

Article

Not peer-reviewed version

Research on the Application of Large Language Models in Classification and Indexing

[Michael Reynolds](#) *

Posted Date: 6 May 2025

doi: [10.20944/preprints202505.0266.v1](https://doi.org/10.20944/preprints202505.0266.v1)

Keywords: Large language models; classification and indexing; topic term extraction; classification label assignment; cross-domain indexing; data privacy



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Research on the Application of Large Language Models in Classification and Indexing

Michael J. Reynolds

University of Michigan, USA; mjreynolds0203@gmail.com

Abstract: the rapid development of the information age, classification and indexing, as crucial components of knowledge organization and information retrieval, face critical challenges in efficiency and accuracy across various fields. Traditional classification and indexing methods exhibit significant shortcomings in semantic understanding, handling diverse datasets, and integrating domain knowledge. Large language models (LLMs), such as the GPT series, offer a new avenue for the intelligent transformation of classification and indexing through their powerful natural language processing capabilities. This study explores the application potential of LLMs in classification and indexing tasks. By analyzing the technical characteristics and current use cases of LLMs, the study investigates their performance in key tasks such as topic term extraction, classification label assignment, and cross-domain indexing support. Additionally, it examines challenges associated with applying LLMs, including matching model performance with classification requirements, addressing data privacy and ethical concerns, and integrating domain knowledge. Finally, by analyzing typical application cases, the study envisions the future development directions of LLMs in the field of classification and indexing. This research aims to provide theoretical foundations and technical support for the practical use of LLMs in classification and indexing tasks.

Keywords: large language models; classification and indexing; topic term extraction; classification label assignment; cross-domain indexing; data privacy

1. Introduction

Classification and indexing are central components of knowledge management and information retrieval, widely applied in fields such as library science, archival management, academic research, and e-commerce. These tasks primarily involve organizing, categorizing, and annotating large volumes of documents, data, or content to improve information retrieval efficiency. However, traditional classification and indexing methods, which often rely on manual annotation or rule-based automation tools, suffer from low efficiency, limited accuracy, and difficulties in handling diverse and complex datasets. These challenges have rendered traditional methods inadequate for meeting modern information processing demands. With the rapid advancements in big data and artificial intelligence technologies, developing efficient and intelligent classification and indexing systems has become a critical goal. In recent years, the emergence of large language models has provided a transformative solution for classification and indexing tasks. As cutting-edge achievements in deep learning, LLMs such as GPT and BERT leverage pre-training and fine-tuning techniques to efficiently understand and generate natural language. Their superior ability to deeply understand semantics and accurately capture contextual meaning significantly outperforms traditional algorithms. These characteristics make LLMs exceptionally effective in tasks like text classification, sentiment analysis, and topic extraction, laying a robust foundation for the intelligent transformation of classification and indexing. However, the practical application of LLMs still faces numerous challenges, including their integration with domain-specific knowledge, balancing performance with computational resources, and addressing data privacy and ethical issues. These problems demand systematic research and innovative solutions. This paper focuses on the application of LLMs in classification and indexing, aiming to explore their technical implementation, practical value, and future potential. It begins by

analyzing the technical principles of LLMs and their current applications in natural language processing. Subsequently, it examines the characteristics and requirements of classification and indexing tasks, as well as the application scenarios of LLMs in topic term extraction, classification label assignment, and cross-domain indexing support. The paper then discusses the challenges LLMs face in practical applications and proposes corresponding solutions. Finally, by exploring typical application cases, it envisions the future development directions of LLMs in the classification and indexing domain. This study seeks to provide both theoretical guidance and practical insights for leveraging LLMs to build more intelligent and efficient classification and indexing systems.

2. Overview of Large Language Models

2.1. Definition and Technical Foundation of Large Language Models

Large language models (LLMs) are natural language processing systems built using deep learning techniques, with a core objective of understanding and generating natural language text. Unlike traditional language models, LLMs typically employ neural networks based on the Transformer architecture, leveraging pre-training and fine-tuning to achieve deep learning of large-scale datasets. During pre-training, LLMs extract linguistic patterns, such as syntactic structures, semantic relationships, and contextual connections, from vast amounts of text data through unsupervised learning. In the fine-tuning stage, supervised learning is used to adapt the models to specific tasks, such as text classification, question answering, and machine translation. The technical foundation of LLMs includes the Transformer architecture, multi-head self-attention mechanisms, and large-scale distributed training. The Transformer architecture replaces the sequential processing nature of traditional recurrent neural networks (RNNs) with efficient parallel computation, significantly improving training efficiency and performance. The multi-head self-attention mechanism, a core component of Transformers, captures semantic relationships between different parts of a sentence, enhancing contextual understanding. Additionally, LLMs often utilize billions or even hundreds of billions of parameters, trained on high-performance computing clusters, providing the technical basis for their exceptional performance in language generation and understanding tasks.

