

软件工程报告

Lab1 软件调研-国产大模型评测

计算机学院 计算机科学与技术

2212452 孟启轩

一、要求

从编程能力、文本生成、知识推理、角色扮演等角度，比较至少3个国产大模型，列出它们的优缺点和特色，并对它们进行排名。

二、选取四个国产大模型

- DeepSeek
- 文心一言
- 通义千问
- Kimi

三、大模型评测

(1) 模型概述

DeepSeek

模型架构：采用混合专家模型（MoE）架构，通过动态路由机制提升特定任务的响应精度，在长文本理解和多轮对话中表现突出。

训练数据侧重：训练数据包含大量学术论文、代码库（如 GitHub）、专业文献，在 STEM（科学、技术、工程、数学）领域回答更具深度。

文心一言

模型架构：基于 Transformer 架构，拥有庞大的参数规模和深度的神经网络结构，能够处理复杂的语言任务，如文本生成、问答、翻译等。

训练数据侧重：训练数据涵盖了大量的文本数据，包括书籍、论文、新闻、网页等，对中文语境有深入的理解和优化，能够生成符合中文表达习惯的高质量文本。

通义千问

模型架构：基于 Transformer 的改进架构，强化多模态融合能力（文本、图像、视频），与阿里云基础设施深度集成。

训练数据侧重：数据融合电商交易记录、多语言商业文档，对中文语境下的市场趋势和商业需求有更深入的理解和优化。

Kimi

模型架构： Kimi 采用了最新的 AI 技术，具备高度的中文和英文对话能力，同时在安全性和准确性上有所突破。其模型架构注重对话的流畅性和内容的准确性。

训练方式： Kimi 的训练方式注重对用户指令的精准理解和执行，以及在特定领域的深度定制化服务。通过对大量对话数据的学习和优化，Kimi 能够更好地理解用户的意图，提供更加准确和有用的回答。

(2) 模型评测

1.编程能力

(1) 代码补全

测试模型是否能够自动补全代码。

- 给你一个整数数组 nums，返回数组中最大数和最小数的最大公约数。两个数的最大公约数是能够被两个数整除的最大正整数。

```
class Solution {  
public:  
    int findGCD(vector<int>& nums) {  
  
    }  
};
```

结果及分析

四个大模型补全的代码都能正确运行，并且结果正确。可见对于简单的代码补全任务，四个大模型的处理能力都比较理想。

</> 代码

C++ ▾ 智能模式

```
1  class Solution {
2  public:
3      int findGCD(vector<int>& nums) {
4          int max_val = nums[0];
5          int min_val = nums[0];
6          // 找出最大值和最小值
7          for (int num : nums) {
8              if (num > max_val) {
9                  max_val = num;
10             }
11             if (num < min_val) {
12                 min_val = num;
13             }
14         }
15         // 计算最大公约数
16         int a = max_val;
17         int b = min_val;
18         while (b != 0) {
19             int temp = b;
20             b = a % b;
21             a = temp;
22         }
23         return a;
24     }
25 };
```

已存储

☒ 测试用例 | >_ 测试结果**通过** 执行用时: 0 ms

- Case 1
- Case 2
- Case 3

(2) 代码生成

测试模型是否能从零编写可运行的代码。

- 问题1：用Python实现一个简单的Web服务器，监听8080端口，并返回'Hello, World!'。

用Python实现一个简单的Web服务器，监听8080端口，并返回'Hello, World!'



已深度思考 (用时 35 秒)

以下是使用Python内置库实现的简单Web服务器：

```
python
from http.server import HTTPServer, BaseHTTPRequestHandler

class SimpleHTTPRequestHandler(BaseHTTPRequestHandler):
    def do_GET(self):
        # 设置响应状态码
        self.send_response(200)

        # 设置响应头
        self.send_header('Content-type', 'text/html')
        self.end_headers()

        # 设置响应内容
        self.wfile.write(b'Hello, World!')

def run_server():
    # 服务器配置
    server_address = ('', 8080) # 空字符串表示监听所有可用接口
    httpd = HTTPServer(server_address, SimpleHTTPRequestHandler)

    print(f"服务器已启动，访问地址：http://localhost:{server_address[1]}")
    try:
        httpd.serve_forever()
    except KeyboardInterrupt:
```

- 问题2：请编写一个Python 多线程爬虫，它可以并发抓取多个网页的 HTML 内容，并将数据存入本地文件。代码要求：使用 requests 获取网页内容（支持 User-Agent 伪装）。使用 ThreadPoolExecutor 并发抓取多个网址，提高爬取速度。使用 BeautifulSoup 解析 HTML（比如提取网页标题）。将爬取的数据存入 data.json 文件，以 JSON 格式存储 URL 和标题。

