DOI: 10.54254/2755-2721/87/20241573

Detection of network false information based on artificial intelligence models

Haoxi Mao

Computer Science, Shanxi Agricultural University, Jinzhong, Shanxi, 030801, China

quxin@ldy.edu.rs

Abstract. There is often some false information in social platforms to mislead public opinion. Due to the rapid development of the Internet, the spread of false information on the Internet has become easier, which has brought many losses to people's economy and life. In this paper, the relevant research on false information based on bidirectional convolutional networks is analyzed, and the method is divided into four stages: data preprocessing, model architecture, training process and prediction process. Then, the relevant research on rumor detection by propagation tree kernel model are analyzed. Finally, this paper delineates a comprehensive framework that amalgamates enhanced Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and a hybridized Black Widow Optimization (BWO) with Moth Optimization Algorithm (MOA) (referred to as HM-BWO) for the accurate detection of network-based false information. This paper analyzes and discusses the challenges of poor universality and insufficient detection speed encountered by the current rumor detection research and puts forward the idea of introducing migration model to solve the problem of poor universality and Spark to solve the problem of insufficient detection speed. This article provides a good overview of the field of online disinformation.

Keywords: machine learning, natural language process, rumor detection.

1. Introduction

Network false information refers to information presented in the form of words, images, audio and video or symbols in the carrier of modern information networks, which is inconsistent with the facts or fabricated, disrupts the political, social, economic and other order, or infringes on the legitimate rights and interests of others. In social media, there is usually some false information to mislead the public opinion, so as to achieve the purpose of obtaining economic benefits or achieving certain political goals. With the progression of the internet era, the substantial rise in the quantity of internet users, and the dissemination of online social networking practices, the propagation of misinformation on the web is more facile than that via conventional news mediums such as newspapers and radio. The nominal expense associated with sustaining social media platforms, coupled with their user-friendly interfaces, exacerbates this developing trend. By comparing the depth, size, maximum breadth and structural virality of the cascade of false information and true information forwarding, it can be observed that false information disseminates significantly more extensively, rapidly, deeply, broadly, and structurally perniciously than true information [1]. The losses caused by false information are huge. For example, BBC News reported on August 12, 2020, that a new study showed that in the first three months of this year alone, false information about the novel coronavirus on social media had led to at least 800 deaths and about 5,800 hospitalizations worldwide [2]. Therefore, research on the detection of false information on the Internet is very necessary. In the face of the massive network information generated every day, traditional manual detection methods cannot accurately distinguish false information, while artificial intelligence has a strong ability to extract text image information feature values and prediction, and can quickly identify false information on the network and reduce the harm brought by false information on the network.

Artificial intelligence has made great progress in recent years. It has a variety of algorithms, such as logistic regression model, decision tree, naive Bayes function, convolutional neural network, etc., which have been applied in finance, medical treatment, games, risk assessment, data analysis and other fields. In journalism, there have been many people using artificial intelligence technology to do related research, and the detection of false information on the Internet is an important direction of journalism. At present, there are many researchers using the relevant model of artificial intelligence to study the detection of false information on the Internet. For instance, the Binary Graph Convolutional Network (Bi-GCN) model delves into these two aspects by employing a concurrent top-down and bottom-up approach to rumors' dissemination. The system employs the Graph Convolutional Network (GCN) in conjunction with a top-down directed graph to elucidate the patterns of rumor propagation, and conversely, integrates the GCN with the inverse graph of rumor propagation to encapsulate the architecture of its dissemination [3]. Furthermore, certain studies have embraced the non-linear structural characteristics of the propagation tree, integrating them with linear features for the purpose of rumor classification [4]. In addition to applications in machine learning, deep learning models are also applicable in pertinent research concerning the detection of false information on the internet. For instance, graph neural networks are employed to acquire the representation of user relevance from the binary graph encapsulating the correlation between users and source tweets [5], and the representation of information propagation using tree structure. This paper then combines the representations learned from these two modules to classify the rumors.

