

# 大语言模型在电力系统中的应用初探

牛泽原, 李嘉媚, 艾芊

(上海交通大学电子信息与电气工程学院, 上海市 闵行区 200240)

## Preliminary Exploration of the Application of Large Language Models in Power Systems

NIU Zeyuan, LI Jiamei, AI Qian

(School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University,  
Minhang District, Shanghai 200240, China)

**ABSTRACT:** Implementing large language models in power systems is becoming increasingly prevalent with pervasive development. This paper presents a preliminary investigation into the applications of large language models in power systems, analyzing their technical advantages and practical value in areas such as prompt engineering, multimodal processing, and model fine-tuning. Through experiments, the study validates the potential and capabilities of large language models in supporting auxiliary operations within power systems and examines their performance across various tasks. The results demonstrate that the large language model can address power system issues such as power prediction, optimal power flow calculation, and specialized knowledge question and answer through the designed prompt engineering framework and model fine-tuning. This provides electrical engineers with an initial reference for utilizing large language models to solve power system problems while laying a solid foundation for broader future applications.

**KEY WORDS:** large language model; power system; prompt engineering; fine-tuning; power prediction; optimal power flow

**摘要:** 大语言模型在电力系统中的应用研究正在广泛深入地开展。该文初步探讨了大语言模型在电力系统中的应用, 分析了其在提示工程、多模态处理、模型微调等方面的技术优势和实践价值。通过实验验证了大语言模型在电力系统辅助运行中的潜力和能力, 并分析了其在不同任务中的表现。研究结果表明, 大语言模型结合提示词框架设计和模型微调, 可以解决电力系统中如功率预测、最优潮流计算和专业知识问答等问题, 给为电气工程师利用大语言模型解决实际问题提供了参考, 也为未来更广泛的应用探索提供了良好基础。

**关键词:** 大语言模型; 电力系统; 提示词工程; 微调; 功率预测; 最优潮流

DOI: 10.13335/j.1000-3673.pst.2024.1483

## 0 引言

随着人工智能技术的快速发展, 大语言模型 (large language model, LLM) 的应用正在广泛深入各个领域。当下电力系统面临着多种挑战, 包括非线性非凸问题的求解<sup>[1]</sup>、海量数据的处理<sup>[2]</sup>、实时监控与故障诊断<sup>[3]</sup>、电力标准数字化<sup>[4]</sup>等, 这些问题使得寻找一种在解决电力系统问题中具备实时处理海量数据能力并且安全高效的工具的需求与日俱增。而现阶段较为成熟的 LLM 及其相关技术, 为解决这一需求提供了新的思路<sup>[5]</sup>。

在有关 LLM 在电力系统中的应用方面, 文献[6]介绍了一个模块化框架, 整合了电力系统与大型语言模型领域的专业知识, 提高了 GPT-4o 在未曾接触过的电力系统仿真工具上的仿真编码准确率, 并验证了其作为科研助手的潜力。文献[7]探讨了如何利用 LLM 帮助电力工程师处理常规和创新的编程任务, 提出了两种框架: 一种是用于常规任务的端到端框架, 适用于已知问题, 如单元组合问题 (unit commitment, UC); 另一种是人机互动框架, 用于解决复杂的问题, 如带有不确定性的调度问题, 文中同时评估了常用 LLM 在编程任务中的成功率、一致性和稳健性, 指出了 LLM 在代码自动生成和错误校正方面的潜力。文献[8]使用 ChatGPT 生成适用于潮流研究的配电系统, 验证了 GPT-4 在提供功能性电路和建议工程解决方案方面的优越性能, 但也指出了其在复杂逻辑问题上的局限性。文献[9]分析了 LLM 在电力系统负荷与新能源发电出力预测、电力系统规划、电力系统运行、电力系统故障诊断与系统恢复、电力市场等领域的潜在应用空间, 同时探讨了电力系统数据的质量与可获取性、输出结果可解释性以及隐私保护等问题。

**基金项目:** 上海市“科技创新行动计划”软科学研究青年项目 (24692105200)。

Project Supported by Shanghai Soft Science Research Youth Project of Science and Technology Innovation Action Plan (No. 24692105200).

在有关 LLM 应用优势和潜力方面,文献[10]作者提出了一种名为 LLM 驱动的进化算法 LMEA (local movement evolutionary algorithm),通过 LLM 在组合优化问题中的应用,减少了对领域知识和人工设计的依赖。文献[11]探讨了 GPT-4V-Vision 在处理多种输入模式、执行复杂任务以及人机交互方面的独特能力,展示了其在开放世界视觉理解、视觉描述、多模态知识、常识、场景文本理解、文档推理、编程、时间推理、抽象推理和情感理解等多个领域的表现。文献[12]提出了多模态基础模型 (GridFM)的概念,用于解决电网面临的复杂性和不确定性问题。

在 LLM 的提示词工程和模型微调方面,文献[13]引入了一个涵盖数据构建、模型训练和评估的通用工具使用框架 ToolLLM,经过微调的 LLaMA 该框架展示了执行复杂指令和泛化到未见过的 API 的显著能力,并表现出与 ChatGPT 相当的性能。Jules White 等人在文献[14]中提出了一种提示词分类框架。文献[15]和[16]提出了关于如何设计和优化输入提示以获取期望的 AI 响应的全面指南,涵盖了基础知识、技术、最佳实践、高级策略以及实际应用案例,旨在帮助用户最大化利用 ChatGPT 的潜力。文献[17]利用微调 LLM 来减少负荷曲线分析中的数据需求,通过一个两阶段的微调策略,展示了如何使用 GPT-3.5 模型恢复电力系统负荷曲线中的缺失数据,并通过实证评估证明该微调模型在准确恢复缺失数据方面的有效性和可比性。

以上研究表明,LLM 作为新的电力系统分析工具,在电力系统中具有良好的应用潜力。基于 LLM 解决电力系统典型业务需求,可以充分利用 LLM 的优势和红利,减少相关人员的冗余工作,提升电力系统智能化水平。基于上述研究和分析,本文将初步探讨 LLM 在电力系统中的应用,选取功率预测、最优潮流计算和专业知识问答三类典型电力系统业务需求进行研究,分析 LLM 提示词工程、多模态处理、模型微调等方面的技术实践和应用价值。实例分析表明,LLM 在经过本文提出的针对电力系统的提示词工程或微调后,拥有正确、高效解决电力系统中特定问题的能力。

