



华院计算
www.UniDT.com



中国信通院

认知智能技术与应用研究报告 (2023年)



华院计算技术(上海)股份有限公司
中国信息通信研究院云计算与大数据研究所
2023年7月



版权声明

本报告版权属于华院计算技术（上海）股份有限公司和中国信息通信研究院云计算与大数据研究所，并受法律保护。转载、摘抄或利用其他方式使用本报告文字或观点的，应注明“来源：华院计算技术（上海）股份有限公司和中国信息通信研究院云计算与大数据研究所”。违反上述声明者，公司和研究院保留追究其法律责任的权利。

前言

随着深度学习技术演进，语音识别、图像识别等技术准确率得到大幅提高，机器的听、看等“感知”能力已经达到甚至超越人类水平，逐渐进入应用成熟期。然而，当机器在面对思考、推荐、决策、联想等复杂“认知”问题时，仍无法完全满足应用要求。伴随全球数字化进程加快，人工智能技术与产业融合程度逐步加深，“能思考，会判断”的认知智能技术逐步应用于智能客服、智能推荐、智能营销、智能分析等诸多场景，不断释放产业应用价值。目前，认知智能正步入发展快车道，大模型应用热潮又为其注入一股新能量，进一步扩张认知智能应用潜力和价值边界。

本报告由华院计算技术(上海)股份有限公司(以下简称“华院”)和中国信息通信研究院云计算与大数据研究所联合撰写，围绕认知智能发展现状、关键技术、典型应用场景和未来发展方向进行阐述，共分为四个部分。**第一部分**聚焦认知智能发展现状，围绕学术环境、技术创新、产业规模等维度进行阐述；**第二部分**聚焦认知智能关键技术，介绍知识数据融合、知识表示推理、认知规划与决策、多模态学习及情感计算五个技术方向；**第三部分**围绕医疗、金融、制造及教育等重点应用领域，分析典型落地场景及应用案例；**第四部分**围绕大模型、生成式人工智能对认知智能的推动作用进行分析，展望可解释、类脑智能等前沿技术发展前景。

由于人工智能仍处于快速发展阶段，我们对认知智能还有待进一步研究，报告如存在不足之处，敬请大家批评指正。

目录

第一章 人工智能迈入“认知智能”新阶段 1

- (一) AI核心要素推动认知智能发展 1
- (二) 认知智能学术关注度迅速提升 2
- (三) 认知智能应用市场前景广阔 4

第二章 认知智能核心技术日益成熟 7

- (一) 多模态学习有效提高信息认知维度 7
- (二) 知识数据融合加强认知基础 8
- (三) 知识表示与推理可优化认知逻辑 9
- (四) 认知规划和决策处理复杂任务 10
- (五) 情感计算优化智能交互体验 13
- (六) 生成式人工智能提高机器创造力 14

第三章 认知智能场景化应用逐渐深化 16

- (一) 认知智能促进医疗服务再升级 16
- (二) 认知智能促进金融科技应用创新 17
- (三) 认知智能加速工业制造智能转型 19
- (四) 认知智能推进教育模式创新优化 20

第四章 认知智能技术与应用方向展望 22

- (一) 多模态大模型拓展认知智能应用边界 22
- (二) 可解释的认知智能增强技术可信度 23
- (三) 类脑智能推动认知智能向意识智能发展 24

参考文献 25

第一章 人工智能迈入“认知智能”新阶段

认知智能的核心是通过模仿人类认知过程，使机器具备理解、推理、解释、联想、演绎等多方面能力，涉及自然语言处理、知识计算、情感计算、类脑智能等多领域技术，可应用于智能推荐、智能营销、风险预测、智能交互等众多场景。近年来，人工智能数据、算法和算力的快速发展，支撑了认知智能技术创新和场景化应用，越来越多研究机构和学者开始关注认知智能理论研究和技术落地，认知智能产业应用前景广阔。

（一）AI核心要素推动认知智能发展

数据要素方面，数据量飞速增长，数据质量持续优化，为认知智能思考和决策提供依据，扩展认知智能应用范围。根据 Statista 测算，到 2035 年，全球数据产生量将达到 2141ZB¹，为认知智能提供充足数据资源储备。此外，多地出台相关政策措施，通过鼓励高质量基础训练数据集开放共享、建设数据训练基地、搭建数据集精细化标注众包服务平台等举措，保障训练数据的高质量供给。优质数据与知识的双重驱动，进一步提升认知能力上限。

算法要素方面，机器学习、深度学习、强化学习等算法创新迭代，推动认知智能在推理及预测方面性能持续提升。机器学习领域，**对比学习**通过比较输入样本和目标样本之间的差异来进行自监督式学习，

¹ ZB：中文名是泽字节，外文名是 Zettabyte，是计算机术语，代表的是十万亿亿字节。

能够发现数据中的潜在结构和特征，从而更好地进行分类、识别或生成等任务；深度学习领域，**元学习**通过从多个任务的学习中汲取经验，进而提高模型在新任务上的泛化能；强化学习领域，**人类反馈强化学习**将实体识别任务转化为强化学习问题，将人类的反馈作为奖励信号，在与人类交互不断学习优化模型，从而提高实体识别的准确率和鲁棒性，推动认知智能向更高水平的理解和预测能力方向发展。此外，由 ChatGPT 引发的大模型热潮为认知智能发展注入新活力，各类算法与大模型结合运用，将在提升认知理解与推理能力中发挥重要作用，进一步扩大认知智能应用范畴。

