2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP) | 978-1-6654-0505-8/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICCWAMTIP51612.2020.9317325

FAKE NEWS DETECTION USING DEEP RECURRENT NEURAL NETWORKS

TAO JIANG[1], JIAN PING LI[1], AMIN UL HAQ[1], ABDUS SABOOR[1]

[1]School of Computer Science and Engineering University of Electronic Science and Technology of China, Chengdu
E-MAIL: 3057861564@qq.com, jpli2222@uestc.edu.cn, khan.amin50@yahoo.com, saboor_khan97@yahoo.com

Abstract:

The widely spread of fake news has significantly impacted our life in politics and economics. To solve this problem, different researchers have proposed various machine learning and deep learning models. However, most of them detect fake news without desired accuracy. Therefore, we proposed a deep learning framework that classifies fake news from real ones with 99.82% accuracy. This BiLSTM model was trained and tested on a fact-checking dataset. Furthermore, we used different model evaluation metrics like precision, recall, F1-meassure, execution time to prove the efficiency of our model.

Keywords:

BiLSTM; Binary classification; Fake news detection; Deep learning

1. Introduction

Fake news has caused too many debates since the 2016 U.S. presidential election campaign. However, the human ability to detect news deception without expert assistance is quite challenging. Even fact-checking websites like Snopes.com or PoliticFact.com can only detect fake news on a small scale because the manual verification is laborious and time-consuming. We should try to solve this problem by using automatic tools to classify this news.

The task of fake news detection is usually confusing with other deception detection tasks like spam detection [1], rumor detection [2] and satire detection [3]. In this paper, we defined fake news detection as a prediction of specific news article that are fabricated intentionally to mislead people.

In some papers, researchers not only use news body to detect fake news but also combine image data, news source and comments to capture the features of fake news. In our paper, we only rely on the article text to capture the linguistic features that represents the fake news writing styles.

Khattar, et al. [4] proposed Multimodal Variational Autoencoder (MVAE) model, which used a bimodal variational autoencoder coupled with a binary classifier for the task of fake news detection. They used multimodal (textual + visual) information from popular microblogging websites: Weibo and Twitter dataset. Eventually they got 0.745 accuracy on twitter dataset and 0.824 accuracy on Weibo dataset. Their model contained three components: Encoder, Decoder and Fake News Detector which use the learned shared representation to predict. They used RNNS with LSTM cells to extract features from textual content and CNNs to train image descriptors. Shu, et al [6] developed a sentence-comment co-attention sub-network to train both news contents and user comments to jointly capture explainable top-k check-worthy sentences and user comments. Their model outperforms seven state-of-the-art fake news detection methods by at least 5.33% in F1-score, better than baselines by 30.7% in Precision. Popat, et al [7] presented DeClarE model, which aggregates signals from external evidence articles. It considered the context of the claim via word embeddings and the web articles captured via a bidirectional LSTM, while using an attention mechanism to focus on parts of the articles according to their relevance to the claim to increase interpretability.

2. Materials and methods

2.1. Data set

In [8], Ozbay et al performed more than twenty machine learning models on ISOT dataset. In this dataset, 44898 article news are contained in total, 21417 of them are real (label class is 1) and 23481 are fake (label class is 0). The dataset features include title, news body, subject, date and label. There are eight different news topics in column 'subject'. To get enough information, we combine the news body and news title to train and test our models. All text data are cleaned and tokenized before feeding them into our BiLSTM models. A little section of ISOT dataset before cleaned and after cleaned can be seen in figure1 and figure 2.

	title	text	subject	date	category
500	Senate panel votes to advance tax bill	WASHINGTON (Reuters) - The U.S. Senate Budget ...	politicsNews	November 28, 2017	1
501	Tillerson 'offended' by claims of State Depart...	WASHINGTON (Reuters) - U.S. Secretary of State...	politicsNews	November 28, 2017	1
502	Trump to make remarks at White House at 3 p.m....	WASHINGTON (Reuters) - U.S. President Donald T...	politicsNews	November 28, 2017	1
503	U.S. budget chief Mulvaney says CFPB staff sho...	WASHINGTON (Reuters) - U.S. budget chief Mick ...	politicsNews	November 28, 2017	1
504	Russian envoy to U.S. to inspect San Francisco...	MOSCOW (Reuters) - Moscow's ambassador to the ...	politicsNews	November 28, 2017	1

Fig.1 A section of uncleaned ISOT dataset

	text	category
500	WASHINGTON (Reuters) U.S. Senate Budget Commit...	1
501	WASHINGTON (Reuters) U.S. Secretary State Rex ...	1
502	WASHINGTON (Reuters) U.S. President Donald Tru...	1
503	WASHINGTON (Reuters) U.S. budget chief Mick Mu...	1
504	MOSCOW (Reuters) Moscow's ambassador United St...	1