The present study endeavors to furnish a thorough summary of the artificial intelligence mechanisms employed in the detection of misinformation across the Internet. It is mainly composed of four parts and the rest is organized as follows. First, in the second part, this review will elaborate on the methods used to detect false information on the Internet in detail. In the third part, the current status and development of the current network false information detection and the challenges it faces. In the last part, I will make a summary of the whole article.

2. Method

2.1. Introduction of machine learning

Network false information detection based on machine learning mainly includes data collection, data preprocessing, feature extraction, selection of learning model, training model, model evaluation, model optimization and practical application shown in Figure 1.

Data collection: Collect large amounts of textual data containing true information and rumors.

Data preprocessing: the text is cleaned, the word is divided, the word is stopped and so on, and the text is transformed into a form suitable for model processing.

Feature extraction: Extract meaningful features from preprocessed text, such as word frequency, part of speech, semantic features, etc.

Select learning model: It is imperative to choose an appropriate machine learning model that aligns with the distinct attributes of the problem at hand.

Training model: Utilize the annotated training data to refine the model.

Model evaluation:Utilize an exclusive test dataset to assess the efficacy of the model.

Model optimization: Adjust and optimize the model based on the evaluation results, such as adjusting parameters, trying different models, or combining multiple models.

Practical application: The optimized model is applied to new text data for rumor detection.

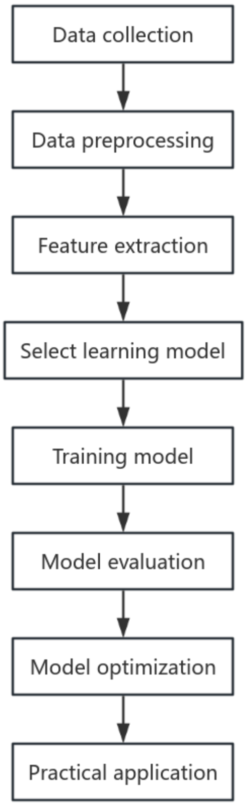

Figure 1. Flow chart of network false information detection based on machine learning (Photo/Picture credit: Original).

2.2. Machine learning models

2.2.1. Bi - GCN

The authors introduce a novel rumor detection approach utilizing a Bidirectional Graph Convolutional Network (GCN), designed to identify rumors on social media platforms. This method integrates the attributes of both rumor propagation and diffusion, conducting an analytical investigation of rumor characteristics through traversal in both upward and downward propagation directions. The methodology is primarily composed of four sequential phases: data preprocessing, architectural design of the model, the training regimen, and the subsequent prediction phase. During the data preprocessing stage, the researcher employed a trio of datasets: Weibo, Twitter15, and Twitter16. These data sets contain information such as users, posts, retweets and reply relationships. The authors extracted the TF-IDF values of the posts as features, and built a propagation structure based on the forward and reply relationships. The authors then divided the data into training sets, verification sets and test sets based on the rumor tags. Using DropEdge during the training phase to avoid overfitting problems. The researchers employed a bidirectional Bi-GCN architecture for the identification of rumors within social media platforms.This architectural design integrates a Top-Down graph Convolutional network (TD-GCN) with a Bottom-Up graph Convolutional network (BU-GCN). The TD-GCN facilitates the dissemination of rumors by enhancing the transmission of information from parent nodes to their descendant nodes. Conversely, the BU-GCN encapsulates rumor propagation by aggregating information from descendant nodes to their parent node. The model splices the output of TD-GCN and BU-GCN to obtain the final rumor detection result [3].

2.2.2. Propagation tree kernel model

The author uses the Propagation Tree Kernel (PTK) and the Context-Sensitive Propagation Tree Kernel (cPTK) to detect rumors.

Two datasets, Twitter15 and Twitter16, were used in the experiment. These two datasets contain a large number of source tweets and their spread structure, and have been labeled as either rumors or non-rumors.The efficacy and preeminence of the proposed propagation tree kernel model, as well as the context-sensitive propagation tree kernel model, have been empirically substantiated through experimental validation on the respective datasets.

