2-2020

Multi-level fine-scaled sentiment sensing with ambivalence handling

Zhaoxia WANG
Singapore Management University, zxwang@smu.edu.sg

Seng-Beng HO

Erik CAMBRIA

International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems

Multi-Level Fine-Scaled Sentiment Sensing with Ambivalence Handling

Zhaoxia Wang[*,§], Seng-Beng Ho[†,¶] and Erik Cambria[‡,‖]

**School of Information Systems,*
Singapore Management University, 81 Victoria Street, Singapore 188065
*†Institute of High Performance Computing, A*STAR,*
1 Fusionopolis Way, Singapore 138632
‡School of Computer Science and Engineering,
Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798
§zxwang@smu.edu.sg
¶hosb@ihpc.a-star.edu.sg
‖cambria@ntu.edu.sg

Received 6 September 2018
Revised 15 May 2019

Social media represent a rich source of information, such as critiques, feedback, and other opinions posted online by Internet users. Such information is typically a good reflection of users' sentiments and attitudes towards various services, topics, or products. Sentiment analysis has become an increasingly important natural language processing (NLP) task to help users make sense of what is happening in the Internet blogosphere and it can be useful for companies as well as public organizations. However, most existing sentiment analysis techniques are only able to analyze data at the aggregate level, merely providing a binary classification (positive vs. negative), and are not able to generate finer characterizations of sentiments as well as emotions involved. This paper describes a new opinion analysis scheme, i.e., a multi-level fine-scaled sentiment sensing with ambivalence handling. The ambivalence handler is presented in detail along with the strength-level tune parameters for analyzing the strength and the fine-scale of both positive and negative sentiments. It is capable of drilling deeper into text in order to reveal multi-level fine-scaled sentiments as well as different types of emotions.

Keywords: Ambivalence sentiment handling; emotion sensing; multi-level fine-scaled sentiment analysis, sentiment strength level; social media analysis

1. Introduction

Social media, such as Twitter, Facebook and Chinese Weibo, are accessed widely by Internet users for a variety of purposes, such as sharing their comments or experiences towards certain products, services or policies. Therefore, the analysis of such social media data can provide opportunities for those who are eager to gauge public opinion about their products or services.

Sentiment analysis of social media text refers to the use of natural language processing technologies to identify or study affective states, subjective information or attitude hidden in the social media text. Good social media sentiment analysis can motivate various corporations to poll timely opinions. Business corporations are eager to understand the market preferences for their products to improve their market share. Consumers would like to use online reviews to help them make better purchase decisions. Similarly, politicians would like to respond accurately to public views of their policies. Consequently, sentiment research has gained much attention in recent years[1] and have been applied in different areas.[2,3]

Sentiment analysis is a branch of affective computing research that aims to classify text (but sometimes also audio and video) according to the conveyed emotions or polarity.[4] Most of the literature is on English language but recently an increasing number of publications is tackling the multilinguality issue.[5] Most commercial off-the-shelf (COTS) sentiment analysis engines are only able to provide the analyzed sentiment polarity at the aggregate level, e.g., positive, negative, or neutral. Some of them even consider sentiment analysis as a mere binary classification problem (positive vs. negative).

Compared to the aggregate-level sentiment analysis, fine-grained sentiment analysis can yield more specific fine-grained results, characterizing emotions into finer subcategories such as anxiety, sadness, and anger for negative emotion, and excitement and happiness for positive emotions.[6] For example, in the text '*What a nice phone, I really happy to have one", the* happiness emotion is sensed and not merely a positive one. In another example, *"Holy shit such brand phone looks stupid even in the TV commercials, I am really angry about it!"*, the angry emotion is expressed along with a sense of disappointment.

Emotion sensing aims to extract a set of more precise emotions within the broad class of positive or negative sentiments. Cambria et al.[7] have presented a concept-level knowledge based for sentiment analysis, named SenticNet, to help understand emotions in informal communication texts. For example, SenticNet classifies positive sentiments into different emotion levels, e.g., ecstasy, joy, trust, admiration, etc. Similarly, negative sentiments are classified into as many levels, e.g., sadness, fear, anger, etc. Also, polarity in SenticNet is not just a binary label (positive vs. negative), but rather a floating number that spans from -1 (extreme negativity) to +1 (extreme positivity), passing through 0 (neutral).

There has been a fair amount of research works related to sentiment analysis and emotion detection,[6-11] and some of them considered the strength on a numeric scale, like SenticNet. In this paper, a new scheme of multi-level fine-scaled sentiment classification with ambivalence handling is described, in which the ambivalence handler method is presented. The strength-level tuning parameters along with k bands of sentiment strengths are described for analyzing the detailed scales of the positive or negative sentiments.

