第六章 文本转换

大语言模型具有强大的文本转换能力,可以实现多语言翻译、拼写纠正、语法调整、格式转换等不同类型的文本转换任务。利用语言模型进行各类转换是它的典型应用之一。

在本章中,我们将介绍如何通过编程调用API接口,使用语言模型实现文本转换功能。通过代码示例,读者可以学习将输入文本转换成所需输出格式的具体方法。

掌握调用大语言模型接口进行文本转换的技能,是开发各种语言类应用的重要一步。文本转换功能的应用场景也非常广泛。相信读者可以在本章的基础上,利用大语言模型轻松开发出转换功能强大的程序。

一、文本翻译

文本翻译是大语言模型的典型应用场景之一。相比于传统统计机器翻译系统,大语言模型翻译更加流畅自然,还原度更高。通过在大规模高质量平行语料上进行 Fine-Tune,大语言模型可以深入学习不同语言间的词汇、语法、语义等层面的对应关系,模拟双语者的转换思维,进行意义传递的精准转换,而非简单的逐词替换。

以英译汉为例,传统统计机器翻译多倾向直接替换英文词汇,语序保持英语结构,容易出现中文词汇使 用不地道、语序不顺畅的现象。而大语言模型可以学习英汉两种语言的语法区别,进行动态的结构转 换。同时,它还可以通过上下文理解原句意图,选择合适的中文词汇进行转换,而非生硬的字面翻译。

大语言模型翻译的这些优势使其生成的中文文本更加地道、流畅,兼具准确的意义表达。利用大语言模型翻译,我们能够打通多语言之间的壁垒,进行更加高质量的跨语言交流。

1.1 翻译为西班牙语

```
from tool import get_completion

prompt = f"""

将以下中文翻译成西班牙语: \
```您好,我想订购一个搅拌机。```
"""

response = get_completion(prompt)
print(response)
```

Hola, me gustaría ordenar una batidora.

### 1.2 识别语种

```
prompt = f"""
请告诉我以下文本是什么语种:
    ```Combien coûte le lampadaire?```
"""
response = get_completion(prompt)
print(response)
```

这段文本是法语。

1.3 多语种翻译

```
prompt = f"""
请将以下文本分别翻译成中文、英文、法语和西班牙语:
    ```I want to order a basketball.``
"""
response = get_completion(prompt)
print(response)
```

```
中文: 我想订购一个篮球。
英文: I want to order a basketball.
法语: Je veux commander un ballon de basket.
西班牙语: Quiero pedir una pelota de baloncesto.
```

### 1.4 同时进行语气转换

```
prompt = f"""
请将以下文本翻译成中文,分别展示成正式与非正式两种语气:
    ```would you like to order a pillow?``
"""
response = get_completion(prompt)
print(response)
```

```
正式语气: 您是否需要订购一个枕头? 非正式语气: 你想要订购一个枕头吗?
```

1.5 通用翻译器

在当今全球化的环境下,不同国家的用户需要频繁进行跨语言交流。但是语言的差异常使交流变得困难。为了打通语言壁垒,实现更便捷的国际商务合作和交流,我们需要一个智能的**通用翻译工具**。该翻译工具需要能够自动识别不同语言文本的语种,无需人工指定。然后它可以将这些不同语言的文本翻译成目标用户的母语。在这种方式下,全球各地的用户都可以轻松获得用自己母语书写的内容。

开发一个识别语种并进行多语种翻译的工具,将大大降低语言障碍带来的交流成本。它将有助于构建一个语言无关的全球化世界,让世界更为紧密地连结在一起。

```
user_messages = [
 "La performance du système est plus lente que d'habitude.", # System
performance is slower than normal
  "Mi monitor tiene pixeles que no se iluminan.",
                                                               # My monitor has
pixels that are not lighting
 "Il mio mouse non funziona",
                                                               # My mouse is not
working
                                                               # My keyboard has
 "Mój klawisz Ctrl jest zepsuty",
a broken control key
 "我的屏幕在闪烁"
                                                             # My screen is
flashing
]
```

```
import time
for issue in user_messages:
```

```
time.sleep(20)
    prompt = f"告诉我以下文本是什么语种,直接输出语种,如法语,无需输出标点符号:

'``{issue}``"
    lang = get_completion(prompt)
    print(f"原始消息 ({lang}): {issue}\n")

prompt = f"""
    将以下消息分别翻译成英文和中文,并写成中文翻译: xxx
英文翻译: yyy
的格式:
    ```{issue}``"""

response = get_completion(prompt)
 print(response, "\n============")
```

```
原始消息 (法语): La performance du système est plus lente que d'habitude.
中文翻译:系统性能比平时慢。
英文翻译: The system performance is slower than usual.

