



T-Bert: A Spam Review Detection Model Combining Group Intelligence and Personalized Sentiment Information

Yue Shang¹ , Meiling Liu¹  , Tiejun Zhao², and Jiyun Zhou³ 

¹ School of Information and Computer Engineering, Northeast Forestry University,
Harbin 150006, China

{1095517318,mlliu}@nefu.edu.cn

² Department of Computer Science, Harbin Institute of Technology,
Harbin 150001, China

tjzhao@hit.edu.cn

³ Lieber Institute, Johns Hopkins University, Baltimore, MD 21218, USA

Abstract. The content of online comments largely affects users' willingness to purchase goods or services. Driven by interests, spam reviews continue to emerge to induce users maliciously. Most of the existing related work is based on the easy-camouflaged feature information, and the deep learning model is rarely used. The BERT model is prominent in various tasks in the NLP field, and whether it can be successfully applied to the spam review identification task has not been verified. In this paper, we propose a new research strategy for this task: the multi-dimensional representation combining group intelligence and users' personalized sentiment information can more effectively detect spam reviews. Through fine-grained sentiment analysis of reviews based on product dimension and user dimension, we effectively acquire group intelligence and user personalized sentiment, respectively; Based on the ability of BERT to model the embedding of text context information, the semantic information is acquired. Finally, the three are combined based on Triple Network structure to detect spam reviews. We conduct a large number of experiments on three public datasets and the recall rate and F1 value both exceed the results of state-of-the-art works, which proves the feasibility and effectiveness of our proposed strategy, and verifies the modeling ability of the BERT in the task of detecting spam reviews.

Keywords: Spam review detection · Sentiment analysis · BERT · Triple network

1 Introduction

Product reviews affect users' shopping behavior. According to the survey, 64% of users will read the reviews before buying goods, 87% of users choose to buy after

Supported by the National Natural Science Foundation of China under Grant 61702091 and the Fundamental Research Funds for the Central Universities under Grant No 2572018BH06.

© Springer Nature Switzerland AG 2021

I. Farkaš et al. (Eds.): ICANN 2021, LNCS 12895, pp. 409–421, 2021.

https://doi.org/10.1007/978-3-030-86383-8_33

reading good reviews, and 80% of users give up after reading bad reviews [1]. Fake reviews refer to reviews that are written to intentionally confuse the consumer [2]. However, research by the Washington Post [3] found that more than 60% of reviews of electronic products on Amazon.com were fake. Therefore, it is very important to automatically identify the authenticity of network platform information and provide users with more authentic information. Review information is text information, so the identification of spam reviews can be regarded as a text classification problem. Li et al. [4] propose a neural network composed of two convolutional layers combined with sentence importance weights for deceptive review detection. Liu et al. [5] based on the combination of bidirectional long short-term memory network and features, carry out fake review detection, which can well learn the long-distance correlation in the sequence.

However, the existing models are at a deadlock in the recognition effect, one of the reasons is that the embedding layer usually provides context-independent word-level features by Word2Vec or Glove models. Moreover, the spam review dataset is too small to implement task-based architecture. Therefore, it has a good potential to further improve the performance to generate context aware word vectors with the help of pre-trained language models on large-scale datasets.

In addition, the experimental results show that emotional features have a good effect in the recognition of fake reviews [6]. After conducting data mining on the Yelp fake review datasets, we find that reviewers usually describe many aspects of the product to express their opinions and convince others. Different users will produce a variety of fine-grained evaluations when evaluating the same product, which reflects the quality of the product in an all-round way. Because the spammers are not personal experience, the non-real information they posted may be different from the public evaluation. For example, for a restaurant, the real users have a negative evaluation on the dishes and a positive evaluation on the drinks. Spammers also have positive comments on drinks, but they are full of praise on dishes. J. Surowiecki points out in the book *The Intelligence of Crowds*: “In the right environment, a group has extraordinary intelligence, and this intelligence often beats the smartest person in the group” [7]. The group’s evaluation of a product aspect can represent the real level of the aspect.

Therefore, if we can mine the potential group intelligence in product reviews, and combine user’s emotional attitude towards product aspects, we can verify whether the user’s emotional attitude is true, and use the public intelligence to detect spam reviews more effectively.

In this paper, we fuse group intelligence with users’ personalized sentiment information and context semantic information to generate multidimensional representations for the identification of spam reviews, and propose a new model Triple BERT (T-Bert) based on the structure of Triple Network [8] and BERT component, which provides a new solution strategy for the task of spam review detection.

