Received: 17 September 2023 | Revised: 10 October 2023 | Accepted: 20 October 2023 | Published online: 26 October 2023

RESEARCH ARTICLE

Artificial Intelligence and Applications
2024, Vol. 2(2) 135–140
DOI: 10.47852/bonviewAIA32021743

BON VIEW PUBLISHING

ERNIE and Multi-Feature Fusion for News Topic Classification

Weisong Chen[1], Boting Liu[2] and Weili Guan[3,*]

[1]*School of Civil Engineering, Sichuan Agricultural University, China*

[2]*School of Computer, Electronics and Information, Guangxi University, China*

[3]*College of Digital Economics, Nanning University, China*

Abstract: Traditional news topic classification methods suffer from inaccurate text semantics, sparse text features, and low classification accuracy. Based on this, this paper proposes a news topic classification method based on Enhanced Language Representation with Informative Entities (ERNIE) and multi-feature fusion. A semantically more accurate representation of text embedding is obtained by ERNIE. In addition, this paper extracts word, context, and key sentence based on the news text. The key sentences of the news are obtained through the TextRank algorithm, which enables the model to focus on the content points of the news. Finally, this paper uses the attention mechanism to realize the fusion of multiple features. The proposed method is experimented on BBC News. The experimental results show that we achieve classification accuracies superior to those of the compared methods, while validating the structural validity of the proposed method. The method in this paper has a positive effect on promoting the research of news topic classification.

Keywords: news topic classification, ERNIE, multi-feature fusion, attention mechanism, TextRank

1. Introduction

With the rapid development of Internet technology, the Internet has gradually become the main platform for the public to read news. However, nowadays, the Internet platform is filled with a large number of unstructured and unthematic news texts, which affects the news reading experience of Internet users to a certain extent. In the news topic classification method, the news text will be transformed into a vector expression understood by the deep learning model. The semantic accuracy and feature strength of the text vectors directly affect the final classification accuracy (Kaur & Bajaj, 2016). Therefore, it is of great practical significance to improve the semantic accuracy and feature strength of the news text to achieve a higher news topic classification accuracy rate and to improve the news reading experience of Internet users.

For text vectorization tools, in 2013 Google team proposed Word2vec (Mallik & Kumar, 2024) word embedding model, Word2vec is based on Continuous Bag-of-Words (CBOW) (Fati et al., 2023) algorithm and Skip-gram (Wang et al., 2023) algorithm to train word vectors. CBOW and Skip-gram algorithm are both based on the idea of center word prediction, and the trained word vectors have contextual information, are more general, and have achieved good results in various tasks. Later, language models such as Global Vectors for Word Representation (GloVe) (Hossain et al., 2023) and FastText (Saravanan et al., 2022) appeared, which further improved the semantic accuracy of word vectors based on Word2vec. However, the word vectors generated by either Word2vec or GloVe and FastText cannot express multiple meanings of a word, and such word vectors have only one vector expression in different contexts, which is obviously not in line with the actual semantics. In recent years, there are large-scale pre-trained language models such as BERT (Alammary, 2022) and XLNet (Wang & Zhang, 2022) based on transformer (Tezgider et al., 2022), which have achieved excellent results in various NLP tasks based on a large amount of corpus pre-training and fine-tuning with the downstream tasks. BERT cleverly applies the attention mechanism (Liu et al., 2022) to make the word vectors have dynamic semantic expressions, which improves the semantic accuracy of the text. However, the masking prediction strategy of BERT makes it inconsistent in the pre-training phase and fine-tuning phase, which affects the performance of the downstream tasks. XLNet improves on BERT by proposing permutation language modeling instead of BERT's autoencoder language modeling (AE) of BERT, which effectively improves the masking inconsistency problem of BERT. However, both BERT and XLNet belong to word masking prediction, and word masking will destroy the wholeness of some words, e.g., masking "Bei" or "Jing" alone will destroy the complete meaning of "Beijing," which will reduce the pre-training. This reduces the semantic accuracy of the text in the pre-training phase. To solve this problem, Baidu proposes Enhanced Representation through Knowledge Integration (ERNIE) (Zhang et al., 2019), a pre-training language model based on full-word masking. Full-word masking retains the original semantics of the words, so the semantic accuracy of the text generated by ERNIE is higher.