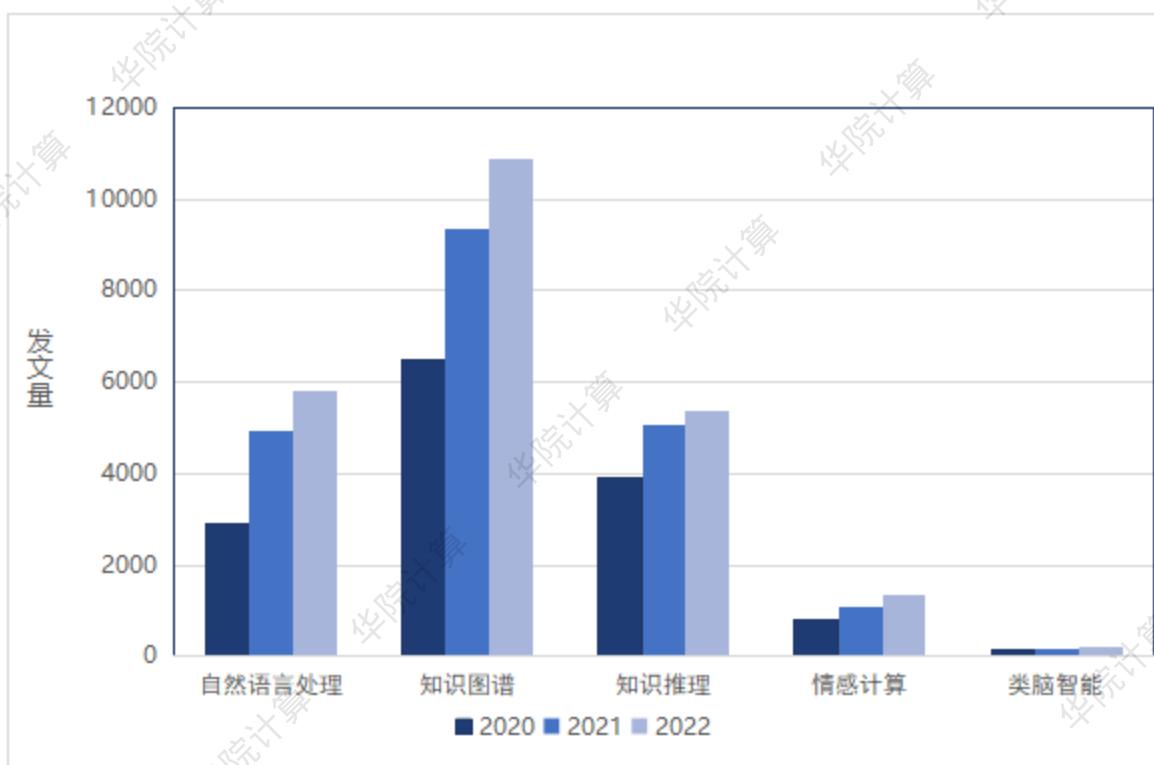
算力要素方面，算力规模持续扩大和计算环境不断创新，为认知智能算法研究及产业落地提供原动力，支撑认知智能实现产业应用。据 IDC 与浪潮信息联合发布《2022-2023 中国人工智能计算力发展评估报告》统计，我国智能算力继续保持快速增长，2023 年智能算力规模预计达到 427 百亿亿次/秒(FLOPS)，较 2022 年提升 59.3%，可以更好支持认知智能在复杂场景的计算速度要求。此外，动态、异构的智能计算环境及服务，以及分布式智能计算（云、边、端融合）²，可在一定程度上响应动态感知、场景应用中数据变化，实现模型快速调整，从而提高认知精度和优化容错能力。

（二）认知智能学术关注度迅速提升

近年来，学术界越来越重视认知智能相关技术理论和实践应用研究，涌现了一批以认知智能专题的学术期刊及国际会议，主要涵盖认

²《人工智能发展对算力及其环境的需求》

知计算、人类工程学、类脑智能等多个领域，旨在加强认知智能领域探讨，对认知智能前沿技术研究具有重要作用。依据 ProQuest 公开数据显示，在学术期刊和国际会议中，认知智能相关论文逐年增长。从自然语言处理、知识图谱、知识推理、情感计算及类脑智能等认知智能维度进行统计，近三年内（2020—2022 年）总发文量近 6 万篇，年发文量逐年提高。



数据来源：ProQuest 公开数据

图 1 2020—2022 年认知智能领域论文数量统计

高校、科研机构和企业积极建设认知智能相关的科研平台和团队，联合多方开展研究。一方面，国内外各大高校和研究机构围绕认知智能、认知计算、类脑智能等方向组建科研项目组及实验室，如斯坦福大学 HAI 研究院组建神经和认知科学项目组、麻省理工学院成立大脑和认知科学学院、清华大学搭建脑与智能实验室、北京大学智能学院

创建认知推理研究组。另一方面，产学研紧密合作推动科研成果高效转化，如中国科学技术大学与科大讯飞联合建立认知智能全国重点实验室，以认知智能技术的广泛和共性需求为牵引，重点开展语义计算、知识建模等认知智能基础理论及技术研究。此外，部分重点行业企业开始加强认知智能应用落地，人民网成立传播内容认知全国重点实验室，专注研究建设可计算、可感知、可认知、可运用的内容传播服务。

（三）认知智能应用市场前景广阔

得益于认知智能应用广度和深度持续增加，认知智能市场规模不断扩增。从全球范围来看，认知智能市场潜力巨大，未来市场规模增长势头迅猛。依据 Precedence Research 研究机构报告显示，按照自然语言处理、机器学习、自动推理、信息检索等技术分类，在医疗、金融、IT、零售等多个领域进行统计，2022 年全球认知智能市场规模达到 326 亿美元，未来 10 年将以 22% 的复合年增长率保持增长，预计 2032 年市场总规模将达到 2381 亿美元。从国内范围来看，国内市场增速高于全球，年增速稳步提升。根据爱分析发布的数据显示，2023 年中国认知智能市场规模为 418.4 亿元人民币，较 2022 年增长 48.5%，预计未来两年，将分别以 50.7%、51.1% 的年增长率实现规模快速扩增。

