

Deepseek企业级Agent项目开发实战

Part 4. Ollama REST API - OpenAI Compatibility

本节内容我们来看一下 OpenAI Compatibility。OpenAI 的 API 接口是大模型应用开发中最常用、且集成度最高的 API 接口规范，其兼容接口主要包括：

- chat/completions
- completions
- models
- embeddings

我们上两节内容中介绍的 /api/generate 和 /api/chat 接口，其实就是 Ollama 兼容 OpenAI 的 REST API 接口的底层实现。其中：

- /api/generate 接口对应 OpenAI 的 completions 接口；
- /api/chat 接口对应 OpenAI 的 chat/completions 接口；

因此我们现在再来看 ollama 中的 OpenAI compatibility 的 API 接口调用，就非常容易理解了。

- 非流式调用

```
from openai import OpenAI # type: ignore

client = OpenAI(
    base_url='http://192.168.110.131:11434/v1/', # 这里修改成可访问的 IP
    api_key='ollama', # 这里随便写，但是api_key字段一定要有
)

chat_completion = client.chat.completions.create(
    model='deepseek-r1:32b',
    messages=[
        {
            'role': 'user',
            'content': '你好，请你介绍一下什么是人工智能？',
        }
    ],
)

print(chat_completion.choices[0].message.content)
```

<think>

好吧，我现在要了解一下什么是人工智能。嗯，首先，这个词听起来有点高科技的感觉，对吧？但具体是什么意思呢？

我记得人工智能常被缩写为AI，可能和计算机有关。是不是说让机器像人一样思考？好像是这样。比如现在的手机、电脑或者机器人能做些智能的事情，可能都属于这个范畴。

那人工智能有什么具体的例子呢？比如说Siri或者小爱同学这些语音助手，它们能听懂我的话并回答问题，是不是也是AI的一部分？还有AlphaGo战胜围棋世界冠军的新闻，这应该就是AI的应用吧。那这些都是在模拟人类的思考和决策过程吗？

然后，听说有两种类型的人工智能，一种是弱人工智能，一种是强人工智能。弱人工智能好像是针对特定任务设计的，比如识别图像或翻译语言，这种不需要具备真正的理解和感受能力，对吧？而强人工智能则更接近人类，有自我意识，这样的机器可能现在还没有出现，还停留在理论上。

那**AI**的工作原理是什么呢？听说涉及到很多算法和大数据。数据越多，模型训练得越好，预测结果也就越准确。比如机器学习，这是一个重要的分支，它是通过让机器从大量数据中找到规律而不是依靠明确的指令来工作的。深度学习则是机器学习的一种方法，使用多层神经网络来模拟人脑的学习方式，处理像图像识别这样的复杂任务。

人工智能的发展历史是怎样的？好像开始于上世纪**50**年代，在达特茅斯会议上首次提出了**AI**的概念。那时候可能还只是初步的想法，后来经历了几起几落，比如**70**年代的冬天是因为技术还不成熟，应用受限。到了现在，计算能力加强了，算法也进步了，加上大数据的应用，**AI**才有了突飞猛进的发展。

那人工智能有哪些应用呢？听说在医疗、金融、交通这些领域都有应用。例如，医疗图像分析可以帮助医生更准确地诊断疾病；金融科技中可能用于检测欺诈行为或者智能投资；自动驾驶汽车也是**AI**的一个重要方向，涉及环境感知和决策系统。还有，个性化推荐在电商和社交媒体中越来越普遍，提升用户体验。

当然，人工智能的发展也带来了一系列挑战。数据隐私和安全问题很重要，因为很多**AI**应用需要收集大量用户数据，如何保护这些数据不被滥用是一大难题。伦理道德方面，比如自动驾驶车遇到紧急情况的决策选择，这不仅关系到算法设计，还涉及到人性化的考量。劳动力替代也是一个讨论热点，担心**AI**会取代许多传统的工作岗位，从而引发社会问题。

总的来说，人工智能正在快速发展，并在多个领域带来变革。它帮助我们解决了许多复杂的问题，提高了生产效率和生活质量。不过，随之而来的挑战也需要被认真对待。未来的发展需要技术、伦理和社会三方面的共同努力，确保人工智能能够造福全人类。

嗯，我可能有些地方理解得不够准确，比如**AI**到底是如何进行推理的？这些算法具体是怎么工作的呢？或许还需要进一步学习相关的课程或者资料来深入理解。

</think>

人工智能（**Artificial Intelligence, AI**）是一种模拟人类智能的技术，使计算机系统能执行复杂的任务，如学习、推理和问题解决。它通过算法处理数据，模仿人类认知功能，应用于各个领域。