The rest of the paper is organized as follows: Section II analyzes common algorithms for sentiment analysis, emotion models and existing emotion sensing technologies; Section III presents the proposed methodology of multi-level fine-scaled sentiment analysis with ambivalence handling; Section IV shows the performance of the proposed method; finally, Section V concludes the paper and proposes possible improvements for future work.

2. Discussion on Existing Sentiment Analysis, Emotion Models and Emotion Sensing Technologies

Sentiment analysis is one of the hottest areas of research in social analytics.[12-14] Typically the various methods of sentiment analysis is broadly categorized into two types: the learning-based methods (e.g., deep learning) and non-learning based methods (e.g., lexical-based methods).[14]

In learning-based methods, known properties derived from labelled training data are used to make predictions about the properties of new data.[12] For the analysis of text data, the relationship between different components of the text segment is derived. A wide variety of learning-based methods, such as Maximum Entropy classifiers,[15] support vector machines,[16,17] Naïve Bayes classifiers,[18,19] and extreme learning machines,[20,21] have been used for sentiment analysis.[21,22]

To be effective and achieve an acceptable accuracy for classification, the learning-based methods typically require a sufficiently large labelled training dataset.[23,24] However, due to the diversity of the contents in typical social media discussions, it is difficult to know *a priori* an adequate size for the training dataset,[6,16] and the labelling task can be costly or even prohibitive.[6,14,16]

In contrast with learning-based methods, the non-learning based methods, such as lexical-based methods do not suffer from this shortcoming of dependency on training data sets.[12-14] It derives the sentiment polarity of a text according to the sentiment or emotion indicators which are lexicons used in the text. D. Mohey and E. M. Hussein delivered a detailed survey on sentiment analysis challenges through analyzing the relationship between the sentiment analysis and the three format sentiments: structured sentiments, semi-structured sentiments and unstructured sentiments.[13] Their results revealed Part-of-speech (POS) tagging and lexicon-based techniques were still the popular approaches.[13] Gonçalves et al. compared eight common lexical-based methods (LIWC, SenticNet, SentiWordNet, PANAS-t, SASA, Happiness Index, Emotions, and SentiStrength) with the aim to find out the most effective method.[14] The study indicated that SentiWordNet possesses the largest coverage, i.e., in terms of the fraction of messages whose sentiment is identified, while LIWC has the highest agreement, i.e., in terms of the fraction of identified sentiments that agrees with the ground truth.[14] However, both SentiWordNet and LIWC methods do not measure the strengths of sentiments in fine scales.

Even though the issue of labelled dataset is not of concern for lexical-based methods, the challenge for these methods is in the creation of dictionaries: How to create specific dictionaries adequate for handling the processing in different domains. Sentiment analysis is closely associated with emotion theories. Generally speaking, sentiment analysis aims to detect the attitude or emotions of a user when they are communicating in certain topics or domains.[6,7] Wang et al. proposed new methods in which various domain knowledge bases as well as topic dictionaries were built to address topic and domain-specific adaption.[6,7]

Their developed methods had the capabilities to derive dominant sentiment valence as well as basic emotions. Morente-Molinera et al. recently presented a new method that leveraged sentiment analysis to create fuzzy ontology.[25] The method proposed can help the computational system to better process the opinion texts with the application of sentiment analysis. The other model proposed by the same author is to leverage sentiment analysis procedures to extract the preference values from the free text used for debate, which can help the group members generate the most convincing decisions with the consideration of everyone's opinion.[26]

Shaver et al.'s approach to emotion analysis assumes that emotions can be grouped into prototypes and that a "whole" emotion is made up of various "emotion parts".[27] In their experiment, they first selected a group of words and had them rated based on whether each word was an emotion. This step resulted in a list of 135 emotion words. An abstract-to-concrete emotion is then developed through a typical prototyping approach, and they discovered six basic emotions on the hierarchy's lowest level: *joy, love, surprise, sadness, anger* and *fear.*

Psychologists Ortony and Turner do not think it is meaningful to treatbasic emotions as psychologically primitive.[28] They proposed that there is a hierarchical structure that organizes all emotions, and each of these emotion is discrete and independent from the others. Hence, according to their view, there is no basic set of emotions that serve as the constituents of others. Ekman stated that there are 6 basic emotions – anger, fear, disgust, joy, sadness and surprise.[29] The idea that there are distinctive facial expressions forms the basis of Ekman's emotion model, and emotions are characterized as discrete, measurable, and physiologically distinct. The fact that in his model emotions are treated as families of related states means that it is consistent with Shaver's model [27]. Ghazi et al. also made use of the model proposed by Ekman to distinguish automatically between prior and contextual emotion words in the context of sentences.[30]

Building on Ekman's biologically oriented view of emotion, Plutchik proposed the idea of "wheel of emotions".[31] A wheel-like diagram of emotions is used to visualize eight basic emotions. These eight primary emotions are grouped into the dimension of positive vs negative basis, e.g., joy versus sadness; anger versus fear; trust versus disgust; and surprise versus anticipation,[31,32] and these are placed on opposite sides of the wheel. This model states that complex emotions are a composition of several basic emotions, and the main idea is consistent with Shaver's model. However, some of the basic emotions defined are different from those of Shaver's.

