Fuzzy-Rough Set Based Multi-labeled Emotion Intensity Analysis for Sentence, Paragraph and Document

Chu Wang[1], Shi Feng[1,2(✉)], Daling Wang[1,2], and Yifei Zhang[1,2]

[1] School of Information Science and Engineering, Northeastern University, Boston, USA
wangchu@research.neu.edu.cn,
{fengshi,wangdaling,zhangyifei}@ise.neu.edu.cn
[2] Key Laboratory of Medical Image Computing, Northeastern University,
Ministry of Education, Shenyang 110819, People's Republic of China

Abstract. Most existing sentiment analysis methods focus on single-label classification, which means only a exclusive sentiment orientation (negative, positive or neutral) or an emotion state (joy, hate, love, sorrow, anxiety, surprise, anger, or expect) is considered for the given text. However, multiple emotions with different intensity may be coexisting in one document, one paragraph or even in one sentence. In this paper, we propose a fuzzy-rough set based approach to detect the multi-labeled emotions and calculate their corresponding intensities in social media text. Using the proposed fuzzy-rough set method, we can simultaneously model multi emotions and their intensities with sentiment words for a sentence, a paragraph, or a document. Experiments on a well-known blog emotion corpus show that our proposed multi-labeled emotion intensity analysis algorithm outperforms baseline methods by a large margin.

Keywords: Multi-labeled emotion analysis · Emotion intensity · Fuzzy-rough set

1 Introduction

Nowadays, a lot of papers have been published for analyzing sentiments in social media documents, and the previous work concerned with following different facets.

Sentiment Orientation. Most previous researches focused on sentiment orientation classification, i.e. classifying the subjective text into two-orientation (*positive* and *negative*), or three-orientation (*positive*, *neutral*, and *negative*).

Fine-Grained Emotion. Different above two or three categories sentiment, many current researches considered more sentiments such as *joy*, *hate*, *love*, *sorrow*, *anxiety*, *surprise*, *anger*, *expect* [5] called as fine-grained emotion.

Project supported by National Natural Science Foundation of China (61370074, 61402091), the Fundamental Research Funds for the Central Universities of China under Grant N140404012.

J. Li et al. (Eds.): NLPCC 2015, LNAI 9362, pp. 444–452, 2015.
DOI: 10.1007/978-3-319-25207-0_41

Previous work mainly focus on single-labeled sentiment analysis for either sentiment orientation or fine-grained emotion. In fact, multiple emotions may be coexisting in just one sentence, paragraph, or document. For the multi-labeled emotion classification, we propose an improved fuzzy-rough set based approach to detect the multi-labeled emotions and calculate their corresponding intensities in social media text. Experiment results using a well annotated blog emotion dataset show that our proposed algorithm significantly outperforms other baselines by a large margin.

2 Related Work

Multi-labeled emotion analysis can be regarded as multi-label learning problem. Zhang, et al. [8] did a comprehensive review for the multi-label learning algorithms. The typical problem transformation algorithm is called "Binary Relevance", which is proposed by Boutell, et al. [1]. Elisseeff, et al. [3] solved the multi-label data problem by improving the kernel SVM.

Although there are a lot of papers for multi-label learning and sentiment analysis, little work has been done for detecting the multi-labeled emotions, especially the emotion intensities in one social media post.

In this paper, we regard the problem as uncertain emotion classification problem and apply fuzzy-rough theory for solving it. Fuzzy-rough set was first proposed by Dubois and Prade in 1990 which based on Rough set [4] and Fuzzy set theory [7] and combined their advantages together [2]. Now, it has been widely used in many uncertain classification problems. Sun and Ma [6] gave an approach to decision making problem by combining the soft set with fuzzy-rough theory.

Although fuzzy-rough theory has been widely used in many applications, there is no existing literature on uncertain emotion classification for social media.

3 Preliminary

Let SM be a social media text document consisted of m sentences, i.e. $SM=\{s_1, s_2, \ldots, s_m\}$, where s_i (i=1, 2, …, m) be ith sentence and contains n_i sentiment words. The sentiment word set $W_i=\{w_1, w_2, \ldots, w_{n_i}\}$ in s_i expresses a multi-labeled emotion set with different intensity $E_i=\{e_j i_j\}$, where e_j (j=1, 2, ..., 8) represents ith emotion in 8 emotions: *joy*, *hate*, *love*, *sorrow*, *anxiety*, *surprise*, *anger*, and *expect*, respectively, and i_j represents the intensity of e_j in s_i. Our task is to detect every emotion $e_j \in E_i$ and calculate its intensity $i_j \in E_i$ for s_i according to recognized W_i. Moreover, for a paragraph containing multi sentences and a document containing multi paragraphs, our task is to do so for every paragraph and the document.

