

# 天翼云主机&deepseek融合产品 推广和应用赋能

天翼云科技有限公司

2025年2月

**1**

**什么是DeepSeek**

**2**

**DeepSeek能做什么**

**3**

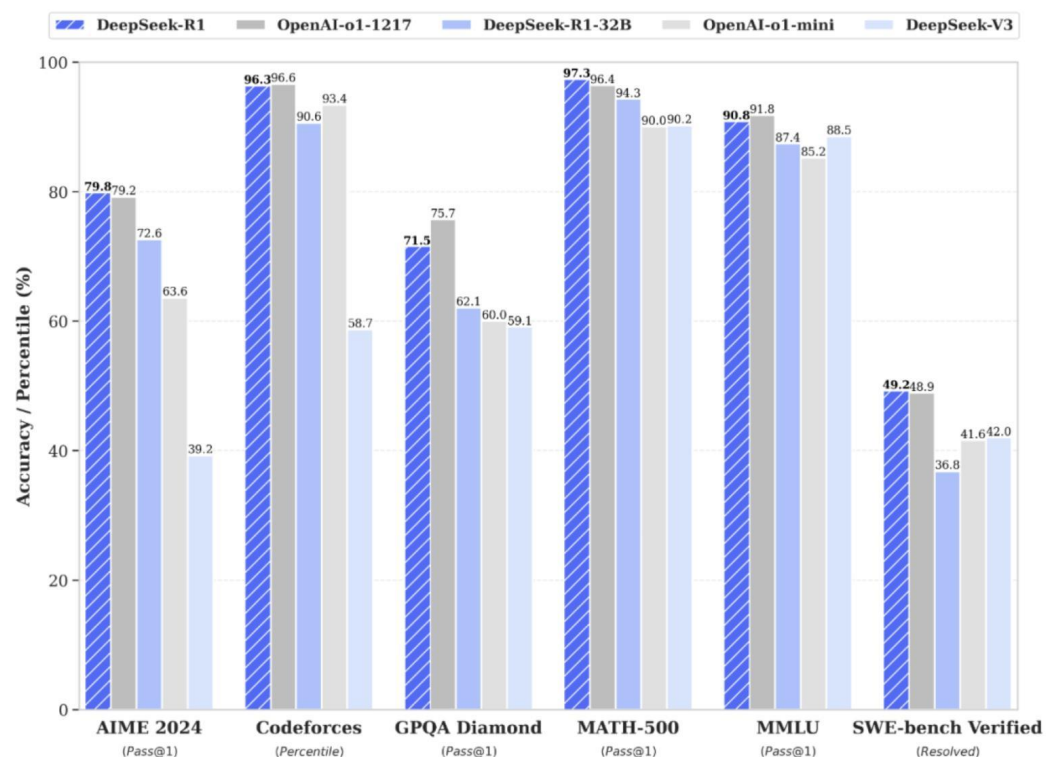
**天翼云DeepSeek云主机**

# 什么是Deepseek?

- DeepSeek 是一个基于 Transformer 架构的大型语言模型 (LLM)，由深度求索 (DeepSeek) 公司开发。它能够处理自然语言理解、生成、翻译、问答等多种任务。在目前大模型主流榜单中，DeepSeek-V3 在开源模型中位列榜首，与世界上最先进的闭源模型不分伯仲。在对话模型典型任务方面的评测效果如下：

## 性能对齐 OpenAI-o1 正式版

DeepSeek-R1 在后训练阶段大规模使用了强化学习技术，在仅有极少标注数据的情况下，极大提升了模型推理能力在数学、代码、自然语言推理等任务上，性能比肩 OpenAI o1 正式版。



