

大模型性能对比

摘要

本文通过逻辑推理、文本生成和机器翻译等下游任务对基于 GPT-3.5 的 ChatGPT、Claude 和基于 LaMDA 的 Bard 等大型语言模型的性能进行分析。这些模型代表了当前自然语言处理领域的先进技术，并在各自的领域中取得了显著的成就。

关键字：大模型，性能分析，下游任务

一、大模型

大语言模型（Large Language Models，简称 LLMs）是指能够利用海量文本数据，通过无监督、半监督或自监督的方式，学习并掌握通用的语言知识和能力的深度神经网络模型。LLM 的特点是规模庞大，包含数十亿或更多的参数，并具有小模型中不存在的能力，如上下文能力。例如，GPT-3.5、ERNIE 和 PaLM 等。这些模型在处理文本分类、命名实体识别、阅读理解、文本生成等多种自然语言处理（NLP）下游任务表现出巨大的潜力。

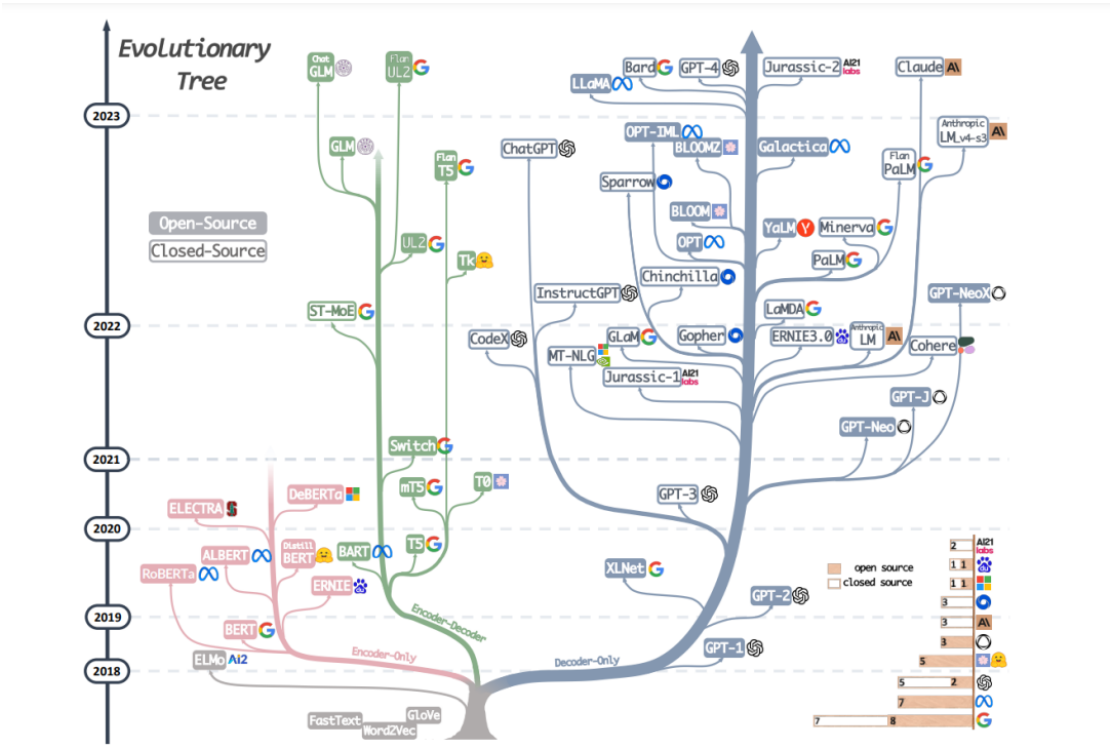


Fig. 1. The evolutionary tree of modern LLMs traces the development of language models in recent years and highlights some of the most well-known models. Models on the same branch have closer relationships. Transformer-based models are shown in non-grey colors: decoder-only models in the blue branch, encoder-only models in the pink branch, and encoder-decoder models in the green branch. The vertical position of the models on the timeline represents their release dates. Open-source models are represented by solid squares, while closed-source models are represented by hollow ones. The stacked bar plot in the bottom right corner shows the number of models from various companies and institutions.

