

# 面向审计行业DeepSeek 大模型操作指南

版本 1.0 | 适用对象：审计从业人员

计算机学院大模型团队提供

2025 年 2 月 8 日

# 目录

|                        |    |
|------------------------|----|
| 1 DeepSeek 基本概况 .....  | 3  |
| 2. DeepSeek 主要版本 ..... | 4  |
| 3. DeepSeek 审计能力 ..... | 5  |
| 4. DeepSeek 部署方法 ..... | 6  |
| 4.1 官方渠道 .....         | 6  |
| 4.1.1 网页版使用 .....      | 6  |
| 4.1.2 手机版使用 .....      | 8  |
| 4.2 第三方渠道 .....        | 8  |
| 4.2.1 硅基流动&华为云 .....   | 8  |
| 4.2.2 纳米 AI 搜索 .....   | 9  |
| 4.2.3 阿里云 .....        | 10 |
| 4.2.4 百度智能云 .....      | 11 |
| 4.2.5 火山引擎 .....       | 11 |
| 4.2.6 其他平台 .....       | 12 |
| 4.3 本地部署 .....         | 12 |
| 4.3.1 下载 ollama .....  | 13 |
| 4.3.2 合适版本安装 .....     | 13 |
| 4.3.3 输入安装代码 .....     | 15 |
| 4.3.4 测试部署模型 .....     | 17 |
| 4.3.5 部署非量化模型 .....    | 18 |
| 5. DeepSeek 审计助手 ..... | 20 |
| 5.1 基础操作场景 .....       | 20 |
| 5.2 审计工作辅助 .....       | 21 |
| 5.3 审计学习考试 .....       | 22 |
| 5.4 其他提示 .....         | 23 |

## 1. DeepSeek 基本概况

DeepSeek 是由杭州深度求索人工智能基础技术研究有限公司（简称“深度求索”）开发的一系列人工智能模型。该模型拥有数以亿计甚至更多的参数，通过在海量文本数据上进行预训练，学习到丰富的语言结构和语义信息；并支持智能对话、准确翻译、创意写作、高效编程、智能解题和文件解读等多种功能。其“深度思考”和“联网搜索”功能能够更全面地理解用户问题并提供准确答案。

杭州深度求索人工智能基础技术研究有限公司公司成立于 2023 年 7 月 17 日，专注于开发先进的大语言模型（LLM）和相关技术。自成立以来，公司在 AI 领域取得了显著成果，主要使用数据蒸馏技术，得到更为精炼、有用的数据。2024 年 1 月 5 日，发布 DeepSeek LLM（深度求索的第一个大模型），目前，DeepSeek-R1、V3、Coder 等系列模型已上线国家超算互联网平台。英伟达称，DeepSeek-R1 是最先进的大语言模型，亚马逊和微软也接入 DeepSeek-R1 模型。DeepSeek 大模型在多个基准测试中表现优异，尤其是在代码和数学任务上，超越了其他开源模型，甚至与领先的闭源模型（如 GPT-4 和 Claude-3.5-Sonnet）不相上下。

DeepSeek 被业界认为“以高性价比著称的 AI 模型服务商”，原因是这家公司的出现极大地降低了大模型训练和应用的成本，如该公司开发的 DeepSeek-V3 训练成本仅 557.6 万美元，而 OpenAI 训练 GPT-4 所花费的成本高达 7800 万美元甚至是 1 亿美元，双方的成本相差至少 10 倍。DeepSeek-V3 在数学、代码能力和中文知识问答方面还超过了 GPT-4，可以说是性价比超高。此外，DeepSeek 团队只有 139 名研发人员，而开发 GPT 的 OpenAI 团队则有 1200 名研究人员。

在审计领域，DeepSeek 大模型能够帮助审计人员高效处理各类多源异构的审计数据、识别风险、提升审计质量；通过自动化的数据处理、智能化的风险识别和定制化的报告生成等功能，帮助审计人员降低人工成本、提高审计质量和效率。

## 2. DeepSeek 主要版本

目前，DeepSeek 的核心版本主要有 DeepSeek-V3、DeepSeek-R1、Janus Pro，表 1 中列出了这 3 个核心版本的特点和适用场景。