2.2. Current Application Scenarios of Large Language Models

LLMs have been widely adopted in natural language processing and have achieved or surpassed human-level performance in many tasks. For example, in text classification tasks, LLMs excel in sentiment analysis, topic classification, and semantic annotation due to their precise contextual comprehension. In generative tasks, they demonstrate outstanding capabilities in text generation, automatic summarization, and machine translation, particularly in generating coherent, logically sound long texts. Furthermore, LLMs are highly effective in open-domain question answering, knowledge retrieval, and dialogue systems, leveraging their robust semantic understanding to handle complex queries and contextual dependencies, thus advancing intelligent question answering and human-computer interaction systems. Beyond traditional natural language processing domains, LLMs are also central to multi-modal learning, serving as the linguistic processing core in conjunction with image, audio, and other data types. This integration drives research and development in multi-modal intelligence. Additionally, LLMs have been extensively applied in specialized fields such as legal document analysis, medical knowledge extraction, and financial data processing, showcasing their flexibility and generality in supporting industry-wide digital transformation. In summary, with their powerful natural language processing capabilities and broad application prospects, LLMs offer both technical feasibility and practical opportunities for classification and indexing tasks. Through further research and performance optimization, their potential in the classification and indexing domain can be fully realized.

3. Characteristics and Requirements of Classification and Indexing

3.1. Core Tasks and Features of Classification and Indexing

Classification and indexing are fundamental tasks in knowledge management and information organization. They aim to logically and structurally organize textual or data resources by extracting topic terms and assigning classification labels. These tasks require a deep understanding of content semantics and precise application of domain knowledge, exhibiting the following distinctive features: Firstly, classification and indexing demand a high level of semantic understanding. The process involves not only identifying explicit information (e.g., keywords and phrases) but also uncovering implicit semantic relationships and contextual logic to accurately reflect the core themes of documents or data. Secondly, the complexity of classification tasks lies in the diversity of resource types and indexing requirements. For example, classification targets may include documents from different fields, social media content, and multi-modal data, each with varying language expressions and structural forms. Lastly, classification and indexing often require deep integration with domain knowledge to meet the standardization and specialization needs of specific industries or disciplines. For instance, medical document indexing must adhere to professional systems such as MeSH (Medical Subject Headings) to ensure scientific validity and usability. In summary, the complexity of classification and indexing arises from the diverse demands of tasks, deep semantic analysis, and close integration with domain knowledge. These features present significant challenges to traditional methods while providing a natural entry point for the adoption of LLMs.

3.2. Technical Requirements and Challenges in Classification and Indexing

Effective classification and indexing rely on efficient and accurate technical support, but traditional approaches face clear limitations in several areas: Firstly, efficiency and scalability pose major challenges. With the explosion of digital content, manual annotation is insufficient for processing vast datasets, while rule-based automation struggles with complex semantics, leading to inefficiencies and quality issues. Secondly, traditional methods lack adaptability to diverse datasets. As classification targets expand from structured to unstructured and multi-modal data, systems face increasing demands for semantic parsing and diversified data handling, which are difficult for conventional approaches to meet. Lastly, standardization and domain-specific adaptability remain critical challenges. Different domains require adherence to specific industry norms and systems, such as academic classification schemes in education, necessitating extensive domain knowledge and corpus resources that traditional methods cannot flexibly accommodate. Given these technical demands and challenges, LLMs provide a breakthrough solution with their superior natural language understanding and generation capabilities. By learning semantic patterns from vast datasets and adapting to diverse tasks through fine-tuning, LLMs offer transformative potential for classification and indexing. Future research should focus on leveraging LLMs to overcome these bottlenecks, enhancing the overall efficiency and intelligence of classification and indexing processes.

