从开普勒到牛顿：如何改进Transformer读懂物理世界

发布于 2026-04-09 13:22:47 151 0 举报

大家好，我是赛博解生酱。

我们总在聊AI的“智能”，以transformer为底座，现在的AI模型已经写上万行严丝合缝的代码，能通过最难的律考和数学；能背下整本《经典力学》，一字不差地默写万有引力公式，能解复杂的天体运动微分方程；然而，一些情况下，AI却连一个小球从斜面滚下的轨迹都预测不准，甚至无法预判“杯子碰歪了会掉在地上摔碎”，“水泼在桌上会顺着桌沿流下来”。

最近，斯坦福大学在Arxiv发表的《From Kepler to Newton》捅破了这层窗户纸：**通用Transformer之所以学不会物理世界的底层规律，不是模型不够大、数据不够多，而是缺了三个极简、却完全贴合物理世界本质的归纳偏置。补上这三个偏置，Transformer就能彻底跳出“曲线拟合”的怪圈，从只会死记轨迹的“数据搬运工”，变成能自主发现物理定律的“AI物理学家”。**

这篇文章，除了论文里的结论，还会掰开揉碎了讲清楚：这三个归纳偏置到底解决了什么问题，为什么之前的Transformer注定学不会牛顿力学，以及这件事，可能会给卡了很久的具身智能、甚至通用 人工智能 ，带来什么样的启发。

开普勒与牛顿：两种智能的本质鸿沟

在技术细节之前，我们先把故事的核心讲透：开普勒和牛顿，到底差在哪？

400多年前，天文学家开普勒拿着老师第谷积累了几十年的行星观测数据，熬了十几年，终于总结出了行星运动三大定律：轨道定律（行星绕太阳走椭圆）、面积定律、周期定律。靠着这三大定律，他能精准预测任何一颗行星在未来任何时刻的位置，误差小到当时的天文仪器都很难测出来。

但直到去世，开普勒都没能回答一个最根本的问题：**为什么行星非要绕着太阳走椭圆，而不是沿着直线飞出太阳系？**

而牛顿，只用一个万有引力公式，就彻底回答了这个问题：

再结合第二定律，他不仅能预测行星的轨迹，更能解释所有天体运动的底层逻辑——行星之所以走椭圆，是因为太阳和行星之间的引力，时时刻刻在改变着行星的运动方向。甚至靠着这套公式，人类在还没通过望远镜看到海王星的时候，就精准算出了它的位置和轨道。

这就是两种智能的本质鸿沟：

> 开普勒的能力，是**拟合现象**：靠着海量历史数据，找到数据里的规律，精准复刻见过的场景，但换个场景就彻底失效；
>
> 牛顿的能力，是**理解本质**：找到支配所有现象的底层、通用、不变的物理规律，哪怕是从未见过的场景，也能靠着规律做出正确的预判。

而此前AI在世界模型上的所有尝试，都困在了“开普勒式智能”里。

2025年Vafa等人的经典实验，已经把这个困境摆到了台面上：他们用GPT-2规模的Transformer，喂了200亿个token的行星轨迹数据，模型的轨迹预测准确率拉到了近乎满分，能画出完美的椭圆轨道。但当研究者打开模型的内部表征，却发现里面完全没有引力、加速度的任何信息——它根本没懂“行星为什么这么转”，只是靠着海量历史数据，拟合出了椭圆的几何形状，一旦给行星的初始速度改一点点，它的预测就会瞬间崩盘，行星要么直接撞向太阳，要么直接飞出太阳系。

这就像一个背熟了题库的学生，原题能秒选答案，只要题目换个数字、改个条件，就直接交白卷。

而斯坦福的这篇论文，就是找到了拦住Transformer从“开普勒”到“牛顿”的三道坎，并用三个极简的归纳偏置，一步一步迈了过去。

三个归纳偏置：让Transformer真正读懂物理世界–

1. 空间平滑性：先给AI一双能看懂“连续空间”的眼睛

Transformer学不会物理世界的第一个致命bug，是**离散token化，彻底打碎了物理空间的连续性。**

行星的位置是二维连续坐标，比如x和y的取值范围是（AU是天文单位，地球到太阳的距离）。但因为Transformer天生是为文本设计的，处理的是离散token，所以研究者做了一个操作：把x和y的取值范围，各自独立分成了7000个均匀的小格子（bin），每个格子对应一个独立的token，每个token都有一个随机初始化的嵌入向量。

