Relationship of media bias and public opinion on climate change

Liu Zhuolong

A0179678B

EE5001

Department of Electrical and Computer Engineering

The Calendar Year: 2018/2019

Abstract

It’s widely known that climate change, also referred to as global warming, is unambiguously happening. And, as the IPCC concluded in 2013, "It is extremely likely that human influence has been the dominant cause of the observed warming since the mid-20th century." However, debate about the causes and the necessity of government action has never stopped. Research shows that it’s related with the political. In this report, we collected and analyzed more than 5000 news articles from 9 newspaper agencies with different political leanings by applying data mining techniques, Natural Language Processing (NLP) techniques and Machine Learning (ML) algorithms. The result shows that political has limiting impact on this topic and most people in US believe that climate change is happening, and we need to tackle it.

Keywords*:* Climate Change; Machine Learning; Natural Language Processing; Data Mining

Declaration

I hereby declare that the project report entitled “Relationship of media bias and public opinion on climate change” submitted by me to Department of Electrical & Computer Engineering, National University of Singapore, of the requirement for the course EE5001 is carried out by me under the guidance of Prof. Jimmy Chih-Hsien Peng. I further declare that the work reported in this project has not been submitted in any other institute or university.

Acknowledgement

I would like to show my deepest gratitude to my supervisor, Prof. Jimmy Chih-Hsien Peng, a respectable, responsible and resourceful scholar, who has provided me with valuable guidance in every stage of the writing of this report. Without his enlightening instruction, impressive kindness and patience, I could not have completed my report. His keen and vigorous academic observation enlightens me not only in this project but also in my future study.

Contents

I. Introduction

The climate change, also referred to as global warming, is the observed century-scale rise in the average temperature of the Earth's climate system and its related effects [1]. Multiple lines of scientific evidence show that the climate system is warming [2][3][4]. Many of the observed changes since the 1950s are unprecedented in the instrumental temperature record, and in paleoclimate proxy records of climate change over thousands to millions of years [5]. It's a common sense that climate change is indeed happening, and it already resulted in many severe disasters like droughts, forest fires and heavy hurricanes.

As the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report concluded in 2013, "It is extremely likely that human influence has been the dominant cause of the observed warming since the mid-20th century." [6] While there has been a widespread consensus in many countries like China and Sweden that climate change is mainly caused by human being's activity and that there is an unquestionable need to take countermeasures, there has been substantial debate about the causes and the necessity of political action in the United States [7]. According to Gallup, a leading polling organization, the public opinion on climate in US has become highly partisan – in 2017 there are 66% Democrats worrying a 'great deal' about global warming or climate change, but only 18% Republicans held the same opinion [8].

To understand the political influence on the public opinion on climate change, this report looks at the opinion of mainstream newspaper agencies with different political leaning in US since it is reasonable to assume that public's opinion is (partly) effected by those newspapers. We collected over 5000 news articles from 9 newspapers with different political bias from January 1, 2016 to August 15, 2018 and analyzed the distribution of these articles. In this report, we divide articles into two categories – articles in one category convince that climate change is happening, it's caused by human activities and government should take action to tackle this problem, which referred to as 'supporting articles' or 'supporting frame', while articles in another category hold opposite opinion and they are referred to as 'counter articles' or 'counter frame'.

To classify the large number of articles, we applied Natural Language Processing (NLP) techniques and Machine Learning (ML) algorithms to train a classification model, which can be used to classify the category of new input articles (i.e. input text data). In this report, we use Long Short-Term Memory networks (LSTM) model to carry out this task, which is proposed by Sepp Hochreiter and Jürgen Schmidhuber [9].

The analysis result shows that 75.02% articles of liberal-bias (democrats) newspapers agencies belong to supporting frame, which did not surprise us. However, result also shows that the percentage of supporting articles in conservative-bias (republican) newspapers is 67.57%. It means that political leaning only has limited impact on mainstream medias. We also analyzed the distribution before and after June 1, 2017, when is the date of US quitting the Paris Agreement. We found that the supporting rate (i.e. percentage of supporting articles) of liberal-bias newspapers increased by 5.65% after this event. And 1.04% decline is observed in conservative-bias medias.

The rest of this report are organized as follows: In Chapter 2, we present literature review and methodology. In Chapter 3 we present results and discussions. And then we conclude this report in Chapter 4.

II. Literature Review and Methodology

This chapter presents literature review and methodology used in this report, mainly consists of three parts: Data Acquisition, Natural Language Processing (NLP) techniques and Machine Learning (ML) algorithms.

The flowchart outlining the major steps are shown in Figure 1. As the diagram illustrates, we first ‘scrape’ data using net spider from the Internet, and then do data pre-processing and word2vec by applying NLP techniques. By doing so, we convert raw data from text to numbers, which can be sent into the neural networks and fit the weights.