Suttles et al. opted to use Plutchik's model over Ekman's, as they felt that the latter focuses more on negative emotions.[33] The four sets of basic bipolar emotions from the eight basic bipolar emotions defined by Plutchik, allow emotion classification to be treated as a binary classification problem, unlike in the case of Ekman's model.[33] On the other hand, Alena et al. leveraged and enhanced on the various above emotion models and incorporated them into a typical lexical approach.[34] They proposed nine basic emotions, i.e., anger, disgust, fear, guilt, interest, joy, sadness, shame and surprise. These nine emotion words were annotated by expert annotators and compiled into an emotion dictionary.[34]

For implementing emotion sensing technologies, researchers can leverage on the above emotion definitions and select different sets of emotions. Also, more specific emotions such as inspired, keen, and hopeless that are not listed and discussed above can be added for the emotion sensing technologies.

According to the above discussion on sentiments and emotions, we can have various kinds of emotions categorized into different basic emotion groups and researchers from different domains can have their unique enhancements of the emotion models to fine tune their intended outcomes.

The review presented in this section reveals a need for research into fine-grained sentiment analysis. The issues such as the handling of the ambivalence sentiment should be considered. Definition of "ambivalence" by Merriam-Webster: simultaneous and contradictory attitudes or feelings (such as attraction and repulsion) toward an object, person, or action. Ambivalence sentiments refer to attitudes or comments towards something or someone that contain both positively and negatively valenced components.[35] Ambivalent sentiment is pervasive, especially in the comments found in various online media, which often include a mixture of positive and negative comments, even though the person posting the comments would like to express just a positive or negative sentiment. However, there were very few publications which discuss ambivalent sentiment. In addition, the definition of the fine-grained sentiment, sentiment strength, and the selection of certain emotions for business use may also need to be considered. We propose to address these in this paper.

3. Proposed Multi-Level Fine-Scaled Sentiment Analysis with Ambivalence Handling

Sentiments and emotions are closely related.[6,7] A sentence with negative sentiment can contain a variety of emotions, such as anger and sadness. In our approach, the classification process is performed in two phases: (1) recognizing the sentence's sentiments with ambivalent handling, and (2) finely identifying the multi-level fine-scaled sentiments as well as the specific emotions involved. For an item of text data including multiple sentences, sensing analysis is performed on individual sentences, and then the analysis is carried out at the paragraph and article levels through one of two "sum" methods. One is simply to count the number of positive and/or negative sentences; the other method is to leverage on the fuzzy sum based on the adaptive fuzzy inference algorithm.[6,7]

3.1. *Ambivalent Sentiment Handling*

The pseudocode to implement sentiment analysis method without considering ambivalence sentiment handling is shown in Algorithm 1 below. It is a simplified implementation version of the previous work.[6,7]

Algorithm 1. Simplified Implementation of Sentiment Analysis

```
"BEGIN
     Read(configuration file)
     Read(lexicons)
     InputText = Read(all input text)
     FOR each inputText:
          cleanText = Cleanup(inputText)
          Opinions = DecomposeIntoOpinions(cleanText)
          FOR each opinion:
               CheckIfQuestion(opinion)
               CheckForExceptionalWords(opinion)
               CheckForMixedEmotions(opinion)
               CheckForNegation(opinion)
               Words = GetAllWords(opinion)
               FOR each word:
                    CalculateScore(word)
                    CombineWithNegation(word)
                    UpdateScore(+veScore, -veScore)
               END FOR LOOP
               AggregateScores(+veScore, veScore)
          END FOR LOOP
          AggregateScores(+veScore, negativeScore)
          Derive sentiment category by examining +veScore
               and -veScore

          OutputResult()
     END FOR LOOP
END
```

