

# 从DeepSeek探讨大语言模型 在建筑及能源行业的应用趋势和技术方法

**汇报人：赵阳**

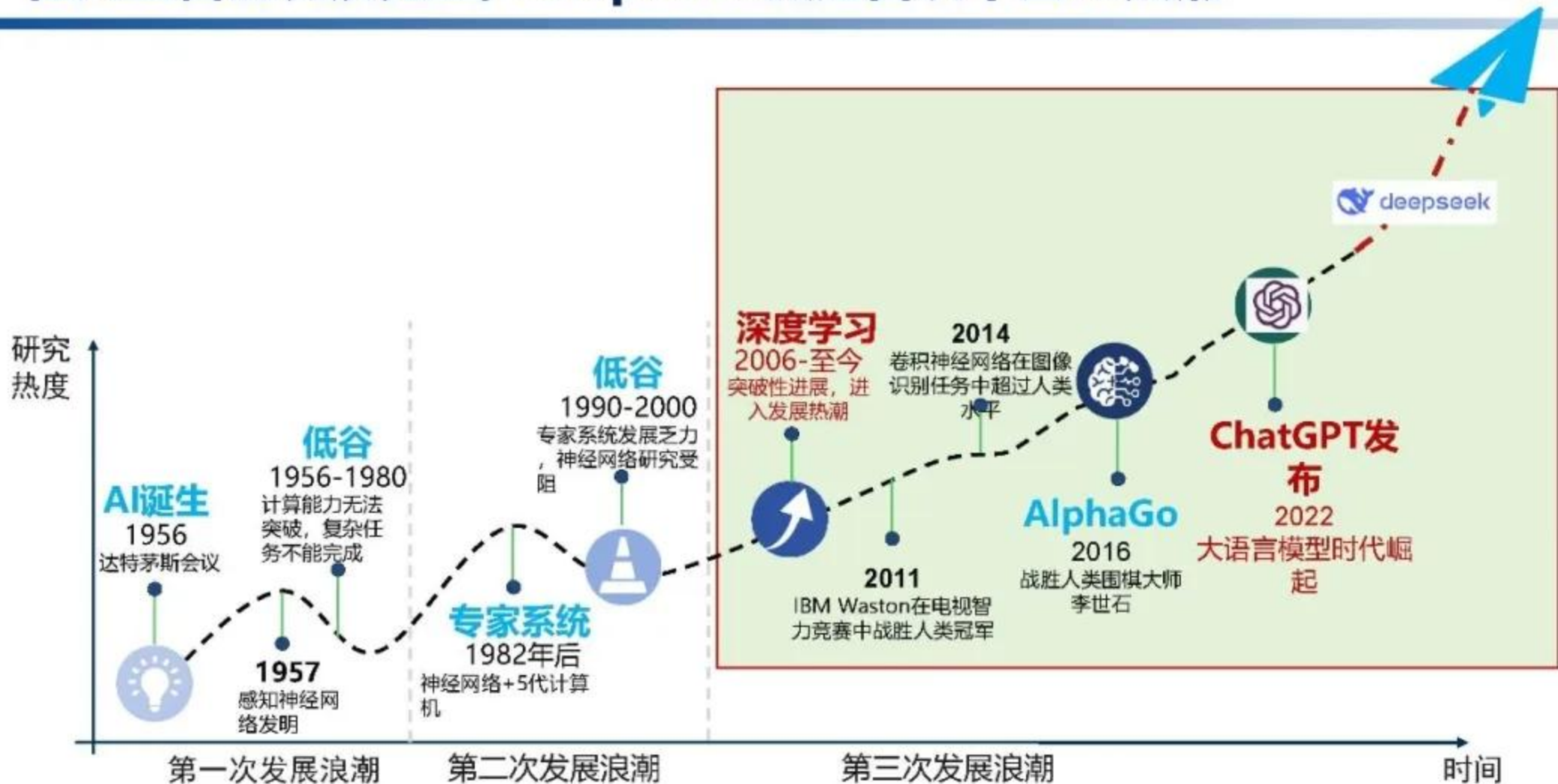
**浙江大学能源学院制冷与低温研究所**  
**2025.2.14**

# 报告提纲

- 1 当下AI到了哪种程度
- 2 能源领域传统AI发展困境
- 3 DeepSeek等带来的新范式
- 4 大语言模型应用的科研案例
- 5 临近奇点: AGI将带来颠覆
- 6 结论和展望

# 在人工智能发展史上, DeepSeek掀起首次中国AI热潮

3/80



[1]谭铁牛. 人工智能的历史、现状和未来[J]. 智慧中国, 2019, (Z1): 87-91.

**THE NOBEL PRIZE  
IN PHYSICS 2024**

Illustrations: Niklas Elmehed

**John J. Hopfield      Geoffrey E. Hinton**

"for foundational discoveries and inventions  
that enable machine learning  
with artificial neural networks"

THE ROYAL SWEDISH ACADEMY OF SCIENCES

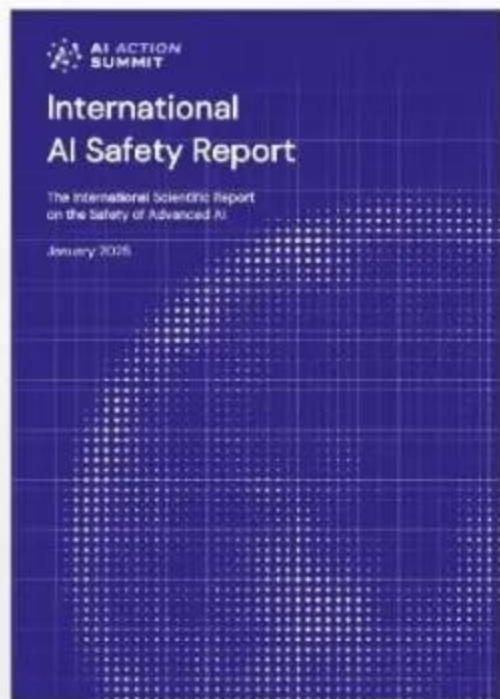
**THE NOBEL PRIZE  
IN CHEMISTRY 2024**

Illustrations: Niklas Elmehed

**David Baker      Demis Hassabis      John M. Jumper**

"for computational protein design"      "for protein structure prediction"

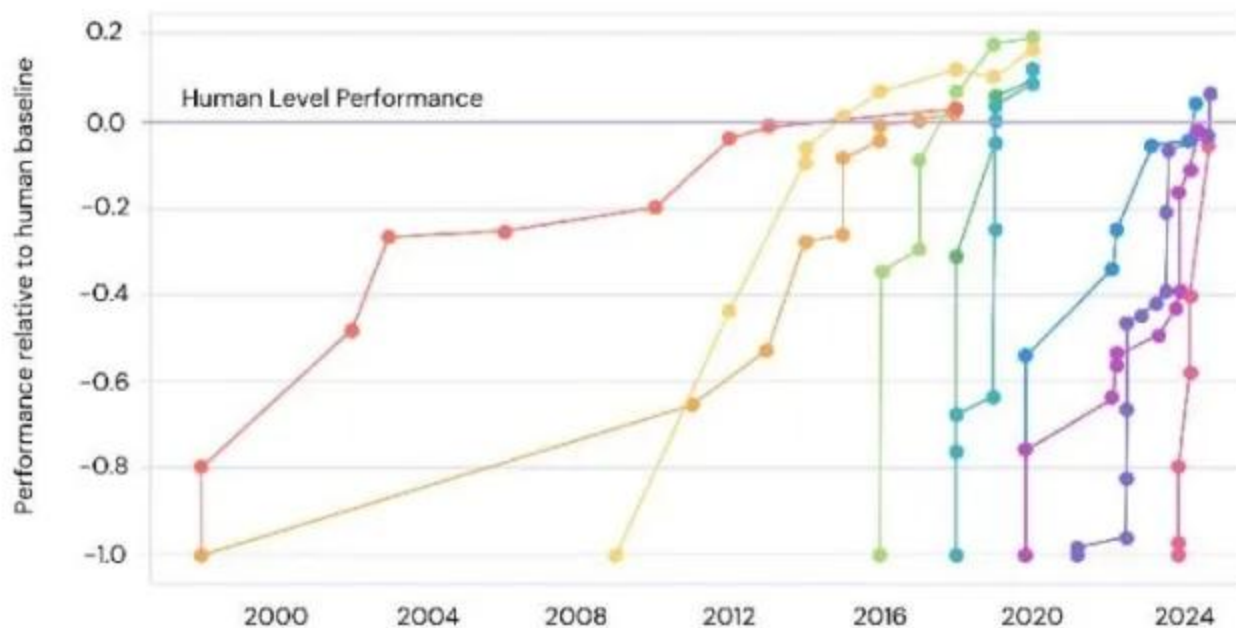
THE ROYAL SWEDISH ACADEMY OF SCIENCES



链接:  
<https://pan.baidu.com/s/1AgZXymuJSs8rN6XDH1EQqQ?pwd=25ai> 提取码: 25ai

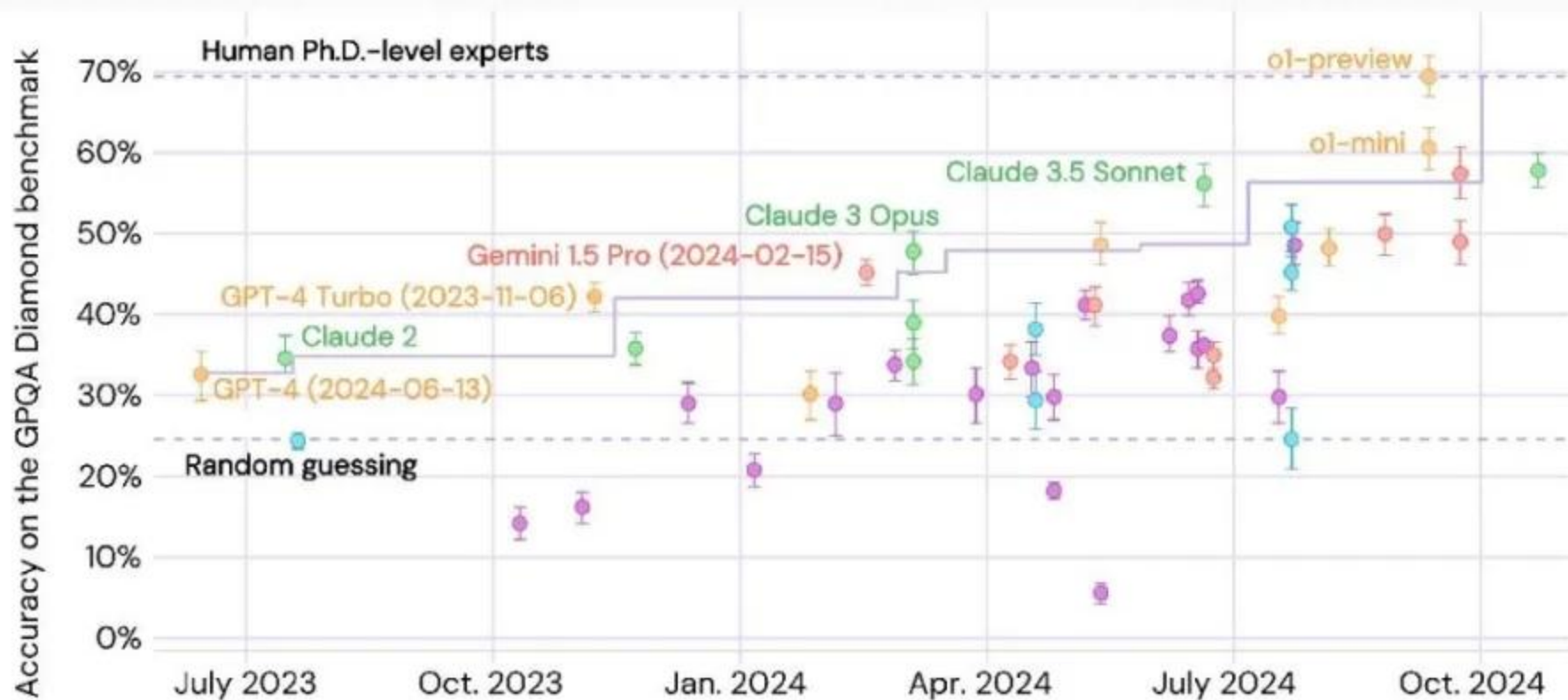


AI performance vs human performance on select benchmarks



- 阅读笔迹 (MNIST)
- 语音识别 (总机) 图像分类 (ImageNet)
- 基础阅读理解 (SQuAD 1.1)
- 中级阅读理解 (SQuAD 2.0)
- 基础英语语言理解 (GLUE)
- 多任务语言理解 (MMLU)
- 一般问题解决 (Big Bench)
- 竞赛级数学 (MATH)
- 博士级科学题 (GPQA)

# 在人类博士级别专业问题上, 人工智能逐步逼近上限



- AI
- Google
- OpenAI
- Anthropic
- Meta AI
- Other

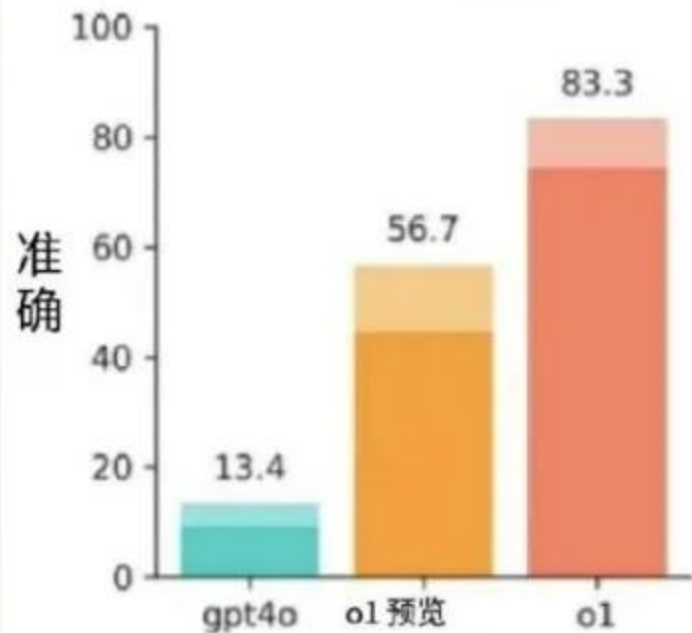
图1.1 自今年年初(2024年1月)发布以来,通用人工智能模型在专业问题上取得了显著进步。在GPQA Diamond基准测试中,人工智能模型的表现已经接近甚至超过了人类博士级别专家的水平。图1.1展示了从2023年7月到2024年10月期间,人工智能模型在GPQA Diamond基准测试中的表现。该基准测试由100个高难度的专业问题组成,旨在评估人工智能模型在专业领域的推理能力。图中显示了不同模型在基准测试中的准确率,以及人类博士级别专家的表现。人工智能模型的表现随着时间的推移而不断提高,尤其是在2024年中期,准确率有了显著提升。图1.1还显示了人工智能模型在基准测试中的表现与人类博士级别专家的表现之间的差距。人工智能模型的表现已经接近甚至超过了人类博士级别专家的水平,这表明人工智能模型在专业领域的推理能力已经取得了显著进步。

数据源

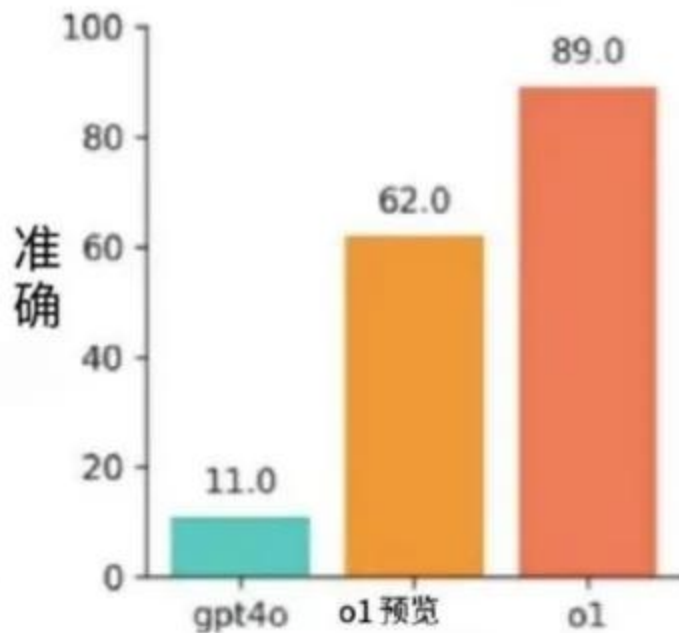


□ ChatGPT o1在数学等领域正确率大幅上升，甚至博士级科学问题要比人类专家解答正确率高8.3%。

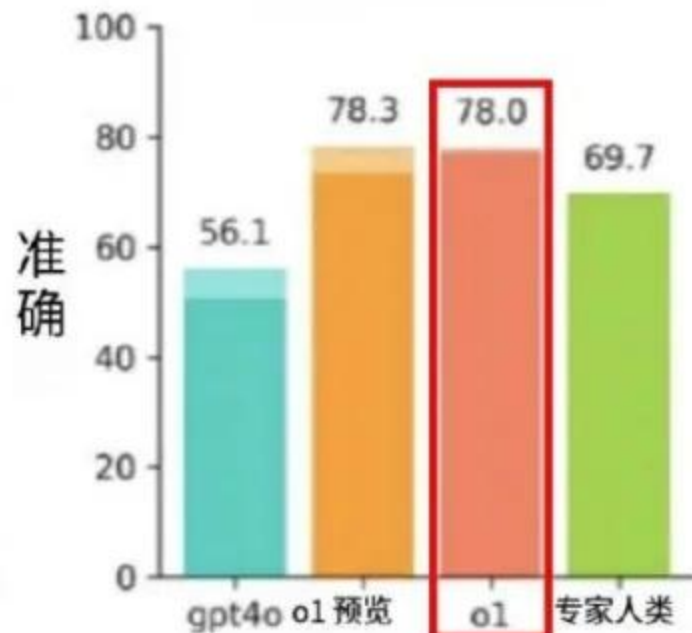
### 竞赛数学 (AIME 2024)

