**3 HAN: A Deep Neural Network for Fake News Detection**

In this article ,we had learned that the fake news can also be detected using a 3 level hierarchal attention network (3HAN). 3HAN has three levels, one each for words, sentences, and the headline, and constructs a news vector: an effective representation of an input news article, by processing an article in an hierarchical bottom-up manner. We will use the headline of the news article to be a distinguishing feature of fake news. 3HAN gives a differential importance to parts of an article, on account of its three layers of attention. We will use following models to create vectors :

* **Bag of words** - The Bag of Words (BoW) model is the simplest form of text representation in numbers.
* **TF-IDF -** Term frequency–inverse document frequency, is a numerical statistic that is intended to reflect how important a word is to a document in a collection
* **Bag of ngrams –** The bag of ngrams model is a sequence of N-words in a sentence. Here, N is an integer which stands for the number of words in the sequence.
* **Bigrams -** If you put N=2 in n-gram, then it is a bi-gram
* **SVM -** Support Vector Machine (SVM) is a supervised machine learning algorithm that can be used for both classification or regression challenges. However, it is mostly used in classification problems.

And for applying machine learning tasks on these models we will use following neural models:

* **GloVe** - GloVe is an unsupervised learning algorithm for obtaining vector representations for words
* **GloVe-Avg**
* **GRU** - A gated recurrent unit (GRU) is part of a specific model of recurrent neural network that intends to use connections through a sequence of nodes to perform machine learning tasks associated with memory and clustering, for instance, in speech recognition.
* **GRU-Avg**
* **HAN-Max** - Max pooling is a sample-based discretization process. The objective is to down-sample an input representation (image, hidden-layer output matrix, etc.), reducing its dimensionality and allowing for assumptions to be made about features contained in the sub-regions binned
* **HAN-Avg** -Average Pooling is a pooling operation that calculates the average value for patches of a feature map, and uses it to create a downsampled (pooled) feature map. It is usually used after a convolutional layer.

**Baseline of Code:**

It is divided into two main folders ie

* **Word Count Based Models:**
* Bag-of-words and its TF-IDF constructs a vocabulary of the most frequent 50,000 words. The count of these words is used as features. The TF-IDF count is used as features in the other model variant.
* Bag-of-ngrams and its TF-IDF uses the count of the 50,000 most frequent ngrams (n <= 5). The features are formed in the same way as above.
* SVM+Bigrams uses the count of the 50,000 most frequent bigrams as features with an SVM classifier.
* **Neural Network Models:**
* GloVe-Avg flattens the article text to a word level granularity as a sequence of words. The GloVe embeddings of all words are averaged to form the feature vector.
* GRU treats the article text as a sequence of words. A GRU with an annotation dimension of 300 is run on the sequence of GloVe word embeddings. The hidden annotation after the last time step is used as the feature vector.
* GRU-Avg runs a GRU on the sequence of word embeddings and returns all hidden annotations at each time step. The average of these hidden annotations is used as the feature vector.
* HAN and Variants include HAN-Ave, Han-Max and HAN. HAN uses a two level hierarchical attention network. HAN-Ave and Han-Max replaces the attention mechanism with average and max pooling for composition respectively. Since the code is not officially released we use our own implementation.

**Models used in Code :**

The models folder contains code for:

**3HAN-Ave:** Vectors are composed using global average pooling

**3HAN-Max:** Vectors are composed using global max pooling

**3HAN:** Our proposed model with attention mechanism

**3HAN+PT:** Variant of 3HAN to include pre-training