4. Exploring the Application of Large Language Models in Classification and Indexing

4.1. Implementation of Automatic Topic Term Extraction

Topic term extraction is a core task of classification and indexing, aimed at extracting keywords or phrases that accurately represent the theme of a text. Traditional methods, which often rely on statistical models or rule-based algorithms, are limited in handling complex semantics and diverse datasets. The introduction of large language models (LLMs) offers a transformative solution to this challenge. LLMs leverage pre-training on large-scale corpora to learn rich semantic associations and linguistic rules, giving them a significant advantage in processing complex semantics and contextual relationships. For instance, in academic literature indexing, LLMs can perform deep semantic

analysis to extract core topic terms with precision, without over-reliance on predefined rules. Moreover, fine-tuning enables LLMs to adapt rapidly to domain-specific requirements, demonstrating high efficiency and accuracy in specialized fields such as medicine and law. Additionally, LLMs excel in dynamic contextual analysis, extracting globally representative topic terms from lengthy texts, which is especially beneficial for cross-domain indexing and multi-modal data processing.

By applying LLMs to topic term extraction, classification and indexing processes experience significant efficiency gains. LLMs also enhance semantic parsing accuracy for complex datasets, addressing the limitations of traditional methods in uncovering implicit semantic information.

4.2. Optimization of Classification Label Assignment

Classification label assignment is another critical task in classification and indexing, aimed at automatically assigning predefined labels based on text content to facilitate efficient document organization and retrieval. Traditional methods, often using rule-based classifiers or basic machine learning models, struggle with multi-label assignment and semantic associations between labels. LLMs demonstrate strong capabilities in classification label assignment. Their semantic understanding and contextual analysis skills allow precise matching between text content and label semantics, ensuring accurate label assignment. LLMs can also handle multi-label scenarios; for example, in e-commerce product categorization, LLMs can analyze product descriptions to assign multiple relevant labels (e.g., "Electronics," "Smart Devices"), improving classification precision. Furthermore, LLMs' adaptability to cross-domain tasks expands possibilities for classification and indexing. In multilingual environments, LLMs can leverage cross-linguistic semantic understanding to achieve unified indexing across different languages. Additionally, LLMs support label expansion with few-shot learning techniques, enabling them to quickly adapt to new classification needs with minimal labeled samples. This facilitates dynamic updates and expansions of label libraries, providing robust technical support.

4.3. Cross-Domain Classification and Indexing Support

As knowledge expands and data types diversify, classification and indexing are increasingly moving toward cross-domain and multi-modal applications. Traditional methods often struggle to adapt to differences in datasets across domains. LLMs, however, exhibit significant advantages in cross-domain indexing support. LLMs balance generality with domain-specific adaptability, rapidly meeting the knowledge requirements of multiple fields through fine-tuning. For instance, in concurrent indexing tasks in education and healthcare, LLMs can seamlessly switch between medical and educational terminology. Moreover, LLMs show strong scalability in handling multi-modal data, such as integrating text and image information for classification and indexing. This capability supports the development of knowledge repositories that combine textual and visual content. Exploring LLM applications in classification and indexing reveals their immense potential in topic term extraction, classification label assignment, and cross-domain indexing support. These applications provide viable solutions to traditional bottlenecks in classification and indexing while laying the foundation for future large-scale knowledge organization.

5. Challenges and Solutions in Applications

5.1. Matching Model Performance with Practical Needs

Despite their strengths, LLMs face mismatches between their performance and practical requirements in classification and indexing tasks. Training and inference for LLMs often demand substantial computational resources, posing challenges for resource-constrained small and medium-sized enterprises. Additionally, while LLMs perform well in semantic understanding and generation, they may struggle with complex domain-specific datasets due to insufficient domain knowledge. For

instance, in specialized areas like medical or legal indexing, LLMs may show limitations in generalization for domain-specific terms. One feasible strategy is to enhance domain adaptability through fine-tuning and knowledge injection techniques. Fine-tuning on domain-specific datasets can enable LLMs to better understand industry terminology and contexts. Additionally, integrating structured domain knowledge via knowledge graphs into the models can improve their understanding of specialized content. Lightweight and efficient optimization, such as adopting smaller yet high-performance model variants, can reduce resource demands and enhance usability.

5.2. Data Privacy and Ethical Issues

Classification and indexing tasks often involve training and inference on large datasets that may contain sensitive or private information, such as user behavior records or medical case data. This raises risks of data breaches or privacy violations. Furthermore, models trained on imbalanced or biased datasets may produce skewed results, affecting the fairness and accuracy of indexing. Addressing these concerns requires robust data privacy protection techniques. Federated learning can enable distributed training without uploading raw data to centralized servers, reducing the risk of data leakage. Strengthening bias mitigation in training data, such as through data augmentation or re-sampling, can effectively reduce biases in model outputs. Establishing clear ethical guidelines and usage standards, such as restricting application scenarios, ensures that LLMs are used ethically and responsibly.

