Proceedings of the 3rd International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/5/20230619

Detecting fake news with deep neural networks

Qiaoyi Huang

Chinese University of Hong Kong, Shenzhen, Guangdong, China

121090208@link.cuhk.edu.cn

Abstract. The raise of social media and the easy access to obtain and spread news has increased the probability of various news dissemination. However, an undetermined factor in usage scenario is the authenticity of spreading news, which has posed detrimental effects on individuals and society in current scenario. It is imperative to develop proper mechanism for fake news detection. These years, with the sharp development of deep learning, referring to the method of training multi-layer artificial neural networks, some researchers have utilized deep learning in fake news detection and classification. Although many studies have achieved a preeminent level and high accuracy, scarcity of review work regarding fake news detection with deep learning is still an important problem. This work urges to derive a comprehensive review of fake news detection and classification techniques for further study in the aspect, and mechanism in various studies for sake of raising accuracy. This work also figure out some possible challenges with existing limitation based on previous research.

1. Introduction

The easy dissemination of information feature of social media and networks has added to exponential growth of various news. For instance, prevalent social networks like Twitter, Facebook, and Instagram trigger the proliferation of commercial advertisements, grapevine news and politics decisions. However, falsification of information also raise with spreading of news. The authenticity of news on social networks is at stake where some people intentionally create or disseminate fake news for sake of reputation or advertisement income, which increase intricacy to detect. In the current scenario, usage of social media networks is prevalent and user number is enormous. People increasingly tend to believe news from social media than from traditional news sources such as newspaper or authority dissemination [1]. According to data from datareportal.com, Facebook had 2.934 billion monthly active users in July 2022. With a large amount of audience and thus large spreading scale, fake news easiliy obtain more supporters and dissemination, which may also lead to further detrimental effects on the society. Meanwhile, in reality, fake news is intricate to detect by audience themselves since the way people respond to facts will be changed by the manipulation and confusion of fake news [2]. Hence, there is a pressing necessary for fake news detection and classification for social networks. Application of fake news detection mechanism can be given in two taxonomy.

Social networks platform can apply automatic detection as a precaution for audience. For instance, for sake of decreasing risk of economic loss. In recent scenario, fake news poses influence on stock market. CHEQ's global reporting with Roberto Cavazos, an executive at the University of Baltimore Cavazos, estimates that marketers tend to lose $1.3 billion this year in online advertisement spending because of news falsification. The post “Can fake news affect the stock market?” figures out that

DOI: 10.54254/2755-2721/5/20230619

deceptive news can impact the stock market. In 2013, 130 billion USD in stock value was wiped out in minutes under the influence of an AP tweet about an "explosion" that injured Barack Obama. The facts show that fake news may strongly influence economy development or cause enormous loss. With application of automatic fake news detection, social networks users can get rid of these kinds of loss more easily.

Numerous studies in the past have used computational methods like natural language processing (NLP) to find anomalies that distinguish between information that is factually accurate and deceptive literature [3]. Natural language processing covers grammatical analysis, textual analysis, discourse interpretation, etc. when seen from the perspective of study material. Natural language processing has a diverse variety of applications from a practical standpoint. Natural language processing is used in a wide variety of contexts today, including machine translation, handwritten and printed character identification, speech recognition, text-to-speech transformation, information retrieval, information extraction, and information filtering, text classification, and opinion mining. Data mining, machine learning, knowledge engineering, and artificial intelligence research all require language processing. When researchers derive several feature clusters, they can utilize NPL in fake news detection by classifying semantic and textual content.

Based on former prepossessing, researchers utilize machine learning methods in fake news detection. Machine learning mechanisms include K-nearest neighbors (KNN), support vector machine (SVM), decision tree (DT), and stochastic gradient descent (SGD), achieving the highest accuracy (92%) with SVM and logistic regression. KNN refers to the mechanism of classification where given a training data set, for a new input instance, find the K closest to the instance (the K neighbors mentioned above) according to some self developed features in the training dataset [4]. The input instance will be than assigned to the class of which most of nearest neighbors referred to. SVM refers to a generalized linear classifier utilizing supervised learning to deal with dual classification with the decision boundary of maximum margin hyperplane for solving the learning sample. It can utilize kernel method to deal with nonlinear regression, which is also an important and prevalent method of kernel learning.

In the prepossessing of classification, the features of the task are constructed manually, and then some machine learning algorithms are used to classify the task. However, there are two problems in these methods. Firstly, the construction of task features depends on people's background knowledge, so the quality of features is unstable. Most of research focus on linguistic content of news without considering internal factors including education level, cultural background, latent value, etc.. Secondly, these features are shallow and mostly based on statistical methods, which cannot well describe the deep semantic information.

