Ezana N. Beyenne

MSDS 453, Section 57 2020

**Assignment 4**: **Fake News Prediction Using** **Neural Network Models**

**Abstract**

Fake news is news deliberately spreading disinformation or hoaxes via social media or traditional print and broadcast platforms. Fake news is a medium like clickbait stories, in that they are published with the intent to sensationalize dishonest outright fabricated headlines to damage a person, entity, an agency or gain financially or politically. Steve Job’s former advisor accused Facebook of failing to police misinformation because it keeps users coming back to the site (Vega, 2020). Fake news has been shown to affect people decisions and I wanted to see if we could develop a Neural Network to correctly distinguish between fake and legitimate news.

**Introduction**

Fake news, (also known as fabricated news), is typically found in traditional news, social media, or fake news websites, presenting news as factually accurate even though it has no basis in fact (Tufekci, 2018). Fabricated news has enormous popular appeal, are unsubstantiated stories, yet consumed by millions of people. Unfortunately, these fabrications are not only limited to politics but are also found in areas like vaccination, nutrition, and stock values. Claire Wardle identifies seven types of fake news (Wardle, 2017):  
1. *Satire or parody*: material with the potential to fool, but no intention to harm.  
2. *False connection*: material with visuals and/or headlines that don’t match the content.  
3. *Misleading content*: material intended to frame an issue or individual.

4. *False context*: material whose real content is shared with false contextual information.

5. *Impostor content*: material whose genuine sources are impersonated with false, made up stories.

6. *Manipulated content*: material with genuine information or images are doctored up to deceive.

7. *Fabricated content*: material is intended to deceive or harm with 100% fake content.

Information shapes our view of the world and allows us to make important decisions based on the information that we have. Information we receive allows us to form ideas about people or situations around us. When false, distorted, or exaggerated information is spread, we make bad decisions or form wrong opinion on such information. Fake news has led to bullying, violence against innocent people, racist ideas, fear-mongering, and it is now shown to have had a major impact on the last American presidential election (30secondes.org).

**Literature review**

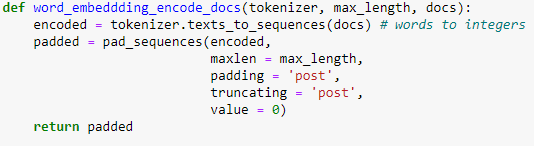
Research using machine learning was a tool anticipated to identify fake news and prevent them from going viral and spreading misinformation (Grothaus, 2019). However, machine learning has been manipulated to easily create fake news without any human intervention, and doing a poor job of identifying fake news (Grothaus, 2019). Research conducted by two MIT doctoral students found that computers could identify machine learning generated text, but they could not identify if that text was true or false. The reason is that machine learning algorithms could easily interpret positive statements, but could not interpret negative statements (Grothaus, 2019). The database used to train machine learning algorithms called Fact Extraction and Verification (FEVER) had some inherent biases when trying to identify fake news. For machine learning algorithms to correctly identify fake news, we must somehow weed out human bias and prejudices during its training phase (Grothaus, 2019). Kaggle also conducted a Fake News challenge, where competitors are tasked with developing a machine learning program to identify articles that might be unreliable fake news articles.

Fake news detection using various Natural Language Processing (NLP) and machine learning is still an active area of research with most of the focus being on social media platforms. The reason for this is that people have moved from traditional print and stand-alone websites to social media platforms (i.e. Twitter, Facebook, etc..) to consume news (B. Parikh, et al, 2020).

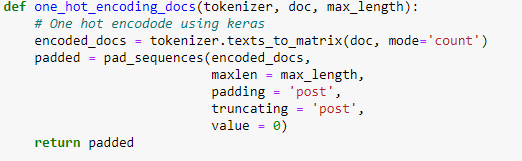
**Methods**

In this study, I used various neural network algorithms to see if I could correctly identify if an article was either fake news or not in the training dataset provided by the Fake News Kaggle challenge. I converted the data in train.csv into both word embedding and one hot encoding vectorization techniques to compare their effectiveness on the 1D CNN, LSTM, GRU and Bidirectional LSTM neural network methods. I then employed a tripartite splitting of the train.csv data to get a 75/15/10 percent train/dev/test split on around 18,250 rows of data after cleanup.