原始消息 (西班牙语): Mi monitor tiene píxeles que no se iluminan.
中文翻译: 我的显示器有一些像素点不亮。
英文翻译: My monitor has pixels that do not light up.

原始消息 (意大利语): Il mio mouse non funziona
中文翻译: 我的鼠标不工作
英文翻译: My mouse is not working

原始消息 (这段文本是波兰语。): Mój klawisz Ctrl jest zepsuty
中文翻译: 我的Ctrl键坏了
英文翻译: My Ctrl key is broken

原始消息 (中文): 我的屏幕在闪烁
中文翻译: 我的屏幕在闪烁
英文翻译: My screen is flickering.

```

# 二、语气与写作风格调整

在写作中,语言语气的选择与受众对象息息相关。比如工作邮件需要使用正式、礼貌的语气和书面词 汇;而与朋友的聊天可以使用更轻松、口语化的语气。

选择恰当的语言风格,让内容更容易被特定受众群体所接受和理解,是技巧娴熟的写作者必备的能力。随着受众群体的变化调整语气也是大语言模型在不同场景中展现智能的一个重要方面。

```
prompt = f"""
将以下文本翻译成商务信函的格式:
```小老弟,我小羊,上回你说咱部门要采购的显示器是多少寸来着?```
"""
response = get_completion(prompt)
print(response)
```

```
尊敬的先生/女士,
我是小羊,我希望能够向您确认一下我们部门需要采购的显示器尺寸是多少寸。上次我们交谈时,您提到了这个问题。
期待您的回复。
谢谢!
此致,
```

三、文件格式转换

大语言模型如 ChatGPT 在不同数据格式之间转换方面表现出色。它可以轻松实现 JSON 到 HTML、XML、Markdown 等格式的相互转化。下面是一个示例,展示如何使用大语言模型**将 JSON 数据转换为 HTML 格式**:

假设我们有一个 JSON 数据,包含餐厅员工的姓名和邮箱信息。现在我们需要将这个 JSON 转换为 HTML 表格格式,以便在网页中展示。在这个案例中,我们就可以使用大语言模型,直接输入JSON 数据,并给出需要转换为 HTML 表格的要求。语言模型会自动解析 JSON 结构,并以 HTML 表格形式输出,完成格式转换的任务。

利用大语言模型强大的格式转换能力,我们可以快速实现各种结构化数据之间的相互转化,大大简化开发流程。掌握这一转换技巧将有助于读者**更高效地处理结构化数据**。

```
data_json = { "resturant employees" :[
          {"name":"Shyam", "email":"shyamjaiswal@gmail.com"},
          {"name":"Bob", "email":"bob32@gmail.com"},
          {"name":"Jai", "email":"jai87@gmail.com"}
]}
```

```
prompt = f"""
将以下Python字典从JSON转换为HTML表格,保留表格标题和列名: {data_json}
"""
response = get_completion(prompt)
print(response)
```

```
</thead>
Shyam
  shyamjaiswal@gmail.com
 Bob
  bob32@gmail.com
 Jai
  jai87@gmail.com
```

