

NLP (CS671A) Project Proposal

Affect in Tweets

Group Number: 9

Group Members: Anshdha (anshdha@iitk.ac.in)
Apurv Gupta (apurv@iitk.ac.in)
Bhanu Yadav (bhanuydv@iitk.ac.in)
Debabrata Ghosh (drtghosh@iitk.ac.in)
Divyat Mahajan (divyatm@iitk.ac.in) (Auditor)
Piyush Bagad (bpiyush@iitk.ac.in)

Mentor: Dr. Harish Karnick (*hk*)

January 19, 2018

1 Problem Statement

The problem is a compilation of an array of tasks intended to automatically determine the intensity of emotions and intensity of sentiment or valence of the tweeters from their tweets. The individual tasks comprise of emotion and intensity regression and ordinal classification as well as multi-label classification of emotions.

2 Dataset

The tasks are a part of SemEval-2018 (CodaLab competition) for which the datasets can be found in the following online repositories:

- competitions.codalab.org/competitions/17751#learn_the_details_-_datasets
- saifmohammad.com/WebPages/TweetEmotionIntensity-dataviz.html

3 Expected Workflow

We will try to cover the the following tasks in our project:

- **Emotion Intensity Regression:** Given a tweet and an emotion, determine the intensity of that emotion that best represents the mental state of the tweeter i.e. a real-valued score between 0 (least of that emotion) and 1 (most of that).
- **Emotion Intensity Ordinal Classification:** Given a tweet and an emotion, classify the tweet into one of four ordinal classes of intensity (anger, fear, joy, and sadness) of that emotion which best represents the mental state of the tweeter.

- **Sentiment Intensity Regression:** Given a tweet, determine the intensity of sentiment or valence that best represents the mental state of the tweeter i.e. a real-valued score between 0 (most negative) and 1 (most positive).
- **Sentiment Analysis Ordinal Classification:** Given a tweet, classify it into one of seven ordinal classes, corresponding to various levels of positive and negative sentiment intensity, that best represents the mental state of the tweeter.
- **Multi-label Classification:** Given a tweet, classify it as ‘neutral or no emotion’ or as one, or more, of eleven given emotions(anger, anticipation, disgust, fear, joy, love, optimism, pessimism, sadness, surprise, trust) that best represent the mental state of the tweeter.
- **Feature Analysis:** For each extracted entity from tweets, we add its semantic concept as an additional feature, and measure the correlation of the representative concept with negative/positive sentiment. We apply this approach to predict sentiment for the dataset.

References

- [1] *WASSA-2017 Shared Task on Emotion Intensity*. Saif M. Mohammad and Felipe Bravo-Marquez. In Proceedings of the EMNLP 2017 Workshop on Computational Approaches to Subjectivity, Sentiment, and Social Media (WASSA), September 2017, Copenhagen, Denmark.
- [2] *Emotion Intensities in Tweets*. Saif M. Mohammad and Felipe Bravo-Marquez. In Proceedings of the sixth joint conference on lexical and computational semantics (*Sem), August 2017, Vancouver, Canada.
- [3] *Semantic Sentiment Analysis of Twitter*. Hassan Saif, Yulan He and Harith Alani. The Semantic Web – ISWC 2012, Springer Berlin Heidelberg, Berlin, Heidelberg, 508–524.
- [4] *Challenges in Sentiment Analysis*. Saif. M. Mohammad A Practical Guide to Sentiment Analysis, Springer, 2016.
- [5] *Stance and Sentiment in Tweets*. Saif M. Mohammad, Parinaz Sobhani, and Svetlana Kiritchenko. Special Section of the ACM Transactions on Internet Technology on Argumentation in Social Media, In Press.