ChinaRxiv

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-202511.00173

Application Prospects, Risks, and Challenges of Large Language Models in the Maritime Domain

Authors: Jiayin Yang, Yin Yong, Gao Zhengli, Zhang Daihui, Wang Wenxin, Jing Qianfeng

Date: 2025-11-15T00:00:00+00:00

Abstract

The advent of large language models constitutes a new information technology revolution, with their applications infusing new vitality across various industries. This paper focuses on the application of large language models in the maritime domain, proposing their prospective applications as well as associated risks and challenges, aiming to provide guidance for the intelligentization and informatization of maritime operations. First, this paper introduces the concept and development trends of large language models. Second, it explores the broad application prospects of large language models in maritime contexts, encompassing maritime education and training, shipboard documentation management, navigational assistance, and the analysis and synthesis of lessons learned from maritime accidents. Concurrently, it identifies critical issues requiring resolution in maritime large language model research, including the provision of medical advice aboard vessels, integration of multimodal data, fulfillment of high data security requirements, delineation of liability boundaries, and challenges in aligning with practical maritime needs and application scenarios. The rapid development of large language models will undoubtedly inject new vitality into the maritime domain.

Full Text

Preamble

Received: June 21, 2025
Funding: National Key R&D Program of China (2022YFB4300803, 2022YFB4301402); Provincial Science and Technology Plan Project (2022JH1/10800096); Dalian Maritime University Teaching Reform Project "Building an Intelligent Teaching Model Based on Knowledge Graphs and AI—A Case Study of Shipbuilding Basics and Ship Principles and Cargo" (Project

No. 2024-12); Dalian Maritime University "Wisdom for the Future" Classroom Development Studio Funding Project (Project No. JGFZS202401).

Authors:
First Author: YANG Jiyin (1999–), male, Ph.D. candidate, primarily engaged in navigation aid research. E-mail: yangjiyin@dlmu.edu.cn
Corresponding Author: JING Qianfeng (1993–), male, Ph.D., primarily engaged in navigation aid and multi-sensor fusion research. E-mail: jqf_{dlmu}@dlmu.edu.cn
Address: No. 1 Lingshui Road, Dalian, Liaoning, China, Postal Code: 116026
Tel: 13149746632

The Prospects and Risk Challenges of Large Language Model Applications in Maritime Navigation

YANG Jiyin, YIN Yong, GAO Zhengli, ZHANG Daihui, WANG Wenxin, JING Qianfeng*
(Navigation College, Dalian Maritime University, Dalian 116026, China)

Abstract

The advent of large language models represents a new information technology revolution, bringing renewed vitality to all industries. This paper focuses on the application of large models in maritime navigation, proposing both prospects and risk challenges to provide references for the intelligentization and informatization of the maritime field. First, we introduce the concept and development trends of large language models. Second, we explore broad application prospects in maritime navigation, including navigation education and training, ship library and data management, ship-assisted navigation, and reflection on maritime accident lessons. Simultaneously, we identify key issues requiring resolution in maritime large language model research, including providing medical advice for ships, integrating multimodal data, meeting high data security requirements, clarifying responsibility boundaries, and addressing the difficulty of aligning with actual navigation needs and application scenarios. The rapid development of large language models will undoubtedly inject new vitality into the maritime domain.

Keywords: Large Language Model (LLM); Intelligent navigation; application prospect; risk challenge

1. Introduction

Since the beginning of the 21st century, artificial intelligence technology has achieved remarkable progress, with deep learning methods significantly improving algorithmic performance across various traditional domains, particularly in natural language processing [1]. In November 2022, OpenAI released ChatGPT (Chat Generative Pre-trained Transformer), a natural language processing tool

based on large language models (LLM), triggering a new wave of artificial intelligence. Compared with traditional chatbots, ChatGPT demonstrates superior language understanding and text generation capabilities, substantially broadening its application scope [2,3]. In scenarios such as text generation, knowledge retrieval, and polishing translation, ChatGPT has reached a professionally usable level. Following ChatGPT' s success, domestic and international teams have successively explored large language models with different implementations and application domains, developing corresponding chatbots such as Huawei' s Pangu series, Google' s Bard [4], Tsinghua University' s ChatGLM [5], Baidu' s Wenxin Yiyan [6], Alibaba' s Tongyi Qianwen [7], and iFlytek' s Spark Model. These teams have deeply explored vertical domains including office work, education, automotive, and healthcare, uncovering further potential of large language models [8].