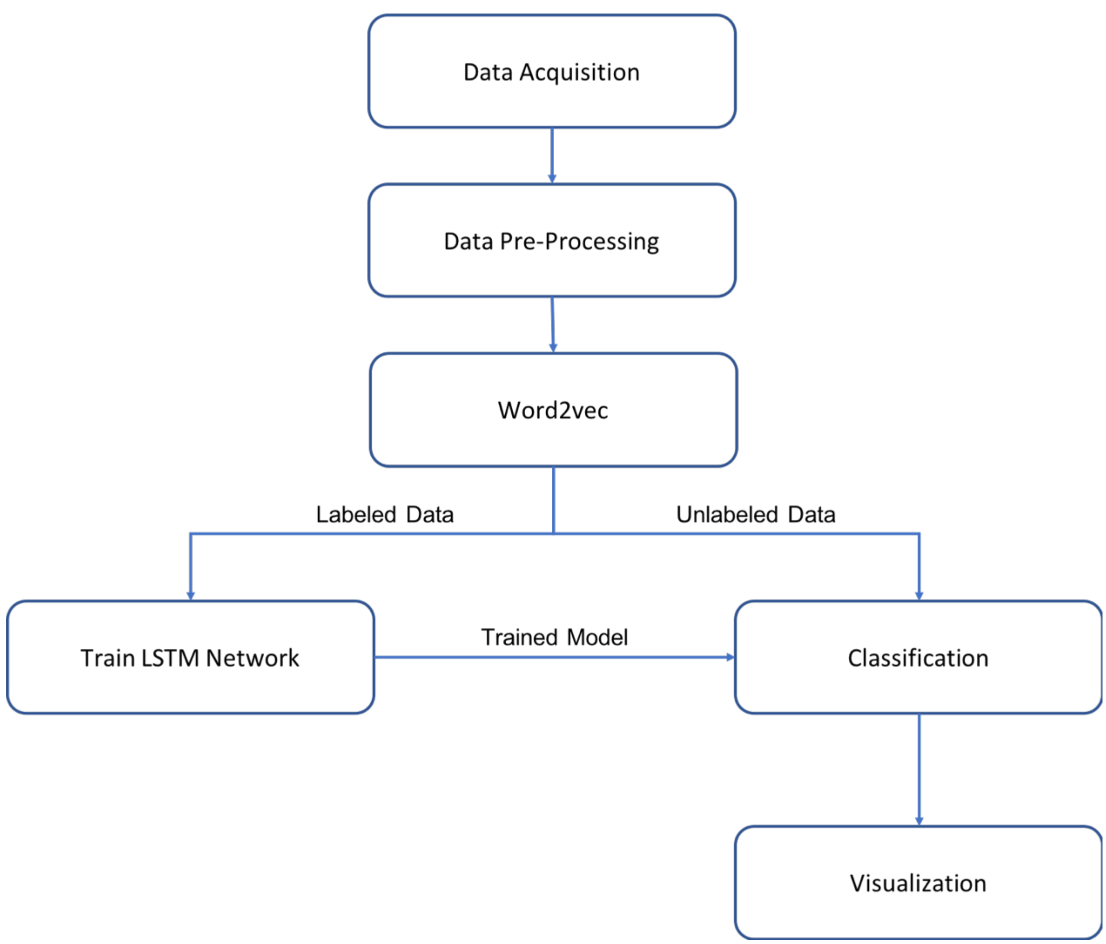


Figure 1. Flowchart of main steps

After transforming text to numbers, we train the Long Short-Term Memory (LSTM) networks – a special kind of RNN, capable of learning long-term dependencies [10]. Finally, we use trained LSTM networks to do classification and then visualize the results.

2.1 Data Acquisition

We use the Python toolkit named Scrapy, a widely used, free and open-source web-crawling framework, as our main tool to scrape data from the Internet. The architecture and data flow of Scrapy are presented in Figure 2. Tasks are firstly sent to the scheduler from spider and then transmit to the downloader according to Depth-First Order (DFO). After downloading webpages, the responses are sent back to the spiders to extract data. Then the item pipeline will filter duplicated items and save them in .json format.

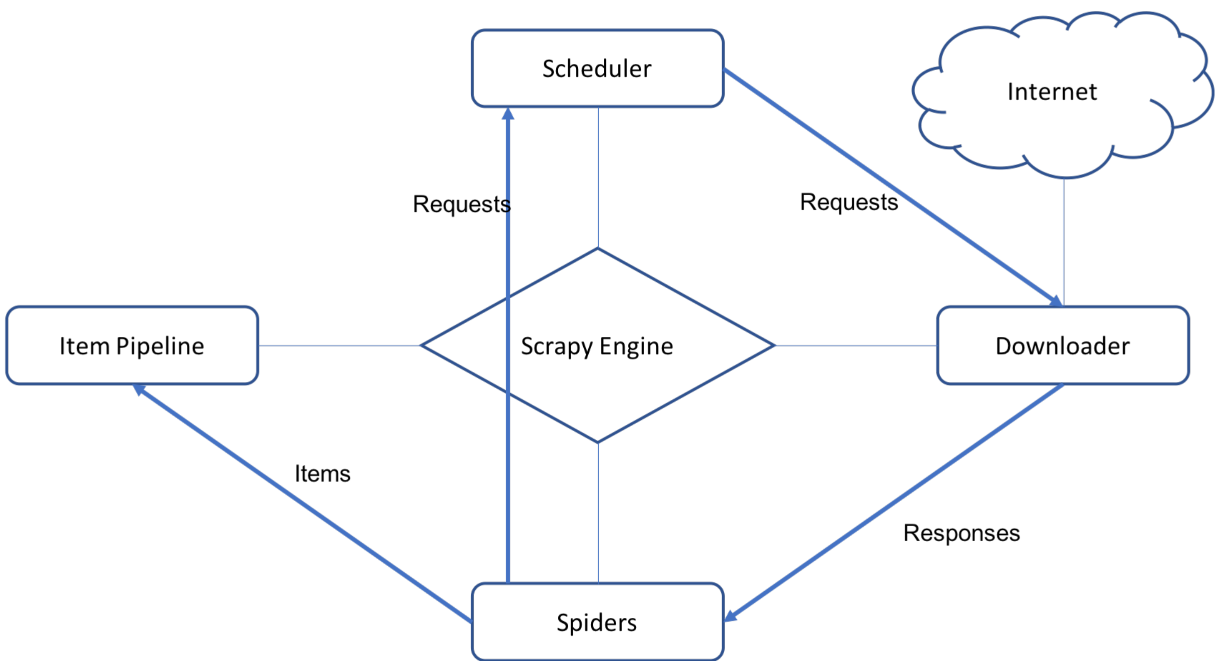


Figure 2. Architecture and data flow of Scrapy

Generally speaking, there are two types of web pages, static and dynamic (client-side dynamic). The static web pages are easy to scrape and parse. What Scrapy need to do is only downloading web pages and extract what we need according to HTML tags. However, lots of web pages are dynamic, whose contents change in response to user behaviors or at specified timing events. Client-side scripting languages like JavaScript, used for Dynamic HTML (DHTML), are frequently used to orchestrate media types (sound, animations, changing text, etc.) of the presentation. However, the Scrapy itself can only parse static web pages. We need another toolkit Splash, a Java rendering service, to co-operate with Scrapy to crawl both static and dynamic web pages.

