



深度学习在核物理中的应用

庞龙刚 周凯 王新年

Deep Learning for Nuclear Physics

PANG Longgang, ZHOU Kai, WANG Xinnian

在线阅读 View online: <https://doi.org/10.11804/NuclPhysRev.37.2019CNPC41>

引用格式:

庞龙刚, 周凯, 王新年. 深度学习在核物理中的应用[J]. *原子核物理评论*, 2020, 37(3):720–726. doi: 10.11804/NuclPhysRev.37.2019CNPC41

PANG Longgang, ZHOU Kai, WANG Xinnian. Deep Learning for Nuclear Physics[J]. *Nuclear Physics Review*, 2020, 37(3):720–726. doi: 10.11804/NuclPhysRev.37.2019CNPC41

您可能感兴趣的其他文章

Articles you may be interested in

用于核物理研究的精密激光光谱技术的发展和展望

Development and Prospect of Precision Laser Spectroscopy Techniques for Nuclear Physics Study
原子核物理评论. 2019, 36(2): 161–169 <https://doi.org/10.11804/NuclPhysRev.36.02.161>

标准对力模型的新迭代方法(英文)

A New Iterative Approach for the Exact Solution of the Standard Pairing Model
原子核物理评论. 2018, 35(4): 518–522 <https://doi.org/10.11804/NuclPhysRev.35.04.518>

致密物质状态方程：中子星与奇异星

Dense Matter Equation of State: Neutron Star and Strange Star
原子核物理评论. 2019, 36(1): 1–36 <https://doi.org/10.11804/NuclPhysRev.36.01.001>

基于不稳定核基本性质测量的原子核结构研究

Study of Nuclear Structure by the Measurement of the Ground State Properties of Unstable Nuclei
原子核物理评论. 2018, 35(4): 382–389 <https://doi.org/10.11804/NuclPhysRev.35.04.382>

核物质四阶对称能中的交换项相关物理

Physics of Fock Terms on Fourth-order Symmetry Energy of Nuclear Matter
原子核物理评论. 2018, 35(4): 549–554 <https://doi.org/10.11804/NuclPhysRev.35.04.549>

格点核子-核子势在核物质中的相对论效应(英文)

Relativistic Effects in Nuclear Matter with Lattice NN Potential
原子核物理评论. 2017, 34(3): 505–508 <https://doi.org/10.11804/NuclPhysRev.34.03.505>

文章编号: 1007-4627(2020)03-0720-07

深度学习在核物理中的应用

庞龙刚^{1,2}, 周凯³, 王新年^{1,2,4}

(1. 华中师范大学夸克与轻子教育部重点实验室, 武汉 430079;

2. 华中师范大学粒子物理研究所, 武汉 430079;

3. 法兰克福高等研究中心, 德国 法兰克福 60438;

4. 劳伦斯伯克利国家实验室, 美国 加州伯克利 94720)

摘要: 深度学习是目前最好的模式识别工具, 预期会在核物理领域帮助科学家从大量复杂数据中寻找与某些物理最相关的特征。本文综述了深度学习技术的分类, 不同数据结构对应的最优神经网络架构, 黑盒模型的可解释性与预测结果的不确定性。介绍了深度学习在核物质状态方程、核结构、原子核质量、衰变与裂变方面的应用, 并展示如何训练神经网络预测原子核质量。结果发现使用实验数据训练的神经网络模型对未参与训练的实验数据拥有良好的预测能力。基于已有的实验数据外推, 神经网络对丰中子的轻原子核质量预测结果与宏观微观液滴模型有较大偏离。此区域可能存在未被宏观微观液滴模型包含的新物理, 需要进一步的实验数据验证。

关键词: 机器学习; 深度学习; 核结构; 核物质状态方程; 核裂变

中图分类号: O152; O414.3

文献标志码: A

DOI: 10.11804/NuclPhysRev.37.2019CNPC41

1 引言

机器学习 (Machine Learning) 是一类计算机算法的统称。这类算法不需要对特定任务编码, 能自动从大量数据中提取特征。深度学习^[1] (Deep learning) 是机器学习的子集, 自动继承了机器学习的定义。深度学习构造了一个网络化、参数化的泛函模块, 使用大量训练样本学习网络参数。传统的机器学习算法在面对海量数据时会出现性能饱和, 深度学习因为使用了层级化的多隐藏层, 可以学习层级化特征, 具有更强的表达能力。深度学习会随着训练数据的增多, 逐渐达到甚至超过人类在同类任务上的表现。不同研究者对深度学习的定义不同, 对于多数研究者来说, 深度学习等价于多隐藏层神经网络。

