

The Strategy of Discriminating False Comments on the Internet by Fusing Probabilistic Topic and Word Vector Models

Fei Long

Foreign Language School, Harbin University of Commerce, China
longfei0420@outlook.com

Abstract: *With the acceleration of the social process of "Internet+", e-commerce has entered an era of rapid development. In response to the subjective judgments in current false information detection methods and the problem of low classification accuracy caused by the failure to extract semantic information hidden in text, this study proposes Latent Dirichlet Allocation - Word vector-Edge Convolutional Neural Network (LW-ECNN) Model. First, the probabilistic topic model is organically combined with the word vector model, and then the edge path structure is proposed to optimize the Convolutional Neural Network (CNN) discriminative classification model. Finally, the data-side optimization module is combined with the edge CNN classifier to form a web-based false comment discrimination model based on the probabilistic topic and word vector models. The results show that the probabilistic topic and word vector-based online false comment discrimination model has good results in detecting false comments, and can provide data reference for false comment detection by false detection-related departments, which is of great practical significance to assist the network environment cleaning work.*

Keywords: *Probabilistic topics, word vectors, online reviews, false detection, CNN.*

Received January 4, 2024; accepted May 15, 2024
<https://doi.org/10.34028/iajit/21/3/14>

1. Introduction

With the speedy growth of computer software and the prevalence of the Web, e-commerce ushered in an era of rapid development [22]. E-commerce has changed people's traditional shopping mode, and the scale of online shopping is expanding, gradually becoming the main shopping method of users nowadays [12]. While online shopping brings convenience to people's shopping, it also brings some problems. The nature of online shopping is different from traditional physical shopping, as buyers cannot directly touch the goods, and often can only understand the goods through merchant descriptions and buyer reviews [11]. In order to increase the sales of their products, some merchants make false propaganda and false reviews for their own products at any cost [29]. False reviews can lead to misjudgment of the product by users, resulting in a poor shopping experience and, in serious cases, affecting the credibility of the e-commerce platform [14]. The current methods for detecting false censorship mostly rely on statistical methods and machine learning techniques to study this issue. This method only utilizes statistical and computer linguistic methods to represent text features, and then uses supervised learners for text classification. Although it can also achieve good results. However, this method has subjective judgments in feature extraction of text and is unable to extract semantic information hidden in the text. The above situation makes it

necessary to further improve the classification accuracy of the model. Therefore, to further improve the accuracy and efficiency of identifying false online comments, help consumers effectively identify real comments, and improve their online shopping experience, this study starts with data processing and proposes a processing method that combines probabilistic topic models and word vector models to classify and optimize online comment data. The innovation of this study lies in:

- 1) Combining the probabilistic topic model with the word vector model, providing a new solution to the problem of large proportion differences in false comment data.
- 2) Exploring the adaptability impact of the Latent Dirichlet Allocation topic-Word vector (LW) module on general classifiers, providing reference for improving the efficiency of network false comment detection.

The research content mainly includes four parts. The second part is a review of the research status of domestic and foreign methods for detecting false online comments; the third part proposes a network false comment discrimination method based on probability topic and word vector. The first section establishes a joint model of probability topic model and word vector, and the second section constructs a network false comment discrimination model based on improved Convolutional Neural Network (CNN); the fourth part

validates the application effect of the proposed Latent Dirichlet Allocation-Word vector-Edge Convolutional Neural Network (LW-ECNN), network false comment discrimination model. The results indicate that the network false comment discrimination model based on probabilistic themes and word vectors has good false comment detection performance.

2. Related Work

The study of online false comment discrimination has become a hot topic in today's society, and many scholars at home and abroad have conducted relevant research. Alsubari *et al.* [3] believed that product reviews on e-commerce platforms had become an important judgment channel for online consumers. This study proposed the term frequency inverse file frequency method to extract features and their representations, and the proposed method outperforms comparable methods in terms of accuracy. Wu *et al.* [27] believed that false comments can seriously disrupt the market of e-commerce trading centers, so a new detection method was proposed to address this issue and achieved excellent results. Fang *et al.* [8] argued that existing false comment detection did not consider the impact of multiple information sources on false comment recognition. In order to overcome these problems, a false comment recognition method based on dynamic knowledge graph was proposed. De *et al.* [4] argued that fake news was not a new concept, but a common phenomenon in today's era. This study identified the main methods currently available for identifying fake news and how these methods can be applied in different situations. Moon *et al.* [23] and his team proposed a learning model that learns words from each category to distinguish between true and false comments for positive and negative evaluations.

Ruan *et al.* [25] argued that artificial false comments mimic real comments in many ways, and therefore propose an artificial false comment detection model. This model combined the AdaBoost model and short-term memory neural network to analyze user account and geographic location information, and achieved an accuracy of over 80% on the Yelp dataset. Kauffmann *et al.* [16] believed that consumer data was extremely important data information, and therefore proposed a framework to automatically analyze these comments. The proposed model utilized sentiment analysis to analyze online comments on well-known e-commerce websites, enabling the use of natural language processing techniques to detect and delete false comments. The empirical results showed that the proposed model had positive implications for more comprehensive decision-making. Budhi *et al.* [5] used machine learning classifiers to detect false comments and introduced two sampling methods to improve the accuracy of false comment classes on balanced datasets. The proposed method can achieve 89% accuracy in

random undersampling and oversampling on CNNs. Mohawesh *et al.* [20] found that online reviews have a significant impact on decision-making in the field of commodity buying and selling. Based on this, this study introduced an analytical model to address the concept drift problem in detection and conducted experiments using multiple real-world datasets. The results showed that there was a significant negative correlation between concept drift and detection performance, which had a positive significance in improving detection performance. Plotkina *et al.* [24] created two sets of comments (false and true) from an experiment of 1041 respondents and compared them with psycholinguistic deception cues. The resulting automated tools explained the value and motivation of comments and detected deceptive comments with an accuracy of 81%. Kolhar [17] compared four supervised machine learning algorithms and found that the Support Vector Machine (SVM) algorithm outperformed other algorithms, achieving high accuracy in text classification and false comment detection.

Through the studies of many research scholars, it can be found that the research on the detection of online fake comments is insufficient at this stage, and there is still room to improve the efficiency of online fake comment detection. Therefore, the present study is based on probabilistic themes and word vectors for the discrimination of online fake comments, which has a certain promotion effect on the development of online fake comment detection.