免费

+

国产

+

开源

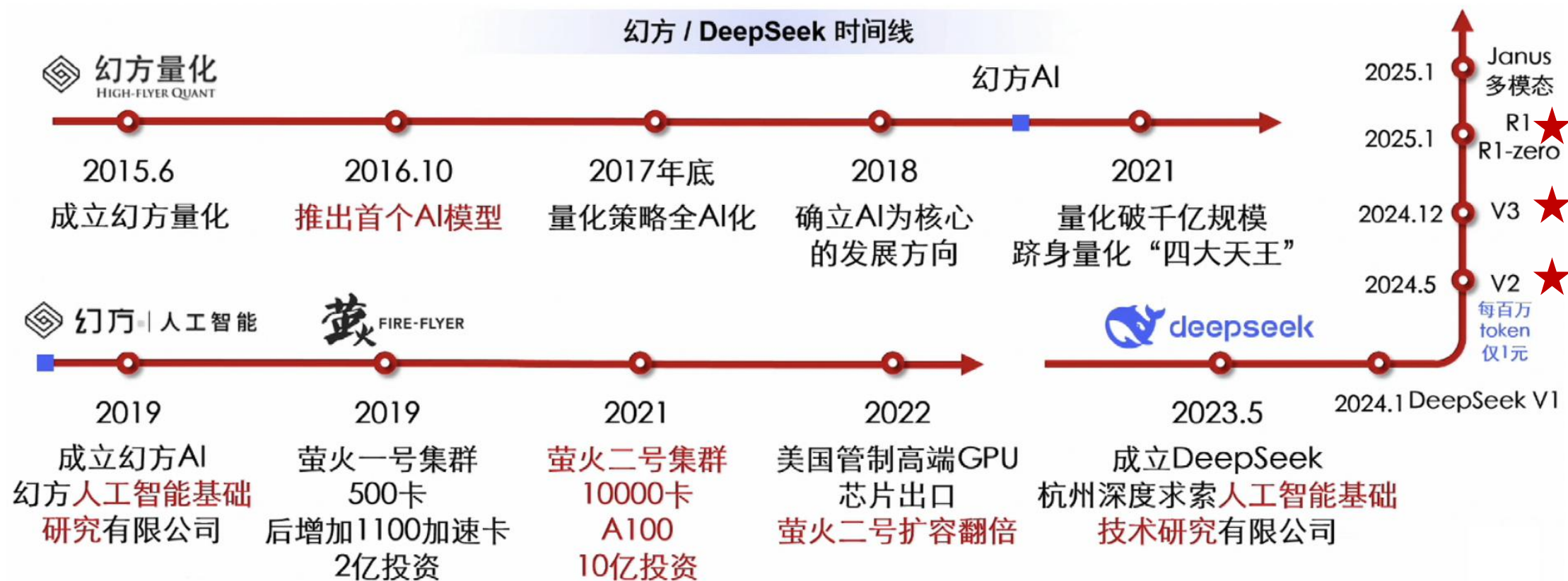
+

强大

+

AI

- DeepSeek (深度求索)，成立于2023年5月，母公司为私募量化剧透“幻方量化”。
- 2024年12月，推出DeepSeek-V3模型，模型参数671B
- 2025年1月，推出 DeepSeek-R1-Zero、DeepSeek-R1，以及基于 Qwen 和 Llama 架构从 DeepSeek-R1 知识蒸馏获得的六个稠密模型(1.5B、7B、8B、14B、32B、70B)，其DeepSeek-R1性能接近OpenAI的GPT-o1。



1

什么是DeepSeek

2

DeepSeek能做什么

3

天翼云DeepSeek云主机

## 智能文献检索与分析

- ✓ **快速定位关键文献：**利用 DeepSeek R1 强大的语义理解能力，快速从海量文献中筛选出与研究主题高度相关的文献，为科研人员节省大量时间，提高文献检索效率
- ✓ **深入分析文献内容：**模型能够对文献进行深度分析，提取关键信息、总结研究方法和结论，帮助科研人员快速把握文献的核心内容，为后续研究提供有力支持
- ✓ **生成文献综述草稿：**根据检索和分析的结果，DeepSeek R1 可以自动生成文献综述的初稿，为科研人员撰写综述文章提供参考和启发，减轻文献整理的工作负担

## 复杂实验数据分析

- **处理大规模实验数据：**天翼云 GPU 云主机的强大算力结合 DeepSeek R1 的数据处理能力，能够高效处理大规模实验数据，快速提取有价值的信息，加速实验结果的分析过程
- **发现数据中的隐藏规律：**模型通过深度学习算法，挖掘实验数据中的潜在规律和关联，为科研人员提供新的研究思路和假设，推动科研创新
- **辅助实验设计优化：**基于对已有实验数据的分析，DeepSeek R1 可以为科研人员提供实验设计的优化建议，如调整实验参数、改进实验方法等，提高实验的成功率和效率

## 高效代码生成与优化

- **快速生成代码片段：**在编程过程中，DeepSeek R1 能够根据科研人员的需求快速生成相应的代码片段，减少手动编写代码的工作量，提高编程效率
- **代码逻辑优化建议：**模型可以对已有的代码进行分析，提出优化建议，帮助科研人员改进代码的逻辑结构，提高代码的可读性和运行效率
- **跨语言代码转换辅助：**对于涉及多种编程语言的研究项目，DeepSeek R1 可以辅助进行跨语言代码转换，降低不同语言之间的转换难度，促进不同技术栈之间的融合





## 01. 智能办公助手

- 自动化文档处理：DeepSeek R1 可以自动读取和理解各类办公文档，如报告、邮件、合同等，提取关键信息并进行分类整理，提高文档处理的效率和准确性
- 会议内容总结与记录：在会议过程中，DeepSeek R1 可以实时记录会议内容，并自动生成会议总结，提炼会议要点和决策事项，方便会后查阅和跟进

## 02. 数据分析与决策支持

- 企业数据洞察：利用 DeepSeek R1 的数据分析能力，对企业内部的海量数据进行挖掘和分析，为企业管理层提供数据驱动的决策支持，帮助企业更好地把握市场趋势和业务发展方向
- 风险评估与预警：模型可以对企业的业务数据进行实时监测和分析，及时发现潜在的风险因素，并发出预警信号，帮助企业提前采取措施，降低风险损失

## 03. 智能客服与客户关系管理

- 知识库构建与维护：企业可以利用 DeepSeek R1 快速构建企业知识库，将分散的知识资源进行整合和分类，方便员工随时查询和学习，提高企业的知识共享和传承效率
- 知识问答与解答：模型能够实时回答员工在工作中遇到的各种知识性问题，提供准确的答案和解决方案，促进企业内部知识的流通和应用



## 智能写作与创意生成

DeepSeek R1 可应用于智能写作、创意绘画、音乐创作等领域，激发用户的创意和灵感。

例如，一些创作者利用 DeepSeek R1 生成创意内容，如诗歌、故事、绘画作品等，为生活和艺术创作带来新的可能性。



## 游戏开发与虚拟现实

天翼云 GPU 云主机的强大图形处理能力结合 DeepSeek R1 的智能交互能力，为游戏开发和虚拟现实应用提供了强大的技术支持。

例如，某游戏开发团队利用 GPU 云主机和 DeepSeek R1 模型，开发出具有智能 NPC 和交互剧情的游戏，提升了玩家的游戏体验。



## 个性化推荐与内容分发

基于 DeepSeek R1 的用户行为分析和内容推荐能力，可实现个性化的内容分发和推荐服务。

例如，某视频平台利用 DeepSeek R1 模型，根据用户的观看历史和偏好，为用户精准推荐视频内容，提高了用户的观看体验和平台的用户粘性。



1

什么是DeepSeek

2

DeepSeek能做什么

3

天翼云DeepSeek云主机

当前，DeepSeek-R1模型已在天翼云多个公有云平台产品上线，包括“息壤一站式智算服务平台”、“科研助手”、“公有云主机DeepSeek镜像”、“天翼云AI云电脑”、“魔乐社区”等，同时支持英伟达及国产昇腾910B算力。

