FAKE NEWS DETECTION USING MACHINE LEARNING

[1]V. Monica,
Scholar, Department of Information Technology, Velagapudi Ramakrishna Siddhartha Engineering College,vuppalamonica@gmail.com
[2]N. Satya Krishna,
Ph.D, Assistant Professor, Department of Information Technology, Velagapudi Ramakrishna Siddhartha Engineering College,satya.krishna.nunna@gmail.com

Abstract

social media, fake news has become a social problem and, in some cases, spreads more and faster than true information. In this paper, I evaluate the performance of an attention mechanism for detecting fake news on two datasets, one containing traditional online news articles and the other news. from various sources. I compared the results on the datasets and the results of Attention Mechanism with LSTM and traditional machine learning methods. it shows the attention mechanism does not work as well as expected. In addition, I made changes to the original attention mechanism using word2vec embedding which works better in this particular case.

4108

DOI Number:10.14704/nq.2022.20.8.NQ44442 **NeuroQuantology 2022; 20(8): 4108-4117**

Introduction

Fake news has quickly become a society problem, being used to propagate false or rumor information in order to change people's behaviour. It has been shown that propagation of fake news has had a non-negligible influence of 2016 US presidential elections. Fake news has also been used in order to influence the referendum in the United Kingdom for the "Brexit". In this paper I experiment the possibility to detect fake news based only on textual information by applying traditional machine learning techniques as well as bidirectional- LSTM and attention mechanism on two different datasets that contain different kinds ofnews. In order to work on fake news detection, it is important to understand what is fake news and how they are characterized. The following is based on Fake News Detection on Social Media: A Data MiningPerspective.

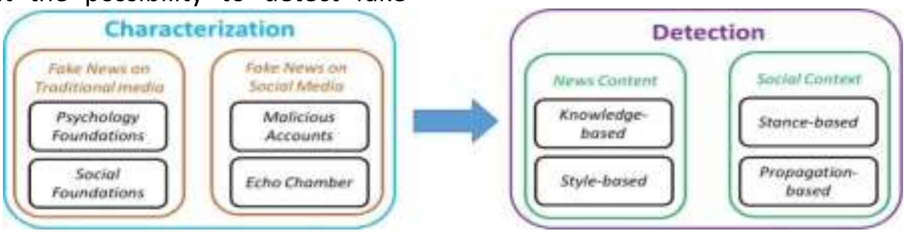

Figure 1: Fake news on social media: from characterization to detection.

1.

1 News ContentFeatures

Features will be extracted from these four basic components, with the mains features being linguistic-based and visual-based. As explained before, fake news is used to influence the consumer, and in order to do that, they often use a specific language in order to attract the readers. On the other hand, non-fake news will mostly stick to a different language register, being more formal. This is linguistic-based features, to which can be added lexical features such as the total number of words, frequency of large words or unique words.

The second features that need to be taken into account are visual features. Indeed, modified images are often used to add more weight to the textual information. For example, the Figure 2is supposed to show the progress of deforestation, but the two images are actually from the same original one, and in addition the WWF logo makes it lookliketobefromatrustedsource.

Figure 2: The two images provided to show deforestation between two dates are from the same image taken at the same time.

1.2 Social ContextFeatures

In the context of news sharing on social media, multiple aspect can be taken into account, such as user aspect, post aspect and group aspect. For instance, it is possible to analyse thebehaviourofspecificusersandusetheirmeta datainordertofindifauserisat risk of trusting or sharing false information. For instance, this metadata can be its centre of interest, its number of followers, or anything that relates to it. It is also possible to extract features from the content using latent Dirichlet allocation (LDA).

- **Expert-oriented**: relies on experts, such as journalists or scientists, to assess the newscontent.
- **Crowdsourcing-oriented**: relies on the wisdom of crowd that says that if a sufficiently large number of persons say that something is false or true then it should be.
- **Computational-oriented**: relies on automatic fact checking, that could be based on external resources such asdbpedia.

These methods all have pros and cons, hiring experts might be costly, and expert are limited in number and might not be able to treat all the news that is produced. In the case of crowdsourcing, it can easily be fooled if enough bad annotators break the system and automatic fact checking might not have thenecessary accuracy.