3. A Method for Discriminating False Comments on the Web by Fusing Probabilistic Themes and Word Vectors

3.1. Probabilistic Topic Model Combined with Word Vector Model

The proportion of false online comments in online comments is relatively low, and the number of online comments is extremely large, requiring the use of machine learning methods to accurately judge online comments [1, 15]. Research has shown that the Latent Dirichlet Allocation (LDA) probabilistic topic model can effectively integrate with user emotions and has good application effects in spam comment detection. Therefore, it can be applied to the field of false comments. In order to further improve the detection efficiency of machine learning, this study starts with data processing and proposes an LDA probability topic model combined with word vector processing methods. The word vector model and LDA model respectively complete the representation of text content from the granularity of semantics and themes, which can enrich the dimensions of text content expansion. Therefore, it has been widely used in the field of text content expansion. The structure of the LDA model is shown in Figure 1.

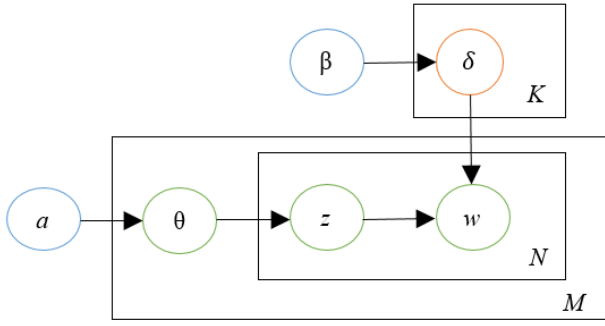


Figure 1. LDA probabilistic subject model structure.

In a document there are several parts, each part consists of N words, i.e., each topic consists of N words [2]. In the Potential LDA model, the probability distribution parameter θ between documents and topics, and the probability distribution parameter δ between topics and words are two core parameters. To determine these parameters, mathematical methods need to be used for derivation and calculation. Through a series of probability and statistical theories, the specific distribution forms of these parameters can be obtained, laying the foundation for modeling the theme of the document. The joint probability distribution of two parameters is shown in Equation (1).

$$P(\theta, \delta, w, z) = \prod_d p(\theta^d) \prod_t p(\delta^t) \prod_d \prod_t (\theta_t^d)^{m_t^d} \prod_d \prod_v (\delta_v^t)^{n_v^t} \quad (1)$$

In Equation (1), w denotes the word in the document, z denotes the topic, d represents the document, t represents the topic, and v denotes the number of words present in the dictionary. Equation (1) is simplified using the Dirichlet distribution property as shown in Equations (2) and (3).

$$D(a_1 \dots a_k) = \frac{\Gamma(a_1) \dots \Gamma(a_k)}{\Gamma(a_1 + \dots + a_k)} \quad (2)$$

$$D(a_1 \dots a_k) = \frac{\Gamma(a)^k}{\Gamma(ka)} \quad (3)$$

In Equation (2), a_1, \dots, a_k represents the Dirichlet parameter. When Equation (2) satisfies $a_1 = a_2 = \dots = a_k$, Equation (3) holds, and using Equation (3) to simplify Equation (1), the result is shown in Equation (4).

$$P(\theta, \delta, w, z) = \prod_d \left[\frac{\Gamma(T\alpha)}{\Gamma(\alpha)^T} \prod_t (\theta_t^d)^{m_t^d + \alpha - 1} \right] \prod_t \left[\frac{\Gamma(T\beta)}{\Gamma(\beta)^V} \prod_v (\delta_v^t)^{n_v^t + \beta - 1} \right] \quad (4)$$

The marginal probability can be derived from Equation (4) as shown in Equation (5).

$$P(z, w) = \prod_d \left[\frac{\Gamma(T\alpha)}{\Gamma(\alpha)^T} \frac{\prod_t \Gamma(m_t^d + \alpha)}{\Gamma(\sum_t m_t^d + T\alpha)} \right] \prod_t \left[\frac{\Gamma(T\beta)}{\Gamma(\beta)^V} \frac{\prod_v \Gamma(n_v^t + \beta)}{\Gamma(\sum_v n_v^t + V\beta)} \right] \quad (5)$$

In Equation (5), m denotes the number of documents, β denotes the hyperparameter of Dirichlet distribution, and the solution models of parameters θ and δ are derived by combining Equation (5) with the full

conditional probability formula, as shown in Equations (6) and (7).

$$\theta_t^d = \frac{\hat{m}_t + \alpha}{\sum_t \hat{m}_t + T\alpha} \quad (6)$$

$$\delta_v^t = \frac{\hat{n}_v + \beta}{\sum_v \hat{n}_v + V\beta} \quad (7)$$

The construction of probabilistic topic model helps to improve the speed and accuracy of data screening processing, and cooperate with the word vector model to generate vector data as the data input of judgment model. The commonly used Word2vec word vector model has two architectural approaches, namely Continuous Bag of Words (CBOW) and Skip-gram, and the model structure of CBOW is shown in Figure 2. Word2Vec is a commonly used model for training word vectors. This tool implements the mapping of words from high-dimensional space to low-dimensional space, which can convert words into vector form and be recognized by computers; And through the calculation of word vectors, the similarity between words can be calculated. The text vector constructed by Word2Vec achieves feature reduction and contextual semantic expression, but lacks global semantic information. The LDA theme model focuses on the construction of the overall semantic features of the text set. Therefore, the two models have good adaptability, and their combination can improve the representation ability of text vectors.

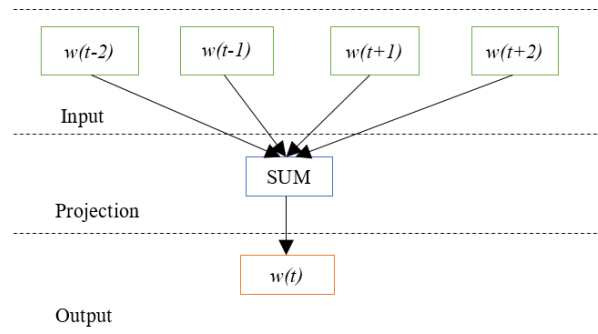


Figure 2. Model structure of CBOW.

