Poster:Research on multi-feature fusion false review detection based on DistilBERT-BiLSTM-CNN

Jing Zhang
Xi'an University of Science and Technology
Lintong Qu, Xian Shi, China

Ke Huang
Xi'an University of Science and Technology
Lintong Qu, Xian Shi, China
huangkeup@163.com

ABSTRACT

In the context of rapid expansion of e-commerce and social media, fake reviews have become a prominent issue, misleading consumers and damaging the reputation of businesses. To more effectively identify and filter these false reviews, this paper introduces a multi-feature fusion strategy and proposes a composite architecture based on DistilBERT-BiLSTM-CNN. Experimental results on the Amazon dataset show that compared to the fake review detection methods of LSTM, HAN, BERT, RoBERTa and ELECTRA models, our model's method improved by approximately 4.2%, 3%, 2.6%, 1.3% and 0.8% respectively, achieving a maximum accuracy of 91.5%.

KEYWORDS

DistilBERT model, multi-feature fusion, false review detection, deep learning, text analysis.

ACM Reference Format:

Jing Zhang and Ke Huang. 2018. Poster:Research on multi-feature fusion false review detection based on DistilBERT-BiLSTM-CNN. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX).* ACM, New York, NY, USA, 2 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

In the field of fake review detection, traditional techniques mainly focus on the analysis of text content. For example, studies have shown that text-based techniques, such as natural language processing (NLP) and machine learning, can effectively identify specific language patterns and inconsistencies in fake reviews[**mir2023online9444444444444**]. However, these methods often overlook other important information behind the reviews, such as user behavior and timestamps. Recent studies have started to focus on multi-dimensional features, including user behavior and temporal factors, to enhance the accuracy and comprehensiveness of detection[**shishah2021fake722222222222**]. Based on these findings, our research proposes a multi-feature fusion strategy. This strategy integrates review text features, reviewer behavior features, and product information features. The advantage of this approach is that it can more comprehensively capture various aspects influencing fake reviews, thereby improving the accuracy and robustness of detection. Additionally, our research also applies the attention mechanism to multi-dimensional features not limited to text content, further enhancing the model's fine-grained analytical capabilities[**qian2021hierarchical83333333333333**].

Multi-feature fusion strategy: This study not only deeply analyzes textual information but also integrates review text features, reviewer behavior features, and product information features into the model. By adopting a multi-feature fusion strategy, the model can more comprehensively capture various aspects influencing fake reviews. Our research expands the application scope of the attention mechanism, applying it not only to textual content but also to other multi-dimensional features. This approach enables the model to more finely allocate weights to information from different sources, thereby improving accuracy and robustness in identifying fake reviews.

DistilBERT-BiLSTM-CNN composite architecture: We effectively combined DistilBERT, BiLSTM, and CNN, three neural network structures, to create a composite architecture. This design not only achieves advanced abstraction processing of textual information but also identifies local features and contextual relationships within the text. Crucially, by using DistilBERT as the pre-trained model, we have reduced the computational burden while maintaining high accuracy of the model.

Next, we will summarize the multi-feature fusion fake review detection method based on DistilBERT-BiLSTM-CNN. The experimental results show that our model has achieved a certain improvement in accuracy compared to other advanced models.

2 SOLUTION OVERVIEW

In this study, we recognize that the context of the review, the subject of the review, and the reviewer may influence the meaning of the same vocabulary in different reviews. Traditional research often adopts a method based on manual feature extraction, which involves first processing word vectors using neural network models to obtain textual representations, and then combining or fusing these representations with other non-semantic features. However, this approach somewhat neglects the deep interactions and dependencies between vocabulary in the text and non-semantic features. To overcome this challenge, we propose a strategy based on the attention mechanism, effectively combining non-semantic features with word vectors. In this strategy, the attention mechanism plays a key role, dynamically assigning weights to each feature, ensuring that each feature interacts appropriately with word vectors during the fusion process. This method not only fuses the information of textual and non-semantic features but also fully considers their

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

interactions, thus providing a more comprehensive and in-depth analytical.

