

澳洲幸运5九码计划

EMCm7DuGMf9IBRLV

澳洲幸运5九码计划MSRA首席研究员刘炜清：为每位研究员提供AI科研助理

2025，随着大语言模型技术的迅猛发展，数据科学领域正经历一场静默的革命。传统的特征工程、模型训练与迭代优化流程，正被智能化的研发助手所改变。

在第三届AIGC产业峰会上，微软亚洲研究院（MSRA）首席研究员刘炜清带来了一项引人瞩目的研究成果——RD-Agent，一个旨在“自动化、增强到重塑”数据科学研究流程的智能系统。

这项研究源于一个现实问题：当大模型浪潮席卷各行各业时，数据科学家们面临着被颠覆还是被赋能的选择题。与其坐等被颠覆，MSRA团队选择主动探索，将Agent技术转化为科研助理，为每位研究员提供一个“虚拟助理”，承担那些高门槛但又重复性强的研发工作。

从最初帮助研究员实现想法的开发助手，到能够自主提出研究方向的研究伙伴，RD-Agent正在以“AI驱动数据驱动AI”的方式，重新定义数据科学的工作流程。这不仅是一项技术创新，更是对未来科研方式的一次大胆探索。

为了完整体现刘炜清的思考，在不改变原意的基础上，量子位对演讲内容进行了编辑整理，希望能给你带来更多启发。

中国AIGC产业峰会是由量子位主办的AI领域前沿峰会，20余位产业代表与会讨论。线下参会观众超千人，线上直播观众320万+，累计曝光2000万+。

话题要点

以下为刘炜清演讲全文：

起源：以自动化为目标

大家早上好，我是来自微软亚洲研究院（MSRA）的刘炜清。今天我给大家带来的研究题目是大语言模型时代下的数据科学新引擎RD-Agent，从自动化、增强到重塑。这三个关键词恰好也是描述了我们RD-Agent的起源、现状和未来。

我们首先从RD-Agent以自动化为目标的起源开始介绍起。首先介绍一下我们RD-Agent背后的团队，我们团队从2017年初开始与金融行业的各个领域不同方向的公司进行深度的产业的科研的一些合作，大家对2017年这个时间有一些感觉的话，这恰好是AlphaGo当年横空出世打败世界冠军，使各行各业有一个担忧，自己的行业会不会被AI所颠覆，同时也会有一些小的期待——会不会自己做的业务有机会能够被AI所赋能的奇妙的一段时间。

我们团队代表MSRA跟金融行业的合作伙伴们一起探讨当时最先进的AI技术，当时是Deep Learning深度学习的技术，看看能不能对金融行业核心的场景业务和问题进行一个智能化的升级。比较幸运的是，我们在多年的努力下面还是有不少研究成果成功的落地，并且在合作伙伴实际产品和业务中间获得不错的效果。

我们做这些合作的过程中会发现，实际场景产业落地中间遇到的挑战和困难，并没有很好地被学术界所广泛关注和很好地解决，我们就会把这样的一类挑战进行抽象，并且尝试对他进行解决，最后以学术论文的方式分享给业界以及学术界，帮助大家更多地关注这些核心的挑战，以及帮助方向持续的演进，做出我们自己的贡献。

做产业落地研究的时候，我们发现这一类的研究直接去做，可能跟真实场景会有很大的GAP。我们需要更好的基础设施、研究框架才能帮助我们做有真实价值的研究，我们以量化研究为例子，当时发现公开可获得的基础设施研究框架都不太能满足需求后，我们自己内部开发了这样的一个研究框架Qlib，并且将它进行开源。比较幸运的是获得社区里比较多的关注和认可，给了我们很大的动力来持续研究和改进它。

回望过去将近十年产业相关的应用和落地的工作，我们主要的工作都是在数据科学和机器学习的范畴。左边的图中所示，我们常用范式就是在业务海量数据中间找出来有价值的特征，利用这些特征我们训练模型对其中的规律进行建模，从而得到智能化的解决方案服务于各种各样业务场景的需求。智能化解决方案的整个开发过程则是以迭代渐进的方式进行的，第一版看看效果怎么样，根据反馈再改进下一版，迭代地得到最终的智能化解决方案。

