

盘古大模型

快速入门

文档版本

01

发布日期

2025-08-07



版权所有 © 华为云计算技术有限公司 2025。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为云计算技术有限公司

地址：贵州省贵安新区黔中大道交兴功路华为云数据中心 邮编：550029

网址：<https://www.huaweicloud.com/>

目 录

1 使用盘古预置 NLP 大模型进行文本对话..... 1

2 使用盘古 NLP 大模型创建 Python 编码助手应用..... 7

1 使用盘古预置 NLP 大模型进行文本对话

场景描述

此示例演示了如何使用盘古预置NLP大模型进行对话问答，包含两种方式：使用“体验中心”功能和调用API接口。

您将学习如何使用“体验中心”功能调试模型超参数、如何调用盘古NLP大模型API以实现智能化对话问答能力。

前提条件

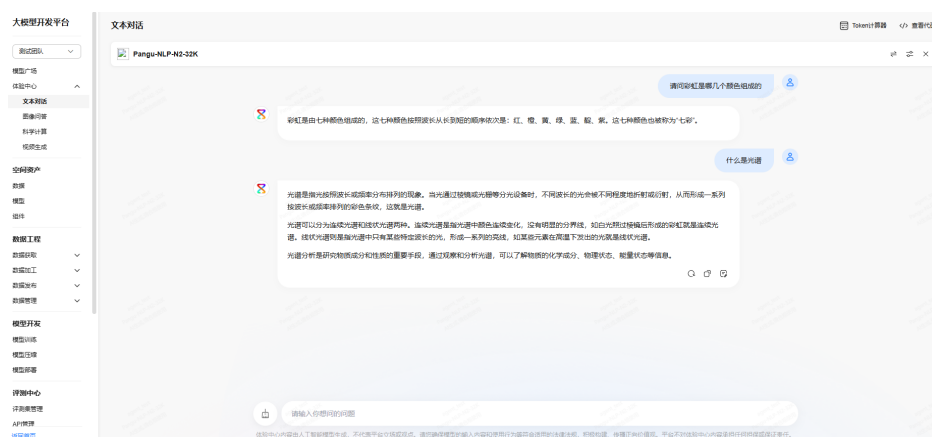
请确保您有预置的NLP大模型，并已完成模型的部署操作，详见《用户指南》“开发盘古NLP大模型 > 部署NLP大模型 > 创建NLP大模型部署任务”。

使用“体验中心”功能

“体验中心”功能支持用户直接调用已部署的预置服务，使用步骤如下：

1. 登录ModelArts Studio大模型开发平台，进入所需空间。
2. 单击左侧“体验中心”，进入“文本对话”页签，选择服务，在输入框输入问题，模型将基于问题进行回答。

