International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Refined Global Word Embeddings Based on Sentiment Concept for Sentiment Analysis

Aditya Kumar Yadav
Department of computer science and engineering,
A. P. J. Abdul Kalam Technical University, 226031, Lucknow

ABSTRACT

Sentiment analysis is a significant area of study in natural language processing that finds extensive use in journalism, politics, and other domains. In sentiment analysis, word embeddings are important. The sentiment lexicons are directly incorporated into conventional word representation using the current senstiment embeddings techniques. This sentiment representation technique is unable to offer precise sentiment information for words in many situations since it can only distinguish the sentiment information of distinct words, not the same word in several settings. To address the issue, this research suggests the sentiment idea. First, based on the word context, we determined the best sentiment idea for words in the Microsoft idea Graph. Then, using the multi-semantics sentiment intensity lexicon that we built in this research to ensure accurate. Embedding of sentiment information and provide more accurate semantics and sentiment representation for words, we extracted the sentiment information of words under the optimal sentiment concept. Ultimately, we integrated enhanced word embedding techniques to produce a more thorough word representation. The validity of the word embeddings method based on the sentiment notion suggested in this work is confirmed by comparing it with traditional and sentiment embeddings methods on six typical datasets.

INTRODUCTION

An automatic technique for obtaining sentiment data from unstructured texts is sentiment analysis. Machine learning, data mining, or natural language processing (NLP) are just a few domains that employ sentiment analysis. An essential stage in sentiment analysis is word vector representation. These days, the most popular word embedding systems are Word2Vec [1] & GloVe [2], both based on dispersed representation. Words with similar contexts have vector representations, based on the theory. which are similar.. Due to its ability to represent texts by capturing several contextual information, it is beneficial for many tasks involving semantic similarity.However, it might have the reverse impact in jobs involving SA. For instance, "cry" & "laugh" possess its same sentence context. "She is crying" & "she is laughing," therefore Word2Vec & GloVe will provide vector representations of "cry" and "laugh" that are similar. However, when analyzed from a sentiment perspective, the sentiment polarities of "laugh" and "cry" are opposed. To address this issue, researchers [3] and [4] enhanced the precision of sentiment analysis by including sentiment data derived from Word2Vec & GloVe.Nonetheless, there are still several issues with an examination of sentiment. Since words can mean different things depending on the context., it typically reflect different moods. In the lines "He bought the blue hat," "He said it was just a blue," and "He is blue that nothing is going to get better," for instance, the word "blue" has distinct meanings and conveys distinct emotions. Since the information about a word's sentiment in different sentences is the same, the current sentiment embedding methods cannot offer precise sentiment information for words that indicate distinct contexts and store the proper embedding. Instead, they immediately embed the sentiment lexicons into word representation. To give more correct semantics and sentiment representation, this study suggests a sentiment idea to deal with the problem. To further improve SA accuracy, we use the word's context to pick the optimum sentiment notion..

The following are this paper's primary contributions:
(1) By putting out the sentiment notion, which aims to accurately incorporate sentiment data along give words more precise semantic and sentiment representations., (2) merging six common sentiment intensity lexicons into a sentiment intensity lexicon that contains both multi-semantic and single-semantic sentiment terms to give more precise sentiment data for terms involving various meanings. (3) The calculation of Refined Global Word Embeddings (RGWE) involves averaging we improved Both Refined-GloVe and Refined-Word2Vec. RGWE incorporates sentiment data from both the inside and the outside, in addition to several position features. This work proposes a sentiment-based word embeddings method, and its validity is confirmed through an experiment on six datasets of varying sizes and categories.

The following is the main body in this paper. The pertinent The work on sentiment analysis word embeddings is presented in Section II. A thorough explanation on the word representation method RGWE suggested in this research is given in Section III. The experimental data have been compared & analyzed in Section IV. The work of this study is summarized in Section V, which also looks ahead to further research.

RELATED WORK

As NLP has developed, scholars worked hard on word embeddings and have turned their attention to examination of sentiment. A modified word embeddings approach determined by Part-of-Speech (POS)tagging technology for sentiment lexicons has been suggested by Rezaeinia et al. Phametal improved the SA efficiency of pre-trained word embeddings; Jiang et al. [5] created the Bag-of-words text representation technique determined by sentiment topic words, those performed well in sentiment analysis, using deep neural networks and sentiment topic words, including context information. [7] presented a mixed model with several CNNs that concentrated on Word2Vec, GloVe, & one-hot character vectors for word embeddings; Zhou et al. In tasks involving aspect sentiment categorization, it performed well. [8] developed a text representation model that considerably used TF-IDF also topic data from LDA of SA to minimize the standard representation model's dimension in word vector space. [9] Han and colleagues created a hybrid NN model. by combining CNN as well as LS for object representation. that combined user and product data; The BERT model, which Devlinetal [10] developed to represent text, can more accurately capture the shifting relationships among words in texts and did well on SA tasks; The issue of data sparsity can be resolved by using a neural topic model, Liu et al. [11] suggested integrating the text's latent topic information into word-level semantic representations. They also introduced a unique topic-word attention method to examine these word semantics from the standpoint of word association with the issue; By putting out a model that incorporated varying degrees of prior information into word embeddings, Li et al. [12] improved sentiment analysis performance; An improved word representation method that Xu et al. introduced weighted word vectors that were generated and had a higher F1 score. [13]; Petersetal. Additionally, the method integrated sentiment analysis into TF-IDF, the conventional algorithm. [14] suggested a deep learning framework-based text representation model, and Hao et al. [15] introduced a cross-domain sentiment organization technique utilizing ran-domembeddings, that yielded good results in the sentiment analysis job while maintaining a similar structure across embedding space; Usama et al. developed a model for representing English texts that included sentiment, syntax, & semantics through the training of several sentiment texts corpora. [16] integrated multilevel features from multiple network architectures and layers to improve sentiment analysis accuracy; Majumder et al. [17] proposed a framework for multitasking learning to restore two-task performance and demonstrated the connection between sentiment categorization and sarcasm detection; Ma et al. [18] suggested extending Targeted aspect-based polarity classification or target-dependent aspect detection were integrated tasks that Sentic LSTM will handle., as well as explicitly integrating explicit and implicit knowledge; Cambria et al. [19] enhanced human-computer connection by using CommonSense Computing to improve computers' ability to perceive and communicate emotions; Akhtar et al. [20] proposed utilizing a multi-layer perceptron network which combed its outputs as part of a stacked ensemble strategy for sentiment intensity prediction. using both traditional feature-based models and deep learning.; Gu et al. [21] improved sentiment analysis performance by adding sentiment intensity scores using sentiment lexicons to pretrained word vectors...Despite the researchers' incorporation of word sentiment information into representation algorithms, it is still impossible to integrate sentiment information precisely. This study presents a method that gives words more precise semantics and sentiment representation..

PROPOSED METHOD

This section will go into greater detail about the RGWE strategy that this study recommends. To create R-Word2Vec and R-GloVe, Word2Vec & GloVe's original word vectors, we first incorporate several attributes, including a sentiment concept, position, sentiment, or POS. The Refined-Word2Vec & Refined-GloVe representations then averaged to get RGWE, which accounts for external and internal sentiment data and other positioning considerations.

