

DeepSeek使用教程蓝皮书

从入门到进阶完整指南



目录CONTENTS

第一章：DeepSeek概述

1.1 DeepSeek核心功能与优势	01
1.2 如何快速上手	04

第二章：提示词使用技巧

2.1 把握DeepSeek-R1提示词的两大核心要素	07
2.2 DeepSeek-R1提示词的具体使用技巧	11
2.3 生活场景实战案例示范	16

第三章：DeepSeek部署教程

3.1本地部署	18
3.2 云端部署	22

第四章 DeepSeek的应用场景和前景

4.1 DeepSeek应用场景	26
4.2 DeepSeek应用前景	30

版权声明

本文件内容主要来源于公开新闻报道。所有新闻内容均基于已公开发表的信息。

本文件内容受版权保护，禁止任何形式的商业性使用。未经 DAC 全球数据资产理事会明确书面授权，任何单位或个人不得以任何方式复制、传播、展示、表演、制作衍生作品或进行其他商业性利用。本声明的解释权归 DAC 全球数据资产理事会 所有。对于违反本声明规定的行为，DAC 全球数据资产理事会 有权采取法律手段维护自身权益。

第一章：DeepSeek 概述

1.1 DeepSeek 核心功能与优势

在人工智能技术高速迭代的时代背景下，推理模型作为核心驱动力，持续重塑各行业的发展格局。杭州深度求索科技推出的新一代推理模型 DeepSeek-R1，凭借卓越的性能表现与创新性技术架构，在核心功能与竞争优势方面实现多维度的突破性进展，成为人工智能领域的重要成果。

1、构建完备推理能力体系，实施参数规模分层设计

DeepSeek-R1 运用独特的参数规模分层设计理念，构建了一套完备的推理能力体系。其最大版本的参数规模高达 6710 亿，庞大的参数规模赋予模型强大的学习与推理能力，使其在复杂任务处理中表现出色。在数学推理领域，于 AIME 测试中，DeepSeek-R1 成功达成 79.8% 的准确率，这一数据表明其具备高效且精准解决高难度数学问题的能力，彰显出深厚的数学推理底蕴。在编程能力方面，DeepSeek-R1 在 Codeforces 竞赛中超越 96.3% 的人类选手。Codeforces 竞赛汇聚全球顶尖编程人才，DeepSeek-R1 在此脱颖而出，充分证明其强大的编程代码生成与逻辑推理能力。

为满足不同用户群体的多元化需求，DeepSeek-R1 推出精简版模型。该精简版模型采用知识蒸馏技术，将大规模模型的知识进行提炼与浓缩，从而实现于日常设备的本地化部署。这一创新设计意义重大，它使得从科研级的复杂计算到消费级的日常应用，均能受益于 DeepSeek-R1 的强大推理能力，形成覆盖范围广泛、层次分明的完整产品矩阵。

2、凸显高性价比优势，实现推理性能与成本的优化平衡

与 OpenAI-O1 系列对比，DeepSeek-R1 在推理性能上毫不逊色。在 MATH-500 测试中，DeepSeek-R1 达到 97.3% 的准确率，与 OpenAI-O1 系列保持同等推理性能水平。然而，DeepSeek-R1 的优势不仅体现于性能，更体现在极具竞争力的性价比方面。通过对强化学习策略的优化，DeepSeek-R1 有效降低成本。其 API 服务成本仅为竞品的三分之一，输入/输出 tokens 单价分别为 1 元/百万和 16 元/百万。在需要高频调用推理服务的商业场景中，这一成本优势具有决定性意义。

以电商领域为例，商家在进行商品推荐、用户行为分析等任务时需频繁调用推理服务，DeepSeek-R1 的低成本特性能够助力商家在确保服务质量的同时，大幅降低运营成本，提升企业盈利能力与市场竞争力。

3、创新双轨训练机制，以前沿技术驱动卓越性能

在技术架构层面，DeepSeek-R1 采用独有的双轨训练机制。一方面，借助长思维链微调技术，模型能够深入学习将复杂问题拆解为简单子问题的方法，进而逐步解决复杂问题，显著提升复杂问题的拆解能力。另一方面，运用无监督强化学习技术，突破传统训练范式的局限。在传统训练模式下，模型通常依赖大量标注数据，而无监督强化学习使 DeepSeek-R1 在有限标注数据的情况下，仍能维持卓越的推理性能。

这种创新的技术路径取得显著成效。在 SWE-bench 等工程类测试中，DeepSeek-R1 成功超越 O1 系列。尤为值得关注的是，它实现推理过程的自我反思与迭代优化。如同人类在解决问题时会不断总结经验、反思方法，DeepSeek-R1 在推理过程中能够持续优化自身推理策略，形成类似人类的问题解决思维链。这一特性使模型在面对复杂多变的问题时，能够更加灵活、高效地寻求解决方案。

4、强化灵活部署能力，拓宽模型应用边界

DeepSeek-R1 在部署灵活性方面表现卓越。它支持 4bit 量化压缩技术，通过对模型参数的压缩，在不影响模型性能的前提下，减少模型的存储空间与计算资源需求。同时，动态模型切换技术的应用，使模型能够依据不同的应用场景与需求，灵活切换不同的模型版本，以实现最佳性能表现。

这种强大的部署灵活性，使 DeepSeek-R1 既能适配云端分布式系统，如在电商数据爬虫架构中，能够高效处理海量数据，为电商企业提供精准的市场数据与用户信息；又能在本地设备实现每秒 60 tokens 的高效响应，在智能座舱交互场景中，能够快速响应用户语音指令，提供流畅的交互体验；在工业级数据处理中，亦能凭借高效的计算能力与灵活的部署方式，满足工业生产对数据处理的高要求。从智能座舱交互到工业级数据处理，DeepSeek-R1 的弹性部署能力显著拓宽其应用边界，为各行业的智能化升级提供有力支撑。

5、促进开源生态与商业服务协同，重塑行业发展格局

在市场应用层面，DeepSeek-R1 通过独特的开源策略构建强大的生态优势。

它将部分版本的模型权重免费开放，这一举措极大降低开发者的准入门槛，吸引全球众多开发者参与基于 DeepSeek-R1 的应用开发。同时，DeepSeek-R1 提供标准化 API 与定制化训练框架的双重支持。对于企业用户而言，标准化 API 使其能够快速接入基础服务，节省开发与成本；定制化训练框架则允许企业针对私有数据进行深度优化，打造契合企业特定需求的 AI 应用。

相较于同类模型，DeepSeek-R1 的显著特征在于成功平衡专业级推理性能与普惠化服务供给。在 LiveBench 测试中，其问题解决率较前代提升 46%，而单位解题成本降至传统方案的 1/30。这种技术突破与商业模式的协同创新，正深刻重塑 AI 推理服务的行业格局。它不仅为开发者和企业提供更强大、更经济、更灵活的 AI 推理解决方案，也推动整个 AI 行业朝着更加开放、创新、普惠的方向发展。

DeepSeek-R1 凭借在参数规模分层设计、性价比优势、双轨训练机制、灵活部署以及开源生态与商业服务协同等多维度的创新与突破，展现出强大的核心功能与竞争优势。展望未来，DeepSeek-R1 有望在更多领域发挥关键作用，推动人工智能技术的持续发展与广泛应用。

