SITCON2019—You are Fake News!

Fake News Detection using Machine Learning

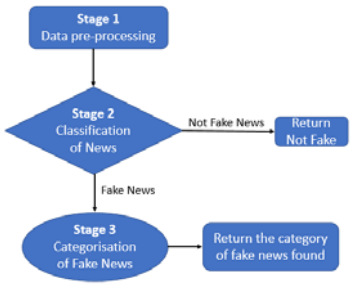
NTU ANTSlab 陳廷易

Abstract

With the rise of social media sites and platforms, it has now become increasingly easy to spread fake news amongst the general public, which might result in averse consequences such as the Washington Pizzeria Attack[1]. It is thought that Machine Learning Techniques could be used to evaluate the truthfulness of a given news article. Using a Convolutional Neural Network and Recurrent Neural Network on a combined dataset, comprised of three different open-source datasets, we evaluate the effectiveness of our model to detect fake news. These findings might be useful for future works into the area of using Machine Learning Techniques to identify fake news.

Implementation Methodology

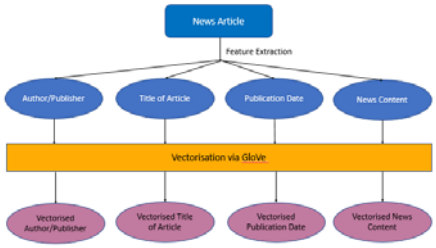
Overview



Our Implementation consists of a 3-stage process as shown in the diagram on the left.

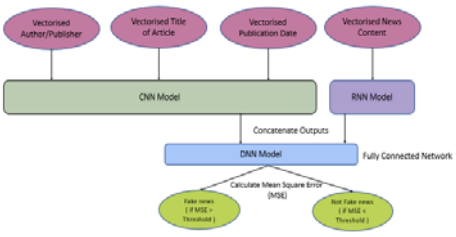
Combining data from the datasets : Liar dataset from UCSB, fake news dataset from Kaggle and dataset published by Signal Media in conjunction with the Recent Trends in News Information Retrieval 2016 conference. We selected the features in common amongst the dataset and used those features for evaluation.

Data pre-processing



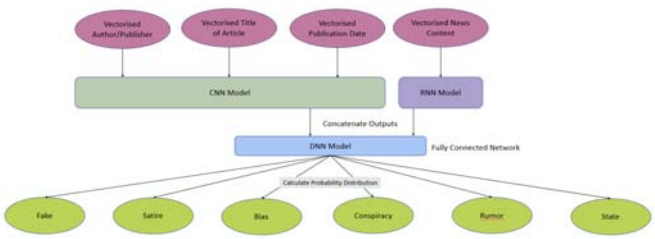
The News Article is then segmented into features and then vectorised as shown in the diagram on the left.

Classification of News

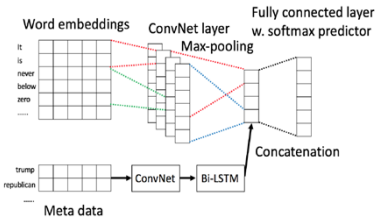


Subsequently, the vectorised features are passed to a CNN or a RNN Model, before it is concatenated together via a DNN Model. DNN will utilise regression to return a Mean Square Error metric that will be used to determine if a given news article is fake or not.

Categorization of Fake News

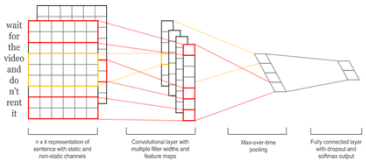


Similarly, we follow the same processes up till the DNN Model. At the DNN model, we will generate a probability distribution that will be used to predict the category of fake news. Another diagram of the process is shown on the right.

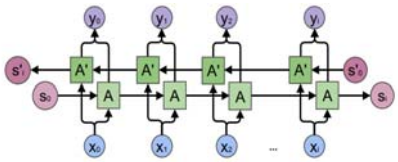


Description of Methods used

- Global Vectors for Word Representation (GloVe)
  GloVe is a pre-trained Word Embedding model which converts words into a vector using the probability of co-occurrence between words to encode the meaning of a given text to determine the relation between words.
- Convolutional Neural Network (CNN)



  Convolutional neural networks (or convnets for short) are used in situations where data can be expressed as a "map" wherein the proximity between two data points indicates how related they are. Neurons don't have to recognize the whole sentences to discover the pattern.
- Recurrent Neural Network (RNN)



  In this project, we utilise a Bidirectional Recurrent Network to split the neurons of a regular RNN into two directions, one for positive time direction (forward states), and another for negative time direction(backward states). Those two states' output are not connected to inputs of the opposite direction states. By using two time directions, input information from the past and future of the current time frame can be used unlike standard RNN which requires the delays for including future information.

Results

As of press time, we were able to obtain an accuracy of 0.4736 on our training data at stage 2. At stage 3, resulting in an accuracy of 0.8662 on the holdout test set. A sample of our results is shown below.

```
Enter the date of the news: 2016-10-27
Enter the author of the news or leave it if unknown: Andrea Lawson Gray
Enter th title of the suspicious news: Comment on HALLOWEEN IN THE CASTRO RETURNS IN 2014!
Enter the article text: It will be recalled that the Halloween bash attracted as many as 500,000 each year and has become a maj
or tourist attraction, second and third only to the Pride parade and Folsom Street Fair. Stabbing and shooting incidents, belie
ved to be perpetrated by straight revelers, prompted city officials to permanently ban the event beginning in 2010.
```

```
The result is:
Most of them are true
It might has bias.
```

Combining news we collected from *New York Times* and *Guardian* of U.S. presidential election in 2016. We believe that given the datasets available to us, we have obtained relatively good results as we have managed to obtain accuracies higher than the ones published in the paper presenting the Liar Dataset[4].

Conclusion

With the news we could access, and using the common features amongst the different datasets. We not only are able to utilise CNN and RNN to make a judgement of the authenticity of a news article, but also the category of fake news a news article might belong to, if it is deemed as fake.

Future Works

- Online Learning for training of model
- Application of model on news apart from the US Presidential Election
- Usage of different models for evaluation
- Utilisation of more features in the model

References

[1] Goldman, C. K. (2016, December 05). In Washington Pizzeria Attack, Fake News Brought Real Guns. Retrieved January 03, 2018, from https://www.nytimes.com/2016/12/05/business/media/comet-ping-pong-pizza-shooting-fake-news-consequences.html
[2] Bidirectional recurrent neural networks. (2017, December 17). Retrieved January 03, 2018, from https://en.wikipedia.org/wiki/Bidirectional_recurrent_neural_networks
[3] 22, 2. I. (n.d.). Hashrocket. Retrieved January 03, 2018, from https://hashrocket.com/blog/posts/a-friendly-introduction-to-convolutional-neural-networks
[4] Wang, W. Y. (2017, May 01). "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection. Retrieved January 03, 2018, from https://arxiv.org/abs/1705.00648