3.1 "WORD2VEC MODEL & GLOVE MODEL

Word2Vec is a popular word embeddings" approach that can mine vast amounts of data in distributing the vector representation of words. Instead of using words to forecast context like the skip-gram model, CBOW model employs context to predict words. These are Word2Vec's two available models. Each of the three layers that comprise the model—the input, projection, and output layers—is capable of accurately representing word embeddings. That which uses the skip-gram model used in this paper to represent text.

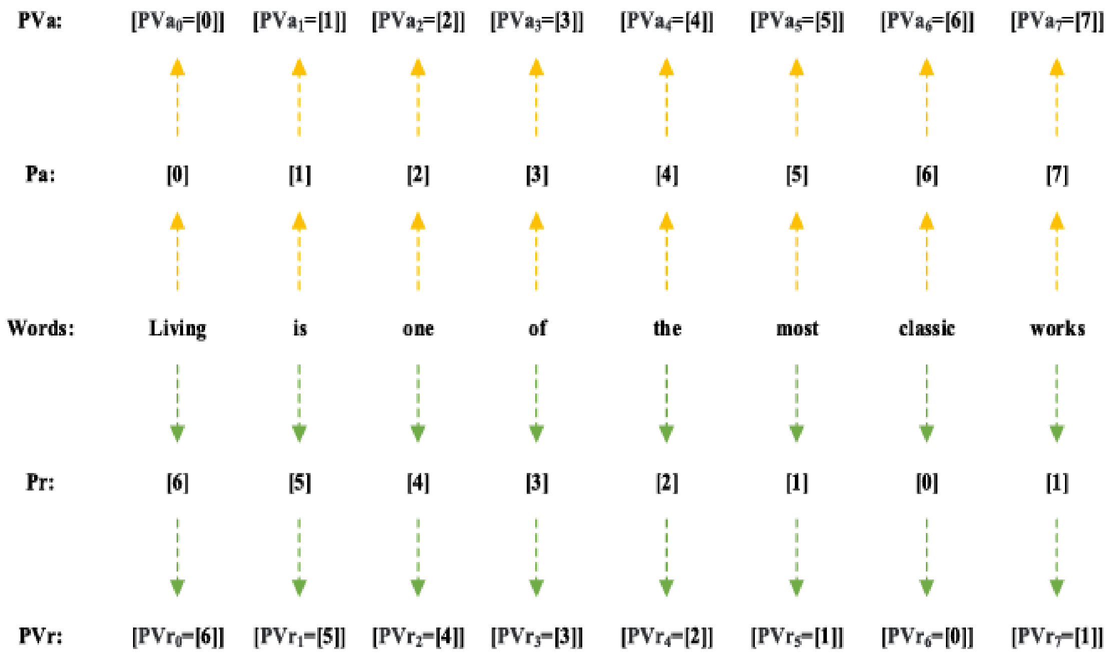


FIGURE1. Positionencodingofwords.

The global co-occurrence matrix provides the basis for the word vector representation in Glove, another well-liked word embeddings paradigm. The farther away two words are in a context, the lower the weight. Considering how two words are separated in the context window, it proposes an attenuation function..

3.2 POS

Word syntactic information can be found in POS tags. For sentiment recognition, the POS data for words is crucial in SA. Generally, distinct POS of words convey distinct moods and semantics. When the POS of "novel" is an adjective, for instance, its meaning is "fancy" and conveys positive sentiment; when it is a noun; it conveys narrative rather than sentiment. The Stanford parser is utilized in this study to link to Word2Vec/GloVe word vectors, either convert POS data to vectors or tag its POS of words.. Thus, the POS properties of words will be included in its Refined-Word2Vec/Refined-GloVe vectors..

3.3 POSITION

Yu et al. [22] examine how position features function in tasks involving sentiment analysis. R-Word2Vec and R-GloVe incorporate the words' relative and absolute position attributes.. The following is the comprehensive description:

3.3.1 ABSOLUTEPOSITION

Words' absolute positions are encoded and then converted to vector representation. Figure 1 illustrates how the sentence "Living is one of the most classic works"'s absolute position Pa=[0,1,2,3,4,5,6,7] is transformed into an absolute position vector PVa= [PVa0, PVa1, PVa2, PVa3, PVa4, PVa5, PVa6,PVa7], and then connects the words' absolute location feature vector using the Refined-Word2Vec vector representation.

3.4 RELATIVE POSITION

To offer more understandable position information, we combine the equally significant relative position data in Refined-Word2Vec with Refined-GloVe's absolute position feature. Regarding the relative location feature, we think that a word nearer a sentiment word makes a greater contribution to the phrase's sentiment assessment. To illustrate, the term "most" within the statement "Living is one in the most classic works" is more closely related to a mood word "classic," which appropriately conveys its sentiment of the sentence. The relative distance between context words and their sentiment word, which is set to 0 in the relative position coding of words, is known as their position. As shown in Figure 1, the line "Living is one of the most classic works" have been transformed into a relative position vector

PVr=[PVr0,PVr1,PVr2,PVr3,PVr4,PVr5,PVr6,PVr7], while the vector of Refined-GloVe is then coupled with the words' feature vector as relative position.

3.5 SENTIMENT LEXICONS

Sentiment intensity lexicons, Emotional lexicons (e.g., Strap-parava and Valitutti [25]) are instances of "binary sentiment lexicons (e.g., Hu and Liu [23]) and multi-classification sentiment lexicons (e.g., Riloff & Wiebe [24])". This uses sentiment polarity lexicons, as opposed to sentiment intensity. Work because they may provide more thorough and detailed sentiment data for words. The FSIL we created and its specifics Table 1 lists the sentiment intensity lexicons we employed..

TABLE1. Details of sentiment intensity lexicons.

Refs.	Lexicon	Size	Score Range
Nielsen FA.[27]	"AFINN	2477	[-5,+5]
Taboada M[28]	SO-COL	6306	[-5,+5]
Mohammad S M[29]	"NRC Hashtag Sentiment Lexicon	54129	[-7,+7]
Zhu X[30]	NRC Emotion Lexicon	62468	[-5,+5]
Cambria E[31]	SenticNet 5	100000	[-1,1]
Baccianella S[32]	Sentiwordnet" 3.0[2]	117659"	[0,1]

Sentiwordnet3.0 provides both, although it does not explicitly provide sentiment scores for each of the six sentiment intensity lexicons selected, it does provide positive and negative scores for each semantic of sentiment words in its interval [0,1]. For every semantic of sentiment words in Sentimentiwordnet 3.0, we calculate the sentiment score using formula (1) [26].:

$$SentiScore = Pos_{score} - Neg_{score} \qquad (1)$$

Possessing the sentiment words' favorable score in a senseman-tics Negscore is the sentiment word's negative score in that semantics, whereas SentiScore is Sentiwordnet 3.0's emotion score for the semantic word for sentiment. The emotion its interval [−1, +1] is then created utilizing the scores of six sentiment intensity lexicons using the normalized technique. In various semantics, we consider the sentiment words' sentiment information. As a result, we analyze the sentiment information and semantics of each sentiment term into six lexicons as follows. ws:

a. Since the sentiment word "w" appears formula (2) is used in several semantic sentiment lexicons to determine an FSIL sentiment score.