In this paper we intend to apply fuzzy-rough set method for solving the challenge problem.

Definition 1: U and W are two finite and nonempty universes. Suppose that R is an arbitrary relation from U to W, the triple (U, W, R) is called a generalized fuzzy approximation space. For any set $A \in F(U)$, the upper and lower approximations of A, $\overline{R}(A)$ and $\underline{R}(A)$, are defined respectively as:

$$\overline{R}(A) = \vee_{y \in W}[R(x, y) \wedge A(y)], x \in U \tag{1}$$

$$\underline{R}(A) = \vee_{y \in W}[1 - R(x, y) \vee A(y)], x \in U \tag{2}$$

The pair ($\overline{R}(A)$, $\underline{R}(A)$) is referred to as a generalized fuzzy rough set, and R is referred to as upper and lower generalized fuzzy rough approximation operators.

We regard the emotion between [0, 1] as fuzzy degree, and use fuzzy-rough set to deal with emotion intensity analysis problem. Our plan is to calculate approximations of each sentiment value to the strongest A, and propose an improved algorithm of fuzzy-rough set.

4 Improving Fuzzy-Rough Set for Emotion Intensity Prediction

Suppose the following $E=\{e_1, e_2 \ldots, e_8\}$ is the set of emotions, and $W=\{w_1, w_2, w_3, w_4, w_5\}$ is the set of sentiment words in a sentence. We first mark the intensity as 0 or 1 in Table 1. The analysis methods are as follows.

Table 1. The Examples of Relationship between Multiple Emotions and Different Words

$w\backslash i\backslash e$	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
w_1	0	1	0	1	0	1	0	0
w_2	1	0	1	0	0	1	0	0
w_3	0	0	1	1	1	0	0	0
w_4	1	0	1	0	1	0	0	0
w_5	1	0	0	0	0	1	0	0

We have this following result: $F(e_1)=\{w_2, w_4, w_5\}$, $F(e_2)=\{w_1\}$, $F(e_3)=\{w_2, w_3, w_4\}$, $F(e_4)=\{w_1, w_3\}$, $F(e_5)=\{w_3, w_4\}$, and $F(e_6)=\{w_1, w_2, w_5\}$. This means that three words in this sentence has the emotion *hate*. Similarly, we can describe the features for other emotions in universe U.

Here we introduce eight basic emotion categories from [5]: *joy*, *hate*, *love*, *sorrow*, *anxiety*, *surprise*, *anger* and *expect*. Each word is associated with eight emotions and ten intensities, and most sentences have more than one emotion word in this situation. We need a new algorithm to estimate the multi-labeled emotional attributes of the whole sentence, paragraph, or document.

Definition 2: Let (F^{-1}, W) be a fuzzy set over E, the triple relation (E, W, F^{-1}) is called as the fuzzy approximation space. For any $A \in F(W)$, the upper and lower approximations of A, $\overline{F}(A)$ and $\underline{F}(A)$ with respect to the fuzzy approximation space (E, W, F^{-1}) are fuzzy sets of U, whose membership functions are defined as:

$$\overline{F}(A)(x) = \vee y \in W[(F^{-1}(x)(y)) \wedge A(y)], x \in E \tag{3}$$

$$\underline{F}(A)(x) = \wedge y \in W[(1 - F^{-1}(x)(y)) \vee A(y)], x \in E \tag{4}$$

In order to show the importance of key words, we divide the [0,1] range at 0.5, For any $A \in F(W)$, the upper approximation and lower approximation of A, $\overline{F}(A)$ and $\underline{F}(A)$, if $F(e)(w)=0$ then $\overline{F}(A)=0$ and $\underline{F}(A)=0$, if $F(e)(w)\neq 0$ then

$$\overline{F}(A)(e) = \vee_{w \in W}[F(e)(w) \wedge A(w)], e \in E \tag{5}$$

$$\underline{F}(A)(e) = \begin{cases} \wedge_{w \in W}[(1 - F(e)(w)) \vee A(w)], & e \in E, F(e)(w) \in [0.5,1] \\ \wedge_{w \in W}[F(e)(w) \wedge A(w)], & e \in E, F(e)(w) \in [0, 0.5) \end{cases} \tag{6}$$