1.2 如何快速上手

1.2.1 访问方式

DeepSeek-R1 提供了网页版和 App 两种使用方式，无论你是电脑党还是手机党，都能找到适合自己的方式。

网页版：直接访问官网 chat.DeepSeek.com 使用，完全免费，无需下载，随时随地都能使用。



App：在各大应用商店搜索“DeepSeek”，认准蓝色鲸鱼图标，下载安装后即可开启 AI 之旅。

1.2.2 基础功能

基础功能进入 DeepSeek-R1 的聊天界面，你会看到两个关键开关：“深度思考”和“联网搜索”。



深度思考：开启这个功能，便是调用 DeepSeek-R1 模型，它会根据你的问题进行深度推理，给出更精准的答案、更有深度的回答。如果不开启，默认情况下，DeepSeek 使用的是 V3 模型。

联网搜索：这个功能让 DeepSeek-R1 能够参考最新的网络信息，确保回答的时效性和准确性。如果不开启，默认情况下，DeepSeek 使用的是截至 2023 年 10 月的训练数据。而且，如果提供的是网络链接，一定要打开这个功能，否则他会告诉你无法访问。

第二章：提示词使用技巧

早期的 AI 模型（如 SFT 模型）像刚学走路的孩子，需要明确的指令和分步骤示范（比如 Chain-of-thoughts 和 Few-shots），才能输出符合预期的内容。但像 GPT-o1、Deepseek-R1 这类经过强化学习训练的模型，已经像学会自主解题的学生——它们通过大量 Q&A 训练出了自己的推理能力。这时候如果还用“第一步、第二步”的详细指令，反而会干扰它们的思考节奏，就像给高中生强行规定用小学算术步骤解题。以基金报告案例为例

► 旧版 SFT 模型适用写法：

“请按以下步骤制作基金分析报告：

先收集基金净值、收益率等数据

用 Excel 计算各项指标

对比同类基金表现

分析市场环境

最后整理成报告”

► 新版 RL 模型适用写法：

“我需要一份某基金的财务分析报告，请包含核心财务指标、市场对比和风险评估等要素”

RL 模型会自动生成完整分析框架（数据收集→指标计算→横向对比→环境分析→报告整合），而分步骤指令反而会限制其发挥已掌握的金融分析能力。就像让大学生用小学算术步骤解题，反而降低效率。

2.1 把握 DeepSeek-R1 提示词的两大核心要素

2.1.1 输入输出设计

DeepSeek-R1 在输入输出设计以用户需求为导向，构建了极为简洁高效的模式。用户在使用时，只需向模型输入清晰明确的问题，无需复杂的格式编排或冗余表述，模型便会依据自身强大的算法和丰富的知识储备，快速处理信息，精准输出精炼且切题的结果。这种直接明了的输入输出设计，不仅大幅减少了用户与模型交互时的理解成本和操作难度，还显著提升了交互效率，让用户能更便捷、快速地获取所需信息，在各类应用场景中都展现出了极高的实用性和易用性。

1、输入设计：简洁明确的任务指令设定

在输入环节，旨在让用户以最简洁的方式传达核心任务目标。就如同给朋友布置任务时，我们只需清晰说明“要做什么”，而无需详述具体的操作步骤。例如，当用户希望进行金融领域的分析时，直接输入“请分析基金投资价值”，模型便能迅速捕捉到核心任务。这与传统的模型使用方式形成鲜明对比，以往在使用某些模型时，用户可能需要像教导实习生一样，详细罗列从数据收集、整理到分析的每一个步骤，过程繁琐且低效。

DeepSeek-R1 这种设计的背后，依托于其强大的深度推理能力。模型能够基于用户输入的简洁指令，自动在海量的知识储备中进行检索和分析，理解任务的深层含义和潜在需求。以委托专业顾问制定家庭理财方案为例，用户只需告知“制定家庭理财方案”，模型就如同专业顾问一样，明白需要综合考虑用户的收入、支出、资产状况、风险承受能力等多方面因素，进而展开全面的分析和规划。

2、输出设计：精准可控的成果交付标准

DeepSeek-R1 在输出环节，对于成果交付标准设定明确。这就像建筑工程验收时，不仅要求完成大楼建造，还需明确达到如 LEED 金级认证标准等具体要求。在与模型交互中，用户通过清晰表述“要什么样的结果”，能够引导模型输出更符合预期的内容。

例如，当用户要求模型“用中学生能理解的体育竞技案例解释金融风险对冲”时，模型接收到这一目标导向的指令后，会充分调动其强大的类比推理能力。它可能会将金融市场中的风险比作体育比赛中的各种不确定因素，如比赛中的对手实力、突发状况等，而将风险对冲策略类比为运动员或团队为应对这些不确定因

素所采取的战术安排，如防守策略、战术调整等。通过这样生动形象的类比，模型能够以中学生易于理解的方式，清晰阐述金融风险对冲这一复杂概念。

这种输出设计，使得用户能够根据具体需求，精确引导模型的输出方向和内容风格。无论是在学术研究、商业分析还是日常创作中，用户都能通过设定明确的输出标准，获得针对性强、质量高的结果。例如在撰写学术论文时，用户可以要求模型“以严谨的学术语言，结合最新的研究成果，阐述人工智能在医疗领域的应用前景”，模型便会遵循这一要求，输出符合学术规范、内容详实的文本。

DeepSeek-R1 的提示词输入输出设计，通过简化输入流程、明确输出标准，降低了用户的使用门槛，提高了交互效率，还充分发挥了模型的深度推理和强大的语言处理能力，为用户带来了更加智能、便捷、高效的使用体验，在众多应用场景中展现出独特的优势和价值。随着技术的不断发展和应用的深入，这种设计理念有望进一步拓展和优化，为人工智能与各行业的融合发展提供更有力的支持。

2.1.2 信息传递原则

在 DeepSeek-R1 的提示词输入输出体系中，动态平衡原则起着关键作用，它涉及任务描述与结果要求两个维度的连续光谱调节，旨在实现用户与模型之间高效、精准的信息传递，从而达成最优的交互效果。

1、任务描述维度：隐式模糊与显式具体的动态调节

在任务描述维度，存在着从隐式模糊到显式具体的递进空间，这一空间为用户提供了多样化的交互选择，以适配不同的应用场景和需求。

(1) 隐式指令：赋予模型探索空间

隐式指令类似于布置开放式研究课题，仅传达宽泛的任务核心，如“分析基金表现”。这种指令方式给予 DeepSeek-R1 充分的探索自由，模型凭借其强大的知识储备和深度推理能力，对任务进行全面的理解与分析。在这一过程中，模型会自主检索相关的金融数据、市场动态以及各类基金评价指标，从多个维度对基金表现展开研究，包括但不限于基金的历史收益情况、风险波动特征、投资策略的有效性等。通过这种方式，模型能够挖掘出潜在的、多层面的信息，为用户提供较为全面且具有探索性的分析结果，适用于用户对基金表现仅有初步了解需求，希望获取广泛信息的场景。