b. The sentiment word "w" also appears within a semantic sentiment lexicon, and it has the same sentiment score and semantics in FSIL.

$$SentiScore_w = \frac{\sum_{r=1}^{R} senti\text{Score}}{R}$$

$SentiScore_w$ is Win's sentiment score. FSIL, where R is the quantity of emoticons, and w is in, $SentiScore_{w\_r}$ in the sentiment lexicon represents the sentiment score of w. r;

c. The sentiment word "w" has several interpretations and appears in one or more sentiment lexicons. Initially, the cosine formula is utilized to ascertain its semantic similarity between several interpretations in w. Next, establish its threshold for semantic similarity H. The semantic representation of SSG is chosen randomly from among the semantics sentiScore in the semantics in FSIL. The average emotion score of the comparable semantic similarity within the semantics group (SSG) has been greater than H..

FSIL is derived from 343,039 emotion words by integrating, deduplicating, combining semantics, and computing six sentient lexicons' sentiment ratings. Of the 172,677 sentiment words in FSIL, 28146 have a single semantic meaning, whereas 144,531 have multiple interpretations..

3.6 SENTIMENT CONCEPT

We compare the words in the sentences with the sentiment words in FSIL to judge whether it is a sentiment word. Words in the sentences are context words except the sentiment words. A word can convey different sentiments depending on its context. A word has multiple semantics and belongs to different sentiment concepts in different contexts. For example, the sentiment concept of pink in sentence I like pink skirts is color , which expresses neutral sentiment. Whereas its sentiment concept in sentence He is the pink in the Foreign Of ce is elite , which expresses positive sentiment. Therefore, it is very important to determine the sentiment concept of words in the Sentiment Analysis tasks, which can determine the sentiment information in different contexts. The senti ment concept library used in this paper is Microsoft Concept Graph.3 For the sentence S w1 w2 wm, we predict the probability distribution of sentiment concept of the word wi, refer to formula (3) [33]:

"P(c|V)= $\frac{\exp(c.V)}{\sum_{c\epsilon c(w)} \exp(c.V)}$

$$p(c \mid V) = \frac{exp(c.V)}{\sum_{c_i \in C(w)} exp(c_i.V)}$$

C(w) are Microsoft Concept Graph" candidate concept set of wi, and Formula (4) determines V, indicating a vector representation of S:

$$V = \frac{1}{m} \sum_{i=1}^{m} e_i$$

After determining the ideal Refined-Word2Vec or Refined-GloVe, respectively, blend external or internal sentiment data under an optimal sentiment concept for words.

3.6.1 INTERNAL SENTIMENT INFORMATION EMBEDDINGS

Figure 2 illustrates how Refined-Word2Vec embeds intrinsic emotion in formation. The steps are as follows in detail:

1www."purl.com/net/sentimentoftweets"

2http://sentiwordnet.isti.cnr.it

3https://concept.research.microsoft.com/

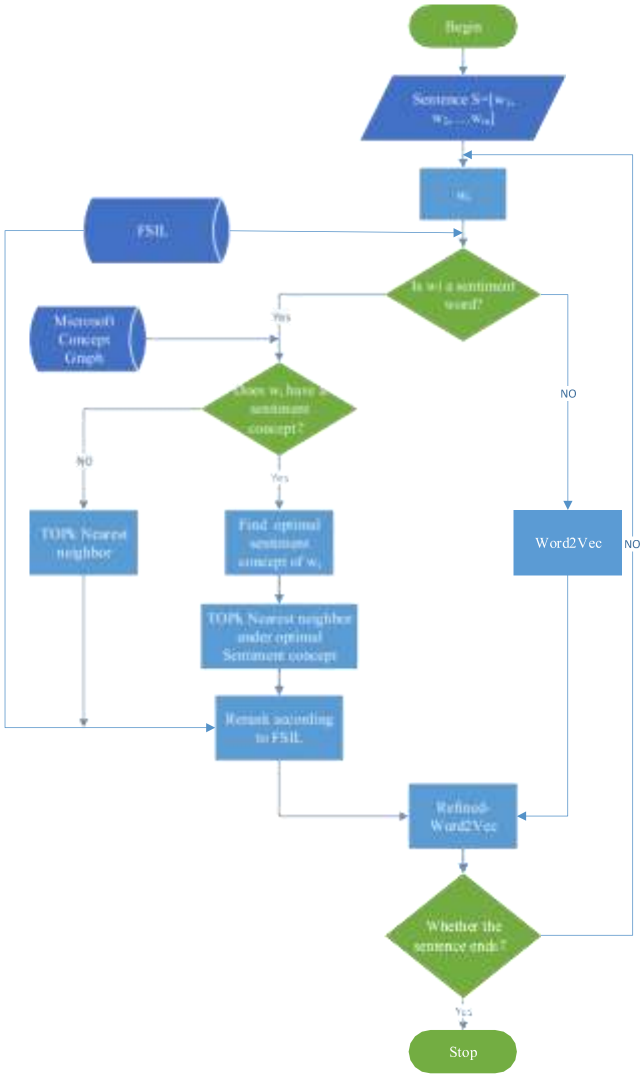


FIGURE2. The flow chart of internal sentiment information embedding.

(1) Examine its FSIL sentiment lexicon to see if w_i are sentiment word;
(2) We identify the best sentiment idea if w_i is a sentiment term. $c_{optimal}$ of w_i from Concept Graph on Microsoft;
(3) Choice TOP_k terms that are most semantically similar under $c_{optimal}$ of wi;
(4) Locate its sentiment intensity rating under $c_{optimal}$ of w_i & TOP_k in FSIL, After that, reorder TOP_k based on how it differs from w_i A smaller difference indicates more sentiments that are comparable to w_i and the better ranking. Two scenarios are considered while calculating TOP_k sentiment intensity score in FSIL:

a. TOP_i is in FSIL. We find the sentiment intensity score under $c_{optimal}$ of TOP_i in FSIL;
b. TOP_i is not in FSIL. We set the sentiment intensity score of TOP_i to 0.

TOP_i are frequently absent from its FSIL due to Microsoft Concept Graph's ability to group related words into concepts.

As a result, when one word is sentimental, it frequently happens that other words within the same concept are too. The purpose of incorporating internal sentiment data into Re ned-Word2Vec refers to formula (6)[21]:

$$\begin{aligned} argmin(V) = argmin \sum_{i=1}^{n} \Big[& \alpha dist\left(v_i^{t+1}, v_i^t\right) \\ & + \beta \sum_{j=1}^{10} w_{ij} dist\left(v_i^{t+1}, v_j^t\right) \Big] \end{aligned}$$

$$dist\left(v_i, v_j\right) = \sum_{d=1}^{D} \left(v_i^d - v_j^d\right)^2$$

n are the quantity of target words which want improvement. The target word's semantic vector distance and refined vector representation are shown in the first part. v_i^{t+1} at step t+1 & its v_i^t at step t throughout its iterative optimization procedure. Second portion shows the weighted sum of the emotion vector distance among v_i^{t+1} of w_i & v_i^t at step t of related words w_j..Using formula (7), we get the distance between the D-dimensional vectors v_i and v_j The divergence degree of v_i^{t+1} & v_j^t , as well as the proximity degree of v_i^{t+1} & v_j^t are controlled by α and β, respectively. The movement direction of vit is controlled by its sentiment contribution w_{ij} of w_j to v_i^{t+1}, while its movement distance of vit is controlled by α/β. v_i^t .