深度学习分为监督学习 (Supervised Learning)、非监督学习 (Un-supervised Learning)、半监督学习 (Semi-supervised Learning)、自监督学习 (Self-supervised Learning) 与强化学习 (Reinforcement Learning)^[2]。强化学习训练智能体, 通过观察环境, 作出动作, 获得奖励, 来与环境进行交互, 目标为最大化未来奖励之和^[3]。强化学习在物理学中应用不是非常广泛。物理学中应用最广泛的是监督学习, 训练深度神经网络, 在数据的输

入和标签之间建立函数映射, 目标是最大化网络预测成功的概率。其次是非监督学习, 通过聚类 (Clustering), 线性降维 (主成分分析 PCA) 或非线性降维 (tSNE), 自编码器 (Auto-Encoder) 以及生成对抗网络 (GAN), 提取数据中最重要的统计信息。最近的 Moco 类算法^[4] 使用对比损失, 通过最小化同类样本, 最大化异类样本潜在表示的夹角, 实现从图像中无监督提取特征, 使得七个下游任务达到与监督学习特征提取相近的准确度。使用较少但未来可能在物理学领域大放异彩的是半监督学习算法, 这种算法仅使用少量标注数据以及大量的未标注数据, 就将图像分类任务做到了与监督学习相近的正确率^[5-6], 一举破解了深度学习需要海量标注数据的魔咒。目前最好的半监督学习算法一般会使用自洽损失函数, 对无标注数据做微调增广, 要求神经网络做出自洽预测。自监督学习根据数据自身的规律, 生成训练样本, 比如在图像中加入噪声, 将原图作为标签, 训练神经网络去噪, 或随机遮挡语句中的词语, 训练神经网络根据上下文预测被遮挡部分^[7]。

一个广泛存在的误解是深度学习非凸优化, 损失函数容易困在局部极小值, 而非全局极小值。在实际应用中, 深度学习在处理各种复杂问题时, 总能找到不错的解。深度学习的三位图灵奖获得者之一 Yann Lecun 解

收稿日期: 2019-12-31; 修改日期: 2020-04-03

基金项目: 国家自然科学基金资助项目 (11861131009)

作者简介: 庞龙刚 (1984-), 男, 河南西峡人, 副教授, 博士, 从事核物理和机器学习研究; E-mail: lgpang@mail.ccnu.edu.cn。

释，一个高维参数空间存在局部极小值点的几率远远低于鞍点 (Saddle Point) 存在的几率。如图 1 所示，鞍点是指损失函数曲面上沿某些参数方向是极小值，沿另一些方向是极大值的点。只要还有一个方向是极大值点，最优参数总能继续向损失函数更低的方向逃逸。对于参数个数 n 为上百万的神经网络，假设沿某个参数方向损失函数是极小值的几率为 0.5，则沿所有方向都是极小值的概率为 0.5^n ，无限接近于 0。高维参数空间中一个零梯度的点为局域极小值的概率和其为鞍点的概率之比等于 $\lim_{n \rightarrow \infty} \left(\frac{0.5^n}{1-0.5^n} \right) = 0$ 。实际训练时，随机梯度下降引入的扰动和动量机制，都可帮助损失函数逃离鞍点。

