**Journal of Computing and Electronic Information Management**
ISSN: 2413-1660 | Vol. 15, No. 3, 2024

# Application of Large Language Models in Power System Operation and Control

**Yiqian Zhang***

Electronic Information Engineering, International Education College, Xiamen University of Technology, Xiamen, Fujian Province, 361000, China

* **Corresponding author Email:** Kricese@163.com

**Abstract:** The introduction of "carbon peak" and "carbon neutrality" targets and the vigorous advancement of national energy market construction, renewable energy sources like wind and solar power are experiencing rapid development. However, this also brings challenges in accounting for uncertainties in various operational scheduling and optimization control processes of the power system, especially for the theoretical control methods. Fortunately, Large language model (LLM) with rapid development has shown promising prospects in the power sector. This review summarizes the application of LLM technology in power system operation and control, outlining the new power system's demand for AI technology, the impact of LLM on system management, and the technological foundation including network architecture, training methods, and data configuration. Finally, it explores the applications of LLM in power system operation and control from the perspectives of generation, transmission, distribution, consumption, and equipment.

**Keywords:** Large language model; Power system; Artificial intelligence.

## 1. Introduction

With the deepening advancement of decarbonization and clean energy in the power industry, increasing the proportion of renewable energy sources, primarily wind and solar power, is a crucial approach to achieving a cleaner and lower-carbon electricity grid transition. Unlike conventional fossil fuels, renewable energy exhibits significant volatility, intermittency, and uncertainty, making it difficult to accurately predict its output curves. To accommodate a high proportion of renewable energy, the power system needs to address uncertainties in various operational scheduling and optimization control processes, which poses significant challenges to the grid's security, stability, and low-carbon operation. Balancing safety and economic considerations while scheduling units to ensure adequate reserve capacity is a key issue for increasing renewable energy integration, promoting low-carbon development in the power sector, and reducing grid operation costs.

To optimize the power source structure, the power system has begun focusing on developing uncertainty optimization scheduling methods that coordinate generation, grid, load, and storage. These methods use distributed resources like energy storage and demand response to smooth out the grid's uncertainty and volatility [1][2]. The multi-dimensional nature of distributed resources adds complexity to existing centralized optimization problems and increases the difficulty of traditional mathematical optimization methods. Compared to traditional centralized approaches, distributed optimization methods enhance the efficiency of solving grid operation and control optimization problems and facilitate the coordinated optimization of various distributed resources. However, distributed optimization methods still rely on accurate generation and load forecasting and mathematical modelling, requiring the decoupling of multiple optimization subproblems and iterative calculations for problem-solving. Traditional mathematical modelling approaches face limitations in modelling accuracy and computational precision, making it challenging to achieve the desired scheduling outcomes when dealing with grid nonlinearity, strong randomness, and uncertainty.

Unlike traditional model-based methods, artificial intelligence technologies that rely on deep learning and reinforcement learning algorithms offer a 'model-free' technical advantage in optimization and control problems. These technologies exhibit strong pattern recognition and rapid decision-making capabilities when assisting grid operators with operational analysis and control decisions. Additionally, with the rapid growth of data and advancements in computational resources, Large language models represented by Chat-GPT have made significant breakthroughs. They have demonstrated tremendous application value across various domains, scenarios, and tasks. The vertical deepening application of these models in the energy and power sectors is showing both urgent demand and broad prospects. LLM technology refers to the use of massive neural network models with numerous parameters to address complex natural language processing and cognitive tasks. Representative examples of these large models include the GPT (Generative Pre-trained Transformer) series, such as GPT-3 and GPT-4, and the BERT (Bidirectional Encoder Representations from Transformers) model. These models, through pre-training on extensive datasets, can acquire rich semantic and linguistic knowledge, enabling capabilities such as text generation, question answering, and semantic understanding. The rise of these large models has led to breakthroughs in language processing, intelligent dialogue, and knowledge reasoning [3][4].

