

Converting Sentiment Annotated Data to Emotion Annotated Data

Manasi Kulkarni^{1,2}

²Computer Engineering and IT
Veermata Jijabai Technological Institute
Mumbai

manasi@cse.iitb.ac.in

Pushpak Bhattacharyya¹

¹Computer Science and Engineering,
Indian Institute of Technology Bombay
Mumbai

pb@cse.iitb.ac.in

Abstract

Existing supervised solutions for emotion classification demand large amount of emotion annotated data. Such resources may not be available for many languages. However, it is common to have sentiment annotated data available in these languages. The sentiment information (+1 or -1) is useful to segregate between positive emotions or negative emotions. In this paper, we propose an unsupervised approach for emotion recognition by taking advantage of the sentiment information. Given a sentence and its sentiment information, recognize the best possible emotion for it. For every sentence, the semantic relatedness between the words from sentence and a set of emotion-specific words is calculated using cosine similarity. An emotion vector representing the emotion score for each emotion category of Ekman's model, is created. It is further improved with the dependency relations and the best possible emotion is predicted. The results show the significant improvement in f-score values for text with sentiment information as input over our baseline as text without sentiment information. We report the weighted f-score on three different datasets with the Ekman's emotion model. This supports that by leveraging the sentiment value, better emotion annotated data can be created.

1 Introduction

An emotion or a feeling represents a state of mind for any person. Various researchers have put-forward classification of emotions into various categories such as Plutchick emotion model with 8 basic emotions (Plutchick 1980), the Ekman's Model with six basic emotions – anger, disgust, fear, happy, surprise, sadness (Ekman 1972) and so on. Users easily share their experiences, opinions, and emotions on various topics, product reviews on social platforms such as Twitter, Facebook, Whatsapp. Understanding the emotions expressed in such short posts can facilitate

many important downstream applications such as an emotion-aware chatbots, analysis of user reviews, personalized recommendations, a help to psychologically ill patients, and so on. Therefore, it is important to develop the effective emotion recognition models to automatically identify emotions from such text or messages.

The task of emotion detection is typically modelled as supervised multi-class classification or multi-labelled classification task. Supervised models need very large annotated data. Such datasets may not be readily available and are costly to obtain (Jianfei Yu, 2018). In case of unavailability of annotated data, unsupervised learning approaches (A Agrawal, 2012; Milagros Fernández-Gavilanes, 2015) can be an ideal solution for the emotion recognition.

However, we assume that text annotated with sentiment information (positive, negative or neutral) is easily available. Sentiment classification predicts positive or negative sentiment polarity of a sentence whereas emotion classification labels the sentence at fine-grain level with one of the emotions, such as happy, surprise, anger, fear, disgust, sadness etc. Happy and surprise emotion are termed as the positive sentiment emotions and anger, disgust, fear, sadness as the negative sentiment emotions. For example in the sentence,

Passed an exam by two points ... (+1)

The sentiment information provided with sentence in above example, helps to confirm that the sentence is made with positive emotion such as happiness, surprise etc rather than the negative emotions.

A close friend of mine have not contacted me long time. ... (-1)

Sentiment value of -1 shows exclusion of positive emotions by reducing chances of emotion recognition system being confused them with positive

emotions. The sentiment information helps to narrow down choices to one of the negative emotions for the sentence and it must be sadness, fear, anger, or disgust.

Therefore, we aim to use this sentiment information along with the sentence to create the emotion labelled dataset by recognizing emotion in unsupervised way. To be precise to create the emotion labelled resources with help of available sentiment labelled resources, as a further fine grained emotion analysis task.

In this paper, we propose an unsupervised approach based on [A Agrawal \(2012\)](#), with our modifications as discussed later in detail. We use the sentiment labelled data that is the sentence and the respective sentiment information as input and recognize the best possible emotion for the same. This approach uses word-embeddings to represent the words in the sentences as well as emotion-specific words. The cosine similarity measure is used to calculate the semantic relatedness between a sentence and the emotion-specific words. The vector representing score for every emotions category is calculated for each sentence and then emotion-score is modified using dependency relations of open-class words from that sentence.

The rest of the paper is organized as follows. The section 2 describes related supervised, unsupervised and hybrid approaches previously proposed in the literature. The section 3 discusses the methodology of proposed system. Section 4 briefs on the experimental set up, datasets used as well as modifications done to the dataset as required for experimentation. The results are discussed in section 5.

