Journal of Electronics and Information Science (2024)
Clausius Scientific Press, Canada
DOI: 10.23977/jeis.2024.090320
ISSN 2371-9524 Vol. 9 Num. 3

Research on the application of computer-aided deep learning model in natural language processing

Xiaokai Jiang, Xuewen Ding*, Chunyu Liu, Xinyi Li, Yuan Zhang, Shaosai Wang

School of Electronic Engineering, Tianjin University of Technology and Education, Tianjin, China

Corresponding author: dingxw@tute.edu.cn

Keywords: Natural language processing; computer-aided; deep learning

Abstract: With the advent of the big data era and significant advancements in computing power, deep learning (DL) technology has achieved breakthrough progress in natural language processing (NLP), particularly in language understanding and generation. This study focuses on the application of neural networks and their variants, such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short Term Memory Network (LSTM), and Gated Recurrent Unit (GRU), in NLP tasks, as well as how the Transformer model leverages self-attention mechanisms to enable parallel processing and effectively capture long-distance dependencies. Additionally, the research explores the importance of pre-training and self-supervised learning in enhancing model generalization and reducing overfitting. In terms of specific applications, this paper provides a detailed analysis of DL models in text classification, sentiment analysis, machine translation, dialogue systems, and question-answering systems. It demonstrates how these models significantly improve the efficiency and effectiveness of NLP tasks through automatic learning of complex features, strong generalization capabilities, and end-to-end learning processes.

1. Introduction

In the research field of natural language processing (NLP), which is full of challenges and opportunities, the language communication barrier between human and machine is gradually being broken by technological innovation. With the advent of the era of big data and the leap of computing power, deep learning (DL) technology has brought unprecedented development opportunities for NLP with its powerful ability of automatic feature extraction and end-to-end learning mode [1-2]. From the intimate service of intelligent customer service to the accurate communication of machine translation, from the delicate insight of emotional analysis to the smooth interaction of dialogue system, DL model is understanding and generating human language in an unprecedented way, greatly expanding the application boundary of NLP.

With the rise of Transformer architecture and the wide application of pre-training language models such as BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer), the processing efficiency and effect of NLP tasks have been significantly improved, which not only promoted the research process in academia, but also brought many innovative applications to industry [3-4]. However, although DL has made remarkable achievements in NLP, there are still many unsolved mysteries behind it, such as the explanatory power of the model, the ability to understand complex language structures and the effectiveness of cross-language transfer learning. This study deeply discusses the latest application progress of computer-aided DL models in NLP, aiming at revealing how these models can realize the intelligence of language understanding and generation by learning the potential laws and patterns in large-scale text data.

2. Core technology of computer-aided DL model in NLP

2.1. Neural networks and variants

In the field of NLP, DL models achieve efficient processing and understanding of language by simulating the working mode of human brain neural networks. Among them, models such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short Term Memory Networks (LSTM), and Gated Recurrent Units (GRU) have played important roles in NLP tasks due to their unique network structures and processing mechanisms [5-7].

CNN has achieved great success in the field of image processing at first (see Figure 1). It automatically extracts features from input data through convolution layer and pooling layer, which are very important for distinguishing the shapes, textures and colors of different objects. CNN is also of great significance in NLP. By transforming the text data into a word vector matrix, CNN can capture the local features of different N-grams, which can effectively extract the key features of the text while retaining the word order information. In the task of text classification, CNN can identify important semantic and structural features in sentences, which is helpful to classify texts quickly and accurately. In addition, in the aspect of emotional analysis, CNN can identify the emotional color implied in the sentence, which has a good effect on understanding the emotional tendency of the text.

