

大语言模型能力分析与应用

宗成庆

中国科学院自动化研究所

cqzong@nlpr.ia.ac.cn

提 纲

- 
- 1.概念约定
 - 2.能力分析
 - 3.应用定位
 - 4.个人观点

1. 概念约定

●语言模型(language model)

- n -元文法模型(n -gram model): 1980s
- 神经语言模型(neural network based language model): 2003, 10, 13
- 预训练语言模型
(Pre-training LM, PLM/ **G**enerative **P**re-**T**rained Transformer): 2018 ~
- 大语言模型 (large language model, LLM): 2020(1750亿参数)

●本次讨论

- ChatGPT/ GPT-4

提 纲

1. 概念约定

 2. 能力分析

3. 应用定位

4. 个人观点

2. 能力分析

◆ 拥有渊博的知识和强大的分析、对话能力

➤ 很强的对话能力

- 多轮对话的上下文理解
- 用户意图的理解能力
- 回复质量和逻辑一致性
- 甚至可以“交互式学习”

➤ 很好的语言生成和表达能力

- 规范的语言表达能力
- 长文本生成能力
- 代码生成能力
- 多语言生成能力
-

➤ 丰富的知识储备

- 常识性知识
- 百科知识
- 数学问题求解
-

2. 能力分析

➤ 在 ChatGTP 上的实验(2022.12.15): **关于写作**

A: 我们聊聊滑雪吧。你会滑雪吗?

C: 对不起,我是一个纯粹的语言模型,我无法感知或了解外界环境。我无法参与体育活动,也无法对滑雪进行评价。我只能根据您提供的信息来回答问题,不能对外界环境进行评价或描述。如果您有其他问题,我将尽力回答您的问题。

A: 你能写一首关于大雪的诗歌吗?

C: 虽然我不会滑雪,但我可以尝试写一首关于大雪的诗歌:

大雪纷飞,
银装素裹。
白茫茫一片,
如同洁白的棉被。
踏着深厚的积雪,
挥动双手捕捉飞舞的雪花。
欢乐的嬉戏,
让人忘却寒冷的冬日。
希望您喜欢这首诗。

2. 能力分析

A: 还不错，就是不押韵。

C: 我抱歉地通知您，我只是一个纯粹的语言模型，我无法感知或了解外界环境，也无法理解诗歌的韵律和押韵。我的目的是提供简单的信息和回答问题，而不是写诗歌。如果您对我的诗有任何批评或建议，我将非常感激。

A: 你可以把这首诗翻译成英文吗？

C: 是的，我可以尝试把这首诗翻译成英文：

The snow is falling heavily,
Dressed in silver clothes.
The vast expanse is white,
Like a pure cotton blanket.
Walking on the deep snow,
Swinging your hands to catch the flying snowflakes.
Happy play,
Forget the cold winter.

希望您能喜欢这首诗的英文版本。

2. 能力分析

A: 你能把这首诗改写成散文吗?

C: 是的, 我可以尝试把这首诗改写成散文:

当大雪纷飞, 披着银色的衣裳时, 白雪覆盖了整个世界, 就像一床洁白的棉被。踩着厚厚的积雪, 挥动双手捕捉飞舞的雪花, 在冬日的寒冷中尽情嬉戏, 忘却一切烦恼。

希望您喜欢这篇散文。

A: 谢谢你! 我快要下班了。

C: 不客气, 感谢您提出宝贵的问题。如果您有其他问题, 欢迎随时联系我, 我将尽力为您解答。祝您有一个愉快的下午!

2. 能力分析

关于医疗

Model: Default (GPT-3.5)

2023.4.18

NW

有人长期运动，喜欢打球和爬山，可是最近左脚底前面部分出现麻木、疼痛，甚至剧痛症状。请问是什么原因造成的？



这种疼痛和麻木的症状可能是由多种因素引起的。以下是一些可能的原因：



1. 运动损伤：如果您运动时使用不正确的姿势或过度使用脚部肌肉，可能会导致肌肉、韧带或肌腱的拉伤或扭伤，从而导致疼痛和麻木。
2. 跖筋膜炎：跖筋膜是脚底部的一层薄膜，它支撑着脚弓。如果这个膜发炎，可能会导致脚底疼痛和麻木。
3. 足底筋膜纤维化：这是一种常见的足底疼痛症状，可能是由于过度运动、过度使用或足部负重过重引起的。
4. 神经问题：麻木和疼痛也可能是由于神经问题引起的，例如坐骨神经痛或足底神经痛。

