PANicDEMIC: Emotions on COVID-19 Over Time Using Natural Language Processing

Team 34: 20160888_ErnChern, 20180755_Duc Khanh Thi Nguyen, 20170906_AssemZhunis 20180745_MaiTungDuong, 20140005_Minsoo Kang

1 Introduction

COVID-19 is a pandemic affecting people emotionally and quantifying the emotional impact of pandemic is challenging. We propose a model to detect emotion patterns in english tweets related to COVID-19 over time, investigate causes and tracking the trend of emotions by classifying tweets to 4 emotion classes: sadness, fear, joy and anger.

2 Approach

In initial pipeline, from collected tweets, we extract features, perform classification of tweets into 4 classes and extract the cause for each class. However, with our available data, the initial pipeline raises a problem because of the absence of labeled dataset for COVID-19 tweets. To solve the problem, we update our pipeline with transfer learning: train feature extractor with the general tweets dataset, combine this with additional COVID-19 features (the cause of emotion we extracted in the initial pipeline) to adapt to our specific target of COVID-19 tweets. **(see Appendix B)**

3 Implementation

3.1 Datasets

- Target dataset (2): Over 30,000 English tweets about COVID-19 per day from March and April retrieved from Kaggle.
- Generic dataset (3): Saif Mohammad dataset is tweets that are pre-labeled into 4 emotions. 75% of the data was used to train the classifier and the rest was reserved for validation. Instances of each categories are counted as: Anger-857, Fear-1147, Joy-823, Sadness-786. Dataset can be considered to be balanced.
- Covid-19 mini-dataset: manually labeled tweets and news headlines, collected with tweepy and by manual work. This dataset was used for testing accuracy of the feedback classifier. Instances of each categories are counted as: Anger-25, Fear-25, Joy-25, Sadness-25

3.2 Feature Extraction

By using Bag-of-words model, we count every occurrence of a word in the tweets, and used TF-IDF vectorizer to give more weight to important words. From generic dataset, emotional words were extracted and given normalized scores based on log likelihood for each emotion (with smoothing). So each words have likelihood score for 4 emotions (classes) with offset to 0 for the least likely class. The scoring formula is as in **Appendix A**.

3.3 Classification

In classifier, we summed the scores of the words per class in a tweet, so each tweet has scores for each emotions. The tweets were then classified to the class with highest score. After cause extraction the classifier was modified to include more relevant features.

3.4 Cause Extraction

After testing the classifier on Target dataset, the most frequent words per classes were vectorized and K-mean clustered. Then the cause for classes were compared to acquire exclusive causes and common between certain classes. These organized cause set was used as parameters of accuracy experiment. Different feature sets and score modification was tried. 450k tweets were used for each testing.

3.5 Feedback

With the more relevant causes extracted, we added the features and added or deducted scores according to how how the feature is expected to affect the likelihood. Through various experiments to improve the classifier, we figured out subtracting score when an unlikely feature appears worked best with reasonable behavior.**(see Appendix C)**

4 Experiments

1. Our classifier (GitHub) was first validated on the generic dataset to evaluate the baseline classifier.
2. The classifier was tested on the COVID-19 minidataset to see if it adapts well to new data.
3. We perform classification on the target Kaggle dataset, extract the cause and feedback to the classifier. The classifier was again tested on the COVID-19 minidataset to see if it feedbacks has added some informative features.
4. We use the adapted classifier to perform classification on the target Kaggle dataset to see the trend

extract the cause.
5. We additionally perform classification on news dataset to compare.

5 Result & Discussion

5.1 Feature Extraction

Emotional words are extracted from generic dataset and visualized using word cloud (see **Appendix D**). Some of the examples are *depression*, *unhappy*, *gloomy* for **Sadness**, *nightmare*, *nervous*, *panic* for **Fear**, *optimism*, *lively*, *glee* for **Joy**, and *offended*, *fury*, *rage* for **Anger**.

5.2 Classification Accuracy with Feedback

In the generic model, the classification accuracy on the generic test dataset is **79%** for CountVectorizer and **80%** for TF-IDF Vectorizer. The confusion matrices are in **Appendix E**.