Table 1. Sources of news articles

Sources	**Numbers (uncleaned)**	**Numbers (cleaned)**	**Political Bias**
CNN	2307	455	Liberal
USA Today	1824	159	Liberal
New York Times	4778	970	Liberal
Houston Chronicle	10657	2552	Liberal
Washington Post	1501	322	Liberal
Dallas Morning News	2079	204	Conservative
Florida Times-Union	778	185	Conservative
Fox News	2035	549	Conservative
Las Vegas Review-Journal	682	124	Conservative
Total	26641	5520	-

After this step, we got 26641 (uncleaned) news articles from January 1, 2016 to August 15, 2018 that matched keyword ‘climate change’ or ‘global warming’. The information of collected data is listed in Table 1.

2.2 Data Pre-processing

As we know, Neural Networks (NNs) can only process real numbers. However, the data we collected in 3.1 is text consists of English letters, spaces and punctuations, which cannot be recognized by NNs. Hence, we need apply some NLP techniques to convert those text to real numbers, which will be introduced in section 3.2.5. And before we do so, we should firstly pre-process the data, make it clean and friendly. In this section, we make use of a Python toolkit named NLTK to implement tokenization, stop words filtering and word normalization.

2.2.1 Data Clean

After observing raw text data, we found that many of them have nothing to do with ‘climate change’ or ‘global warming’, although these keywords appeared in those articles. The reason is that some news in other topics may mentioned ‘climate change’ or ‘global warming’ once or twice but did not really discuss on them.

To throw away those unwanted data, we searched keywords like ‘climate change’, ‘greenhouse gases’ and ‘sea level rise’ on headlines and the first 300 characters of each text. Then we only save those data who satisfied the requirement. After data clean, we now have 5520 cleaned data, which is presented in Table 1.

2.2.2 Data Labeling and Balancing

Training data is extremely important in machine learning because it is used for learning [11]. The training data must be balanced and representative so that the parameters of network could be fitted well enough to separate incoming data to expected categories.

Usually, in sentiment analysis applications, people use existing dataset like SemEvel or IMDB to train the model to predict human beings' emotion. However, peoples' opinion on climate change is much complicated and ambiguous than somebody like or dislike a movie. We had tried our best to search for existing dataset in this domain, but it seems like nobody has ever published a well labelled news dataset about people's opinion on climate change or global warming. Hence, we have to label the collected data manually. We spent a lot of time to do this and finally we got 356 labelled data with 282 belong to supporting frame and 72 belong to the counter frame. We know that the amount of data is far from enough, but it is the best we can do constrained by limiting time and resources.

Obviously, the labelled data is unbalanced (282 vs. 72). Unbalanced training data will result in definite bias when we train the networks. Although the unbalanced data might give us higher overall accuracy due to the test set and incoming data could also be highly unbalanced, it will still lead to lower accuracy when predict counter frame articles, which will result in inaccuracy and unreliable. In order to make our result more objective and credible, we chose to balance our labelled data by using oversampling techniques. In the meantime, we also want to 'create' more new data from existing data to increase the quantity of training data.

To solve this problem, we first randomly selected 16 samples from each category and build the test dataset to make sure the test set is independent with training set. Next, we split every remaining article into two parts and combined each part with other split articles in the same category. For instance, we split article A and B into A1, A2 and B1, B2, respectively then we combine A1 with B2 and A2 with B1. After doing this, we got 532 climate change frame data and 112 counter frame data. Then, we simply copy and paste the counter-frame data until we also have 532 counter frame data. At last, we have 1064 balanced training data.

2.2.3 Tokenization and Stop Words Filtering

Tokenization, or lexical analysis is the process of converting a sequence of characters (such as in a computer program or web page) into a sequence of tokens (strings with an assigned and thus identified meaning) [12]. For example, there is a sentence "The quick brown fox jumps over the lazy dog", after tokenization, it becomes a list of single words: ["The", "quick", "brown", "fox", "jumps", "over", "the", "lazy", "dog"], which is much easier to process.

The next step is to filter stop words in articles. Stop words are those words who have few effects on the training progress, such as 'is', 'am', 'I'. This step is actually reducing the dimension of dataset and speed up the training process.

2.2.4 Word Normalization

Word normalization is a process by which text is transformed in some way to make it consistent in a way which it might not have been before [13]. Generally speaking, word normalization is a technique that transform different words with the same meaning into the same format. For example, we want to match 'U.S.A', 'US', 'U.S'. and 'USA', so we apply word normalization and convert all of them into 'US'. Also, we want to match 'like', 'likes' and 'liking', so we transform them into 'lik', which might not exist in the dictionary. By doing this, we eliminate the ambiguous caused by synonyms and different formats of the same word.

Case folding

As its name tells, case folding reduces all letters to lower case. By doing case folding, for example, our program can identify ‘Fed’ and ‘fed’ as the same word. It is obviously meaningful. It’s also important to note that some words like ‘US’ should not be transferred into lower case ‘us’, since they have totally different meaning.

Stemming

In morphology, words consist of two parts, stems and affixes. Stems are the core meaning-bearing units, such as ‘stem’ in ‘stems’ and ‘mean’ in ‘meaning’. And the affixes are bits and pieces that adhere to stems often with grammatical functions [14], like ‘s’ and ‘ing’ in ‘stem’ and ‘meaning’, respectively.

Stemming is the process of reducing inflected (or sometimes derived) words to their word stem [14]. By stemming, we can significantly reduce the vocabulary size, which will be built to map words to numbers. Porter algorithm is one of the most common English stemmers, which was proposed by Porter Martin in [15]. It contains series of mapping rules from inflected words to their stem. Additionally, it’s not necessary for stemming to use original stem as the result. For instance, Porter’s algorithm will reduce ‘argue’, ‘argued’, ‘argues’ and ‘arguing’ to stem ‘argu’, but not ‘argue’.