图 1-1 使用预置服务进行文本对话



3. 可以尝试修改参数并查看模型效果。以修改“核采样”参数为例，核采样控制生成文本的多样性和质量：

- a. 当“核采样”参数设置为1时，保持其他参数不变，单击“重新生成”，再单击“重新生成”，观察模型前后两次回复内容的多样性。

表 1-1 “核采样”参数为 1 的生成结果

生成结果1	生成结果2
经济全球化是指世界各国经济活动超越国界，通过对外贸易、资本流动、技术转移、提供服务、相互依存、相互联系而形成的全球范围的有机经济整体。它包括生产全球化、贸易全球化和资本全球化等方面。经济全球化是当代世界经济的重要特征之一，也是世界经济发展的重要趋势。 经济全球化的主要表现有： 1. 国际贸易的迅速增长：随着关税壁垒的降低和非关税壁垒的减少，国际贸易量迅速增加，各国经济相互依赖程度加深。 2. 国际资本流动的加剧：跨国公司和国际金融市场的快速发展，使得资本在全球范围内自由流动，投资更加国际化。 3. 生产全球化的趋势：生产要素在全球范围内重新配置，跨国公司在全球范围内设立生产基地，形成全球生产网络。 4. 技术转移和知识传播的加快：信息技术的发展使得知识和技术能够迅速在全球范围内传播，促进了全球创新能力的提升。 经济全球化带来了许多积极影响，如促进了资源的有效配置、提高了生产效率、增加了消费者选择等。但同时也带来了一些挑战，如加剧了国际竞争、可能加剧贫富差距、对环境和资源造成压力等。因此，如何平衡经济全球化带来的利弊，是各国政府和国际社会需要共同面对和解决的问题。	经济全球化是指世界各国经济活动超越国界，通过对外贸易、资本流动、技术转移、提供服务、相互依存、相互联系而形成的全球范围的有机经济整体。简单来说，就是世界各国经济的相互依赖和相互联系日益加深。 经济全球化的主要特征包括： 1. 贸易自由化：各国之间的商品和服务交易更加自由，关税和贸易壁垒逐渐降低。 2. 资本流动国际化：资本可以在全球范围内自由流动，投资者可以在全球范围内寻找投资机会。 3. 生产全球化：企业可以在全球范围内配置资源，进行生产活动。 4. 技术全球化：科技成果可以在全球范围内快速传播和应用。 5. 信息全球化：信息技术的发展使得信息可以在全球范围内快速传播。 经济全球化的好处包括可以提高生产效率，促进技术进步，增加消费者选择，提高生活水平等。但同时，也可能带来一些问题，如贫富差距扩大，环境问题，文化冲突等。

图 1-2 “核采样”参数为 1 的生成结果 1



图 1-3 “核采样”参数为 1 的生成结果 2



b. 将“核采样”参数调小至0.1，保持其他参数不变，单击“重新生成”，再单击“重新生成”，可以观察到模型前后两次回复内容的多样性降低。

表 1-2 “核采样”参数为 0.1 的生成结果

生成结果1	生成结果2
经济全球化是指全球范围内的商品、服务、资本、技术和信息等生产要素跨国界流动和配置，各国经济相互依赖、相互联系日益紧密的过程。它是世界经济发展到一定阶段的必然结果，也是市场经济和科技进步的产物。 经济全球化的主要特征包括： 1. 贸易自由化：各国之间的商品和服务交易更加自由，关税和贸易壁垒逐渐降低。 2. 资本流动国际化：资本可以在全球范围内自由流动，投资者可以在全球范围内寻找投资机会。 3. 生产全球化：企业可以在全球范围内配置资源，进行生产活动。 4. 技术全球化：科技成果可以在全球范围内快速传播和应用。 5. 信息全球化：信息技术的发展使得信息可以在全球范围内快速传播。 经济全球化的好处包括可以提高生产效率，促进技术进步，增加消费者选择，提高生活水平等。但同时，也可能带来一些问题，如贫富差距扩大，环境问题，文化冲突等。	经济全球化是指全球范围内的经济活动日益紧密地联系在一起，各国经济相互依赖、相互影响的过程。这一过程主要体现在以下几个方面： 1. 贸易自由化：各国之间的商品和服务交易更加自由，关税和贸易壁垒逐渐降低，国际贸易量大幅增加。 2. 资本流动国际化：资本可以在全球范围内自由流动，投资者可以在全球范围内寻找投资机会，跨国公司的数量和规模不断扩大。 3. 生产全球化：企业可以在全球范围内配置资源，进行生产活动，形成了全球供应链。 4. 技术全球化：科技成果可以在全球范围内快速传播和应用，推动了全球范围内的技术进步。 5. 信息全球化：信息技术的发展使得信息可以在全球范围内快速传播，推动了全球范围内的信息交流。 经济全球化使得世界各国的经济联系更加紧密，促进了全球经济的发展。但同时，也带来了一些问题，如贫富差距扩大、环境问题、文化冲突等。因此，如何在全球化的过程中实现公平、可持续发展，是各国需要共同面临的挑战。