主要类型包括弱人工智能，专注于特定任务，与强人工智能不同，后者是具备自我意识的通用型**AI**，尚未实现。

AI的工作原理涉及机器学习和深度学习，利用大量数据分析规律，提升准确性。发展历史始于**1956**年，经历起伏后，在计算能力和大数据推动下迅速进步。

应用广泛，涵盖医疗、金融、交通等，如诊断辅助和自动驾驶。同时面临技术和伦理挑战，需平衡创新与社会影响。

总体而言，**AI**正改变世界，带来机遇与挑战，需多方面协作确保其为人类福祉服务。

• 流式调用

```
from openai import OpenAI

client = OpenAI(
    base_url='http://192.168.110.131:11434/v1/',
    api_key='ollama',
)

messages = [
    {
```

```

        'role': 'user',
        'content': '你好, 请你介绍一下什么是人工智能?',
    }
]

try:
    # 调用聊天接口
    stream = client.chat.completions.create(
        model='deepseek-r1:32b',
        messages=messages,
        stream=True
    )

    # 处理流式响应
    for chunk in stream:
        if chunk.choices[0].delta.content is not None:
            print(chunk.choices[0].delta.content, end='', flush=True)

except Exception as e:
    print(f"发生错误: {str(e)}")

```

<think>

嗯, 我现在想了解一下什么是人工智能。这个话题好像挺热的, 但具体是什么呢? 我记得以前看到过一些关于机器学习和深度学习的東西, 不过可能还有其他方面。

首先, 人工智能听起来像是让机器像人一样思考或者做出决策。以前科幻电影里经常出现机器人取代人类的情节, 这应该就是人工智能的一种表现吧。那是不是说所有有智能的机器都是人工智能呢? 比如手机里的语音助手, 它们能和我对话, 可能也是用了AI技术。

我还记得有个词叫做“机器学习”, 好像这是实现人工智能的一种方法。机器学习让计算机通过数据自己学习, 而不是被明确编程去做某件事。这就像是教小孩认字, 不是直接告诉每个字的规则, 而是让他们多看多练, 自己记住这些字一样。那深度学习又是怎么回事呢? 是不是更复杂一点的技术, 比如用神经网络来模拟人的大脑?

然后还有自然语言处理和计算机视觉, 这些都是人工智能的应用领域吧。自然语言处理就是让机器理解人类的语言, 比如翻译软件或者聊天机器人。计算机视觉则是让机器能够识别图像或视频内容, 例如人脸识别或者自动驾驶中的目标检测。

人工智能好像有不同的层次, 有弱人工智能 (Narrow AI) 和强人工智能 (General AI)。弱人工智能只能执行特定任务, 比如说Siri只会回答问题或设置闹钟, 而不能像人类那样全面思考。强人工智能则理论上可以处理各种任务, 像电影里的机器人那样, 但现实里好像还没实现。

应用方面, 除了刚才提到的语音助手和推荐系统, 医疗诊断也是个例子。比如AI帮助医生分析X光片或者CT扫描结果, 提高诊断准确率。自动驾驶汽车也是一个典型的应用, 依靠大量的传感器和算法来做出驾驶决策。

伦理问题可能涉及隐私、就业影响以及AI失控的可能性吧。隐私的话, 比如面部识别可能会被用来监视人们; 就业方面, 自动化可能导致一些岗位的消失; 至于AI失控, 虽然现在看起来不太现实, 但长远来看还是需要考虑的风险。

人工智能的发展离不开大数据和计算能力的进步, 这些给了算法更多的数据去学习和处理更复杂的问题。这也让模型变得更大、更高效, 推动了技术的发展。

学习人工智能的话, 似乎得掌握编程语言如Python或Java, 数学知识比如线性代数和微积分也很重要。机器学习的基本概念和一些框架, 如TensorFlow或者PyTorch, 也是必须要了解的工具。如果我能掌握这些基础, 可能就能开始做些AI项目了。

不过，我还不清楚各个技术之间的区别，比如机器学习和深度学习是什么关系？它们是不是包含的关系呢？另外，数据在其中扮演什么角色？为什么数据对训练模型这么重要？

还有，伦理方面的问题真的有很多需要考虑的。比如隐私保护，怎么平衡数据利用和个人权利？就业方面，如何应对自动化带来的失业问题？这些都是社会需要解决的挑战。

总的来说，人工智能是一个涵盖广泛领域的技术，它让机器能够模拟人类智能进行学习、推理和决策。虽然前景广阔，但也带来了许多需要解决的问题。作为一个刚开始学的人，我可能要从基础开始，一步步深入理解各个组成部分和技术细节。

</think>

人工智能（**Artificial Intelligence, AI**）是指计算机系统模拟人类智能的能力，涵盖学习、推理、问题解决、感知和语言处理等领域。以下是对**AI**的详细介绍：

- 基本概念**：
 - 人工智能使机器能够执行通常需要人类 **intelligence** 的任务，如语音识别和决策制定。
- 技术基础**：
 - 机器学习（Machine Learning）**：计算机通过数据自主学习，而非遵循明确指令。
 - 深度学习（Deep Learning）**：属于机器学习的一部分，利用多层神经网络模拟人脑处理信息方式。
 - 自然语言处理（NLP）**：理解和生成人类语言，应用于聊天机器人和翻译工具。
 - 计算机视觉（Computer Vision）**：使计算机识别和分析图像、视频内容。
- 层次与应用**：
 - 弱人工智能（窄人工智能, Narrow AI）**：专注于特定任务，如语音助手。
 - 强人工智能（通用人工智能, AGI）**：具备全面人类智能的理论能力，尚未实现。
 - 应用领域包括自动驾驶、医疗诊断、推荐系统和图像分类。
- 伦理与挑战**：
 - 涉及隐私保护、就业影响、**AI**失控风险以及算法偏见等伦理问题。
- 发展驱动因素**：
 - 依赖大数据和计算能力进步，推动模型复杂性和处理效率提升。
- 学习路径**：
 - 学习编程语言如**Python**，数学基础（线性代数、微积分），掌握机器学习框架如**TensorFlow**或**PyTorch**。