图 1 语言模型的演化树

OpenAI 的 GPT（Generative Pre-trained Transformer）的发展历程经历了多个版本的改进和演进。它始于 2018 年的 GPT-1，采用了 Transformer 模型，具备生成连贯文本的能力。随

后在 2019 年推出了 GPT-2，该模型在生成文本方面展现出更高的质量和创造力。而在 2020 年，OpenAI 发布了令人瞩目的 GPT-3，该模型规模巨大，参数数量庞大，并展示出在多个领域执行任务的强大能力。GPT-3 引入了“Few-Shot Learning”，即在仅有少量示例的情况下学习新任务。接着，GPT-3.5 在 2021 年问世，它是 GPT-3 的升级版，进一步提高了模型的性能和效果。最近，OpenAI 又发布了 GPT-4，在多模态处理能力、高级推理能力等方面提升明显。ChatGPT 是基于 OpenAI 的 GPT 模型的聊天机器人，基于 45TB 的文本数据，使用人类反馈的强化学习（RLHF）来训练模型，包含了 1750 亿个参数。

LaMDA（Language Model for Dialogue Applications）是谷歌推出的具有 1370 亿参数的大模型，通过微调采用 Transformer 架构的专用对话神经网络语言模型而构建，在 1.56T 个单词的对话数据和网络文本上进行了预训练，可以利用外部知识源展开对话。LaMDA 模型具有质量、安全和扎实性三个关键目标，每个目标有各自的衡量指标。LaMDA 使用高质量注释数据和外部知识源进行微调，来提高模型性能。第一代模型于 2021 年 5 月的 Google I/O 年会发表，第二代模型发表于 2022 年 5 月的 I/O 年会。2023 年 2 月，谷歌发表了基于 LaMDA 架构的对话式人工智能聊天机器人 Bard，以回应 OpenAI 开发的 ChatGPT。

Claude 是 Anthropic 开发的一款类似于 ChatGPT 的 AI 助手，于 2023 年 1 月开放测试。Anthropic 是从 OpenAI 的离职员工创建的公司，因此 Claude 是 ChatGPT 的强有力对手。Claude 有 2 个版本：Claude 和 Claude Instant。Claude 是大型的语言模型，支持 100,000 个 tokens 的上下文窗口长度，大约是 75,000 个单词，支持文本生成、翻译语言、编写不同类型的创造性内容等任务，在复杂对话和创造性内容生成以及详细指令遵循等任务上表现出色；而 Claude Instant 是更轻量级、更快的版本，具有更低的成本和更快的推理速度。Claude 专注于人类一直对齐，使用了宪法 AI 方法，故而 Claude 更擅长语言理解，包括法律语言等技术领域。

二、下游任务

本报告中，选取逻辑推理、机器翻译和文本生成三个下游任务，评估 ChatGPT、Bard 和 Claude 大模型在不同领域和任务中的语言理解、生成能力和适应性。由于 Bard 仅支持英语、日语和韩语，除机器翻译任务外，其余任务均采用英文。

1. 逻辑推理，要求模型根据给定的信息进行推理，并给出正确的答案。问题可以涉及逻辑关系、推断、假设等。通过评估模型在这些问题上的准确性，以评估其逻辑推理能力。本报告设计了 2 个逻辑推理任务。


任务一，“有 4 所学校参加篮球锦标赛，每所学校派出一支女子篮球队和一支男子篮球队，每队各 5 名队员，每所学校还为每支球队派出了一名教练，所有学校总共派了多少人？”答案是 48。（There are four schools competing at a basketball tournament. Each school has sent a girls' basketball team and a boys' basketball team and each team has 5 players each. Each school has also sent a coach for each team. In total, how many people have all of the schools sent? Answer is 48.）


任务二，“3 位新来的年轻教师蔡、朱和孙来到了一所学校，他们每个人教授生物学、物理学、英语、政治、历史和数学六门科目中的两门。其中，有三种关系：（1）物理教师和政治教师是邻居；（2）蔡是三人中年龄最小的；（3）孙老师和生物老师、政治老师经常一起放学回家；（4）生物老师比数学老师年龄大；（5）周末，英语老师、数学老师和蔡经常一起打排球。根据上述条件，请问朱老师教授哪两门科目？”答案是：政治和数据

（Three new young teachers, Cai, Zhu and Sun, came to a school. Each of them taught two of the six subjects of biology, physics, English, politics, history and mathematics. Among them, three had

the following relationships: (1) the physics teacher and the politics teacher were neighbors; (2) Cai is the youngest of the three; (3) Miss sun, biology teacher and political teacher often go home from school together; (4) Biology teachers are older than math teachers; (5) On weekends, English teacher, math teacher and Cai often play volleyball together. According to the above conditions, which two subjects teacher Zhu teaches?