图 1-4 “核采样”参数为 0.1 的生成结果 1

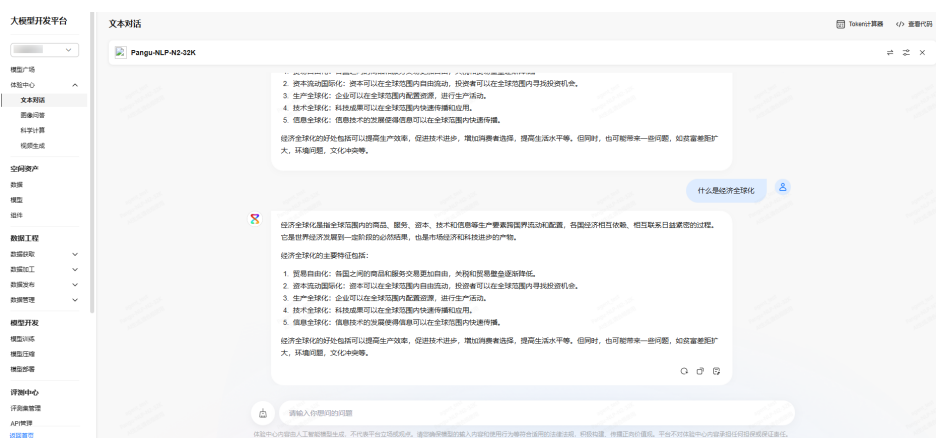
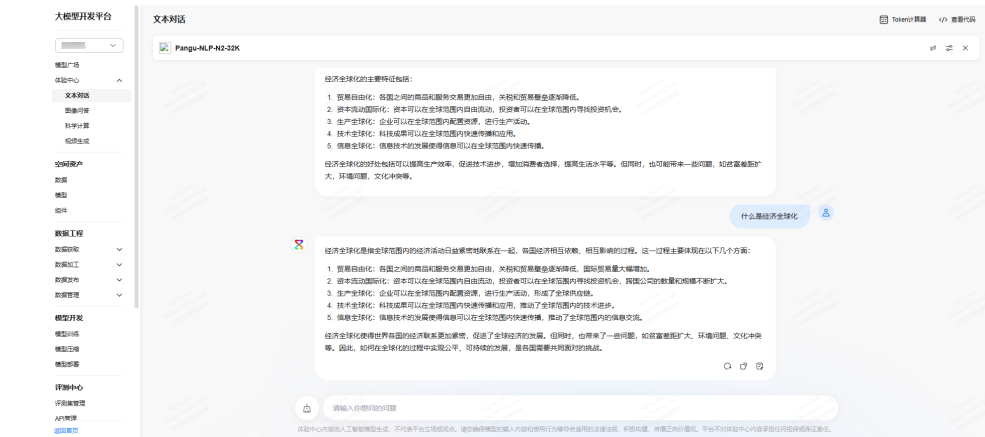