Fig.2 A section of cleaned ISOT dataset

2.2. Bidirectional Recurrent neural networks

Deep leaning neural neworks are used in many applications [9-11]. However, bidirectional recurrent is a common variant of recurrent neural networks. It performs better on some tasks than normal RNNs and usually is used in natural language processing. RNN relies on sequence or time series to capture pattern. Therefore, when people disrupt or reverse time steps, the representations from the data can be significantly different. Based on this characteristic of RNN, bidirectional RNN like BiLSTM, are proposed. BiRNNs contain two normal RNNs, each of them processes the data in one direction and combine their representations at last. Generally speaking, bidirectional RNNs can capture some patterns that are ignored by normal RNNs.

2.3. Text tokenization

Preprocessing usually means tokenization, and weighting words in the document. To transform tokenized text into vectors, word embedding is often used. For word sequences, pretrained embedding neural networks like word2vec and glove are commonly used. In our work, we used glove to get the pretrained weighted word vectors. Considering the computation time, we limit the vocabulary to only consider the top 10000 terms across the entire corpus. We also padded every instance into 300 words long by using keras.

2.4. proposed bidirectional recurrent neural network

In the embedding layer, input_dim is the length of selected vocabulary, namely 10000. The input length is 300 just like the length of our padded instances. Output dimensionality is 100. To avoid overfitting, we also used dropout to reduce the parameters. Other detailed parameters can be seen in table 1.

Table 1 BiLSTM model structure

layer type	output shape	parameters
embedding_1	(None,300,100)	1000000
bidirectional_1	(None,300,64)	34048
bidirectional_2	(None,128)	66048
dense_1	(None,32)	4128
dense_2	(None,1)	33

2.5. Validation method

To select the best models and parameters, we used hold-out cross validation [18-19] methods on our dataset. The datasets are splited into two parts: training dataset and test dataset. Training data accounts for 80% of original dataset and the remaining data are used for testing.

2.6. Evaluation metrics

We employed different evaluation metrics to evaluate our models, such as precision, recall, accuracy and F1-meassure [12-15]. These metrics can be computed by using equations (1), (2), (3), (4).

$$\text{Accuracy}= \frac{TP+TN}{TP+TN+FP+FN} \text{x}100\% \tag{1}$$

$$\text{Recall} = \frac{TP}{TP+FN} \text{ x}100\% \tag{2}$$

$$\text{F1-score} = 2\text{x}\frac{(precision)(recall)}{precision+recall}\text{x}100\% \quad (3)$$

$$\text{Precision} = \frac{Tp}{Tp+FP}\ \text{x}100\% \quad (4)$$

2.7. Proposed model procedures

The whole detection procedures are described in table 2.

Table 2 Pseudo-code of proposed fake news detection model

Step1: Begin
Step2: Pre-processing of fake news data set;
Step3: Train BiLSTM model on training data set with different parameters;
Step4: Validated BiLSTM model on testing data set;
Step5: Computes different model performance criterion;
Step6: End

3. Experimental results analysis

Our model has been trained with 80% data. The remaining essential parameters about our model have been shown in table 2. The model performance on the fake news dataset in accuracy, precision, recall, F1-meassure has been computed in table 3. The whole experiment is carried out by using python configuration of Intel Core TM i5-2410M, 4GB random access memory with 640BG hard drive, window 10. According to the results in table 3, we can see that the model performance is extraordinary. The model obtained 99.82 % accuracy, 100% recall, 100% precision, 100% F1-score. The training loss and validation loss and training accuracy and validation accuracy of our model have been shown in figure 3. The results proved that our model is of great quality and can be used to detect fake news.

Table 3 Model essential parameters

epochs	callbacks	Optimizer	loss	Output activation function
15	ReduceLROnPlateau ((monitor='val_accuracy', patience=2,verbose=1, factor=0.5,min_lr=0.001))	Adam(lr=0.01)	binary_crossentropy	sigmoid

