

## 第三章 评估输入——分类

在本章中，我们将重点探讨评估输入任务的重要性，这关乎到整个系统的质量和安全性。

在处理不同情况下的多个独立指令集的任务时，首先对查询类型进行分类，并以此为基础确定要使用哪些指令，具有诸多优势。这可以通过定义固定类别和硬编码与处理特定类别任务相关的指令来实现。例如，在构建客户服务助手时，对查询类型进行分类并根据分类确定要使用的指令可能非常关键。具体来说，如果用户要求关闭其账户，那么二级指令可能是添加有关如何关闭账户的额外说明；如果用户询问特定产品信息，则二级指令可能会提供更多的产品信息。

```
delimiter = "####"
```

在这个例子中，我们使用系统消息（system\_message）作为整个系统的全局指导，并选择使用“#”作为分隔符。分隔符是用来区分指令或输出中不同部分的工具，它可以帮助模型更好地识别各个部分，从而提高系统在执行特定任务时的准确性和效率。“#”也是一个理想的分隔符，因为它可以被视为一个单独的 token。

这是我们定义的系统消息，我们正在以下面的方式询问模型。

```
system_message = f"""
你将获得客户服务查询。
每个客户服务查询都将用{delimiter}字符分隔。
将每个查询分类到一个主要类别和一个次要类别中。
以 JSON 格式提供你的输出，包含以下键：primary 和 secondary。

主要类别：计费（Billing）、技术支持（Technical Support）、账户管理（Account Management）
或一般咨询（General Inquiry）。

计费次要类别：
取消订阅或升级（Unsubscribe or upgrade）
添加付款方式（Add a payment method）
收费解释（Explanation for charge）
争议费用（Dispute a charge）

技术支持次要类别：
常规故障排除（General troubleshooting）
设备兼容性（Device compatibility）
软件更新（Software updates）

账户管理次要类别：
重置密码（Password reset）
更新个人信息（Update personal information）
关闭账户（Close account）
账户安全（Account security）

一般咨询次要类别：
产品信息（Product information）
定价（Pricing）
反馈（Feedback）
与人工对话（Speak to a human）

"""
```

了解了系统消息后，现在让我们来看一个用户消息（user message）的例子。

```
user_message = f"""\n
我希望你删除我的个人资料和所有用户数据。"""
```

首先，将这个用户消息格式化为一个消息列表，并将系统消息和用户消息之间使用 "#####" 进行分隔。

```
messages = [
    {'role': 'system',
     'content': system_message},
    {'role': 'user',
     'content': f"{delimiter}{user_message}{delimiter}"},
]
```

如果让你来判断，下面这句话属于哪个类别：“我想让您删除我的个人资料。我们思考一下，这句话似乎看上去属于“账户管理（Account Management）”或者属于“关闭账户（Close account）”。

让我们看看模型是如何思考的：

```
from tool import get_completion_from_messages

response = get_completion_from_messages(messages)
print(response)
```

```
{
  "primary": "账户管理",
  "secondary": "关闭账户"
}
```

模型的分类是将“账户管理”作为“primary”，“关闭账户”作为“secondary”。

请求结构化输出（如 JSON）的好处是，您可以轻松地将其读入某个对象中，例如 Python 中的字典。如果您使用其他语言，也可以转换为其他对象，然后输入到后续步骤中。

下面让我们再看一个例子：

```
用户消息：“告诉我更多关于你们的平板电脑的信息”
```

我们运用相同的消息列表来获取模型的响应，然后打印出来。

```
user_message = f"""\n
告诉我更多有关你们的平板电脑的信息"""
messages = [
    {'role': 'system',
     'content': system_message},
    {'role': 'user',
     'content': f"{delimiter}{user_message}{delimiter}"},
]
response = get_completion_from_messages(messages)
print(response)
```

```
{
  "primary": "一般咨询",
  "secondary": "产品信息"
}
```

这里返回了另一个分类结果，并且看起来似乎是正确的。因此，根据客户咨询的分类，我们现在可以提供一套更具体的指令来处理后续步骤。在这种情况下，我们可能会添加关于平板电脑的额外信息，而在其他情况下，我们可能希望提供关闭账户的链接或类似的内容。这里返回了另一个分类结果，并且看起来应该是正确的。

在下一章中，我们将探讨更多关于评估输入的方法，特别是如何确保用户以负责任的方式使用系统。

## 英文版

```
system_message = f"""
You will be provided with customer service queries. \
The customer service query will be delimited with \
{delimiter} characters.
Classify each query into a primary category \
and a secondary category.
Provide your output in json format with the \
keys: primary and secondary.
```

```
Primary categories: Billing, Technical Support, \
Account Management, or General Inquiry.
```

```
Billing secondary categories:
Unsubscribe or upgrade
Add a payment method
Explanation for charge
Dispute a charge
```

```
Technical Support secondary categories:
General troubleshooting
Device compatibility
Software updates
```

```
Account Management secondary categories:
Password reset
Update personal information
Close account
Account security
```

```
General Inquiry secondary categories:
Product information
Pricing
Feedback
Speak to a human
```

```
"""
```

```
user_message = f"""
I want you to delete my profile and all of my user data"""
```

```
messages = [  
    {'role': 'system',  
     'content': system_message},  
    {'role': 'user',  
     'content': f"{delimiter}{user_message}{delimiter}"},  
]
```

```
response = get_completion_from_messages(messages)  
print(response)
```

```
{  
    "primary": "Account Management",  
    "secondary": "Close account"  
}
```

```
user_message = f"""\  
Tell me more about your flat screen tvs"""  
messages = [  
    {'role': 'system',  
     'content': system_message},  
    {'role': 'user',  
     'content': f"{delimiter}{user_message}{delimiter}"},  
]  
response = get_completion_from_messages(messages)  
print(response)
```

```
{  
    "primary": "General Inquiry",  
    "secondary": "Product information"  
}
```