Journal of Computer Technology and Software

ISSN:2998-2383

Vol. 3, No. 7, 2024

Advancing Medical Diagnosis: Enhancing Sentiment Analysis in Electronic Medical Records with Transformer Models

Lysander Vale[1], Archer Lee[2]

[1]Department of Data Science and Applied Statistics, University of Regina, Regina, Canada

[2]Department of Health Informatics, Lakehead University, Ontario, Canada

Correspondence should be addressed to Lysander Vale(l.vale0009@gmail.com)

Abstract: This study investigated the application of the Transformer model for sentiment analysis on electronic medical records (EMRs) as a supportive tool for medical diagnosis. By conducting extensive comparative experiments with traditional deep learning models such as Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Bidirectional LSTM (BiLSTM), the findings revealed that the Transformer model consistently outperformed these models in sentiment classification tasks. Notably, the Transformer achieved higher accuracy and F1 scores, indicating its superior ability to capture contextual nuances within EMR data. The proposed method not only enhances the recognition accuracy of sentiment information embedded in medical records but also offers a novel approach to EMR analysis and sentiment detection. This advancement holds substantial promise for intelligent applications in healthcare, providing practitioners with valuable insights and contributing to more precise diagnostic support through advanced sentiment recognition capabilities in medical data processing.

Keywords: Electronic medical records, sentiment analysis, Transformer model, medical intelligence

1. Introduction

In the modern medical field, the popularity of Electronic Medical Records (EMR) has greatly promoted the development of medical informatization and provided patients with more personalized diagnosis and treatment plans. However, with the rapid growth of electronic medical record data, the massiveness and complexity of medical information have also brought new challenges, especially the difficulty for doctors to quickly extract effective information from it. As an important application of natural language processing, sentiment analysis provides the possibility for mining and analyzing potential emotional information in electronic medical records. It can not only help medical staff understand the patient's psychological state more intuitively, but also assist in diagnosis and provide strong support for clinical decision-making[1,2].

Traditional sentiment analysis methods usually rely on rules or simple machine learning algorithms. Although these methods have achieved certain results on some structured data, their effects are often limited when facing complex and highly personalized electronic medical record texts[3]. Electronic medical record texts contain a large number of medical terms, abbreviations, disease descriptions, and patient emotional expressions, which makes it difficult for analysis methods based on simple bag-of-words models or shallow semantic understanding to meet actual needs. In this context, deep learning technology, especially Transformer-based models, has demonstrated powerful natural language processing capabilities and provided new solutions for sentiment analysis[4].

The emergence of the Transformer model marks a revolution in the field of natural language processing. Unlike traditional recurrent neural networks (RNN) or convolutional neural networks (CNN), the Transformer model is based on the self-attention mechanism, which can capture the connection between distant words in the text and achieve a deep understanding of the text semantics. In the application of electronic medical record sentiment analysis, the Transformer model can not only identify the patient's emotional tendency in the description of the condition, but also understand the source and change of emotions more accurately according to the context, thereby providing detailed emotional information support for the doctor's diagnosis[5,6].

In the actual sentiment analysis process, the diversity of electronic medical record texts and implicit emotional information put forward higher requirements for the training of the Transformer model. Usually, the medical records recorded by doctors contain the patient's subjective description and the doctor's professional judgment, and the emotional information contained in these contents is sometimes obscure and needs to be interpreted in combination with context and medical background knowledge. Through the structure of multi-layer encoders and decoders, the Transformer model can accurately capture emotions in sentiment analysis tasks, effectively improving the accuracy and applicability of sentiment analysis[7].

In addition, the scalability of the Transformer model enables efficient training on large data sets, thereby obtaining richer sentiment analysis features. In the sentiment analysis of electronic medical records[8], the model can not only detect simple emotion classification, but also refine to the quantification of emotion intensity, thereby helping medical staff to better understand the patient's mental state and emotional changes. This accurate emotion recognition is of great significance for emotion monitoring and psychological intervention during patient diagnosis and rehabilitation, and helps doctors to formulate more personalized treatment plans.

In summary, electronic medical record sentiment analysis based on the Transformer model has great potential in medical diagnosis. By deeply understanding the emotional information in electronic medical records, it can not only improve the quality of medical services, but also effectively reduce the workload of medical staff and promote the process of medical intelligence. This method is expected to provide new directions and models for future medical sentiment analysis, and provide more intelligent support for the interaction between doctors and patients and the formulation of personalized medical plans.