2.2.5 Word2vec

Word2vec is a group of related models that are used to produce word embeddings, which is a technique for mapping words or phrases from the vocabulary to vectors of real numbers. These models are shallow, two-layer neural networks that are trained to reconstruct linguistic contexts of words [16].

Word2vec comprises continuous bag-of-word (CBOW) and skip-gram (SG). The embeddings are created via a two-layer neural network, where the model (CBOW or SG) corresponds to the prediction task. For CBOW, the objective is to predict a word given its context (here context refers to some window of words). For SG, the objective is to predict the context given the word. The final output of the model is only important for training reasons; once training has converged, the output vector is ignored. It is actually the encoding in the hidden layer that is used as the vector representation for the text [17].

This algorithm makes use of a binary classification objective (i.e. Logistic Regression) to discriminate the real target word w_t from k imaginary noise words $\widetilde{w}$. This progress is illustrated below in Figure 3 [17] CBOW model (For SG model the direction is simply inverted).

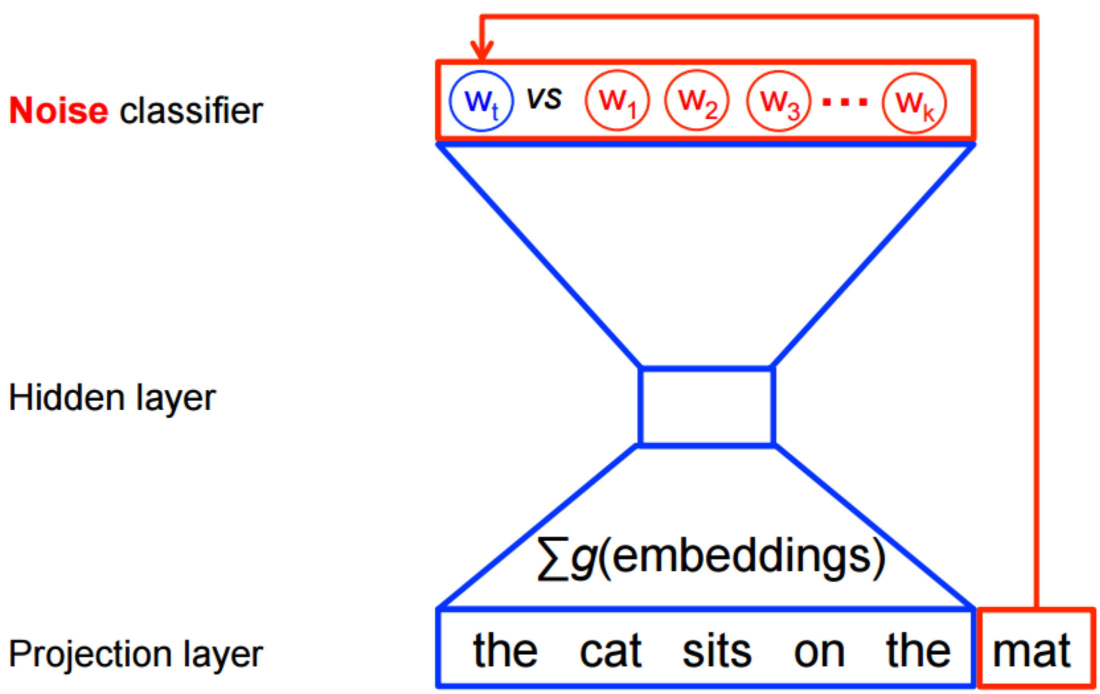


Figure 3. CBOW word2vec model [17]

Our objective is to maximize the Noise-Contrastive Estimation (NCE) loss:

$$J_{NEG} = logQ_\theta(D = 1|h, w_t) + k\mathbb{E}[\log Q_\theta(D = 0|h, \widetilde{w})]$$

Where $\widetilde{w} \sim P_{noise}$ and $Q_\theta(D|h, w)$ is the binary logistic regression probability of seeing context h appears before word w in the dataset D. And the outcome of the hidden layer is extracted and used as the vector representation for the text.

2.3 Model Training

In this section, we'll introduce two widely used neural network models, referred to as Multi-layer Perceptron (MLP) and Long Short-Term Memory networks (LSTM). All implementation in this section is done by using TensorFlow Keras, a powerful high-level API for building and training machine learning models.