As explained earlier, fake news usually tries to influence consumer behaviour, and thus generally use a specific style in order to play on the emotion. These methods are called deception-oriented stylometric methods. The second method is called objectivity-oriented approaches and tries to capture the objectivity of the texts or headlines.

These kind of style is mostly used by partisan articlesoryellowjournalism,thatis,websitesth atrelyoneye-catchingheadlineswithout reporting any useful information. An example of these kind of headline couldbe This kind of headline plays on the curiosity of the reader that would click to read the news.

The last features that have not been used yet are social media features. There are two approaches to use these features: stance-based and propagation-based.

- **Stance-based** approaches use implicit or explicit representation. For instance, explicit representation might be positive or negative votes on social media. Implicit representation needs to be extracted from the post itself.
- **Propagation-based** approaches use features related to sharing such as the number ofretweets on twitter.

Figure 3: Different approaches to fake news detection.

When it comes to state of the art for text classification, it includes Long short-term memory (LSTM), Attention Mechanism, indrnn, Attention-Based Bidirection Hierarchical Attention Networks for Text Classification, Adversarial Training Methods For Supervised Text Classification, Convolutional Neural Net- works for Sentence Classification and RMDL: Random Multimodel Deep Learning for Classification.

2Literature Survey

Reis et al. [13] use machine learning techniques on buzz feed article related to US election. The evaluated algorithm are k-Nearest Neighbours, Naive-Bayes, Random Forests, SVM with RBF kernel and XG Boost. Pérez-Rosas etal. [14]used almost thesamesetoffeaturesbutused linear SVM as a model and worked on a different dataset. Ruchansky et al.[15] used a hybrid network, merging news content features and meta-data such as social engagement in a single network. To do so, they used an RNN for extracting temporal features of news content and a fully connected network in the case of social features. The results of the two networks are then concatenated and use for final classification.

Tacchini et al. [16] focus on using social network features in order to improve the relia- bility of their detector. The dataset was collected using Facebook Graph API, collection pages from two main categories: scientific news and conspiracy news. They used logistic regression and harmonic algorithm to classify news in categories hoax and non-hoax. Harmonic Algorithm is a method that allows transferring information across users who liked some commonposts.

Thorne et al.[17] worked on Fake News Challenge by proposing a stack of different clas- sifiers: a multilayer perceptron with relu activation on average of word2vec for headline and tf-idf vectors for the article body, average word2vec for headlines and article body, tf-idf bigrams and unigram on article body, logistic regression with L2 regularization and concatenation of word2vec for headlines and article body with MLP and dropout.Finally, a gradient boosted tree is used for the final classification. Yang et al. [18] used a CNN with images contained in article in order to make the classification. They used kaggle fake news dataset1, in addition they scrapped real news from trusted source such as [4] New York Times and Washington Post. Their network is made of two branches: one text branch and one image branch.

We have seen in the previous sections that most of the related works focus on improving the prediction quality by adding additional features. The fact is that these features are not always available, for instance some article may not contain images. There is also the fact that using social media information is problematic because it is easy to create a new account on these media and fool the detection system. That's why I chose to focus on the article body only and see if it is possible to accurately detect fake news.

3 Data Exploration

A good starting point for the analysis is to make some data exploration of the data set.
The first thing to be done is statistical analysis such as counting the number of texts per
class or counting the number of words per sentence. Then it is possible to try to get an
insight of the data distribution by making dimensionality reduction and plotting data in
2D.