支持产品平台	部署方式	上线功能	使用模式	模型版本	参数规模
息壤一站式智算服务平台	公有云/私有化	推理	API调用/体验中心	DeepSeek-R1	671B
科研助手	公有云	推理	应用商城/Web页面交互	DeepSeek-R1	7B
公有云主机	公有云	推理	DeepSeek镜像加载	DeepSeek-R1	7B
AI云电脑	公有云	推理	AI应用中心模型调用	DeepSeek-R1	32B
魔乐社区	公有云	推理	线上体验/开源下载	DeepSeek-R1	671B
息壤智算一体机 DeepSeek版	私有化	训练/推理	私有化部署	DeepSeek-R1	视配置而定

## GPU/CPU 选择

- 推荐使用 NVIDIA 高性能 GPU（如 H800、A100、V100、A10等）或国产化GPU（如昇腾910B），若无 GPU，可选择多核处理器（如 Intel Xeon）。
- GPU 加速可显著提升模型运行效率，CPU 配置需满足最低 8GB 内存要求。

## 内存与存储

- 至少需要 8GB 内存，建议 32GB 或更高，以应对复杂任务。
- 存储空间需充足，建议使用 SSD 硬盘，以提高模型加载和数据处理效率。

## 网络环境

- 确保网络连接稳定，带宽充足，便于模型文件下载和后续数据传输。
- 若网络不稳定，可尝试使用加速下载工具或镜像站点。



## 操作系统

- 推荐使用 Linux 操作系统（如 Ubuntu 20.04 或更高版本），其兼容性和性能更优。
- Windows 用户也可通过 Ollama 完成部署，但需注意部分命令和路径差异。

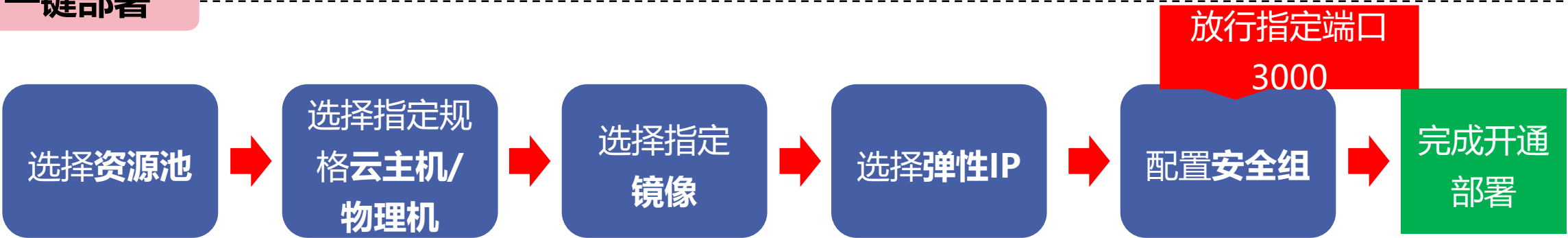
## Python 环境

- 安装 Python 3.8 或更高版本，确保安装必要的 Python 库（如 TensorFlow、PyTorch 等）
- 确保安装支持 DeepSeek R1 的框架，如基于 PyTorch 的实现。

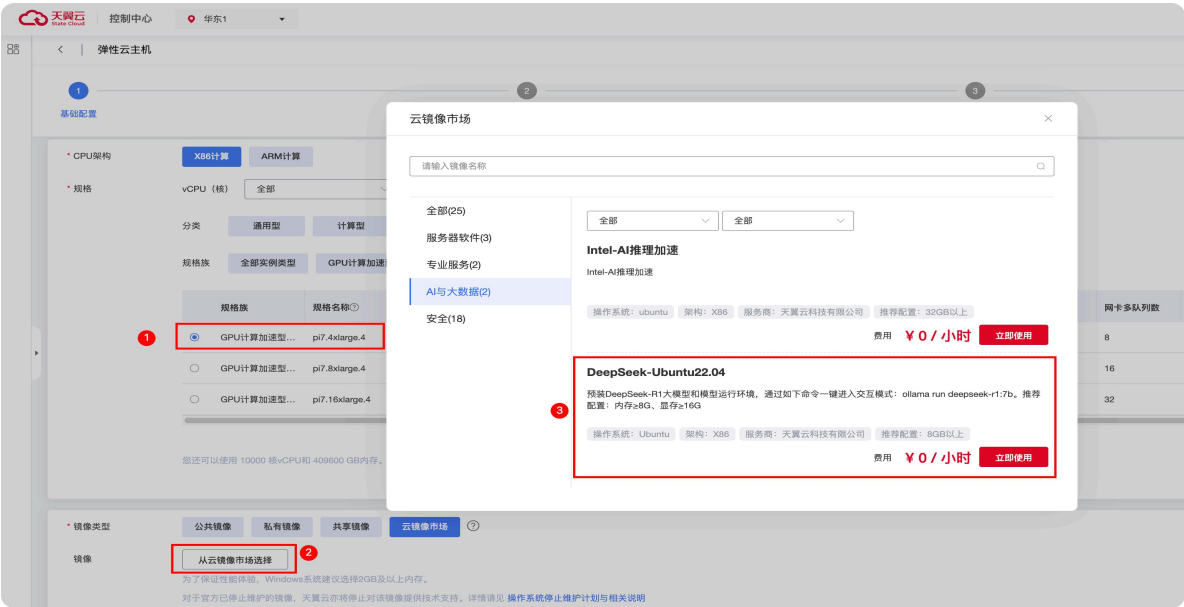
## 模型文件准备

- 访问 DeepSeek 官方 GitHub 仓库或 Ollama 官网，根据需求选择不同版本的模型文件。
- 模型参数量从 1.5B 到 671B 不等，可根据硬件配置选择合适版本。