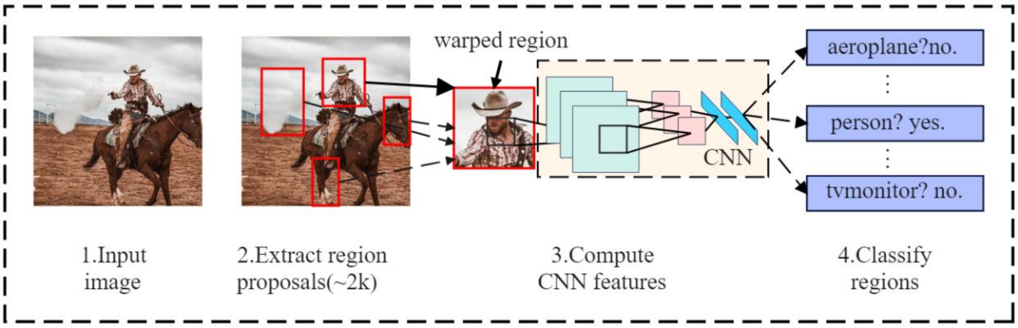

Figure 1: A network for computing CNN features

RNN is a special neural network structure, and its network structure is shown in Figure 2, which has the ability to process sequence data. Different from the traditional neural network, RNN has a circular structure, and can share information with the same parameters at each time step, thus capturing the context of the input sequence. This feature makes RNN perform well in NLP tasks, especially when dealing with time-dependent tasks, such as language modeling, machine translation and named entity recognition. RNN provides strong support for these tasks by learning context to predict the next word or character.

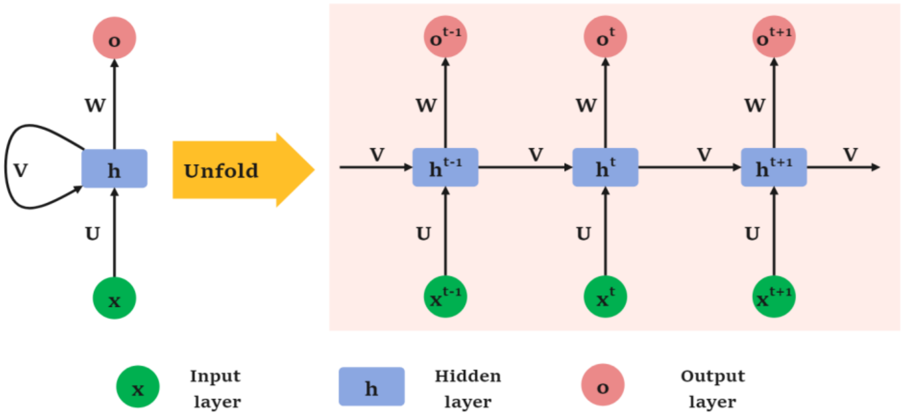

Figure 2: RNN network structure

LSTM is a variant of RNN, which aims to solve the problems of gradient disappearance and gradient explosion that are easy to occur when standard RNN processes long sequence data. By introducing gating mechanism (including forgetting gate, input gate and output gate) and memory unit, LSTM can selectively retain or forget the information in the sequence, thus effectively learning long-term dependence. The LSTM unit structure is shown in Figure 3. In NLP, LSTM is widely used in language modeling, machine translation, sentiment analysis and other fields, which significantly improves the performance and stability of the model.

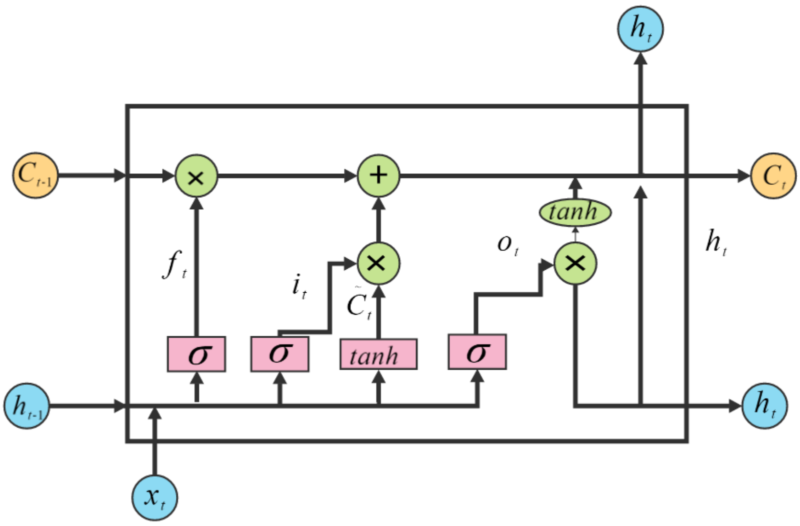

Figure 3: LSTM cell structure

GRU is another common RNN variant (see Figure 4), similar to LSTM, but with fewer parameters and simpler structure. GRU controls the flow of information by resetting the gate and updating the gate, and realizes a similar memory and forgetting mechanism. Because GRU is more efficient, GRU may perform better than LSTM in some NLP tasks. Especially in the scenes that need quick response or limited resources, GRU has become a more attractive choice.

