Comparative Analysis of Lexicon-Based Emotion Recognition of Text

Anima Pradhan, Manas Ranjan Senapati, and Pradip Kumar Sahu

1 Introduction

Human emotions are classified into two types: verbal and non-verbal. Verbal emotions are expressed in the form of speech, sounds, or texts, whereas non-verbal emotions come through facial expression, body movement, or hand gestures. Understanding the emotions of a person by analysing his/her feelings or thoughts written in texts is quite a challenging task. This is because most of the time emotional words are not used to express the emotions. Hence the system needs to analyse the texts, interpret, and predict the perception of concepts to identifying human emotions such as joy, anger, and fear

Human–computer interaction plays a significant role in recognizing emotions in the text [1, 2]. Nowadays various social networking sites, such as news, blogs, and discussing forum allows people to share views as emotions, sentiments, and opinions. Quite a few researchers are of the opinion that recognizing emotions is a more important task than identifying sentiment polarity. More than one emotion may be categorized into the same sentiment polarity, i.e. positive, negative, or neutral that can influence the sentence differently. For example, "I was scar" (FEAR) and "The morning newspaper has not arrived yet" (ANGER) come under the negative polarity. Both sentences convey different types of information to the decision-makers from the perspective of emotions [3]. Therefore, researchers have proposed emotion recognition task using emotion-word lexicon decisions [4] and machine learning methods [5].

The emotional analysis is a fine-grained model and is known as a natural evolution of sentiment analysis. Several articles have been written about sentiment analysis with a limited amount of work focusing on emotion recognition from texts. Emo-

A. Pradhan (✉) · M. R. Senapati · P. K. Sahu
Department of Information Technology, Veer Surendra Sai University of Technology, Burla, Odisha, India
e-mail: animap2011_phdit@vssut.ac.in

tion recognition has many applications such as stock prediction [6], advertisement or product recommender systems [7], political speech [8] influenced by people's emotions, marketing strategies [9] of a company based on consumer's emotions, etc. Generally, there are three labels, namely positive, negative, or neutral to represent sentiments. However, at the same time for emotions, a distinct number of representations exist such as "Plutchik's wheel of emotions" [10] with eight emotions (joy, surprise, trust, sadness, fear, anger, anticipation, and disgust) or "Ekman's" [11] six emotions (sadness, fear, happiness, disgust, anger, and surprise). "WordNet-Affect (WNA)" [12] and "NRC word-emotion lexicon" [13] include handcrafted emotion lexicons which associates between words and emotions identified by "Plutchik" and "Ekman".

Though various number word-emotion lexicons have been developed for English, the size of emotion lexicons is still small than sentiment lexicons. Another challenging task is to create high-quality and high-precision emotion lexicons for the researchers."Depechemood" is one of the largest emotion lexica, which generate numerical scores for various emotion automatically. Later, an extended version of "Depechemood", is developed known as "Depechemood++ (DM++)", to improve the performance in terms of coverage and precision using simple techniques. Here the data is directly feed into the lexicon and it interprets associated emotions to score automatically rather than to only label them.

Therefore, "DM++" is focused on emotion recognition on textual information and compares the performance with another emotion lexicon "NRC", which is also publicly available on the web. To extract emotions, techniques of "Natural Language Processing (NLP)" are applied and implemented on Python language version 3.6.

The organization of this paper is as follows. In Sect. 2, related work on "machine learning" and "lexicon-based" approach for emotion recognition is presented. In Sect. 3, detail of our research method for automatic emotion classification is explained. Result is evaluated in Sect. 4 and conclusion of the paper is presented in Sect. 5.

2   Related Work

In this section, a review of the research effort to detect emotions made by different researchers is presented. Based on the two popular techniques, the review is divided into a "machine learning" and "lexicon-based approach". In a machine, it depends on the availability of the word-emotion pair in the respective lexicon [14], whereas the domain-independent nature of "lexicon-based approaches" makes it training dependent.

