Multi-Label Emotion Classification on Trending Tweets using Machine Learning

Simon Islam and Animesh Chandra Roy

Department of Computer Science & Engineering
Chittagong University of Engineering & Technology
Chittagong-4349, Bangladesh
u1504062@student.cuet.ac.bd, animeshroycse@gmail.com

1. INTRODUCTION

As of October 2020, Twitter has 340 million active users. People express their views on Twitter using microblogs called tweets. Thus, it allows users to express their thoughts more thoroughly. Twitter keeps the track of current world using its Trending feature. It detects trends by calculating the volume of a keyword over a given period. Training a model that can classify the emotions in the trending tweets will give us an understanding of the mass mindset in real-time.

1.1 Motivations

Most of the work done in the past regarding tweets classification has either been on sentiment analysis or multi-class classification. But humans are complex beings. We don't think in one emotion at a time. Human thoughts can consist of multiple emotions. Thus, a tweet posted by a human can also contain multiple emotions. For this reason, we need to perform multi-label emotion classification of trending tweets.

1.2 Objectives

The objectives of this work are as follows:

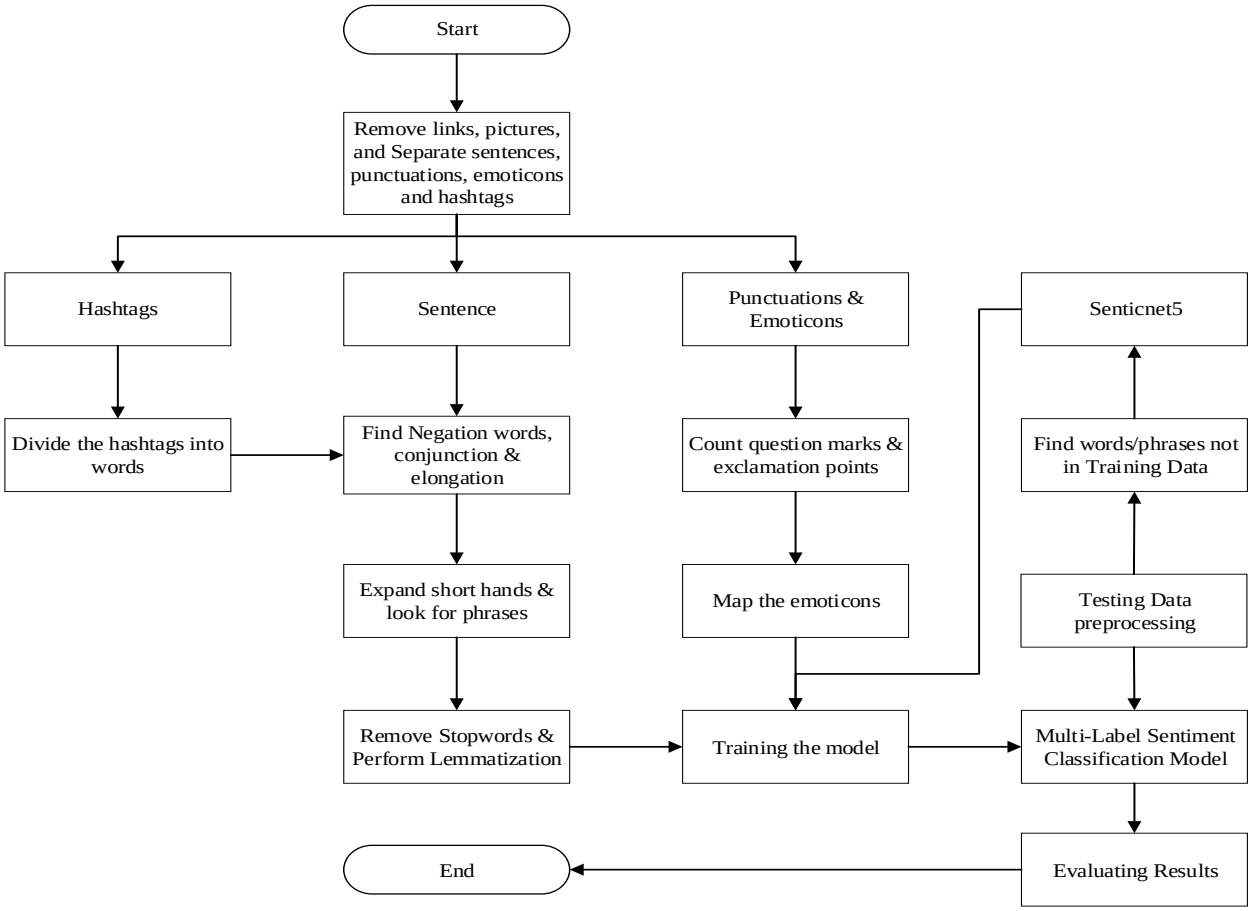
- To extract real-time trending tweets from Twitter
- To perform Multi-Label Emotion Analysis of the tweets (sadness, disgust, joy, interest, admiration, anger, surprise & fear)