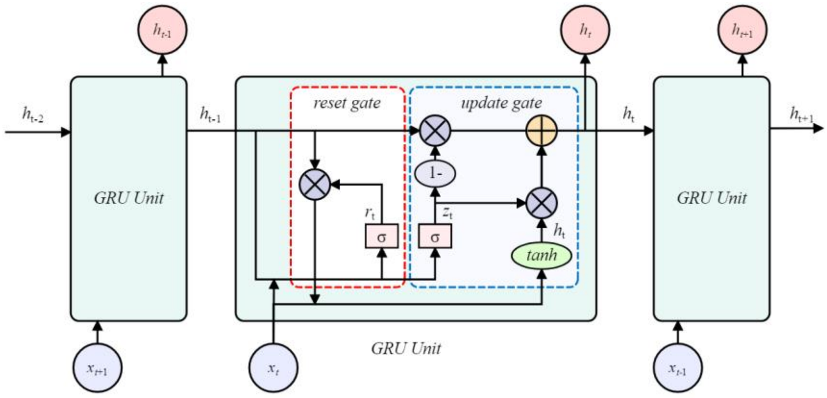

Figure 4: Schematic diagram of GRU module

2.2. Transformer model

Transformer model is a DL model, which was proposed by Vaswani et al. in 2017 and is mainly used for NLP tasks. It abandons the traditional RNN and CNN structure and is completely based on the self-attention mechanism [8]. This design enables Transformer to process serial data in parallel, which greatly improves the calculation efficiency. The Transformer model structure is shown in Figure 5.

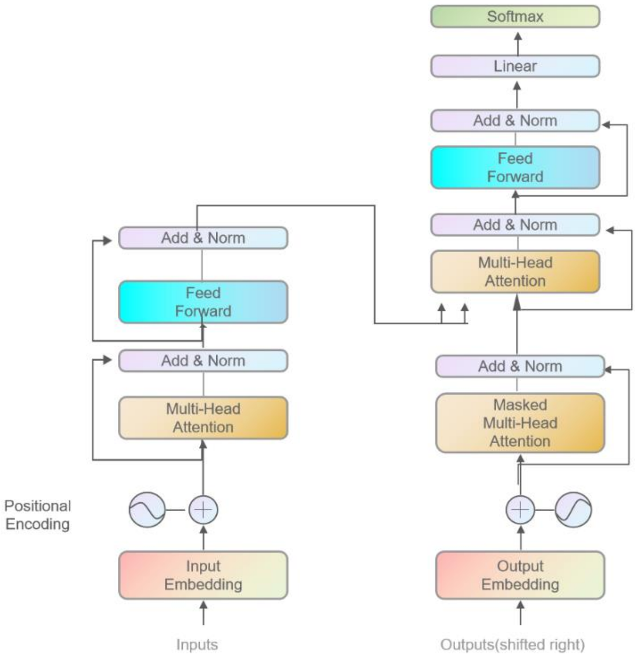

Figure 5: Transformer model structure

Self-attention mechanism is the core of Transformer, which allows the model to consider the information of the whole sequence when dealing with each element of the sequence [9]. Specifically, the self-attention mechanism obtains the representation of each position by calculating the attention weights between the elements in each position in the sequence and all other position elements, and then weighting and summing the elements according to these weights. Transformer model has obvious advantages, including its parallel processing ability supported by self-attention mechanism, which allows the model to process all elements in the sequence at the same time, which RNN cannot do; It can effectively capture long-distance dependencies and is very important for parsing complex language structures. In addition, by increasing the number of encoder and decoder layers, the model shows a high degree of flexibility and expansibility, thus enhancing the expressive ability.

Representative language models based on Transformer include: BERT, which was put forward by Google in 2018, pre-trained the deep bidirectional model by jointly using context information, and performed well on multiple NLP tasks; GPT, introduced by OpenAI in the same year, uses unsupervised learning autoregressive method to generate text, which is outstanding in the generation task; T5 (Text-to-Text Transfer Transformer), launched by Google in 2019, transforms all NLP tasks into a unified text-to-text format, covering translation, summarization and classification.

2.3. Pre-training and self-supervised learning

Pre-training technology is an important factor for the success of DL model, especially in NLP field. By pre-training on large-scale text data sets, the model can learn the general features, patterns and grammatical rules of the language, laying the foundation for fine-tuning for specific tasks, enhancing the generalization ability of the model and reducing over-fitting. As a form of pre-training, self-supervised learning allows the model to learn automatically from unlabeled data by designing prediction tasks (such as masking language model tasks), which reduces the dependence on a large number of manually labeled data, helps to capture the internal structure and complex relationship of data, and is especially suitable for dealing with large-scale unlabeled text data such as the Internet.