Deep learning can raise its accuracy automatically with a large scale of data while machine learning methods requires intricate methods to improve performance. The figure below can show the comparison between deep learning and traditional machine learning mechanisms. When considering fake news detection and classification, a large amount of data containing textual and semantic information can be passed to deep learning algorithm to improve performance. Deep learning fit the feature of large data scale of fake news detection.

In order to create complex features for traditional machine learning methods, reseachers first undertake a thorough exploratory data analysis on the dataset. This is followed by a straightforward dimensionality reduction procedure. The machine learning algorithm must be given the best functionality, which must be carefully chosen. When employing deep learning networks, this is not necessary as good performance is frequently obtained by simply feeding the data to the network. This entirely eliminates the time-consuming and difficult feature engineering phase from the process. Since latent audience attitudes, semantic content, and textual material are present in the data set, there will be a substantial number of complicated features throughout the fake news detection and classification process [5].

Compared to conventional machine learning methods, deep learning approaches are more easily applied to many domains and applications. Transfer learning makes it possible for deep networks that

DOI: 10.54254/2755-2721/5/20230619

already exist to be applied to other applications within the same area. For instance, feature extraction front-ends for object identification and segmentation networks frequently make use of pre-trained image classification networks. The training of the entire model is facilitated by using these pre-trained networks as front-ends, which frequently enables superior performance in a shorter amount of time.

Furthermore, the fundamental concepts and methods of deep learning applied in other fields tend to be extremely portable. Learning how to apply deep networks to natural language processing once the fundamental deep learning theory in the field of speech recognition is understood is not too difficult because the baseline knowledge is fairly comparable. Contrary to popular belief, classical machine learning involves feature engineering and domain- and application-specific machine learning approaches in order to develop high-performance machine learning models. For many domains and applications, the knowledge base of classical machine learning is quite variable, and frequently necessitates in-depth specialist research in each specific field. As a result, deep learning offers a considerable advantage when solving linguistic issues and performs better when classifying semantic content, which satisfies the demand for false news detection.

Although the deep neural network is applied to fake news detection, it solves many problems and improves the performance of the task. However, there are very few reviews on this task. As a result, many beginners are unable to get a good understanding of the latest progress of the task and what the latest methods have been proposed. In addition, the review can be seen as a phased summary that contributes to the development of subsequent research-based work. Hence, we combed and summarized the development of this task in recent years. We distinguish the definition of fake news detection tasks, and then introduce the commonly used data sets and evaluation metrics of tasks. In addition, we explain the mainstream deep learning methods used for the task. Some of the latest approaches and potential challenges was covered in this paper.

2. Background

2.1. Task Definition

Several possible taxonomy of fake news is shown as below [5]:

Clickbait: includes news articles that are produced to raise interest of audience where authenticity can not be promised, which are to increase web clicks and popularity to these pages and increase revenue from third party organizations by receiving advertisements.

Propaganda: This kind of news usually has a certain political motivation, which tend to raise the user's agenda and influence the user's position.

Comments: The author combines some facts to prove their points and to influence the reader's opinion about some recent event. For example, campaigning for an opinion in a public event in order to support one side.

Satire: These articles are for entertaining purposes, but the language has features such as irony, which cannot be machine learned with simple linguistic features, and can easily cause a larger workload.

Rumor: Usually created during a public event, it is a mixture of true and false information and may be used to cause anxiety or promote behavior to achieve a certain goal. For instance, spreading news of knockdowns during the pandemic encouraged people to buy supplies in order to drive up prices and benefit from it.

2.2. Taxonomy Based on Possessing Type

Vision-based: These news items use additional graphics, such as modified photographs, edited videos, or a combination of both, as content [6].

User-based: These false reports are created by fictitious accounts and are directed towards a particular group of people. They may have been produced with the intention of influencing particular demographics based on age ranges, worldviews, cultural norms, and nationalities. Special elements or common sense may have been added to the news to increase its credibility [7].

DOI: 10.54254/2755-2721/5/20230619

Knowledge-based: These articles present a scientific explanation for a particular unresolved issue and persuade the reader that it is accurate. For instance, using disinfectants or natural medicines that raise blood sugar levels in the body can be useful in treating sickness.

2.3. Datasets

News Feed on Buzz: Facebook posts from the week before the US presidential election are included in the dataset. A reporter verifies the accuracy of every post. Since multimedia content is not included in the dataset, it is only used for content-based detection techniques [8].

LIAR: The PolitiFact.com fact-checking website provided the dataset. There are 12836 phrasal phrases in it, and they are categorized into true, most true, half-true, almost true, false, and pants.

Grammar: The dataset includes 330 talks worth of linguistically similar Twitter exchanges about rumor tweets.