On the other hand, testing the classifier on COVID-19 mini-dataset got **83%** accuracy before feedback. After feedback, the accuracy of classification has reached **84%**. The confusion matrices are in **Appendix E**.

5.3 Emotions in Tweets

From the proportion of each emotion in tweets over time (see **Appendix F**), we can see that anger tweets are dominant, as more than half of the tweets are classified as anger. Fear follows next. By looking closer at the trend of anger tweets, we notice that there is a sharp increase during the week between April 13 and 19. This may be explained by the sudden rise of death cases in the United States during the week. Another possible explanation of the trend can be unfair treatment during the critical situation, according to some of the top anger tweets during the week:

- *This makes me so angry I can't even express how angry I am COVID19 TrumpOwnsEveryDeath RacismIsAVirus* [anger score: 2.09]
- *All Black People being Quarantined. AfricansinChina chineseracism RacismFromChina RacistChina RacismInChina* [anger score: 2.07]

5.4 Emotions in News

The proportion of emotions in the news is similar to tweets (see **Appendix F**). Anger is dominant and fear follows. However, there is a peak in anger emotion during the week between April 6 and 12. By looking at the top anger news during the week, we found news related to the launch of mobile tracking by the South African government for COVID-19 (news link). This has brought worries that the government will use the tracking system to spy on people(news link). Privacy concerns can be one of the reasons that cause anger in people.

5.5 Cause Extraction

The result of the cause extraction for each emotion is shown as word clusters in **Appendix G**.

Anger: One of the clusters is related to **quarantine**, with words such as *lockdown*, *quarantine* and *stayathome*. Another notable cluster is related to **government**, with words such as *government*, *trump* and *response*. We infer governments' failures in response have sparked anger in people. This aligns with the guess in explanation of the trend in anger tweets.

Fear: An exclusive cluster is related to **death**, with words such as *death*, *confirmed* and *cases*. This result is as expected, that people are frightened by the confirmed and death cases during the pandemic.

Joy: Although it is not a common emotion during pandemic, the result shows a cluster of **positivity**, with words such as *hope*, *thank* and *love*. We can see that some people are spreading positive thoughts such as hope and gratitude during the pandemic.

Sadness: One of the clusters is related to **frontline workers**, with words such as *workers*, *medical* and *patients*. Another cluster is from **empathy**, with words such as *lost*, *understand* and *die*. We can see that the sacrifice of frontline workers and the feeling of losing loved ones made people sad.

6 Conclusion

To sum up, anger stays dominant in emotions on COVID-19. The main sources of anger may lie in the government response and policies. However, further investigation is needed to support our hypothesis. The next step may include classifying results by location and analyzing not only new titles but also the content of the news. In this case, building a news-specific model may be required.

Overall, by using a simple and interpretable model in Python and NLTK, our project may contribute in providing an insight to take into account the mental state of the country, while imposing some policies and setting recovery strategies during hard times of pandemic.

References

[1] GitHub link to the project repository: https://github.com/john-mai-2605/PANicDEMIC

[2] Kaggle data set for COVID-19 tweets can be found on https://www.kaggle.com/smid80/coronavirus-covid19-tweets-late-april

[3] Saif Mohammad pre-labeled dataset on 4 emotions http://saifmohammad.com/WebDocs/EmoInt%20Train%20Data/

[4] Julie Beck's Article : New research says there are only four emotions. https://www.theatlantic.com/health/archive/2014/02/new-research-says-there-are-only-four-emotions/283560/

[5] Agencies. 2020. South Africa launches mobile tracking of those with coronavirus. April. https://www.today.ng/news/africa/south-africa-launches-mobile-tracking-coronavirus-289989

[6] Norma Young. 2020. South Africans are worried the government will use coronavirus phone tracking to spy on them. April. https://qz.com/africa/1834409/coronavirus-south-africans-are-worried-about-cellphone-privacy/

