主流大模型体验

**选取模型：**

DeepSeek-V3 GPT4o mini Llama3 Claude3.5

**写作能力：**

***分别测试了中英文的说明类文本和文学性文本生成***

**DeepSeek：**

在中文理解和生成方面表现出色，领先于其他模型。这表明其在理解和生成中文文本方面具有强大的能力。

**GPT：**

在生成流畅、连贯的英文文本方面表现优异，其在多种写作任务中，如文章生成和对话系统中，均展现出高水平的写作能力。​

**Llama：**

在多语言文本生成方面具有优势，其在多种语言的文本生成任务中表现良好，尤其在处理多语言混合文本时表现突出。​

**Claude：**

在生成创意和文学性文本方面表现出色。其能够生成富有创意和文学性的内容，适用于需要创意写作的场景。​

**数学问题解决能力：**

***测试了高等数学和线性代数相关问题***

**DeepSeek：**

高等数学：在处理高等数学问题时表现出色，显著高于其他模型。

线性代数：在解决线性代数相关问题时，能够准确处理矩阵运算、特征值计算等任务，显示出强大的线性代数处理能力。 ​

**GPT：**

高等数学：在高等数学问题上表现一般。​在复杂的数学推理任务中，GPT可能需要更多的推理步骤，且准确率相对较低。​

线性代数：能够处理基本的向量运算和矩阵乘法，但在解决复杂问题时，经常出现错误的解答。​

**Llama：**

高等数学：在高等数学问题上表现中等。​在严谨证明方面存在困难，对于复杂定理仅能给出零散或模糊的论述。

线性代数：能够处理一般的线性代数问题，但在面对复杂的矩阵运算时，其表现不佳。​

**Claude：**

高等数学：在高等数学问题上表现良好。​

线性代数：能够有效处理线性代数问题。然而，在处理复杂的线性代数问题时，可能需要更多的推理和计算。

**总结：**

DeepSeek：在写作、解决数学问题方面均表现出色，适合需要综合能力的应用场景。​

GPT：在英文写作和对话系统中表现优异，但在数学推理方面存在一定局限。​

Llama：在多语言文本生成和一般复杂度的数学推理任务中表现良好，适合处理多语言和一般复杂度任务，不适合用于数学推理方面。

Claude：在创意写作方面具有优势，但在数学推理方面表现良好，可以正确解决大部分问题。