ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

A Survey on Satirical News Detection and Analysis using Attention Mechanism and Linguistic Features

Gayathri G. Nair[1], Raseek C[2]

PG Student, Dept of Computer Science and Engineering, Government Engineering College, Palakkad, Kerala, India[1]

Assistant Professor, Department of Computer Science and Engineering, Government Engineering College, Palakkad, Kerala, India[2]

Abstract: Satirical news is a type of parody presented in the format of news. The word satire itself means the use of humor, irony, exaggeration, or ridicule to expose and criticize peoples stupidity or vices. The satirical news is particularly used in the context of contemporary politics and other topical issues. It is a double edged sword, on one hand it is entertainment and on the other hand it is deceptive and harmful. And it could be potentially misleading, in either of the following cases. First one is Satirical cues are too subtle to be recognized. And the other one is reader lack the contextual or cultural background. Satirical news detection is considered as a problem since not everyone can recognize it as a satire one. And the spreading of the false news may lead to hurt the credibility and trust in the social media websites. Existing works only consider document level features to detect the satire, which could be limited. This paper considers all paragraph-level linguistic features to find the satire by incorporating neural network and attention mechanisms. There exist different methods to find out whether the article is satire or not. In this paper the comparative study on the different approaches for the detection of the satire is performed. Its also investigate the difference between paragraph-level features and document-level features.

Keywords: Satirical News Detection, Natural Language Processing (NLP), Deep Learning, Attention Mechanism

I. INTRODUCTION

In the epoch, the use of social media sites like Facebook, twitter is increasing day by day. Throughout the whole social media, the leverages of different kind of news becoming resilient on their platforms. The audience for the identical platform of news has associate large quantity of increase in the recent years. We all know that even when the true news spreads across the social media platforms, there is a huge amount of hoax news are also getting in them. Fake news is the category which contain false information. The fake news aims to cheat a person who reads the one. The fake news can be any of the subsequent category: Sarcasm, Satire, Humor, Rumor, etc. This paper finds out the existence of satire in the context of fake news. That is to avoid the fake news spreading in the social media platforms, as an initial step it is needed to find out the satirical news. Before aiming to the satirical news detection we have to know what is the word satire means. Satire is a powerful art form which has the ability to point out the deficiencies in certain human behaviors and the social issues which result from them in such a way that they become absurd, even hilarious, which is therefore entertaining and reaches a wide audience. So the goal of this study is to compare totally different satirical news detection ways.

Satirical news is taken into account to be entertainment. However, its tasking acknowledge the recognize the satire if the satirical cues are too subtle to be unmasked and the reader lacks the contextual or cultural background. Assuming readers interpret satirical news as true news, there is not much difference between satirical news and fake news in terms of the consequence, which may hurt the credibility of the media and the trust in the society. In fact, it is reported in the Guardian that people may believe satirical news and spread them to the public regardless of the ridiculous content[1]. It is also concluded that fake news is similar to satirical news via a thorough comparison among true news, fake news, and satirical news[2]. This paper focuses on the study of satirical news detection to make sure the trustworthiness of online news and forestall the spreading of potential misleading information. On social networks, the reach and effects of information spread occur at such a quick pace and so amplified that distorted, inaccurate or false information acquires a tremendous potential to cause real world impacts, within minutes, for millions of users. Recently, many

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

public concerning this downside and a few approaches to mitigate the problem were expressed. In the era of the Internet, online journalism is now a common practice. Online news articles have a significant contribution in keeping people informed about what is happening in the world. The usage of web to spread news comes with the disadvantage of deception. The presence of deceptive and misleading news articles has been around for a while. Although some news articles often have a disclaimer about it being fake, many other don't and thus readers could be led to believe them to be true. This leads to spread of misinformation, which may also start off a rumour. The importance of the detection of deceptive news is increasing rapidly, as more and more people start relying on online news as their major source of news.