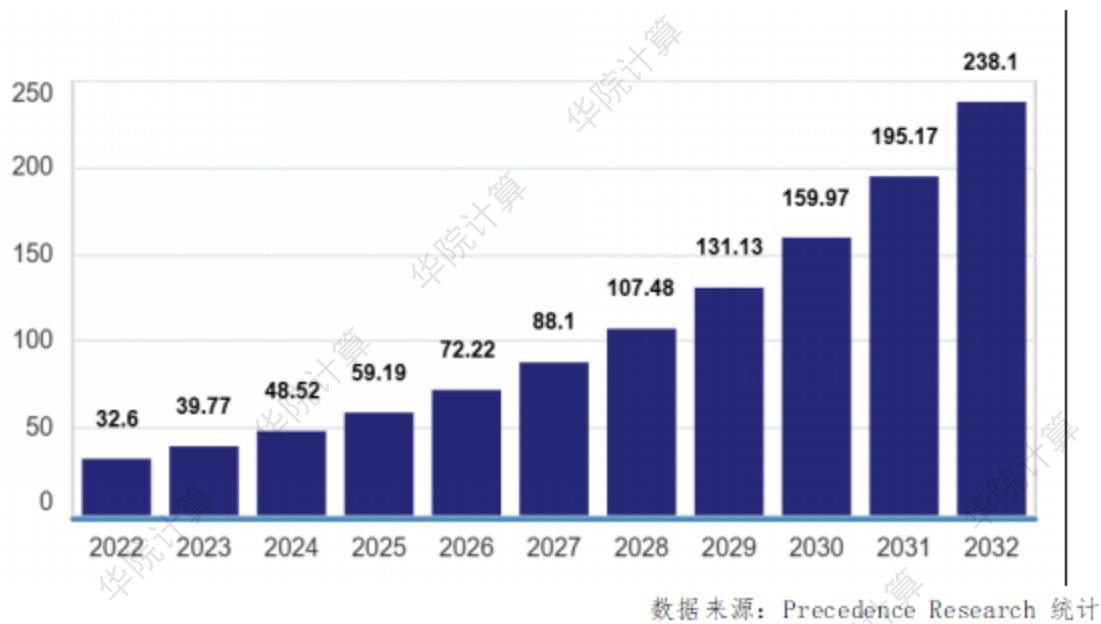
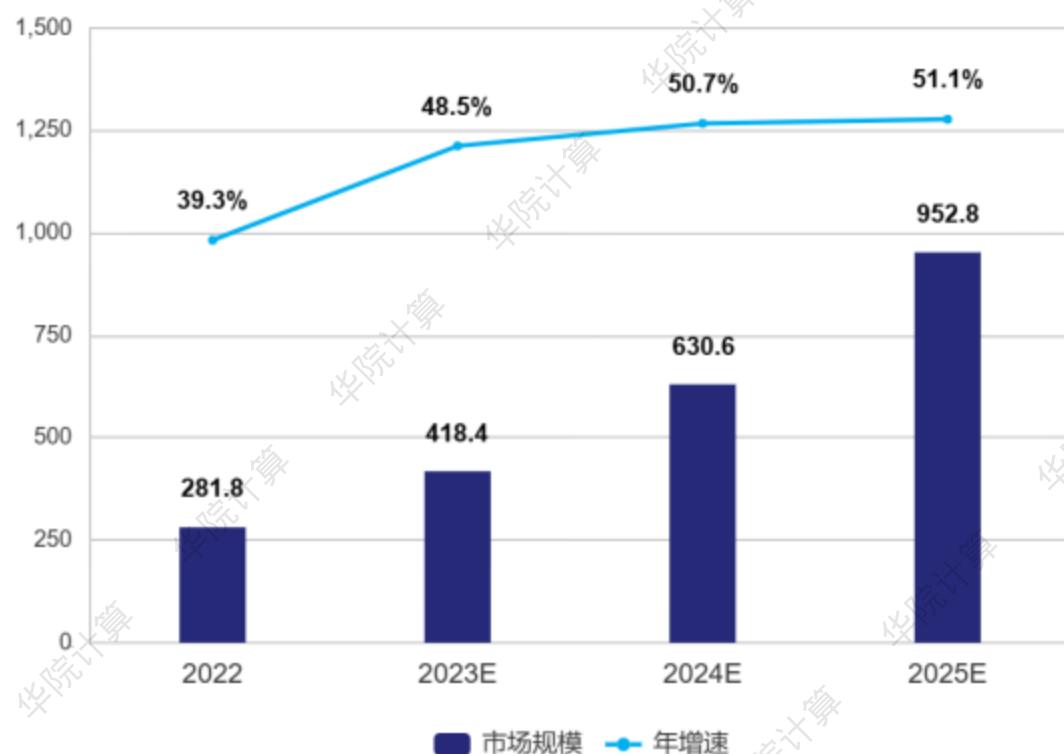


图 2 2022—2032 年全球认知智能市场规模（单位：十亿美元）



数据来源：《2023 爱分析·认知智能厂商全景报告》

图 3 2022—2025 年中国认知智能市场规模（单位：亿元人民币）

国外企业积极布局认知智能产品服务和解决方案，与场景化需求深度融合。Microsoft 以 Azure 认知服务平台为行业开发者用户提供服务，推出具备智能对话、智能推荐、操作执行等能力的全新智能助

手 Copilot 助力用户提高生产力；Google 组建 Google DeepMind 研究团队，将致力于基于深度学习的认知智能技术研究与产品开发；SparkCognition 推出的基于认知的风险预测系统，在能源、航空航天、国防和安全等领域实现落地；CognitiveScale 提供的智能决策平台 Cortex AI，以可视化方式提供模型和业务操作，用户可自行参与个性化决策方案制定。