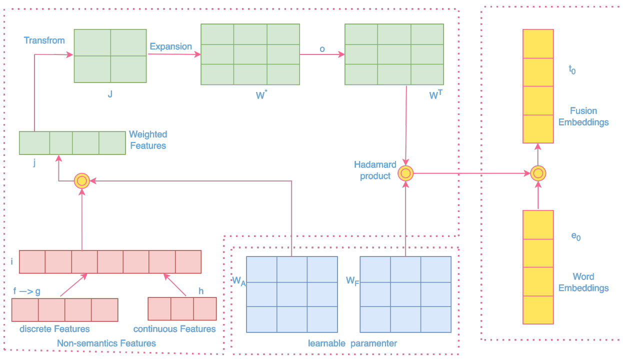

Figure 1: Framework Diagram of Multi-feature Fusion Strategy Based on Parameter Attention Mechanism.

Our method is shown in Fig.1 as taking non-semantic features and word sequences as inputs, with the output being word vectors fused with features. To incorporate non-semantic features into the word vectors of the text e_0, we use an attention function to process discrete features f and continuous features h. Discrete features f are transformed into continuous vectors through mapping. Assuming the set of all possible discrete feature values f_i is V_i, each possible value $u \in V_i$ is transformed into a real number vector $P(u) \in R^d$ in the mapping relationship P, where the hyperparameter d defines the dimension of the vector.

After concatenating the vector representations of continuous and discrete features, a feature vector $i = [h_1, h_2, s, g_1, g_2, s]$ is formed, which is then input into a fully connected layer:

$$j = W_A i + b \tag{1}$$

To generate the final fused feature vector t_i, assuming its dimension is n, the shape of the weight matrix W_F is $m \times n$, then the shape of the input vector W^+ should also be $m \times n$. First, the input vector is reshaped into a matrix of $p \times p$, where p and q are factors of m and n, respectively, and p and q are much smaller than m and n. Then, the matrix is repeatedly concatenated to obtain a matrix of $m \times n$:

$$W^* = \begin{bmatrix} J & S & J \\ S & S & S \\ J & S & J \end{bmatrix} \tag{2}$$

And it is compressed to the interval [0, 1]:

$$W^+ = \sigma(W^+) \in [0, 1]^{m \times n} \tag{3}$$

Finally, the compressed matrix is subjected to Hadamard operation with the learnable weight matrix, and then input into another fully connected layer, thereby obtaining a vector fused with feature information and semantic word information, represented as:

$$t_i = (W^+ \cdot W_F) e_i + b \tag{4}$$

In this study, we developed a fake review detection model based on DistilBERT-BiLSTM-CNN, which mainly consists of two core parts: the word vector representation module and the deep text feature extraction and classification module. The workflow of the model is as follows: First, a BERT-based pretrained model is used to convert text into word vectors for preliminary feature processing. This step lays the foundation for deep feature extraction. Then, the DistilBERT-BiLSTM-CNN network is utilized to extract deep contextual features and local key information of the text. The BiLSTM part enhances the understanding of text sequences, while the CNN part focuses on extracting key features from these sequences. Finally, after processing through a fully connected layer, softmax regression is used for classification, generating the final classification feature vector and outputting the predicted category.

Through this structure, the model effectively combines the powerful semantic understanding capabilities of BERT, the sequence processing ability of BiLSTM, and the local feature extraction ability of CNN, thereby achieving higher accuracy and efficiency in fake review detection tasks.

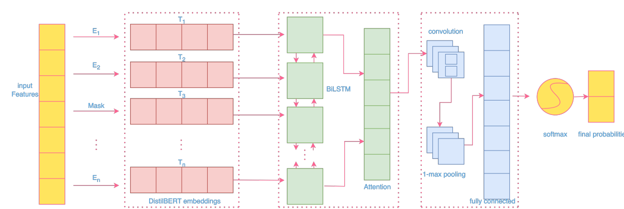

Figure 2: DistilBERT-BiLSTM-CNN Model Framework Diagram.

We have constructed a hybrid network structure based on DistilBERT-BiLSTM-CNN, As shown in Fig.2

3 CONCLUDING REMARKS

This study proposed an innovative method for detecting fake reviews, implementing a multi-feature fusion strategy within the DistilBERT-BiLSTM-CNN architecture. Through extensive experiments conducted on the Amazon dataset, we have made several key findings, highlighting the effectiveness of integrating a diverse set of features (including review text features, reviewer behavior features, and product information features).

4 ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (Grant No.61902311, 51904226, 52274137, 62172330), in part by the Post doctoral Research Foundation of China(Grant No.2019M663801), the Key R&D Foundation of Shaanxi Province(Grant No.2021SF-479), and the Scientific Research Project of Shaanxi Provincial Education Department (No.22JK0459).

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009