2 Related Work

2.1 Spam Review Detection

The task of spam review detection began in 2007 [9]. Spam review detection is a specific application of the general problem of deception detection, mainly using text and behavioral features. Behavior features include the number of good/bad reviews [10], the frequency of comments [12], etc.; text features include the length of the comment text [10,11], various vocabulary and syntactic features [13], etc. In addition, some works combine text features with behavior features. Wang et al. [14] combined the two as sentence representation based on CNN model, which solved the cold start problem in spam review detection. Wang et al. [15] used MLP to obtain user behavior features and CNN to obtain text language features, and combined them based on attention neural network to identify fake reviews. Yuan et al. [16] used hierarchical fusion attention mechanism to generate fusion text representation from the perspective of user and product, and based on TransH algorithm to model the relationship among user, product and review text, to generate more reliable review representation.

Previous work mostly based on Word2Vec or GloVe for word vector representation, but it is not enough to capture the complex semantic relevance in sentences. Recently, pre-trained language models such as ELMo and BERT have been shown to be effective in generating context-aware word vectors with the potential to further improve performance, and have been shown to be effective in a number of natural language processing applications, so far, however, no work has been done to apply BERT to the spam review detection task. In this paper, we use BERT as the basic model to construct the word vector and verify the performance of BERT model in this task.

2.2 Fine-Grained Sentiment Analysis

Fine-grained sentiment analysis is a challenging and significant subtask in sentiment analysis. Fine-grained sentiment analysis [17] aims to identify the sentiment polarity of specific aspects. This task enables users to evaluate the comprehensive sentiments of all aspects of a given product or service and have a more comprehensive understanding of its quality [18]. Fine-grained sentiment analysis can be subdivided into three categories: the first one is to detect the polarity of sentiment corresponding to a given aspect in a sentence [19,20], but it is difficult to be applied because fine-grained aspects need to be labeled in advance; the second is Aspect-oriented Opinion Words Extraction (AOWE) [21,22], which aims to extract the opinion words corresponding to a given aspect from the sentence; the last is End-to-End Aspect-based Sentiment Analysis (E2E-ABSA), whose goal is to jointly detect aspect terms/categories and corresponding aspect sentiment.

On the one hand, existing studies [6,25] show that using sentiment features can effectively identify fake reviews; on the other hand, as mentioned in the Sect. 1, in order to integrate group intelligence into the model and further improve the reliability of spam review detection, we conduct E2E-ABSA on spam

review data, so as to obtain the group sentiment corresponding to all aspects of the product. Meanwhile this measures the degree of deviation of user’s sentiment from the public’s sentiment. These two are used as the auxiliary information of spam review detection, which provides a new solution for this task.

3 Methodology

The structure of T-Bert is shown in Fig. 1. We regard spam review detection task as a binary task. Firstly, for each user’s (or product’s) reviews, we conduct fine-grained sentiment analysis on the sentences, and cluster the fine-grained aspects to get the user’s (or product’s) sentiment tendency in each fine-grained aspect, that is, the group sentiment tendency $G_i = \{a_{i1}, a_{i2}, a_{i3}, a_{i4}, a_{i5}, a_{i6}\}$ of product P_i , and the personal sentiment tendency $S_i = \{b_{i1}, b_{i2}, b_{i3}, b_{i4}, b_{i5}, b_{i6}\}$ of user U_j to product P_i . Secondly, given an input sentence $X_i = \{x_{i1}, x_{i2}, \dots, x_{iT}\}$ of length T , we encode it with the BERT component of the L Transformer layers to get a contextualized sentence representation $E^L = \{e_1^L, e_2^L, \dots, e_T^L\} \in \mathbb{R}^{T \times D}$, where D represents the dimension of the vector. Finally, we combine G_i, S_i and E^L to identify spam reviews. Our goal is to determine whether X_i is a spam review.

