

# 项目1

## 初识人工智能

本项目从短视频智能推荐等生活场景切入,引导大家认识人工智能在生活中的广泛应用,进而深入解析人工智能的定义、发展历程及核心技术。同时,详细介绍 GPT、文心一言和 DeepSeek 等主流大模型,帮助大家建立起对大模型的初步认知,激发学习兴趣。

### 案例导入

#### 短视频应用如何“猜你喜欢”

在数字时代的浪潮中,短视频以其碎片化、高趣味性的特点,迅速占领了大众的娱乐与信息获取阵地。2026年4月15日发布的《中国网络视听发展研究报告(2026)》显示,截至2025年12月,我国网络视听用户规模达10.99亿人,网络视听用户人均单日使用时长为201分钟(3小时21分钟),其中微短剧表现尤为突出,人均单日使用时长达到129分钟。这意味着平均每个用户每天都要在短视频平台上完成数十次甚至上百次的内容浏览动作。在这看似轻松的滑动指尖背后,支撑起庞大用户群体持续沉浸体验的核心技术,正是堪称“隐形大脑”的智能推荐系统。

当你打开任意一款短视频应用 App,无须烦琐的搜索操作,首页呈现的内容便能精准戳中你的兴趣点:喜欢美食的人会刷到各地特色小吃教程,关注科技的用户能看到最新数码产品测评,宝妈群体则常收到育儿经验分享。这种“打开即喜欢”的神奇体验并非偶然,而是人工智能技术与大数据分析深度融合的必然结果。

例如,清晨6:30,大学生小赵的手机闹钟刚响,他习惯性地伸手划开屏幕,点开了常用的短视频 App。首页第一条内容便精准命中了他的晨间习惯:一段59秒的篮球战术解析短视频,来自他关注已久的篮球教学博主。这种精准的时间场景匹配并非巧合,系统通过长期记录发现,小赵连续三个月每天清晨都会花5~10分钟观看篮球相关内容,于是自动将优质篮球内容前置到他的晨间推荐队列中。观看过程中,小赵对视频中提到的“挡拆战术细节”部分反复回看了三次,每次都拖动进度条重新观看关键画面,这个行为被系统实时捕捉并标记为“深度兴趣点”。当视频播放结束时,屏幕立刻弹出了两条关联内容:一条是该博主讲解同类战术的进阶视频;另一条则是 NBA 比赛中运用该战术的经典集锦。这种即时响应机制,让小赵的兴趣探索形成了连贯的体验链条。

通过这一案例,我们不仅能揭开“猜你喜欢”的技术面纱,更能直观感受大模型技术如何将复杂的算法逻辑转化为通俗易懂的用户体验,为后续深入学习人工智能基础与大模型应用奠定实践认知基础。

## 任务 1.1 理解推荐系统背后的智能逻辑

看似简单的滑动体验背后，是一套由服务器集群、算法模型和数据处理中心构成的精密智能系统在 24 小时持续运转。推荐系统之所以能精准捕捉用户喜好，其核心逻辑建立在“数据收集—特征提取—智能匹配—系统迭代”的四层闭环机制之上，如图 1-1 所示。而大模型技术的引入则让这套机制的效率和精度实现了质的飞跃。对于初学者而言，理解这套智能逻辑的构成要素和运转规律，是迈入人工智能世界的重要一步。



图 1-1 推荐系统背后的智能逻辑

### 1.1.1 数据收集：构建用户兴趣的“原材料仓库”

推荐系统运转的第一步，是建立全维度的用户行为数据收集系统，这如同为智能推荐铺设“信息高速公路”。现代短视频 App 的数据采集能力已达到惊人的细致程度，形成了覆盖用户互动全流程的“数据感知网络”。

在显性互动行为层面，系统会精确记录用户的每一次主动操作：点赞行为被标记为“明确正向反馈”，且长按点赞与快速点赞会被区分为不同强度的喜好信号；评论内容会通过自然语言处理技术进行情感分析，带有积极词汇的评论会被赋予更高权重；转发行为被视为“强烈推荐意愿”，系统会据此推断该内容的传播价值；收藏操作则被标记为“长期兴趣点”，相关内容会在后续同类场景中优先推荐。这些显性行为数据构成了用户兴趣分析的基础素材。

