

# 评人工智能如何走向新阶段？

陆首群 2019.8.8 (后略有修改)

机器学习/深度学习取得巨大进步，今天已成为引领人工智能发展引擎，这时机器学习/深度学习的运作空间其重点在语音识别、图像识别、自动驾驶以及AlphaGO 围棋等领域，但机器学习/深度学习是有缺陷的，它本质上是一项黑盒子（暗箱操作）技术或盲模型，其训练过程是不可解释、不可理解和不可控的，为了推动人工智能大发展，机器学习/深度学习必须克服其自身的缺陷，打破黑盒子痼疾，实现可解释的人工智能。最近不少机构的研究团队正在研发可解释的机器学习/深度学习技术，并尝试建立可解释的机器学习/深度学习模型。

为了探索人工智能发展目标，比较有前途的是从异步脉冲神经网络出发，这是类似生物的神经网络，最近发布的一款基于脉冲神经网络的神经拟态计算系统（也称为类脑计算），这款计算系统包含 1 亿神经元，它的出现颠覆了传统计算机的冯-诺依曼架构，实现人脑的智能功能。

得知识者得天下！真正的人工智能是由知识驱动的，通过新知识工程，要实现可解释的人工智能还有最后一公里的差距，其短板表现在大规模语义网络（知识图谱）的能力不足，一批研发团队正在对提升大规模语义网络中的理解和解释两个核心能力的研究，并对不被其概括进来的常识进行推理攻关。

除从上述三路出发探索人工智能的发展目标外，目前伴随着国内外 10 几个实例的脑机接口，也是提高人工智能的一条新路。

# 人工智能走向何方？

陆首群 2020.6.12（取跟贴535文字）

评“人工智能如何走向新阶段？”-535条跟贴为你揭密！

自从我们于2019年8月初提出“评人工智能如何走向新阶段？”这个命题以来，迄今10个多月了，我们陆续收到国内外传递来的跟贴535条。

我们开辟了一个讨论的平台（自由论坛），期望国内外的专家和草根不拘一格都可以发送跟贴而来：针对人工智能的发展各抒其见、发表评论（我们不在乎评浅、评错）。为避免“某国”对此敏感议题过度干预，我们一开始就公开申明：由国内外传递人工智能发展前沿的信息（跟贴）仅限于开源的或公开发表的。

在535条跟贴中，有一些是人工智能大师的言论（或者是有人转载以跟贴方式传递的）。谈到“人工智能走向何方？”大家可从535条跟贴中找到答案。所以这批跟贴弥足珍贵！

人工智能走向何方？在跟贴447中谈到，人工智能未来发展的第一目标是人类智能或接近（类似）人类智能。

也就是说，人工智能的发展任务是实现可解释的人工智能，或者说是让计算机建立像人脑那样的认知智能。

跟贴447还谈到，目前从三条路出发正在探索人工智能发展目标：

一是从机器学习/深度学习模型出发，

二是从异步脉冲神经网络出发，

三是从新知识工程、大规模语义网络出发。

其实在535条跟贴中还有伴随着十几个实例的脑机接口一条路，在我们这里可称为：

四是从脑机接口出发。

### （一），从机器学习/深度学习模型出发

众所周知，机器学习/深度学习取得巨大进步，今天已成为引领人工智能大发展的引擎，但机器学习/深度学习是有缺陷的，它的黑盒子特性使其行为背后缺乏透明度，用户很难理解机器学习/深度学习模型是如何做出决策的，其不可理解、不可解释的缺陷限制了它的扩大、提升应用场景，有人甚至说机器学习/深度学习（技术）已达天花板。

在 527 跟贴中，图灵奖得主、贝叶斯网之父 Judea pearl 谈到，当前机器学习理论有局限性，完全以统计学或盲模型（即黑盒子）的方式运行，所以不能成为强 AI 的基础。

在 528 跟贴中也谈到，深度学习算法并不完美，有待继续加强理论研究。

可喜的是有一些跟贴（如 480、517、529、530）正在研发可解释的机器学习模型。在跟贴 517 中马格德堡大学的一个研发团队正在进行可解释人工智能技术的研究，他们采用神经科学的方法打破黑盒子，在跟贴 530 中，发表了可解释机器学习技术，他们打破黑盒子，克服机器学习模型不可解释缺陷，创建可解释机器学习新模型。

总的来说，可解释的机器学习/深度学习技术已开始突破，可解释的机器学习/深度学习新模型正在建立。

### （二），从异步脉冲神经网络出发

英特尔最近发布了神经拟态计算的最新进展，这是由异步脉冲神经网络出发，实行类脑计算的重大突破。

脉冲神经网络（SNNs）是神经拟态计算中一种全新的模型，以其 Loihi 芯片为基础的神经拟态计算系统 PohoikiSprings 包含 1 亿神经元。“神经拟态计算”或“类脑计算”指的是模拟人脑神经机制和运行方式有关计算。

神经拟态计算可以模仿人脑中自然神经元网络的方式将计算模块重新分布。颠覆传统冯-诺伊曼的硬件和软件架构，实现人脑的智能功能。神经拟态计算芯

片体积小、功耗低、符合生物进化最本质优势。神经拟态计算在算法和芯片设计上可实现以低一千倍以下的功耗去完成同样效果的模型训练。

神经拟态计算的优势是功耗低、神经元的智能性和自主性（最大优势），不是单纯解决一个数据训练、模式识别问题，而是解决多模态感知、非结构化信息白的感知和推理。

神经拟态计算被认为引领下一代人工智能的主流计算模式。

神经拟态计算还发于发展的初始阶段，当前的重点是抓应用突破、应用落地（为此英特尔成立开源的神经拟态研究社区 INRC）。神经拟态计算未来前景的切入点：①非结构化数据、实时性要求高的场景，②多模态、实时的场景（如机器人、无人机和需要持续学习、自适应学习的场景）。

（三），从知识工程、大规模语义网络出发

如何通过新知识工程实现可解释的人工智能？

目前通过知识工程尚难实现可解释的人工智能，还差最后一公里。最后一公里的短板是大规模语义网络（知识图谱）能力不足（图灵奖得主、人工智能大师 YoshuaBengio 就指出 IBM Watson 的大规模语义网络能力不足），需要进一步提升其中的理解和解释两个核心能力，另外要得决大规模语义网络尚不能概括进来的常识等推理攻关问题。

## 国内外跟贴留言：486-535条

486，谷歌的中国工程师团队提出颠覆性的自动驾驶行为预测算法模型。

在复杂的交通场景中，自动驾驶面对周围的车辆和行人，如果能在其行进中在未来 5 秒内提高行为预测，必将加速自动驾驶的发展。

行为预测的难点在于周围行人、车辆的不确定性和各种规则之外的行为，难以进行确定性预测。其难点尤其在遇到盲点与遮挡时，更是如此。

对于无人车，周围的环境大致分两类：一是地图特征，其中包括车道线、斑马线、红绿灯、速度标示、停车指示牌等固有道路要素，二是无人车周围的物体运动轨迹。

传统的自动驾驶行为预测做法的缺点是：①将物体渲染到图片上，从时空看这是一个缺乏效率的方式，②采用卷积神经网络，这在自动驾驶行为预测方面存在局限性，因为它不适合应对长距离的道路信息。谷歌的一个中国工程师团队，在 Waymo 协助下，提出了一个全新自动驾驶行为预测模型 VectorNet。