2. RELATED WORK

In recent years, multi-label classification has started to gain traction. Liu et al. [1] used multi-label classification approach on three collections of tweets based on various incidents. Bouazizi et al. [2] used 3 pairs of opposite sentiments to label their tweets. They didn't consider statements of opposite polarity. Cabrera-Diego et al. [3] different lexica and two multi-label algorithms to classify their *JIRA Issue Tracker* comments. In our work, we want our classifier to be able to label emotions without depending on a specific topic. It may not be as accurate as those trained on specific topics. But we hope to use it for a general purpose like determining the emotions of trending tweets.

3. METHODOLOGY

To collect trending tweets in real-time, we will use Twitter API. The collected tweets will be sent to our multi-label emotion classifier to classify them as various emotions. The classifier will be trained using a random set of tweets extracted from *Sentiment140* [4] dataset. The tweets will be classified by hand and preprocessed before being used to train the classifier. The words/phrases present in the test set but not in the training set will be added to the training set to boost it's classifying power. *Senticnet5* [5] will be used for this purpose.



4. RESULTS

So far we completed the assignment of emotions to the tweets and used the binary relevance technique to create and evaluate our model. We used Random Forest Classifier and SVM in binary relevance. A total of 7701 tweets were evaluated by hand to assign the appropriate emotions. In the future, we wish to extend it to 10000. We also want to use other algorithms like Label Power Set, Classifier Chains, and HOMER.

4.1 Equations

There are 5 criteria we want to base our evaluation on. They are Hamming Loss (HL), Subset Accuracy (SA), Macro F1, Micro F1, and average Accuracy.

$$HL = \frac{1}{|N|.|L|}\sum_{i=1}^{|N|}\sum_{j=1}^{|L|} xor(y_{i,j}, z_{i,j}) \quad [1]$$

$$Precision = \frac{|T \cap P|}{|P|}, Recall = \frac{|T \cap P|}{|T|} \;\&\; F1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad [2]$$

Here, N is the total number of data samples, L is the number of labels, y is the target and z is prediction. P and T are sets of predicted labels and true labels respectively.

4.2 Tables and figures

Here Table 1, 2 & 3 shows distribution of labels, no of labels present and evaluations respectively.

Table 1. Distribution of labels in dataset

Name of Emotion	No of Tweets
Joy	1907
Sadness	3675
Anger	844
Disgust	1320
Admiration	559
Surprise	801
Interest	1653
Fear	612

Table 2. Number of present labels in tweets

Number of Labels	Number of Tweets
0	312
1	4056
2	2764
3	498
4+	71
Total	7701

Table 3. Evaluations

Evaluation	Binary Relevance	
	RFC	SVM
Hamming Loss	0.249	0.279
Subset Accuracy	0.714	0.668
Micro F1	0.832	0.812
Macro F1	0.827	0.807
Average Accuracy	0.750	0.721

5. DISCUSSIONS

In this work, we tried to determine the emotions being shared on Twitter. People's reactions to trending topics will help us understand the future impact of it. It will give real-time reactions before any detailed and thorough results are available. Thus, it can be used to start any follow-up plans in the right direction.

6. CONCLUSION

Here, we trained multi-label emotion classifiers to evaluate trending tweets. The work is still in progress. It has a low hamming loss indicating its success in detection. We hope that upon further improvements, it will perform all the tasks described in the methodology and will yield an even more successful outcome.

REFERENCES

[1]. S. M. Liu and J. H. Chen, "A multi-label classification based approach for sentiment classification," *Expert Syst. Appl.*, vol. 42, no. 3, pp. 1083–1093, 2015, doi: 10.1016/j.eswa.2014.08.036.

[2]. M. Bouazizi and T. Ohtsuki, "Multi-class sentiment analysis on twitter: Classification performance and challenges," *Big Data Min. Anal.*, vol. 2, no. 3, pp. 181–194, 2019, doi: 10.26599/BDMA.2019.9020002.

[3]. L. A. Cabrera-Diego, N. Bessis, and I. Korkontzelos, "Classifying emotions in Stack Overflow and JIRA using a multi-label approach," *Knowledge-Based Syst.*, vol. 195, p. 105633, 2020, doi: 10.1016/j.knosys.2020.105633.

[4]. A. Go, R. Bhayani, and L. Huang, "Twitter Sentiment Classification using Distant Supervision," *Processing*, vol., pp. 1–6, 2009.

[5]. E. Cambria, S. Poria, D. Hazarika, and K. Kwok, "SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.