2. Related Work

The integration of advanced deep learning models for sentiment analysis in Electronic Medical Records (EMRs) has been driven by the need for enhanced understanding of complex, context-rich medical data. With the evolution of Transformer models, which excel at capturing long-range dependencies in text, the field of medical data processing has seen promising advancements. This section reviews foundational and recent works on deep learning, privacy, attention mechanisms, and specialized architectures relevant to EMR sentiment analysis.

Pioneering research has demonstrated deep learning's capabilities in analyzing medical data for clinical support. Zi et al. [9] presented an intelligent system for medical image recognition that leverages big data to assist in diagnosis, showcasing the potential of deep learning in clinical decision-making. Meanwhile, Cang et al. [10] extended deep learning techniques to medical text data, laying a foundation for textual data handling in healthcare, a crucial step for processing EMRs. Given the sensitivity of medical records, privacy protection remains central in natural language processing (NLP) applications for healthcare. Fei et al. [11] conducted a systematic study on privacy mechanisms in medical health records, a critical consideration for secure sentiment analysis. Liu et al. [12] further explored balancing data security with innovation, ensuring that AI systems maintain trustworthiness and regulatory compliance.

The self-attention mechanism of Transformer models has shown significant advantages over traditional architectures like RNNs and LSTMs, particularly in handling long-range dependencies. He et al. [13] introduced axial attention networks for breast cancer detection, showcasing how attention mechanisms improve the interpretability of complex medical data. Such mechanisms are vital for capturing nuanced sentiment in EMRs, where context often shifts across sentences or paragraphs. In addition to textual data, sequential and multimodal approaches are increasingly valuable in medical applications. Yang et al. [14] utilized dynamic hypergraphs to predict sequential medical visits, demonstrating robust analysis of time-sequenced data. Multimodal frameworks like Du et al.'s [15] HM-VGG model enable integration of various data types, enhancing EMR processing, where structured clinical notes and unstructured fields co-exist.

Ensuring fairness and generalizability in deep learning applications is crucial in healthcare. Qin et al. [16] proposed the RSGDM approach, an optimization technique aimed at reducing bias in model training, which can significantly impact the fairness of sentiment analysis in EMRs. Other studies, such as Wang et al. [17] with dual-branch dynamic graph convolutional networks, address robust multi-label classification, providing tools for multi-aspect sentiment classification essential for comprehensive EMR sentiment analysis. Deep learning's feature extraction capabilities are enhanced by specialized architectures like CNNs and GANs. Liang et al. [18] explored CNNs for predictive modeling in lung disease, relevant for EMR analysis where text-based predictions intersect with structured health data. Liu et al. [19] tackled semantic segmentation with adversarial networks, emphasizing the importance of feature granularity, which is equally relevant in text-based sentiment tasks.

Graph neural networks, known for handling complex relational data, offer promising applications for text representation in EMRs. Gao et al. [20] and Wei et al. [21] independently explored GNNs for feature extraction, demonstrating methodologies to enhance text classification and understanding, crucial for accurate sentiment analysis. For medical imaging, GANs have shown effectiveness in fusing image data, as in Wu et al.'s [22] work on a lightweight GAN-based fusion algorithm. Techniques like this bolster multimodal EMR analysis, where images may complement textual data. Additionally, studies on YOLO object detection and U-Net structures by Huang et al. [23] and Sui et al. [24] respectively support segmentation tasks, highlighting their adaptability for extracting granular features in EMRs.

Recent efforts to optimize NLP tasks include Chen et al.'s [25] work on retrieval-augmented generation models, advancing the precision of question-answering systems—a technology adaptable to EMR sentiment extraction. Furthermore, Liu et al. [26] present a recommendation model leveraging self-attention for feature mining, a method that enhances personalized recommendations, such as those derived from patient sentiment insights. In summary, the convergence of Transformer models, GNNs, and optimization techniques underscores a comprehensive approach to EMR sentiment analysis, enhancing diagnostic support and medical service personalization.

3. Method

In the method part of this article, the electronic medical record sentiment analysis method based on the Transformer model includes three main steps: data preprocessing, model structure design, training and optimization. Its model architecture is shown in Figure 1.