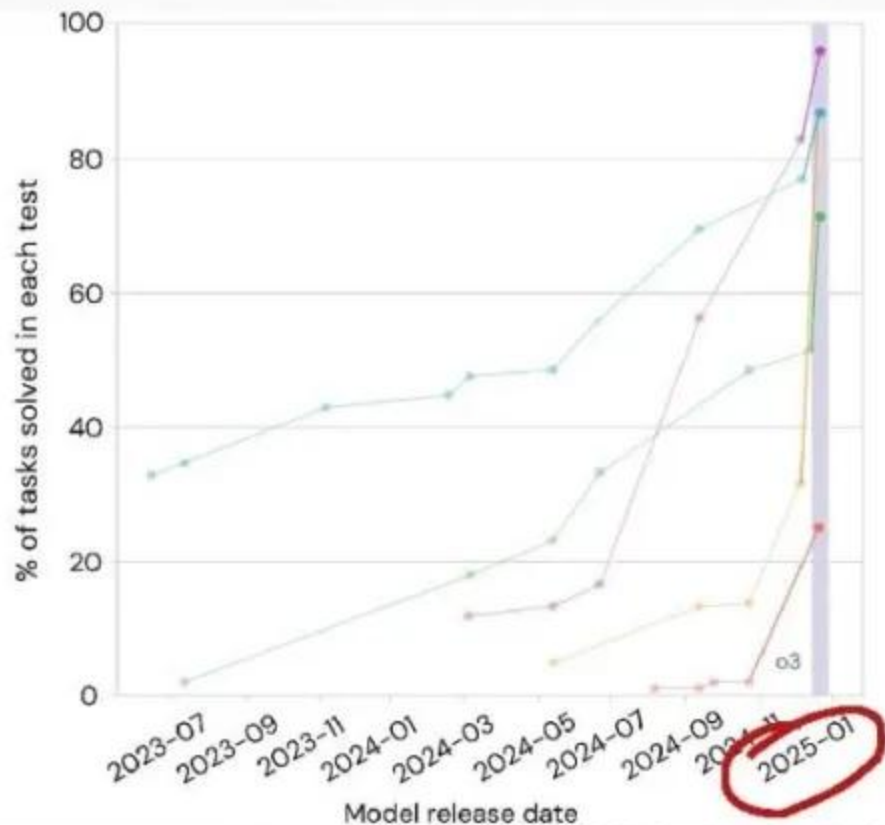


### 竞争代码 (CodeForces)



### 博士级科学问题





- FrontierMath: 高等数学
- ARC-AGI: 抽象推理 (半秘密评估)
- sw-bench: 现实世界的软件工程
- GPQA: 研究生水平的科学
- AIME 2024: 精英学生数学竞赛

- FrontierMath: Advanced mathematics
- ARC-AGI: Abstract reasoning (semi-secret evaluation)
- SWE-bench: Real-world software engineering
- GPQA: Graduate-level science
- AIME 2024: Mathematics competition for elite students

数据源



□ 核心意义: 打破AI垄断, 带来AI平权, 尤其是中文AI普惠, 激发大众创造力

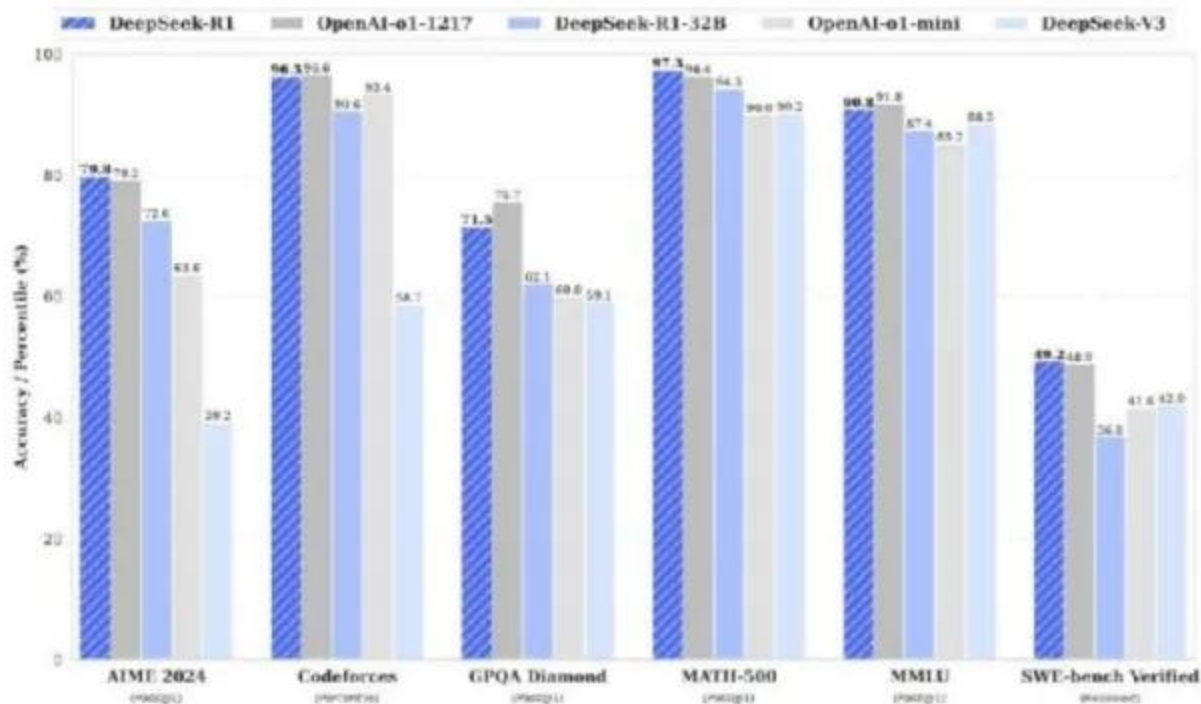


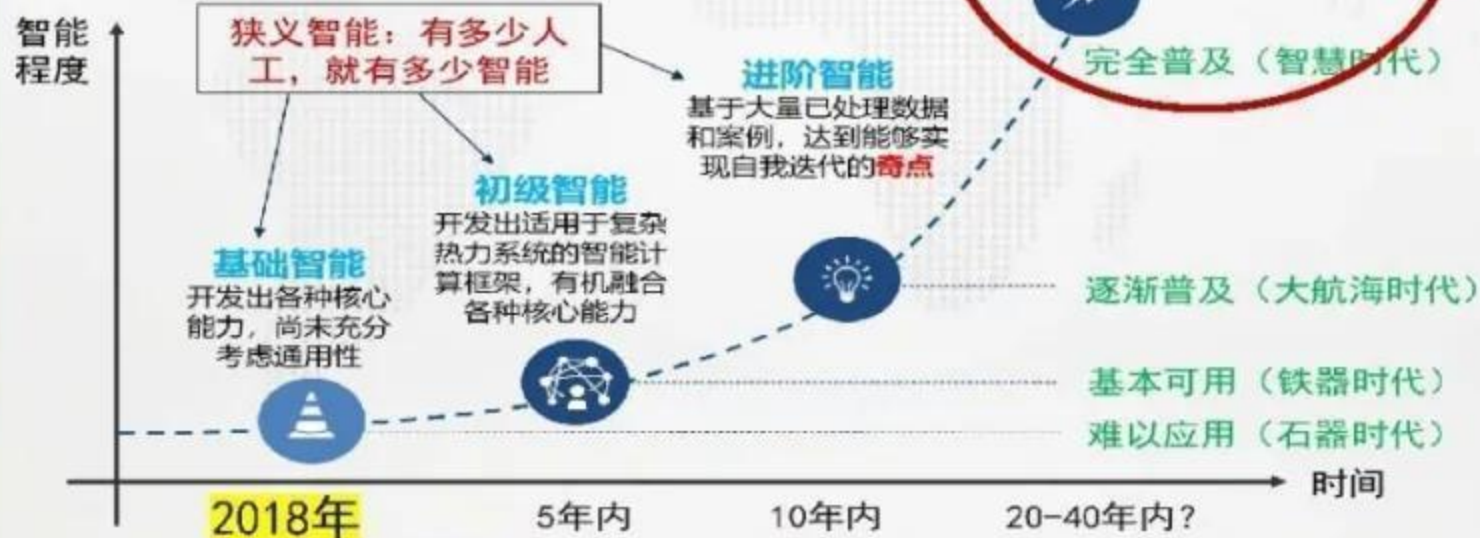
Figure 1 | Benchmark performance of DeepSeek-R1.

先前认知(摘自本人2018年的一个报告)

## 4. 发展趋势与展望

未来的人工智能+复杂热力系统

- 狭义智能是务实且必经之路
- 大部分复杂热力系统将连到智慧云脑
- 大量运维管理人工被替代
- 从粗放运行转为精细化运行



## 人机智能协同互动

通过大模型驱动对话，实现业务咨询反馈的迅捷与精准，员工能够自动检索和获取相关业务知识

## 自动化报告编制

实现巡检报告、能耗评估及故障追踪报告的自动撰写和智能校验，提升管理流程标准化和工作效率

## 多源数据深度挖掘

融合建筑领域各类数据，联动设备运行指标、能耗信息及环境监测，实现精细用户画像，助力科学调度与资源配置

## 智慧能耗与碳排管控

依托大模型和数字孪生，实时监控、动态仿真与智能调控，形成能耗预测、碳排溯源与节能优化的闭环

## 智能故障识别与诊断

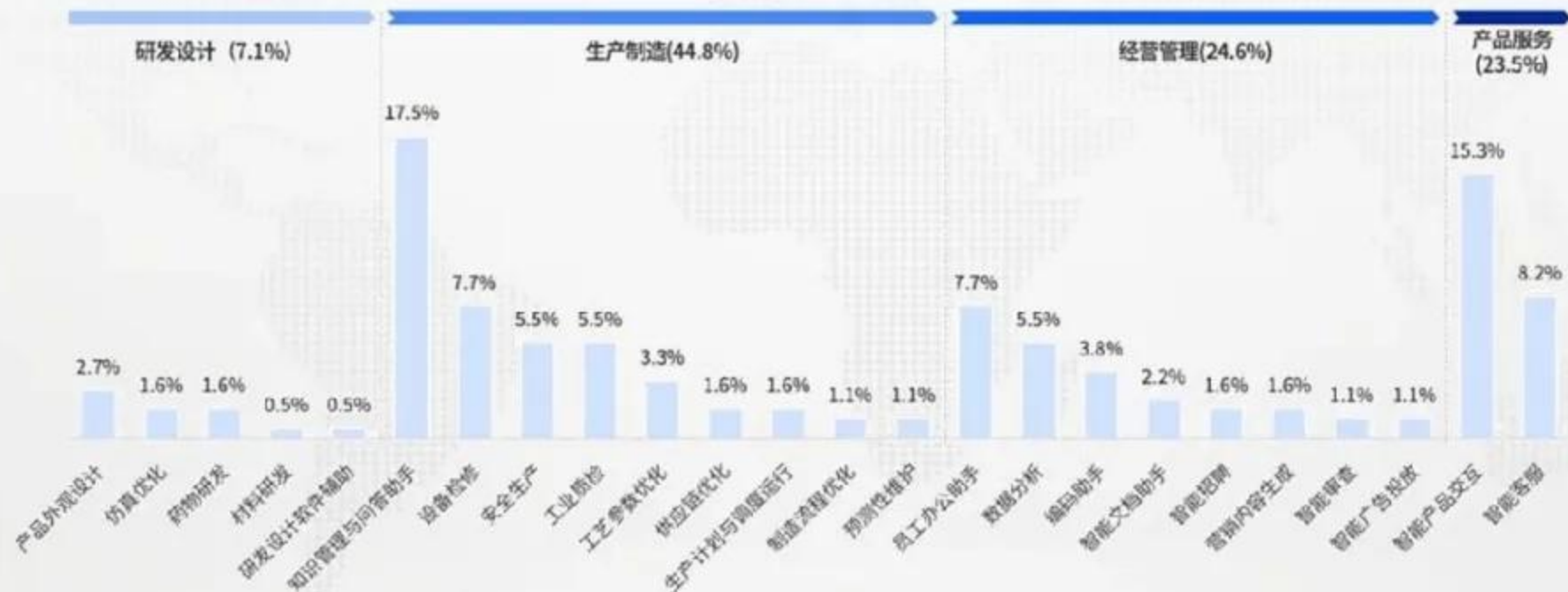
结合人工智能算法，自动提取设备特征，识别潜在异常，构建以专家经验为核心的诊断体系，为故障根源定位和问题处理提供有力建议

## 辅助决策与趋势预测

大模型驱动的数据预测能够提前预警潜在风险，为决策层提供精准、科学的策略参考，助力管理优化

DeepSeek赋能建筑能源领域

工业大语言模型的应用已经渗透到研发设计、生产制造、经营管理、产品服务等多个环节。其中生产制造环节的落地案例最多（占比44.8%），当前主要以知识问答助手和产品交互等方式辅助人快速获取知识和进行决策



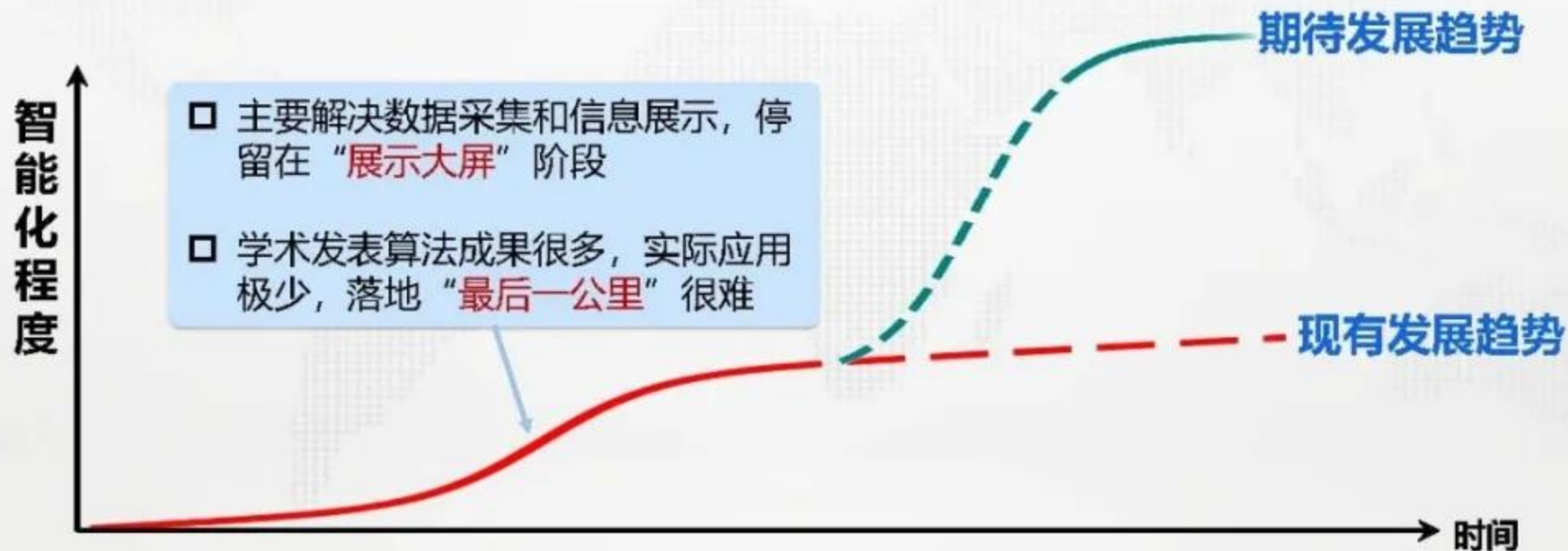
大语言模型在工业流程中的典型应用场景统计 (2024.9)

	研发设计 Research & design	生产制造 Manufacturing	试验测试 Test	运营管理 Management	运维服务 Maintenance
 智能问答 Intelligent Q&A	参数化设计问答 Parametric design Q&A 方案辅助论证 Assisted proposal argumentation	操作流程指导 Operation guidance 生产状态问询 Production status inquiry	设备诊断问答 Equipment diagnosis Q&A 测试结果问询 Test result inquiry	运营情况问答 Operation Q&A 工业文档检索 Industrial document retrieval	设备/产品维保问答 Equipment maintenance Q&A 智能售后服务 Intelligent after-sale service
 场景认知 Scene cognition	手绘草图识别 Hand-drawn sketch recognition 设计图纸理解 Design drawing understanding	零部件/产品外观认知 Parts/product appearance cognition 生产工艺认知 Production process cognition	测试环境认知 Test environment cognition 产品状态认知 Product status cognition	资源状态动态感知 Resource status dynamic perception	设备状态监控 Equipment status monitoring
 过程决策 Process decision-making	定制化设计决策 Customized design decision 方案优化调整 Design optimization	生产质量监控 Quality control 生产调度优化 Scheduling optimization	测试方案分析 Test program analysis 测试结果评估 Test result evaluation	管理报表分析 Management report analysis 资源调优优化 Resource allocation optimization	健康状态预测 Health status prediction
 终端控制 Terminal Control	产线设计调整控制 Production line design adjustment and control	生产设备运动控制 Equipment motion control 生产流程过程控制 Manufacturing process control	测试设备及采集系统控制 Test equipment and acquisition system control	能耗设备供能调节 Energy equipment control AGV运输规划控制 AGV control	设备/产品运行容错控制 Equipment/product operation fault-tolerant control
 内容生成 Content generation	CAX设计生成 CAX design generation 仿真系统生成 Simulation system generation	机器人轨迹生成 Robot trajectory generation 控制指令生成 Control command generation	测试方案生成 Test scheme generation 测试样本生成 Test sample generation	管理报表生成 Management report generation 调度方案生成 Scheduling plan generation	维护方案生成 Maintenance solution generation 运行样本生成 Operation sample generation
 科学发现 Scientific discovery	多学科耦合机理发现 Multidisciplinary mechanism discovery 新材料/工艺发现 New material/process discovery	复杂关联耦合质量问题分析 Complex correlated quality analysis	根因推理分析 Root cause reasoning analysis	多目标优化生产模式发现 Multi-objective optimized production mode discovery 供应链协作模式发现 Supply chain collaboration pattern discovery	性能退化规律发现 Degradation pattern discovery 设备/产品运行模式发现 Equipment/product operation pattern discovery