As can be seen from Figure 2, the input of CBOW consists of feature words and their upper and lower associated words, and the output is a word vector of a specific word. The CBOW model is a typical bag-of-words model with equal word vector distance between the input and output, and the complexity is calculated as shown in Equation (8).

$$q = g * h + h * j \quad (8)$$

In Equation (8), g denotes the length of the window in Input, h denotes the dimension of Projection, and j denotes the number of words in the training dictionary. The model structure of Skip-gram is shown in Figure 3.

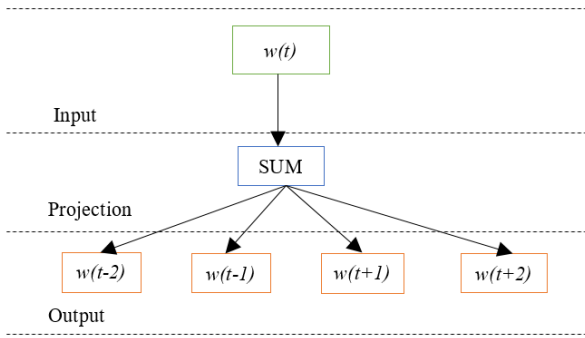


Figure 3. Model structure of skip-gram.

As can be seen from Figure 3, the model structure of Skip-gram is just the opposite of the CBOW model, where its input is the word vector of specific words and its output is the upward and downward associated word vector of feature words. This change leads to its higher computational complexity, but the results are more accurate, and the complexity of the Skip-gram model is shown in Equation (9).

$$q = z * (h + h * j) \tag{9}$$

To obtain more accurate prediction results, the Skip-gram model is chosen as the training model for the word vector model in the study. In this model, the comment data is accepted by the output layer, and the training is done by the implicit layer to derive the feature word vectors of $w(t)$, and the output of the neighboring word vectors is done in the output layer. The LDA is combined with the Word2vec model to form the L-W word vector processing module to optimize the data processing of the network false comment discrimination model, and its processing flow is shown in Figure 4.

See Figure 4, the L-W word vector processing module first accepts pre-cleaned data, then performs LDA modeling of the real comments to derive the $max(\theta)$ parameters and generates new real comments. Then the fake comments are trained with the generated new real comments using the Word2vec module and word vectors are generated. Finally, the processed

feature vector data is fed into the discriminative classification model along with the $max(\theta)$ parameters.

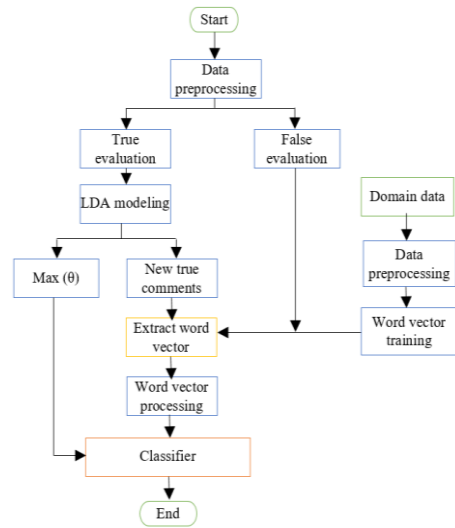


Figure 4. Processing flow chart of L-W word vector module.

In addition to optimization at the data processing end, the study also proposes optimization strategies for the classification algorithm model to improve the accuracy of the discriminative model. Convolution possesses a strong learning classification capability and is often applied in various feature classification processing tasks [18, 26]. Human learning usually involves several processes, including preview, learning, review, consolidation, and reinforcement, in order to proficiently apply knowledge. However, using ordinary convolutional neural network structures can only extract data features through convolutional kernels, and gradually reduce errors through backpropagation algorithms, lacking the ability to relearn the data. Therefore, the network structure of convolutional neural networks has been improved by adding a side path to the network to enhance its ability for later reinforcement learning. On top of the original CNN network structure, add a side path learning network to enhance its relearning ability. The optimized ECNN structure is shown in Figure 5.

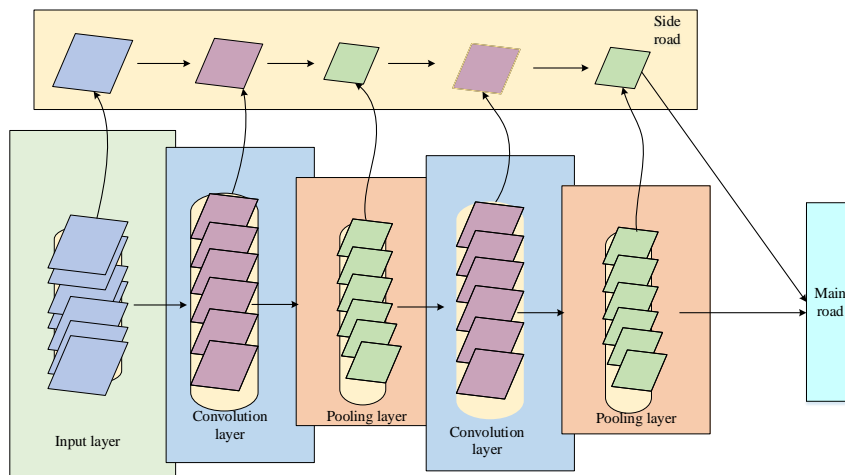


Figure 5. ECNN model structure diagram.