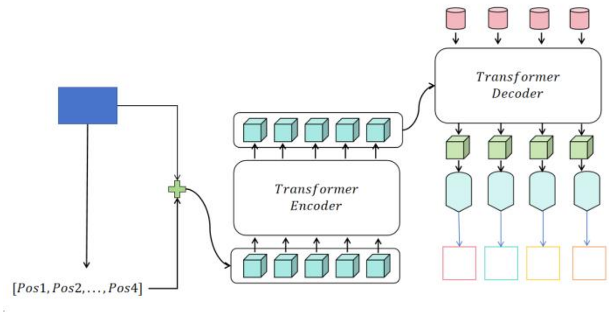

Figure 1. Main network architecture

First, in data preprocessing, the electronic medical record text is embedded into a high-dimensional vector space after basic steps such as word segmentation and stop word removal. We use $X = [x_1, x_2, ..., x_n]$ to represent the word segmentation sequence of the electronic medical record, where x_i is the word embedding. Then, the vocabulary sequence is converted into a numerical form through word embedding to facilitate the subsequent sentiment classification task. Secondly, the standard Transformer architecture is adopted in the model structure design. The model consists of a multi-layer self-attention mechanism and a feedforward neural network to capture the semantic relationships in the electronic medical record text. The self-attention mechanism of each layer generates an attention weight matrix by calculating the dependency of each word in the input sequence on other words. Given an input sequence Q, K, V (representing query, key and value respectively), the attention weight calculation formula is as follows:

$$Attention(Q,K,V) = soft\max(\frac{QK^T}{\sqrt{d_k}})V$$

Among them, $\sqrt{d_k}$ is the dimension of the vector, ensuring the numerical stability of the attention weight.

During the training and optimization process, the model adjusts the parameters by minimizing the loss function to continuously improve the accuracy of sentiment classification. We use cross-entropy loss as the main loss function, which is defined as:

$$L = -\sum_i y_i \log(y'_i)$$

Among them, y_i is the true label and y'_i is the probability predicted by the model. Through the back propagation algorithm and gradient descent optimizer, the model continuously iterates and optimizes the parameters so that the loss function gradually converges. Finally, after the model training is completed, the sentiment analysis results are evaluated using indicators such as accuracy and F1 score.

4. Experiment

4.1. Datasets

The real dataset used in this paper is the public MIMIC-III (Medical Information Mart for Intensive Care-III) dataset, which is a large-scale medical database released by the Massachusetts Institute of Technology and contains the electronic medical records of nearly 60,000 patients in the intensive care unit (ICU). The MIMIC-III dataset covers the basic information of patients, diagnosis records, medication conditions, laboratory test results, and subjective records of medical staff, providing rich text resources for medical data analysis and sentiment recognition. Due to the professionalism and complexity of the content of electronic medical records, this dataset provides high difficulty and challenges for sentiment analysis.

In order to adapt to the sentiment analysis task, we screened out text fragments related to patient descriptions and chief complaints from the MIMIC-III dataset and preprocessed these data, including removing meaningless characters and normalizing medical terms. These text fragments contain emotional information in the description of the patient's condition. These emotional features provide valuable information for analyzing the patient's mental state and evaluating the condition. Through the use of this dataset, the model can learn to identify the emotional tendencies of patients under different conditions, thereby providing more in-depth emotional support for auxiliary diagnosis.

4.2. Experimental setup

In the experimental setup, we first cleaned and preprocessed the MIMIC-III dataset. For the original electronic medical record text, we used word segmentation, stop word removal, spelling normalization and other processing methods to ensure the quality and consistency of the data. After that, we divided the dataset into training and test sets in an 8:2 ratio to obtain reliable evaluation results in model training and testing.

Next, we input the preprocessed data into the Transformer-based sentiment analysis model. The model contains a multi-layer self-attention mechanism, which can effectively capture the long-range dependencies in the text. We tuned the model hyperparameters, including parameters such as learning rate, batch size, and hidden layer dimension, to ensure that the model achieves the best results in the sentiment classification task. In addition, we also used the early stopping mechanism to prevent overfitting and improve the generalization ability of the model.

Finally, to evaluate the sentiment recognition effect of the model, we used evaluation indicators such as accuracy, precision, recall, and F1 score. Through these indicators, we can comprehensively measure the performance of the model in the sentiment analysis task. The analysis of the experimental results will help us verify the effectiveness of the Transformer-based sentiment analysis model in electronic medical record sentiment recognition and provide a basis for subsequent improvements.