国内企业进一步探索产品落地路径，在服务方式及应用场景方面寻求突破。一方面，头部企业积极探索以大模型为底座的应用新范式。百度搭建知识中台为行业提供智能搜索、智能推荐等解决方案，以文心大模型为底座，推动能力升级和行业赋能；科大讯飞推出认知大模型，以通用大模型能力为底座赋能教育、医疗等行业。另一方面，腰部企业优化自身技术与产品服务能力，进一步探索认知智能应用方向。华院计算构建底层智能引擎，为社会治理、工业、金融、零售等行业提供智能化产品和服务，在智能配煤、智能推荐等方向形成独有优势；竹间智能的基于认知智能的知识工程平台 Gemini，具备认知搜索、知识推理、辅助决策等多项能力；小 i 机器人将认知智能用于元宇宙交互场景，构建认知推理体系，优化人机交互体验，推进现实与虚拟结合。

第二章 认知智能核心技术日益成熟

面对应用场景中理解、推荐、预测、交互等各类认知需求，以知识和数据共同驱动的认知能力尤为重要。多模态学习通过从多种数据中提取信息，从而提高机器的理解能力。知识数据融合、知识表示与推理、认知规划和决策是认知推理的核心技术链，通过分析、推理、决策寻得问题最优解，实现复杂认知任务处理。情感计算、生成式人工智能等核心技术可提升机器对多维度信息的理解能力，进一步优化人机交互体验，提升机器想象力与创造力。

（一）多模态学习有效提高信息认知维度

多模态学习是一种旨在处理和理解多种模态的数据的机器学习方法，通过挖掘来自文本、图像、视频、音频等多种不同模态数据中的信息，实现各个模态间信息交流和转换，利用多模态数据之间的互补性，可学习到更好的特征表示，进一步提高机器对复杂场景的认知能力。多模态学习，目前主要用于在自然语言处理、计算机视觉、语音识别等任务，以及视频片段检索、图像描述生成、情感分析等场景。

多模态学习可通过模态特征提取、多模态对齐和多模态融合等方式实现。多模态特征提取，主要是从图像、音频和文本等多个不同的信息来源中提取特征，以获得更全面、更准确的数据，提高模型的性能与准确性^[4]。多模态特征提取需要对不同的数据类型进行分析，从而提取有用的特征信息，并进行合并来形成一个统一的特征表示。多模态对齐，主要将来自不同感官模态或者不同数据源的信息进行整合，

以实现信息的一致性和互通性。在计算机视觉中，多模态对齐可以将图像和文字进行对齐，以实现智能图像搜索或者图片中物体的识别。在自然语言处理中，多模态对齐可以将文本和音频进行对齐，以便于文本转语音或语音转文本。**多模态融合**，主要将多种不同的数据源（如文本、音频、图像和视频等）进行整合，综合利用多种数据来源的信息，进一步提高分析和推断的准确性和可靠性。在**多模态融合问题上**，可分为模型无关的方法、基于模型的方法。模型无关方法是传统多模态融合的基本思路，可以具体分为特征融合、决策融合和混合方法，此类方法的实现与具体模型无关，系统可利用单模态机器学习方法简单组合来构建。基于模型的方法，为多模态融合设计了特殊结构，有图模型、神经网络、多核学习等。

（二）知识数据融合加强认知基础

知识是认知的基础，数据是人工智能的重要“原料”，加强知识与数据融合运用能够提升机器的基本认知能力。一方面，数据中蕴含自然规律与行为模式，可以通过学习和训练，对规律和模式进行自动提取，从而实现智能化任务处理；另一方面，知识中蕴含领域背景与专业技能，可以提供逻辑结构和规则方法，将其以特定方式表示，可为认知推理提供依据。在数据不充分情况下，知识数据融合技术可以将先验知识或外部知识集成到数据模型中，从而克服数据不足、数据噪声、数据偏差等问题，提升模型的鲁棒性。

依据知识融入数据的方式，知识数据融合通常可分为浅层融合和深层融合。浅层融合是将知识以统计签名或规则的形式加入模型中，

有助于提高算法的求解效率和准确性，可基于领域知识和专家规则进行知识数据融合，建立系统化的模型，用于支持各种决策的制定和问题的解决。如将心理学领域知识与数据融合，打造分维模型，基于人的内在因素对人的行为进行建模，从而实现用户画像分析。心理学知识引入可以输入数据与输出数据之间的桥梁，引导模型更快更好地学习到输入与输出数据之间的内在联系，从而进一步探寻用户真实意图。深度融合是将知识以图结构或逻辑形式嵌入到模型中，有助于增强模型对知识的理解，可通过基于代数方程和知识图谱进行数据知识融合，代数方程可用于学习先验知识和知识数据增强，知识图谱中的结构化的符号知识与深度学习模型的学习能力相结合，以提供更丰富和全面的智能应用。如利用 ATOMIC 知识图库与数据模型融合的方式，可打造具备同理心回复的对话系统³。

（三）知识表示与推理可优化认知逻辑

知识表示是一种计算机可接受的用于描述知识的数据结构，合理的表示方式可在不确定性数据中寻求更优解。知识图谱是最主要的知识表示方式之一，其本质是一种揭示实体之间关系的语义网络，可形式化地描述现实世界的事物及其相互关系。知识图谱中，以实体表示具有独立存在和特定属性的对象，如人、地点、事件等；以关系表示实体之间的联系，如出生于、属于、发生在等；将实体和关系连接，形成包含多个三元组的复杂的网络结构，用来表示丰富

³ 《Improving Empathetic Dialogue Generation by Dynamically Infusing Commonsense Knowledge》

的语义信息和逻辑关系。然而，在现实认知场景下数据知识所蕴含的信息存在不确定性。为表征不确定的动态知识，可利用概率性知识图谱来表示更符合真实世界的知识属性。以常识图谱和行为概率性知识图谱引入对话生成模型的方法，可使问答机器人具备同理心，形成内在性格一致性，从而大幅度提升问答质量。

