Fake News Detection Based on a Bi-directional LSTM with CNN

Yu Ji(✉)

University of California, Santa Cruz, Santa Cruz, CA 95060, USA
yji34@ucsc.edu

Abstract. The misleading information brought by fake news has troubled our society for a long time. Recently, the increasing spreading rate of fake news has more severe consequences than ever in the past. Many types of neural networks have been applied to solve fake news detection and other natural language problems during these years. Nevertheless, due to the limitation of each structure, a hybrid neural network would often achieve a preferable accuracy. In this paper, a novel deep neural network is proposed for the fake news detection problem based on Convolutional Neural Network and Bi-directional Long Short Term Memory network. Sequential information will be captured by using Bi-LSTM and hidden features will be captured at a detailed level using CNN. The model will be tested on large-scale datasets, which demonstrated better performance than conventional neural networks.

Keywords: Machine learning · Deep learning · Convolution network · Bi-LSTM · Natural language processing

1 Introduction

The rapid growth of social media platforms indeed brings information accessibility to their users. Unfortunately, fake news has also been wildly spreading via traditional print and online social media, which obtain economic or political benefits while manipulates its readers. Therefore, fake news detection becomes a practical natural language processing (NLP) problem to prevent people from misleading by false information.

Although detecting fake news is a long-discovered problem in human life, people could only define the trueness of a piece of news with their mind most time during the past centuries. However, the development of machine learning and deep learning algorithm has brought a new perspective into this field. In the past few years, there have been various experiments and approaches in NLP problems based on machine learning and deep learning algorithms. Naïve Bayes was chosen to perform the classification as one of the most traditional machine learning methods. Nonetheless, this experiment has been done by Granik in 2017 has a great start, it still leaves some space for improvement. [1] Then, other machine learning models are also introduced by Ahman. Etc, such as linear SVM, logistic regression, multiplayer perceptron, K nearest neighbors, and random forests. [2] These models have been tested with multiple datasets. Even though the logistic regression model outperformance other methods, all

W. Cao et al. (Eds.): CONF-CDS 2021, CCIS 1513, pp. 36–44, 2021.
https://doi.org/10.1007/978-981-16-8885-0_4

of them achieve less accuracy than ensemble techniques, deep learning methods. According to Yang. Etc., TI-CNN model, which extracts explicit and latent features, can identify fake news effectively. [3] However, their model relies on the combined information from both the text and image. There also have been researches on formats of the dataset. Dong, etc. introduces his CNN-based model to solve fake news detection problem seeing as semi-supervised learning. [4] By only providing part of the label, the model spends a large amount of time on calculation. Aldwairi and Alwahedi approach fake news detection in social media networks with various groups of datasets which are distinguished by numbers and other marks. [5] On the other hand, LSTM and other RNN models are tested by Shen and Zhang on the sentence classification task. [6] Later on, many hybrid models are introduced by others in the NLP area. Okoro, etc. achieved great accuracy by combining the human literacy news detection tool and the machine linguistic and network-based approaches. [7] Zhang, etc. have developed their deep diffusive neural network to solve the same problem. [8] In addition, a combination of CNN and RNN approach is proposed by Zhang, etc. called RCNN. [9] Meanwhile, Bahad, etc. also have built their model using Bi-directional LSTM network. [10] Based on that, a more advanced hybrid network is generated from the combination of a CNN with an attention mechanism and a Bi-LSTM network [11].

In this paper, a hybrid neural network is applied to the fake news detection problem. The part of Bi-LSTM will firstly uncover the sequential information in both forward and backward order. The part of CNN will be used to reveal the hidden connections at a detailed level. This proposed hybrid neural network model will solve the fake news detection problem in a more effective way.

The rest of this paper is structured as follows: Sect. 2 is the summary of related works in the field. Section 3 demonstrates the neural network models that have been developed and used for comparison. Section 4 concludes our work.

2 Method

2.1 An Overview of the Proposed BiLSTM-CNN Algorithm

The proposed approach consists of two main parts, which are text preprocessing and deep neural network. The input datasets are two sets of classified news sets, either true or fake. Removing stopwords and punctuations of the dataset is the first step of the preprocessed pipeline. Then embedding dictionary is set up using Word2Vec algorithm. Next is the neural network, shown in Fig. 1. The generated matrix representation of sentences produced by the embedding layer will be sent into a hyper neural network structure. Sequential information will be captured by the first level of structure, Bi-LSTM. And the next level neural network CNN will extract the hidden connection of the words.