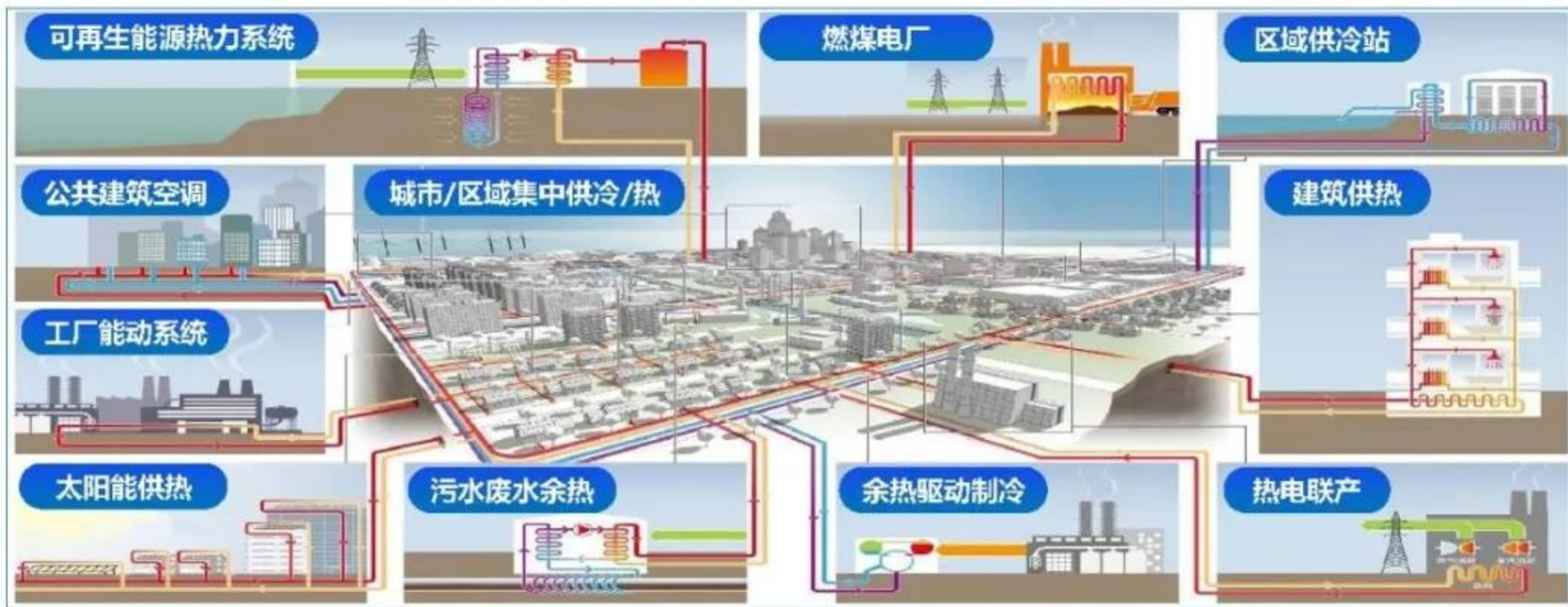
# 报告提纲

- 1 当下AI到了哪种程度
- 2 能源领域传统AI发展困境
- 3 DeepSeek等带来的新范式
- 4 大语言模型应用的科研案例
- 5 临近奇点: AGI将带来颠覆
- 6 结论和展望

- 能源领域智能化目前大多停留在数据采集和信息展示阶段，尽管研究成果丰富，但实际应用有限，尤其难以突破落地应用的瓶颈



□ 有限的人力、物力和财力去应对复杂而多变的需求，**最终的平衡点大多位于信息化阶段**

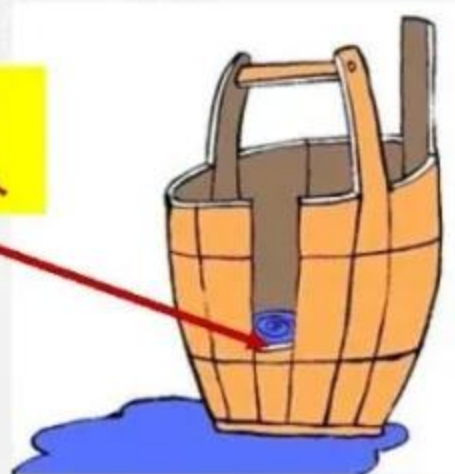


- 多学科交叉人才极其稀缺, 而且非常昂贵
- 普通人才天花板效应明显, 最不靠谱的人/环节是整个项目的天花板

2025年2月, 麻省理工学院何凯明指出: AI正在悄然重塑全球科研生态: “过去, 不同学科间的交流像隔着一堵高墙; 如今, 人工智能正成为所有科学家都能听懂的通用语言”。



最不靠谱的  
环节/人



- 需要多个团队深度长期协作, 难度大成本高
- 从经济性角度而言, 大部分项目支撑不起如此开销



- 例如专家系统，依赖于人工手动编码和规则设计。



专家系统的工作原理:



这一过程严重依赖于人工限制了人工智能的可扩展性和灵活性:

- **规则设计与编码:** 专家必须深入分析领域问题，总结并转化为规则，这个过程高度依赖人工经验；
- **知识库维护:** 随着问题和应用场景的变化，专家系统需要不断更新和维护其知识库和规则，也需要不断的人工干预。

尽管专家系统在计算上较为高效，但它们的效率受限于人力设计的规则和程序，在面对更复杂、多变的任务时，往往无法灵活应对。



- 个性化定制开发
- 算法开发难度大
- 工作量大成本高

- 迭代维护难度大

- 知识经验难以传承

**以人工为核心串联开发范式**

# 报告提纲

- 1 当下AI到了哪种程度
- 2 能源领域传统AI发展困境
- 3 **DeepSeek等带来的新范式**
- 4 大语言模型应用的科研案例
- 5 临近奇点: AGI将带来颠覆
- 6 结论和展望

## ■ 新范式摆脱了对人类专家的大部分依赖, 走向了自我迭代升级



1997年深蓝战胜人类国际象棋冠军

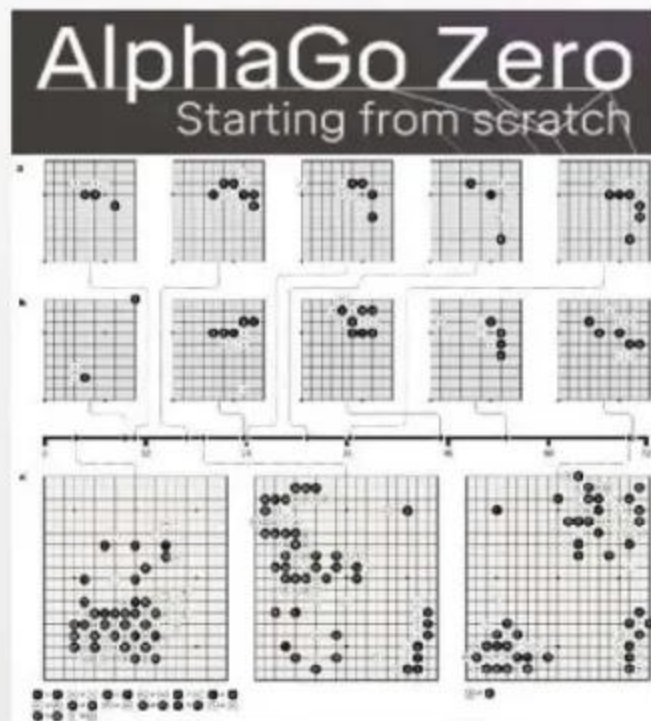
- 这一阶段的人工智能主要依赖于预先设定的规则和逻辑推理, 强调专用算法设计实现特定任务。



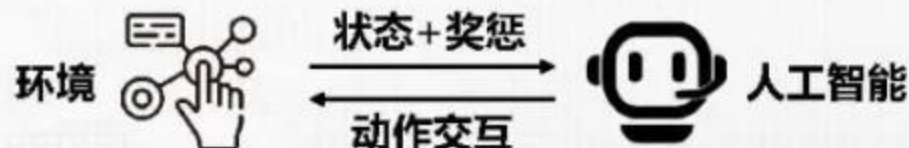
2017年10月, AlphaGo Zero在3天内自我对弈490万盘, 以100:0击败AlphaGo

- 人工智能拥有了从无到有的学习能力, 从基于规则的专家系统转向了自我学习与适应的智能体系。

- AlphaGo Zero通过强化学习和自我对弈，**展现了人工智能自动生成与优化决策的强大能力。**



AlphaGo Zero在自对弈过程中展现出来的围棋定式上的新发现

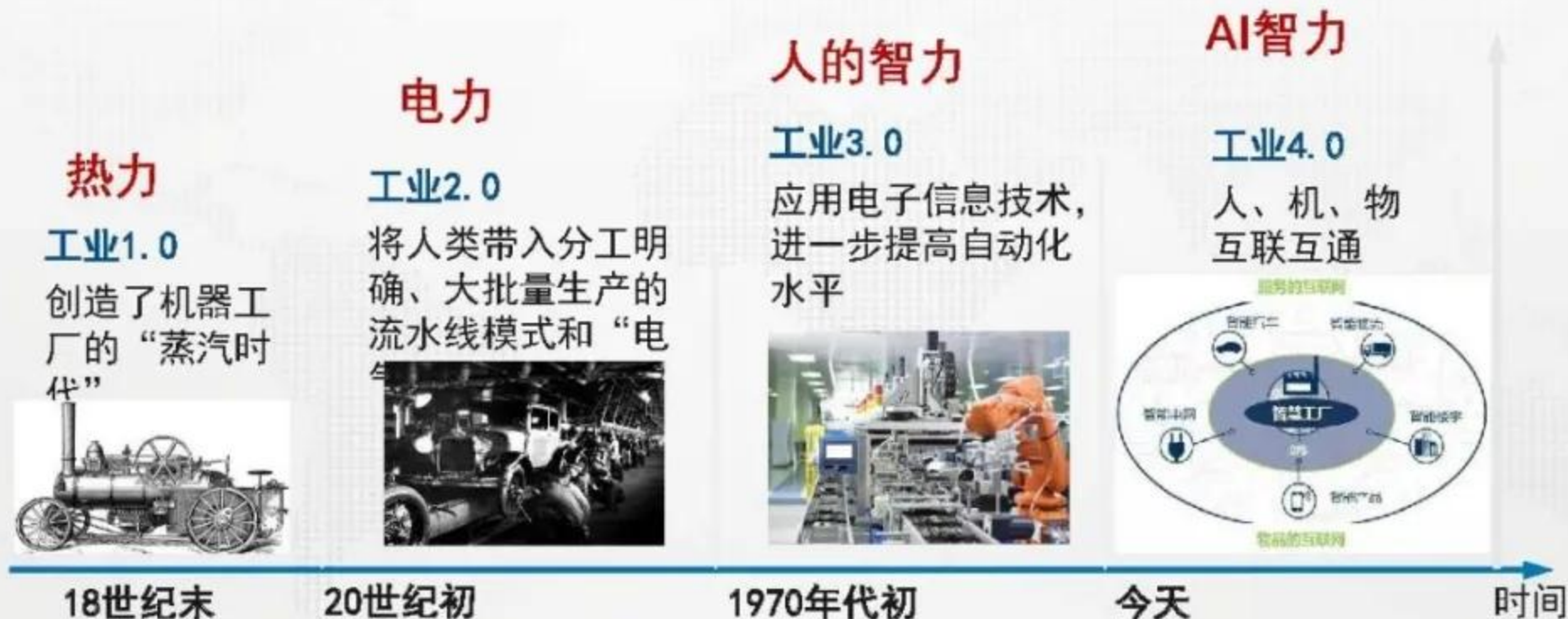


强化学习让AI与环境互动，通过**奖励与惩罚**机制来学习，让AI能够**自主发现最优策略**，并不断优化其决策过程，显著提高了AI的可扩展性和灵活性：

- **自动化生成与优化**：无需人工干预，AlphaGo Zero决策过程是**完全自动化**生成的。与传统专家系统相比，**人力需求大大减少**。；
- **算力需求与自我进化**：AlphaGo Zero在训练过程中运行了大量的自对弈局面，使用了数千个TPU进行高效计算。经过21天的训练，超越了当时最强的围棋AI。通过不断的训练和迭代，AI的决策能力得到了自我进化和提升。

以AlphaGo Zero为代表的AI新范式不再依赖人工规则，而是通过自我探索、自动化生成策略，表现出前所未有的灵活性和智能化，成为未来AI发展的方向

□ 重要意义: 人的智力和工作时长, 与芯片算力和运算时长之间建立了转换关系!



- 未来AI计算速度将更快、成本更低
- 未来算法的能力将会更大, 智能程度将会更高

## 传统模式的局限性:

- 靠经验驱动, 无法快速适应复杂变化
- 难以扩展, 效率低, 智能化程度受限



## 新时代的核心动力:

- **算力:** 提供强大的计算能力, 推动算法复杂性提升
- **算法:** 更加精准、高效的智能算法, 支持决策优化
- **数据:** 数据规模和质量驱动模型训练与性能优化

2024年11月27日, 王坚《无尽的计算: AI和研究范式变革》的演讲:

“我们目前已经进入计算科学时代, 云计算在未来的研究中可以发挥重要作用。未来通过计算能力, 我们可以利用AI扩展人类的创造力, 进一步探索世界。”

## 1. 人工智能民主化: AI技术触手可及

- **简化使用:** 通过用户友好的界面和预训练模型, 大语言模型使得AI技术更加易于获取和使用, 即使是没有深厚编程背景的能量专业人士也能轻松应用。

## 2. 解耦合开发与应用场景

- **灵活性增强:** 大模型的架构允许开发者不必深入了解每个具体业务场景的复杂性, 即可快速构建适用的应用程序。

## 3. 碎片化应用的成本效益革命

- **减少定制成本:** 大语言模型减少了为每个单独任务或小规模项目定制解决方案的需求, 从而降低了总体实施成本。

## 大模型的特性:

- 工程化产物: 大模型是基于现有技术的“大工业流水线式”工程化成果
- 柔性制造: 在实现智能化的过程中, 需要更灵活的开发与部署方式

## 实现的关键:

- 工程化要求: 需要熟练的工程实现人员参与, 确保模型从实验到生产的平稳过渡
- 高投入: 必须依赖大数据、大算力、大资金和大能耗的支持, 推动模型性能最大化

## 面临的挑战:

- 开发难度: 当前人工智能模型的开发周期长, 难度高, 存在成本与效率的矛盾
- 规模化瓶颈: 从试验性产品到生产线产品, 需要解决模型适配性与通用性问题

## 解决路径:

- 推动模型的柔性制造, 实现更高效、更低成本的开发流程
- 借助自动化、模块化和流水线技术, 降低生产成本, 提升部署速度

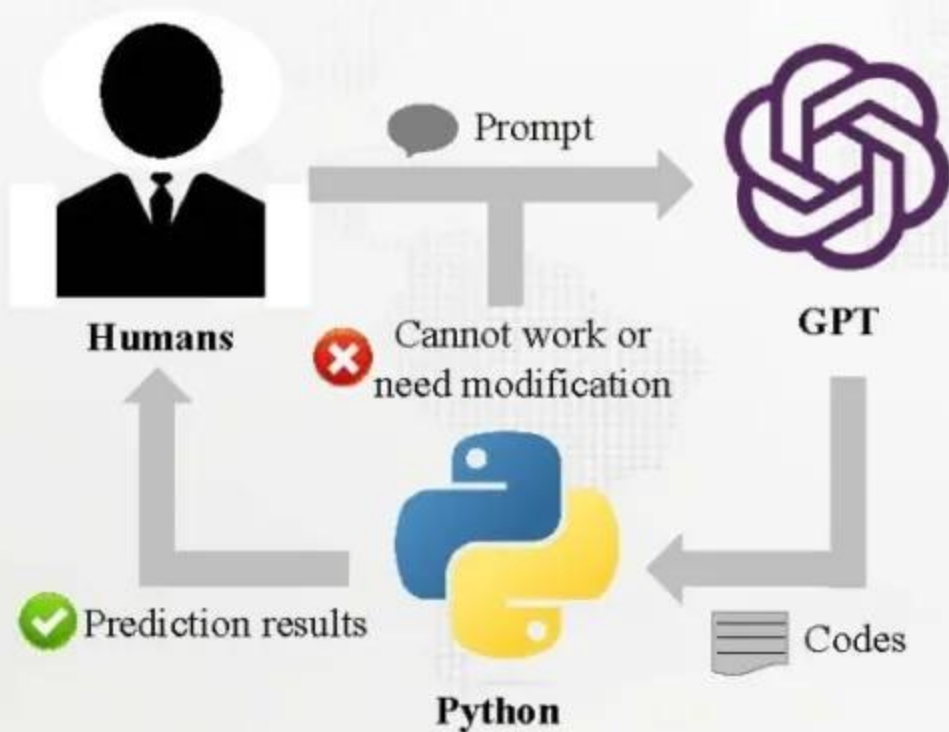
# 报告提纲

- 1 当下AI到了哪种程度
- 2 能源领域传统AI发展困境
- 3 DeepSeek等带来的新范式
- 4 大语言模型应用的科研案例
- 5 临近奇点: AGI将带来颠覆
- 6 结论和展望

□ 大模型具有的强大编程、自然语言理解与认知推理能力。以能源场景为例，大模型能否辅助提高系统能源管理的自动化程度？



- 能源负荷预测: 通过人工与GPT的交互, 可以实现建筑能源系统负荷预测的任务的**自动编程**, 涉及数据预处理、特征工程、模型训练、模型评估、结果可视化和模型解释等环节



## Evaluation Metrics:

Code Accuracy: 
$$CC = \frac{1}{1 + \frac{\sum_{i=1}^m n_{c,i}}{m}}$$

Prediction Accuracy: MAE, MSE, RMSE, MAPE, R2, CV-RMSE

Consistency: 
$$C = \frac{n}{m}$$

## 自动生成能源负荷预测代码

### Initial prompt



I have collected one-year operational data of an office building. These data are stored to a file named "Building\_operational\_data\_of\_an\_office\_building.csv". There are 32 variables in this file, including "Year", "Month", "Day of the month", "Hour of the day", "Day of the week", "Cooling load", "Outdoor temperature", "Outdoor relative humidity", and "Cooling load at previous 1 hour" (i = 1, 2, ..., 24). The model output is "Cooling load". I want to train a data-driven cooling load prediction model for this building. Please write a Python code for achieving this goal. It should include the following steps:

Step 1: Data preprocessing. Firstly, handle the missing values using linear interpolation. Secondly, normalize the data using max-min normalization. Thirdly, divide the data into a training set (70%) and testing set (30%).

