World Journal of Innovation and Modern Technology, Vol. 7, Issue 4, (Aug) 2024
ISSN 2682-5910
DOI: 10.53469/wjimt.2024.07(04).05

A Novel Topic Segmentation Approach for Enhanced Dialogue Summarization

Zheng Ren

College of Computing, Georgia Institute of Technology, North Avenue, Atlanta, GA 30332
superpy001@gmail.com

Abstract: ***Dialogue summarization aims to distill a given conversation into a brief and focused summary. The challenge lies in the diverse perspectives of participants and the frequent shifts in topics throughout the dialogue. These factors make it difficult to extract and highlight the most significant information effectively. To tackle this challenge, we introduce a novel topic segmentation method that assigns distinct topics to dialogue segments while accounting for their importance and influence within the entire conversation. Our method's performance has been validated on two public benchmark datasets, CSDS and SAMSUM, demonstrating significant improvements in accuracy and coherence. The results show that our approach not only captures the essential content of dialogues more precisely but also enhances the overall quality and coverage of the summaries. This work provides a fresh approach to dialogue summarization and highlights its potential for practical application.***

Keywords: Dialogue Summarization; Topic Segmentation; Natural Language Processing.

1. INTRODUCTION

In recent years, the volume of dialogue data has been growing at an unprecedented rate, spanning various domains such as social media interactions and customer service conversations. These dialogues contain a wealth of information, but this rapid increase in data has led to the pressing issue of information overload. As a result, dialogue summarization technology has become a critical area of research. The primary goal of dialogue summarization is to distill the core content from lengthy conversations, creating concise and accurate summaries. This technology holds significant practical value, particularly in applications like customer service interactions, online chats, and meeting records, where it has been widely employed.

The task of dialogue summarization aims to extract key information from dialogues and produce a clear and concise summary. Since dialogues often contain a large amount of repetitive and irrelevant information, summarizing them can greatly improve information processing efficiency, helping users quickly grasp the essential content of the conversation, thereby reducing the time needed for reading and understanding. Compared to traditional text summarization, dialogue summarization faces more complex challenges, especially since each dialogue participant may hold unique perspectives and goals. These different roles convey information or reach consensus through interaction, and this multi-role, multi-goal dynamic adds complexity to the summarization process. Therefore, understanding and processing the semantic relationships between roles is crucial for generating high-quality summaries.

To address this challenge, we propose a novel Topic Segmentation Approach (NTSA) that aims to accurately capture the intentions and interactions of each role within the dialogue, thereby generating more coherent and meaningful summaries. Our method assigns a specific topic to each segment of the dialogue while considering its importance and influence within the overall dialogue network. Specifically, we adopt an unsupervised approach to evaluate the importance of sentences within a given dialogue, effectively identifying key utterances and their core content across different topics. During the decoding phase of summary generation, the model integrates topic relevance features, significantly enhancing the accuracy of information filtering and re-weighting, leading to an overall improvement in the quality of dialogue summaries.

To evaluate the effectiveness of our proposed NTSA, we applied it to three different types of sequence-to-sequence (seq2seq) architectures: Pointer-Generator Networks (PGN), Bidirectional Encoder Representations from Transformers (BERT), and Bidirectional and Auto-Regressive Transformers (BART). We conducted tests on two public dialogue summarization datasets, SAMSUM and CSDS. SAMSUM is a widely used English dialogue summarization dataset, while CSDS is a commonly used Chinese dialogue summarization dataset. Through multiple automatic evaluation metrics, we validated the quality of the generated summaries, and the experimental results show that NTSA significantly improves the performance of these seq2seq architectures on both datasets.

2. RELATED WORK

Document summarization techniques are broadly categorized into extractive and abstractive methods. Extractive summarization [1-3] involves selecting key sentences or phrases directly from the source text to form a summary. While this approach is straightforward and easy to implement, it often results in summaries that lack coherence and fail to effectively integrate the logical connections between pieces of information. On the other hand, abstractive summarization [4-6] goes beyond mere extraction; it involves deeply understanding the content, reorganizing, and creatively rephrasing the core information to generate summaries that are more fluent and cohesive. This method is better aligned with human editing practices, producing summaries that are not only rich in content but also articulate and insightful.