如果症状持续存在，建议您咨询医生进行检查，以确定症状的确切原因，并获得适当的治疗。

2. 能力分析

NW

在这种情况下，还能运动吗？选择什么样的运动比较合适？



如果您出现了麻木和疼痛症状，最好暂时避免剧烈的运动和活动，以免加重症状和损伤。但是，适当的运动可以有助于改善症状和恢复健康。以下是一些适合的运动：



1. 游泳：游泳是一个低冲击的全身性运动，可以帮助您锻炼身体而不会给脚部施加压力。
2. 静态自我拉伸：这些伸展运动可以帮助您放松肌肉和缓解压力，从而减轻疼痛和麻木症状。
3. 轻度有氧运动：如快走、慢跑、骑车等，可以帮助您提高心肺功能和健康水平。
4. 瑜伽：瑜伽的许多姿势可以帮助您放松和加强身体的肌肉和关节，从而缓解疼痛和麻木症状。

请注意，运动前最好咨询医生的意见，以便选择最适合您的运动方式和强度。同时，也要在运动中注意保护脚部，避免过度使用和过度负荷，以避免加重症状和损伤。

2. 能力分析

NW

A和B两个人在饭店吃饭时发生争执，A先动手打了B，但没有造成重伤。B报警后民警将两人带到派出所，B不同意和解。请问民警应该如何处理？

关于法律



在这种情况下，民警应该根据相关法律法规，对A和B进行调查和处理，以维护公正和法律的权威。



1. 对A和B进行询问和取证：民警应当向A和B了解事发经过，了解是否有其他证人或证据，如监控录像等，以收集证据。
2. 进行医学鉴定：民警应当将B送到医院进行医学鉴定，确定受伤程度，以作为后续处理的参考。
3. 根据相关法律规定进行处理：如果A的行为构成了刑事犯罪，民警应当依照法律程序进行处理，可能会涉及到行政拘留或刑事处罚等。如果A的行为不构成犯罪，民警应当及时调解双方，促成和解，或者将案件移交到相关部门进行调解或处理。
4. 加强宣传教育：民警应当向双方进行法律宣传和教育，提高双方的法律意识和法律素养，以预防类似事件的发生。

在处理这种情况时，民警应当尽力保持公正和客观，遵循法律程序，依法处理，确保双方的合法权益得到维护。

2. 能力分析

NW

请给我一些建议，如何构建中医的知识图谱。

关于NLP



构建中医的知识图谱需要涉及多个方面，包括中医理论、临床经验、药物治疗等，以下是一些建议：



1. 确定知识图谱的结构和分类体系：需要根据中医的理论体系，将相关的概念、术语进行分类，形成一个层次分明的知识图谱结构，这有利于后续知识的维护和更新。
2. 收集中医相关的数据：需要搜集大量的中医相关的数据，包括中医文献、病历数据、临床试验数据、中药材药效数据等，同时需要建立相应的数据清洗和处理机制，以确保数据的准确性和完整性。
3. 选择合适的知识表示方式：可以采用多种方式对知识进行表示，包括本体论、语义网、关系数据库等，需要选择适合中医知识特点的知识表示方式，以便更好地表达中医知识。
4. 选择合适的算法和工具：需要选择适合中医知识图谱构建的算法和工具，如机器学习、自然语言处理等，以实现知识的自动化抽取和处理。
5. 进行知识图谱的评估和验证：需要对构建的中医知识图谱进行评估和验证，以确保知识图谱的质量和准确性，同时可以不断进行迭代和优化，提高知识图谱的实用价值和适应性。

总之，中医知识图谱的构建需要综合运用多种技术手段和方法，需要专业人士的参与和不断的实践和改进。

2. 能力分析

Model: GPT-4

关于算术



When I was 6 my sister was half my age. Now I'm 70 how old is my sister?