深度思考 请编写一个Python 多线程爬虫，它可以并发抓取多个网页的 HTML 内容，并将数据存入本地文件。
代码要求：
使用 requests 获取网页内容（支持 User-Agent 伪装）。
使用 ThreadPoolExecutor 并发抓取多个网址，提高爬取速度。
使用 BeautifulSoup 解析 HTML（比如提取网页标题）。
将爬取的数据存入 data.json 文件，以 JSON 格式存储 URL 和标题。



已完成深度思考 (用时14秒)

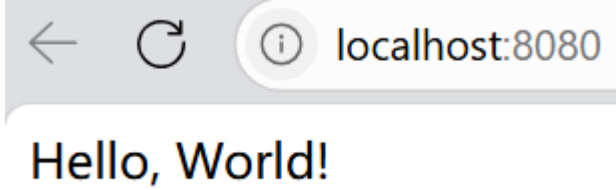
以下是使用 Python 实现的多线程爬虫示例，满足你的所有要求：

```
Python
import json
import requests
from concurrent.futures import ThreadPoolExecutor
from bs4 import BeautifulSoup

class WebCrawler:
    def __init__(self, urls, max_workers=5):
        self.urls = urls
        self.max_workers = max_workers
        self.user_agent = 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/1
        self.results = []
```

结果及分析

- 对于问题1：四个大模型给出的代码都能正确运行并且结果正确。



- 对于问题2：DeepSeek和通义千问给出的代码能够正确运行，并且爬取到的网页结果也符合要求。但是，文心一言和Kimi给出的代码有误，并不能直接运行。

```
[Running] python -u "f:\study\软件工程\大模型调研\test.py"
获取 https://www.stackoverflow.com 时出错: 403 Client Error: Forbidden
获取 https://www.pandas.pydata.org 时出错: HTTPConnectionPool(host
NameResolutionError("<urllib3.connection.HTTPSConnection object at
failed"))
Traceback (most recent call last):
  File "f:\study\软件工程\大模型调研\test.py", line 74, in <module>
    main()
  File "f:\study\软件工程\大模型调研\test.py", line 60, in main
    for future in concurrent.futures.as_completed(future_to_url):
                        ^^^^^^^^^^^^^
NameError: name 'concurrent' is not defined
```

由此对比，对于简单任务，四个大模型的处理都比较理想。但对于较为复杂的任务，DeepSeek和通义千问的处理能力较强，并且相关的一些注意事项也给出了，而文心一言和Kimi的处理能力较弱，并不能一次完成对于任务的正确分析处理。

(3) 代码问答

测试模型是否能够回答编程相关的问题。

- 这个函数的时间复杂度是多少？

```
def factorial(n):
    return 1 if n == 0 else n * factorial(n - 1)
```

结果及分析

def factorial(n):
 return 1 if n == 0 else n * factorial(n - 1)

这个函数的时间复杂度是多少？

已完成推理

分析递归阶乘时间复杂度

这个递归阶乘函数的时间复杂度是 $O(n)$ 。

分析过程：

- 基本情况：**当 $n == 0$ 时，函数直接返回 1，这是一个常数时间操作，记为 $O(1)$ 。
- 递归情况：**当 $n > 0$ 时，函数会递归调用 `factorial(n - 1)`，然后将结果乘以 n 。每次递归调用都会将问题规模减小 1。
- 递归深度：**递归调用的深度为 n ，因为从 n 开始，每次减 1，直到达到基本情况 $n = 0$ 。
- 每层操作：**在每一层递归中，除了递归调用本身，其他操作（如乘法和减法）都是常数时间，记为 $O(1)$ 。

总结：

递归阶乘函数的时间复杂度可以表示为：

$$T(n) = T(n - 1) + O(1)$$

展开后为：

$$T(n) = O(1) + O(1) + \dots + O(1) \text{ (共 } n \text{ 次)}$$

因此，总的时间复杂度是 $O(n)$ 。

复制

再试一次

分享

👍👎

这个函数的空间复杂度是多少？>

有没有非递归的阶乘算法？>

四个大模型都能够回答正确，并且给出了时间复杂度。可见对于代码问答任务，四个大模型的处理能力都比较理想。此外，DeepSeek和通义千问也给出了空间复杂度和其他的相关分析，文心一言和Kimi仅仅回答了该问题，不过Kimi和通义千问在回答下方有用户可能接下来要问的问题（如上图所示），用户直接点击即可再次提问。

