

从“傻白甜”到“杠精”

AI 有了认知智能萌芽

1

从理解语义到“读心术”,认知智能初现

“你告诉一个 AI 自己有减肥计划, AI 如果每天建议你吃沙拉,你就会说它不好,不符合心意;如果它天天让你吃蛋糕,你也会说它不好,不符合减肥需求。”上海纽约大学商学部主任陈宇新教授举了个连真正的人类都难以把握的两难抉择,充分说明理解语义并不一定能够真正把握要义。

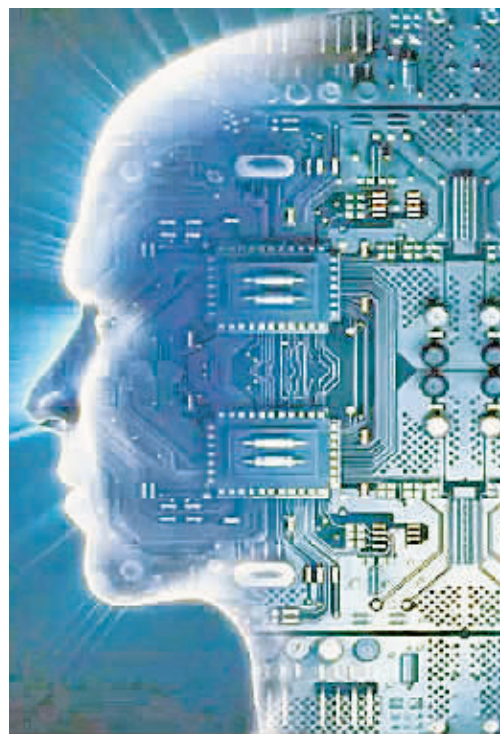
“虽然主人说这是我想要的,但 AI 做了之后,主人仍会说这不是他想要的……”拥有“认知智能”的未来 AI 将充分理解这句堪比绕口令的实质,陈宇新认为,当数据足够多,算法足够强大,模型足够优秀之后, AI 将从理解语义到拥有“读心术”,分析推测出人类当时可能都不确定的真正想法,在人心矛盾时给出“挠到痒处”的答案。

将这个“减肥吃什么”的难题放大来看,就是对正反的“两难”进行一系列诸如身体指征、爱好、时间等多维度

的数据综合分析后,用到大量的人类学、心理学、社会学的内容和方法,通过一系列探索性地问答获得数据,给出结果。

这种探索性的分析将使 AI 得到进化,并拥有更广阔的“职业”选择。据介绍,美国有公司做出了第一个人工智能律师 ROSS, 美国一个律师以往用上百上千个小时做一个案例的分析和调查,现在用了 ROSS 和人一起合作,时间减少到 28 分钟,ROSS, 可以提供正反双方的意见让律师做决策。

“正反意见的博弈还可以帮高层做决策,可以通过一个机器人提供正反两面意见的辩论,帮助高层综合分析做更好的决策。”谢少毅认为,这方面的研究才刚刚开始。进化的 AI 将被用在更多需要做信息分析或做决策的领域,例如高级销售、经济分析师、企业高级管理人员等。



数据融合包括拥有海量数据和实现高效融合两方面。 据网络

2

数据动态融合, AI 走进认知智能的“基石”

从海量数据中“捞出”与当前任务关联的数据,是 AI 拥有综合分析能力的第一步。百分点董事长兼 CEO 苏萌表示, AI 走进认知智能首先建立在数据融合的基础上。

以辩论项目为例,沃森化身的这位“女性”辩手,和另一位人类辩手就“是否应该把体育博彩合法化”的议题展开辩论时,能在听懂人类长达 4 分钟的快速陈述(700—900 个单词)后,进行迅速反应,搜集资料、选取角度,给出旁征博引的辩驳;在短兵相接的短辩交锋中也能很快理解对方的观点、组织语句并做出有针对性的论述。

苏萌表示,数据目前来看仍然是人工智能的基石。不久前,德国商业软件巨头

SAP 以 80 亿美元收购全球最大在线调查公司 Qualtrics,说明了数据融合的重要性。2017 年,百分点也并购了中国最大的在线调研服务商极速洞察,使得百分点同时拥有高达 380 万的中国最大用户样本库,将双方的行为数据和态度数据进行融合。

数据融合包括拥有海量数据和实现高效融合两方面。

苏萌表示,数据融合指的是打破业务系统的烟囱融合海量的多元异构的数据,它既是一次企业内外部数据的融合,也是大数据的融合,同时也是历史数据和实时数据的融合,只有数据融合了之后,才能洞见真相避免偏见,才能明智决策避免武断。

“知识图谱是我们让机器去了解、认识

人类世界的核心,它把我们的知识构建成网状的知识结构,再通过人机交互直接输出。”苏萌进一步解释,数据融合意味着知识需要不断更新,但在更新一个知识点的时候往往需要更新整个知识系统,操作非常繁琐。