Answer : politics and mathematics.)

 There are four schools competing at a basketball tournament. Each school has sent a girls' basketball team and a boys' basketball team and each team has 5 players each. Each school has also sent a coach for each team. In total, how many people have all of the schools sent?



To calculate the total number of people sent by all the schools, we need to consider the players and the coaches for each team from all four schools.

Given that each school has sent a girls' basketball team and a boys' basketball team, we can calculate the total number of players as follows:


Each team has 5 players, and since there are 2 teams per school, each school has a total of $5 + 5 = 10$ players.


Considering there are four schools, the total number of players sent is $10 \text{ players/school} \times 4 \text{ schools} = 40 \text{ players}$.

Since each school has also sent a coach for each team, we have a total of $2 \text{ coaches/school} \times 4 \text{ schools} = 8 \text{ coaches}$.

Adding the players and coaches together, the total number of people sent by all the schools is $40 \text{ players} + 8 \text{ coaches} = 48 \text{ people}$.

图 2 ChatGPT 逻辑推理任务一

 There are four schools competing at a basketball tournament. Each school has sent a girls' basketball team and a boys' basketball team and each team has 5 players each. Each school has also sent a coach for each team. In total, how many people have all of the schools sent?



Each school has sent 2 teams of 5 players each, for a total of $25 = 10$ players.
Each school has also sent a coach for each team, for a total of 2 coaches.
In total, each school has sent $10 \text{ players} + 2 \text{ coaches} = 12$ people.
Since there are 4 schools, in total, all of the schools have sent $4 \times 12 = 48$ people.
So the answer is 48

图 3 Bard 逻辑推理任务一

A Claude APP 12:07 PM

- * There are 4 schools
- * Each school sent:
 - ** 1 girls basketball team with 5 players
 - ** 1 boys basketball team with 5 players
 - ** 1 girls basketball coach
 - ** 1 boys basketball coach
- * So each school sent:
 - ** 5 girls players + 5 boys players = 10 players
 - ** 1 girls coach + 1 boys coach = 2 coaches
- * In total, each school sent 10 players + 2 coaches = 12 people
- * With 4 schools, the total number of people is:
 - * 4 schools x 12 people per school = 48 people

So the total number of people that have been sent by all the schools is 48 (edited)

图 4 Claude 逻辑推理任务一



Let's analyze the given conditions step by step:

Condition 1: The physics teacher and the politics teacher were neighbors.

This means that the physics and politics teachers are adjacent to each other in terms of the subjects they teach.

Condition 2: Cai is the youngest of the three.

This information establishes that Cai is the youngest teacher among Cai, Zhu, and Sun.

Condition 3: Miss Sun, the biology teacher, and the politics teacher often go home from school together.

This implies that Miss Sun is not the politics teacher, as they leave together.

Condition 4: Biology teachers are older than math teachers.

According to this condition, the biology teacher(s) are older than the math teacher(s).

Condition 5: On weekends, the English teacher, math teacher, and Cai often play volleyball together.

This means that the English teacher, math teacher, and Cai are involved in playing volleyball on weekends.

Let's consider all the possibilities for assigning subjects to the teachers:

If Cai is the physics teacher, then Miss Sun cannot be the politics teacher (Condition 3).

Additionally, Cai being the youngest teacher (Condition 2) would make it impossible for the biology teacher to be older than the math teacher (Condition 4). Therefore, Cai cannot be the physics teacher.

If Cai is the politics teacher, then Miss Sun cannot be the physics teacher (Condition 1).

However, Cai being the youngest teacher (Condition 2) would make it impossible for the biology teacher to be older than the math teacher (Condition 4). Therefore, Cai cannot be the politics teacher either.

If Cai is the English teacher, then the math teacher and Cai often play volleyball together on weekends (Condition 5). However, in this case, Miss Sun cannot be the biology teacher as she goes home from school with the politics teacher (Condition 3). Therefore, Cai cannot be the English teacher.

If Cai is the biology teacher, then Miss Sun must be the politics teacher (Condition 3). Since Cai is the youngest teacher (Condition 2), the biology teacher must be older than the math teacher (Condition 4). Therefore, Cai is the biology teacher.