In sentiment analysis problem defined in this paper, due to the fuzzy mathematics, all the emotional intensities of words are marked between 0 and 1. Here A is very important, and we choose the strongest emotion intensity of each sentiment word as the value of A. Because it is the highest value of each emotion, if other words is more approximate to A, the stronger of the emotion intensity of the sentiment words is. So we only need to calculate the approximation of each word. Then we construct the decision object A on the evaluation of the words universe W.

Take one sentence for example: “我在6点之前就睡不着了，因为激动，马上就能领略九寨沟的秀丽风光了”. The emotional value mark of this sentence is: *joy*=0.8, *love*=0.5 and *expect*=0.5. We calculate $\overline{F}(A)$ and $\underline{F}(A)$ with Formula (5) and (6), the calculated value of our algorithm is the same order as the result annotated by human which is showed in Table 2.

Table 2. The Examples of Weight Computation Based on Our Improving Fuzzy-Rough Set

	激动	领略	秀丽	风光	睡	$\overline{F}(A)$	$\underline{F}(A)$	*value*
Joy	0.6	0.6	0	0	0.5	0.6	0.5	1.1
Hate	0	0	0	0	0	0	0	0
Love	0	0.6	0.6	0.4	0	0.6	0.4	1
Sorrow	0	0	0	0	0	0	0	0
Anxiety	0	0	0	0	0	0	0	0
Surprise	0	0	0	0	0	0	0	0
Anger	0	0	0	0	0	0	0	0
Expect	0.5	0	0	0	0.5	0.5	0.5	1
A	0.6	0.6	0.6	0.4	0.5			

In sentiment analysis, we always focus on the extraction of sentiment words, such as adjectives/adverbs, verbs used to qualify nouns. In Table 2, there are eight emotions, and each sentiment word may have more than one basic emotions. The value of word is between 0 and 1, and 0 means the words do not have the emotional attribute. As the value growing approach to 1, the intensity of the emotion of the word is getting bigger. This feature of the value applies to fuzzy set theory, so we can use our improving fuzzy-rough set to do our work.

Our improving method has several advantages. Firstly, it is suitable to calculate the intensity of the text which between [0, 1]. Secondly, it considers the weight of the emotion value not only from the value of the words but also the influence of the whole words. Thirdly, it focuses on a human logic that emotion strength is not simply the sum. For example, three weak sentiment words usually neither form a powerful sentiment value nor get a weaker one.

As we explained above, our method is suitable to solve this problem. In Table 2, we get the result in [0, 2] (e.g. 1.1), but the emotion value of the sentence annotated by people is between [0, 1]. So we still need to build a regression algorithm to make our predictive method completed. In this paper, we choose Linear Regression Algorithm to approach our target. In summary, we give the algorithm for Intensity Prediction based on our improving fuzzy-rough set model as follows.

Algorithm 1. Intensity Prediction Algorithm

Input: fuzzy set (F, W) // W is the sentiment words set, and F is the fuzzy relation
Output: emotion intensity
Description:

1. Read every sentiment word $w \in W$;
2. Compute strongest emotional intensity object A as:
$$A=\sum_{i=1}^{W}\frac{maxF(w_i)}{w_i}, w_i \in W, i.e., A(w_i)=max\{F(e_j)(w_i)\big| e_j \in E\}$$
3. Calculate the improving fuzzy-rough upper approximation $\overline{F}(A)$ and fuzzy-rough lower approximation $\underline{F}(A)$ // see formula (5) and (6);
4. Calculate the choice value $m=\overline{F}(A)(e_i)+\underline{F}(A)(e_i)$, $e_i \in E$;
// E is the universe of the emotion
5. Normalize m into [0, 1] with Linear Regression Algorithm;
6. Return m as emotion intensity;

5 Experiments

5.1 Dataset and Evaluation Metric

We use Changqin Quan's [5] dataset to evaluate our proposed method. The corpus contains 1,487 documents, with 11,953 paragraphs, 38,051 sentences, and 971,628 Chinese words. And each of them is annotated with 8 kinds of emotion [5]. We use cross-validation to validate our algorithm. The whole dataset is divided into 5 parts on average, and each time there are four training sets and one testing set.