3.2. Web False Comment Discriminant Model Design

As shown in Figure 5, the side-path network is combined by layer-by-layer extraction of the original CNN network, so the side-path structure has the same convolutional layer and pooling layer as the main path. After the word vector is input into the ECNN discriminative model, it is convolved by the convolutional layer to produce multiple feature maps, each of which consists of several neurons. After several convolutional kernel operations for extraction, deeper features of the data are obtained. Let the data dimension of the word vector be D , which is processed and input into the matrix E of the ECNN model and the matrix F for $a \times a$. Let the weight matrix of the convolution kernel be F , its size be $c \times c$, the parametric term be b_1 , and the step size of each convolution operation be taken as 1. After the convolution operation, the feature matrix H , which is calculated as shown in Equation (10).

$$H_{ij} = R[\sum_{i=1}^d \sum_{j=1}^d (E_{ij} F_{ij}) + b_1] \quad (10)$$

In Equation (10), H_{ij} represents an element value in the matrix H , R represents the ReLu function, and E_{ij} and F_{ij} represent the corresponding element values in E and F , respectively. The size of the feature matrix H is calculated as shown in Equation (11).

$$z = (a - c + 1) \times (a - c + 1) \quad (11)$$

In Equation (11), z represents the size of the feature matrix H . The convolutional neural network performs convolutional processing by scanning the data layer by layer through the perceptual field of view. The receptive field is the area corresponding to a certain step of the convolution kernel each time it moves. The acceptance field coupled with the weight sharing mechanism greatly reduces the parameters in the training process. The number of parameters to be trained after each layer of convolution is set to np , which is calculated as shown in Equation (12).

$$np = (in_{cm} \times s_{ck} + 1) \times ou_{cm} \quad (12)$$

In Equation (12), in_{cm} denotes the number of feature maps at the input, s_{ck} denotes the convolutional kernel size, and ou_{cm} denotes the number of feature maps at the output. After the convolutional layer, the data has been processed, the feature extraction has been completed, but its output neurons have not been reduced, so it still needs to be reduced by the pooling layer [7]. The pooling layer is also connected to the previous feature map through a local region, which is the pooling window, also called the pooling core [13, 28]. The structure of the alternating connection of the convolutional and pooling layers allows the data to continuously extract features and then reduce the dimensionality to obtain the size of the pooled feature map S_{pl} , as shown in Equation (13).

$$S_{pl} = \frac{S_{cko}}{S_{pw}} \quad (13)$$

In Equation (13), S_{cko} denotes the size of the feature map output from the convolution kernel, and S_{pw} denotes the pooling window size [6, 21]. The study uses hybrid pooling as shown in Equation (14).

$$L_{ij} = \alpha(\max_{i=1, j=1} (F_{ij}) + b_2) + (1 - \alpha) \left(\frac{1}{c} (\sum_{i=1}^c \sum_{j=1}^c F_{ij}) + b_2 \right) \quad (14)$$

In Equation (14), L denotes the pooled eigenface matrix, F denotes the pooling function, c denotes the pooling window size, and α denotes a random number in the interval of $[0, 1]$. Since each convolutional layer is equipped with multiple convolutional kernels, each convolutional kernel generates a feature map that extracts data features from multiple layers and then filters the information further through the pooling layer [10, 19]. The ReLu activation function is usually used in the fully connected layer to avoid overfitting and make the network more robust. Finally, the connected information is transferred to the output layer, and then the softmax classifier parameters are set to output the results [9]. The data processing module is combined with the classification discriminator module to obtain the web false comment discriminator model based on probabilistic topics and word vectors, and its structure is shown in Figure 6.

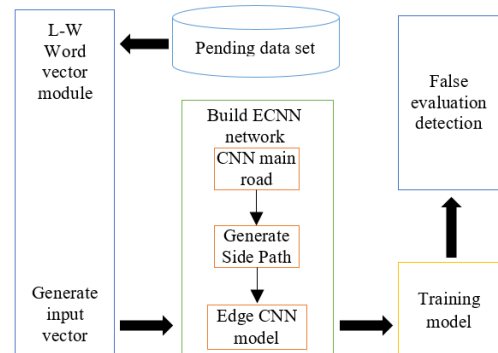


Figure 6. A discriminant model of false internet comments based on probability subject and word vector.

As can be seen from Figure 6, the L-W word vector module first processes the original dataset to generate word vector data and feeds it into the CNN discriminator model, which constructs the basic model of the main path and generates the side learning model, which is constructed into the ECNN discriminator model. Finally, the ECNN model is then trained and parameter tuned to complete the detection of false comments.

4. Experiments and Analysis of the Web False Comment Discriminative Model

The experiment was conducted under Windows 10 operating system, using Pycharm development tool and

Python development language to build the model and complete the validation and analysis of the model, the dataset used e-commerce reviews Amazon and restaurant reviews Yelp. Amazon dataset after data pre-processing, containing 6739 fake reviews and 56871 real reviews, a total of 63610 reviews. It was randomly split in the proportions of 8:2, of which 80% was used for testing and 20% for verifying. The download link for the Amazon dataset is: <https://github.com/niainio/amazon/index.com>. The download link for the Yelp dataset is: <https://www.yelp.com/dataset/documentation/main>.

The validity of the L-W word vector module was first verified by taking word vector dimension values of 50, 75, 100, 125, 150, 175, and 200, respectively, to observe the effect of the number of word vectors selected on the accuracy of the model, and the results obtained are shown in Figure 7.

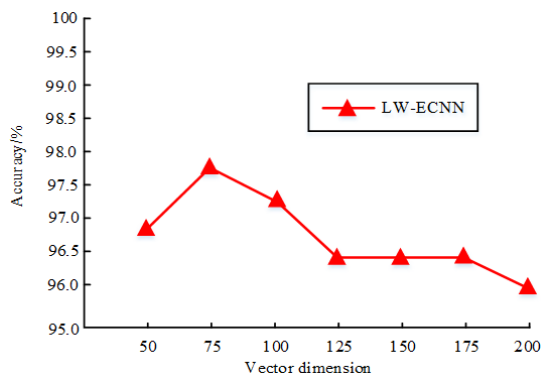


Figure 7. Experimental results under different vector dimensions.

See Figure 7, the model achieves the highest discriminative accuracy of 97.8% at a vector dimension of 75. The accuracy of the model first increases as the vector dimension increases, reaching a peak when the vector dimension increases to about 75. Subsequently, the model accuracy decreases as the vector dimension increases. The model accuracy is more stable at the vector dimension of 125-175, and this condition may be related to the insufficient number of experiments sets and the insufficient content of the experiment sets. Based on the experimental results, the study set the word vector dimension of the L-W model to 75, and continued to observe the effect of the number of topic-words selected on the LDA modeling, comparing the logarithmic perplexity indexes of different number of topics when the number of topic-words was 20, 25, 30 and 35, and the results are shown in Figure 8.