人工智能是一门广泛且快速发展的技术领域，既带来创新机遇，也引发伦理和社会挑战。理解其基础知识和应用有助于更好地把握这一领域的未来发展。

在 **OpenAI Compatibility** 规范下，目前 **Ollama** 支持的模型参数如下：

支持的功能

功能	描述
聊天完成	Chat completions
流媒体	Streaming
JSON模式	JSON mode
可再现的输出	Reproducible outputs
视觉	Vision
工具	Tools

支持的请求字段

请求字段	描述
model	模型
messages	消息
frequency_penalty	频率惩罚
presence_penalty	存在惩罚
response_format	响应格式
seed	种子
stop	停止
stream	流式输出
stream_options	流式选项
include_usage	包含使用情况
temperature	温度
top_p	Top-p 采样
max_tokens	最大令牌数
tools	工具

```
from openai import OpenAI

client = OpenAI(
    base_url='http://192.168.110.131:11434/v1/',
    api_key='ollama',
)
```

```

messages = [
    {
        'role': 'user',
        'content': '你好，请你介绍一下什么是人工智能？',
    }
]

try:
    # 调用聊天接口
    stream = client.chat.completions.create(
        model='deepseek-r1:32b',
        messages=messages,
        stream=True,
        temperature=0.7,
        max_tokens=1024,
    )

    # 处理流式响应
    for chunk in stream:
        if chunk.choices[0].delta.content is not None:
            print(chunk.choices[0].delta.content, end='', flush=True)

except Exception as e:
    print(f"发生错误: {str(e)}")

```

<think>

好的，我现在要帮用户介绍什么是人工智能。首先，人工智能这个词大家可能听过不少，但具体是什么意思呢？我应该从基本概念开始讲起。

我记得AI英文是Artificial Intelligence，中文翻译过来就是人工智能。它主要是模拟人类的智能行为，比如学习、推理和解决问题这些。接下来，我可以分几个方面详细说明，这样结构更清晰。

第一部分可以介绍定义和发展背景。人工智能不仅仅是计算机程序，而是能够执行复杂任务的技术集合。还要提到它的起源，比如1956年的达特茅斯会议，还有近年来技术进步带来的突破，比如深度学习和神经网络的应用。

第二部分讲核心能力比较合适。这部分要涵盖感知与理解、思考与推理、学习与适应以及行动与创造这几个方面。每个点都要简单解释一下，让用户明白AI具体能做什么。

第三部分是应用领域，用户可能想知道AI在哪些地方用到了。比如自然语言处理、计算机视觉、机器学习和机器人这些，都是比较热门的应用方向。

第四部分需要说明人工智能的特点。数据驱动、适应性强、自动化程度高、多学科交叉这些都是重要的特点，要逐一解释清楚，让用户明白为什么AI会有这样的表现。

最后，在结尾部分可以总结一下，指出AI已经渗透到各个领域，并且未来可能会带来更多的变化和挑战。

整体上，我要确保内容全面，但不过于技术化，用通俗易懂的语言来描述。这样用户即使没有相关背景也能理解什么是人工智能。

</think>

当然！人工智能（Artificial Intelligence，简称AI）是指由人创造的能够执行通常需要人类智能的任务的系统或机器。这些任务包括学习、推理、问题解决、感知和语言理解等。

一、定义与发展背景

人工智能的目标是模拟人类的智能行为，使其能够完成复杂的任务。它起源于20世纪50年代，近年来随着计算能力的提升和大数据的应用，取得了显著进展。

二、核心能力

1. ****感知与理解****: 通过传感器或摄像头捕捉信息，并进行分析。
2. ****思考与推理****: 基于已有知识进行判断和决策。
3. ****学习与适应****: 从经验中获取知识并优化性能。
4. ****行动与创造****: 根据任务要求执行操作或生成内容。

三、应用领域

1. ****自然语言处理（NLP）****: 如聊天机器人、翻译工具。
2. ****计算机视觉****: 图像识别、面部识别等技术。
3. ****机器学习****: 通过数据训练模型进行预测和分类。
4. ****机器人与自动化****: 用于制造业和服务行业的智能设备。

四、特点

1. ****数据驱动****: 依赖大量数据进行训练。
2. ****适应性强****: 能根据新信息调整策略。
3. ****自动化程度高****: 减少人工干预，提高效率。
4. ****多学科交叉****: 融合计算机科学、心理学和数学等知识。

五、总结

人工智能正在改变我们的生活方式，从日常生活到工业生产都有广泛应用。随着技术进步，它将继续推动社会变革，但也带来了隐私保护等挑战。