Journal of Science and Engineering VOL. 1, NO. 1 (OCTOBER 2024) PP. 13-18

English Text Sentiment Analysis Network Based on CNN and U-Net

Weirong Zhang
School of Information Engineering
Weifang Vocational College
Weifang, Shandong, China

Jinfeng Wang
School of Information Engineering
Weifang Vocational College
Weifang, Shandong, China

Editor: Shibin David

Abstract

Sentiment orientation analysis of English text is a core issue within the realm of natural language processing. Traditional methods of word segmentation often encounter ambiguity when processing English language texts. In light of this, the present study introduces an innovative approach to English text sentiment analysis that utilizes a convolutional neural network (CNN) coupled with a U-network (U-Net). This method employs parallel convolutional layers to grasp the relationships and interactions between word vectors, which are then fed into a hierarchical attention network based on the U-Net to ascertain the sentiment polarity. The experimental outcomes demonstrate that the model achieves better accuracy on English review datasets, outperforming many existing sentiment analysis models.

Keywords:

English text sentiment analysis, CNN, U-Net

1 Introduction

Sentiment analysis, alternately termed opinion mining, pertains to the extraction of individuals' emotions, viewpoints, assessments, sentiments, and emotional states regarding services, products, entities, individuals, issues, events, subjects, and their characteristics (Gong et al. (2024); Alfreihat et al. (2024)). Analyzing the emotional inclination from textual data is a facet of sentiment analysis, aimed at discerning the author's emotional stance towards the subject matter discussed within the text. The knowledge-based technique, which was initially widespread in this domain, necessitates the formulation of intricate rules to enhance the computer's comprehension of human language (Delgadillo et al. (2024)). Consequently, this techniques is effective only on a limited dataset. As textual data proliferates, it becomes infeasible to process texts using the knowledge-based technique. Since the 1990s, machine learning techniques have started to gain prominence in the realm of textual sentiment analysis.

Nevertheless, machine learning methods mentioned above belong to shallow learning, characterized by relatively straightforward functional models and computational methods

(Najafabadi (2024)). This simplicity hinders their capacity to capture complex functions with limited sample sizes and computational resources, resulting in poor generalization capabilities and the labor-intensive requirement of manually selecting numerous data features. These limitations have led to a bottleneck for machine learning techniques in this particular endeavor.

Deep learning is capable of automatically extracting crucial features and their representations from raw data, effectively handling a range of sophisticated tasks. It offers significant advantages in terms of model building, representational power, and optimization capabilities. Among the deep learning models, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have emerged as particularly prominent. CNNs are adept at capturing localized structural information from data, while RNNs are designed to handle information that is structured in a sequence or series.

Over the recent years, hybrid models that integrate the architectures of both CNNs and RNNs have been developed, yielding remarkable outcomes in the realm of text sentiment analysis. The attention mechanism, a cutting-edge advancement in deep learning, enables the capture of the most salient textual features and enhances the model's structural efficiency. Employing deep learning models for the analysis of emotional inclination in text is currently a trending research avenue.

The present study posits that the emotional inclination of a text is shaped at both the sentence level and the word or character level. Initially, a text is an assembly of sentences, each carrying a distinct weight in contributing to the overall sentiment analysis. For instance, within a text that generally conveys a positive sentiment, a few sentences with negative connotations may not significantly impact the overall assessment and do not necessarily alter the final judgment. In a similar vein, sentences are constructed from words or characters, each exerting a varying degree of influence on the determination of the sentence's emotional tone. Current models seldom delve into the emotional nuances of text from this perspective, often neglecting the impact of the text's structural hierarchy and contextual content on sentiment analysis outcomes. To address this, the paper proposes a hierarchical sentiment analysis model that incorporates an attention mechanism. This mechanism filters out the most influential text information on sentiment analysis results from two different levels.