Table 4 Model performance with multiple metrics

Accuracy(%)	precision(%)	recall(%)	F1-score(%)	Time(m)
99.82	100	100	100	50

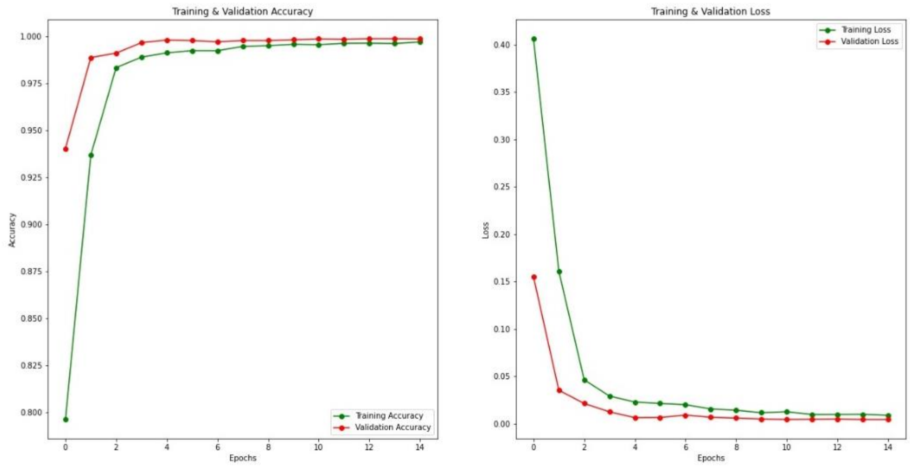

Fig 3. Model loss and accuracy curve

4. Conclusion

Our bidirectional recurrent neural network model can detect fake news with extremely high accuracy. In reality, we can use our model to automate the news validation process and help us save a lot time and free individuals from laborious fact-checking work. Our model only has two BiLSTM layers and two dense layers but achieved great performance with 99.82% in accuracy, 100% in precision, recall and f1-score.

In future work, we may try to combine other models like CNN, GRU, LSTM with our models to see if integrated layers can improve the results. And normally speaking, data size is always the bottleneck of machine learning and deep learning tasks, so we will try to collect more data. All news in our dataset are written in English, so we also will try to apply our model in non-English dataset.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 61370073), the National High Technology Research and Development Program of China (Grant No. 2007AA01Z423), the project of Science and Technology Department of Sichuan Province, the project of Science and Technology Department of Chongqing Municipality, Chengdu Chengdian Network Technology Co., Ltd., Chengdu Civil-military Integration Project Management Co., Ltd., and Sichuan Yin Ten Gu Technology Co., Ltd.

References

[1] Asghar, Muhammad Zubair, et al. "Opinion spam detection framework using hybrid classification scheme." Soft computing 24.5 (2020): 3475-3498.
[2] Li, Quanzhi, Qiong Zhang, and Luo Si. "Rumor detection by exploiting user credibility information, attention and multi-task learning." Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. 2019.
[3] Toçoğlu, Mansur Alp, and Aytuğ Onan. "Satire detection in Turkish news articles: a machine learning approach." International Conference on Big Data Innovations and Applications. Springer, Cham, 2019.
[4] Khattar, Dhruv, et al. "Mvae: Multimodal variational autoencoder for fake news detection." The World Wide Web Conference. 2019.
[5] Shu, Kai , et al. "dEFEND: Explainable Fake News Detection." KDD 2019.
[6] Popat, Kashyap , et al. "DeClarE: Debunking Fake News and False Claims using Evidence-Aware Deep Learning." (2018).
[7] Ozbay, Feyza Altunbey , and B. Alatas . "Fake news detection within online social media using supervised artificial intelligence algorithms." Physica A: Statal Mechanics and its Applications 540(2020).
[8] Zhang, Malu, Hong Qu, Xiurui Xie, and Jürgen Kurths. "Supervised learning in spiking neural networks with noise-threshold." Neurocomputing 219 (2017): 333-349.
[9] Zhang, Malu, Hong Qu, Ammar Belatreche, Yi Chen, and Zhang Yi. "A highly effective and robust membrane potential-driven supervised learning method for spiking neurons." IEEE transactions on neural networks and learning systems 30, no. 1 (2018): 123-137.
[10] Zhang, Malu, Hong Qu, Ammar Belatreche, and Xiurui Xie. "EMPD: An efficient membrane potential driven supervised learning algorithm for spiking neurons." IEEE Transactions on Cognitive and Developmental Systems 10, no. 2 (2017): 151-162.
[11] Ul Haq, Amin, et al. "A novel integrated diagnosis method for breast cancer detection." Journal of Intelligent & Fuzzy Systems Preprint (2020): 1-16.
[12] Haq, Amin Ul, et al. "Feature selection based on L1-norm support vector machine and effective recognition system for Parkinson's disease using voice recordings." IEEE access 7 (2019): 37718-37734.
[13] Haq, Amin Ul, et al. "Comparative analysis of the classification performance of machine learning classifiers and deep neural network classifier for prediction of Parkinson disease." 2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP). IEEE, 2018.
[14] HAQ, AMIN UL, et al. "Identifying the Predictive Capability of Machine Learning Classifiers for Designing Heart Disease Detection System." 2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing. IEEE, 2019.
[15] HAQ, AMIN UL, et al. "Identifying the Predictive Capability of Machine Learning Classifiers for Designing Heart Disease Detection System." 2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing. IEEE, 2019.