*Corresponding author: Weili Guan, College of Digital Economics, Nanning University, China. Email: 2113391032@st.gxu.edu.cn

ERNIE introduces external entity knowledge to enhance the prediction effect of pre-training. Based on the above, ERNIE is chosen as the vectorization tool for word sequences instead of Word2vec and BERT models.

In text classification tasks, a greater number of features imply a larger set of elements that the model can consider. Satisfactory predictions can be achieved by equipping the model with an ample supply of text features. In the context of the Chinese language, a word serves as the fundamental unit of a word and represents the smallest unit with semantic significance (Liu et al., 2023). Therefore, through the utilization of words, we can address semantic limitations and augment the textual attributes of news articles. Furthermore, identifying the key sentences within news articles allows us to swiftly grasp the core points of events. Similarly, the incorporation of key sentences into the training of deep learning models follows the same principle. By introducing key sentence features from news articles, we enhance the model's capacity to comprehend news content and facilitate more precise predictions regarding news topic categories. Based on the above, we propose a new method for categorizing news topics. The contribution of this paper is shown below.

(1) We replace traditional language models such as Word2vec with ERNIE. More semantically, accurate text feature vectors are obtained by ERNIE.
(2) This paper obtains the key sentences of each news paragraph through the TextRank algorithm to allow the model to capture the main points of the news to understand the text.
(3) We fused word, context, and key sentence with attention mechanism to achieve importance differentiation between multiple features.

2. Literature Review

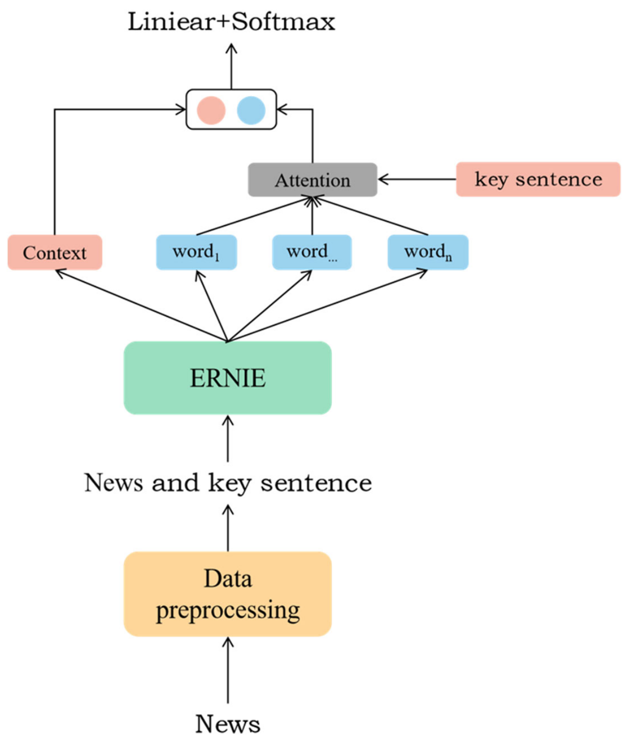
News topic classification is a hot issue in NLP, and many researchers have worked to achieve higher news classification accuracy. The literature by Lin et al. (2021) proposes a news topic classification method based on BERT, which achieves better classification accuracy than models such as recurrent neural network (RNN). The literature by Dogru et al. (2021) proposes a news classification method based on doc2vec and convolutional neural network (CNN). The feature vectors of news text are obtained by Doc2vec, and the text features are further extracted by CNN (Soni et al., 2023). The literature by Agarwal et al. (2023) proposes a news classifier based on machine learning algorithms. The news text is wordized by term frequency-inverse document frequency (Zhou, 2022) and then machine learning algorithms such as random forest (Jalal et al., 2022) are used to get the prediction results. The literature by Bei et al. (2023) proposes a news topic classification method based on Word2vec and CNN. Firstly, the news text is vectorized by Word2vec features, and then the features are further selected by CNN, and finally the classification results are obtained. The literature by Wang (2022) proposes a recurrent convolutional neural network (RCNN)-based news classification method. Firstly, the news text is passed through N-gram to get the text feature vector, and then input into RCNN (Lai et al., 2015) to get the predicted categories. The literature by Zhang et al. (2022a) proposes a news text classification method based on BERT with feature projection network. The news text context information is extracted by BERT, and the news classification effect is strengthened by combining the feature projection network. The literature by Zhang et al. (2022b) proposes a poetic sentiment analysis method based on ERNIE and multi-feature fusion, which introduces BERT and ERNIE to extract text features, respectively, and finally multi-feature fusion to enrich text feature information. The literature by Bao et al. (2022) proposed a text classification method based on ERNIE. The text was vectorized by ERNIE, then passed through the multi-channel network composed of Deep Pyramid CNN (DPCNN) and Bi-directional Long Short-Term Memory (BiLSTM), and finally fused the output of each channel to achieve multi-feature fusion. The literature by Li et al. (2021) proposes a text classification method based on ERNIE and RNN and further extracts text context information through RNN.