The combination of pre-training and self-supervised learning creates an efficient learning paradigm for NLP models. By performing self-supervised tasks such as Masked Language Modeling (MLM) and Next Sentence Prediction (NSP), models can learn deep structural and semantic features of language from large-scale unlabeled text during the pre-training phase [10-11]. This learning approach enables pre-trained models like BERT and GPT to acquire robust context-sensitive representations, providing a solid foundation for addressing various downstream tasks.

3. The concrete application of DL model in NLP

3.1. Text classification and sentiment analysis

Text classification is the basic task of NLP, which involves assigning text to predefined categories. DL models, such as FastText, TextCNN, TextRNN and BERT, have shown remarkable advantages in text classification [12-13]. FastText is suitable for the rapid training of large-scale data sets, TextCNN is good at short and medium text classification, TextRNN can capture long-distance dependence, and BERT performs well in various tasks through fine-tuning. The advantages of these models are automatic learning of complex features, strong generalization ability and end-to-end learning process, and direct output of classification results from the original text.

Sentiment analysis is a key task in NLP, which aims to identify and classify the emotional tendency of texts. DL models, such as LSTM, GRU, attention mechanism and BERT, have played an important role in this field. LSTM and GRU can effectively capture emotional changes, and attention mechanism can improve the recognition accuracy of key information, while BERT can capture deep semantics through pre-training and fine-tuning, which is suitable for fine-grained analysis. The advantages of these models are that they can understand the context, support multi-task learning and adapt to specific tasks quickly through transfer learning.

3.2. Machine translation

The DL model adopts an end-to-end translation method in machine translation, which directly translates from the source language to the target language, avoiding the intermediate representation or phased processing in the traditional methods, and capturing the complex mapping between the source language and the target language through deep neural networks. The end-to-end model often uses a sequence-to-sequence (Seq2Seq) framework, which consists of an encoder and a decoder. The former converts an input sequence into a context vector, and the latter generates an output sequence based on this vector [14]. Attention mechanism, as a key component of DL model, enables the model to focus on the most relevant part of the input sequence when processing the sequence, and enhances the accuracy of the alignment between the source language and the target language. For example, Transformer model uses self-attention mechanism to achieve efficient parallel processing and global dependency capture.

In recent years, large-scale pre-training language models, such as BERT, GPT and its variant XLM, have learned rich language expressions through unsupervised pre-training on large corpora, and demonstrated excellent performance through fine-tuning on specific tasks. By optimizing BERT and adopting Byte-Pair Encoding (BPE) technology, XLM has increased vocabulary sharing among different languages and further improved the translation effect. DL model, especially the model based on Transformer, has become the standard in the field of machine translation, which significantly improves the quality and practicability of translation. For example, Google's neural machine translation system proves the effectiveness and superiority of DL method with its high-quality translation results, especially in the processing of long sentences and complex sentence patterns.

3.3. Dialogue system and question answering system

The dialogue system uses DL model to realize natural and smooth text or voice communication, and its core technologies include Seq2Seq model, attention mechanism, reinforcement learning (RL) and multi-round dialogue management [15]. Seq2Seq model processes the dialogue history through the encoder-decoder architecture to generate coherent related replies; Attention mechanism enables the model to focus on different parts of input and improve the accuracy of understanding and response; RL optimizes the recovery strategy of the model through the reward and punishment mechanism, making it closer to human conversation habits; Multi-round dialogue management ensures the coherence and naturalness of dialogue by tracking the dialogue state and predicting the user's intention, and the application of structures such as LSTM and RNN further enhances the effect of multi-round dialogue.

Question answering system is a computer program that can directly provide accurate answers rather than simple web links, and DL has played a key role in its development [16]. Through semantic matching, context understanding and answer generation, DL models, such as CNN, RNN, BERT and graph neural network, can not only capture the semantic relationship between questions and answers to achieve accurate matching, but also analyze the context information of complex questions and even directly generate answers that conform to grammatical rules, thus greatly improving the accuracy and efficiency of the question answering system.