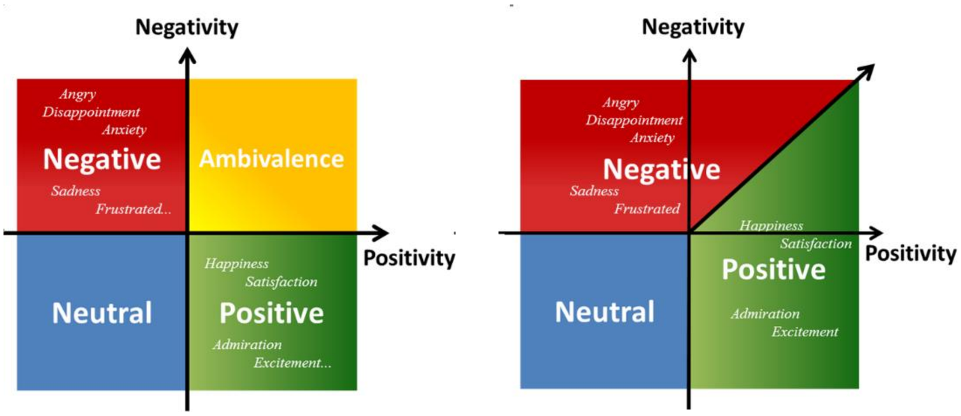
P/S: There are 6 sentiment categories in the outcomes:
Neutral: if +veScore and –veScore are both 0
Postiive: if +veScore is > 0 and –veScore is 0
Negative: +veScore is 0 and –veScore is > 0
Mixed-Neutral: if +veScore equals to –veScore ,and not 0
Mixed-Positive: if +veScore > –veScore and –veScore is not 0

Mixed-Negative: if –veScore > +veScore and +veScore is not 0

Table 1 shows the output of Algorithm 1, which includes the six sentiment categories: Positive, Negative, Neutral, Mixed-Neutral, Mixed-Negative, and Mixed-Positive, which may be collapsed into 4 sentiment categories: Positive, Negative, Neutral and Ambivalence. Ambivalence handling can then be performed to further reduce the 4 sentiment categories into 3 categories as shown in Fig. 1.

Table 1. Sentiments with Ambivalence Catergory

Sentiment definition as well as categories		Meaning
6 Categories	4 Categories	
Neutral	Neutral	Neither positive nor negative
Negative	Negative	Contains only negative sentiments
Positive	Positive	Contains only positive sentiments
Mixed-Negative		Contains both positive and negative sentiments, but with a stronger weightage of negative sentiments
Mixed-Positive	Ambivalence	Contains both positive and negative sentiments, but with a stronger weightage of positive sentiments
Mixed-Neutral		Contains both positive and negative sentiments with equal weightage of each



(1) The 4 categories (2) The three categories

Fig. 1. Illustration of ambivalence handling.

The transformation in Fig. 1 is based on the transferring rules shown in Table 2. The final sentiment category is based on the sentiment weightage of valence and this is to be calibrated using a survey. Each text item with ambivalent sentiment is further classified into positive, negative or neutral categories based on the transferring rules designed as well as by 12 volunteers. For example, the text “This brand phone is expensive, but I still want to buy it since it is really good” has been classified as positive sentiment based on the transferring rules and it is also classified as positive sentiment by all the 12 volunteers. The results obtained using transferring rules is consistent with the reasoning results obtained by human volunteers as shown in Table 2.

Table 2. Rule for Handling with Ambivalent Text

Ambivalence handler according to the rules			**Survey result**
Ambivalence outputs	Final outputs	Example	(N(Pos), N(Neg), N(E))
Mixed-Positive, with a stronger weightage of positive	Positive	This brand phone is expensive, but I still want to buy it since it is really good.	(12,0,0)
Mixed-Negative, with a stronger weightage of negative	Negative	This brand phone is good, but I do not want to buy one since it is so expensive.	(0,11,1)
Mixed-Neutral, with equal weightage of each	Positive, Negative or Neutral**	This brand phone is really good. I do not buy it because it is expensive.	(2,1,9)

**Survey rules: It is assuming that the ambivalent outputs can be forced into positive, negative or neutral according to the weightage of positive or negative.

Rules can be easily designed to force the Mixed-Positive and Mixed-Negative into Positive or Negative respectively as shown in Table 2. For Mixed-Neutral, according to the survey results, it can be forced into Positive, Negative or Neutral according to the exact requirement the user wants to focus on.

3.2. *Multi-Level Fine-Scaled Sentiment Analysis*

Good fine-grained sentiment analysis will need to consider the sentiment scales. It is easy for human-beings to understand that "Very good" represents stronger positive sentiment than "Good"; and that "Best" represents stronger positive sentiment than "Very good". Therefore, to mimic what human beings can do for fine-grained sentiment analysis, we designed an advanced linguistic processing method by proposing the strength-level tuning parameters, A, B, C and D, to modify the sentiment strengths, varying from strongest, stronger, baseline to below-baseline:

- Strongest Strength-level Tuning Parameter A (the Highest Enhancer Parameter A): enhancer, or amplifier key indicators that include "most", "surprisingly", "extremely", "super", "stunningly", and others.
- Stronger Strength-level Tuning Parameter B (the Comparison Enhancer Parameter B): enhancer, or amplifier key indicators that include "pretty", "very", and others.
- Baseline Strength-level Tuning Parameter C (Common Parameter C): No enhancers, or amplifiers found in the text, and no reducers or diminishers found in the text.
- Below-baseline Strength-level Tuning Parameter D (Reducer Parameter D): reducer or diminisher key indicators that include "minor", "mini" and others.