2.1 Machine Learning Approach

"Machine learning approaches", such as supervised and unsupervised learning depend on the various classifiers. "Plutchik's wheel of emotions" is classified using different classifiers ("Logistic Regression", "Bayesian", "Support Vector Machines (SVM)", and "Random Forest"), and their performances are compared [5]. Another study compared three machine learning classifiers, "SVM", "Decision Tree", and "Naive Bayes" to a lexicon-based approach ("NRC lexicon"). Some studies demonstrated the results using the "Naive Bayes" classification algorithm in emotion detection [15, 16]. Other studies classified emotions using the "SVM machine learning classification algorithm" [13, 17–19].

2.2 Lexicon-Based Approach

"Lexicon-based approaches" use single or multiple lexical resources to detect emotions. The most popular lexicon "WordNet Effect" was developed [16] by tagging effective synsets with "Ekman's" six basic emotions with its meaning in English "WordNet". It contains 2874 synsets and 4787 words. Though the "WordNet effect" is of limited size, its quality is good as it was created and validated manually. "NRC Emotion lexicon", the largest annotated emotion lexicon [20], contains 14,200 unigram words obtained from Google n-gram corpus accompanied by "Plutchik's eight emotions". "DepecheMood" [21] was created automatically by extracting social media data from "rappler.com", which were crowd annotated news articles accompanied "Rappler's Mood meter" that allowed the users to share their feelings/emotions about the articles they are reading. The lexicon consists of 37K words with seven emotion scores (afraid, inspired, sad, angry, annoyed, don't care, happy, and amused)."DepecheMood++" is a high-precision/high coverage lexicon and extended version of "DepecheMood" used in domain-specific tasks [22].

3 Automatic Emotion Classification

Here, a brief description of the process on how to collect, annotate the data set, and compare the publicly available lexicons and to apply NLP techniques on "NRC" and "DepecheMood++" is given.

3.1 Data Source

International Survey on Emotion Antecedents and Reactions (ISEAR) sentence-label emotion data set consists of 7666 sentences is used in the experiment. It is the collection of news headlines from news websites and newspapers. This data set consists of seven emotion classes: joy, disgust, anger, fear, shame, surprise, and sadness. The data set which is in a CSV file and labelled with emotions is extracted using Pandas dependency. The extracted data is then used to show the average percentage of votes for each emotion. Here, Joy has a higher percentage of votes as reported in Table 1.

First, the emotion matrix *Emotion_matrix* is built using "DepecheMood++" emotion lexicon, which provides the voting percentage of each sentence in the eight emotion labels: happy, angry, amused, don't care, afraid, annoyed, inspired, and sad. Then, each document is Part of Speech (PoS) tagged and the nouns, adjectives, and verbs are extracted, which are later lemmatized and the lists of lemmas feed into the lexicon to compute the emotion score for each emotion label.

Mathematically, it was written as follows:

Let D be a set of documents represented as follows: $Dn = \{d1, d2, ...dn\}$ where n is total number of documents, $E(Di) = \{$basic emotion assigned to document$\}$ and $Em = \{e1, e2, ...em\}$ be the list of emotion labels represented as follows: ["AFRAID" ,"AMUSED","ANGRY" ,"ANNOYED" ,"DONT_CARE" ,"HAPPY", "INSPIRED" ,"SAD"].

Based on "Rappler's mood meter", the lexicon contains eight mood-related words. The technique is applied on the data set which consists of seven emotion classes. Out of the eight mood-related word used in "Rappler's mood meter", four words like happy, angry, sad, and afraid are replaced with joy, anger, sadness, and fear for its applicability on the dataset is being used. The rest of the four emotions Amused, Annoyed, Don't Care, and Inspired are discarded as it is not available in the data set that is being used in the experiment. Even though the emotion words are discarded but still the technique has assigned some emotion score because another similar word is used in the sentence. A part of the matrix generated by this process is given in Table 2.

