Machine Learning in Astronomy: Possibilities and Pitfalls *Proceedings IAU Symposium No. 368 2022 ,,,*
A. Mahabal C. Fluke & J. McIver eds.
© 2022年 国际天文学联合会 DOI：00.0000/X0000000000000000X

arXiv:2302.04324v1 [astro-ph.CO] 8 Feb 2023

宇宙学中的深度机器学习：演进还是革命?

Ofer Lahav[1]

[2]D物理与天文学系，伦敦大学学院，伦敦 WC1E 6BT，英国 K

摘要。 机器学习（ML）能否带来根本性发现，并解决宇宙学中的未解难题？对宇宙当前组成的详细观测与宇宙学常数 Λ & 冷暗物质模型相符，但哈勃常数和团簇度因子的观测结果之间仍存在一些尚未解决的不一致（“张力”）。为进一步理解这些问题，对数十亿个星系及其他探测对象开展的大规模巡天需要采用新的统计方法。近年来，机器学习的强大能力，尤其是“深度学习”，已在天体分类、光度红移、异常检测、增强模拟以及推断宇宙学参数等方面得到证明。本文认为，较为传统的“浅层学习”（即通过预处理进行特征提取）实际上相当深入，因为它融入了人类知识；而“深度学习”则可能被视为一个黑箱，除非辅以可解释性工具。机器学习在宇宙学中的“杀手级应用”尚未出现。本文还探讨了如何以新的方式培养下一代科学家，以应对即将到来的数据密集型科学挑战。最后，本文要求聊天机器人 ChatGPT 回答文章标题所提出的问题。 †

关键词。方法：数据分析 – 方法：统计学 – 技术：星系

1. 引言

天文数据呈指数级增长，令人瞩目：例如，星系样本从20世纪80年代的数千个和数万个（CfA、IRAS），发展到20世纪90年代及21世纪初的数百万个（2dF、斯隆数字化巡天、重子振荡光谱巡天），再到如今的3亿个（DES），以及将于2023—2024年启动的新一轮巡天（Rubin-LSST、Euclid）所观测的数十亿个星系。下文图1展示了部分巡天的数据采集速率、星系数量、成本和研究人员数量。

科学与技术各领域的另一场变革，则是人工智能（AI）和机器学习（ML）的发展。就术语而言，机器学习是人工智能的一个子集，而深度学习（DL）是机器学习中的一种新方法（见图2）。如今，人工智能与机器学习在天文学中是一个十分活跃的领域，每周都有大量论文发表，短短的报告无法一一展示。我在此冒昧介绍多年来与我的博士生和博士后共同开展的研究，并以此为例，探讨人工智能与机器学习在天文学及其他领域的应用。如今，机器学习在天文学中的应用十分普遍，但仍有一些悬而未决的问题：这些方法在多大程度上是“黑箱”，它们与传统统计学有何联系，以及如何实施这些方法。近期的综合性综述包括：Baron（2019）关于机器学习在天文学中的应用，Carleo等（2019）关于机器学习在物理科学各领域的应用，以及Huertas-Company与Lanusse（2022）关于深度学习在星系巡天中的应用，涵盖天体分类、星系性质研究、发现、模拟仿真和推断。

本文的结构如下。本文将在当前及未来星系巡天的背景下回顾 Λ & 冷暗物质（ΛCDM）模型的现状，这些巡天将绘制数十亿

† 本文基于在英国皇家天文学会乔治·达尔文讲座（2020年）、于韩国举行的国际天文学联合会第368届研讨会（2022年）上的全体会议报告，以及在希伯来大学举办的“天体物理学与宇宙学中的未解问题”研讨会（2022年）上的报告。

Survey	Data per night/day	Galaxies	Cost	Scientists
DES	1 TeraB	~300 Million (all observed)	~$70M	~400
DESI	40 GigaB	~35 Million (12M observed)	~$70M	~600
Rubin-LSST	15 TeraB	~Billions	~$1.0B	~1000
Euclid	850 GigaB	~Billions	~$1.5B	~1500
SKA	1 PetaB	~Billions	~$1.3B	~1000

图 1. 星系巡天示例：数据速率、星系数量、资金资源和人力资本。DES 的观测已完成，DESI 正在收集数据，而 Rubin-LSST、Euclid 和平方公里阵列（SKA）将在未来几年投入运行。

星系。本文介绍人工智能与机器学习在宇宙学和天体物理学中的应用实例：（i）天体分类；（ii）光度红移；（iii）暗物质制图；以及（iv）基于模拟推断银河系和仙女座星系的质量及宇宙学参数。本文还回顾了如何将新一代天文学家培养成数据科学家，并展望人工智能与机器学习在天文学中的前景，同时以伦敦大学学院数据密集型科学博士培训中心（CDT）为例。

2. 背景：宇宙学模型的现状

近年来，爱因斯坦广义相对论不断取得成功：从测量宇宙学常数 Λ（或其变体），到探测以光速传播的引力波（GW）（LIGO合作组 2017a）。具体而言，近期宇宙学测量结果强烈支持一种“共识”模型：当前宇宙是平坦的，其中约有 5% 的重子、25% 的冷暗物质和 70% 的暗能量，另有少量有质量中微子（例如，普朗克合作组 2018年；暗能量巡天合作组 2021年；综述参见 Lahav与Liddle 2022年）。现代观测证据认为，暗能量的发现归功于 20 世纪 90 年代末的 Ia型超新星巡天（Perlmutter等（1999）；Riess等（1998））， 此后其他观测探针也提供了支持，尽管早在此前十年，支持暗能量存在的证据就已日益增多。暗能量的概念是爱因斯坦宇宙学常数的一种变体， Λ；而类似 Λ 的线性力这一设想可追溯至牛顿《自然哲学的数学原理》（例如 Calder 和 Lahav 2008）。这一 ΛCDM 范式及其扩展引出了关于宇宙起源的根本问题。通常，