## 一键部署



就是标准创建云主机流程, 只需选择指定镜像, 开机自动完成部署与服务启动



DeepSeek R1镜像可按需上线到资源池

在云镜像市场中选择预置了DeepSeek R1模型的DeepSeek-Ubuntu22.04镜像

模型	推荐最低配置	
	推荐规格	GPU配置
DeepSeek-R1-Distill-Qwen-1.5B	c7.xlarge.2(4核8GB),系统盘60GB SSD	-
DeepSeek-R1-Distill-Qwen-7B	c7.2xlarge.4(8核32GB), 系统盘60GB SSD	-
DeepSeek-R1-Distill-Qwen-1.5B	pi2.2xlarge.4(8核32GB), 系统盘60GB SSD	1卡 T4 (显存16G)
DeepSeek-R1-Distill-Qwen-7B	pi7.4xlarge.4(16核64GB), 系统盘60GB SSD	1卡 A10 (显存24G)
DeepSeek-R1-Distill-Llama-8B	pi7.4xlarge.4(16核64GB), 系统盘60GB SSD	1卡 A10 (显存24G)
DeepSeek-R1-Distill-Qwen-14B	p8a.6xlarge.4(24核96GB), 系统盘100GB SSD	1卡 A100 (显存40G)
DeepSeek-R1-Distill-Qwen-32B	p8a.12xlarge.4(48核192GB), 系统盘100GB SSD	2卡 A100 (显存80G)
DeepSeek-R1-Distill-Llama-70B	p8a.12xlarge.4(96核384GB), 系统盘100GB SSD	4卡 A100 (显存160G)
DeepSeek-R1、DeepSeek-V3满血版模型的参数量较大 (671B)	4台 physical.acas910b.2xlarge11(4路48核48线程) 系统盘: 2*480GB(SSD), 数据盘: 2*3200GB(NVMeSSD)	32卡 昇腾 910B(显存2048GB)
	1台 physical.h8ns.6xlarge28(2路48核48线程) 系统盘: 1* 480GB(M.2), 数据盘: 2*3200GB(NVMeSSD)	8卡 NVIDIA H800(显存640GB)

备注:

1. 预装镜像:DeepSeek-Ubuntu22.04(预装DeepSeek-R1:7B): 推荐配置: 内存≥8G、显存≥16G
2. 折扣价: 普通云主机是按照25折进行计算, GPU云主机按照45折进行计算, 可根据实际需求进行申请折扣; GPU裸金属为对声出货价, 实际价格可走OA进行申请对应折扣
3. 此报价云主机只包含CPU、内存、GPU、系统盘, 不包含EIP价格, 网络需根据客户需求进行添加; GPU裸金属也不包含EIP价格



天翼云提供丰富的IaaS资源，支撑用户开通部署deepseek大模型！



01

丰富的资源种类，资源池遍布全国

- 云主机提供：T4、V100、A10、A100等规格，根据需求订购不同规格，部署按需版本大模型；
- 裸金属提供：A100、A800、H800、L40s、昇腾910B，满足英伟达和纯国产化的deepseek部署需求。

02

保障资源容量

- 天翼云资源A10GPU资源超1000卡可售规模；
- 各类智算GPU裸金属仅250台可售。

天翼云各类匹配deepseek资源一览表  
由于deepseek热度大，资源容量动态变化，随时更新  
GPU裸金属资源在下单前提前确认

资源类型	GPU卡型号	单卡显存	全国可售量（卡）	官网产品大类
云主机	T4	16G	410	弹性云主机
	A10	24G	1091	
	A100	40G	22	
	3060	12G	540	科研助手
	T4	16G	838	
裸金属	A10	24G	292	物理机
	A100	40G	48	
	A800	40G	440	
	A800	80G	32	
	L40s	48	520	
	H800	80G	88	
	910B	64G	824	科研助手
	A100	40G	126	

构建三线支撑矩阵，快速拉通天翼云技术支撑人员  
保障一线的各类资源、交付等需求

400  
热线  
电话  
支  
撑



天翼云工单系统



产品	支撑矩阵		
	一线	二线	三线
公有云主机	13地市天翼云公有云支撑团队 (云地分、集软)	郭文杰、王信飞、张顺	方明海、杜雅红
AI云电脑		郭文杰、王信飞、张顺	黎瑞聪、杨龙
息壤一站式智算服务平台		郭文杰、王信飞、张顺	王磊、郭旻旻
科研助手		郭文杰、王信飞、张顺	
息壤智算一体机 DeepSeek版		郭文杰、王信飞、张顺	
华东大区分工： 郭文杰：苏州、镇江、淮安、宿迁、连云港 王信飞：无锡、南通、盐城、扬州 张顺：南京、常州、徐州、泰州			

天翼云官网首页置顶页面，一键直达

天翼云

活动 产品 解决方案 应用商城 合作伙伴 开发者 智算云 支持与服务 了解天翼云

Q 关系数据库SQL Serv... 文档 控制中心 备案 管理中心 郭\*杰

免费体验DeepSeek，上天翼云息壤

新老用户均可免费体验2500万Tokens，限时两周

立即前往

新春年货节

爆款云主机2核4G限时秒杀，85.5元/3个月起！

息壤高校钜惠活动

多款产品享受线上订购超值特惠

免费试用

多款云产品免费试用

天翼云脑AOne

连接、保护、办公，All in One！

稳定、安全、可信的云产品

提供多种预置规格，一键订购

免费体验 科研助手 智算底座 智算赋能 AI云电脑 魔乐社区

高性能IaaS助力 AI 深度应用

提供多种类型的云主机、裸金属资源，即开即用，实现业务分钟级部署，满足个人开发者或企业业务部署的需求 [更多DeepSeek云主机使用说明](#)