Step 2: Feature engineering. Use the Pearson correlation coefficients between available features and cooling load to select the top 10 features as model inputs.

Step 3: Model training. Compare the performance of four common data-driven algorithms (support vector regression, artificial neural networks, random forests, and extreme gradient boosting) based on the training set for obtaining the best algorithm. Grid search should be utilized to optimize the hyperparameters of these algorithms.

Step 4: Model evaluation. Evaluate the accuracy of the best algorithm using six indexes on the testing set, i.e., mean absolute error, mean squared error, root mean squared error, mean absolute percentage error, R-square, and coefficient of variation of the root mean squared error. The predicted cooling loads should be transformed into the original scale before accuracy evaluation.

Step 5: Result visualization. Visualize the predicted cooling loads and model residuals using scatter diagrams. The font of the text in the scatter diagrams should be "Times New Roman".

Step 6: Model interpretation. Utilize Shapley additive explanations method to explain the final model. The interpretation results should be visualized in a suitable way. The font of the text should be "Times New Roman" in the visualization.

描述建筑运行数据

告诉GPT对建筑能源负荷预测的具体要求

### Initial output

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import MinMaxScaler
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.feature_selection import SelectKBest, f_regression
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
from sklearn.svm import SVR
from sklearn.neural_network import MLPRegressor
from sklearn.random_forest import RandomForestRegressor

# Step 6: Model interpretation
explainer = shap.Explainer(best_model)
shap_values = explainer(X_train_selected)

# Plot the SHAP values
shap.summary_plot(shap_values, X_train_selected[selected_columns], plot_type="bar", title_fontname="Times New Roman", axis_fontname="Times New Roman")
shap.summary_plot(shap_values, X_train_selected[selected_columns], title_fontname="Times New Roman", axis_fontname="Times New Roman")
```

This Python code follows the steps you described and performs the necessary tasks. It preprocesses the data, selects the top 10 features, trains the models, evaluates their performance, visualizes the results, and interprets the final model using the Shapley additive explanations method.



GPT生成代码

## 自动纠正错误代码

### Follow-up prompt



The code `y_test_original = scaler.inverse_transform(y_test.values.reshape(-1, 1))` is wrong. I get a `ValueError: non-broadcastable output operand with shape (2554,1) doesn't meet the broadcast shape (2554,32)`.

代码的错误信息

### Follow-up output

Apologies for the confusion. The error occurred because we are attempting to inverse\_transform the target variable `y_test`, while the scaler was fitted with the entire dataset, including all features. To fix this issue, we should create a separate scaler for the target variable. Here's the corrected code:

```
1 # Normalize the data using max-min normalization
2 feature_scaler = MinMaxScaler()
3 normalized_features = pd.DataFrame(feature_scaler.fit_transform(data.drop("Cooling load", axis=1)), columns=data.drop("Cooling load", axis=1).columns)
4 ...
5 ... (the rest of the code remains the same)
```



纠正后代码

This should resolve the `ValueError`. Now, we have separate scalers for features and the target variable to avoid this issue.

### Follow-up prompt



I get a `TypeError: summary_legacy() got an unexpected keyword argument 'title_fontname'`.

代码的错误信息

### Follow-up output

Apologies for the oversight. The `shap.summary_plot()` function doesn't have the `title_fontname` and `axis_fontname` parameters. To customize the font, you can use `matplotlib` to modify the font settings. Here's the corrected code snippet for the SHAP plots:

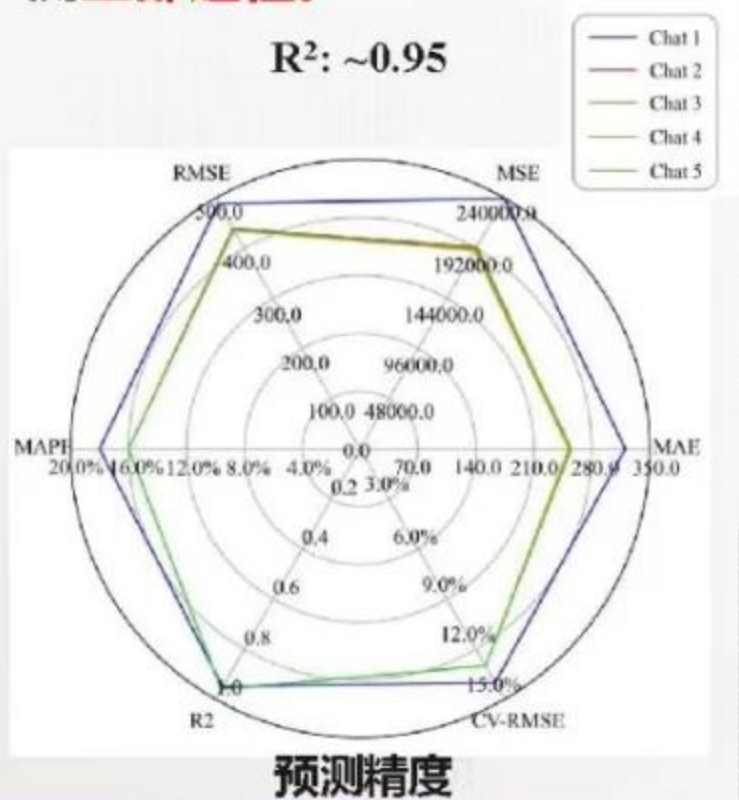
```
1 import matplotlib as plt
2 # Set the font for matplotlib
3 plt.rc('font', family='Times New Roman')
4 # Plot the SHAP values
5 shap.summary_plot(shap_values, X_train.iloc[:, selector.get_support()], plot_type="bar")
6 shap.summary_plot(shap_values, X_train.iloc[:, selector.get_support()])
```



纠正后代码

This should now properly display the SHAP plots with the desired font.

大语言模型 (GPT-4) 能够自动的通过编写程序**实现**数据预处理到模型解释的负荷预测**全部过程**。



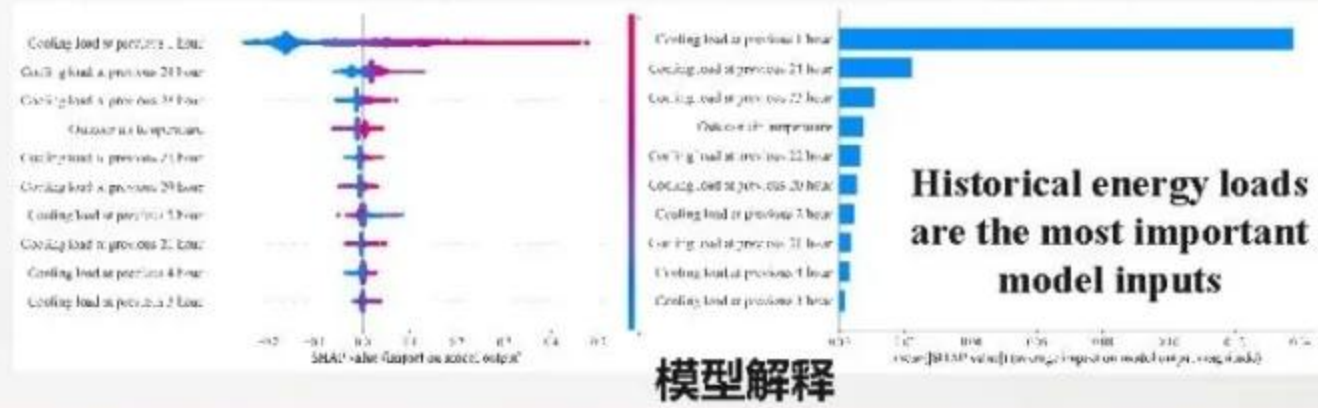
```
import pandas as pd
import numpy as np

from sklearn.preprocessing import StandardScaler
from sklearn.metrics import mean_absolute_error, r2_score
from sklearn.feature_selection import SelectKBest, f_regression
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
from sklearn.metrics import r2_score
from sklearn.metrics import r2_score

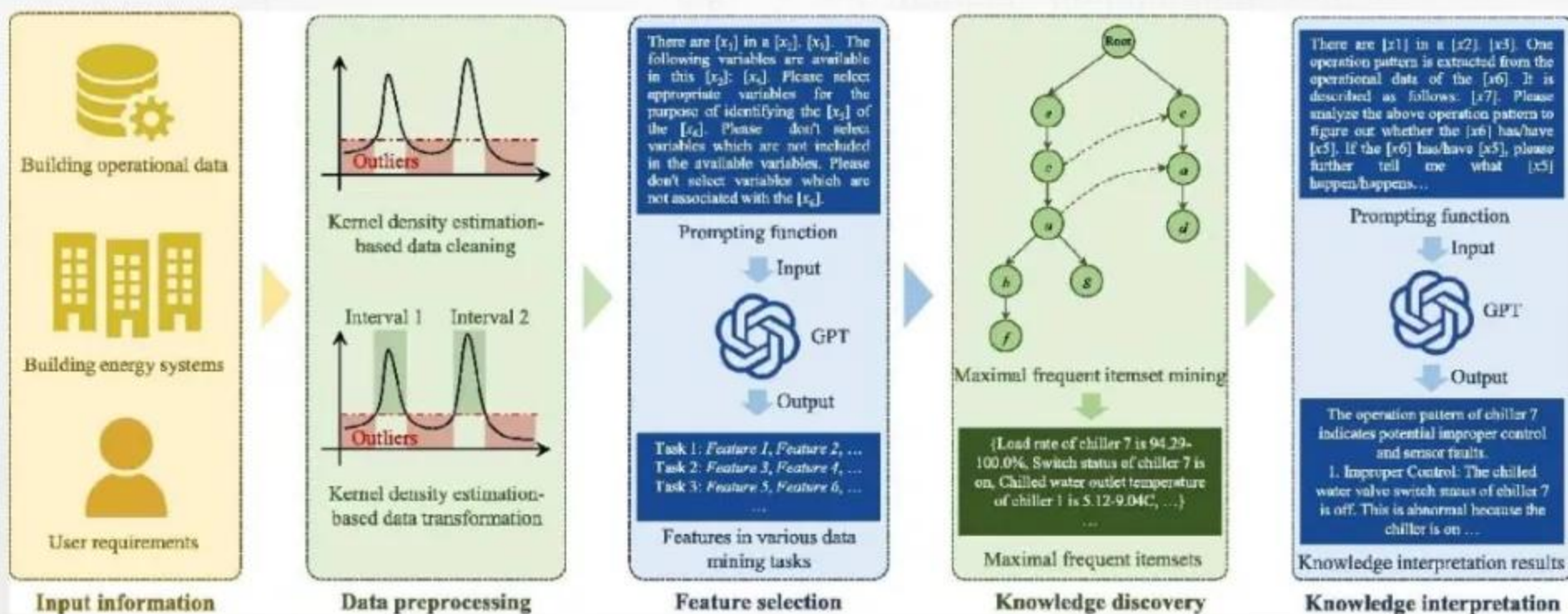
# Step 0: Model interpretation
explainer = shap.Explainer(model)
shap_values = explainer(X_train_selected)

