ReAcclimate: The new Climate Change Lexicon

Rui Aguiar(raguiar2),
Anthony Carrington(acarring),
Harold Wang(haroldw)
Stanford University

Abstract

To tackle the issue of addressing climate change, our team developed a novel model that predicts the best climate change words to use to obtain the highest engagement improvement. Our data set consisted of over 2M tweets about climate change related topics, with a tweet's engagement measured by a function of its number of retweets and likes. Since we wanted to generate cross-platform engagement, we first used K-Means to separate all of the users in our data-set into groups. We considered these groups separately and trained different networks for each in our model. We then utilized a LSTM neural network to predict a tweet's engagement by its input text and features about the user. We then used this model to make substitutions to see which alternative words maximized the engagement delta.

1 Introduction

Climate change is a generation-defining issue with many devastating implications for the future. However, despite scientists' longtime warnings about the dangers of climate change, many politicians and Americans have not been receptive to any sort of productive discussion on the subject. Our project hopes to help solve this problem by discovering the effective climate-change language that drives the most engagement in a cross-group setting.

Subtle differences in expression can have a profound impact on listeners' sentiment of a conversation. A lexicon to guide our communication can reduce friction and help to propagate ideas to a larger population base. Climate change researchers, activists, and legislators need this guideline to spread the awareness on this life-threatening issue. In the 1980s, the Republican party hired a team of top linguists to define a lexicon of terms that would drive engagement from both parties and swing votes towards their viewpoint. ReAcclimate is an application project that aims to similarly construct a climate change lexicon by using ML techniques to find the most engaging climate change words.

Since climate change is such a divisive issue today, we first wanted to separated users by their average climate change sentiment to see if language and engagement behave any differently among the groups, since ultimately our goal was to find the best climate words for everyone, including naysayers. In order to do this, we grouped using K-Means clustering, and trained a LSTM neural network for each group to predict engagement.

After implementing our prediction model, we were then able to make language substitutions within input tweets to see which alternative words and phrases invoked the most positive change in engagement. To be explicit, our research had two phases. In the first phase, we took in as input all of a users tweets and then used K-Means clustering to output labels that assigned them to one of two groups. In the second phase, we trained two neural networks (one for each group). Summed up: the input is all tweets from users in a group, the network is an RNN with LSTM cells, and the output is the predicted engagement of a tweet.

CS230: Deep Learning, Fall 2018, Stanford University, CA.

2 Related work

Much prior research has been done trying to understand the effects of climate change on human sentiment. Emily M. Cody [1] analyzed tweets from 2008 to 2014 to determine the correlation between climate change focused tweets and user happiness. The paper concluded that negative events such as natural disasters typically decrease user happiness while positive events like climate rallies prompts positive sentiments. Others like Xiaoran An [2] have completed similar studies and found a strong correlation between major climate events and climate change sentiment, with peoples' sentiment strengthening in bouts of extreme weather. We utilized these researchers' work to motivate our interest in considering users' climate change sentiment when evaluating their engagement. Since climate change seems to have distinct and polarizing effects on people's sentiment, we hypothesized that the best language that should be used to address climate change would vary based on a users' overall climate change sentiment. We therefore ran two models on two separate user groups that were K-means clustered by their average sentiment.

Additionally, prior studies have shown the strength of using LSTM neural networks for NLP tasks like text classification. In a sentiment classification study of 1.6M tweets, Wang. et. al displayed that the LSTM network outperformed other classifiers like SVM and Naive Bayes due to its ability to capture word order, and long term dependencies [7]. Moreover, in a hashtag recommendation study, LSTMs achieved higher accuracy over the Random Forest model for their ability to encode word features that also contain the semantics of sentences and sentence relations [8]. A related deep-learning project on maximizing twitter impact, measured by the number of retweets, showed a top accuracy of 61% for predicting engagement using RNN with LSTM cells [6]. Although this was better than human performance, their model suffered from over-fitting issues. We sought to improve upon their model by increasing the size of our dataset of tweets, because we believe the primary weakness of their model was predicting per account based on 3,000 total tweets, which didn't generalize.