The above methods utilize machine learning or deep learning to accomplish news topic classification. However, the feature representation of the language model they apply is not strong enough. Moreover, there is no combination of more features to enhance the textual features of news. Based on this, the proposed method combines ERNIE and multi-feature fusion, which is a new method for news topic classification.

3. Proposed Methodology

The structure of the proposed method is shown in Figure 1. The proposed method is mainly divided into data preprocessing layer, ERNIE layer, and multi-feature fusion layer. The details of each part will be presented subsequently. In this paper, the reason for combining ERNIE with multiple features is to enhance the quality of text representation while introducing multiple features to increase the information content of the text. This allows the model to have a more comprehensive understanding of news text and ultimately make more accurate determinations of news topics.

Figure 1
ERNIE and multi-feature fusion for news topic classification



3.1. Data preprocessing

First, the noise words within the news text are removed. Second, the key sentences of the news text are extracted using the TextRank algorithm, which is a text summarization algorithm, and its core algorithmic idea is to use the graph network to generate the weighted graph nodes of each word. If there is a relationship between two words, edges are established between the two word nodes, and the weights of each node are updated iteratively during training, and the update formula for the weights of each node is as follows.

$$WS\{V_i\} = (1-d) + d^* \sum_{v_i \in InV_i} \frac{w_{ji}}{\sum_{v_k \in OutV_j} w_{jk}} WS\{V_j\} \tag{1}$$

where $WS\{V_i\}$ and $WS\{V_j\}$ represent the weight values of word i and word j. V_i and V_j represent the nodes of i and j in the graph. InV_i and InV_j represent the in-degree set of i and the out-degree set of j, respectively. d is the damping factor, which is usually set to 0.85, indicating that the probability that the point points to another node is 85%. In fact key sentences are the highest ranked sentences in TextRank. All sentences are constructed into a graph by TextRank, which is then trained to constantly update the importance of the sentences, ultimately resulting in a sentence importance ranking. The formula for calculating the similarity between each sentence node is as follows:

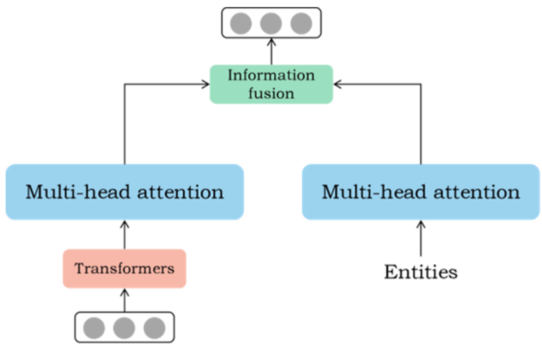
$$\text{similarity}(S_i,S_j) = \frac{|\{w_k\}|w_k \in S_i \,\&\, w_k \in S_j|}{\log(|S_i|) + \log(|S_j|)} \tag{2}$$

where S_i and S_j are the two sentence nodes and w_k is the word between the two sentences; the whole Equation (2) is calculating the content repetition between the two sentences. After the TextRank algorithm processed to get the key sentences of each news text, here the key sentence with the highest weight value is selected.