In this paper, the sentiment contribution w_{ij} of w_j to v_i^{t+1} is calculated by using formula(8):

$$w_{ij} = \frac{1}{e^t} 0 \leq t \ll 2$$

t is the absolute difference between w_i and w_{ij} sentiment intensity. Our reasoning led to the construction of formula (8): to v_i^{t+1} , the emotion contribution w_{ij} of w_j diminishes as the absolute difference t grows. With a smaller absolute difference from w_i, similar words w_j provide more v_i^{t+1} than those with a larger absolute difference. As seen in formula (8), the target word's improved vector representation is more complete than the sentiment information of similar terms..

3.6.2 EXTERNAL SENTIMENT INFORMATION EMBEDDINGS

Refined-Glove incorporates external sentiment data using formula (9) These meanings and sentimental information about words are contained in the first half. The words' sentiment notion information is part of the second section to restrict these meanings or emotion ranges..

$$v_i=\gamma_i e_{ig}+\gamma e_c \qquad (9)$$

e_{ig} indicates the sentiment weight, and γi represents the original vector which GloVe used to demonstrate w_i .. The ideal sentiment concept $c_{optimal}$ for w_i is represented by vector ec, whereas the sentiment concept weight is denoted by γ. The sentiment weight formula we created, which was influenced by Xu et al. [13], is displayed in formulas (10) (11):

$$\gamma_i=TF-IDF_i*\tau \qquad (10)$$

TF-IDF remains its approach for calculating weight in text classification, which is most commonly employed.. We believe that the emotion weight increases with an sentiment intensity of the words..

TABLE2. “Detailed statistics of the experimental datasets.

Dataset	Train	Valid	Test	Total	Classes	Balance
SemEval[10](Nakov etal,2013)”	9684	1654	3813	15151	3(positive/neutral/negative)	No
SST1[11](Socher etal ,2013)	8544	1101	2210	11855	5(very positive/positive/ neutral/negative/very negative)	No
SST2(Socher etal,2013)	6920	872	1821	9613	2(positive/negative)	No
IMDB[12](Pang etal, 2005)	40000	5000	5000	50000	2(positive/negative)	Yes
Amazon[13](health)	8000	1000	1000	10000	2(positive/negative)	Yes
Yelp 2014[14](Restaurant)	3072	384	384	3840	5(very positive/positive/ neutral/negative/very negative))	Yes

Various sentiment contribution values (τ) are set depending on how strongly the words express a sentiment. The matching sentiment contribution value is then determined using a formula once we have the sentiment intensity under the coptimal of terms. (11):

$$\tau = \begin{cases} 1 & |SentiScore_w| = 0 \\ 6/5 & 0\,|SentiScore_w| < 0.2 \\ 7/5 & 0.2 \leq |SentiScore_w| < 0.4 \\ 8/5 & 0.4 \leq |SentiScore_w| < 0.6 \\ 9/5 & 0.6 \leq |SentiScore_w| < 0.8 \\ 2 & 0.8 \leq |SentiScore_w| \leq 1 \end{cases}$$

$SentiScore_w$ "is the sentiment score of *w* in FSIL.

3.7 REFINED GLOBAL WORD EMBEDDINGS"

The representation for the two different vectors after words has been represented via Refined-Word2Vec or Refined-GloVe is averaged to determine RGWE. This are derived from our analysis of

(1) In Refined-Word2Vec, words' absolute position and relative location features are integrated.. Different position features are integrated by Refined-GloVe, respectively, to get a more thorough representation of position features;

(2) In order to achieve a more thorough sentiment feature representation, RGWE combines the internal and exterior sentiment features, which are incorporated in Refined-Word2Vec and Refined-GloVe, respectively.

EXPERIMENT

4.1 DATASETS

Six accessible, traditional public datasets are chosen to assess how well the RGWE suggested in Sentiment Analysis tasks performs. Table 2 displays the dataset details. We conduct experiments following train/valid/test in the typical train/valid/test SemEval [34], SST1 [35], & SST2 [35] datasets. Regarding databases like IMDB [36], Yelp2014 (restaurant) [38], or Amazon (health) [37], that are totally balanced but lack standard train/valid/test, we explore by stratifying sample using 8:1:1 to generate its equivalent train/valid/test. To prevent technical terminology from influencing sentiment analysis, we also ensure that the training and test sets' intersection is not empty.

4.2 EXPERIMENT SETTING

4.2.1 DATAPREPROCESSING

We carry out the following general preprocessing for datasets: 1. Remove special characters and English words; 2. Remove stopwords and words with fewer than five frequencies; 3. Convert all uppercase to lowercase; 4. Extend abbreviations 10 to make sure sentiment terms are present in FSIL; 5. Stemming; 6. Text segmentation; 7. POStagging. 11. We don't remove the short sentences since we believe that some of them, like "verygood," "itistoobad," etc., convey sentiment.

4.2.2 WORD EMBEDDINGS METHODS

Word embeddings methods for comparison:

Traditional word embeddings: Word2Vec[12](skip-gram) andGloVe13;

Sentiment embeddings : SSWE[14];

Refined Embeddings : Refined-Word2Vec, Refined-GloVe, Seninfo+TF-IDF[13], Re(GLOVE) [21], and the RGWE that we suggested.

Pre-trained on 300-dimension training datasets are Seninfo+TF-IDF, Re(GLOVE), SSWE, GloVe, Refined-GloVe, Refined-Word2Vec, & RGWE. Vectors for words can randomly assigned for terms that have not been used in previously trained

4.2.3 DEEP LEARNING METHODS

There are 3 widely used DL techniques are chosen to examine sentiment texts:

Convolution Neural Networks (CNN): CNN uses a convolutional layer to extract local feature information from texts. The convolution filter's area sizes were 2, 3, 4, or 60 filters for every region size..

Bidirectional Long Short Term Memory Network (Bi-LSTM): records its preceding contextual feature information in texts both forward or backward, avoiding the problems of gradient fading along bursting. Two 128-unit hidden network units are employed..
Bidirectional Gated Recurrent Units (Bi-GRU): are variation in Bi-LSTM with fewer parameters and a simpler structure, lowering the gating based on Bi-LSTM. Two ×128 hidden network units are used.

9 http://nlp.stanford.edu/IR-book/html/htmledition/dropping-commonterms-stop-words-1.html

10 http://www.noslang.com/dictionary

11 https://nlp.stanford.edu/software/tagger.shtml

12 https://code.google.com/archive/p/Word2Vec/

13 http://nlp.stanford.edu/projects/glove/

14 http://ir.hit.edu.cn/~dytang/

TABLE3. “The F1-score of different embedding methods on datasets with CNN.