## 1 大语言模型的优势

LLM 作为电力系统中可以利用的工具,其主要优势体现在语义理解和学习能力、调用外部工具的扩展能力以及与多模态处理系统的集成能力。这些优势赋予了 LLM 处理电力系统中问题的能力。

### 1.1 语言模型：记忆与学习

作为 LLM,其显著的优势之一就是其对于文本的理解能力和具有记忆<sup>[18]</sup>、学习<sup>[19]</sup>的对话能力。LLM 能够基于输入(Prompts)给出语义上合理的输出(Responses)。尽管预训练数据集包含的内容尚未明确,但目前的 LLM 有能力为电气工程领域的特定问题提供逻辑的回答<sup>[20]</sup>。该能力的主要组成部分可能来自于大量带有记忆的模型参数的自然结果,转换器架构中的高效处理功能允许模型高效检索这些已记忆的信息。

在对话型 LLM 中,如 ChatGPT,用户在一次对话中的上下文中提到的信息同样会被模型学习、记忆,以更好地对用户的输入做出反应。

### 1.2 工具插入：功能扩展

LLM 解决电力系统中的问题时,调用外部工具,如求解器等,则可以大幅度提高其回答的科学性<sup>[21]</sup>。在现有的 LLM 中,如 ChatGPT4,已经具有嵌入式的工具,如 Python 执行器、PSS/E、PSCAD、PowerWorld 和 CyME 等,这些工具可以由 LLM 调用来解决复杂问题。

### 1.3 信息输入：多模态集成

现阶段,通过集成多模态模型,对话型 LLM 能够接受的信息已经不止于文本,包括但不限于以下代码文件输入: .jpg、.png 图片信息输入, .csv、.xlsx 表格数据输入以及.py<sup>[22]</sup>。在电力系统中,数据类型可以包括传感器数据、监控图像、设备故障日志、潮流分析的数学模型代码等。LLM 能够通过语义嵌入<sup>[23]</sup>将这些不同类型的数据整合在一起,使其能够理解电力系统中来自不同来源的信息,从而做出情境化处理。

考虑到电力系统中的数据隐私性<sup>[24]</sup>,本文从“数据”和“数据描述”角度,论述了如何在使用过程中最大程度上保护数据隐私和安全。

### 1.4 本文研究框架

基于上述特性,本文的研究从探究如何构建高效的 LLM 提示词框架、如何通过微调以构建专业知识程度较高的大模型的研究方向出发,提出了 3 个任务进行实例研究分析:

- 1) Task1, 电力系统功率预测。
- 2) Task2, 电力系统最优潮流计算。
- 3) Task3, 电力系统专业知识回答。

本文的整体研究框架如图 1 所示。其中 Task1 和 Task2 的研究框架如图 1(a)所示,主要基于现有的 LLM 开展提示词工程测试研究,目标是探索通过 LLM 解决电力系统中常见问题的方法,例如功