(2) 显式指令：提升输出针对性

显式指令则如同设计实验参数，通过精确限定任务范围和关键要素，使模型输出更具针对性。例如，“对比 2022-2024 年中美新能源基金夏普比率与最大回撤”。在此指令中，明确了时间范围（2022-2024 年）、研究对象（中美新能源基金）以及具体的分析指标（夏普比率与最大回撤）。DeepSeek-R1 接收到此类指令后，能够聚焦于特定的数据和分析方向，快速筛选出相关信息，避免在无关信息上的无效运算，从而高效地生成满足用户特定需求的结果。这种指令方式适用于用户对分析内容有明确要求，需要获取精准、特定信息的场景，如在投资决策过程中，投资者需要基于特定时间段和特定指标对基金进行对比分析，以做出合理的投资选择。

2、结果要求维度：抽象-具体轴的精准校准

在结果要求维度，沿着抽象-具体轴进行校准是实现精确信息传递的另一关键。

（1）抽象目标：依赖模型自主构建框架

抽象目标的设定类似要求“提供专业级行业分析”，这主要依赖 DeepSeek-R1 的知识储备和对行业的理解来自主构建分析框架。模型会根据任务描述和自身对行业的认知，确定分析的关键要点和逻辑结构。例如，在进行金融行业分析时，模型可能会从宏观经济环境、行业政策法规、市场竞争格局、主要企业发展态势等多个层面展开，运用专业的金融分析方法和理论，对行业的现状和未来发展趋势进行综合评估。这种方式能够充分发挥模型的专业能力和知识整合能力，为用户提供一个全面、系统的行业分析框架，适用于用户对分析结果的结构和内容没有具体要求，仅希望获取专业视角的整体分析的场景。

（2）具体目标：引导输出结构化形态

具体目标则如“生成含五年收益波动率、管理费比较等 10 项指标的对比矩阵”，通过明确的结构化要素引导模型输出特定形态的结果。在这种情况下，用户清晰地定义了所需结果的结构和内容要素，DeepSeek-R1 会按照这些要求，有针对性地组织数据和信息，生成符合要求的对比矩阵。这种方式能够确保输出结果的格式和内容满足用户的特定需求，便于用户直接应用于实际工作或决策中，如在投资研究报告撰写过程中，需要以特定的表格形式呈现基金对比数据，以便直观地进行分析 and 比较。

3、信息传递的整体动态平衡：类比导航系统的精准交互

DeepSeek-R1 提示词信息传递的动态平衡原则，可类比为使用导航软件。在导航过程中，用户只需准确输入起点和终点信息，导航系统便会依据算法自动规划出合理的路线。若用户在导航过程中频繁进行不必要的干预，如不断指挥“下个路口左转”，反而可能干扰系统的正常运算逻辑，导致导航结果的混乱。在与 DeepSeek-R1 交互时，用户应明确告知任务描述和结果要求，例如“我要从北京南站到首都机场 T3，要避开拥堵路段”，其中“从北京南站到首都机场 T3”相当于任务描述，“避开拥堵路段”则类似结果要求。用户通过这种清晰、准确的信息传递方式，将任务交给 DeepSeek-R1 的算法进行处理，模型便能依据自身的能力和算法，生成满足用户需求的高质量输出结果，实现高效、精准的交互。

DeepSeek-R1 的提示词信息传递原则，通过在任务描述和结果要求两个维度上的动态平衡调节，为用户提供了灵活且精准的交互方式。这一原则不仅充分发挥了模型的强大能力，还提升了用户与模型交互的效率和质量，在金融分析、学术研究、商业决策等众多领域展现出显著的应用价值，为推动人工智能在各行业的深度应用奠定了坚实基础。

2.2 DeepSeek-R1 提示词的具体使用技巧

在运用 DeepSeek-R1 进行任务处理时，编写清晰、具体的提示语是确保输出高质量结果的核心要素，其关键在于精准明确问题以及清晰界定期望的结果。掌握科学有效的提示词使用技巧，能够充分发挥 DeepSeek-R1 的强大功能，实现用户与模型之间的高效交互。

1、精准识别任务类型，适配提示词形式

(1) 隐式任务：引导模型自主探索

隐式任务的目标通常较为模糊，这就需要 DeepSeek-R1 凭借自身的算法和知识储备，对任务的具体要求进行自主推断。例如，当用户输入“探讨科技发展趋势”时，这属于典型的隐式任务。此时，模型会从多个维度展开探索，包括但不限于当前热门的科技领域，如人工智能、量子计算、生物技术等，以及这些领域在技术突破、市场应用、政策导向等方面的发展趋势。为了引导模型更全面且深入地探索，用户可以在提示词中适当补充背景信息或限定一定的探索范围，如“基于当前全球经济格局，探讨未来五年内可能对经济产生重大影响的科技发展趋势”。这样的提示词能够在给予模型探索自由的同时，避免其探索方向过于发散，从而获取更具针对性和价值的信息。

(2) 显式任务：精确传达任务指令

显式任务具有明确的目标，且对具体步骤和结果都有详细说明。以“计算 2023 年全年沪深 300 成分股的平均市盈率，并按照市盈率从低到高进行排序”为例，这一任务明确了时间范围（2023 年全年）、研究对象（沪深 300 成分股）、具体操作（计算平均市盈率）以及结果呈现方式（按照市盈率从低到高排序）。在处理显式任务时，用户务必确保提示词的准确性和完整性，避免出现歧义或信息缺失。每一个关键要素都应清晰表述，DeepSeek-R1 才能快速准确地理解任务要求，高效筛选和处理相关数据，输出符合要求的结果。

(3) 复合任务：合理拆解与整合

复合任务包含多个子任务，每个子任务可能具有不同的具体要求。比如，“为一家计划进入新能源汽车市场的企业制定市场战略，包括市场调研分析、目标客户群体定位、竞争对手分析以及初步的产品推广策略”。这一任务涵盖了市场调研、客户定位、竞品分析和推广策略等多个子任务。在编写提示词时，用户需要

将复合任务合理拆解，明确每个子任务的具体要求和先后顺序。可以采用结构化的表述方式，如“首先，进行新能源汽车市场的全面调研分析，包括市场规模、增长趋势、政策环境等；其次，基于调研结果，精准定位目标客户群体，分析其消费特征和需求；然后，对主要竞争对手进行详细分析，包括其产品特点、市场份额、竞争优势等；最后，根据前面的分析结果，制定初步的产品推广策略”。通过这种方式，DeepSeek-R1 能够有条不紊地完成各个子任务，并将结果进行有效整合，形成完整的市场战略方案。

2、根据任务类型，明确结果预期

（1）抽象结果：给予模型发挥空间

当任务预期的是抽象结果时，结果通常较为宽泛，不包含具体细节。例如，要求 DeepSeek-R1 “对当前教育行业的变革进行分析”，模型可能会从教育理念的转变、教学模式的创新、技术应用对教育的影响等多个宏观层面进行阐述。在这种情况下，用户可以在提示词中适当引导模型的分析方向，如“从社会发展和技术进步的角度，对当前教育行业的变革进行深入分析”，但无需对结果的具体内容和形式进行过多限制，让模型能够充分发挥其知识整合和分析能力，输出具有宏观视野和深度洞察的内容。

（2）具体结果：清晰界定形式与结构

对于期望获得具体结果的任务，结果具有明确的形式和结构，如数据表格、图表等。比如，“生成一份近五年中国智能手机市场份额的年度对比表格，包含主要手机品牌及其市场份额数据”。在编写此类提示词时，用户必须清晰界定结果的具体形式和结构，包括表格的列标题（年份、手机品牌、市场份额等）、数据的准确性要求以及数据来源的限定等。这样，DeepSeek-R1 能够根据明确的要求，准确地收集和整理数据，生成符合标准的数据表格，方便用户直接使用或进一步分析。