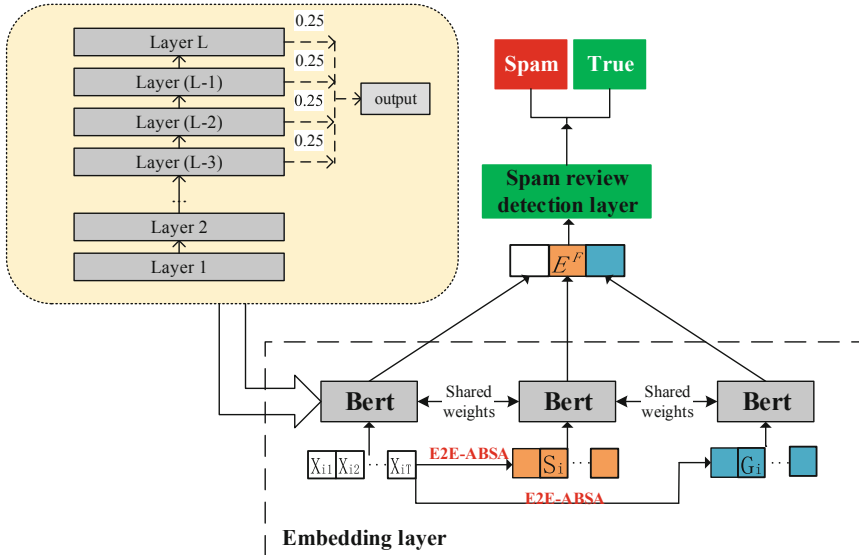


Fig. 1. Overall structure of T-Bert model. S_i stands for the user’s personalized sentiment, G_i stands for group intelligence.

3.1 Group Intelligence and User Personalized Sentiment

We extract the fine-grained aspects that users are concerned about from the reviews. The fine-grained aspect refers to the product attributes that contained in the user’s review. Since the fine-grained aspect is not marked in the spam review dataset, the amount of data that is too large to be manually marked, and it is difficult to define the marking standard, we use the transfer learning method to mark the fine-grained information. This research is based on the Yelp dataset, which includes restaurant reviews and a small number of hotel reviews. Therefore, we use the method in work [23] to train the fine-grained sentiment analysis model based on the data of SemEval 2016 [26], and use the model to label the Yelp dataset. Each review in the dataset is annotated to get a triple information $(A_i, W_i, POS/NEG)$, that is, the fine-grained aspect $A_i = \{A_{i1}, A_{i2}, \dots, A_{in}\}$ referred to in sentence X_i , and the corresponding sentiment word W_i for each fine-grained aspect A_{ix} , as well as this group of fine-grained sentiment tendency POS/NEG , POS represents positive sentiment, NEG represents negative sentiment. In order to obtain group intelligence and user personalized sentiment, we further analyze the annotated fine-grained sentiment information.

We use the labeling standards in the SemEval dataset to divide the fine-grained aspects into 6 categories: *restaurant*, *food*, *drink*, *service*, *ambience*, and *location*. First, de-duplicate and merge the fine-grained aspect words contained in all review sentences to obtain the fine-grained aspect word set ASP . Perform word frequency statistics on ASP , and select 10 seed words in each category to form a seed word set \bar{A} in order from highest to bottom. Second, use the Word2Vec model to train the Yelp review dataset to obtain the word vector model. Finally, based on the word vector model, the similarity between each seed word in each category in \bar{A} and ASP_i is calculated. If the average similarity is greater than the threshold α , then ASP_i belongs to this category. As shown in Table 1, the fine-grained aspect word set \bar{A} divided into 6 categories is generated according to the above steps.

Table 1. Part of the fine-grained aspect word set \bar{A}

Restaurant	Food	Drinks	Service	Ambience	Location
Dish	Meat	Wine	Services	Atmosphere	Downtown
Travelocity	Meal	Wines	Waitstaff	Environment	Branch
Hostess	Pancakes	Vino	Staff	Setting	Outpost
Meal	Dessert	Chardonnay	Serve	Interior	Hotspot
Product	Desserts	Prosecco	Servers	Atomsphere	Avenue
Stew	Seafood	Saki	Natured	Atomosphere	Bucktown
Adjacent	Dishes	Shiraz	Courteous	Decoration	Situated

We determine the sentiment polarity of each category in the review sentence based on simple rules. For example, in the category of *food*, if the

number of positive fine-grained words is greater than the number of negative fine-grained words, the sentiment of *food* is positive, and vice versa. From the product dimension, perform fine-grained sentiment analysis and clustering on all review information of the product P_i to obtain its group sentiment feature $G_i = \{a_{i1}, a_{i2}, a_{i3}, a_{i4}, a_{i5}, a_{i6}\}$, Where a_{i1} represents a certain category of group sentiment polarity. From the user dimension, the fine-grained sentiment analysis results of U_j 's evaluation of P_i are clustered according to \hat{A} , which is regarded as user's personalized sentiment feature $S_i = \{b_{i1}, b_{i2}, b_{i3}, b_{i4}, b_{i5}, b_{i6}\}$, where b_{ix} represents a certain category of user sentiment polarity.