他们提出了一种抽象化认识周围环境信息的做法：用向量（Vector）来简化地表达地图信息和移动物体，这一做法抛弃了传统的用图片渲染的方式，达到了降低数据量、计算量的效果，即 VectorNet 模型的做法，无需将环境信息渲染成图片，而表达成抽象、简化的向量形式（模型计算速度比卷积神经快一个数量级）。在向量化的基础上，该模型在所有向量之间添加了语义关系，让自动驾驶车辆不仅能看到环境信息，更能进一步理解环境中不同要素之间的关系。在自动驾驶的语境下，对要素之间的关系的认识可以帮助进行行为预测。所以 VectorNet 在实际应用中具备很强实用性，可提升自动驾驶行为测试精准度。经 Waymo 实测，该模型技术提高了行为预测的精准度，比现有方法提升了近 20%，而在占用内存和计算量上则减少了约 8 成。

目前该论文已被计算机视觉领域三大国际顶级会议之一的 CVPR 接收。

487, 优选脉冲神经网络 (SNN) 的不同训练方法以改善其特性 (计算效率、准确性、推理延时)。

脉冲神经网络 (SNN), 通常被称为第三代神经网络, 它使用稀疏脉冲事件在时间上对输入信息进行编码, 可以利用它们来实现认知任务的更高计算效率。但考虑到最新的模拟神经网络 (ANN) 所带来的准确性的飞跃, SNN 训练算法的成熟度要低得多, 从而导致了 SNN 和 ANN 之间的准确性差距。论文提出了不同的 SNN 训练方法, 其生物保真度不相同, 并评估了它们在复杂图像识别数据集上的功效。首先, 论文针对 SNN 中无监督的表示学习提出基于生物学上合理的峰值定时依赖可塑性 (STDP) 的确定性和随机算法。论文中对 CIFAR-10 数据集的分析表明, 基于 STDP 的学习规则使卷积层可以使用更少的训练示例来自学习底层输入特征。但是基于 STDP 的学习在适用浅层 SNN ( $\leq 4$  层) 时会受到论文提出了一种转换方法, 以将现成的经过训练的 ANN 映射到 SNN, 以实现节能推断。论文在 ImageNet 上展示了 VGG16-SNN 的 69.96% 的准确性。但是, ANN 到 SNN 的转换会导致较高的推理延迟, 以实现最佳精度。为了使推理延迟最小化, 论文提出了基于峰值的误差反向传播算法, 该算法使用脉冲神经元的可微近似。论文在 CIFAR-10 上进行的初步实验表明, 与转换后的 SNN 相比, 基于脉冲的错误反向传播可以有效地捕获时间统计信息, 从而将推理延迟降低多达 8 倍。

本论文发表在《IEEE》期刊上

<https://ieeexplore.ieee.org/abstract/document/9053914>

488, 麻省理工开发基于人工智能的快速筛选方法用于开发新材料及应用于复杂材料分析。

美国麻省理工学院研究人员开发了一种快速筛选系统, 以测试数十亿种潜在材料的抗断裂性。对于开发新材料或保护涂层的工程师而言, 有数十亿种不同的可能性可供选择。在实验室中, 如果使用计算机模拟确定材料的确切属性, 验

证过程可能要花费数小时、数天甚至更多的时间。现在麻省理工学院开发的一种基于人工智能的方法可以将测试时间减少到几毫秒，从而可以快速筛选候选材料。这项工作的重点是分析裂纹在材料分子结构中的传播方式来预测材料的破裂或断裂方式。记录了有序晶体结构的相对取向和材料各自对断裂的反应方式，然后逐个模拟材料如何破裂。最后将模拟的大量数据输入到 AI 系统中，训练可以发现潜在物理原理并预测不属于训练集的新材料性能模型，对于分子动力学的单个模拟，运行模拟需要花费几个小时，但是在此人工智能预测中，只需花 10 毫秒即可完成所有模式预测，还能显示裂纹是如何一步一步形成的。研究人员认为，该开发方法具有通用性，AI 模型可以应用于更为复杂的材料分析。

据《MIT-News》报导

489, 用脉冲神经网络解决一类约束非 Lipschitz 的优化问题。本文针对一类约束非 Lipschitz 优化问题，提出了一个脉冲神经网络 (SNN)，其中目标函数是一个非光滑、非凸函数和一个非 Lipschitz 函数的和，其可行集为一个闭合凸集。使用平滑近似技术，通过微分方程，对提出的神经网络进行建模，可以轻松实现。在可行集中目标函数有界条件下，我们证明了可行集中任何初始点的 SNN 解的全局存在性和一致有界性。在平滑函数的 Lipschitz 属性下提供了 SNN 解决方案的唯一性。我们表明，SNN 解的任何累积点都是优化问题的固定点。数值结果包括图像复原，盲源分离，变量选择和最小化条件数，以说明理论结果并显示 SNN 的效率。

490, 研发基于深度强化学习新算法对交互式医疗图像分割动态过程建模、决策、求解。来自上海交大和华东师大的团队研发了一个基于深度强化学习的新算法 Itek-MRL，将交互式医疗图像分割的动态过程建模成一个马尔科夫决策过程，

然后用深度强化学习求解。该算法从整体上考虑分割更新序列，充分挖掘了交互分割前后的关联。由于体素级的图像分割标注任务造成单智能体的强化学习算法遭遇探索空间爆炸问题，因此该团队采用多智能体模型，将每个图像体素看作一个智能体。通过让所有智能体共享同一个行为策略，将探索空间减小到了可行的范围。多智能体模型还能同时实现图像体素的合作交互，从而有效地捕捉分割任务中体素之间的依赖性。

491，谷歌宣布召开第七届精细视觉分类研讨会。为了帮助支持细粒度的视觉分类这一领域的进一步发展，谷歌赞助并共同组织了第七届细粒度视觉分类研讨会（FGVC7），与计算机视觉和模式识别（CVPR）关联，该研讨会将作为虚拟聚会于 2020 年 6 月 19 日举行。谷歌强调了今年的世界级细粒度挑战赛内容，包括预测果树病到预测时尚特质。谷歌邀请了来自世界各地的计算机视觉研究人员参加此次研究会。

492，在计算机视觉领域应对尺度变化的目标检测新算法。目标检测是计算机视觉领域的重要研究方向。为了更好地解决物体检测中的尺度问题，近日商汤 EIG 算法中台团队提出应对尺度变化的目标检测新算法。

特征金字塔是解决物体尺度问题的重要手段，但是不同 level 的特征图其实存在较大的语义差距。论文提出轻量级金字塔卷积 PConv，对特征金字塔进行 3-D 卷积来挖掘内在尺度的关联性。同时，论文提出尺度均衡的金字塔卷积 SEPC，以减少特征金字塔与高斯金字塔间的差异。这一改进提高了检测器应对尺度变化的鲁棒性，提升了 SOTA Single-Stage 目标检测算法的性能，而且几乎不影响推理速度。



