Summary of a Neural Net Application to Fake News Detection

Zach Quinn

Bellevue University**Problem Being Solved**

The focus of this research summary is the research paper FAGON: Fake News Detection Model Using Grammatical Transformation on Deep Neural Network, by researchers Youngkyung Seo, Seong-Soo Han, You-Boo Jeon, and Chang Sung-Jeong. In the course of three experiments, the researchers attempt a novel application of text-based and classification-based neural network architectures. The researchers attempted to solve the problem of the proliferation of Korean language fake news. Korean language news stories are harder to classify as real or fake news because of how its sentences are structured (Seo et al., 2019).

In English sentences can really only be structured in two ways and still be understandable, a construction prioritizing a sentence’s subject (active voice) or a structure prioritizing a sentence’s object (passive voice). In Korean, sentences can be structured in a multitude of ways and still retain the same meaning. The spread of misinformation is especially threatening in Korea because, given its geographic location. It is nestled between two hostile nations, China and North Korea, which have a history of seeking to undermine its comparably democratic society. Therefore, in order to avoid being vulnerable to disinformation campaigns by enemy states, it is vital for South Korea to develop a reliable means of detecting fake news in the Korean language.

**Plain-Language Explanation**

The researchers employed both a CNN (convolution neural network) classifier and LTSM (long short-term memory) network architecture and RNNGs (Recurrent Neural Network Grammars). To put it plainly, the module has two objectives. The first is to identify and break down similar sentences into a format that a computer model can easily intake. Then a generated model compares the individual components of the broken down sentence to other broken down sentences in the set of given data (which contains the texts of thousands of news stories). In order to properly train the model, the set of broken down sentences (our data set) will be split into two collections of broken down sentences. The researchers will write code to create a process that will compare the collections and then evaluate the results. Since most data is generated as integers (real numbers), values closest to zero will determine whether a story is true and values closest to one will determine if a story is false, which constitutes a process called labeling. If this process is successful, then the model can be applied to any number of data sets that contain Korean language news stories and, theoretically, yield the same results.

**Why a neural net?**

A big part of this particular experiment (and many data science experiments) is determining what process (or model) helps the computer best fulfill the task we assign it. In this experiment, the researchers compared several different approaches, including cosine similarity, a support vector machine and convolution neural network. While these methods sound complex the task they are performing is deceptively simple: Helping us classify whether or not a news story is objectively (based on what the machine has ‘learned’) fake or real. A neural net is ideal for any classification problem because the alternative to machine learning is reading every news story and manually determining whether or not it is fake news. Instead, a neural net allows us to ‘train’ a computer to instantly search through a large amount of data and, essentially, learn in a similar manner that a human would: By trial and error.

**Length and complexity of model training**

The training of the solution took three different experiments comparing the utility of SVM, cosine similarity, and CNN to determine which model represented the best choice as a classifier. The number of training data set values was 35,144, the number of validation data set values was 10,040, and the number of test data set attributes was 5,021 for this classifier experiment. The researchers evaluated the sentence generator model, classification model and conducted a final test for model accuracy.

**Comparative Results**

The first experiment was to compare a model both with and without grammatical transformation. Ultimately, the grammatical transformation model scored a higher training rate than non-grammatical transformation models with an accuracy rate of 0.72 out of 1. The second experiment sought to determine the best classifier. In the end the authors’ novel implementation FAGON (Fake News Detection Model Using Grammatical Transformation on Deep Neural Network) prevailed. At least in instances of Korean language articles, the authors’ novel classifier was more suitable for the experiment than established SVM, cosine similarity and CNN methodologies. The third experiment endeavored to determine the number of layers for CNN classification. The authors determined that incrementing the number of layers on the classification models results in less loss resulting in a model that is trained better.

**Questions for the authors:**

* How might you apply this model to classification problems dealing with English, which is a language spoken frequently in South Korea?
* What visualizations could you generate to better demonstrate the comparison between methodologies for a non-technical audience?
* How did you determine how to split your train/test data?
* What other classifiers were considered? For instance, did you consider a Passive Aggressive Classifier?
* What other metrics might reassure peers about the effectiveness of this classifier? Have you generated a confusion matrix or determined k-means/cross-validation scores?

**References**

Seo, Y., Han, S., Jeon, Y., Jeong, C. (2019). FAGON: Fake News Detection Model Using Grammatical Transformation on Deep Neural Network. *KS|| Transactions on Internet and Information Systems*, 13(10). p. 4958-4970.

URL (for instructor reference): <https://eds.b.ebscohost.com/eds/detail/detail?vid=1&sid=b0210934-8683-42b9-87d8-1fcc6dcbc8da%40pdc-v-sessmgr03&bdata=JnNpdGU9ZWRzLWxpdmU%3d#AN=139496512&db=iih>