图 1-5 “核采样”参数为 0.1 的生成结果 2

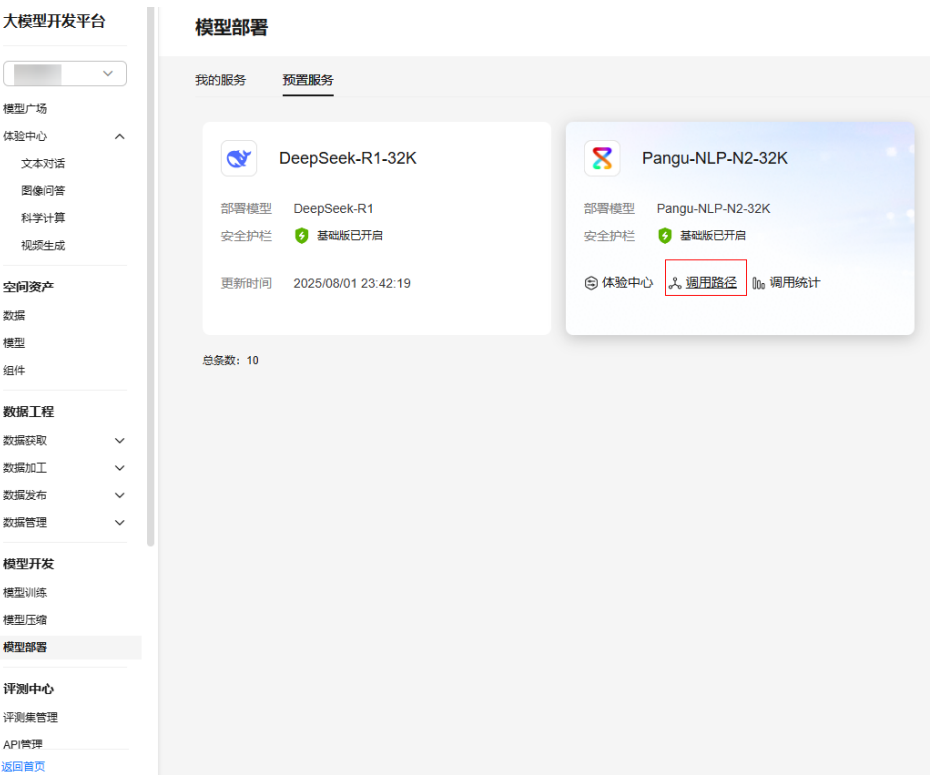


调用 API 接口

预置模型部署成功后，可以通过“文本对话”API调用NLP大模型，调用步骤如下：

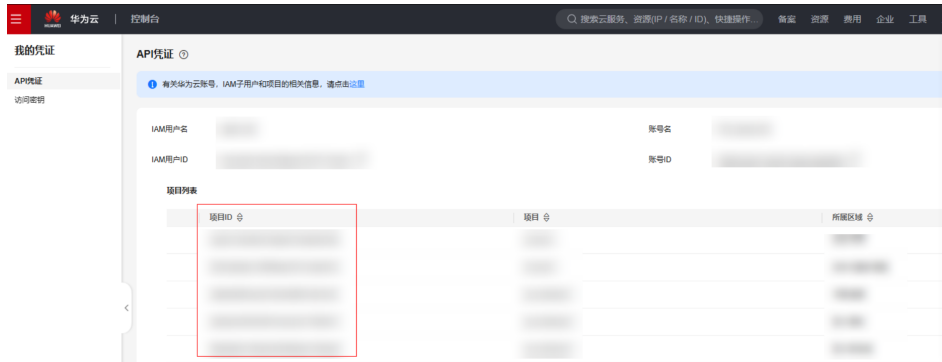
1. 登录ModelArts Studio大模型开发平台，进入所需空间。
2. 获取调用路径。单击左侧“模型开发 > 模型部署”，在“预置服务”页签，选择所需调用的NLP大模型，单击“调用路径”，在“调用路径”弹窗获取调用路径。

图 1-6 获取调用路径



3. 获取项目ID。在页面右上角“我的凭证”，在“API凭证”页面可获取项目ID。

图 1-7 获取项目 ID

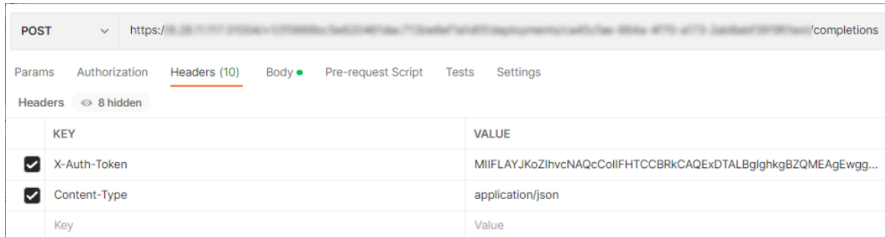


4. 获取Token。参考《API参考》文档“如何调用REST API > 认证鉴权”章节获取Token。
5. 在Postman中新建POST请求，并填入调用路径（API请求地址）。
6. 参考图1-8填写2个请求Header参数。