“出版行业或者媒体,甚至公安行业,所涉及的知识都需要实时动态理解,某一个知识不是固定不变的,所以知识图谱未来的发展方向是动态知识图谱。”苏萌解释,这个世界由很多的东西构成,比如人、事、地,这些可以理解为本体,本体发生变化或者本体间的关系发生变化后,动态知识图谱能够迅速自动重构,高效地实现对一个行业的知识图谱的构建。

3

攻坚在路上,来自小数据的挑战

新的领域,没有海量样本、训练速度不够,是不是每遇到一个特殊场景,都要重新建模,输入大量的样本让机器重新学习一次呢?条件并不允许。

“很多客户的实际应用场景是数据量不够甚至是比较少的,在这种情况下我们结合迁移学习技术,解决了小样本的模型训练问题,标注量只有以往的 10%。”百分点首席算法科学家苏海波说。

从小数据中获得学习能力,也是人工智能走向认知智能的一个重要部分。人工智能专家经常用猫举例:小孩子认得猫只要指给他看一两次就够了, AI 却需要成百上千张图片,才能掌握猫的特征。

为此,人工智能科学家尝试结合迁移学习开发一系列算法,节约人工标注样本的时间,让模型在少量的标注数据上,也能取得好的效果,可以理解为赋予 AI“举一

反三”的能力。“例如像电商评论情感分类这样一个任务,用传统的深度学习模型需要数万条数据才能达到 85% 左右的效果,但是如果采用深度迁移学习技术,数百条数据就能达到同样的效果。”苏海波说,我们判断认知智能未来会迎来黄金十年的发展,为此百分点成立了认知智能实验室。除了深度迁移学习技术,实验室还会重点研发多语种自然语言处理技术,帮助认知智能实现跨种族、国际化。

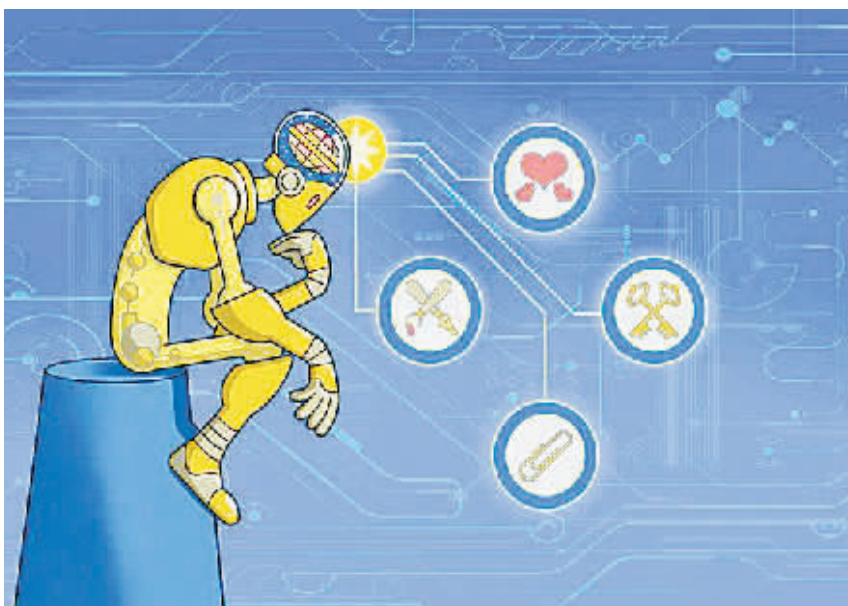
“未来我们会与各大高校和研究机构开展合作,建立联合实验室,共同探索更多前沿的认知智能技术,包括各个重点行业的知识图谱构建、自动问答等等。基于这些技术,实验室会研发出更多行业落地的应用产品,为客户创造价值,用认知智能推动社会进步。”苏海波说。

据科技日报

每个人身体里都有一个“杠精”:“这可不一定”“我可不这么认为”“那可说不准”……这些“杠”词要么在心里默念、要么在口中碎碎念、要么大声讲出来,无论是否表达,“杠”的意识出现均是人类独立思维的表现,所基于的正是人类的综合分析能力。

如今,人工智能(AI)也开始“杠”得有条理、有逻辑、有思路起来。日前,在百分点举办的无界智能发布会现场,IBM 大中华区全球信息科技服务部总经理谢少毅介绍了人工智能做辩论的项目,“杠精”们大爱的口头禅出现在了 IBM 的人工智能系统?“Project Debater”口中,这背后是机器学习、自然语言处理、海量数据搜索处理、识别情绪等大量人工智能技术的应用。

越来越多的人工智能研究者发明新的技术让 AI 向人类“偷师”,一定程度上学得触类旁通、独立思考的能力。从“傻白甜”到“杠精”,AI 向真实的“人”又贴近了一步,从“感知智能”慢慢走向“认知智能”,哪些技术是 AI 不可或缺的? 还有哪些仍待攻坚呢?



从小数据中获得学习能力,也是人工智能走向认知智能的一个重要部分。据网络