

CSE 578 PROJECT PORTFOLIO

SANIKA SHAH

TITLE:

PEARL - Personal Emotion Analysis, Reasoning, and Learning.

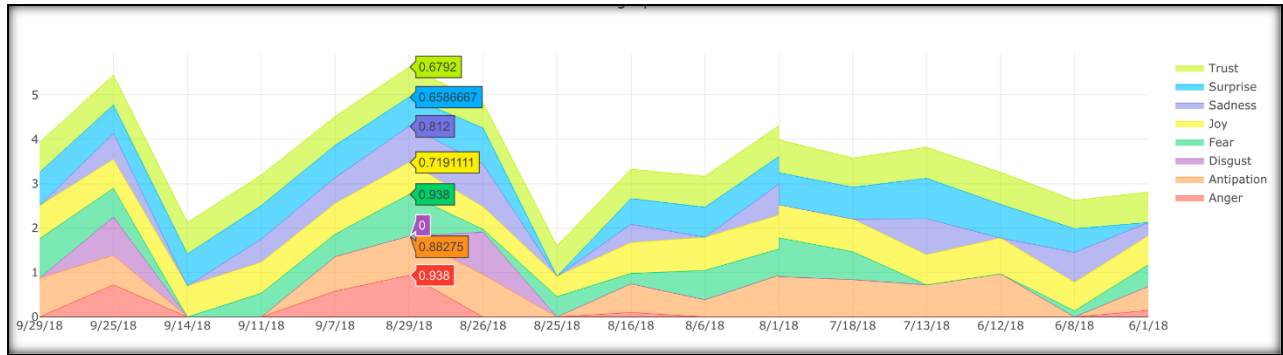
INTRODUCTION:

It is a timeline based, visual analytic system which visualizes a person's emotional profile based on their past tweets. The tweets are classified into 8 categories using Plutchik model. Another model called VAD model is used to show Valence, Arousal and Dominance values of the words that belong to the tweet. Given a username to the system, it'll display the graphs which will summarize the emotional profile of the user. The primary visualization shows the trend of all 8 emotions over time and other graphs show detailed view of tweets and the words that led to the emotions along with their VAD values.

IMPLEMENTATION:

- We collected the tweets of popular users such as Narendra Modi, Donald Trump, Barack Obama from June 2018 to September 2018 in a csv file which consisted of username, tweet, and date of the tweet.
- We used NRC dataset to get Plutchik emotion and VAD scores of words.

- Lexicon based approach was used to classify tweets into one of the emotions of Plutchik model and assigned probabilities to those.
- Similarly, we assigned VAD values for each word of a tweet and the average values were then taken for the tweet to find the valence, arousal and dominance of tweet.
- The final csv contains the username, tweet, date and time of tweet, probabilities of each emotion and their VAD values.
- For the primary visualization, we used D3 and javascript to implement an interactive timeline graph to plot the valence values of each emotion over the months June to September. This chart is shown in below figure.



Emotional profile view based on the twitter handle of Narendra Modi

RESULTS:

The above graph shows how the emotions vary with time based on tweets. Other graphs show detailed view of the tweets which includes raw tweet view that gives the tweets for that period of time, scatterplot which shows wordwise emotions and VAD values. We have added one extension graph which shows number of tweets per emotion and its trend with time.

MY CONTRIBUTION:

I contributed the following:

- Preprocessing and classifying data.
- I used NRC dataset to get wordwise emotions and VAD values.
- I classified the tweets into 8 categories of Plutchik model using Lexicon based approach

- I also calculated VAD values for tweets by averaging the values for words
- I visualized the scatterplot of Valence and Arousal values for words of the tweet using d3.js

LESSONS LEARNED:

- Data preprocessing using python which involved lot of file handling
- Data classification using Lexicon based approach
- Use of D3 for visualizing graphs

TEAM MEMBERS

- Sanika Shah
ssshah33@asu.edu
- Divya Prakash Sivakumar
dsivaku2@asu.edu
- Mugdha Kolhe
mkolhe@asu.edu
- Tanmay Manolkar
tmanolka@asu.edu

REFERENCES:

- Jian Zhao, Liang Gou, Fei Wang, and Michelle Zhou. PEARL: An Interactive Visual Analytic Tool for Understanding Personal Emotion Style Derived from Social Media. IEEE Conference on Visual Analytics Science and Technology (VAST), 2014.