将上述 HTML 代码展示出来如下:

```
from IPython.display import display, Markdown, Latex, HTML, JSON
display(HTML(response))
```

resturant employees

name	email
Shyam	shyamjaiswal@gmail.com
Bob	bob32@gmail.com
Jai	jai87@gmail.com

四、拼写及语法纠正

在使用非母语撰写时,拼写和语法错误比较常见,进行校对尤为重要。例如在论坛发帖或撰写英语论文时,校对文本可以大大提高内容质量。

利用大语言模型进行自动校对可以极大地降低人工校对的工作量。下面是一个示例,展示如何使用大语言模型检查句子的拼写和语法错误。

假设我们有一系列英语句子,其中部分句子存在错误。我们可以遍历每个句子,要求语言模型进行检查,如果句子正确就输出"未发现错误",如果有错误就输出修改后的正确版本。

通过这种方式,大语言模型可以快速自动校对大量文本内容,定位拼写和语法问题。这极大地减轻了人工校对的负担,同时也确保了文本质量。利用语言模型的校对功能来提高写作效率,是每一位非母语写作者都可以采用的有效方法。

```
text = [
   "The girl with the black and white puppies have a ball.", # The girl has a
ball.
   "Yolanda has her notebook.", # ok
   "Its going to be a long day. Does the car need it's oil changed?", # Homonyms
   "Their goes my freedom. There going to bring they're suitcases.", # Homonyms
   "Your going to need you're notebook.", # Homonyms
   "That medicine effects my ability to sleep. Have you heard of the butterfly
affect?", # Homonyms
   "This phrase is to cherck chatGPT for spelling abilitty" # spelling
]
```

```
for i in range(len(text)):
    time.sleep(20)
    prompt = f"""请校对并更正以下文本,注意纠正文本保持原始语种,无需输出原始文本。
    如果您没有发现任何错误,请说"未发现错误"。

例如:
    输入: I are happy.
    输出: I am happy.
    ```{text[i]}```""
 response = get_completion(prompt)
 print(i, response)
```

```
0 The girl with the black and white puppies has a ball.
1 Yolanda has her notebook.
2 It's going to be a long day. Does the car need its oil changed?
3 Their goes my freedom. There going to bring their suitcases.
4 You're going to need your notebook.
5 That medicine affects my ability to sleep. Have you heard of the butterfly effect?
6 This phrase is to check chatGPT for spelling ability.
```

下面是一个使用大语言模型进行语法纠错的简单示例,类似于Grammarly(一个语法纠正和校对的工具)的功能。

输入一段关于熊猫玩偶的评价文字,语言模型会自动校对文本中的语法错误,输出修改后的正确版本。 这里使用的Prompt比较简单直接,只要求进行语法纠正。我们也可以通过扩展Prompt,同时请求语言 模型调整文本的语气、行文风格等。

```
text = f"""

Got this for my daughter for her birthday cuz she keeps taking \
mine from my room. Yes, adults also like pandas too. She takes \
it everywhere with her, and it's super soft and cute. One of the \
ears is a bit lower than the other, and I don't think that was \
designed to be asymmetrical. It's a bit small for what I paid for it \
though. I think there might be other options that are bigger for \
the same price. It arrived a day earlier than expected, so I got \
to play with it myself before I gave it to my daughter.
"""

prompt = f"校对并更正以下商品评论: ```{text}```"
response = get_completion(prompt)
```

```
print(response)
```

I got this for my daughter's birthday because she keeps taking mine from my room. Yes, adults also like pandas too. She takes it everywhere with her, and it's super soft and cute. However, one of the ears is a bit lower than the other, and I don't think that was designed to be asymmetrical. It's also a bit smaller than I expected for the price. I think there might be other options that are bigger for the same price. On the bright side, it arrived a day earlier than expected, so I got to play with it myself before giving it to my daughter.

引入 Redlines 包,详细显示并对比纠错过程:

# 如未安装redlines, 需先安装

diff = Redlines(text, response)

display(Markdown(diff.output\_markdown))

got to play with it myself before I gave giving it to my daughter.

```
!pip3.8 install redlines

from redlines import Redlines
from IPython.display import display, Markdown
```

Got-I got this for my daughter for her daughter's birthday cuz because she keeps taking mine from my room. Yes, adults also like pandas too. She takes it everywhere with her, and it's super soft and cute. One cute. However, one of the ears is a bit lower than the other, and I don't think that was designed to be asymmetrical. It's also a bit small-smaller than I expected for what I paid for it though. the price. I think there might be other options that are bigger for the same price. It-price. On the bright side, it arrived a day earlier than expected, so I