表 1 DeepSeek 核心版本与适用场景

| 模型版本               | 发行时间       | 模型大小 | 核心能力                        | 适用场景示例               |
|--------------------|------------|------|-----------------------------|----------------------|
| DeepSeek-V3        | 2024-12-26 | 671B | 通用自然语言处理（NLP），支持长文本理解、多语言交互 | 合同条款解析、政策法规匹配、审计报告生成 |
| DeepSeek-R1        | 2025-1-20  | 671B | 复杂逻辑推理，强化数学与代码生成能力          | 财务数据分析、异常检测、风险建模     |
| DeepSeek Janus Pro | 2025-1-28  | 7B   | 多模态（文本+图像+表格）理解与生成          | 票据识别、图表数据关联分析、可视化报告  |

这三个版本原始模型权重已经在 hugging face 上开源，用户可以免费下载。国内使用可以通过其镜像 (<https://hf-mirror.com/>) 获取。DeepSeek-V3 和 DeepSeek-R1 的模型参数量较大，达到了 671B，直接部署这两个模型需要 1.3~2 TB（FP16）的显存支持（如 128 卡 H100 的集群）。为方便一般用户本地使用，DeepSeek 团队使用 Qwen2.5 和 Llama3.3，以 DeepSeek-R1 为教师模型，蒸馏了 6 款小模型，包含 1.5B~70B 在内共有 6 个尺寸，如表 2 所示。

表 2 DeepSeek-R1 蒸馏的 6 个尺寸的模型

| 蒸馏的模型                         | 基座模型                                   | 下载地址                        |
|-------------------------------|----------------------------------------|-----------------------------|
| DeepSeek-R1-Distill-Qwen-1.5B | <a href="#">Qwen2.5-Math-1.5B</a>      | <a href="#">HuggingFace</a> |
| DeepSeek-R1-Distill-Qwen-7B   | <a href="#">Qwen2.5-Math-7B</a>        | <a href="#">HuggingFace</a> |
| DeepSeek-R1-Distill-Llama-8B  | <a href="#">Llama-3.1-8B</a>           | <a href="#">HuggingFace</a> |
| DeepSeek-R1-Distill-Qwen-14B  | <a href="#">Qwen2.5-14B</a>            | <a href="#">HuggingFace</a> |
| DeepSeek-R1-Distill-Qwen-32B  | <a href="#">Qwen2.5-32B</a>            | <a href="#">HuggingFace</a> |
| DeepSeek-R1-Distill-Llama-70B | <a href="#">Llama-3.3-70B-Instruct</a> | <a href="#">HuggingFace</a> |

即使经过了蒸馏，7B 模型也需要 20~25G 的显存，即使是 24G 的 4090 显卡，部署也存在一定的风险。为此，在个人使用时，很多会将这类模型进行进一步量

化，以缩减模型大小，ollama 官方拉取的 DeepSeek 模型即是通过 4bit 量化后的模型。

*这里需要注意：无论是模型蒸馏还是量化，都会或多或少降低模型的能力。*

### 3. DeepSeek 审计能力

#### （一）数据采集与预处理

DeepSeek 支持多种数据源的接入，包括财务系统、ERP 系统和数据库等，确保数据获取的全面性。

通过数据清洗、缺失值填补、异常值检测和格式转换等操作，DeepSeek 能够自动清洗、转换和整合数据，确保数据质量，并将不同来源的数据统一格式化，为后续分析提供高质量的数据基础。

#### （二）数据分析与挖掘

DeepSeek 提供多种数据分析工具，如趋势分析、比率分析和异常检测等，帮助审计人员快速识别潜在的风险区域。DeepSeek 还可以进行时序分析，揭示财务数据中的潜在问题。

DeepSeek 利用机器学习算法识别潜在风险和异常交易，通过结合历史数据训练风险识别模型，实时监控异常交易、非正常模式和潜在的舞弊行为。

#### （三）支持自定义分析模型

用户可以根据具体审计需求自定义分析模型，针对特定场景（如税务审计、资产管理审计等）设定独特的分析参数。

#### （四）风险识别与评估

DeepSeek 基于预设规则和机器学习模型识别潜在风险领域，通过预设的审计规则和数据驱动的机器学习模型，自动识别潜在风险区域，帮助审计人员发现财务漏洞、操作风险或法律风险。