According to the description above, the Strength-level Tuning Parameters, ***A***, ***B***, ***C*** and ***D***, are designed to modify the sentiment strengths, varying from strongest, stronger, baseline to below-baseline. The normalized values of the 4 Strength-level Tuning Parameters are A, B, C and D, satisfying:

$$0<D<C<B<A\leq 1 \tag{1}$$

The modification is applied based on the presence or absence of a prescribed set of keywords. We use examples shown in Table 3 below to illustrate how the 4 strength-level tuning parameters work.

Table 3. Illustrations of how the 4 strength-level tuning parameters works.

Items of text data and initial sentiment score $p(s)$ before tuning*		The sentiment score $p_f(s)$ after tuning by the strength-level tuning parameters		
Item	$p(s)$	Key indicators	ts	$p_f(s)=p(s).ts$
…It is slightly positive …	1	slightly	ts=D=0.25	0.25
…it is positive…	1	no key indicator found in the text	ts=C=0.50	0.5
…it is very positive…	1	very	ts=B=0.75	0.75
…it is super positive…	1	super	ts=A=1.00	1.0

*Assume the initial sentiment score are coarse-grained, positive sentiment score is 1, negative sentiment score is -1, and neutral is 0.

After the strength-level tuning parameter is performed on the initial sentiment scores p(s), the new sentiment score pf(s) reflects the strength of the sentiment, but the sentiment polarities do not change as shown in Table 3 above.

Misspelling and emotionally exaggerated words in the sentence have also been handled as described in the previous work [6 7]. For example, the words, gooooood", and "gooood" will be treated as "very good", a strengthened form, rather than "good".

Since individual words or phrases with a negation prefix, e.g., "not" before "bad", imply a negation of the sentiment, linguistic patterns are also introduced in the matching process to rectify the sentiments. Similarly, such patterns are used to rectify the sentiment strength of the text, in the cases where peculiar words preceding or proceeding from the sentiment phrases are discovered. For example, "fucking good" actually implies a strengthened form of "good".

In the previous research [36], for a target word or sentence, ***s,*** the obtained polarity score ***p***(s), which represents the strength of the sentiment, will satisfy the condition [36]:

$$p(s) \in [-1, 1] \tag{2}$$

where values are discrete: -1 is for negative, 0 is for neutral and +1 is for positive.

The modified sentiment score, p_f(s) of the word or sentence *s*, is obtained by applying the strength parameter $t_s \in (0, 1]$ on the polarity score, i.e.,

$$pf\,(s) = p(s)\ ts \tag{3}$$

where values of $p_f(s)$ is still bounded in [-1,1], since $p(s) \in [-1, 1]$ and $t_s \in (0, 1]$, then $p_f(s) \in [-1, 1]$.

Supposing the positive and negative regions can each be divided into *k* bands of sentiment scales, there will be *2k* +1 classifications, which include the neutral class of *p* (*s*)=0. Fig. 2 shows examples of the multi-level fine-scaled levels for *k*=3, 5 and 10. Qualitative measures may be assigned to the different bands, which may not be equal in size, according to the requirements of the industry to give a more intuitive representation of the sentiment.

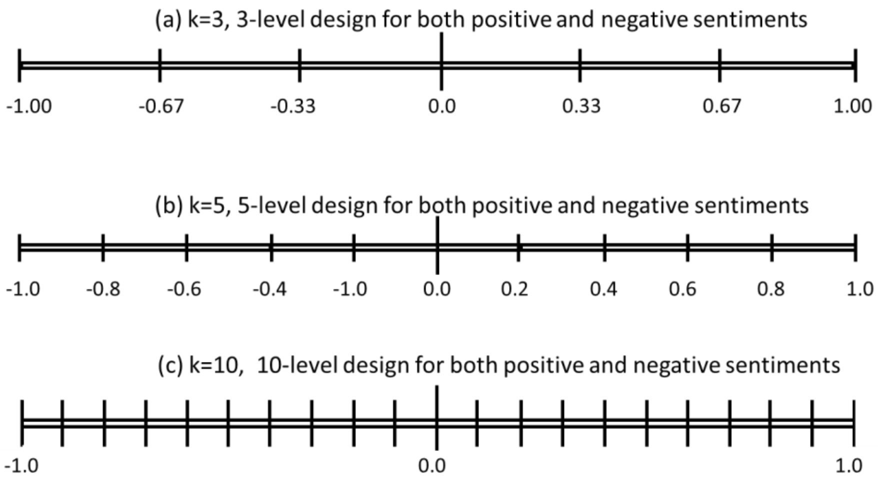


Fig. 2. Illustration of fine-grained sentiment levels.