Method	Dataset	SemEval	SST1”	SSt2	IMDB	Amazon	Yelp 2014
Conventional Embeddings	Word2Vec[1]	62.3	45.7	84.1	84.5	84.4	42.1
	GloVe[2]	63.2	46.8	84.9	85.5	85.2	43.1
Sentiment Embeddings	SSWE	64.1	47.1	87.6	87.6	87.7	43.5
	Senifo+TF-IDF[13]	66.7	49.1	88.8	89.0	89.0	45.4
	Re(GloVe)[21]	68.2	50.2	89.6	89.6	89.6	46.1
Refined Embeddings	Refined-Word2Vec	66.8	49.5	89.2	89.2	88.8	45.9
	Refined-GloVe	68.3	50.5	89.7	89.7	89.6	46.3
	RGWE	69.1	50.8	90.1	90.1	89.9	46.9

Furthermore, all of these datasets we selected are short texts; 97.8% of sentences are less than 110 words long, while the longest sentence is 326. We therefore decided to put the maximum length at 110. Less than 110-word sentences will have no vector. We use dropout and set it to 0.5 to prevent overfitting. Softmax is the classification function, and Tanh is the buried layer's activation function..

4.3 EXPERIMENT RESULTS

Following validation set verification, we set the ideal k to 10. For datasets SemEval, Amazon, and Yelp 2014, the optimal α:β=0.03 and for datasets SST1, SST2, and IMDB, the optimal α:β=0.1, respectively. As the evaluation indicator, we use the F1 Score, which may completely quantify performance. The average F1 Score after ten runs on test sets represents the experimental outcomes.

4.3.1 "COMPARISON OF DIFFERENT WORD EMBEDDINGS METHODS

Seninfo+TF-IDF, Re(GLOVE), RGWE, GloVe, SSWE, Refined-Word2Vec, Refined-GloVe, or Word2Vec" experimental findings have been compared in Table 3.

Table 3 shows that Seninfo+TF-IDF and SSWE, two sentiment embedding techniques, perform better on datasets than standard embedding techniques. This demonstrates the significance of Seninfo+TF-IDF and SSWE, which use sentiment information about words. Hence, sentiment analysis uses sentiment information. In this study, Refined-Word2Vec and Refined-GloVe outperform Seninfo+TF-IDF and SSWE. The reason is that more specific sentiment features can be obtained by integrating into sentiment intensity lexicon FSIL, which comprises 172677 sentiment phrases. The emotion features and sentiment concept features of words are present in R-Word2Vec & R-GloVe, which are more effective at expressing the actual sentiment of words in sentences. On the SemEval, SST1, and Yelp2014 datasets, Refined-GloVe outperforms Re(GloVe) by a small margin. The sentiment data use different methods is integrated using Refined-GloVe or Re(GloVe).

[10] http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=28685

[11] https://nlp.stanford.edu/sentiment/

[12] https://www.imdb.com/

[13] http://snap.stanford.edu/data/amazon-meta.html

[14] https://www.yelp.com/

However, words' sentiment idea information is absent from Re(GloVe). On datasets with several classifications, the distinction is more noticeable. The RGWE approach performs the best. Average F1 values for CNN's fine-grained classification, ternary classification, and binary classification are 48.85%, 69.1%, & 89.86%, respectively. Because of this, RGWE incorporates both internal and external emotion data in addition to other position features.

4.3.2 "COMPARISON OF DIFFERENT DEEP LEARNING METHODS

The experimental" outcomes of several DL techniques in conjunction with RGWE on datasets are displayed in Table 4. Clearly, RGWE & Bi-GRU work better together than RGWE or Bi-LSTM. GRU are easier to converge than LSTM since it is an LSTM variation with fewer parameters. Although GRU outperforms LSTM on small-scale datasets, LSTM outperforms GRU on large-scale datasets. It is challenging to find large publically accessible sentiment analysis datasets, however. Our experimental data sets show that GRUp outperforms LSTM because the representative and conventional sentiment analysis datasets we select have an inadequate scale..

4.3.3 "OPTIMAL SEMANTIC SIMILARITY THRESHOLD H

Finding its sentiment" information of words in different SSGs is difficult when building because identical semantics may emerge in distinct SSGs if a semantic similarity threshold H is set too high; conversely, dissimilar semantics may appear into same SSG if a threshold H is set too low. This is why the sentiment intensity lexicon on FSIL is important. Making it impossible to discern the sentiment data with several meanings contained in a single SSG. To identify the ideal threshold H, we analyze how well various thresholds perform on various datasets.

Figure 3 displays the effectiveness of various threshold H on datasets. The ideal threshold H is found to be between 0.7 and 0.8. H=0.78 is the value we set in this paper.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

4.3.4 "OPTIMAL SENTIMENT CONCEPT WEIGHT γ"

In R-GloVe, a sentiment concept's contribution to sentiment analysis is measured using its sentiment concept weight, or γ. If γ is too little, it will not be able to capture the differences between various sentiment concepts; if it is too large, the contribution will be exaggerated along with sentiment analysis accuracy will be diminished. We evaluate how well various sentiment concept weights perform to determine the ideal sentiment concept weight on datasets with H 0.78.

TABLE4. "The F1-score of different deep learning methods on datasets with RGWE.

Method	Dataset	SemEval	SST1	SST2	IMDB	Amazon	Yelp 2014
CNN	RGWE	69.1	50.7	89.6	90.1	89.9	46.9
Bi-LSTM	RGWE	70.0	51.6	90.3	91.0	90.6	47.3
Bi-GRU	RGWE	70.8	52.3	91.2	91.3	91.1	47.9"