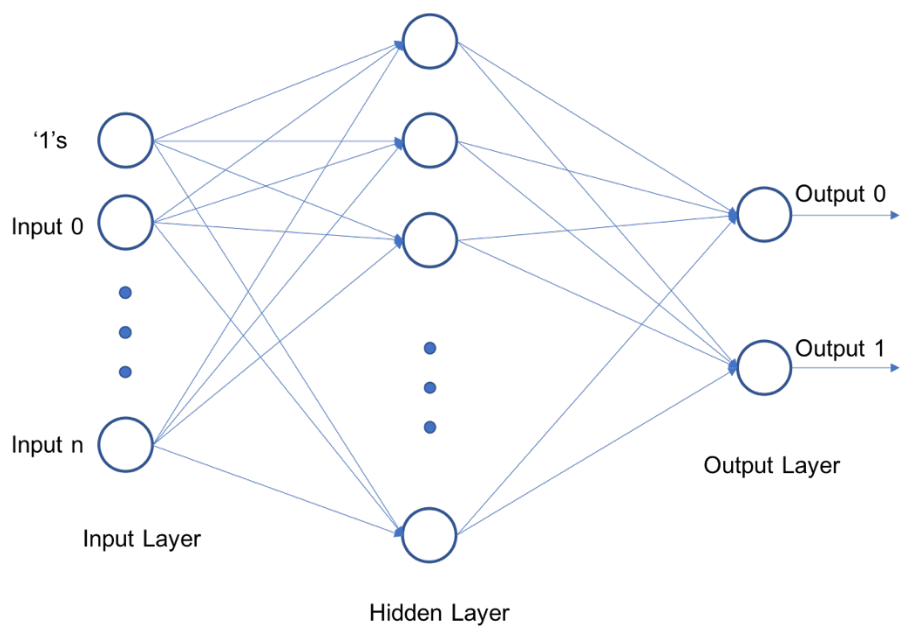


Figure 4 Structure of Multi-layer Perceptron

2.3.1 Multi-layer Perceptron

A Multi-layer Perceptron (MLP) is a class of feedforward artificial neural network. An MLP consists of, at least, three layers of nodes: an input layer, a hidden layer and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised learning technique called backpropagation for training [18] [19]. Its multiple layers and non-linear activation distinguish MLP from a linear perceptron. It can distinguish data that is not linearly separable [20].

The structure of MLP is shown in Figure 4, the output of each single node is given by $\phi(v)$:

$$\phi(v) = activation(v) = sigmoid(XW + b)$$

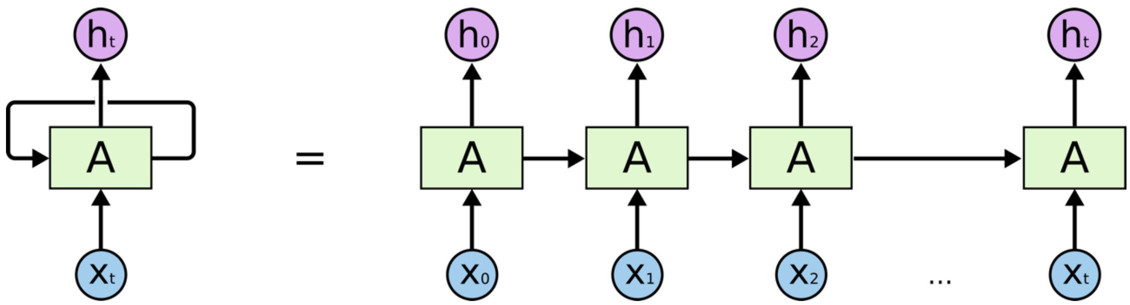
Where the XW is matrix product of input data ($1 \times n$ row vector, n is the number of inputs) and weights ($n \times 1$ column vector), b is the bias. The final prediction result will be the node number of largest outcome in output layer, for example if $output_1 = 0.8$ and $output_0 < 0.8$, then the input article will be labelled as label '1', i.e. supporting article.

2.3.2 LSTM Networks

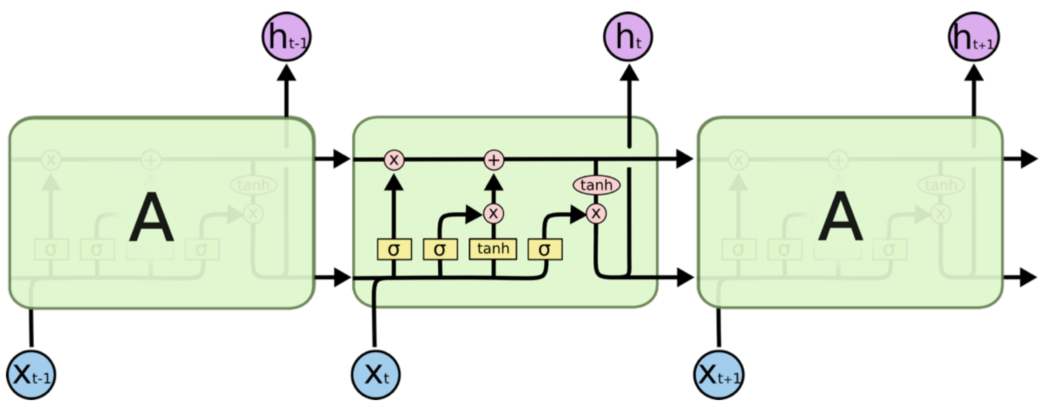
Long Short-Term Memory networks – usually just called "LSTMs" – are a special kind of Recurrent Neural Network (RNN), capable of learning long-term dependencies [10]. The structure of LSTM is presented as below in Figure 5 [10].

As Figure 5(a) shows, RNN is a loop-like or equivalent china-like network, which looks at the input x_t and outputs h_t and send some information back to the chunk to make information to persist. Hence, the RNN is an ideal structure to handle continuous data like videos and articles since information from previous time unit can be passed to the next time unit, while traditional neural networks cannot do this. In Figure 5(b) is the inside structure for LSTM, after training and fitting, those sigmoid and tanh functions can extract useful information and throw out unwanted part of data and then make precise classification.

In this report, we try both MLP and LSTM and got 63.51% and 71.25% accuracy, respectively. As expected, the LSTM performs better since our input data is news articles, which are continuous. Hence, we choose LSTM as our classification model and all the results presented in this report will base on LSTM.



(a) Structure of Recurrent Neural Network



(b) The repeating module in LSTM

Figure 5. Structure of Long Short-Term Memory networks [10]

III. Results and Discussion

This chapter presents the classification results. The discussion will focus on the overall distribution of supporting rates of each (liberal or conservative) bias. From the overall distribution we can analyze the impact of political leaning on the public opinion on climate change or global warming.

Besides, we also try to find the influence of the significant event – US quitted the Paris Agreement in June 1, 2017, which has far-reaching implications on the field of climate change.

To get better result, we ran each trial 10 times and use the average value as the result. All results present in this section is the average result of 10-time trials.