知识推理是认知的基本手段，让机器模拟人类推理方式寻求推理路径，合适的推理逻辑可优化认知能力。知识推理是依据推理策略、利用形式化的知识、进行机器思维和问题求解的过程，可实现已有知识基础上推测出未知知识。知识推理可根据不同组织方式进行分类，**一是以逻辑表达式方式推理**，逻辑符号可以表示事实、规则和结论之间的关系，通过推理引擎进行推理和推断，从而得出新的结论和知识；**二是以知识图谱方式推理**，知识图谱可以将不同领域、不同类型的知识进行整合和优化，以关系路径作为特征进行推理，预测实体间可能存在的某种特定联系；**三是以规则引擎方式推理**，规则引擎是知识推理技术的核心，规则引擎可以分为前向推理和后向推理两种方式，前向推理是从已知的事实和规则推断新的结论和知识，后向推理是从已知的结论和知识推断出对应的事实和规则。

（四）认知规划和决策处理复杂任务

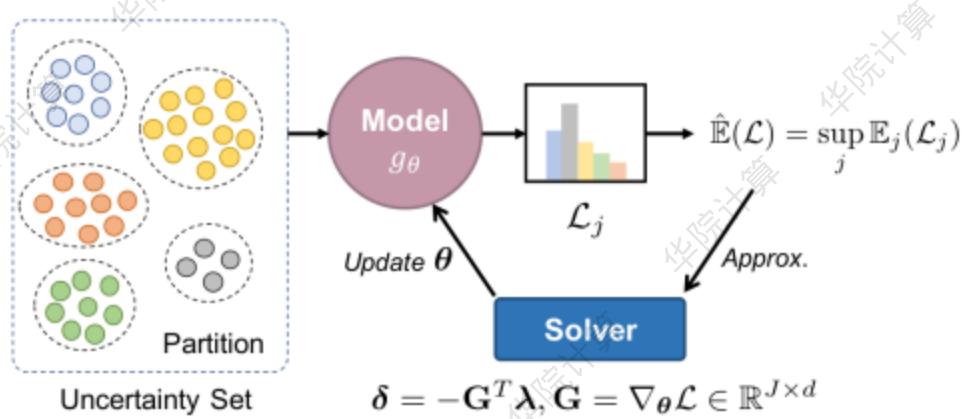
认知规划和决策是认知智能的核心环节之一，可帮助机器更好理解复杂认知任务，通过确定最优行动计划以达到预期目的。在认

知规划和决策中，规划可为了实现特定任务目标，依据资源、时间、能力和条件等因素构建行动方案；决策可在不确定性的情况下，通过对信息的处理和分析确定最优化的行动路线，以达到预期的成效。认知规划和决策可对复杂任务进行全局分析，规划和决策将复杂任务拆解为更小的、可管理的问题，分析相关目标、决策因素和变量之间的相互作用关系，用于理解问题的本质，并为每个子问题制定具体的步骤，通过逐步解决小的问题来建立全局认知，进而实现复杂问题优化求解。认知规划和决策可以帮助解决复杂任务中的不确定问题，一是可以将复杂问题进行分解，根据具体任务进行资源分配及限制；二是可对各类因素进行分析并逐步调整行动计划，以应对问题不确定性和时变性。

深度学习和强化学习在机器解决复杂规划与决策问题中发挥重要作用。一方面，深度学习可用于处理大量且复杂的数据，有助于训练规划的预测模型和分配资源。深度学习具备的深度层次结构可以捕获数据中的复杂的关系，用来训练预测模型处理无序、非线性和复杂数据的规划问题，以便预测一组决策的结果并确定可行解，避免搜索所有决策空间。另一方面，强化学习可用于规划实时性、不确定性任务类型，帮助模型自适应调整策略。强化学习的核心是奖励和反馈，强调的是如何在不完全了解问题的情况下，通过与环境交互来学习行动策略，收集环境反馈并根据反馈调整其行为。在认知规划决策中，强化学习可以帮助模型学习最优决策方案并让模型逐渐掌握整个规划问题的策略，被广泛应用于解决智能交通和机

器人场景的规划与决策问题中。此外，深度学习和强化学习的结合还可以优势互补，为机器解决复杂规划与决策问题提供了新的思路与方向，例如谷歌大脑发布的深度强化学习(Deep Reinforcement Learning)。

针对因数据偏态分布导致的规划和决策结果偏差问题，偏态学习可以有效提升模型鲁棒性。在实际问题中，由于正负样本不均衡、样本量少、样本含噪声等原因，数据偏态分布是非常常见的问题。例如在金融领域中，欺诈交易的样本数量往往远小于正常交易的样本数量，因此导致模型对欺诈交易的识别能力较弱。为了解决这个问题，可构建分布鲁棒的机器学习框架，并设计非独立同分布情形下的损失函数^{[1][2]}，此方案可以有效解决实际应用场景中普遍存在的样本非独立同分布性问题。从实验结果来看，在机器视觉领域，分布鲁棒的机器学习框架在收敛速度方面达到经典算法的10倍。在强化学习领域，实验证明，分布鲁棒的机器学习框架在收敛速度方面比经典算法提升50%，在得分方面分别比经典算法提升5%—10%⁴。



⁴ 《M²DQN: A Robust Method for Accelerating Deep Q-learning Network》

图4 基于非线性期望的鲁棒机器学习框架^[2]

（五）情感计算优化智能交互体验

情感计算的目的是通过赋予机器识别、理解和表达人类情感的能力，来建立和谐人机环境，使机器具备更高、更全面的认知能力。情感计算通过情感识别和情感合成使机器具备情感的理解和表达能力，有了情感，机器可以像人类对内部和外部刺激做出相应的反应，有助于使得机器像人一样进行自然亲近和生动交互。

