

创建文本对话请求

POSThttps://apis.iflow.cn/v1/chat/completions

Content-Type: application/json

请求参数说明

Authorization

类型string

位置header

是否必填是

描述

使用以下格式进行身份验证：

Bearer <your api key>

（访问 心流 官网登陆获取API KEY）。

LLM 模型

参数名	类型	是否必填	默认值	描述
messages	object	是	-	构成当前对话的消息列表。
messages.content	string	是	中国大模型行业2025年将会迎来哪些机遇和挑战?	消息的内容。
messages.role	enum<string>	是	user	消息作者的角色。可选值：user, assistant, system
model	enum<string>	是	tstars2.0	对应的模型名称。为更好的提升服务质量,我们将不定期对本服务提供的模型做相关变更,包括但不限于模型上下线、模型服务能力调整,我们会在可行的情况下以公告、消息推送等适当的方式进行通知。支持的模型请参考快速开始页面。
frequency_penalty	number	否	0.5	调整生成 token 的频率惩罚, 用于控制重复性。
max_tokens	integer	否	512	生成的最大 token 数量。取值范围：1 < x < 8192
n	integer	否	1	返回的生成结果数量。
response_format	object	否	-	指定模型输出格式的对象。
response_format.type	string	否	-	响应格式的类型。
stop	string	null	否	-
stream	boolean	否	false	如果设置为 true, token 将作为服务器发送事件 (SSE) 逐步返回。
temperature	number	否	0.7	控制响应的随机性。值越低, 输出越确定; 值越高, 输出越随机。

参数名	类型	是否必填	默认值	描述
<code>tools</code>	<code>object</code>	否	-	模型可能调用的工具列表。目前仅支持函数作为工具。使用此参数提供一个函数列表，模型可能会为其生成 JSON 输入。最多支持 128 个函数。
<code>tools.function</code>	<code>object</code>	否	-	函数对象。
<code>tools.function.name</code>	<code>string</code>	否	-	要调用的函数名称。必须由字母、数字、下划线或短横线组成，最大长度为 64。
<code>tools.function.description</code>	<code>string</code>	否	-	函数的描述，用于模型选择何时以及如何调用该函数。
<code>tools.function.parameters</code>	<code>object</code>	否	-	函数接受的参数，描述为 JSON Schema 对象。如果不指定参数，则定义了一个空参数列表的函数。
<code>tools.function.strict</code>	<code>boolean</code>	<code>null</code>	否	<code>false</code>
<code>tools.type</code>	<code>enum&lt;string&gt;</code>	否	<code>function</code>	工具的类型。目前仅支持 <code>function</code> 。
<code>top_k</code>	<code>number</code>	否	<code>50</code>	限制 token 选择范围为前 k 个候选。
<code>top_p</code>	<code>number</code>	否	<code>0.7</code>	核采样参数，用于根据累积概率动态调整每个预测 token 的选择范围。

请求举例

CURL Python Javascript Java

```
curl --request POST \
  --url https://apis.iflow.cn/v1/chat/completions \
  --header 'Authorization: Bearer <iflow API KEY>' \
  --header 'Content-Type: application/json' \
  --data '{
    "model": "tstars2.0",
    "messages": [
      {
        "role": "user",
        "content": "中国大模型行业2025年将会迎来哪些机遇和挑战？"
      }
    ],
    "stream": false,
    "max_tokens": 512,
    "stop": [
      "null"
    ],
    "temperature": 0.7,
    "top_p": 0.7,
    "top_k": 50,
    "frequency_penalty": 0.5,
    "n": 1,
    "response_format": {
      "type": "text"
    },
    "tools": [
      {
        "type": "function",
        "function": {
          "description": "<string>",
          "name": "<string>",
          "parameters": {},
          "strict": false
        }
      }
    ]
  },'
```

响应参数

非流式输出    流式输出

参数名	类型	是否必填	默认值	描述
choices	object[]	是	-	模型生成的选择列表。
choices.finish_reason	enum<string>	否	-	生成结束的原因。可选值： - stop : 自然结束。 - eos : 到达句子结束符。 - length : 达到最大 token 长度限制。 - tool_calls : 调用了工具（如函数）。
choices.message	object	是	-	模型返回的消息对象。
created	integer	是	-	响应生成的时间戳。
id	string	是	-	响应的唯一标识符。
model	string	是	-	使用的模型名称。
object	enum<string>	是	-	响应类型。可选值： - chat.completion : 表示这是一个聊天完成响应。
tool_calls	object[]	否	-	模型生成的工具调用，例如函数调用。
tool_calls.function	object	否	-	模型调用的函数。
tool_calls.function.arguments	string	否	-	函数调用的参数，由模型以 JSON 格式生成。注意：模型生成的 JSON 可能无效，或者可能会生成不属于函数定义的参数。在调用函数前，请在代码中验证这些参数。
tool_calls.function.name	string	否	-	要调用的函数名称。
tool_calls.id	string	否	-	工具调用的唯一标识符。
tool_calls.type	enum<string>	否	-	工具的类型。目前仅支持 function。 。可选值： - function : 表示这是一个函数调用。
usage	object	是	-	Token 使用情况统计。
usage.completion_tokens	integer	是	-	完成部分使用的 token 数量。
usage.prompt_tokens	integer	是	-	提示部分使用的 token 数量。
usage.total_tokens	integer	是	-	总共使用的 token 数量。

响应信息

非流式    流式

```
{
  "id": "<string>",
  "choices": [
    {
      "message": {
        "role": "assistant",
        "content": "<string>",
        "reasoning_content": "<string>"
      },
      "finish_reason": "stop"
    }
  ],
  "tool_calls": [
    {
      "id": "<string>",
      "type": "function",
      "function": {
        "name": "<string>",
```

```
        "arguments": "<string>"
      }
    }
  ],
  "usage": {
    "prompt_tokens": 123,
    "completion_tokens": 123,
    "total_tokens": 123
  },
  "created": 123,
  "model": "<string>",
  "object": "chat.completion"
}
```