This works uses multiple corpus in order to train and test different models. The main
corpus used for training is called Fake News Corpus[29]. This corpus has been automatically
crawled using opensources.co labels. In other words, domains have been labelled
with one or more labels in

Because Fake News Corpus is the main dataset, the data exploration will start withthis dataset. And the first thing is to count the number of items per class. Before startingthe analysis, it is needed to clean up the dataset. As it is originally given in a large 30GBCSV file, the first step is to put everything in a database in order to be able to retrieveonly wanted a piece of information. In order to do so, the file has been red line by line.It appears that some of the lines are badly formated, preventing them from being readcorrectly, in this case they are dropped without being put in the database. Also, eachline that is a duplicate of a line already red is also dropped. The second step in cleaningthe set consists of some more duplicate removal. Indeed, dropping same lines removeonly exact duplicate. It appears that some news does have the same content, with slightvariation in the title, or a different author. In order to remove the duplicate, each text ishashed using SHA256 and those hash a compared, removing duplicates and keeping only

4 Methodology

LSTM or Long Short-Term Memory[8] is a kind of recurrent neural network that fits well to temporal or sequential input such as texts. A RNN is a type of neural network where the hidden state is fed in a loop with the sequential inputs. There are usually shown as unrolled version of Each of the Xi being one value in the sequence. In this case, Xi values are word vectors. There are two possibilities, either use pre-trained vector with word2vec or make Xi inputs a parameter to learn in the same way as it works for the Word2Vec algorithm, having a one-hot encoding of the word and a matrix of weights to tune. Each method will be used. Recurrent Neural Networks do not work very well with long-term dependencies, that is why LSTM have been introduced. It is made of an input gate, an output gate and a forget gate that are combined

$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f)$ (1)

$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i)$ (2)

$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o)$ (3)

$c_t = f_t \circ c_{t-1} + i_t \circ \sigma_c(W_c x_t + U_c h_{t-1} + b_c)$ (4)

$h_t = o_t \circ \sigma_h(c_t)$ (5)

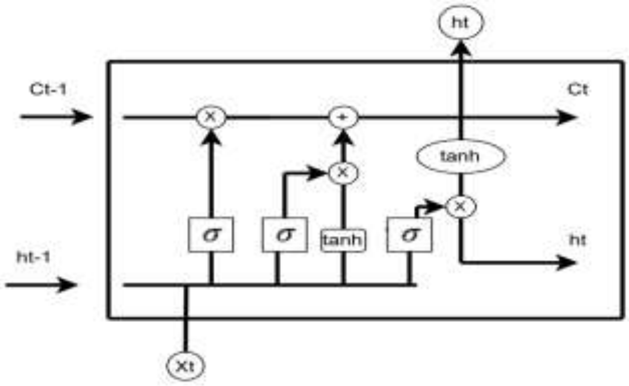

Figure 4 Bidirectional LSTM

Figure 4shows how it works. A bidirectional LSTM works the same way, but the input is fed in the two directions, from the start to the end and from the end to the start.

Attention mechanism adds an extra layer between LSTM outputs and the final output of the network. It merges word-level features into sentence features using a weight vector.

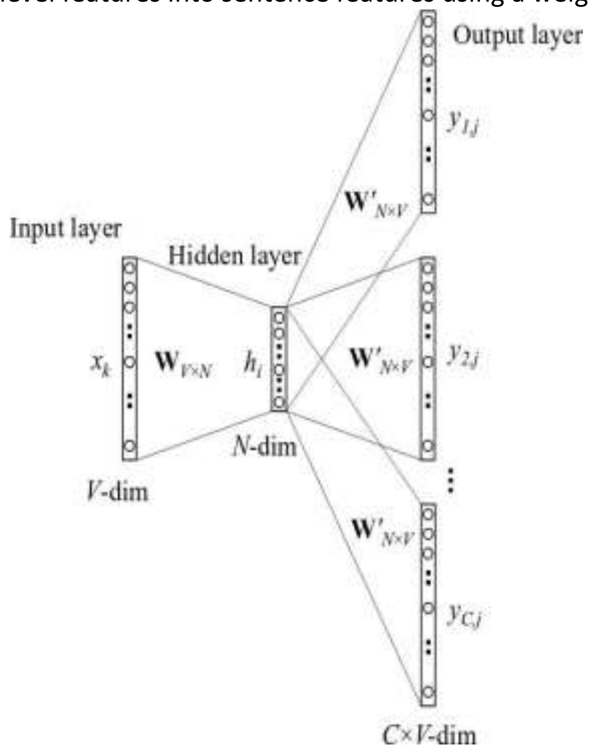

Figure 5 Skip-gram model with multiple outputs.