2.4. Metrics

Accuracy is culculated based on confusion matrix as shown below, where TN, TP, FN, FP refer to true negative, true positive, false negative and false positive results:

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{1}$$

Precision is a measure of how many samples that are anticipated to be positive really turn out to be positive. Predicting a positive class can be done in one of two ways: either by doing a positive class prediction (TP) or by doing a negative class prediction (FP).

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Recall is the percentage of the positively anticipated positive class that actually is positively predicated. Personal understanding: The proportion of the population's actual positive classes out of all expected positive classes:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

3. Recent Researches

3.1. Fake News Detection based on 3V Properties Using Deep Learning

Nowadays, some users spread fake news and lead other users to believe them intentionally in order to gain link and reputation. These affects the environment of social media. To defect fake news, in terms of the user's profile, the paper offers some fresh information, like the user's collaboration with others in making a certain choice [9]. The spread of information and the diffusion process make it extremely difficult to identify these contents quickly, emphasizing the necessity for automatic false news detection. This study employs deep learning algorithms to identify bogus news on Facebook and in the Chrome environment with greater accuracy than prior learning. Three main reasons that fake news are relatively hard to detect. Firstly, the use of internet trigger the fast spread of fake news and also some influence of famous people or companies with high reputation will lead to the belief of some fake news. Secondly, fake news would try to mock true news by containing some true evidence to support some wrong results with more support from academician and researchers. Thirdly, fake news are spread through multiple posts. The paper tries to distinguish fake news using 3V property, which is velocity, volume and veracity. Previous research did not focus on volume of fake news and its influence, which also mainly used predictive models to do the classification. To achieve goal of detecting some characteristics of fake news, for detecting fake news in the Facebook platform on a user's home page, a fake news detection framework is suggested. The author gathered both true and false news from many profiles that are involved in news dissemination. For enhanced decision-making,

DOI: 10.54254/2755-2721/5/20230619

the gathered data is analyzed using a number of machine learning(ML) and deep learning-based tools and algorithms.

3.2. Fake News Detection based on UPFD Methods fusing Endogenous and Exougenous Information

The majority of fake news detection research mainly focus on the content of fake news or the surrounding exogenous factors that raise deceptive signals [10]. This paper focus more on preference of users and judge whether users tends to spread this fake news or not. The confirmation bias theory suggests that people tend to believe news that are consistent with their existing knowledge. In response to the limitations of a large number of studies on people's different preferences, this paper investigates the novel problem of using user preferences for fake news detection. A novel framework, UPFD, is proposed to simultaneously capture various signals from user preferences by joint content and graphical modeling. Previous fake news detection techniques is mainly consist of two main methods. Firstly, fact checking approach, which is the most direct and common approach, with the disadvatage of the large amount of need to distinguish evidence from domain experts. Secondly, the computational methods by using deep learning to detect fake news. This techniques only focus on news content or user exogenous context. Previous social or psychological research focus on user endogenous factors like preference. And the confirmation bias theory suggests that people tend to believe news that consistent with their existing knowledge.

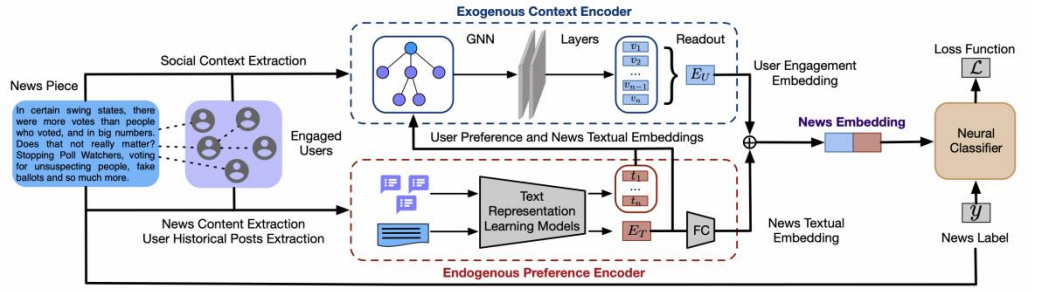

Figure 1. The architecture of the UPFD framework.

The user preference-aware UPFD framework for fake news detection. We extract the exogenous context from the news article and its engaged social media users as a news propagation graph and encode the endogenous data based on user historical posts and the news texts. A GNN encoder is used to combine the endogenous and exogenous data. The neural classifier receives the final news embedding, which is made up of user engagement embedding and news textual embedding, in order to predict the veracity of the news.