2.2 Text Preprocessing

Preprocessing the unstructured data, such as text, is one of the cores of natural language processing. The machine learning models often require the matrix form of the input from both extraction and manipulation of the texts. Specifically, removing punctuation, lowercase conversion, and word embedding are applied in the experiment. In addition, removing the Stopwords is also used to perform data cleaning. Stopwords, such as "the, and, who," often appears in English sentences but with limited information knowledge. The information density will remain the same without them. In this experiment with a Python3 environment, the English Stopwords sublibrary in the Natural Language Toolkit (NLTK) is applied.

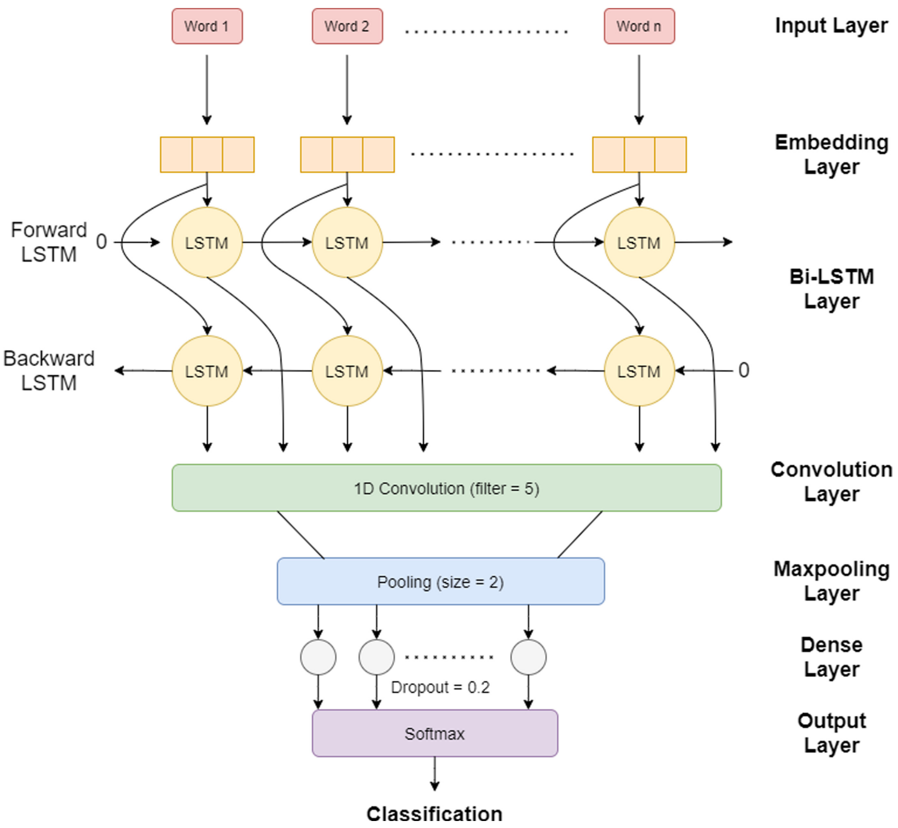

Fig. 1. This is an architecture diagram of the proposed model. The input information will stream through the Bi-LSTM part, which consists of the embedding layer and Bi-LSTM layer. Then the information will be used to adjust the connection in the CNN structure. After all, the last layer will use a SoftMax function to output classification.

2.3 Sequential Information Captured by Bi-LSTM

The Bi-LSTM model used contains a word embedding layer and a two-part LSTM structure. The embedding layer will convert each word to a vector representation according to a pre-learned algorithm. The similar meaning of the words will result in a closer distance in the vector space. Specifically, the continuous skip-gram model (Fig. 2) is chosen from Word2Vec algorithms is used in this experiment. As a reversed structure of the CBOW algorithm, skip-gram is wildly used to solve unsupervised learning problems by finding the most related words for a given word. Under this structure, input is designed to be the target word and output is designed to be the context word.