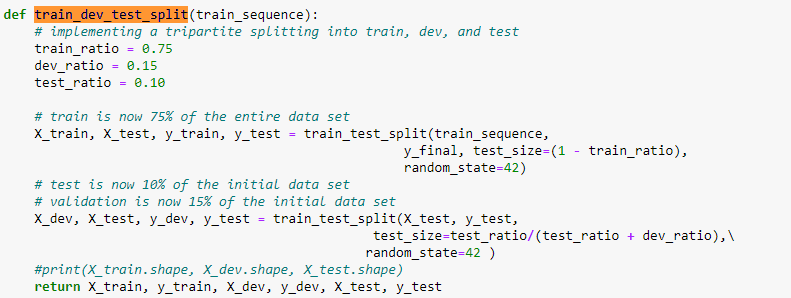
**Table 1:** Word Embedding code



**Table 2:** One hot embedding code



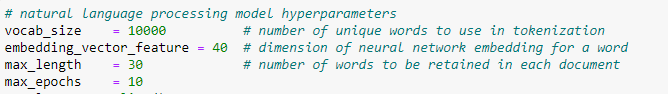
**Table 3**: Train dev test split code



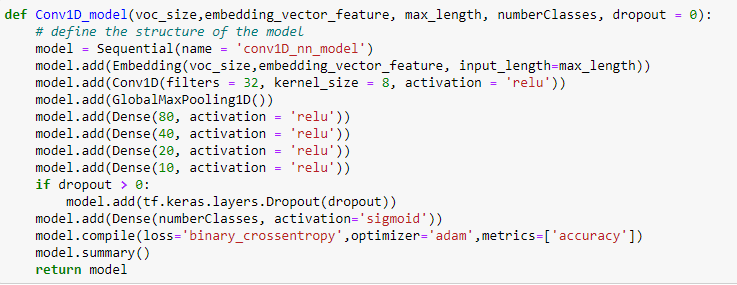
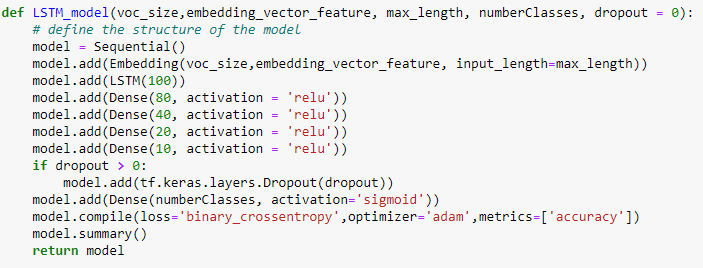
In addition to using a factorial design with the alternative methods mentioned above using both word embedding and one hot encoding vectorization techniques, I also added alternative settings to the neural network methods by adding dropout regularizations. The dropout regularization of 25%, and 50 % on the neural networks 1D CNN, LSTM, GRU and Bidirectional LSTM where conducted on both vectorization techniques.

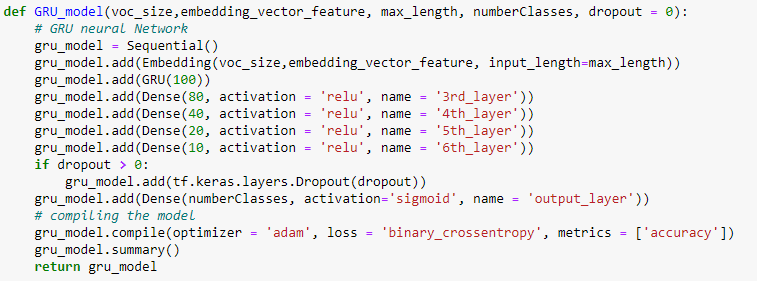
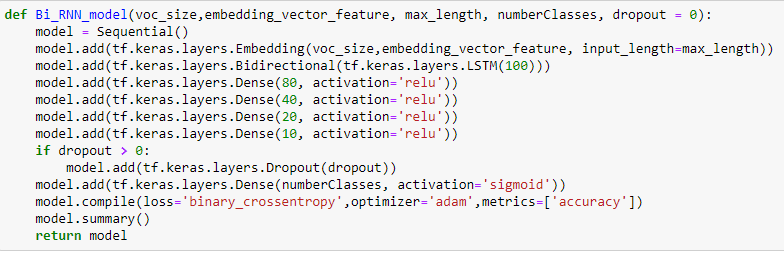
Early stopping was used to monitor the performance of the neural networks during the training phase in order to reduce over-fitting and improve generalization. I also was able to provide graphs of training and development set accuracy and loss, in addition to a graph of the area under the ROC curve. Since this was a binary classifcation problem ( it is fake news or legitimate news), I employed the sigmoid activation function in the output layer of the neural networks, with a loss of binary crossentropy.

