Fake News Detection Using Various Datasets

Shamir Towsif Khandaker
Computer Science
University of Louisiana at Lafayette
Lafayette, U.S.A
C00423766@louisiana.edu
https://github.com/shamirtowsif/DataMiningProject.git

I. Introduction

In this project I compare different deep learning models for detecting fake news detection. For my dataset I am using [1], [2] and [3]. [1] is a "publicly available dataset for fake news detection", [2] is a dataset comprising "a complete sample of news published in Facebook from 9 news agencies over a week close to the 2016 U.S. election from September 19 to 23 and September 26 and 27" while [3] is a dataset "collected from a browser extension called BS detector developed for checking news veracity". I have compared six different deep learning models which are RNN, Bidirectional RNN, LSTM, Bidirectional LSTM, GRU and Bidirectional GRU. I have calculated accuracy for all of them. For [1], I am getting the best accuracy for Bidirectional GRU and the worst result for RNN. For [3] , I am getting the best accuracy for RNN and the worst result for GRU. I couldn't run any deep learning model for [2]. So I trained the dataset on a fully connected model. So here I got only one accuracy.

There has been an explosion of fake news, rumors and misinformation with the advent of digital age. It has seriously undermined trust in many public and private institutions which has for years served as the pillars of democracy, accountability, freedom of speech and freedom of the press. Fake news had a deep impact upon the very contentious "Brexit" referendum and it has also made the 2016 U.S. presidential election very controversial. Not to mention the enormous impact it has been having in recent years in destroying the communal fabric of emerging new and democratic countries such as Pakistan and Bangladesh. One statistic of the impact that fake news is having in recent times is that during the 2016 U.S. presidential elections there were many fake news circulating on the social media and among them the top twenty "generated 8,711,000 shares, reactions, and comments on Facebook, ironically, larger than the total of 7,367,000 for the top twenty most-discussed election stories posted by 19 major news websites" [4]. If these alone doesn't alarm society then we should also mention the enormous impact it can have on the economy as well. Previously there was a fake new claiming that former U.S. President Barack Obama was killed in an explosion and that news wiped out $130 in stock value [4].

Fake news and rumors are not a new phenomenon. Since the invention of the printing press there has always been people who wanted to use these types of technology to deceive people, to mobilize them for various political as well as economic purposes. So the obvious question that arises is that why has the impact of fake news been so severe recently? Its not a simple question to answer. One reason is the internet. Internet has connected diverse groups of people across the world. People and communities who have lived isolated from each other have suddenly discovered that they are connected to global world. Many good things have come out of this such as this global connectivity has increased the transaction of goods. knowledge and ideas. But one of the bad byproduct of this is that it has enabled some people to deliberately deceive people and disseminate their toxic and poisonous views. But far more bigger reason is the emergence of the social media such as Facebook and Twitter. Facebook and Twitter has created huge chunks of isolated echo chambers where people from across the world, unmoored from their local traditions, are extremely susceptible to fake news and radical and toxic ideas. In these echo chambers "biased information is often amplified and reinforced" [4].

The rise of fake news has made the dissemination of it a very lucrative offer as well. According to NBC news dozens of teenagers from Velas, Macedonia has joined a campaign of disseminating fake news where each person has "earned at least $60,000 in the past six months – far outstripping their parents' income and transforming his prospects in a town where the average annual wage is $4,800" [4]. Which begs another question? It's common sense that many individuals, especially youth, are attracted to radical ideas and ideologies? But come fake news sources, sources which designs news specifically designed to deceive their readers, have gained their trust? Many sociological and psychological factors come into play. As well as the reasons listed above, another reason is that " human ability to detect deception is only slightly better than chance: typical accuracy rates are in the 55%-58% range, with a mean accuracy of 54% over 1,000 participants in over 100 experiments" [4]. Which is why we need dynamic and robust methods to detect fake news to help the people distinguish between truth and falsehood.

A. *What is fake news?*

Before we can discuss how we can detect fake news we have to properly define it. There is no universal definition of fake news. But the definition that we are concerned with is -

An news article which contains false information specifically inserted for malicious intent.