（3）结构化结果：规范模块内容与格式

结构化结果包含多个模块或部分，每个部分都有明确的内容和格式要求。以撰写一份商业计划书为例，要求 DeepSeek-R1 “生成一份包含项目概述、市场分析、产品或服务介绍、营销策略、财务规划和风险评估等模块的商业计划书，每个模块内容不少于 500 字，采用严谨的学术格式，引用权威数据和案例进行论证”。

在这种情况下，用户不仅要明确各个模块的内容要求，还要规范格式，如字体、字号、段落间距、引用格式等。通过详细的提示词设定，DeepSeek-R1 能够按照既定的结构和要求，生成内容丰富、格式规范的商业计划书，满足用户在实际商业活动中的需求。

3、编写提示语的技巧

（1）使用量词和限定词

量词：在提示词中运用“5 种”“10 个”等量词，能够有效突破思维惯性。例如，当要求 DeepSeek-R1 生成解决某类问题的方案时，使用“提供 5 种不同的创新解决方案”，可促使模型从不同角度思考，生成多样化的解决方案，避免方案的单一性和局限性。这是因为模型在处理任务时，会依据量词的限定，全面调动其知识储备，进行更广泛的搜索和推理，从而挖掘出更多潜在的解决方案。

限定词：通过领域限定词能够保持思维的发散方向，避免思维过于发散或过于狭窄。例如，“在远程办公场景下，探讨提高团队协作效率的方法”，“在远程办公场景下”这一限定词明确了问题的应用场景，使模型能够聚焦于该场景下的相关因素，如远程沟通工具的使用、虚拟团队管理等，生成更具针对性的内容。又如，“基于 SaaS 平台，分析企业数字化转型的策略”，“基于 SaaS 平台”限定了分析的基础平台，模型会围绕 SaaS 平台的特点、优势以及在企业数字化转型中的作用等方面展开分析，输出更贴合实际需求的内容。

（2）结构化指令

模块化设计：对于复合任务，将任务分解为多个模块，每个模块设定明确的要求，有助于提高模型处理任务的效率和准确性。例如，“生成含三大模块（宏观影响、企业对比、投资建议）的 Markdown 文档”，这种模块化的指令使模型清晰地了解任务的组成部分和各部分的具体要求，能够有针对性地进行内容生成。在生成过程中，模型会分别对宏观影响、企业对比和投资建议这三个模块进行独立的分析和内容创作，然后按照 Markdown 文档的格式要求进行整合，最终生成结构清晰、内容完整的文档。

具体格式要求：明确输出的具体格式，如 PDF、Excel 表格、Mermaid 图等，能让模型生成符合特定格式要求的结果，方便用户使用。例如，“制作包含时间序列折线图和表格的 PDF”，用户明确要求输出为 PDF 格式，且包含时间序

列折线图和数据表格。模型在处理任务时，会根据 PDF 格式的规范，生成相应的图表和数据，并进行合理的排版，确保生成的 PDF 文件能够满足用户在展示、打印或进一步分析等方面的需求。

（3）明确任务步骤

分步指导：对于复杂的任务，分步骤给出指导，每一步骤都有具体的要求，能够帮助模型更好地理解任务流程，提高任务处理的准确性。例如，“‘按’概念-公式-应用’结构排列，对机器学习中的决策树算法进行详细阐述”，模型会按照这一结构要求，首先对决策树算法的概念进行清晰的解释，然后给出相关的公式推导，最后结合实际案例说明其应用场景和方法。通过这种分步指导的方式，模型能够有条不紊地完成复杂任务，生成逻辑连贯、内容详实的结果。

示例参考：提供示例或参考模板，能帮助 DeepSeek-R1 更好地理解任务要求，生成更符合用户期望的内容。例如，“用厨房用具类比说明欧姆定律，包含至少 3 个生活案例”，通过给出“用厨房用具类比”这一示例和“至少 3 个生活案例”的要求，模型能够明确任务的表达方式和内容丰富度要求。模型会将欧姆定律中的电流、电压、电阻等概念与厨房用具中的水流、水压、水管粗细等进行类比，然后结合具体的厨房场景，如使用水龙头调节水流大小类比调节电压改变电流等，生成生动形象、易于理解的内容。

（4）专业术语和标准

专业术语：在提示词中使用专业术语和行业标准，能够确保输出的专业性和准确性。例如，在医疗领域，“使用最新临床指南标准，对某疾病的诊断和治疗方案进行分析”，模型会依据最新的临床指南标准，运用专业的医学术语，对疾病的诊断方法、治疗原则、药物选择等进行准确的分析和阐述。这使得生成的内容能够满足医疗专业人员的需求，为临床决策提供可靠的参考。

技术规范：对于技术任务，明确技术规范 and 标准至关重要。例如，“处理 $n \leq 100$ 的输入，设计一个高效的算法”，“处理 $n \leq 100$ 的输入”明确了算法的输入范围，模型在设计算法时会考虑在这一输入范围内的效率和准确性。又如，“提供状态码 415/503 应对方案”，模型会针对 HTTP 状态码 415（不支持的媒体类型）和 503（服务不可用），按照相关的网络技术规范，给出相应的故障排查和解决方法，确保生成的应对方案符合技术标准和实际应用需求。

在使用 DeepSeek-R1 时，用户通过精准识别任务类型，合理编写提示词，并根据任务类型明确结果预期，同时运用上述编写提示语的技巧，能够实现与模型的高效交互，充分发挥其强大的功能，获得高质量的输出结果。这不仅有助于提升工作效率和决策质量，还能推动 DeepSeek-R1 在更多领域的深入应用和创新发展。

2.3 生活场景实战案例示范

1、金融投资分析场景

(1) 隐式任务+抽象结果

Q: "评估特斯拉股票投资价值"

A: "给出专业风险评估报告"

(2) 显式任务+抽象结果

Q: "对比特斯拉与比亚迪 2023 年 Q3 财报中的研发投入占比"

A: "用汽车制造业的行业标准进行分析"

(3) 显式任务+具体结果

Q: "提取近五年苹果公司股息支付率的季度数据"

A: "制作包含时间序列折线图和数据表格的 PDF"

(4) 复合任务+结构化结果

Q: "分析美联储加息对科技股的影响，对比微软/谷歌/亚马逊的 β 系数变化"

"

A: "生成含三大模块（宏观影响、企业对比、投资建议）的 Markdown 文档，
每模块不超过 300 字"

2、教育辅导场景

(1) 隐式任务+抽象结果

Q: "解释光速不变原理"

A: "让初中生能理解基本概念"

(2) 显式任务+抽象结果

Q: "用厨房用具类比说明欧姆定律"

A: "包含至少 3 个生活案例"

(3) 显式任务+具体结果

Q: "设计三角函数入门课程的 5 个知识点"

A: "按'概念-公式-应用'结构排列"

(4) 复合任务+结构化结果

Q: "创建针对高考物理的 30 天复习计划"

A: "输出 Excel 表格，含每日主题/重点公式/典型例题编号，按力学-电磁"

学-热学模块划分”

3、医疗咨询场景

(1) 隐式任务+抽象结果

Q: "分析我的体检报告异常项"

