Sentence-Level Emotion Detection from Text Based on Semantic Rules

Dibyendu Seal, Uttam K. Roy and Rohini Basak

Abstract Detecting emotion from text has become an interesting topic in the field of natural language processing. Emotion detection aims to detect and recognize types of emotion from various sources such as text, facial expression and gestures, and speech. This paper proposes an efficient emotion detection technique by searching emotional words from a pre-defined emotional keyword database. The method analyzes the emotion words and phrasal verbs, also considers negation words and exhibits better performance than recent approaches.

Keywords Emotion detection · NLP · Natural language toolkit

1 Introduction

Emotion is essential to all aspects of human lives. It helps us in decision-making, affects our social relationships and shapes our daily behavior. Detecting emotion is helpful in human–computer interaction (HCI). With the wide use of social media like Twitter, Facebook, Blogs, product reviews and other public forum posts, there is a need to develop an algorithm to identify and analyze emotions available in text.

Emotion can be broadly categorized into six (6): happy, sad, anger, disgust, fear and surprise [1]. In this work, we have considered these six emotions to be detected from sentences which are supposed to consist of emotional words. Phrasal verbs and negative words have been analyzed to fine-tune the result. The model has been tested with the standard dataset and exhibits superior performance to recent works.

D. Seal (✉) · U. K. Roy · R. Basak
Jadavpur University, Kolkata, India
e-mail: dibyendu.seal@gmail.com

U. K. Roy
e-mail: royuttam@gmail.com

R. Basak
e-mail: visitrohinihere@gmail.com

M. Tuba et al. (eds.), *Information and Communication Technology for Sustainable Development*, Advances in Intelligent Systems and Computing 933,
https://doi.org/10.1007/978-981-13-7166-0_42

Section 2 provides recent works in this area; Sect. 3 describes the emotion detection algorithm in detail, and in Sect. 4 we have listed the results and the issues we faced while implementation.

2 Related Work

Although there are various sources of emotion such as text, facial expression and gestures, and speech, detecting emotion from text is most challenging. Emotion detection from text relies on the emotional keyword analysis model.

In one of the primitive works [2] by Cecilia Ovesdotter et al., the authors tried to classify the emotional affinity of sentences in the children's fairy tales for future use of text-to-speech synthesis. In text-based emotion prediction (TEP) task, they tried to classify emotions from children's fairy tales using supervised machine learning and got an encouraging result in a first set of experiments.

As per Chomsky and Saussure [3], sentences can be broken down into two groups of words: function words and content words. Function words include nouns, adjectives, verbs and adverbs, while content words are prepositions, conjunctions and auxiliary verbs. Function words have the most influence on emotion of the text. These words supply basic emotion values or connection to the sentence where content words contain zero of minimum weightage.

Tao [4] analyzed the functional words and classified those into an emotional keyword, modifier words and metaphor words. Emotional keywords are those which can directly relate to the emotion, and modifier words increase or decrease or negate the emotion. Metaphor words are those which are not directly linked to an emotion. For example, 'kindness' always shows the gentle and positive emotion, 'joy.'

Ezhilarasi and Minu [5] tried to extract emotion using WordNet and its word construction. They collected the synsets (inner meaning) of the input from various domains to create an emotional ontology which is further used to classify complex sentences. However, their work toward creating the emotional ontology was limited to affected verb only.

Das and Bandyopadhyay [6] described the method of text emotion identification depending on emotion tagged word. They used WordNet for preprocessing and the conditional random field (CRF)-based word-level emotion classification. They proposed methods for handling negative words in sentence-level emotion detection.

3 Proposed Work

Possibility of extracting emotion or any other computational information from a sentence or text would not be very meaningful until some fundamental artificial intelligence is applied to text. The overall emotion extraction process broadly consists

of text preprocessing, keyword extraction from sentences and keyword analysis to find the emotional affinity of the keyword (Fig. 1).

3.1 Sentence Preprocessing

Text preprocessing performs the following tasks:

- Break the complex and compound sentence to a collection of simple sentences.
- Extract negative and high-intensity words from the sentence, e.g., '*Not*' being considered as negative word whereas '*Very*' being considered as intensity word.
- Parse the input text and extract necessary keywords using POS tagging.
- Extract phrasal verb from the sentence using trigrams and assign special tags to the phrasal verb in order to process them separately further.