高性能算力解锁AI新境界

支持用户快速搭建DeepSeek-R1推理、微调、训练环境。在金融、医疗、法律等垂直领域，通过微调训练，使模型更好地理解领域术语和知识；在企业内部数据上微调模型，避免敏感数据泄露，同时提升模型效果。

产品优势:

- 限时免费体验
- 超低价格折扣
- 详细的部署指导
- 丰富的产品规格

DeepSeek C7 云主机

Ollama + Open WebUI + DeepSeek R1-7B，直接对话，无Token焦虑

区域	华东1-可用区1
规格选择	8核32G
系统盘类型	高IO
系统盘大小 (GB)	60
带宽 (M)	2
时长	1个月
0.00 元	立省1,096.00元
立即购买 已抢购100.00%	

DeepSeek A10 GPU 云主机

Ollama + Open WebUI + DeepSeek R1-7B，直接对话，无Token焦虑

区域	华东1-可用区2
规格选择	16核64G
系统盘类型	通用型SSD
系统盘大小 (GB)	60
带宽 (M)	10
时长	1个月
1,901.37 元	立省2,852.06元
立即购买	

预置DeepSeek A10云主机，使用推荐配置，快速完成订购，一键运行Deepseek-R1模型

快速体验

科研助手

智算底座

### 高性能算力解锁AI新境界

支持用户快速搭建DeepSeek-R1推理、微调、训练环境。在金融、医疗、法律等垂直领域，通过微调训练，使模型更好地理解领域术语和知识；在企业内部数据上微调模型，避免敏感数据泄露，同时提升模型效果。

**产品优势：**

- 限时免费体验
- 超低价格折扣
- 详细的部署指导
- 丰富的产品规格

DeepSeek C7 云主机

Ollama + Open WebUI + DeepToken焦虑

区域

规格选择

系统盘类型

系统盘大小 (GB)

带宽 (M)

时长

0.00 元

立省1,096.00元

立即购买 已

DeepSeek 昇腾裸金

Ollama + Open WebUI + DeepToken焦虑

DeepSeek A10 GPU 云主机

区域

规格选择

镜像

系统盘类型

系统盘大小 (GB)

数据盘类型

数据盘大小 (GB)

带宽 (M)