DeepSeek 会对识别出的风险进行评估和排序，根据风险的严重程度、发生概率和影响范围，自动评估并排序，帮助审计人员优先处理最关键的风险点。

(五) 审计证据收集与管理

通过 DeepSeek 的自动化分析，系统能够生成详细的审计底稿，包括审计过程、分析方法、数据来源及审计结果等内容，确保审计工作的透明性和可追溯性。

DeepSeek 支持审计证据的电子化存储和管理，审计证据以电子形式存储，支持文档管理、版本控制和权限管理，方便审计人员快速查阅和追溯。

(六) 可视化与报告生成

DeepSeek 提供丰富的可视化图表，包括图表、热力图和流程图等，帮助审计人员直观展示分析结果，提升报告的可读性和说服力。

系统能够自动生成标准化的审计报告，包含详细的数据分析结果、风险评估和审计结论等内容，显著减少报告编写时间。

4. DeepSeek 部署方法

使用 DeepSeek 主要有三种渠道：官方渠道、第三方渠道、本地部署。这三种渠道各自特点如表 3 所示。

表 3 DeepSeek 三种使用渠道对比

| 渠道    | 优点                      | 缺点                               |
|-------|-------------------------|----------------------------------|
| 官方渠道  | 功能齐全、操作简单（联网搜索/跨设备同步）   | 高峰期易崩溃，取决于流量，看运气                 |
| 第三方渠道 | 规避官方崩溃风险，国产 GPU 加速或白嫖算力 | 功能受限（如对话记录不保存），需实名认证/复杂配置        |
| 本地部署  | 隐私性强、永久离线，定制化模型选择       | 依赖硬件性能（需高配电脑），技术门槛较高，大部分部署的是蒸馏版本 |

4.1 官方渠道

DeepSeek 官方分为网页版和移动版，网页版用户点击“开始对话”并注册后即可使用；移动版需通过手机下载注册后使用，两者功能相同。

4.1.1 网页版使用

访问链接：<https://chat.DeepSeek.com/>，可以在任何设备和浏览器打开。之前从未登录过的用户需要进行登陆，使用手机号、微信或者邮箱登陆即可。如图 1

所示，输入自己的手机号，点击发送验证码，然后通过接收到的验证码登录即可。



图 1 DeepSeek 注册页面

登录成功后，进入图2 所示界面，然后点击“开始对话”就可以使用。



图 2 DeepSeek 官方网站主界面

不过需要注意，那就是如何选择 V3 还是 R1 模型，可以参考下图。此外还

可根据需要，选择是否勾选“联网搜索”。



图 3 不同版本的 DeepSeek 选择

### 4.1.2 手机版使用

手机版的使用和电脑版基本一样，根据需要是否激活 R1 即可。唯一不同的是需要在手机安装对应的App。安装方法如下：

方法 1：进入 DeepSeek 官网后，将鼠标移动至“获取手机 APP”处，扫描弹出的 APP 下载二维码（<https://download.DeepSeek.com/app/>），然后选选择对应的下载渠道即可。

方法 2：在手机自带的应用商城中，搜索DeepSeek，点击下载安装即可。

## 4.2 第三方渠道

随着 DeepSeek 模型迅速走红，官方平台面临访问量激增的压力，经常遇到服务拥堵的情况。不过，国内主流云计算平台已全面接入 DeepSeek 模型，为用户提供 了稳定可靠的替代方案。这些非官方 渠道提供 了 三类模型选择：DeepSeek-V3 模型、完整版 DeepSeek-R1 模型(671B 参数)、轻量级 DeepSeek-R1 模型（参数规模从 1.5B 到 70B 不等）。其中，完整版 R1 模型保留了全部 671B 参数，能发挥出最佳性能，但对计算资源要求较高，通常需要支付一定费用。轻量级模型则通过知识蒸馏技术，在保持核心功能的同时大幅降低了参数规模，可在普通算力环境下流畅运行，为用户提供了更灵活的选择。

### 4.2.1 硅基流动&华为云

硅基流动与华为云团队联合首发并上线了基于华为云昇腾云服务的 DeepSeek R1/V3，推理服务目前支持 V3 和 R1 大模型，以及多款 R1 蒸馏小模型。

在硅基流动的一站式大模型云服务平台 SiliconCloud 上（网址为：<https://siliconflow.cn/zh-cn/>），用户注册后可以在网页右侧选择 DeepSeek-R1 等模型进行体验使用，如图4 所示。