Finally, the output of sentence processing module, i.e., extracted keyword, phrasal verbs, special emotion words, negation and intensity words, is passed to the keyword analysis module. On a high level, the module works as below (Fig. 2).

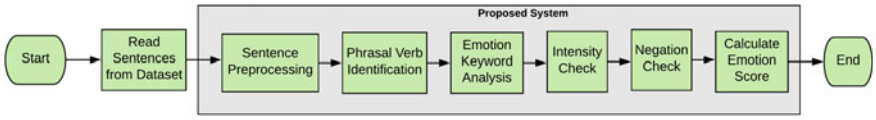

Fig. 1 Emotion detection process flow

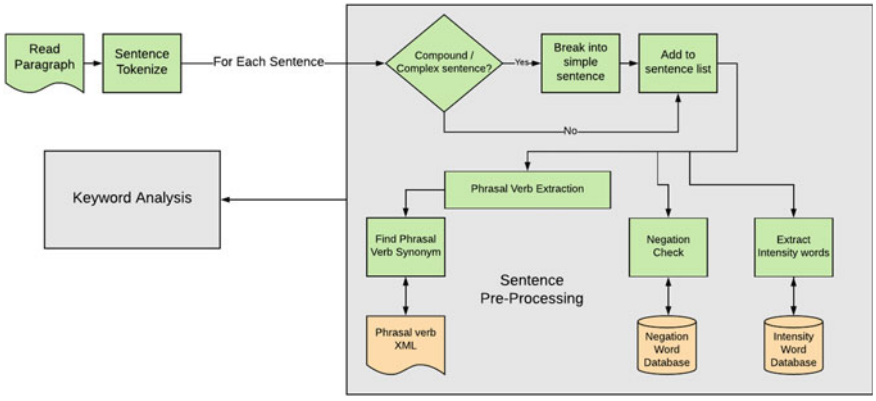

Fig. 2 Sentence preprocessing

3.2 *Phrasal Verb Analysis*

Phrasal verbs play a significant role in emotion detection. We analyzed the ISER emotion dataset and found around 50 phrasal verbs which can directly be assigned to emotion class. No other recent emotion detection work had taken phrasal verbs into consideration. https://www.powerthesaurus.org is an excellent resource of English words and their synonyms. We constructed an exhaustive list of phrasal verbs (around 1200 phrasal verbs) and created a database of synonyms of each phrasal verb. For example, '*Cheer up when the night comes, because mornings always give you another chance.*'—the phrasal verb '*Cheer up*' gives a positive meaning of the sentence. The following phrasal verbs show different emotions:

Anger. Put back, lie around, come along, pass away, go around, go away, talk back, scare away, get back, give away, walk away, get along;
Fear. Pass away, move away, get away, go in, walk back, walk away, go back (Fig. 3).

The output XML contains the phrasal verb and their synonyms as given below (Fig. 4).

Extraction of PV from Text
There are two methods being used to extract the PV from the input text.

- Direct comparison of PV from the phrasal verb database is done to find out the words which are used 'as-is' in the text. If the match is found, we further look into the emotion database for matching synonyms in keyword analysis module.

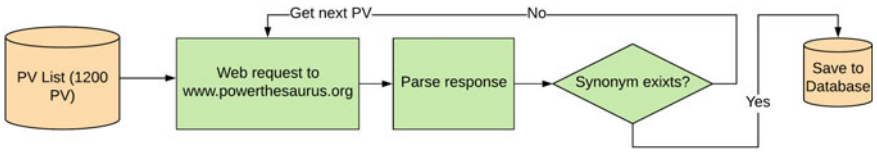

Fig. 3 Phrasal verb synonym extraction process

Fig. 4 Phrasal verb synonym

```
<PVList>
  <Verb>Account_for<syn>explain</syn></Verb>
  <Verb>Account_for<syn>justify</syn></Verb>
  <Verb>Account_for<syn>rationalize</syn></Verb>
  <Verb>Account_for<syn>elucidate</syn></Verb>
  <Verb>Account_for<syn>explicate</syn></Verb>
  <Verb>Account_for<syn>describe</syn></Verb>
  <Verb>Account_for<syn>attribute</syn></Verb>
  <Verb>Account_for<syn>clarify</syn></Verb>
  <Verb>Account_for<syn>number</syn></Verb>
  <Verb>Account_for<syn>cover</syn></Verb>
  <Verb>Account_for<syn>expound</syn></Verb>
  <Verb>Account_for<syn>furnish</syn></Verb>
  <Verb>Account_for<syn>impart</syn></Verb>
  <Verb>Account_for<syn>rationalise</syn></Verb>
  <Verb>Account_for<syn>solve</syn></Verb>
  <Verb>Account_for<syn>spread</syn></Verb>
  <Verb>Account_for<syn>vindicate</syn></Verb>
  <Verb>Account_for<syn>accredit</syn></Verb>
</PVList>
```

- Trigrams are useful when the phrasal verb is being used along with an additional word in between, e.g., your behavior has pissed me off. From this, pissed me off trigram matches with PV list and the synonyms are looked into emotion database for matching keyword.