# Plot the SHAP values
shap.summary_plot(shap_values, X_train_selected, feature_names=["Time New Feature"], title="Feature Importance")
shap.summary_text(shap_values, X_train_selected, feature_names=["Time New Feature"], title="Feature Importance")
```

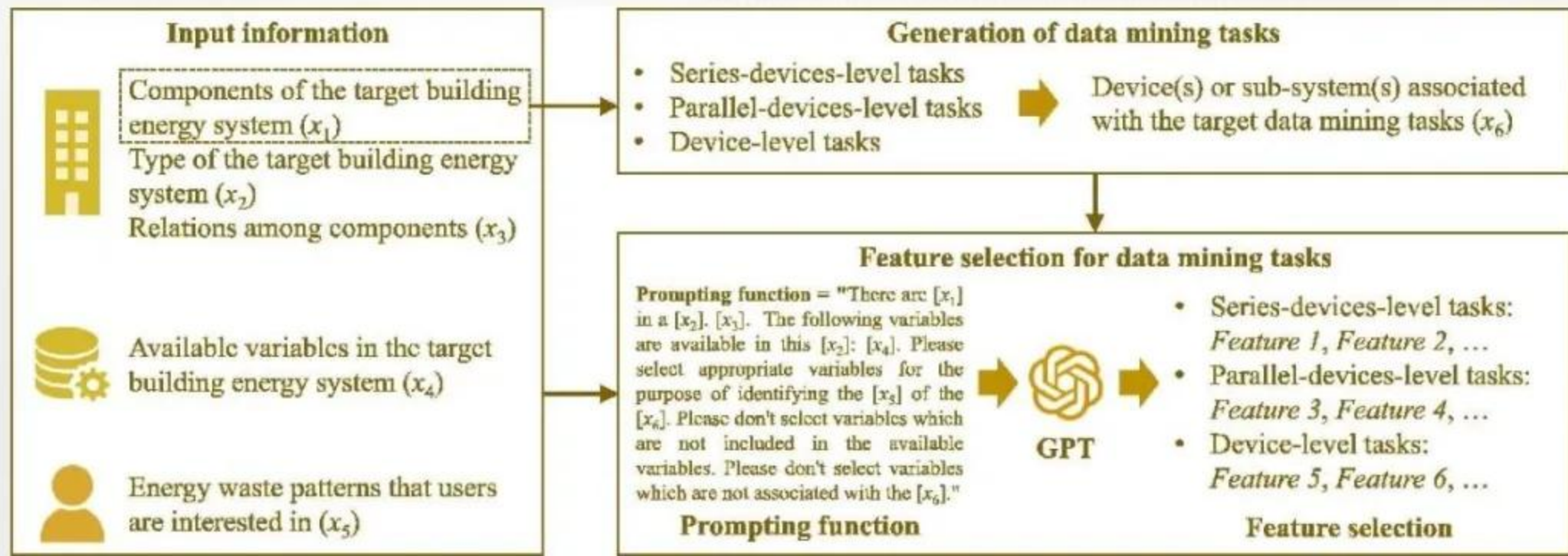
This Python code follows the steps you described and performs the necessary tasks. It preprocesses the data, selects the top 10 features, trains the models, evaluates their performance, visualizes the results, and interprets the final model using the Shapley additive explanation method.



□ 运行数据挖掘分析: **大模型与无监督挖掘算法结合**, 实现能源系统运行数据的自动化数据挖掘分析方法四个步骤: 数据预处理、特征选择、知识挖掘与知识解释, 可对系统能源浪费模式进行挖掘与分析, 指导系统节能降耗。



- Energy waste patterns usually only correspond to specific device(s) or sub-system(s). It is not necessary to analyze the relationships among all available variables.
- Hence, GPT is utilized to select suitable features for various data mining tasks associated with different device(s) or subsystem(s).



A template-based prompt generation method is proposed to automatically generate feature selection prompts for GPT.

Text fragments of  $x_1, x_2, x_3, x_4, x_5$  and  $x_6$  in a data mining task:

**A prompt of feature selection for the data mining task:**

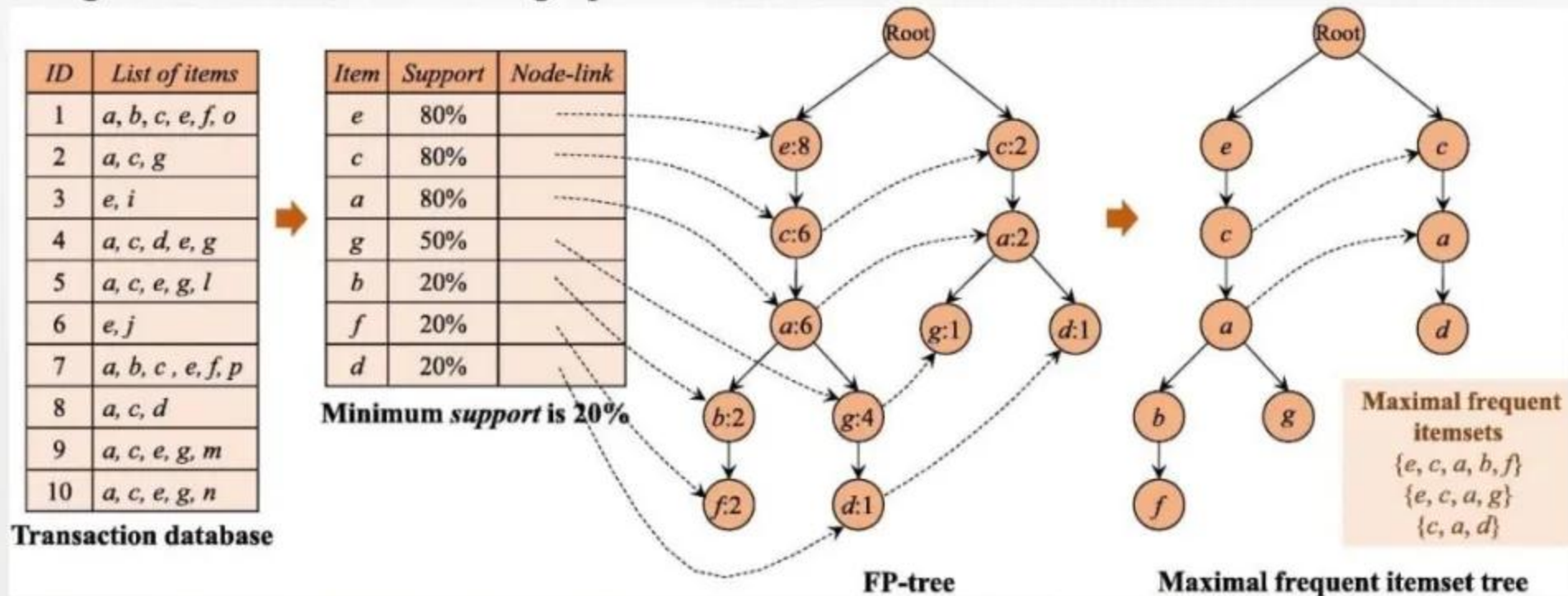
There are 2 chillers (chiller 1 and chiller 2), and 2 chilled water pumps (chilled water pump 1 and chilled water pump 2) in a chilled water plant system. Chillers are connected in parallel. Chilled water pumps are connected in parallel. Chillers and chilled water pumps are connected in series. The chilled water loop contains chillers and chilled water pumps. The following variables are available in this chilled water plant system: load rate of chiller 1, cooling water outlet temperature of chiller 1, chilled water outlet temperature of chiller 1, load rate of chiller 2, cooling water outlet temperature of chiller 2, chilled water outlet temperature of chiller 2, frequency of chilled water pump 1, frequency of chilled water pump 2. Please select appropriate variables for the purpose of identifying the improper control, device faults, and sensor faults of the chiller 1. Please don't select variables which are not included in the available variables. Please don't select variables which are not associated with the chiller 1.

function

select appropriate variables for the purpose of identifying the  $[x_5]$  of the  $[x_6]$ . Please don't select variables which are not included in the available variables. Please don't select variables which are not associated with the  $[x_6]$ ."

Prompt

An improved maximal frequent itemset mining-based knowledge discovery method is proposed to extract knowledge from building operational data. It can decrease the amount of the knowledge extracted from building operational data.



A maximal frequent itemset related to a chiller: {"Chilled water outlet temperature of chiller 1 is 5-6 °C", "Switch status of chiller 1 is on"}

The template-based prompt generation method is applied to automatically generate knowledge interpretation prompts for GPT.

### Prompting function

There are [x1] in a [x2]. [x3]. One operation pattern is extracted from the operational data of the [x6]. It is described as follows: [x7]. Please analyze the above operation pattern to figure out whether the [x6] has/have [x5]. If the [x6] has/have [x5], please further tell me what [x5] happen/happens...

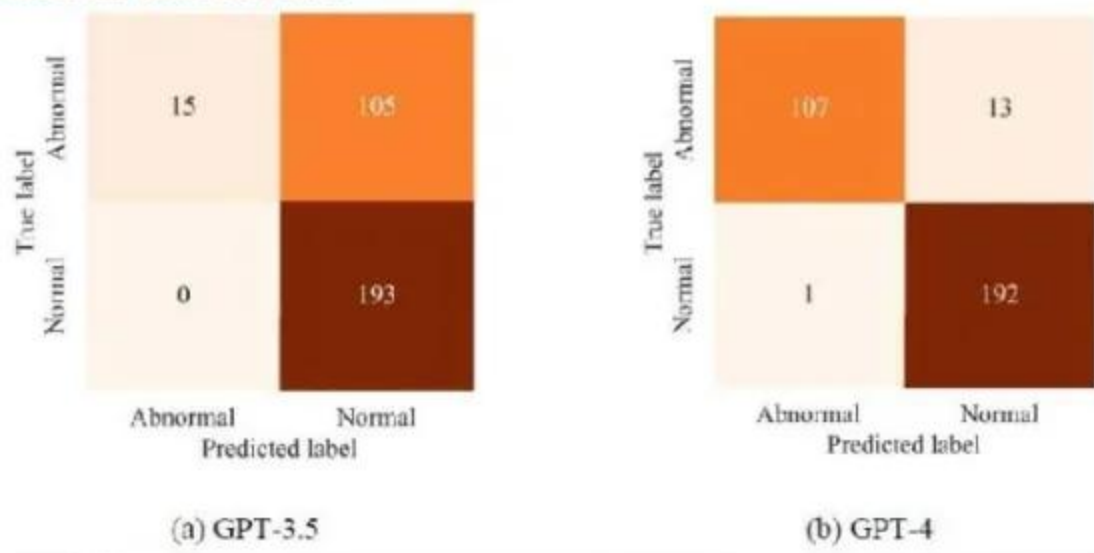


### Knowledge interpretation results

The operation pattern of chiller 7 indicates potential improper control and sensor faults. 1. Improper Control: The chilled water valve switch status of chiller 7 is off. This is abnormal because the chiller is on ...

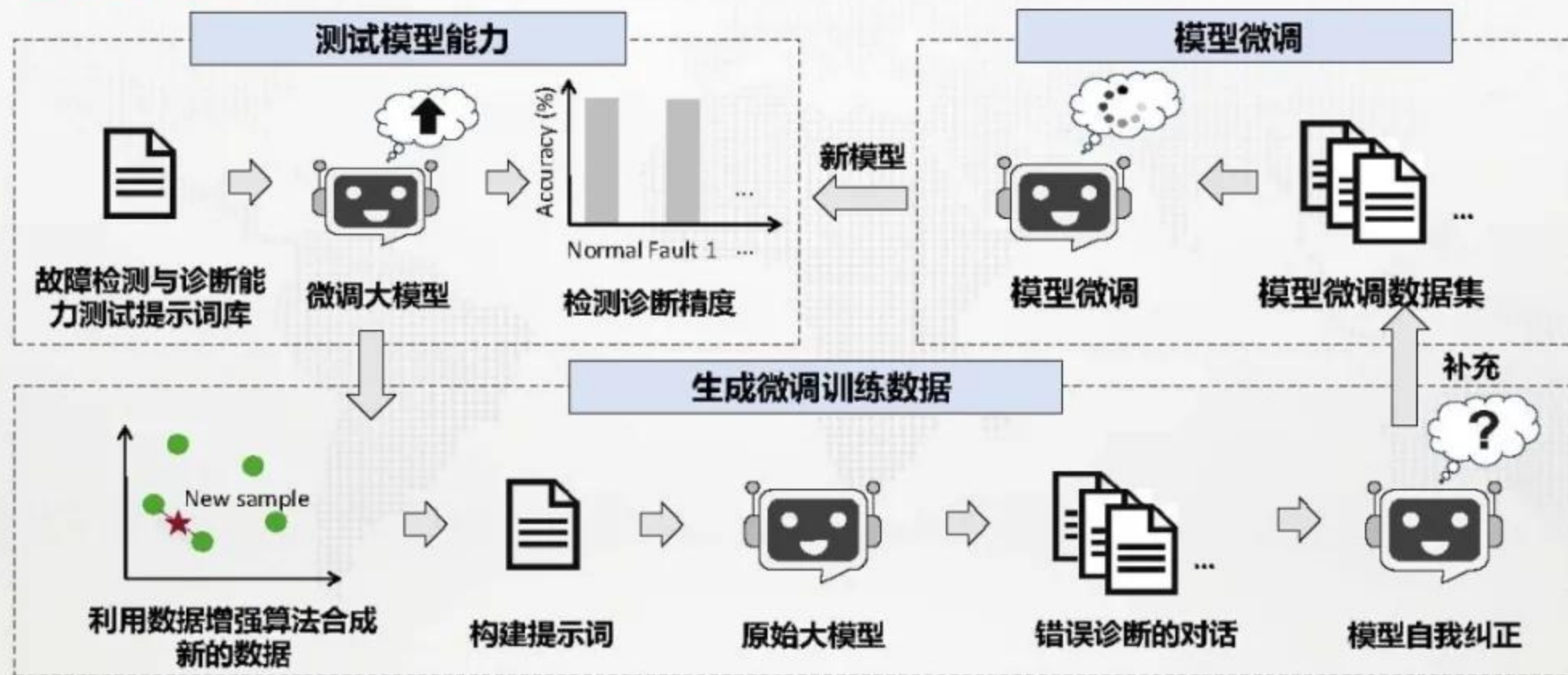
Insert the maximal frequent itemset into the slot [x7].

- 313 maximal frequent itemsets are extracted by the knowledge discovery method.