3 Data set and Features

We used Tweepy [3], a Python Twitter API wrapper, and tweet IDs from George Washington University researchers [5] to collect over 2M climate change tweets created between September 21, 2017 and May 17, 2019. They were found using the following keywords and hashtags: climate change, global warming, climate hoax, #climatechange, #climatechangeisreal, #actonclimate, #globalwarming, #climatechangehoax, #climatedeniers, #climatechangeisfalse, #globalwarminghoax, #climatechangenotreal.

We preprocessed tweets by removing emojis, URLs, spaces, and stop words. In addition, all texts are stored in lower case and vectroized using the pre-trained GloVe matrix. We optimized the vector size using the validation set and the final embedding size is shown in Table 2.

Table 1: Data Features

Attribute	**Type**	**Note**	**Example**
userId	String		SenWarren
followerCnt	int	input	5,051,195
text	String	input	we want to live in a world with clean air ...
favoriteCnt	int	Label (engagement)	20,053
retweetCnt	int	Label (engagement)	3,226
sentiment	Float	K-Mean Cluster	0.74

We divide the dataset into train (80%), validation (10%), and test set (10%). There are 1,765,288 training examples, 220,661 test and validation examples. Each data point has 6 features as listed in the table above.

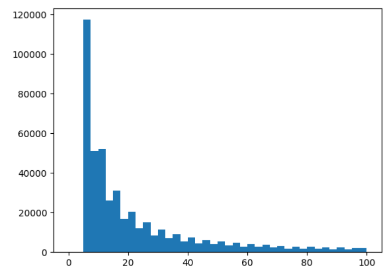
3.1 Engagement Score

We used the number of favorites and retweets to define a single evaluation metric, the engagement score, to capture user responses. Retweets have a weight ratio of 3 to reflect the fact that retweets convey a greater liking than favorites. The engagement score definition is shown below:

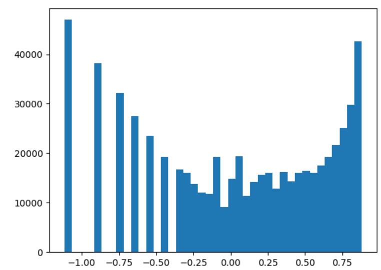
$$E(N_{favorites}, N_{retweet}) = N_{favorites} + 3 * N_{retweet} \tag{1}$$

The engagement score has a power-log distribution where most engagement is 0. This behavior is expected since most tweets have no favorites nor retweets. In response, we applied the Box-Cox transformation to adjust the distribution so it is more uniformly distributed between -1 and 1.

$$E(N_{favorites}, N_{retweet}) = \frac{-1}{0.6}((N_{favorites} + 3 * N_{retweet})^{-0.6} - 1) \tag{2}$$



(a) Original Engagement Score Distribution



(b) Transformed Score Distribution

4 Models and Analysis

4.1 K-Means Clustering

We first used K-Means to cluster users who made climate-change related tweets into groups based on the average sentiment of all of their climate change tweets. We did this based on our understanding that sentiment about climate change varies tremendously, so trying to capture the best climate change language for different user groups was important to us. After trying out different K values to understand how many inherent sentiment clusters exist in our dataset, we found that a K value of 2 provides the best distortion value and allows us to have a diverse set of users in each group. In summary, our algorithm first generates the average sentiment for user j across all their climate change tweets. We then initialize centroids μ_1, μ_2 to zero, and repeat the k-means algorithm: $c^{(i)} := argmin_j \|x^{(i)} - \mu_j\|^2$, then $\mu_j = \sum_{i=1}^{n} 1\{c^i = j\}x^i / \sum_{i=1}^{n} 1\{c^i = j\}$ until convergence.

4.2 Engagement Prediction Model

The engagement model predicts the engagement score by taking in the tweet text with relevant user information. Given the nature of our dataset, we selected the user's follower count since it has an affect on engagement, regardless of language.