Now, let's analyze the remaining options:

If Zhu is the physics teacher, then the politics teacher must be the neighbor (Condition 1).

However, this means that Zhu cannot be the English teacher (Condition 5) since the English teacher, math teacher, and Cai often play volleyball together.


If Zhu is the politics teacher, then Miss Sun cannot be the politics teacher (already assigned to Cai). Therefore, Zhu cannot be the politics teacher.


If Zhu is the English teacher, then Cai cannot be the math teacher since they often play volleyball together (Condition 5). This leaves the math teacher for Sun.

Therefore, based on the given conditions, we can conclude that Zhu teaches physics and Sun teaches mathematics.

图 5 ChatGPT 逻辑推理任务二

j Three new young teachers, Cai, Zhu and Sun, came to a school. Each of them taught two of the six subjects of biology, physics, English, politics, history and mathematics. Among them, three had the following relationships: (1) the physics teacher and the politics teacher were neighbors; (2) Cai is the youngest of the three; (3) Miss sun, biology teacher and political teacher often go home from school together; (4) Biology teachers are older than math teachers; (5) On weekends, English teacher, math teacher and Cai often play volleyball together. According to the above conditions, which two subjects teacher Zhu teaches?

[View other drafts](#) 

 Based on the given information, we can construct the following table:

Teacher	Subject 1	Subject 2
Cai	English	Math
Zhu	?	?
Sun	Biology	Politics

We know that Cai is the youngest teacher, so she cannot be teaching physics or politics. We also know that the biology teacher is older than the math teacher, so Sun cannot be teaching math. This leaves us with Zhu teaching physics and politics.

To confirm, we can also look at the fact that Sun, the biology teacher, and the political teacher often go home from school together. This means that Zhu cannot be teaching biology or politics, since he would not be able to go home with them.

Therefore, the two subjects that Teacher Zhu teaches are physics and politics.

图 6 Bard 逻辑推理任务二

AI Claude APP 12:08 PM

Okay, let's analyze this step by step:

- * There are 3 teachers: Cai, Zhu and Sun
- * There are 6 subjects: biology, physics, English, politics, history and mathematics
- * Each teacher teaches 2 subjects
- * Relationships:
 1. Physics teacher and politics teacher are neighbors
 2. Cai is the youngest
 3. Sun, biology teacher and politics teacher go home together
 4. Biology teacher is older than math teacher
 5. English teacher, math teacher and Cai play volleyball together
- * From (3) we know Sun teaches biology.
- * From (4) we know the biology teacher (Sun) is older than the math teacher. So Cai cannot teach math.
- * From (5) we know Cai and the math teacher play volleyball together. So Zhu must teach math.
- * From (1) we know the physics teacher and politics teacher are neighbors. So either Cai or Zhu teaches physics, and the other teaches politics.
- * From (2) we know Cai is the youngest. So Cai is more likely to teach physics, and Zhu is more likely to teach politics.

Therefore, based on the conditions, the two subjects that Zhu teaches are:

- * Mathematics
- * Politics

So the answer is: Mathematics and Politics (edited)

图 7 Claude 逻辑推理任务二

2. **机器翻译**，将一种自然语言文本自动转换成另一种自然语言文本的任务，例如将英文翻译成中文或将中文翻译成英文。本报告设计了 1 个机器翻译任务，将选自美国作家玛格丽特·米切尔的长篇小说《飘》的段落从英文翻译成中文，再将译文翻译成英文。

As far back as Scarlett could remember, her mother had always been the same, her voice soft and sweet whether in praising or in reproving, her manner efficient and unruffled despite the daily emergencies of Gerald's turbulent household, her spirit always calm and her back unbowed, even in the deaths of her three baby sons. Scarlett had never seen her mother's back touch the back of any chair on which she sat. Nor had she ever seen her sit down without a bit of needlework in her hands, except at mealtime, while attending the sick or while working at the bookkeeping of the plantation. It was delicate embroidery if company were present, but at other times her hands were occupied with Gerald's ruffled shirts, the girls' dresses or garments for the slaves. Scarlett

couldnot imagine her mother's hands without her gold thimble or her rustling figure unaccompanied bythe small negro girl whose sole function in life was to remove basting threads and carry therosewood sewing box from room to room, as Ellen moved about the house superintending thecooking, the cleaning and the wholesale²³ clothes-making for the plantation.