3.2. ERNIE

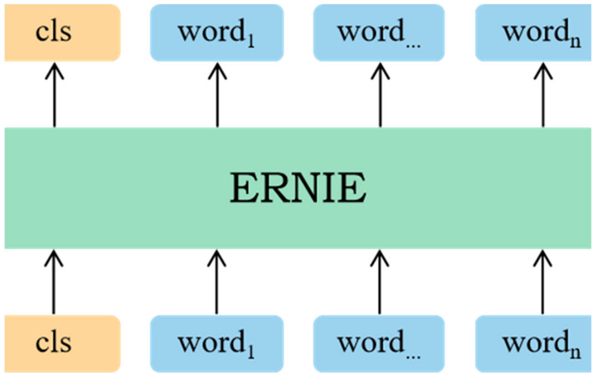
ERNIE is a large language model that evolved from BERT. Comparing BERT and XLNet, the biggest improvement of ERNIE is based on whole-word training and the introduction of entity knowledge. Whole-word training is when pre-training tasks are performed based on predictions based on all words instead of individual words. Whole-word pre-training avoids destroying the overall meaning of words, which makes ERNIE have semantically more accurate textual representations. In addition, entity information fusion is a very important innovation, which can help the model to understand the real meaning of entity words as shown in Figure 2.

Figure 2
ERNIE's knowledge fusion



As shown in Figure 2, the entity information is the personal name or place name with special significance. ERNIE integrates the external entity information with the prediction task after extracting the information from the attention of multiple people, thus enhancing the semantic representation ability of ERNIE. In this paper, ERNIE is used as a vectorization tool for word features, and contextual information of word sequence is extracted and input to the next layer, as shown in Figure 3.

Figure 3
ERNIE's news text processing



As shown in Figure 3, news text is entered into ERNIE in terms of words. cls are special characters that represent the contextual meaning of the current text. Therefore, the cls vector is used as the context vector in this paper. Compared to BERT, ERNIE focuses on improvements during the pre-training stage, so ERNIE has a parameter count similar to BERT, with approximately over 100 million parameters.

3.3. Multi-feature fusion

First, we build an attention network based on key sentences. The formula is shown below.

$$a = \text{Softmax}\left(\frac{QK^T}{d}\right) \tag{3}$$

$$C = A \otimes V \tag{4}$$

In Equation (3), Q is the key sentence vector, K is the news content vector, d is the scaling factor, and Softmax is the normalization function. Equation (3) is used to calculate the attention score of news text based on key sentences, and then combined with Equation (4) to achieve word feature fusion. Finally, the context vector of news is combined with the content vector of news after attention update. To better explain the process of multi-feature fusion in this paper, the relevant algorithms are shown in Table 1.

4. Experimental Result

4.1. Experimental datasets

In this paper, comparison and ablation experiments were carried out on BBC News (Greene & Cunningham, 2006) and R8. The detailed information of the dataset is shown in Table 2.

Table 1
Multi-feature fusion

Algorithm 1 Multi-feature fusion

Input: Word sequence vector($Word_{Vec}$), key sentence vector(Key_{Vec}), Context vector($Context_{Vec}$).
Output: Multi-feature fusion vector (MF_{Vec}).
for epoch ← 0 to Epoch **do**
d ← Dimensions of WordVec
Q ← WordVec
K ← transpose WordVec
V ← WordVec
Content ← Softmax((QK)/d)V
MFVec ← concat(Content, ContextVec)
end for
return MFVec

Table 2
Experimental datasets information

Datasets	Train	Test	Val	No. of classes
BBC News	1225	600	400	5
R8	5485	2189	7688	8

4.2. Experimental setting

The proposed method has a learning rate of 2e-5. Constrained by device resources, in order to train the model more efficiently, we chose a batch size of 32. In the BBC News dataset, around 80% of the documents have a length of around 500. Therefore, this paper will uniformly trim the documents in the BBC News dataset to a length of 500. Similarly, the document length for the R8 dataset is 65. When the loss of the validation set does not decrease for 10 consecutive batches, the training is stopped.