So for a news article to be fake news, it is not enough for it to contain false information because many honest journalists sometimes get their information wrong. It might be categorized as false news but it is not fake news. But for a false news, the false information inside it has to be inserted there for malicious intention. So when we are detecting fake news it is not enough to say that the news article contains false information. We also have to know the intention of the author of the article. And it is precisely this last part which has made fake news detection such a difficult task. In this paper we trying to figure out whether various deep learning models are capable of detecting the malicious intent of the user or not as well as which models perform better.

B. Background

As we have discussed above, there are two features that define a news article as fake news. These two features are 1) false information and 2) malicious intent. So we can divide our fake news detection model into two separate modules. They are 1) fact-checking module and 2) intent detection module

C. Fact checking module

Fact checking modules are used to detect the accuracy of various claims in a news article. There are three ways to do fact checking. These ways are 1) expert oriented fact checking 2) crowd sourcing oriented fact checking and 3) automated fact checking.

1) Expert Oriented Fact Checking: In expert oriented fact checking there is a domain expert who goes through the news articles thoroughly and then labels whether the claims made in the article are true or not. Expert oriented articles are the best way for checking. But they take a lot of time and they are also not scalable. There many online fact checking websites. In Table: I, we have shown a comparison of various fact checking websites.

	Topics Covered	Content Analyzed	Assessment Labels
PolitiFact	American Politics	Statements	True; Mostly true; Half true; Mostly false; False; Pants on fire
The Washington Post Fact Checker	American politics	Statements and claims	One pinocchio; Two pinocchio; Three pinocchio; Four pinocchio; The Geppetto checkmark; An upside-down Pinocchio; Verdict pending
FactCheck	American Politics	TV ads, debates, speeches, interviews and news	True; No Evidence False
Snopes	Politics and other social and topical issues	News articles and videos	True; Mostly true; Mixture; Mostly false; False; Unproven; Outdated; Miscaptioned; Correct attribution; Misattributed; Scam; Legend
TruthOrFiction	Politics, religion, nature, aviation, food, medical, etc.	Email rumors	Truth; Fiction; etc.
FullFact	Economy, health, education, crime, immigration, law	Articles	Ambiguity (no clear labels)
HoaxSlayer	Ambiguity	Articles and messages	Hoaxes, scams, malware, bogus warning, fake news, misleading, true, humour, spams, etc.

TABLE I: Comparison of various expert oriented fact checking websites

2) Crowd Sourcing Oriented Fact Checking: Crowd sourcing is another way to do fact checking. Crowd sourcing basically asks the crowd to label the news article. Then this methods aggregates the labels into a single label. The most popular crowd sourcing website that we know of is called *https://www.fiskkit.com*. This method is much more scalable and robust than expert oriented methods. But the result it provides is highly unrelieable.

3) Automated Fact Checking: Automatic fact checking does two things. They are 1) extract verifiable claims 2) verifying those claims. There can be many things written in a news article. Every sentence in the article shouldn't be verified to begin with. Every sentence in the article might not even be verified to begin with. The first task of the automated fact checker is to extract verifiable claims from the article. The second task is to verify the claims. We can verify the data against some external resources. We can use two types of external resources. They are *open web* and *knowledge graph*. Open web has been work on [5]. There are many knowledge graphs available right now. Two of the most famous ones are *DBPedia* and *Google Relation Extraction Corpus*.

D. Intent Detection Module

It is very difficult to detect an author's intent from the text that he writes. Especially when the text that he writes is specifically written to deceive the readers. Nonetheless in recent years there has been many style based intent detection deep learning models that have been developed. These models can be divided into 1) deception oriented models and 2) objectivity oriented models.

1) Deception Oriented Models: Many advanced Natural Language Processing models have been used for deception detection. These models search for two things. They are 1) deep syntax 2) rhetorical structure. Probabilistic context free grammar can be used to detect deep syntax within sentences. Rhetorical Structures can be utilised to capture the differences between deceptive and truthful sentences [6].

2) Objectivity Oriented Models: Objectivity oriented models try to extract style signals which can mean that the article isn't objective enough. There have been studies on hyper partisan news sources showing that articles from these news sources have particular linguistic features. There have been many studies on trying to recognize on clickbait titles as well which might mean that the author who has written that title has less interest in informing the audience and more interest sensationalism [6].