3.2 Triple Bert

The BERT model is a new language model that uses bidirectional Transformers for pre-training on a large number of corpora, and performs amazingly in many tasks in the NLP field. We built a spam review detection model T-Bert based on the Triple Network framework and BERT.

Embedding Layer. We use the BERT component as the embedding layer of the T-Bert model. For each token X_{it} in sentence X_i , We add token embedding, segment embedding and position embedding to e_t , $t \in [1, T]$ to form the input feature $E^0 = \{e_1, e_2, \dots, e_T\}$ of the first branch of the embedding layer. Then L transformer layers are introduced to refine the token-level features layer by layer. Finally, the output E^L obtained by splicing the last four layers is the representation of the review sentence X_i .

$$E^L = 0.25 \times E^{L-1} + 0.25 \times E^{L-2} + 0.25 \times E^{L-3} + 0.25 \times E^{L-4} \quad (1)$$

In order to combine group intelligence, user personalized sentiment and text information for spam review detection, we use BERT component to transform the two dimensions of sentiment information constructed in Sect. 3.1. First, the two features G_i and S_i are Onehot mapped and normalized. Then, we pack each feature value in the S_i of the P_i as $E^{s0} = \{e_{s1}, e_{s2}, \dots, e_{s12}\}$, where $e_{st}, t \in [1, 12]$ is the combination of the token embedding, segment embedding, and position embedding corresponding to the input feature token. This is the second branch of the embedding layer. The input feature $E^{g0} = \{e_{g1}, e_{g2}, \dots, e_{g12}\}$ of the third branch of embedding layer is generated in the same way. Note that the BERT components of the first branch, the second branch and the third branch share weights. The calculation process is as shown below, where $E^{gl} \in \mathbb{R}^{12 \times D}$, $E^{sl} \in \mathbb{R}^{12 \times D}$ are the representation of group intelligence feature G_i and user sentiment feature S_i respectively.

$$\begin{aligned} E^{gl} &= \text{Transformer}_l(E^{g^{l-1}}), \\ E^{sl} &= \text{Transformer}_l(E^{s^{l-1}}). \end{aligned} \quad (2)$$

Spam Review Detection Layer. In order to identify spam reviews, we build four different spam review detection layers on the embedding layer to classify the feature representations obtained before. We concatenate E^L , E^{gl} and E^{sl} to form the input $E^F \in \mathbb{R}^{(T+24) \times D}$ of spam review detection layer.

Linear The obtained E^F is input into a max pooling layer. The most distinctive features in each sentence can be selected to form a sentence representation $h_L \in \mathbb{R}^D$, and then input into the linear classification layer. Finally, softmax function is used to calculate the probability of classification category as follow:

$$\begin{aligned} h_L &= \max_{dim=1} (E^F), \\ P &= \text{softmax}(h_L W_L), \end{aligned} \quad (3)$$

where $W_L \in \mathbb{R}^{D \times C}$, C is the number of categories.

Bidirectional Long Short-Term Memory (BiLSTM). BiLSTM is a combination of forward LSTM and backward LSTM, which can better capture bidirectional semantic dependencies. Input the obtained E^F into BiLSTM to obtain the task-specific hidden representation $h \in \mathbb{R}^{2H}$, where H is the hidden layer size in BiLSTM, and then obtain the predicted value P through the softmax function:

$$\begin{aligned} h &= \text{BiLSTM}(E^F) = [\vec{h}, \overleftarrow{h}], \\ P &= \text{softmax}(h W_2). \end{aligned} \quad (4)$$

Attention Network. The attention mechanism in seq2seq breaks the limitation that the encoder can only use the final single vector result, so that the model focuses on the input information that is more important for the output information. We use the attention mechanism to calculate E^F , extract the implicit features in sentences, focus on the words that are important for classification, and generate a specific representation $h_A \in \mathbb{R}^D$ of this task.

$$\begin{aligned} h_A &= \beta E^F, \\ \beta &= \frac{\exp(E_i')}{\sum_{n=1}^{T+24} E_n}, \\ E' &= \tanh(E^F W_a), \end{aligned} \quad (5)$$

where β is the score function that determines the importance of the words in the whole sentence, $W_a \in \mathbb{R}^{D \times D}$ is the transformation matrix. Similarly, a linear layer with softmax activation as before is stacked on the designed attention layer to output the prediction.