2 Related Work

The state of the art approaches for emotion Recognition task is supervised approach. The labeled training data is a crucial resource required for building such systems. Due to the lack of a large human annotated datasets, many emotion classification tasks have been performed on text data gathered from social media such as twitter, and the hash-tags, emojis or emoticons are used as the emotional labels for the same.

As the unsupervised approaches do not need the annotated dataset, different unsupervised approaches are also performed by researchers.

[A Agrawal \(2012\)](#) found open-class words which they named as NAVA words that is Noun, Adjective, Verb, Adverb words. Pointwise Mutual Information (PMI) based model with syntactic dependencies is used to perform emotion recognition. [Shoushan Li and Zhou \(2015\)](#) created a Dependence Factor Graph (DFG) as learning model based on label dependence and context dependence. The hybrid approaches use the unsupervised approaches for feature creation, pattern extraction which are later used by supervised classification models for emotion classification. [Carlos Argueta \(2015\)](#) had proposed an unsupervised graph-based approach for bootstrapping the Twitter-specific emotion-bearing patterns and then used them for classification task. [Li and Xu \(2014\)](#) used predefined linguistic patterns to extract emotion causes and considered them as features for classification using SVM.

[Jianfei Yu \(2018\)](#) have used transfer learning approach for sentiment classification task and then emotion classification task. Also few researchers have contributed towards creation of emotion-aware embedding. Distant supervision and Recurrent Neural Network (RNN)-based approach is proposed for learning emotion-enriched representations. ([Ameeta Agrawal, 2018](#))

3 Proposed System

The emotion recognition framework for unsupervised approach is as shown in Figure 1. The input sentence is pre-processed to get the open-class words from that sentence. The second step is to compute the semantic relatedness using cosine similarity between word embedding of words in sentence and emotion-specific words. The module three modifies the emotion score for every emotion from vector computed at module 2. Later, module 4 averages over emotion vectors of all words of a sentence to find resulting emotion present in that sentence.

Let S be the sentence, $S = \langle w_1, w_2, \dots, w_n \rangle$ and S_s be the respective sentiment value (+1 or -1 or 0). Let E be the set of possible emotions from selected emotion model such as $E = \{e_1, e_2, \dots, e_m\}$. To every emotion category, we have assigned few affect bearing words which represent that emotion. Table-1 shows few affect-bearing words used for each emotion category. The aim is to predict the best possible emotion E_s belongs to

E for the sentence S . Eventually, one of the emotions from the Ekman’s emotion model- Anger, Disgust, Fear, Happy, Sadness and Surprise will be predicted for every sentence.

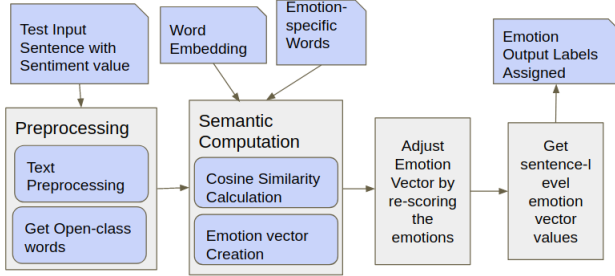


Figure 1: Overview of System

3.1 Computing Semantic Relatedness using Cosine Similarity

We have used pre-trained word embedding as they better represent co-occurrence information of words. The words in a given sentence and the emotion-specific words are represented using their respective word embedding and the semantic relatedness between them is found using cosine similarity. Let A and B be word vector representation for 2 words then:

$$sim(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\| \cdot \|B\|}$$

3.2 Computing Vector of Scores for Emotion Categories

The emotion score vector for every open-class words $\{w_1, w_2, \dots, w_n\}$ of a sentence is created. The length of the emotion vector is six values corresponding to six emotions of Ekman’s model.