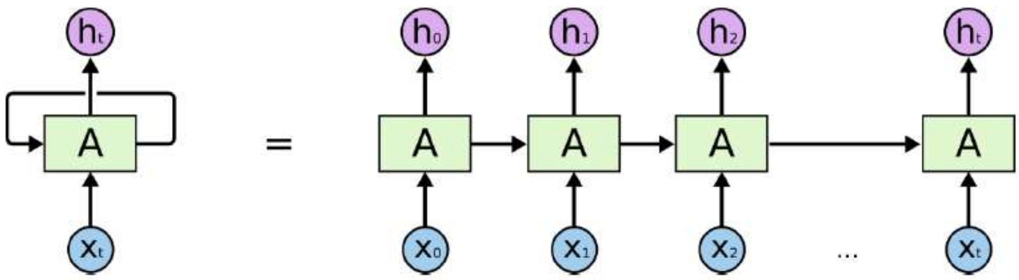

Figure 6: Unrolled RNN (Understanding LSTM Networks, https://colah.github.i o/posts/2015-08-Understanding-LSTMs/)

Results

As explained earlier, both models have been trained using different embedding: the first one being pre-trained word2vec vectors of size 300 and the second one being a tunable parameter with different embedding size.

When it comes to LSTM trained on liar-liar dataset, it simply does not works. It classi- fies almost all the texts as being from the same class. Although, it reaches a good score on the training data, it does not manage to generalize correctly. Figure 7shows the recall, precision and f1-score and loss for training and testing set of the best models for the LSTM using word2vec. We can see that even if the training score increase, the testing