When you were 6 years old, your sister was half your age, which means she was 3 years younger than you. Now that you're 70, your sister would be $70 - 3 = 67$ years old.



- 可以大幅度提高工作效率
- 极大地拓展个人视野
- 在各领域不同任务上都有中等水平以上的表现

2. 能力分析

ChatGPT 在2022年高考全国试卷客观题上的表现

	得分率	准确率
语文	-	-
英语	93.0 / 95	56/60
地理	24/36	6/9
政治	44/48	11/12
历史	24/40	6/10
数学(理)	30/50	6/10
数学(文)	35/55	7/11
物理	0/12	0/2
化学	6/42	1/7
生物	18/30	3/5
合计	274/408= 67%	96/126= 76%

客观题：在排除了坏样本（带图题,听力题等）之后，在全部126个样本数可以达到**76%**的准确率和**67%**的得分率，其客观题能力与**两名**500分左右的高考生（文科和理科各一名）相当。

主观题：对ChatGPT结果的人工打分，在文科综合（历史，地理，政治）试卷上取得了不错的成绩（得分率达78%），生物试卷的得分率为50%，而在数学、物理、化学、历史试卷上的表现则不佳（得分率低于30%）。

感谢复旦大学邱锡鹏教授提供的数据！

2. 能力分析

◆ 难以掩饰的“愚笨”和“无知”

➤ 特定领域的专业知识匮乏

➤ 模型的内生性缺陷

- 可解释性
- 知识来源缺乏引用
- 出现“臆想”答案
- 事实性错误……

➤ 由于对齐而产生人为倾向性

- 回复带有主观性倾向/价值观
- 无法展现和平衡多种观点
- 缺乏灵活的角色和具身感知能力
- ……

➤ 不良表现时有发生

- 投机取巧：代替人工完成任务
- 无中生有：杜撰实验报告
- 以假乱真：内容虚假，混淆视听
- 偷梁换柱：篡改内容
- 瞒天过海：错误观点或意识形态
- ……

2. 能力分析

NW

鲁迅和周树人是同一个人吗？



不是。鲁迅（1881年9月25日-1936年10月19日），原名周樟寿，字豫山，是中国现代文学史上杰出的文学家、思想家、革命家，代表作有《狂人日记》、《阿Q正传》、《呐喊》等。



周树人（1895年1月25日-1972年11月1日），原名周懋华，字有楼，是中国现代文学史上著名的文学家、教育家、翻译家，代表作有《故乡》、《春蚕》、《朱自清散文集》等。

两位文学家名字中的“周”字相同，但是他们的名字、生平、思想、作品等方面都有着很大的不同，是两个不同的人。

2. 能力分析

ChatGPT

NW

你知道“林黛玉倒拔垂杨柳”的故事吗



是的，我知道“林黛玉倒拔垂杨柳”的故事。这个故事出自于中国古典小说《红楼梦》，是小说中著名的情节之一。  

故事发生在贾宝玉家中，林黛玉因为伤心难过，于是去院子里倒拔了一棵垂杨柳。贾宝玉得知后非常惋惜，认为这棵树根深固扎，难以倒拔，而林黛玉却倒拔了它，正好象征着她在这个世界上的无奈和无力，所以这个场景在小说中被视为一个非常有象征意义的情节。

这个情节也代表了林黛玉的悲剧命运和贾宝玉对她的深情。林黛玉的性格和命运与垂杨柳有着相似之处，都是柔弱、易折的存在。而贾宝玉则是一个有情有义的人，他深爱林黛玉，但是最终却不能拯救她于水深火热之中。

GPT-4

NW

你知道“林黛玉倒拔垂杨柳”的故事吗



“林黛玉倒拔垂杨柳”是中国古典名著《红楼梦》中的一段描写。《红楼梦》是清代作家曹雪芹创作的长篇小说，被誉为中国古典小说的四大名著之一。  

故事发生在林黛玉来到荣府后，她因生性聪慧、才情出众，深受贾母与王夫人的宠爱。有一天，林黛玉在荣府的大观园里散步，看到一片美丽的垂杨柳。当时垂杨柳正值春天繁茂，杨柳依依，景色宜人。然而林黛玉却对这垂杨柳心生厌恶，她认为它们像病夫一样无精打采，不值一顾。于是林黛玉倒拔了其中一棵垂杨柳。