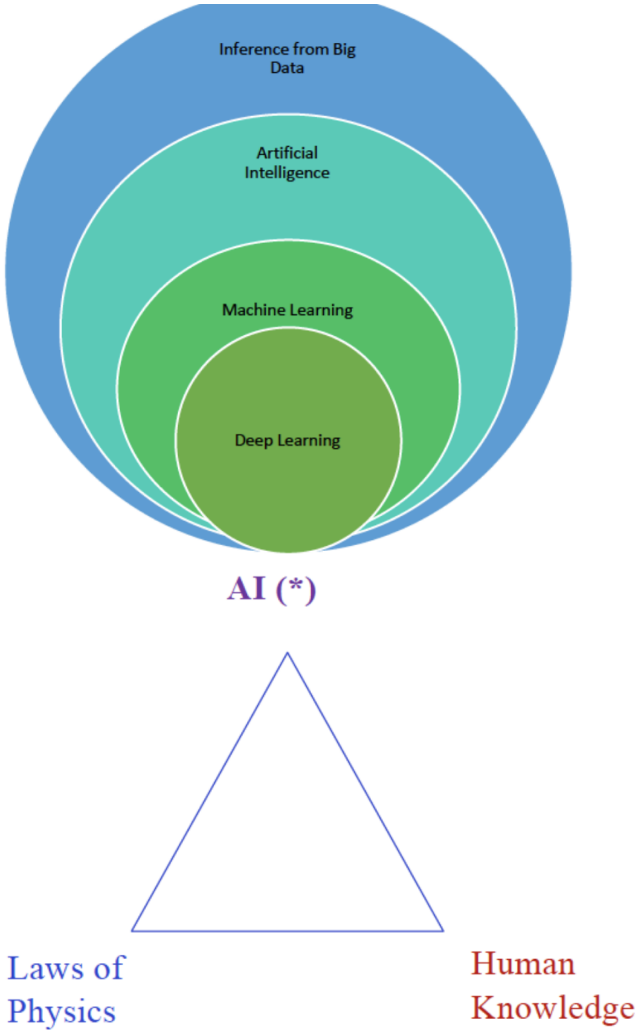
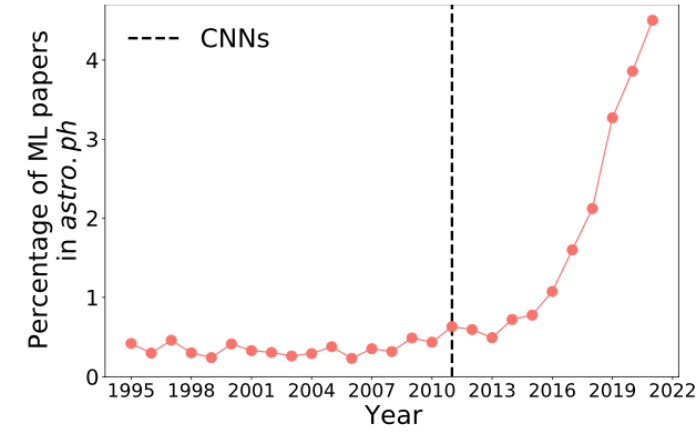


图 2. 上：各子领域之间的关系：统计推断、人工智能、机器学习和深度学习。下：（1）物理学的“客观”定律、（2）当前的人类知识，以及（3）能够根据新数据更新人类知识的增强智能之间的相互作用。

暗能量由状态方程参数 w 定量描述，它是压强与密度之比，即 $w = P/\rho$。 $w = -1$ 的情形与广义相对论中的爱因斯坦宇宙学常数完全对应。但 w 可能随宇宙时期而变化，例如在标量场的情形下。

本质上， w 同时影响宇宙几何和结构增长率。这些效应可以通过一系列宇宙学探针观测到，包括宇宙微波背景（CMB）、星系聚类、星系团、弱引力透镜和Ia型超新星。例如，Ia型超新星充当“标准烛光”，而重子声学振荡（BAO）则作为“标准尺”。当前的宇宙学模型 ΛCDM在拟合大多数观测结果方面取得了极为成功的表现。然而，根据宇宙微波背景测量得到的哈勃常数 $H_0 = 67.4 \pm 0.5\,\text{km/sec/Mpc}$