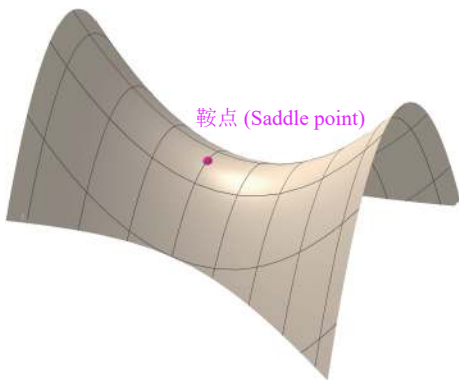


图 1 (在线彩图) 鞍点 (Saddle point) 示意图在高维参数空间，损失函数曲面上存在大量鞍点。

如图 2 所示，最简单的神经网络架构是密连神经网络，隐藏层的每个神经元与它前一层以及后一层的所有神经元都有连接，这种架构又被称为多层感知机 (Mul-

tilayer Perceptron)。这种神经网络架构简单，但网络参数众多，容易过拟合 (Overfitting) 到训练数据。限制每个神经元与相邻层仅存在局部连接，且不同神经元共享权重，可大大减少神经网络的参数个数，这种神经网络架构称作卷积神经网络 (Convolutional Neural Network)^[8-9]。深度卷积神经网络因为卷积核的局部连接，共享权重，平移不变性等特征，成为计算机视觉领域最强大的模式识别算法。卷积神经网络适用于可以转化为图像的输入数据。如果输入是树结构，则递归神经网络 (Recursive Neural Network) 是最适用的算法^[10]。如果输入数据是由很多顶点和边构成的有向图或无向图，则图神经网络 (Graph Neural Network，又称 Message Passing Network) 是适用的算法^[11-12]。递归神经网络和图神经网络被广泛应用于高能部分子喷注分类^[13-15]。如果输入是一大堆粒子，每个粒子包含四动量和种类信息，则点云神经网络 (Point Cloud Network) 是适用的架构^[16]。点云神经网络为每个粒子构造一个密连网络，不同粒子的密连网络共享同样的网络架构和参数，将粒子云从 m 维输入空间映射到 n 维潜在空间，然后对每个维度上的所有粒子实施重排不变操作 (Permutation Invariant Operations)，构造粒子在高维空间的关联。点云神经网络不仅用于重离子碰撞中核物质状态方程的识别^[17]，也用于喷注分类^[18]。如果输入数据是特征列表，可能非神经网络算法，比如提升树 (XGBoost) 是最适用的算法^[19]。选择机器学习算法时，应对不同的数据格式使用适合的神经网络架构。

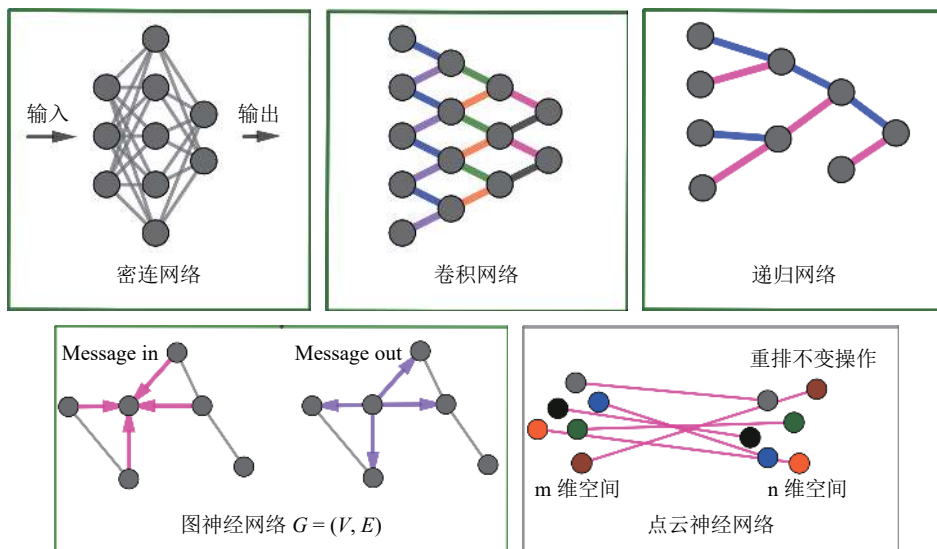


图 2 (在线彩图) 不同神经网络架构的对比

密连网络中每个神经元与相邻层所有神经元有连接，参数众多。卷积网络通过局部连接和共享权重，减少参数个数。递归神经网络适用于树结构的输入数据。图神经网络的输入包括节点和边，通过消息收集与分发更新节点与连边的参数。点云神经网络适用于核碰撞产生的存在于动量空间的大量粒子。除了密连网络与图神经网络，其他连线颜色一致的代表共享网络参数。

将深度学习技术应用于物理学研究时,其可解释性(Interpretability)^[20]与预测的准确度和不确定性同等重要。机器学习分为黑盒子(Black Box)与白盒子(White Box)模型。简单的线性分类器(Linear Classifier),单棵决策树(Decision Tree)属于白盒子模型。深度神经网络,由多棵决策树组成的随机森林(Random Forest)和提升树(Gradient Boosting Trees),则属于黑盒子模型。白盒子模型决策过程未超出人类理解能力范围,而黑盒子模型决策过程是复杂的非线性函数,人类在可解释性研究的帮助下只能部分理解。可解释性研究分为全局解释和局部解释。全局解释的目标是理解神经网络的每个神经元学到了什么特征,适用所有的输入样本。全局解释使用梯度上升(Gradient Ascent),将一张空白图片逐步更新以最大程度激活某个神经元,从而得到每个神经元学到的特征^[21]。局部解释的目标是理解神经网络为何对特定的输入作出当前判断。局部解释有很多种方法。第一种是局部替换法(Surrogate Method)^[22],将图片中的某个像素或超像素(颜色相近的相邻像素)替换成噪声,将修改后的图片输入到神经网络,观测神经网络预测结果发生的变化。预测结果改变的幅度,反映了被替换像素的重要程度。逐像素替换则会绘制出单张图片的重要性地图。第二种是基于梯度的显著法(Saliency Map)^[23]。假设神经网络是线性模型,则神经网络的输出对输入图片的梯度表示图片中不同区域的重要性。因为神经网络高度非线性,这种近似有很多问题。考虑到深度卷积神经网络的深层一般学习到抽象的、整体的概念,并大致保留相对位置信息,将高层的特征地图映射到输入图像上,则会提供整体概念上的重要性区域,这种方法称为类激活地图(Class Activation Map)^[24]。将高层特征与神经网络输出的相关性层层传递到输入图片,绘制出更精确的重要性地图,称为层间相关性传递算法(Layer-wise Relevance Propagation)^[25]。这些局域解释性方法也可以理解为神经网络的注意力机制。注意力区域提供了物理研究中非常重要的最相关相空间分布区域,可以启发科学家寻找更好的物理量。