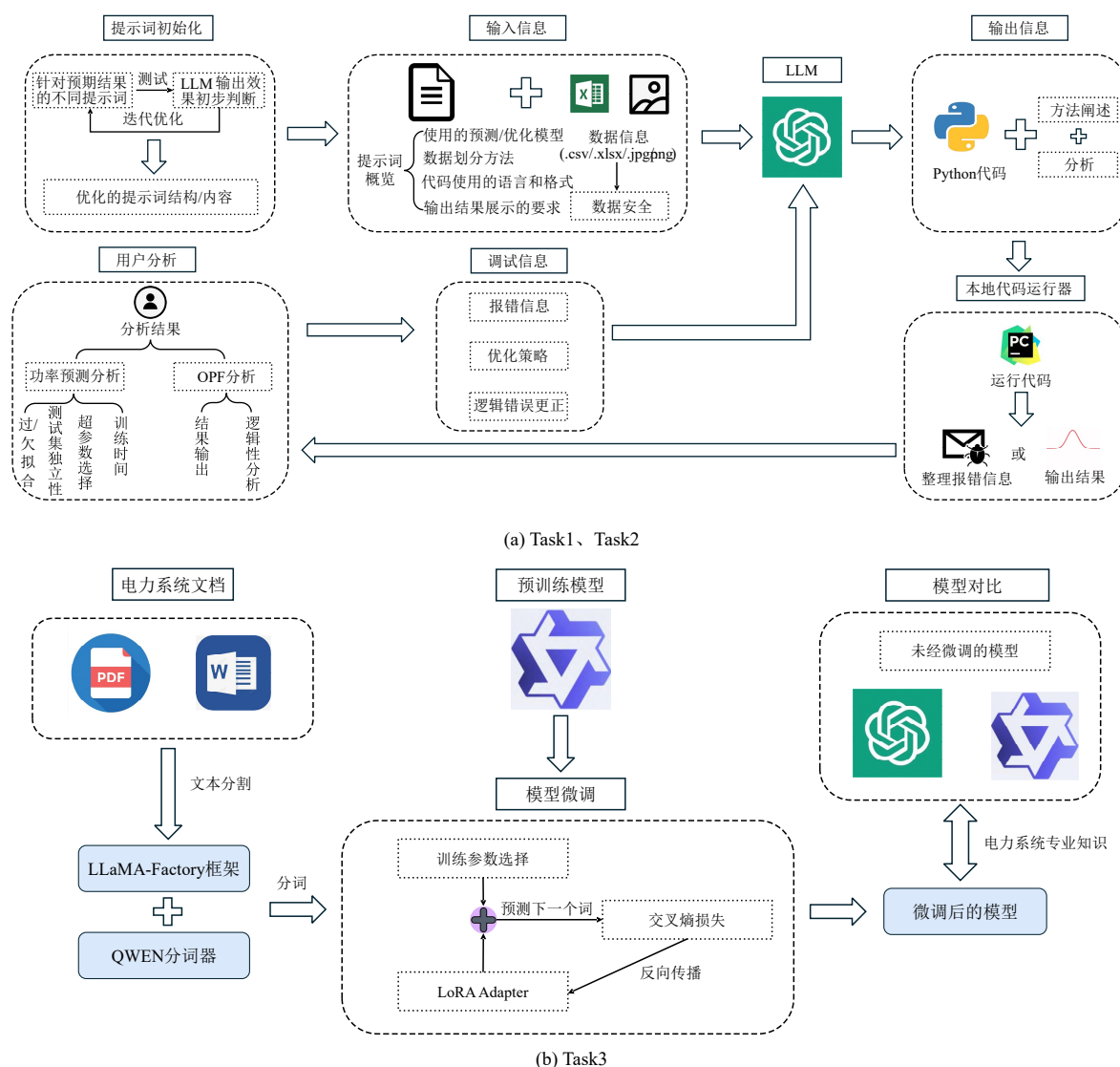


图1 本文研究框架  
Fig. 1 Research structure

率预测和最优潮流(optimal power flow, OPF)计算问题。Task3 的研究框架如图 1(b)所示,旨在对 LLM 进行微调,进而构建拥有专业知识图谱的电力系统行业大模型。

在 Task1 和 Task2 中,首先通过提示词优化,利用 LLM 生成优化的信息结构内容。输入信息包括各种格式的数据文件(如.csv、.xlsx、.jpg、.png),这些数据经过预处理和优化后,输入到 LLM 中。LLM 输出信息被传递到 Python 代码中,用于方法验证和分析,最终输出结果可以是运行代码、生成报告或图表等形式。此外,用户分析结果可以用于进一步的需求信息调整,包括聚焦信息、优化策略和逻辑错误修正。

在 Task3 中,主要关注模型微调过程。本文使用 LLaMA-Factory 框架<sup>[25]</sup>进行模型微调,结合 Qwen 模型自带的分词器 QWEN Tokenizer<sup>[26]</sup>来进行文本预处理。模型微调基于 LoRA Adapter 进行,

通过训练参数迭代并利用交叉熵损失进行反向传播来优化模型。最后,微调后的模型与未微调的基础模型进行比较,确保电力系统专业知识的准确应用。

## 2 提示词工程与模型微调

提示词工程是通过精心设计和调整输入提示词来优化 LLM 的输出,使其更符合用户的需求和期望<sup>[27]</sup>。大模型微调则是通过在特定任务或领域的数据上进一步训练预先训练好的 LLM,使其在该任务或领域上表现得更准确和有效<sup>[28]</sup>。两者都可以在不大幅度增加计算资源的情况下,提高模型的应用效果。

### 2.1 提示词工程

在本文中使用的提示词模式框架如表 1 所示。每个类别对应不同的提示词内容。输入语义理解类别解释了 LLM 如何理解用户输入。输出定制

表 1 提示词模式框架  
Table 1 Prompts pattern framework

模式	提示词内容
输入语义理解	LLM 如何理解用户输入
输出定制	约束 LLM 生成的输出的格式
错误识别	LLM 检查输出中的错误
提示改进	LLM 提高输出的质量
上下文控制	指定 LLM 关注的上下文信息

类别则关注约束和定制 LLM 生成的输出格式。错误识别类别专注于检查 LLM 输出中的错误。提示改进类别旨在提高 LLM 输出的质量。最后，上下文控制类别则用于指定 LLM 关注的上下文信息。该框架帮助用户根据不同需求优化提示词，从而提升 LLM 的使用效果。

2.2 模型微调

基于适配器(Adapter)的模型微调<sup>[29]</sup>的优势在于：不必对 LLM 本身的参数进行训练，而是只训练适配器层，进而能够降低训练成本。其实现形式如图 2 所示。

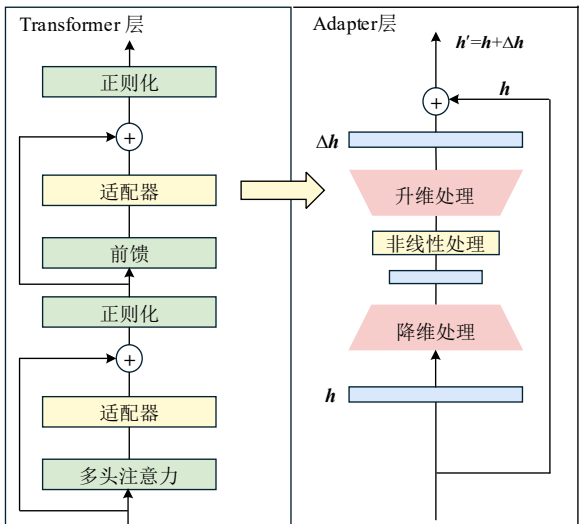


图 2 基于适配器的模型微调实现

Fig. 2 Adapter-based model fine-tuning implementation

图 2 中左侧是 Transformer 层的架构<sup>[30]</sup>，包括正则化、多头注意力机制、前馈网络和适配器。其中正则化用于防止模型过拟合，通过加入额外约束提升模型的泛化能力；多头注意力机制允许模型在不同的子空间中关注输入序列的不同部分，从而捕捉更多细节和关联；前馈网络则是对每个位置的表示进行非线性变换，进一步提升表示能力；适配器在不改变模型主要结构的情况下，添加额外层次进行特征调整和微调，以便适应特定任务需求。右侧是 Adapter 层的详细处理过程。Adapter 层在 Transformer 层的基础上，增加了升维处理、非线性处理和降维处理的步骤。

$h$  表示 Transformer 层的输入向量，来自上一层

的输出或模型的初始输入； $h'$  表示通过适配器层(Adapter 层)后的输出向量； $\Delta h$  表示适配器层计算得到的增量向量。它是通过适配器层中的升维处理(up-projection)、非线性处理和降维处理(down-projection)生成的向量，用于对原始输入  $h$  进行调整或增强。

本文使用 LoRA 方法<sup>[31]</sup>对经过预训练的大模型进行微调。LoRA 可以看作是一种特殊的 Adapter，其核心思想为通过低秩分解的方式减少参数量，引入的参数变化隐含在分解矩阵中，如图 3 所示。