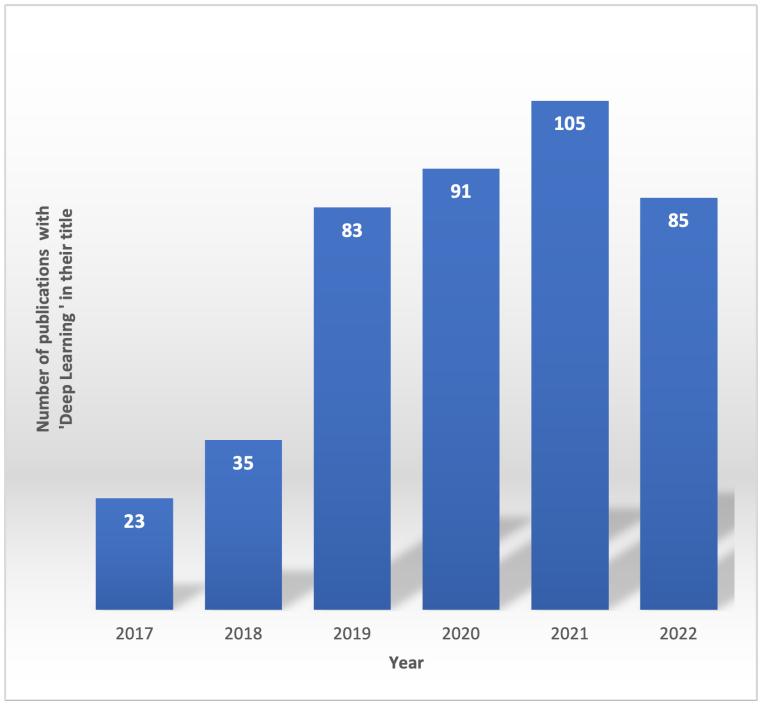


图 3。 上图：标题或摘要中报告至少一种机器学习技术的天体物理学（arXiv.org）论文所占的百分比。早在 1995 年之前，就已经出现了一些关于天文学机器学习的论文（见正文）。竖线表示卷积神经网络的普及。来源：D. Piras，博士论文（博士培训中心-数据密集型科学/伦敦大学学院）。下图：每年标题中包含“深度学习”的天体物理学 arXiv 论文数量。

（普朗克合作组 2018年）与本地测得的 $H_0 = 73.2 \pm 1.3$ km/sec/Mpc（Riess等人 2021年）之间存在约 H_0 的张力，这可能迫使我们重新考虑标准 ΛCDM模型（综述见Shah等人 2021年）。未知的系统误差可能是造成这种张力的原因，因此我们需要利用 H_0 的独立探针（例如引力波“明 sirens”和“暗 sirens”）来确认我们的宇宙模型是否陷入困境。此外，在团块度参数 S_8方面也存在约 2σ 的张力，具体而言，弱引力透镜测得的质量涨落

三项不同的巡天结果似乎比普朗克宇宙微波背景辐射的测量结果略显“平滑”。机器学习能否帮助解决宇宙学中的这些问题？下面将讨论机器学习应用中的一些挑战和实例。

作为巡天项目的一个例子，我们现在重点介绍暗能量巡天（DES），我自2004年以来一直参与其中。DES于2019年完成了六个观测季，共计758个夜晚。DES利用位于智利的4米布兰科望远镜上的大视场相机，通过五个滤光片对5000平方度的天空进行了成像。该巡天生成了一个包含3亿个带有光度红移的星系和1亿颗恒星的目录。此外，一项覆盖27平方度的时域巡天搜索获得了数千个Ia型超新星及其他暂现源的样本。DES 合作组（2021年）测得 $w = -1.03 \pm 0.03$ （68%置信水平）（结合DES弱引力透镜与星系聚类、普朗克数据及其他探针），结果与爱因斯坦宇宙学常数一致。DES还带来了意料之外的科学成果，包括发现太阳系中的遥远天体、银河系的矮星系、高红移类星体，以及探测到来自LIGO引力波双中子星事件GW170817的光学闪光。从2003年的构想到建设和观测，再到2019年的首批成果，DES的历程载于《DES专著》（Lahav等编，2020年）。

下一代巡天项目“第四阶段”（Rubin-LSST、DESI、Euclid、Roman-WFIRST）的科学目标更为宏大。一个引人深思的问题是：为了测量诸如 w 这样的宇宙学参数（例如 Lahav & Silk 2021），何时才该停止？简而言之，这取决于替代理论有哪些。当然，提高测量的准确度和精度是有价值的，因为这可能会揭示意想不到的结果。探测偏离 ΛCDM 的现象并测量中微子质量，有望成为整个物理学领域的里程碑式测量。

3. 深度学习与浅层学习

天体物理学、高能物理学，以及科学、技术与医学的几乎所有领域，都正在经历一场“工业革命”，其特点是数据集规模庞大、模型复杂。例如，前文提到的星系巡天项目将在未来十年内为数十亿个星系编目。同样，粒子物理学实验（如欧洲核子研究中心的实验）也会生成巨量数据集。这些数据集需要全新的数据分析方法。值得庆幸的是，统计学家和计算机科学家发明了强大的方法，并将其广泛应用于不同领域。近年来，机器学习技术尤其受到广泛关注。常用的方法包括人工神经网络（ANN，见例如图 7）、提升决策树（BDT）和支持向量机（SVM）。各种方法已有实现，例如 Scikit-Learn† 和 TensorFlow‡等软件包。近年来，机器学习领域的一项重大突破是“深度学习”（DL），例如参见 Le Cun 等人（2015）和 Goodfellow 等人（2016）。深度学习并非先进行预处理，再执行特征提取。例如，一些深度学习算法采用由多个层（“滤光片”）组成的级联结构，让全部数据“自己说话”（见例如图 5）。深度学习既可用于监督学习，也可用于无监督学习。一种广受欢迎的深度学习实现是卷积神经网络（CNN）。深度学习无疑是天体物理学领域的一项重要新趋势（图 3）：论文标题中包含深度学习的论文数量从 2017 年的 23 篇增至 2021年的 105 篇。深度学习的一项新颖应用是探测强引力透镜（Metcalf 等人，2019）；在这方面，卷积神经网络的表现优于目视检查和其他机器学习算法。