2 深度学习在核物理中的应用

当前整个人类社会正经历第三次人工智能热潮。深度学习技术不仅引发了这次热潮,而且正引领并推动整个社会向更加智能化的方向发展。机器视觉、自然语言处理、语音识别、机器人控制、决策制定等等,深度学习进入的每一个领域,很快就会在学科最前沿超越传统

方法,甚至是超越人类。将深度学习技术应用于物理学,解决人类尚未解决的科学难题,是当前国际上前沿的交叉学科研究方法。接下来通过研究动机简单介绍深度学习在核物理领域的应用场景。

高能核碰撞是一个复杂的多阶段、多物理的非平衡态动力学演化过程。这些物理过程包含相对论效应引起的色玻璃凝聚与胶子饱和、夸克胶子等离子体(QGP)的产生和膨胀、强子化以及强子的散射与共振衰变等。实验最后测量到的,是大量末态粒子在动量空间的分布,而大家感兴趣的是初态核结构,核物质状态方程以及QGP的性质。从末态复杂的数据中提取对初态核结构以及状态方程最敏感的物理量非常困难。首先,QGP火球的寿命只有 10^{-23} s左右,半径也只有10 fm,无法直接探测。其次,耗散相对论流体力学与强子输运产生熵,存在信息丢失,无法保证重要信息编码在QGP复杂的动力学演化中,并在末态粒子分布中幸存了下来。使用深度学习一方面可以验证重要信息是否编码在复杂的末态数据中,另一方面可以从末态数据完全或部分解码初态信息,并提供末态数据中最重要的相空间区域。

2.1 使用相对论重离子碰撞提取核物质状态方程

高能核碰撞对应量子色动力学(QCD)相图的高温重子化学势附近区域。重子化学势表示正反物质比例,在宇宙早期和高能核碰撞中,产生了几乎等量的正反物质并以QGP的形态存在,重子化学势接近于零。格点QCD预言,在这个区域,QGP到强子共振气体的转变是光滑渐变。在中低能核碰撞,碰撞重合区域沉积了大量重子,正物质起主导,对应着高重子化学势,QGP到强子共振气体可能经历了一阶相变。QCD相变临界点将光滑渐变区与一阶相变区分隔开来。当前在中能与低能区做核碰撞的束流能量扫描,其最重要的物理目标就是寻找QCD相变临界点。

描述QGP时空演化的相对论流体力学对状态方程非常敏感。如果QGP到强子共振气体的转变是一阶相变,则压强作为能量密度的函数在Maxwell构造下有一个平台结构。此平台结构意味着QGP火球有一个压强梯度为零的壳层,在此壳层内,声速为零。由于QGP的膨胀由压强梯度驱动,因此在此壳层内QGP膨胀加速度为零,时空演化相对于其他区域缓慢。如果QGP到强子共振气体的转变是光滑渐变,则QGP火球内不存在压强梯度为零的壳层,时空演化当与一阶相变的状态方程不同。通过粒子的横动量谱与椭圆流区分光滑渐变与一阶相变的状态方程非常困难。使用深度卷积神经网络,将不同状态方程下相对论流体力学演化末态的粒子

分布作为神经网络的输入，将演化使用的核物质状态方程的种类作为标签，做监督学习，则会将寻找QCD相变临界点的任务转化为两个相变区域的分类问题^[26]。

2.2 使用相对论重离子碰撞提取原子核变形因子

原子核的结构影响着高能核碰撞的初态能量沉积以及能量密度分布在几何空间的各向异性。Voloshin^[27]在2010年的文章中展示，金核金核与铀核铀核的对心碰撞事例产生的粒子多重数与各向异性流的斜率不一致，使用传统的蒙特卡罗Glauber初态能量密度沉积机制无法描述此实验数据，而基于胶子饱和与组分夸克模型的能量密度沉积机制能够描述实验数据。因此，碰撞初态原子核的结构帮助排除了一种曾经广泛使用的能量密度沉积机制^[28]。

