

# 行业研报

## 人工智能行业系列（五）：预训练模型在AI产业生态中的地位与价值

Industry Research Report: Artificial Intelligence Industry Series (5) - The Status and Value of Pre-training Models in the AI Industry Ecology

業界調査レポート：人工知能業界シリーズ（5）——AI業界のエコロジーにおける事前トレーニングモデルのステータスと価値

报告标签：人工智能、机器学习、迁移学习、预训练模型

主笔人：谢子博

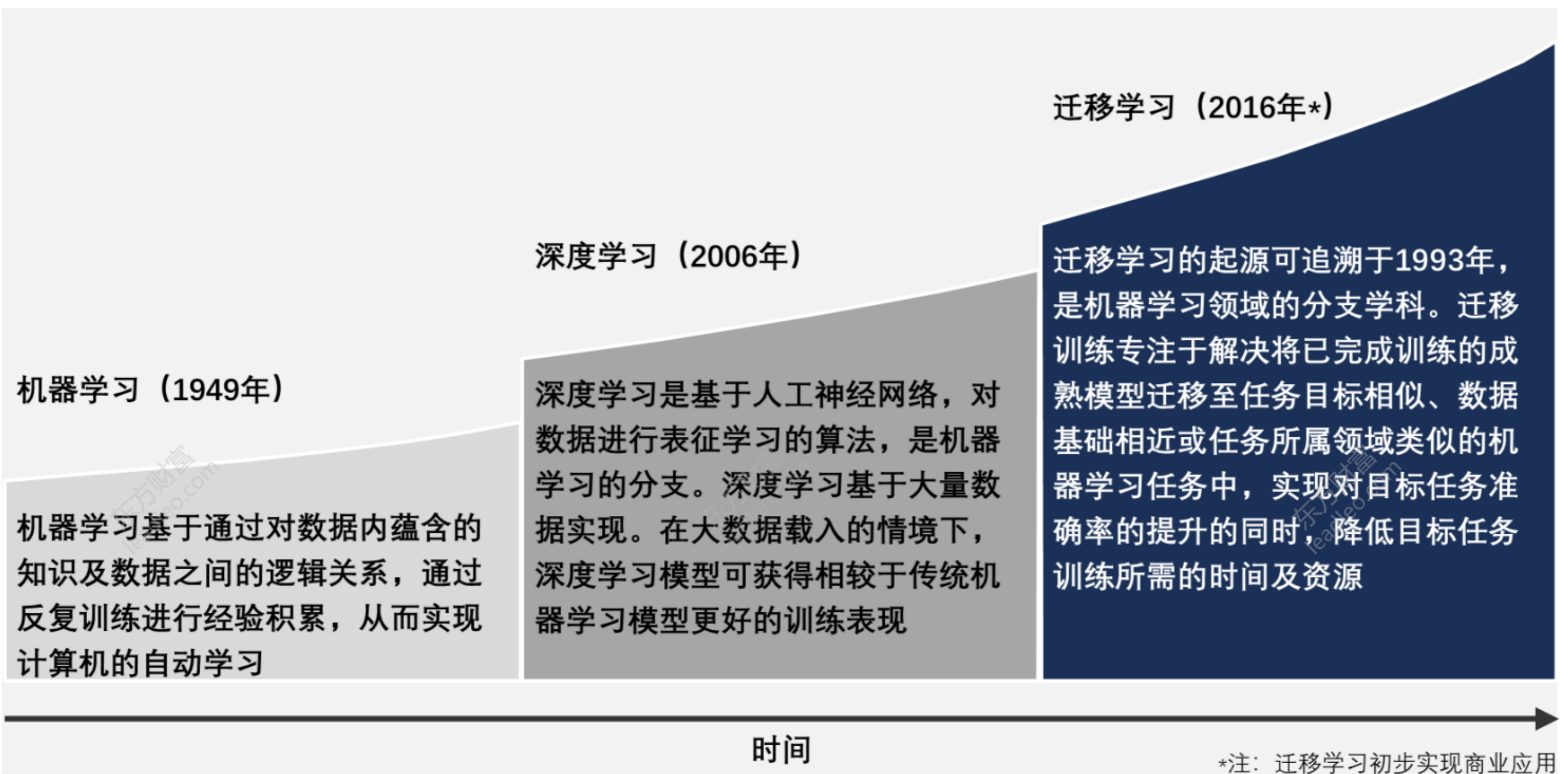
# Q1: 迁移学习的出现, 为人工智能行业的发展带来什么?