I initially had an area under the curve score (AUC) of 0.5, but in order to resolve that I played around with the hyperparameters of shown in table 4 and also increasing the number of hidden layers in the neural networks. I also wanted to see what the ideal number of epochs were, and with the early stopping taking place, I settled on ten epochs. The GRU netural network with one hot encoding could have used more hidden layers, since it still had an AUC score around 0.5, but I hesitated because the GRU with word embedding had a higher score.

**Table 4:** Hyperparameters used in the neural networks  


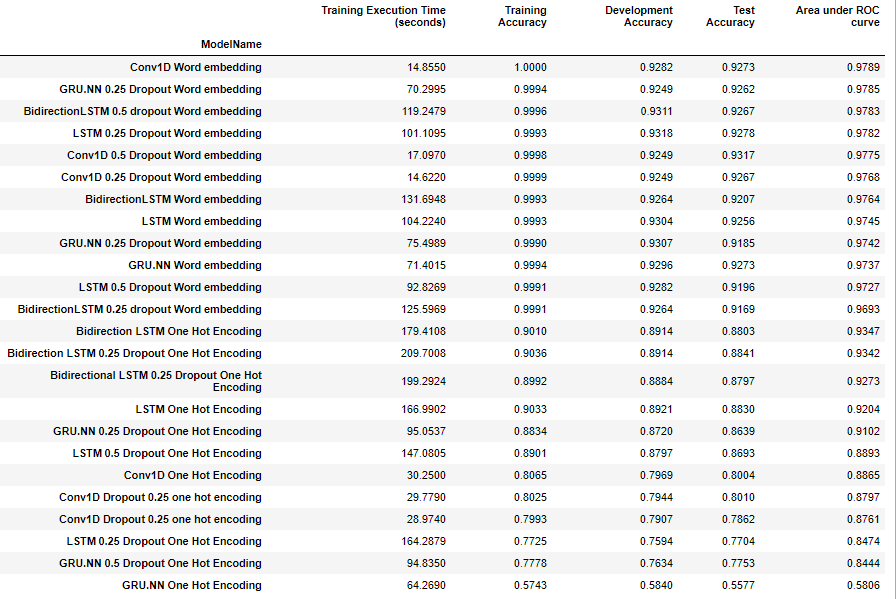
I used Keras Tensorflow 2.1.0 to build the four neural network functions with dropout being passed in as an optional parameter. The setup of the various neural networks is show in the tables below, this is after spending time trying to find the right combinations of layers and structures.