同质异位素 Zirconium($Z=40, A=96$)与 Ruthenium($Z=44, A=96$)这两种原子核质子数不同，核子数相等，最早被用在德国重离子研究中心(GSI)实验中研究核碰撞中的重子阻停。因为质子数不同，两种碰撞系统产生的电磁场有很大差别。又因为核子数相等，预期两个系统产生近似的集体流背景。2010年Voloshin^[27]建议使用此碰撞系统研究手征磁化效应。使用密度泛函，湖州师范学院的研究者计算了两种同质异位素的质子与中子在核内的分布。使用相对论流体力学，他们发现核碰撞产生的磁场与初态的几何空间各项异性对核子密度分布非常敏感，可用来区分基于密度泛函计算的核子分布与Woods-Saxon核子分布^[29]。

高能核碰撞中的一些未解之谜可能需要更精细的原子核结构来解释。其中一个谜题是超对心核碰撞的二阶谐振流与三阶谐振流的相对比值。实验测得在0~1%中心度时，二阶与三阶谐振流的数值基本相等。但描述高能核碰撞最成功的相对论流体力学模型，因为剪切黏滞对高阶谐振流的压低更为强烈，无法同时描述实验测量的二阶与三阶谐振流。传统的蒙特卡罗初始条件比较简化，只考虑了单核子的Woods-Saxon分布，而忽略了核子核子之间的关联。2014年的一篇文章考虑了核子核子之间的关联，发现能够部分解决超对心碰撞之谜^[30]。

核结构为相对论重离子碰撞末态粒子在动量空间的分布带来可观效应。从核碰撞末态复杂的数据，能否提取参与碰撞原子核的结构信息，比如核变形因子、质子中子分布以及核子核子关联？变形核的形状可以由变形的Woods-Saxon分布近似描述，改变变形因子，原子核的形状从南瓜变化为纺锤形。变形核碰撞时，由于随机欧拉旋转，为各向异性流带来相比于球形核碰撞更

大的方差。变形核的碰撞产生复杂的粒子多重数与各向异性流之间的关联。机器学习中的深度神经网络被验证可使用核碰撞的蒙特卡罗模拟，从这些复杂的关联中，提取核碰撞初态的原子核变形因子的绝对值^[31]。

2.3 使用深度学习加速相对论流体力学模拟

使用相对论流体力学模拟高能核碰撞需要极大的CPU或GPU计算开销。这非常不利于高能核碰撞领域积攒大规模的蒙特卡罗模拟数据，研究末态可观测量对各种物理参数的依赖。北京大学课题组使用StackedUNet近似相对论流体力学方程的解^[32]，无论是定性还是定量分析，都得到很好的结果。同时，基于神经网络的求解方案比基于差分近似的数值解方案加速了600倍。即便考虑到CPU版本的流体力学到GPU版本时的60倍加速，基于神经网络的求解方案仍然要快10倍左右。

除了基于UNET这样的图像翻译算法，还有一种算法可以用来求解偏微分方程，这种算法被称为“懂物理的神经网络”(Physics Informed Deep Neural Network)^[33]。这种方法将神经网络用作物理模型的一个模块，通过梯度下降和误差的向后传递更新神经网络参数，从而使得整个物理模型成为数据驱动模型，使解析解未知或数值解未知的部分更加健壮。这种方法也被用作数值求解偏微分方程，一般偏微分方程分为初值问题与边值问题。神经网络提供了参数化的偏微分方程变分解，通过最小化变分解对偏微分方程本身以及对初始条件和边界条件的违背程度，学习偏微分方程的近似解。因为深度神经网络可以近似任何连续函数，神经网络对输入的各阶微分可解析计算，这种方法有很广的应用前景。

3 深度学习在低能核物理中的应用

3.1 使用贝叶斯神经网络研究原子核质量与核裂变碎片产额分布

传统的神经网络不能给出预测结果的不确定性。将传统神经网络神经元之间的连接权重从固定的数值变为概率分布，则会得到贝叶斯神经网络架构(Bayes Neural Network, BNN)。BNN将网络参数的分布视为先验，计算网络输出与标签之间的似然函数，通过贝叶斯公式学习后验分布，使用后验分布对神经网络参数的积分，得到神经网络预测的均值和不确定性。如果使用简单的高斯函数先验，则BNN相当于使用两倍于传统神经网络的参数，学到了无穷多神经网络的系综。最近，BNN被广泛用于低能核物理研究领域，精确预测原子

核质量，贝塔衰变寿命以及原子核的裂变碎片产额分布 [34-37]。这对于理解原子核内核子的多体相互作用与核反应至关重要。

贝叶斯神经网络的缺点之一在于训练过程复杂，需要高昂的计算花销。一种有效的替代方法是对传统神经网络使用蒙特卡罗 Dropout。Dropout 指的是按概率随机丢掉部分神经元，在训练时相当于得到多个神经网络，形成系综。预测时打开 Dropout，对同一个输入计算得到多个输出，其均值与方差被证明可以用来量化模型的不确定性 [38-39]。这种方法在不降低预测精度和泛化能力，不增加计算开销的同时，提供了模型预测的不确定性。