尽管深度学习方法的性能令人印象深刻，但理解深度学习算法仍然很重要，不能只是把它们当作“黑箱”来使用。可以通过在玩具模型上运行机器学习算法来做到这一点，也可以使用 *explainability* 和 *interpretability* 工具（见下文）。我们还认为，机器学习中传统的特征提取预处理绝非“浅层”。人类在物理学和天文学方面的丰富知识影响了此类特征的选择。

†http://scikit-learn.org/stable/ ‡ https://www.tensorflow.org/

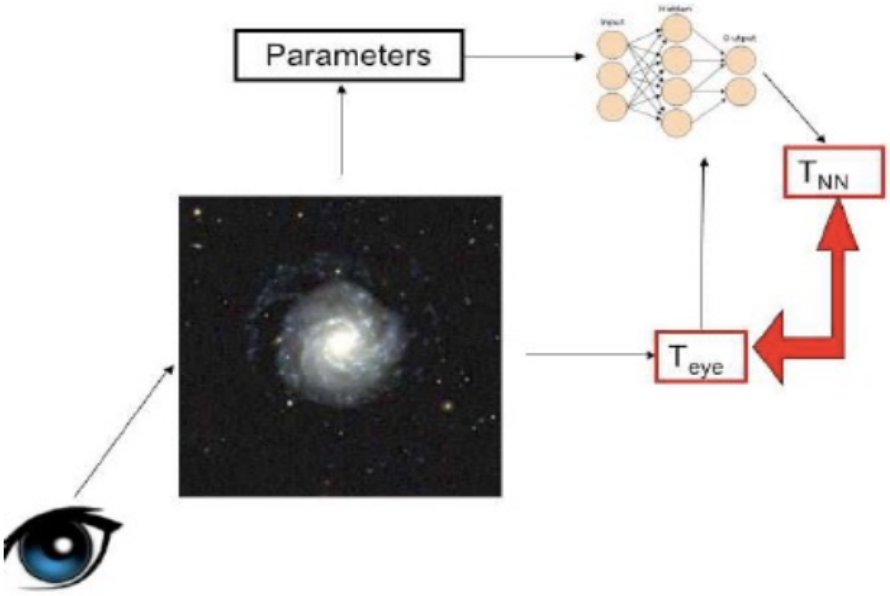


图 4. 使用人工神经网络进行星系分类的监督学习示意图。训练集的分类通过目视检查整幅图像完成，而训练和测试只使用从星系图像中经预处理提取的十来个特征。图片来源：M. Banerji 等人（2010年）。

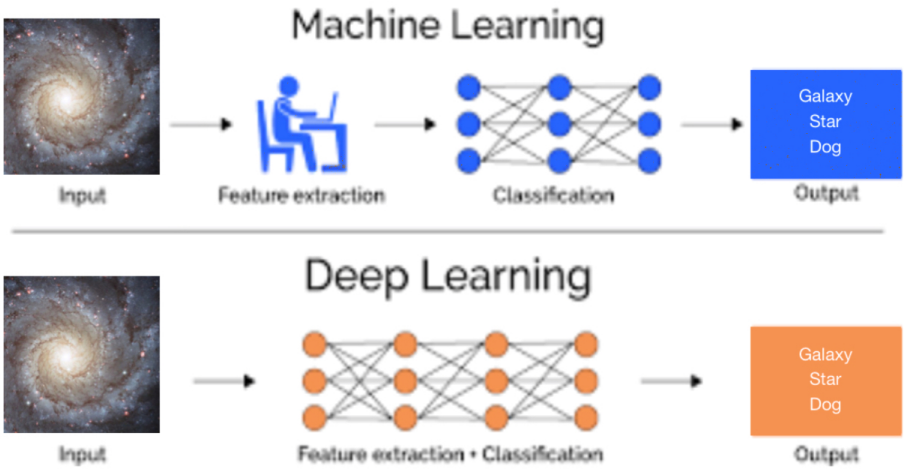


图 5. “浅层学习”（基于预处理特征提取）与“深度学习”的对比。图片来源：S. M. Lahav，改编自维基百科

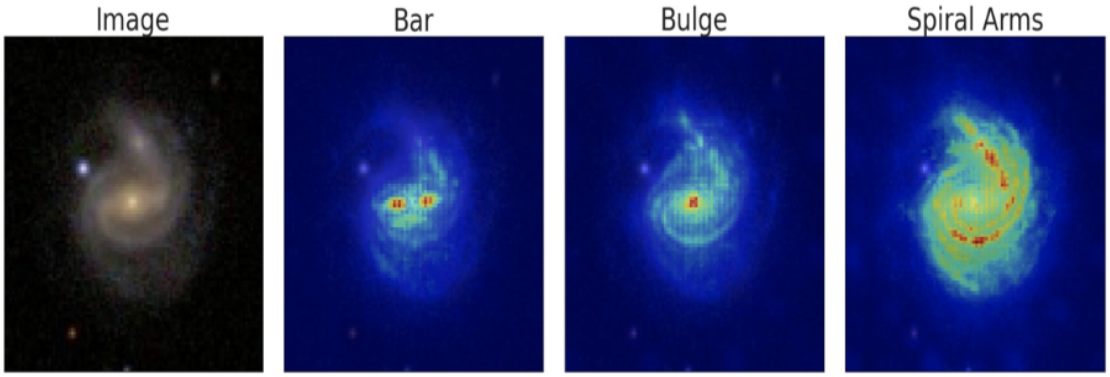


图6。 可解释人工智能示例：左侧星系图像对应的三张显著性图，显示了每个像素对目视分类特征（棒、核球和旋臂）的敏感程度，所用工具为可解释人工智能工具（Bhambra 等，2022）。