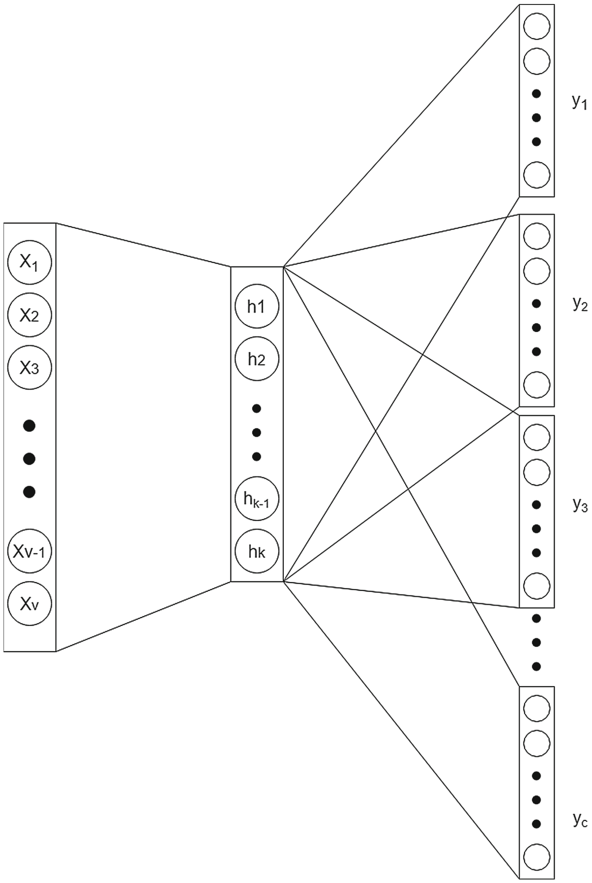

Fig. 2. This is a structural representation of the Skip-Gram Model. The x_i is the input given. The middle layer is the only one hidden layer. There is no activation function side of the hidden layer. The output layer on the right side computes the output y_i from the dot product of the hidden layer h_i and weight matrix.

In the experiment, each piece of news contains a sequence of words $w_1, w_2, \ldots, w_n$. Based on that, the average log probability will be used to predict context words given the target word within a limited training context with the size c.

$$\frac{1}{I}\sum\nolimits_{i=1}^{I}\sum\nolimits_{-c\leq j\leq c, j\neq 0} log\, P(W_{i+j}|W_j) \tag{1}$$

The other part of the model is the Bi-LSTM layer. This layer consists of two LSTM layers, from both forwards and backwards. LSTM can handle long-term dependency because both cell states and hidden states are included in an LSTM structure, where a vanilla RNN only contains a hidden state. The other reason for choosing the LSTM unit instead of RNN is to prevent the gradient vanishing situation. In addition to a single layer, Bi-LSTM is used to represent both sequential information in forward and backward, which are connected to the same output layer.

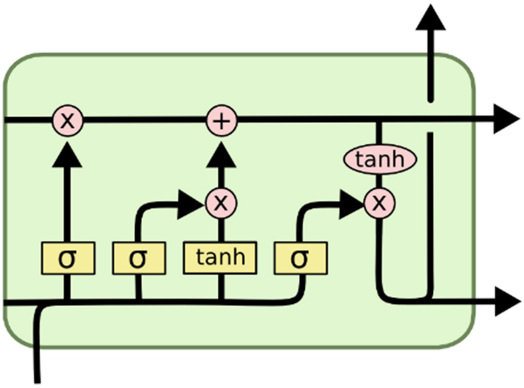

Fig. 3. A Complete LSTM unit [12]

As shown in Fig. 3, the current LSTM unit can be calculated as follows at a time t:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{2}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$C_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{4}$$

$$C_t = f_t * C_{t-1} + i_t * C_t \tag{5}$$

$$o_t = \sigma(W_0 \cdot [h_{t-1}, x_t] + b_0) \tag{6}$$

$$h_t = o_t * tanh(C_t) \tag{7}$$

In those equations, σ represents the sigmoid activation function. Weights are updated according to the flowchart from the bottom left to the top right. X_t is the input vector at the time t; W_i, W_f, W_C, W_o represent the weight associated with X_t in different gates; h_{t-1}, h_t represent the hidden state; b_f, b_C, b_o represent the bias offset of each gate. And * is the elementwise multiplication.