举个最直观的例子： x=0.001和x=0.002，在物理世界里是无限接近的两个点，代表行星几乎没动；但如果这两个数刚好落在了两个相邻的格子里，在Transformer眼里，它们就是两个完全独立、毫无关联的token——因为嵌入向量是随机初始化的，在训练之前，这两个向量的相似度，和x=0.001与x=100的向量相似度没有任何区别。

这就像我们教一个孩子认空间，却把一张完整的世界地图，剪成了7000×7000个碎纸片，每个纸片上只写一个数字，然后让孩子靠这些碎纸片，去理解“两个点离得近是什么意思”。

赛博解生
作者相关精选

从开普勒到牛顿：如何改进Transformer读懂物理世界

结果是，哪怕用了200亿token训练，模型的x和y坐标的线性探测，也只有0.86。

这个数字意味着什么？

模型只能捕捉到粗粒度的空间差异，比如行星在第一象限还是第三象限，但是完全丢失了细粒度的空间连续性；

真实世界里一个完美的圆形轨道，在模型的嵌入空间里，变成了四个碎片化的点云，象限之间的全局结构勉强能保留，但每个象限里的局部结构完全扭曲、混乱；

最致命的是，引力的大小和行星到太阳的距离平方成反比，连“距离”都算不准，模型根本不可能学会万有引力定律。

解决方案和关键实验细节

针对这个问题，论文提出了两个可落地的解决方案，同时推导出了空间映射的缩放定律，彻底量化了token化的危害。

第一个方案：缩小词汇量（格子数量）论文里做了一组对照实验，固定训练数据量，把词汇量V从10000降到1000、再降到100，结果发现，词汇量越小，模型学到的空间映射越准。当V从7000降到128时，空间映射的直接从0.86涨到了0.99以上。

同时论文推导出了空间映射质量的缩放定律，公式如下：

这个公式里，D是训练token数量，V是词汇量。它告诉我们一个反直觉的结论：**词汇量对空间映射的影响，比训练数据量更大。**意味着，词汇量翻一倍，训练数据量要翻不止一倍，才能维持空间映射的质量。这也是为什么哪怕用了200亿token，7000的词汇量也让模型学不到完美的空间表征。

当然，词汇量也不能无限缩小，否则格子太粗，坐标的精度就不够了，预测轨迹自然会有误差。所以需要在空间映射质量和预测精度之间，找一个平衡点。

第二个方案：彻底抛弃离散token化，用连续回归替代分类任务这是更根本的解决方案：既然token化会破坏空间连续性，那我们就不用token了，直接把连续的x/y坐标作为模型的输入和输出，把原本的“下一个token预测（分类任务）”，改成“下一个状态预测（回归任务）”。

这个改动的好处是显而易见的：物理空间里越近的点，在输入里天然就越近，模型不需要再从离散的token里，重新学习空间结构，空间平滑性从一开始就被天然保证了。

这一步，直接解决了Transformer“看不懂空间”的核心问题，让模型从“看碎纸片”，变成了“看完整的地图”，为后续学习物理规律，打下了最基础的地基。

2. 空间稳定性：驯服自回归的误差累积，让AI学会在噪声里修正自己

解决了空间平滑性的问题，我们马上就遇到了第二个经典难题：**连续回归的自回归预测，会出现致命的误差累积。**

我们预测行星轨迹，是一个典型的自回归过程：给模型前50个时刻的行星坐标，让它预测第51个点；然后把预测出来的第51个点，当作已知的上下文，再预测第52个点；以此类推，一步步预测出后面的所有轨迹。

这个过程里，只要某一步的预测有一点点误差，这个误差就会被带到下一步的预测里，像滚雪球一样越滚越大。尤其是连续回归任务，模型的输出是无界的，误差一旦出现，就可能无限放大；而之前的离散分类任务，输出只能是7000个格子里的一个，相当于有一个硬约束，哪怕预测错了，也不会出现离谱的数值，天然有一定的“纠错能力”。