Appendix A: Scoring

$$raw(w,c) = log\ P(w|c) = log \frac{count(w,c)+1}{count(c) + vocab\_size}$$

$$score(w,c) = raw(w,c) - min_{c'}(raw(w,c'))$$

Appendix B: Pipeline

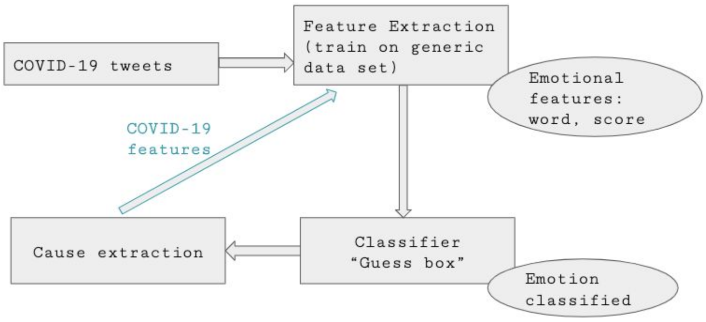

Figure 1: Final pipeline

Appendix C: Accuracy Experiment for Cause Feedback

Settings for 85% accuracy was deemed less credible and the setting wasn't used

Score for the features\Features added	Exclusive Cause	Common to other three	Inclusive Cause
-0.5		0.83	0.84
-0.3	0.83	0.83	0.85
-0.1	0.83	0.83	0.85
0.1	0.83	0.83	0.82
0.3	0.83	0.84	0.82
0.5	0.83	0.84	0.82
1	0.82	0.84	0.81