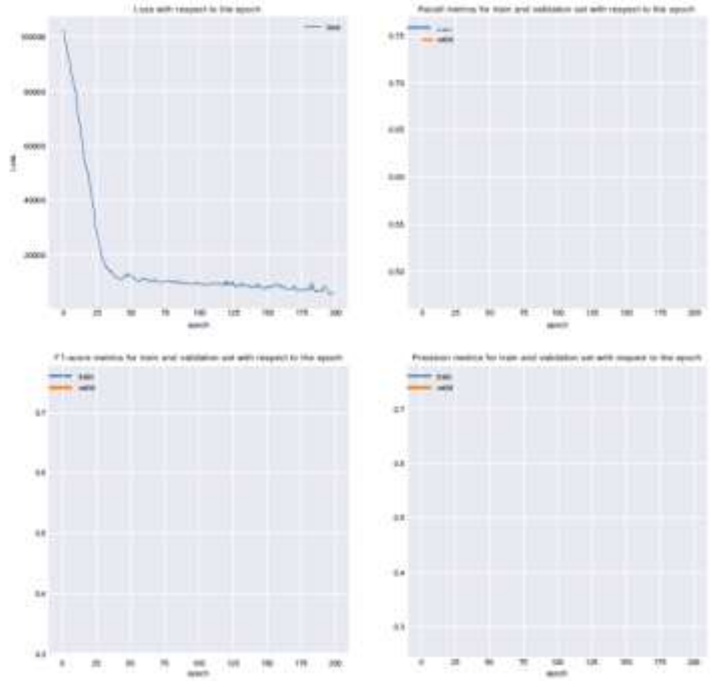

Figure 7: Best LSTM With word2vec

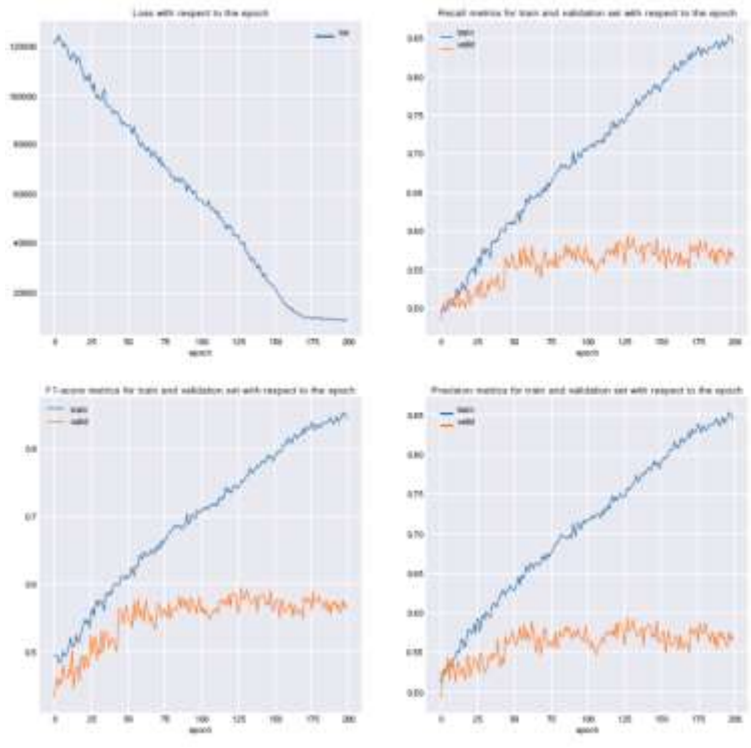

Figure 8: Best LSTM with word embedding as tunable parameters.

