CS 2731
Introduction to Natural Language Processing
<u>Project Proposal</u>

Submitted by:

HuiHui Xu
Raphael Fernandes
Sanya Bathla Taneja

MOTIVATION:

Social networks, like Twitter, are one of the main platforms where individuals express their opinions, feelings and thoughts on a variety of topics. These messages give invaluable information of individual and the general public, like individual mental states as well as the emotion of a larger group.

There has been extensive work in sentiment analysis, particularly using Twitter data. Analysis of sentiment of messages posted on Twitter (tweets) and users of the microblogging platform is a fascinating question for researchers, evidenced by the fact that Twitter Sentiment Analysis has been one of the major tasks in each SemEval Shared Tasks since 2013 (Rosenthal, Nakov, et al.; Rosenthal, Mohammad, et al.; Nakov et al.; Rosenthal, Farra, et al.; Mohammad et al.). SemEval is the International Workshop on Semantic Evaluation, formerly SensEval.

Sentiment can be categorized as negative, positive, and neutral. However, emotions have more fine-grained classifications and these vary according to the annotated dataset. Emotion identification and classification in tweets is an interesting subproblem of the overarching Twitter Sentiment Analysis goal, which looks into more fine-grained emotion classes within the positive and negative sentiment categories. The range of emotions used for this task has varied depending on the study as well as the dataset used due to subjectivity in recognizing emotions in short texts. Common choices of emotions include the six emotions mentioned by Ekman (fear, anger, disgust, sadness, happiness and surprise in Ekman's Atlas of Emotions) and eight basic emotions defined by Plutchik (joy, trust, fear, surprise, sadness, anticipation, anger, and disgust in Plutchik's "wheel of emotions").

PROBLEM STATEMENT:

We are going to tackle the problem of emotion classification in English tweets and explore the underlying multiclass, multilabel nature of the classification. We propose to analyze the features (such as individual words, hashtags and emoticons) that contribute to different emotions and emotion intensity in the tweets using our dataset as well as through sentiment lexicons.

Our dataset further contains 11 different emotions. As emotion classification is highly subjective, we believe that there are certain emotions that are more similar and not so easily distinguishable such as joy and surprise, anger and disgust, and optimism and joy. We propose to investigate how the pairs of emotions are similar to each other in the way they manifest in the dataset.

POTENTIAL APPLICATIONS:

Natural language applications in commerce, public health, disaster management, and public policy can benefit from knowing the affectual states of people—both the categories and tthe he intensities of the emotions they feel.

MAIN APPROACH:

- Use Random Forest ensemble classifier for the multi-class, multi-label emotion classification task.
- Random Forest further provides important features for emotion classification. The advantage lies in the fact that this algorithm classifies the data using features that are are human readable and thus can give insight into the classification beyond the usual metrics.
- After recognizing these features, an analysis of their properties such as valence, arousal and dominance can be undertaken.
- We will further explore possible classification models using different combination of emotions that have higher correlation as opposed to other emotions, and observe whether this improves upon using less categories of emotions than earlier.

In order to use Random Forest, we are going to transform out multi-label problem into single label problem. Binary relevance is a way to do that. For example, we will pair attributes and labels into number of groups that is equivalent to the number of labels.

EXPERIMENTAL SETUP:

➢ **Dataset**

- SemEval-2018 Task 1: Affect in Tweets (Mohammad et al.), subtask E-c (Emotion Classification).
- The dataset consists of 10,983 English tweets with each tweet containing one or more sentiment labels. This is further divides in 6838 in train, 886 in dev set, and 3259 tweets in the test set.
- The sentiments are divided into anger, anticipation, disgust, fear, joy, love, optimism, sadness, surprise, and trust. The set of emotions includes the eight basic emotions as per Plutchik (1980), as well as a few other emotions that are common in tweets (love, optimism, and pessimism).

➢ **Task**

- We plan to train a random forest classifier with our train corpus, using the dev set for hyperparameter tuning.
- The important features will be retrieved, analyzed and visualized using a decision tree.
- After acquiring these features we will check their arousal, dominance and valence in different lexicons: MPQA (Deng and Wiebe) and NRC VAD Lexicons (Mohammad)

➢ **Metrics**

- **Jaccard index**

We use Jaccard index to compare the similarity and diversity of sample sets. This is the commonly used metric for multilabel classification and is also the official metric in the competition (Mohammad et al.).

- **F1 Micro**
  For computing the micro precision all true positives and false positives of all tweets are summed up. These sums are used to calculate a single micro precision value.

- **F1 Macro**
  The macro precision is the average of these single per-tweet precisions. Further on the macro F1-measure is also calculated as the average of the single per-tweet F1-measures.