Figure 2: Accuracy Experiments Results for Different Settings

Appendix D: Word Clouds from Feature Extraction

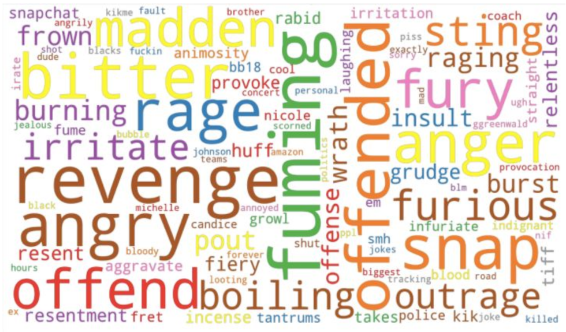

Figure 2: Anger

Figure 3: Fear

Figure 4: Joy

Figure 5: Sadness

Appendix E: Confusion Matrices of Classification

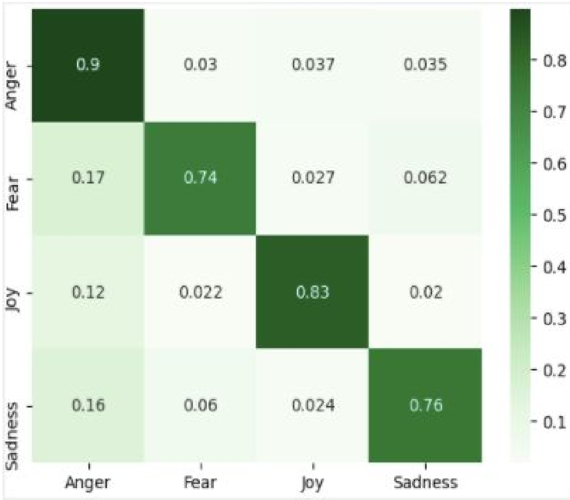

Figure 6: Generic Model on general Tweets dataset

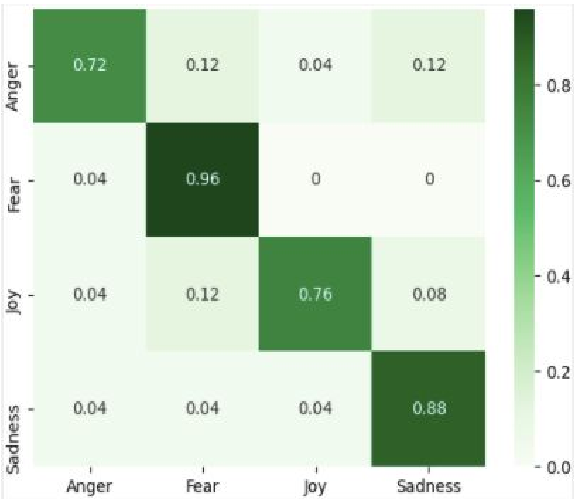

Figure 7: Adapted Model on COVID-19 mini-dataset before feedback

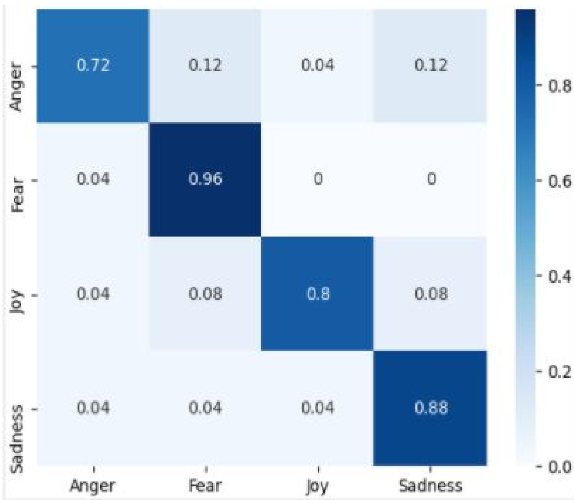