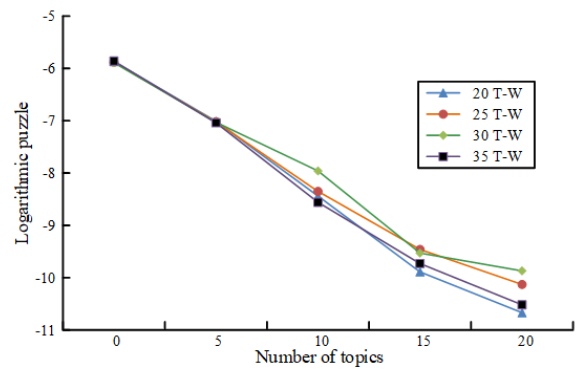


Figure 8. The influence of the number of topics on logarithmic puzzle under different topic-word numbers.

See Figure 8, when the number of topics is lower than 5, there is almost no difference in the logarithmic perplexity of each topic-word setting. When the number of topics is 5-15, 30T-W is in the dominant position, followed by 25T-W. When the number of topics exceeds 15, 25T-W achieves some advantages over 30T-W, and both outperform the other settings, and when the number of topics reaches 20, it is still 30T-W that has the best performance. Therefore, the optimal topic-word of the model is set to 30. After completing the model configuration, the proposed model is compared with the commonly used classification models SVM, LR and MLP to observe the differences in the performance indexes of Accuracy, F1-score, Precision and Recall under the same experimental conditions. The results of the before-and-after comparison experiments for each model to access the data processing module of L-W word vector model are shown in Figure 9.

See Figure 9, the Accuracy, Precision, Recall, and F1-score performance metrics of each classification algorithm are improved after accessing the L-W word vector module. In the Accuracy metric, the L-W module improves the SVM and CNN algorithm models the most, followed by MLP. In the Precision metric, the L-W module improves the SVM algorithm model the most, and improves the LR model the least. In the Recall metric, the L-W module has a large impact on each model and substantially improves the recall of the classification algorithm model. In the F1-score metric, the L-W model has the largest improvement on the SVM model, followed by the CNN model. The above results indicate that the L-W word vector model proposed in the study can effectively process the data information and substantially improve the accuracy and efficiency of each classification algorithm. For the ECNN algorithm model proposed in the study, there is a 0.34 improvement in Accuracy metric, 0.18 improvement in Precision metric, 0.24 improvement in Recall metric, and 0.14 improvement in F1-score metric. The results obtained from the experiments comparing the LW-ECNN model proposed in the study with other false comment discrimination models under different data sets are shown in Table 1.

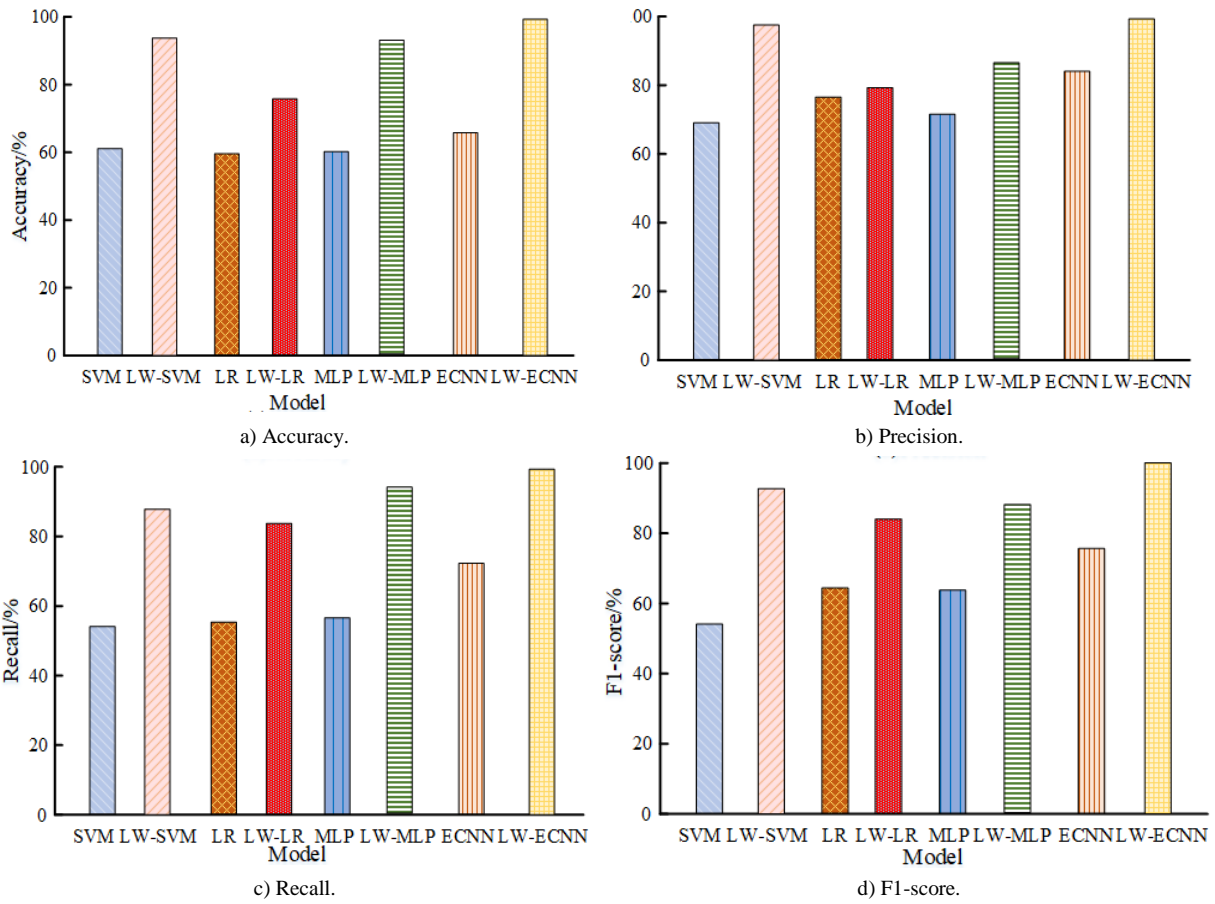


Figure 9. The performance of each model changes before and after the L-W word vector module is connected.

Table 1. Performance comparison of various false comment discrimination models under different data sets.

