

# AI 推理解决方案镜像使用文档

本文档以 Ubuntu 20.04.5 LTS 系统预装 AIGC/LLM 场景的 AI 推理解决方案为例，介绍如何使用预装 AI 推理解决方案镜像的 GPU 实例进行 AIGC/LLM 场景试用。本文档示例是在阿里云 gn7i 实例上完成的。

## 前提条件

创建一台 ecs.gn7i-c16g1.4xlarge 实例，具体操作请参见：[https://help.aliyun.com/document\\_detail/60149.html](https://help.aliyun.com/document_detail/60149.html)。请选择镜像市场中的“ai-inference-solution”镜像。

## 操作步骤

1. 远程连接 ECS 实例。连接方式请参见：[https://help.aliyun.com/document\\_detail/71529.html](https://help.aliyun.com/document_detail/71529.html)。
2. 下载相关的模型，然后运行镜像内应用，在 GPU 云服务器执行 AI 推理。