这也是为什么之前的研究者普遍认为：“离散分类比连续回归，更适合轨迹预测任务”。

这个问题在实验里的具体表现

论文里做了一组基础实验：用不加任何优化的连续回归模型，以前50个真实坐标为上下文，自回归预测后50个点。结果是：

前3步的预测误差很小，轨迹和真实轨道几乎重合；

从第5步开始，误差快速放大，轨迹开始偏离椭圆；

到第20步的时候，预测的轨迹已经彻底崩盘，行星要么直接冲向坐标原点（太阳），要么直接飞出了的范围，和真实轨道没有任何关系了。

这就像一个新手司机开车，方向盘稍微打偏了一点，就慌了神，越修正越歪，最后直接冲出了马路。

论文里的解决方案：带噪声的上下文训练

针对误差累积，论文里用了一个极简、却极其有效的方法：**在训练的时候，给输入的历史坐标，加入可控的高斯噪声，强迫模型学会在有误差的输入下，依然做出正确的预测。**

我们把这个方法的数学形式写出来，会看得更清楚：

公式里的，是标准高斯噪声，是噪声的强度；

简单说，就是训练的时候，我们故意给每一个历史坐标，都加一点随机的“小抖动”，模拟推理时的预测误差；

模型要想让损失函数最小，就必须学会忽略这些噪声，甚至修正这些误差，依然预测出正确的下一个坐标。

这个方法的本质，是在训练的时候，就提前让模型适应“输入有误差”的场景，学会了误差自修正的能力。就像老司机开车，哪怕路面有坑、方向盘晃了一下，也能轻松修正方向，不会因为一点小意外就崩盘。

关键实验结果

论文里测试了不同噪声强度的效果，结果非常清晰：

当（不加噪声）时，模型预测50步后的平均距离误差，达到了几十AU，完全不可用；

目录

赛博解生
作者相关精选

从开普勒到牛顿：如何改进Transformer读懂物理世界

合轨迹预测”的结论。

这一步，让Transformer拥有了在真实世界里稳定预测的能力——毕竟真实世界里，永远没有完美无噪声的传感器数据，永远有各种意外和扰动，一个不会修正误差的模型，永远没法在真实世界里落地。

3. 时间局部性：最关键的一跃，逼AI从“描轨道”变成“懂引力”

前两个归纳偏置，让AI能画出更完美、更稳定的椭圆轨道，成了更优秀的开普勒，但它依然不懂“行星为什么会这么转”。真正让它蜕变成牛顿的，是第三个归纳偏置——**时间局部性**。

先给大家介绍，什么是牛顿力学里的时间局部性，这是整个问题的核心。

牛顿第二定律，是一个二阶常微分方程，而加速度，是位置对时间的二阶导数：

这个方程有一个极其关键的数学性质：**只要我们知道了某一时刻，行星的位置和速度，就能唯一确定未来任意时刻，行星的运动轨迹**。

用通俗的话讲：你要算一个小球下一秒会飞到哪，只需要知道它“现在在哪”和“现在飞得多快、往哪个方向飞”，完全不需要知道它10秒前、1分钟前、甚至1小时前在哪。

这就是时间局部性的本质：**未来的状态，只由最近的两个瞬时状态决定，和更早的历史毫无关系**。而要确定这两个状态，对应的上下文长度，严格等于2。

但之前的Transformer，用的是全局上下文长度（比如L=100），自注意力机制允许模型关注历史上所有的轨迹点。这就给了模型走捷径的机会：它根本不需要去学什么引力、加速度，只需要用前100个点，拟合出椭圆的几何参数——半长轴、半短轴、椭圆的焦点位置、拉普拉斯–龙格–楞次（LRL）向量，然后用这个椭圆，去“描”出后面的点。

这就是典型的开普勒式工作：先确定轨道的形状，再沿着轨道描点，完全不需要懂轨道背后的物理规律。

论文的核心操作：严格限制注意力的上下文窗口
论文里的解决方案，简单到极致：**把Transformer的注意力窗口，严格限制在最近的2个时间步，模型只能用当前和上一个时刻的坐标，去预测下一个时刻的坐标**。

这一下，就把模型走捷径的路彻底堵死了。它没法再用100个历史点去拟合椭圆的形状，只能用两个瞬时的坐标点，去寻找决定轨迹变化的底层规律。

实验结果：牛顿力学在模型里自发涌现了
论文里用线性探测的方法，去看模型的隐藏状态里，到底有没有学到引力相关的物理量，包括引力的大小、x方向的引力分量、y方向的引力分量，同时也探测了椭圆轨道的几何参数。