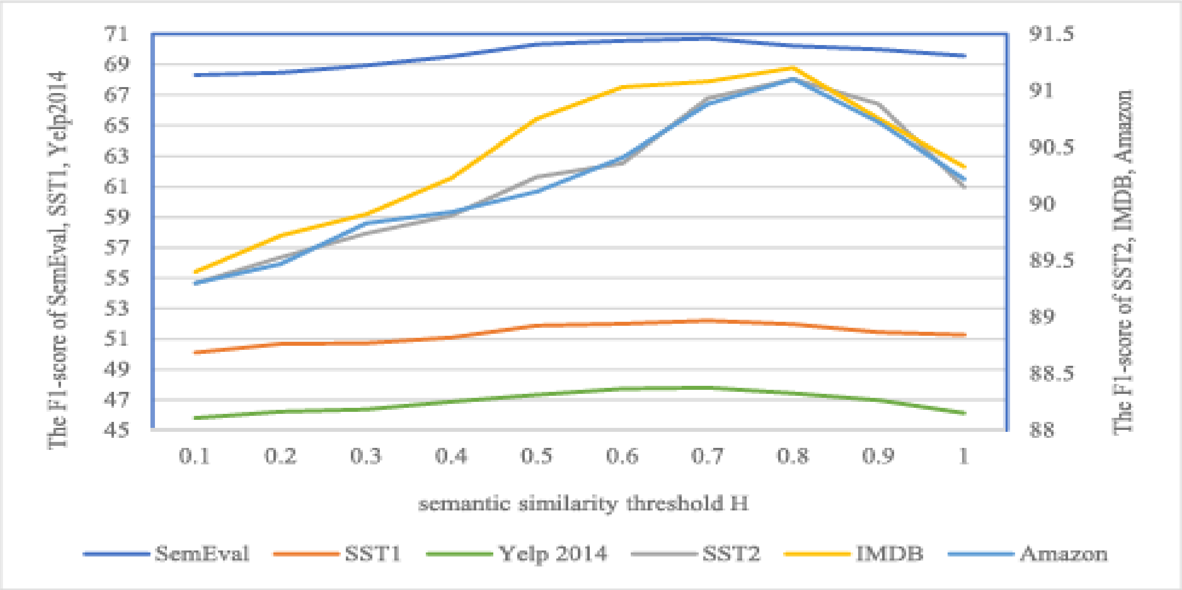


FIGURE3. Performance comparison of different threshold H on datasets with Bi-GRU

Figure 4 illustrates how various weight γ values work on datasets in Refined-GloVe. It is evident indicates 0.8 to 1.0 is the optimal range for sentiment concept weight. In this work, γ=0.9 is set..

4.3.5 "THE INFLUENCE OF SENTIMENT CONCEPT ON SENTIMENT ANALYSIS

We" assess how sentiment analysis affects sentiment. Based on Word2Vec and GloVe, the vector representations incorporating features other than sentiment ideas are Word2Vec+sen or GloVe+sen, respectively. The average for the Word2Vec or GloVe combination is called RGWE1. RGWE1+sen is the symbol denoting the average for Word2Vec+sen & GloVe+sen.

Simply using sentiment lexicon is not as effective as embedding sentiment features along with sentiment concept features, as Table 5 shows. Refined-GloVe with sentiment concept's F1-score is 1.6%, 1.2%, 0.8%, 1.1%, 0.9%, and 1.1% higher than GloVe+sen on six datasets, respectively. On six datasets, F1-score in RGWE within sentiment concept is 1.5%, 1.3%, 0.8%, 1.2%, 0.8%, and 1.6% higher than RGWE1+sen, respectively, demonstrating the significance of sentiment concept in SA. Furthermore, we discovered that sentiment concepts have a stronger influence on fine-grained classification compared to other classifications. This is because several categories (negative and very negative) can be expressed similarly, while fine-grained sentiment classification is more specific. The sentiment information is more significant when differentiating between words with different meanings. IMDB is more prone to sentiment notions in binary classification datasets. We consider that there is more information on IMDB and a wider range of opinions.

4.3.6 "THE INFLUENCE OF DIFFERENT TYPES SENTIMENT LEXICONS ON SENTIMENT ANALYSIS"

It is possible to obtain sentiment information about words by embedding sentiment polarity or sentiment intensity lexicons into word representations. A comparison is made between embedding sentiment intensity lexicons vs sentiment polarity lexicons. terms with an emotion intensity higher than zero are classified as positive terms in the FSIL to construct the fusion sentiment polarity lexicon (FSPL). Conversely, negative keywords are those whose sentiment intensity is less than zero.

The embedding of FSIL performs better than FSPL, as illustrated in Figure 5, since FSIL offers more comprehensive sentiment information for words than FSPL, which just distinguishes sentiment polarities.

4.3.7 "THE INFLUENCE OF DIFFERENT SIZE SENTIMENT LEXICONS ON SENTIMENT ANALYSIS

Lexicon size" is growing sequentially by AFINN-FSIL. We compare how lexicon size affects sentiment analysis via seven different-sized sentiment intensity lexicons.

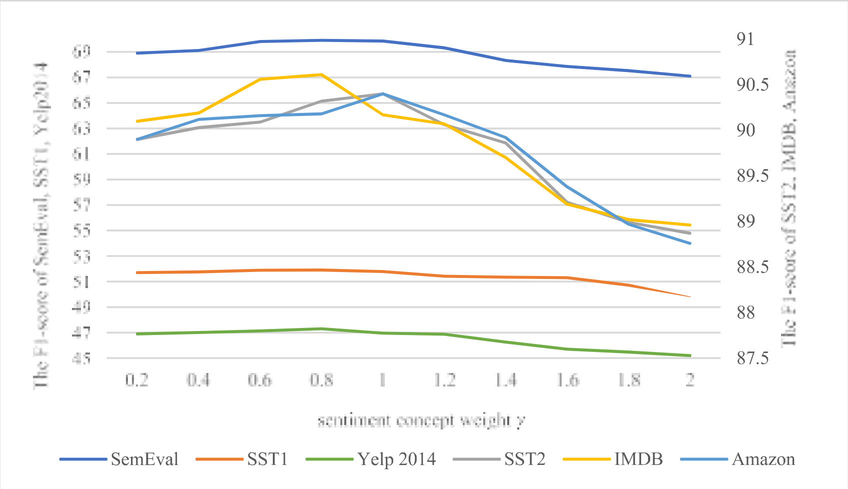


FIGURE4: "Performance comparison of different weight γ on datasets with Bi-GRU".

TABLE5. "The F1-score of sentiment concept on sentiment analysis with Bi-GRU".

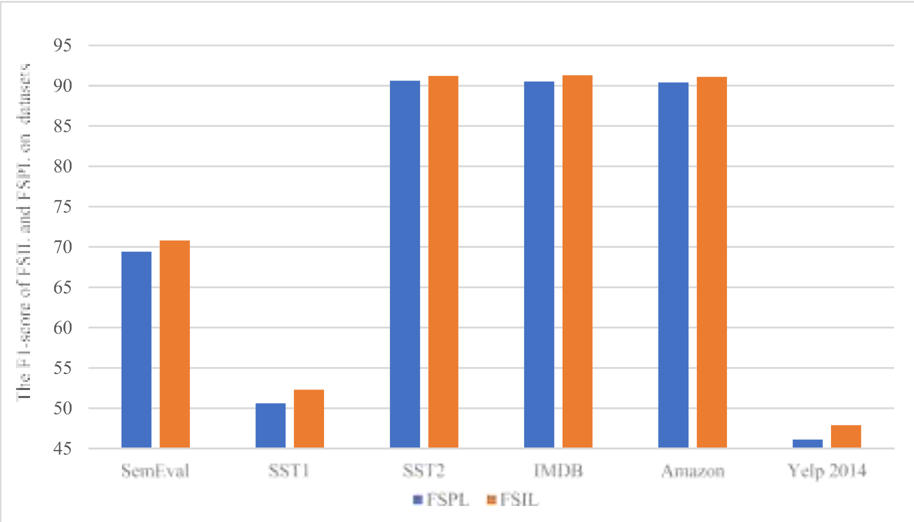
Method	SemEval	SST1	SST2	IMDB	Amazon	Yelp 2014
Word2Vec	65.1	48.7	87.4	87.7	87.5	43.9
Word2Vec+sen	67.3	50.3	88.7	88.9	88.8	45.3
RefinedWord2Vec	68.9	51.6	89.6	90.1	89.7	46.5
GloVe	65.9	59.5	87.9	88.2	87.7	45.0
GloVe+sen	68.3	50.7	89.6	89.7	89.5	46.2
Refined-GloVe	69.9	51.9	90.4	90.8	90.4	47.3
RGWE1	65.4	49.2	87.6	88.5	87.3	44.3
RGWE1+sen	69.3	51.0	90.4	90.1	90.3	46.3
RGWE	70.8	52.3	91.2	91.3	91.1	47.9

FIGURE5. "Performance of FSIL and FSPLon different datasets with Bi-GRU".