Today, large language models have been widely applied across thousands of industries. In the shipping industry, China' s Ministry of Transport and six other ministries jointly issued the "Guiding Opinions on Intelligent Shipping Development." Intelligent shipping represents a new modern shipping business form that deeply integrates traditional shipping elements with modern information, communication, sensing, and artificial intelligence technologies. Large language models constitute an indispensable key technology for realizing intelligent shipping. This paper focuses on the application of large models in the shipping industry, proposing nine application scenarios including maritime education and training, navigation book and data management, and collision avoidance decision-making assistance. Simultaneously, we emphasize the risks and challenges brought by large models as an emerging technology, hoping to provide references for maritime researchers investigating large model applications in the navigation field.

2. Development History of Large Language Models

Large language models are considered one of the key technologies for future artificial general intelligence [9]. A language model (LM) is a technique for modeling the generation probability of word sequences, with its primary goal being to estimate the probability distribution of the next word or character given a context [10]. Language models can measure the linguistic plausibility of a sentence and are applied in various natural language processing (NLP) tasks. Language models play an important role in NLP, with two illustrative applications being: (1) Natural language generation evaluation: In tasks such as machine translation and text summarization, language models can assess whether generated sentences conform to human language patterns by calculating sentence probabilities. Sentences with higher probabilities in the language model are considered more consistent with human language habits, thereby improving the quality of natural language generation systems. (2) Corpus statistical analysis: Language models can analyze large corpora, learning relationships between words and contextual information through statistical modeling. These statistical methods

help us understand language patterns and characteristics, providing linguistic knowledge for other natural language understanding models. For instance, language models can calculate word co-occurrence frequencies and predict the next possible word, assisting in information retrieval and text classification tasks.

Language models can be categorized into statistical language models and neural network language models. Statistical language models use n-gram statistical methods to improve performance in retrieval and other tasks [11], while neural network language models can effectively utilize distributed word vectors to model contextual relationships in language.

2.1 Large-Scale Pretrained Language Models

In 2017, Google introduced the Transformer, a feature extractor employing self-attention mechanisms [12]. This innovation replaced previously common structures such as convolutional neural networks (CNN) and recurrent neural networks (RNN) in language model research. The Transformer architecture enhanced natural language understanding capabilities and enabled deeper semantic representation. By adopting self-attention mechanisms and stacked encoder-decoder structures, this architecture significantly surpassed CNNs and other deep learning models in semantic feature extraction and task-specific feature extraction. Its advantage lies in simultaneously considering both global information and local dependencies in text, thereby better capturing contextual and semantic connections.

The Transformer model consists of an encoder and a decoder, where the encoder processes input information and the decoder transforms encoded information into output results. Compared with traditional methods, the Transformer model can attend to contextual information simultaneously and trains more efficiently. Based on the innovative Transformer concept, a series of derivative methods have emerged, including Bidirectional and Auto-Regressive Transformers (BART) and Text-to-Text Transfer Transformer (T5) [13]. BART expands the model' s reasoning space by introducing noise to decompose original training texts, while T5 uses a complete Transformer structure as a pretrained language model, converting various problems into a text-to-text format.

Currently, Transformer-based language models follow two main technical paradigms: pretraining + fine-tuning (represented by BERT) and pretraining + prompting (represented by ChatGPT). Additionally, a series of new models using only partial Transformer structures have emerged, including encoder-based and decoder-based models. These developments have created vigorous growth opportunities for the large-scale pretrained language model field. Their relationships are illustrated in [Figure 1: see original paper].

2.2 Trends in Large Language Model Development

Multimodality: Large language models are beginning to integrate multiple data types including text, speech, video, and images, gradually evolving into

more complex models that exhibit higher-level intelligence.

Rationalization of Model Scale: Initially, large language models improved performance by increasing model scale, but the trend is shifting toward balancing model scale with economic costs.

Shift from Model Research to Model Application: Large language models are no longer confined to basic model research but are actively developing model ecosystems and expanding applications in domains such as government affairs.

3. Applications of Large Language Models in Maritime Navigation

3.1 Maritime Education and Training

ChatGPT can generate teaching content such as lecture notes, textbooks, and case analyses to meet maritime education needs [14,15]. It can also provide personalized learning resources based on students' needs and knowledge levels, helping them better understand and master maritime knowledge [16]. Furthermore, it can automatically generate maritime examination questions, reducing teachers' workload while ensuring question quality and accuracy. Additionally, it can recommend relevant learning resources based on students' learning situations and preferences, enabling personalized learning path recommendations.