更值得关注的是隐性行为信号的捕捉能力。系统通过埋点技术记录用户的每一个细微动作：观看时长是重要的衡量指标，当用户观看某类视频的平均时长超过平台均值 2 倍时，会被判定为“深度兴趣”；滑动速度蕴含丰富信息，每秒滑动超过 1.5 次被视为“快速筛选”，低于 0.5 次则表示“谨慎浏览”；重复观看行为会被重点标记，同一视频观看三次以上通常意味着该内容具有较高价值；甚至连暂停位置都暗藏玄机，视频前 1/3 处暂停多为内容吸引后 1/3 处暂停可能表示兴趣衰减。

设备与环境数据同样是重要的分析维度。系统会记录用户的设备型号（推断消费能力）、网络环境（Wi-Fi 环境下推荐高清长视频，移动网络则优先推送标清短视频）、地理位置（结合 GPS 信息推送本地内容）、使用时段（区分通勤、午休、睡前等场景）等数据。在获得用户授权的情况下，部分 App 还会通过传感器收集更细致的环境信息，如光线强度（夜间自动减少刺眼画面内容）、运动状态（步行时推荐竖屏单手操作内容）等。

这些多维度的数据会通过加密传输技术实时上传至后台服务器，形成动态更新的用户行为日志。行业数据显示，每个活跃用户每天产生的行为数据可达 7000~9000 条，这些数据经过清洗和结构化处理后，存储在分布式数据库中，为后续的兴趣分析提供海量素材。值

得注意的是,正规平台都会采用数据脱敏技术,对用户隐私信息进行加密处理,在保障推荐效果的同时,保护用户数据安全。

### 1.1.2 特征提取:给内容贴上“数字标签”

当海量用户行为数据被收集后,推荐系统的第二步工作便是对内容进行全方位的特征提取,如同为每条短视频贴上密密麻麻的“数字标签”,这是实现精准推荐的关键环节。在大模型技术应用之前,特征提取主要依赖人工标注和简单的机器识别,效率和精度都十分有限;而随着大模型技术的普及,特征提取已进入全自动化、多维度的智能时代。

基础特征层是内容标注的起点。每条短视频首先会被标记基础属性:视频时长被精确到秒(如15秒、60秒、3分钟等不同挡位),发布时间精确到分钟级(用于判断内容新鲜度),创作者信息包括粉丝量、领域标签、历史作品风格等(帮助识别优质创作者),视频格式则包括分辨率、横竖屏比例、是否带字幕等(适配不同设备播放需求)。这些基础特征虽然简单,却是实现精准推荐的基础维度,例如系统会根据用户习惯优先推荐特定时长的视频。

内容特征层构成了特征提取的核心。在视觉识别方面,系统通过计算机视觉技术对画面进行逐帧分析,识别出人物(性别、年龄、职业特征)、场景(室内、户外、城市、乡村)、物品(汽车、美食、数码产品等)、动作(运动、烹饪、教学等)等视觉元素,甚至能识别画面风格(写实、动画、复古等)。在音频处理方面,通过语音转文字技术将视频中的语音内容转化为文本,再通过关键词提取技术识别核心话题;同时对背景音乐进行风格分析(流行、古典、摇滚等)和情感识别(欢快、悲伤、激昂等)。

语义特征层体现了大模型技术的优势。传统特征提取只能识别孤立的关键词,而基于大模型的语义理解技术能把握内容的深层含义:通过自然语言处理技术分析视频文案的主题思想,如“教你三分钟做早餐”的核心主题是“简易早餐教程”;通过逻辑关系识别理解内容结构,如“先准备食材,再开火烹饪”的步骤关系;甚至能识别出内容中的隐性信息,如从“熬夜后如何护肤”中推断出“关注健康养生”的潜在主题。这种深层语义理解能力,让特征提取从“看表面”升级为“懂内涵”。

情感特征层让推荐更具“温度”。系统通过多模态情感分析技术,综合画面色调(暖色调传递温馨感)、语音语调(语速、音量、音调变化)、文字情感倾向(积极、消极、中性词汇占比)等多方面信息,为视频标注情感标签,如“励志正能量”“轻松搞笑”“科普严谨”“温馨治愈”等。对于职业教育相关内容,还能识别出“实操性强”“讲解清晰”“案例丰富”等更细致的评价维度,帮助推荐更符合学习需求的内容。