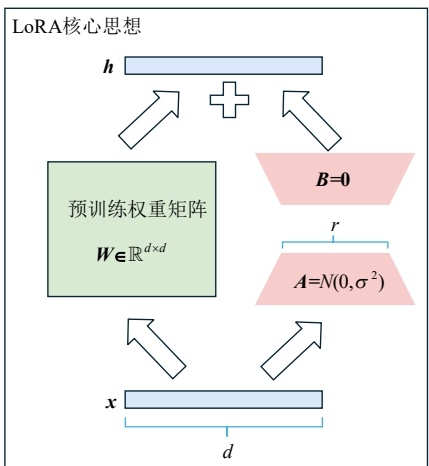


图 3 LoRA 微调的核心思想

Fig. 3 Core idea of fine-tuning based on LoRA

其目标函数如下所示：

$$\max_{\Theta} \sum_{(x,y) \in Z} \sum_{t=1}^{|y|} \lg[P_{\Phi_0 + \Delta\Phi(\Theta)}(y_t | x, y_{<t})] \quad (1)$$

式中： $\Theta$  为用于编码任务特定参数增量  $\Delta\Phi$  的小尺寸参数集合，优化目标是找到最佳的  $\Theta$ ； $Z$  为训练使用的数据集； $x$  为输入序列； $y$  为目标(输出)序列； $\Phi_0$  为预训练模型的初始权重。

传播矩阵的定义如下所示：

$$h = W_0 x + \Delta W x = W_0 x + B A x \quad (2)$$

式中： $W_0$  为预训练模型中的初始权重矩阵，维度为  $\mathbb{R}^{d \times k}$ ； $\Delta W$  为权重更新矩阵； $B$  为低秩分解中的矩阵，维度为  $\mathbb{R}^{d \times r}$ ， $r \ll d, k$ ； $A$  也为低秩分解中的矩阵，维度为  $\mathbb{R}^{r \times k}$ ； $A$  和  $B$  都是可训练的参数。

3 实例分析

为了客观地评价不同提示词框架和不同模型在处理任务时的效果，本文构建了一种多角度的综合性提示词初步评价体系，如表 2 所示。评价体系包含 4 个主要指标：生成代码的可行性、代码完成目标任务的情况、代码运行速度和结果中是否存在逻辑性错误。

最后根据加权计算总得分并给出初步评价：



表 2 提示词初步评价体系  
Table 2 Preliminary evaluation system for prompts

指标	评价	评分	综合权重/%	
			Task1	Task2
生成代码的可行性	调试次数≤1	100	20	30
	调试次数≤3	80		
	4≤调试次数≤10	60		
	调试次数>10 或失败	0		
代码完成目标任务的情况	完成度 100%	100	40	40
	80%≤完成度≤100%	80		
	60%≤完成度≤80%	60		
	完成度<60%	0		
代码运行速度	80%~100%	100	20	0
	60%~80%	80		
	<60%	60		
结果中是否存在逻辑性错误	否	100	20	30
	是	0		

表 3 不同提示词框架分析(Task1)  
Table 3 Analysis of different prompt frames (Taks1)

测试序号	[数据文件]	数据描述	本地运行	LSTM	Pytorch	数据划分	[LSTM 模型表述]	初步评价		
								ChatGPT-4o	ChatGPT-4	ChatGPT-3.5
1	1	0	0	0	0	0	0	D	D	D
2	1	0	1	0	0	0	0	D	D	D
3	1	0	1	1	0	0	0	C	C	D
4	1	0	1	1	1	0	0	B	C	C
5	1	0	1	1	1	1	0	A	B	C
6	1	0	1	1	1	1	1	B	A	C
7	0	1	1	1	1	1	1	B	A	C

注：在 ChatGPT-3.5 中，[数据文件]和[LSTM 模型表述]均以文本的形式输入。

每个测试序号下的每个模型进行独立的 10 次实验，以避免偶然性带来的错误评价。在保持其他提示词不变的前提下，对其中“LSTM”和“Pytorch”进行了额外的分析。表 4 展示了基于 ChatGPT-4 模型的测试结果。文章使用的 LSTM 和 GRU 模型参考文献[32]。

表 4 特别提示词分析  
Table 4 Analysis of special prompts

测试序号	LSTM		Pytorch		初步评价
	LSTM	GRU	Pytorch	TensorFlow	
1	0	1	0	1	D
2	1	0	0	1	C
3	0	1	1	0	B
4	1	0	1	0	A

基于表 3 和表 4 得到的研究结果，本文在算例中使用的提示词框架确定为：[ChatGPT-4o: [数据文件]，“可视化”，“本地运行”，“LSTM”，“Pytorch”，“数据划分”]，[ChatGPT-4: “数据描述”，“可视化”，“本地运行”，“LSTM”，“Pytorch”，“数据划分”，[LSTM 模型表述]]。由于 ChatGPT-3.5 的表现不能达到完成任务的目标，因此放弃使用 ChatGPT-3.5。

图 4 展示了 ChatGPT-4 基于优化提示词的测试

0~60 分为 D，61~80 分为 C，81~90 分为 B，91~100 分为 A。

### 3.1 Task1：功率预测

在电力系统运行中，准确预测未来一段时间内的功率需求对电力调度至关重要。本任务的目标是使用 LLM 结合历史功率负荷数据和时间等多维度输入特征，构建并优化深度学习模型进行短期功率预测。为保护用户隐私，本任务在提示词工程中通过定义数据结构，使用户可在不上传具体数据的前提下实现功率预测。

在提示词初始化的测试过程中，各提示词对输出结果的影响分析如表 3 所示。0 代表在提示词框架中没有包含该提示词，1 代表包含了该提示词。

Input: [LSTM.pdf]

User: 1) 我在本地存储了名为 Power.csv 文件，其中数据的行列数分别为[1, 4000]。请你将数据集 Power.csv 中的数据按照6:2:2的比例划分为训练集、验证集、测试集。  
2) 使用LSTM的方法进行预测，LSTM 的数学模型如文件 LSTM.pdf 所示，并在测试集上检验模型的预测效果。  
3) 使用Pycharm的代码格式，调用pytorch包，每个epoch输出训练MAE和RMSE，最后在测试集上可视化预测结果与实际数据的对比，并输出代码的总运行时间。

ChatGPT : To get started with your LSTM prediction task using the Power.csv dataset, we'll first load the data, then split it into training, validation, and test sets with a 60:20:20 ratio. After that, we'll set up the LSTM model according to the formulas you provided...;

Here is the complete PyTorch code to accomplish this:

```
import pandas as pd
import numpy as np
import torch...;
```

This script covers data loading, preprocessing, LSTM model setup, training with validation, and final visualization of results...