493, 采用深层聚合算子 (DLA) 处理不同尺度目标解决多目标跟踪任务。华中科技大学和微软亚洲研究院的研究人员将目标检测和重识别融合进一个框架, 帮助解决多目标跟踪任务, 同时提出了一个简单而有效的新方法。该方法以 30fps 运行在公开数据集上的性能超越了之前 SOTA 结果。采用 anchor-free 目标检测方法, 估计高分辨率特征图上的目标中心。去掉锚点这一操作可以缓解歧义问题, 使用高分辨率特征图可以帮助 Re-ID 特征与目标中心更好地对齐。然后, 添加并行分支来估计像素级 Re-ID 特征, 这类特征用于预测目标的 id。具体而言, 学习既能减少计算时间又能提升特征匹配稳健性的低维 Re-ID 特征。在这一步中, 研究者用深层聚合算子 (DeepLayerAggregation, DLA) 来武装主干网络 ResNet-34, 从而融合来自多个层的特征, 处理不同尺度的目标。

494, 深度学习在图上的发展——图神经网络 (GNN)。不久前在 ICLR2020 大会上, Deepmind 的研究人员 Petar Veličković 报告了用于算法推理的图表示学习。由于图的不规则性, 传统的深度学习算法在处理复杂的图时面临着巨大的挑战。图表示学习是深度学习的一个方向, 是一种无监督学习, 其核心是研究图数据的表示, 图神经网络 (GNN) 给图表示学习带来了新的建模方法。

为什么要通过图表示学习——图神经网络 (GNN) 来发展深度学习算法用以外理图问题? 因为很多数据是图结构 (如社交网络、经济网络、生物医学网络、信息网络、互联网、神经网络等), GNN 一一也适用于这些网络。

十年来研究人员一直为克服 GNN 模型表现力不足致力于提升其表现力的研究, 近来本文研究者开发了多任务基准, 以确保 GNN 能同时理解多种特性, Petar 表示, 这是解决复杂图形问题的基础。

495, 自动驾驶行为预测算法模型。当前, 自动驾驶技术取得了很大进展, 但欲在 L5 等级路况下行驶, 乃至在全路况条件下行驶, 还是未定之数!

今天，全球自动驾驶技术的两家领先的企业：Waymo 和百度拟采取两种不同的发展思路。

#### ①Waymo 在谷歌帮助下提出颠覆性的自动驾驶行为预测算法模型

他们提出一个全新的 VectorNet 模型，该模型用向量来简化表达地图信息和移动物体（这是一种抽象化认识周围环境信息的做法），在所有向量间添加了语义关系，让自动驾驶车辆不仅能看到环境信息，更能进一步理解环境中不同要素间的关系，在自动驾驶语境下，对要素间的认识可帮助行为预测并大大提高预测精确度。该论文已被计算机视觉领域三大国际顶会之一的 CVPR 接收。

#### ②百度尝试采取另一种思路，被称为“中国方案”。

该方案不再训练自动驾驶汽车在现有城市环境中进行导航，而是对现有城市或街道路侧进行数字化改造，以适应并促进自动驾驶技术的发展。

#### 496，获物流试运行许可这个“空中飞的”厉害了！

据悉，亿航智能（一家智能自动驾驶飞行器科技企业）正式对外宣布，已获得中国民用航空局颁颁发的全球首个自动驾驶飞行器（AAV）亿航 216 的物流试运行许可，意味着亿航智能可率先开展 150 公斤以上大载重空中物流商业试运行，这个“空中飞的”厉害了！

#### 497，百度刷新人脸伪造检测数据集识别“换脸”诡计。

近日，百度在权威 DeepFake 防伪数据集 FaceForensicsBenchmark (FFB) 上刷新记录，以整体检测准确率 0.821 的成绩超越 SOTA，领跑行业。

FFB 是由德国慕尼黑工业大学 (TUM) 联合谷歌等共同发布的大规模人脸伪造数据集。FFB 拥有四种先进的人脸伪造技术：Face2Face、FaceSwap、DeepFakes 和 NeuralTextures。这些伪造出来的人脸图像自然逼真，人眼几乎无法辨别真伪，极大地提高了任务难度。

百度将自研的 LearningGeneralizedSpoofCue 算法引入人脸合成图像鉴别任务，提高了整体检测准确率，并登顶 FFB 榜单。

随着 DeepFake、FaceSwap 等人脸编辑及生成技术的发展，极大的推动了如虚拟主播、人脸合成等新兴娱乐文化的应用，但同时也给人脸安全带来了潜在威胁，例如通过“换脸”技术进行色情视频生成、虚拟政治人物讲话等，对社会造成严重不良影响。目前 AI 换脸已形成一个黑产业链：上游提供软件技术，中游提供视频、照片定制，下游售卖成品视频（或做全产业链的生意），已造成严重危害。此外，在一些使用人脸注册识别场景下，可能存在不法分子通过上传 AI 合成人脸图像以企图蒙混过关。一系列“换脸”技术合成的人脸图像很难通过肉眼辨别真伪。

498，自然语言处理必须解决以下问题：

- ① 言上歧义，
- ② 语言的鲁棒性，
- ③ 知识依赖问题（语义取决于一些背景知识），
- ④ 上下文。

499，ToP 深度学习论文。

在 GitHub 上有一个关于深度学习论文阅读路线图的存储库，包含 100 篇 ToP 深度学习论文，涉及自然语言处理、机器人、语义图像分割等多个领域。

500，神经拟态计算 VS 深度学习。

三位 AI 专家（杜克大学陈怡然、浙江大学唐华锦、英特尔宋继强）谈基于脉冲神经网络的神经拟态计算与深度学习的关系

宋：神经拟态计算和深度学习的关系是兼容并蓄不是取代。对深度学习已有的擅长，如模拟人类视觉或自然语言交互的任务，让深度学习的网络去模拟。在其他方面，如英特尔的 Loihi 芯片做了嗅觉方面的工作，还有机器人操控，多模态甚至跨模态之间的知识存储，可用神经拟态计算去实现。

唐：在一些特殊场景中，如并不需要精确的计算结果，而需在一个实时环境中给出一个鲁棒响应时，神经拟态计算有绝对优势。

陈：目前两者实现的功能没有特别大的不同。神经拟态计算具有鲁棒性及实时性优势，这些只是在深度学习上提升而不是技术上突破。

宋：英特尔成立神经拟态研究社区（INRC），这是开源开放的，要在应用方面突破，抓应用落地。

神经拟态计算未来前景的切入点：

①一个是非结构化数据、实时性要求高的场景，②多模态、实时的场景（如机器人、无人机、需要持续学习自适应学习的场景）。

## 501，深度学习和推理

如何让深度学习模型鲁棒运行和推理？

为何说采用 MegEngine “训练推理一体化”的范式后，可为深度学习引入推理创造条件？

如果深度学习实现推理后，能否改变其黑盒子操作的不可解释为可？

## 502，国内外跟贴：8个脑机接口案例；

①跟贴 12 条，2019 年 8 月发布，美国卡内基梅隆大学脑机接口案例，

②跟贴 51 条，2019 年发布，俄罗斯“脑机接口”公司（Neurobotics）和莫斯科物理技术学院（MIPT）脑机接口案例，

③52 条跟贴，2019 年 7 月发布，美国脸书和加州大学旧金山分校（UCSF）脑机接口案例，

④175 条跟贴，2019 年发布，美国特斯拉“脑机接口”公司脑机接口案例，

⑤308 条跟贴，2020 年 1 月发布，中国浙江大学求是高等研究院脑机接口案例，  
⑥363 条跟贴，2019 年发布，美国加州大学伯克利分校脑机接口案例，  
⑦378 条跟贴，2019 年发布，美国加州大学旧金山分校脑机接口案例，  
⑧386 条跟贴，2020 年 1 月发布，中国天津大学脑科学与类脑研究中心脑机接口案例，  
2016 年，对宇航员进行无创脑机接口测试。