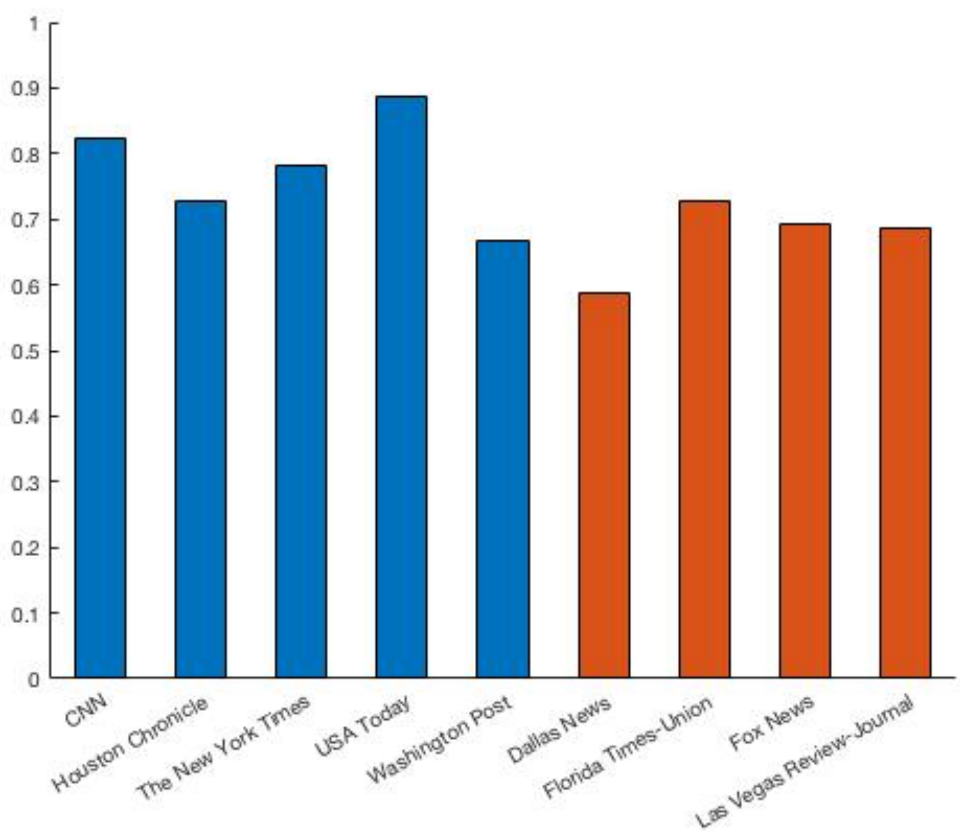


Figure 6. Distribution of supporting rate for each news agencies

3.1 Overall Analysis and Discussion

Figure 6 and Table 2 shows the positive percentage of each news source, where blue bars are news sources with liberal-leaning and red bars are those who have conservative-wing views. From this bar graph we can know that, although those conservative-wing biased news sources have slightly lower supporting rate, the lowest one, the Dallas News, still has 58.74% articles said that climate change is happening, and we should try our best to fight it.

Table 2. Experimental results of articles' attitude towards climate change

	Total Number	**Supporting Number**	**Supporting rate**	**Political Leaning**
Washington Post	291	194.5	66.84%	Liberal
CNN	436	358.4	82.2%	Liberal
USA Todays	156	138.4	88.72%	Liberal
New York Times	928	725.2	78.15%	Liberal
Houston Chronicle	2468	1793.5	72.67%	Liberal
Liberal	4279	3210	**75.02%**	Liberal
Fox News	495	343.3	69.35%	Conservative
Dallas Morning News	182	106.9	58.74%	Conservative
Las Vegas Review-Journal	89	61.1	68.65%	Conservative
Florida Times-Union	119	86.7	72.86%	Conservative
Conservative	885	598	**67.57%**	Conservative
Total	5164	3808	73.74%	–

In another way, even liberal-bias news agents show higher supporting rate of climate change, there still exist some news sources which have relatively low supporting rate. For instance, Washington Post has only 66.84% percentage of articles belong to climate change frame.

We also plot the average supporting rate of liberal-bias agents and conservative-bias agents, which is shown in Figure 7. The liberal-leaning news agents give us 75.02% supporting rate, which is 7.45% higher than the supporting rate of conservative-leaning news agents (67.57%).

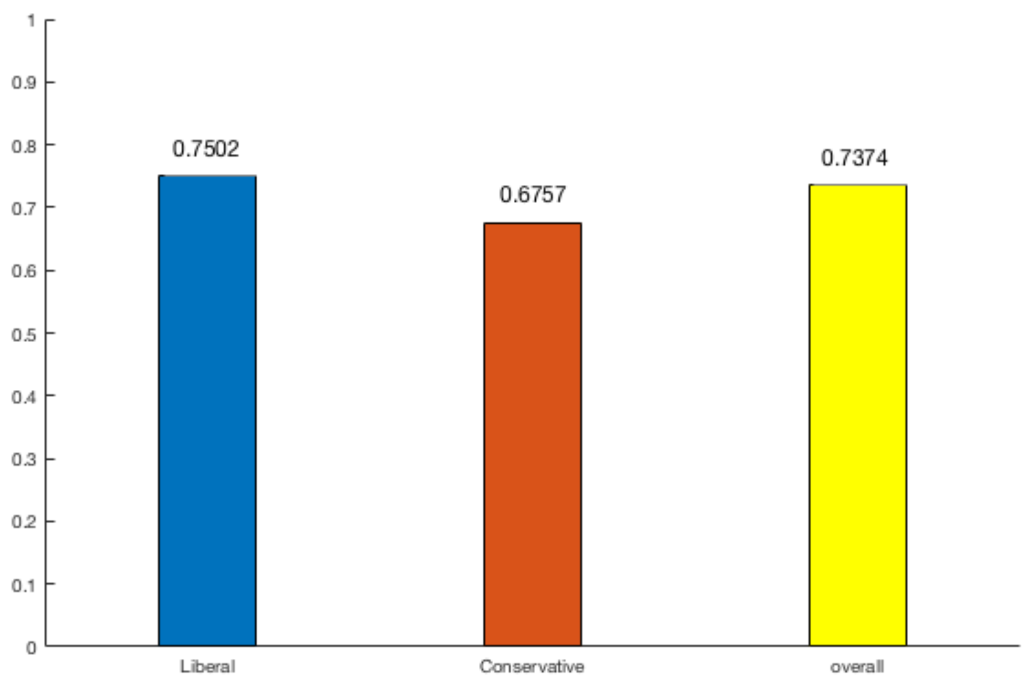


Figure 7. Supporting rate of each bias

It's reasonable to say that acknowledging climate change is a common sense among U.S. news sources. All the news sources show more than 50% overall supporting rate from January 2016 to August 2018, and the average supporting rate of all these agents is 73.74%, which is given in Table 2. This truth tells us most people in U.S. trust in scientists' research and clearly know and acknowledge that global warming is happening and it's threating human beings, even though Trump already make U.S. quit the Paris Agreement and said that climate change is a 'hoax'.