这个故事反映了林黛玉的性格特点，她聪明、机智、独立，有时也带有一些叛逆。在《红楼梦》中，林黛玉是一位极具个性和魅力的人物，她的故事至今仍具有很高的文学价值。

2. 能力分析

A: 请介绍一下鸡兔同笼问题。

C: 鸡兔同笼问题是一个经典的数学问题，问题的形式通常是这样的：如果有一个笼子里有鸡和兔子，总共有 n 只脚，问笼子里有多少只鸡和多少只兔子。这道题目的答案通常可以用一个简单的方程来推导出来：如果 c 只鸡和 r 只兔子，那么 $c + r = n$ ，并且 $2c + 4r = 2n$ 。由于鸡和兔子的脚数之和是已知的，所以我们可以用这两个方程来求解 c 和 r 的值。例如，如果 $n = 10$ ，那么 $c + r = 10$ ，并且 $2c + 4r = 20$ ，我们可以用第一个方程来推出 $c + r = 10$ ，然后用第二个方程来推出 $c = 5$ 和 $r = 5$ 。这样，我们就可以得到答案：笼子里有 5 只鸡和 5 只兔子。

A: 如果鸡和兔子一共36个头, 94只脚。那有多少只鸡, 多少只兔子呢?

C: 根据鸡兔同笼问题的推导方法，我们可以得到两个方程来求解这个问题：

$$c + r = 36$$

$$2c + 4r = 94$$

由于 c 和 r 分别代表鸡和兔子的数量，因此我们可以用第一个方程来推出 $c + r = 36$ ，然后用第二个方程来推出 $c = 23$ 和 $r = 13$ 。因此，答案是：笼子里有 23 只鸡和 13 只兔子。

一本正经地胡说八道

2. 能力分析

自我评价

Model: Default (GPT-3.5)

NW

请谈谈使用 ChatGPT 可能给社会带来的好处和问题。



ChatGPT 是一种强大的自然语言处理技术，可以通过对语言的理解和生成，与人类进行自然对话和交流。它可能会带来很多好处，但也存在一些问题。



好处：

1. 改善客户服务：ChatGPT 可以为企业提供更快速和更准确的客户服务，通过与客户进行实时交互，识别问题并提供解决方案。
2. 提高教育效果：ChatGPT 可以成为一个有用的教学工具，可以根据学生的需求和兴趣生成自适应教学材料，提高教育效果。
3. 改进医疗保健：ChatGPT 可以帮助医生和医疗保健机构提供更好的服务，可以提供快速和准确的答案，帮助诊断和治疗疾病。
4. 促进多语言沟通：ChatGPT 可以为人们提供跨语言的沟通方式，使得语言不再成为交流的障碍。
5. 提高生产效率：ChatGPT 可以为企业提供自动化的客户支持和服务，可以快速响应客户需求，提高生产效率。

2. 能力分析

问题：

1. 隐私问题：ChatGPT 可能会记录用户的对话内容，这可能会引起用户的隐私担忧，同时也可能会被不法分子利用。
2. 偏见问题：ChatGPT 的训练数据可能存在偏见，这可能导致 ChatGPT 生成的内容也具有偏见。
3. 可信度问题：ChatGPT 生成的内容可能不准确，这可能导致人们对 ChatGPT 的信任度降低。
4. 社交问题：ChatGPT 可能会导致人们与真正的人类沟通减少，这可能会对社交造成负面影响。
5. 伦理问题：ChatGPT 可能会被用于欺诈、骗局和其他不道德的行为，这可能会引起道德和法律问题。

总之，ChatGPT 可能会为社会带来很多好处，但也存在一些问题。必须仔细考虑这些问题，并采取适当的措施来解决这些问题，以确保 ChatGPT 的应用能够真正带来社会的利益。