Additionally, the representation of word vectors plays a crucial role in text classification tasks. Recently, research into the granularity of word vectors has become increasingly refined, with some studies focusing on character-level analyses. English text is primarily represented at the character level, with words composed of 26 letters, where individual letters typically lack intrinsic meaning. Due to the unique characteristics of the English language, most text classification tasks involve word segmentation processes. However, this segmentation often results in fixed combinations of letters, which can lead to ambiguities and may not accurately reflect the correct forms of words. To address this issue, this paper employs the parallel convolutional layers of a convolutional neural network to learn word-level features from English text, bypassing the need for parsing trees and other syntactic analysis methods. This approach avoids the complexities associated with linguistic knowledge and intricate data preprocessing. Experimental results indicate that using the trained word-level vectors as the original features yields better performance compared to other methods.

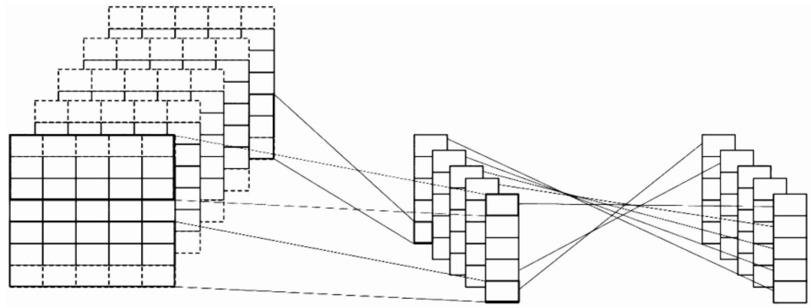

Figure 1: CNN convolutional layer and the convolution operation

2 Method

The operational flow of the model presented in this paper can be summarized as follows: Initially, convolutional neural networks are utilized to transform word vectors into a novel vector space, capturing the positional and contextual data of the words. Subsequently, a hierarchical attention mechanism is applied to discern the sequential information of sentences and texts, as well as their relative significance in determining the overall sentiment. In essence, the model integrates various deep learning methodologies to their fullest extent, taking into account both the local and global aspects of the text. This approach circumvents information loss and effectively identifies the most influential data for the outcome determination.

2.1 Parallel Convolutional Layers

Originally deployed in the realm of computer vision, CNN has, in recent years, demonstrated impressive performance in text classification as well. During text processing, conventional CNN models typically transform words into vector form, apply convolutional kernels of varying sizes and numbers to these vectors element-wise, and produce the final output after a sequence of operations such as convolution, pooling, dropout regularization. CNNs are adept at capturing local interrelationships between words within text, which is commonly employed in sentiment analysis tasks. As the objective of employing CNN in this study is to extract features from words within a single sentence for input into the subsequent layer's structure, only the convolution operation is utilized, which is illustrated in Figure 1.

2.2 Hierarchical Attention Network

The attention mechanism made its debut in machine translation tasks in 2014, and following its evolution over time, a variety of distinct variants have emerged.

Given learnable weight matrices W, U and V and the bias b, the influence score from the j-th input to the i-th output e_{ij} corresponding to h_j and s_{i-1} is computed via:

$$e_{ij} = V tanh(W h_j + U s_{i-1} + b). \tag{1}$$

where $tanh$ stands for the hyperbolic tangent function.

On the basis of e_{ij}, the attention a_{ij} is calculated by:

$$a_{ij} = \frac{exp(e_{ij})}{\sum_{k=1}^{T} exp(e_{ik})}. \tag{2}$$

where T is the total number of input data.

The attention a_{ij} indicates the semantic contribution of the corresponding word in a sentence. In a similar way, the semantic contribution of the corresponding sentence in the text can be measured. The sentiment polarity can be obtained by the softmax output of the weight sum between the semantic contribution and the corresponding features.

3 Experiments

3.1 Experimental Settings

Datasets, i.e., online movie review dataset (OMR) including 39800 reviews and online shopping review set (OSR) including 55700 reviews, are labeled by the inferred emotional scores on the basis of review scores, where review scores ranging 1-5 corresponds to labels ranging 0-1, with 0 indicating negative sentiment and 1 indicating positive sentiment.