II. Methodology

I approached the problem as a text classification problem. I wanted to use deep learning models for this text classification problem. The first model to be used on text was RNN since RNN has an “internal state memory” that can process sequence of inputs. So that makes it ideal for processing text [7]. Here on 1 an RNN model is shown.

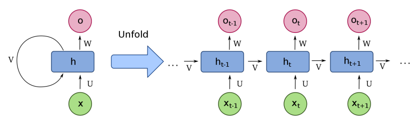

Fig. 1: RNN Model

I expect LSTM and GRU to perform the better than RNN. There are feedbacks loops in the recurrent layer of all RNNs which lets these models hold “memory” over time. But, because of the exponential decay of the gradient, it can be difficult to train standard RNNs to solve problems that require learning longterm temporal dependencies. LSTM network is a type of RNN that uses special units in addition to standard units.This special unit is called a ‘memory cell’ that can maintain information in memory for long periods of time. A set of gates is used to control when information enters the memory, when it’s output, and when it’s forgotten. This architecture lets them learn longer-term dependencies. GRUs are similar to LSTMs, but use a simplified structure. GRUs are used a lot recently since it is computationally faster. But the performance of GRU is almost same as LSTM. Here on 2 and 3, an LSTM and a GRU model is shown.

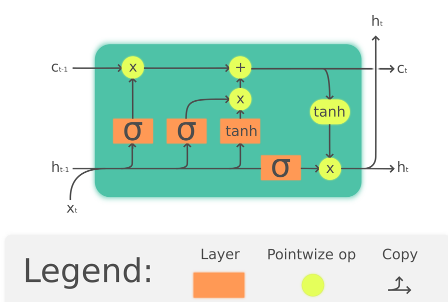

Fig. 2: LSTM Model

One of the big problems of fake news detection is that there aren’t very good publicly available datasets round for running deep learning models. We could find four publicly available datasets. These datasets are 1) BuzzFeedNews 2) Liar liar dataset 3) BS Detector and 4) Credbank.

A. Dataset

1) BuzzFeedNews: People have collected news from 9 different news agencies published in face book around the

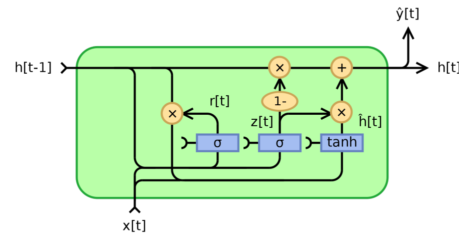

Fig. 3: GRU Model

Datasets Statistics	
Training set size	10,269
Validation set size	1,284
Testing set size	1,283
Avg. statement length (tokens)	17.9
Top-3 Speaker Affiliations	
Democrats	4,150
Republicans	5,687
None (e.g. FB Posts)	2,185

TABLE II: The LIAR Dataset Statistics

2016 presidential election. ”from September 19 to 23 and September 26 and 27”. These articles were fact-checked and labeled by 5 Buzzfeed journalists. It also contains ”linked articles, attached media, and relevant metadata. It contains 1,627 articles–826 mainstream, 356 left-wing, and 545 right-wing articles” [2].

2) LIAR: This dataset was created from the PolitiFact website which we have discussed earlier. ”It includes 12,836 human-labeled short statements, which are sampled from various contexts, such as news releases, TV or radio interviews, campaign speeches, etc.” This dataset has been labeled with six different labels. They are 1) True 2) Mostly True 3) Half True 4) Barely True 5) False and 6) Pants on Fire. These dataset has not only claims from a news article but also various metadata for these claims such as the speaker of the claim, the job of the speaker, party of the speaker, the state from which the article has been originated, the subject of the claim and the venue from which the speaker originated the claim. These are all metadata for the claim [1].

3) BS Detector: This dataset was created and labeled by a browser extension. When you visit a website, the browser extension visits all the methods of the website and checks them against a manually compiled website list. And this is how the browser extension has created the fake new list. That means that the labelling has not been done by an expert unlike the previous ones [3].

4) Cred Bank: This is a crowd sourced dataset ”of approximately 60 million tweets that cover 96 days starting from October 2015.” It has been categorized info 1000 news events and it has been labeled by Amazon Mechanical Turk [8].