时长

台数

3.99折

1,901.37 元

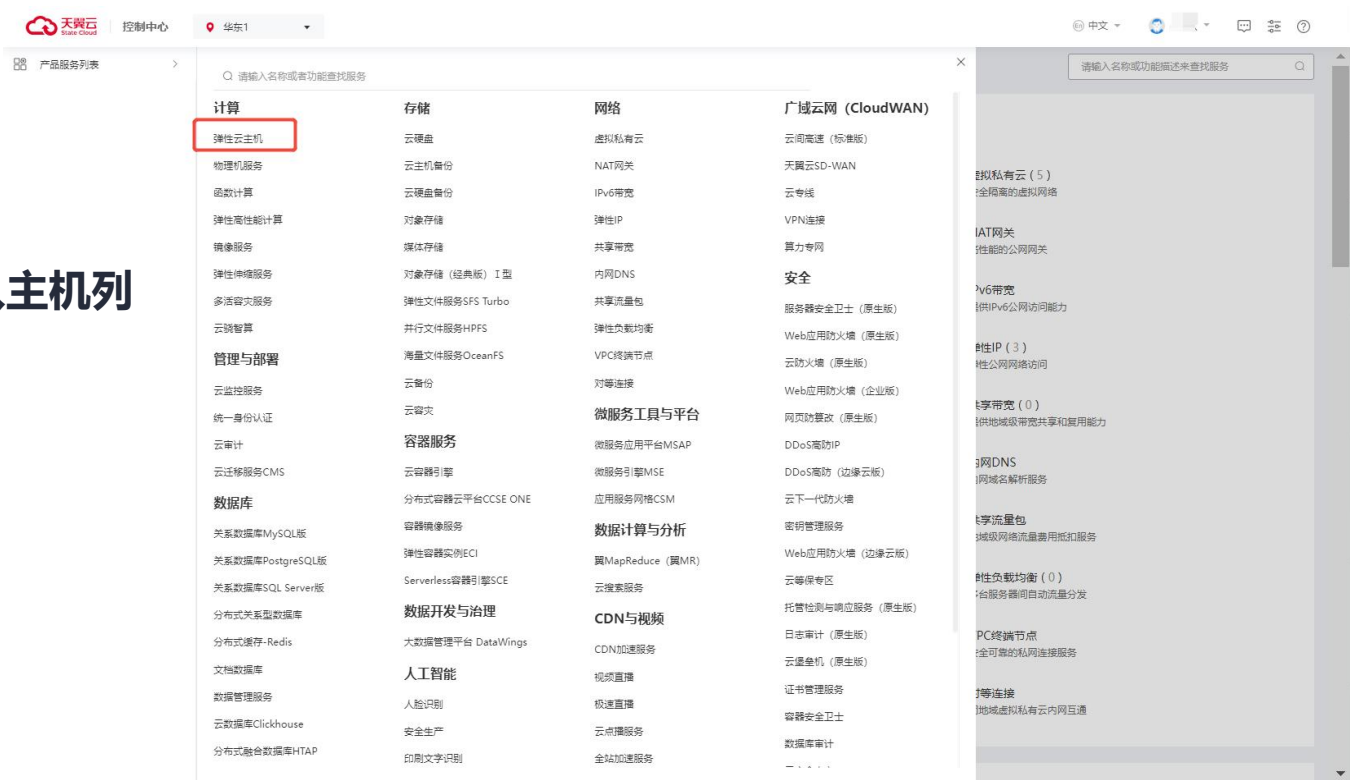
立省2,852.06元

立即购买

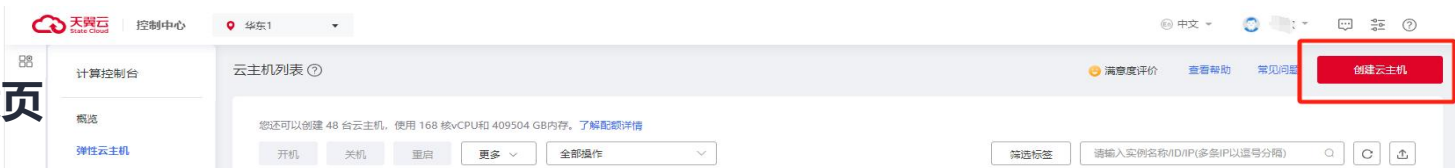
# 自动化部署deepseek云主机步骤 (1/6)