A: "给出专业医学建议"

(2) 显式任务+抽象结果

Q: "对比幽门螺杆菌感染不同治疗方案的优劣"

A: "使用最新临床指南标准"

(3) 显式任务+具体结果

Q: "列出高血压患者饮食禁忌的 20 种食物"

A: "按钠含量/胆固醇/嘌呤等指标分类"

(4) 复合任务+结构化结果

Q: "根据患者数据（男/55 岁/BMI28/空腹血糖 7.8）制定健康管理方案"

A: "包含运动处方（频率/强度/时长）、膳食模板（早中晚热量分配）、监测指标（每周检测项目）"

4、代码开发场景

(1) 隐式任务+抽象结果

Q: "开发 TODO 列表应用"

A: "实现基本功能"

(2) 显式任务+抽象结果

Q: "用 React 构建支持分类标签的 TODO 应用"

A: "遵循组件化设计原则"

(3) 显式任务+具体结果

Q: "编写 Python 函数计算斐波那契数列第 N 项"

A: "包含时间/空间复杂度注释，处理 $n \leq 100$ 的输入"

(4) 复合任务+结构化结果

Q: "创建分布式爬虫系统抓取电商价格数据"

A: "输出包含架构图（Mermaid 格式）、核心类伪代码（Java/Python 双版本）、异常处理流程（状态码 415/503 应对方案）的技术文档”。

第三章：DeepSeek 部署教程

3.1 本地部署

3.1.1 选择适合你的 DeepSeek 模型大小

在部署前，要根据自己的硬件配置和使用场景选择合适的模型大小，具体参考如下：

模型版本	CPU	内存	显卡	存储	网络
DeepSeek-R1-1.5B	任意四核处理器	8GB	无需 GPU	12GB	本地部署无需高带宽，云端需 $\geq 100\text{Mbps}$
DeepSeek-R1-7B	Ryzen7 或更高	16GB	RTX3060 (12GB) 或更高	80GB	本地部署无需高带宽，云端需 $\geq 100\text{Mbps}$
DeepSeek-R1-14B	i9-13900K 或更高	32GB	RTX4090 (24GB) 或更高	200GB	本地部署无需高带宽，云端需 $\geq 100\text{Mbps}$
DeepSeek-R1-32B	Xeon8 核 +128GB 或更高	64GB	2-4 张 A100 80GB 或更高	320GB	多卡需 PCIe4.0 互 联，云端部署建议 $\geq 1\text{Gbps}$ 带宽
DeepSeek-R1-70B	Xeon8 核 +128GB 或更高	128GB	8+张 A100/ H100, 显存 $\geq 80\text{GB}/\text{卡}$	500GB+	NVIDIA NVLink/InfiniBand ($\geq 200\text{Gbps}$) 互联

3.1.2 安装 Ollama

Ollama 是一个本地模型运行工具，可以方便地下载和运行各种开源模型，正好适合我们的 DeepSeek 。

下载地址：<https://ollama.com/>

根据你电脑的操作系统选择下载的 Ollama 版本，并进行安装



Get up and running with large language models.

Run Llama 3.3, DeepSeek-R1, Phi-4, Mistral, Gemma 2, and other models, locally.

Download ↓

Available for macOS,
Linux, and Windows

Download Ollama


macOS


Linux


Windows

Download for Windows

Requires Windows 10 or later

3.1.3 下载并运行 deepseek 模型

按 Win 键输入 cmd 打开命令提示符并输入【Ollama】后按下回车，如下图所示即为安装成功。




```
Microsoft Windows [版本 10.0.26100.2894]
(c) Microsoft Corporation. 保留所有权利。

C:\Users\qeeyou>ollama
Usage:
  ollama [flags]
  ollama [command]

Available Commands:
  serve      Start ollama
  create     Create a model from a Modelfile
  show       Show information for a model
  run        Run a model
  stop       Stop a running model
  pull       Pull a model from a registry
  push       Push a model to a registry
  list       List models
  ps         List running models
  cp         Copy a model
  rm         Remove a model
  help       Help about any command

Flags:
  -h, --help      help for ollama
  -v, --version    Show version information

Use "ollama [command] --help" for more information about a command.
```

安装命令如下：安装完成后出现【success】则代表安装成功

1、内存低于 4GB 安装命令如下：

```
ollama run deepseek-r1:1.5b
```

2、内存在 8GB-12GB 安装命令如下：

```
ollama run deepseek-r1:7b
```

```
ollama run deepseek-r1:8b
```

3、内存在 12GB 以上安装命令如下：

```
ollama run deepseek-r1:14b
```

4、内存在 32GB 以上安装命令如下：

```
ollama run deepseek-r1:32b
```

5、内存在 70GB 以上安装命令如下：

```
ollama run deepseek-r1:70b
```

注意：内存越高，效果越好，但占用储存会更多。

```
>>> 如何使用deepseek
<think>

</think>

DeepSeek是一个专业的智能助手，它能够理解和执行多种编程语言和逻辑推理任务。以下是一些常见的问题和回答：

1. **编程问题**：如果你有特定的编程问题，可以提供代码或描述，我可以帮你编写解决方案。
2. **数学问题**：如果你需要解答数学问题，我会仔细分析并给出详细的步骤解释。
3. **逻辑推理**：对于复杂的逻辑推理问题，我能够逐步分解并逐步求解。
4. **自然语言理解与生成**：如果涉及自然语言处理或对话，我可以理解用户需求并提供相应的帮助和建议。
5. **可视化工具**：有些问题需要通过图形化工具来展示结果，我会准备一些适合的可视化资源。

如果你有具体的问题或任务，请提供更多细节，我会尽力为你解答！
```

3.2 云端部署

如果你电脑配置不够,又想直接在客户端应用里体验 DeepSeek-R1&V3 模型,可在本地安装以下产品,接入 SiliconCloud 或者阿里云百炼等 API 后(可自定义添加这两款模型),即可体验 DeepSeek-R1&V3。