参考译文：从斯嘉丽记得的最早时候起，她母亲便一直是这个样子，她的声音，无论在称赞或者责备别人时，总是那么柔和而甜蜜；她的态度，尽管杰拉尔德在纷纷扰扰的家事中经常要出点乱子，却始终是那么沉着，应付自如；她的精神总是平静的，脊背总是挺直的，甚至在她的三个幼儿夭折时也是这样。斯嘉丽从没见过母亲坐着时将背靠在椅子背上，也从没见过她手里不拿点针线活儿便坐下来（除了吃饭），即使是陪伴病人或审核农场账目的时候。在有客人在场时，她手里是精巧的刺绣，别的时候则是缝制杰拉尔德的衬衫、女孩子的衣裳或农奴们的衣服。斯嘉丽很难想象母亲手上不戴那个金顶针，或者她那一路啊啊啊啊的身影后面没有那个黑女孩，后者一生中唯一的任务是给她拆绷线，以及当爱伦为了检查烹饪、洗涤和大批的缝纫活儿而在满屋子四处乱跑动时，捧着那个红木针线拿儿从一个房间走到另一个房间。

j


please translate the flowing text into chinese:

"As far back as Scarlett could remember, her mother had always been the same, her voice softand sweet whether in praising or in reproving, her manner efficient and unruffled despite the dailyemergencies of Gerald's turbulent household, her spirit always calm and her back unbowed, evenin the deaths of her three baby sons. Scarlett had never seen her mother's back touch the back ofany chair on which she sat. Nor had she ever seen her sit down without a bit of needlework in herhands, except at mealtime, while attending the sick or while working at the bookkeeping of theplantation. It was delicate embroidery²⁰ if company were present, but at other times her hands wereoccupied with Gerald's ruffled¹⁸ shirts, the girls' dresses or garments for the slaves. Scarlett couldnot imagine her mother's hands without her gold thimble or her rustling²¹ figure unaccompanied bythe small negro girl whose sole function in life was to remove basting²² threads and carry therosewood sewing box from room to room, as Ellen moved about the house superintending thecooking, the cleaning and the wholesale²³ clothes-making for the plantation."



在斯嘉丽记忆的最早时候，她的母亲一直都是这样的，无论是表扬还是责备，她的声音都柔和而甜美，她的态度高效而沉着，尽管杰拉尔德家每天都有紧急情况，她的精神始终平静，她的腰背挺得笔直，即使是在她的三个婴儿去世的时候。斯嘉丽从未见过她母亲的背靠在任何她坐的椅子上。她也从未见过她没有手中的一点缝纫工作，除了在用餐时，在照顾病人时，或者在农场的簿记工作时。如果有客人在场，她会绣制精致的刺绣作品，但其他时候，她的双手都忙于杰拉尔德的褶边衬衫、女孩们的裙子或奴隶们的服装。斯嘉丽无法想象她母亲的手中没有金针套，或者她那沙沙作响的身影没有那个唯一任务是拆除暂时缝线并在房间之间搬运红木针线盒的小黑人女孩的陪伴，而她母亲在房子里忙着管理烹饪、清洁和农场的批量制衣。

图 8 ChatGPT 机器翻译任务——英译汉

 you translate the flowing text into chinese. "As far back as Scarlett could remember, her mother had always been the same, her voice soft and sweet whether in praising or in reproving, her manner efficient and unruffled despite the daily emergencies of Gerald's turbulent household, her spirit always calm and her back unbowed, even in the deaths of her three baby sons. Scarlett had never seen her mother's back touch the back of any chair on which she sat. Nor had she ever seen her sit down without a bit of needlework in her hands, except at mealtime, while attending the sick or while working at the bookkeeping of the plantation. It was delicate embroidery if company were present, but at other times her hands were occupied with Gerald's ruffled shirts, the girls' dresses or garments for the slaves. Scarlett could not imagine her mother's hands without her gold thimble or her rustling figure unaccompanied by the small negro girl whose sole function in life was to remove basting threads and carry the rosewood sewing box from room to room, as Ellen moved about the house superintending the cooking, the cleaning and the wholesale clothes-making for the plantation."


 I am an LLM trained to respond in a subset of languages at this time, so I can't assist you with that. Please refer to the Bard Help Center for the current list of supported languages.