这个示例展示了如何利用语言模型强大的语言处理能力实现自动化的语法纠错。类似的方法可以运用于 校对各类文本内容,大幅减轻人工校对的工作量,同时确保文本语法准确。掌握运用语言模型进行语法 纠正的技巧,将使我们的写作更加高效和准确。

### 五、综合样例

语言模型具有强大的组合转换能力,可以通过一个Prompt同时实现多种转换,大幅简化工作流程。

下面是一个示例,展示了如何使用一个Prompt,同时对一段文本进行翻译、拼写纠正、语气调整和格式转换等操作。

```
prompt = f"""
针对以下三个反引号之间的英文评论文本,
首先进行拼写及语法纠错,
然后将其转化成中文,
再将其转化成优质淘宝评论的风格,从各种角度出发,分别说明产品的优点与缺点,并进行总结。
润色一下描述,使评论更具有吸引力。
输出结果格式为:
【优点】xxx
【缺点】xxx
【缺点】xxx
【缺点】xxx
【i转】xxx
【i转】xxx
**注意,只需填写xxx部分,并分段输出。
**将结果输出成Markdown格式。
{text}*
```

```
response = get_completion(prompt)
display(Markdown(response))
```

#### 【优点】

- 超级柔软可爱, 女儿生日礼物非常受欢迎。
- 成人也喜欢熊猫,我也很喜欢它。
- 提前一天到货,让我有时间玩一下。

#### 【缺点】

- 一只耳朵比另一只低,不对称。
- 价格有点贵,但尺寸有点小,可能有更大的同价位选择。

#### 【总结】

这只熊猫玩具非常适合作为生日礼物,柔软可爱,深受孩子喜欢。虽然价格有点贵,但尺寸有点小,不对称的设计也有点让人失望。如果你想要更大的同价位选择,可能需要考虑其他选项。总的来说,这是一款不错的熊猫玩具,值得购买。

通过这个例子,我们可以看到大语言模型可以流畅地处理多个转换要求,实现中文翻译、拼写纠正、语气升级和格式转换等功能。

利用大语言模型强大的组合转换能力,我们可以避免多次调用模型来进行不同转换,极大地简化了工作流程。这种一次性实现多种转换的方法,可以广泛应用于文本处理与转换的场景中。

# 六、英文版

#### 1.1 翻译为西班牙语

```
prompt = f"""
Translate the following English text to Spanish: \
```Hi, I would like to order a blender```
"""
response = get_completion(prompt)
print(response)
```

Hola, me gustaría ordenar una licuadora.

1.2 识别语种

This language is French.

1.3 多语种翻译

```
prompt = f"""
Translate the following text to French and Spanish
and English pirate: \
```I want to order a basketball```
"""
response = get_completion(prompt)
print(response)
```

```
French: ```Je veux commander un ballon de basket```
Spanish: ```Quiero ordenar una pelota de baloncesto```
English: ```I want to order a basketball```
```

#### 1.4 同时进行语气转换

```
prompt = f"""
Translate the following text to Spanish in both the \
formal and informal forms:
'would you like to order a pillow?'
"""
response = get_completion(prompt)
print(response)
```

```
Formal: ¿Le gustaría ordenar una almohada?
Informal: ¿Te gustaría ordenar una almohada?
```

#### 1.5 通用翻译器

```
user_messages = [
"La performance du système est plus lente que d'habitude.", # System
performance is slower than normal
"Mi monitor tiene píxeles que no se iluminan.", # My monitor has
pixels that are not lighting
"Il mio mouse non funziona", # My mouse is not
working
"Mój klawisz Ctrl jest zepsuty", # My keyboard has
a broken control key
"我的屏幕在闪烁" # My screen is
flashing
]
```

```
for issue in user_messages:
 prompt = f"Tell me what language this is: ```{issue}```"
 lang = get_completion(prompt)
 print(f"Original message ({lang}): {issue}")

 prompt = f"""
 Translate the following text to English \
 and Korean: ```{issue}```
 """
 response = get_completion(prompt)
 print(response, "\n")
```

```
Original message (The language is French.): La performance du système est plus
lente que d'habitude.
The performance of the system is slower than usual.
시스템의 성능이 평소보다 느립니다.
Original message (The language is Spanish.): Mi monitor tiene píxeles que no se
English: "My monitor has pixels that do not light up."
Korean: "내 모니터에는 밝아지지 않는 픽셀이 있습니다."
Original message (The language is Italian.): Il mio mouse non funziona
English: "My mouse is not working."
Korean: "내 마우스가 작동하지 않습니다."
Original message (The language is Polish.): Mój klawisz Ctrl jest zepsuty
English: "My Ctrl key is broken"
Korean: "내 Ctrl 키가 고장 났어요"
Original message (The language is Chinese.): 我的屏幕在闪烁
English: My screen is flickering.
Korean: 내 화면이 깜박거립니다.
```