The task of satirical news detection is very extremely advanced. Within the daily task we all know that to decide the judgment on the value of information about a news is not performed by machine. The machine solely decide where and to whom the news may be displayed and them holds on the control of the spreading of news, not the content of them. Freedom of speech must be protected in the least value, and that includes taking into consideration distinct types of publications, which may be humoristic, sarcastic, Satirical or just conveying simple opinions on something, even if they are only based on personal beliefs. In this paper it describes different approaches for the task of satirical news detection. Since the satirical clues and spreads across the document it is important to consider both the paragraph level features rather than considering the document level features.

This paper is organized as follows: Section II gives a formal definition of the satirical news detection. Section II also discusses about various classification (machine learning) approaches, feature based approaches and neural network approaches for satirical news detection and provides a comparison between them. Section III gives brief concluding comments.

II. SATIRICAL NEWS DETECTION SYSTEM

The goal of the satirical news is to cheat someone by provide false information and which is represented in a comedial way. The task of understanding the honestness of a news is the main objective of the problem. Figure 1 describes the over all intention of the proposed system to find out whether it is a satire or not. The problem of satirical detection can be formally defined as: Given a news document N, the goal of satirical news detection is to model and discover the satire content in N and to classify it is either satire or not.

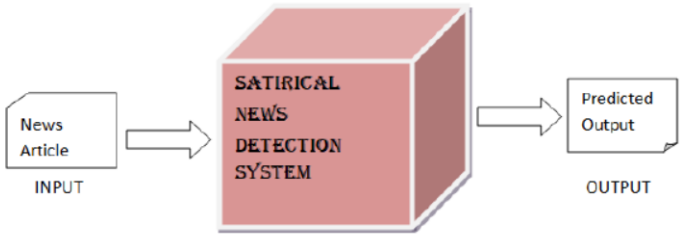

Fig. 1. Overall of the proposed system

A. Classification Based on Approaches

There are different approaches for the detection of satirical news. We categorize related works into four categories: content-based detection for news genre, truth verification and truthfulness evaluation, deception detection, and identification of highly attended component using attention mechanism. Figure 2 shows an overall classification of the methodology.

Burfoot et. al.[2]introduces the novel task of deciding whether a news wire article is true or satirical. The system experiment with SVMs, feature scaling, and a number of lexical and semantic feature types, and achieve promising results over the task. This paper describes a way for filtering satirical news articles from true news wire documents. They outline a satirical article as one which collectively exposes real-world individuals, organizations and events to ridicule. The contributions of this analysis are: Firstly they introduces a novel task to the podium of computational linguistics and machine learning, and make available a standardized data set for research on satire detection; and

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

secondly the authors develop a method which is adept at identifying satire based on simple bag-of words features, and further extend it to include richer features.