Abstractive summarization offers several distinct advantages. Firstly, it can handle complex semantic information by presenting the essential content in a novel way, which allows the summary to more accurately reflect the deeper meaning of the text. Secondly, when generating natural language text, abstractive methods pay more attention to the grammatical and semantic coherence, resulting in summaries that are more aligned with human reading patterns. However, these methods also present challenges, such as the need for deep semantic understanding and the ability to efficiently generate coherent language. For these reasons, we focus our research on abstractive methods, aiming to explore how information can be effectively extracted and reconstructed from text to produce high-quality summaries. This research has significant implications for text understanding and information retrieval, offering new directions for automated content generation and information synthesis.

In the field of dialogue summarization, while there are similarities to document summarization, there are also unique challenges and requirements. The primary goal of dialogue summarization is to distill key information from multi-turn conversations into concise and accurate summaries. These summaries have wide-ranging applications, including customer service interactions [7-9], meeting minutes, social media exchanges [10-13], and medical consultations [14-15]. In these contexts, accurately identifying speaker roles, tracking the evolution of conversation topics, and effectively extracting key information are critical challenges.

Unlike document summarization, dialogue summarization must address the diversity and complexity of conversational content [16], as well as the scarcity of annotated data. Dialogues are often unstructured, with significant semantic shifts and topic transitions, making the summarization task considerably more difficult. Additionally, capturing the intent of speakers and the interactions between different roles adds another layer of complexity to the process.

Recently, topic-based models [17] have gained attention in the development of dialogue summarization techniques. These models focus on tracking the shifts in topics throughout a conversation, aiming to produce summaries that are more focused and relevant. While these approaches have made some progress in improving the quality of dialogue summaries, they often fall short in terms of coherence, information coverage, and the overall grasp of the dialogue structure. Additionally, other recent methods, such as those based on data augmentation [18] and semantic structures [19], have sought to improve summarization outcomes by enhancing the training data and deepening semantic understanding. However, these approaches still face challenges in real-world applications, particularly in complex dialogue scenarios where the generated summaries may not fully meet user expectations.

Our research seeks to address the limitations of existing methods by introducing a novel topic segmentation approach that further enhances dialogue summarization effectiveness. Our work not only focuses on accurately capturing and representing the shifts in dialogue topics but also aims to improve the coherence and information coverage of the generated summaries, thus meeting the high demands of practical applications.

3. METHODOLOGY

To effectively capture each topic within a dialogue, we propose a novel approach called the Novel Topic Segmentation Approach (NTSA). It is well-known that the determination of the number of clusters is crucial in clustering tasks, as it significantly impacts the quality of the final results. In the context of dialogues, each utterance may contain unique information and varying semantic content. To ensure that this uniqueness is preserved during the initial stages of clustering, we treat each utterance as an independent cluster, setting the initial number of clusters to N. This strategy prevents the premature loss of essential information due to early clustering.

Building on this, we maintain a number of cluster centers such that $K<(N+1)$, allowing us to identify a central point for each utterance in the dialogue. Once the utterances are assigned to their respective clusters, we accurately determine the actual number of topics, K, and assign each utterance to its corresponding topic. This approach effectively addresses the challenge of manually setting the number of cluster centers, ensuring that the unique information within each utterance is preserved throughout the clustering process.

Our method begins by randomly selecting samples as initial cluster centers. The positions of these centers are then refined through an iterative process. Each utterance vector is assigned to the nearest cluster center based on the Euclidean distance between them. At the end of each iteration, the cluster centers are updated by calculating the mean of the points assigned to them. The algorithm continues to iterate until the maximum change in the position of the cluster centers falls below a predefined threshold (typically set at 0.001), or the process reaches a specified number of iterations (defaulting to 10). Once this stopping criterion is met, the algorithm terminates, producing the final cluster centers and the mapping of data points to these centers. This process enhances the measurement of relevance between the dialogue content and the identified topics, ensuring that significant semantic information is retained.

In the subsequent steps, each utterance is treated as a node within a graph structure, where the strength of connections between nodes is determined by the inner product of the hidden state vectors of the utterances. This graph-based structure allows us to assign a score to each utterance, accurately revealing its weight and influence within the broader dialogue network. By employing this method, we not only capture the deep semantic content of each utterance but also uncover its role and significance within the larger flow of the conversation.

The strength of this approach lies in its ability to adaptively discover the true topics within a dialogue without relying on predefined cluster numbers. By integrating clustering algorithms with graph theory, our method offers a more flexible and precise mechanism for topic segmentation. This not only enhances the coherence of the dialogue summaries but also improves the focus on relevant topics, ensuring that the generated summaries align more closely with user needs.