4. 天体分类

先谈一点个人经历：20世纪90年代初，我在剑桥的天文学研究所工作时，对机器学习产生了兴趣。我和同事讨论过，能否让计算机而非人类来对星系进行分类。一次偶然的机会，我读到一本关于人工神经网络的书，随后与 Michael Storrie-Lombardi 和 Laerte Sodre 合作开展研究，之后 Avi Naim 的博士研究也延续了这项工作（Storrie-Lombardi 等，1992；Lahav 等，1995；Naim 等，1995）。我们设计了一项实验，请6位分类“专家”给840张星系图像贴上标签（这些图像由剑桥自动照相底片测量机扫描）。随后，我们训练一个人工神经网络，利用十几个特征预测德·沃库勒 T 型分类系统中的15种类型。该人工神经网络的分类结果与人工分类相差不超过2个类型。这项工作在当时并非主流，但后来被拓展应用于星系动物园的分类（例如 Banerji 等，2010；Walmsley 等，2020）。一个有趣的问题是，完整图像中的哪些特征会让人判断某个天体属于这一类型而非另一类型。可解释人工智能工具（如 SmoothGrad）为解决这一问题迈出了一步。Bhambra 等（2022）通过计算每个像素处目视分类 T 型相对于像素强度的平滑梯度（绕过算法的内部架构），得出了显著性图（见图6），以此展示了这一方法。

另一个重要领域是星系光谱的分类。相关方法包括主成分分析（例如 Madgwick 等，2003），以及熵和相对熵等信息论度量（Slonim 等，2000；Ferreras 等，2022）。大多数方法都表明，包含数千个波段的光谱，其信息量可以压缩约 100 倍。先进的机器学习方法能否揭示更多信息，还有待观察。

分类也可以在时域中进行。例如，Lochner 等（2016）展示了如何利用 DES 模拟的 Ia 型超新星（SN）光变曲线，从 SN 光变曲线中提取特征（使用 SALT2、参数模型和小波），并运用五种机器学习方法（朴素贝叶斯、K 近邻、支持向量机、人工神经网络和提升决策树）对其进行分类。通过曲线下面积（AUC），他们发现，例如，使用提升决策树算法进行小波特征提取时，AUC 可达 0.98，而 1 代表完美分类。由此发展出了 SNMachine 软件包，供 Rubin-LSST 使用；预计 Rubin-LSST 将在 10 年的观测中探测到数百万颗超新星。另一个例子是利用机器学习在引力波数据中搜寻核心坍缩超新星（Iess 等，2023）。

di

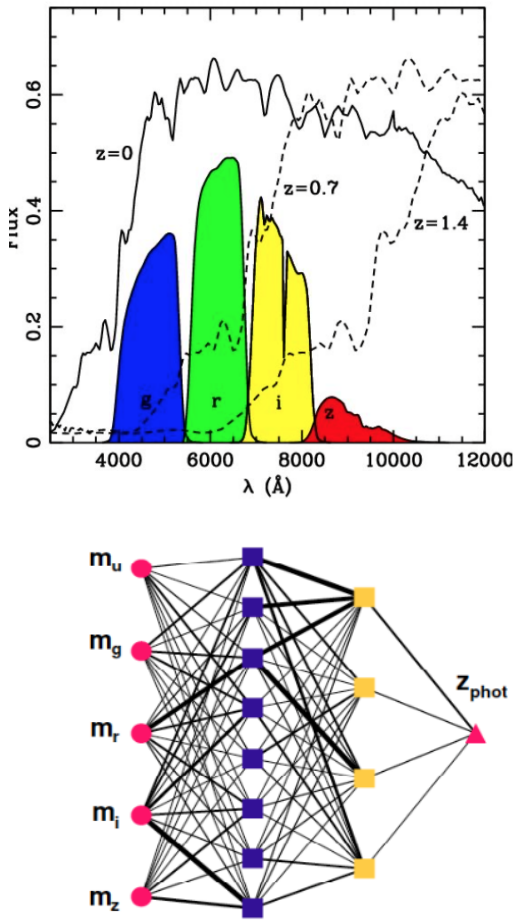


图 7. 上图：典型星系光谱（通量随波长变化），展示了在三个不同红移下由四个滤光片 g、r、i、z 采集的样本。图片来源：维基百科。下图：一个人工神经网络配置，以 5 个星等（每个滤光片对应一个）为输入，以红移为输出。图片来源：Sadeh 等人（2016 年）。

5. 光度红移

测量星系的距离对于星系聚集或弱引力透镜等宇宙学研究至关重要。理想情况下，人们希望为每个星系测得光谱红移。但实际上，获取光谱需要很长的曝光时间，因此可以改为进行多波段测光。光度红移（“光度红移”）不如光谱红移精确，但光度红移星系的数量远多于光谱红移星系，因此仍有可能获得可靠的宇宙学结果。DES、KiDS、HSC、Rubin-LSST、Euclid 和 Roman-WFIRST 等光度巡天都高度依赖光度红移。可以说，推导光度红移令宇宙学家十分头疼（他们更愿意使用光谱红移），却让机器学习者乐在其中！事实上，光度红移可以视为一个逆问题：从通过多个滤光片测得的通量推导红移（见图 7）及其对应的概率分布函数。文献中已有十多种推导光度红移的方法。广义而言，这些方法可分为