In the course of the experiment, the author delineates the dissemination of every individual tweet as a hierarchical tree framework. Herein, the root node corresponds to the originating tweet, while the leaf nodes signify the audience's responses to the post. Furthermore, the directed edges encapsulate the inter-node responsive relationships.Subsequently, the higher-order structure of diverse rumor types is delineated through the computation of the likeness between their propagation trees.Certainly, the Pattern Tree Kernel (PTK) or its compressed variant (cPTK) is utilized to measure the similarity within these propagation trees. Subsequently, these trees are incorporated as features into a kernel-based Support Vector Machine (SVM) classifier for the intent of classification. Within the confines of a multi-classification endeavor, the one-to-many classification strategy is adopted, wherein the category garnering the highest likelihood is nominated as the predictive outcome [4].

2.2.3. ICNN

In the current research, they present a novel hybrid deep learning architecture for the detection of fake news. This framework integrates Enhanced Convolutional Neural Networks (ECNN), Long Short-term Memory Networks (LSTM), and a hybridized approach incorporating the Black Widow Optimization (BWO) and Moth Optimization (MO) algorithms to facilitate the automated detection and categorization of fraudulent news content prevalent on social media platforms. The data utilized in the analysis is sourced from the Fake News Challenges (FNC) repository, as well as the KDnuggets and ISOT datasets.

In this experiment, the authors used an improved convolutional neural network structure called ICNN. The Integrated Convolutional and Recurrent Neural Network (ICNN) harnesses the merits of both convolutional and recurrent neural networks, facilitating the concurrent processing of textual spatial and temporal information. The architecture of the Integrated Convolutional Neural Network (ICNN) comprises an input layer, a convolutional layer, a pooling layer, a recurrent layer, a fully connected layer, and an output layer. The input layer is tasked with processing textual data, whereas the convolutional and pooling layers are responsible for the extraction of textual features. The recurrent layer applies recurrent processing on the extracted features, thereby enhancing the network's ability to model temporal dependencies [5].

3. Discussion

3.1. Limitations and challenges

3.1.1. Generality

Generally, the generality of the model is strongly related to the training set used in the training model. For example, the data set of a social platform is used to train the false information detection model, and the trained model is used to detect false information on another platform, which will fail to accurately detect false information. Because their data distribution is different, such as video data, text data, user comment data etc, there are great differences, resulting in poor universality of detection models. However, due to the poor universality of the detection model, it is required to train the model separately for different models and different language countries, resulting in a great increase in cost. Therefore, how to train the detection model with certain universality or easy migration through the field of easy data collection for cross-domain, cross-platform and cross-source information detection is a huge challenge that cannot be avoided in the application of false detection technology.

3.1.2. Speed

In the real-time application environment, information detection is faced with massive data flow, and false information spreads viral far faster than true information. For the release of some information, time sensitivity is very important. The core of human curiosity and the pursuit of novelty constitutes the foundational factor that ascertains the timeliness value of news content. Daily, a vast quantity of information traverses social media platforms, and there is a prevalent preference among internet users for the most recent updates, which in turn intensifies the necessity for the immediate identification and verification of misinformation. At the same time, there are great requirements for the speed of detecting false information. Training models to detect faster is a big challenge.

3.2. Future prospects

3.2.1. Transfer learning

Transfer learning is a kind of learning method in machine learning. Transfer learning facilitates the application of established learning models within diverse, yet interconnected, environments. In conventional machine learning approaches [6-8], the models lack sufficient flexibility, resulting in suboptimal performance when addressing variations in data distribution, dimensionality, and model output alterations. Transfer learning mitigates these constraints by incorporating knowledge from the source domain, thereby enhancing the modeling process under varying conditions of data distribution, feature dimensions, and model output dynamics [6]. The combination of transfer learning and false information detection model can effectively solve the problem of poor generalization of the model.

3.2.2. Spark

Apache Spark is an expedient big data processing engine that excels in distributed computing environments, enabling the efficient manipulation of massive datasets across clusters. This technology has garnered significant popularity in recent times. This framework is poised to supersede Hadoop, as depicted in Figure 2. Its primary benefits include rapid processing, user-friendliness, and exceptional adaptability. Its paramount feature is its expediency, boasting speeds that are 100 times faster in memory and 10 times faster on disk when compared to HadoopMapReduce [9, 10].