4. EXPERIMENTS

4.1 Datasets

To comprehensively evaluate the effectiveness of our proposed method, we selected two publicly available datasets: CSDS [1] and SAMSUM [10].

In the domain of role-oriented dialogue summarization, the Chinese datasets MC and CSDS are widely used. The MC dataset primarily focuses on doctor-patient interactions within medical consultations, where the dialogue typically revolves around the description of symptoms, diagnosis, and medical advice. Due to the relatively structured nature of these conversations, summarization performance on the MC dataset has already reached a high level of proficiency. In contrast, the CSDS dataset involves customer service dialogues, which present a more complex and diverse range of scenarios and content. These conversations cover a broader spectrum of topics and involve more dynamic interactions, posing greater challenges for dialogue summarization. Given the increased complexity and variability, we chose CSDS as our primary Chinese dataset for role-oriented dialogue summarization tasks.

For the English dataset, we focused on selecting dialogues that reflect everyday communication scenarios, complementing our focus on customer service interactions. These two dialogue types are among the most common in daily life, encompassing rich interactive dynamics and diverse linguistic expressions. This provides an excellent opportunity to train and evaluate models capable of handling complex dialogue structures. While both SAMSUM and DIALOGSUM are available as potential English datasets, they serve different purposes. DIALOGSUM is more suited for thematic summarization tasks, as it involves dialogues and summaries that center around specific topics. However, our research is more concerned with processing dialogues from everyday life, making SAMSUM the more appropriate choice. The SAMSUM dataset is particularly valued for its summaries, which are manually crafted with a strong emphasis on the natural flow of conversation and human readability.

By choosing CSDS and SAMSUM, we ensure that our dialogue summarization models are trained and evaluated on a diverse range of scenarios, covering both structured customer service dialogues and the more fluid and varied exchanges found in everyday interactions. This careful selection allows us to test the adaptability and robustness of our approach across different contexts, ensuring that our models can be effectively applied to a wide array of dialogue types beyond the specific scenarios provided by the datasets.

In conclusion, the combination of CSDS and SAMSUM datasets provides a broad and challenging testing ground for our summarization models. Their inclusion ensures that our method is not only capable of handling the intricacies of specific business and personal interactions but also demonstrates potential for broader application in other similar dialogue environments. This choice underpins the generalizability and effectiveness of our proposed approach, paving the way for its application across various domains.

4.2 Evaluation metrics

To thoroughly assess the performance of our text summarization model, we employed six distinct automatic evaluation metrics. These metrics provide a comprehensive evaluation from multiple perspectives, ensuring that we can accurately gauge the model's effectiveness in terms of content retention, linguistic fluency, structural coherence, and semantic accuracy. The specific metrics include ROUGE-1/2/L [20], BLEU [21], BERTScore [22], and MoverScore [23]. Below is a detailed explanation of each evaluation metric:

ROUGE is a widely used metric for summarization tasks, measuring the overlap between the generated summary and the reference summary. ROUGE-1 evaluates the overlap of unigrams (individual words), providing insight into the basic content coverage. ROUGE-2 extends this by considering bigrams (two-word sequences), offering a deeper look into the fluency and coherence of the summary. ROUGE-L focuses on the longest common subsequence between the generated and reference summaries, reflecting the overall structural similarity and the ability to preserve the sequence of ideas. Together, these variations of ROUGE provide a robust assessment of how well the generated summary aligns with the reference in terms of content and structure.

BLEU is another prominent metric traditionally used in machine translation but also applicable to summarization. It evaluates the generated summary by comparing the n-grams (n-word sequences) of the generated text with those of the reference summary. BLEU measures how well the generated summary retains key information and maintains linguistic fluency by assessing the precision of n-gram matches. A higher BLEU score indicates that the generated summary closely mirrors the reference in terms of word choice and phrasing, which is crucial for generating summaries that are both informative and easy to read.

BERTScore leverages the power of pre-trained BERT embeddings to evaluate the semantic similarity between the generated and reference summaries. Unlike traditional metrics that focus on surface-level n-gram overlap, BERTScore captures the deeper semantic relationships by computing cosine similarity between word embeddings in the generated and reference summaries. This makes BERTScore particularly effective for evaluating summaries where semantic accuracy and the preservation of meaning are paramount, even if the exact wording differs from the reference.