#### 2.1 语气风格调整

```
prompt = f"""
Translate the following from slang to a business letter:
'Dude, This is Joe, check out this spec on this standing lamp.'
"""
response = get_completion(prompt)
print(response)
```

```
Dear Sir/Madam,

I hope this letter finds you well. My name is Joe, and I am writing to bring your attention to a specification document regarding a standing lamp.

I kindly request that you take a moment to review the attached document, as it provides detailed information about the features and specifications of the aforementioned standing lamp.

Thank you for your time and consideration. I look forward to discussing this further with you.

Yours sincerely, Joe
```

#### 3.1 文件格式转换

```
data_json = { "resturant employees" :[
 {"name":"Shyam", "email":"shyamjaiswal@gmail.com"},
 {"name":"Bob", "email":"bob32@gmail.com"},
 {"name":"Jai", "email":"jai87@gmail.com"}
]}
```

```
prompt = f"""
Translate the following python dictionary from JSON to an HTML \
table with column headers and title: {data_json}
"""
response = get_completion(prompt)
print(response)
```

```
<!DOCTYPE html>
<html>
<head>
<style>
table {
 font-family: arial, sans-serif;
 border-collapse: collapse;
 width: 100%;
}
td, th {
 border: 1px solid #dddddd;
 text-align: left;
 padding: 8px;
}
tr:nth-child(even) {
 background-color: #dddddd;
}
</style>
</head>
<body>
```

```
<h2>Restaurant Employees</h2>
Name
 Email
Shyam
 shyamjaiswal@gmail.com
Bob
 bob32@gmail.com
Jai
 jai87@gmail.com
</body>
</html>
```

```
from IPython.display import display, Markdown, Latex, HTML, JSON
display(HTML(response))

td, th {
 border: 1px solid #dddddd;
 text-align: left;
 padding: 8px;
}

tr:nth-child(even) {
 background-color: #dddddd;
}
```

# **Restaurant Employees**

Name	Email
Shyam	shyamjaiswal@gmail.com
Bob	bob32@gmail.com
Jai	jai87@gmail.com

#### 4.1 拼写及语法纠错

```
text = [
 "The girl with the black and white puppies have a ball.", # The girl has a
ball.
 "Yolanda has her notebook.", # ok
 "Its going to be a long day. Does the car need it's oil changed?", # Homonyms
 "Their goes my freedom. There going to bring they're suitcases.", # Homonyms
 "Your going to need you're notebook.", # Homonyms
 "That medicine effects my ability to sleep. Have you heard of the butterfly
affect?", # Homonyms
 "This phrase is to cherck chatGPT for spelling abilitty" # spelling
]
```

```
for t in text:
 prompt = f"""Proofread and correct the following text
 and rewrite the corrected version. If you don't find
 and errors, just say "No errors found". Don't use
 any punctuation around the text:
    ```{t}``"""
    response = get_completion(prompt)
    print(response)
```

```
The girl with the black and white puppies has a ball.

No errors found.

It's going to be a long day. Does the car need its oil changed?

There goes my freedom. They're going to bring their suitcases.

You're going to need your notebook.

That medicine affects my ability to sleep. Have you heard of the butterfly effect?

This phrase is to check chatGPT for spelling ability.
```

```
text = f"""
Got this for my daughter for her birthday cuz she keeps taking \
mine from my room. Yes, adults also like pandas too. She takes \
it everywhere with her, and it's super soft and cute. One of the \
ears is a bit lower than the other, and I don't think that was \
designed to be asymmetrical. It's a bit small for what I paid for it \
though. I think there might be other options that are bigger for \
the same price. It arrived a day earlier than expected, so I got \
to play with it myself before I gave it to my daughter.
"""
```

```
prompt = f"proofread and correct this review: ```{text}```"
response = get_completion(prompt)
print(response)
```