特征提取的过程是动态迭代的。每条新视频上传后,首先会经过自动化特征提取系统进行初步标注,生成基础标签;随着用户互动数据的积累,系统会根据实际观看反馈优化特征标签,如发现标注为“美食教程”的视频实际观看者多为健身人群,则会补充“低脂食谱”标签;同时定期通过人工审核校准特征提取的准确性,形成“机器自动标注—用户反馈优化—人工审核校准”的闭环机制。据统计,一条热门短视频的特征标签可达数百个,这些标签共同构成了内容的“数字身份证”,为后续的智能匹配提供精准依据。

### 1.1.3 智能匹配：算法如何“牵线搭桥”

当用户行为数据与内容特征标签都已准备就绪，推荐系统便进入最核心的智能匹配环节。如果说数据收集是“原材料采购”，特征提取是“产品加工”，那么智能匹配就是“精准销售”，它如同一个永不疲倦的“智能红娘”，持续计算用户兴趣与内容特征的匹配度，最终为每个用户生成独一无二的推荐列表。

传统推荐算法主要依靠两种经典思路。协同过滤算法如同“朋友推荐”，通过分析用户群体的行为模式，找到与目标用户兴趣相似的“邻居用户群”，将这些相似用户喜欢的内容推荐给目标用户。例如发现喜欢篮球视频的用户中有 80% 也喜欢健身内容，系统就会向新的篮球爱好者推荐健身视频。内容推荐算法则如同“同类推荐”，基于用户历史喜欢的内容特征，推荐具有相似特征的其他内容，例如，用户喜欢过“烤箱美食教程”，就推荐带有“烘焙”“甜品制作”标签的内容。这两种算法各有优势，协同过滤擅长发现潜在兴趣，内容推荐则更稳定可靠。

大模型技术的加入为推荐算法带来了革命性突破，主要体现在三个关键能力上。

一是上下文理解能力。传统算法只能孤立分析单条内容的匹配度，而基于大模型的推荐系统能理解用户兴趣的连贯性和演变规律。例如，当用户先观看“篮球规则入门”，再浏览“篮球基础运球”，系统会推断出“篮球学习进阶”的兴趣路径，进而推荐“篮球投篮技巧”的后续内容，形成连贯的学习链条；当用户从“考研复习规划”转向“公务员备考经验”，系统能识别出“职业规划探索”的潜在主题，避免推荐方向混乱。

二是跨领域关联能力。大模型凭借海量知识储备，能发现看似无关的兴趣点之间的潜在联系。例如通过分析发现“喜欢户外露营”的用户中，有 65% 同时关注“天文摄影”，系统会向露营爱好者推荐星空拍摄教程；识别出“关注职业教育”的用户往往也关心“技能证书考试”，从而实现跨领域内容推荐。这种关联能力打破了传统推荐的“领域壁垒”，既丰富了推荐多样性，又能发现用户自己都未察觉的潜在兴趣。

三是实时学习与动态调整能力。传统推荐算法的更新周期较长，往往需要数小时甚至数天才能根据用户行为变化调整策略，容易导致推荐内容过时。而基于大模型的实时推理技术，能在分钟级甚至秒级时间内响应用户行为变化：当用户连续划过某类内容时，系统立即降低该类内容权重；当用户对新领域内容表现出兴趣时，快速增加相关推荐占比。这种“即时反馈—快速调整”机制，让推荐系统始终保持对用户兴趣的敏锐捕捉。

智能匹配的过程还涉及复杂的推荐策略优化。系统需要在多种目标之间找到平衡：既要保证推荐精准度（用户喜欢的内容占比），又要保持内容多样性（避免信息茧房）；既要推荐热门优质内容，又要扶持新兴创作者；既要满足短期兴趣需求，又要引导长期健康的内容消费。为此，推荐系统通常采用多目标优化算法，为不同目标设置动态权重：例如新用户初期优先保证多样性，老用户则侧重精准度；工作日侧重实用内容，周末则增加娱乐内容占比。

推荐结果的最终呈现还需考虑场景适配因素。系统会根据用户当前的使用场景动态调整推荐策略：在移动网络环境下优先推荐短时长视频，减少用户流量消耗；在夜间模式下推荐亮度较低的视频内容，保护用户视力；在碎片化时间（如通勤）下推荐独立完整的短视频，在整块时间（如周末）则推荐系列化内容。这种“千人千面”且“因时制宜”的智能匹配机