As mentioned previously, we leveraged the GloVe embedding matrix to vectorize the text and feed the vector into a two layer LSTM network. An LSTM network is a recurrent neural network that can capture information effectively from sequential inputs. An LSTM cell consists of input, output, and forget gates that learn to focus on important words within a text body and embed salient features into the latent space. The engagement prediction model feeds LSTM outputs into a fully connected network. The fully connected network concatenates latent features with the user follower count to predict the user engagement score. The model architecture is shown in the figure below:

We selected mean-square-error $1/n \sum_{i=1}^{n} (\hat{y}^i - y^i)^2$ as the loss function since the engagement score is continuous and well distributed over the input range. We used the Adam optimizer with default parameters ($\alpha = 10^{-3}$) to train our model.

$$
\begin{aligned}
i &= \sigma(x_t U^i + s_{t-1} W^i) \\
f &= \sigma(x_t U^f + s_{t-1} W^f) \\
o &= \sigma(x_t U^o + s_{t-1} W^o) \\
g &= \ tanh(x_t U^g + s_{t-1} W^g) \\
c_t &= c_{t-1} \circ f + g \circ i \\
s_t &= \tanh(c_t) \circ o
\end{aligned}
$$

Figure 2: LSTM cell equations

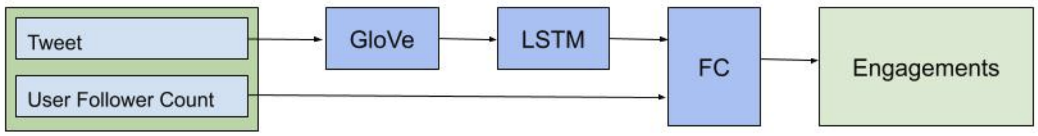


Figure 3: Engagement Prediction Model Architecture

We optimized hyperparameter values by sweeping valid hyper-parameter combinations select the one with the lowest validation loss. The final values are shown in the table below.

Table 2: Model Hyper-Parameters

Hyperparameter Name	**Acceptable Values**	**Optimal Value**
GloVe Embedding Size	[20, 50, 100, 200]	100
LSTM Layer Count	[1,2,3]	2
LSTM Hidden Dimension Size	[8,32,64,128,256]	256
FC Layer Count	[1,4,8,10]	4
Activation Function	[ReLu, Tanh, Sigmoid, Leaky Relu]	Relu

4.3 Word Selection

After building our predictive engagement model, our goal was to find the best alternative words could be utilized to increase engagement. We did this by:

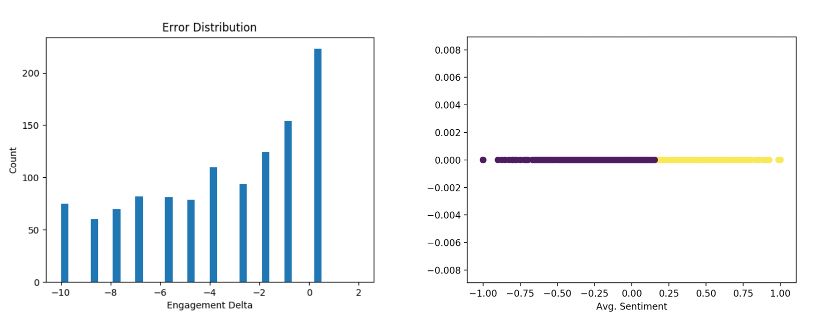
1. Iterating through unique words in the text corps dictionary
2. For each word, finding 5 closest alternatives in the Glove embedding
3. For each alternative, iterating through all tweets that contain the word and substituting the word with the alternative
4. Recomputing the engagement score with the new string
5. Recording the top 8 alternatives with the highest engagement score delta

5 Results and Discussion

To evaluate our unsupervised clustering of users (by sentiment), we measured the distortion across a train/val/test split and checked for over-fitting. The distortion on our training set of users for K-Means was 384.08, the distortion on our val set was 190.06, and our test set distortion was 180.13. This implies that the unsupervised learning model is not over-fitting our dataset. We also measured the cutoff sentiment to understand how our groups were being formed. The cutoff sentiment for our two user groups was 0.16, which indicates that our unsupervised model naturally grouped anything from negative to slightly positive together, and moderately positive or above in a separate group.

We quantified our engagement score model by measuring the validation loss and average prediction error. To prevent overfitting, we used regularization techniques and early stopping. The final engagement score prediction model test set error distribution is shown in the figure below. The majority of tweets have a small engagement error (below 2). The average prediction error is 2.2.

The top 10 alternative words that have the highest positive cross-group engagement delta are shown in the table below. More generally, these are the word substitutions that, accounting for both sentiment groups, would stimulate the highest predicted engagement improvement.