The process replaces phrasal verbs and assigns special tag: *_EmotionPV_ <Seq No>*, for the subsequent process to understand the occurrence of phrasal verbs.

3.3 Emotional Keywords

First step of emotion detection is discovering keywords and phrases that associate with emotion. A list of words and corresponding emotion is called emotion lexicon. This list starts with root emotion words and is further extended using the WordNet synonyms. We added emotion synonyms to the list from WordNet synonyms and online dictionary (https://www.powerthesaurus.org).

The final list contained several hundred emotion words, for example.

Anger words. Furious, irritated, indignant, shout, mad, disturbed, annoyed, etc.;
Fear words. Anxiety, fright, worry, terror, dismay, horror, etc.

3.4 Keyword Analysis

Keyword analysis module takes the keyword from the sentence as input, looks for the keyword in the emotion lexicon and assigns a score. This module also gets synonyms for the phrasal verbs from phrasal verb synonym list. This module first does the stemming and POS tagging of the tokenized words to get root and restrict pronoun, preposition and interjections to get into the synonym lookup process as those words do not contribute to the emotion directly. This module considers noun, adjective, verb and adverb words for emotion detection keywords being searched into 'emotional keyword' lists and synonyms of phrasal verbs being searched from the phrasal verb xml, and the synonyms are further searched into 'emotional keyword' lists. The overall architecture of keyword analysis is given as (Fig. 5):

The emotion database contains a large number of words that affects the emotion directly. However, it is impossible to create a comprehensive list of such words. Not only that, many words indirectly represent to an emotion. For example, '*Win*' does not define an emotion by its meaning, but any sentence containing '*Win*' has a greater possibility of having joyous character rather than anything other. Our algorithm searches for the key-word in the emotion database and if keyword being matched, then we assign a point to the corresponding emotional score. If the match is not found, then we first extract the synonyms of the word from WordNet lexical database [7] and further search synonyms in the emotion database. We repeated this process for three levels to get significant coverage of keywords. However in each level, we reduced

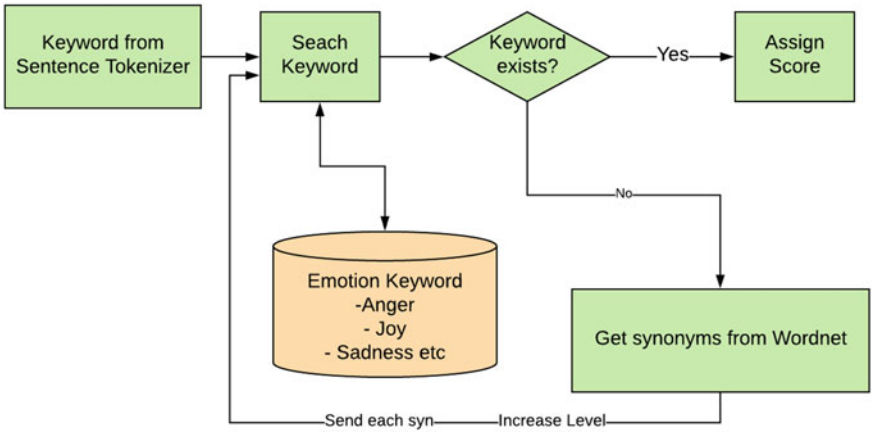

Fig. 5 Keyword analysis

the weight by 33% as meaning gets changed significantly in each level, and beyond third level, the process yields erroneous results.

```
Iteration = 0
Weight = 1
function CheckEmotion:
for keyword in keyWordList:

  If Iteration < 3
  Iteration ++
  Check keyword exists in Emotion Database
    Yes:
      Emotion Score += Weight (Full weightage given)
    NO:
      Get synonyms List

Weight = Weight*0.66 (33% reduction)
GOTO function
```