图 4 使用 ChatGPT-4 时的提示词和输出结果  
Fig. 4 Prompts and results using ChatGPT-4

记录。用户提供了具体要求，包括数据集的划分比例、使用 LSTM 模型进行预测以及在 PyCharm IDE 中调用 PyTorch 包进行代码的实现，并要求输出每个 epoch 的平均绝对误差(mean absolute error, MAE)和均方根误差(root mean squared error, RMSE)以及总运行时间。ChatGPT-4 响应了这些要求，提供了详细的代码步骤，从数据加载、预处理、模型设置、训练到最终结果的可视化。附录 A 图 A1 和附录 A 图 A2 分别展示了 ChatGPT-4 在优化过程中和报错反馈的过程中的问答流程。

在优化过程中，用户要求优化隐藏层数量、历史输入/预测输出长度、学习率、训练周期和批量大小，以提高预测效果和代码运行速度。ChatGPT-4 响应了这些需求，提供了优化后的参数设置，并建议用户应用这些更改。

在报错反馈的流程中，ChatGPT-4 根据报错信息识别出输入数据的维度问题，并建议在 DataLoader 中设置正确的 batch\_size，同时调整数据的维度以符合模型的期望。ChatGPT-4 提供了修改后的代码，确保正确处理数据维度问题，并请求用户尝试运行新版本的代码。

图 5 为对应输出的代码的运行结果，即预测结果的可视化。可以看出，在较为简单的时序预测任务中，模型表现效果良好。

表 5 中展示了模型在解决功率预测问题时的模型测试数据。各个模型的调试次数上限设置为 10 次，对于-refine 测试而言，调试次数代表优化迭代次数(取最优结果对应的调试次数)。对于非-refine

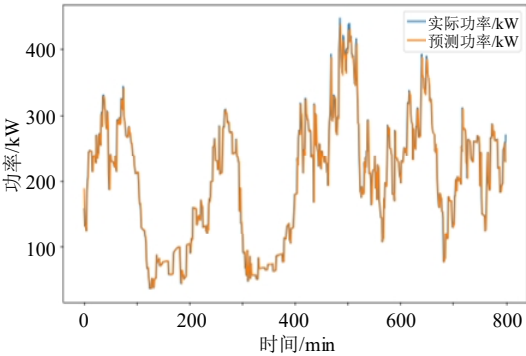


图 5 ChatGPT-4 生成代码本地运行结果  
Fig. 5 Results of the code generated by ChatGPT-4

测试而言，调试次数代表模型正确 Debug 时经过的询问次数。在“功率预测”的任务中，ChatGPT-4 的表现总体来讲优于 ChatGPT-4o。ChatGPT-4o 生成的代码可行性较好，通过较少次数的调试即可满足要求，但整体的预测效果和代码的优化能力不如 ChatGPT-4。这是因为在 ChatGPT-4 的输入中包含了 LSTM 的数学模型，因此生成的代码在处理效率上表现得更加优异。而在提示词预测试中，向 ChatGPT-4o 输入 LSTM 的数学模型却导致了预测效果的下降，因此，在该特定任务中，ChatGPT-4 的表现优于 ChatGPT-4o。

3.2 Task2: 电力系统 OPF 计算

最优潮流计算是电力系统优化的重要组成部分，旨在通过优化发电机功率输出和潮流分配，以最小化运行成本，同时满足电力系统的物理和安全约束。本任务的目标是利用 LLM，生成最优潮流计算的 Python 代码，并调用求解器进行求解。

表 5 各模型效果对比(Task1) Table 5 Comparison of each model(Task1)									
LLM 模型	调试次数	MAE	RMSE	运行时间/s	由 ChatGPT 生成的神经网络参数				
					隐藏层数量	历史数据输入长度	学习率	训练周期	批量大小
ChatGPT-4o	0	13.51	22.21	35.03	64	10	0.01	100	32
ChatGPT-4o-refine	3	13.29	21.75	22.42	128	20	0.005	50	64
ChatGPT-4	1	12.75	21.40	16.25	50	1	0.001	50	64
ChatGPT-4-refine	5	11.98	21.03	9.64	100	1	0.005	100	128

数据来源：基于文献[33]中的改进 IEEE RTS 24 节点系统。为了简化潮流计算，对数据进行了筛选，筛选后的数据中包含：火发电机组的技术参数，火电机组的成本参数，线路的起始点、终止点、线路电阻及损耗成本系数和潮流限制以及各节点的负荷信息。本节中使用开源的求解器 IPOPT。提示词框架分析结果见表 6。根据表 6 所示，本文选择的提示词框架为[ChatGPT-4o/ChatGPT-4: [数据文件]，“数据描述”，“IPOPT”，[目标函数]]。

在本文中考虑的优化问题的表述如附录 A 中

式(A1)—(A8)所示。

附录 A 图 A3 展示了用户请求和 ChatGPT-4o 的响应过程，用户要求处理 IEEE 24 节点电力系统的数据，并使用 Ipopt 求解最优潮流问题。具体任务包括读取 Excel 文件中的电力系统数据、提取火电机组的最小和最大输出功率、成本参数(a, b, c)以及其他必要的数据的描述，并基于 PyCharm 格式编写代码，使用 Ipopt 求解器进行优化求解。ChatGPT-4o 通过 pandas 库加载 Excel 文件，提取并预览相关数据，确保准确性。随后，提供了在