结果完全印证了论文的核心猜想：

1. **当上下文长度=100时**：模型对引力相关量的线性探测只有0.9左右，而对椭圆轨道参数的探测达到了0.998。模型完美学会了开普勒的几何方法，却几乎没懂牛顿的力学规律。
2. **当上下文长度=2时**：模型对引力相关量的线性探测直接飙升到了0.999，几乎完美复刻了万有引力公式；而对椭圆轨道参数的探测，降到了0.9左右。模型彻底抛弃了曲线拟合的捷径，自发学会了牛顿力学的底层规律。

更有意思的是，论文里发现了一个清晰的“相变过程”：随着上下文长度从2逐步增加到10、20、50、100，模型的牛顿力场表征精度持续下降，而开普勒轨道参数的表征精度持续上升。上下文长度每增加一点，模型就更偏向“开普勒”一点，更远离“牛顿”一点。

为什么会这样？答案很简单： 当上下文只有2个点的时候，模型唯一能做的，就是计算两个点之间的位置变化，得到速度；再计算速度的变化，得到加速度；而加速度，直接对应着引力。它必须学会，才能准确预测下一个点的位置。

而当上下文足够长的时候，模型总能找到更简单的捷径：用全局的点拟合椭圆，根本不需要费力去学习底层的物理规律。就像我在之前的文章里提到的“深度陷阱”：参数越多、上下文越长的模型，越容易找到死记硬背的捷径，反而不会去学习通用的、底层的结构。

这一步，是最关键的一跃。Transformer终于从“只会描述现象”，变成了“能理解底层规律”，真正读懂了物理世界。

架构的本质：归纳偏置，就是你在流形上定的“游戏规则”

到这里，我们可以回答一个行业里争论了很久的问题：为什么CNN、RNN、Transformer在不同任务上表现天差地别？它们的本质差异到底是什么？

答案就是：**归纳偏置的差异，也就是你给流形上的点云，定义了什么样的连边规则和局部算子**。

在流形篇里写过一句话，在这里依然是核心准则：**我们不是在高维欧氏空间里乱拟合函数，而是在未知流形上，定义一种局部算子和连边规则，然后靠堆叠和学习，得到全局的计算模式。**

我们可以把主流架构的归纳偏置，用几何语言彻底拆解：

1. **CNN：硬编码的空间局部性**卷积的本质，是流形上的固定邻域共享算子。它的归纳偏置是：流形在各处的局部几何近似平稳，同一个滤波器可以在整个空间复用。对应的连边规则，是只和卷积核大小内的近邻点连边，严格限制了局部性。这种硬编码的局部性，让CNN在图像任务上样本效率极高，但当语义流形的局部结构差异很大时，共享权值反而成了约束。

目录

赛博解生
作者相关精选

从开普勒到牛顿：如何改进Transformer读懂物理世界

相似性。但这也是它此前学不会物理世界的根源：它允许任意两个点之间建立强连接，也就是在流形上开了无数个虫洞。在行星轨迹预测里，模型可以直接给100步前的初始点一个高权重，用全局的点云拟合椭圆，根本不需要去学习局部的动力学算子。

而这篇论文的核心贡献，就是给Transformer的自注意力，加上了时间局部性的约束，把它的连边规则，从“全局任意连”，拉回了“只和最近两个时间步连”。这一下，就把Transformer从一个全局曲线拟合器，逼成了一个局部动力学算子的学习者。

读懂物理世界，才是具身智能的真正破局点

讲完了论文的技术细节，也许会有疑问：不就是让AI学会了牛顿力学吗？这东西我们几百年前就知道了，有什么大不了的？

其实，这篇论文的价值，远不止于让AI学会了一个万有引力公式，也给卡了很多年的具身智能，找到了潜在的破局之路。

整个具身智能行业，都困在一个死循环里： 我们能做出能后空翻的波士顿动力机器人，能做出能走能跳的特斯拉Optimus，能在仿真环境里，让机器人练几百万、几千万次，学会开门、倒水、抓取物体。但这些机器人，一到真实的家庭环境里，就彻底拉胯了。

仿真里练了几万次倒水的机器人，到了真实世界里，杯子换了个形状、桌面晃了一下、水流快了一点，就直接把水洒了一地；练了无数次开门的机器人，遇到一个没见过的门把手，就直接束手无策。