## ■ 迁移学习促进工业等垂直应用领域人工智能技术的深化应用

人工智能机器学习的发展可追溯至19世纪中期, 旨在通过对数据内涵的信息与数据之间的逻辑关系的理解实现计算机对数据的自主学习及推演。深度学习是机器学习的分支, 旨在通过更多的数据, 利用人工神经网络模型实现对更为抽象、深层次的逻辑信息进行学习。由于深度学习的实现需基于大量训练数据实现, 深度学习的面世快速推进了AI技术在互联网、金融等数据基础庞大的领域中的应用。然而, 深度学习在推进工业、医疗等领域中AI深化应用的难度较大。阻碍因素主要有二: (1) 该类领域中各细分领域差异化较大, 各项目目标相似度较低, 而基于深度学习的AI开发项目所需花费的时间及资金成本较高, 难以实现规模化的商业应用; (2) 各细分领域可用于训练的数据量较小, 基于深度学习的AI项目准确率难以达到应用所需的标准。

迁移学习致力于将已实现应用的成熟模型的开发经验迁移至新的AI开发项目中, 从而实现在降低训练资源需求的同时, 提升人工智能项目的表现。成本的降低与项目表现的提升使得AI技术工业、医疗等领域的应用难度有效降低。

图表1: 人工智能发展历程



来源: 头豹研究院

## Q2: 迁移学习是否提供了人工智能开发新途径, 改变了人工智能项目开发传统流程?

### ■ 迁移学习打破传统AI零基础开发困境, 助力实现AI技术规模化应用

传统人工智能项目的开发需由数据团队完成数据的采集与处理、由AI开发团队完成从模型搭建、训练、调优等多个步骤。由于多数开发步骤均需要AI开发团队从零开始, 多数工程均具有开发周期较长、开发难度较大等特点。随着人工智能行业的发展, 产业链分化逐渐清晰, 专注于数据准备工作的细分行业析出, 降低AI开发核心团队工作压力。此外, 学术界与产业界对算法模型的迭代与优化, 为AI开发团队提供更多可借鉴的理论知识。然而, 细分行业的析出与知识体系的借鉴难以改变AI开发团队从零开始的开发流程现状, 难以切实解决AI技术深化应用受限的发展困境。

迁移学习的面世为AI开发团队提供了更多的开发路径。通过对成熟模型训练过程中的数据二次分析及筛选, 或是特征体系的保留, 亦或是成熟模型神经网络架构及参数的保留, 迁移学习有效降低AI开发工作的难度, 简化AI开发流程整体的工作流程。迁移学习是推动AI技术实现规模化应用的有效途径。