Convolutional Neural Network (CNN). The CNN model proved to be effective for NLP and achieved excellent results in semantic analysis [27]. In this paper, we use the convolution kernel of the CNN layer to perform a convolution operation on the review sentence representation E^L to obtain the hidden features $O_i \in \mathbb{R}^{f \times (T-k+1)}$ in the text.

$$\begin{aligned} O_i &= W \cdot E_{i:i+k-1}^L, \\ V_c &= \max_{0 \leq i \leq T-k} (O_i). \end{aligned} \quad (6)$$

where f is the channel for the convolution and k is the width of the convolution kernel. \cdot represents the dot product operation of the matrix, $i = 0, 1, 2, \dots, T - k$ and $W \in \mathbb{R}^{k \times D}$. The convolution core is repeatedly applied for the convolution operation and fed into the max pooling layer for filtering features.

Above is a process of feature extraction by a filter. In this paper, m filters of different sizes are used to extract as many features as possible, and then these features are spliced to get the review representation $h_{c1} \in \mathbb{R}^{m \times f}$. Then we combine the filtered token level text features with the sentence level output of BERT model and E^{tl} , E^{sl} to get the final sentence representation $h_c \in \mathbb{R}^{m \times (f+3D)}$. Finally, h_c is input into the linear layer with softmax activation function to get the classification result.

4 Experiment

4.1 Datasets and the Evaluation Metrics

In order to verify the effectiveness of the model, we conducted experiments on three public datasets: YelpChi [10], YelpNYC and YelpZIP [11]. The data are real business reviews of restaurants and hotels from different areas of the Yelp website. It can be found that the average sentence length of real review sentence is longer than that of spam review sentence because it involves fine-grained aspects description. There is no significant difference between spam reviews and real reviews when observed from sentence-level sentiment analysis.

We used precision, recall and F1 scores to evaluate the effectiveness of the model. The precision reflects the correctness of the model in predicting spam reviews, and the recall reflects the proportion of correctly predicted spam reviews by the model in all spam reviews. F1 score is the harmonic mean of precision and recall.

4.2 Baselines and Implementation Detail

In the comparison experiment, we compare the BERT-based model with several advanced methods in existence. ABNN [15] is a neural network based on attention mechanism, which uses MLP to obtain user behavior features and CNN to obtain text language features, and combines the two based on attention to identify spam reviews. HFAN [8] is a hierarchical fusion of attention among users, reviews and products to get a comment representation that integrates the three to classify comments. DFFNN [14] is a deep feedforward neural network, which combines bag-of-word/n-gram feature, word embedding and multiple emotion indicators of the review sentence as representation. In addition, we also compare the modeling effect of several spam review detection layers with different network structures and the influence of different sentiment features on the detection ability.

In the embedding layer, we use the pre-trained ‘‘BERT-base-uncasd’’ model, where the number of transformer layers $L = 12$, the hidden size $D = 768$, that is,

the sentence representation dimension is 768 and the sentence length $T = 200$. In the spam review detection layer, the learning rate is set to $2e - 5$, the dropout rate is set to 0.5 and the training batch size is 128. The hidden layer dimension of BiLSTM is set to 300. In convolution neural network, the size of convolution kernel channel is $f = 50$, and the width of convolution kernel increases from 1 to 11. A total of 11 filters with different sizes are used.

4.3 Results and Analysis

The Embedding Effect of the BERT Model: The experimental results are shown in Table 2 below. Compared with other methods without BERT model, BERT + Linear is not as good as the best model when using only text information as detection feature, however, the recall rate and F1 value are slightly different from other models that use a variety of information, which validates the performance of BERT model in the task of detecting spam reviews. It shows that the BERT model encoded by the association between any two tokens can generate a review representation with rich contextual information for the spam review detection layer.