DL model is widely used in intelligent customer service, voice assistant and automatic question answering system. Intelligent customer service uses DL to understand user consultation and provide 24-hour service to improve customer satisfaction and reduce operating costs; Voice assistants such as Siri and Alexa realize speech recognition, natural language understanding and dialogue management through DL, and execute user instructions; The automatic question answering system provides professional knowledge consultation in financial, medical and educational industries, which can accurately understand professional questions and give reliable answers, such as assisting doctors to quickly find medical records, diagnose diseases or recommend treatment plans.

4. Conclusion

Computer-aided DL model has shown remarkable application potential and practical results in NLP field. By simulating the working mode of human brain neural network, these models can process and understand languages efficiently, especially in tasks such as text classification, sentiment analysis, machine translation and dialogue system. In particular, the Transformer model and its derived pre-training language models such as BERT and GPT, through the self-attention mechanism and pre-training of large-scale text data, not only improve the computational efficiency, but also enhance the model's ability to capture complex language structures and long-distance dependencies. In addition, combined with self-supervised learning technology, these models can effectively learn on unlabeled large-scale text data, further improving their generalization ability and reducing the risk of over-fitting. Therefore, the computer-aided DL model provides a powerful tool for understanding and generating human language, which greatly broadens the application prospect of NLP.

References

[1] Dhopavkar, G., Welekar, R. R., & Ingole, Piyush K.Vaidya, ChanduWankhade, Shalini VaibhavVasgi, Bharati P. (2023). Optimizing resource allocation in cloud for large-scale deep learning models in natural language processing. journal of electrical systems, 19(3), 62-77.

[2] Shahin, N., & Ismail, L. (2024). From rule-based models to deep learning transformers architectures for natural language processing and sign language translation systems: survey, taxonomy and performance evaluation. Artificial Intelligence Review, 57(10), 1-51.

[3] Al-Makhadmeh, Z., & Tolba, A. (2020). Automatic hate speech detection using killer natural language processing optimizing ensemble deep learning approach. Computing, 102(2), 501-522.

[4] Guo, J. (2022). Deep learning approach to text analysis for human emotion detection from big data. Journal of Intelligent Systems, 31(1), 113-126.

[5] Boulieris, P., Pavlopoulos, J., Xenos, A., & Vassalos, V. (2024). Fraud detection with natural language processing. Machine Learning, 113(8), 5087-5108.

[6] Zhang, L., Yang, X., Li, S., Liao, T., & Pan, G. (2022). Answering medical questions in chinese using automatically mined knowledge and deep neural networks: an end-to-end solution. BMC bioinformatics, 23(1), 136.

[7] Sawant, A. A., & Devanbu, P. (2021). Naturally!: how breakthroughs in natural language processing can dramatically help developers. IEEE Software, 38(5), 118-123.

[8] Yang, D., Zhu, D., & Wan, G. F. (2024). Semantic similarity caculating based on bert. Journal of Electrical Systems, 20(2), 73-79.

[9] Liu, H., Jun, G., & Zheng, Y. (2021). Chinese named entity recognition model based on bert. MATEC Web of Conferences, 336(3), 06021.

[10] Alyafeai, Z., Al-Shaibani, M. S., Ghaleb, M., & Ahmad, I. (2023). Evaluating various tokenizers for arabic text classification. Neural Processing Letters, 55(3), 2911-2933.

[11] Shao, Q., Zhang, W., & Wang, S. (2023). End-to-end aspect category sentiment analysis based on type graph convolutional networks. high technology letters, 29(3), 325-334.

[12] Wang, W., Lee, C. M., & Liu, T. P. W. (2023). An empirical study of cyclical learning rate on neural machine translation. Natural language engineering, 29(2), 316-336.

[13] Zou, A., Wu, X., Li, X., Zhang, T., Cui, F., & Xu, J. (2024). Curriculum pre-training for stylized neural machine translation. Applied Intelligence, 54(17-18), 7958-7968.

[14] Maruf, S., Saleh, F., & Haffari, G. (2021). A survey on document-level neural machine translation: methods and evaluation. ACM Computing Surveys, 54(2), 1-36.

[15] Yang, M., Liu, S., Chen, K., Zhang, H., Zhao, E., & Zhao, T. (2020). A hierarchical clustering approach to fuzzy semantic representation of rare words in neural machine translation. IEEE Transactions on Fuzzy Systems, 28(5), 992-1002.

[16] Wang, J., & Dong, Y. (2020). Measurement of text similarity: a survey. Information (Switzerland), 11(9), 421.