图表2: 零基础学习开发流程与迁移学习算法开发流程对比



来源: 头豹研究院

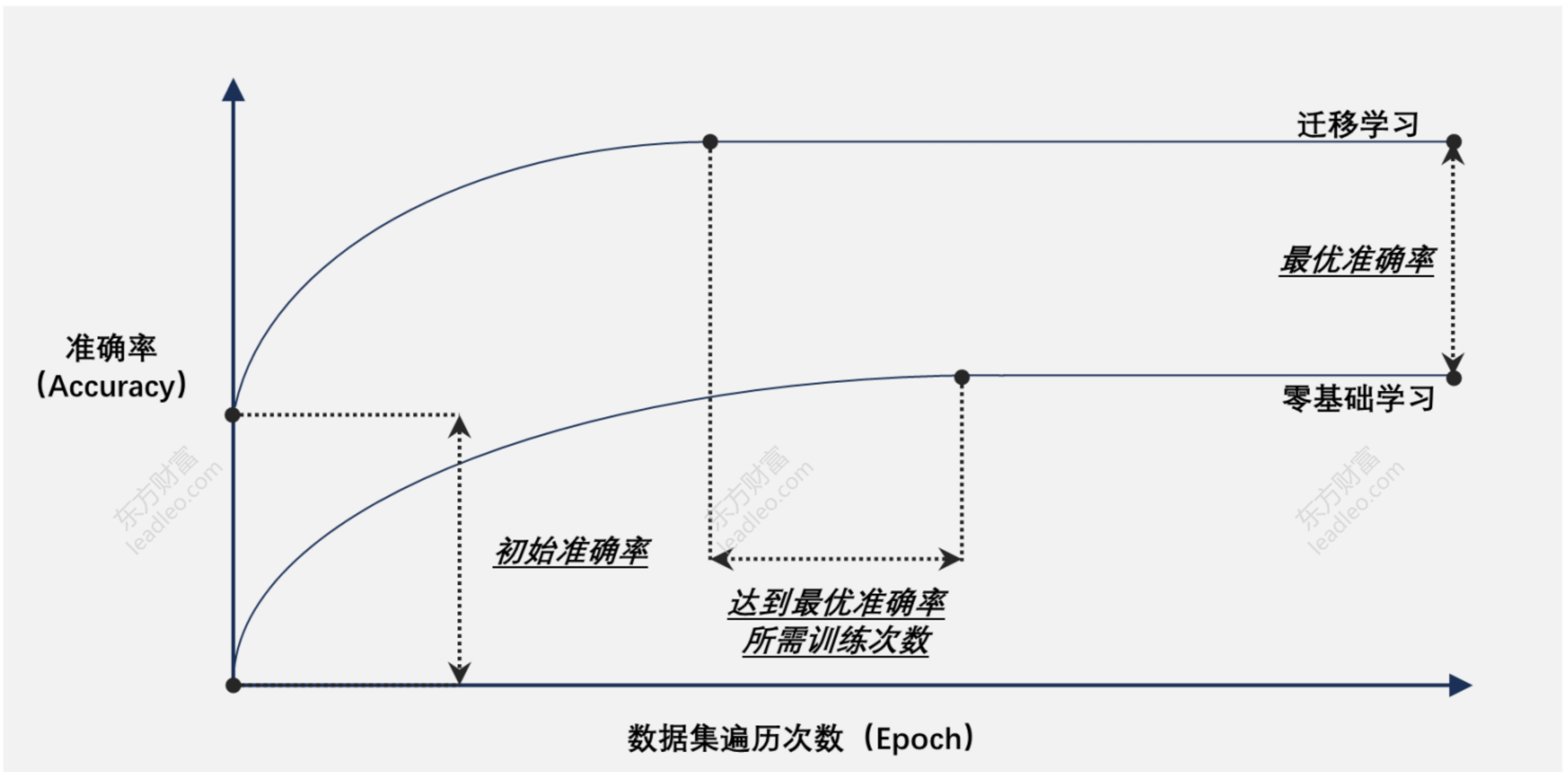
### Q3: 相对于零基础学习, 迁移学习对人工智能项目的开发起到了什么作用?

■ 迁移学习凭借对成熟项目开发经验的借鉴, 实现对开发效率的提升及模型效果的优化

相较于零基础学习, 迁移学习借鉴对成熟AI项目开发经验, 实现对AI开发资源利用效率的优化及AI项目表现的提升。迁移学习对AI开发流程的优化可直观体现在算法模型的训练过程中, 表现在以下两个层面:

- **模型初始准确率:** 相较于从零开始的模型训练方式, 迁移学习可保留成熟模型网络架构及参数, 在模型初始化训练阶段获得更好的模型表现。此外, 模型初始准确率的提升也有效缩短了模型训练过程中达到最优准确率所需的训练周期, 实现对训练资源(硬件算力)利用率的优化;
- **模型最终准确率:** 预训练模型(PTMs)在训练过程中所用的数据量较大, 模型层数更为复杂。在开发经验及应用实例较少的场景下, PTMs的应用对AI开发项目具有显著借鉴意义, 使模型的构建与开发更为成熟。相对于零基础开发, 基于PTMs的迁移学习可实现对模型表现的优化。