Table 2. Experimental results of single BERT using only text information

	YelpChi			YelpNYC			YelpZIP		
	P	R	F1	P	R	F1	P	R	F1
ABNN	69.23	54.51	60.99	70.67	54.97	61.84	73.34	57.42	64.41
HFAN	73.03	57.21	64.16	74.65	59.42	66.17	78.45	65.49	71.39
DFFN	71.34	53.95	61.44	71.96	54.07	61.75	72.69	56.93	63.85
Bert+Linear	68.56	59.58	63.76	70.49	60.83	65.30	74.58	64.33	69.08
Bert+ATT	69.36	59.69	64.16	71.18	60.38	65.34	74.96	65.57	69.95
Bert+LSTM	70.09	60.30	64.83	71.95	61.87	66.53	75.38	66.18	70.48
Bert+CNN	71.51	60.73	65.68	72.37	62.37	67.00	76.23	67.39	71.54

Performance of Different Spam Review Detection Layers: The experimental results are shown in Table 2. When only text information is used as the clue of spam review detection, the precision, recall and F1 values of BERT + ATT, BERT + LSTM and BERT + CNN are higher than those of BERT + Linear. Therefore, the use of more powerful network structure can bring better effect for the spam review detection task than only using the linear layer. This result shows that merging context information is helpful to sequence modeling and can provide more effective sentence representation for text classification tasks.

Performance of Different Sentiment Information: The results are shown in Fig. 2(a), Fig. 2(b) and Fig. 2(c). S-BERT refers to Siamese BERT, which takes text information as the input of the first branch of the embedding layer and user personalized sentiment feature as the input of the second branch. The rest of the model structure is the same as that of T-Bert. As Fig. 2(a) shown, when different features are used as the potential thread for spam review detection, the precision of using sentiment features is not greatly improved compared with using only text information, but the recall rate and F1 value are greatly improved, which indicates that when fine-grained sentiment information is fused, the ability of the model to identify spam reviews is improved. From Fig. 2(b) and Fig. 2(c), we can find that the detection ability is further improved by combining the product and user dimensions, that is, combining the group intelligence with the user’s personalized sentiment, which verifies our previous hypothesis that the effective use of group intelligence can better detect spam reviews.

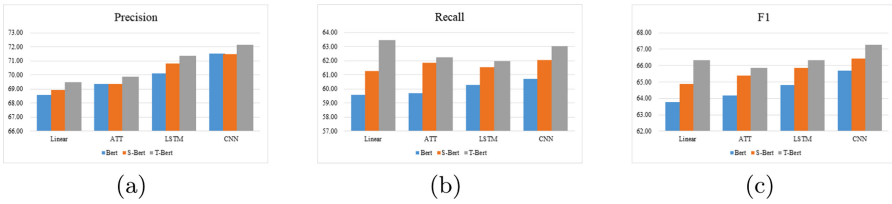


Fig. 2. Performance of different models on YelpChi dataset using different sentiment features.

Comparing with Table 2, it can be seen that the recall rate and F1 value of T-Bert have been greatly improved. Compared with the existing technology, the average recall rate and F1 value of the three data sets have been improved by 4.6% and 2.4% respectively. The experimental results verify the effectiveness and feasibility of the proposed strategy. But there is still room for improvement, the improvement of model’s precision is not so good. The reason is that: in order to obtain fine-grained aspect information annotation, transfer learning method is used. However, the accuracy of annotation can not reach 100%. The result of annotation further affects the accuracy of subsequent spam review detection. How to further improve the effect of the model is our next research plan.

5 Conclusion

In this paper, we propose a new research strategy for spam review detection task, and verify the effectiveness of BERT component in this task. Specifically, we propose a strategy to improve the effectiveness of spam review detection by using group intelligence and user personalized sentiment information. In order to effectively use the intelligence of the group, we combine the group intelligence and

the user personalized sentiment information with the text information to generate multidimensional representation, and propose a new model Triple BERT (T-Bert) based on the structure of Triple Network and BERT component. We explore the use of the BERT model as the embedding layer to generate review representations with rich contextual information, and to couple the BERT component with multiple neural models, a large number of experiments are carried out on three benchmark datasets to verify the effectiveness of the strategy proposed in this paper. The results show that BERT performs well in the task of spam review detection and improves the effectiveness of the T-Bert model.