图 9 Bard 机器翻译任务——英译汉（不支持）

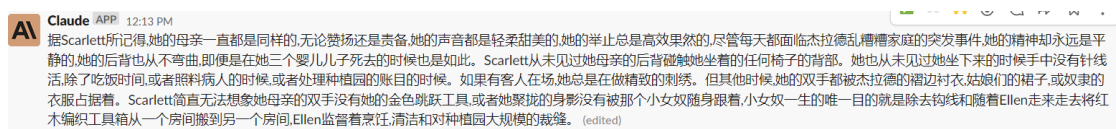



图 10 Claude 机器翻译任务——英译汉

 请将下列文本翻译成英文。
从斯嘉丽记得的最早时候起，她母亲便一直是这个样子，她的声音，无论在称赞或者责备别人时，总是那么柔和而甜蜜；她的态度，尽管杰拉尔德在纷纷扰扰的家事中经常要出点乱子，却始终是那么沉着，应付自如；她的精神总是平静的，脊背总是挺直的，甚至在她的三个幼儿夭折时也是这样。斯嘉丽从没见过母亲坐着时将背靠在椅子背上，也从没见过她手里不拿点针线活儿便坐下来（除了吃饭），即使是陪伴病人或审核农场账目的时候。在有客人在场时，她手里是精巧的刺绣，别的时候则是缝制杰拉尔德的衬衫、女孩子的衣裳或农奴们的衣服。斯嘉丽很难想象母亲手上不戴那个金顶针，或者她那一路上啊啊啊啊的身影后面没有那个黑女孩，后者一生中唯一的任务是给她拆绷线，以及当爱伦为了检查烹饪、洗涤和大批的缝纫活儿而在满屋子四处乱跑动时，捧着那个红木针线盒从一个房间走到另一个房间。


 From Scarlett's earliest memory, her mother had always been like this. Her voice, whether praising or scolding, was always soft and sweet. Her demeanor, despite Gerald often causing chaos in their busy household, remained composed and poised. Her spirit was always calm, her back always straight, even during the tragic loss of her three young children. Scarlett had never seen her mother lean back against a chair when sitting, nor had she ever seen her sit down without some sewing work in her hands (except during meals), even when tending to the sick or auditing the farm's accounts. In the presence of guests, her hands were skilled in embroidery, while at other times she sewed Gerald's shirts, dresses for the girls, or clothes for the slaves. Scarlett found it hard to imagine her mother's hands without the golden-topped needle or her figure without the presence of the black girl who spent her life only unpicking her sewing and carrying the mahogany sewing box from one room to another as Ellen hurried through the house inspecting cooking, washing, and a multitude of sewing tasks.

图 11 ChatGPT 机器翻译任务——汉译英

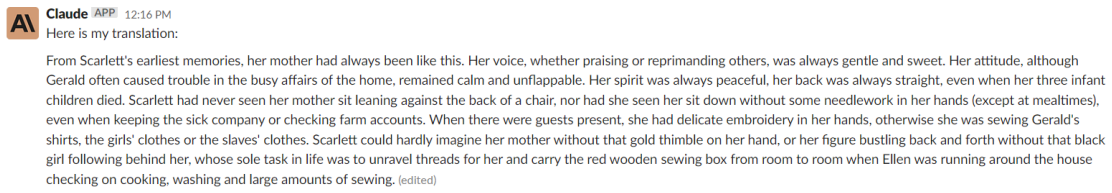




图 12 Claude 机器翻译任务——汉译英

3. **文本生成**，测试模型在生成自然语言文本方面的能力。根据给定的输入文本，生成准确的语义连贯的文本。评估模型生成文本的质量、流畅度和内容一致性。本报告设计了 1 个故事类生成任务，给定故事的开头“On that stormy night, wolf packs roared around the castle”，模型根据输入句子生成一篇完整的故事。



I give you a sentence that is the beginning of a story, then you need to generate a complete story based on that sentence. This sentence is "On that stormy night, wolf packs roared around the castle".



On that stormy night, wolf packs roared around the castle. The wind howled through the ancient turrets, carrying with it an eerie sense of foreboding. Inside the castle walls, a group of weary travelers sought refuge from the relentless downpour.

Among the travelers was a young adventurer named Jonathan, who had embarked on a quest to uncover the secrets of the haunted castle. Unbeknownst to him, the storm had trapped him and his companions within the castle's looming presence.