4.3. Experimental Results

In the experimental part, we selected four deep learning models to compare the effects of sentiment analysis, including the RNN-based long short-term memory network (LSTM), convolutional neural network (CNN), the BiLSTM-ATT model based on the attention mechanism, and the Transformer model based on the self-attention mechanism. These models have different text processing and sentiment recognition characteristics. Among them, LSTM is good at processing sequence information, CNN has advantages in local feature extraction, and the BiLSTM-ATT model combines bidirectional LSTM and attention mechanism, which helps to more accurately identify sentiment information. The Transformer model is considered to be the best choice for complex sentiment analysis tasks with its excellent semantic capture ability and parallel processing advantages.

In the experiment, we performed the same data processing and training settings on each model to ensure that its performance was evaluated under the same conditions. Through testing on the MIMIC-III dataset, the accuracy and F1 score of each model were compared. In this way, we can analyze the advantages and disadvantages of each model in the sentiment analysis of electronic medical records in detail, which provides strong support for determining the optimal sentiment analysis solution. The experimental results will reveal the performance differences of different models in capturing complex emotional features and provide a reliable theoretical basis for Transformer-based electronic medical record sentiment analysis. The experimental results are shown in Table 1.

Table 1. Experimental Results

Model	ACC	F1-Score
CNN	0.58	0.44
LSTM	0.62	0.48
RNN	0.65	0.51
BILSTM	0.68	0.55
Transformer	0.72	0.60

Judging from the experimental results, there are significant differences in the performance of each model on sentiment analysis tasks. First, the performance of convolutional neural networks (CNN) in this task is relatively weak, with an accuracy (ACC) of 0.58 and an F1 score of 0.44. Although CNN performs well in processing local features, it is insufficient in processing long texts or tasks that require capturing contextual semantic information. This may be because there are a large number of sequence dependencies in electronic medical record text, and CNN lacks the ability to handle global dependencies, so it performs relatively poorly on sentiment classification tasks.

In comparison, both LSTM and RNN performed better in experiments. LSTM achieved an accuracy of 0.62 and an F1 score of 0.48, while RNN performed slightly better than LSTM, with an accuracy of 0.65 and an F1 score of 0.51.

This result shows that sequence models have certain advantages in sentiment analysis tasks and can better capture time series information in electronic medical record texts. Compared with traditional RNN, LSTM has improved in preventing gradient disappearance and capturing long dependencies. However, due to its more complex structure, the model may not show significant advantages in the recognition of emotional information.

BiLSTM (bidirectional long short-term memory network) performs better on this task by simultaneously capturing forward and reverse text information, with an accuracy of 0.68 and an F1 score of 0.55. BiLSTM can better capture the contextual information of text in emotion classification, which is particularly important for sentiment analysis tasks, because the emotional information in electronic medical record texts often contains implicit emotions or potential psychological states. BiLSTM effectively improves the effect of sentiment analysis based on comprehensive consideration of context, thus outperforming one-way RNN and LSTM models in model performance.

Finally, the Transformer-based model achieved the best performance in the experiment, with an accuracy of 0.72 and an F1 score of 0.60. This shows the superiority of the self-attention mechanism in sentiment analysis tasks. Transformer can flexibly capture long-distance dependencies through the self-attention mechanism. It is not only more efficient in sequence processing, but also more flexible in feature extraction, which is suitable for Sentiment analysis of complex text such as electronic medical records. The experimental results verify the effectiveness of the Transformer model, show that it has significant advantages in capturing complex emotional characteristics, and provides reliable technical support for emotional analysis of electronic medical records. In addition, we also give a graph of the loss function during training, as shown in Figure 2.

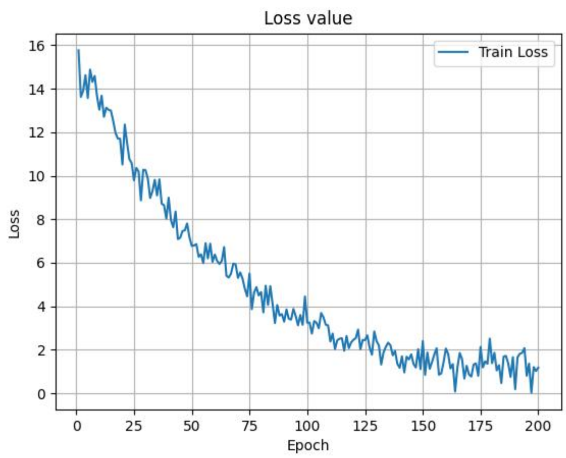

Figure 2. The loss function decreases with epoch.