大模型客户端应用: ChatBox、Cherry Studio、OneAPI、LobeChat、NextChat

代码生成应用: Cursor、Windsurf、Cline

大模型应用开发平台: Dify

AI 知识库: Obsidian AI、FastGPT

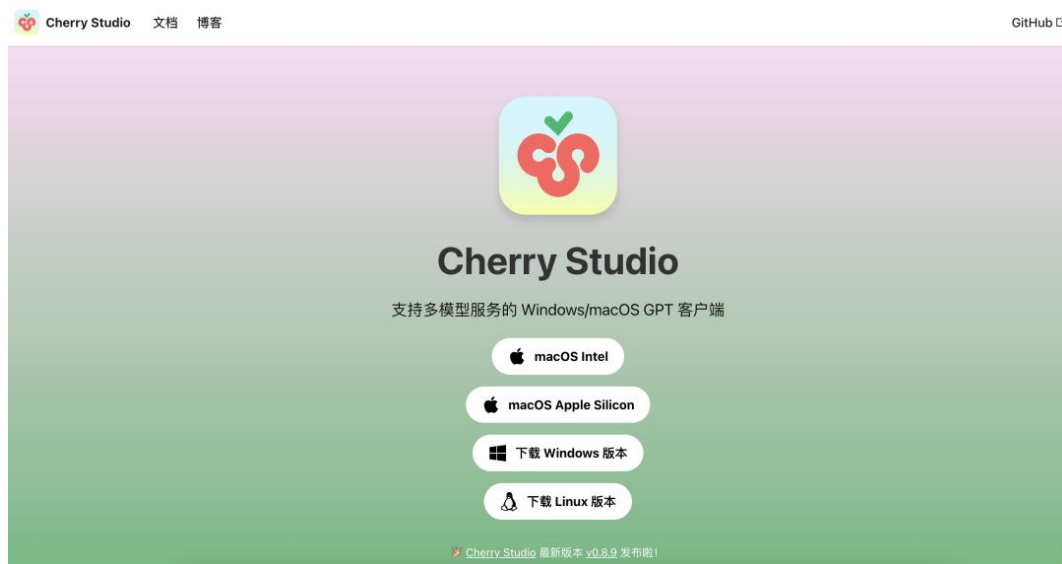
翻译插件: 沉浸式翻译、欧路词典

接下来以 Cherry Studio 为例,来学习如何部署

3.2.1 安装 Cherry Studio

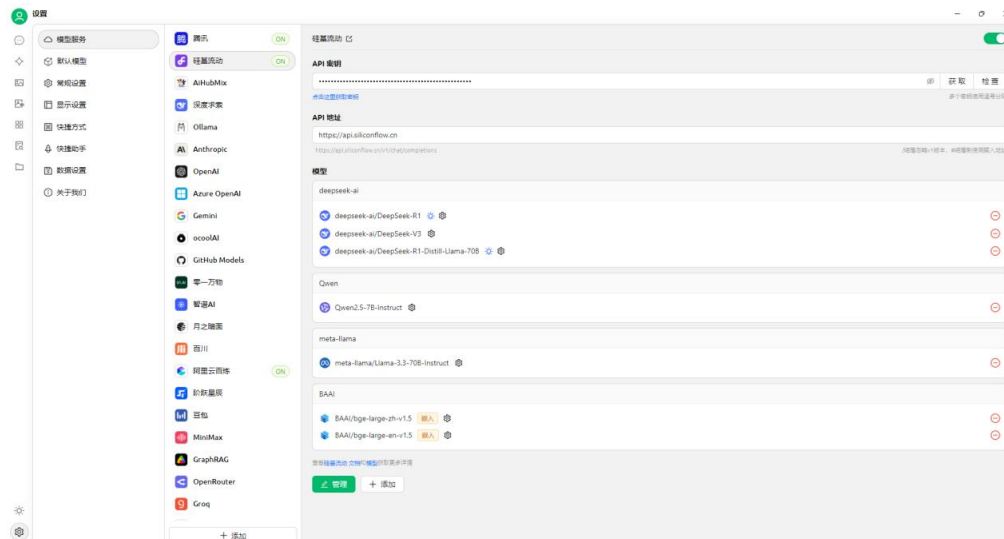
Cherry Studio 是一款支持多模型服务的桌面客户端,内置了超过 30 个行业的智能助手,旨在帮助用户在多种场景下提升工作效率。它兼容 Windows、Mac 和 Linux 系统,无需复杂设置即可使用。

进入 Cherry Studio 官网(<https://cherry-ai.com/>),按步骤下载安装对应版本的 Cherry Studio 即可直接使用。



3.2.2 配置 SiliconCloud 等线上模型服务

1、打开并点击 Cherry Studio 左下角的“设置”，在模型服务中选择默认的【硅基流动】。



2、注册硅基流动

访问下面的链接注册硅基流动 <https://siliconflow.cn/zh-cn/注册> 即可免费获得价值 14 元的 2000 万 token 的额度，可以用于后续模型使用。

欢迎登录SiliconFlow

+86 您的手机号

短信验证码 获取验证码

邀请码 (选填)

注册 / 登录

邮箱登录

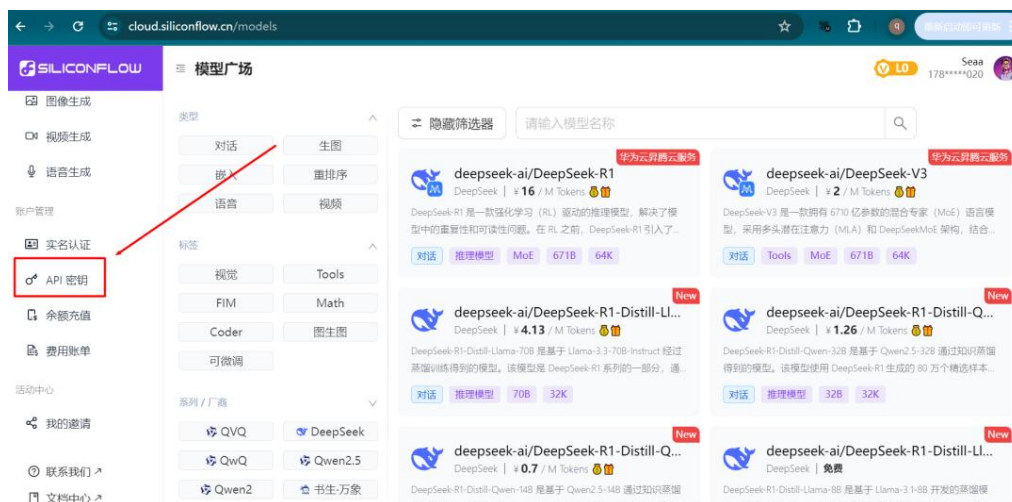
其它登录方式

GitHub

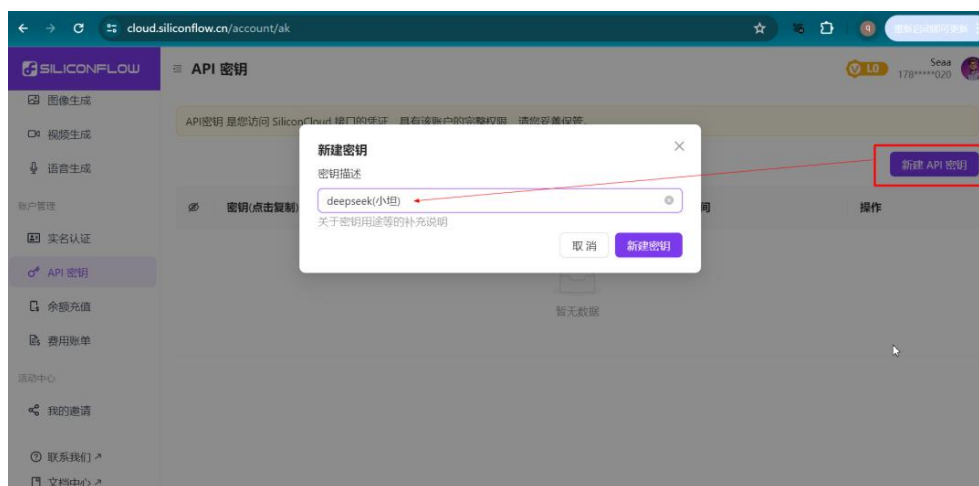
Google

注册登录即表示同意 [用户协议](#) 和 [隐私政策](#)