The first step, f_t, is used to forget the information sent from the last round. Then, i_t and C_t are combined to determine the information stored in the current cell state. This is also as known as updating the cell state. After C_t, the tanh layer, is used to update the current one, it also has to update the old cell state C_{t-1} to C_t by multiplying f_t and adding the new candidate values. Later, a sigmoid layer and a tanh layer will decide the final output of a LSTM cell. Eventually, a complete Bi-LSTM structure will have its result by calculating the elementwise sum of h_i forward and h_i backward.

2.4 Convolutional Neural Network

After the weights are passed from the Bi-LSTM layers, a traditional convolutional neural network (CNN) is utilized to finish the true or fake news categorization. Each convolutional layer will decrease the window size to generate a new feature representation of the information. And a max-pooling layer will concatenate and capture the hidden representation. A combination of both layer structures is applied to make the final prediction with a SoftMax activation function (Fig. 4).

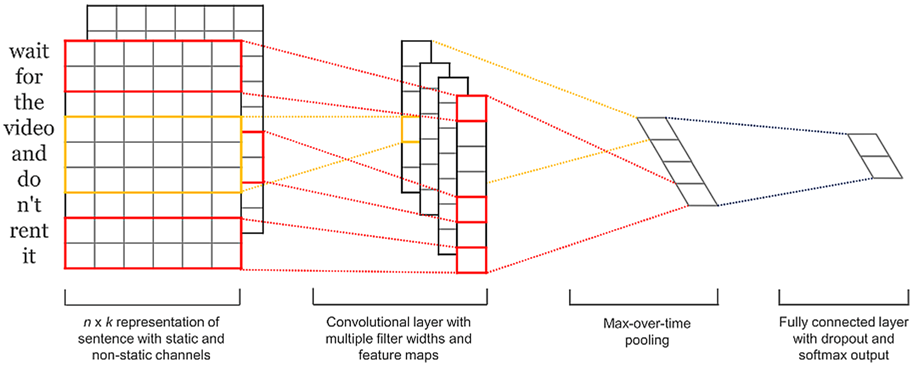

Fig. 4. A Complete CNN structure [13]

3 Experiment and Result

All codes are written in Python 3.7 using TensorFlow 1.14. All experiments have been performed on a Core processor Intel CPU i7-7800X 3.5 GHz with 32 GB RAM and a 1080Ti graphic card. The dataset is obtained from Fake and real new Dataset at Kaggle.

Dataset: The dataset is from the Kaggle website. [14] As one of the most popular datasets used in the NLP research, this dataset contains the fake and real news from during the American presidential election period in 2016. There are two pre-labeled files, true and fake accordingly, containing more than 20000 pieces of news each. Title, Text, Subject, and Date, a total of four attributes are introduced in each news. The content under the Text attribute is used as the input of the experiment. Also, 0 is given to the FAKE set, and 1 is given to the TRUE set as the label to distinguish two sets.

The experiment has 100 as the batch size, 10 as the epochs, and 0.2 as the validation split rate. The performance accuracy of the models is evaluated according to the prediction accuracy, whether the news is correctly classified as TRUE or FAKE (Figs. 5 and 6) (Table 1).

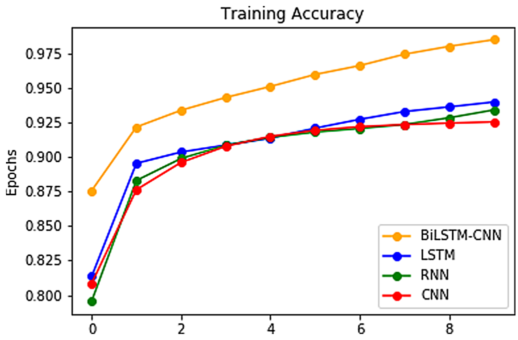

Fig. 5. Training accuracy among all four tested models

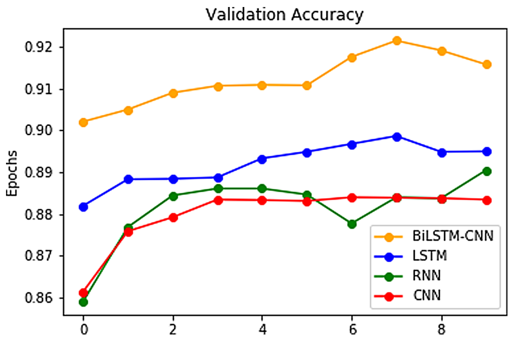

Fig. 6. Validation accuracy among all four tested models

Table 1. Evaluation of the proposed model and some common neural networks (Proposed Model is in bold)