AI technology combining ChatGPT with metaverse, virtual reality, and augmented reality can drive innovation in maritime education [17,18]. Students can simulate actual maritime operations and situation handling through virtual scenarios, interact with virtual crew members [19], and conduct real-time maritime practice training and decision-making exercises. In this process, ChatGPT serves as a virtual mentor for students, providing guidance, feedback, and answers to questions, thereby enhancing learning effectiveness and practical abilities.

Maritime education involves extensive knowledge domains, and large language models can help students quickly acquire and understand complex maritime knowledge. Through natural language processing technology, large language models generate teaching materials, provide intelligent Q&A, and offer personalized learning suggestions, achieving personalized and efficient maritime education. Simultaneously, they assist maritime instructors in teaching evaluation and knowledge point mastery. By analyzing students' learning situations and knowledge levels, teachers can better guide students and improve the quality of maritime teaching.

3.2 Ship Library and Data Management

Navigational books and materials are voluminous and complex, making effective management and knowledge retrieval challenging. For instance, when seeking

operational steps for activating a specific function in navigation equipment manuals, ChatGPT enables users to input the manual into a large language model and obtain knowledge through questioning, providing a simple and rapid solution.

Moreover, ChatGPT can automate reading and knowledge extraction from maritime literature [20]. It can read maritime documents and generate concise summaries that capture main content, helping researchers quickly grasp key points and saving reading time. Additionally, based on user questions, ChatGPT can extract relevant knowledge from maritime literature and provide answers [21], assisting researchers in rapidly locating and utilizing knowledge within documents [22].

Large language models demonstrate significant advantages in processing massive amounts of maritime literature, voyage records, and navigation reports [23]. They can efficiently perform text summarization, entity recognition, and relationship extraction, helping maritime researchers quickly extract valuable information from large text corpora. Such automated text processing accelerates maritime knowledge accumulation and exchange. Furthermore, large language models hold potential advantages in knowledge graph construction and maritime knowledge base updating. By processing multi-source maritime data, large language models can help build more comprehensive and accurate knowledge graphs and knowledge bases, promoting in-depth maritime research.

3.3 Ship-Assisted Navigation

ChatGPT large language models can provide decision-making suggestions for ship navigation. Maritime navigation is highly complex, requiring captains to develop personalized maneuvering and decision-making plans based on different situations. Large language models possess powerful predictive and classification capabilities, and when combined with specific ship information and rich navigation data, they can provide customized navigation decision support to improve operational effectiveness. presents a case where the authors used ChatGPT to obtain collision avoidance decisions.

3.4 Maritime Accident Lesson Reflection and Summarization

Combined with speech recognition technology, ChatGPT can converse with crew members in dialogue form to obtain detailed information about ship accidents. Simultaneously, large models can assist in tasks such as determining responsibility attribution based on specific on-site conditions, as shown in . After maritime accidents occur, ChatGPT can dialogue with crew members, asking questions about crew conditions, weather conditions, and crossing traffic situations. It can collect, classify, and integrate accident information from multiple sources faster than humans, enabling maritime regulatory authorities to more comprehensively understand accident circumstances and providing a foundation for subsequent accident lesson summarization and responsibility allocation.

3.5 Providing Medical Advice for Crew

Research on large language models in the medical field is developing rapidly. Medical domain large models such as Bencao and Mingyi can provide more professional medical advice, and specialized large language models for medical vertical domains, such as MedChat for vertigo treatment, are also rapidly evolving.

Large language models can play the role of professional doctors, providing physical and mental health support for crew members [24]. For crew members' potential questions regarding medical conditions, treatment methods, and drug side effects, ChatGPT can provide answers and explain medical terminology [25], enabling them to accurately administer medication and promptly assist injured personnel without professional medical support. Additionally, ChatGPT can provide health education [27] and rehabilitation guidance [28] to crew members, helping them better manage their recovery process and accelerate healing [29].

3.6 Maritime Law Domain

Large language models can combine maritime cases with admiralty law to analyze cases, sort out case clues, and provide decision support for judges.

3.7 Communication Message Understanding

Large language models are not limited to text tasks. Through multimodal development, they can combine with speech recognition tasks. By inputting VHF maritime communication audio, large models can organize maritime situations, help officers enhance situational awareness, and automatically generate VHF logs, providing more support for navigation records.

3.8 Maritime Environment Perception

Combined with image recognition technology, large language models can analyze information about sea conditions, obstacles, and vessels to organize maritime traffic situations, predict collision risks, and provide reasonable navigation decisions based on the International Regulations for Preventing Collisions at Sea (COLREGs). By integrating ship historical data and existing weather forecast information into text input for large models, they can perform reasoning based on databases and historical routes to provide optimized navigation routes.