情感识别可挖掘出有关情感的特征数据与内在情绪状态的映射关系，通过交互设备获取的图像、视频、文本等多模态数据，帮助机器更好地识别用户情感。目前，情感识别主要分为单模态专家模型和多模态情感识别大模型。单模态情感识别主要通过对视觉信号、语音信号和文本信号分别构建专家模型，提取面部表情特征、声音特征、文本特征等，通过提升各专家模型的表征能力辅助情感识别决策。多模态情感识别涉及对多个模态信号进行处理、各个模态情感特征的学习、多模态特征之间的融合、各模态之间的联合建模等，是情感计算中的热点研究问题之一。其利用检测、跟踪、姿态估计等算法，得到相关人脸微表情、人体姿态、声音、对话文本等数据，获取各个模态的对人物情感的决策结果，利用多模态融合策略，提升情感识别精度。此外，结合大模型应用趋势，多模态情感识别大模型构建了连续学习框架，结合网上公开数据和现有智能生成模型提供增量学习能力，持续扩充大模型的数据源，进而提升鲁棒性和域泛化能力^[5]。

可控情感生成，能依据给定的输入来控制生成情感，可将情感融入文字、语音或视频等内容，实现表达方式优化。可控情感生成技术，通过输入特定的关键词或者表情符号，让虚拟角色或聊天机器人表达出不同的情感，如喜悦、愤怒、沮丧等，从而实现更加真实、情感化的交流体验，提高生成内容的表现力，主要分为三种实现方式。**一是基于规则的方法**，使用基于规则的形式化语言来描述情感表达方式，利用符号逻辑运算来生成文本或语音。该方法需要依靠人工编写规则，因此可扩展性和灵活性受到一定限制。**二是基于统计模型的方法**，这种方法利用统计分析技术，从大规模数据中学习情感表达方式和情感模式，并根据所需情感类型生成文本或语音。**三是基于深度学习方法**，这种方法使用神经网络，可以自动学习情感表达方式和情感模式，具有高可扩展性、高效性和高自适应性的优点。目前，可控情感生成主要被应用于游戏、虚拟现实、聊天机器人等领域中。

（六）生成式人工智能提高机器创造力

生成式人工智能利用深度学习方法，可依照输入数据中的条件或限制生成所需新数据。生成式人工智能的核心技术是生成式模型，如生成式对抗网络（GAN）、变分自编码器（VAE）和扩散模型（Diffusion Models）等。此类模型能够捕捉到输入数据集的潜在结构，进一步发现规律，提高机器想象力。同时，扩散模型在生成式人工智能中的重要性尤为凸显，扩散模型是指利用随机过程中的马尔可夫链来建模数据生成过程。具体而言，扩散模型将数据视为一个随机变量，并通过迭代步骤改变数据的分布，使其逐渐接近目标分布。扩

散模型以多次迭代方式，生成与目标数据分布相似的新样本，从而实现自动创建文本、图像、音频和视频等内容，可用于自动化广告制作、新闻报道、视频生产、内容写作等诸多场景。

生成式人工智能具有强大的创造力和推理能力，可以进一步提高认知智能应用价值。在认知科学和心理学中，人类的创造力和想象力被认为是理解人类认知的重要因素之一。生成式人工智能通过学习数据集中的规律和模式，能够创造、发现和实现规律，从而实现对复杂问题的快速理解和创造性解决方案的生成。作为想象力和创造力的重要来源，生成式人工智能以造新事物的能力，已在许多领域中发挥重要作用，如 Stable Diffusion 在图像生成、视频生成等领域中表现出色，可以生成高质量、逼真的图像、视频以及其他类型的数据，其应用涵盖了艺术、娱乐、虚拟现实以及计算机图形学等领域；DALL-E2 可以根据与输入的文本生成相关的图像。通过对大规模数据集进行预训练，DALL-E2 能够生成具备创造力和多样性的图像。它在图像生成、创造性设计等领域具有重要的应用潜力。

第三章 认知智能场景化应用逐渐深化

随着技术的发展、产业格局的演变和市场需求的变化，智能预测、辅助决策、智能推荐等认知应用在医疗、金融、制造、教育等行业得到进一步落地。认知智能在各行业中的应用日益广泛和深入，成为行业智能化转型升级的重要引擎和动力。

（一）认知智能促进医疗服务再升级

认知智能与医疗领域场景深度融合，逐渐寻得更广阔的落地空间。一方面，认知智能可以通过医疗数据找到诊断逻辑，通过结合医生经验，模拟医生问诊方式，辅助医生做出更精准的诊断；另一方面，认知智能技术可以对患者信息进行分析及归纳，从而为患者提供个性化的治疗方案，智能化分析可优化医疗资源，提高治疗效率，进一步提升患者满意度。

在诊前评估场景，认知智能通过对患者信息、病情和症状自动分析，进而得出初步评估结果，为后续诊疗提供参考，同时可提供个性化的诊疗建议和推荐，可提升诊疗效率，并节约医疗资源，从而进一步提高患者满意度。例如，华院以数字人、智能对话、智能分析等技术构建的虚拟认知评估师一体机，可用于智能化认知评估测试。依据经典的认知功能评测量表，设计虚拟人对话交互，使得患者和老年群体可以通过与虚拟评估师对话来完成标准的认知评估测试。

在疾病预测场景，认知智能通过分析大量的医学数据和患者信息，辅助医生进行疾病预测，为患者制定个性化的治疗方案，从而提高治

疗效果和患者的生存率，可应用于各种疾病的诊断和治疗，如心血管疾病、癌症、糖尿病等。例如，麻省理工开发 Sybil 智能系统可帮助人群预测 1~6 年患肺癌的风险；IBM Watson for Oncology 可以通过分析患者的病历，提供个性化的癌症治疗方案。