By particular, Florida Times-Union and Las Vegas Review Journal made endorsement of Trump in the 2016 United States Presidential election. However, they didn't show their supporting at least in this climate change topic.

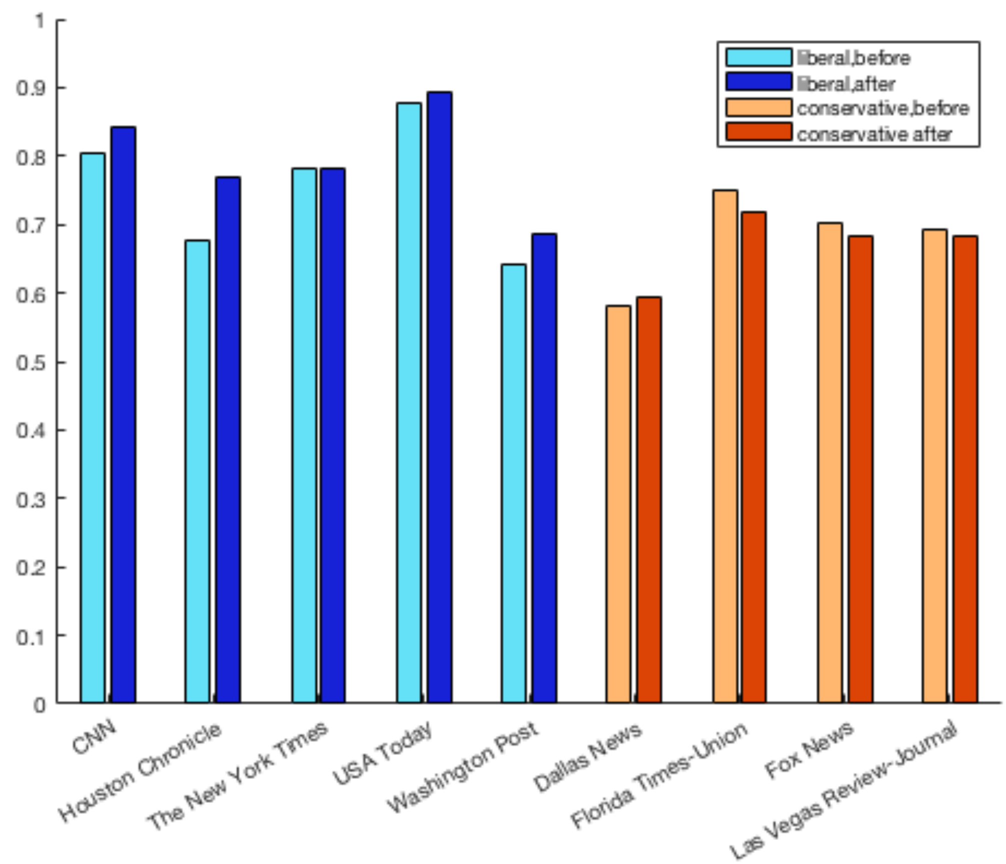


Figure 8. Supporting rate of each agencies before/after US quitting the Paris Agreement

3.2 Impact of quitting Paris accord

In this section, we will concentrate on the influence of quitting Paris accord, which is an international significant event made by Trump and his cabinets. Affected by this event, the eyes of the world were on the climate change during the whole June 2017. Hence, it makes sense for us to analyze the impact on the public opinions on climate change. Figure 8 shows the supporting rate of each agents before and after U.S. quitting the Paris accord.

Table 3. Experimental data of impact of quitting Paris Agreement

	Total Number	**Supporting Number**	**Supporting rate**	**Political Leaning**
Washington Post	118	75.7	64.15%	Liberal
	173	118.8	68.67%	
CNN	236	189.9	80.47%	Liberal
	200	168.5	84.25%	
USA Todays	64	56.2	87.81%	Liberal
	92	82.2	89.35%	
New York Times	456	355.9	78.05%	Liberal
	472	369.3	78.24%	
Houston Chronicle	1120	758	67.68%	Liberal
	1348	1035.5	76.82%	
Liberal	1994	1435.7	72.00%	Liberal
	2285	1774.3	77.65%	
Fox News	285	200.1	70.21%	Conservative
	210	143.2	68.19%	
Dallas Morning News	90	52.3	58.11%	Conservative
	92	54.6	59.35%	
Las Vegas Review-Journal	39	27	69.23%	Conservative
	50	34.1	68.20%	

Florida Times-Union	35	26.3	75.14	Conservative
	84	60.4	71.90%	
Conservative	449	305.7	68.08%	Conservative
	436	292.3	67.04%	
Total	2443	1741.4	71.28%	–
	2721	2066.6	75.95%	

From the diagram and Table 3 we can find that for every liberal-bias news agents, the supporting rates after U.S. quitting the Paris accord are higher than before. In which, the New York Times has lowest growth from 78.05% to 78.24% and Houston Chronicle give us the highest growth from 67.68% to 76.82. For the conservative-bias agents, all agents show downtrend except the Dallas New, whose supporting rate went up to 59.35% from 58.11%.