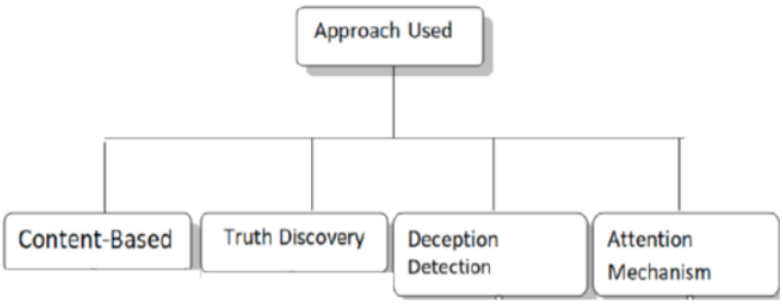

Fig. 2. Classification Based on Approaches

The planned classification in [2] supplements the bag of-words model with feature weighting, using the two methods: Binary feature weights, Bi-normal separation feature scaling[2]. Under Binary feature weights scheme all features are given the same weight, regardless of how many times they appear in each article. The topic and sentiment classification examples cited found binary features gave better performance than other alternatives. In Bi-normal separation feature scaling: BNS [2,4] has been shown to outperform other established feature representation schemes on a wide range of text classification tasks. This superiority is especially pronounced for collections with a low proportion of positive class instances. BNS produces the highest weights for features that are strongly correlated with either the negative or positive class. Features that occur evenly across the training instances are given the lowest weight. This behaviour is particularly helpful for features that correlate with the negative class in a negatively-skewed classification task, so in their case BNS should assist the classifier in making use of features that identify true articles. The authors aim in [3] is to see whether an author has tried to hide his writing style in a written document[3]. In the ancient authorship recognition, authorship of a document is decided using linguistic features of an authors writing style. In deceptive writing, when an author is deliberately hiding his regular writing style, authorship attribution[3] fails as a result of deceptive document lacks stylistic similarity with the authors regular writing style. Though identifying correct authorship of a deceptive document is hard, the goal is to see if it is possible to discriminate deceptive documents from regular documents. To detect adversarial writing, it need to identify a set of discriminating features that distinguish deceptive writing from regular writing. After determining these features, supervised learning techniques can be used to train and generate classifiers to classify new writing samples.

The performance of stylometry methods depends on the mix of the selected features and analytical techniques. They explored three feature sets to spot stylistic deception. This work[4] analyse the information credibility of news propagated through Twitter, a popular micro blogging service[4]. Previous research has shown that most of the messages posted on Twitter are truthful, but the service is also used to spread misinformation and false rumours, often unintentionally. On this paper they specialise on automatic methods for assessing the credibility of a given set of tweets. Specifically, analyse micro blog postings related to trending[4][5] topics, and classify them as credible or not credible, based on features extracted from them. They use features from users posting and re-posting (re-tweeting) behaviour, from the text of the posts, and from citations to external sources. The main hypothesis is that the level of credibility of information disseminated through social media can be estimated automatically. They believe that there are several factors that can be observed in the social media platform itself, and that are useful to asses information credibility. These factors include:

- The reactions that certain topics generate and the emotion conveyed by users discussing the topic: e.g. if they use opinion expressions that represent positive or negative sentiments about the topic.
- The level of certainty of users propagating the information: e.g. if they question the information that is given to them, or not.
- The external sources cited: e.g. if they cite a specific URL with the information they are propagating, and if that source is a popular domain or not.
- Characteristics of the users that propagate the information, e.g. the number of followers that each user has in the platform.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

In Fan Yang et.al.[4],they first present their 4-level hierarchical neural network[4][6] and make a case how linguistic features can be embedded in the network to reveal the difference between paragraph level and document level. Then they describe the linguistic features. The authors build the model in a hierarchy of character-word-paragraph-document. The general overview of the model can be viewed in Figure 3.

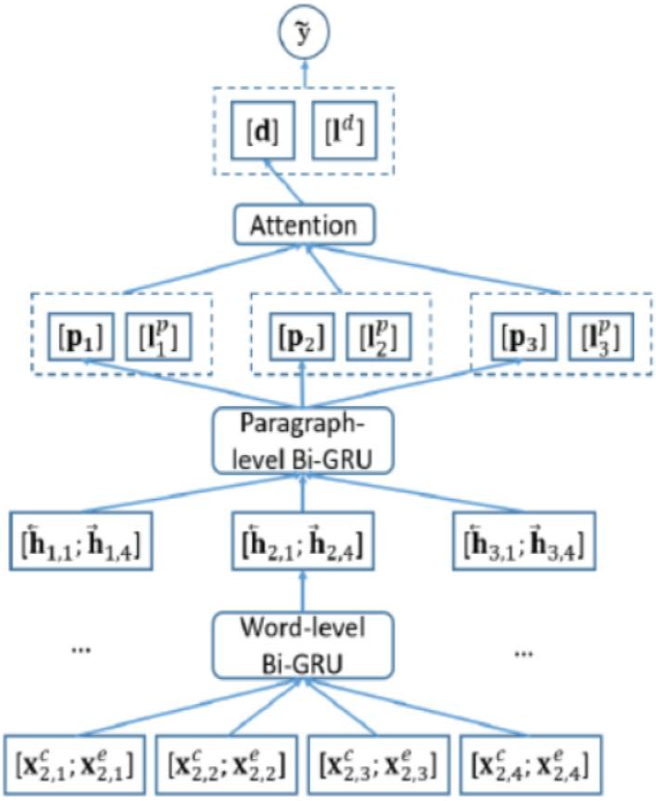

Fig. 3. Overview of the model used by[2]