(5) 注释生成

测试模型是否能够合理生成注释，帮助理解核心代码。

- 一个完整的对于双向循环链表的插入、合并和拆分的程序。请大模型给出核心逻辑的注释。

结果及分析

文心 X1

```
return 0;
}
```

帮我代码的核心逻辑部分写注释，帮助我理解



思考和行动过程

```
cpp
1  #include<iostream>
2  using namespace std;
3
4  // 双向循环链表节点模板类
5  template <typename T>
6  class node
7  {
8  public:
9      node<T>* next;    // 指向下一个节点的指针
10     node<T>* last;    // 指向前一个节点的指针
11     int data;         // 节点存储的数据
12
13     // 构造函数：初始化节点数据，并设置前后指针
14     node(int p, node<T>* head, node<T>* la)
15     {
16         data = p;
17         next = head;    // 初始时next指向头节点
18         last = la;     // 初始时last指向前一个节点（在插入时会调整）
19     }
20 };
```

四个大模型给出的注释都很合理，都很好地解释了源程序。DeepSeek和文心一言还举了一个例子进行程序执行流程的演示，通义千问则给出了一些问题和优化建议，而Kimi只回答了问题，其他的相关问题需要用户再进行询问。

2.文本生成

(1) 改写润色

测试模型是否能够有效地改写、润色文本。

因此，在信息时代与生物技术时代，新闻工作者面临很多挑战。传统新闻工作者缺乏人工智能和生物技术相关的知识和技能，难以适应快速变化的技术环境。信息发达的今天，网络不是法外之地，如何确保数据隐私，信息的正确性，以及伦理道德问题。在涉及复杂的报道时，信息技术的应用可能导致公众对新闻的真实性产生怀疑。当然，挑战与机遇并存。利用AI技术，新闻机构可以根据用户兴趣和行为提供个性化的新闻推荐，提高用户参与度和满意度。AI还可以自动化处理大量的数据分析和内容生成工作，使新闻工作者有更多时间进行深度报道和一线调查。虚拟现实、增强现实等技术的应用，为新闻报道提供了全新的形式和体验，可能会开辟“新形式新闻”。通过信息技术，新闻机构可以更便捷地进行全球信息采集和合作，共享资源和数据，使新闻更国际化。

请帮我润色这段话，优化句子流畅度、增强逻辑性、提升表达质量。

结果及分析

深度思考 因此，在信息时代与生物技术时代，新闻工作者面临很多挑战。传统新闻工作者缺乏人工智能和生物技术相关的知识和技能，难以适应快速变化的技术环境。信息发达的今天，网络不是法外之地，如何确保数据隐私，信息的正确性，以及伦理道德问题。在涉及复杂的报道时，信息技术的应用可能导致公众对新闻的真实性产生怀疑。当然，挑战与机遇并存。利用AI技术，新闻机构可以根据用户兴趣和行为提供个性化的新闻推荐，提高用户参与度和满意度。AI还可以自动化处理大量的数据分析和内容生成工作，使新闻工作者有更多时间进行深度报道和一线调查。虚拟现实、增强现实等技术的应用，为新闻报道提供了全新的形式和体验，可能会开辟“新形式新闻”。通过信息技术，新闻机构可以更便捷地进行全球信息采集和合作，共享资源和数据，使新闻更国际化。

请帮我润色这段话，优化句子流畅度、增强逻辑性、提升表达质量。



已完成深度思考（用时7秒）

以下是润色后的版本，在保持原意的基础上优化了逻辑结构、语言流畅度和表达的专业性：

在信息时代与生物技术时代交汇的当下，新闻工作者正面临前所未有的挑战与机遇。首先，技术变革对传统新闻从业者的专业能力提出了更高要求。由于缺乏人工智能、生物技术等领域的专业知识与技能，部分从业者难以适应技术驱动的新闻生产模式，尤其在处理涉及基因编辑、算法伦理等跨学科议题时，容易陷入信息壁垒。其次，数字技术的普及带来了新的伦理困境：网络空间并非法外之地，如何在海量数据中保障隐私安全、确保信息准确性、维护伦理道德，成为新闻机构必须直面的课题。此外，技术应用可能引发公众对新闻真实性的质疑——当深度伪造（Deepfake）、自动化内容生成等技术被滥用时，信息的真实性与可信度面临严峻考验。

然而，技术革新同样为新闻行业开辟了突破性的发展路径。人工智能技术的深度应用，使新闻机构能够通过用户画像与行为分析实现精准的内容推荐，从而提升用户参与度与满意度。同时，AI驱动的数据分析与自动化内容生产工具，显著减轻了记者的重复性劳动，使其得以将更多精力投入深度调查与独家报道。更具革命性的是，虚拟现实（VR）、增强现实（AR）等技术正重塑新闻呈现形式，通过沉浸式报道让用户“身临其境”，这或将催生“元新闻”（Metajournalism）等全新新闻形态。此外，数字化协作平台的普及，使全球新闻机构能够突破地域限制，实现资源与数据的实时共享，推动新闻报

深度思考 基于QwQ推理模型，支持自动联网搜索

四个大模型都给出了改写后的文本，并且都能提升句子的流畅度、增强逻辑性、提升表达质量。但是DeepSeek和文心一言改变了文章结构，加入了一些小标题，尤其是DeepSeek，当然也可能是我给的段落专业性比较强。


(2) 翻译

测试模型是否能够将文本翻译成目标语言。


Plastic surgery makes people more confident and gives them advantages. Not all people are born beautiful. Plastic surgery can change an ugly duckling into a beautiful swan. Good-looking people often have more opportunities. Plastic surgery can not only make people more beautiful but can also give them a competitive advantage. But plastic surgery cannot guarantee success and involves high risks.