### 503，忆阻性神经混合芯片加速推动人工智能技术发展

近来全球开展忆阻性神经混合芯片的研发十分火红。

来自俄罗斯的一则报导：俄罗斯罗巴切夫斯基州立大学与多国科学家合作，提出了一种“忆阻性神经混合芯片”（Memristive Neuro hybrid chip）的概念，芯片可用于紧凑型生物传感器和神经假体，该概念是基于神经细胞和微流体技术组合的现有和前瞻性解决方案，使“空间有序活动神经网络”的植入成为可能。

其实研发忆阻器技术在欧美早已开始了！忆阻器芯片具有非线性电阻记忆的独特性，在模拟信号处理系统具有广阔前景，此外还可用作电生理活动传感器，发挥信息积累和非易失性储存功能。

在清华大学自主研发的第二枚类脑芯片（发表在《自然》杂志上）也显示了忆阻器技术。今年 2 月 27 日由钱鹤、吴华强教授领衔的团队研发出一款基于多阵列忆阻器存算一体系统，在处理或训练卷积神经网络时能效优于 GPU（以前沿用的图形处理器芯片）高两个数量级。

### 504，DeepMind 最新力作：分布式强化学习框架 Acme，智能体并行性加强

近年来，在深度学习技术和算力提升的双重加持下，强化学习已经在众多复杂的 AI 挑战中取得了辉煌战绩。无论是象棋、围棋，还是王者荣耀，以及各类经典游戏，强化学习的表现都是足以令人叹服。

但深度强化学习在带来开创性进展同时也带来一些挑战：这些进步常常以底层强化学习算法的规模及复杂性为代价，并使得已公开的强化学习算法或 idea 变得难以实现。

为了解决强化学习算法由单进程原型到分布式系统扩展过程中智能体的重新部署问题，DeepMind 推动了一种新的分布式强化学习框架“Acme”。

505，神经调节算法——受人脑神经调节机制启发的人工智能新颖算法。

神经科学领域所发现的认知机理首次在多任务环境下发现了新的算法应用。

为利用人工智能更好地模拟人脑神经调节功能机制的关键打开了新视野。

尽管近年来人工智能领域取得巨大进步，但离人类智能还很远。当前的人工智能技术允许训练计算机智能体可以专门有针对性地对它们进行人工训练，以更好地执行某项特定任务。

人类通过使用一生中获得的技能，能够非常有效地适应新情况。例如一个学会在客厅走路的孩子也会很快学会在花园里走路。因为所学习得到的走路技能与大脑突触的可塑性有关，突触可塑性改变了神经元之间的联系。在客厅学习行走技能与在花园中行走所需快速适应技能则与神经调节过程相关。神经调节通过化学神经调节介质调整神经元本身的输入——输出特性：

模拟人脑神经突触的可塑性是当今人工智能所有最新进展的基础。

然而到目前为止，还没有科学研究成果提出将神经调节机制引入到人工神经网络的方法。

为了攻克这个难题，一个由人工智能专家和神经科学家组成的团队开发了一种基于人脑功能机制的新算法，称为神经调节算法。这个新颖而卓越的算法可以创建能够执行训练期间未遇到的任务的智能体，使智能主体能够自动适应未知情况。

在神经科学中，神经调节指一种神经传导过程。在此过程中，一个特定神经元使用一个或多个神经传导物质来控制一系列神经元。被一小群神经元覆盖住的神经传递介质会在神经系统中大范围地被释放出来，进而影响到许多的神经元。

506，日本发明的人工智能机器人系统的确令人赞叹。

众所周知，日本曾推出美女机器人（几经修改），能够聊天，富有表情，其外形犹如东方美女，其动作与特征与人类相差无几！

最近日本东京大学又发明了一款机器人，名为 Tengoro，模仿人类动作，其“肌肉”比人类结实、柔软（采用 16 个驱动器操纵各种动作，像人类肌肉那样收缩），其灵活性打破世界记录（实测时比人类灵活 6 倍），运动时还会流汗（采用 108 个电池，供流汗表情需要，还供机器人以能量）。日刊报导“震撼世界”！

507，清华大学与京东数科同台共秀仿生手。

AI 机器人如何实现“心灵手巧”。

6 月 4 日京东数科 AI 产业仿生手研发团队与清华大学人工智能研究院智能机器人研究中心同台共秀仿生手。京东数科介绍其自主研究成果（可穿戴 AI 仿生手）可在 0.5 秒内快速识别和响应肌电信号，准确率接近 100%，基于此助力特种人群劳动效率提升 50%。

清华孙富春主任说，灵巧手蓬勃发展主要得益于触觉传感器、人工智能的助推。从迟钝的机械手进化为像人手一样灵巧且感知外界，并能通过脑机接口将感知到的事物输送到大脑，依照大脑反馈的指令执行。他提到未来发展共融的概念，即机器人应实现与人、环境的共融，包括与人的协作。机器人的手也应像人手一样，能够通过肌肉纤维进行控制骨头模式。

## 508, 华米 AI 医疗提高心脏健康监测水平

近来, 华米研发 BioTrackerPPG 高精度生物追踪光学传感器, 发布 RealBeatsAI 生物数据引擎算法, 实现了 PPG 心律不齐 (含房颤) 本地实时甄别、ECG 心律不齐 (含房颤) 本地实时甄别等功能, 大幅度提高了心脏健康监测的精确率和效率。

509, 2019 年中、美、日三国顶级计算机科学家数据一家美国机构统计, 美国 680 人 (占 46.3%), 中国 480 人 (占 32.7%), 日本 310 人 (占 20.4%)。

据一家日本机构统计, 美国 632 人 (占 44.3%), 中国 500 人 (占 35.1%), 日本 294 人 (占 20.6%)。

美日两国机构的统计数据差不多。

## 510, 减少训练人工智能巨大能源消耗。

MIT 减少训练人工智能神经网络巨大的能源消耗 (释放 60 多万磅二氧化碳) 将其降到 1/1300。

为了释放和提高人工智能算力, 深度神经网络模型需要训练。大规模训练运行 AI 需要把海量的计算机服务器集中起来消耗巨大的电力来支持程序运转, 这通常会排放大量的二氧化碳。

MIT 研究员 Amherst 研究发现 (2019.6), 训练运行一款自然语言处理 AI 模型可以释放超过 62.6 万磅二氧化碳。

MIT 研发团队在 4 月份发表的论文中提出一种训练运行神经网络的新方法, 该方法在调试运行神经网络时释放的二氧化碳仅为当前所用方法释放的 1/1300。神经网络是模仿人脑功能的一套算法。