3.3. *Emotion Sensing*

The previous research does not identify specific emotions. In order to assess the emotions within the sentences, a set of prescribed emotion categories need to be defined. It is best to define these basic categories according to the domain or industry on which the sentiment analysis is performed. From these basic categories, there will be many words or phrases that fit each of the basic emotion categories. Depending on the domain and industrial usage, the list of words and phrases are identified to form an emotion dictionary that we built.

The words or phrases in the text to be analyzed will then be matched against the emotion dictionary from which the basic emotion will be identified. This is applied to all text with either positive or negative sentiment. Similarly, linguistic patterns are also leveraged in the matching process to rectify the emotion outputs of the text.

The level of sentiment and emotion resolution can be refined by increasing the number of sentiment bands as well as increasing the number of basic emotion categories and correspondingly modifying the emotion dictionary. The final outcome of any examined text is the modified sentiment score and the set of discovered basic emotions.

4. Performance Testing with Discussion

In this section, we discuss the testing of our algorithms on a set of social comments which is downloaded from the website: http://www.glassdoor.com/. It provides employee reviews on various companies. 5000 entries were collected for each of the two sentiment categories – positive (Pros) and negative (Cons).

Since the data downloaded are only labelled with two categories of positive or negative sentiment, we applied the proposed ambivalence handler to handle the ambivalent sentiments as discussed in Section 3-A. In order to keep the companies' names as well as the analysis results confidential to respect their privacy, we do not mention the names of the companies involved but only used the data to test the performance of our sensing technology. The results with ambivalence and without ambivalence are shown in Table 4.

As shown in Table 4, even without the ambivalence handler, the proposed method, with accuracy of 78.56%, is already significantly outperforming the Stanford NLP method, which has accuracy 69.36%. All the performance measures, i.e., accuracy, precision, recall and F1 of the proposed method, are higher than those of Stanford NLP. With the ambivalence handler, the performance increased from 78.56% to 82.61%. This case study results demonstrate the merit of the proposed methods.

To business corporations, multi-level fine-scaled sentiment and emotion sensing can yield a deeper insight into their business performance. Knowledge of the scales of sentiments and detailed emotions of customers will facilitate a targeted response to handle a complaint well.

To our knowledge, this paper is the first work to propose ambivalence handling method for sentiment analysis. Ambivalence sentiment is pervasive, especially in the comments found in various online media, which often include a mixture of positive and negative comments, even though the person posting the comments would like to express just a positive or negative sentiment. This is one major reason why our proposed ambivalence handling system is significant.

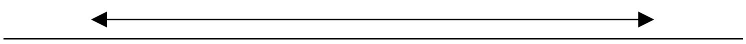
Table 4. Performance for analyzing social comments

Method	Proposed method		Stanford NLP
	With ambivalence handling	**Without ambivalence handling	
Accuracy %	82.61	78.56	69.60
Precision %	84.63	79.47	69.62
Recall %	82.61	78.56	69.60
F1 %	82.35	78.39	69.59

** Without ambivalence handling the ambivalence output are treated as false classifications.

In addition, comparing the results with the ambivalence handling and without the ambivalence handling, both the 4 matrix Accuracy, Precision, Recall and F1 are increased. This demonstrated that ambivalence handling enhances the performance of the system.

For multi-level sensing, we use the sentiment scores in Table 3 above as test cases to demonstrate the possible results. We select 4 level outputs as shown in Fig. 3. Then, the sentiment scores listed in the right column in Table 3 will be from slightly positive level 1, to strongly positive level 4 as shown in table above.



Slightly	Level 1	Level 2	Level 3	Level 4	Strongly
Positive	0-0.25	0.25-0.5	0.5-0.75	0.75-1.0	Positive
	0.25	0.5	0.75	1.0	

Fig. 3. The results when a 4-level output is selected.

When we select 10 level outputs, based on the sentiment score values, the sentiment in Table 3 will fall in different strength levels: positive level 3, level 5, level 8 and level 10 as shown in Fig. 4 below:

Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8	Level 9	Level 10
0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	04-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1.0
		0.25		0.5			0.75		1.0

Fig. 4. The results when a 10-level output is selected.

Therefore, dividing the sentiment levels into different bands, the different fine-grained sentiment analysis results can be presented accordingly.

We are not able to compare our results with the benchmarks because currently there are no benchmarks available for such multi-level sentiment analysis.