The emotion vector for a word w_i is computed by finding cosine similarity of word w_i with every emotion-specific word from each emotion category. Let there be m emotion categories and $\{EW_1, EW_2, \dots, EW_m\}$ be sets of l emotion-specific words for each emotion e_j . Then the emotion score ES for w_i is calculated as:

$$ES(w_i, e_j) = \sum_{k=1}^l sim(w_i, EW_j^k) \quad (1)$$

$$\forall j = 1 \dots m$$

An emotion score vector EV for every word is created using a sentiment value and an emotion scores of corresponding emotions with given

Emotion	Emotion Words
Anger	anger, angry, annoy, irritate, frustrate
Disgust	disgust, hate, dislike, ill, sick
Fear	fear, worry, terrify, afraid, frighten
Happiness	happiness, happy, love, joy, glad
Sadness	sadness, sad, hurt, cry, bad
Surprise	surprise, amazing, astonish, wonderful, incredible

Table 1: Few affect-bearing words used

sentiment value. So, an emotion-score vector for word w_i is,

$$EV_{w_i} = \{ES(w_i, e_1), ES(w_i, e_2), \dots, ES(w_i, e_m)\} \quad (2)$$

Here, the sentiment value plays a crucial role. As we consider the Ekman’s emotion model, happy and surprise emotions are emotion with positive sentiment. The emotions such as fear, disgust, angry and sadness are emotions with negative sentiment. Hence, if $sentiment = +1$ then,

$$EV_{w_i} = \{ES(w_i, happy), ES(w_i, surprise), 0, 0, 0, 0\} \quad (3)$$

if $sentiment = -1$ then,

$$EV_{w_i} = \{0, 0, ES(w_i, anger), ES(w_i, disgust), ES(w_i, fear), ES(w_i, sadness)\} \quad (4)$$

3.3 Re-scoring Scores in Emotion Vector of a Word

With the intuition that dependency relationship can contribute more towards emotion detection, we use these relations of open class words to modify the emotion vector EV of the dependent word. The Stanford’s coreNLP dependency parser is used for finding dependencies between the open-class words that is noun, adjective, verb and adverb which are considered for further processing.

Let $sd(w_1, w_2)$ be a syntactic dependency relation between word w_1 as dependent and word w_2 as modifier. For example, in adjectival modifier relation $amod(life, happy)$, dependent word is $life$ and modifier word is $happy$. We find these syntactic dependencies of sentence at the time of pre-processing itself and use dependency relations for the re-scoring purpose.

Let D be the dependent word from sentence S and M be the respective modifier word from sentence S . Then the emotion vector D_p of p^{th} word is modified by taking average over emotion vectors of the dependent word D_p and its modifier word M_p . This will help in strengthening every emotion score related to that word w_i of sentence S . (A Agrawal, 2012)

$$EV_{D_p} = \frac{EV_{D_p} + EV_{M_p}}{2} \quad (5)$$

3.4 Processing Emojis

With growing usage of social media, many times, text messages are accompanied with suitable emoji. Hence, emoji as input can contribute towards detecting emotion. Every emoji is being assigned CLDR short name by Unicode Common Locale Data Repository to describe that emoji. For example, *grinning face*, *beaming face with smiling eyes* and so on. Same procedure as mentioned in section 3.1 to 3.3 is followed for creating emotion vector M_{EV} for every emoji.

3.5 Computing Emotion Vector for Sentence

The emotion vector S_{EV} for the sentence S is calculated by taking average over emotion vectors EV_{w_i} of all words from that sentence, and emoji emotions vector M_{EV} .

3.6 Resultant Emotion Prediction

The emotion vector of text S is :

$$S_{EV} = \langle S_{e_1}, S_{e_2}, \dots, S_{e_m} \rangle$$

where the emotion vector for emoji, if present in sentence is:

$$M_{EV} = \langle M_{e_1}, M_{e_2}, \dots, M_{e_m} \rangle$$

and

$$S_{EV} = \frac{1}{n} \sum_{i=1}^n EV_{w_i} + M_{EV} \quad (6)$$

the best possible predicted emotion E_s for S as:

$$E_s = \operatorname{argmax}(S_{e_i}) \quad (7)$$

$$\forall i = 1 \dots m$$

4 Experimental Setup

The datasets used for testing and recognizing emotions are ISEAR dataset (ise), Twitter Emotion Corpus (Mohammad, 2012) and Semeval-2018 Affect in Tweets English Test dataset (Saif M. Mohammad, 2018).

ISEAR dataset (ise) The ‘‘International Survey on Emotion Antecedents and Reactions’’ dataset published by Scherer and Wallbott is built by collecting questionnaires answered by people with different cultural backgrounds (Bostan and Klinger, 2018). A total of 7,665 sentences labeled with single emotions. The labels are joy, fear, anger, sadness, disgust, shame, and guilt.

Twitter Emotion Corpus (Mohammad, 2012) is prepared with emotion-word hashtags as emotion labels. These are termed as noisy labels as labelled by users. This corpus contains 21050 sentences labelled with one of the emotions from Ekman’s emotion model.