Table 1 Average percentage of votes for each emotion in dataset

Emotion	Votes	Emotion	Votes
Joy	0.145	Sadness	0.140
Fear	0.144	Disgust	0.142
Anger	0.143	Shame	0.142
Guilt	0.140		

Table 2 An excerpt of the *Emotion_matrix*

Sent_id	AFRAID	AMUSED	ANGRY	ANNOYED	DON'T_CARE	HAPPY	INSPIRED	SAD
doc_id_01	**0.074**	0.146	**0.115**	0.138	0.125	**0.095**	0.214	**0.091**
doc_id_01	**0.079**	0.138	**0.110**	0.128	0.159	**0.101**	0.171	**0.114**
doc_id_03	**0.091**	0.128	**0.116**	0.122	0.116	**0.084**	0.215	**0.127**
doc_id_04	**0.103**	0.141	**0.134**	0.131	0.131	**0.108**	0.127	**0.14**
doc_id_05	**0.084**	0.164	**0.107**	0.132	0.124	**0.097**	0. 0.207	**0.084**

The bold value signify the probability value of common emotions in both the dataset and Rappler's mood meter

4 Evaluation

Experiments on the data set is performed using several benchmark algorithms. For all the experiments, the data labelled with Joy, Angry, Sadness, and Fear are considered.

The correlation between the emotion score extracted from *Emotion_matrix* is compared with the predicted score for the ISEAR data set using "Pearson's correlation". The result obtained from the correlation analysis is given in Table 3. It can be verified that for "NRC" correlation score is low for emotions like fear and anger, whereas it is high for joy and sad. Similarly, for "DM++", all the four emotions correlation score are high. The result shows that "DM++" outperformed the "NRC". To carry out the classification for the each emotion, emotion scores are normalized between 0 to 1 using the formula given below:

$$e' = \frac{(e - \min(e))}{(\max(e) - \min(e))} \tag{1}$$

The normalized emotion score is then converted into a binary representation. If the score is more than 0.5, changed into 1 otherwise 0. For evaluation, F1-Measure is employed, and the results obtained are given in Table 4.

The classification accuracy for the corpus using "Naive Bayes'", "Logistic Regression", "Support Vector Machine", and "Gaussian Naive bayes" as applied on "DM++" and "NRC" lexicons is given in Table 5. The accuracy of PoS@token and lemma is compared with a popular word lexicon "NRC".

Table 3 Pearson correlation score between predicted and word lexicon

Emotion	DM++ @Lemma	DM++ @PoS	NRC @token
Fear	0.129	**0.280**	0.047
Anger	0.134	**0.205**	0.085
Joy	0.199	**0.304**	0.228
Sad	0.242	**0.289**	0.141

Table 4 F1-Measure results for emotion classification

Emotion	DM++@Lemma	DM++@PoS	NRC@token
Fear	0.382	**0.339**	**0.339**
Anger	0.419	**0.432**	0.429
Joy	0.496	**0.525**	0.394
Sad	**0.534**	0.453	0.529

The bold value signify the highest F1-Measure

Table 5 Comparison of classification results in terms of accuracy over all emotions, NB, LR, SVM, KNN and GNB using DM++ and NRC word lexicon

Lexicon	Accuracy				
	NB	LR	SVM	KNN	GNB
DM++@lemma	0.26	0.43	0.39	0.53	0.43
DM++@PoS_token	0.30	**0.48**	**0.45**	**0.57**	**0.47**
NRC token	**0.40**	0.40	0.40	0.17	0.39

The bold value signify the highest accuracy

5 Conclusion

Emotion detection is one of the important fields for researchers in various applications. There are several works that have been proposed in emotion detection from audio and facial information. On the other hand, emotion detection from textual information is an interesting and novel research area. Therefore, a lexicon-based emotion detection system is focused to identify emotions from text. In an emotion recognition task, two word-emotion lexicons "NRC" and "Depechemood++" have shown their skills in identifying emotions from ISEAR data set. The classification accuracy was considered to evaluate the performance of five machine learning algorithms like "Naive Baye's", "Logistic Regression", "K-Nearest Neighbours", "Support Vector Machine", and "Gaussian Naive Bayes" classifiers. The experimental results based on the ISEAR corpus indicate that there are some distinct differences between the performances of the "DM++" and "NRC" lexicons. The performance of "NRC" is better in "NB", whereas "Depechemood++" performed better in "LR", "SVM", "KNN", and "GNB" algorithm.