Character-Level Encoder[1][7] use Convolutional Neural Networks (CNN) to encode word representation from characters. CNN is effective in extracting morphological information and name entities (Ma and Hovy, 2016), both of which are common in news. Every word is presented as a sequence of n characters and each character is embedded into a low-dimension vector. The sequence of characters c is brought to the network. A convolution operation with a filter Wc is applied and moved along the sequence. Max pooling is performed to select the most important feature generated by the previous operation.

During peer review process or early stage of research paper writing, metadata such as author name, year of publication, venue, etc. are not available. This content based method [2] for citation recommendation which remains robust when metadata are missing for query documents. This system introduces a model for recommending citation called citeomatic. Which is useful for the researchers to find the existing work of research topic. The input to this model can be query document or URI or details of paper which includes metadata.

B. Classification Based on Features

This section describes different types of features supposed to embody characteristics of satire news documents. Burfoot et.al[2] uses three different types of features as, headline features, Profanity and Slang. Most of the articles in the corpus have a headline as their first line. To a personality's reader, the overwhelming of the satire documents in their corpus are immediately recognizable as such from the headline alone, suggesting that their classifiers may get something out of having the headline contents explicitly identified in the feature vector. To this current finish, they add an additional feature for each unigram appearing on the first line of an article. In this way the heading tokens are represented twice: once in the overall set of unigrams in the article, and once in the set of heading unigrams.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

Profanity:[2] True news articles terribly include a verbal quote which contains offensive language, but in practically all other cases it is incumbent on journalists and editors to keep their language clean. A review of the corpus shows that this is not the case with satirical news, which occasionally uses profanity as a humorous device. Slang:[2] As with profanity, it is intuitively true that true news articles tend to avoid slang. An impressionistic review of the corpus suggests that informal language is much more common to satirical articles.

The performance of stylometry ways depends on the mix of the selected features and analytical techniques. The authors in [3]explored three feature sets to identify stylistic deception. Zheng et al. proposed the Write prints features that can represent an authors writing style in relatively short documents, especially in online messages[3]. These sink features are not unique to this work, but rather represent a superset of the features used in the stylometry literature. The proposed system used a partial set of the Whiteprints features. Their adaptation of the Write prints features consists of three kinds of features: lexical, syntactic, and content specific.

Lexical features: These features include both character based and word-based features[3]. These features represent an authors lexicon-related writing style: his vocabulary and character choice. The feature set includes total characters, special character usage, and several word-level features such as total words, characters per word, frequency of large words, unique words. Syntactic features: Each author organizes sentences differently. Syntactic features represent an authors sentence-level style. These features include frequency of function words, punctuation and parts-of-speech(POS) tagging. They use the list of function words from LIWC 2007 [6]. Content Specific features: Content specific features refer to keywords for a specific topic. These have been found to improve performance of authorship recognition in a known context [1].

The authors in Castello et. al[4] includes four different types of features. The Message-based features consider characteristics of messages, these features can be Twitter independent or Twitter dependent. Twitter-independent features include: the length of a message, whether or not the text contains exclamation or question marks and the number of positive/negative sentiment words in a message. Twitter-dependent features include features such as: if the tweet contains a hash tag, and if the message is a re-tweet. User-based features consider characteristics of the users[6] which post messages, such as: registration age, number of followers, number of followees (friends in Twitter), and the number of tweets the user has authored in the past.