Semeval-2018 Affect in Tweets English Test dataset (Saif M. Mohammad, 2018) is gold standard multi-labelled dataset with 3259 tweets annotated with multiple emotions. The emotion labels are anger, anticipation, disgust, fear, joy, love, optimism, pessimism, sadness, surprise, trust. Every emotion is labelled as 0 or 1 to show presence of that emotion. If all are 0 then tweet is considered to be neutral.

4.1 Modification in Datasets for Testing

Few modifications are incorporated before using them for testing and experiments.

4.1.1 Mapping of Emotion Labels to Ekman’s Emotion Model

Not all of these datasets are labelled with the Ekman’s emotions. The researchers follow different emotion models such as the Plutchick model, the Parrot’s emotion model. Also few researchers use emotion categorization as per requirement of the system and the data. Hence, we have mapped these emotion labels to one of the best suitable Ekman’s emotion as shown in Table-2

This mapping is coarse-grain mapping as the Ekman’s model represents six basic emotions - Happy, Surprise, Anger, Disgust, Fear and Sadness and all other emotions can be mapped in these emotions directly.

4.1.2 Labeling Datasets with Sentiment Values

Not all the above mentioned datasets are annotated with sentiment values. Hence, to illustrate this problem definition, we labelled these datasets with sentiment value, based on already available emotion labels to the sentences.

Sr No	Dataset Name	Happy	Surprise	Anger	Fear	Disgust	Sadness
1	ISEAR dataset	Happy	–	Anger	Fear	Disgust	Sadness
2	Semeval-2018 Task-1 Affects in Tweets English Dataset	Anticipation, Joy, Love, Optimism, Trust	Surprise	Anger	Worry	–	Pessimism, Sadness
3	Twitter Emotion Corpus	Joy	Surprise	Anger	Fear	Disgust	Sadness

Table 2: Mapping of original emotion labels to Ekman’s emotions

The sentences from datasets with positive emotions such as happy, love, joy, surprise etc are labelled with positive (+1) sentiment value. And the sentences with negative emotions such as anger, disgust, fear, sadness etc are mapped to negative (-1) sentiment value.

Now, the datasets are in the required format for further processing and testing. The format of every testing example is:

< sentence, sentiment value, emotion >.

While testing the system, the sentence and the sentiment value from modified datasets are considered as input and best possible emotion is recognized. Later, these predicted emotions are compared with these emotion labels for checking the accuracy of the system.

4.2 Experiments

- The sentence is pre-processed to remove the stopwords, hyperlinks, hashtags, usernames and the special characters if any. Also part of speech tagging is done to obtain open-class words that is noun, verb, adjective, and adverb. The NLTK PoS tagger and the Wordnet word categories are used to perform the same. As the closed-class words do not contribute towards emotions, they are not considered for further processing. The syntactic dependencies are retrieved for given input sentence using Stanford coreNLP dependency parser.
- We have experimented on different datasets mentioned in Table-2 using different pre-trained word embeddings such as Google Word2Vec (Tomas Mikolov, 2013), Glove (Jeffrey Pennington and Manning, 2014), and FastText (Joulin et al., 2016)

- The experiments are performed on the text with sentiment information and without sentiment information too. The weighted F-score, precision and recall are used as metrics to evaluate the accuracy of system.

5 Results and Discussions

As shown in Table-3, the experiments are performed in two different ways, first by considering only text/sentences as input and secondly by considering text and its sentiment value as input. The experiments conducted with only sentences as input, serves here as baseline against experiments using sentences with sentiment information.

The sentence with respective sentiment information as input shows significant improvement in weighted F-score value. The results are shown in the Table-3. It is observed that Google Word2Vec word vectors performs better than other word embedding.

The Semeval-2018 Task-1 dataset (Saif M. Mohammad, 2018) is multi-labelled dataset. The annotated emotions are assigned independently. This task is multi-class emotion recognition so we consider prediction 'correct' even if one of the assigned emotions is predicted by our system.

It is visible in Table-4 that sentences with sentiment value as input, the F-score of every individual emotion category has been improved drastically. This shows better prospects for such emotion recognition and conversion process for new resource creation with fine-grained labeling from the sentiment to the emotion.

The recall values for datasets using Google Word2Vec are shown in Table-5 for illustration.