Model	Accuracy	Precision	Recall	F1 Score
CNN	0.883	0.893	0.891	0.892
Vanilla RNN	0.890	0.899	0.903	0.901
LSTM	0.895	0.909	0.908	0.908
BiLSTM-CNN	**0.916**	**0.918**	**0.914**	**0.916**

Four models are tested in the experiment: CNN, RNN, LSTM, BiLSTM-CNN (proposed model). Overall, the new BiLSTM-CNN model solves this and returns a better prediction accuracy, precision, recall, and F1 score than all other models.

4 Conclusion and Future Work

In this paper, fake news is classified by a complex deep learning structure. Firstly, the Word2Vec word embedding layer is to produce the matrix representation of the input news. Then, the deep learning models learn and predict the label of given news. Given an evaluation and a comparison, BiLSTM-CNN outperforms CNN, vanilla RNN, and LSTM. Both advantages of Bi-LSTM and CNN are included in this hyper model. Specifically, the hidden connection between sentences and sequential information from forward and backward are captured. As a result, the proposed model shows its outstanding performance in a task for long-term-text categorical classification.

The dataset used in this paper focuses on the news related to the reports of the American presidential election. Although the designed model and classification receive a preferable accuracy on this dataset, it does not represent it is eligible to solve problems in other fields. Thus, future work will extend the topic and languages in fake news detection problems. If possible, the tone of reporting the news should also be considered as a parameter in the analysis.

References

1. Granik, M., Mesyura, V.: Fake new detection using naïve Bayes classifier. In: 2017 IEEE Ukraine Conference on Electrical and Computer Engineering (UKRCON), pp. 900–903. IEEE, Kyiv, Ukraine (2017)
2. Ahmad, I., Yousaf, M., Yousaf, S., Ahmad, M.O..: Fake news detection using machine learning ensemble methods. Complexity in Deep Neural Networks 1(1), (2020)
3. Yang, Y., Zheng, L., Zhang, J., Cui, Q., Li, Z., Yu, P.S.: TI-CNN: convolutional neural networks for fake news detection. arXiv (2018)
4. Dong, X., Victor, U., Qian, L.: Two-path deep semisupervised learning for timely fake news detection. IEEE Trans. Comput. Soc. Syst. **7**(6), 1386–1398 (2020)
5. Aldwairi, M., Alwahedi, A.: Detecting fake news in social media networks. Procedia Comput. Sci. **141**, 215–222 (2018)
6. Shen, L., Zhang, J.: Empirical evaluation of RNN architectures on sentence classification task. arXiv (2019)
7. Okoro, E.M., Abara, B.A., Umagba, A.O., Ajonye, A.A. Isa, Z.S.: A hybrid approach to fake news detection on social media. Nigerian Journal of Technology **37**(2), 454–462 (2018)
8. Zhang, J., Dong, B, Philip, S.Y.: FAKEDETECTOR: effective fake news detection with deep diffusive neural network. arXiv (2018)
9. Zhang, X., Chen, F., Huang, R.: A combination of RNN and CNN for attention-based relation classification. Procedia Comput. Sci. **131**(1), 911–917 (2018)
10. Bahad, P., Saxena, P., Kamal, R.: Fake news detection using bi-directional LSTM recurrent neural network. Procedia Comput. Sci. **165**, 74–82 (2019)

11. Xu, T., Du, Y., Fu, C., Xie, C.: Incorporating forward and backward instances in a bi-lstm-cnn model for relation classification. In: 2018 IEEE 4th International Conference on Computer and Communications, pp. 2133–2137. IEEE, Chengdu, China (2018)
12. Colah Homepage. https://colah.github.io/posts/2015-08-Understanding-LSTMs. Accessed 24 Mar 2021
13. WILDML Homepage. http://www.wildml.com/2015/12/implementing-a-cnn-for-text-classification-in-tensorflow/. Accessed 24 Mar 2021
14. Kaggle Homepage. https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset. Accessed 30 Dec 2020