4.3. Baselines

4.3.1. Machine learning

This paper uses random forest, k-nearest neighbor (KNN), and support vector machine (SVM) in machine learning as comparison methods. They are commonly used text classification algorithms with excellent classification performance.

4.3.2. Deep learning

In this paper, CNN (Kim, 2014), BiGRU (Zhou & Bian, 2019), and FastText (Joulin et al., 2016) are selected as comparison methods. Both CNN and BiGRU are classic text classification algorithms, which are widely used in various text classification tasks until today.

4.3.3. Related methods

In order to highlight the superiority of ERNIE, BERT is chosen as the comparison method. Methods related to the classification of recent news topics were selected as comparison methods, namely RCNN (Wang, 2022), Doc2vec+CNN (Dogru et al., 2021), and Word2vec+CNN (Bei et al., 2023).

4.4. Results

The accuracy rate, recall rate, and F1 values of each method are shown in Table 3. The following is a calculation of the macro average on the test set.

Table 3
Experimental result (%)

Method	BBC News Accuracy	BBC News Recall	BBC News F1-value	R8 Accuracy	R8 Recall	R8 F1-value
RF	89.12	89.44	89.33	93.69	93.88	93.72
KNN	88.64	88.59	88.63	94.11	94.56	94.38
SVM	93.61	93.54	93.66	94.28	97.33	97.30
CNN	92.84	92.65	92.88	95.76	95.60	95.71
BiGRU	93.65	93.42	93.55	96.32	96.45	96.40
FastText	92.64	92.32	92.70	96.13	96.20	96.18
BERT	96.24	95.68	95.90	97.85	97.80	97.82
RCNN	94.54	95.26	94.88	96.82	96.77	96.80
Doc2vec +CNN	94.89	95.15	95.06	96.55	96.42	96.48
Word2vec +CNN	95.06	95.22	95.31	97.12	97.33	97.26
Proposed method	97.74	97.65	97.82	98.31	98.36	98.34

As shown in Table 3, the proposed method achieves the best news classification effect on BBC News, with the accuracy rate, recall rate, and F1 value reaching 97.74%, 97.65%, and 97.82%, respectively. The proposed method achieves the best news classification effect on R8, with the accuracy rate, recall rate, and F1 value reaching 98.31%, 98.36%, and 98.34%, respectively. Especially on BBC News, the accuracy of the proposed method is 3.2% higher than that of RCNN. Compared with Doc2vec+CNN, the accuracy is 2.85% higher. Compared with Word2vec+CNN, the accuracy is 2.68% higher. The above fully demonstrates the superiority of the proposed method. In addition, we also found that the classification performance of deep learning methods is significantly better than that of machine learning methods. This shows that the feature vector and feature extraction capabilities of deep learning methods are superior to machine learning methods. Next, we conducted ablation experiments, as shown in Table 4.

Table 4
Ablation result (%)

Method	BBC News Accuracy	BBC News Recall	BBC News F1-value	R8 Accuracy	R8 Recall	R8 F1-value
ERNIE	96.55	96.20	96.43	98.06	98.10	98.08
ERNIE + Word	97.12	97.32	97.22	98.16	98.11	98.13
ERNIE + Word + Att	97.55	97.43	97.38	98.21	98.25	98.23
ERNIE + Word + Att (Key sentence)	97.74	97.65	97.82	98.31	98.36	98.34

Compared with ERNIE and BERT, the accuracy is 0.31% and 0.21% higher, indicating that ERNIE has a higher semantic accuracy. Compared to ERNIE, ERNIE+Word has an accuracy rate that is 0.57% higher, which shows that combining word features helps enrich news content. Compared to ERNIE+Word, ERNIE+Word +Att has an accuracy rate that is 0.43% higher. This indicates that the attention mechanism has effectively extracted the important information from the news text. Compared with ERNIE+Word+Att, the attention mechanism based on key sentences has an accuracy rate that is 0.19% higher. The key sentence features enable the model to focus on the relatively important parts of the news text. The above fully demonstrates the effectiveness of the proposed method.