Fig. 6 illustrates the impact of sentiment lexicon size on sentiment analysis by showing that the F1-score rises as sentiment lexicon size grows. Figure 6 further shows that the embedding in FSIL outperforms all other datasets. It is because: (1) FSIL represents the largest sentiment lexicon among the seven. (17267); (2) FSIL offers more comprehensive sentiment data for terms in various situations, which can be examined by examining the multi-semantics and varying sentiment intensities found in 83.7% of sentiment words.

4.4 ERROR ANALYSIS

Some of the sentences in our experiment had erroneous analytical findings. We draw the following 2 conclusions:



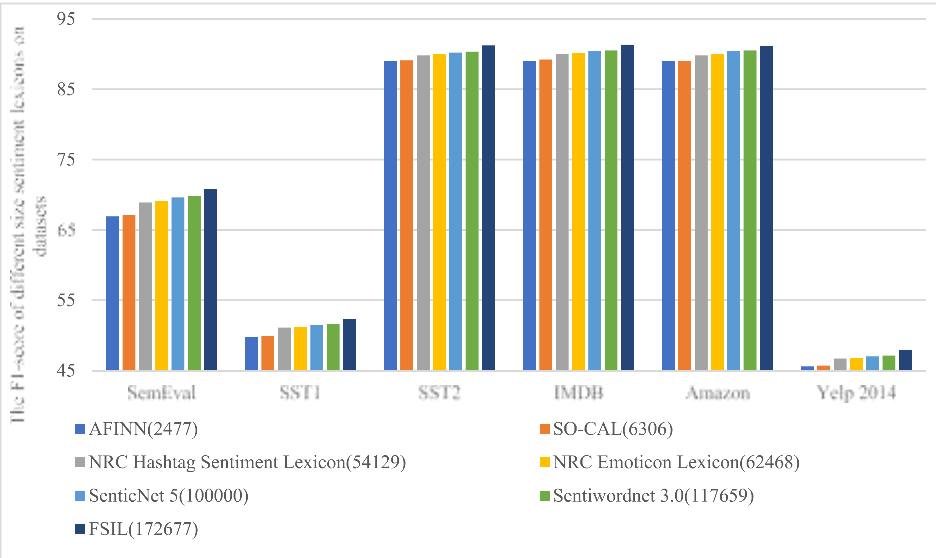


FIGURE6. "Performance of different size sentiment lexicons on datasets with Bi-GRU".

(1) In FSIL, nouns, verbs, adjectives, and adverbs make up 63%, 6%, 21%, and 10% of the total. That is, 37% of the feeling words in FSIL are non-nouns, and 63% are nouns. However, since the Microsoft Concept Graph, it has nouns for ideas or instances (words under concept); sentiment concepts only become discernible for noun sentiment words.. This indicates that the appropriate sentiment notion is absent from 37% of the words in FSIL.

(2) The claim that FSIL contains 28146 single-semantic sentiment words and 144531 multi-semantic sentiment words is untrue. Because single-semantic sentiment words communicate neutral emotion or have additional semantics that are not present in FSIL, they may be multi-semantic.

CONCLUSION

SA have been utilized in numerous fields since it introduced NLP technology, and researching word embedding strategies is crucial for SA tasks since the effectiveness of SA is more significantly impacted by word embedding quality.. The RGWE approach, which is founded on the sentiment notion, is proposed in this study to solve the issue of current word representation algorithms' inability to include sentiment information reliably in tasks involving sentiment analysis. We provide more accurate sentiment & semantic representations for words and select the optimal sentiment notion for words based on different situations. By averaging R-Word2Vec along with R-GloVe, RGWE increases the precision in sentiment analysis by using external as well as internal sentiment information in addition to different position features. Its RGWE's authenticity is verified by comparing it with sentiment & conventional embedding methods on representative datasets. Nevertheless, the Microsoft Concept Graph's ideas & examples are restricted to nouns; a feeling concept for adjectives and Verbs,

6. REFERENCES

[1] T. Mikolov, K. Chen, G. Corrado, and J. Dean, ‘‘Efficient estimation of word representations in vector space,’’ in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2013, pp. 1–12.

[2] J. Pennington, R. Socher, and C. Manning, ‘‘Glove: Global vectors for word representation,’’ in *Proc. Conf. Empirical Methods Natural Lang.Process. (EMNLP)*, 2014, pp. 1532–1543.

[3] D. Tang, F. Wei, B. Qin, N. Yang, T. Liu, and M. Zhou, ‘‘Sentiment Embeddings with Applications to Sentiment Analysis,’’ IEEE Trans.*Knowl. Data Eng.*, vol. 28, no. 2, pp. 496–509, Feb. 2016.

[4] Y. Ren, Y. Zhang, M. Zhang, and D. Ji, ‘‘Improving Twitter sentiment classification using topic-enriched multi-prototypical word embeddings,’’ in Proc. *AAAI Conf. Artif. Intell.*, 2016, pp. 3038–3044.

[5] Z. Jiang, S. Gao, and L. Chen, ‘‘Study on text representation method based on deep learning and topic information,’’ Computing,vol 102,no.3,
pp.623–642, Sep.2019.

[6] S. M. Rezaeinia, R. Rahmani, A. Ghodsi, and H. Veisi, ‘‘Sentiment analysis based on improved pre-trained word embeddings,’’ *Expert Syst. Appl.*,vol. 117, pp. 139–147, Mar. 2019.

[7] D.-H. Pham and A.-C. Le,‘‘Exploiting multiple word embeddings and one-hot character vectors for aspect-based sentiment analysis,’’Int. J. Approx.*Reasoning*, vol. 103, pp. 1–10, Dec. 2018.

[8] W. Zhou, H. Wang, and H. Sun, ‘‘A method of short text representationbasedonthefeatureprobabilityembeddedvector,’’*Sensors*,vol.19,no.17,
pp.185–209, Aug 2019.

[9] H. Han, X. Bai, and P. Li, ‘‘Augmented sentiment representation by learning context information,’’ Neural Comput.*Appl.*,vol.31,no.12,
pp.8475–8482, Dec.2019.

[10] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, ‘‘BERT: Pre-trainingof deep bidirectional transformers for language understanding,’’ in *Proc.North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.*,2019, pp. 4171–4186.

[11] W. Liu, G.Cao,andJ. Yin, Bi-level attention model for sentiment analysis of short texts,’’ *IEEE Access*, vol. 7, pp. 119813–119822, Sep. 2019.