5.3. Deep Integration with Domain Knowledge

While LLMs excel in generalization and semantic understanding, effective classification and indexing often require integration with domain-specific knowledge. For example, medical classification must adhere to strict terminological systems, while educational indexing may require alignment with curriculum standards and academic taxonomies. Without deep domain knowledge integration, LLMs may produce inaccurate or non-standard results. The key to solving this problem lies in deep domain knowledge integration. Knowledge distillation techniques can incorporate expert knowledge into model weights, enhancing domain-specific adaptability in indexing tasks. External knowledge engines, such as embedding knowledge graphs into the inference process, provide dynamic support for domain knowledge. Few-shot learning techniques, leveraging a small number of labeled samples, enable LLMs to quickly adapt to new domain tasks, further improving flexibility and accuracy. In summary, the primary challenges in applying LLMs to classification and indexing include mismatches between model performance and practical needs, data privacy and ethical concerns, and insufficient domain knowledge integration. By optimizing model performance, incorporating data protection technologies, and enhancing domain knowledge integration, these challenges can be effectively addressed. These solutions not only improve the efficiency and quality of classification and indexing but also lay a foundation for continuous model optimization and innovation.

6. Application Prospects and Practical Cases

6.1. Analysis of Typical Application Scenarios

The application of large language models (LLMs) in classification and indexing has demonstrated immense potential across various fields, particularly in document management, knowledge retrieval, and business categorization. In the domain of libraries and document management, LLMs can achieve rapid classification and retrieval of vast volumes of documents through automated topic extraction and classification label assignment. For instance, in university libraries, LLMs can automatically assign topic tags based on textual content and categorize documents into corresponding academic directories, significantly improving the efficiency of document organization and the accuracy of retrieval. In academic research, especially in

multidisciplinary studies, LLMs can support cross-domain indexing tasks through deep semantic understanding. For example, research papers that address both artificial intelligence and healthcare often challenge traditional classification methods to index multiple disciplines simultaneously. LLMs, however, can identify semantic associations across disciplines based on content, enabling effective cross-domain categorization. Additionally, in the e-commerce sector, classification and indexing are extensively used in product categorization and recommendation systems. LLMs can swiftly analyze product descriptions and assign accurate multi-label classifications (e.g., "Electronics," "Home Appliances"), providing users with more precise search results and recommendations. These typical scenarios illustrate that LLMs not only enhance the efficiency of classification and indexing but also expand the boundaries of its applications, offering robust technical support for the intelligent transformation of various industries.

6.2. Future Development Directions for Large Language Models

Looking ahead, the development of LLMs in classification and indexing will further promote the intelligence and automation of indexing tasks while unlocking more possibilities for innovation in knowledge management and information retrieval. First, in improving indexing accuracy, future LLMs may incorporate more refined fine-tuning techniques and few-shot learning capabilities, enabling precise adaptation to multi-domain and multilingual data and addressing the generalization challenges currently faced in domain-specific indexing. Second, in the realm of multi-modal indexing, as technological breakthroughs continue, LLMs are expected to integrate with non-textual data such as visual and auditory information, paving the way for intelligent indexing systems based on multi-modal data. For example, in the media industry, LLMs could simultaneously process textual and image data to assign comprehensive classification labels to news content. Furthermore, LLMs will achieve greater explainability and standardization in indexing results through the support of knowledge graphs. By deeply integrating with knowledge graphs, LLMs can conduct indexing based on standardized domain knowledge, reducing ambiguities and biases in classification results. Meanwhile, advances in distributed computing and federated learning technologies will reduce the resource demands of LLMs, making them more accessible to a broader range of enterprises and institutions. In practical terms, future LLMs are likely to be widely applied in more industries, such as legal document management, medical case classification, and financial data analysis. These fields involve highly complex and specialized data, imposing higher requirements on classification and indexing tasks. The application of LLMs will not only enhance efficiency in these areas but also promote data structuring and standardization, driving intelligent management and sharing of domain knowledge. In conclusion, the application prospects of LLMs in classification and indexing are vast. Practical cases in document management, e-commerce, and cross-domain research demonstrate that this technology is gradually transforming traditional indexing methods. In the future, as technology continues to evolve and innovate, LLMs will play an increasingly significant role in advancing the intelligence, standardization, and diversification of classification and indexing, bringing profound changes to the fields of knowledge management and information retrieval.