/	Amazon				Yelp			
	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
LW-SVM	97.6	96.3	89.4	87.6	95.3	94.6	88.4	86.1
LW-ECNN	98.9	97.4	93.5	92.6	98.4	96.8	92.5	93.3
LW-CNN	96.5	93.2	86.7	84.2	94.3	90.5	86.7	88.4
LW-ACNN	98.3	96.8	94.6	91.5	97.8	95.4	91.6	91.7

The LW-ECNN model has 98.9% accuracy, 97.4% precision, 93.5% recall and 92.6% F1 value for Amazon dataset, which are 2.4%, 4.2%, 6.8% and 8.4% higher than LW-CNN model respectively. The LW-ECNN model has 98.4% accuracy, 96.8% precision, 92.5% recall, and 93.3% F1 value under the Yelp dataset, which are 4.1%, 6.3%, 5.8%, and 4.9% higher than the LW-CNN model, respectively. This result shows that the proposed ECNN model effectively improves the algorithmic accuracy of the CNN model. The performance of each algorithmic model under the Amazon dataset is slightly better than that under the Yelp dataset, and the situation is related to the previous parameter tuning settings, where the tuning parameters are more inclined to discriminate false reviews for the Amazon dataset. The test results under the Amazon dataset in Table 1 are visualized and analyzed as shown in Figure 10.

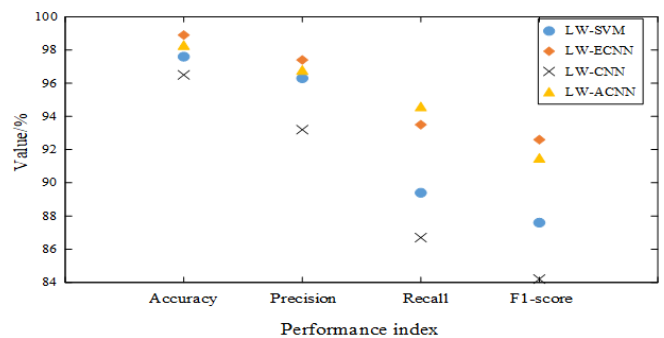


Figure 10. Performance comparison of different models under Amazon dataset.

See Figure 10, the LW-ECNN model proposed by the study performs best in all metrics, and only the Recall metric in the Amazon dataset test is lower than the LW-ACNN model, mainly because the LW-ACNN model combines the attention mechanism and carries out the first and last segmentation processing, which improves the recognition efficiency at the expense of certain accuracy, and also has a higher recall. However, the method is inferior to the LW-ECNN model proposed in

the study in terms of overall accuracy. The LW-ECNN model has an accuracy rate of 98.9% with the Amazon dataset, 2.4% more accurate than the LW-CNN and 1.3% more accurate than the LW-SVM model. The accuracy rate of the LW-ECNN model is 97.4%, 4.2% more accurate than the LW-CNN model and 1.1% more accurate than the LW-SVM model. The recall rate of the LW-ECNN model is 93.5%, 6.8% more than the LW-CNN model and 4.1% more than the LW-SVM model. The F1 value of the LW-ECNN model is 92.6%, 8.4% more than the LW-CNN model and 5.0% more than the LW-SVM model. The test results under the Yelp dataset in Table 1 were visualized and analyzed as shown in Figure 11.

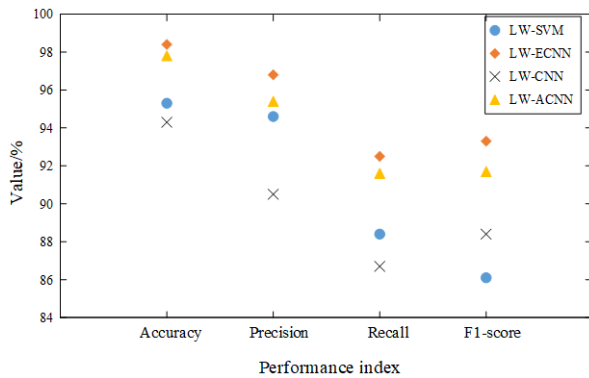


Figure 11. Performance comparison of different models under Yelp dataset.

See Figure 11, the LW-ECNN proposed in the study performs best in all indicators in the Yelp dataset test, and this result shows that the proposed ECNN effectively improves the algorithmic accuracy of the CNN model. The LW-ECNN model has an accuracy rate of 98.4% with the Yelp dataset, 4.1% more than the LW-CNN model and 3.1% more than the LW-SVM model. The accuracy rate of the LW-ECNN model is 96.8%, 6.3% more than the LW-CNN and 4.1% more than the LW-SVM. The recall rate of the LW-ECNN is 92.5%, 5.8% more than the LW-CNN and 4.1% more than the LW-SVM model. The F1 value of the LW-ECNN is 93.3%, 4.9% more than the LW-CNN model and 4.1% more than the LW-SVM model. 5.8% higher than the LW-CNN model and 4.1% more than the LW-SVM model. The F1 value of the LW-ECNN model is 93.3%, 4.9% more than the LW-CNN model and 7.2% more than the LW-SVM. The results show that the proposed ECNN model improves the performance of online fake comment detection in terms of both data processing and classifier effectiveness.

In order to further verify the performance of the LW-ECNN model, this study selected three other popular models for comparative verification. The comparative models come from three types of false comment detection models in references [6, 21, 28], which are the DRI-RCNN model based on recursive convolutional neural networks, the New Integrated Model (NIM), and the Unsupervised and End to End Model (UEM) based

on unsupervised learning. The experiment was conducted on the Yelp dataset, and the final results are shown in Figure 12. In the figure, the detection accuracy and F1 value of the LW-ECNN model are higher than those of the DRI-RCNN, NIM, and UEM models. This further validates the superior performance of the LW-ECNN model in detecting false comments.

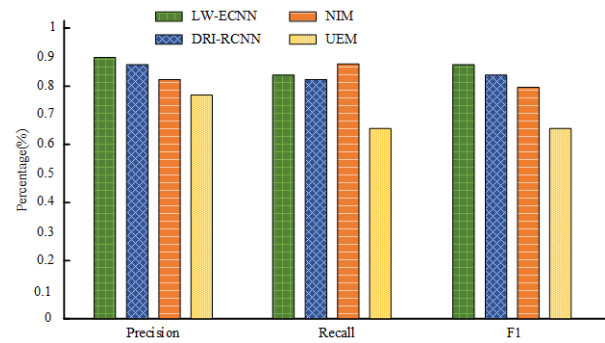


Figure 12. Validation results of four methods on Yelp dataset.