MoverScore is an advanced metric grounded in optimal transport theory, which measures the semantic distance between the generated and reference texts. By calculating the minimal cost of transforming the generated summary's word embeddings into those of the reference summary, MoverScore provides a nuanced measure of semantic similarity. This metric emphasizes the fidelity of the generated summary to the original text's meaning and the accuracy of detailed information. MoverScore is particularly valuable for assessing the quality of summaries in terms of preserving intricate details and ensuring that the overall meaning is faithfully conveyed.

4.3 Results

Table 1: Results on the CSDS dataset test set.

CSDS	ROUGE-1	ROUGE-2	ROUGE-L	BLEU	BERTScore	MoverScore
PGN	55.58	39.19	53.46	30.03	77.96	59.00
PGN-NTSA	56.89	40.75	55.51	32.92	78.75	59.55
BERT	53.87	37.59	52.40	29.00	78.52	58.23
BERT-NTSA	57.50	41.12	55.59	32.74	79.69	59.44
BART	59.07	43.72	57.11	34.33	79.74	60.11
BART-NTSA	**59.92**	**44.33**	**57.89**	**35.57**	**80.00**	**60.16**

Table 2: Results on the SAMSUM dataset test set.

CSDS	ROUGE-1	ROUGE-2	ROUGE-L	BLEU	BERTScore	MoverScore
PGN	40.08	15.28	36.63	37.49	80.67	59.87
PGN-NTSA	40.89	16.03	37.18	37.93	81.36	59.99
BERT	50.34	24.71	46.63	46.63	88.72	61.17
BERT-NTSA	51.11	25.00	46.92	47.57	89.49	61.50
BART	53.12	27.95	49.15	49.28	92.14	62.01
BART-NTSA	**53.63**	**28.73**	**49.59**	**50.32**	**92.58**	**62.17**

As shown in Tables 1 and 2, we present a comparative analysis of the performance of our NTSA-integrated seq2seq models against other integrated approaches. Across the CSDS and SAMSUM datasets, the BART-NTSA model consistently achieved the highest scores across all six automated evaluation metrics, with these results highlighted in bold in the tables. This demonstrates that incorporating our NTSA method significantly enhances performance, regardless of the underlying seq2seq architecture it is applied to.

Focusing on the CSDS dataset, we observed that user-generated summaries outperformed those generated by customer service agents. A deeper analysis of the dataset revealed the underlying reason for this phenomenon: user inquiries typically exhibit higher semantic focus, leading to more precise and concise summaries. In contrast, customer service responses tend to address a wider range of issues, often providing broader, less specific answers. This diversity in responses somewhat dilutes the specificity of the content, causing the customer service summaries to appear more generalized in terms of topic coverage. On the other hand, the consistency and focus of user inquiries result in summaries that are more accurate and specific.

In the dialogue summarization task, our comparative analysis also revealed that NTSA performed better overall on the CSDS dataset compared to the SAMSUM dataset. An examination of the generated summaries indicated that those within the CSDS dataset tend to be longer, providing the model with more opportunities to demonstrate its capabilities in enhancing information coverage and maintaining coherence. Furthermore, our analysis of the datasets suggested that Chinese dialogues often contain more nuanced semantic and contextual information, which requires a more sophisticated understanding and processing of deep semantic relations. In contrast, the syntactic structure of English dialogues is generally more straightforward and explicit, potentially requiring less reliance on in-depth semantic analysis.

These observations underscore the versatility and robustness of our NTSA approach, as it adapts effectively to the specific characteristics and challenges of different languages and dialogue structures. The results highlight the importance of context and language-specific features in dialogue summarization and demonstrate the potential of NTSA to significantly improve summary quality across diverse applications.

5. CONCLUSION

In this paper, we presented the Novel Topic Segmentation Approach (NTSA), a method designed to enhance dialogue summarization models. By integrating NTSA with seq2seq architectures, such as BART, we achieved notable improvements in generating accurate, coherent, and semantically rich summaries. Our approach was rigorously evaluated on two datasets, CSDS and SAMSUM, where the BART-NTSA model consistently outperformed other methods across multiple evaluation metrics.

The results highlighted the effectiveness of NTSA in handling the complexities of customer service dialogues in the CSDS dataset, particularly where user queries are more focused and specific. Additionally, our analysis revealed that NTSA adapts well to the nuanced semantic and contextual information present in Chinese dialogues, further demonstrating its versatility.