In this experiment, we use 4 evaluation metrics based on [8] to test our method. Let $S = \{(x_i, Y_i)| 1 \le i \le p)\}$ be the test set and h be the learned multi-label classifier.

Subset Accuracy: $$subsetacc_s(h)=\frac{1}{p}\sum_{i=1}^{p}[| h(x_i=Y_i) |] \quad (7)$$

The subset accuracy evaluates the fraction of correctly classified examples, i.e. the predicted label set is identical to the ground-truth label set. It tends to be overly strict especially when the size of label space is large.

Hamming Loss: $$hloss_s(h) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{q}|h(x_i)\Delta Y_i| \quad (8)$$

where Δ stands for the symmetric difference between two sets. The hamming loss evaluates the fraction of misclassified instance-label pairs, i.e. a relevant label is missed or an irrelevant is predicted. In this paper, we need to improve the evaluation metric. In the formula below, we try to measure the difference between the value we predict and the real one:

$$hloss_s(h) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{q}|h(x_i) - \Delta Y_i| \quad (9)$$

One-error: $$One-error_s(h) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{q}[[\arg\max_{y\in Y} f(x_i, y)] \notin Y_i] \quad (10)$$

The one-error evaluates the fraction of examples whose top-ranked label is not in the relevant label set. In our paper, it means the strongest emotion of the eight we predict is wrong. As we can see, the value of the One-error is the fewer, the better.

Average Precision

$$avgprec_s(h) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{|Y_i|}\sum_{y\in Y_i}\frac{|\{y' \mid rank_f(x_i, y') \le rank_f(x_i, y), y' \in Y_i\}|}{rank_f(x_i, y)} \quad (11)$$

The average precision evaluates the average fraction of relevant labels ranked higher than a particular label $y\in Y_i$. In this paper, it means whether the descending order of the value of each emotion is right. This metric is the larger, the better.

5.2 Experiment Results

As few methods were proposed for multi-labeled emotion intensity, we compare our method with the classic Fuzzy union one. It is defined as: $(A\cup B)(x)=\max(A(x), B(x))$ for all $x\in X$. Taking the value of *joy* from Table 2 for example: *Joy*(Fuzzy union)=max(激动(*joy*), 领略(*joy*), 睡(*joy*))=max(0.6, 0.6, 0.5)=0.6.

In our experiment, it means each kind of the 8 emotions takes its strongest value of all words. And we have also introduced a Naïve Bayes method to compare with our algorithm. In NB method, we assumed that every emotion is independent from each other. Taking the value of *joy* from Table 2 for example: *Joy*(Naïve Bayes)=*P*(激动|*joy*)**P*(领略|*joy*)**P*(睡|*joy*)**P*(*joy*)=6/17*6/17*5/17*17/43=0.01448.

Firstly, we evaluate the label prediction accuracy with Subset Accuracy. Using our method, the percent of emotions figured out in the paper is showed in Table 3.

Table 3. Label Prediction Accuracy

	Document	Paragraph	Sentence
Subset Accuracy	99.76%	98.61%	96.38%

Next we compare our multi-labeled method with two algorithms, the first one is multi-label learning method (with no label intensity), and the other one is Fuzzy union. As we all know, in multi-label learning setting, the emotion value of any sentiment word is 0 or 1. In the fuzzy mathematics, fuzzy union is a very commonly used method. In this experiment it means that we use the max value of each emotion in every sentiment words in the sentence. We take an average result values of 5 cross validation, and the results at three different textual levels are shown in Table 4.

Table 4. Emotion Label Intensity Analysis of Hamming Loss

	Hloss (single-labeled)	*Hloss* (Fuzzy union)	*Hloss* (NB)	*Hloss* (ours)
Document	0.09845	0.06853	0.21445	**0.03929**
Paragraph	0.10659	0.05117	0.15133	**0.02598**
Sentence	0.11179	0.03927	0.10320	**0.01767**

Hamming Loss is a measure of the value of general accuracy. *Hloss* is the fewer, the better. So as we can see, our method is better than the others.