现在我们到了大模型的时代，这一波技术的浪潮与之前AlphaGo带来的那一波狂热相比，也带来了各行各业的新一轮的思考，有没有可能这个行业会被颠覆，或者自己的业务有没有可能被赋能。这一波技术潮流从我们自己的体感来看，会感觉我们自己本领域的研究人员第一波受到的冲击其实很大的。大语言模型出来以后，很多持续稳步发展多年的研究领域，已观察到慢慢有些式微，面临很多挑战。这时候我们自己也要考虑，我们自己的研究领域和方向这些东西是被颠覆还是赋能？与其等着被颠覆，我们尝试能不能用大语言模型对我们自己的研究进行赋能，就好像我们之前研发的Qlib基础设施框架帮助我们更好地做相关研究一样，这就是我们RD-Agent最初以研发自动化为目标的设计初衷。

在这样的初衷下面，理想情况下有了大语言模型，有了Agent，是不是能够让每一个研究员、每一个数据科学家都能有一个Agent为代表的科研助理或者研究实习生，它来承担我们日常工作中间需要做非常多的重复、但是又有高门槛的工作。

具体来说，人人都有科研助理意味着人人都是老板，老板怎么来做这样的一件事情呢？我有大概的想法，同时我手头上有一堆数据，我交给科研助理Agent，帮我实现一下看看想法怎么样。

当我们有了正确代码的实现，训练好了一个模型，同时对它进行正确的全面的评测以后，如果第一版效果还不错，老板英明！第一版效果太好也会想想是不是我有一些Test Data甚至Label都Leak了，我们得check一下有没有bug；效果太差的话，当然得看一看是什么原因，是不是训练的时候loss都飞掉了。往往这些检查验证实现的过程是需要多轮迭代才能得到最终的可信结果的，这个过程需要比较高的门槛，同时也需要非常繁重又相对重复的劳动。

经过几个月的努力我们构建了Agent工具，帮助我们自己在日常研究工作中间很大的效率的提升，减轻了很多细节实现的负担。当我们已经验证了能够赋能我们自己研究的时候，我们就在想这个Agent有没有可能真正赋能产业，赋能实际问题，我们找到我们合作伙伴看他们场景中间有没有类似的问题，我们最终找到了论文或者研究报告的自动实现或者验证的问题，这个问题在与合作伙伴的沟通中间发现，在他们日常的研发工作中间每天都在发生的，比如说看到一个公开或者半公开的研究报告或者论文，大概描述了一个新颖的方法。同时在他们论文所在的那个数据集，假设叫数据集A上效果非常好，但是在他们自己自有的数据跟论文中所使用的的数据不同，场景也有稍微的区别，那它效果怎么样？这就需要重新实现这个方法看看在我自己这边好不好使，往往这样的工作并不是所有都能找到开源的代码，这时候需要自己来复现重新验证。这类的工作其实占用了他们研发过程中间非常多的精力和时间。经过一段时间

的努力我们会发现RD-Agent确实能够很大地帮助到他们真实的日常工作，让这个事情能够自动化提升研发效率。

现状：为研究员提供AI科研助理 / 为各行业提供AI自动研发团队

具体来看，我们选择的是一类属于Feature Generation的具体问题，这是我们统计分析出来实现每一个idea所需要写的核心代码，其实也就是几十行的体量，我们发现如果使用原生大模型做这件事情几乎不可能在直出的实践中做对，经过很多改进和尝试，RD-Agent系统能达到80、90%可用的效果。当然我们也观察到了带reasoning能力的大模型出来以后，模型的代码实现能力得到很大的提升，但是依然很难做到一次就对，依然需要采用我们这种多次迭代渐进的方式才能做对。

有了第一阶段对自己的研究过程进行赋能，以及对实际产业界的任务进行了赋能以外，我们不太担心会被颠覆了，我们心思开始活络了，能不能做更有价值的一件事情呢？现阶段我们目标构建一个通用的数据科学或者机器学习的Agent，目的是为了增强人类专家的能力和产出。怎么做呢？在这一阶段我们的目标是自动地做新场景和新问题，并且持续的改进当下的解决方案，而不仅仅是当科研助理，仅仅是实现人类专家给出来的idea。

