore, November 06 - 10, 2017, pages 797–806, 2017.

Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. 2017. Show and Tell: Lessons Learned from the 2015 MSCOCO Image Captioning Challenge. IEEE Trans. Anal Pattern. Mach. Intell. 39, 4 (2017), 652–663.

Potthast, et al. (2017). "A Stylometric Inquiry into Hyperpartisan and Fake news." arXiv preprint arXiv:1702.05638 (2017).

Shearer and Gottfried. 2017. News Use Across Social Media Platforms 2017. (2017).

Symeon Papadopoulos Duc-Tien Dang-Nguyen Giulia Boato Michael Riegler Yiannis Kompatsiaris et al. Christina Boididou, Katerina Andreadou. 2015. Verifying Multimedia Use at MediaEval 2015. In MediaEval.

Tanushree Mitra and Eric Gilbert. 2015. CREDBANK: A Large-Scale Social Media Corpus With Associated Credibility Annotations. In Proceedings of the Ninth International Conference on Web and Social Media, ICWSM 2015, University of Oxford, Oxford, UK, May 26-29, 2015. 258–267.

William Yang Wang. 2017. "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 2: Short Papers. 422–426.

Wilson, JQ, & Kelling, GL (1982). Broken windows: Police and neighborhood security. Monthly Atlantic, 249, 29e38.

Wortley, R., & Mazerolle, L. (2012). Environmental criminology and crime analysis. Hoboken: Taylor and Francis.

Xin Liu, Pingjun Zou, Weishan Zhang, Jiehan Zhou, Changying Dai, Feng Wang, Xiaomiao Zhang, "CPSFS: A Credible Personalized Spam Filtering Scheme by Crowdsourcing", Wireless Communications and Mobile Computing, vol. 2017, Article ID 1457870, 9 pages, 2017. https://doi.org/10.1155/2017/1457870

Yimin Chen, Niall J. Conroy, and Victoria L. Rubin. Misleading online content: Recognizing clickbait as ”false news”. In Proceedings of the 2015 ACM Workshop on Multimodal Deception

Detection, WMDD@ICMI 2015, Seattle, Washington, USA, November 13, 2015, pages 15–19, 2015.

Zhang, Y., Guan, J., Zhou, S., & Zhang, Z. (2013). Rumor evolution in social networks. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 87 (3), Article 032133. (2013). http://discovery.ucl.ac.uk/1387962/1/PhysRevE.87.032133.pdf

Zhiwei Jin, Juan Cao, Yongdong Zhang, and Jiebo Luo. News verification by exploiting conflicting social viewpoints in microblogs. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA., pages 2972–2978, 2016.