- GPT-3.5 and GPT-4 are utilized to analyze the 313 maximal frequent itemsets.
- The detection accuracy of GPT-4 is **89.17% for energy waste patterns** and **99.48% for normal operation patterns**.

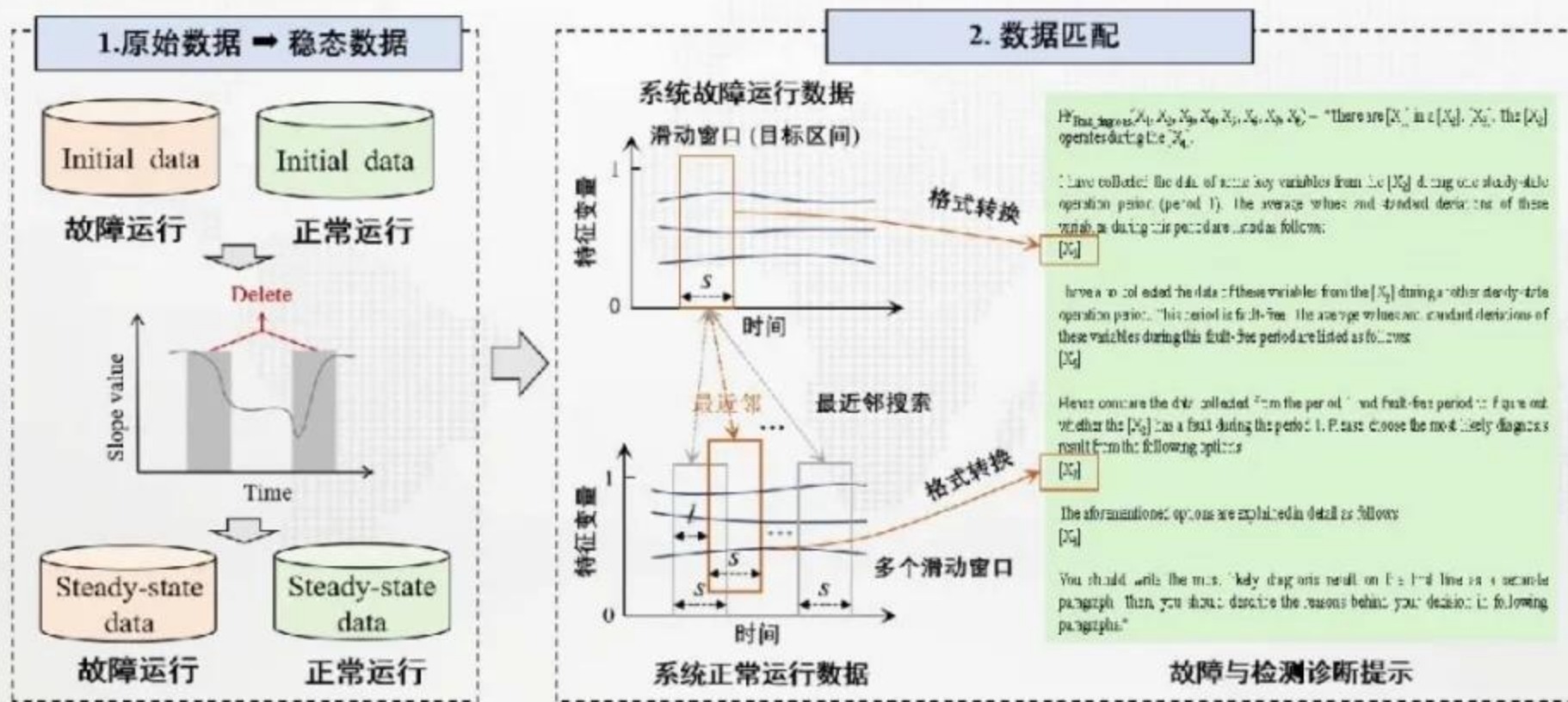


Confusion matrixes of (a) GPT-3.5 and (b) GPT-4.

故障检测诊断: 基于标注数据, 结合模型自纠正与数据增强算法, 实现大语言模型的**自动化监督微调训练**方法, 提高了LLMs在执行能源系统故障诊断任务中的准确性。



**第一步：对原始数据进行预处理，并基于室外温度、相对湿度等特征变量匹配与当前故障阶段最相似的无故障运行阶段，形成系统状态对比，构建数据形式提示。**



## 基于模板格式，自动填充相应内容，批量化生成用于大语言模型故障检测与诊断的提示词

$PF_{\text{Fault\_diagnosis}}(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8) =$  "There are  $[X_1]$  in a  $[X_2]$ .  $[X_3]$ . The  $[X_2]$  operates during the  $[X_4]$ .

I have collected the data of some key variables from the  $[X_2]$  during one steady-state operation period (period 1). The average values and standard deviations of these variables during this period are listed as follows:  
 $[X_5]$

I have also collected the data of these variables from the  $[X_2]$  during another steady-state operation period. This period is fault-free. The average values and standard deviations of these variables during this fault-free period are listed as follows:  
 $[X_6]$

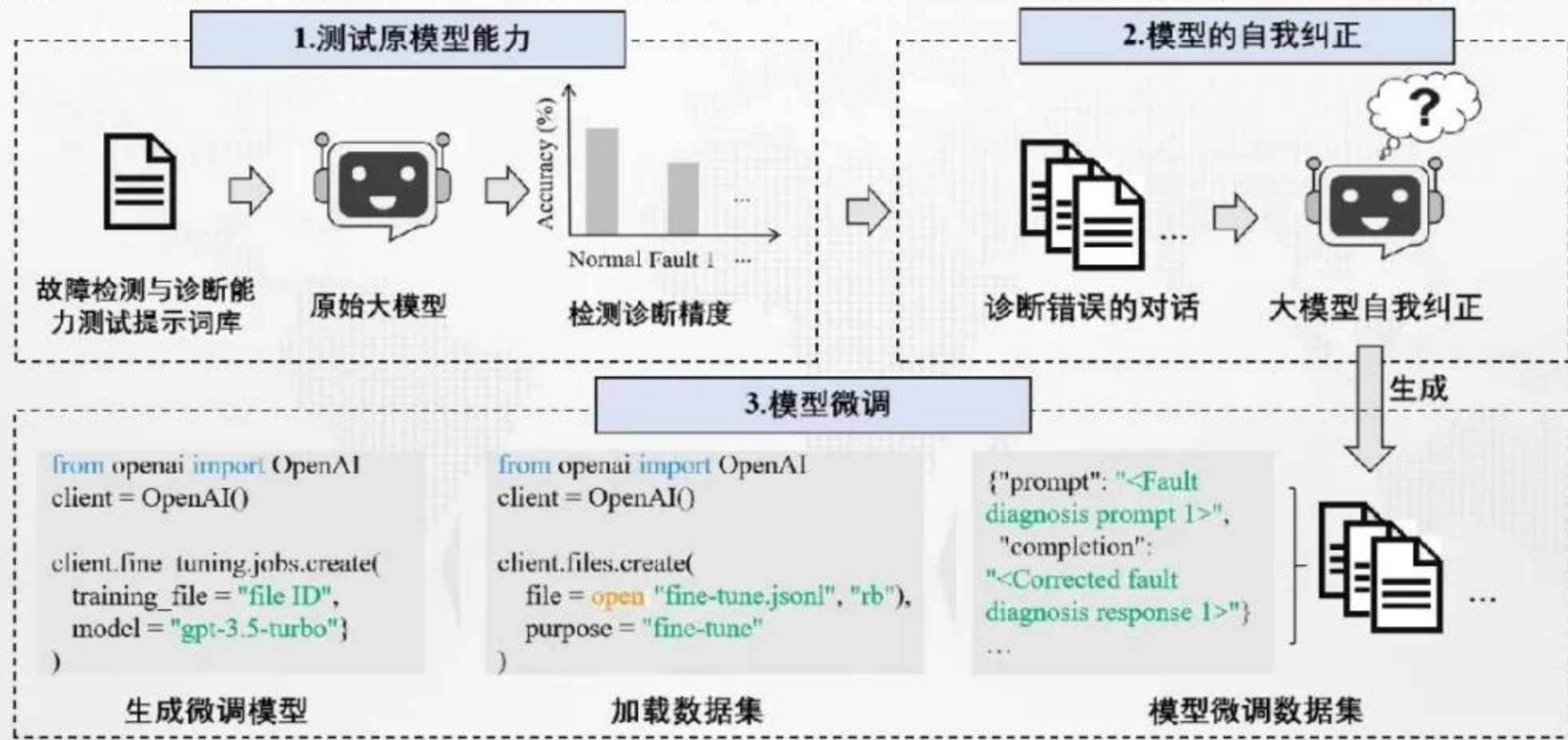
Please compare the data collected from the period 1 and fault-free period to figure out whether the  $[X_2]$  has a fault during the period 1. Please choose the most likely diagnosis result from the following options:  
 $[X_7]$

The aforementioned options are explained in detail as follows:  
 $[X_8]$

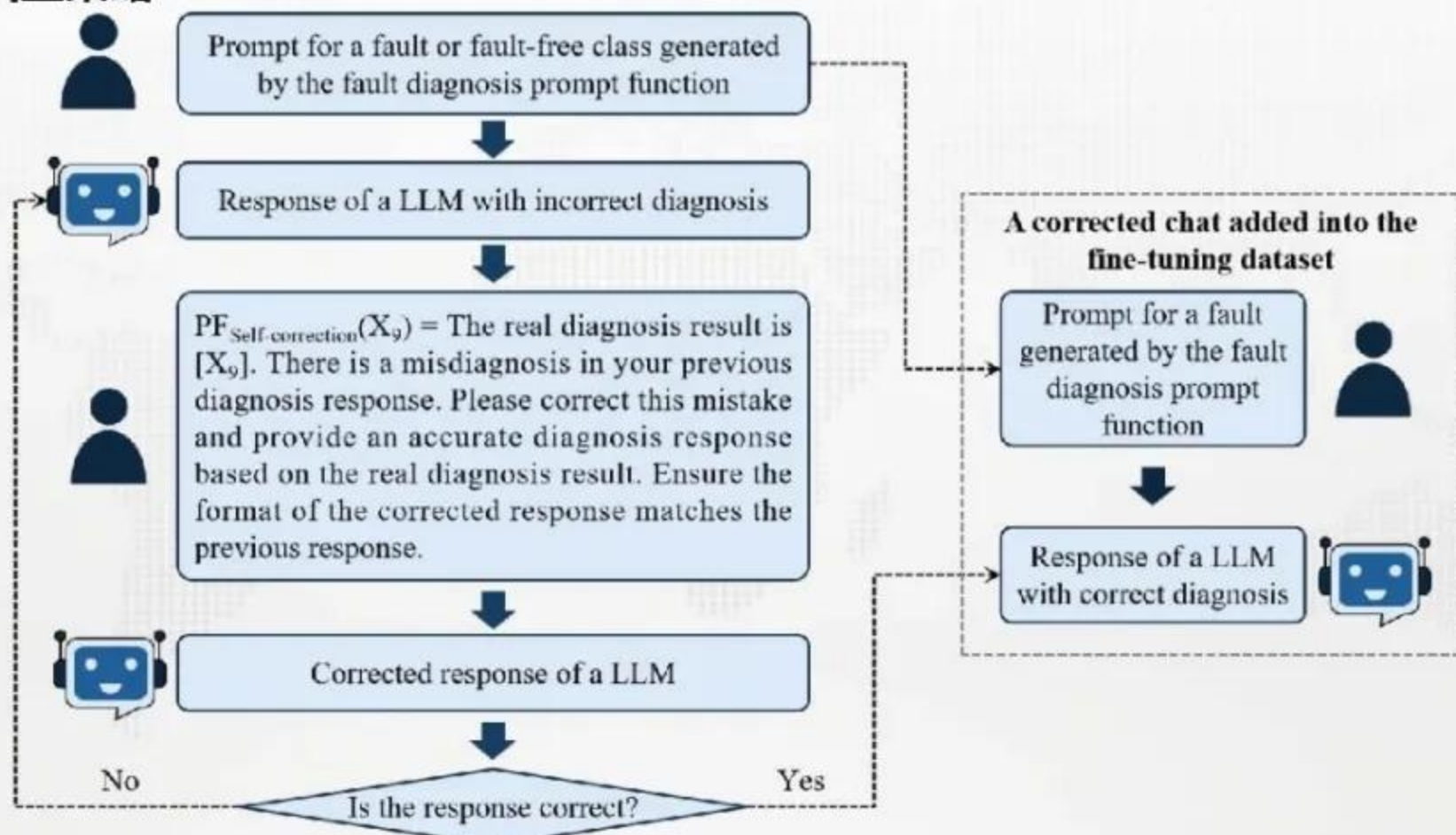
You should write the most likely diagnosis result on the first line as a separate paragraph. Then, you should describe the reasons behind your decision in following paragraphs."

NO.	待填充内容
X1	系统组成形式（设备、组件）
X2	系统的类型
X3	设备之间的关系（串联、并联...）
X4	系统的运行模式（制冷、制热...）
X5	系统关键数据变量在一段时间内的平均值和标准偏差
X6	系统关键数据变量在正常运行时的平均值和标准偏差
X7	系统可能存在的状态选项
X8	回答格式要求

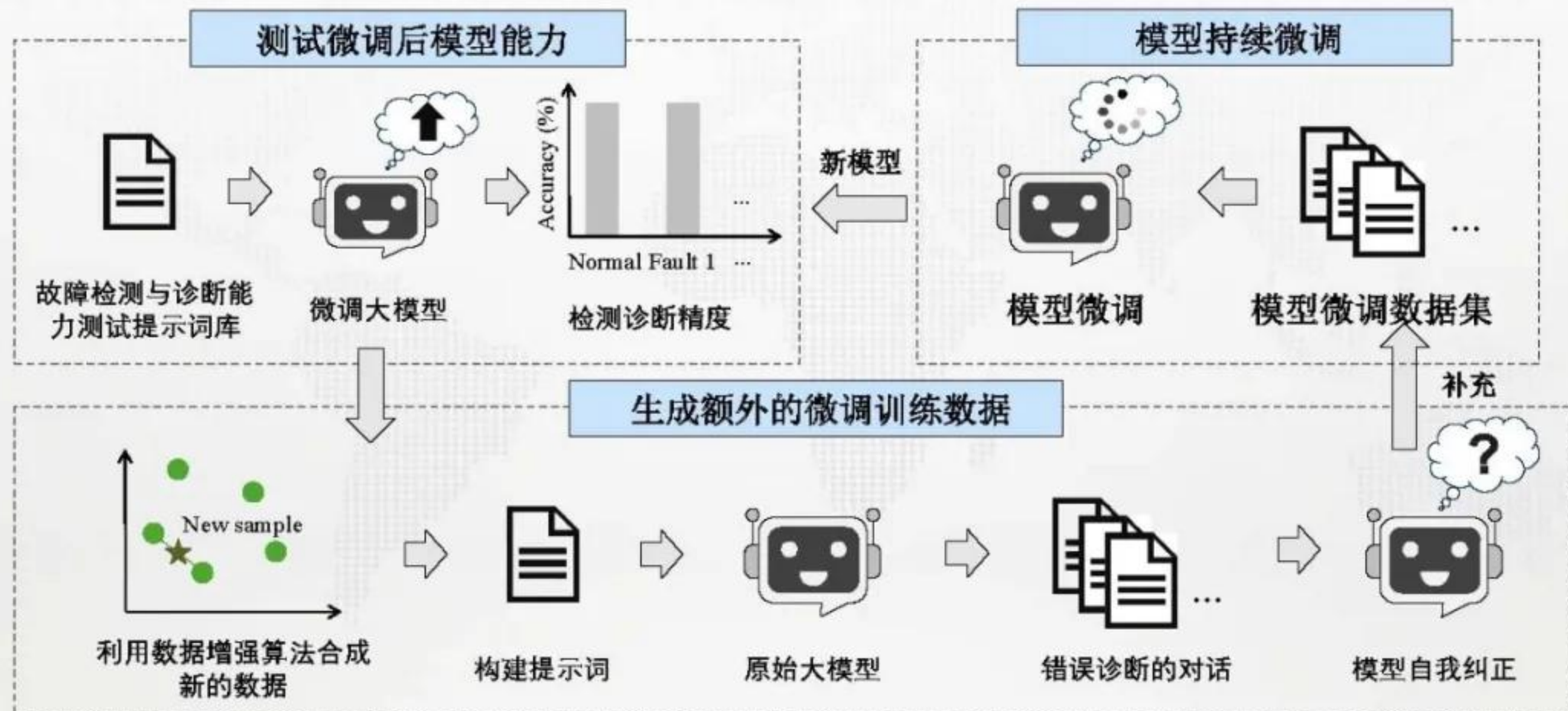
对原始模型进行能力测试, 结合标注信息使其自纠正错误诊断内容, 构建语料, 微调大模型



## 大模型的自我纠正策略:



评估微调后模型能力, 利用数据增强合成新的训练数据, 对模型偏弱的能力进行针对性提升。



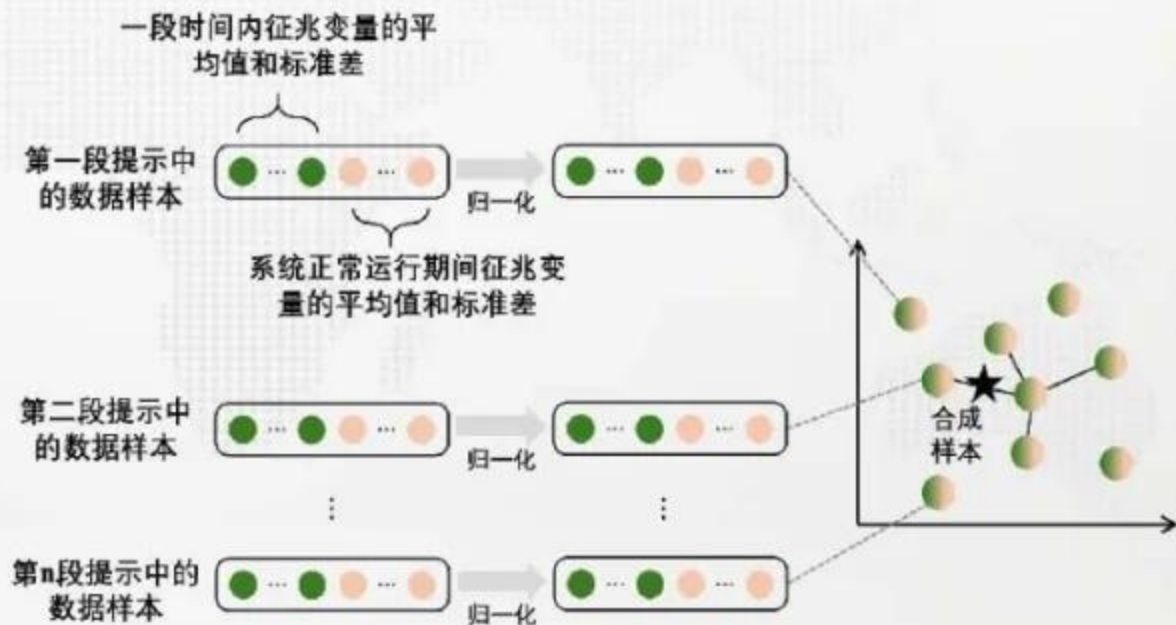
## 基于诊断准确率自适应调整每类样本需新增的提示量，并基于SMOTE算法合成新样本

□ 根据评估结果对**诊断准确率较低**的故障类别按照比例生成额外提示，以提高训练量

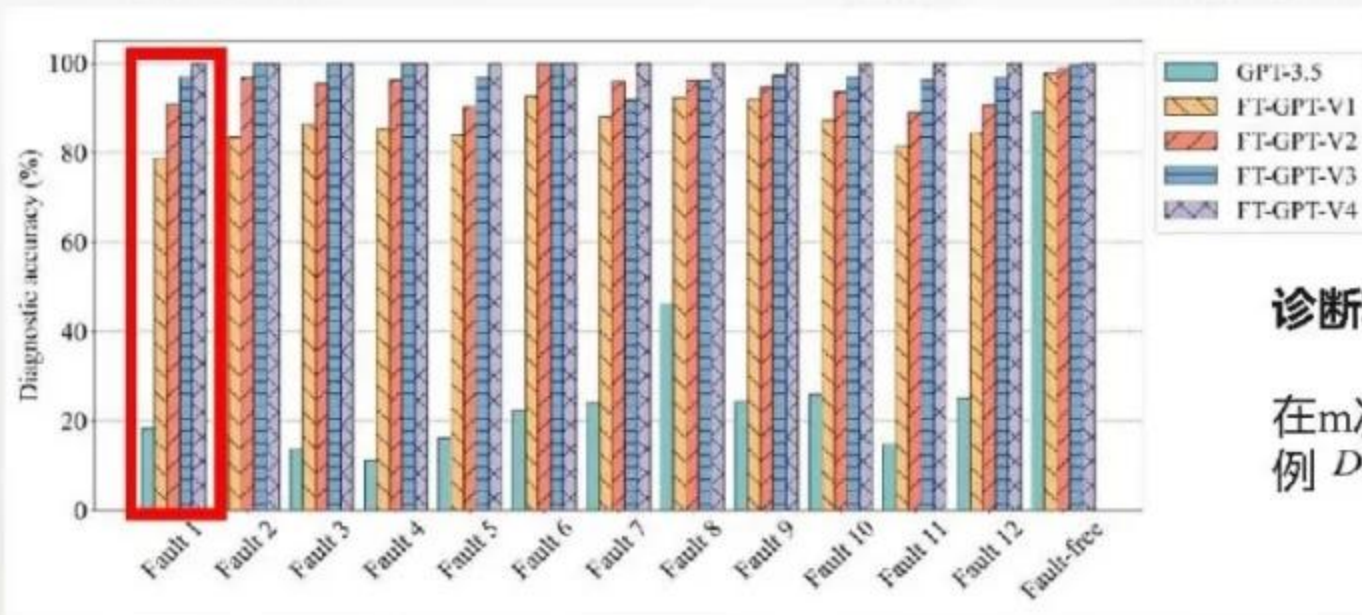
$$N_{\text{add}}^i = N_{\text{total}} \times r \times \frac{1 - P_i}{\sum_j^{N_3} (1 - P_j)}$$

- $N_{\text{add}}^i$  为第*i*个类生成的额外微调对话的数量
- $N_3$  故障类和无故障类的数量
- $N_{\text{total}}$  原始微调数据集中的对话数
- $r$  为需生成的新的提示量增加的百分比
- $P_i$  优化后的模型在第*i*类的诊断准确率
- $P_j$  优化后的模型在第*j*类的诊断准确率

□ 基于已有检测与诊断提示数据样本，**利用SMOTE算法**，合成每类别提示对应数量的新的提示样本。



- 原始GPT-3.5模型和四种微调GPT-3.5模型对RP-1312数据集中12种故障和无故障的最低诊断准确率分别为11.1 %、78.8 %、88.9 %、92.0 %和100.0 %，能够实现高精度的故障检测诊断。
- 最终的微调后模型较原GPT-3.5模型对13类别的平均诊断准确率提高了73.3 %。



### 诊断精度

在m次对话中模型正确诊断的比例  $DiA = \frac{\sum_{i=1}^m count_{i,j}}{m}$

2024-2025期间浙大EnergyX课题组共发表大语言模型相关研究SCI论文5篇，涉及大语言模型在建筑能源系统中的大数据分析、故障检测诊断、负荷预测等场景的应用与能力评估。

序号	题目	年份	期刊
1	Generative pre-trained transformers (GPT)-based automated data mining for building energy management: Advantages, limitations and the future	2024	Energy and Built Environment
2	Evaluation of large language models (LLMs) on the mastery of knowledge and skills in the heating, ventilation and air conditioning (HVAC) industry	2024	Energy and Built Environment
3	Automated data mining framework for building energy conservation aided by generative pre-trained transformers (GPT)	2024	Energy and Buildings
4	Domain-specific large language models for fault diagnosis of heating, ventilation, and air conditioning systems by labeled-data-supervised fine-tuning	2024	Applied Energy
5	Automated data-driven building energy load prediction method based on generative pre-trained transformers (GPT)	2025	Energy



本二维码文件永久有效  
微信扫码即可获取文件

博士生



鲁洁



章文恺



冯晨昕



张健



郑泽宇

硕士生



王嘉茜



贺佳宁



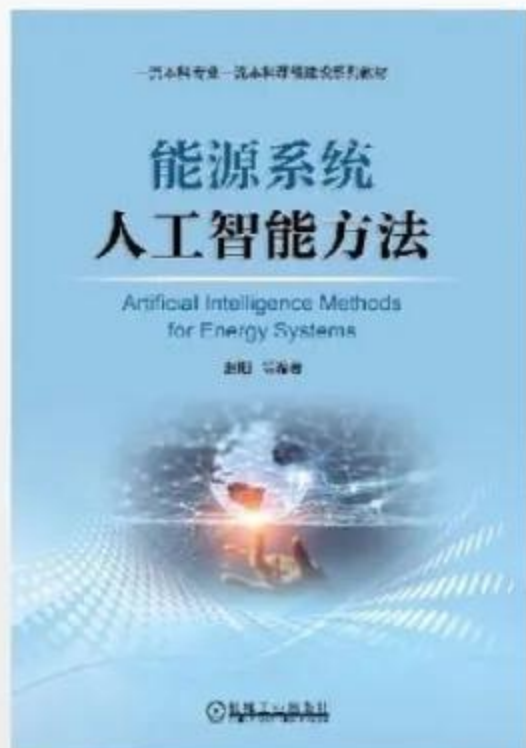
蒋睿勤

博士后



章超波

荷兰埃因霍温理工大学



机械出版社, 2023.12出版

课件地址:

<https://pan.baidu.com/s/1cuIBC3QKGuTyNhXfjMiJyA?pwd=az3g>

提取码: az3g

<input type="checkbox"/> 文件名	大小