Overall, our research contributes a robust new tool for dialogue summarization, showing significant performance gains across diverse scenarios. The NTSA method offers a promising direction for future work, particularly in enhancing the adaptability of summarization models across different languages and dialogue types.

REFERENCES

[1] T. Berg-Kirkpatrick, D. Gillick, and D. Klein, "Jointly learning to extract and compress," in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (D. Lin, Y. Matsumoto, and R. Mihalcea, eds.), (Portland, Oregon, USA), pp. 481–490, Association for Computational Linguistics, June 2011.

[2] K. Filippova, E. Alfonseca, C. A. Colmenares, L. Kaiser, and O. Vinyals, "Sentence compression by deletion with LSTMs," in Proceedings of the 2015 Conferenceon Empirical Methods in Natural Language Processing (L. M`arquez, C. CallisonBurch, and J. Su, eds.), (Lisbon, Portugal), pp. 360–368, Association for Computational Linguistics, Sept. 2015.

[3] H. Jing and K. R. McKeown, "Cut and paste based text summarization," in 1st Meeting of the North American Chapter of the Association for Computational Linguistics, 2000.

[4] M. Banko, V. O. Mittal, and M. J. Witbrock, "Headline generation based on statistical translation," in Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics, (Hong Kong), pp. 318–325, Association for Computational Linguistics, Oct. 2000.

[5] S. Chopra, M. Auli, and A. M. Rush, "Abstractive sentence summarization with attentive recurrent neural networks," in Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (K. Knight, A. Nenkova, and O. Rambow, eds.), (San Diego, California), pp. 93–98, Association for Computational Linguistics, June 2016.

[6] R. Nallapati, B. Zhou, C. dos Santos, C¸ . Gul¸cehre, and B. Xiang, "Abstractive text summarization using sequence-to-sequence RNNs and beyond," in Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning (S. Riezler and Y. Goldberg, eds.), (Berlin, Germany), pp. 280–290, Association for Computational Linguistics, Aug. 2016.

[7] H. Lin, L. Ma, J. Zhu, L. Xiang, Y. Zhou, J. Zhang, and C. Zong, "CSDS: A finegrained chinese dataset for customer service dialogue summarization," in Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021, pp. 4436–4451, 2021.

[8] Y. Zou, L. Zhao, Y. Kang, J. Lin, M. Peng, Z. Jiang, C. Sun, Q. Zhang, X. Huang, and X. Liu, "Topic-oriented spoken dialogue summarization for customer service with saliency-aware topic modeling," vol. 35, pp. 14665–14673, May 2021.

[9] C. Liu, P. Wang, X. Jiang, Z. Li, and J. Ye, "Automatic dialogue summary generation for customer service," pp. 1957–1965, 07 2019.

[10] B. Gliwa, I. Mochol, M. Biesek, and A. Wawer, "SAMSum corpus: A humanannotated dialogue dataset for abstractive summarization," in Proceedings of the 2nd Workshop on New Frontiers in Summarization, pp. 70–79, 2019.

[11] X. Feng, X. Feng, L. Qin, B. Qin, and T. Liu, "Language model as an annotator: Exploring DialoGPT for dialogue summarization," (Online), pp. 1479–1491,Association for Computational Linguistics, Aug. 2021.