5. Conclusion

This paper introduces a novel approach to classifying news topics. Our method incorporates word features and key sentence features, enhancing the textual analysis of news articles. Word features complement the overall text analysis, while key sentence features direct the model's attention toward crucial segments of the news content. An attention mechanism facilitates knowledge fusion among these multiple features. The experimental results demonstrate the superiority of our proposed method over the comparison method when applied to the BBC News dataset. These findings unequivocally highlight the effectiveness of our approach, offering a fresh perspective on the task of news topic classification. The limitation of this paper lies in the lack of validation of the proposed method's performance on other text classification tasks, which will be the focus of our future work.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in https://doi.org/10.1145/1143844.1143892 and https://github.com/yao8839836/text_gcn.

References

Agarwal, J., Christa, S., Pai, A., Kumar, M. A., & Prasad, G. (2023). Machine learning application for news text classification. In *2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 463–466. https://doi.org/10.1109/Confluence56041.2023.10048856

Alammary, A. S. (2022). BERT models for Arabic text classification: A systematic review. *Applied Sciences*, *12*(11), 5720. https://doi.org/10.3390/app12115720

Bao, D., Qin, D., Hong, L., & Zhan, S. (2022). Multi-channel text classification model based on ERNIE. In *Proceedings of the 2022 11th International Conference on Computing and Pattern Recognition*, 321–327. https://doi.org/10.1145/3581807.3581853

Bei, X., Ping, L. S., & En, Z. M. (2023). Research on news text classification based on TextCNN. In *Proceedings of International Conference on Electronic Information Engineering and Data Processing, 12700*, 127002B. https://doi.org/10.1117/12.2682270

Dogru, H. B., Tilki, S., Jamil, A., & Hameed, A. A. (2021). Deep learning-based classification of news texts using Doc2Vec model. In *2021 1st International Conference on Artificial Intelligence and Data Analytics*, 91–96. https://doi.org/10.1109/CAIDA51941.2021.9425290

Fati, S. M., Muneer, A., Alwadain, A., & Balogun, A. O. (2023). Cyberbullying detection on Twitter using deep learning-based attention mechanisms and continuous bag of words feature extraction. *Mathematics*, *11*(16), 3567. https://doi.org/10.3390/math11163567

Greene, D., & Cunningham, P. (2006). Practical solutions to the problem of diagonal dominance in kernel document clustering. In *Proceedings of the 23rd International Conference on Machine Learning*, 377–384. https://doi.org/10.1145/1143844.1143892

Hossain, A., Konok, U. H., Islam, R., Ruhani, R. M. K., Musfikin, R., Uddin, M. M., . . . , & Tuhin, R. A. (2023). Utilizing GloVe embeddings for deep learning-based analysis of research paper abstracts. In *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications*, 1–6. https://doi.org/10.1109/HORA58378.2023.10156746

Jalal, N., Mehmood, A., Choi, G. S., & Ashraf, I. (2022). A novel improved random forest for text classification using feature ranking and optimal number of trees. *Journal of King Saud University-Computer and Information Sciences*, *34*(6), 2733–2742. https://doi.org/10.1016/j.jksuci.2022.03.012

Joulin, A., Grave, E., Bojanowski, P., Douze, M., Jégou, H., & Mikolov, T. (2016). FastText.zip: Compressing text classification models. *arXiv Preprint: 1612.03651.*

Kaur, G., & Bajaj, K. (2016). News classification and its techniques: A review. *IOSR Journal of Computer Engineering*, *18*(1), 22–26.

Kim, Y. (2014). Convolutional neural networks for sentence classification. *arXiv Preprint:1408.5882*.