According to the above, LLMs are effective tools for handling natural language and multimodal big data, and their deepening application in the energy and power sectors also reveals broad development prospects. In the field of power system operation and control, the current paradigm of human-machine hybrid intelligence involves using human cognitive intelligence to handle all tasks and business at various levels of semantic knowledge. On this basis, clearly defined

numerical problems, including simulation, computation, optimization, and decision-making, are addressed using simulators, optimizers, and AI algorithms. However, this paradigm will be disrupted by AI large models. Artificial intelligence will partially replace abstract and semantic-level issues that previously relied solely on human cognitive intelligence. Consequently, the intelligence of grid operation and control will advance significantly, leading to substantial transformation.

This paper will first introduce what is large language mode and how does it be used in vertical industry technologies. It will then provide a comprehensive overview and outlook on the development of new theoretical methods and engineering applications of LLMs in power system.

# 2. Concept of LLM

## 2.1. The architecture of LLM

Transformer-based large language models have made significant advances in the field of natural language processing [5]. Through unsupervised pre-training and autoregressive generation, these models can learn rich semantic information from vast corpora, achieving impressive performance across various NLP downstream tasks. However, real-world tasks often involve complex information from multiple modalities, such as text, vision, and audio. Current large language models are limited to understanding textual information, which greatly restricts their task handling capabilities and variety. Therefore, multimodal large models that support unified knowledge representation of text, images, and audio represent a major direction for future advancements in artificial intelligence.

Prior to the explosion of LLM technology, the technical approaches for multimodal large models primarily focused on developing general foundational architectures capable of handling multiple modalities of data and their downstream tasks, as illustrated in Figure 1. For instance, in the case of image and text modalities, early multimodal models predominantly employed architectures combining text encoders with pre-trained visual object detectors [7]. However, since the pre-trained object detectors no longer participated in end-to-end training, it became challenging to align the distribution of text and image information in the latent space. The Transformer architecture demonstrate substantial representational and modelling capabilities in the image domain. Consequently, subsequent multimodal models shifted towards unified architectures based on Transformer encoding. These models have evolved from being capable of only text-image understanding and discriminative tasks to accommodating generative tasks as well.

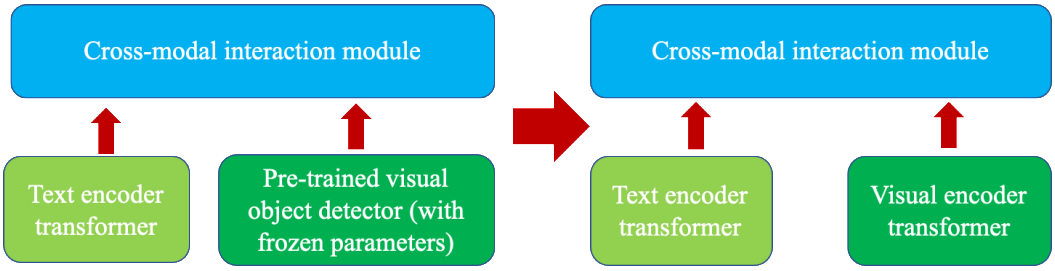


**Figure 1.** General architecture for multimodal models

Following the explosion of LLMs, the parameter scale of large models has surpassed the hundred-billion mark, making end-to-end training of multimodal models that integrate tens-of-billion parameter pre-trained LLMs extremely costly. Consequently, multimodal large model architectures like BLIP-2, which use frozen parameters of LLMs as the backbone, have been proposed [8]. As shown in Figure 2, the BLIP-2 architecture freezes the parameters of pre-trained image encoders and LLMs, and introduces a lightweight Query Transformer (Q-Former). This approach uses a two-stage training process—representation learning and generation learning—to bridge the modality gap between visual features and LLM text features. It achieves significant improvements in downstream multimodal task accuracy with a small number of parameters and has become a key reference architecture for subsequent multimodal large models.