为什么？因为现在的具身智能，走的还是“开普勒式”的老路：靠海量数据，拟合“看到什么画面，就做什么动作”的映射关系，根本不懂背后的物理规律。

它不知道“杯子倾斜超过45度，水就会洒出来”，不知道“手用的力气太大，杯子会被捏碎”，不知道“地面有水，走路会打滑”。它只是在背仿真里练过的动作，一旦场景和训练里不一样，就彻底失效了。

而这篇论文，指明了潜在的方向：**真正的具身智能，必须有一个牛顿式的世界模型。**

这个世界模型，不是对海量观测数据的记忆，而是对物理世界底层规律的理解。它不需要在仿真里练几百万次倒水，就能知道“杯子倾斜的角度，决定了水流的大小”；不需要练无数次走路，就能知道“踩到滑的地面，要减小步幅、放慢速度”；哪怕遇到一个完全没见过的场景，也能靠着对物理规律的理解，做出正确的决策。

这才是机器人和人类一样，能在真实世界里灵活行动的核心。我们人类从生下来，就不是靠背无数个场景的动作来生存的，而是在和世界的交互里，慢慢理解了空间、时间、力、因果这些底层规律，然后靠着这些规律，应对所有从未见过的场景。

现在行业里已经有了很多相关的尝试：谷歌DeepMind把物理世界模型融入机器人控制，让机器人在真实环境里的泛化能力提升了数倍；英伟达的Isaac平台，开始用带物理归纳偏置的世界模型，做机器人的预训练；国内很多机器人公司，也开始抛弃“纯数据驱动”的老路，把物理先验融入模型训练。

而这篇论文的价值，就是证明了：**我们不需要给模型硬编码复杂的物理公式，只需要给它注入符合世界本质的极简归纳偏置，它就能自发学会底层的物理规律。**这给具身智能的落地，提供了一条可复制、可扩展的路径。

关于AGI的一个思考：智能的本质，也许是对世界本质的理解

最后，我想聊一聊关于通用人工智能（AGI）的思考。

现在整个行业，都陷入了“Scaling Law”的执念里：觉得模型越大、数据越多，AI就越智能，就能离AGI越近。但这篇论文，给我们泼了一盆冷水：**靠堆数据、堆参数，永远堆不出真正的AGI。**

Scaling Law能让模型记住更多的知识，拟合更复杂的曲线，复刻更多见过的场景，但它永远没法让模型，自动发现世界的底层规律。就像你给一个人看一辈子的行星轨迹，他也未必能发现万有引力；但给一个孩子扔几次石头，他就能懂重力的存在。

区别在哪里？在于孩子天生就带着符合这个世界本质的归纳偏置：空间是连续的，时间是局部的，动作的结果是可重复的，有因必有果。这些刻在人类基因里的**归纳偏置**，让我们能从有限的交互里，快速理解世界的底层规律。

而AI的发展，也是一样的道理。

从开普勒到牛顿，是AI从“拟合数据”到“理解世界”的第一步。这一步，我们给AI注入了物理世界的归纳偏置，让它读懂了经典力学的规律。

而未来，我们要做的，是把这种归纳偏置，从物理世界，扩展到更广阔的领域：从物理世界的因果规律，到人类社会的运行规则，到心理世界的情绪逻辑，再到数学世界的抽象规律。

真正的AGI，从来不是一个能背下所有知识的百科全书，也不是一个能复刻所有见过场景的拟合器。它应该像人类一样，能从有限的经验里，发现世界的底层规律，能在从未见过的场景里，做出正确的决策，能从一个领域的规律里，迁移到另一个完全陌生的领域。

而这一切的起点，就是让AI真正读懂我们身处的这个物理世界。真正的具身智能，它的灵魂，必须是一个牛顿式的世界模型。这个世界模型，不是对观测数据的记忆，而是对物理世界流形的完整表征：

代码语言：javascript AI代码解释

目录

赛博解生
作者相关精选

从开普勒到牛顿：如何改进Transformer读懂物理世界

只有这样的机器人，才能像人类一样，在从未见过的场景里，做出正确的决策。它不用在仿真里练过一万次倒水，就能知道“杯子倾斜太厉害，水会洒出来”；不用练过无数次走路，就能知道“踩到滑的地面，要放慢脚步”。因为它懂了物理世界的底层规则，而不是背下了所有场景的应对动作。