We found that most researchers tended to study the emotions directly, without the help of any sentiment or tone of the overall sentence. However, sentiment and emotion are closely related. For example, anger and sadness emotions are always negative sentiments and will never be positive. Therefore, associating emotion with sentiment is promising research. In our research, we perform sentiment analysis first, and then leverage the emotion dictionaries to further classify the positive sentiment to detailed positive emotions and negative sentiment to detailed negative emotions. In this research, the system can output the 6 often used emotion categories: Anxiety, Anger, Sadness, Satisfaction, Happiness, and Excitement, which are obtained from the categories in the emotion dictionary that we built.

We are not discussing the emotion results further here because there is no benchmark that we can use for comparison. There have been no existing methods available till now which can deliver detailed emotions. However, the ambivalence handling and emotion identification are all required by industry partners. This is also one of the motivations of this research work.

5. Conclusion

In this paper, we proposed a multi-level fine-scaled sentiment analysis with ambivalence handling method. We devised strength-level tuning parameters for measuring the scales of both positive and negative sentiment. Using web social comments, the results demonstrate good sensing capability and significantly better classification performance, compared to that of Stanford NLP. It was also shown that the ambivalence handler significantly increased the overall performance of the algorithms.

Moving forward, several potential improvements can be made on this research. The proposed sentiment strength analysis and emotion sensing method should be tested using more extensive ground-truth data. Furthermore, we are currently exploring ways to enhance the fine-grained scale of sentiment and emotion categories based on the needs of specific industries. In addition, we are in the process of building a set of extensive and reliable knowledge bases as well as the benchmark data for the calibration of future models.

References

1. E. Cambria, H. Wang, and B. White, Guest Editorial: Big social data analysis, *Knowledge- Based Syst.* **69** (2014) 1−2.
2. A. R. Alaei, S. Becken and B. Stantic, Sentiment analysis in tourism: capitalizing on big data, *J. Travel Res*. **58**(2) (2019) 175−191.
3. A. Murugan and H. T. Chelsey, Sentiment analysis of movie reviews using R, *Practical Text Analytics* (2019) 193−220.
4. A. Hussain and E. Cambria, Semi-supervised learning for big social data analysis, *Neurocomputing* **275** (2018) 1662−1673
5. S. L. Lo, E. Cambria, R. Chiong and D. Cornforth, Multilingual sentiment analysis: from formal to informal and scarce resource languages, *Artif. Intell. Rev.* **48**(4) (2017) 499−527.
6. Z. Wang, J. C. Tong and D. Chan, Issues of social data analytics with a new method for sentiment analysis of social media data, *IEEE 6th International Conference on Cloud Computing Technology and Science* (2014) 899−904.
7. E. Cambria, Y. Li, F. Xing, S. Poria, and K. Kwok, SenticNet 6: Ensemble application of symbolic and subsymbolic AI for sentiment analysis, CIKM (2020).
8. Y. Susanto, A. Livingstone, B. C. Ng, and E. Cambria, The hourglass model revisited, *IEEE Intelligent Systems* **35**(5) (2020).
9. I. Chaturvedi, E. Ragusa, P. Gastaldo, R. Zunino and E. Cambria, Bayesian network based extreme learning machine for subjectivity detection, *J. Franklin Inst.* **355**(4) (2018) 1780− 1797.
10. Y. Xia, E. Cambria, A. Hussain, and H. Zhao, Word polarity disambiguation using Bayesian model and opinion-level features, *Cognitive Computation* **7**(3) (2015) 369−380.
11. E. Cambria, M. Grassi, A. Hussain, and C. Havasi, Sentic computing for social media marketing, *Multimedia Tools and Applications* **59**(2) (2012) 557−577.
12. Z. Wang, J. C. Tong, X. Xin and H. C. Chin, Anomaly detection through enhanced sentiment analysis on social media data, *International Conference on Cloud Computing Technology and Science* (2014), pp. 917−922.
13. D. Mohey and E. M. Hussein, A survey on sentiment analysis challenges, *J. King Saud Univ.- Eng. Sci.* **30**(4) (2018) 330−338.
14. P. Gonçalves and M. Araújo, Comparing and combining sentiment analysis methods, *ACM Conf. Online Soc. Networks* (2013) 27−38.
15. H. Ji, H. Deng and J. Han, Uncertainty reduction for knowledge discovery and information extraction on the World Wide Web, *Proceedings of the IEEE* **100**(9) (2012) 2658−2674.
16. B. Gokaraju, S. Durbha, R. King, S. Member and N. Younan, A machine learning based spatio-temporal data mining approach for detection of harmful algal blooms in the Gulf of Mexico, *IEEE J. Sel. Top Appl. Earth Obs. Remote Sens*. **4**(3) (2011) 710−720.