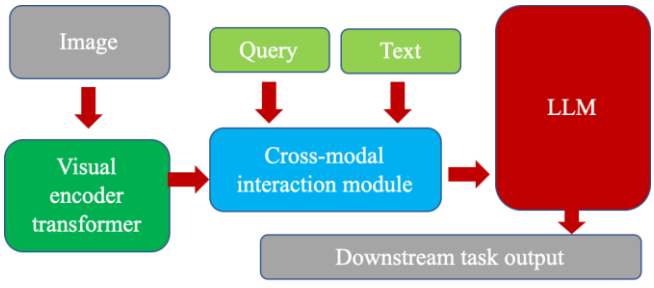


**Figure 2.** BLIP-2 multimodal large model architecture

## 2.2. Loss function of LLM

Using image and text modalities as an example, the downstream tasks of multimodal models primarily include image captioning, visual question answering, image generation, and cross-modal retrieval. For these tasks, the mainstream loss functions used during the training phase of multimodal large models generally fall into the following four categories [9].

Image-text contrastive (ITC) loss: ITC is based on the idea of contrastive learning. It learns shared representations of images and texts by maximizing the similarity of positive pairs (images and texts with the same content) and minimizing the similarity of negative pairs (images and texts with different content). The goal of the ITC loss function is to map similar images and texts to close regions in the representation space, thereby achieving cross-modal semantic alignment and relevance modelling

Image-text matching (ITM) loss: ITM loss aims to measure the semantic matching between images and texts. It constructs a binary classification task to maximize the matching relationship between image-text pairs. Compared to ITC, ITM loss places greater emphasis on the alignment between specific image-text pairs.

Masked language modelling (MLM) loss: MLM involves randomly selecting and masking certain words in the input sequence, and then requiring the model to predict the masked words based on the context. By forcing the model to learn semantic and syntactic information from the context, MLM

helps the model develop more comprehensive language representations. This self-supervised learning approach enhances the model's understanding and generation capabilities, providing valuable pre-trained features for various downstream NLP tasks.

Language modelling (LM) loss: The goal of LM is to predict the probability distribution of the next word or character given the context. By maximizing the probability of correctly predicting the next word, LM enables the model to learn the probability distribution and structure of language, thereby enhancing its understanding and generation capabilities. This provides the model with fundamental abilities for modelling and generating language sequences

# 3. Technical path for adapting LLM to vertical industries

The technical pathways for adapting AI large models to vertical industries include data collection and preprocessing, industry-specific knowledge construction, model selection and fine-tuning (industry-specific model development), intelligent task decomposition and execution for complex tasks, and integration and deployment—such as AutoGPT. These technical pathways and methods provide guidance and reference for the application of AI large models in vertical industries [10].

Data Collection and Preprocessing: For specific vertical industries, collect large-scale relevant data based on general AI large models, including text, images, audio, etc. Data quality and suitability are crucial for model performance, necessitating preprocessing steps such as data cleaning, denoising, and annotation.

Vertical Industry Knowledge Construction: Vertical industries often have unique domain knowledge, including specialized terminology, rules, and constraints. Integrating domain knowledge into AI large models can enhance their effectiveness in the industry. Common methods include manual data annotation, knowledge graph construction, and expert knowledge extraction.

Model Selection and Fine-Tuning: Start with a pre-trained model and fine-tune it using industry-specific data to adapt it to particular tasks and scenarios. Fine-tuning involves adjusting hyperparameters, training set sampling strategies, and optimization algorithms. This process benefits from reinforcement learning based on human feedback and multimodal knowledge graph techniques, leveraging domain experts' knowledge to ensure the model meets industry needs.

Intelligent Task Decomposition and Execution: Vertical industries have specific tasks and requirements that may require customizing AI large models. For example, adding specific category labels for text classification tasks or additional convolutional layers or classifiers for image processing tasks. This customization helps the model better meet the specific requirements of vertical industry tasks.