图表3: 迁移学习与零基础学习对比



来源: 头豹研究院

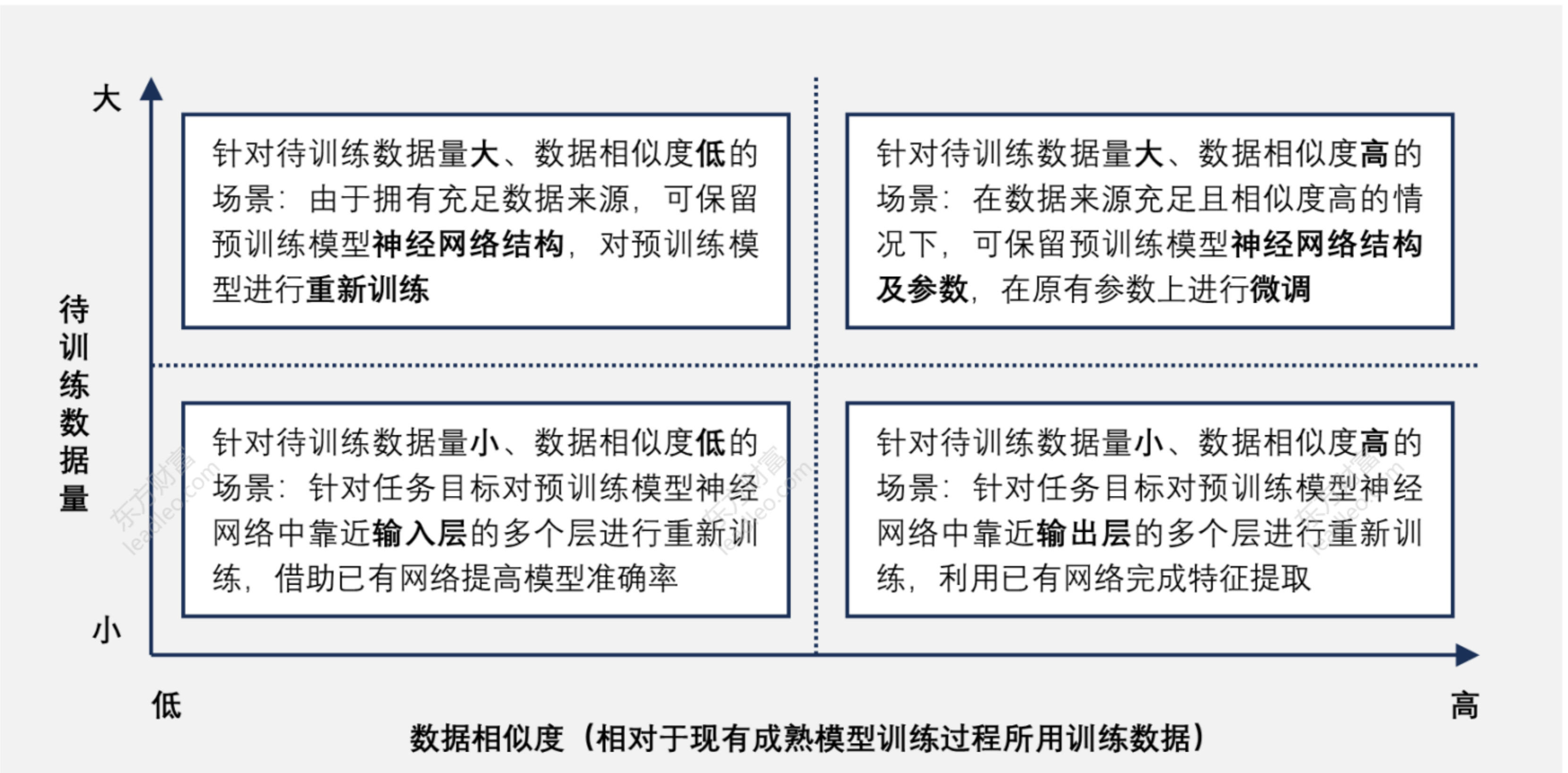
# Q4: 预训练模型 (PTMs) 如何解决差异化数据条件下的任务需求? 迁移学习将在人工智能未来的发展路径上扮演怎样的角色?

