“SUMMER INTERNSHIP PROJECT REPORT”

FAKE NEWS DETECTION

Submitted By

ABHINOV PHUKON

ROLL NO. 172010007003

B.E COMPUTER SCIENCE

Academy Session- 2017-2021

Barak Valley Engineering College

Nirala, Karimganj - 788701, Assam, INDIA

September 2020

TABLE OF CONTENTS

ABSTRACT

Information preciseness on Internet is an increasingly important concern but web-scale data hampers the ability to identify, evaluate and correct such data, or so called "fake news". An important goal in improving the credibility of news is to identify the fake news timely. In this project, I propose to study the fake news detection problem. The objective of this project is to build a classifier that can predict whether a piece of news will be labeled as REAL or FAKE based only its content, thereby approaching the problem from a purely NLP perspective. For this purpose, a Convolutional Neural Network architecture is explored. Extensive experiments have been done previously on fake news detection which helped me compare my model with several state-of-the-art models such as bidirectional LSTM, RNN, etc., and the results of my model with 96% accuracy on the test dataset have demonstrated the effectiveness of the trained model.

INTRODUCTION

In recent years, fake news have been appearing in large numbers and widespread in the online world both in news media and social networking sites. It is typically generated for commercial interests to attract viewers and collect advertising revenue. However, people and groups with potentially malicious agendas have been known to initiate fake news in order to influence events and policies around the world. For instance, during the 2016 US president election, various kinds of fake news about the candidates widely spread through both official news media and the online social networks. Erroneous information was written to motivate the voters' irrational emotion and enthusiasm. Such kinds of fake news sometimes can bring about devastating effects, and an important goal in improving the credibility of these news is to identify the fake news timely.

While there exist tools and products to detect sources of fake news (e.g., whether a web site publishes misleading news), I have approached this problem as an instance of text classification. Text classification has a rich research history within the NLP community, and an equally impressive array of practical applications to showcase its importance.

METHODOLOGY ADOPTED

The dataset used consisted of 6335 news articles. It has title, text and ground-truth labels for FAKE and REAL news. I approached this problem as an instance of text classification, using only the content of the article as the source of features. Instead of the title, the content of the news article was used for classification from the dataset. With a thorough investigation of the given data, the most commonly occurring words in these fake news were determined and a vocabulary was created. These were then integer-encoded and fed into the model architecture during training. The model embeds the integer-encoded words into a vector space. Along with it hidden patterns or sequences of the words used in fake news are captured by the convolutional layers and the final output is determined by a sigmoid nonlinearity function. Thus the model is trained using the labels specified for each observation.

Reason behind this adopted strategy: Research shows that CNNs are very effective on several NLP tasks, for instance, semantic parsing, sentence modeling, and other traditional NLP tasks. From an NLP perspective, approaching this problem as an instance of text classification and using only the content of the article as the source of features, offers an interesting and valuable opportunity to identify patterns that can be coded in a classifier. Doing so afforded me to focus on Natural Language Processing related algorithms such as the CNN algorithm described above. Besides the words, there also exist some hidden patterns or sequences of the words used in fake news, which can be captured by the convolutional layers of my model.

The results of my model with 96% accuracy on the test dataset have demonstrated the effectiveness of the trained model.

IMPLEMENTATION

The dataset used consists of 6335 news articles. It has title, text and label (containing ground-truth labels for FAKE and REAL news) as columns. The data was read using the pandas library. Only the content of the "text" part of the dataset was used as features and the title of the news was left out. The data was then cleaned with various methods like removing punctuations and symbols, converting each word to lower case, etc. The dataset was then split into training, validation and test sets using the sklearn library.

The content of 'text' was then tokenized using Tokenizer module of the Tensorflow framework and thus a vocabulary with size 10000 consisting of the most common words was created along with an out-of-vocabulary token. Then the content was converted into padded sequences using the tokenizer instance. This data preparation step was done in order to integer-encode the words so that they may be used by the training model.

The Training and validation padded sequences were then assigned to train_x and validate_x inputs to be used while training the model. The labels were then assigned to the train_y and validate_y targets and also were first binary encoded.

The model architecture was then defined with a word embedding layer that takes input sequences of length the size of the defined vocabulary and embeds them into a vector space of size 16. Then convolution layer with ReLU units and max pooling layer were added, whose output is finally sent to a sigmoid nonlinearity to predict the label. The embedding is learned along with the model itself while training.

The loss function was set to ‘binary crossentrophy’ with optimization function Adam and a learning rate of 0.001. The model was then trained for 10 epochs with 99.98% accuracy on the training set and 95.48% validation accuracy.

CONCLUSION

The plots of the entire training history shows there was no significant overfitting. Then the test accuracy was calculated using the test set which came out to be 96.17%. The confusion matrix was then calculated with 10 false positive and 7 false negative out of 444 observations (articles).

In recent years, the explosive growth in fake news and its erosion to democracy, justice, and public trust has increased the demand for fake news detection and intervention. It is now viewed as one of the greatest threats to democracy, journalism, and freedom of expression.

Researchers have been using deep learning with the large amount of data available in social and news media to increase in intelligence of the models. Word embedding for extracting features or cues that distinguish relations between words in syntactic and semantic form have shown significant improvement in model training.

A complete, production-quality classifier will incorporate many different features beyond the vectors corresponding to the words in the text. For fake news detection, we can add as features the source of the news, including any associated URLs, the topic (e.g., science, politics, sports, etc.), publishing medium (blog, print, social media), country or geographic region of origin, publication year, as well as linguistic features not exploited in my model such as the use of capitalization, fraction of words that are proper nouns (using gazetteers), and others.