Fake News Detection Analysis

Adithiyya Kishoore, Gaurav Madarkal

University of Colorado Boulder

Abstract. Fake news is a major issue in the current time, it has caused serious problems on a global scale in the recent past. This paper talks about some of the existing solutions, issues with such solutions and improvements that can be made for a better outcome.

Keywords: Fake News, Deep Learning, Text Summarization, Artificial Intelligence

1 Introduction

1.1 Motivation

What is fake news? It is the intentional and calculated presentation of false information or misinformation for the express purpose of misleading the reader. Even though this seems like a small issue at first, it can have catastrophic effects when deployed on a grand scale. One need not look further than the recent elections, where both parties were negatively affected by this, and more importantly the common man.

This became a major cause for concern with the advent of social media. When every single platform started to become a news provider, or rather fake news, to be accurate, with it's own set of sources, whose veracity cannot be objectively judged. As each and every news article cannot be checked for truth fullness on a large scale, the idea of coming up with an automated solution came into picture using the technique of deep learning and AI.

If we are to successfully overcome the predicament, the average consumer would be able to effectively make the right decisions that could lead to a better life for himself and his fellow neighbour.

1.2 Issues with Existing Solutions

The state of the art solutions that have been deployed and tested till now have all either categorized a news articles in a binary fashion. However, we cannot label news as fake or otherwise with little knowledge and context of the truth. And most of the fake news that gets proliferated has a semblance of truth to it, and requires common sense to judge correctly. As existing solutions are mostly keyword based, AI can easily be fooled. In fact, other AI have been deployed to generate fake news. This is the weakness in the existing approach.

1.3 Proposal

Our approach involves providing meta information to the algorithm along with the actual content of the document. This information could involve the page-rank of the website hosting the document along with the page-ranks of the websites cited in the article.

We would parse the articles cited in the document and compare the similitude between them. This would combat the issue of spurious websites citing popular articles whose content barely resembles the content in the article being judged.

Another weapon which we could employ in our arsenal is to compare the article with other popular articles that have been reported by verified publishers on the same topic.

The goal of this endeavour is to apply a confidence to the articles instead of taking a binary classification approach.

1.4 Deep-Dive of the Dataset

The dataset that is being primarily used is the *Fake and Real News Dataset* from Kaggle. It consists of 17903 unique articles.
The dataset columns are distributed into *title*, *text*, *subject*, and *date of publication.*

Some insights into the data appear to reveal that the character length of fake news articles are approximately twice in length to news that is true. While this does not lead to anything conclusive, it provides a possible metric that could be considered as one of our independent features.

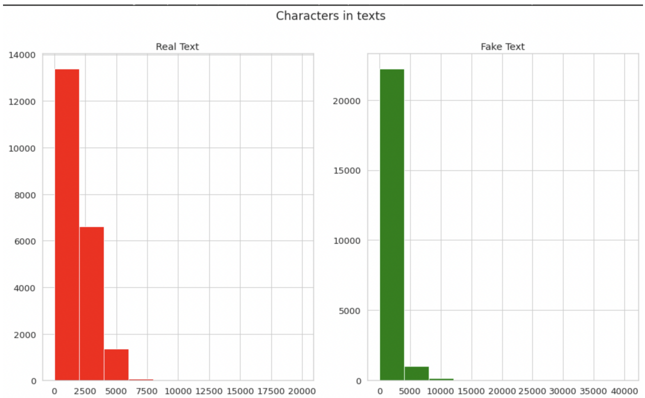

Fig. 1. Fake v Real Text - Article Lengths

2 Related Work

The core idea leveraged in this project comes from the paper *Attention is all you Need* which is one of the breakthrough papers detailing the method of self-attention to generate context while creating the embeddings that are passed to the deep-learning model.

Related works to solve this problem revolves around Linguistic approaches [1] which involve Data Representation, Deep Syntax, Semantic Analysis[2]. This project aims to leverage these approaches to better analyze the article under question. The aim here is to not only consider the current article but also involve the related articles and judge it based on what is happening in reality.

This topic seems very similar to the popular problem of "Spam or Ham"[3], but the major difference is that Spam detection does not require contextual information, the algorithm need not understand what is the truth in order to decide whether a mail is spam or not.