Attention Mechanism

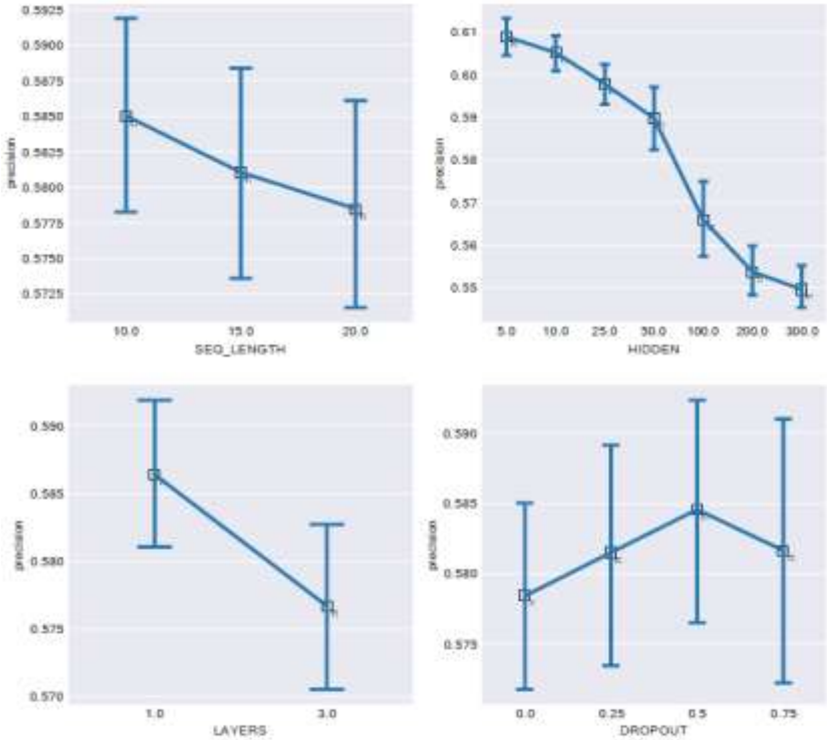

Figure 9: Confidence Interval of Precision for Each Parameter Value

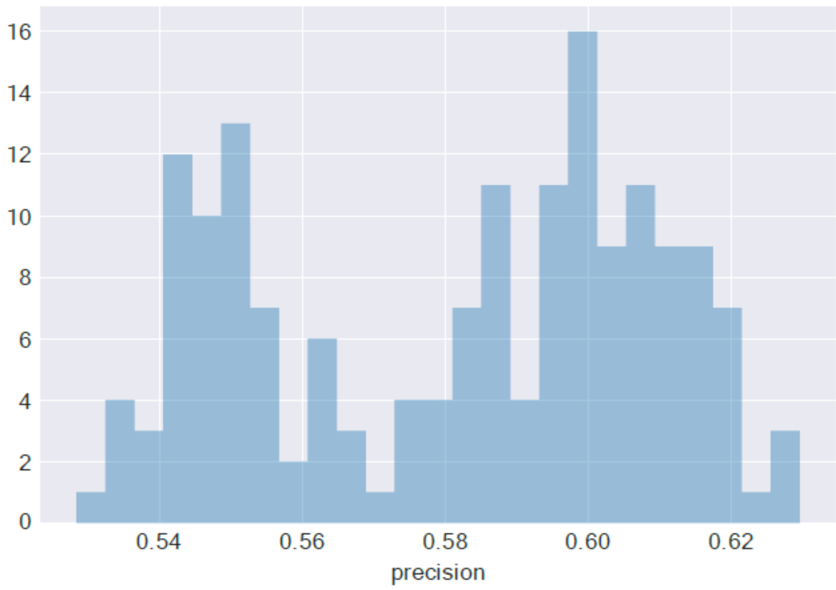

Figure 10: Distribution of the precision of best epochs for all the models trained with word2vec embedding.

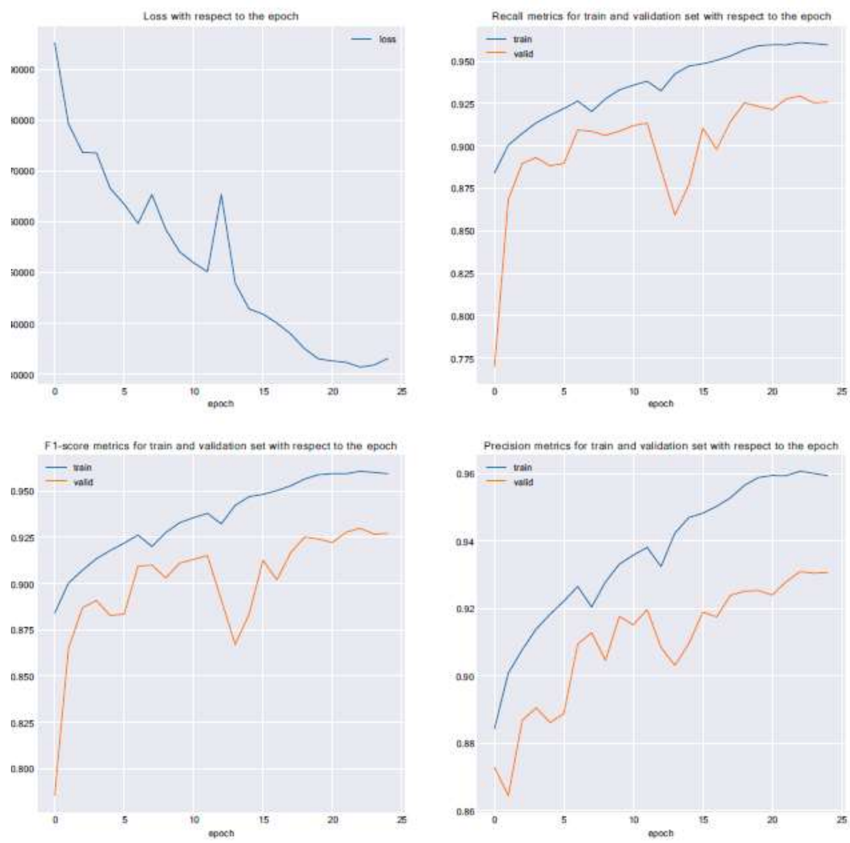

Figure 11Training plots of the best LSTM using word2vec embedding

The same way as in the models will be trained on the parameters that produced the best results on the training set, and trained on training and validation set, and tested on testing set. The parameters used for training are given at Table 1. The results for all four models

model	embedding size	Sequence Length	num hiddens	dropout	Early Stop
LSTM	300	10	50	0.75	126
LSTM + word2vec	300	10	50	0.0	160
Attention	10	20	10	0.75	400
Attention + word2vec	300	20	5	0.75	25

Table 1: Parameters used for training

Conclusion

In this chapter I have investigated how state-of-the-art deep learning models work on fake news detection, and it shows that for the particular case of fake news detection it does not outperform traditional machine learning methods. I have also made some addition to the original model that improves the performances by a few percent by replacing the tunable word embedding by constant one using word2vec. It shows out that it helps re-duceoverfittingandincreaseresultonthetesting set. A hypothesis to explain why these two deep learning methods do not works as well as machine learning methods is the fact that in this case text are required to be the same size. Which means that some of them require some padding and the other are srunk. In the second case, information islost.In addition, it shows that Liar-Liar Corpus is hard to work on, with 60% precision, when Fake News Corpus still have good results.