Integration and Deployment - AutoGPT: After training and customizing the model, integrate and deploy it within the vertical industry's systems. This includes developing and debugging interfaces to ensure stable operation in practical applications while considering the model's performance and efficiency to meet industry needs.

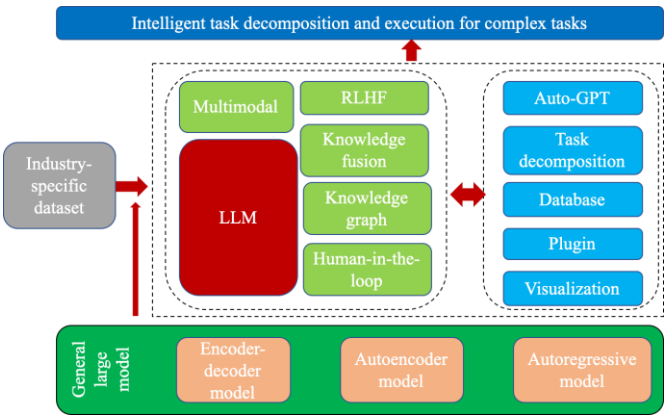


**Figure 3.** Technical pathways for adapting LLM to vertical industries

# 4. Applications of LLM in power system

## 4.1. Equipment fault diagnosis

Current power generation equipment fault records are mostly logged by on-site personnel, leading to issues with irregular data formats, lengthy text, excessive details, and difficulty in quickly obtaining solutions. By integrating the general semantic understanding capabilities of multimodal large models with historical fault data and operational procedures, we can develop an industry-specific multimodal large model for the power generation sector. This model would assist business personnel in quickly diagnosing equipment issues and recommending corrective actions, thereby reducing unplanned downtime.

Specifically, we will first address the issues present in current power generation operations, such as data redundancy, significant data gaps, diverse sources, and complex structures, by organizing and constructing a fault diagnosis and decision support corpus. This involves researching unified information representation techniques for multi-source heterogeneous data to achieve standardized management of equipment data. Secondly, in the context of adapting large models with billions of parameters for industry-specific tasks, traditional full-parameter fine-tuning methods present challenges such as high training costs and extensive computational time. Therefore, we will develop lightweight fine-tuning methods based on the organized multimodal power generation corpus to minimize the number of fine-tuning parameters and computational complexity, thereby enhancing the performance of downstream diagnostic decision tasks. Furthermore, to effectively integrate the high-precision simulations of existing mechanistic models with the rapid generalization and inference capabilities of large models, we will investigate data-mechanism fusion-driven fault diagnosis algorithms that combine mechanistic model information. This approach aims to achieve high-precision and rapid diagnosis of equipment faults. Finally, to improve the large model's semantic understanding and response accuracy for complex business problems, we will explore multimodal human-machine interaction interface design methods for various business scenarios. This will facilitate rapid evaluation of the content generated by large models by business experts. Based on interaction records, we will investigate reinforcement learning techniques informed by expert feedback to continuously enhance the model's ability to analyse and handle complex tasks, ensuring the controllability of

decisions. The specific technical plan is illustrated in Figure 4.

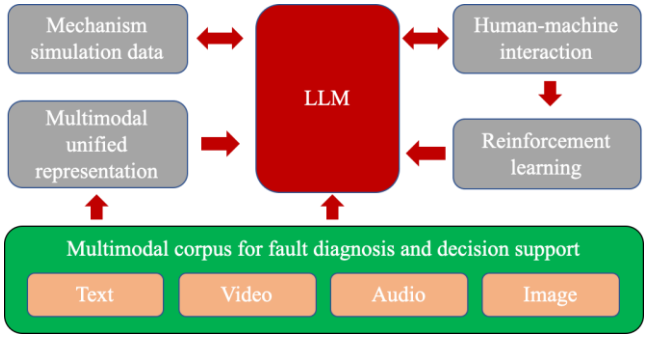


**Figure 4.** Application of LLM in fault diagnosis of equipment

## 4.2. Transmission network operation control

The operation of large power grids is complex, with control and regulation involving numerous specialized business areas and substantial electrical computations. This results in high demands for the timeliness and effectiveness of strategy formulation. In the context of new power systems, the challenges faced by main grid dispatch and control are becoming increasingly severe. Meanwhile, the power grid, over long periods of operation, accumulates vast amounts of status data, procedural experience, and control records. Fully leveraging this multi-source heterogeneous data can greatly enhance the current control mode, which relies on dispatcher experience.