登录后在左边的导航栏找到 API 密钥。



点右上角的新建 API 密钥。密钥描述随便写，可以建多个密钥，别忘记这是啥就行。



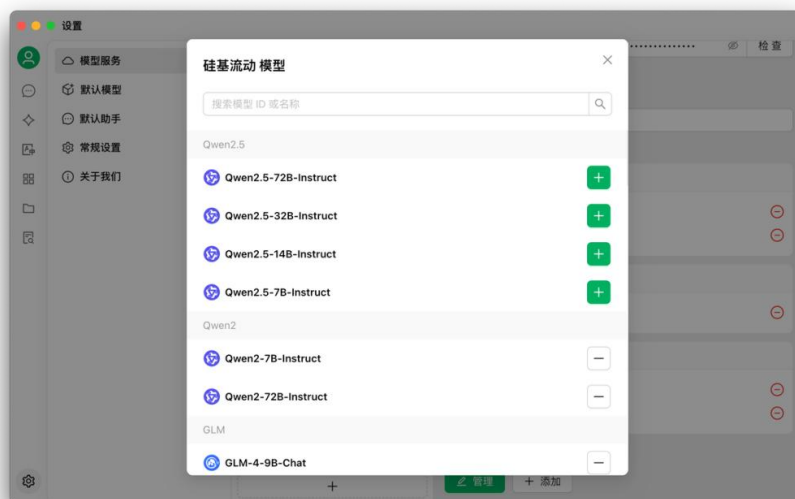
新建完成后会得到一个加密的 API Key 了，就是你的密钥，点击密码就能复制，后面会用到这个 API Key。



点击 API 密钥新建或复制已有密钥，并将密钥填入 Cherry Studio 的 API 密钥。



点击管理/添加模型。可选择添加 SiliconCloud 上的其他各类模型。



点击左侧菜单栏的“对话”按钮，然后在输入框内输入文字即可开始聊天。
你可以选择顶部菜单中的模型名字切换模型。



第四章 DeepSeek 的应用场景和前景

4.1 DeepSeek 应用场景

DeepSeek-R1 以其卓越的性能、创新的技术架构和出色的性价比，在多个行业实现了规模化应用接入，构建起一个覆盖消费电子、云计算、金融、通信等多领域的立体化生态布局，展现出强大的技术渗透力和生态扩展性，成为推动各行业智能化变革的重要力量。

1、消费电子领域：主流品牌深度融合，打造智能交互新体验

在消费电子领域，DeepSeek-R1 的应用成果显著。华为率先接入，开启了该模型在消费电子产品中的应用先河。随后，OPPO、荣耀、魅族、vivo 四大手机厂商迅速跟进，相继完成系统级整合，深度挖掘 DeepSeek-R1 的技术优势，为用户带来智能化交互升级。

OPPO Find N5 借助小布助手，依托 DeepSeek-R1 强大的自然语言处理能力，实现了精准的语音唤醒功能。用户只需说出唤醒词，手机便能迅速响应，执行各类指令，大大提升了操作便捷性。同时，其本地文件生成功能利用模型的智能创作能力，能够根据用户输入的简单提示，生成结构清晰、内容丰富的文档、邮件等，满足用户多样化的创作需求，有效提升了工作效率。

荣耀 MagicOS8.0 用户可通过 YOYO 智能体商店直接调用 DeepSeek-R1 的模型服务。无论是智能问答、信息检索，还是复杂的内容创作，用户都能在 YOYO 智能体商店中一站式完成。这不仅丰富了用户的交互体验，还为荣耀手机在智能助手领域树立了差异化竞争优势，进一步提升了产品的市场竞争力。

魅族 21 系列及 AR 眼镜产品线也已启动 DeepSeek-R1 的功能上线。在魅族 21 系列手机中，模型赋能语音助手，使其具备更智能的语义理解和交互能力，能够根据用户的语境和习惯，提供更加个性化的服务。在 AR 眼镜产品线中，DeepSeek-R1 助力实现实时的图像识别与信息交互，为用户带来沉浸式的增强现实体验，推动 AR 技术在消费级产品中的普及和应用。

目前，DeepSeek-R1 已形成覆盖国内主流手机品牌（除小米外）的移动端应用矩阵，构建起一个庞大的智能交互生态系统，推动消费电子行业朝着智能化、

个性化方向加速发展。

2、云计算领域：十大云平台深度对接，赋能云服务智能化升级

在云计算领域，华为云、腾讯云、阿里云、百度智能云、移动云、中国联通、京东云等十大云平台已完成与 DeepSeek-R1 的深度技术对接，共同探索智能化云服务的新范式。

移动云依托其分布广泛的 13 个智算中心，为 DeepSeek-R1 提供蒸馏与智能体编排服务。智算中心强大的计算能力，能够对模型进行高效的优化蒸馏，在保证模型性能的前提下，减少模型的计算资源需求，使其能够更加高效地运行。同时，智能体编排服务根据不同用户的业务需求和场景特点，灵活组合 DeepSeek-R1 的各项功能，为用户提供定制化的智能解决方案，满足企业多样化的业务需求。

京东云通过言犀平台实现了 DeepSeek-R1 的一键部署，极大地简化了企业接入模型的流程。企业只需通过言犀平台，即可快速将 DeepSeek-R1 集成到自己的业务系统中，无需复杂的技术搭建和配置，便能享受其强大的智能服务，大幅降低了企业使用人工智能技术的门槛，加速企业数字化转型进程。

中国联通星罗平台支持私有化与公有化场景的灵活切换，满足不同企业对数据安全性和隐私性的不同需求。对于对数据安全要求极高的金融、政府等行业客户，星罗平台提供私有化部署方案，确保企业数据在本地安全可控；对于一般性企业，公有化场景则提供便捷、高效的智能服务，降低企业运营成本。通过这种灵活的部署方式，中国联通能够为不同类型的客户提供定制化的云计算服务，进一步拓展了 DeepSeek-R1 的应用场景。

通过与十大云平台的紧密合作，DeepSeek-R1 为云计算服务注入了强大的智能动力，推动云服务向智能化、个性化、定制化方向迈进，助力企业在数字化时代实现更高效的发展。

3、金融行业：头部券商集中接入，加速金融智能化转型

金融行业对 DeepSeek-R1 的应用呈现爆发式增长态势。春节前后，包括华安证券、东兴证券、光大证券、中信建投证券等在内的 13 家头部券商已完成 DeepSeek-R1 的本地化部署，将其深度融入金融业务的核心环节。

在智能投研方面，DeepSeek-R1 强大的数据分析和预测能力，能够对海量的

金融市场数据进行实时分析和挖掘，为投资决策提供全面、准确的市场洞察和风险评估。通过对宏观经济数据、行业动态、公司财务报表等多维度数据的深度分析，模型能够快速生成投资研究报告，为分析师提供高效的研究辅助工具，大大提升了投研效率和质量。

在风险建模领域，DeepSeek-R1 利用其卓越的机器学习能力，构建更加精准的风险预测模型。通过对历史数据和实时市场数据的学习，模型能够准确识别潜在的风险因素，提前预警风险事件，帮助金融机构及时采取风险控制措施，降低风险损失。例如，在信用风险评估中，模型能够综合考虑客户的信用记录、财务状况、行业风险等因素，给出更加准确的信用评分，为金融机构的信贷决策提供有力支持。

随着越来越多的金融机构接入 DeepSeek-R1，其在金融行业的应用场景将不断拓展，有望推动金融行业实现智能化、数字化转型，提升金融服务的效率和质量，为实体经济的发展提供更有力的支持。