- 单击“产品服务列表>弹性云主机”，进入主机列表页



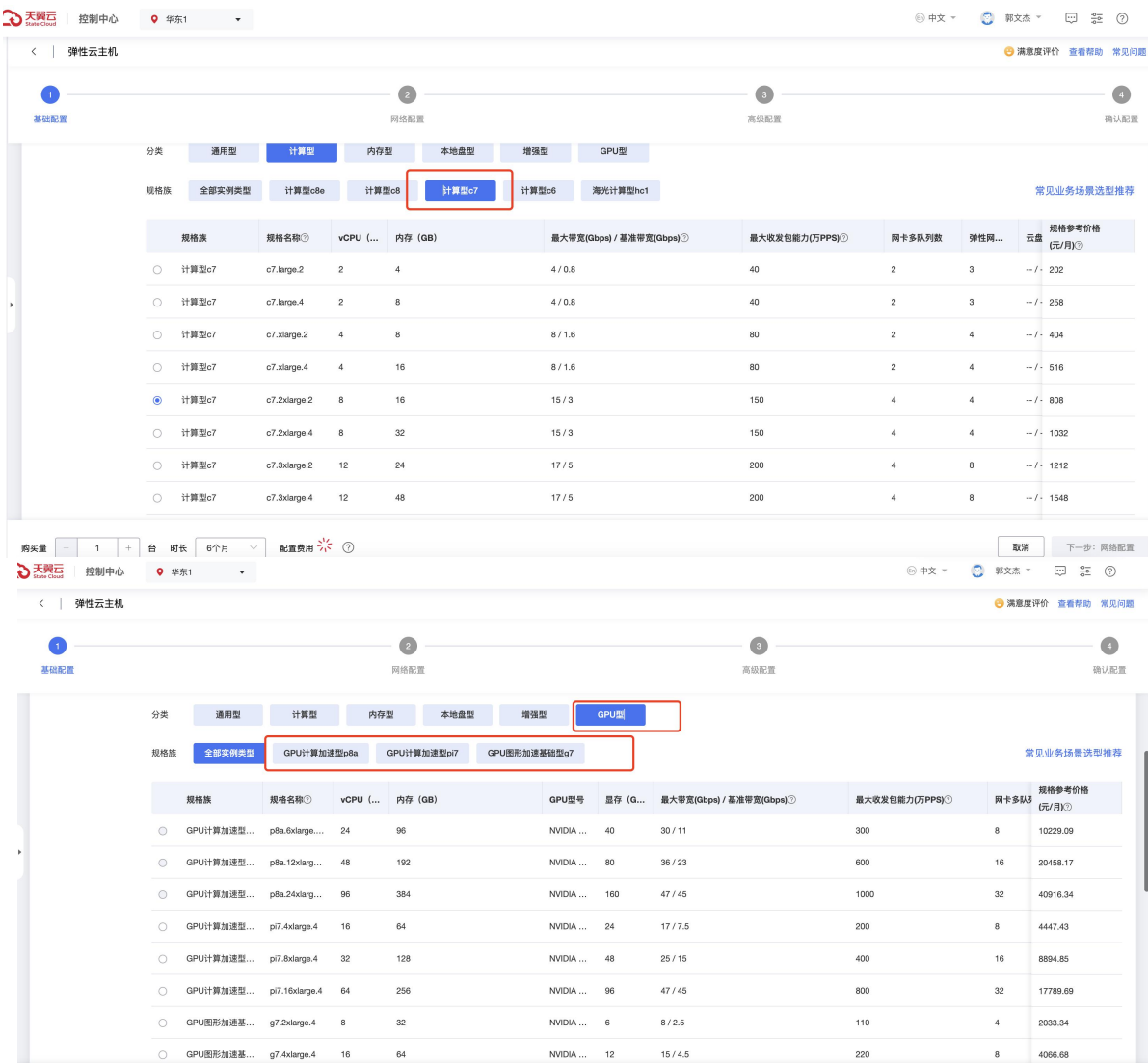
- 单击“创建云主机”，进入弹性云主机创建页



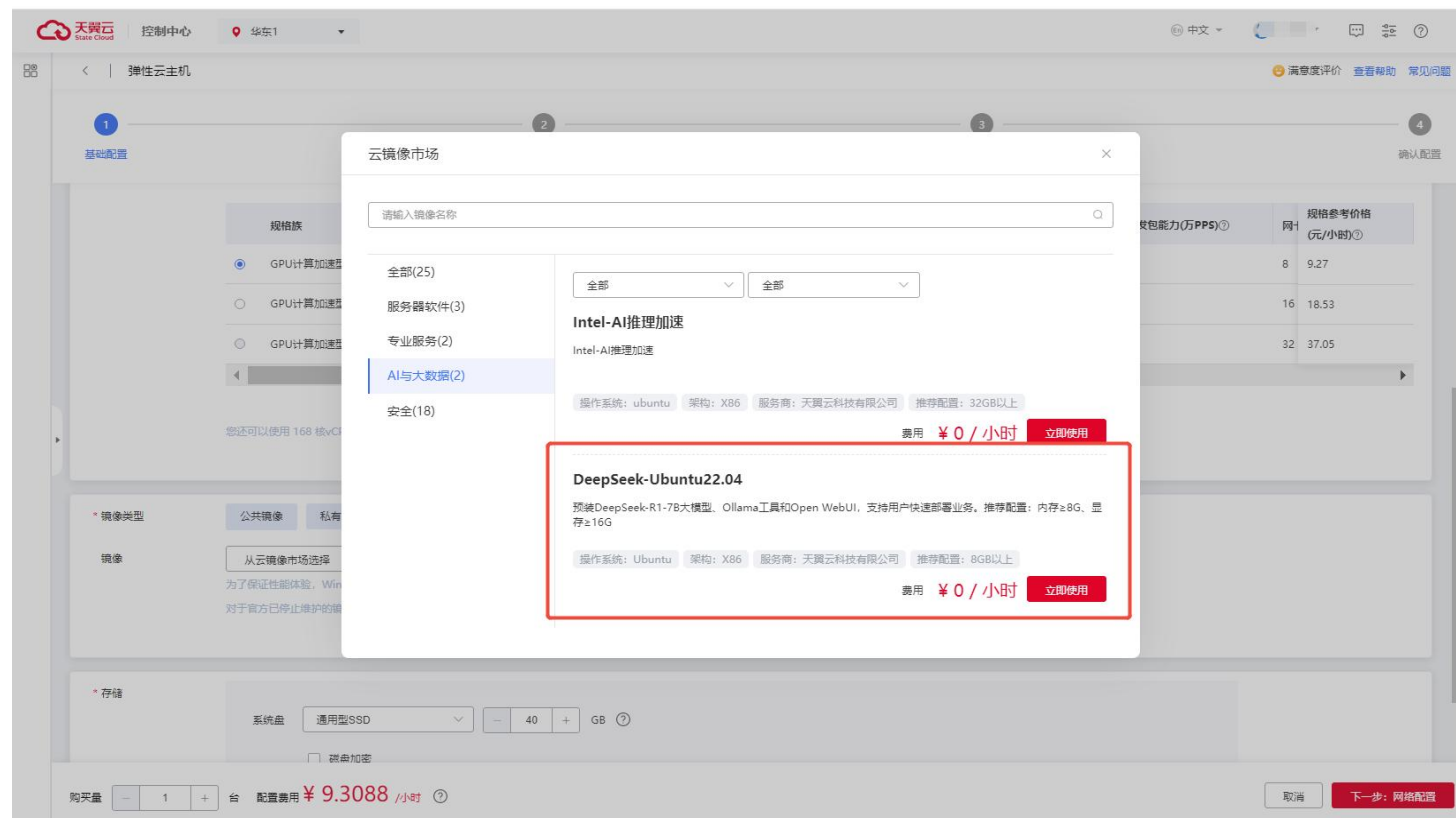


# 自动化部署deepseek云主机步骤 (2/6)

- 根据业务需求配置“计费模式”、“地域”、“企业项目”、“虚拟私有云”、“实例名称”、“主机名称”等。
- 选择规格。此处选择"CPU架构"为"X86"、"分类"为"GPU型"、“规格族”为"GPU计算加速型pi7"、“规格”为"pi7.4xlarge.4"。
- 同时针对想要尝鲜测试用户，可以使用C7系列，规格不低于4C8G的云主机，也可安装Deepseek专有系统



- 选择镜像。“镜像类型”选择“镜像市场”，在云镜像市场中选择预置了DeepSeek R1模型的DeepSeek-Ubuntu22.04镜像。
- 当前暂时支持ubuntu版本的deepseek镜像，如客户有其他系统镜像需求，可以通过支撑矩阵的联系人，传递需求，协助客户完成镜像加载。



# 自动化部署deepseek云主机步骤 (4/6)

- **网络及高级配置：**设置网络，包括"网卡"、"安全组"，同时配备 "弹性IP" 用于下载和访问模型；设置高级配置，包括"登录方式"、"云主机组"、"用户数据"；
- **确认并支付即可开通。**

天翼云 控制中心 华东1

弹性云主机

1 基础配置 2 网络配置 3 高级配置 4 确认配置

**配置**

**基础配置**

云主机名称	ecm-149e	地域	华东-华东1	计费模式	包年/包月
企业项目	default	虚拟私有云	vpc-df8f(15...)	镜像	DeepSeek-Ubuntu22.04
规格	GPU计算加速型pi7   pi7.4xlarge.4   16核 64G   GPU NVIDIA ...	系统盘	通用型SSD 60GB 未加密	可用区	随机分配

**网络配置**

安全组	Default-Security-Group	弹性IP	自动分配	IP版本	IPv4
带宽	12Mbps	主网卡	subnet-df8f(15...4)		

**高级配置**

登录方式	密码	云主机组	—	委托策略名称	—
标签键 (标签值)	—				

您还可以创建 48 台云主机。

购买时长

1个月 2个月 3个月 4个月 5个月 6个月 7个月 8个月 9个月 10个月 11个月 1年 2年 3年 4年 5年

☐ 启用自动续订

购买量

1

选择使用已有弹性公网IP或给网卡分配固定IP地址后，不能批量创建云主机

协议

☒ 我已阅读并同意相关协议 《天翼云弹性云主机服务协议》，《天翼云弹性云主机服务等级协议》，《天翼云硬盘服务协议》，《天翼云弹性IP服务协议》，《弹性云主机产品个人信息保护声明》

配置费用 ¥ 4841.43

上一步 立即购买

- 天翼云deepseek镜像，预置ollama工具，并预置了Open WebUI，通过web界面进行模型交互；
- 确保web界面正常打开，需要确保安全组策略中放行3000端口。