From this image, you can see the changing trend of the loss value (Loss value) with the number of training rounds (Epoch) during the training process. The figure shows that the loss value drops rapidly in the early stages of training, which shows that the model has learned more effective features at the beginning and can quickly reduce the training error. Within the first 50 epochs, the loss value decreased rapidly, gradually decreasing from the initial 16 to close to 4, showing that the training effect of the model was better. As training progresses, the decline in loss values gradually slows down, but the overall decline shows a stable downward trend, indicating that the model is constantly optimizing its parameters and adapting to the training data.

When approaching 200 epochs, the loss value tends to be stable, about 2, which shows that the training effect of the model tends to be stable, and there is no obvious sign of overfitting. This phenomenon indicates that the model may be close to convergence, that is, the error on the current training set has reached a relatively low level and no longer decreases significantly. This convergence trend also implies that the adjustment space of the model parameters is smaller, indicating that the model has reached a better state under the current settings. Overall, this image reflects that the model's training process proceeds steadily and performs well as it gradually stabilizes.

5. Conclusion

This paper verifies the advantages of the Transformer model in emotion recognition by comparing the performance of multiple deep learning models on the task of electronic medical record sentiment analysis. The experimental results show that the Transformer model performs well in processing complex texts and capturing long-distance semantic relationships, and its accuracy and F1 score are higher than those of models such as LSTM, RNN, and BiLSTM. This shows that the Transformer model has high reliability in identifying emotional information in electronic medical records, providing strong support for clinical diagnosis and medical services.

In summary, the Transformer-based sentiment analysis method provides an innovative solution for the intelligent processing of electronic medical records. This method can not only help medical staff quickly obtain patients' emotional information, but also assist in the assessment of psychological status, which is of great significance to personalized medical services. Future research can further optimize the model performance and combine more real medical data sets to further improve the practicality and robustness of the model.

References

[1] Wang, M., Xie, Y., Liu, J., Li, A., Chen, L., Stromberg, A., ... & Wang, C. (2024). A Probabilistic Approach to Estimate the Temporal Order of Pathway Mutations Accounting for Intra-Tumor Heterogeneity. Cancers, 16(13), 2488.

[2] Hossain E, Rana R, Higgins N, et al. Natural language processing in electronic health records in relation to healthcare decision-making: a systematic review[J]. Computers in biology and medicine, 2023, 155: 106649.

[3] Nerella S, Bandyopadhyay S, Zhang J, et al. Transformers in healthcare: A survey[J]. arXiv preprint arXiv:2307.00067, 2023.

[4] Antikainen E, Linnosmaa J, Umer A, et al. Transformers for cardiac patient mortality risk prediction from heterogeneous electronic health records[J]. Scientific Reports, 2023, 13(1): 3517.

[5] Yang Z, Mitra A, Liu W, et al. TransformEHR: transformer-based encoder-decoder generative model to enhance prediction of disease outcomes using electronic health records[J]. Nature communications, 2023, 14(1): 7857.

[6] Denecke K, Reichenpfader D. Sentiment analysis of clinical narratives: a scoping review[J]. Journal of Biomedical Informatics, 2023, 140: 104336.

[7] Cao, J., Xu, R., Lin, X., Qin, F., Peng, Y., & Shao, Y. (2023). Adaptive receptive field U-shaped temporal convolutional network for vulgar action segmentation. Neural Computing and Applications, 35(13), 9593-9606.

[8] Chen, B., Qin, F., Shao, Y., Cao, J., Peng, Y., & Ge, R. (2023). Fine-grained imbalanced leukocyte classification with global-local attention transformer. Journal of King Saud University-Computer and Information Sciences, 35(8), 101661.

[9] Zi, Y., Cheng, X., Mei, T., Wang, Q., Gao, Z., and Yang, H., "Research on Intelligent System of Medical Image Recognition and Disease Diagnosis Based on Big Data", Proceedings of the 2024 IEEE 2nd International Conference on Image Processing and Computer Applications (ICIPCA), pp. 825-830, 2024.