Topic-based features are aggregates computed from the previous two feature sets; for example, the fraction of tweets that contain URLs, the fraction of tweets with hash tags and the fraction of sentiment positive and negative in a set. Propagation-based features consider characteristics related to the propagation tree that can be built from the retweets of a message. These includes features such as the depth of the re-tweet tree, or the number of initial tweets of a topic (it has been observed that this influences the impact of a message, e.g. in [5]).

Linguistic features have been successfully applied to expose differences between deceptive and genuine content, so subsume most of the features in previous works. The idea of explaining fictitious content is extended here to reveal how satirical news differs from true news. They divide the linguistic features into four families and compute them separately for paragraph and document. Psycholinguistic Features: Psychological differences are useful for problem, because professional journalists tend to express opinion conservatively to avoid unnecessary arguments. On the contrary, satirical news includes aggressive language for the and accuracy while satirical news is related to emotional cognition. To capture the above observations, employ Linguistic Inquiry and Word Count (LIWC) (Pennebaker et al., 2007) as psycholinguistic dictionary. Each category of LIWC is one independent feature and valued by its frequency.

CONCLUSION AND FUTURE WORK

The various methods for the satirical news detection is studied in its resilient means. And classified the previous works into four categories: content-based detection for news genre, truth verification and truthfulness evaluation, deception detection, and identification of extremely attended component using attention mechanism. Even though the satire detection had been studied widely, the same system has to improved to include more features. The satire, sarcasm, humour, rumour and irony detection can be offer a helping hand to the fake news detection which is an emerging and interesting field in Natural language processing. This study leads and acts as a baseline approach to the fake news detection downside.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

PCON-2019

National Conference and Seminar on Innovations in Engineering & Technology

Govt Polytechnic College, Palakkad

Vol. 8, Special Issue 1, January 2019

REFERENCES

[1]. Fan Yang, Arjun Mukherjee and Eduard Gragut, Satirical News Detection and Analysis using Attention Mechanism and Linguistic Features, Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, September 711, 2018. Association for Computational Linguistics.
[2]. Clint Burfoot and Timothy Baldwin. 2016. Automatic satire detection: Are you having a laugh? In Proceedings of the ACL-IJCNLP 2013 conference short papers. Association for Computational Linguistics.
[3]. Sadia Afroz, Michael Brennan, and Rachel Greenstadt.2014. Detecting hoaxes, frauds, and deception in writing style online. In 2012 IEEE Symposium on Security and Privacy. IEEE.
[4]. Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2016. Information credibility on twitter. In Proceedings of the 20th international conference on World wide web. ACM.
[5]. Francesco Barbieri, Francesco Ronzano, and Horacio Saggion. 2015. Do we criticise (and laugh) in the same way? automatic detection of multi-lingual satirical news in twitter. In IJCAI.
[6]. Yimin Chen, Niall J Conroy, and Victoria L Rubin. 2015. Misleading online content: Recognizing clickbait as false news. In Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection. ACM.
[7]. Rada Mihalcea and Carlo Strapparava. 2009. The lie detector: Explorations in the automatic recognition of deceptive language. In Proceedings of the ACL-IJCNLP 2009 Conference Short Papers. Association for Computational Linguistics.
[8]. Martin Potthast, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff, and Benno Stein. 2017. A stylometric inquiry into hyperpartisan and fake news.arXiv preprint arXiv.
[9]. Victoria Rubin, Niall Conroy, Yimin Chen, and Sarah Cornwell. 2016. Fake news or truth? using satirical cues to detect potentially misleading news. In Proceedings of the Second Workshop on Computational Approaches to Deception Detection, San Diego, California. Association for Computational Linguistics.
[10]. Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016b. Hierarchical attention networks for document classification. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.
[11]. Antonio Reyes, Paolo Rosso, and Davide Buscaldi.2015. From humor recognition to irony detection: The figurative language of social media. Data Knowledge Engineering, 74.
[12]. Tim Rocktaschel, Edward Grefenstette, Karl Moritz Herman, and Phil Blunsom. 2015. Reasoning about entailment with neural attention. arXiv, preprint arXiv.