3.2 深度学习拟合原子核质量

在原子核质量预测的有限程液滴模型中，基态质量与质子数 Z 和中子数 N 的关系通过复杂的表面能、体积能、库仑能、自旋轨道耦合修正以及对修正等构建 [40]。

基于 RBF 和 BNN 算法，文献 [34, 36] 已经将原子核质量拟合到了前所未有的精度。这一节简单举例，对比提升树 (Gradient Boosting Trees) 和深度神经网络 (Deep Neural Network)，在拟合原子核基态质量盈余上的表现。使用的算法分别由 Sklearn [41] 和 Tensorflow [42] 中的 Keras 模块构造，原子核质量数据由文献 [40] 提供。使用全连接神经网络，输入层两个神经元，对应归一化的 (Z, N) ，输出层一个神经元，表示原子核质量，中间 3 个隐藏层，神经元个数分别为 (1024, 256, 128)。使用 ReLU 激活函数和 Adam(lr=0.001, beta_1=0.9, beta_2=0.999) 优化器，训练 100 个 Epoch。

图 3 使用实验上测量得到的原子核中的 90%(1934 样本) 做训练，10%(215 个样本) 做测试，可以看到神经网络在 215 个测试样本上的预测误差为 1.6 MeV。在这个小数据集上，神经网络的泛化能力仍然高于传统的机器学习算法提升树。

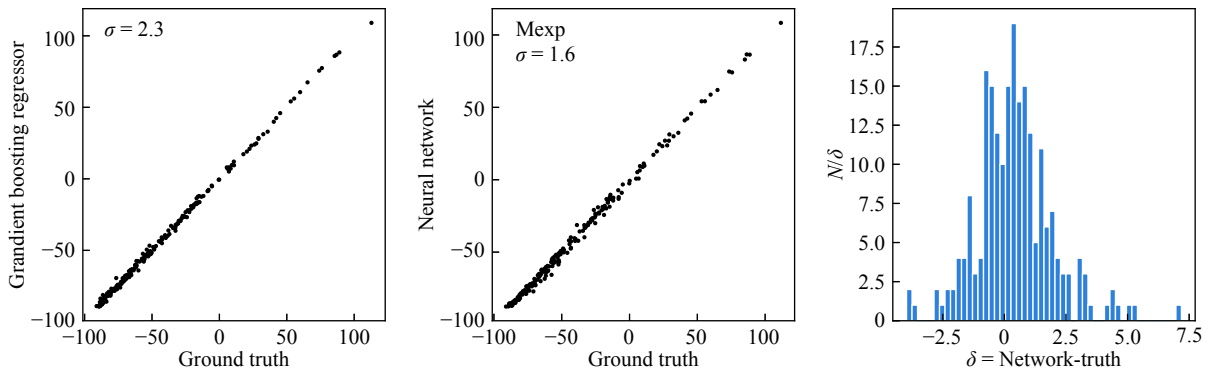


图 3 (在线彩图) 分别使用提升树 (a) 和 4 层神经网络 (b)，对 1934 个实验数据样本做训练，215 个实验数据样本做测试，得到的预测结果及误差，(c) 为神经网络预测结果与真值差值的分布

图 4 使用实验数据做训练，对没有测量数据的原子核进行预测，并与理论预测结果对比。提升树在训练区间之外的预测结果与宏观微观液滴模型预测结果相差甚远，基于神经网络的模型比提升树的结果略好，但与理

论预测结果的均方差 σ 也高达 65 MeV 以上。从结果来看，提升树倾向于将新的重核的质量盈余约束在训练数据的质量区间 (<150 MeV)。神经网络对超出训练数据的重原子核，也倾向于给出更小的质量，但最大预测质