I have chosen to run our experiments on the “LIAR LIAR”, “Buzzfeed”” and “BS Detector” dataset. In II we are showing the statistics for the “LIAR LIAR” dataset.

In III we are showing the statistics for the “Buzzfeed”

Datasets Statistics	
Training set size	1369
Validation set size	456
Testing set size	457

TABLE III: The LIAR Dataset Statistics

Datasets Statistics	
Training set size	7799
Validation set size	2600
Testing set size	2600

TABLE IV: The BS Detector Dataset Statistics

dataset.

In IV we are showing the statistics for the “BS Detector” dataset.

B. *Experiments*

We are comparing six different deep models. These models are RNN, Bidirectional RNN, LSTM, Bidirectional LSTM, GRU and Bidirectional GRU. We are creating one-hot vectors for input. Then we decided to use the *glove.6B.100d* to create our embeddings. For the experiments in the “Liar Liar” dataset we have also used the Parts of Speech of the words in the claims as inputs as well. And also we have used Syntactic dependency for our inputs. For the experiments in “Liar Liar” and “BS Detector” dataset we have used the Adam optimizer and built in keras models. We are using 30 epochs. Our size of the hidden layer is 100. Unfortunately, in regards to the Buzzfeed dataset, there was no text for me to run any deep learning model. So I used a feed forward network.

III. Results

For “Liar Liar” dataset we can see that we get the best accuracy for Bidirectional GRU and the worst result for RNN. All of our results are in Table: V.

For “BS Detector” dataset we can see that we get the best accuracy for RNN and the worst result for GRU . All of our results are in Table: VI.

For “Buzzfeed” dataset our results are in Table: VII.

Models	Accuracy
RNN	21.55
Bidirectional RNN	22.18
LSTM	21.70
Bidirectional LSTM	22.34
GRU	23.28
Bidirectional GRU	23.36

TABLE V: The LIAR Dataset Results

Models	Accuracy
RNN	86.81
Bidirectional RNN	85.69
LSTM	86.27
Bidirectional LSTM	86.58
GRU	84.46
Bidirectional GRU	85.88

TABLE VI: The BS Detector Results

Models	Accuracy
Fully Connected Network	72.87

TABLE VII: The Buzzfeed Result

IV. Future Work

In future we want to run CNN and Hybrid CNN and LSTM on this dataset. We also intent to Hybrid LSTM and Capsule Neural Network on this dataset.

V. Conclusion

In this work we compare different deep learning models for the task of fake news detection. But since we are not checking the veracity of the claims, that is why we are not getting very good accuracy.

References

[1] W. Y. Wang, “”liar, liar pants on fire”: A new benchmark dataset for fake news detection,” *CoRR*, vol. abs/1705.00648, 2017. [Online]. Available: http://arxiv.org/abs/1705.00648

[2] M. Potthast, J. Kiesel, K. Reinartz, J. Bevendorff, and B. Stein, “A stylometric inquiry into hyperpartisan and fake news,” 2017.

[3] “Getting real about fake news,” https://www.kaggle.com/mrisdal/fake-news.

[4] X. Zhou and R. Zafarani, “Fake news: A survey of research, detection methods, and opportunities,” *CoRR*, vol. abs/1812.00315, 2018. [Online]. Available: http://arxiv.org/abs/1812.00315

[5] M. Banko, M. J. Cafarella, S. Soderland, M. Broadhead, and O. Etzioni, “Open information extraction from the web.” in *Ijcai*, vol. 7, 2007, pp. 2670–2676.

[6] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake news detection on social media: A data mining perspective,” *CoRR*, vol. abs/1708.01967, 2017. [Online]. Available: http://arxiv.org/abs/1708.01967

[7] “Understanding rnn and lstm,” https://towardsdatascience.com/understanding-rnn-and-lstm-f7cdf6dfc14e.

[8] T. Mitra and E. Gilbert, “Credbank: A large-scale social media corpus with associated credibility annotations,” in *ICWSM*, 2015.

VI. Appendix

All source code is in *https://github.com/shamirtowsif/DataMiningProject*.