制,最终实现了我们所体验到的“猜你喜欢”的神奇效果。

#### 1.1.4 系统迭代：推荐如何越用越“懂你”

智能推荐系统并非一成不变的固定程序,而是一个持续进化的“学习型系统”,它通过“冷启动—兴趣探索—精准推荐—持续优化”的全生命周期迭代过程,不断提升对用户兴趣的理解精度,这个动态进化过程体现了人工智能“从陌生到熟悉”的学习逻辑。

冷启动阶段是推荐系统与新用户的“初次见面”。当用户首次下载 App 或注册账号时,系统缺乏历史行为数据,这时会通过多种策略实现初步推荐:基于用户注册时填写的兴趣标签(如“美食”“科技”)推送相关内容;根据设备信息(如学生群体常用中端机型)推断用户属性;参考同地区、同年龄段用户的普遍兴趣;同时保证内容的多样性,涵盖热门推荐、不同领域精选内容等,通过广泛撒网收集初始反馈数据。这个阶段的推荐精准度通常在 30%~40%,但为后续兴趣探索奠定了基础。

兴趣探索阶段是系统“认识用户”的关键期。在用户使用的前 2~4 周,系统会采用“试探—反馈—调整”的策略:首先推送覆盖多个领域的内容,记录用户对不同领域的互动数据(如观看时长、互动率);然后根据初步反馈增加兴趣领域内容占比,同时保留一定比例的新领域内容进行试探;通过多轮迭代逐渐聚焦核心兴趣,同时发现潜在兴趣点。例如发现用户对“科技”领域的互动率最高,进一步细分是“数码产品”还是“科学科普”;在探索过程中发现用户对“科技+生活”的跨界内容特别感兴趣,从而优化推荐方向。这个阶段的精准度通常能提升至 50%~60%。

精准推荐阶段体现系统的成熟度。当用户使用超过一个月,积累了足够的行为数据后,系统已构建起相对完善的用户画像,进入精准推荐阶段:基于稳定的兴趣标签推荐高匹配度内容;通过兴趣权重动态调整不同领域的内容占比,如核心兴趣占 60%,次要兴趣占 30%;同时建立“兴趣衰减模型”,对长期未互动的兴趣领域逐渐降低权重,避免推荐过时内容。这个阶段的推荐精准度可提升至 70%以上,用户会明显感受到“系统越来越懂我”。

持续优化机制保证系统“与时俱进”。推荐系统通过 A/B 测试不断优化算法:将用户随机分为两组,对实验组推送新算法生成的推荐列表,对照组使用原有算法,通过对比两组的互动数据(点击率、完播率、留存率等)评估新算法效果;定期分析用户反馈,如“不感兴趣”举报内容的特征,针对性优化推荐策略;跟踪外部热点事件(如体育赛事、节日活动)及时调整推荐方向;同时借鉴行业前沿技术,如引入更先进的大模型提升语义理解能力。这种持续优化机制让系统始终保持进化活力。

值得注意的是,推荐系统在迭代过程中会特别警惕“信息茧房”问题——过度聚焦单一兴趣导致用户视野受限。为此,系统会通过多种策略保持推荐多样性:强制保留 10%~20%的探索性内容,来自用户较少接触的领域;定期推送跨领域关联内容,拓展兴趣边界;设置“发现页”专门推荐新兴领域优质内容;基于用户的“偶尔兴趣”适当拓展相关内容。这种“精准匹配为主,适度探索为辅”的平衡策略,让推荐系统既能满足用户核心需求,又能带来意外发现的惊喜。

从冷启动的陌生试探到精准推荐的默契配合,短视频推荐系统的迭代过程生动展现了人工智能“学习—进化”的本质特征。这个过程中,数据是“养料”,算法是“大脑”,用户反馈是“方向”,三者共同作用让推荐系统不断成长,最终实现了我们所体验到的“越用越懂你”的

智能服务效果。通过理解这个动态迭代过程,我们能更清晰地把握人工智能技术“从数据到智能”的转化逻辑,为深入学习人工智能基础与大模型应用打开认知窗口。

## 任务 1.2 理解人工智能的概念

通过短视频推荐系统的案例,我们已经直观感受到了人工智能技术的应用价值。但要真正理解这项技术,还需要从基础概念入手,建立对人工智能的系统认知。本节将带你走进人工智能的世界,了解其定义内涵、发展脉络、核心分支及大模型在其中的重要地位,为后续学习提供理论框架。