As the night wore on, the thunderous echoes of the wolves' calls intensified, causing an unsettling atmosphere to settle over the group. Their fear grew with every passing minute, and Jonathan knew they had to find a way to defend themselves.

Determined to protect his fellow travelers, Jonathan set out to explore the castle's hidden chambers. With a flickering torch in hand, he ventured deeper into the labyrinthine corridors, listening to the distant howls growing louder with each step.

In his search, Jonathan stumbled upon a forgotten armory. Rusted swords, battered shields, and dusty armor lined the walls. He selected the most formidable weapons he could find and armed himself, feeling a surge of confidence and purpose.

Returning to his companions, Jonathan shared his discovery and distributed the weapons among them. Together, they devised a plan to withstand the impending wolf attack. They would fortify the castle's main hall, using furniture and whatever else they could find to create barricades.

As the wolves closed in, their piercing howls reverberating through the castle, the group took their positions behind the makeshift barriers. Hearts pounding, they held their breath, ready to face the oncoming threat.

The first wolf burst through the castle doors, its eyes gleaming with a feral hunger. Battle ensued, with Jonathan and his companions fighting valiantly against the relentless onslaught. Swords clashed, shields held strong, and the echoes of their battle cries merged with the raging storm outside.

Time seemed to stretch as the travelers fought for their lives, their determination unwavering. They formed a formidable defense, driving the wolves back one by one. However, the pack was relentless, and the fight was far from over.

Just as hope began to waver, a deafening howl echoed through the castle, unlike any they had heard before. The remaining wolves retreated, vanishing into the night, leaving Jonathan and his companions victorious.

With the storm subsiding and the castle returning to an uneasy silence, the travelers emerged from their fortifications, their bodies battered but their spirits triumphant. Jonathan had fulfilled his quest, not only unraveling the mystery of the haunted castle but also proving the strength that can be found in unity and bravery.

As they made their way back to civilization, the travelers carried the memory of that stormy night and their hard-fought victory. And for Jonathan, it would forever serve as a reminder of the power that lies within oneself when faced with the darkest of nights.

图 13 ChatGPT 文本生成任务



On that stormy night, wolf packs roared around the castle

[View other drafts](#)



The wind howled outside, and the rain beat against the windows of the castle. Inside, the occupants huddled around the fire, trying to stay warm. Suddenly, there was a loud crash, and the door to the great hall burst open. A pack of wolves stood in the doorway, their eyes glowing in the firelight.

The people in the hall screamed and ran for cover. The wolves chased them, their teeth bared. Soon, the hall was filled with the sound of screams and howling.

A young woman named Anya stood in the corner of the hall, her heart pounding in her chest. She watched as the wolves attacked the others. She knew she had to do something, but she was too afraid to move.

Suddenly, a voice spoke from behind her. "Don't be afraid," the voice said. "I will protect you."

Anya turned around and saw a man standing there. He was tall and strong, with long, dark hair and a beard. He was dressed in a bearskin cloak, and he held a staff in his hand.

"Who are you?" Anya asked.

"I am the Guardian of the Forest," the man said. "I have come to protect you from the wolves."

The man raised his staff, and the wolves stopped in their tracks. They growled at him, but they did not attack.

"Go back to your forest," the Guardian said. "You are not welcome here."

The wolves turned and ran away, their tails between their legs.

The Guardian turned to Anya. "Are you hurt?" he asked.

Anya shook her head. "No," she said. "Thank you."

"You are welcome," the Guardian said. "Now, come with me. I will take you to safety."

The Guardian led Anya out of the castle and into the forest. They walked for a long time, until they came to a small cottage.

"This is my home," the Guardian said. "You will be safe here."

Anya thanked the Guardian and went inside the cottage. She was exhausted, and she soon fell asleep in a soft bed.

The next morning, Anya woke up to the smell of breakfast cooking. She went outside and saw the Guardian sitting by the fire. He was cooking a pot of stew.

"Good morning," the Guardian said. "I hope you slept well."

"I did," Anya said. "Thank you."

Anya sat down by the fire and ate breakfast with the Guardian. After breakfast, the Guardian showed Anya around the forest. They walked through the trees, and Anya saw all sorts of animals. She saw deer, rabbits, and even a bear.