国际数据公司 2018 年度发布的白皮书预计, 到 2025 年全球数据量将达到 175zb, 这些数据都需要处理, 而处理它们必然要耗费能源。



MIT 研发团队还提出 13 种利用 AI 帮助人类适应气候变化并减轻不利影响的方法，包括 AI 能帮助人类预测可再生能源的供求关系，规划大规模的碳排放项目。

#### 511，哈佛大学微型机器人 HAMR—Jr

最近在国际机器人与自动化会议（ICRA2020）上隆重登场。

哈佛大学约翰-保尔森工程与应用科学院研制的哈佛动态微型机器人 HAMR—Jr 打造出迄今为止最小（重量 0.32g）、速度最快（每秒可运行 14 个体长）的机器人，可以奔跑、跳跃、携带重物（背负 3.5g）和快速转弯，是最小的“四足机器人”，在抢救、探测及医疗手术等领域具有无限潜力。

#### 512，人工智能垃圾分拣机器人在北京迎来国内首次落地。

不知疲倦的人工智能捷米机器人进行混合可回收垃圾精准分类，把一线工人从肮脏枯燥的工作中解放出来。

现场工作人员通过大数据对机器人进行训练，让机器人具有类似人脑的判断能力，从而对废弃物进行快速准确识别和分类。

机器人根据客户要求将不同物品不同材质的废弃物分成 PET、HDPE、PP、玻璃、利乐包、易拉罐、衣服、鞋子、纸张等 20 多种可回收物。

现场的智能分拣机器人还可以对传送带上密集的废纸和其他二维物体进行分拣，如挑出废纸中塑料袋、塑料膜等杂物，大大提高了废纸纯度。人工智能垃圾分拣机器人每分钟可以分拣 70 多件物品，速度约为普通人的 2 倍，还可 7 天 X24 小时连续工作。未来废弃物回收工作完全可实现无人化。

513，狗能嗅出新冠患者气味。据香港《南华早报》6 月 6 日报导，法国阿尔福国立兽医学院多米尼克-格朗让教授的研发团队称：狗能嗅出感染了新冠病毒患者的气味，成功率在 83%~100%之间。另据美国《科学》周刊报导，狗的鼻子比

人类灵敏上亿倍。此前研究表明，狗能嗅出多种疾病气味，包括糖尿病和某些癌症。

514，李开复谈人工智能。

在创新工场人工智能训练营 DeeCamp 开营仪式上讲话（摘录）

这一次 DeeCamp200 余名来自 MIT、CMM、清华、北大、北航、中科院大学等同学线上汇集，远程学习协作，以后在医疗、教育、自动驾驶等赛道做出可落地的 AI 项目 Demo。

AI 变化迅速，如何进入世界创造价值？如何找到创业机会？

他谈到中国如何崛起成为 AI 超级大国？

中国 AI 崛起原因有 6 点：

①大批优秀的年轻 AI 工程师②坚韧的企业家③产品创新急起直追④高度互联网化产生海量数据⑤中国 AI 投资全球领先⑥有利技术发展政策

AI 是在互联网应用上打响第一炮，最牛的互联网公司都在中美，这些互联网有海量数据可做 AI，中国创业者更拼搏（美国创业者更绅士）。

中国创业者借鉴海外模式到本土创新，现在产品走向世界：如快手、拼多多、VIPkid、今日头条、蚂蚁金服等。快手从四、五线城市走出来，在美国不可能出现。

在自动驾驶方面，美国策略是不改变道路，中国则为无人驾驶打造新的公路，新基建的赋能提升安全度，保证安全度足够让无人车上路，进而累计新的数据，让整个技术迭代的链条滚动下去。中美不同政策：美国不改变基础设施的无人驾驶政策，中国为无人驾驶打造新城市、新公路。

对于未来 AI 各领域，中国公司有相当大的发展空间，现在是 AI 落地的时代，AI 已从少数精英“发明期”走向遍地开花“应用期”。这一变革的临界点是深度学习，深度学习打破了从“不可用”到“可用”的界限，把 AI 推进到应用期

（注：这点业界有不同识，AI 基础理论与 AI 应用落地同时并举，在抓应用落地时要防止炒作泡沫化）。

李谈：此后的 Transformer、强化学习、迁移学习、GAN 等技术虽有很大贡献但不会像深度学习一样有如此革命性进展。

AI 发明期带来了红利，创造了很多成功的创业家，如旷视的印奇、地平线的余凯等，在 AI 应用期，很难期待只凭一个新技术就能创造一个旷视商汤，更要看具体的商业价值，在应用期 AI 最大的机会就是帮助传统行业提升，对 AI 落地，如何选择商业模式很重要。

他说，学者和创业者思维模式不同，做学术界的科学家，像 Geoffrey Hinton 那样，做科学型的企业创始人，除科研外还要进行企业/行业探索、市场开拓、把技术产品化、市场化，像普林斯顿李凯教授那样，在创业成功后再回校研究，也可像李飞飞那样立足于学术界做兼职参与谷歌的 AI。

他举例：2011 年成立旷视，处在 AI 发明期，聚拔了一大群清华姚班人才，从游戏开始，花三年探索商业模式。

到 2017 年成立创新工场子公司创新奇智，直接开始做商业落地，在创始团队 6 位高管中，4 位是商业背景出身，2 位是技术背景出身，这家公司在创立第三年收入已超过 8 亿元（成为创新工场投出的第 6 家独角兽）。

他说，我们也有一些不成功的例子：有一批医疗影像读片公司，创始人都是技术出身，他们不懂医院业务，找不到商业模式。还创办了一家教育公司做了非常复杂的实体 AI 机器人，机器人老师可在黑板上写字，可与孩子们交流，但机器人老师缺乏市场需求，还不如直接做一个卡通形象老师借屏幕和孩子们交流。李强调 AI 赋能产业定制化。

515, AI 视觉全球挑战赛支付宝今年夺冠。

由全球最大机器学习平台 Kaggle 承办的 CVPR2020FGVC (Fine-Grained Visual Categorization, 细粒度图像分类) 挑战赛, 支付宝从 1317 支参赛队伍中胜出夺冠!

由于大量信息干扰的存在, 细粒度图像分类是计算机视觉领域极其挑战的方向, 在实际应用中对于提升交易风险辨识度有极高帮助。

安全科技是支付宝优势, 目前已沉淀一套深度数据挖掘、机器学习算法应用及 AI 自动建模经验, 还具备自动学习能力。支付宝自研智能风控引擎 AlphaRisk 是全球最先进的风控系统之一, 在其保护下, 交易资损率不到 0.00001% (远低于行业平均水平), 交易风险率低至千分之 0.64, 领先国际同行 Paypal 2.3 万倍。

516, 人工智能顶级研究者在哪里?