Got this for my daughter for her birthday because she keeps taking mine from my room. Yes, adults also like pandas too. She takes it everywhere with her, and it's super soft and cute. However, one of the ears is a bit lower than the other, and I don't think that was designed to be asymmetrical. Additionally, it's a bit small for what I paid for it. I believe there might be other options that are bigger for the same price. On the positive side, it arrived a day earlier than expected, so I got to play with it myself before I gave it to my daughter.

```
from redlines import Redlines
from IPython.display import display, Markdown

diff = Redlines(text,response)
display(Markdown(diff.output_markdown))
```

Got this for my daughter for her birthday cuz-because she keeps taking mine from my room. Yes, adults also like pandas too. She takes it everywhere with her, and it's super soft and cute. One cute. However, one of the ears is a bit lower than the other, and I don't think that was designed to be asymmetrical. It's Additionally, it's a bit small for what I paid for it though. it. I think believe there might be other options that are bigger for the same price. It price. On the positive side, it arrived a day earlier than expected, so I got to play with it myself before I gave it to my daughter.

daughter.

5.1 综合样例

```
text = f"""
Got this for my daughter for her birthday cuz she keeps taking \
mine from my room. Yes, adults also like pandas too. She takes \
it everywhere with her, and it's super soft and cute. One of the \
ears is a bit lower than the other, and I don't think that was \
designed to be asymmetrical. It's a bit small for what I paid for it \
though. I think there might be other options that are bigger for \
the same price. It arrived a day earlier than expected, so I got \
to play with it myself before I gave it to my daughter.
"""
```

```
prompt = f"""
proofread and correct this review. Make it more compelling.
Ensure it follows APA style guide and targets an advanced reader.
Output in markdown format.
Text: ```{text}``
"""

# 校对注: APA style guide是APA Style Guide是一套用于心理学和相关领域的研究论文写作和格式化的规则。
# 它包括了文本的缩略版,旨在快速阅读,包括引用、解释和参考列表,
# 其详细内容可参考: https://apastyle.apa.org/about-apa-style
# 下一单元格内的汉化prompt内容由译者进行了本地化处理,仅供参考
response = get_completion(prompt)
display(Markdown(response))
```

Title: A Delightful Gift for Panda Enthusiasts: A Review of the Soft and Adorable Panda Plush Toy

Reviewer: [Your Name]

I recently purchased this charming panda plush toy as a birthday gift for my daughter, who has a penchant for "borrowing" my belongings from time to time. As an adult, I must admit that I too have fallen under the spell of these lovable creatures. This review aims to provide an in-depth analysis of the product, catering to advanced readers who appreciate a comprehensive evaluation.

First and foremost, the softness and cuteness of this panda plush toy are simply unparalleled. Its irresistibly plush exterior makes it a joy to touch and hold, ensuring a delightful sensory experience for both children and adults alike. The attention to detail is evident, with its endearing features capturing the essence of a real panda. However, it is worth noting that one of the ears appears to be slightly asymmetrical, which may not have been an intentional design choice.

While the overall quality of the product is commendable, I must express my slight disappointment regarding its size in relation to its price. Considering the investment made, I expected a larger plush toy. It is worth exploring alternative options that offer a more substantial size for the same price point. Nevertheless, this minor setback does not overshadow the toy's undeniable appeal and charm.

In terms of delivery, I was pleasantly surprised to receive the panda plush toy a day earlier than anticipated. This unexpected early arrival allowed me to indulge in some personal playtime with the toy before presenting it to my daughter. Such promptness in delivery is a testament to the seller's efficiency and commitment to customer satisfaction.

In conclusion, this panda plush toy is a delightful gift for both children and adults who appreciate the enchanting allure of these beloved creatures. Its softness, cuteness, and attention to detail make it a truly captivating addition to any collection. While the size may not fully justify the price, the overall quality and prompt delivery make it a worthwhile purchase. I highly recommend this panda plush toy to anyone seeking a charming and endearing companion.

Word Count: 305 words