Figure 8: Adapted Model on COVID-19 mini-dataset after feedback

Appendix F: Emotion Trend from March to April

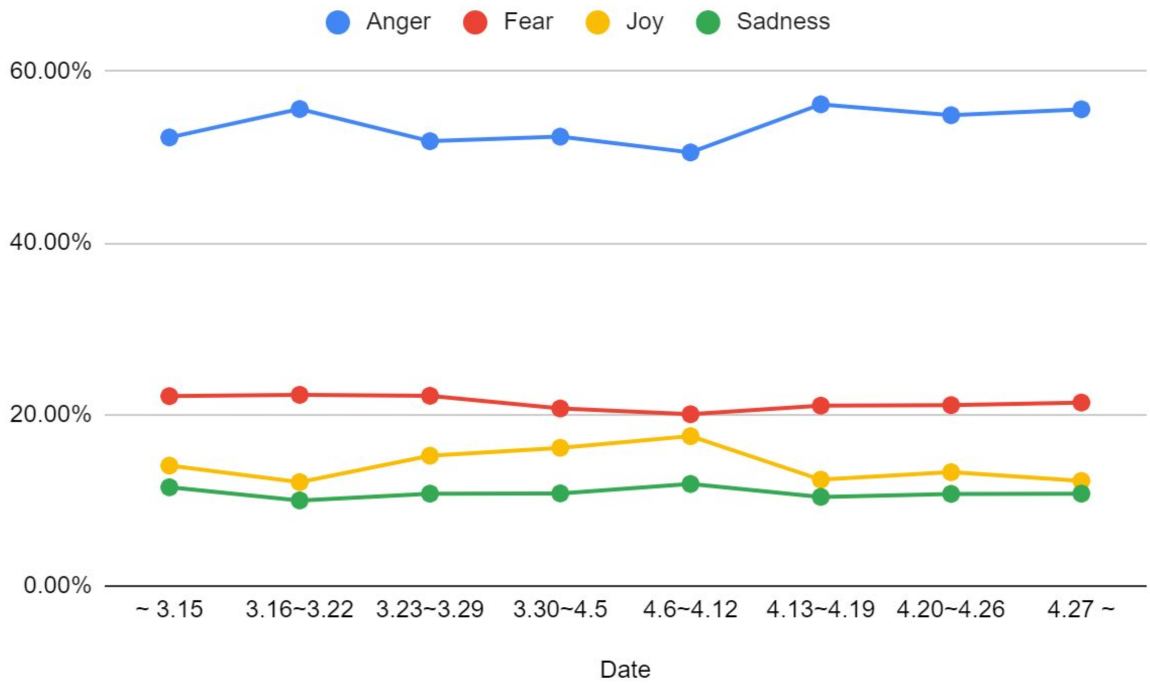

Figure 9: Proportions of emotions in tweets

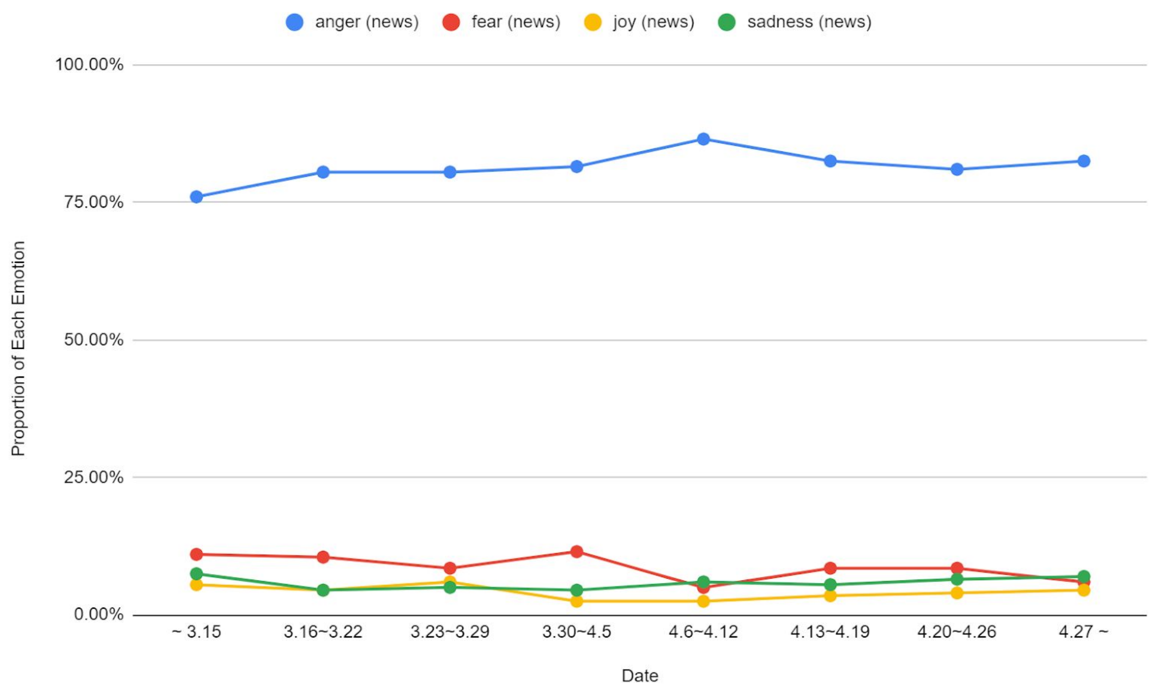

Figure 10: Proportions of emotions in news