Lai, S., Xu, L., Liu, K., & Zhao, J. (2015). Recurrent convolutional neural networks for text classification. *Proceedings of the AAAI Conference on Artificial Intelligence, 29*(1). https://doi.org/10.1609/aaai.v29i1.9513

Li, J., Zhang, D., & Wulamu, A. (2021). Chinese text classification based on ERNIE-RNN. In *2021 2nd International Conference on Electronics, Communications and Information Technology*, 368–372. https://doi.org/10.1109/CECIT53797.2021.00072

Lin, D., Wang, H., Liu, M., & Li, P. (2021). News text classification based on bidirectional encoder representation from transformers. In *2021 International Conference on Artificial Intelligence, Big Data and Algorithms*, 137–140. https://doi.org/10.1109/CAIBDA53561.2021.00036

Liu, B., Guan, W., Yang, C., & Fang, Z. (2023). Effective method for making Chinese word vector dynamic. *Journal of Intelligent & Fuzzy Systems*, *45*(1), 941–952. https://doi.org/10.3233/JIFS-224052

Liu, Y., Li, P., & Hu, X. (2022). Combining context-relevant features with multi-stage attention network for short text classification. *Computer Speech & Language*, *71*, 101268. https://doi.org/10.1016/j.csl.2021.101268

Mallik, A., & Kumar, S. (2024). Word2Vec and LSTM based deep learning technique for context-free fake news detection. *Multimedia Tools and Applications, 83*, 919–940. https://doi.org/10.1007/s11042-023-15364-3

Saravanan, T., Jhaideep, T., & Bindu, N. H. (2022). Detecting depression using Hybrid models created using Google's BERT and Facebook's Fast Text Algorithms. In *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering*, 415–421. https://doi.org/10.1109/ICACITE53722.2022.9823581

Soni, S., Chouhan, S. S., & Rathore, S. S. (2023). TextConvoNet: A convolutional neural network based architecture for text classification. *Applied Intelligence*, *53*(11), 14249–14268. https://doi.org/10.1007/s10489-022-04221-9

Tezgider, M., Yildiz, B., & Aydin, G. (2022). Text classification using improved bidirectional transformer. *Concurrency and Computation: Practice and Experience*, *34*(9), e6486. https://doi.org/10.1002/cpe.6486.

Wang, C., & Zhang, F. (2022). The performance of improved XLNet on text classification. In *Proceedings of the Third International Conference on Artificial Intelligence and Electromechanical Automation, 12329*, 1232900. https://doi.org/10.1117/12.2646785

Wang, X. (2022). Research on news text classification based on RCNN. *Frontiers in Computing and Intelligent Systems*, *2*(2), 58–62. https://doi.org/10.54097/fcis.v2i2.4086

Wang, X., Zhao, H., & Chen, H. (2023). Improved Skip-gram based on graph structure information. *Sensors*, *23*(14), 6527. https://doi.org/10.3390/s23146527

Zhang, H., Zeng, C., Pan, L., Hao, R., Wen, C., & He, P. (2022a). News topic text classification method based on BERT and feature projection network. *Journal of Computer Applications*, *42*(4), 1116–1124. https://doi.org/10.11772/j.issn.1001-9081.2021071257

Zhang, L., Wu, Y., Chu, Q., Li, P., Wang, G., Zhang, W., . . . , & Li, Y. (2022b). SA-Model: Multi-feature fusion poetic sentiment analysis based on a hybrid word vector model. In *2022 5th International Conference on Pattern Recognition and Artificial Intelligence*, 984–988. https://doi.org/10.1109/PRAI55851.2022.9904158

Zhang, Z., Han, X., Liu, Z., Jiang, X., Sun, M., & Liu, Q. (2019). ERNIE: Enhanced language representation with informative entities. *arXiv Preprint: 1905.07129*.

Zhou, H. (2022). Research of text classification based on TF-IDF and CNN-LSTM. *Journal of Physics: Conference Series, 2171*(1), 012021. https://doi.org/10.1088/1742-6596/2171/1/012021

Zhou, L., & Bian, X. (2019). Improved text sentiment classification method based on BiGRU-Attention. *Journal of Physics: Conference Series, 1345*(3), 032097. http://doi.org/10.1088/1742-6596/1345/3/032097

How to Cite: Chen, W., Liu, B., & Guan, W. (2024). ERNIE and Multi-Feature Fusion for News Topic Classification. *Artificial Intelligence and Applications*, *2*(2), 135–140. https://doi.org/10.47852/bonviewAIA32021743