## ■ 预训练模型开辟模型开发及训练新途径，推动人工智能规模化应用进程

基于预训练模型的迁移学习可通过对预训练模型不同程度的保留及迁移，实现对差异化场景下的开发需求，减少模型训练所需求的资源及时间。针对新任务可用训练数据量较小的场景，AI开发团队无需重新构建神经网络模型，可通过对预训练模型中部分神经网络层级的保留，并针对任务目标对剩余层级进行适应性修改及重新训练，借助成熟的神经网络架构完成模型的开发与训练。针对新任务可用训练数据量较大的场景，AI开发团队可保留预训练模型的神经网络架构，并针对新任务可用数据与预训练模型所用数据的相似度对模型参数进行取舍，减少模型训练所需要的时间。

迁移学习与预训练模型的出现为AI项目的开发开辟了新的途径，有效降低了针对新兴领域或专业领域的AI项目在项目开发层面的难度及门槛，推动各应用领域智能化转型进程加速落地。从短期来看，迁移学习将是推动工业、医疗等垂直应用领域人工智能技术规模化应用的核心动力。

图表4: 预训练模型迁移训练四大应用途径



来源：头豹研究院

# 方法论

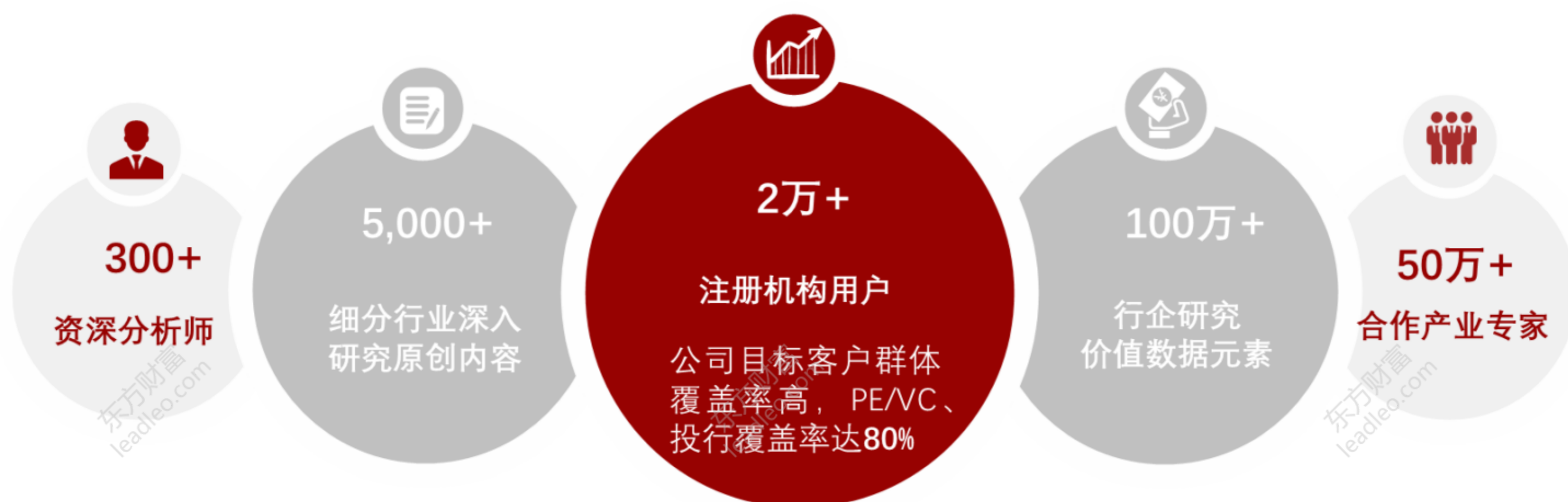
- ◆ 头豹研究院布局中国市场，深入研究10大行业，54个垂直行业的市场变化，已经积累了近50万行业研究样本，完成近10,000多个独立的研究咨询项目。
- ◆ 研究院依托中国活跃的经济环境，从人工智能、迁移学习、预训练模型等领域着手，研究内容覆盖整个行业的发展周期，伴随着行业中企业的创立，发展，扩张，到企业走向上市及上市后的成熟期，研究院的各行业研究员探索和评估行业中多变的产业模式，企业的商业模式和运营模式，以专业的视野解读行业的沿革。
- ◆ 研究院融合传统与新型的研究方法，采用自主研发的算法，结合行业交叉的大数据，以多元化的调研方法，挖掘定量数据背后的逻辑，分析定性内容背后的观点，客观和真实地阐述行业的现状，前瞻性地预测行业未来的发展趋势，在研究院的每一份研究报告中，完整地呈现行业的过去，现在和未来。
- ◆ 研究院密切关注行业发展最新动向，报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入，保持不断更新与优化。
- ◆ 研究院秉承匠心研究，砥砺前行的宗旨，从战略的角度分析行业，从执行的层面阅读行业，为每一个行业的报告阅读者提供值得品鉴的研究报告。