翻译这段话

结果及分析


 Plastic surgery makes people more confident and gives them advantages. Not all people are born beautiful. Plastic surgery can change an ugly duckling into a beautiful swan. Good-looking people often have more opportunities. Plastic surgery can not only make people more beautiful but can also give them a competitive advantage. But plastic surgery cannot guarantee success and involves high risks.


翻译这段话





思考和行动过程 


整形手术能够重塑自信并创造竞争优势。先天容貌的局限性使得并非所有人都能拥有理想外貌，而医疗整形技术恰似现代版的“丑小鸭变天鹅”童话，为求美者提供了改写基因剧本的可能。在颜值经济时代，出众的容貌往往能转化为职场机遇、社交红利乃至心理优势，这种对美学资本的投入正被越来越多人视为值得尝试的人生策略。然而需要清醒认识到，手术刀下的美丽蜕变并非万能通行证——任何医疗行为都伴随着不可预知的风险，从麻醉意外到术后感染，从效果未达预期到心理落差，整形手术的风险系数始终如影随形。在追求外貌升级的同时，理性评估自身需求、审慎选择医疗机构、充分了解潜在并发症，才是对生命美学负责任的态度。

由文心  生成

 重新生成









四个大模型都给出了翻译后的文本，并且都能将文本翻译成目标语言。但是通义千问和Kimi的翻译结果不够准确，太过于简洁，并且翻译后的文本不够流畅，比较生硬。而DeepSeek和文心一言的翻译更加生动丰富，更贴合原文。

(3) 总结摘要

测试模型能否准确提取文章核心内容。

我将下面这则新闻交给大模型进行总结与摘要。

走失二十多年的谢浩男回家！“悬赏千万寻子”父亲举办迎亲宴[<https://news.qq.com/rain/a/20250316A02XR500>]

结果及分析

写在最后：愿天下无拐，愿所有亲人终得团圆

谢爸爸的坚持让我们看到，亲情的力量足以跨越时间，打破命运的不公。然而，还有无数个家庭在煎熬中等待，他们的故事同样值得被关注。

愿所有失散的孩子都能早日回家，愿天下再无“谢爸爸”的悲剧上演，愿每一位父母都能牵起孩子的手，一生相伴，永不分离。

请帮我生成上述文章的总结摘要



- ✓ 已完成推理
- ✓ 生成谢爸爸寻子故事总结

总结摘要：

文章讲述了深圳商人谢岳（谢爸爸）的儿子谢浩男在1998年被拐，24年后父子终于团圆的故事。谢浩男2岁时在深圳白石洲被附近卖水果的流动商贩拐走，人贩子将他藏匿三天后带至河南。谢爸爸为寻子卖掉房子创业，历经24年终于在警方帮助下找回儿子。重逢时，父子均穿紫色衣服，谢浩男已成长为双一流高校研究生及科研人才，谢爸爸则事业有成。文章还提到两人相似的外貌特征及重逢后的温馨场景，同时呼吁社会关注打拐寻亲，强化法律打击、利用科技助力寻亲并鼓励社会支持，希望天下无拐、亲人团圆。



四个大模型都给出了总结与摘要，并且都能准确提取文章核心内容。但是DeepSeek、文心一言和通义千问的摘要较为生硬，更像一则学术报告，而Kimi的摘要则比较贴合实际。

3.知识推理

(1) 问题解答

测试模型是否能够回答知识与生活中的问题。

- 问题1：如何求 $y=Ae^x$ 的反函数（A是正常数）
- 问题2：如果我晚上睡得很晚，但第二天又必须早起，怎样才能在短时间内恢复精力？

结果及分析

如何求 $y=Ae^x$ 的反函数 (A是正常数)



已深度思考 (用时 134 秒)

首先，我们从原函数 $y = Ae^x$ 出发，其中 A 是一个正常数。

1. 解出 x 表示为 y 的函数：

$$y = Ae^x$$

两边同时除以 A ：

$$\frac{y}{A} = e^x$$

取自然对数：

$$\ln\left(\frac{y}{A}\right) = x$$

因此，得到 $x = \ln\left(\frac{y}{A}\right)$ 。

2. 交换变量 x 和 y ：

将 x 和 y 交换，得到反函数：

$$y = \ln\left(\frac{x}{A}\right)$$

3. 验证反函数的正确性：

- 原函数在 $x = 0$ 时， $y = A$ 。反函数在 $x = A$ 时， $y = \ln\left(\frac{A}{A}\right) = \ln(1) = 0$ ，正确。
- 原函数在 $x = 1$ 时， $y = Ae$ 。反函数在 $x = Ae$ 时， $y = \ln\left(\frac{Ae}{A}\right) = \ln(e) = 1$ ，正确。

