PROJECT PROPOSAL

PROJECT SYSTEM

PEARL - An Interactive Visual Analytic Tool for Understanding Personal Emotion Style Derived from Social Media

TEAM MEMBERS

1) Name: Sanika Shah
   Email: ssshah33@asu.edu

2) Name: Divya Prakash Sivakumar
   Email: dsivaku2@asu.edu

3) Name: Mugdha Kolhe
   Email: mkolhe@asu.edu

4) Name: Tanmay Manolkar
   Email: tmanolka@asu.edu

SUMMARY

The PEARL system is an interactive timeline based visual analytic tool which allows users to view examine a person's emotional style derived from his tweets. The system consists of three main components: text preprocessor, emotion analysis engine, and visualizer.

For a given twitter user id, all the tweets of that user were collected. The text preprocessor then cleans the raw tweet using tokenization and stemming to extract keywords. These keywords are then the input to the emotion analysis engine which will then derive the emotions based on two emotional models, namely: Plutchik's discrete *categorical model* to classify emotional states into eight primary emotions and also characterize all the detected emotions or moods by the continuous *dimensional model* with three measurements (VAD): valence, arousal, and dominance.

For the visualisation, the system consists of an emotional profile overview, a detailed emotion timeline view, a mood word view, and a raw tweets view. In order to inspect the emotional profile, an interactive legend is also provided along with an action menu. Critical points in the emotional profile are also shown.

FEATURES WE WON'T IMPLEMENT

- Instead of raw tweets view we will be listing the sentimental classification for each tweet.
- Informative tooltips that shows the words tweeted by the user.
- Tooltips displayed on timeline graph (circle points) that shows tweet wise VAD model values.
- E,O,R action buttons which shows the extreme motions, emotional outlook and resilience.

SYSTEM EXTENSIONS

- Adding a pie chart to the system which gives relative percentage of tweets for different sentiments in the given time period which can be selected from the timeline graph.
- Planning to implement a machine learning approach for sentiment analysis instead of lexicon based approach given in the paper which has certain limitations like classification of negative sentences (eg: I am not happy)

ADDITIONAL DATASETS

https://www.kaggle.com/kazanova/sentiment140

EXPECTED TIMELINES

Sept 22 - Exploring Twitter api (Sanika, Mugdha), visualization tools (Divya Prakash, Tanmay)
Sept 30 - Tweets preprocessing - tokenization(Mugdha,Tanmay), stemming(Sanika), metadata extraction (Divya Prakash)
Oct 20 - Detecting affective expressions in tweets using models (Sanika, Mugdha), classification/clustering of emotions/moods(Divya Prakash, Tanmay)
Oct 22 - Submit short milestone report - ALL
Nov 15 - Graph Visualisation - ALL
Nov 20 - Elevator pitch - ALL
Nov 27 - Poster sessions - ALL
Nov 30 - Submit project and report - ALL