PyCharm 中使用 Ipopt 求解器的基本代码, 包含必要的库导入和初始设置。

将 ChatGPT 生成的代码的运行结果与经过检验正确的基于 MATLAB+Gurobi 的参考最优潮流计算结果进行对比, 如表 7 所示。在最优潮流成本方面, ChatGPT-4o 生成的结果为 58630.09 美元, 与参考计算结果 58630.10 美元几乎一致。发电机的功率输出与参考计算结果几乎完全相符, 在功率分配方面, ChatGPT-4o 能够精确地模拟实际系统的运

行。然而, 在线路潮流计算中, 两种方法存在略微差异, 具体体现在某些线路的潮流值上。这些差异主要是由于使用不同的求解器导致的, 求解器在处理数值计算时可能会采用不同的算法或近似方法, 从而引起细微的数值差异。这些微小的差异在实际应用中通常是可以接受的, 不会对整体优化结果和系统运行产生显著影响。总体来看, ChatGPT-4o 在处理最优潮流计算问题时展示出了较高的精度和一致性。

表 6 不同提示词框架分析(Task2)  
Table 6 Analysis of different prompt frames (Task 2)

测试序号	[数据文件]	数据描述	IPOPT	[目标函数]	初步评价		
					ChatGPT-4o	ChatGPT-4	ChatGPT-3.5
1	1	0	0	0	D	D	D
2	1	0	1	0	C	C	C
3	1	0	1	1	B	B	C
4	0	1	1	1	A	B	B
5	1	1	1	1	A	A	B

注: 在 ChatGPT-3.5 中, [数据文件]和[目标函数]均以文本的形式输入。

表 7 最优潮流计算结果对比  
Table 7 Comparison of OPF calculation results

OPF 计算方式	最优潮流时的 成本/\$	发电机功率/MW				线路潮流/MW			
		G1	G11	G21	G31	L1	L11	L21	L31
ChatGPT-4o Generation	58630.09	16	48.33	155	155	27.48	21.71	-278.21	-190.1
OPF 计算参考	58630.1	16	48.33	155	155	27.46	21.71	-278.14	-190.11

3.3 Task3: 专业知识回答

电力系统中包含大量复杂的专业知识, 电力工程师在日常工作中需要快速、准确地获取相关领域的知识解答。本任务的目标是构建电力系统专业知识的问答, 并通过微调构建一个适用于电力系统领域的专业知识问答的 LLM。

本文使用 LLaMA-Factory 框架进行模型的微调, 在本地部署的大模型为“通义千问-1.8B-Chat”。

训练使用的数据集如图 6 所示。

每个数据点包括指令(Instruction)、输入(Input)和输出(Output)3 部分。指令部分要求模型解释特定问题的答案, 输入部分列出具体问题, 例如“虚拟电厂有哪些控制方式?”等, 输出部分则提供了详细的回答内容, 旨在帮助模型学习如何根据指令生成准确的回答。

微调过程中的损失表现如附录 A 图 A4 所示。

微调后的结果如图 7 所示。用于测试的问题是独立于训练集之外的测试集。对于“什么是虚拟电厂?”这个问题的回答。Qwen 简要地描述了虚拟电厂作为一种利用可再生能源替代传统电厂的系统。Fine-tuned Qwen 则提供了更详细的分类, 解释了虚拟电厂的定义和作用, 包括分布式能源的集合、对电网的综合管理和提高电力效率的方法。ChatGPT-4o

NO.1	Instruction: 请解释以下问题的答案。	Input: 虚拟电厂有哪些控制方式?	Output: 1) 集中控制方式: 虚拟电厂通过控制协调中心(Control Coordination Center, CCC)可以完全...
NO.2	Instruction: 请解释以下问题的答案。	Input: 虚拟电厂有哪些种类?	Output: 1) 商业型虚拟电厂 CVPP: CVPP从商业收益的角度出发, 不考虑配电网的影响...
NO.3	Instruction: 请解释以下问题的答案。	Input: 虚拟电厂是怎样运行的?	Output: 虚拟电厂的目标是将分散的分布式能源优化组合成有机整体, 使其作为一个特别的电厂参与电力市场...
NO.n	...		

图 6 微调使用的数据集示例

Fig. 6 Example of dataset used for fine-tuning

以更加全面的方式介绍了虚拟电厂, 通过信息通信技术整合各种能源资源, 形成类似于传统电厂的大型发电系统, 并强调了其智能调度和优化控制的优势。

结果显示经过微调后的“通义千问-1.8B-Chat”模型在专业知识回答能力方面不亚于 ChatGPT-4o。




User:  什么是虚拟电厂?		
<b>Qwen:</b> 虚拟电厂是一种利用可再生能源(如太阳能、风能、水能等)来发电的发电系统。它是一种利用电力系统中的虚拟能源来替代传统能源发电系统,以实现更高效的能源利用和减少碳排放。	<b>Fine-tuned Qwen:</b> 目前对于虚拟电厂并没有统一的定义,虚拟电厂主要有以下几类定义:1)虚拟电厂是一系列分布式能源的集合,以传统发电厂的角色参与电力系统的运行;2)虚拟电厂是对电网中各种能源进行综合管理的软件系统;3)虚拟电厂也包括能效电厂,通过减少终端用电设备和装置的用电需求的方式来产生“富余”的电能,即通过在用电需求方安装一些提高用电效能的设备,达到建设实际电厂的效果...	<b>ChatGPT-4o:</b> 虚拟电厂 (Virtual Power Plant, VPP)是通过信息通信技术将分布式能源资源(如太阳能发电、风能发电、小型水电、燃气轮机、储能系统等)整合在一起,形成一个类似于传统电厂的大型发电系统。VPP通过智能调度和优化控制,实现对这些分布式能源资源的协同管理和调度,从而提高能源利用效率和电网稳定性。虚拟电厂的主要特点和优势包括:分布式能源整合: VPP能够将分散的能源资源...