7. Conclusion

Large language models, with their superior natural language processing capabilities, provide efficient and intelligent solutions for classification and indexing tasks. This paper explored their potential applications in topic term extraction, classification label assignment, and cross-domain indexing support, while analyzing challenges such as performance-demand alignment, data privacy, and domain knowledge integration. Corresponding solutions were proposed to address these issues. Through the study of practical cases, the significant advantages of LLMs in fields such as document management and e-commerce were demonstrated. Looking forward, as technology continues to advance, LLMs will play a greater role in improving indexing accuracy, supporting multi-modal data, and enhancing standardization, injecting new momentum into the intelligent transformation of knowledge management and information retrieval.

References

1. Wu Z. Deep learning with improved metaheuristic optimization for traffic flow prediction[J]. *Journal of Computer Science and Technology Studies*, 2024, 6(4): 47-53.
2. Thirunavukarasu, Arun James, et al. "Large language models in medicine." *Nature medicine* 29.8 (2023): 1930-1940.
3. Zhang J, Xiang A, Cheng Y, et al. Research on Detection of Floating Objects in River and Lake Based on AI Image Recognition[J]. *Journal of Artificial Intelligence Practice*, 2024, 7(2): 97-106.
4. Zhao Y, Hu B, Wang S. Prediction of brent crude oil price based on lstm model under the background of low-carbon transition[J]. *arXiv preprint arXiv:2409.12376*, 2024.
5. Yu Q, Wang S, Tao Y. Enhancing anti-money laundering detection with self-attention graph neural networks[C]//SHS Web of Conferences. EDP Sciences, 2025, 213: 01016.
6. Wang, G. Zhang, Y. Zhao, F. Lai, W. Cui, J. Xue, Q. Wang, H. Zhang, and Y. Lin, "Rpf-eld: Regional prior fusion using early and late distillation for breast cancer recognition in ultrasound images," in 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2024, pp. 2605-2612.
7. Mo K, Chu L, Zhang X, et al. Dral: Deep reinforcement adaptive learning for multi-uavs navigation in unknown indoor environment[J]. *arXiv preprint arXiv:2409.03930*, 2024.
8. Min, Liu, et al. "Financial Prediction Using DeepFM: Loan Repayment with Attention and Hybrid Loss." 2024 5th International Conference on Machine Learning and Computer Application (ICMLCA). IEEE, 2024.
9. Ma D, Wang M, Xiang A, et al. Transformer-Based Classification Outcome Prediction for Multimodal Stroke Treatment[J]. *arXiv preprint arXiv:2404.12634*, 2024
10. Li X, Cao H, Zhang Z, et al. Artistic Neural Style Transfer Algorithms with Activation Smoothing[J]. *arXiv preprint arXiv:2411.08014*, 2024.
11. Guo H, Zhang Y, Chen L, et al. Research on vehicle detection based on improved YOLOv8 network[J]. *arXiv preprint arXiv:2501.00300*, 2024.
12. Diao, Su, et al. "Ventilator pressure prediction using recurrent neural network." *arXiv preprint arXiv:2410.06552* (2024).
13. Demszky, Dorottya, et al. "Using large language models in psychology." *Nature Reviews Psychology* 2.11 (2023): 688-701.
14. Cheng Y, Yang Q, Wang L, et al. Research on Credit Risk Early Warning Model of Commercial Banks Based on Neural Network Algorithm[J]. *arXiv preprint arXiv:2405.10762*, 2024.
15. Tang, Xirui, et al. "Research on heterogeneous computation resource allocation based on data-driven method." 2024 6th International Conference on Data-driven Optimization of Complex Systems (DOCS). IEEE, 2024.
16. Wang T, Cai X, Xu Q. Energy Market Price Forecasting and Financial Technology Risk Management Based on Generative AI[J]. *Applied and Computational Engineering*, 2024, 100: 29-34.
17. Tan C, Zhang W, Qi Z, et al. Generating Multimodal Images with GAN: Integrating Text, Image, and Style[J]. *arXiv preprint arXiv:2501.02167*, 2025.
18. Yan, Hao, et al. "Research on image generation optimization based deep learning." *Proceedings of the International Conference on Machine Learning, Pattern Recognition and Automation Engineering*. 2024.
19. Tan C, Li X, Wang X, et al. Real-time Video Target Tracking Algorithm Utilizing Convolutional Neural Networks (CNN)[C]//2024 4th International Conference on Electronic Information Engineering and Computer (EIECT). IEEE, 2024: 847-851.
20. Yang H, Wang L, Zhang J, et al. Research on Edge Detection of LiDAR Images Based on Artificial Intelligence Technology[J]. *arXiv preprint arXiv:2406.09773*, 2024.
21. Xiang A, Qi Z, Wang H, et al. A Multimodal Fusion Network For Student Emotion Recognition Based on Transformer and Tensor Product[J]. *arXiv preprint arXiv:2403.08511*, 2024.
22. Zhao Y, Hu B, Wang S. Prediction of brent crude oil price based on lstm model under the background of low-carbon transition[J]. *arXiv preprint arXiv:2409.12376*, 2024.
23. Xiang A, Zhang J, Yang Q, et al. Research on splicing image detection algorithms based on natural image statistical characteristics[J]. *arXiv preprint arXiv:2404.16296*, 2024.