人工智能正在成为国际竞争新焦点。

据美国保尔森基金会 (Paulson Institute) 下属的麦克罗波洛智库 (Macropolo) 公布的一项名为“全球人工智能人才追踪”的调查报告: 美国聘用了全球 60% 的顶级研究人员, 中国占 10.6%, 欧洲占 10.2%。

值得注意的是这些人才的来源:

美国本土研究者只占 31%, 27% 的 AI 研究者来自中国 (这是 AI 顶级研究人才最大来源国), 来自欧洲占 11%, 来自印度占 11%, 来自伊朗占 4%, 来自加拿大占 3%, 其他占 13%。

即: 美国人工智能领域顶级研究者中有 1/3 来自中国, 全球顶尖人工智能研究者超过一半在美国研究。

为了更加精准地评估顶级人工智能学者的流动趋势, 麦克罗波洛智库选择了人工智能领域的顶级会议 NeurIPS, 该会议主要关注神经网络和深度学习方面的理论进展, 这两个领域被视为推动人工智能最新发展的重要领域。

在 2019 年年底的 NeurIPS 大会上，共有 15920 位研究者提交了 6614 篇论文，最终接收率为 21.6%。从 1428 篇被接收的论文中抽取 175 篇论文的 675 名作者的信息，结果显示，有近 1/3 的研究者来自中国（在中国完成大学本科阶段的学习），同时发现这些人大多数在美国生活、为美国公司和大学工作。

517，神经科学打开了人工智能的黑盒子。

马格德堡大学（Otto von Guericke University）人工智能实验室 Sebastian Stober 团队研发在认知神经科学启发下的可解释人工智能技术的研究项目（CogXAI），即将运用认知神经科学的方法来分析人工神经网络，从而更好地理解它们的工作方式。

他们认为，人工神经网络 ANNs 是一种受自然大脑结构启发的自学习智能系统，它就像生物神经系统一样，能够通过实例学习来独立解决复杂的问题。

在人类大脑中，这些网络是数以百万计的神经细胞通过化学信号和电信号相互交流组成的，而人工神经网络可以被理解为计算机程序，但由于其强大的学习能力和灵活性，近年来人工神经网络在“深度学习”下已成为智能系统开发的热门选择。

科学家通过大脑研究对人类大脑的学习行为有重要发现，而人工神经可以借鉴用来获得快速有效的学习行为，并研究 ANN 的自学习系统如何做出预测，和/或为什么会出错。

Stober 说，对自然大脑研究已经超过 50 年了，然而目前这种潜力很少被用于 AI 架构的开发。通过将神经科学的方法转移到人工神经网络的研究中，他们的学习过程也将变得更加透明和容易理解。通过这种方式，在学习过程的早期阶段就可以识别出人工神经元的故障，并在训练中加以纠正。

人工神经网络的正在迅速发展，通过使用高性能计算机，越来越多的人工神经元可以用于学习。然而这些网络日益复杂，甚至连专家也难以理解它们的内部

流程和决策。如果我们希望未来能够安全使用人工智能就必须全面了解它的工作原理。

518, 俄美中的自动驾驶技术; 自动驾驶技术日趋进步与成熟。不同级别的无人驾驶汽车有什么区别?

根据美国高速交通安全管理局 (NHTSA) 规定, 无人驾驶技术共分 6 个等级, 分别是 0—5。L0-L1 级完全需要人类驾驶员, L2-L3 级可实现半自动无人驾驶, L4-L5 级可实现全自动驾驶, 区别在于 L4 级需要在特定的道路和天气下, 而 L5 级可以适应全地形。

为了实现真正的无人驾驶, 世界多国开始在无人驾驶技术研发、无人车辆制造方面展开布局, 并正在走向实用。

近年来俄罗斯大力发展无人驾驶技术。目前约有 130 辆配备了激光雷达、传感器、摄像头的无人驾驶汽车正在进行道路测试, 除城市交通外, 也期待在远程货运方面发挥作用, 今年 5 月, 俄罗斯卡车企业卡玛斯与俄天然气工业公司借助导航和物体识别设备共同完成一项无人驾驶货车在北极行转 2500 公里的测试。美国硅谷有无人驾驶“五大家族”之称的 Zoox、Waymo、Cruise、ArgoAI、Aurora 在自动驾驶领域展开激烈竞争。

在中国, 百度、腾讯、蔚来、四维图新、小马智行、北京三快等公司推出数十辆无人驾驶汽车, 进入了相应的路测和后续性能完善阶段。

国内无人驾驶技术研发主要围绕两方面展开: 一是单车智能方案, 通过在车上安装多个雷达传感器来实现对周围环境的感知, 二是车路协同方案, 由车辆及与车辆周边环境实现数据交互, 从而完成对周围环境的感知。

2018 年 5 月, 领骏科技发布了国内首个具备“量产产品形态”的 L4 级无人车, 搭载其自主研发的新一代自动驾驶系统, 目标直指自动驾驶的商业化落地量产, 领骏科技还推出了高性能自动驾驶仿真平台。2019 年初, “济南 5G 通信智能网

联汽车测试道路启动活动”在山东省内第一条智能网联汽车测试道路举行，活动现场展出了“全球首台 L4 级无人驾驶电动卡车 HOWO — T5G”、“特定区域低速自动驾驶公交车”两款智能网联产品。

无人驾驶作为伴随 5G、云计算、大数据、人工智能等新一轮技术发展衍生的一项前沿科技，不仅有助于降低物流成本、提升货物运送效率，更可能引发未来商业模式变革。

519, 脸书 Deepfake 检测挑战赛。

6 月 13 日，脸书宣布其首届 Deepfake 检测“以假乱真”挑战赛结果：获胜算法准确率 65.18% 发现非常逼真的 Deepfake 虚假内容。

Deepfake 检测挑战赛旨在寻找能够识别出基于人工智能的优秀的视频算法。

2114 名参赛者提交了 3.5 万个 Deepfake 检测算法参加此次挑战赛，他们需要从约 10 万个短片的数据集中识别假视频。脸书雇佣了 3000 多名演员来制作这些视频，并利用 AI 进行修改，将其他演员的脸粘贴到视频中。

脸书 CTOMikeSchroepfer 认为，这次挑战赛将为研究人员创造一个基准，以指导他们未来的工作。

520, 中国推数字货币 DCEP, 《福布斯》(Forbes) 如何看?

中国人民银行的数字货币 (DCEP) 已到试验阶段，不断引起国内外区块链和金融科技专家广泛关注。

摘录《福布斯》两位撰稿人的评论。

探讨 DCEP 落地将会对中国国内和国际货币格局造成何种影响？数字人民币会否推动国际货币去中心化？是否会掀起新一轮货币革命？电子支付领军企业是否会受到冲击？

DCEP 未来发展前景如何？

传统观点：以电子货币形式出现的人民币不会对国际货币格局造成额外影响。

RogerHuang（资深区块链经济撰稿人）反驳了这一观点，他认为此观点忽视了电子货币独立于经济政治基础的技术特性及其关键的战略意义，DCEP 将使人民币可以在多种货币环境下交互流通（这是包括美元在内的其他货币无法实现的），DCEP 作为一种可在中国、非洲及一带一路各国流通的数字货币，将可能打开大片市场。

Huang 还谈到，数字货币是中国在新技术领域的新探索，DCEP 作为由政府支持的国际流通数字货币，可作为私人组织和各国政府参考的模式，有望推动全球货币数字化发展进程。