通过验证，确认反函数的正确性。

最终答案：

$$y = \ln\left(\frac{x}{A}\right)$$



如果我晚上睡得很晚，但第二天又必须早起，怎样才能在短时间内恢复精力？



思考和行动过程

当面临短期睡眠不足的紧急情况时，可通过以下科学策略快速恢复精力，同时尽量减少对健康的影响：

1. 黄金20分钟「碎片补觉」

- 原理：**美国睡眠基金会研究表明，20分钟小睡能提升34%警觉度，超过40分钟则可能进入深度睡眠导致醒后更疲惫。
- 操作：**设置20分钟闹钟，选择完全黑暗环境，用白噪音机屏蔽干扰，避免进入深度睡眠阶段。

2. 生物钟重启法

- 光照疗法：**晨起后立即用10000勒克斯光照灯照射视网膜20分钟，抑制褪黑素分泌，激活交感神经。
- 冷水刺激：**用15°C冷水洗脸30秒，通过潜水反射刺激交感神经，提升心率和代谢率。

3. 精准咖啡因摄入

- **剂量控制**：按体重计算摄入量（建议2-4mg/kg），例如60kg成人饮用1-2杯美式咖啡。
- **黄金时段**：在醒来后90分钟内摄入，此时皮质醇水平较低，咖啡因效果最佳。

4. 应急营养补充

- **快速供能组合**：香蕉（快速碳水）+杏仁（健康脂肪）+黑巧克力（黄烷醇）三重组合，既能稳定血糖又提升认知表现。
- **电解质补充**：饮用含钠、钾的椰子水，缓解脱水引起的疲劳感。

5. 微运动激活

- **5分钟HIIT**：进行1分钟高抬腿+40秒休息，重复3组，促进内啡肽分泌。

深度思考已开启。输入你的问题，帮你深度分析和解答

四个大模型都给出了合理的答案。这个方面看不出多大区别。

(2) 内容分析

测试模型是否能够对文本进行内容分析。

- 我给了大模型一则俄乌停火的分析，让大模型总结观点。

结果及分析

- **美国的战略考量**：通过将乌克兰纳入关键矿产交易框架，美国可能试图将乌克兰的经济安全与自身利益绑定，同时减少直接军事承诺。
- **泽连斯基的困境**：在特朗普政策反复无常的背景下，泽连斯基需在施压俄罗斯与维护自身利益间谨慎平衡，未来可能面临更多外交挑战。

核心矛盾与争议点

- **短期停火 vs 长期安全**：乌克兰寻求明确军事保障，但美国更倾向通过经济合作间接实现安全目标。
- **欧洲角色的缺失**：乌克兰希望欧洲参与和平进程，但美国未明确支持，可能削弱协议的国际合法性。
- **特朗普的不可预测性**：其对乌政策的反复（如暂停援助、重启合作）增加了局势的不确定性，泽连斯基需持续应对。

总结

文章揭示了美乌在沙特会谈中达成的停火提议及联合声明，但强调了协议的模糊性、双方关系的复杂性，以及未来谈判中需解决的关键问题。核心矛盾在于美国试图通过经济手段而非军事承诺推动和平，而乌克兰则需在压力下应对特朗普的政策不确定性，同时争取国际支持。最终，和平进程能否持续取决于俄罗斯的反应及美乌能否弥合战略分歧。



四个大模型给出的分析结果都贴合原文观点，并且都给出了更深层次的思考，分析了俄乌美三方各自的目的与压力，比较全面。

(3) 逻辑推理

测试模型进行逻辑推理的准确性与合理性。

- 三个小伙子同时爱上了一个姑娘，为了决定他们谁能娶这个姑娘，他们决定用手枪进行一次决斗。小李的命中率是30%，小黄比他好些，命中率是50%，最出色的枪手是小林，他从不失误，命中率是100%。由于这个显而易见的事实，为公平起见，他们决定按这样的顺序：小李先开枪，小黄第二，小林最后。然后这样循环，直到他们只剩一个人。那么这三个人中谁活下来的机会最大呢？他们都应该采取什么样的策略？ 正确答案应该是小黄。

结果及分析

- 小黄的生存路径：
 - 小李击杀小林后，对决小李（53.85%生存）。
 - 小黄击杀小林后，对决小李（53.85%生存）。
 - 总生存概率 $\approx 35\%$ 。
- 小林的生存路径：
 - 仅在小李、小黄均未击中后存活，随后击杀小黄并与小李对决（70%生存）。
 - 总生存概率 $\approx 24.5\%$ 。