<input type="checkbox"/>  第一章 绪论.pptx	7.4M
<input type="checkbox"/>  第五章 优化方法.pptx	12.2M
<input type="checkbox"/>  第四章 监督学习的方法下.pptx	7.7M
<input type="checkbox"/>  第四章 监督学习方法上.pptx	22.4M
<input type="checkbox"/>  第二章 无监督学习.pptx	10.4M
<input type="checkbox"/>  第三章 数据预处理.pptx	4M

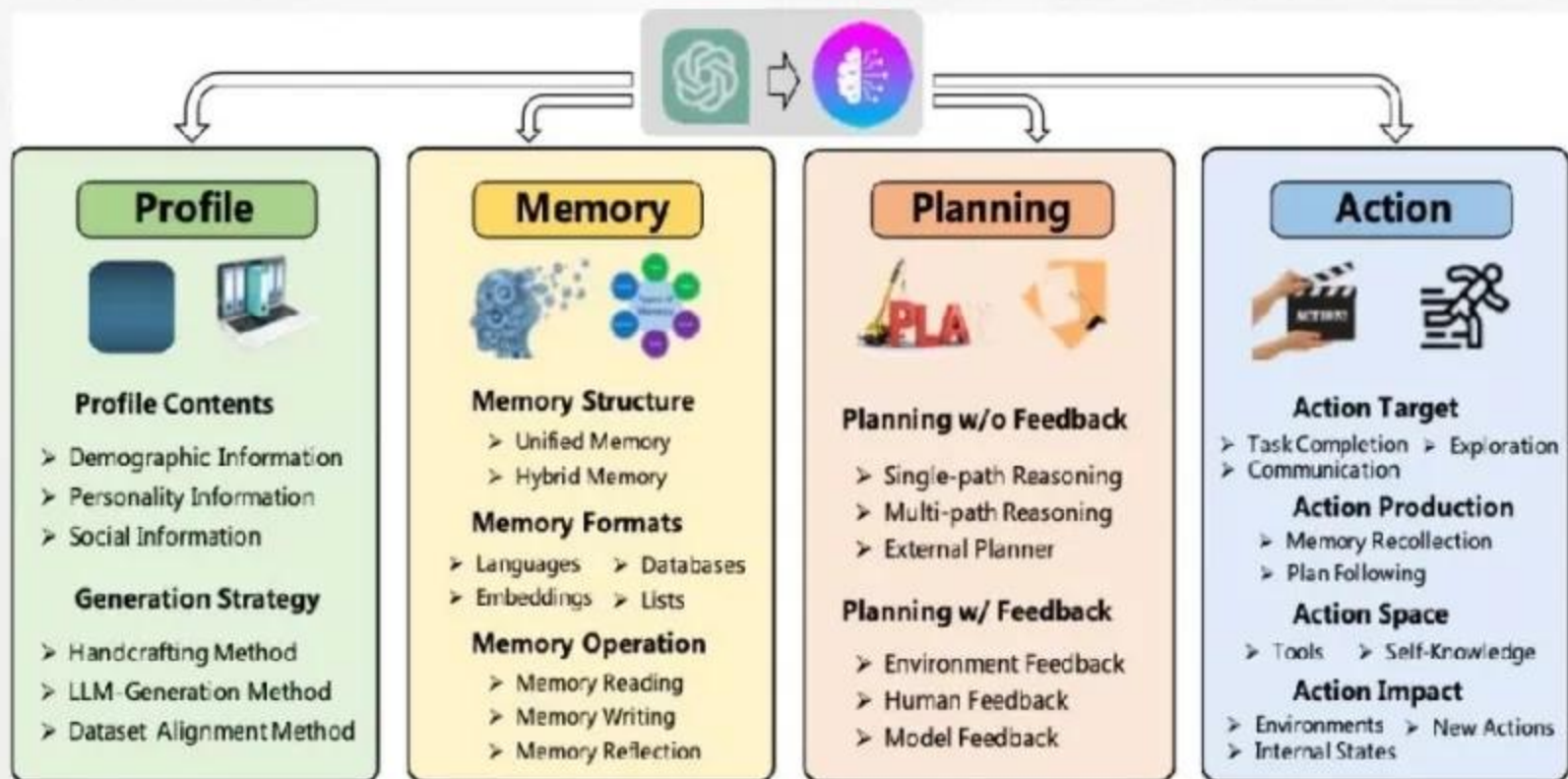
公众号EnergyX



# 报告提纲

- 1 当下AI到了哪种程度
- 2 能源领域传统AI发展困境
- 3 DeepSeek等带来的新范式
- 4 大语言模型应用的科研案例
- 5 临近奇点: AGI将带来颠覆
- 6 结论和展望

2025年2月10日, OpenAI 联合创始人兼首席执行官 Sam Altman 说:“这种 Agent 最终将能够胜任顶级公司中, 拥有数年经验的软件工程师所能完成的大部分工作, 处理时长不超过几天的任务。”

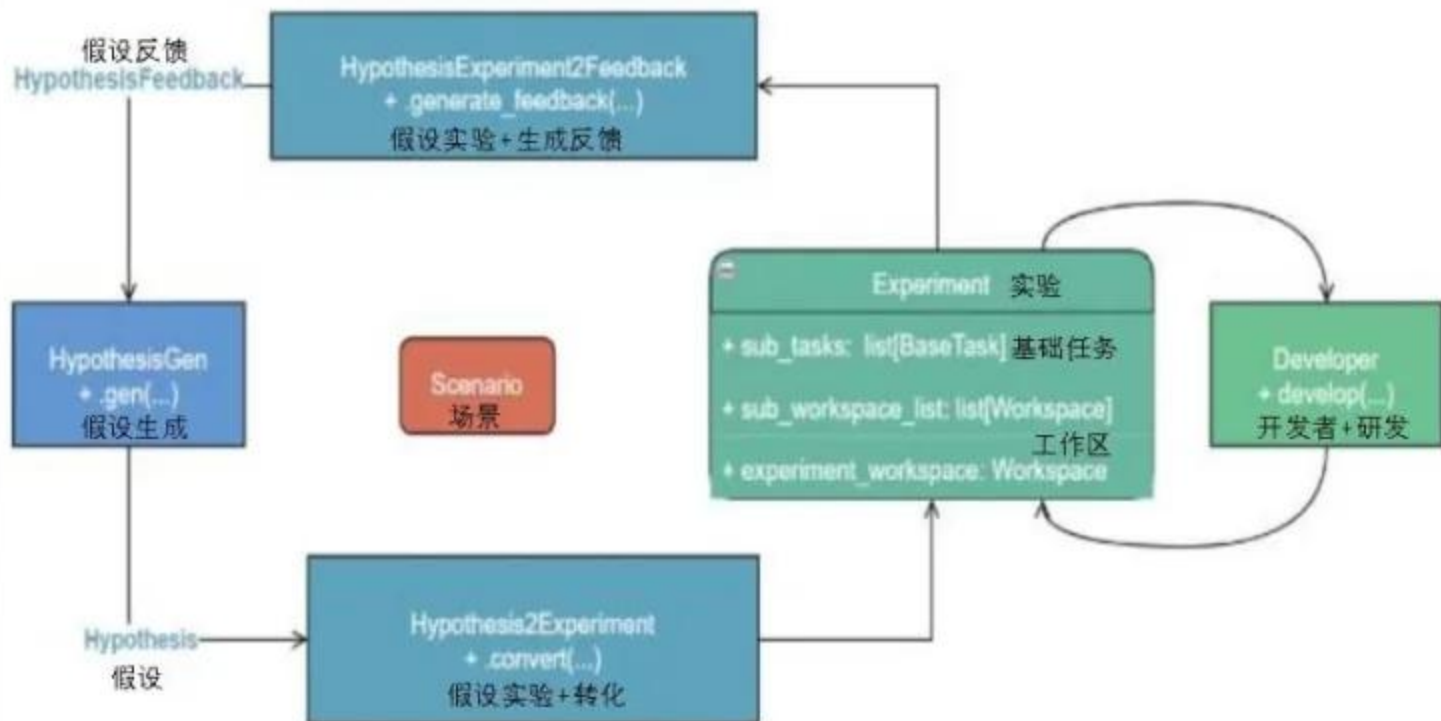


2024年11月28日, 微软开源RDAgent, 致力于打造一个工业研发过程的自动化助手, 专注于数据驱动场景, 通过提出新想法和实施它们来简化模型和数据的开发。



1. **创意与研究:** 从创意到研究提案
2. **假设与实验:** 形成假设进行实验
3. **能力开发:** 实现创意的能力
4. **反馈与实施:** 收集反馈, 工作空间实施
5. **应用场景:** 金融医疗等领域应用

从假设生成到实验反馈的各个环节, 每个步骤都是实验和开发流程中的关键环节, 确保从创意到实现的转化和优化。



**1.动态学习与适应:** Agent通过不断学习和适应, 优化决策过程, 提升问题解决能力。

**2.跨领域应用能力:** Agent能够灵活应用于金融、医疗等多个领域, 满足不同场景需求。

**3.集成开发与反馈:** Agent在集成环境中开发, 通过反馈循环实现持续迭代和性能提升。

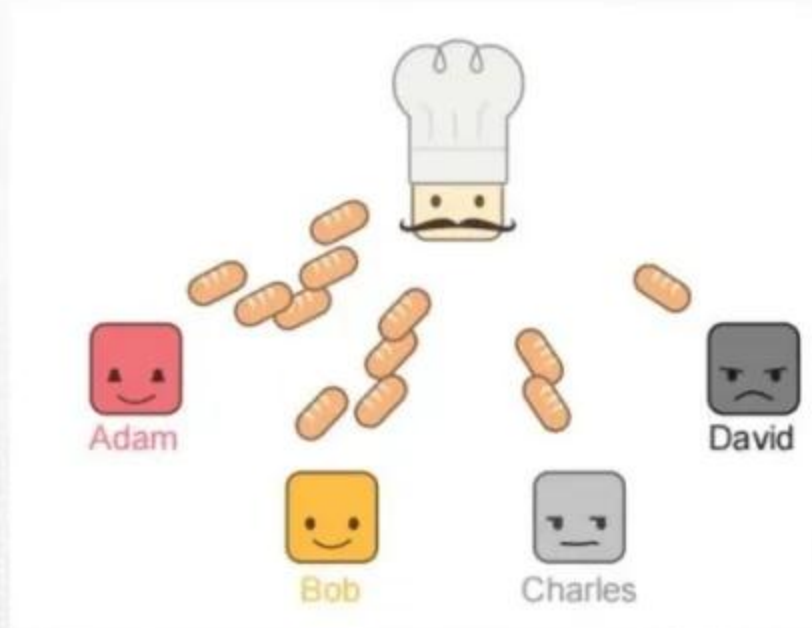
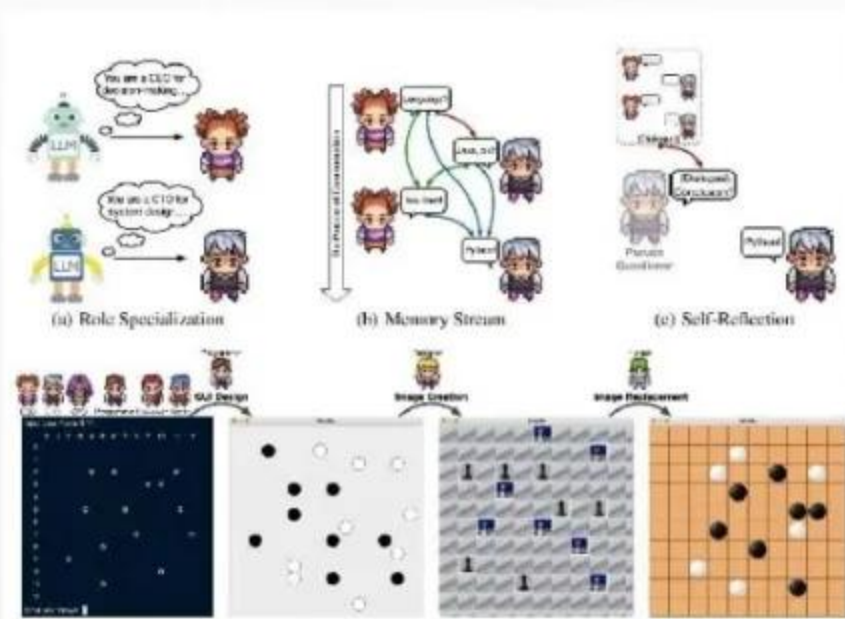


**1.信息提取:** 从报告和论文中提取研究方法和模型。

**2.方法实现:** 将提取的方法转化为可执行的代码和模型。

**3.结果分析:** 执行代码, 获取并分析研究结果, 为发展提供依据。

研发过程概述: 研究人员阅读论文和报告, 提取可实现的方法 (通常制定为数学公式或模型架构), 以寻求潜在的研究方向。然后, 他们准确地实施这些方法, 以获得进一步分析和开发的结果。



- **任务细化与分配:** 由GPT担任游戏公司内各个关键角色, 每个角色均被赋予明确且对应的职责。
- **协作与互动:** 各个角色之间建立高效的沟通机制, 通过交流, 提供必要的支持与反馈。
- **信息整合与决策制定:** 基于收集到的信息, 决策者进行总结与分析, 以形成全面的决策依据。
- **循环优化与迭代:** 流程采取循环迭代的方式进行, 确保问题得到持续的优化与解决。

AlphaGo Zero的成功不仅是算力+算法+数据, 还有数字化的棋盘!

2016年3月, AlphaGo以**4: 1**  
赢了人类围棋世界冠军



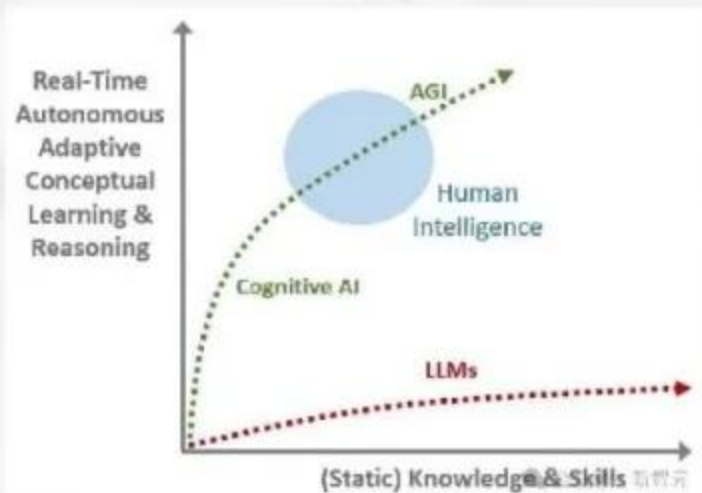
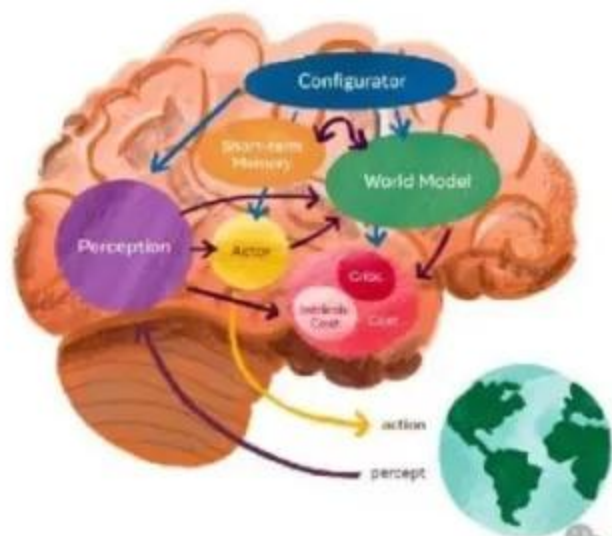
2017年10月, AlphaGo Zero在3天内  
自我对弈490万盘, 以**100:0**击败  
AlphaGo



一些学者坚信世界有一种「世界模型」，并着力开发一种新的类似大脑的AI架构，是通过更真实地模拟现实世界（多模态输入训练）来解决当前系统的局限性，例如幻觉和逻辑上的缺陷。这也是想要AI接近人类智力水平，需要像婴儿一样学习世界运作的方式。



2018图灵奖得主  
卷积神经网络创始人  
Yann LeCun



反馈: GPT输出的文字和文字答案之间的差距



**生命1.0**

反馈: 世界模型的反馈, 学到因果!

**迭代进化!**

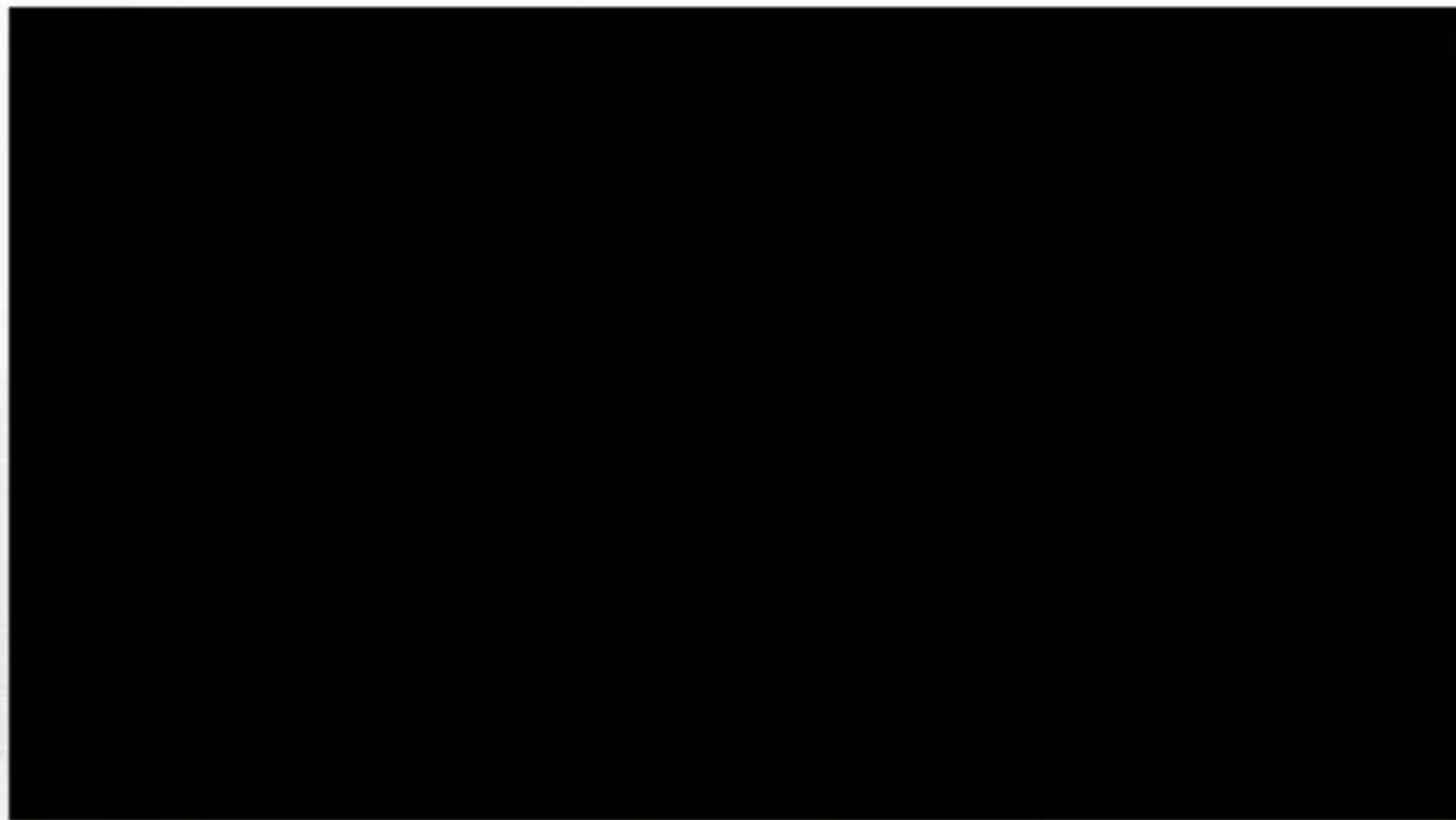


**生命2.0**

Prompt: Drone view of waves crashing against the rugged cliffs along Big Sur's garay point beach. The crashing blue waters create white-tipped waves, while the golden light of the setting sun illuminates the rocky shore. A small island with a lighthouse sits in the distance, and green shrubbery covers the cliff's edge. The steep drop from the road down to the beach is a dramatic feat, with the cliff's edges jutting out over the sea. This is a view that captures the raw beauty of the coast and the rugged landscape of the Pacific Coast Highway.

提示词：俯瞰无人机视角，海浪拍打着比克瑟尔加雷岬海滩的崎岖悬崖。湛蓝的海水激起白色的浪花，而落日的金色光辉照亮了岩石海岸。远处有一个带灯塔的小岛，悬崖边缘覆盖着绿色的灌木。公路到海滩之间的陡峭落差极具戏剧性，悬崖的边缘向海洋突出。画面捕捉了海岸线的原始之美以及太平洋沿岸公路崎岖壮丽的自然风光。







**这里不是人类文字!**



这里不是人类文字!

超智能机器可以设计出更好的机器, 毫无疑问, 这会带来 '智能爆炸', 而人类的智能将远远落后。因此, 第一台超智能机器将是人类最后需要发明的机器。

— 欧文·约翰·古德, 1965 年

反馈: 世界模型的反馈, 学到因果!



**完成进化闭环!**

## 平台简介

面向未来智能化需求, 历时八年花费千万, 开发新一代数字孪生平台、SCADA和边缘控制器

- 以知识图谱为核心, 贯通物理仿真、AI、故障诊断和群控, 代码量逾百万行
- 支持冷、热、气(汽)大规模机理和数据驱动混合仿真
- 能够自动生成节能控制策略, 并进一步自动生成SCADA群控软件, 节能性好, 且降低自控开发量90%以上

## 数字孪生云平台



领域知识  deepseek、数据驱动建模

故障诊断模块

AI建模 (脚本嵌入)

## 现场群控

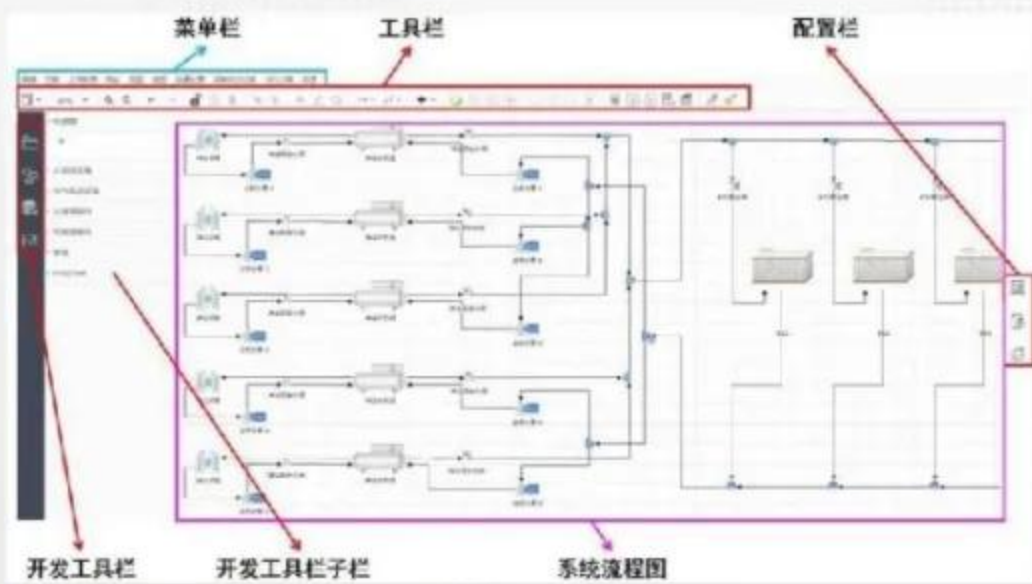


SCADA群控软件

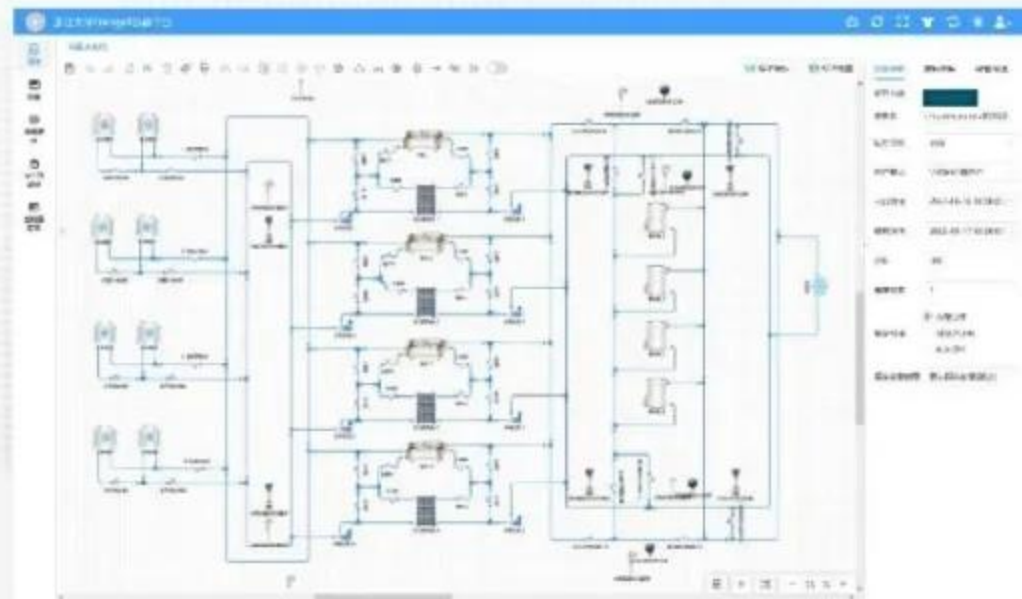


边缘控制器

- 需要深度工匠精神, 要有迭代思维, 年复一年升级迭代, 人才是关键
- 群控软件和仿真软件开发难度、周期和代价非常高



EnergyX 1.0版本



EnergyX 2.0版本

- 赵阳博士有**十年**国内外带软件开发团队经验, 具有出色软件架构能力
- 长期在开发一线, **有死磕的工匠精神、永不放弃的热情、有核心团队**

赵阳博士亲自架构和开发90%内核代码, 基于Rust语言



第一代: 2018



第二代: 2021



第三代: 2022



第四代: 2023



第五代: 2024

<b>从技术出发找场景</b>	<b>不合理的期望管理</b>
<b>大模型团队不了解业务</b>	<b>缺乏清晰的产品/技术路线</b>
<b>技术方案过度设计</b>	<b>缺乏对技术发展的认知</b>
<b>沉迷于模型训练</b>	<b>高估项目的成功率</b>
<b>低估评估的重要性</b>	<b>忽略微调的隐藏成本</b>

# 报告提纲

- 1 当下AI到了哪种程度
- 2 能源领域传统AI发展困境
- 3 DeepSeek等带来的新范式
- 4 大语言模型应用的科研案例
- 5 临近奇点: AGI将带来颠覆
- 6 未来展望和结论

## ■ AI领域各路大咖预测: 5-10年内, 人类将迎来AGI的时代。

2024年11月9日



OpenAI CEO  
Sam Altman

「通往AGI只差最后一步,  
2025年实现AGI」

- 通往AGI的道路现已清晰可见;
- 当智能和能源极为充足时, 物理学中的所有难题都将得到解决。

2023年6月15日



NVIDIA CEO  
黄仁勋

「5年实现AGI,  
10年算力提高100万倍」