4. Problems to be Addressed in Maritime Large Language Models

4.1 Enriching Multimodal Data

Humans use multiple senses to receive multimodal information, enabling effective cognition and task processing. The same applies to large language models. Currently, large language models primarily rely on textual information,

and multimodal information input can supplement knowledge and enable more tasks. Presently, large language model applications in the maritime domain concentrate on text processing scenarios, yet maritime data encompasses not only text but also video, images, and voice data. Domestic and international products such as GPT-4, KOSMOS-1 [30], and Chinese products like "Wenxin Yiyan" and "Tongyi Qianwen" already support multimodal data processing. Future research in the maritime domain must investigate large language model applications for processing video, image, and voice data types.

4.2 Meeting High Data Security Requirements

Since maritime data involves personal privacy and commercial confidentiality, data privacy and security issues must be considered when applying maritime large language models. These concerns include data leakage, unauthorized data access, and data misuse. Maritime data contains sensitive ship information such as MMSI, ship call signs, and trajectory data, necessitating effective protective measures. Solutions include using data encryption technologies to protect data transmission and storage, establishing strict data access and permission control mechanisms to authorize only specific personnel to access sensitive data, and formulating appropriate data sharing policies and contracts that clearly define data usage purposes and limitations while maintaining consistency with relevant laws and regulations. Additionally, large language models may generate large amounts of convincing but unsourced content within short timeframes, making responses appear highly credible while actually lacking reliable foundations, which can seriously mislead users with lower critical judgment capabilities. Therefore, enhanced monitoring and auditing of data security are required to promptly detect and respond to security vulnerabilities and threats.

4.3 Clarifying Responsibility Boundaries

In ship management, the captain, as the primary decision-maker, bears responsibility for decisions made. However, when captains use large language models, they may face situations where responsibility is difficult to define. First, large language model outputs are unpredictable, making it challenging for captains to make reasonable decisions based on these outputs. Second, large language models are trained on existing data, essentially learning existing knowledge and performing reasoning, which makes it difficult for them to make rational judgments about emergency handling and future trend prediction in maritime scenarios. Finally, ship decision-making is based on the principle of "maximizing ship interests," while large language models represent machine will; whether they can reflect the captain' s will remains to be verified. The "Interim Measures for the Management of Generative Artificial Intelligence Services" stipulates that providers should "improve the accuracy and reliability of generated content and take effective measures to enhance the transparency of generative artificial intelligence services." In practical large language model usage, responsibility boundaries must be clarified, the alignment between model decisions

and human decisions must be reasonably defined, machine will and captain will must be correctly distinguished, and maximum effort must be made to enable large language models to make decisions consistent with the public will of the ship. Researching human-in-the-loop decision-making processes is a promising approach to ensure final decisions are made by humans, avoiding errors and information mismatches caused by AI hallucinations.

4.4 Disconnect from Actual Maritime Needs and Application Scenarios, Difficult Implementation

The maritime domain is highly complex and specialized, requiring in-depth maritime knowledge and professional judgment. While large language models may demonstrate excellent general language understanding capabilities, their understanding of maritime domain-specific knowledge and professionalism remains limited. This makes applying large language models to decision support and on-site guidance in maritime scenarios challenging, as illustrated by the experiments referenced in .

5. Conclusion

Large language models represented by ChatGPT and DeepSeek possess application potential in numerous maritime-related domains including maritime education, navigation book and data management, assisted navigation, and maritime accident lesson summarization, with the capability to reshape maritime industry service models. Although facing many challenges in technology, law, security, and social issues, the advent of large language models has already initiated major transformations in the maritime domain. Only by keeping pace with technological development trends can we ride the wave and achieve success.

References

[1] Wang Y, Hu M, Ta N, et al. Large Language Models and Their Applications in Government Affairs[J]. Journal of Tsinghua University (Science and Technology), 2024, 64(04): 649-658.
[2] Mann D L. Artificial Intelligence Discusses the Role of Artificial Intelligence in Translational Medicine[J]. JACC: Basic to Translational Science, 2023, 8(2): 221-223.
[3] Zhou C, Li Q, Li C, et al. A comprehensive survey on pretrained foundation models: A history from bert to Chatgpt[J]. International Journal of Machine Learning and Cybernetics, 2024:
[4] Lin T, Wang Y, Liu X, et al. A Survey of Transformers[J]. CoRR, 2021, abs/2106.04554.
[5] Tay Y, Dehghani M, Bahri D, et al. Efficient Transformers: A Survey[J]. CoRR, 2020, abs/2009.06732.
[6] Radford A, Narasimhan K, Salimans T, et al. Improving language understanding by generative Pre-Training[J]. 2018. San Francisco, CA, USA, 2018.