Appendix G: Emotion Clusters

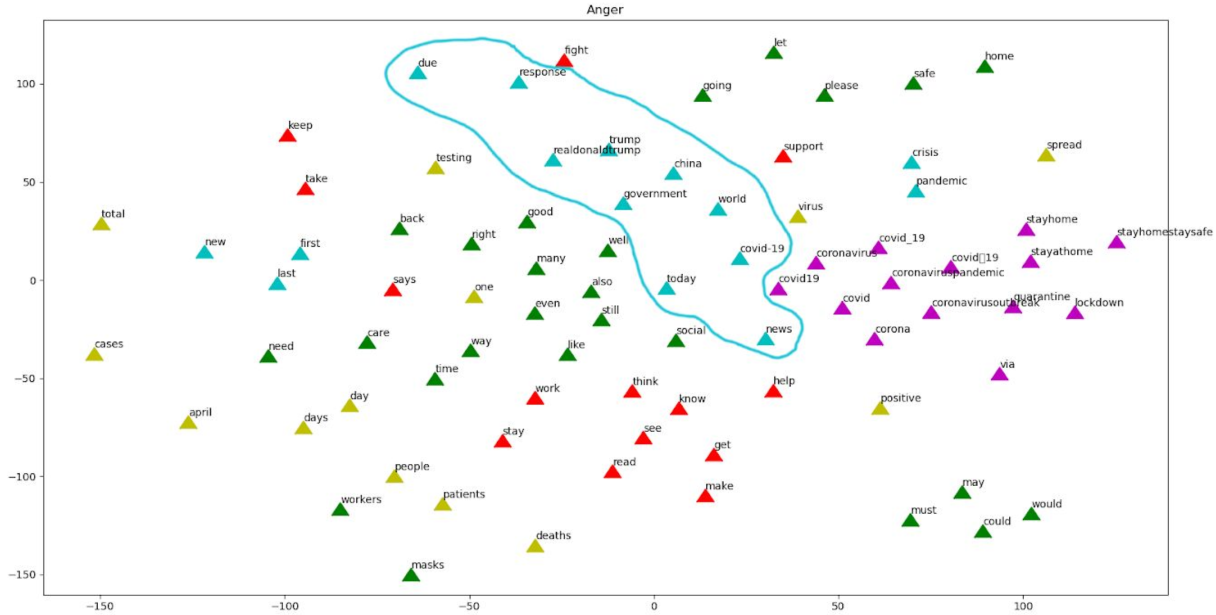

Figure 11: Anger

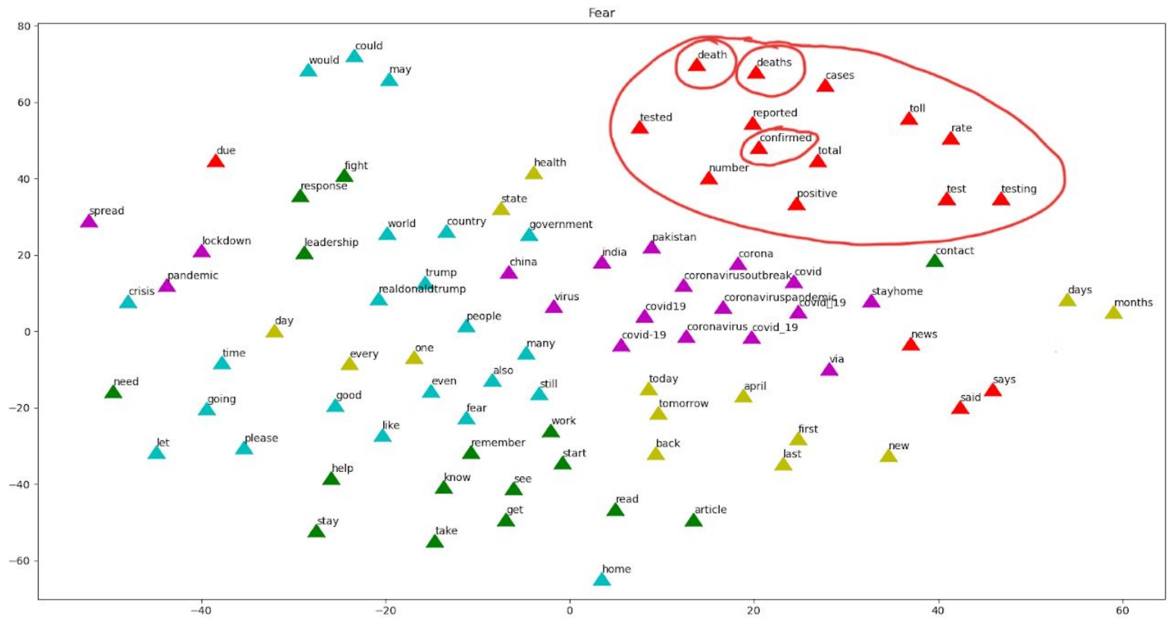

Figure 12: Fear

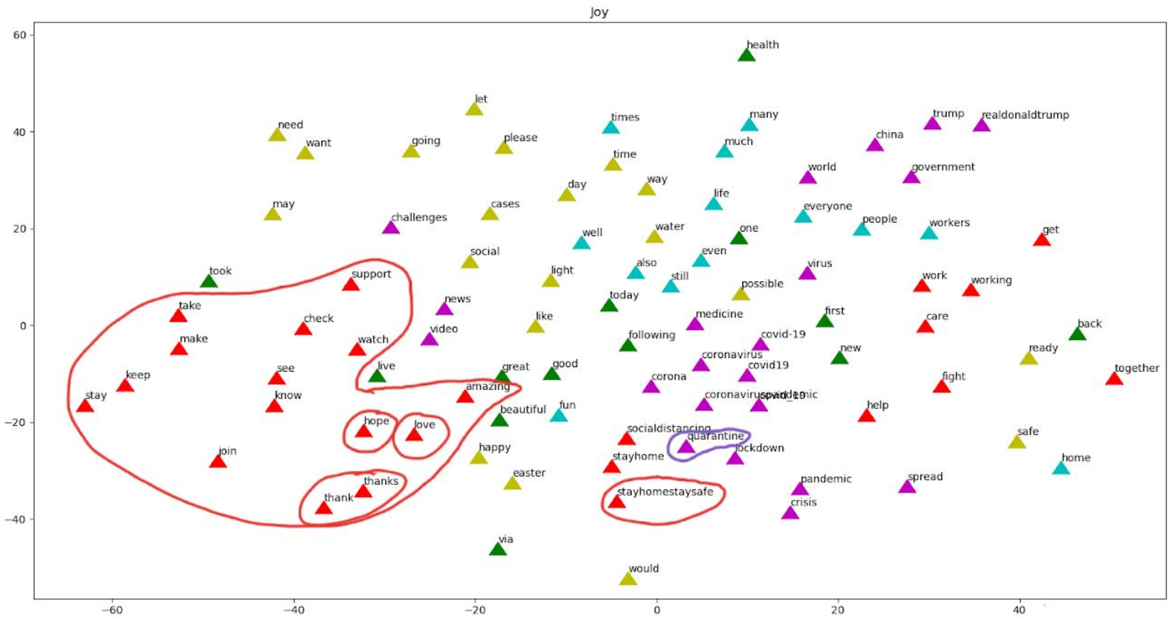

Figure 13: Joy

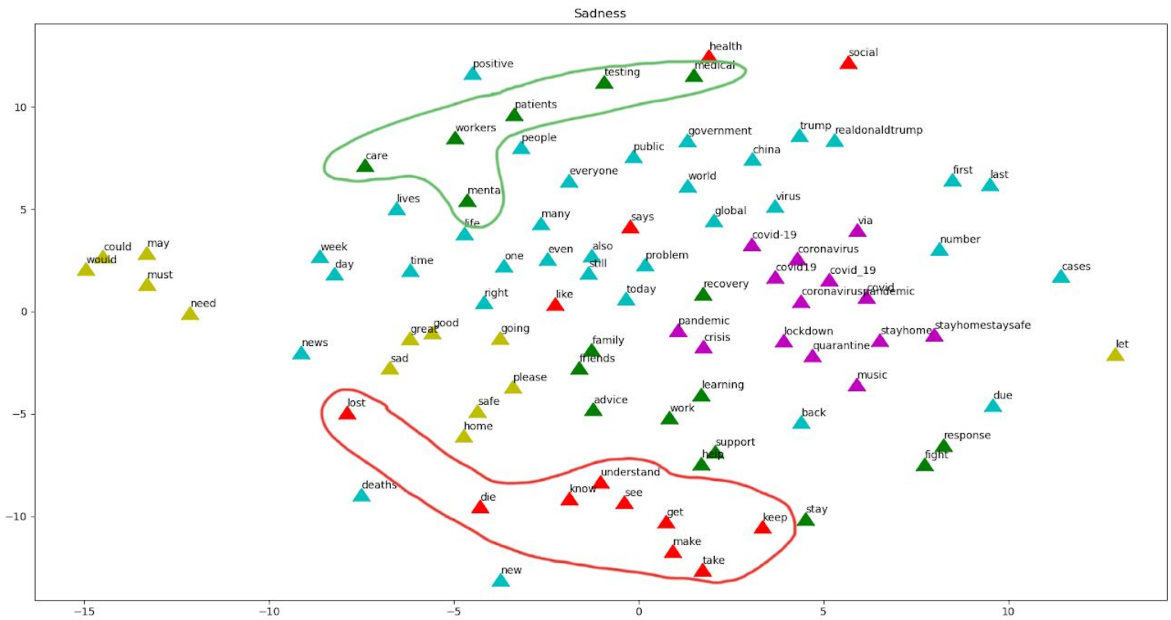

Figure 14: Sadness