[12] Y.Li,Q.Pan,T.Yang,S.Wang,J.Tang, and
E.Cambria,‘‘Learningwordrepresentationsforsentimentanalysis,’’*Cognit.Comput.*,vol.9,no.6,
pp.843–851,Dec.2017.

[13] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, ‘‘Sentiment analysis ofcommenttextsbasedonBiLSTM,’’*IEEEAccess*,vol.7,pp.51522–51532,Jan.2019.

[14] M.E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark,K.Lee, and
L. Zettlemoyer, ‘‘Deep contextualized word representations,’’ *J. Assoc.Comput. Linguistics*, vol. 1, pp. 2227–2237, Mar. 2018.

[15] Y. Hao, T.Mu, R.Hong, M.Wang, X.Liu, and J.Y. Goulermas, ‘‘Cross-domain sentiment encoding through stochastic word embedding,’’ IEEE Trans. *Knowl. Data Eng.*, vol. 32, no. 10, pp. 1909–1922, Oct. 2020.

[16] M. Usama, W. Xiao, B. Ahmad, J. Wan, M. M. Hassan, and A. Alelaiwi, ‘‘Deep learning based weighted feature fusion approach for sentiment analysis,’’ *IEEE Access*, vol. 7, pp. 140252–140260, Sep. 2019.

[17] N. Majumder, S. Poria, H. Peng, N. Chhaya, E. Cambria, A. Gelbukh, and E. Cambria, ‘‘Sentiment and sarcasm classification with multitask learning,’’ *IEEE Intell. Syst.*, vol. 34, no. 3, pp. 38–43, May 2019.

[18] Y. Ma, H. Peng, T. Khan, E. Cambria, and A. Hussain, ‘‘SenticLSTM: A hybrid network for targeted aspect-based sentiment analysis,’’ *Cognit.Comput.*, vol. 10, no. 4, pp. 639–650, Aug. 2018.

[19] E. Cambria, A. Hussain, C. Havasi and C. Eckl, ‘‘Sentic computing: Exploitation of common sense for the development of emotion-sensitivesystems,’’ in *Development of Multimodal Interfaces: Active Listeningand Synchrony* (Lecture Notes in Computer Science), vol. 5967. Berlin, Germany: Springer, 2010, doi: 10.1007/978-3-642-12397-9_12.

[20] M.S.Akhtar,A.Ekbal,andE.Cambria,‘‘Howintenseareyou? Predictingintensitiesofemotionsandsentimentsusingstackedensemble[applicationnotes],’’ *IEEE Comput. Intell. Mag.*, vol. 15, no. 1, pp. 64–75, Feb. 2020.
[21] S. Gu, L. Zhang, Y. Hou, and Y. Song, ‘‘A position-aware bidirectional attention network for aspect-level sentiment analysis,’’ in *Proc. Int. Conf.Comput. Linguistics*, 2018, pp. 774–784.
[22] L.-C.Yu, J.Wang, K.R.Lai,andX.Zhang, ‘‘Refiningwordembeddingsusing intensity scores for sentiment analysis,’’ *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 3, pp. 671–681, Mar. 2018.
[23] M.HuandB.Liu,‘‘Miningandsummarizingcustomerreviews,’’in*Proc.10thACMSIGKDDInt.Conf.Knowl.DiscoveryDataMining*,2004, pp.168–177.

[24] E.RiloffandJ.Wiebe, ‘‘Learningextractionpatternsforsubjectiveexpres-sions,’’ in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2003, pp.105–112.

[25] C. Strapparava and A. Valitutti, ‘‘WordNet affect: An affective extension of WordNet,’’ in *Proc. Conf. LREC*, 2004, pp. 1083–1086.
[26] F.H.Khan,U.Qamar,andS.Bashir,‘‘Asemi-supervisedapproachtosen-timentanalysisusingrevisedsentimentstrengthbasedonSentiWordNet,’’*Knowl. Inf. Syst.*, vol. 51, no. 3, pp. 851–872, 2017.
[27] F. A. Nielsen, ‘‘A new ANEW: Evaluation of a word list for sentiment analysis microblogs,’’ in *Proc. ESWC Workshop Making Sense Microposts, Big Things Come Small Packages*, 2011, pp. 93–98.
[28] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, ‘‘Lexicon-basedmethodsforsentimentanalysis,’’*Comput.Linguistics*,vol.37,no.2, pp.267–307,2011.

[29] S. M. Mohammad, S. Kiritchenko, and X. Zhu, ‘‘NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets,’’ in *Proc. 7th Int.Workshop Semantic Eval. Exercises (SemEval)*, Aug. 2013, pp. 116–122.
[30] X. Zhu, S. Kiritchenko, and S. Mohammad, ‘‘NRC-Canada-2014: Recent improvements in the sentiment analysis of tweets,’’ in *Proc. Int. Conf.Comput. Linguistics*, 2014, pp. 443–447.
[31] E. Cambria, D. Hazarika, K. Kwok, and S. Poria, ‘‘SenticNet 5: Discovering conceptual primitives for sentiment analysis using context embeddings,’’ in *Proc. Nat. Conf. Artif. Intell.*, 2018, pp. 1795–1802.
[32] S. Baccianella, A. Esuli, and F. Sebastiani,‘‘SentiWordNet3.0: An enhanced lexical resource for sentiment analysis and opinion mining,’’in *Proc. Lang. Resour. Eval.*, 2010, pp. 17–23.
[33] J. Cheng, Z. Wang, J.-R. Wen, J. Yan, and Z. Chen, ‘‘Contextual textunderstanding in distributional semantic space,’’ in *Proc. 24th ACM Int.Conf. Inf. Knowl. Manage.*, Oct. 2015, pp. 133–142.
[34] P. Navkov, S. Rosenthal, Z. Kozareva, V. Stoynov, A. Ritter, and T. Wilson, Semeval-2013task2: Sentiment analysis in Twitter,’’ in Proc.*SemEval*,2013, pp. 312–320.
[35] R. Socher, A.Perelygin, J.Wu,J.Chuang,C.D.Manning, A.Ng,and C. Potts, ‘‘Recursive deep models for semantic compositionality over sentiment treebank,’’ in *Proc. Conf. Empirical Methods Natural Lang.Process. (EMNLP)*, 2013, pp. 1631–1642.

[36] B. Pang and L. Lee, ‘‘Seeing stars: Exploiting class relationships for sentiment categorization with respect to ratings scales,’’ in Proc. 43rd Annu.*Meeting Assoc. Comput. Linguistics (ACL)*, 2005, pp. 115–124.
[37] K. Baktha and B. K. Tripathy,‘‘Investigation of recurrent neural networks in the field of sentiment analysis,’’ in *Proc. Int. Conf. Commun. SignalProcess. (ICCSP)*, Apr. 2017, pp. 2047–2050.
[38] S. Kiritchenko, X. Zhu, C. Cherry, and S. Mohammad, ‘‘NRC-Canada-2014:Detectingaspectsandsentimentincustomerreviews,’’in*Proc.8thInt. Workshop Semantic Eval. (SemEval)*, 2014, pp. 437–442.