### 1.2.1 人工智能的定义与发展历程

人工智能(artificial intelligence, AI)是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。简单来说,它是让机器具备类似人类的感知、思考、学习和决策能力的技术集合。其核心目标包括模拟人类的感知能力(如视觉、听觉识别)、思考能力(如逻辑推理、问题解决)、学习能力(如从经验中改进性能)和行为能力(如与环境交互)。

人工智能的发展历程犹如一部技术进化史,可大致分为四个关键阶段,如图 1-2 所示。



图 1-2 人工智能发展的四个阶段

第一阶段是萌芽期(20 世纪 50—70 年代),1956 年,达特茅斯会议首次提出“人工智能”概念,标志着该学科正式诞生。这一时期的研究聚焦于逻辑推理和问题求解,如艾伦·纽厄尔和赫伯特·西蒙开发的“逻辑理论家”程序能证明数学定理,但受限于计算能力,早期系统仅能处理简单任务,20 世纪 60 年代末遭遇第一次“AI 寒冬”。

第二阶段是专家系统时代(20 世纪 80 年代),随着计算机性能提升,人工智能进入实用化探索期。专家系统通过将领域专家知识编码成规则库,在医疗诊断(如 MYCIN 系统)、化学分析(如 DENDRAL 系统)等领域取得成功。但由于开发成本高、适应性差,20 世纪 80 年代末因投资撤离陷入第二次“AI 寒冬”,人们意识到仅靠手工编码知识难以实现通用智能。

第三阶段是机器学习崛起期(20 世纪 90 年代—2010 年),算法与数据的结合催生新突破。统计学习方法逐渐主导研究,支持向量机、决策树等算法在分类、回归任务中表现优异。2001 年,IBM“沃森”击败人类冠军。2009 年,谷歌自动驾驶汽车上路测试,标志着 AI 开始从实验室走向实际应用。这一时期的里程碑是 2012 年亚历克斯·克里泽夫斯基团队用深度卷积神经网络在 ImageNet 竞赛中大幅超越传统方法事件,至此,深度学习时代正式开启。



微课:专家系统

第四阶段是大模型爆发期(2010年至今),以2017年Transformer架构提出为标志,人工智能进入规模化发展新阶段。2020年,GPT-3以1750亿参数规模展现出强大的通用能力。2022年,ChatGPT引发全球AI热潮。2023年,多模态大模型实现文本、图像、音频的统一理解。2025年,国产化DeepSeek模型的出现,点燃了大模型开源的浪潮,促进了人工智能大模型技术在日常生活中的广泛应用。如今,人工智能已渗透到医疗、教育、制造等千行百业,从实验室技术演变为推动社会变革的核心力量。

### 1.2.2 人工智能的主要特征

人工智能的主要特征包括可学习性、可自主性、可感知性、高效性、自适应性、纠错性和协同性,如图1-3所示。

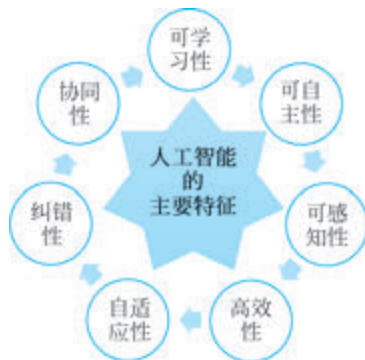


图 1-3 人工智能的主要特征

#### 1. 可学习性

人工智能的核心特点之一是强调学习和演化。这意味着 AI 不仅可以从数据中自主地创建算法或模型,还能预测或识别未来的数据。这种学习能力使机器能够从多方面充分地学习,而不需要人类直接告诉它如何操作。例如,机器学习技术可以训练自学习算法,并应用于新的数据集,从而提高预测精度和准确性。此外,基于学习能力的人工智能技术还可以实现人类的自主学习和进化,进一步促进 AI 的可靠性和实用性。

#### 2. 可自主性

人工智能的自主性体现在其能够独立地进行学习、推理和决策,而不需要人类的干预或指示。这种自主性是通过机器的认知、分析和选择最佳决策方案来实现的,跨越了从人类思维到机器思维的界限。为了实现这一功能,机器必须首先具备感知、推理和学习能力,以便在执行决策过程中进行复杂的数据分析和治理。例如,人工智能系统可以利用监督学习方法预测股票或房价,并在遇到新信息或数据时自主调整算法或模型,以更好地适应其数据和环境。