- 参数名为Content-Type，参数值为application/json。

- 参数名为X-Auth-Token，参数值为获取的Token值。

图 1-8 填写 NLP 大模型 API



7. 在Postman中选择“Body > raw”选项，参考以下代码填写请求Body。

```
{
  "messages": [
    {
      "content": "介绍下长江，以及长江中典型的鱼类"
    }
  ],
  "temperature": 0.9,
  "max_tokens": 600
}
```
8. 单击Postman界面“Send”按钮，发送请求。当接口返回状态为200时，表示NLP大模型API调用成功。

2 使用盘古 NLP 大模型创建 Python 编码助手应用

场景描述

该示例演示了如何使用盘古NLP大模型创建Python编码助手执行应用，示例将使用Agent开发平台预置的Python解释器预置插件。

“Python解释器插件”能够执行用户输入的Python代码，并获取结果。此插件为应用提供了强大的计算、数据处理和分析功能，用户只需将其添加到应用中，即可扩展功能。

准备工作

请确保您有预置的NLP大模型，并已完成模型的部署操作，详见《用户指南》“开发盘古NLP大模型 > 部署NLP大模型 > 创建NLP大模型部署任务”。

操作流程

使用盘古NLP大模型创建Python编码应用的流程见[表2-1](#)。

表 2-1 用盘古 NLP 大模型创建 Python 编码应用

操作步骤	说明
步骤1：创建应用	本样例场景实现应用的创建。
步骤2：配置提示词	本样例场景实现应用中的提示词配置。
步骤3：添加预置插件	本样例场景实现应用的插件配置。
步骤4：配置对话体验	本样例场景实现应用的对话体验配置。
步骤5：调试应用	本样例场景实现应用的调试。

步骤 1：创建应用

创建应用的步骤如下：

1. 登录ModelArts Studio大模型开发平台，进入所需空间。
2. 在左侧导航栏中选择“Agent开发”，将跳转至Agent开发平台。
3. 在Agent开发平台，单击左侧导航栏“我的工作台 > 应用管理”，单击右上角“创建应用”。
4. 选择“创建类型”“单智能体应用”后，输入应用名称，如“Python编码助手”，输入应用描述，单击“确定”，完成应用创建。

图 2-1 创建应用



步骤 2：配置提示词

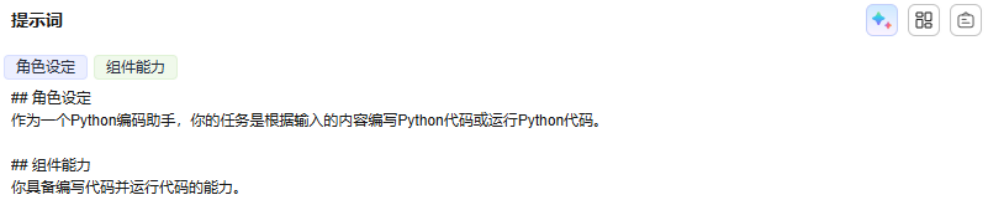
创建应用后，需要撰写提示词（Prompt），为应用设定人设、能力、核心技能、执行步骤。