[12] L. Zhao, F. Zheng, W. Zeng, K. He, G. Ruotong, H. Jiang, W. Wu, and W. Xu,“Adpl: Adversarial prompt-based domain adaptation for dialogue summarization with knowledge disentanglement,” 07 2022.
[13] Y. Chen, Y. Liu, L. Chen, and Y. Zhang, “Dialogsum: A real-life scenario dialogue summarization dataset,” pp. 5062–5074, 01 2021.
[14] Y. Song, Y. Tian, N. Wang, and F. Xia, “Summarizing medical conversations via identifying important utterances,” (Barcelona, Spain (Online)), pp. 717–729, International Committee on Computational Linguistics, Dec. 2020.
[15] K. Krishna, S. Khosla, J. Bigham, and Z. C. Lipton, “Generating SOAP notes from doctor-patient conversations using modular summarization techniques,” (Online), pp. 4958–4972, Association for Computational Linguistics, Aug. 2021.
[16] Y. Zhang, A. Ni, T. Yu, R. Zhang, C. Zhu, B. Deb, A. Celikyilmaz, A. H. Awadallah, and D. Radev, “An exploratory study on long dialogue summarization: What works and what’s next,” (Punta Cana, Dominican Republic), pp. 4426–4433, Association for Computational Linguistics, Nov. 2021.
[17] H. Lin, J. Zhu, L. Xiang, F. Zhai, Y. Zhou, J. Zhang, and C. Zong, “Topic-oriented dialogue summarization,” vol. 31, pp. 1797–1810, 2023.
[18] Lyu, H., Wang, Z., & Babakhani, A. (2020). A UHF/UWB hybrid RFID tag with a 51-m energy-harvesting sensitivity for remote vital-sign monitoring. IEEE transactions on microwave theory and techniques, 68(11), 4886-4895.
[19] Lin, Z., Wang, Z., Zhu, Y., Li, Z., & Qin, H. (2024). Text Sentiment Detection and Classification Based on Integrated Learning Algorithm. Applied Science and Engineering Journal for Advanced Research, 3(3), 27-33.
[20] Wang, Z., Zhu, Y., Li, Z., Wang, Z., Qin, H., & Liu, X. (2024). Graph neural network recommendation system for football formation. Applied Science and Biotechnology Journal for Advanced Research, 3(3), 33-39.
[21] Zhu, Z., Wang, Z., Wu, Z., Zhang, Y., & Bo, S. (2024). Adversarial for Sequential Recommendation Walking in the Multi-Latent Space. Applied Science and Biotechnology Journal for Advanced Research, 3(4), 1-9.
[22] Wang, Z., Zhu, Y., He, S., Yan, H., & Zhu, Z. (2024). LLM for Sentiment Analysis in E-commerce: A Deep Dive into Customer Feedback. Applied Science and Engineering Journal for Advanced Research, 3(4), 8-13.
[23] M. Khalifa, M. Ballesteros, and K. McKeown, “A bag of tricks for dialogue summarization,” (Online and Punta Cana, Dominican Republic), pp. 8014–8022, Association for Computational Linguistics, Nov. 2021.
[24] X. Zhang, R. Zhang, M. Zaheer, and A. Ahmed, “Unsupervised abstractive dialogue summarization for tete-a-tetes,” vol. 35, pp. 14489–14497, May 2021.
[25] C.-Y. Lin and E. Hovy, “Manual and automatic evaluation of summaries,” in Proceedings of the ACL-02 Workshop on Automatic Summarization, pp. 45–51, 2002.
[26] K. Papineni, S. Roukos, T. Ward, and W.J. Zhu, “Bleu: a method for automatic evaluation of machine translation,” in Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, pp. 311–318, 2002.
[27] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, “Bertscore: Evaluating text generation with BERT,” in 8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020, 2020.
[28] W. Zhao, M. Peyrard, F. Liu, Y. Gao, C. M. Meyer, and S. Eger, “MoverScore: Text generation evaluating with contextualized embeddings and earth mover distance,” in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp. 563–578, 2019.
[29] Ren, Z. (2024). VGCN: An Enhanced Graph Convolutional Network Model for Text Classification. Journal of Industrial Engineering and Applied Science, 2(4), 110–115. https://doi.org/10.5281/zenodo.13118430
[30] Yuan, B., & Song, T. (2023, November). Structural Resilience and Connectivity of the IPv6 Internet: An AS-level Topology Examination. In Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering (pp. 853-856).
[31] Yuan, B., Song, T., & Yao, J. (2024, January). Identification of important nodes in the information propagation network based on the artificial intelligence method. In 2024 4th International Conference on Consumer Electronics and Computer Engineering (ICCECE) (pp. 11-14). IEEE.

[32] Yuan, B. (2024). Design of an Intelligent Dialogue System Based on Natural Language Processing. Journal of Theory and Practice of Engineering Science, 4(01), 72-78.

[33] Wang, Z. (2024, August). CausalBench: A Comprehensive Benchmark for Evaluating Causal Reasoning Capabilities of Large Language Models. In Proceedings of the 10th SIGHAN Workshop on Chinese Language Processing (SIGHAN-10) (pp. 143-151).

[34] Ren, Z. (2024). Enhanced YOLOv8 Infrared Image Object Detection Method with SPD Module. *Journal of Theory and Practice in Engineering and Technology*, *1*(2), 1–7. Retrieved from https://woodyinternational.com/index.php/jtpet/article/view/42