Fake News Detection on Social Network

Yiyang Sun[1], Handong Xu[2]

[1]New York University Shanghai [2]New York University Shanghai

Abstract

We designed a text CNN model to detect fake political news from social media. By using the word frequency dictionary, the model outperformed traditional detection systems in terms of both accuracy and running time.

Introduction

People nowadays can easily retweet or repost messages on social media. Without checking the authenticity of the content, it becomes very easy for people to share fake news to their friends by mistake. Fake news on social networks like Twitter, Weibo, Wechat, etc can misguided lots of people and have caused great damage to the society. For example, in the year of 2018, a news wrote that a group of Russian hackers invaded the U.S. electricity grid during the winter. But it turned out this article was not written based on solid facts but vague hypothesis. After realizing the consequence of fake news, many companies nowadays starts to hire people to label fake news manually. However, such a job wastes a lot of time and resources and cannot detect fake news when they first come out online. To detect fake news more effectively and quickly, many algorithms have been designed with the help of deep learning. And many states of art results have been realized which encourage our further research into this area.

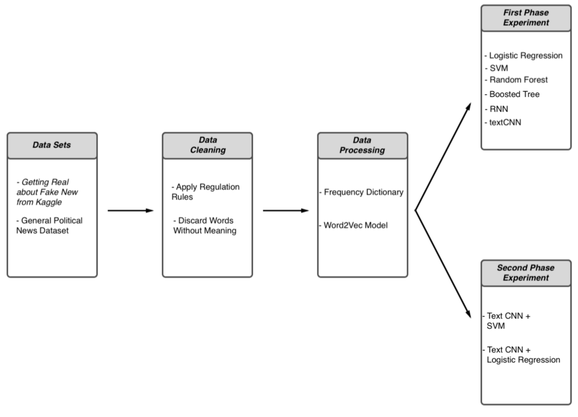

Figure 1: Overall Experiment Process

Methods Used

The following methods were used to complete the data preprocessing:

- word frequency dictionay
- word2vec embedding

The following methods were used to compare the accuracy:

1. Logistic Regression
2. Support Vector Machine
3. Random Forest
4. Boosted Tree
5. Recurrent Neural Network - Long Short-Term Memory
6. Text Convolutional Neural Network

Important Result

- By comparing the overall accuracy and running time, we can notice that the RNN model and our textCNN model with frequency dictionary had the best result. And the running time of textCNN is only 1/10 of the running time with RNN.
- In terms of the precision and recall score, both RNN and textCNN performed very well on dataset1. While with a limited number of data on dataset2, textCNN is not able to label the fake news very well compared with RNN.

First Phase Experiment

Kim designed a CNN model which consists of one-dimensional convolutional layers and max pooling layers[1]. Inspired by Kim's model, we designed our own text-CNN model which can classify news texts with many sentences into the category of true or fake. First, we chose to use the vectors from frequency dictionary instead of the embedded vectors from word2vec method. In order to make the task more similar to a image detection problem, we reshaped the vectors into 4 dimensions which are number of samples, lens of word vector, embedding dimensions and a constant. Then, we applied a two-dimensional convolution matrix instead of one which was supposed to be more sensitive to the pattern. Moreover, to avoid the loss of information during the max pooling process, we chose to use k-max-pooling instead of the traditional 1-max-pooling method.

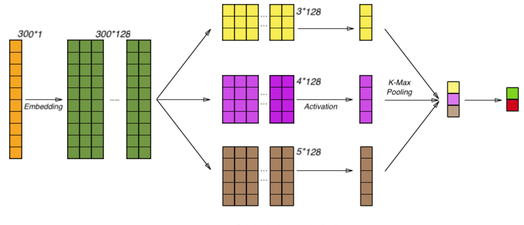

Figure 3: textCNN Model Structure

Second Phase Experiment

The result from the first phase experiment proved that our text-CNN model was effective. However, the result was obtained from training on only one dataset. When we tried to apply the model to a different dataset, the model could not classify the news correctly even though both datasets were related to political news. This could indicate that our model overfit to one specific dataset and loss the capability of detecting general patterns in political news. So we tried to ensemble the models from different methods and datasets.

Accuracy	Without Ensemble-DataSet 1	With Ensemble-DataSet 1	Without Ensemble-DataSet 2	With Ensemble-DataSet 2
0.9*freq-textCNN + 0.1*freq-SVM	98.84%/50.25%	98.63%	50.00%/53.66%	51.21%
0.9*freq-textCNN + 0.1*w2v-SVM	98.84%/89.15%	98.68%	50.00%/78.69%	50.00%
0.6*freq-textCNN + 0.4*w2v-LR	98.84%/88.98%	98.58%	50.00%/78.53%	77.16%
0.6*freq-textCNN + 0.4*freq-LR	98.84%/64.99%	98.63%	50.00%/61.81%	61.80%

Figure 2: Ensemble Result

Result

After the data preprocessing, we can see that the dataset focused mainly on the US presidential election. And there is a strong correlation between the fake news and the word 'trump' from our prediction result.

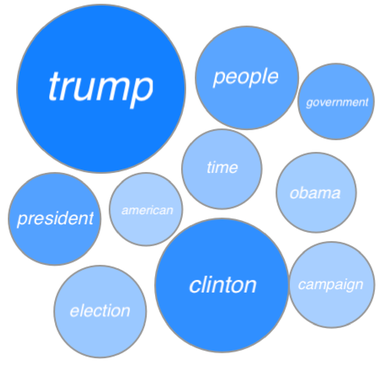

Figure 4: Data Preprocessing Result

Conclusion

During the experiment, we found out that text CNN gave a higher accuracy and auc value than traditional machine learning algorithms for both the word2vec or frequency dictionary methods. Meanwhile, textCNN and LSTM have both returned good results. Also, compared with the LSTM model, text CNN is not able to detect false news very well with a small number of data and the false positive rate is relatively high. However, the run time of textCNN was 10 times faster than that of LSTM. The experiment results showed that textCNN could accurately and efficiently predict the US president election dataset based on only the tile and text of the news.

Future Work

The dataset we used contains mostly US election news with its title and content. With the new round of US election approaching, we plan to obtain more data on the new US president election. Meanwhile, we want to get the propagation structure of these news to perfect the model. So far, researchers only focus on one of the two stages to detect fake news. By running models on both stages, hopefully, the accuracy will increase because more information are added to the model.

References

[1] Yoon Kim.
Convolutional neural networks for sentence classification.
In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751, Doha, Qatar, October 2014. Association for Computational Linguistics.

Acknowledgements

We would like to express our very great appreciation to Professor Ratan Dey for his valuable and constructive suggestions during the planning and development of this research work. Also we are particularly grateful for the assistance given by Professor Olivier Marin.

Contact Information

- Email: ys2380@nyu.edu
- Phone: +86 189 6434 6858

NYU 上海 SHANGHAI 纽约大学