可分为模板方法和训练方法，后者可利用规模较小的光谱红移数据集来训练机器学习工具。大约二十年前，我们团队开发了一套人工神经网络软件包，最初是 ANNz（Collister 和 Lahav，2004），后来发展为 ANNz2（Sadeh 等，2016）。我们还尝试通过加入星系结构参数（Soo 等，2018），以及采用深度学习方法分析完整的星系图像（Henghes 等，2022；该方法沿袭 Pasquet 等，2019），来提高光度红移的准确度。此外，还可以同时利用光度数据生成光度红移和恒星质量（Mucesh 等，2021）。我们也研究了训练数据集规模及其他因素对光度红移机器学习算法可扩展性的影响（Henghes 等，2022）。有关解决光度红移问题的其他机器学习方法（例如自组织映射），请参阅 Huertas-Company 与 Lanusse（2022）的综述。机器学习光度红移在未来许多年里仍将是一个活跃的研究领域，适用于 Rubin-LSST、Euclid 及其他大型光度巡天。

6. 基于模拟的星系质量、暗物质图和宇宙学参数推断

大规模模拟和机器学习算法的发展为宇宙学研究开辟了许多新的方向。以下是几个例子：

(i) 银河系与仙女座星系的质量总和：最初，人们通过经典的“时间论证”（Kahn & Waltjer 1959；Lynden-Bell 1981）对其进行估算，该论证假设两个点状天体相互靠近。其总质量根据观测到的当前距离、相互靠近的速度以及宇宙的年龄推算得出。但实际上，这两个星系都是本星系群的一部分，本星系群中有数十个星系，且本身位于宇宙网之中。为了将这种结构和动力学的复杂性纳入考量，可以将观测到的双体系统与大量模拟星系对进行比较。McLeod 等人（2017）使用 30,000 个模拟星系晕对训练了人工神经网络；Lemos 等人（2021）则采用密度估计无似然推断，分析了从 ΛCDM模拟中提取的 200 万个星系晕对。后者研究得出，密度对比度为 200 的范围内所包含的总质量为 $M_{200} = 4.6 \times 10^{12}\ ^{+2.3}_{-1.8}$（68% 置信水平）太阳质量。该估计值与时间论证的结果一致，但略高于其他方法得出的各星系质量之和。得益于所采用的先进机器学习方法，误差范围也得到了更精细的估算。

（ii）从弱引力透镜重建暗物质图：利用因弱引力透镜而发生畸变的遥远星系图像（“宇宙剪切”），推导投影质量分布这一逆问题，最早由 Kaiser 和 Squires（1993）提出的经典算法加以解决。图8展示了一个重建质量图的示例（使用维纳滤波器）。近年来，人们采用了多种方法解决这一重建问题，其中包括一种深度学习方法（Jeffrey 等，2021b）。该算法通过约 10^5 次模拟学习宇宙剪切与投影质量之间的关系。随后，将观测到的剪切（例如，DES在5000平方度天空范围内获得的1亿个星系形状测量）输入算法，以生成暗物质图。这些技术还可应用于Euclid和Rubin-LSST（两者各自覆盖约20000平方度的天空，预计观测到10亿个星系）。这类暗物质图对于与星系图中的星系团和空洞进行比较（Hang 等，准备中），以及与大尺度结构的其他多波段观测结果进行比较，都非常有用。更一般地说，图对于在“场级别”研究结构非常有用，可超越通常使用的二点统计。星系图和暗物质图可用于N点统计（例如 Gaudi 等，2019）、概率分布函数（例如 Clerkin 等，2016）以及图统计学，例如最小生成树（Naidoo 等，2020）。在小尺度上，弱引力透镜图能够揭示晕和星系团中潜在的暗物质分布。这可用于研究重子和中微子对晕子结构的影响，以及暗物质粒子相互作用对正在合并的星系团的影响。另一个很有前景的方向是利用深度学习在图上进行训练，以预测宇宙学参数（Kacprzak 和 Fluri，2022；Jeffrey 等，2021a），甚至通过显著性图（类似于图6）解释图中哪些区域能够为我们提供关于特定宇宙学和天体物理参