Accuracy (Acc) and F1 score (Fs) are utilized to evaluate the performance of the model. Accuracy indicates the correct ratio of model analysis, which is computed via:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}. \tag{3}$$

where TP (True Positive) is the number of data with both positive labeled polarity and positive predicted polarity. TN (True Negative) is the number of data with both negative labeled polarity and negative predicted polarity. FP (False Positive) is the number of data with negative labeled polarity and positive predicted polarity. FN (False Negative) is the number of data with positive labeled polarity and negative predicted polarity. F1 score balances the evaluation from precision P and recall R:

$$Fs = \frac{2PR}{P + R}. \tag{4}$$

where $P = \frac{TP}{TP+FP}$ and $R = \frac{TP}{TP+FN}$.

Existing emotion analysis methods are regarded as the comparison methods, including ABCDM (Basiri et al. (2021)), AC-BiLSTM (Liu and Guo (2019)), CRNN (Wang et al. (2016)), and DPCNN (Johnson and Zhang (2017)). The results of comparison experiments are shown in Table 1, which demonstrates that the proposed model has the superior on the English text sentiment analysis task.

4 Conclusion

In response to the challenge of English text sentiment analysis, this study introduces an advanced sentiment analysis model that integrates CNN and U-Net into a framework. Initially, the CNN model is deployed to acquire a semantic representation vector of the text. Subsequently, the U-Net based hierarchical attention network is harnessed to distill semantic from various features of the text. Ultimately, the model conducts classification

Table 1: The results of comparison experiments

Method	*Acc* on OMR (%)	*Fs* on OMR (%)	*Acc* on OSR (%)	*Fs* on OSR (%)
ABCDM	90.90	90.59	93.25	93.29
AC-BiLSTM	84.76	84.33	91.15	91.08
CRNN	84.16	83.40	91.20	91.19
DPCNN	80.21	79.64	87.67	87.74
Proposed	93.16	93.17	94.78	94.92

and outputs the emotional category. To substantiate the efficacy of the proposed model, comparative experiments have been conducted using two distinct datasets: one comprising movie reviews and the other consisting of online shopping reviews. The outcomes demonstrate that the classification performance of our model surpasses that of other models under comparison. While the model has delivered promising results, it is accompanied by lengthy training durations. For future endeavors, efforts should be directed towards significantly decreasing the model's training time without compromising its performance.

Acknowledgments and Disclosure of Funding

The authors declare that there is no conflict of interest and no funding.

References

Manar Alfreihat, Omar Saad Almousa, Yahya M. Tashtoush, Anas M. R. Alsobeh, Khalid Mansour, and Hazem Migdady. Emo-sl framework: Emoji sentiment lexicon using text-based features and machine learning for sentiment analysis. *IEEE Access*, 12:81793–81812, 2024.

Mohammad Ehsan Basiri, Shahla Nemati, Moloud Abdar, Erik Cambria, and U. Rajendra Acharya. ABCDM: an attention-based bidirectional CNN-RNN deep model for sentiment analysis. *Future Generation Computer Systems*, 115:279–294, 2021.

Josiel Delgadillo, Johnson Kinyua, and Charles Mutigwe. Finsosent: Advancing financial market sentiment analysis through pretrained large language models. *Big Data and Cognitive Computing*, 8(8):87, 2024.

Lianting Gong, Xingzhou He, and Jianzhong Yang. An image-text sentiment analysis method using multi-channel multi-modal joint learning. *Applied Artificial Intelligence*, 38(1), 2024.

Rie Johnson and Tong Zhang. Deep pyramid convolutional neural networks for text categorization. In Regina Barzilay and Min-Yen Kan, editors, *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, pages 562–570, Vancouver, Canada, 2017. Association for Computational Linguistics.

Gang Liu and Jiabao Guo. Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing*, 337:325–338, 2019.

Maryam Khanian Najafabadi. Sentiment analysis incorporating convolutional neural network into hidden markov model. *Computational Intelligence*, 40(2), 2024.

Xingyou Wang, Weijie Jiang, and Zhiyong Luo. Combination of convolutional and recurrent neural network for sentiment analysis of short texts. In *Proceedings of the 26th International Conference on Computational Linguistics*, pages 2428–2437, Osaka, Japan, 2016. The COLING 2016 Organizing Committee.