对于AI来说，也是一样。从开普勒到牛顿，是AI从“拟合数据”到“理解世界”的第一步。而从理解物理世界，到拥有因果推理、类比创造、自主决策的能力，就是AI通往通用智能的完整路径。

模型 数据 机器人 解决方案 论文

目录

评论

期待你的精彩评论

0/300 发表

推荐阅读

编辑精选文章

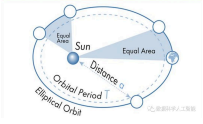
换一批

Java与Go差别在哪，谁要被时代抛…	12380	MCP协议详解：一文读懂跨时代的…	30888
一文掌握 MCP 上下文协议：从理论…	11378	LLM 效果不好？可能是 Prompt 写…	6386
鹅厂写码13年，我总结的程序员高效…	11443	进程，线程，协程 – 你了解多少?	8684

从开普勒三大定律到大数据分析

大数据 数据分析 http https 网络安全

开普勒定律是德国天文学家开普勒提出的关于行星运动的三大定律。这三大定律又分别称为椭圆定律、面积定律和调和定律，内容如下：



数据科学人工智能 2022/03/31 2.8K 0

继Ilya之后，KAN一作再发檄文：Scaling终将撞铁壁！

模型 数据 压缩 scaling 基础

继Ilya之后，柯尔莫哥洛夫-阿诺德网络KAN一作向Scaling Law发出最新檄文！

新智元 2026/01/13 154 0

宇宙认知的迭代路径：读《星空的琴弦》

交个朋友

加入腾讯云官网粉丝站

蹲全网底价单品 享第一手活动信息

赛博解生
作者相关精选

从开普勒到牛顿：如何改进Transformer读懂物理世界

人类绝望，机器接盘：用AI自动发现三体的守恒定律！北大校友与《生命3.0》作者共同杰作

机器学习 神经网络 深度学习 人工智能 编程算法

熟悉《三体》的科幻爱好者们都知道，三体人所在行星围绕着三颗恒星运行。不仅行星轨道极其不稳定，连三颗恒星之间的相对位置也变化无穷。所以，三体人经常要面临灭绝性的气候，不是严寒就…

AI科技评论 2021/05/19 956 0

北大物理学院欧阳颀院士：成为科学家的五大要素

其他

“科学的道路漫长而艰辛，要能持之以恒的坚持，需要做到兴趣驱动而非职业（收入）驱动，问题驱动而非学科驱动， 科学趣味驱动而非发表论文（SCI）驱动。”7月23日，北京大学物理学院教授欧阳颀院士在第18期理解未来讲座上，做了以“科学、科学…

大数据文摘 2018/05/22 980 0

给GNN一堆数据，它自己发现了万有引力定律

编程算法 机器学习 神经网络 深度学习 人工智能

来源：机器之心本文约2100字，建议阅读5分钟如果牛顿没被苹果砸中，GNN 和符号回归也能发现万有引力定律？ 机器学习 (ML) 推动了科学的巨大进步，从粒子物理学到结构生物学再到宇宙学，机器学习能够在大型数据集中学习特征，对不同的对象…

数据派THU 2022/03/25 479 0

牛顿棺材板快盖不住了：用深度神经网络解决三体问题，提速一亿倍

神经网络

刘慈欣在为自己的科幻小说起名为《三体》时，他早已知道“三体”本身就是一个不可回答的问题。

量子位 2019/10/31 498 0

人工智能在天体物理中的应用

神经网络 人工智能 深度学习 大数据

编者按：从人类诞生的那一刻起，人们对宇宙奥秘的求索就从未停止。今天，天文学已经进入了一个具有多波段、多信使的海量观测数据的黄金时期，人工智能技术将对天文领域产生深远影响。近日…

AI科技评论 2019/10/10 1.8K 0

给GNN一堆数据，它自己发现了万有引力定律

编程算法 机器学习 神经网络 深度学习 人工智能

机器之心报道 编辑：小舟、陈萍 如果牛顿没被苹果砸中，GNN 和符号回归也能发现万有引力定律？ 机器学习 (ML) 推动了科学的巨大进步，从粒子物理学到结构生物学再到宇宙学，机器学习能够在大型数据集中学习特征，对不同的对象进行分类，并执…

