intro:

Global warming is an important area of discussion and research in the world today. Though it is widely believed to exist, there still are people who do not accept its validity. There are also groups of people that admit that they think global warming is true, but regard it as an unsubstantial threat to themselves and the world.

People’s perception of global warming is important to consider when looking at potential solutions to this worldwide problem. This paper uses social media to gain an insight on people’s opinion of global warming, the effects of political events, natural disasters and other international events, and the effect these events can have on people’s perceptions. More specifically, this paper studies what type of events raise the most awareness and can change a person’s beliefs of global warming.

method:

The first step in this project was to collect data. Using the Twitter API, a program collected and stored data about tweets with any of the keywords “climate change”, etc from April seventh 2015 to “enter current date”. The data included the tweet, the date and time of the tweet, the username, the location of the user, etc.

The next step was to create a machine learning algorithm to classify the tweets. There were three classifications of tweets, pro global warming, anti global warming and neutral. The pro class included tweets for people who believe in the global warming issue and are raising awareness. The anti class is for people who believe global warming is false, or is not a pressing issue. The neutral class included tweets that had little to do with global warming. For example, the algorithm would catch a tweet that said, “We need to change the climate of our workplace.” This tweet and others like it with no direct connection to global warming would be included in this class.

The machine learning algorithms used were naive bayes and support vector machines (SVM). By manually searching through a few thousand tweets, 464 pro, anti and neutral tweets each were found for the training set. Due to the similarity of the tweets, it was difficult to discern between anti and pro tweets, especially due to the use of sarcasm, but the program had a success rate of approximately 80-85%. (talk about 60:40 or 70:30) The program then took in the millions of tweets collected and returned the sentiment of each tweet.

The next step was to determine which events were causing the daily fluctuations of pro and anti global warming tweets. Multiple news websites were used to conclude which events were causing influxes in tweets. These events included extreme weather conditions like heat records, natural disasters, political debates, business deals, charity events and more. The daily tweets were examined as well to see which events in a day caused more substantial numbers. Through this, a rough timeline of climate change events was created.

(talk about signal analysis and polarity)