Futureworks

Basing fake news detection only on supervised models on text have shown not to be enough in all the cases. In order to solve this problem, most of the research focus on additional information such as author information. I think the most successful approach would be automatic fact checking model, that is, compelling the model with some kind of knowledge base, the purpose of the model would then be to extract information for the text and verify the information in the database. The problem with this approach would be that the knowledge base would need to be constantly and manually update to stay up to date.

References

1. Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 elec- tion.InJournalofEconomicPerspective,volume31,2017.
2. Sepp Hochreiter and Jürgen Schmidhuber.Longshort-term memory. Neural Com- putation, 9:1735–1780,1997.
3. Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter, 19(1):22–36,2017.
4. David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. Journal of machine Learning research, 3(Jan):993–1022, 2003.
5. Craig Silverman and Lawrence Alexander. How teens in the balkans are duping trump supporters with fake news. Buzzfeed News, 3,2016.
6. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and IlliaPolosukhin. Attention is all you need. In NIPS, 2017.
7. Shuai Li, Wanqing Li, Chris Cook, Ce Zhu, and Yanbo Gao. Independently recurrent neural network (indrnn): Building a longer and deeperrnn.
8. Peng Zhou, Wei Shi, Jun Tian, Zhenyu Qi, Bingchen Li, Hongwei Hao, and Bo Xu. Attention-based bidirectional long short-term memory networks for relation classi- fication. In Proceedings of the 54th Annual Meeting of the Association for Com- putational Linguistics (Volume 2: Short Papers), pages 207–212, Berlin, Germany, August 2016. Association forComputational Linguistics.
9. Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. Hierarchical Attention Networks for Document Classification. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computa- tional Linguistics: Human Language Technologies, pages 1480–1489, San Diego, Cal- ifornia, 2016. Association for ComputationalLinguistics.
10. TakeruMiyato, Andrew M. Dai, and Ian Goodfellow. Adversarial Training Meth- ods for Semi-Supervised Text Classification. arXiv:1605.07725 [cs, stat], May 2016. arXiv:1605.07725.
11. Yoon Kim. Convolutional Neural Networks for Sentence classification. arXiv:1408.5882 [cs], August 2014. arXiv:1408.5882.
12. Kamran Kowsari, MojtabaHeidarysafa, Donald E. Brown, Kiana Jafari Meimandi, and Laura E. Barnes. RMDL: Random Multimodel Deep Learning for Classification. Proceedings of the 2nd International Conference on Information System and Data Mining - ICISDM '18, pages 19–28, 2018. arXiv:1805.01890.
13. JulioCSReis,AndréCorreia,FabrícioMurai,AdrianoVeloso,FabrícioBenevenuto, and Erik Cambria. Supervised learning for fake news detection. IEEE Intelligent Systems, 34(2):76–81,2019.
14. VernicaPrez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. Au- tomatic detection of fakenews.
15. Natali Ruchansky, SungyongSeo, and Yan Liu. Csi: A hybrid deep model for fake news detection. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pages 797–806. ACM,2017.
16. Eugenio Tacchini, Gabriele Ballarin, Marco L. Della Vedova, Stefano Moret, andLuca de Alfaro. Some like it hoax: Automated fake news detection in socialnetworks.

17. James Thorne, Mingjie Chen, GiorgosMyrianthous, Jiashu Pu, Xiaoxuan Wang, and Andreas Vlachos. Fake news stance detection using stacked ensemble of classifiers. In Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism, pages 80–83,2017.
18. Yang Yang, Lei Zheng, Jiawei Zhang, QingcaiCui, Zhoujun Li, and Philip S. Yu. Ti- cnn: Convolutional neural networks for fake newsdetection