17. T. Wilson, J. Wiebe and P. Hoffmann, Recognizing contextual polarity: an exploration of features for phrase-level sentiment analysis, *Assoc. Comput. Linguist.* **35**(3) (2009) 399−433.
18. J. Ortigosa-Hernández, J. D. Rodríguez, L. Alzate, M. Lucania, I. Inza and J. A. Lozano, Approaching sentiment analysis by using semi-supervised learning of multi-dimensional classifiers, *Neurocomputing* **92** (2012) 98−115.
19. X. Glorot, A. Bordes and Y. Bengio, Domain adaptation for large-scale sentiment classification: a deep learning approach, *Int. Conf. Mach. Learn.* (2011) 513−520.
20. G.-B. Huang, E. Cambria, K.-A. Toh, B. Widrow and Z. Xu, New trends of learning in computational intelligence [guest editorial], *IEEE Computational Intelligence Magazine* **10**(2) (2015) 16−17.
21. Z. Wang and Y. Parth, Extreme learning machine for multi-class sentiment classification of tweets, *Proc. ELM-2015* **1** (2016) 1−11.
22. Z. Wang, J. C. Tong and H. C. Chin, Enhancing machine-learning methods for sentiment classification of web data, *Asia Inf. Retr. Symp.* **8870** (2014) 394−405.
23. B. Pang, L. Lee and S. Vaithyanathan, Thumbs up? sentiment classification using machine learning techniques, *ACL-02 Conf. Empir. methods Nat. Lang. Process Assoc. Comput. Linguist.* **10** (2002), pp. 79−86.
24. E. Haddi, X. Liu, and Y. Shi, The role of text pre-processing in sentiment analysis, *Procedia Comput. Sci.* **17** (2013) 26−32.
25. J. A. Morente-Molinera, G. Kou, C. Pang, F. J. Cabrerizo and E. Herrera-Viedma, An automatic procedure to create fuzzy ontologies from users' opinions using sentiment analysis procedures and multi-granular fuzzy linguistic modelling methods, *Inf. Sci.* **476** (2019) 222− 238.
26. J. A. Morente-Molinera, G. Kou, K. Samuylov, R. Ureña, and E. Herrera-Viedma, Carrying out consensual group decision making processes under social networks using sentiment analysis over comparative expressions, *Knowledge-Based Syst.* **165** (2019) 335−345.
27. P. Shaver, J. Schwartz, D. Kirson and C. O'Connor, Emotion knowledge: further exploration of a prototype approach, *J Pers. Soc. Psychol.* **52**(6) (1987) 1061−1086.
28. A. Ortony and T. J. Turner, What's basic about basic emotions? *Psychol. Rev.* **97**(3) (1990) 315−331.
29. P. Ekman, An argument for basic emotions, *Cogn. Emot.* **6** (1992) 169−200.
30. D. Ghazi, D. Inkpen and S. Szpakowicz, Prior and contextual emotion of words in sentential context, *Comput. Speech Lang.* **28**(1) (2014) 76−92.
31. D. Chafale and A. Pimpalkar, Review on developing corpora for sentiment analysis using Plutchik's wheel of emotions with fuzzy logic, *Int J. Comput. Sci. Eng.* **2**(10) (2014) 14−18.
32. R. Plutchik, The nature of emotions: human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice, *Am. Sci.* **89**(4) (2001) 344−350.
33. J. Suttles and N. Ide, Distant supervision for emotion classification with discrete binary values, *International Conference on Intelligent Text Processing and Computational Linguistics* (2013) 121−136.
34. A. Neviarouskaya, H. Prendinger and M. Ishizuka. Textual affect sensing for sociable and expressive online communication, *International Conference on Affective Computing and Intelligent Interaction* (2007) 218−229.
35. T. Vogel and M. Wanke, *Attitudes and Attitude Change* (Psychology Press, 2016)
36. Z. Wang, J. C. Tong, P. Ruan, and F. Li, Lexicon knowledge extraction with sentiment polarity computation, *IEEE 16th International Conference on Data Mining Workshops* (*ICDMW*) (2016), pp. 978−983.
37. E. Cambria, A. Hussain, C. Havasi, and C. Eckl, Sentic Computing: Exploitation of common sense for the development of emotion-sensitive systems, *Development of Multimodal Interfaces: Active Listening and Synchrony* (2010) 148−156.

38. E. Cambria, A. Hussain, C. Havasi, and C. Eckl, SenticSpace: Visualizing opinions and sentiments in a multi-dimensional vector space, *Knowledge-Based and Intelligent Information and Engineering Systems* (2010) 385−393.
39. E. Cambria, T. Mazzocco, A. Hussain, and C. Eckl, Sentic Medoids: Organizing affective common sense knowledge in a multi-dimensional vector space, *Advances in Neural Networks* (2011) 601−610.
40. E. Cambria, T. Benson, C. Eckl, and A. Hussain, Sentic PROMs: Application of sentic computing to the development of a novel unified framework for measuring health-care quality, *Expert Systems with Applications* **39**(12) (2012) 10533−10543.