➤ **Baseline**

The random baseline is provided for the multi-label classification for all three metrics mentioned above. We are also provided with the results of the winning teams of the original competition.

➤ **Work plan:**

1. Preprocess the data set and do Random Forest hyperparameter tuning (1 week)
2. Get and analyze features (1 week)
3. Find features in lexicon (2 weeks)
4. Similarity of emotions (2 weeks)
5. Report (1 week)

RELATED WORK

Previous work, both in and outside of the SemEval competition, has focused heavily on classifying tweets into one of the three possible categories: positive sentiment, negative sentiment, or no sentiment (neutral) using a wide range of models from supervised machine learning (Chikersal et al.) and more recently deep learning methods (Hanafy et al.). The unique qualities of Twitter data such as the limited text length, use of emojis, hashtags and other social media components make sentiment analysis of tweets a challenging problem. There have also been various studies exploring the impact of emoticons on sentiment of tweets (Li et al.).

Emotion classification is a multiclass, multilabel classification problem. There have been several efforts to classify emotions in tweets (Cbalabantaray et al.; Suttles and Ide) as well as analyze the linguistic features involved in tweets with different emotions (Mohammad and Bravo-Marquez; Hasan et al.). Nevertheless, it remains an intriguing problem to solve with different sets of emotions available for use, multiple methods for classification, and scope to find interesting patterns in text that help in emotion detection.

This project is an extension of the SemEval-2018 Task 1: Affect in Tweets SubTask E-c which classifies tweets as 'neutral or no emotion' or as one, or more, of eleven given emotions that

best represent the mental state of the tweeter (Mohammad et al.). The winning team for this task was NTUA-SLP that used deep attentive RNN's and transfer learning for emotion classification given the dataset (Baziotis et al.).

REFERENCES

Baziotis, Christos, et al. *NTUA-SLP at SemEval-2018 Task 1: Predicting Affective Content in Tweets with Deep Attentive RNNs and Transfer Learning*. https://arxiv.org/pdf/1804.06658.pdf.

Cbalabantaray, R., et al. "Multi-Class Twitter Emotion Classification: A New Approach." *International Journal of Applied Information Systems*, vol. 4, no. 1, Sept. 2012, pp. 48–53, doi:10.5120/ijais12-450651.

Chikersal, Prerna, et al. *SeNTU: Sentiment Analysis of Tweets by Combining a Rule-Based Classifier with Supervised Learning*. 2015, http://url.com.

Deng, Lingjia, and Janyce Wiebe. *MPQA 3.0: An Entity/Event-Level Sentiment Corpus*. http://mpqa.cs.pitt.edu.

Hanafy, Mohammad, et al. *Combining Classical and Deep Learning Methods for Twitter Sentiment Analysis*. Springer, Cham, 2018, pp. 281–92, doi:10.1007/978-3-319-99978-4_22.

Hasan, Maryam, et al. *EMOTEX: Detecting Emotions in Twitter Messages*. http://www.liwc.net/.

Li, Mengdi, et al. "Multi-Class Twitter Sentiment Classification with Emojis." *Industrial Management & Data Systems*, vol. 118, no. 9, Emerald Publishing Limited , Oct. 2018, pp. 1804–20, doi:10.1108/IMDS-12-2017-0582.

Mohammad, Saif M. *Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words*. http://saifmohammad.com/WebPages/nrc-vad.html.

Mohammad, Saif M., et al. *SemEval-2018 Task 1: Affect in Tweets*. https://competitions.codalab.org/competitions/17751.

Mohammad, Saif M., and Felipe Bravo-Marquez. *Emotion Intensities in Tweets*. https://github.com/felipebravom/AffectiveTweets.

Nakov, Preslav, et al. *SemEval-2016 Task 4: Sentiment Analysis in Twitter*. https://github.com/aritter/twitter.

Rosenthal, Sara, Preslav Nakov, et al. *SemEval-2014 Task 9: Sentiment Analysis in Twitter*. 2014, https://dev.twitter.com.

Rosenthal, Sara, Saif M. Mohammad, et al. *SemEval-2015 Task 10: Sentiment Analysis in Twitter*. 2015, http://www.purl.com/net/lexicons.

Rosenthal, Sara, Noura Farra, et al. *SemEval-2017 Task 4: Sentiment Analysis in Twitter*. https://trends24.in/.

Suttles, Jared, and Nancy Ide. *Distant Supervision for Emotion Classification with Discrete Binary Values*. Springer, Berlin, Heidelberg, 2013, pp. 121–36, doi:10.1007/978-3-642-37256-8_11.