DCEP 是否影响支付宝和微信支付发展。

ZennonKapron（金融科技公司创始人）认为，DCEP 只会助益这些市场化应用，不会取代。

目前支付宝和微信支付占中国 90%的数字支付市场，占中国零售支出 20%~25%，并同时经营财富管理和小额贷款等许多其他金融产品和服务。这两个平台可谓是中国数字金融生态系统的基石。而央行没有打算以 DCEP 取代支付宝和微信支付，相反支付宝和微信支付可能会与 DCEP 整合，以巩固他们在数字金融领域的地位。

521，人工智能的走向。得知识者得天下。

认知智能是机器智能化的关键。

通过知识工程使机器实现可解释的人工智能；

(1)，数据、知识双驱动

①数据分类（结构化、非结构化、半结构化数据）——>数据获取——>

②知识表示（基于符号的、基于向量的）——>知识获取——>



①、②——>知识融合（数据、知识双驱动）——>知识推理——>知识图谱（大规模语义网络）——>实现可理解、可解释的人工智能？

通过知识工程让机器像人类那样具有思考或认知能力，即让机器具有对数据、语言、图片、影像和文本的理解、推理、解释、归纳、演绎的能力

(2)，欲实现可解释的人工智能要依靠大规模语义网络。

近来，自然语言处理——>大规模语义网络取得很大进步，但还存在一些机器难以识别的知识，如常识、背景知识、专业知识、专家经验、隐性知识等（这是人工智能在学习训练中不可或缺的）；再者，大规模语义网络还不够成熟、完善，用以支持实现可解释的人工智能其能力尚嫌不足！

522，机器学习在地质研究中再发威力！

马里兰大学地质系 Kim 团队使用一种 Sequencer 的机器学习算法，结合地震学的方法，探索在地核（由熔融的液态铁构成）与地幔边界周围地质运动和变化情况，为进一步解释地球是如何形成的，这个热岩区域是如何随时间推移变化的，以及预测未来的地震。

523，深度学习是有缺陷的

机器学习/深度学习是一种强大的人工智能算法，是当前引领人工智能大发展的引擎，但深度学习也是有缺陷的，黑盒子技术使其训练过程不可解释。

524，说“深度学习有缺陷”需商榷。

说“深度学习是有缺陷的，黑盒子技术使其训练过程不可解释”，这个问题需要商榷。

深度学习是当前引领人工智能大发展的引擎，就像其他复杂的工程系统一样，它目前还是黑盒子，我们还没能理解它为什么能够取得这么好的效果。世上有很多重大发明，发明时都是黑盒子，人们并不明白原理。从指南针发明到理解

为什么会指南，经历 800 多年，……在经历了一段时间后从不理解到理解，这是正常过程，因此“黑盒子、不可解释”，不是深度学习的缺陷，是下一步需要解决的理论问题。

525，深度学习是有缺陷的，它并非是实现人工智能的一条完美的路径

美国一位 AI 专家在《IEEE Spectrum》（2019.4）上撰文质疑 IBM Watson 研发 AI 医疗应用前景时提出：“深度学习也是有缺陷的”，“深度学习是实现人工智能的路径之一，而非一条完美的路径”。“深度学习本质上是一项黑盒子技术，其训练过程具有难以解释、不可控的特点”。

526，深度学习最大的问题是不可解释和不可理解。

我记得张钹老师去年 8 月 14 日的谈话，其中提到：现在深度学习本质上是基于概率统计的学习，而概率统计最大的问题是不可解释和不可理解。

527，当前机器学习理论有局限性完全以盲模型的方式运行。

图灵奖得主、贝叶斯网络之父 Judea pearl 在 2018 年的论文中谈到，当前机器学习理论有局限性，完全以统计学或盲模型（即黑盒子）的方式运行，所以不能成为强 AI 的基础。

528，深度学习算法目前并不完美有待继续加强理论研究。

国外一位 AI 大师论文中谈到，近年来以深度学习算法为代表的 AI 技术快速发展，在计算机视觉、语音识别、语义理解等领域实现了突破，但其算法并不完美，有待继续加强理论研究。

529，打破黑盒子使深度学习转变为可解释。

深度学习本质上是一项黑盒子技术，其训练过程是不可理解、不可解释的。

科学的发展，打破了黑盒子这把锁，将转变深度学习模型原来的不可理解、不可解释为可理解、可解释。

据我们收到一些跟贴反映，有关机构/人士正在潜心研究这项破解黑盒子的理论。

530，如何通过知识工程实现可解释的人工智能。

目前，通过知识工程尚难实现可解释的人工智能（或者说，机器还不能实现像人类那样的认知智能）。

最后一公里的短板在哪里？短板在知识图谱或大规模语义网络。

知识图谱即为一种大规模语义网络。

大规模语义网络是在大数据时代体现新知识工程的核心技术。

自然语言处理（NLP）取得很大进步，人工智能资深专家吴恩达说，2019年是自然语言处理飞跃的一年。

我们来看一下自然语言研究的发展轨迹：自然语言处理（NLP）系统——>自然语言理解系统（具备一定的理解和解释的能力）或语义网络——>大规模语义网络（包括语言建模和训练模式）。

迄今大规模语义网络还不够完善。人工智能大师 YoshuaBeno 认为，NLP 虽然取得较大进步，但与人类的认知能力相差还甚远。让大规模语义网络来支持实现可解释的人工智能其能力尚嫌不足。

在跟贴 84 中，谈到如何完善大规模语义网络，如何提升其中的理解和解释两个核心能力（从现有一定的基础上提升）。

还有不少知识是大规模语义网络所不能概括的，如常识，常识是难以定义、表达、表征的，目前的大规模语义网络尚不包括常识。除常识外，还有背景知识、专业知识、专家经验、隐性知识等，也不能被大规模语义网络所概括。

跟贴 457 主要谈到华盛顿大学叶锦才研发团队关于常识推理攻关研究的进展。

IBM Watson 在人工智能医疗诊断中，提出具人（embodiment）的概念，要求医学科学家与临床医生沟通，取得共识，以此来克服缺乏常识的困难。

531，通用语言模型/人工智能知识增强语义理解框架/预训练模型。

2019 年 NLP 飞跃，研究人员在让机器理解自然语言方面取得飞跃，通过对巨大的、未标记的数据集进行预训练，使得通用语言新模型可以熟练掌握自然语言（通过在专门语料库上进行微调来掌握给定任务），使得语言模型变得精通语言。

谷歌于 2018 年发布 BERT 通用语言模型或预训练模型。

微软于 2019.6 发布 MT - DNN 预训练模型。

OpenAI 于 2019.2 发布 GPT - 2 通用语言模型。

百度于 2019 年发布 ERNIE 人工智能知识增强语义理解框架。

卡内基梅隆大学、脸书均发表了预训练模型。

这些 BERT、GPT 模型均采用 Transfromer 基础架构。在处理机器翻译、文本摘要、语音识别、问答系统等多种语言任务时，只需进行微调，即可达到 SOTA（最优）性能。