Sr No	Word Embedding	ISEAR dataset		Twitter Emotion Corpus		Semeval-18 Task-1 Dataset	
		w/o Sentiment Value	with Sentiment Value	w/o Sentiment Value	with Sentiment Value	w/o Sentiment Value	with Sentiment Value
1	Google Word Vectors	0.37	0.56	0.32	0.52	0.52	0.76
2	GLOVE vectors	0.23	0.33	0.25	0.45	0.38	0.67
3	Fast Text Word Vectors	0.34	0.49	0.30	0.49	0.51	0.74

Table 3: Weighted F-score using different word embedding and with / without Sentiment Value

Sr No	Method (Input)	Anger	Disgust	Fear	Happy	Sadness	Surprise
ISEAR Dataset							
1	Sentence	0.35	0.27	0.43	0.45	0.33	–
2	Sentence and Sentiment Value	0.45	0.39	0.51	0.97	0.52	–
Twitter Emotion Corpus							
3	Sentence	0.21	0.11	0.28	0.55	0.14	0.10
4	Sentence and Sentiment Value	0.32	0.18	0.42	0.79	0.55	0.15
Semeval-2018 Task-1 Affects in Tweets Dataset							
5	Sentence	0.40	0.43	0.39	0.66	0.51	0.21
6	Sentence and Sentiment Value	0.67	0.65	0.53	0.92	0.76	0.32

Table 4: Emotion category-wise F-score for emotion recognition using Google Word2Vec vectors

Recall values in case of the method with sentiment information has increased by approximately 50% than method without sentiment values. This shows significant improvement in correctly predicting emotion on use of sentiment information.

The Table-6 illustrates the confusion matrices for results of emotion recognition on ISEAR dataset. It can be seen that positive emotions and negative emotions are rarely confused with each other by using sentiment information. The recall and precision is also increased for every emotion. Yet emotions belonging to the same sentiment value need to be achieved with better accuracy. The emotion 'surprise' is not part of emotion labels for ISEAR so 0s in row for 'surprise'.

Conclusion

The proposed system suggests the way for creation of a resource from the available resources. The use of more easily available sentiment labelled data for creating emotion annotated data is significant. The use of sentiment information for recognizing the emotion is good example of fine-grain labeling

task.

The proposed approach shows much better accuracy for text labelled with sentiment value than the baseline as text without sentiment information. The use of sentiment information helps to segregate at initial level between emotions with the different polarity.

As the word vectors are based on distributional hypothesis, they may have higher cosine similarity for opposite words, for example, 'happy' and 'sad'. The synonyms may have very low cosine similarity value. This can affect overall accuracy of the system. The rare words may not contribute much and very common words may get very high cosine similarity with opposite words too. Hence, it is necessary to select better list of emotion-specific words. More processing and linguistic information may be added to improve the accuracy of this system.

Sr No	Method (Input)	Anger	Disgust	Fear	Happy	Sadness	Surprise
ISEAR Dataset							
1	Sentence	0.29	0.19	0.37	0.81	0.26	–
2	Sentence and Sentiment Value	0.43	0.26	0.52	0.94	0.65	–
Twitter Emotion Corpus							
3	Sentence	0.22	0.12	0.21	0.78	0.10	0.10
4	Sentence and Sentiment Value	0.37	0.20	0.34	0.94	0.55	0.10
Semeval-2018 Task-1 Affects in Tweets Dataset							
5	Sentence	0.30	0.33	0.49	0.76	0.45	0.89
6	Sentence and Sentiment Value	0.59	0.57	0.70	0.92	0.82	0.92

Table 5: Emotion category-wise Recall values for emotion recognition using Google Word2Vec vectors