在辅助诊断场景，认知智能用于基于医学图像的诊断辅助系统可以自动分析 CT 和 MRI 影像，可识别病变和辅助诊断，提高诊断速度和准确度。一项针对基于深度学习技术的影像诊断系统的研究表明，该系统可以将医生对 CT 和 MRI 影像的分析速度提高了 30% 至 50%。同时，智能辅助诊断可通过分析患者的病史、症状等信息，帮助医生排除一些不可能的病因，缩小诊断范围，提高诊断准确率和效率。

（二）认知智能促进金融科技应用创新

认知智能不断拓宽金融业务领域边界，深入场景赋能。随着金融数据爆发式增长，数据解析和判断需求衍生出具有金融领域自身特性的融合应用新场景，如智能营销、智能风控、智能理赔、诈骗分析、智能投顾等。认知智能通过对金融数据进行结构化提取和智能化分析，构建金融行业知识图谱，满足金融机构当前面临的海量化、碎片化、多元化的金融服务需求，助力金融业显著降本增效，提升服务质量。

在客户服务场景，目前金融机构的智能客服或厅堂机器人存在交互模板化问题，在与人交互中通常可满足标准化问答，但对于个性化问题的解决响应程度不高；此外，由于认知能力有限，对于问题深层次原因、回应客户情绪、寻找潜在营销机会的能力尚有待提升。认知

智能可结合客服人员的经验，对金融机构的整体数据和知识进行整合，完善智能客服系统。一方面将客服人员的精力集中在解决重要客户或疑难问题上，另一方面提升客服的营销和意见反馈能力，进一步推动客服中心向利润中心转型。

在信贷管理场景，当前金融机构的信贷管理日益依赖数据模型，用于模型训练的数据需由专业人员审批完成，以此方式构建的模型，因存在数据滞后性，所以虽可反映过去的风险特征，却无法预测新风险。认知智能可结合金融机构专业数据审批人员的经验和知识，完成数据自动审批，避免因人工审批而造成模型优化空档期，提升模型的敏捷性和适应性。

在风险管理场景，传统的风险管理往往基于单种风险、单个风险点来进行识别和预防措施，认知智能可将金融机构中经验丰富的风险管理专家的相关经验进行提炼和融入，将风险防控由点及面，放大专家的业务能力，如利用认知智能中的知识图谱，通过关系网络进行团伙识别。

在内容营销场景，认知智能可通过销售数据分析、需求挖掘、营销内容推荐等能力，以智能化应用解决营销内容生产门槛高，宣传口径难以统一等难题。例如，华院以智能推荐、内容生成、数字人等技术构建的“营销内容平台”，平台以智能推荐技术结合保险行业专家知识生成合规且精准的营销文案，依靠数字人实现低成本、高仿真复刻代理人形象与声音，自动化一键生成营销视频。

（三）认知智能加速工业制造智能转型

制造业作为国民经济的重要支柱行业，正积极响应数字化转型浪潮。认知智能赋能价值已渗透到产品设计、生产规划、生产实施、生产交付等诸多重要生产环节，为制造业数字化转型提供了更为广阔的空间和方向。

在产品设计环节，认知智能可实现基于知识的快速方案设计。工业设计技术门槛高，需要设计师具备专业的技术和丰富的经验，在实际生产中，由于人力、时间、成本等因素的限制，很难实现高效、高质的工业设计。认知智能可通过学习和掌握设计规律和趋势，从而生成更加科学合理的设计方案，同时，智能化设计可进行版本快速迭代。

在生产规划环节，认知智能可将专家经验与投产数据融合分析，优化生产资源分配，促进绿色生产。在煤焦钢产业，配煤制定一方面需综合考虑生产安全及成本控制，另一方面高度依赖配煤师经验。认知智能可融合专家经验，并综合分析资源与生产数据，实现生产资源优化。如华院“智能化配煤系统”，以偏态学习算法为基础结合炼焦煤资源数据推荐配煤方案，在保持焦炭质量稳定及生产能效基础上，不断优化焦煤配比，同时，将焦炭质量预测的误差控制在焦化厂可接受的小波动率内，防止因误差变化频率较大而造成的生产风险。

在生产实施环节，认知智能可在设备维护、质量检测中起到重要预警及监督作用。生产工艺优化和质量检测中起到，设备维护中，利用历史数据，结合设备运行状态，及时预测预警运行故障风险，辅助诊断故障原因分析并指导维修维护；质量检测中，利用视觉计算及认知

分析，在精密产品上以远超人类的视觉能力发现微小缺陷，并通过认知模型做出风险预测提示及可修复评估建议。

在生产交付环节，认知智能可通过在物流仓储、营销推荐、售前售后等方面通过数据打通和流转，提供整体交付方案。认知智能合理规划物流及仓储资源，结合地理位置、客户偏好、广告活动等多种数据，锁定客群，制定精准营销策略，借助历史数据等实现更优的售后服务效果。

(四) 认知智能推进教育模式创新优化

认知智能与教育融合，在教学评估、课后辅导、辅助学习等多方面促进教育方式创新变革。在教育领域，认知智能一方面能为学习者提供个性化的支持和帮助，提高学习效果；另一方面能通过革新教学模式，减轻教师教学负担。更进一步，认知智能可以通过改变和拓展教学模式，引发领域与行业变革催生教育生态升级。

在教学评估方面，认知智能可以分析学生的学习过程，包括互动内容、学习行为和测试结果等，为学生个体提供即时支持或反馈，并为教师提供改进教学内容和策略的建议，从而给每一位学生提供细致的学习状况评估和能力评估。通过动态学习诊断、反馈与资源推荐的自适应学习机制，可以适应学生动态变化的学习需求，从而打破标准化的教育限制，释放出学生的创造力与活力。

在学习辅导方面，认知智能可以通过分析教学内容及专业知识，为学习者提供多学科学习辅导服务。如在数学学习中，借助认知智能

技术可以解答学生的问题，用自然易懂的语言讲解有关公式、定理和实例，并且具备步骤批改、错因诊断、交互式学习等个性化辅导能力；在语言学习辅导中，认知智能一方面可以提供字词句、中心思想、语言表达、写作技法批改；另一方面可用于沉浸式的对话练习、实时互动和语音评测。