发布通用语言模型或预训练模型用于对 NLP 或语义网络进行预测训练，用以对 NLP 或语义网络进行微调，以完善语义网络。

532，谷歌发布最新“天马”模型，自动生成文本摘要已超越人类。

在完善大规模语义网络的路上，谷歌 AI 团队研发出了 PEGASU（天马）模型，极大地优化了语义网络的训练效率，从“间隙句子”到文本摘要，小样本性能更佳。研究人员设定下游任务为“提取文本摘要”，而预训练目标为“生成间隙句子”（Gap sentences Generation）。研究人员在输入文档中删除一些完整句子，天马模型的任务就是恢复它们，整个过程像填空一样，这项看似人类都无

法定成的任务真的实现了。在文本上游的这些数据集十分丰富，包括文章、科学论文、专利、短篇小说、电子邮件、法律文件和使用说明书等，为了确保实验结果的准确性，除采用 ROUGE 标准衡量外，还采用人工评估方法（类似于图灵测试）。

533，打破黑盒子克服机器学习模型不可解释缺陷创建可解释的机器学习新模型  
机器学习/深度学习取得巨大进步，今天已成为引领人工智能大发展的引擎，但机器学习/深度学习是有缺陷的，它的黑盒子特性使其行为背后缺乏透明度，用户很难理解机器学习/深度学习模型是如何做出决策的，其不可理解、不可解释的缺陷限制了它扩大、提升应用。

对待机器学习/深度学习模型的发展有两种态度，一种是“深度学习（技术）已达天花板”，“其致命的不可解释缺陷将使其停止不前”，“人工智能发展只能捨此另觅出路”；另一种是“机器学习/深度学习潜力很大”，“决不放弃改造可解释的机器学习模型”！

目前不少人工智能专家正致力于改造机器学习/深度学习模型，致力于研究可解释的机器学习/深度学习技术，可以预见将为机器学习/深度学习发展带来转机！  
此处介绍一篇论文“可解释机器学习技术”，刊载于 Communications of the ACM, “Techniques for interpretable machine Learning”, 2020, 63 (1) : 68-77。

可解释机器学习使得机器学习模型能够以易于理解的方式向用户解释或呈现其行为，我们称这种特性为可解释性（interpretability）或者解释性（explainability）。

534，语言模型 GPT 跨界 CV:

OpenAI 揭示算力 Transformer 具有通用性。

如何完善大规模语义网络。

生成文字的 AI 和生成图像的 AI 竟然是同一个-人工智能要实现通用 (AGI) 了吗?

语言模型 GPT 现在跨界了。它要进军 CV 领域-以 imageGPT 名义处理 “图像任务” 我们在研究自然语言生成，它却已经能生成图像了，下一步是否要生成音乐-GPT-Music? !

535, 评 “人工智能如何走向新阶段?” -535 条跟贴为你揭密!

自从我们于 2019 年 8 月初提出 “评人工智能如何走向新阶段?” 这个命题以来, 迄今 10 个多月了, 我们陆续收到国内外传递来的跟贴 535 条。

我们开辟了一个讨论的平台 (自由论坛), 期望国内外的专家和草根不拘一格都可以发送跟贴而来: 针对人工智能的发展各抒其见、发表评论 (我们不在乎评浅、评错)。为避免 “某国” 对此敏感议题过度干预, 我们一开始就公开申明: 由国内外传递人工智能发展前沿的信息 (跟贴) 仅限于开源的或公开发表的。

在 535 条跟贴中, 有一些是人工智能大师的言论 (或者是有人转载以跟贴方式传递的)。谈到 “人工智能走向何方?” 大家可从 535 条跟贴中找到答案。所以这批跟贴弥足珍贵!

人工智能走向何方? 在跟贴 447 中谈到, 人工智能未来发展的第一目标是人类智能或接近 (类似) 人类智能。

也就是说, 人工智能的发展任务是实现可解释的人工智能, 或者说是让计算机建立像人脑那样的认知智能。

跟贴 447 还谈到, 目前从三条路出发正在探索人工智能发展目标:

一是从机器学习/深度学习模型出发,

二是从异步脉冲神经网络出发,

三是从新知识工程、大规模语义网络出发。

其实在 535 条跟贴中还有伴随着十几个实例的脑机接口一条路，在我们这里可称为：

四是从脑机接口出发。

（一），从机器学习/深度学习模型出发

众所周知，机器学习/深度学习取得巨大进步，今天已成为引领人工智能大发展的引擎，但机器学习/深度学习是有缺陷的，它的黑盒子特性使其行为背后缺乏透明度，用户很难理解机器学习/深度学习模型是如何做出决策的，其不可理解、不可解释的缺陷限制了它的扩大、提升应用场景，有人甚至说机器学习/深度学习（技术）已达天花板。

在 527 跟贴中，图灵奖得主、贝叶斯网之父 Judea pearl 谈到，当前机器学习理论有局限性，完全以统计学或盲模型（即黑盒子）的方式运行，所以不能成为强 AI 的基础。

在 528 跟贴中也谈到，深度学习算法并不完美，有待继续加强理论研究。

可喜的是有一些跟贴（如 480、517、529、530）正在研发可解释的机器学习模型。在跟贴 517 中马格德堡大学的一个研发团队正在进行可解释人工智能技术的研究，他们采用神经科学的方法打破黑盒子，在跟贴 530 中，发表了可解释机器学习技术，他们打破黑盒子，克服机器学习模型不可解释缺陷，创建可解释机器学习新模型。

总的来说，可解释的机器学习/深度学习技术已开始突破，可解释的机器学习/深度学习新模型正在建立。

（二），从异步脉冲神经网络出发

英特尔最近发布了神经拟态计算的最新进展，这是由异步脉冲神经网络出发，实行类脑计算的重大突破。

脉冲神经网络（SNNs）是神经拟态计算中一种全新的模型，以其 Loihi 芯片为基础的神经拟态计算系统 PohoikiSprings 包含 1 亿神经元。“神经拟态计算”或“类脑计算”指的是模拟人脑神经机制和运行方式有关计算。

神经拟态计算可以模仿人脑中自然神经元网络的方式将计算模块重新分布。颠覆传统冯-诺伊曼的硬件和软件架构，实现人脑的智能功能。神经拟态计算芯片体积小、功耗低、符合生物进化最本质优势。神经拟态计算在算法和芯片设计上可实现以低一千倍以下的功耗去完成同样效果的模型训练。

神经拟态计算的优势是功耗低、神经元的智能性和自主性（最大优势），不是单纯解决一个数据训练、模式识别问题，而是解决多模态感知、非结构化信息白的感知和推理。

神经拟态计算被认为引领下一代人工智能的主流计算模式。

神经拟态计算还处于发展的初始阶段，当前的重点是抓应用突破、应用落地（为此英特尔成立开源的神经拟态研究社区 INRC）。神经拟态计算未来前景的切入点：①非结构化数据、实时性要求高的场景，②多模态、实时的场景（如机器人、无人机和需要持续学习、自适应学习的场景）。

（三），从知识工程、大规模语义网络出发

如何通过新知识工程实现可解释的人工智能？

目前通过知识工程尚难实现可解释的人工智能，还差最后一公里。最后一公里的短板是大规模语义网络（知识图谱）能力不足（图灵奖得主、人工智能大师 YoshuaBengio 就指出 IBM Watson 的大规模语义网络能力不足），需要进一步提升其中的理解和解释两个核心能力，另外要得决大规模语义网络尚不能概括进来的常识等推理攻关问题。