One-error is a method to measure whether the maximum value we predict is one of the final results, and the results we get are shown as Table 5. Average precision is widely used in many areas. According to the measure of Average precision [8], the results of the experiments at three levels of text are showed in Table 6.

Table 5. Label Intensity Analysis of One-error

	One-error (Fuzzy union)	*One-error* (NB)	*One-error* (ours)
Document	0.78649	0.76689	**0.04764**
Paragraph	0.58395	0.58807	**0.06903**
Sentence	0.84421	0.37633	**0.07377**

Table 6. Label Intensity Analysis of Average precision

	Avgprec (Fuzzy union)	*Avgprec* (NB)	*Avgprec* (ours)
Document	0.67249	0.36975	**0.74374**
Paragraph	0.71452	0.44310	**0.77940**
Sentence	0.70985	0.58617	**0.95494**

According to Formula (10) and (11), one-error is the fewer, the better. In contrast, average precision is the larger, the better. So we can see our algorithm is significantly better than others.

5.3 Discussion

Our method has several advantages: Firstly, the traditional multi-label learning methods only consider about whether the sentiment words in this sentence have this emotion. It cannot show the emotion intensity. The fuzzy union only consider the strongest one emotion, which means it has to ignore some most of sentiment words. This strategy is less objective. In Naïve Bayes method, we presume that 8 emotions are independent, but the human emotional logic is not completely independent or simply multiplied. Secondly, our method not only considers every emotion value in each sentiment word, but also improves fuzzy-rough logic to fix them together. Thirdly, using the method, we both considered the value of the sentiment and the importance of the words. Specifically, although the emotion value in one kind of emotion may be the same, but with the influence of other kinds of emotion which belonging to different words, the sentiment value of the test result may be different.

On the other hand, the fuzzy logic is more suitable for this situation not only because the value of emotion area is between 0~1, but also considered the human language logic. As we know, a sentence may have some key emotion words, and the common feature of these key words is that their emotion intensities are always the largest. Maybe there are many other sentiment words that have different intensities, but finally, the key words dominate the emotions. Taking this sentence for example: “我游览了沈阳，中街的热闹，世园会的恬静，怪坡的奇妙使我爱上了这座城市”. In this sentence, these three Chinese words “热闹”, “恬静”, “奇妙” all get sentiment of *joy*, the value of them are *joy*=0.5, *joy*=0.3, *joy*=0.4. But the key word is “爱”, whose value is *love*=0.9. Although *joy* has larger sum of words and emotion value, we can still see that the main emotion of this sentence is *love*. This is the human logic. It is neither the sum of words nor the sentiment value. Our improved method inherits the advantage of fuzzy logic, which perfectly matches the emotion logic in this situation.

6 Conclusion and Future Work

In this paper, we propose a new way to solve multi-label emotions intensity analysis problem. For this new field of emotion analysis, we leverage the advanced fuzzy-rough set theory. The proposed algorithm could simultaneously detect multi-labeled emotions and their intensity at sentence, paragraph, and document level. The experiment results demonstrate the effectiveness of our proposed model and algorithm.

In the future, we would like to build the model of multi-labeled emotion and intensity analysis on microblog, website reviews, or other social media. The adverbs and negative words can be further taken into consideration that can improve the result.

References

1. Boutell, R., Luo, J., Shen, X., Brown, C.: Learning multi-label scene classification. Pattern Recognition (PR) **37**(9), 1757–1771 (2004)
2. Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. International Journal of General Systems **17**(2–3), 191–209 (1990)
3. Elisseeff, A., Weston, J.: A kernel method for multi-labelled classification. In: NIPS, pp. 681–687 (2001)
4. Pawlak, Z.: Rough sets. International Journal of Parallel Programming (IJPP) **11**(5), 341–356 (1982)
5. Quan, C., Ren, F.: A blog emotion corpus for emotional expression analysis in Chinese. Computer Speech & Language (CSL) **24**(4), 726–749 (2010)
6. Sun, B., Ma, W.: Soft fuzzy rough sets and its application in decision making. Artif. Intell. Rev. (AIR) **41**(1), 67–80 (2014)
7. Zadeh, L.: Fuzzy sets information and control. IANDC **8**(3), 338–353 (1965)
8. Zhang, M., Zhou, Z.: A review on multi-label learning algorithms. IEEE Trans. Knowl. Data Eng. (TKDE) **26**(8), 1819–1837 (2014)