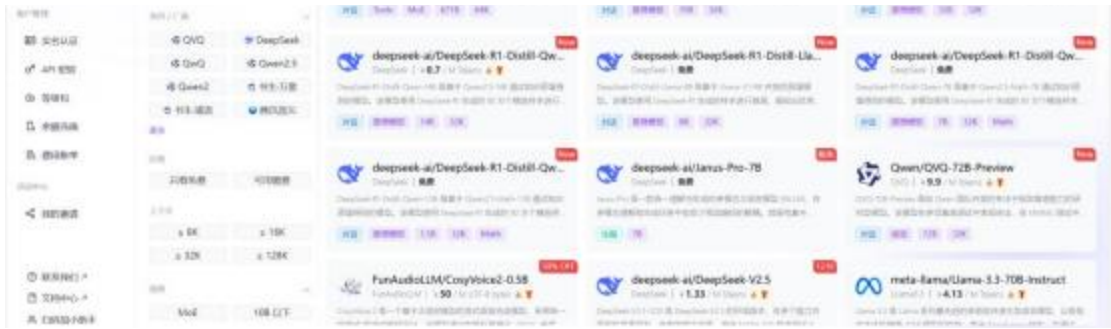


图 4 硅基流动模型广场主界面

#### 4.2.2 纳米 AI 搜索

360 宣布在其旗下纳米 AI 搜索中开通“DeepSeek 高速专线”，用户可在手机应用商店中下载并安装“纳米 AI 搜索”，注册登录后点击底部“大模型”，进入如图 5 所示左边画面。随后选择“DeepSeek-R1-满血版高速专线”，即可进入图 5 所示右边画面。接着就可以在下方的输入框中输入你需要的问题了。



图 5 360 纳米 AI 搜索 app

### 4.2.3 阿里云

阿里云 PAI Model Gallery 支持用户通过云平台一键部署 DeepSeek-V3 、 DeepSeek-R1 模型及其蒸馏版本。用户可以根据业务需求选择部署不同参数量的模型。具体请参考链接：

<https://help.aliyun.com/zh/pai/user-guide/one-click-deployment-DeepSeek-v3-model>



图 6 阿里云模型界面

#### 4.2.4 百度智能云

百度智能云提供两种使用 DeepSeek 模型的途径：一是在模型广场调用 DeepSeek 的 V3 和 R1 模型的 API；二是在体验中心与这两款模型直接对话。其使用网址为：[https://cloud.baidu.com/product-s/qianfan\\_modelbuilder](https://cloud.baidu.com/product-s/qianfan_modelbuilder)。



图 7 百度智能云会话界面

#### 4.2.5 火山引擎

火山引擎支持 V3/R1 等不同规模的 DeepSeek 开源模型，用户可以通过两种方式使用这些模型。一是在火山引擎机器学习平台 veMLP 中部署，适用于自己

进行模型定制、部署、推理的企业；二是在火山方舟中调用模型，适用于期望通过 API 快速集成预训练模型的企业， 目前已经支持 4 个模型版本。其访问的网络地址为：<https://www.volcengine.com/product/ark>



图 8 火山引擎模型界面

4.2.6 其他平台

- (1) 无问芯穹（ 目前提供蒸馏版 R1）：<https://cloud.infini-ai.com/genstudio>
- (2) 国家超算互联网平台（ 目前提供蒸馏版 R1）：<https://chat.scnet.cn/>
- (3) 腾讯云 TI-ONE：<https://console.cloud.tencent.com/tione/v2/aimarket>
- (4) 秘塔 AI 搜索：<https://metaso.cn/>

4.3 本地部署

这是最保险的一种方法，即在本地电脑上安装 DeepSeek 。只要电脑正常运行，用户就可以随时使用 DeepSeek。

但这种方法有一个缺点，那就是一般来说个人电脑的性能有限，只能装“蒸馏版”的 DeepSeek ，这个版本的 DeepSeek 需要占用的电脑资源要比满血版少的多， 同样性能也差不少。

这种方式适合的人群为：1 需要保证数据安全，不能联网 ；2 对于性能的忍受度较高 ；3 只需要普通的 AI 功能。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。  
。如要下载或阅读全文，请访问：

<https://d.book118.com/867154161121010035>