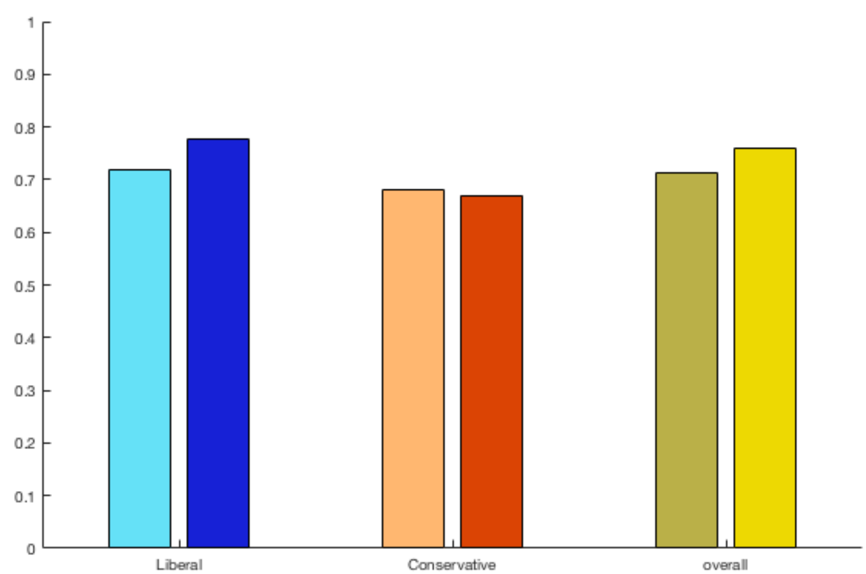


Figure 9. Supporting rate of each bias before/after US quitting the Paris Agreement

(left bar is before and right is after)

The overall trends are shown in Figure 9. As we observed in Figure 8, the supporting rate of liberal-bias news agents has uptrend, which rise from 72% to 77.65%. And the supporting rates of conservative-bias agents went down from 68.08% to 67.04%.

In conclusion, under the impact of quitting Paris Agreement, liberal-bias news agents show uptrend with 5.65% and conservative-bias show downtrend of 1.04%. The fact is that the overall trend goes up due to this event. It's reasonable to say that, although Trump and his cabinets consider climate change is not a threat to human beings, most publics and newspaper agencies still convince that government need to take action on climate change.

IV. Conclusion

This report makes use of net spider techniques to scrape news articles from the Internet. And then apply Natural Language Processing techniques and Machine Learning algorithms to analyze the text data in order to find out the relationship between media bias and public opinion on climate change.

The only assumption we made is that people will read news and can be influenced by them. The results show that political has impact on the attitude of public towards climate change, but the influence is faint. We also conclude that quitting the Paris Agreement increased the overall supporting rate of these newspaper agencies.

There still remains some problems to be solved. In the future, we want to collect more data, especially well labelled data, to train a better neural network model. Hence, the classification can be more precise and maybe we can find more patterns behind those news articles.

Reference

[1] Shaftel, Holly (January 2016). "What's in a name? Weather, global warming and climate change". NASA Climate Change: Vital Signs of the Planet.

[2] Hartmann, Dennis L.; Klein Tank, Albert M. G.; Rusticucci, Matilde (2013). "2: Observations: Atmosphere and Surface". IPCC WGI AR5 (Report). p. 198.

[3] "Myths vs. Facts: Denial of Petitions for Reconsideration of the Endangerment and Cause or Contribute Findings for Greenhouse Gases under Section 202(a) of the Clean Air Act". U.S. Environmental Protection Agency.

[4] "Climate change evidence: How do we know?". Climate Change: Vital Signs of the Planet.

[5] "IPCC, Climate Change 2013: The Physical Science Basis – Summary for Policymakers (AR5 WG1)". Intergovernmental Panel on Climate Change. p. 4.

[6] "IPCC, Climate Change 2013: The Physical Science Basis – Summary for Policymakers (AR5 WG1)". Intergovernmental Panel on Climate Change. p. 17.

[7] Adam Shehata & David Nicolas Hopmann (2012) FRAMING CLIMATE CHANGE, Journalism Studies, 13:2, 175-192, DOI: 10.1080/1461670X.2011.646396

[8] Roger Pielke Jr. Climate Change as Symbolic Politics in the United States, IEEJ Energy Journal Special Issue October 2017

[9] Sepp Hochreiter; Jürgen Schmidhuber (1997). "Long short-term memory". Neural Computation. 9 (8): 1735–1780. doi:10.1162/neco.1997.9.8.1735. PMID 9377276.

[10] Christopher Olah, Understanding LSTM Networks, 27 August, 2015
https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[11] M. McCandless, E. Hatcher, and O. Gespodnetic. Lucene in Action, Second Edition. Manning, 2010.

[12] "Anatomy of a Compiler and The Tokenizer". www.cs.man.ac.uk.

[13] Richard Sproat and Steven Bedrick (September 2011). "CS506/606: Txt Nrmlztn"

[14] Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition. Daniel Jurafsky & James H. Martin. p. 3-4

[15] Porter, Martin F. (1980); An Algorithm for Suffix Stripping, Program, 14(3): 130–137

[16] Mikolov, Tomas; et al. "Efficient Estimation of Word Representations in Vector Space". arXiv:1301.3781.

[17] TensorFlow, Vector Representations of Words,
https://www.tensorflow.org/tutorials/representation/word2vec

[18] Rosenblatt, Frank. x. Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Spartan Books, Washington DC, 1961

[19] Rumelhart, David E., Geoffrey E. Hinton, and R. J. Williams. "Learning Internal Representations by Error Propagation". David E. Rumelhart, James L. McClelland, and the PDP research group. (editors), Parallel distributed processing: Explorations in the microstructure of cognition, Volume 1: Foundation. MIT Press, 1986.

[20] Cybenko, G. 1989. Approximation by superpositions of a sigmoidal function Mathematics of Control, Signals, and Systems, 2(4), 303–314.