[7] Radford A, Wu J, Child R, et al. Language models are unsupervised multitask Learners[J]. OpenAI Blog, 2019, 1(8): 9.
[8] Qin S, Zheng Z, Gu Y, et al. Application Testing and Discussion of Large Language Models in Construction Engineering[J]. Industrial Construction, 2023, 53(09): 162-169.
[9] Xu Y, Hu L, Zhao J, et al. Technical Application Prospects and Risk Challenges of Large Language Models[J]. Computer Applications, 2024, 44(06): 1655-1662.
[10] Ma W, Gong M, Dai H, et al. A Review of Applications of Large Language Models Represented by ChatGPT in Clinical Medicine[J]. Journal of Medical Informatics, 2023, 44(07): 9-17.
[11] Hu Z, Yang R, Zhu J, et al. Research and Application Development of Large Language Models in the Medical Field[J]. Artificial Intelligence, 2023(04): 10-19.
[12] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you Need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 6000-6010.
[13] Raffel C, Shazeer N, Roberts A, et al. Exploring the limits of transfer learning with a unified Text-to-text Transformer[J]. J. Mach. Learn. Res., 2020, 21(1).
[14] Baï doo-Anu D, Ansah L O. Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning[J]. Journal of AI, [no date], 7(1): 52-62.
[15] Opara E, Mfon-Ette Theresa A, Aduke T C. ChatGPT for teaching, learning and research: Prospects and Challenges[J]. Opara Emmanuel Chinonso, Adalikwu Mfon-Ette Theresa, Tolorunleke Caroline Aduke (2023). ChatGPT for Teaching, Learning and Research: Prospects and Challenges. Glob Acad J Humanit Soc Sci, 2023, 5.
[16] Zhai X. Chatgpt and ai: The game changer for Education[J]. Zhai, X.(2023). ChatGPT: Reforming Education on Five Aspects. Shanghai Education, 2023: 16-17.
[17] Talan T, Kalinkara Y. The Role of Artificial Intelligence in Higher Education: ChatGPT Assessment for Anatomy Course[J]. Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi, 2023, 7(1): 33-40.
[18] Paranjape K, Schinkel M, Panday R N, et al. Introducing Artificial Intelligence Training in Medical Education[J]. JMIR Medical Education, 2019, 5(2): e16048.
[19] Seetharaman R. Revolutionizing Medical Education: Can ChatGPT Boost Subjective Learning and Expression?[J]. Journal of Medical Systems, 2023, 47(1): 61.
[20] Arif T B, Munaf U, Ul-Haque I. The future of medical education and research: Is ChatGPT a blessing or blight in Disguise?[J]. Medical Education Online, 2023, 28(1): 2181052.
[21] Vaishya R, Misra A, Vaish A. ChatGPT: Is this version good for healthcare and Research?[J]. Diabetes & Metabolic Syndrome: Clinical Research &

Reviews, 2023, 17(4): 102744.
[22] Gao C, Howard F, Markov N, et al. Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human Reviewers[M]. 2022.
[23] Manning C D. Human Language Understanding & Reasoning[J]. Daedalus, 2022, 151:
[24] Sallam M. The utility of ChatGPT as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations[J]. MedRxiv, 2023: 2023–02.
[25] Javaid M, Haleem A, Singh R P. ChatGPT for healthcare services: An emerging stage for an innovative Perspective[J]. BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 2023, 3(1): 100105.
[26] Lyu Q, Tan J, Zapadka M E, et al. Translating radiology reports into plain language using ChatGPT and GPT-4 with prompt learning: Results, limitations, and Potential[J]. Visual Computing for Industry, Biomedicine, and Art, 2023, 6(1): 9.
[27] Sallam M, Salim N, Barakat M, et al. ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and Limitations[J]. Narra J, 2023, 3(1): e103–e103.
[28] Cheng K, Sun Z, He Y, et al. The potential impact of ChatGPT/GPT-4 on surgery: Will it topple the profession of Surgeons?[J]. International Journal of Surgery, 2023, 109(5): 1545.
[29] Singhal K, Azizi S, Tu T, et al. Large language models encode clinical Knowledge[J]. Nature, 2023, 620(7972): 172–180.
[30] Gülsoylu E, Koch P, Yildiz M, et al. Image and AIS Data Fusion Technique for Maritime Computer Vision Applications[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops. 2024: 859–868.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.