The emotion score is calculated as below:

$$\text{Emotion}_{\text{angry}} = \sum \text{Weight of Angry for All the Keywords}$$

$$\text{Emotion}_{\text{sad}} = \sum \text{Weight of Sadness for All the Keywords}$$

$$\text{Emotion}_{\text{joy}} = \sum \text{Weight of Joy for All the Keywords}$$

$$\text{CalculatedEmotion} = \text{Max}\left(\text{Emotion}_{\text{angry}}, \text{Emotion}_{\text{sad}}, \text{Emotion}_{\text{joy}}\right)$$

Table 1 Effect of negative and intensity words on emotions

Emotion	Negative words	Intensity words	Both
Angry	Neutral	Disgust	Neutral
Sadness	Neutral	Sadness	Neutral
Fear	Fear	Fear	Fear
Joy	Anger	Joy	Anger
Surprise	Sadness	Fear	Neutral

3.5 Intensity Negation and Check

Emotion detection from the text algorithm considers negative and intensity words in sentence and changes the assigned score derived from the keyword analysis. Negative word changes the emotion significantly, whereas intensity words intensify the meaning. A list of negative and intensity keywords is created, and we find the occurrence of the words in the list.

Negative words. No, not, negative, neither, nor, least, none, nothing;
Intensity words. Very, big, deep, high (Table 1).

4 Experimental Results

We used emotion dataset provided by 'International Survey on Emotion Detection Antecedents and Reactions' (ISEAR) containing around 7500 sentences affecting anger, joy, sadness, fear, disgust and guilt. We created individual emotion dataset containing emotional keywords and observed an overall accuracy of 65%. The current model computes 'neutral' emotion to provide more realistic output. 'Neutral' emotion was not classified in ISEAR dataset and that also reduced the accuracy. To justify the consistency and correctness of the algorithm, we validated our result w.r.t to the result obtained by Das et al. [6].

a = no. of sentences correctly assigned to the emotion;
b = no. of sentences incorrectly assigned to the emotion;
c = total number of sentences;
d = no. of sentences with 'neutral' emotion (Table 2).

4.1 Analysis and Discussion

The stemming algorithms are not robust. Porter stemmer failed to stem certain words. For example, 'startle' (which means surprise/fear) becomes start and '*Sorry*' becomes '*Sorri*'. WordNet synsets are not proper in many times. For example, synonym of

Table 2 Test result comparison

Metrics	Expression	Current experiment result (%)	Conditional random field framework [6] (%)
Accuracy	a/c	65	60.47
Precession (P)	$a/(a + b)$	70.11	67.95
Recall (R)	$(a + d)/(a + b + d)$	62.67	65.11
F1 score	2*(Recall * Precision)/(Recall + Precision)	66.18	Not found

'*Last*' is death which belongs to fear emotion. Few sentences are not grammatically correct; rather, it was more aligned toward the style of speaking of an individual. Current algorithm searches each of the synonyms in emotion database. However, there should be a suitable mechanism to select the orders of those words. Idioms that describe an emotion could not be constructed in this experiment.

5 Conclusion

This paper describes the emotion detection process by searching keywords from emotion database. The proposed system improves the performance over existing methods, though not all the limitations of the existing systems are resolved. The emotion database still does not contain an exhaustive list of emotional keywords, and we are still analyzing more resources to improve accuracy.

References

1. Ekman, P.: Emotions Revealed: Understanding Faces and Feelings, 1st edn. Weidenfeld & Nicolson (8 May 2003). ISBN-13: 978-0297607571
2. Alm, C.O., Roth, D., Sprout, R.: Emotions from text: machine learning for text-based emotion prediction. In: HLT '05 Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, pp. 579–586
3. Saussure and Chomsky: The Distinctness of their 'Identical' views, Magyar Nyelv (Hungarian Language), pp. 420–430 (2006)
4. Tao, J.: Context Based Emotion Detection from Text Input. Interspeech 2004, National Pattern Recognition Laboratory, Chinese Academy of Sciences, pp. 1337–1340
5. Ezhilarasi, R., Minu, R.I.: Automatic emotion recognition and classification. In: international conference on modelling, optimization and computing (ICMOC–2012), pp. 21–26
6. Das, D., Bandyopadhyay, S.: Sentence level emotion tagging: 2009. In: 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, pp. 145–162. https://doi.org/10.1109/acii.2009.5349598
7. WordNet. [Online]. Available: http://wordnet.princeton.edu/