结论：

小李活下来的机会最大（约40.5%），其次是黄（35%），最后是林（24.5%）。

最优策略总结：

- 小李：第一枪必须射击小林，后续继续攻击存活的最大威胁。
- 小黄：若小林存活，优先射击小林；若小林死亡，攻击小李。
- 小林：每轮直接击杀当前威胁最大的对手（优先小黄，其次小李）。

小李



四个大模型都给出了推测，各自的推测也都比较合理，至少能够自圆其说，但只有Kimi的推测是正确的。

4.联网搜索

对于比较新的消息进行联网搜索。

- 请你进行联网搜索，帮我分析一下最新款帕萨特和最新款迈腾的优缺点，对比一下买哪一个更好，我预算充足。

结果及分析

14 / 18

通义

对话

效率

智能体

笔记库

建议：选装静音胎、动力总成终身质保（4S店提到的购车福利）。

• 预算充足且偏好操控与质感：迈腾380TSI顶配

原因：底盘更运动，设计更显豪华，适合注重驾驶体验的用户。

5. 注意事项

1. 试驾必做：两车驾驶质感差异明显，尤其是底盘调校和动力响应。

2. 配置选装：帕萨特Pro版需选装关键配置（如哈曼卡顿），迈腾则需高配车型才有全面功能。

3. 保养成本：德系车保养费用较高，需预留预算（参考知识库“后期保养成本高”）。

总结

• 帕萨特：配置全能、适合家庭与商务，适合追求“一步到位”的用户。

• 迈腾：操控扎实、设计更显稳重，适合注重驾驶质感与传统德系风格的用户。

最终选择取决于您更看重配置丰富度还是驾驶体验！

参考来源 · 10 >

分享

帕萨特**与**迈腾**的性能如何？

帕萨特**与**迈腾**的燃油经济性如何？

帕萨特**与**迈腾**的售后网络如何？

深度思考

基于QwQ推理模型，支持自动联网搜索

千事不决问通义

服务生成的所有内容均由人工智能模型生成，其生成内容的准确性和完整性无法保证，不代表我们的态度或观点

参考来源 · 10

1 【大众帕萨特2025款怎么样】新...
前排还是坐在车子后排给人的感觉都是比较舒服的，乘坐空间坐的很舒服，部空间却没有想象中的那么紧凑，后备箱也有...
太平洋汽车网

2 2025帕萨特避坑必看！车主血泪...
哈曼卡顿音响+后排头枕音响：同级独有，长途旅行听歌堪比音乐会现场。座椅通风：南方车主必选！实测夏季特温速...
百家号

3 【图】帕萨特2025款优点_口碑...
【最满意点】1. 众帕萨特 满 箱 能跑 燃 经济 2. 众帕萨特 观 看 起 观 感 熟 又 失 暴 满 满 感 3. 众帕萨特 同 非 常 宽 敞 日 常 使...
汽车之家

4 新年开门红！帕萨特2025出...
新年伊始，上汽大众帕萨特家族便交出了一份令人瞩目的成绩单。据乘联分会公布的最新销量数据显示，2025年1月，上汽大众...
<https://m.163.com/dy/article/IJQER...>

5 上汽大众帕萨特2025出...
你以为十五万只能买家用代步车？那就大错特错了！最近，上汽大众帕萨特家族推出的2025出...
<https://mparticle.uc.cn/article.html?...>

Python多线程爬虫抓取HTML并存储

敞舒适，后排腿部空间表现优异。

动力性能

• 帕萨特：动力强劲，油门调校相对激进，起步加速迅猛，发动机转速提升快，变速箱响应迅速，能带来较强的推背感，但持续行驶时动力略显疲软。

• 迈腾：动力输出线性，加速过程沉稳，持续性好，低速响应也不错，驾驶感受流畅，中后段加速动力可观，更注重稳定性和舒适性。

配置丰富度

• 帕萨特：配置丰富实用，如360°全景影像、无线充电、智能泊车辅助系统等高科技配置，提升了驾驶的便捷性和安全性。

• 迈腾：配置也较为丰富，如15英寸中控触控屏、11.6英寸副驾娱乐屏、10.25英寸全液晶仪表盘、哈曼卡顿音响系统等，科技感强，但部分配置如座椅记忆功能等不如帕萨特。

性价比

• 帕萨特：在同等配置下，价格更具竞争力，终端优惠力度较大，消费者能以较低价格获得配置丰富、品质出色的德系B级车。

• 迈腾：价格相对坚挺，在性价比方面略逊一筹。

综合来看，帕萨特在外观豪华感、内饰质感、配置丰富度及性价比方面更具优势，适合注重豪华体验和性价比的消费者；迈腾则在操控稳定性、动力线性度和内饰科技感方面表现更好，适合追求驾驶乐趣和科技氛围的消费者。