[10] Cang, Y., Zhong, Y., Ji, R., Liang, Y., Lei, Y., and Wang, J., "Leveraging Deep Learning Techniques for Enhanced Analysis of Medical Textual Data", Proceedings of the 2024 IEEE 2nd International Conference on Sensors, Electronics and Computer Engineering (ICSECE), pp. 1259-1263, 2024.

[11] Fei, X., Chai, S., He, W., Dai, L., Xu, R., and Cai, L., "A Systematic Study on the Privacy Protection Mechanism of Natural Language Processing in Medical Health Records", Proceedings of the 2024 IEEE 2nd International Conference on Sensors, Electronics and Computer Engineering (ICSECE), pp. 1819-1824, 2024.

[12] Liu, S., Liu, G., Zhu, B., Luo, Y., Wu, L., and Wang, R., "Balancing Innovation and Privacy: Data Security Strategies in Natural Language Processing Applications", arXiv preprint arXiv:2410.08553, 2024.

[13] He, W., Bao, R., Cang, Y., Wei, J., Zhang, Y., and Hu, J., "Axial attention transformer networks: A new frontier in breast cancer detection", arXiv preprint arXiv:2409.12347, 2024.

[14] Yang, W., Wu, Z., Zheng, Z., Zhang, B., Bo, S., and Yang, Y., "Dynamic Hypergraph-Enhanced Prediction of Sequential Medical Visits", arXiv preprint arXiv:2408.07084, 2024.

[15] Du, J., Cang, Y., Zhou, T., Hu, J., and He, W., "Deep Learning with HM-VGG: AI Strategies for Multi-modal Image Analysis", arXiv preprint arXiv:2410.24046, 2024.

[16] Qin, H., Zheng, H., Wang, B., Wu, Z., Liu, B., and Yang, Y., "Reducing Bias in Deep Learning Optimization: The RSGDM Approach", arXiv preprint arXiv:2409.15314, 2024.

[17] Wang, B., Zheng, H., Liang, Y., Huang, G., and Du, J., "Dual-Branch Dynamic Graph Convolutional Network for Robust Multi-Label Image Classification", International Journal of Innovative Research in Computer Science & Technology, vol. 12, no. 5, pp. 94-99, 2024.

[18] Liang, Y., Liu, X., Xia, H., Cang, Y., Zheng, Z., and Yang, Y., "Convolutional neural networks for predictive modeling of lung disease", arXiv preprint arXiv:2408.12605, 2024.

[19] Liu, H., Zhang, B., Xiang, Y., Hu, Y., Shen, A., and Lin, Y., "Adversarial Neural Networks in Medical Imaging Advancements and Challenges in Semantic Segmentation", arXiv preprint arXiv:2410.13099, 2024.

[20] Gao, E., Yang, H., Sun, D., Xia, H., Ma, Y., and Zhu, Y., "Text classification optimization algorithm based on graph neural network", arXiv preprint arXiv:2408.15257, 2024.

[21] Wei, J., Liu, Y., Huang, X., Zhang, X., Liu, W., and Yan, X., "Self-Supervised Graph Neural Networks for Enhanced Feature Extraction in Heterogeneous Information Networks", arXiv preprint arXiv:2410.17617, 2024.

[22] Wu, Z., Gong, H., Chen, J., Yuru, Z., Tan, L., and Shi, G., "A Lightweight GAN-Based Image Fusion Algorithm for Visible and Infrared Images", arXiv preprint arXiv:2409.15332, 2024.

[23] Huang, G., Shen, A., Hu, Y., Du, J., Hu, J., and Liang, Y., "Optimizing YOLOv5s Object Detection through Knowledge Distillation algorithm", arXiv preprint arXiv:2410.12259, 2024.

[24] Sui, M., Hu, J., Zhou, T., Liu, Z., Wen, L., and Du, J., "Deep Learning-Based Channel Squeeze U-Structure for Lung Nodule Detection and Segmentation", arXiv preprint arXiv:2409.13868, 2024.

[25] Chen, J., Bao, R., Zheng, H., Qi, Z., Wei, J., and Hu, J., "Optimizing Retrieval-Augmented Generation with Elasticsearch for Enhanced Question-Answering Systems", arXiv preprint arXiv:2410.14167, 2024.

[26] Liu, W., Wang, R., Luo, Y., Wei, J., Zhao, Z., and Huang, J., "A Recommendation Model Utilizing Separation Embedding

and Self-Attention for Feature Mining", arXiv preprint arXiv:2410.15026, 2024.