References

1. Yuming, L., Xiaoling, W., Tao, Z., et al.: Review of research on quality inspection and control of user reviews. *J. Softw.* **03**, 506–527 (2014)
2. Ott, M., Choi, Y., Cardie, C., Hancock, J.T.: Finding deceptive opinion spam by any stretch of the imagination. In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pp. 309–319 (2011)
3. WASHINGTON POST. <https://www.washingtonpost.com/business/economy/how-merchants-secretly-use-facebook-to-flood-amazon-with-fake-reviews>. Accessed 23 Apr 2018
4. Li, L., Qin, B., Ren, W., Liu, T.: Document representation and feature combination for deceptive spam review detection. *Neurocomputing* **254**(254), 33–41 (2017)
5. Liu, W., Jing, W., Li, Y.: Incorporating feature representation into BiLSTM for deceptive review detection. *Computing* **102**(3), 701–715 (2019). <https://doi.org/10.1007/s00607-019-00763-y>
6. Hajek, P., Barushka, A., Munk, M.: Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining. *Neural Comput. Appl.* **32**(23), 17259–17274 (2020). <https://doi.org/10.1007/s00521-020-04757-2>
7. Surowiecki, J.: *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations* (2004)
8. Schroff, F., Kalenichenko, D., Philbin, J.: FaceNet: a unified embedding for face recognition and clustering. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815–823 (2015)
9. Jindal, N., Liu, B.: Review spam detection. In: *Proceedings of the 16th International Conference on World Wide Web*, pp. 1189–1190 (2007)
10. Mukherjee, A., Venkataraman, V., Liu, B., Glance, N. S.: What yelp fake review filter might be doing. In: *7th International AAAI Conference on Weblogs and Social Media, ICWSM 2013*, pp. 409–418 (2013)
11. Rayana, S., Akoglu, L.: Collective opinion spam detection: bridging review networks and metadata. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 985–994 (2015)
12. Kc, S., Mukherjee, A.: On the temporal dynamics of opinion spamming: case studies on yelp. In: *Proceedings of the 25th International Conference on World Wide Web*, pp. 369–379 (2016)

13. Dewang, R.K., Singh, A.K.: Identification of fake reviews using new set of lexical and syntactic features. In: Proceedings of the Sixth International Conference on Computer and Communication Technology 2015, pp. 115–119 (2015)
14. Wang, X., Liu, K., Zhao, J.: Handling cold-start problem in review spam detection by jointly embedding texts and behaviors. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (vol. 1: Long Papers), pp. 366–376 (2017)
15. Wang, X., Liu, K., Zhao, J.: Detecting deceptive review spam via attention-based neural networks. In: National CCF Conference on Natural Language Processing and Chinese Computing, pp. 866–876 (2017)
16. Yuan, C., Zhou, W., Ma, Q., Lv, S., Han, J., Hu, S.: Learning review representations from user and product level information for spam detection. In: 2019 IEEE International Conference on Data Mining (ICDM), pp. 1444–1449 (2019)
17. Jo, Y., Oh, A. H.: Aspect and sentiment unification model for online review analysis. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, pp. 815–824 (2011)
18. Sun, C., Huang, L., Qiu, X.: Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, vol. 1 (Long and Short Papers), pp. 380–385 (2019)
19. Dong, L., Wei, F., Tan, C., Tang, D., Zhou, M., Xu, K.: Adaptive recursive neural network for target-dependent twitter sentiment classification. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (vol. 2: Short Papers), pp. 49–54 (2014)
20. Tang, D., Qin, B., Feng, X., Liu, T.: Effective LSTMs for target-dependent sentiment classification. In: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp. 3298–3307 (2016)
21. Hu, M., Peng, Y., Huang, Z., Li, D., Lv, Y.: Open-domain targeted sentiment analysis via span-based extraction and classification. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 537–546 (2019)
22. Zhang, C., Li, Q., Song, D.: Aspect-based sentiment classification with aspect-specific graph convolutional networks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp. 4567–4577 (2019)
23. Zhang, C., Li, Q., Song, D., Wang, B.: A multi-task learning framework for opinion triplet extraction. In: Findings of the Association for Computational Linguistics: EMNLP 2020, pp. 819–828 (2020)
24. Li, X., Bing, L., Zhang, W., Lam, W.: Exploiting BERT for end-to-end aspect-based sentiment analysis. In: Proceedings of the 5th Workshop on Noisy User-Generated Text (W-NUT 2019), pp. 34–41 (2019)
25. Melleng, A., Jurek-Loughrey, A., Padmanabhan, D.: Sentiment and Emotion based text representation for fake reviews detection. In: Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019), pp. 750–757 (2019)

26. Pontiki, M., et al.: SemEval-2016 task 5: aspect based sentiment analysis. In: Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), pp. 19–30 (2016)
27. Yih, W., He, X., Meek, C.: Semantic parsing for single-relation question answering. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (vol. 2: Short Papers), pp. 643–648 (2014)