复制

再试一次

分享

帕萨特和迈腾的油耗表现如何？

帕萨特和迈腾的售后服务怎么样？

迈腾的科技配置有哪些亮点？

输入您的问题，帮你深度解答

k1.5 长思考

联网搜索

内容由 AI 大模型生成，请仔细甄别

© 2025 北京月之暗面科技有限公司 | 京ICP备2023011302号-3 | 京B2-20240852 | 京公网安备11010802043150号

网页搜索

自动展开

全网最低价 限时抢购

更多图片

2024/08/19

图2024款帕萨特优势解析 - 汽车之家论坛

总的来说，2024款帕萨特在外观设计、动力性能、内饰舒适、配置功能、安全性能等方面都表现出色。虽然售价略高，但性价比仍然较高。如果你正在寻找一款时尚 ...

2023/10/30

大众帕萨特优缺点 - 网通社

驾驶感受动力很足哦，2.0T，操控也可以，指哪打哪，就是过街道时要注意车速，哈哈哈。carplay+倾斜的中控屏真的绝配，对了，座椅通风很有用，免得屁股出汗，小风很 ...

【帕萨特怎么样】帕萨特怎么样_缺点_优点_口碑_帕萨特值得买吗

外观：外观方面我选的星空，但是开了一个多月以后觉得不如经典的致好看，还是经典的耐看。其次就是刹车灯太寒酸了，亮起来只有一点点，真的很难受，不知道为什么不直接全亮...

【帕萨特怎么样】大众帕萨特好不好_优点_缺点 - 太平洋汽车

四个大模型都能进行联网搜索，获取较新的讯息，从而给出比较详细的对比，也都在各自的页面列出了联网搜索信息的来源，方便用户查看具体的网页资讯。

5.角色扮演

- 假设你是诸葛亮，你会采用魏延的子午谷奇谋吗？

结果及分析

15 / 18

假设你是诸葛亮，你会采用魏延的子午谷奇谋吗？



已深度思考 (用时 22 秒)

臣亮闻魏文长子午谷之谋，沉吟良久，终以八字定论：“奇正相生，险中求胜；然时不我与，慎终如始。”

昔者文长执戟请命，欲效淮阴侯暗度陈仓之计，亲率精兵五千，负粮十日，直取长安。其策如流星贯月，锐不可当。然亮抚图思之，有三不可：

****一曰天时违逆。****曹魏方强，关中守将夏侯楙虽怯，然长安城高池深。纵使文长突至城下，焉知司马仲达不在雒阳虎视？若其轻骑驰援，则我孤军困于坚城之下，必成垓下之围。昔韩信出陈仓能胜，因三秦离心；今中原归心曹氏，岂可同日而语？

****二曰地理凶险。****子午谷栈道绵延六百余里，夏秋多雨则山洪暴发，冬春积雪则人马难行。昔丞相征南中，亲见毒泉瘴气之害，今若五千精锐困于绝谷，不战自损者恐半。此非畏险，实惜将士性命如金。

****三曰人和未至。****文长骁勇，然性同燎原之火。若使其独领偏师，恐重现关羽失荆州之祸。况先帝新丧，朝中黄皓渐起，若前线有失，则成都谗言必如潮涌。亮非疑文长之忠，实忧庙堂之谗也。

然亮亦非全然不用奇谋。建兴六年出祁山，使赵云为疑兵据箕谷，正合孙子"以正合，以奇胜"之道。唯子午谷之策，乃以国运为注，非万全之算。昔先帝托孤时执亮手曰："嗣子可辅则辅之，如其不才，君可自取。"亮岂敢以先帝基业赌一役乎？

今观后世兵家争执此事，多如弈棋者论局。若处楚汉之际，此计当行；然值三国鼎立，唯有步步为营。亮之不用子午谷，非无胆略，实因季汉输不起一败。后世杜工部诗云"出师未捷身先死"，此中辛酸，岂好弄险者能知？

若天假十年，待凉州尽附，陇右粮足，或可遣文长另率一军出子午，与主力成钳形之势。然此时此地，亮宁取街亭之稳，不赌子午之奇。此非怯战，实为《隆中对》"待天下有变"四字，守之三十年不敢忘也。

DeepSeek和文心一言都带入了情景当中，将自己作为诸葛亮进行分析，回答也以诸葛亮的口吻，并采用文言文，文心一言还在旁白加上了动作，更加生动。而通义千问和Kimi则没有带入情景，只给出了对于问题的分析。