在辅助学习方面，认知智能可为学生提供智能答题、多元的学习内容推荐等各类学习支持功能。例如，智能助学平台 CENTURY TECH，以智能推荐、行为分析的能力帮助学生提高居家学习的质量。该学习平台，依据学生行为数据分析，融合教育领域和神经科学专业知识，来推测学生“更需要”的学习内容，构建个性化的学习计划；同时平台可以通过学习数据，定位学习过程中的能力短板和知识盲区，并为学习者个人学习建议以及个性化内容和途径。

第四章 认知智能技术与应用方向展望

多模态大模型基于自有的微调属性和多模态数据理解融合能力将提升认知的准确性和可靠性，推动认知智能应用落地。生成式人工智能通过提高问题解决效率、提升方案创造性，进一步优化认知智能应用能力。此外，加强用户对决策过程理解和信任对于认知智能技术应用也非常关键，需要在技术和合规两方面“双管齐下”。长远来看，认知智能不是人工智能应用的最终阶段，人工智能将向更具“意识”方向发展。

（一）多模态大模型拓展认知智能应用边界

多模态大模型将延伸出更加强大的认知能力，深入贯通行业应用。在认知智能应用中主要存在两方面挑战，**一是**如何以更高效、更广泛的方式实现落地，**二是**如何在复杂场景中提升理解、推理、决策等认知能力。多模态大模型为应对复杂多变的信息及泛化落地提供了有效的解决方案，**一方面**，大模型自有的微调属性，已经训练好的大模型上进行调整，以适应新的任务或数据集，以此支持模型快速响应不同应用场景；**另一方面**，多模态大模型能够自动学习和提取多模态数据中的特征和关系，对多种语言、图像、音频等多种形式的信息进行联合理解，通过增强认知的多模态关联技术，提高不同模态任务信息的处理能力，为许多任务提供更准确、更可靠的结果。可见，多模态大模型可进一步推动认知智能落地，为实现多模态认知智能提供了驱动力。

（二）可解释的认知智能增强技术可信赖度

加强认知智能的可解释性，提升用户对决策过程理解和信任。可解释的认知智能在处理数据和做出决策的同时，需要向用户提供清晰、透明和易于理解的解释，一方面帮助用户理解系统是如何工作的、为什么做出特定的决策、哪些因素影响了其行为；另一方面使用户自行规划能力优化方向，执行优化方案，从而为用户提供信心，提高产品信任度。但是，实现认知智能的可解释性并非易事，不仅需要提出具备可解释性能力的模型和可以解释复杂模型的分析技术方法，还需要在保护用户隐私和提高透明度之间找到平衡，确保向用户提供关于决策过程的详细信息，同时要防止滥用这些信息以免危害用户的隐私和利益。

确定认知智能的解释性目标，建立统一的评价标准体系。针对不同的应用场景和用户受众所需要解释的内容，需要建立统一的评判指标，主要包括信息性、可移植性、可访问性等。一是信息性，要求算法可提供大量有关决策目标的信息，将用户的决定与模型的解决方案建立联系，使得用户理解模型内部的决策机制；二是可移植性，要求算法在不同的场景和数据都具有良好的应用效果，检验模型应用的边界值；三是可访问性，要求用非专业的解释方式来进行算法的解释，保证非专业人员也能明白算法的决策过程，降低用户对算法提出改进意见时的技术知识门槛。

（三）类脑智能推动认知智能向意识智能发展

认知智能和脑科学之间存在密切的关系，交叉研究成为必然趋势。

人脑能够自组织地协同数百项认知功能，灵活适应复杂多变的环境，为让机器进一步接近人脑认知能力，如何整合多尺度生物信息来构建具有生物合理性和计算高效性的神经网络模型，是类脑智能和计算神经科学领域共同关注和面临的重要挑战。类脑智能主要受脑神经机制和认知行为机制启发，以数据为基础，以计算为手段，利用脑认知功能模拟、脑连接图谱技术、类脑智能体验证等手段，通过软硬件协同的方式模拟人类认知，实现认知机理、学习能力、记忆存储等功能。

基于类脑智能的意识计算，将推动认知智能向意识层面发展。意识是大脑所有活动的总和，类脑的意识体系综合连接、监控、协调、管控各个类脑能力的运作，从类脑的感知、记忆、学习、理解、认知、逻辑、情感、沟通到最终的意识体系是逐步往上构建起来的。对于认知智能中的类脑意识体系，是在对大脑的结构、功能、机制方面的研究基础之上进行构建。目前，意识计算领域理论研究还处于起步阶段，华院已经启动利用计算机科学和人工智能模拟人类意识过程进而增强机器“意识”方面的研究工作。

参考文献

- [1] Xu Q, Xuan X. Nonlinear regression without iid assumption[J]. Probability, Uncertainty and Quantitative Risk, 2019, 4: 1–15.
- [2] Zhang Z, Zou Y, Lai J, et al. M2DQN: A Robust Method for Accelerating Deep Q-learning Network[J]. International Conference on Machine Learning and Computing (ICMLC). 2022.
- [3] Xu Q, Xuan X. A Robust Algorithm for Minimax Problems with Convergence Analysis. *preprint*. 2023.
- [4] Xiaomei Wang, Xuli Shen, Qing Xu, Shenghua Gao. TransEmo: Hierarchical Context Interaction and Reasoning for Emotion Recognition with Transformers. *preprint*. 2023.
- [5] Xuli Shen, Xiaomei Wang, Qing Xu, Weilin Shen, Hua Cai , Weifeng Ge , Shenghua Gao, Xiangyang Xue. EmoTAG: Manipulating Emotional Talking Head via 3DMM Regularized Neural Radiance Fields. *preprint*. 2023.