24. Xiang A, Huang B, Guo X, et al. A neural matrix decomposition recommender system model based on the multimodal large language model[J]. arXiv preprint arXiv:2407.08942, 2024.
25. Shih K, Han Y, Tan L. Recommendation System in Advertising and Streaming Media: Unsupervised Data Enhancement Sequence Suggestions[J]. arXiv preprint arXiv:2504.08740, 2025.
26. Wu, X., Sun, Y., & Liu, X. (2024). Multi-Class Classification of Breast Cancer Gene Expression Using PCA and XGBoost. Preprints. <https://doi.org/10.20944/preprints202410.1775.v1>
27. Wu Z, Wang X, Huang S, et al. Research on prediction recommendation system based on improved markov model[J]. Advances in Computer, Signals and Systems, 2024, 8(5): 87-97.
28. Ozkaya, Ipek. "Application of large language models to software engineering tasks: Opportunities, risks, and implications." IEEE Software 40.3 (2023): 4-8.
29. Zhang W, Huang J, Wang R, et al. Integration of Mamba and Transformer--MAT for Long-Short Range Time Series Forecasting with Application to Weather Dynamics[J]. arXiv preprint arXiv:2409.08530, 2024.
30. Meyer, Jesse G., et al. "ChatGPT and large language models in academia: opportunities and challenges." BioData Mining 16.1 (2023): 20.
31. Shi X, Tao Y, Lin S C. Deep Neural Network-Based Prediction of B-Cell Epitopes for SARS-CoV and SARS-CoV-2: Enhancing Vaccine Design through Machine Learning[J]. arXiv preprint arXiv:2412.00109, 2024.
32. Zhao R, Hao Y, Li X. Business Analysis: User Attitude Evaluation and Prediction Based on Hotel User Reviews and Text Mining[J]. arXiv preprint arXiv:2412.16744, 2024.
33. Rillig, Matthias C., et al. "Risks and benefits of large language models for the environment." Environmental Science & Technology 57.9 (2023): 3464-3466.
34. Mo K, Chu L, Zhang X, et al. DRAL: Deep reinforcement adaptive learning for multi-UAVs navigation in unknown indoor environment[J]. arXiv preprint arXiv:2409.03930, 2024.
35. Ziang H, Zhang J, Li L. Framework for lung CT image segmentation based on UNet++[J]. arXiv preprint arXiv:2501.02428, 2025.
36. Teubner, Timm, et al. "Welcome to the era of chatgpt et al. the prospects of large language models." Business & Information Systems Engineering 65.2 (2023): 95-101.
37. Gao, Dawei, et al. "Synaptic resistor circuits based on Al oxide and Ti silicide for concurrent learning and signal processing in artificial intelligence systems." Advanced Materials 35.15 (2023): 2210484.
38. Wu Z. Mpgaan: Effective and efficient heterogeneous information network classification[J]. Journal of Computer Science and Technology Studies, 2024, 6(4): 8-16.
39. Kasneci, Enkelejda, et al. "ChatGPT for good? On opportunities and challenges of large language models for education." Learning and individual differences 103 (2023): 102274.
40. Wang L, Cheng Y, Xiang A, et al. Application of Natural Language Processing in Financial Risk Detection[J]. arXiv preprint arXiv:2406.09765, 2024.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.