Table 2. Actual application effect of LW-ECNN model.

Model	Accuracy (%)	Non-fake			Fake		
		P (%)	R (%)	F1 (%)	P (%)	R (%)	F1 (%)
NIM	83.11	80.23	79.11	79.98	79.12	79.03	79.16
UEM	84.11	83.11	82.14	80.12	85.11	81.45	80.46
DRI-RCNN	95.12	94.11	90.12	89.96	94.11	90.12	89.96
LW-ECNN	97.26	96.58	92.23	92.59	97.11	95.11	94.11

To highlight the practical application performance of the LW-ECNN model, this study took a certain e-commerce platform as an example and collected 500 real comments and 800 false comments for performance testing. The test results are shown in Table 2. From the data in the Table 2, it can be seen that the LW-ECNN model has a detection accuracy of 97.26% for both false and non-false comments, which is better than the other three models. In addition, the LW-ECNN model showed significant advantages in the P-value, R-value, and F1 value of two types of comment detection. As a result, the LW-ECNN model has shown strong applicability in the detection of actual false comments.

5. Conclusions

This study focuses on the problem of detecting fake comments on the Web, applying probabilistic themes combined with word vector models to optimize the comment data processing, and applying split convolutional neural network classification models to optimize the classifier accuracy. The results show that the proposed L-W word vector model substantially improves the accuracy of the ECNN algorithm model by 0.34 in Accuracy metric, 0.18 in Precision metric, 0.24 in Recall metric, and 0.14 in F1-score metric. The proposed LW-ECNN model improves the accuracy of the ECNN algorithm model in both the proposed LW-ECNN model performs best in all the metrics with an average correct rate of 98.65% in both test datasets. Compared with the current popular DRI-RCNN model, NIM, and UEM models, the LW-ECNN model has

higher detection accuracy and F1 value than the first three. The study explores the possibility of improving the detection of online fake comments in terms of data processing and classifier performance, and provides ideas for improving the effectiveness of online fake comment detection. There are also some shortcomings in this study, such as the dataset used in the experiment is not sufficient, and it is expected that more datasets will be used to validate the study in the subsequent research to improve the accuracy and practicality of the study.

Conflict of Interest

The authors have no relevant financial or non-financial interests to disclose.

Data Availability Statement

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Fundings

The research is supported by: 2022 “Innovation” Project Support Plan for Teachers of Harbin University of Commerce. (Project No.: 22YYB252); Ministry of Education Humanities and Social Sciences Youth Foundation for Research Funds: A Study on Discourse Semantic Representation System and Resource Construction for Language (No. 23YJCZH150).

References

- [1] Aljadani E., Assiri F., and Alshutayri A., “Detecting Spam Reviews in Arabic by Deep Learning,” *The International Arab Journal of Information Technology*, vol. 21, no. 3, pp. 495-505, 2024. <https://doi.org/10.34028/iajit/21/3/12>
- [2] Alsharif N., “Fake Opinion Detection in an e-Commerce Business Based on a Long-Short Memory Algorithm,” *Soft Computing*, vol. 26, no. 16, pp. 7847-7854, 2022. <https://doi.org/10.1007/s00500-022-06806-5>
- [3] Alsubari S., Deshmukh S., Alqarni A., Alsharif N., and Aldhyani T., “Data Analytics for The Identification of Fake Reviews Using Supervised Learning,” *Computers, Materials and Continua*, vol. 70, no. 2, pp. 3189-3204, 2022. <https://doi.org/10.32604/cmc.2022.019625>
- [4] Beer D. and Matthee M., “Approaches to Identify Fake News: A Systematic Literature Review,” in *Proceedings of the Integrated Science in Digital Age*, Batumi, pp. 13-22, 2020. https://doi.org/10.1007/978-3-030-49264-9_2
- [5] Budhi G., Chiong R., Wang Z., and Dhakal S., “Using a Hybrid Content-Based and Behaviour-Based Featuring Approach in a Parallel Environment to Detect Fake Reviews,” *Electronic Commerce Research and Applications*, vol. 47, no. 1, pp. 101048, 2021. <https://doi.org/10.1016/j.elerap.2021.101048>
- [6] Cao C., Li S., Yu S., and Chen Z., “Fake Reviewer Group Detection in Online Review Systems,” in *Proceedings of the International Conference on Data Mining Workshops*, pp. 935-942, 2021. DOI: 10.1109/ICDMW53433.2021.00122
- [7] Chen L. and Zhu H., “Behavior Prediction Based on a Commodity Utility-Behavior Sequence Model,” *Machine Learning with Applications*, vol. 9, pp. 100314, 2022. <https://doi.org/10.1016/j.mlwa.2022.100314>
- [8] Fang Y., Wang H., Zhao L., Yu F., and Wang C., “Dynamic Knowledge Graph Based Fake Review Detection,” *Applied Intelligence*, vol. 50, pp. 4281-4295, 2020. <https://doi.org/10.1007/s10489-020-01761-w>
- [9] Fowler A. and Montagnes B., “Distinguishing between False Positives and Genuine Results: The Case of Irrelevant Events and Elections,” *The Journal of Politics*, vol. 85, no. 1, pp. 304-309, 2023. <https://doi.org/10.1086/719636>
- [10] Hameleers M., Meer T., and Vliegthart R., “Civilized Truths, Hateful Lies? Incivility and Hate Speech in False Information-Evidence from Fact-Checked Statements in the US. Information,” *Communication and Society*, vol. 25, no. 11, pp. 1596-1613, 2022. <https://doi.org/10.1080/1369118X.2021.1874038>
- [11] He S., Hollenbeck B., and Proserpio D., “The Market for Fake Reviews,” *Marketing Science*, vol. 41, no. 5, pp. 896-921, 2022. <https://doi.org/10.1287/mksc.2022.1353>
- [12] Himangshu P. and Nikolaev A., “Fake Review Detection on Online E-commerce Platforms: A Systematic Literature Review,” *Data Mining and Knowledge Discovery*, vol. 35, no. 5, pp. 1830-1881, 2021. <https://doi.org/10.1007/s10618-021-00772-6>
- [13] Hou J. and Zhu A., “Fake Online Review Recognition Algorithm and Optimisation Research Based on Deep Learning,” *Applied Mathematics and Nonlinear Sciences*, vol. 7, no. 2, pp. 861-874, 2021. <https://doi.org/10.2478/amns.2021.2.00170>
- [14] Hui L., “A Review of Research on Identification of False Reviews in E-Commerce,” *Journal of Management and Humanity Resources*, vol. 3, pp. 9-15, 2020. <http://dx.doi.org/10.22457/jmhr.v03a02102>
- [15] Jain P., Pamula R., and Srivastava G., “A Systematic Literature Review on Machine Learning Applications for Consumer Sentiment Analysis Using Online Reviews,” *Computer Science Review*, vol. 41, pp. 100413, 2021. <https://doi.org/10.1016/j.cosrev.2021.100413>