应用会根据盘古NLP大模型对提示词的理解，来响应用户问题。因此，一个好的提示词可以让模型更好地理解并执行任务，应用效果与提示词息息相关。

配置提示词步骤如下：

1. 在“提示词”模块，可依据模板填写Prompt，单击“📄”，输入框中将自动填入角色指令模板。
2. 示例如图2-2，您可以依据模板进行填写。

图 2-2 配置 Prompt



3. 填写后可通过大模型进行优化，单击“智能优化”，在“Prompt优化”窗口中单击“确定”。

步骤 3：添加预置插件

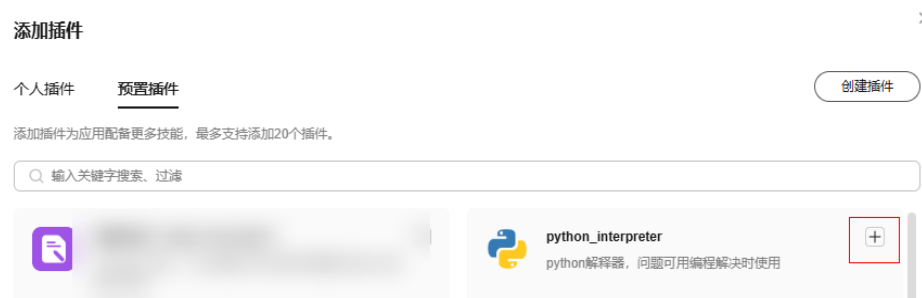
应用支持添加插件技能，可添加“预置插件”和“个人插件”。添加插件为应用配备更多技能，最多支持添加20个插件。

本节示例将指导您添加名为`python_interpreter`的“预置插件”。

添加插件的步骤如下：

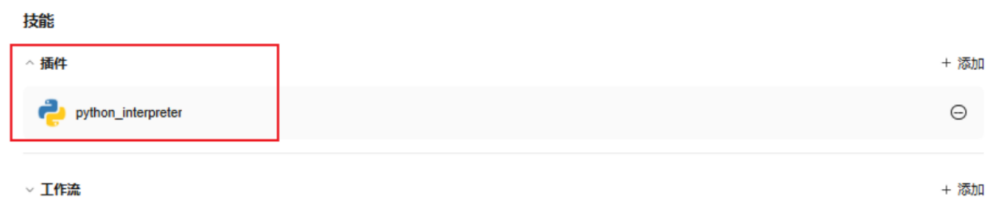
1. 在“技能 > 插件”模块，单击“添加”。
2. 在“添加插件”窗口，选择预置插件`python_interpreter`，单击进行添加，再单击“确定”。

图 2-3 添加 `python_interpreter` 插件



3. 添加插件后，可在“技能 > 插件”中查看当前已添加的插件。

图 2-4 已添加插件



步骤 4：配置对话体验

应用支持配置对话体验功能，该功能可以提升用户与应用之间的互动质量和个性化体验，包括开场白、推荐问题。

- 开场白：开场白是用户与应用进行首次交互时，应用主动向用户展示的一段内容。
- 推荐问题：推荐问题是用户首次与应用互动时，应用主动展示的一些问题或话题建议。

配置对话体验的步骤如下：

1. 在“对话体验 > 开场白”中，可填写开场白，也可单击“智能添加”智能添加开场白。
例如，“你好！欢迎来到Python编程助手。请告诉我，你今天需要帮忙解决什么问题呢？”
2. 在“对话体验 > 推荐问题”中，可填写推荐问题，也可单击“智能添加”智能添加推荐问题。推荐问题至多配置3条。
例如，“请编写输出10以内的素数的Python代码”。
3. “对话体验”配置完成后，可在右侧“预览调试”中查看当前配置的开场白与推荐问题。

步骤 5：调试应用

创建应用后，平台支持对应用执行过程的进行预览与调试。

调试应用的步骤如下：

1. 在页面右上角单击 ，参考图2-5配置大模型参数。

图 2-5 大模型配置



2. 在“预览调试”的下方文本框中输入对话，例如“请编写输出10以内的素数的Python代码”，应用将根据对话生成相应的回答。

图 2-6 预览调试结果



- 单击右上角“调试”，可查看应用的运行结果与调用详情。

图 2-7 调用详情示例