As illustrated in Figure 5, for the dispatch and control of the main grid, a general large model can be trained using multi-source heterogeneous data, such as grid control procedures, contingency plans, historical operation data, and offline simulation data. This control model, acting as the control brain, can infer specific control tasks and objectives based on the grid's operational status, combined with dispatcher judgments and requirements. It matches with the control knowledge base to provide corresponding specific strategy recommendation. By integrating and coordinating various types of control business knowledge, the control model clarifies control tasks with its powerful analytical capabilities. It interacts at a high level with dispatchers for new power system control at the task level, while maintaining compatibility with existing dispatch system functions, thereby improving control efficiency.

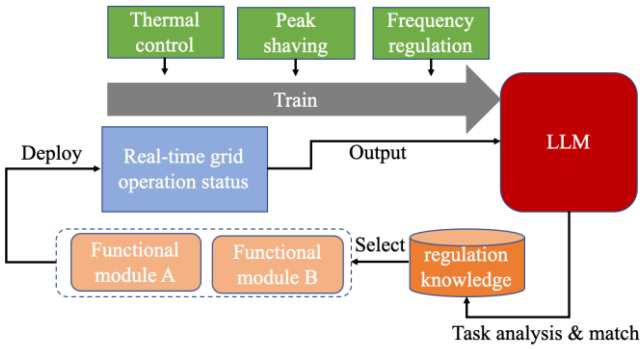


**Figure 5.** Framework of the main network control LLM

## 4.3. Distribution network operation control

Compared to transmission networks, distribution networks have a more flexible structure but lower reliability, making their control tasks more complex. Traditional control issues in distribution networks mainly rely on methods such as model predictive control and optimization theory. With the accelerated construction of new power systems and the large-scale integration of high proportions of renewable energy into distribution networks, innovations in renewable energy generation, grid integration, energy storage, and electric vehicle technologies have been driven forward. At the same time, the inherent volatility, intermittency, and uncertainty of renewable energy present new challenges to the stability and power quality of distribution networks. In the context of high proportions of renewable energy integration, new issues have emerged in distribution networks, such as reverse power flow, insufficient regulation capacity, and deterioration of power quality.

As shown in Figure 6, first, a distribution network control knowledge base is constructed using prior knowledge such as distribution network control procedures, contingency plans, and historical incident records, based on technologies like knowledge graphs. Distribution networks, being directly user-facing, exhibit typical characteristics of massive, multi-source, and heterogeneous data. If all control-related information is built into a knowledge graph, it might lead to challenges such as high construction difficulty, large storage requirements, and limited retrieval speed. Vector databases, through machine learning algorithms, convert unstructured data such as text and images into vector embeddings, storing and managing related data while enabling fast information retrieval based on similarity, thus achieving efficient storage and utilization of massive data. Therefore, control data can be vectorized using natural language processing and an industry-specific vector database can be constructed. Both knowledge graphs and vector databases are methods for organizing and representing information and can be used as input for fine-tuning large models. Knowledge graphs offer natural human-machine interaction advantages in decision support through relational data visualization, while vector databases provide a universal data format for large models to understand business contexts through vector representation, acting as memory and storage modules that drive model training and inference. Ultimately, on the basis of a general-purpose large model, model fine-tuning with distribution network control data leads to the development of a distribution network control large model.