- [16] Kauffmann E., Peral J., Gil D., Ferrández A., and Sellers R., "A Framework for Big Data Analytics in Commercial Social Networks: a Case Study on Sentiment Analysis and Fake Review Detection for Marketing Decision-Making," *Industrial Marketing Management*, vol. 90, pp. 523-537, 2020.
<https://doi.org/10.1016/j.indmarman.2019.08.003>
- [17] Kolhar M., "E-Commerce Review System to Detect False Reviews," *Science and Engineering Ethics*, vol. 24, pp. 1577-1588, 2018.
<https://doi.org/10.1007/s11948-017-9959-2>
- [18] Kumar N. and Kumar R., "The Application of Artificial Intelligence in Electronic Commerce," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 12, pp. 1679-1682, 2021.
- [19] Li L., Fan L., Atreja S., and Hemphill L., "'HOT' ChatGPT: The Promise of ChatGPT in Detecting and Discriminating Hateful, Offensive, and Toxic Comments on Social Media," *ACM Transactions*, vol. 18, no. 2, pp. 1-36, 2024.
<https://doi.org/10.1145/3643829>
- [20] Mohawesh R., Tran S., Ollington R., and Xu S., "Analysis of Concept Drift in Fake Reviews Detection," *Expert Systems with Applications*, vol. 169, pp. 114318, 2021.
<https://doi.org/10.1016/j.eswa.2020.114318>
- [21] Mohawesh R., Xu S., Springer M., Al-Hawawreh M., and Maqsood S., "Fake or Genuine? Contextualised Text Representation for Fake Review Detection," *arXiv preprint*, arXiv: 2112.14343, 2021.
<https://doi.org/10.48550/arXiv.2112.14343>
- [22] Mohawesh R., Xu S., Tran S., Ollington R., and Springer M., "Fake Reviews Detection: A Survey," *IEEE Access*, vol. 9, pp. 65771-65802, 2021.
<https://doi.org/10.1109/ACCESS.2021.3075573>
- [23] Moon S., Kim M., and Iacobucci D., "Content Analysis of Fake Consumer Reviews by Survey-Based Text Categorization," *International Journal of Research in Marketing*, vol. 38, no. 2, pp. 343-364, 2021.
<https://doi.org/10.1016/j.ijresmar.2020.08.001>
- [24] Plotkina D., Munzel A., and Pallud J., "Illusions of Truth-Experimental Insights into Human and Algorithmic Detections of Fake Online Reviews," *Journal of Business Research*, vol. 109, pp. 511-523, 2020.
<https://doi.org/10.1016/j.jbusres.2018.12.009>
- [25] Ruan N., Deng R., and Su C., "GADM: Manual Fake Review Detection for O2O Commercial Platforms," *Computers and Security*, vol. 88, pp. 101657, 2020.
<https://doi.org/10.1016/j.cose.2019.101657>
- [26] Wang X., Zhou T., Wang X., and Fang Y., "Harshness-Aware Sentiment Mining Framework for Product Review," *Expert Systems with Applications*, vol. 187, pp. 115887, 2022.
<https://doi.org/10.1016/j.eswa.2021.115887>
- [27] Wu Y., Ngai E., Wu P., and Wu C., "Fake Online Reviews: Literature Review, Synthesis, and Directions for Future Research," *Decision Support Systems*, vol. 132, pp. 113280, 2020.
<https://doi.org/10.1016/j.dss.2020.113280>
- [28] Zhang W., Du Y., Yoshida T., and Wang Q., "DRI-RCNN: An Approach to Deceptive Review Identification Using Recurrent Convolutional Neural Network," *Information Processing and Management*, vol. 54, no. 4, pp. 576-592, 2018.
<https://doi.org/10.1016/j.ipm.2018.03.007>
- [29] Zhou X. and Zafarani R., "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities," *ACM Computing Surveys*, vol. 53, no. 5, pp. 3395046, 2020.
<https://doi.org/10.1145/3395046>



Fei Long, born on March 11, 1983, female, native to Heilongjiang, 2003-2010 Harbin Engineering University Bachelor's and Master's degree in Linguistics and Applied Linguistics. 2015-2018 PhD in Linguistics and Applied Linguistics,

Computational Linguistics, Communication University of China. 2010 present Associate Professor at Harbin University of Commerce. Publication status of academic papers, publication status of academic works and textbooks, research status of scientific research projects, patent status, and academic awards status. Papers: Analysis of the Development Trends of Language Science in the Era of Artificial Intelligence, Chinese Journal of Education (CSSCI), March 2022. Data Security Analysis and Research Based on Big Data Technology, Chinese Journal of Security Sciences (Chinese Core), September 2022. The Time Course of Chinese Compound Word Recognition as Revealed by ERP data Language, Recognition and Neuroscience (SSCI search), August 2022. Academic works: A Study on the Computability of Discourse Meaning [M] Beijing: People's Daily Publishing House, May 2020. (1/1). Linguistics and Innovative Translation [M] Beijing: Modern Publishing House 2016.4. (2/3). Projects: Research on the Intentional Inference Mechanism of Business News Discourse Based on Cognitive Intelligence, a Key Research Project for Economic and Social Development in Heilongjiang Province. Speech Semantic Representation and Resource Construction for Language Intelligence, Ministry of Education Humanities and Social Sciences Youth Project. From the perspective of discourse research on language policy of publicity and reporting in the context of the "the Belt and Road", Heilongjiang Philosophy and Social Sciences Project.