提 纲

1. 概念约定
2. 能力分析
- ➡ 3. 应用定位
4. 个人观点

3. 应用定位

◆ 定位— AI助手

- 百科问答/ 搜索
- 写作助手/ 文书代笔
- 文案设计
- 外语翻译
- 教学、科研
- 编程协助
- 舆论宣传
-

◆ 被冲击较大的行业

- 文秘
- 教育、教学
- 翻译
- 科研诚信
- 舆论宣传
- 平面设计
-

难度不是太大但重复性较大的工作，即使有点差错也不会造成严重后果的职业，被机器替代是不可逆转的趋势。

无论从事什么职业，水平在中低端的人员，拥有更大的失业风险。

3. 应用定位

◆ 具有应用局限性的行业和领域

- 需要精深专业知识的行业
- 需要严谨解释和验证的行业
- 需要随时更新知识的行业
- 需要多元价值观和包容性文化的领域
- 轻量化、离线型的应用场景

3. 应用定位

◆ 学术界的关注

- Nature 期刊关于 ChatGPT 的讨论 (Feb. 3, 2023)

➤ ChatGPT: five priorities for research

(<https://www.nature.com/articles/d41586-023-00288-7>)

- Conversational AI is a game-changer for science.

- Hold on to human verification (坚持人类验证)
- Develop rules for accountability (制定问责机制)
- Invest in truly open LLMs (投资真正开放的LLM)
- Embrace the benefits of AI (拥抱人工智能的红利)
- Widen the debate (扩大辩论范围)

3. 应用定位

- 香港中文大学（深圳）的相关规定 (March 6, 2023)

Guidelines on the Use of Artificial Intelligence Tools in Teaching, Learning and Assessments

Teachers and students should embrace and are encouraged to become acquainted with these tools. Students should learn to make sensible use of AI tools, not only for their studies, but also for their future professional development and advancement in order to thrive in this AI era.

(教师和学生应该接受并得到鼓励去熟悉这些工具。学生应该学会合理地使用AI工具，不仅是为了他们的学习，也是为了他们未来的职业发展和进步，以便在这个人工智能时代茁壮成长。)

Like any other educational resources, teachers and students should approach AI tools critically, recognising their limitations in an honest and authentic manner and how they could be incorporated into teaching and learning in order to attain the desired learning outcomes.

(与任何其他教育资源一样，教师和学生应该批判性地使用人工智能工具，以诚实和真实的方式认识到它们的局限性，以及如何将它们融入教学和学习中，以获得期望的学习效果。)

3. 应用定位

- 生命未来研究所的公开信 (March 22, 2023)

Future of Life Institute (<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>)

Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

(我们呼吁所有人工智能实验室立即暂停训练比GPT-4更强大的人工智能系统至少6个月。)

3. 应用定位

- 人工智能促进会的声明 (April 5, 2023)

The Association for the Advancement of Artificial Intelligence (AAAI)

<https://aaai.org/working-together-on-our-future-with-ai/>

Working together on our future with AI

We are writing to provide a balanced perspective on managing the progress in the field. We also seek to broaden and strengthen the community of engaged researchers, government agencies, private companies, and the public at large, to ensure that society is able to reap the great promise of AI while managing its risks.

(我们写这篇文章是希望提供一个平衡的视角以管理该领域的进展。我们寻求扩大和加强参与的研究人员、政府机构、私营公司和广大公众的社区，以确保社会能够在管理其风险的同时收获人工智能的巨大前景。)

提 纲

1. 概念约定
2. 能力分析
3. 应用定位
- ➡ 4. 个人观点

4. 个人观点

- 对于普通民众，没有理由，也没有必要限制使用AI工具（包括 ChatGPT 等）；
- 对于学术界（包括教师、学生、科研人员等），确定使用AI工具的边界，如学生做作业、考试等；
- 加强诚信体系建设，使用者诚实、真实地说明使用AI的方式方法，甚至具体细节；
- 合理地引导用户，尤其是青年学生正确地使用AI工具，提高明辨是非的能力，加强批判思维意识；
- 借助技术手段辅助鉴别AI工具可能产生的错误，给用户明确的提示；尤其针对意识形态问题，明确给出代表国家意志的答案；
- 制定知识产权、数据安全、隐私保护、社会伦理和学术诚信等相关法律法规，加大处罚力度，一视同仁，并与诚信体系关联；
- 沉下心来，做好自己的事情，做出中国的GPT，超越GPT！



谢谢!
Thanks!