4、通信基础设施领域：三大运营商全面接入，提升通信智能化水平

在通信基础设施层面，中国移动、中国电信、中国联通三大基础电信企业已全面接入 DeepSeek-R1 模型，结合自身的 5G 网络与云平台能力，在多个领域实现了技术突破。

在客服系统优化方面，DeepSeek-R1 的自然语言处理能力使得客服机器人能够更加准确地理解用户的问题和需求，提供更加智能、高效的服务。通过对用户语音和文字信息的实时分析，客服机器人能够快速给出准确的解答和解决方案，大大缩短了用户等待时间，提升了客户满意度。同时，模型还能够对客服对话数据进行分析，挖掘用户需求和潜在问题，为运营商优化服务策略提供数据支持。

在网络智能运维方面，DeepSeek-R1 利用其强大的数据分析和预测能力，对网络运行数据进行实时监测和分析，提前预测网络故障和性能瓶颈，实现网络的智能化运维。通过对网络流量、设备状态、用户行为等多维度数据的分析，模型能够及时发现网络中的异常情况，并提供相应的优化建议，帮助运营商提高网络的稳定性和可靠性，降低运维成本。

此外，三大运营商还为 DeepSeek-R1 提供专属算力方案，确保模型在通信领域的高效运行。通过与运营商的合作，DeepSeek-R1 能够更好地服务于通信行业，

推动通信基础设施的智能化升级，为 5G 时代的各种应用场景提供更强大的技术支持。

国家超算互联网平台的上线标志着 DeepSeek-R1 进入“国家队”应用序列，支持 6710 亿参数满血版部署，进一步发挥其在大规模数据处理和复杂计算方面的优势。同时，DeepSeek-R1 与平头哥玄铁芯片完成 RISC-V 架构适配，为工业级应用场景的拓展奠定了坚实基础，有望在工业自动化、智能制造等领域发挥重要作用。

4.2 DeepSeek 应用前景

在数字化转型加速的当下，数据已然成为企业与组织的核心资产，全过程数据资产管理对于释放数据价值、驱动精准业务决策起着决定性作用。DeepSeek-R1 作为先进的人工智能模型，以其卓越的参数规模、强大的推理能力和高效的学习机制，在全过程数据资产管理中展现出极具潜力的应用前景。

1、数据收集阶段：智能精准采集与筛选

数据收集是全过程数据资产管理的基础环节，DeepSeek-R1 凭借强大的自然语言处理能力，能够从海量且复杂的数据源中实现智能精准采集与筛选。面对网络舆情、金融市场交易数据、电商用户行为日志等多源异构数据，DeepSeek-R1 可深入理解数据源的结构和语义内容，依据企业业务目标自动筛选出关键数据。

在金融行业，DeepSeek-R1 能实时监控全球财经新闻、社交媒体上的金融讨论群组以及各国监管政策发布平台等信息源。通过对这些数据的实时分析，精准提取利率变动趋势、行业政策调整以及新兴投资机会等关键信息，为金融机构的投资决策提供时效性强、准确性高的数据支持。在电商领域，DeepSeek-R1 可从数以亿计的用户评论和浏览记录中，挖掘出用户对产品功能、外观、价格等方面的反馈，以及潜在的消费行为模式，助力电商企业精准把握市场需求和用户偏好，优化产品选品和营销策略。

2、数据存储阶段：语义理解与智能分类

传统的数据存储方式多依赖简单的关键词匹配或基础元数据分类，难以挖掘数据的深层价值。DeepSeek-R1 在数据存储阶段可基于深度语义理解，实现精准的语义标注与智能分类。

以医疗行业为例，面对海量的病历数据，DeepSeek-R1 能够深入分析病历中的症状描述、诊断依据、治疗方案以及康复记录等内容，按照疾病种类、严重程度、治疗阶段等多个维度进行精细分类标注。这不仅极大提高了数据存储的效率，减少存储空间浪费，还为后续的数据检索和分析提供了便利，确保医疗数据在临床研究、疾病预测等方面的准确性和可用性，为医疗决策提供有力支持。

3、数据分析阶段：深度关联分析与精准预测

数据分析是挖掘数据价值的核心环节，DeepSeek-R1 强大的推理和学习能力使其在多源异构数据的深度关联分析与精准预测方面表现卓越。

在制造业中，DeepSeek-R1 可以整合生产过程中的设备运行数据、原材料质量数据、环境参数数据以及产品质量检测数据等。通过对这些数据的深度分析，挖掘出影响产品质量的关键因素，如设备的关键运行参数阈值、原材料的特定质量指标对产品性能的影响等。同时，基于历史数据和实时监测数据，精准预测设备故障的发生概率，帮助企业提前制定维护计划，降低生产中断风险，提高生产效率和产品质量。

在市场营销领域，DeepSeek-R1 通过分析用户的人口统计学数据、消费历史数据、社交媒体互动数据以及地理位置数据等，构建 360 度精准用户画像。利用这些画像，DeepSeek-R1 能够准确预测用户的购买行为和偏好变化，为企业制定个性化的营销活动提供科学依据，提高营销效果和投资回报率。

4、数据应用阶段：智能决策支持与服务优化升级

在数据应用阶段，DeepSeek-R1 为企业提供全方位的智能决策支持和服务优化升级方案。

在企业战略决策层面，DeepSeek-R1 基于对市场动态、竞争对手态势以及自身业务数据的综合分析，为企业提供战略规划建议。例如，通过对行业趋势、技术创新方向以及市场竞争格局的分析，帮助企业确定业务拓展方向、新产品研发策略以及资源配置方案，把握市场机遇，应对竞争挑战。

在客户服务方面，DeepSeek-R1 作为智能客服的核心引擎，根据客户的问题、历史交互记录以及用户画像数据，快速准确地提供个性化解决方案。在通信行业，DeepSeek-R1 可分析用户的通信行为数据，如通话时长、流量使用习惯、套餐偏好等，帮助运营商优化套餐设计，推出更符合用户需求的通信服务套餐，提高用户满意度和忠诚度，进而增加业务收入。

5、数据管理全流程：自动化与智能化深度融合

DeepSeek-R1 能够深度融入数据管理全流程，实现自动化与智能化的深度融合。它可以根据预设的业务规则和数据分析模型，自动完成数据的清洗、转换、整合等繁琐操作，减少人工干预，降低人为错误，提高数据处理的效率和准确性。

同时，DeepSeek-R1 具备实时监测数据质量和安全性的能力。通过对数据的实时分析，及时发现数据异常波动、数据缺失、数据重复等质量问题，以及潜在的数据安全风险，如数据泄露风险、非法访问行为等，并提供相应的预警信息和

解决方案。在数据合规管理方面，DeepSeek-R1 可根据不同行业的法规政策要求，对数据的收集、存储、使用和传输等环节进行合规性检查，确保企业的数据管理活动严格遵守法律法规，避免法律风险。

DeepSeek-R1 在全过程数据资产管理中具有广泛且深入的应用潜力。通过与数据资产管理的各个环节紧密结合，DeepSeek-R1 能够帮助企业充分挖掘数据价值，提升业务运营效率和竞争力，推动企业在数字化时代实现可持续发展。随着技术的不断创新和应用场景的持续拓展，DeepSeek-R1 有望在数据资产管理领域发挥更为关键的引领作用，推动行业智能化变革迈向新的高度。