# 法律声明

- ◆ 本报告著作权归头豹所有，未经书面许可，任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得头豹同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“头豹研究院”，且不得对本报告进行任何有悖原意的引用、删节或修改。
- ◆ 本报告分析师具有专业研究能力，保证报告数据均来自合法合规渠道，观点产出及数据分析基于分析师对行业的客观理解，本报告不受任何第三方授意或影响。
- ◆ 本报告所涉及的观点或信息仅供参考，不构成任何证券或基金投资建议。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告或证券研究报告。在法律许可的情况下，头豹可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。
- ◆ 本报告的部分信息来源于公开资料，头豹对该等信息的准确性、完整性或可靠性不做任何保证。本报告所载的资料、意见及推测仅反映头豹于发布本报告当日的判断，过往报告中的描述不应作为日后的表现依据。在不同时期，头豹可发出与本报告所载资料、意见及推测不一致的报告或文章。头豹均不保证本报告所含信息保持在最新状态。同时，头豹对本报告所含信息可在不发出通知的情形下做出修改，读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。

# 头豹研究院简介

- ◆ 头豹是中国领先的原创行企研究内容平台和新型企业服务提供商。围绕“协助企业加速资本价值的挖掘、提升、传播”这一核心目标，头豹打造了一系列产品及解决方案，包括：**报告/数据库服务、行企研报服务、微估值及微尽调自动化产品、财务顾问服务、PR及IR服务**，以及其他企业为基础，利用大数据、区块链和人工智能等技术，围绕产业焦点、热点问题，基于丰富案例和海量数据，通过开放合作的增长咨询服务等
- ◆ 头豹致力于以优质商业资源共享研究平台，汇集各界智慧，推动产业健康、有序、可持续发展



## 四大核心服务

### 研究咨询服务

为企业提供**定制化报告服务、管理咨询、战略调整**等服务

### 行业排名、展会宣传

行业峰会策划、**奖项评选**、行业白皮书等服务

### 企业价值增长服务

为处于不同发展阶段的企业，提供与之推广需求相对应的“**内容+渠道投放**”一站式服务

### 园区规划、产业规划

地方产业规划，**园区企业孵化**服务



# 报告阅读渠道

头豹官网 —— [www.leadleo.com](http://www.leadleo.com) 阅读更多报告

头豹小程序 —— 微信小程序搜索“头豹”、手机扫上方二维码阅读研报



添加右侧头豹分析师微信，身份认证后邀您进入行研报告分享交流微信群



详情咨询



客服电话

400-072-5588



上海

王先生： 13611634866

李女士： 13061967127



深圳

李先生： 18916233114

李女士： 18049912451



南京

杨先生： 13120628075

唐先生： 18014813521

# 头豹 Project Navigator 领航者计划介绍



备注：活动解释权均归头豹所有，活动细则将根据实际情况作出调整。

# 头豹 Project Navigator 领航者计划与商业服务

- 头豹以**研报服务**为切入点，根据企业不同发展阶段的资本价值需求，以**传播服务、FA服务、资源对接、IPO服务、市值管理**为基础，提供适合的**商业管家服务解决方案**



扫描上方二维码  
**联系客服报名加入**

备注：活动解释权均归头豹所有，活动细则将根据实际情况作出调整。

# 读完报告有问题？

## 快，问头豹！你的智能随身专家



扫码二维码  
即刻联系你的智能随身专家

千元预算的  
高效率轻咨询服务

### STEP04 专业高效解答

书面反馈、分析师专访、  
专家专访等多元化反馈方式

### STEP03 解答方案生成

大数据×定制调研  
迅速生成解答方案

### STEP02 云研究院后援

云研究院7×24待命  
随时评估解答方案

### STEP01 智能拆解提问

人工智能NLP技术  
精准拆解用户提问



www.leadleo.com  
400-072-5588