References

1. Abdul-Mageed M, Ungar L (2017) Emonet: fine-grained emotion detection with gated recurrent neural networks. In: Proceedings of the 55th annual meeting of the association for computational linguistics, vol 1: Long papers, pp 718–728
2. Alm CO, Roth D, Sproat R (2005) Emotions from text: machine learning for text-based emotion prediction. In: Proceedings of the conference on human language technology and empirical methods in natural language processing, pp 579–586

3. Raghunathan R, Pham MT (1999) All negative moods are not equal: motivational influences of anxiety and sadness on decision making. Organ Behav Hum Decis Process 79(1):56–77
4. Meo R, Sulis E (2017) Processing affect in social media: a comparison of methods to distinguish emotions in Tweets. ACM Trans Internet Technology (TOIT) 17(1):1–25
5. Mohammad SM, Turney P (2013) Crowdsourcing a word-emotion association lexicon. Comput Intell 29(3):436–465
6. Bollen J, Mao H, Zeng X (2011) Twitter mood predicts the stock market. J Comput Sci 2(1):1–8
7. Mohammad SM, Yang TW (2013) Tracking sentiment in mail: how genders differ on emotional axes. In: Proceedings of the 2nd workshop on computational approaches to subjectivity and sentiment analysis. Association for computational linguistics, pp 70–79
8. Pang B, Lee L (2008) Opinion mining and sentiment analysis. Found Trends Inf Retrieval 2(1–2):1–135
9. Bougie R, Pieters R, Zeelenberg M (2003) Angry customers don't come back, they get back: the experience and behavioral implications of anger and dissatisfaction in services. J Acad Mark Sci 31(4):377–393
10. Plutchik R (1994) The psychology and biology of emotion. HarperCollins College Publishers
11. Ekman P (1992) An argument for basic emotions. Cogn Emotion 6(3–4):169–200
12. Strapparava C, Valitutti A (2004) Wordnet affect: an affective extension of wordnet. In Lrec, vol 4, pp 1083–1086
13. Mohammad SM, Zhu X, Kiritchenko S, Martin J (2015) Sentiment, emotion, purpose, and style in electoral tweets. Inf Process Manag 51(4):480–499
14. Koumpouri A, Mporas I, Megalooikonomou V (2015) Evaluation of four approaches for "Sentiment Analysis on Movie Reviews" The Kaggle competition. In: Proceedings of the 16th international conference on engineering applications of neural networks (INNS), pp 1–5
15. Krishnan H, Elayidom MS, Santhanakrishnan T (2017) Emotion detection of Tweets using Naïve Bayes Classifier. Int J Eng Technol Sci Res 4(11):457–462
16. Strapparava C, Mihalcea R (2008) Learning to identify emotions in text. In: Proceedings of the 2008 ACM symposium on applied computing, pp 1556–1560
17. Li W, Xu H (2014) Text-based emotion classification using emotion cause extraction. Expert Syst Appl 41(4):1742–1749
18. Roberts K, Roach MA, Johnson J, Guthrie J, Harabagiu SM (2012) EmpaTweet: annotating and detecting emotions on Twitter. In: Lrec, vol 12, pp 3806–3813
19. Mike T, Kevan B, Georgios P, Di C, Arvid K (2010) Sentiment in short strength detection informal text. J Am Soc Inf Sci Technol 61(12):2544–2558
20. Luyckx K, Vaassen F, Peersman C, Daelemans W (2012) Fine-grained emotion detection in suicide notes: a thresholding approach to multi-label classification. Biomed Inf Insights 5(Suppl. 1):61–69
21. Staiano J, Guerini M (2014) Depechemood: a lexicon for emotion analysis from crowd-annotated news. In: Proceedings of the 52nd annual meeting of the association for computational linguistics, association for computational linguistics, pp 427–433
22. Araque O, Gatti L, Staiano J, Guerini M (2019) Depechemood++: a bilingual emotion lexicon built through simple yet powerful techniques. IEEE Trans Affect Comput 13(1):496–507