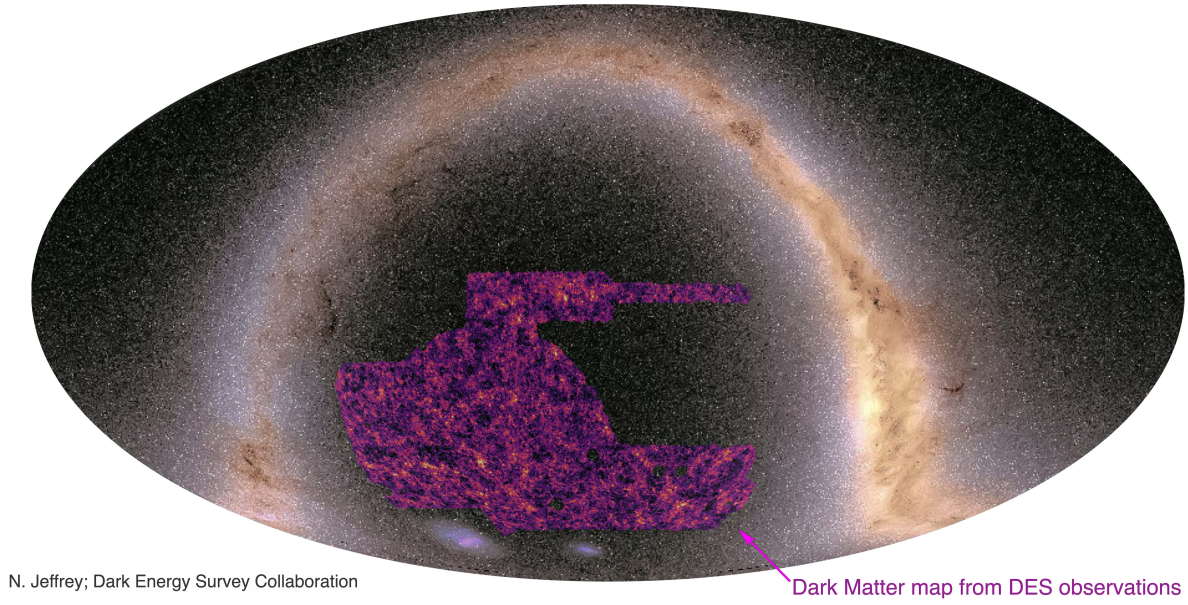


图 8。 利用 DES 第三年弱引力透镜数据绘制的质量分布图，使用了在 5000 平方度巡天区域内测得形状的 1 亿个星系。高密度和低密度区域分别以亮色和暗色显示。图中还展示了盖亚拍摄的银河系图像，以标示天空方位。图片来源：Nial Jeffrey 与 DES 合作组（2021年）。

et 数的最多信息）。模拟尽可能真实且系统误差得到恰当建模，这一点非常重要。

(iii) 机器学习在引力物理学中的应用： 另一个不断发展的领域是对 N 体模拟和流体动力学模拟的分析。例如，Lucie-Smith 等人（2018年）利用机器学习研究模拟初始条件与最终暗晕之间的关系。通过在少量“昂贵”的模拟上进行训练，机器学习可以增强大量“廉价”模拟的效果，例如 Dai 与 Seljak（2021年）提出的拉格朗日深度学习。一个更具雄心的目标是让机器学习算法发现新的物理规律，例如重新发现太阳系的轨道力学（Lemos 等人，2022年）。

7. 培养新一代天文学家成为数据科学家

快速发展的人工智能技术要求采用新型训练方式，使学员能够在天体物理学（高能物理学或物理学的其他分支）和数据密集型科学（DIS）领域发展双重职业。例如，英国的 STFC 已启动数据密集型科学博士培训中心（CDT）新计划。伦敦大学学院的 CDT-DIS† 每年培养约 10 名博士生。这是一个为期 4 年的博士项目，其中包括在约 30 家私营、公共或第三部门的数据密集型科学机构之一进行为期 6 个月的实习，例如从面向时尚零售的自然语言处理，到用于核聚变的机器学习。该博士项目现已纳入伦敦大学学院数据密集型科学与产业中心（DISI）；该中心还开设为期一年的理学硕士项目，招收约 70 名学生。该中心与国际天文学和高能物理项目以及其他大学保持着紧密联系。在英国牛顿基金的支持下，我们还与约旦‡合作开展数据密集型科学培训项目。

† https://www.ucl.ac.uk/data-intensive-science-cdt/ ‡ https://www.ucl.ac.uk/astrophysics/research/ucl-jordan-dis-collaboration

8. 展望

人工智能正在改变世界。我们已经进入这样一个阶段：如果没有某种形式的机器学习，就无法分析宇宙学实验。机器学习无疑提高了生产力。相对较新的深度学习已经被证明对天文学中的许多问题非常有用，例如天体分类、光度红移、异常检测、增强模拟以及宇宙学参数的推断。然而，我认为，更为传统的“浅层学习”（即结合预处理特征提取）实际上相当深入，因为它引入了人类知识；而“深度学习”可能会被视为一个黑箱，除非辅以可解释性工具。人工智能、物理定律与人类知识之间相互关联，并且彼此促进、共同发展。正如一些人所倡导的那样，与其将人工智能视为“人工智能”，不如将其视为“增强智能”。在这一语境下，人工智能是在增强人类智能，而不是取代人类智能（例如 Lochner 与 Bessett，2021年）。

还有其他挑战：我们能否在不把机器学习算法仅仅当作“黑箱”的情况下，对其进行解释或阐释？如何尽量减少训练集不完整所导致的偏差？能否纳入已有的物理学知识？能否发现新的物理学规律？如何更好地培养新一代天文学家，使他们同时成为数据科学家？这引领我们迈向下一阶段：将新颖的人工智能和机器学习方法应用于星系巡天（DES、DESI、Rubin-LSST、Euclid 等），其发现潜力巨大，既能探测标准ΛCDM模型中的细微偏差、测量中微子质量，也有可能发现新的天体种群。或许可以说，我们仍缺少机器学习在宇宙学中的“杀手级应用”，类似于机器学习近期在蛋白质折叠领域取得的成功†。

最后，我向人工智能聊天机器人 ChatGPT‡ 提出了本文标题中的问题：“宇宙学中的深度机器学习是一种演进还是一场革命？”聊天机器人回答道，

Deep machine learning has the potential to be both an evolution and a revolution in the field of cosmology. On one hand

it represents an evolution in the sense that it builds upon and extends existing methods and approaches that have been used in cosmology for decades. For example

machine learning algorithms can be used to analyse and interpret large datasets in ways that are more efficient and accurate than traditional methods.