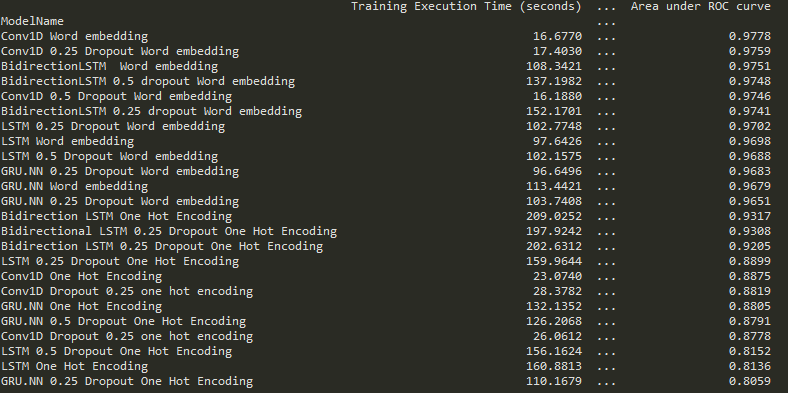
**Table 5:** Convolution 1D model  
  
**Table 6:** LSTM model  


**Table 7:** GRU model  
  
**Table 8:** Bidirectional LSTM  


**Results** The word vectorization technique that ended up with the best performance was the word embedding when compared to the one hot encoding. Additionally, from the results in Table 9, it looks like one hot encoding had the longest training execution times, but not the top test accuracy and area under the curve results. Surprisingly, the Conv1D word embedding model has the top area under the ROC curve and test accuracy scores. The results of the top neural network methods differ from the run using Jupyter notebook vs the Spyder IDE. It would be interesting to see after multiple runs, if the two editor’s results begin to converge. The pdf of the AUC and graphs are shown in the *Result\_Diagrams* folder and the results are stored in the *NN\_Results.csv* file.

**Table 9:** Results of all the models in Jupyter notebook



**Table 10:** Results of all the models in Spyder notebook  


**Conclusions**

The results of this neural network experiment do provide some really encouraging results in identifying fake or real news. It does show that word embedding along with dropout and early stopping can really create models that are good at identifying fake news, but we would have to have the social media platforms or new dissemination sites filter all the news articles through machine learning algorithms before displaying the data. In addition, they need to find a way to stop the viral spread of the fabricated news and allow the models to learn in real-time.

Fake news detection is a difficult classification technique since biases and prejudices are being introduced during training phases. Even with enough data, I do not think that we would be able to achieve high accuracy, because the language and techniques used to disseminate fabricated news are constantly evolving. It is a cat and mouse game, and the mouse seems to be winning due to their shear volume they produce.

**Works Cited**

Tufekci, Zeynep. (2018). It’s the (Democracy-Poisoning) Golden Age of Free Speech. Retrieved from <https://www.wired.com/story/free-speech-issue-tech-turmoil-new-censorship/>

Vega, Nicolas. (2020) Facebook “peddling in an addictive drug called anger”: Steve Jobs adviser. Retrieved from <https://nypost.com/2020/06/12/facebook-peddling-in-an-addictive-drug-called-anger-steve-jobs-advisor/>

Lazer, David M. J.; Baum, Matthew A.; Benkler, Yochai; Berinsky, Adam J.; Greenhill, Kelly M.; Menczer, Filippo; Metzger, Miriam J.; Nyhan, Brendan; Pennycook, Gordon. (March 9, 2018). “The science of fake news”.   
Retrieved from <https://science.sciencemag.org/content/359/6380/1094>

Wardle, Claire. (2017). Fake news. It’s complicated.

Retrieved from <https://firstdraftnews.org/latest/fake-news-complicated/>

30secondes.org. Impacts of Fake News.

Retrieved from <https://30secondes.org/en/module/impacts-of-fake-news/>

Grothaus, Michael. (2019). Machine learning isn’t effective at identifying fake news.

Retrieved from   
<https://www.fastcompany.com/90417625/machine-learning-isnt-effective-at-identifying-fake-news>

Kaggle. Fake-news. Retrieved from <https://www.kaggle.com/c/fake-news>

B. Parikh, V. Patil and P. K. Atrey, "On the Origin, Proliferation and Tone of Fake News," 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), San Jose, CA, USA, 2019, pp. 135-140, doi: 10.1109/MIPR.2019.00031.

Hamdi, Tarek. (2020). Top Research About Fake News Detection 2019. Retrieved from  
<https://www.kaggle.com/c/nlp-getting-started/discussion/123454>

**Folder Structure:**

1. **Data**: contains the train.csv
2. **Results\_Diagrams**: Contains the pdf diagrams of the AUC under ROC, the Training and dev set accuracy and loss.
3. **NN\_Results.csv**: contains the results from Table 9 as a csv file.
4. Jupyter notebook version of the code
5. Python version of the code

**Jupyter notebook code**   