#### 3. 可感知性

人工智能的动态感知能力是其核心特点之一。这种能力允许 AI 系统通过传感器等设备对外界环境(包括人类)进行感知。与人类相似,AI 可以通过听觉、视觉、嗅觉和触觉接收环境中的信息。这些感知到的信息可以转化为文字、语音、表情和动作,从而与人类进行交互和互动。这种交互方式使机器越来越“理解”人类,与人类形成优势互补的合作关系。

#### 4. 高效性

人工智能可以在短时间内完成大量工作。例如,在金融领域,AI 能快速分析市场数据,为投资者提供明智的决策建议。此外,AI 还能自动化执行复杂和烦琐的任务,如制造业中的装配和检测,从而大大提高工作效率。在医疗领域,AI 能迅速、准确地诊断疾病,极大地减少了医生的诊断时间。

## 5. 自适应性

自适应性是人工智能系统的一大特点,意味着它们能够根据不同的环境和任务进行调整,以适应各种情况。这种适应性不仅表现在对环境的感知上,还体现在系统能够自我调整行为以更好地完成任务。例如,在智能家居中,AI可以根据家庭成员的需求自动调整温度和光线,为家庭创造一个更舒适的环境。这种自适应性是基于对环境的深入理解和自我调整的能力。

## 6. 纠错性

人工智能模型基于先前收集的信息,通过特定的算法进行决策,从而大大降低了误差。相比之下,人类在执行任务时,由于各种原因如注意力不集中,容易出错。例如,尽管在线超市 Ocado 使用自动化机器控制数千个机器人以协调商品物流和配送,但该系统仍不是完全无故障的。因此,人工智能不仅可以减少人为失误,还能提高工作效率和准确性。

## 7. 协同性

人工智能的协同性体现在人机协同和各子系统间的协作。人机协同意味着机器和人分别从事自己更擅长的领域,如智能机器擅长枯燥、重复的工作,而人类则胜任创新性工作及人文关怀性工作。这种协同方式确保了工作效率和效果的最大化。同时,多领域专家系统可以协作求解单领域或单个专家系统无法解决的问题,从而提高求解能力和扩大应用领域。这种跨领域的协同工作方式为解决复杂问题提供了新的途径。

### 1.2.3 人工智能的三大学派

人工智能涉及计算机科学、心理学、哲学和语言学等多个学科。可以说几乎是自然科学和社会科学的所有学科,其范围已远远超出了计算机科学的范畴,人工智能与思维科学的关系是实践和理论的关系,人工智能处于思维科学的技术应用层次,是它的一个应用分支。

从思维观点看,人工智能不仅限于逻辑思维,要考虑形象思维、灵感思维才能促进人工智能的突破性发展。数学常被认为是多种学科的基础科学,数学也进入语言、思维领域,人工智能学科也必须借用数学工具,数学不仅在标准逻辑、模糊数学等范围发挥作用,数学进入人工智能学科,它们将互相促进而更快发展。

在人工智能的发展过程中,涌现了从不同的学科背景出发的三大学派,即符号主义学派、连接主义学派和行为主义学派。

#### 1. 符号主义学派

符号主义学派又称符号学派、逻辑学派、心理学派或计算机学派。该学派认为人类认知和思维的基本单元是符号,认知过程就是在符号表示上的运算。符号主义的核心在于用计算机的符号操作来模拟人类的认知过程,从而实现人工智能。符号主义在模糊推理、专家系统、知识图谱构建等领域有广泛应用。在人工智能早期发展中占据主导地位。

该学派的代表人物有西蒙、纽厄尔、明斯基等。该学派认为人工智能源于数学逻辑,其实是模拟人的抽象逻辑思维,用符号描述人类的认知过程,包括知识表示、决策树算法等。

20 世纪 70 年代出现了具备专业知识和逻辑推断能力的专家系统,推动了人工智能的工程应用。但是,高性能个人计算机的普及应用以及专家系统成本的居高不下,使符号主义学派在人工智能领域主导地位逐渐被连接主义学派取代。

## 2. 连接主义学派

连接主义学派又称仿生学派、生理学派、连接学派,包含感知器、人工神经网络、深度学习等技术,当前占据主导地位。连接主义者认为,人类思维的本质在于神经元之间的连接和交互,因此他们试图建立基于神经元模型的学习算法来模拟人类的学习过程。连接主义的代表性算法是深度学习算法,广泛应用于图像识别、语音识别和自然语言处理等领域。