On the other hand, deep machine learning also has the potential to be a revolution in cosmology by enabling the discovery of new patterns and insights that were previously hidden or difficult to discern. For example machine learning algorithms can be used to identify complex correlations

and patterns in cosmological data that may not be easily identifiable through traditional methods of analysis. In this way

deep machine learning has the potential to fundamentally change the way that cosmologists approach and study the universe.”

希望 ChatGPT 最后一句中的预测是正确的！

致谢 感谢多年来我的博士生、博士后和合作者为本文所述工作作出的贡献。感谢 ERC高级资助、STFC 综合资助以及 CDT-DIS资助对这项研究的支持，并感谢在牛津大学万灵学院获得的访问学者奖学金。

参考文献

Banerji, 等，2010年，皇家天文学会月刊，406，342 Baron, D., 等，2019，arXiv:1904.07248 Bhambra, P., 等，2022年，皇家天文学会月刊，511，5032 Bishop, C., 2006, *Pattern recognition and Machine Learning*，施普林格 Calder, L. & Lahav, O., 2008，英国皇家天文学会《天文学与地球物理学》，49，1.13 Carleo, G., 等，2019，《现代物理评论》，91，045002

† https://www.deepmind.com/research/highlighted-research/alphafold ‡ chat.openai.com

Chang, C., 等，2016，皇家天文学会月刊，459，3203 Clerkin, L., 等，2017，皇家天文学会月刊，466，1444 Collister, A. & Lahav, O., 2004，太平洋天文学会刊，116，345 Dai, B. & Seljak, U., 2021年，美国国家科学院院刊，118 (16) e2020324118 暗能量巡天合作组，2016，皇家天文学会月刊，460，1270 暗能量巡天合作组，2022年，《物理评论D》，105，023520 Elgaroy, O,. 等，2002，《物理评论快报》，89，061301 Ferreras, I., 等，2022，RASTI（即将出版），arXiv:2208.05489 Goodfellow, I., 等，2016, *Deep Learning*，麻省理工学院出版社 Gualdi, D., 等，2019，皇家天文学会月刊，484，3713 Henghes, B., 等，2021年，皇家天文学会月刊，505，4847 Henghes, B., 等，2022年，皇家天文学会月刊，512，1696 Huertas-Company, M. & Lanusse, F., 2022，arXiv: 2210.01813 Iess, A., 等，2022，《天文学与天体物理学》，669，A42 Jeffrey, N., 等，2021a，皇家天文学会月刊，501，954 Jeffrey, N., 等，2021b，皇家天文学会月刊，505，4626 Kacprzak, T. & Fluri, J., 2022，arXiv:2203.096116 Kahn, F.D., & Woltjer, L., 1959，《天体物理学杂志》，130，705 Kaiser, N. & Squires G., 1993，《天体物理学杂志》，404，441 Lahav, O., 等，1995，《科学》，267，859 Lahav, O. & Liddle, A., 2021，《粒子物理学评论》，arXiv:2201.08666 Lahav, O. & Silk, J., 2021，《自然 · 天文学》，5，855 Lahav, O., 等（编），2020 年，*The Dark Energy Survey: The Story of a Cosmological Experi-ment*，世界科学出版社 Le Cun, Y., 等，2015，《自然》，521，43 Lemos, P., 等，2021，《物理评论D》，103，023009 Lemos, P., 等，2022，arXiv:2202.02306 LIGO合作组，2017a，《天体物理学杂志》，848，L12 LIGO合作组，2017b，《自然》，551，85 Lochner, M., 等，2016，《天体物理学杂志增刊》，225，31 Lochner, M. & Bassett, B.A., 2021，《天文学与计算》，36，100481 Lucie-Smith, L., 等，2018，皇家天文学会月刊，479，3405 Lynden-Bell, D., 1981，《天文台》，101，111 Madgwick, D.S., 等，2003，皇家天文学会月刊，343，871 McLeod, M., 等，2017，《宇宙学与天体粒子物理学杂志》，12 03 Metcalf, R.B., 等，2019，《天文学与天体物理学》，625，A119 Mucesh, S., 等，2021，皇家天文学会月刊，502，2770 Naidoo, K., 等，2020，MNARS，491，1709 Naim, A., 等 1995，皇家天文学会月刊，275，567 Pasquet, J., 等，2019，A & A，621，A26 Perlmutter, S., 等 1999，《天体物理学杂志》，517，565 普朗克合作组，2018，arXiv:1807.06209 Riess , A.. 等，1998，《天文学杂志》，116，1009 Riess, A., 等，2021，《天体物理学杂志》，853，126 Sadeh, I., 等.. 2016，太平洋天文学会刊，128，4502 Shah, P., 等，2021，《天文学与天体物理学评论》，29，9 Slonim, N., 等 2001，皇家天文学会月刊，323，270 Soo, J., 等，2018，皇家天文学会月刊，475，3613 Storrie-Lombardi, M..C., 等，1992，皇家天文学会月刊，259，8p Walmsley , M., 等，2020，皇家天文学会月刊，491，1554