(a) Engagement prediction error distribution (b) K-means clustering output

Table 3: Model Recommendation

Engagement Delta	Original Word	Suggested Word
1792	threatens	endanger
1536	approval	blessing
1280	immediate	contiguous
1280	contaminated	pollute
1024	costing	cost
1024	spare	bare
1024	greening	rejuvenation
896	options	choice

The recommended substitutions made by the network generally follow our intuition with a few exceptions. The network recommends words that are positive, empowering, and hopeful in nature, such as blessing and rejuvenation, rather over-used words that sounds formal and administrative, such as approval and greening. This is interesting because it indicates that politicians should potentially focus less on the dire consequences of climate change and put more emphasis on climate change solutions. Indeed, "choice" being a recommended word conveys the significance of having the autonomy and willpower to act. Furthermore, substitutions like "endanger" for "threatens" and "pollute" for "contaminated" may be the result of humans associating "endanger" and "pollute" with more negative, human-caused phenomena. It's possible the tweets with these words provoked a stronger reaction from people reading them, increasing their engagement.

6 Conclusion/Future Work

Our model and pipeline output made many plausible substitutions for words related to climate change. In our pipeline, we found that a critical step was the box-cox transformation on our input distribution. This allowed us to prevent overfitting by predicting 0 for a greatly skewed engagement distribution. This allowed our RNN to work well and accuratley predict engagement scores.

There is a lot of potential future work to be done with this project. We realized that user specific information can help to improve the network's ability to predict engagement. Factors such as the user's geographic location, his or her political view, and his or her followers' demographic information can significantly impact user response. We can enhance our data source to incorporate these features as inputs.

We also believe that our model can be replicated in the future to more specifically address substitutions for key climate words and phrases used today, such as testing out improvements for "climate crisis," "environmental collapse," etc. Currently our model is limited to word substitutions and finding the highest delta of engagement regardless of the frequency of the word used. More manual substitution labelling could be performed to run these tests on more popular words and phrases, and we provide a framework in this project to do so.

7 Contributions

In this project, Anthony primarily worked on data collection and the preprocessing pipeline for tweets/engagement scoring. Rui primarily worked on the unsupervised learning part and the predictions of the engagement scores given the shift in vocabulary. Harold worked on the RNN models and generating the engagement score and distribution. We would like to thank Blue Sheffer and Zachary Ostroff for their consistent help throughout the process of this project.

The project github link can be found here: https://github.com/raguiar2/cs229_reacclimate

References

[1] Emily M. Cody. (2018). Climate Change Sentiment on Twitter: An Unsolicited Public Opinion Poll, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0136092

[2] Xiaoran An: Tracking Climate Change Opinions from Twitter Data, https://pdfs.semanticscholar.org/0a20/18c2a701d72d0ded2a9f58faf49f34099e81.pdf

[3] Tweepy v 3.5.0. https://www.tweepy.org/

[4] Trevino, A. (2016, December 06). Introduction to K-means Clustering. Retrieved from https://www.datascience.com/blog/k-means-clustering

[5] George Washington University. (2019, May 20). TweetSets. Retrieved from https://tweetsets.library.gwu.edu/

[6] Zhao Chen, Alexander Hristov, Darvin Yi. A Deep Learning Analytic Suite for Maximizing Twitter Impact, https://cs224d.stanford.edu/reports/ChenHristov.pdfc

[7] Xin Wang, Yuanchao Liu, Chengjie Sun, Baoxun Wang and Xiaolong Wang. Harbin Institute of Technology. (2015). Predicting Polarities of Tweets by Composing Word Embeddings with Long Short-Term Memory, Harbin Institute of Technology, https://www.aclweb.org/anthology/P15-1130

[8] J. Li, H. Xu, X. He, J. Deng and X. Sun. (2016). "Tweet modeling with LSTM recurrent neural networks for hashtag recommendation," 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, 2016, pp. 1570-1577, http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=7727385isnumber=7726591

[9] Britz, Denny, Recurrent Neural Network Tutorial, Part 4 – Implementing a GRU/LSTM RNN with Python and Theano, http://www.wildml.com/2015/10/recurrent-neural-network-tutorial-part-4-implementing-a-grulstm-rnn-with-python-and-theano/