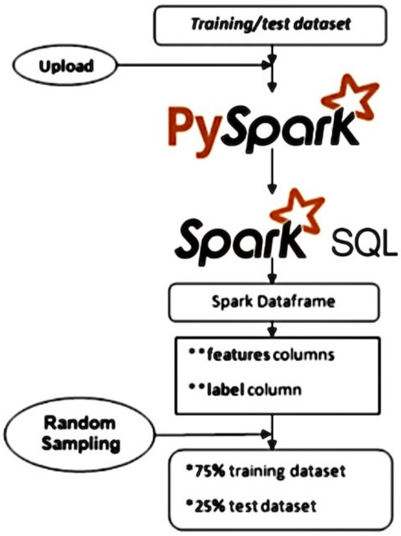

Figure 2. Structure based on spark SQL training and testing data sets [7].

In recent years, many studies have used Apache Spark to conduct data analysis and processing in the data preprocessing stage of network false information detection, but few have combined it with model training. In the future, it is very promising to combine the model training of rumor Apache Spark and network false information detection to improve the speed of training model and model optimization.

4. Conclusion

This manuscript has provided an exhaustive summary of the methodologies employed in the detection of misinformation on the internet. In this paper, researches on the detection of network false information by Bi-GCN, Propagation tree kernel model and ICNN are investigated respectively, and their specific processes and algorithms are explained in detail. After discussion and analysis, it could be found that in the field of network false information detection, there are mainly two problems: poor universality of detection model and insufficient detection speed. To solve these two problems, this paper proposed a solution that uses Spark platform to solve the insufficient speed of model detection and uses transfer learning method to solve the poor universality of detection model. In the future, the further study plans to apply the proposed hypothesis to serve the research of network false information detection.

References

[1] Vosoughi S, Roy D & Aral S 2018 The spread of true and false news online Science vol 359 (6380) pp 1146-1151

[2] Sina Science and Technology 2020 Eating garlic to protect against COVID-19? Study: Nearly 5,800 people were hospitalized worldwide because of misinformation related to COVID-19, at least 800 deaths [EB/OL]

[3] Bian T, Xiao X, Xu T, Zhao P, Huang W, Rong Y & Huang J 2020 Rumor Detection on Social Media with Bi-Directional Graph Convolutional Networks The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)

[4] Hamidian S, Diab MT 2019 Rumor detection and classification for twitter data. arXiv preprint arXiv:1912.08926

[5] Narang P, Singh AV & Monga H 2022 Hybrid Metaheuristic Approach for Detection of Fake News on Social Media International Journal of Performability Engineering vol 18 no 6 June pp 434-443

[6] Pan SJ & Yang Q 2010 A Survey on Transfer Learning IEEE Transactions on Knowledge and Data Engineering vol 22 no 10 pp 1345-1359

[7] Qiu Y, Hui Y, Zhao P, Wang M, Guo S, Dai B, Dou J, Bhattacharya S & Yu J 2024 The employment of domain adaptation strategy for improving the applicability of neural network-based coke quality prediction for smart cokemaking process Fuel Sep 15 vol 372 p 132162

[8] Ma Y, Chen S, Ermon S & Lobell DB 2024 Transfer learning in environmental remote sensing Remote Sensing of Environment Feb 1 vol 301 p 113924

[9] Madani Y, Erritali M & Bouikhalene B 2021 Fake News Detection Approach Using Parallel Predictive Models and Spark to Avoid Misinformation Related to Covid-19 Epidemic In: Gherabi N & Kacprzyk J (eds) Intelligent Systems in Big Data, Semantic Web and Machine Learning Advances in Intelligent Systems and Computing vol 1344 Springer, Cham

[10] Öztürk MM 2024 Tuning parameters of Apache Spark with Gauss–Pareto-based multi-objective optimization Knowledge and Information Systems Feb vol 66 (2) pp 1065-1090