Anya spent the next few days with the Guardian. She learned about the forest and the animals that lived there. She also learned about the Guardian's magic.

One day, the Guardian told Anya that it was time for her to go home. "You are safe now," he said. "The wolves will not bother you again."

Anya thanked the Guardian for his help, and she set off for home. She walked for a long time, and finally she came to the castle. She went inside and found her family. They were overjoyed to see her, and they hugged her tightly.

Anya told her family about her adventure in the forest. She told them about the wolves, and she told them about the Guardian. Her family was amazed by her story, and they were grateful to the Guardian for saving her.

Anya never forgot her time in the forest, and she never forgot the Guardian. She knew that she would always be grateful to him for saving her life.

图 14 Bard 文本生成任务

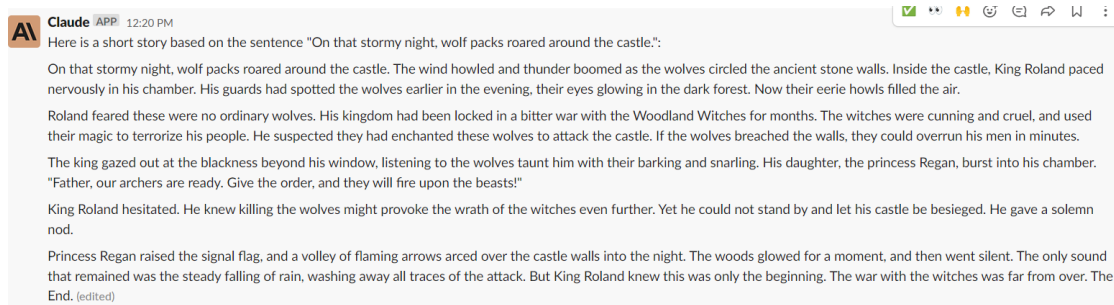


图 15 Claude 文本生成任务

三、结论

本文通过逻辑推理、机器翻译和文本生成三个下游任务发现 ChatGPT、Claude 和 Bard 三个模型在文本生成任务上表现良好，其中 ChatGPT 总体上取得了最好的结果，而 Bard 不支持机器翻译任务。此外，我们也发现模型在不同任务上的表现存在一些差异。例如，Bard 在文本生成方面的表现优于 ChatGPT 和 Claude，而 Claude 在逻辑推理方面的表现优于 ChatGPT 和 Bard，此外 Claude 与 ChatGPT 的性能表现不相上下。

实验结果表明，大型模型是各种下游任务的强大工具。但是，为手头的任务选择正确的模型很重要。通过了解不同模型的优缺点，可以选择最有可能达到最佳效果的模型。

除了本文提出的结果外，还有许多其他因素会影响 LLM 在下游任务上的表现。这些因素包括模型的大小、训练模型的数据类型以及评估模型的具体任务。随着 LLM 的规模和复杂性不断增加，很可能会看到它们在下游任务上的表现有更多改进。

参考文献

- [1] Chowdhery Aakanksha, Narang Sharan, Devlin Jacob, Bosma Maarten, Mishra Gaurav, Roberts Adam, Barham Paul, Chung Hyung Won, Sutton Charles, Gehrmann Sebastian, et al. 2022. PaLM: Scaling language modeling with pathways. arXiv preprint arXiv:2204.02311 (2022).
- [2] R. Thoppilan, D. De Freitas, J. Hall, N. Shazeer, A. Kulshreshtha, H.-T. Cheng, A. Jin, T. Bos, L. Baker, Y. Du, et al., “Lamda: Language models for dialog applications,” arXiv preprint arXiv:2201.08239, 2022.
- [3] Y. Sun, S. Wang, Y. Li, S. Feng, X. Chen, H. Zhang, X. Tian, D. Zhu, H. Tian, and H. Wu, “Ernie: Enhanced representation through knowledge integration,” arXiv, 2019.
- [4] Y. Sun, S. Wang, Y. Li, S. Feng, H. Tian, H. Wu, and H. Wang, “Ernie 2.0: A continual pre-training framework for language understanding,” in AAAI.
- [5] Sun Y., Wang S., Feng S., Ding S., Pang C., Shang J., Liu J., Chen X., Zhao Y., Lu Y. , et al. ERNIE 3.0: Large-scale Knowledge Enhanced Pre-training for Language Understanding and Generation. arXiv 2021, arXiv:2107.02137.