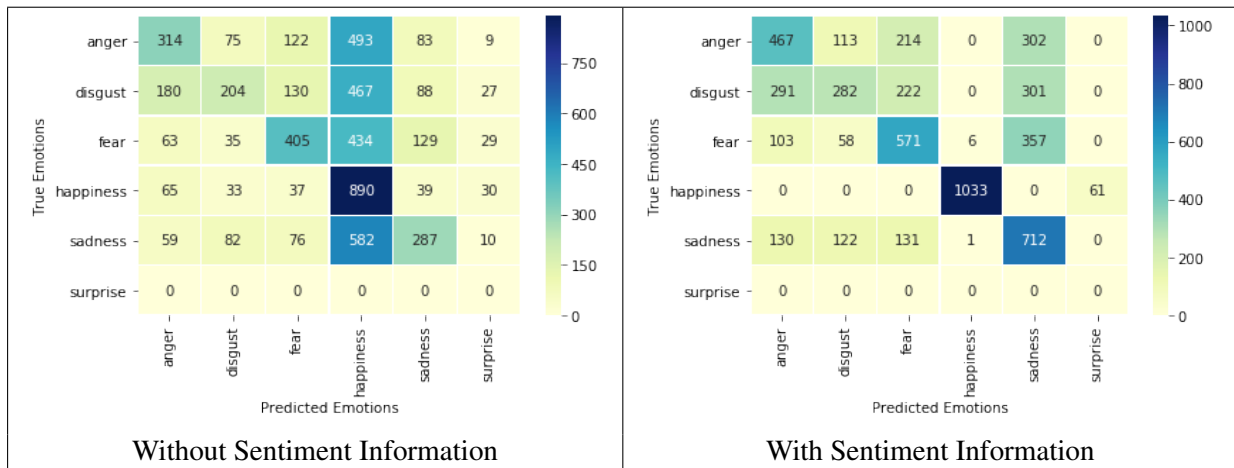


Table 6: Confusion Matrix for ISEAR dataset using Google Word2Vec Vector

References

- Aijun An A Agrawal. 2012. Unsupervised emotion detection from text using semantic and syntactic relations. *IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology*, 1:346–353.
- Manos Papagelis Ameeta Agrawal, Aijun An. 2018. Learning emotion-enriched word representations. *The 27th International Conference on Computational Linguistics, Santa Fe, New Mexico, USA*, page 950–961.
- Wlodek Zadrozny Armin Seyeditabari, Narges Tabari. 2018. Emotion detection in text: a review. *arXiv:1806.00674v1 [cs.CL]*, 2nd Jun 2018.
- Laura-Ana-Maria Bostan and Roman Klinger. 2018. An analysis of annotated corpora for emotion classification in text. *The 27th International Conference on Computational Linguistics, Santa Fe, New Mexico, USA*, page 2104–2119.
- Yi-Shin Chen Carlos Argueta, Elvis Saravia. 2015. Unsupervised graph-based patterns extraction for emotion classification. *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 336–341.
- Al Moatassime Hassan Douiji yasmina, Mousannif Harjar. 2016. Using youtube comments for text-based emotion recognition. *The 7th International Conference on Ambient Systems, Networks and Technologies (ANT)*, pages 292–299.
- Nancy Ide Jared Suttles. Distant supervision for emotion classification with discrete binary values. *CI-CLING, March 2013*.
- Richard Socher Jeffrey Pennington and Christopher D. Manning. 2014. Glove: Global vectors for word representation.
- Jing Jiang Pradeep Karuturi-William Brendel Jianfei Yu, Luis Marujo. 2018. Improving multilabel emotion classification via sentiment classification with dual attention transfer network. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, page 1097–1102.
- Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2016. Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.

- Weiyuan Li and Hua Xu. 2014. Text-based emotion classification using emotion cause extraction. *Expert Systems with Applications*, 41(4):1742–1749.
- Jonathan Juncal-Martínez Enrique Costa-Montenegro Milagros Fernández-Gavilanes, Tamara Álvarez-López. 2015. Gti: An unsupervised approach for sentiment analysis in twitter. *9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 533–538.
- Kiritchenko Svetlana Mohammad, Saif M. 2018. Understanding emotions: A dataset of tweets to study interactions between affect categories. In *11th Edition of the Language Resources and Evaluation Conference (LREC)*, Miyazaki, Japan.
- Saif Mohammad. 2012. emotional tweets. *The First Joint Conference on Lexical and Computational Semantics (*Sem)*, Montreal, Canada., page 246–255.
- Mohammad Salameh-Svetlana Kiritchenko Saif M. Mohammad, Felipe Bravo-Marquez. 2018. Semeval-2018 Task 1: Affect in tweets. *The International Workshop on Semantic Evaluation (SemEval-2018)*, New Orleans, LA, USA, page 246–255.
- Rong Wang Shoushan Li, Lei Huang and Guodong Zhou. 2015. Sentence-level emotion classification with label and context dependence. *53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, 1:1045–1053.
- Kai Chen Greg Corrado-Jeffrey Dean Tomas Mikolov, Ilya Sutskever. 2013. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*.