- 未来的AI将会自我生成数据来进行强化学习;
- 要让AI完美地生成图像与视频, 必须要让它基于物理学, 创建一个世界模型。



DeepMind CEO  
Demis Hassabis

「我们距离AGI  
还有大约10年的时间」

- 未来的AI系统需要更主动、更智能, 能够像AlphaGo一样进行规划和推理, 并在现实世界中采取行动。

2024年10月21日



2018图灵奖得主  
卷积神经网络创始人  
Yann LeCun

「AGI将在5到10年降临,  
JEPARoute将代替LLM」

- LeCun坚信有一种世界模型, 并着力开发一种类似大脑的AI架构, 通过模拟现实世界来解决幻觉和逻辑上的缺陷。

~2024年11月29日

2025年2月10日, OpenAI 联合创始人兼首席执行官 Sam Altman 在其社交媒体发布重磅文章《Three Observations》



我们持续见证着人工智能开发的飞速进展。以下是我们对人工智能经济学的三个观察：

1. 人工智能模型的智能水平，大致与训练和运行它所消耗的资源量（取对数后）成正比。这些资源主要包括训练所需的算力、数据和推理算力。看起来，只要持续投入资金，就能获得持续且可预测的智能提升；
2. 预测这一趋势的规模定律，在多个数量级上都得到了验证。使用同等水平人工智能的成本，大约每 12 个月下降 10 倍。价格降低显著促进了使用量的增加。从 2023 年初 GPT-4 的 token 成本，到 2024 年年中 GPT-4o 的 token 成本变化中，我们可以看到这一点——期间每 token 的价格下降了约 150 倍。摩尔定律曾以每 18 个月翻倍的速度改变世界；而人工智能成本下降的速度，比这还要快得不可思议。
3. 智能水平的线性提升，其社会经济价值的增长本质上是超指数级的。由此推断，我们认为在可预见的未来，没有理由停止对人工智能进行指数级增长的投资。



**任何在 2035 年的人，都应能调动相当于 2025 年全体人类的智力资源。**

每个人都应有权获取无限的“天才”力量，并能随心所欲地加以运用。如今，世界上存在着大量才华横溢却缺乏充分发展资源的群体。如果我们能改变这种状况，由此释放出的创造力，将为我们所有人带来巨大的福祉。



- 随着通用人工智能变得更有能力，额外风险的证据正在逐渐出现。这些风险包括大规模劳动力市场影响、人工智能黑客攻击或生物攻击，以及社会失去对通用人工智能的控制。专家们对这些风险的现有证据有不同的解读：一些人认为这些风险还需要几十年的时间，而另一些人则认为通用人工智能可能在未来几年内导致社会规模的危害。通用人工智能能力的最新进展——特别是在科学推理和编程测试方面——为人工智能支持的黑客攻击和生物攻击等潜在风险提供了新的证据，导致一家主要的人工智能公司将其对生物风险的评估从最佳模型的“低”提高到“中”。

As general-purpose AI becomes more capable, evidence of additional risks is gradually emerging. These include risks such as large-scale labour market impacts, AI-enabled hacking or biological attacks, and society losing control over general-purpose AI. Experts interpret the existing evidence on these risks differently: some think that such risks are decades away, while others think that general-purpose AI could lead to societal-scale harm within the next few years. Recent advances in general-purpose AI capabilities – particularly in tests of scientific reasoning and programming – have generated new evidence for potential risks such as AI-enabled hacking and biological attacks, leading one major AI company to increase its assessment of biological risk from its best model from ‘low’ to ‘medium’.

## ■ AGI将带来人类历史上最深远的变革

### ■ AGI将10-1000倍提高生产效率, 包括开发出能源领域专用人工智能的效率(个人观点)

#### ➢ 提升效率与创造力

AGI能够快速解决复杂问题, 推动科研创新  
创意生成时间被大大压缩, 缺乏判断力易被淘汰!

#### ➢ 经济结构快速调整

自动化可能导致大量传统职业消失, 尤其是基础性和中等技能岗位, 对社会经济带来冲击。

#### ➢ 解放劳动力

AGI可以解放人类从事重复性、危险性、高强度的工作, 使人类专注于更有创造力或更具意义的活动, 解放劳动力。

#### ➢ 伦理和数据安全问题

在军事、金融和医疗等关键领域, AGI的应用如果没有充分的监管, 会带来巨大的安全隐患。

#### ➢ 社会公平与教育普及

AGI可能弥补教育资源不均问题, 为偏远地区或欠发达地区提供个性化学习、医疗和公共服务。

#### ➢ 市场竞争与商业模式

领先企业可能因率先采用AGI技术而获得不可替代的竞争优势, 重新定义市场格局, 而落后的企业可能面临淘汰。



AGI  
+  
世界模型

- 人工智能是新一轮科技革命和产业变革的核心驱动力。加快发展新一代人工智能技术，将决定国家在未来科技与产业领域的战略地位
- 制造业的数字化、网络化、智能化转型，将全面提升研发设计、生产制造、经营管理与运维服务的效率与质量。企业通过智能化升级，增强综合竞争力，抢占未来产业发展制高点
- 每一次科技与产业革命都形成了符合当时需求的新质生产力。历史证明，技术革命催生了全新的产业格局，推动生产方式的优化与系统性重构

AI 是一场通用技术革命，预计将对各行各业带来深远的影响，推动全社会生产力和竞争力的提升！

- 人工智能是能源领域一种新的生产力，正在重新定义行业工作方式和技術路径
- 通过大模型与智能体的结合，能源行业从数据收集到优化决策的智能化
- 大模型的强大推理能力为复杂问题提供精准分析，而智能体的灵活适配性保证了任务的高效执行，能够大幅提升了行业智能程度
- 通过实时数据的循环优化和反馈机制，人工智能持续推动能源体系的数字化与智能化变革，为未来的可持续发展提供了强大动力

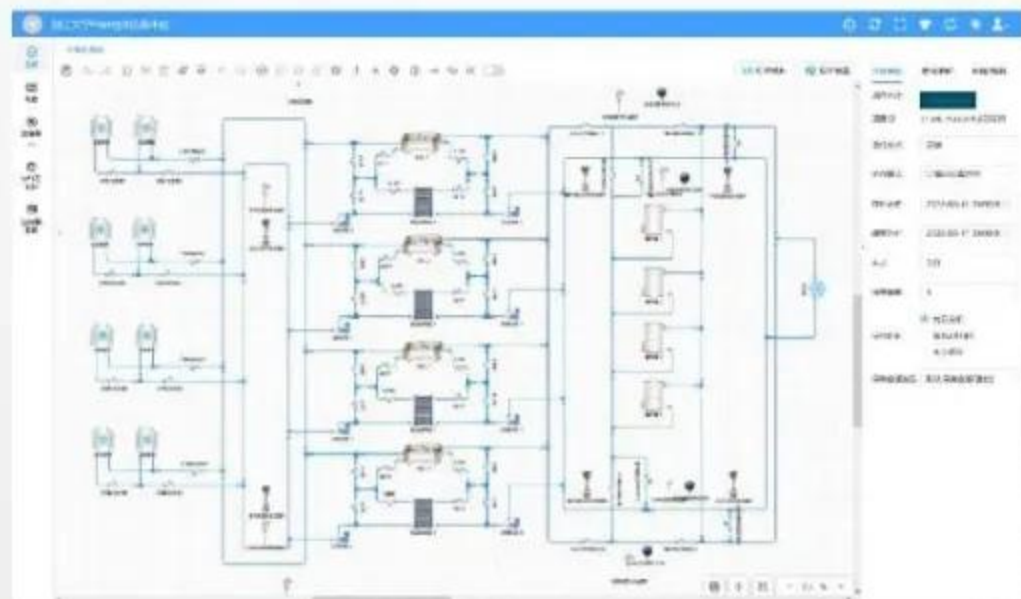
- 对于即将到来的 AGI, 无论是 OpenAI 还是 DeepSeek, 以及这个世界, 都还没有做好充分准备
- 这些人工智能在数量和质量上都将超越人类, 会以前所未有的速度处理信息, 并生成超出人类理解的新解决方案。
- 有可能在短短的 10 年内, 人工智能就会复杂和超级智能到我们甚至无法理解它是如何运作的地步。
- 和届时的 AGI 相比, 我们将会像幼儿园的孩子想努力拿到一个数学博士学位。

**人们不仅容易高估未来1-2年的趋势,**

**而且更容易低估未来10年的趋势!**

- 组织机构和社会都存在巨大的**思维惯性和发展惯性**
- 阶跃式的突变发生时, **真理往往掌握在少数人手里**
- 当看不懂的时候, **先去做**, 和做得明白的人/单位做朋友, 慢慢就懂了

**看不见、看不起、看不懂、来不及**



EnergyX 平台

公众号



**敬请各位专家指正!**



**汇报人：赵阳**

**浙江大学能源学院制冷与低温研究所**  
**2025.2.14**



赵阳 博士、研究员

浙江大学能源工程学院

## 研究方向：复杂热力系统数字孪生、故障检测诊断、智能控制

- 中组部“千人计划”青年专家
- 浙江省特聘专家, 中控集团智库专家
- 连续三年入选斯坦福全球顶尖2%科学家“年度影响力”榜单
- Energy and Built Environment, Building Simulation编委
- 主持国家重点研发中国芬兰合作能效诊断项目1项
- 主持中国荷兰自然科学基金委国际合作智慧能源项目1项
- 主持热力系统故障诊断相关自然科学基金2项
- 主持浙江省尖兵领雁”复杂场景高端工业运维软件”重点项目
- 华为数据中心故障诊断算法及代码开发项目

能源领域第一本  
人工智能教材

### 相关研究成果

专利成果：获得授权国家发明专利28件，登记软件著作权1项

论文成果：发表SCI论文83篇, 迄今所有论文被SCI引用5000余次, H-index 24

- 一作/通讯作者论文篇均被引47次
- ESI高被引论文3篇 (工程领域前1%), 1篇被引400次以上, 16篇被引100次以上