To keep users engaged and maintain retention rate a number of fake news writers use provocative language for their own nefarious purposes, under this assumption a preliminary analysis can be done on the initial content of the article to determine if a further analysis is required. The text summarizing[4] techniques can be used to generate a gist of the article and run an initial analysis.

For the purpose of this project a very popular Kaggle data set will be used [5], this data set is quite comprehensive in terms of the variety of the data and quantity for training. Most of the related work in this field has been around a binary classification as the end result, but this project aims to turn that into a confidence prediction algorithm, meaning, the end result would give an insight about how fake (evaluation metric) an article under question is.

3 Methods

In order to solve the issue of fake news, we have to breakdown the problem into its sub-parts.

These sub-parts include: initial analysis of the data-set, general methods of cleansing the data-set, and specific methods of cleaning to be employed based on the initial data-set analysis. Next, we would extract features which would involve selecting a tokenizer for our purposes and building the model.

If we are to choose a model or a tokenizer like BERT, we would also have to manipulate the input structure, as these models require input in a specific format.

An approach, which is available to solve this problem in a rudimentary manner is using LSTM, with embeddings that are readily available like GloVe. LSTM was a good approach to implement as text requires context to be judged and as LSTM captures the sequential information from the previous hidden layers, due to its RNN-like architecture.

However, as the input got larger, because of the typical length of a news article and also as parsing the text within the hyperlink in the document, the amount of time and memory required increased significantly.

Keeping in mind the above challenges of using LSTM, and the advantages of using self-attention, Transformer based approaches were chosen. The most popular among these were BERT, which also had a few variant models that were built on top of it like RoBERTa. Further exploration of these models and methods of tokenization revealed that RoBERTa used dynamic masking techniques where the masking is done during training unlike BERT where this is bounded, also RoBERTa was trained on a bigger batch size which improved perplexity on masked language modelling objective as well as end-task accuracy.

4 Experimental Design

The experiment consists of independent and dependant variables, with the independent variable being the text document and the dependant variable is the classification of the text into fake or real categories.

The experiment could be varied to include any combination of the peripheral independent columns like *title*, or *subject* along with the *text* to train the model.

The model will be trained with a variety of hyper-parameters and the best model will be chosen for further fine-tuning. Hyperparameters like loss functions (*Categorical_CrossEntropy*), activation functions (*Sigmoid, Relu*), Regularization Techniques (*DropOut, L1-Norm*), Optimizers (*SGD, Adam*), and the number of hidden layers (*about 3 to 5*).

The core layers that define the model would consists of Bi-Directional LSTM layers and in further models - transformer layers from BERT and RoBERTa.

All of the following above models would run on GPUs (which is the industry standard) on Azure which have been designed from the ground-up to accelerate the training process.

The result of a prediction would be a confidence score for both fake and real news, which would be a more practical measure of truthfulness for the end user, instead of a just a binary result. This would contain more information that the user could utilize to arrive at a better conclusion.

To analyse our implemented model which is entirely trained on past instances of fake or real news, the model will be tested on currently trending topics which are labelled as fake from trusted sources of media. An F1 score will be used as the target metric because it values both precision and recall with equal weightage.

References

1. Kelly Stahl, "Fake news detection in social media", 2018, Department of Mathematics and Department of Computer Sciences, California State University
2. M. I. Prabha and G. Umarani Srikanth, "Survey of Sentiment Analysis Using Deep Learning Techniques," 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), 2019, pp. 1-9, doi: 10.1109/ICIICT1.2019.8741438.
3. A. Karim, S. Azam, B. Shanmugam and K. Kannoorpatti, "Efficient Clustering of Emails Into Spam and Ham: The Foundational Study of a Comprehensive Unsupervised Framework," in IEEE Access, vol. 8, pp. 154759-154788, 2020, doi: 10.1109/ACCESS.2020.3017082.

4. N. S. Shirwandkar and S. Kulkarni, ”Extractive Text Summarization Using Deep Learning,” 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697465.
5. Fake News Dataset, Pankaj Gupta, https://www.kaggle.com/datasets/pnkjgpt/fake-news-dataset
6. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin, ”Attention Is All You Need”, 2017