该学派的代表人物有罗森布拉特等,他们认为人工智能源于仿生学,应以工程技术手段模拟人脑神经系统的结构和功能。早在 1943 年,美国心理学家麦卡洛克和数学家皮茨就提出了利用神经网络对信息进行处理的数学模型——MP 模型,自此人类开启了对神经网络的研究。1982 年霍普菲尔德神经网络模型和 1986 年 BP 神经网络模型的提出,使神经网络的理论研究取得重大突破。2006 年,连接主义的领军学者辛顿教授提出深度学习算法,大大提高了神经网络的学习训练能力。

## 3. 行为主义学派

行为主义学派又称进化主义或控制论学派,包含控制论、马尔可夫决策过程、强化学习等技术。行为学派认为智能取决于感知与行为,以及对外界环境的自适应能力。行为主义的代表性应用包括遗传算法、增强学习等,用于教 AI 打游戏、捉迷藏等。

近年来,随着 AlphaGo 取得的重大突破,人工智能吸引了广泛关注。该学派认为人工智能源于控制论,智能行为产生的基础是“从感知到行动”的反应机制,智能行为是在与外界环境交互作用中表现出来的。行为学派代表人物有萨顿等,其代表观点是智能体通过与环境进行交互获得智能,如感知器等。

在人工智能发展历程中,符号主义学派、连接主义学派和行为主义学派不仅先后在各自领域取得了成果,也逐步走向了相互借鉴和融合发展的道路。最近十年,连接主义学派得到快速发展,并在现实生活中得到广泛应用。表 1-1 总结了三大学派的主要技术关注点。

表 1-1 三大学派的主要技术关注点

类 型	符号主义学派	连接主义学派	行为主义学派
其他名称	逻辑学派、心理学派、计算机学派	仿生学派、生理学派	进化主义、控制论学派
研究方向	抽象思维:用公理和逻辑体系构建人工智能体系	形象思维:利用神经网络的连接机制实施学习	感知思维:模拟人的行为实施智能控制
主要技术	模糊推理、专家系统、知识图谱	机器学习、深度学习、AI 大模型	遗传算法、机器人技术、增强学习
代表人物	西蒙、纽厄尔、明斯基	霍普菲尔德、辛顿	萨顿

### 1.2.4 人工智能的演进:从弱 AI 到强 AI

人工智能是一门交叉学科,包含多个核心分支,这些分支从不同维度模拟人类智能,共

同推动技术从“弱 AI”向“强 AI”演进。人工智能按照其智能程度可以分为弱人工智能、强人工智能和超人工智能三个层次,如图 1-4 所示。



图 1-4 人工智能三个层次

(1) 弱人工智能(narrow AI)是指擅长解决特定领域问题的人工智能。例如,能战胜国际象棋世界冠军的人工智能 AlphaGo,它只会下国际象棋,如果问它怎样更好地在硬盘上存储数据,它就无法回答。

(2) 强人工智能(general AI)也称通用人工智能,是指在任何领域都能够胜任人类所有工作的人工智能。它能够进行思考、计划、解决问题、抽象思维、理解复杂理念、快速学习和从经验中学习等操作,并且和人类一样得心应手。图 1-5 是机器人参与踢足球的一个示例。



图 1-5 机器人踢足球

(3) 超人工智能(super AI)是一种超越人类的智能。牛津哲学家、知名人工智能思想家尼克·博斯特罗姆把超人工智能定义为“在几乎所有领域都比最聪明的人类大脑都聪明很多,包括科学创新、通讯和社交技能”。

弱人工智能是指专注于特定领域的人工智能系统,只能在设计范围内完成任务,是当前技术的主流形态;而强人工智能是指具备与人类相当的通用智能,能理解、学习任何人类可完成的智力任务,目前处于理论研究及实践应用阶段。而超人工智能还处于理论研究阶段。

机器学习是人工智能的核心分支,让计算机能从数据中自动学习规律。它分为监督学习(从标注数据中学习输入输出映射关系,如垃圾邮件分类)、无监督学习(从无标注数据中发现隐藏结构,如用户分群)、强化学习(通过与环境交互试错学习最优策略,如 AlphaGo 下棋)三大类。深度学习作为机器学习的子集,基于深度神经网络模拟人脑神经元连接结构,在图像识别、自然语言处理等领域取得突破性进展,是大模型技术的基础。