图7 微调后的“通义千问-1.8B-Chat”模型与其他模型对比

Fig. 7 Comparison of fine-tuned "Qwen-1.8B-Chat" model with other models

相比之下,通义千问具有本地部署和本地训练的优势,从而最大限度地确保了模型的可用性和数据隐私的安全性。此外,微调过程并未显著增加模型的规模和复杂性,微调后的通义千问模型能够以相对较小的模型规模高效地完成专业知识的问答任务。因此,大模型的微调可以在电力系统行业大模型的构建中发挥作用<sup>[34]</sup>。

#### 4 总结与展望

本文探讨了 LLM 在电力系统中的应用,分析了其在提示词工程、多模态处理和模型微调等方面的技术优势和实际应用价值。文章通过具体实例展示了 LLM 在电力系统运行中的应用能力和潜力,特别是在功率预测、最优潮流计算和电力系统专业知识问答等任务中。研究结果表明,LLM 通过设计的提示词框架和模型微调,可以解决电力系统中的问题,为解决电力系统复杂问题提供新的、基于 LLM 的解决思路。

后续的研究围绕“将模型从通用的 LLM 升级到行业专用的 LLM”开展,需要的技术工作为:“预训练+微调”的大模型学习框架[35]和检索增强生成(Retrieval-Augmented Generation, RAG)技术[36]的应用。本文的工作可以作为后续电力系统行业大模型研究的参考。

未来,随着 LLM 在电力系统中的应用不断深入,可以期待其在多个方面带来显著的改进和创新。首先,LLM 的提示工程和多模态处理能力将进一步优化电力系统的故障诊断和优化调度,提高系

统的运行效率和可靠性。通过不断优化提示词框架和模型参数,LLM 在功率预测、负荷管理和能源分配等任务中的表现将更加精准和高效。此外,LLM 与高级电力系统仿真工具的结合,如 PSS/E、PSCAD 等,将大幅提升电力系统仿真和分析的自动化水平,减少对人工干预的依赖。微调技术将进一步增强 LLM 在特定电力系统任务中的适应性和性能。通过对电力系统领域的专用数据进行微调,模型可以更加准确地理解和处理相关问题,从而提高预测和分析的精度。同时,RAG 技术的应用可以使 LLM 在处理复杂任务时,动态地检索外部知识库,从而提供更为准确和详尽的答案。随着 LLM 在隐私保护和数据安全方面技术的进步,其在处理电力系统敏感数据时的应用将更加广泛和安全。LLM 在处理多模态数据方面的能力将促进其在电力系统中更加复杂和多样化的应用,包括图像、文本和时间序列数据的综合分析。这些研究方向和技术创新将推动电力系统智能化和自动化的不断发展,为实现更加稳定、可靠和高效的电力系统运行提供坚实的技术支持。

附录见本刊网络版(<http://www.dwjs.com.cn/CN/1000-3673/current.shtml>)。

#### 参考文献

- [1] 黄铭浩,唐坤杰,董树锋,等.基于碳流理论考虑用户侧碳排放配额的最优潮流[J].电网技术,2023,47(7):2703-2712.  
HUANG Minghao, TANG Kunjie, DONG Shufeng, et al. Optimal power flow considering user-side carbon emission allowances based on carbon flow theory[J]. Power System Technology, 2023, 47(7):