机器之心 2022/03/09 655 0

物理学,心理学,神经科学教授跨界对话:脑科学仍处于“前开普勒时代”

大数据 人工智能 机器学习

大数据文摘作品，未经授权禁止转载，转载具体要求见文末。 导读：人工智能正给经济社会带来巨大变化，而它本身尽管风头正盛，依然存在自身发展的瓶颈：机器学习不灵活，需要较多人工干预或大量标记样本；人工智能的不同模态和认知功能之间交…

大数据文摘 2018/05/22 1K 0

12位古代数学家的现代化成就

目录

交个朋友

赛博解生
作者相关精选

从开普勒到牛顿：如何改进Transformer读懂物理世界

AI大模型真的能理解物理现实么？论基础模型中的通用表征融合与物理实相的涌现

腾讯云架构师技术同盟 腾讯云 tvp AIGC 人工智能 芯片

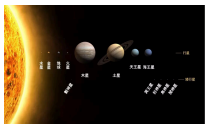
人工智能正在经历一场从任务特定模型向通用基础模型的范式转移。在语言和视觉领域，这一转变已催生了对表征融合现象的深刻观察——即随着模型性能的提升，它们对世界的内在表达趋于一致。MIT最近发表了篇论文《Universally Converging...

走向未来 2026/01/12 222 0

谈论AI之前，你搞懂人类了吗？（颠覆认知）

greenplum

导读：当前，人工智能应用在中国又一次火爆。无独有偶，美国电视剧《西部世界》第二季的第一集一经播出就引起热议。一时间，人和人工智能这个话题又重新被辩论。



IT阅读排行榜 2019/09/17 766 0

微积分的发现是人类精神的最高胜利

数学

在一切理论成就中，未必再有什么像17世纪下半叶微积分的发现那样被看作人类精神的最高胜利了，如果在某个地方我们看到人类精神的纯粹的和唯一的功绩，那正是在这里。——恩格斯

程序员小猿 2021/01/19 804 0

从物理动力学视角，重新理解智能、超级智能与当前大模型的真实水平

系统 框架 论文 模型 统计

这段时间，我反复研读了2026年2月发布在arXiv上，由韩国电子通信研究院Byung Gyu Chae撰写的《Emergence of Superintelligence from Collective Near-Critical Dynamics in Reentrant Neura...

赛博解生 2026/04/09 166 0

Sora不懂物理世界，翻车神图全网爆笑！LeCun马斯克DeepMind大佬激辩世界模型

视频 数据 神经网络 论文 模型

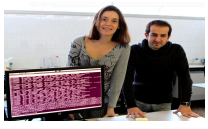
它被抬到半空中时，桌子上就忽然出现了一滩平整的红色玻璃，随后玻璃杯被摔到桌子上，和这滩玻璃融为一体。

新智元 2024/02/26 305 0

揭示世界本质的「机器科学家」，比深度神经网络还强?

编程算法 https 网络安全 神经网络 机器学习

我们正处于“GoPro 物理学”的风口浪尖。无论摄像机聚焦于什么事件，算法都可以识别其中潜在的物理方程。 作者 | Charlie Wood 编译 | 王玥、刘冰一 编辑 | 陈彩娴 2017 年，西北大学化学与生物...



AI科技评论 2022/05/12 509 0

何恺明团队新突破："物理直觉"重构AI视觉系统，去噪神经网络让机器看懂世界规律

深度学习 计算机视觉 机器学习

在计算机视觉领域，何恺明团队再次引领技术浪潮。他们最新提出的去噪哈密顿网络（Denoising Hamiltonian Network, DHN），首次将物理规律与去噪技术深度融合，赋予AI系统“物理直觉”。这...

CoovallyAIHub 2025/03/14 688 0

MIT惊人神作：AI独立提出哈密顿物理！0先验知识，一天破译人类百年理论

目录

交个朋友

加入腾讯云官网粉丝站

蹲全网底价单品 享第一手活动信息

赛博解生

作者相关精选

从开普勒到牛顿：如何改进Transformer读懂物理世界

SPARTAN：基于稀疏Transformer的局部因果学习框架

数据 对象 工作 框架 模型

SPARTAN: A Sparse Transformer Learning Local Causation

CreateAMind 2026/03/11 149 0

目录

交个朋友

加入腾讯云官网粉丝站

蹲全网底价单品 享第一手活动信息