四、总结与思考

DeepSeek

- **优势：**
 - 任务处理均衡：在各种任务（代码生成、文本生成、知识推理等）中表现出色，适用于广泛的应用场景。
 - 专业领域优势：在某些专业任务（如编程）中表现突出，具有较强的知识深度。
 - 回答全面，贴合用户需求：不仅能直接回答问题，还能提供相关背景知识和扩展信息，提升用户体验。
- **不足：**
 - 服务器稳定性有待提升：偶尔会遇到服务器繁忙，影响用户体验。

- 深度思考时间较长：在复杂推理或深度分析时响应速度较慢，可能影响交互流畅性。

- **改进建议：**

- 优化服务器资源分配：提升服务器负载能力，减少因流量过大导致的访问问题，提高系统的可用性和稳定性。
- 改进推理效率：通过优化算法或模型架构（如减少冗余计算、提升缓存机制），加快深度思考任务的响应速度，提高用户体验。
- 增强个性化推荐：根据用户使用习惯调整回答策略，例如提供更简明或更详细的回答模式，提升互动体验。

文心一言

- **优点：**

- 内容丰富，符合需求：能够生成较为详细、丰富的内容，能够较好地贴合用户的需求。
- 文本生成能力较强：在需要长文本生成的任务中表现优秀，具备一定的语言组织能力。

- **不足：**

- 回答卡顿：在生成回答的最后阶段容易出现卡住的情况，可能需要用户手动干预（等待或点击“停止生成”）。
- 响应速度慢：思考时间较长，生成回答的速度相对较慢，影响使用体验。

- **改进建议：**

- 优化生成速度：通过模型优化或缓存机制，减少长时间思考的情况，提高响应速度。在后台进行预计算，提高复杂问题的处理效率。
- 改善生成流畅性：解决回答卡顿的问题，例如优化推理过程，减少生成过程中断的情况。

通义千问

- **优点：**

- 任务处理能力强：能够稳定、高效地完成各种任务，服务可靠。
- 响应速度快：相比其他模型，深度思考时间短，能快速提供答案。
- 交互体验好：在回答下方提供用户可能接下来要问的问题，提高连续对话的便利性。
- 回答质量较高：能提供全面的答案，并补充相关知识点，增强用户理解。
- 专业领域表现优秀：在处理复杂任务或专业问题时，回答准确且速度快，优势明显。

- **不足：**

- 文本生成能力较弱：在创意写作、文章续写、摘要生成等方面表现不佳。

- **改进建议：**

- 优化文本生成能力：训练更大规模的文本数据集，特别是文学、创意写作、新闻摘要等高质量数据。
- 增强个性化与风格多样性：支持不同写作风格（如正式、幽默、诗意等），提升生成内容的适应性。

Kimi

- **优点:**
 - 对话体验流畅自然：能够进行高质量的互动，沟通顺畅。
 - 回答简洁直接：适合快速获取信息，不拖泥带水。
 - 推荐后续问题：能预测用户可能的下一个问题，提高交互效率。
 - 逻辑推理能力较强：在推理类问题上表现优越。
- **不足:**
 - 回答不够全面：一问一答的模式可能导致信息缺失，用户需要多次追问才能获得完整答案。
 - 专业领域表现一般：在处理复杂或专业性较强的问题时，理解能力和回答质量有待提高。
- **改进建议:**
 - 增强答案的全面性：可以在回答时提供更多背景信息、举例说明，避免用户反复追问。
 - 提升专业领域表现：优化对技术、医疗、法律等领域的理解，增强知识库，并提供参考来源。

五、未来展望

未来，国产大模型将朝着**更智能、更自主、更高效、更安全**的方向发展。在自主研发方面，模型将强化底层架构优化，提升国产化能力；在应用拓展上，多模态融合（文本、图像、音频等）将推动更多行业落地；在知识运用上，结合知识图谱与逻辑推理，提升专业性与精准性。同时，轻量化技术将降低成本，实现端侧部署，使AI更普及。安全性方面，隐私保护与内容可控性将进一步加强。随着技术突破与产业深度融合，国产大模型将成为全球AI竞争的重要力量，赋能社会各行各业，助推智能时代加速到来。

六、个人意愿

我对于大模型的使用需求比较高，综合使用下来，我个人感觉通义千问和DeepSeek给我的使用感受比较好，几乎所有的任务都处理的比较合理准确。文心一言和Kimi表现一般，对于我的需求没有很亮眼的表现。

要排名的话，**通义千问**我觉得是目前最强国产大模型，文心一言是真的体验极差，每次回答都特别慢还容易卡住。

排名:

- 1.通义千问
- 2.DeepSeek
- 3.Kimi
- 4.文心一言

七、参考

<https://blog.csdn.net/BuluAI/article/details/145615978>

https://blog.csdn.net/m0_69644950/article/details/145619316

<https://blog.csdn.net/dongwuming/article/details/10005149>

本文仅代表个人观点，所采用的数据与信息并不全面。