

OpenAI学习

OpenAI SDK 使用

提纲

- 概要介绍
- 一、文生文
- 二、文生音
- 三、音生文
- 四、Assistants
- 五、Tokens计算工具
- 六、遇到的问题
- 七、参考资料

概要介绍

- OpenAI SDK提供与OpenAI API官方文档相契合的函数，方便开发者使用。也可直接使用官方API进行调用。
- OpenAI SDK可以根据不同场景使用不同的能力，如：文本生成模型、助理、代币计算、嵌入。可以快速集成到应用程序。
- 此分享旨在对于OpenAI技术预研做个总结。
- 技术预研使用OpenAI SDK版本为v1.17.8，GPT模型使用的是GPT-3.5，文本生成模型使用gpt-3.5-turbo-1106

一、文生文

- OpenAI 的文本生成模型经过训练可以理解自然语言、代码和图像。这些模型提供文本输出来响应其输入。这些模型的输入也称为“提示”。设计提示本质上是如何“编程”大型语言模型，通常是通过提供说明或一些如何成功完成任务的示例。
- 模型gpt-3.5-turbo-1106 的上下文限制16,385 个代币，指的是文本输入和文本输出的总数量不能超过16,385 个代币。最新的 GPT-3.5 Turbo 模型具有改进的指令跟踪、JSON 模式、可重现的输出、并行函数调用等。最多返回 4,096 个输出标记
- 代码示例

二、文生音

- 音频 API 提供speech基于我们的TTS（文本转语音）模型的端点。它带有 6 种内置声音，可用于：
 - 叙述一篇书面博客文章
 - 制作多种语言的语音音频
 - 使用流式传输提供实时音频输出
- 语音合成
 - tts-1,最新的文本转语音模型，针对速度进行了优化。
 - tts-1-hd（高清版）最新的文本转语音模型，针对质量进行了优化。
- 代码示例

三、音生文

- 音频 API 提供两个语音转文本端点，transcriptions并且translations基于我们最先进的开源大型 v2 Whisper 模型。它们可用于：
 - 将音频转录成音频所使用的任何语言。
 - 将音频翻译并转录成英语（目前仅支持英文）。
 - 文件上传当前限制为 25 MB，并且支持以下输入文件类型：mp3、mp4、mpeg、mpga、m4a、wav和webm。

四、Assistants

- Assistants API 允许您在自己的应用程序中构建 AI 助手。助手有指令，可以利用模型、工具和知识来响应用户查询。Assistants API 目前支持三种类型的工具：代码解释器、检索和函数调用。
- Assistants API 旨在帮助开发人员构建能够执行各种任务的强大 AI 助手。
- 助手工具
 - 代码解释器
 - 知识检索
 - 函数调用
 - 支持的文件
- 代码示例

五、Tokens计算工具

- 这个主要是为文本生成模型服务的，用来计算文本的token长度，避免超长。可以在文本输入后，进行计算token长度，如果token超长直接提示错误，也节省了API调用的资源。
- tiktoken-go是官方提供的go库，其算法和官方文档一致，可在应用程序中使用。

六、遇到的问题

- openai的api是否有token长度限制，长文本可以通过api调用实现吗？
- 助手中的函数调用，这个函数在哪里实现，openai是如何执行自定义的函数的？
- 为什么助手功能的函数调用的结果原样发送给AI，AI返回的结果与函数调用返回的结果不同？如何优化让其趋于一致？

八、参考资料

- openai官方文档
 - <https://platform.openai.com/docs/introduction>
- openai官方API文档
 - <https://platform.openai.com/docs/api-reference>
- openai sdk文档
 - <https://github.com/sashabaranov/go-openai?tab=readme-ov-file>
- openai sdk示例代码
 - <https://github.com/likai1130/openai-sdk-experiment>
- tokens计算方式
 - <https://github.com/pkoukk/tiktoken-go#counting-tokens-for-chat-api-calls>
- OpenAI开发者论坛
 - <https://community.openai.com/>

谢谢