- 2703-2712(in Chinese).
- [2] 康重庆, 杜尔顺, 郭鸿业, 等. 新型电力系统的六要素分析[J]. 电网技术, 2023, 47(5): 1741-1750.
- KANG Chongqing, DU Ershun, GUO Hongye, et al. Primary exploration of six essential factors in new power system[J]. Power System Technology, 2023, 47(5): 1741-1750(in Chinese).
- [3] 李旭斌, 田付强, 郭亦可. 新型电力系统中电力设备健康管理及智能运维关键技术探究[J]. 电网技术, 2023, 47(9): 3710-3726.
- LI Xubin, TIAN Fuqiang, GUO Yike. Key technologies for health management and intelligent operation and maintenance of power equipment in new power systems[J]. Power System Technology, 2023, 47(9): 3710-3726(in Chinese).
- [4] 马超, 宋琛. 电力标准数字化: 概念、核心挑战、治理路线图及发展趋势[J]. 电网技术, 2024, 48(2): 480-497.
- MA Chao, SONG Chen. Digital transformation of standards in electric power field: concept, core challenges, governance roadmap and development trend[J]. Power System Technology, 2024, 48(2): 480-497(in Chinese).
- [5] 蒲天骄, 赵琦, 王新迎. 电力人工智能技术研究框架、应用现状及展望[J/OL]. 电网技术, 2024: 1-21[2024-11-01]. <http://kns.cnki.net/kcms/detail/11.2410.tm.20241024.1500.002.html>.
- PU Tianjiao, ZHAO Qi, WANG Xinying. Technology framework, application status and prospects on electric power artificial intelligence[J/OL]. Power System Technology, 2024: 1-21[2024-11-01]. <http://kns.cnki.net/kcms/detail/11.2410.tm.20241024.1500.002.html>(in Chinese).
- [6] JIA Mengshuo, CUI Zeyu, GABRIELA H. Enabling large language models to perform power system simulations with previously unseen tools: a case of Daline[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2406.17215>.
- [7] LI Ran, PU Chuanqing, TAO Junyi, et al. LLM-based frameworks for power engineering from routine to novel tasks[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2305.11202>.
- [8] BONADIA R S, TRINDADE F C L, FREITAS W, et al. On the potential of ChatGPT to generate distribution systems for load flow studies using OpenDSS[J]. IEEE Transactions on Power Systems, 2023, 38(6): 5965-5968.
- [9] 赵俊华, 文福拴, 黄建伟, 等. 基于大语言模型的电力系统通用人工智能展望: 理论与应用[J]. 电力系统自动化, 2024, 48(6): 13-28.
- ZHAO Junhua, WEN Fushuan, HUANG Jianwei, et al. Prospect of artificial general intelligence for power systems based on large language model: theory and applications[J]. Automation of Electric Power Systems, 2024, 48(6): 13-28(in Chinese).
- [10] LIU Shengcai, CHEN Caishun, QU Xinghua, et al. Large language models as evolutionary optimizers[C]//Proceedings of 2024 IEEE Congress on Evolutionary Computation (CEC). Yokohama: IEEE, 2024: 1-8.
- [11] YANG Zhengyuan, LI Linjie, LIN K, et al. The dawn of LMMs: preliminary explorations with GPT-4V (ision) [EB/OL]. [2024-08-05]. <https://www.stableaiprompts.com/wp-content/uploads/2023/10/Chatgpt-Updates.pdf>.
- [12] HAMANN H F, BRUNSCHWILER T, GJORGIEV B, et al. Foundation models for the electric power grid[J]. Joule, 2024, 8(12): 3245-3258.
- [13] QIN Yujia, LIANG Shihao, YE Yining, et al. ToolLLM: facilitating large language models to master 16000+ real-world APIs[C]//Proceedings of the Twelfth International Conference on Learning Representations. Vienna: ICLR, 2024.
- [14] WHITE J, FU Quchen, HAYS S, et al. A prompt pattern catalog to enhance prompt engineering with ChatGPT[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2302.11382>.
- [15] EKIN S. Prompt engineering for ChatGPT: a quick guide to techniques, tips, and best practices[J/OL]. Authorea Preprints. [2024-08-05]. <https://www.techrxiv.org/doi/full/10.36227/techrxiv.22683919.v2>.
- [16] GIRAY L. Prompt engineering with ChatGPT: a guide for academic writers[J]. Annals of Biomedical Engineering, 2023, 51(12): 2629-2633.
- [17] HU Yi, KIM H, YE Kai, et al. Applying fine-tuned LLMs for reducing data needs in load profile analysis[J]. Applied Energy, 2025, 377: 124666.
- [18] KWON W, LI Zhuohan, ZHUANG Siyuan, et al. Efficient memory management for large language model serving with pagedattention [C]//Proceedings of the 29th Symposium on Operating Systems Principles. Koblenz: Association for Computing Machinery, 2023: 611-626.
- [19] AN Shengnan, MA Zexiong, LIN Zeqi, et al. Learning from mistakes makes LLM better reasoner[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2310.20689>.
- [20] MAJUMDER S, DONG Lin, DOUDI F, et al. Exploring the capabilities and limitations of large language models in the electric energy sector[J]. Joule, 2024, 8(6): 1544-1549.
- [21] LU Qiuyu, FANG Jiawei, YAO Zhihao, et al. Enabling generative design tools with LLM agents for mechanical computation devices: a case study[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2405.17837>.
- [22] HU Wenbo, XU Yifan, LI Yi, et al. BLIVA: a simple multimodal LLM for better handling of text-rich visual questions[C]//Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence. Vancouver: AAAI Press, 2024: 2256-2264.
- [23] LI Bohan, ZHOU Hao, HE Junxian, et al. On the sentence embeddings from pre-trained language models[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2011.05864>.
- [24] YAO Yifan, DUAN Jinhao, XU Kaidi, et al. A survey on large language model (LLM) security and privacy: the good, the bad, and the ugly[J]. High-Confidence Computing, 2024, 4(2): 100211.
- [25] ZHENG Yaowei, ZHANG Richong, ZHANG Junhao, et al. LlamaFactory: unified efficient fine-tuning of 100+ language models [C]//Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations). Bangkok: Association for Computational Linguistics, 2024: 400-410.
- [26] BAI Jinze, BAI Shuai, CHU Yunfei, et al. Qwen technical report [EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2309.16609>.
- [27] MARVIN G, HELLEN N, JJINGO D, et al. Prompt engineering in large language models[C]//Proceedings of the ICDICI: International Conference on Data Intelligence and Cognitive Informatics. Tirunelveli: Springer Nature Singapore, 2023: 387-402.
- [28] HU Zhiqiang, WANG Lei, LAN Yihuai, et al. LLM-adapters: an adapter family for parameter-efficient fine-tuning of large language models[C]//Proceedings of 2023 Conference on Empirical Methods in Natural Language Processing. Singapore: Association for Computational Linguistics, 2023: 5254-5276.

- [29] HOULSBY N, GIURGIU A, JASTRZEBSKI S, et al. Parameter-efficient transfer learning for NLP[C]//Proceedings of the 36th International Conference on Machine Learning. Long Beach: PMLR, 2019: 2790-2799.
- [30] HAN Kai, XIAO An, WU Enhua, et al. Transformer in transformer [C]//Proceedings of the 35th Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2021: 15908-15919.
- [31] HU E J, SHEN Yelong, WALLIS P, et al. Lora: low-rank adaptation of large language models[C]//Proceedings of the Tenth International Conference on Learning Representations. Virtual: ICLR, 2021.
- [32] FU Rui, ZHANG Zuo, LI Li. Using LSTM and GRU neural network methods for traffic flow prediction[C]//Proceedings of 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC). Wuhan: IEEE, 2016: 324-328.
- [33] CHEN Xianbang, YANG Yafei, LIU Yikui, et al. Feature-driven economic improvement for network-constrained unit commitment: a closed-loop predict-and-optimize framework[J]. IEEE Transactions on Power Systems, 2022, 37(4): 3104-3118.
- [34] WANG Chen, LIU Yanyi, GUO Tiezheng, et al. Systems engineering issues for industry applications of large language model[J]. Applied Soft Computing, 2024, 151: 111165.
- [35] HUANG Shaomang, PAN Jianfeng, ZHENG Hanzhong. CCoE: a compact LLM with collaboration of experts[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2407.11686>.
- [36] FAN Wenqi, DING Yujuan, NING Liangbo, et al. A survey on rag meeting LLMs: towards retrieval-augmented large language models [C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Barcelona: Association for Computing Machinery, 2024: 6491-6501.



牛泽原

在线出版日期: 2024-11-29。

收稿日期: 2024-08-07。

作者简介:

牛泽原(2001), 男, 硕士研究生, 研究方向为人工智能与电力系统、虚拟电厂, E-mail: nzy0720@sjtu.edu.cn;

李嘉媚(1996), 女, 通信作者, 博士, 助理研究员, 研究方向为人工智能与电力系统、虚拟电厂等, E-mail: lijiaimei@sjtu.edu.cn;

艾芊(1969), 男, 博士, 教授, 博士生导师, 主要研究方向为虚拟电厂规划与运行、大数据与数字孪生, E-mail: aiqian@sjtu.edu.cn。

(责任编辑 徐梅)