蓝色部分可以看到，刚刚我们作为科研助理Agent的部分，我们叫做Development Agent，它是需要大概的idea的描述，之前是依赖人类专家给这样的描述，我是不是也能依赖Agent给出研发的idea呢？橙色部分是我们新加入的Research Agent，它直接以当前我们的任务和场景的描述作为输入，来迭代产生idea，并且根据Development Agent来产生的当前idea下最终模型的真实效果来进行下一轮idea迭代的依据，这就是我们整体的设计。希望在很少的专家介入或者没有领域专家介入的情况下也能自动给出智能化的解决方案，这时候我们就有了RD-Agent当前的整个框架。

这个框架我们在去年底进行了开源，比较短的时间内获得比较好的关注，我们取了一个slogan，AI Drives Data-Driven AI。

在这样的问題下面，我们先做了一个简短的视频介绍我们RD-Agent。

未来：重塑数据科学

看了录屏以后相信大家对于RD-Agent有了更直观的感觉，刚刚说到现阶段的目标是研发增强，怎么增强人类专家呢，光靠原生大语言模型做不到，那么大语言模型做不到什么我们就补什么，比如我们组里有Kaggle Grand Master，我们来看看语言模型做不到，但是他能做到的，我们通过引入领域知识和经验并且优化我们算法使我们的Agent更像人类专家而不仅仅是直接原生大语言模型的能力，这个技术路线很有效但是也有瓶颈，因为人类专家做这一类问題的时候也是有瓶颈的，光靠模仿是很难超越的。

比如我们现在就发现，当迭代到某一轮数，解决方案还不错的时候，Agent给出5个改进路线，人类专家看的时候觉得每个路线都很有道理，我们只有试一试才知道最终的结果，试完之后发现有些性能有增强，有些对性能反而有很大的损害。我们有没有可能通过数据驱动的方式把所有探索的过程记录下来，比如在什么情况下尝试了什么idea，这样的idea最终对性能的提升有没有帮助，并且基于这些数据来增强我们大模型或者Agent相关的能力，做到比人类专家对于什么样的idea效果更好，有更直观更准的判断。这样的话避免我们只是模仿专家但是很难超过专家的瓶颈，从而让我们有机会超过专家达到更好的效果。

当下我们正在快速迭代快速演进我们的相关设计，同时我们也在同步验证当前RD-Agent的能力，由于时间问題我没有办法把细节进行展开，我们在一些Kaggle比赛中间进行验证，我们直观感受是这个RD-Agent已经能够达到初阶从业者水平，比我们接触到很多在校学生们能力显得更强一些。

未来我们希望我们RD-Agent方向是什么？我们奔着自主发现新方法的方向进行努力，如果我们能做到这个方向就是能够重塑Data Science或者是Machine Learning领域。听说围棋的棋手们现在也都跟着AI学习怎么更好地下棋，就是重塑这个领域。

怎么做到这件事情？我们尝试设计三阶段循序渐进的方式，从最开始自动化为初衷，我们为专家减少脏活累活，从而得到更高效的专注于创新。当下我们做增强的事情，我们其实是尝试比专家能够更快地找到我们的方法和场景和数据更好地匹配，更快地找到更优的方法。未来我们期望通过不同的方法在各个场景、数据上表现的观察、分析和理解改进现有的方法或者发明新的方法。真的到了这一步的时候我们自己的领域肯定是被颠覆了，但由于我们能够有能力发明更好的方法，相信这也能够扎扎实实地赋能到各行各业，看上去是挺美好的一个未来。以上就是我关于RD-Agent工作的介绍，由于时间问题很难把很多细节进行展开，也很难介绍未来的计划，如果大家对我们工作感兴趣或者对我们未来的发展想有一个关注的话，欢迎大家关注我们开源项目RD-Agent。

今天我的演讲就是这些，谢谢大家。

澳洲168幸运5开奖结果

澳洲幸运10计划精准计划

极速赛车计划数据大全

168飞艇计划全天预测永久免费

澳洲幸运十官方开奖

澳洲幸运5 稳赢

澳洲幸运10正规吗百度知道

澳洲幸运5官方开奖直播官方app介绍

澳洲5开庄算账机器人

澳洲幸运10精准一期计划

澳洲10开奖记录论

澳洲十全计划网页版

澳洲pk时全天免费计划app

168极速赛车计划软件

澳洲幸运10开奖结果历史记录查询

168澳洲幸运10历史开奖查询

免费168飞艇计划网页版

澳洲幸运10走势分析