添加规则

\* IP版本: ☒ IPv4 ☐ IPv6

\* 方向: ☒ 入方向 ☐ 出方向

\* 授权策略: ☒ 允许 ☐ 拒绝

\* 优先级:  1  ⓘ

\* 协议:    
TCP、UDP报文分片后, 分片不带有端口信息, 需要将端口范围指定为1-65535, 不进行端口过滤。

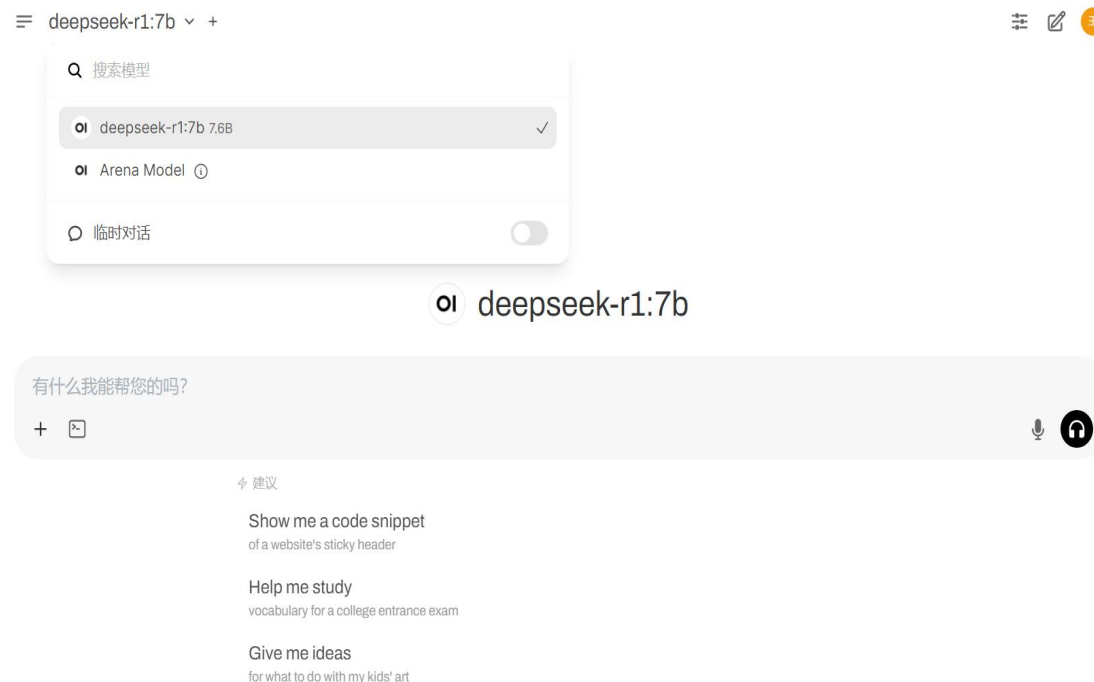
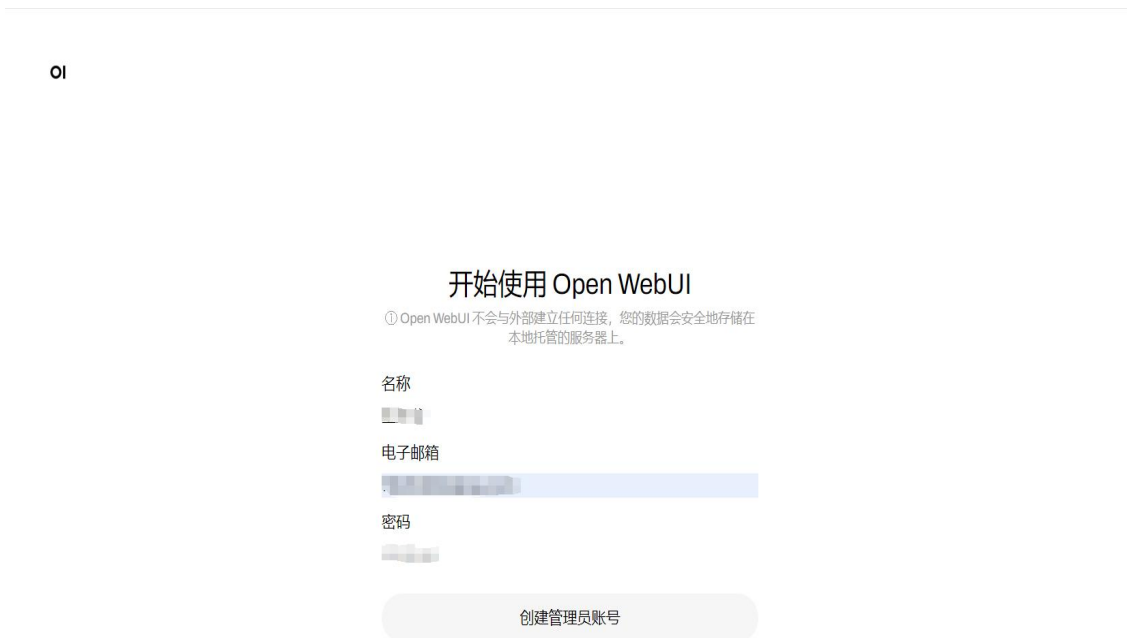
\* 端口范围:

\* 源地址: ☒ IP ☐ 安全组 ☐ 前缀列表  
 ⓘ

描述:

# 自动化部署deepseek云主机步骤（6/6）

- 访问DeepSeek模型的可视化界面。登录地址为：<http://{公网ip地址}:3000>。
- 镜像集成了可视化界面，完成管理员账户注册即可，并可以开始进行DeepSeek模型体验





# 感谢聆听