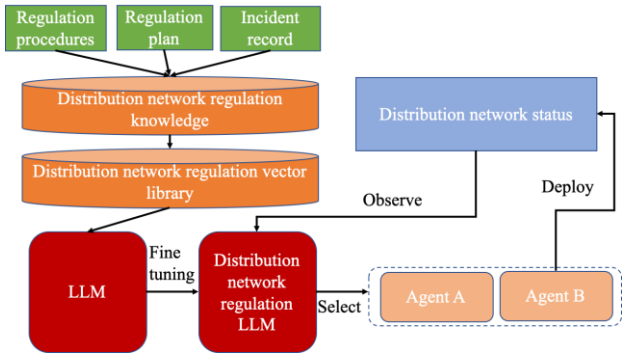


**Figure 6.** Distribution network regulation LLM

# 5. Conclusions

Artificial intelligence large models are disruptive and transformative technologies that have recently emerged and will change the future development and application paradigm of AI technologies. Currently, AI large models are continuously being improved. First, large models will possess more comprehensive multimodal characteristics. In the

industrial field, the incorporation of multi-source heterogeneous big data will enrich multimodal features, making perception and expression pathways more diverse. Second, as large models integrate with intelligent hardware, such as the Internet of Things (IoT) and robotics, large-scale information infrastructure will achieve higher levels of abstract thinking and semantic processing, advancing traditional information infrastructure to a new stage of intelligence. Furthermore, the next step in the development of large models will include self-understanding of tasks, self-iteration, and self-improvement functions, gradually enhancing and accumulating intelligence beyond basic capabilities. These three aspects will guide the future development of large models and can also serve as important directions for the intelligent development of new power system operation and scheduling. The emergence of AI large models offers unprecedented opportunities for the intelligent control of power system operations, and it is hoped that this paper will contribute valuable insights into AI large model-based theories, methods, and engineering applications for grid operation control.

# References


[1] Liu, X., & Xu, W. (2010). Economic load dispatch constrained by wind power availability: A here-and-now approach. IEEE Transactions on sustainable energy, 1(1), 2-9.

[2] Alam, M. S., Al-Ismail, F. S., Salem, A., & Abido, M. A. (2020). High-level penetration of renewable energy sources into grid utility: Challenges and solutions. IEEE access, 8, 190277-190299.

[3] Lu, J., Batra, D., Parikh, D., & Lee, S. (2019). Vilbert: Pretraining task-agnostic visiolinguistic representations for vision-and-language tasks. Advances in neural information processing systems, 32.

[4] Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt. arXiv preprint arXiv:2303.04226.

[5] Driess, D., Xia, F., Sajjadi, M. S., Lynch, C., Chowdhery, A., Ichter, B., ... & Florence, P. (2023). Palm-e: An embodied multimodal language model. arXiv preprint arXiv:2303.03378.

[6] Yue, X., Ni, Y., Zhang, K., Zheng, T., Liu, R., Zhang, G., ... & Chen, W. (2024). Mmmu: A massive multi-discipline multimodal understanding and reasoning benchmark for expert agi. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9556-9567).

[7] Yang, Z., Li, L., Lin, K., Wang, J., Lin, C. C., Liu, Z., & Wang, L. (2023). The dawn of lmms: Preliminary explorations with gpt-4v (ision). arXiv preprint arXiv:2309.17421, 9(1), 1.

[8] Xue, L., Gao, M., Xing, C., Martín-Martín, R., Wu, J., Xiong, C., ... & Savarese, S. (2023). Ulip: Learning a unified representation of language, images, and point clouds for 3d understanding. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 1179-1189).

[9] Yu, W., Yang, Z., Li, L., Wang, J., Lin, K., Liu, Z., ... & Wang, L. (2023). Mm-vet: Evaluating large multimodal models for integrated capabilities. arXiv preprint arXiv:2308.02490.

[10] Shanahan, M. (2024). Talking about large language models. Communications of the ACM, 67(2), 68-79.