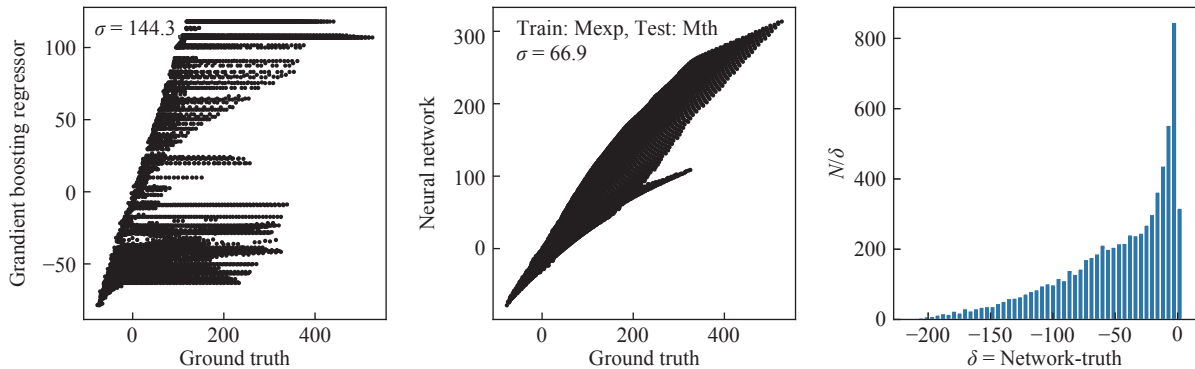


图 4 (在线彩图) 分别使用提升树 (a) 和 4 层神经网络 (b)，对 1934 个实验数据样本做训练，7169 个无实验测量数据的样本做测试，得到的预测结果及与理论预言对比的误差，(c) 为神经网络预测结果与理论预言差值的分布

量并不局限于训练数据。最有启发性的结果在于，图 4 中神经网络的预测结果，与理论预言在 200 MeV 的时候明显分成两支。检查这两个分支对应的原子核的 (N, Z) 。发现下方分支对应 N 大于 $3Z$ 的轻核，比如 $(N, Z)=(72, 21)$, $(79, 24)$, $(65, 18)$ ，而上方分枝对应 N 小于或约等于 $2Z$ 的重核，比如 $(N, Z)=(175, 75)$, $(185, 82)$, $(190, 87)$ 。在实验测量的原子核质量数据中，缺少下方分枝中极端丰中子的轻原子核。基于已有的实验数据外推，神经网络对这些丰中子的轻原子核作

出了不同于宏观微观液滴模型的预测。

图 5 使用宏观微观模型的计算结果作为训练和测试数据。神经网络给出的预测结果与理论预言的误差为 1.4 MeV 左右。神经网络从理论预测数据中学到了各种能量对原子核质量的总贡献。在神经网络学习的过程中，并没有对不同的 N, Z 计算核变形因子，也未基于不同的变形核计算库仑能和壳修正。使用机器学习技术对大量的核反应实验数据分析，或可发掘出未被理论模型捕获的物理。

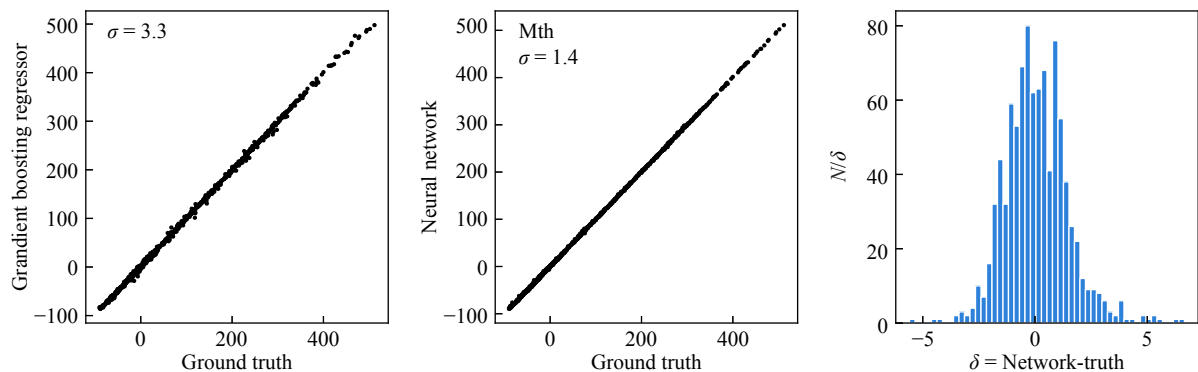


图 5 (在线彩图) 分别使用提升树 (a) 和 4 层神经网络 (b)，对 8386 个理论计算样本做训练，932 个理论计算样本做测试，得到的机器学习预测结果及其与理论预言的绝对误差，(c) 为神经网络预测结果与理论预言差值的分布

4 总结

人工智能领域正因为深度学习技术而发生快速而深刻的变革。监督、非监督、半监督、自监督、强化学习等种种学习方法；密连、卷积、递归、图卷积、点云、贝叶斯等种种神经网络架构；梯度上升、局部替换、类激活地图、层间相关性传递等等可解释性方法；BNN、蒙特卡罗 Dropout 等多网络系统估计模型误差的方法，可以从各个方面协助核物理领域的科学家，从大量核碰撞、核质量与核裂变数据中提取核物质状态方程、核结构等信息，加深对核子之间多体相互作用和裂变动力学机制的理解。使用监督学习在复杂的末态数据与物理模型的参数空间建立映射，可以帮助寻找对某些物理属性最为相关的特征。这种自动化的知识发现能力，对于物理信号的获取有很强的启发作用。