3.3. Fake News Detection Based on LSTM Mechanism

Fake news interrupts the domain building process. It is significant to develop an efficient and useful mechanism to detect fake news [11]. The study develops a language model that extracts syntactic, grammatical, emotional, and readability features of specific news to depict language-driven concepts. The study used a neural-based sequential algorithm to previously detect fake news. Finally, the average accuracy of fake news detection and classification based on linguistic feature driven model reaches 86%. The sequence of neural model results and the model utilizing machine learning and the fake news detection model based on the LSTM word embedded are compared. The comparison results show that the feature-based sequential model can achieve comparable evaluation performance in a short time. Features that require careful observation of news content are still under-explored. The researchers did not learn the importance of each specific characteristics. Along this line, this study developed features to determine whether new features should be used. In fact, learning model depends on the adequacy of the target data set. For example, by real-time streaming data collection for the particular scenario news may need a learning model based on memory. On the other hand, changing

DOI: 10.54254/2755-2721/5/20230619

and diverse content can provide better results for neural models based on semantic features. Therefore, the capability of the model depends on the content of the target dataset and its characteristics. This research use deep learning model developed its own linguistic features, and based on the comparison with deep learning model, measuring the effects of extracting features of language.

4. Challenges

Although deep learning has been utilized in fake news detection and had achieved high accuracy, there are still some latent problems for detection to be solved.

Language Uniqueness: Deep learning networks perform well in processing text recognition and translation, but each language has a unique form of expression, and the same text information may convey different meanings in different moods. For example, the use of implicit irony in Chinese. At the same time, language development has the characteristics of The Times, and different buzzwords may appear in different periods. Researchers need to constantly improve the learning data and content of deep learning networks. For example, the Chinese word "white" in the novel Coronavirus19 refers to medical staff, and "rebel" refers to a person who provides material medical support to people regardless of danger. These are both of The Times and cannot be easily understood.

User uniqueness: The same social platform users have different level of education, family background, cultural background, etc., for the same false news identification ability has a bigger difference. Some of the low degree of false news may not be deep learning algorithms to identify false news, but for the low level of knowledge of people still has the propagation force, influence, can cause bad effects to the social development. Therefore, it is still necessary to improve the accuracy and sensitivity of fake news identification, so as to stop before wider spreading or to take preventive methods.

5. Conclusion

The authenticity of the spread of information is one of the important indicators of the authority of social platforms, and it is also an important factor affecting people's life and behavior. In recent years, the spread of some false information has brought adverse effects to individuals and society. Therefore, it is imperative to establish a correct mechanism for fake news detection. Some researchers have used deep learning for the detection and classification of fake news as a result of the rapid development of this field. Despite the outstanding level and high accuracy attained by numerous research, there is still a need for more review work on deep learning-based fake news identification.This paper provides a comprehensive profile of false news detection and classification technologies, and summarizes the mechanisms of various types of research, expecting to provide reference for further research in this field and provide a more detailed description for new scholars entering this field. Meanwhile, based on previous research, the possible challenges under the existing constraints are pointed out. There are still many practical problems in the application of fake news identification mechanism to real social platforms, and people need to invest more research in this field.

References

[1] Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. ACM SIGKDD explorations newsletter, 19(1), 22-36.

[2] Guo, B., Ding, Y., Yao, L., Liang, Y., & Yu, Z. (2020). The future of false information detection on social media: New perspectives and trends. ACM Computing Surveys (CSUR), 53(4), 1-36.

[3] Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. Information Processing & Management, 57(2), 102025.

[4] Ismail, S., & Latif, R. A. (2013). Authenticity issues of social media: credibility, quality and reality.

[5] Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake news detection using machine learning ensemble methods. Complexity, 2020.

DOI: 10.54254/2755-2721/5/20230619

[6] Jain, A., & Kasbe, A. (2018, February). Fake news detection. In 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) (pp. 1-5). IEEE.

[7] Kaliyar, R. K., Goswami, A., Narang, P., & Sinha, S. (2020). FNDNet – a deep convolutional neural network for fake news detection. Cognitive Systems Research, 61, 32-44.

[8] Shu, K., Cui, L., Wang, S., Lee, D., & Liu, H. (2019, July). defend: Explainable fake news detection. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 395-405).

[9] Parikh, S. B., & Atrey, P. K. (2018, April). Media-rich fake news detection: A survey. In 2018 IEEE conference on multimedia information processing and retrieval (MIPR) (pp. 436-441). IEEE.

[10] Karimi, H., & Tang, J. (2019). Learning hierarchical discourse-level structure for fake news detection. arXiv preprint arXiv:1903.07389.

[11] Nguyen, V. H., Sugiyama, K., Nakov, P., & Kan, M. Y. (2020, October). Fang: Leveraging social context for fake news detection using graph representation. In Proceedings of the 29th ACM international conference on information & knowledge management (pp. 1165-1174).