参考文献：

- [1] LECUN Y, BENGIO Y, HINTON G. *Nature*, 2015, 521: 436.
- [2] GOODFELLOW I, BENGIO Y, COURVILLE A. *Deep Learning*[M]. Cambridge: MIT Press, 2016.
- [3] SUTTON R S, BARTO A G. *Reinforcement Learning: An Introduction*[M]. 2nd ed. Cambridge: MIT Press, 2018.
- [4] HE K, FAN HQ, WU Y X, et al. arXiv: 1911.05722.
- [5] BERTHELOT D, CARLINI N, GOODFELLOW I, et al. arXiv: 1905.02249.
- [6] XIE Q, DAI Z, HOVY E, et al. arXiv: 1904.12848.
- [7] JASON. *Awesome Self-Supervised Learning*[EB/OL]. [2020-02-21]. <https://github.com/jason718/awesome-self-supervised-learning>.
- [8] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. *ImageNet Classification with Deep Convolutional Neural Networks*[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems, 2012, Lake Tahoe, Nevad. Red Hook, New York: Curran Associates Inc, 2012: 1097.
- [9] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. arXiv: 1207.0580.
- [10] SOCHER R, LIN C C, MANNING C, NG A Y. *Parsing Natural Scenes and Natural Language with Recursive Neural Networks*[C]//Proceedings of the 28th International Conference on Machine Learning, ICML 2011. Washington: Omnipress, 2011.
- [11] KIPF T N, WELING M. arXiv: 1609.02907.
- [12] GILMER J, SCHOENHOLZ S S, RILEY P F, et al. arXiv: 1704.01212.
- [13] LOUPPE G, CHO K, BECOT C, et al. *Journal of High Energy Physics*, 2019(01): 057.
- [14] CHENG T. *Comput Softw Big Sci*, 2018, 1: 2.
- [15] HENRION I, BREHMER J, BRUNA J, et al. *Neural Message Passing for Jet Physics*(2017NIPS)[EB/OL]. [2020-02-21]. https://dl4physicsciences.github.io/files/nips_dlps_2017_29.pdf.
- [16] QI C R, SU H, MO K, et al. arXiv: 1612.00593.

- [17] STEINHEIMER J, PANG L G, ZHOU K, et al. *Journal of High Energy Physics*, 2019(12): 122.
- [18] KOMISKE P T, METODIEV E M, THALER J. *Journal of High Energy Physics*, 2018(04): 013.
- [19] CHEN T, GUESTRIN C. arXiv: 1603.02754.
- [20] CHRISTOPH M. Interpretable machine learning. A Guide for Making Black Box Models Explainable, 2019[EB/OL]. [2020-02-21]. <https://christophm.github.io/interpretable-ml-book/>.
- [21] OLAH C, SATYANARAYAN A, JOHNSON I. *Distill*, 2018, 3: 10.
- [22] RIBEIRO, TULIO M, SINGH S, et al. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August, 2016[C]. San Francisco: Association for Computing Machinery New York United States, 2016.
- [23] SIMONYAN K, VEDALDI A, ZISSERMAN A. arXiv: 1312.6034
- [24] ZHOU B, KHOSLA A, LAPEDRIZA A, et al. arXiv: 1512.04150.
- [25] BACH S, BINDER A, MONTAVON G, et al. *PLOS ONE*, 2015, 10(7): e0130140.
- [26] PANG L G, ZHOU K, SU N, et al. *Nature Comm*, 2018, 9: 210.
- [27] VOLOSHIN S A. *Phys Rev Lett*, 2010, 105: 172301.
- [28] ADAMCZYK L, ADKINS J K, AGAKISHIEV G, et al. *Phys Rev Lett*, 2015, 115: 222301.
- [29] XU H J, WANG X B, LI H L, et al. *Phys Rev Lett*, 2018, 121: 022301.
- [30] DENICOL G S, GALE C, JEON S, et al. arXiv: 1406.7792 [nucl-th]
- [31] PANG L G, ZHOU K, WANG X N. arXiv: 1906.06429 [nucl-th].
- [32] HUANG H, XIAO B, XIONG H, et al. *Nucl Phys A*, 2019, 982: 927.
- [33] RAISSI M, PERDIKARIS P, KARNIADAKIS G E. arXiv: 1711.10561.
- [34] NIU Z M, LIANG H Z. *Physics Letters B*, 2018, 778: 002.
- [35] NIU Z M, LIANG H Z, SUN B H, et al. *Physical Review C*, 2019, 99: 064307.
- [36] NIU Z M, FANG J, NIU Y F. *Physical Review C*, 2019, 100: 054311.
- [37] WANG Z A, PEI J, LIU Y, QIANG Y. *Phys Rev Lett*, 2019, 123: 122501.
- [38] GAL Y, GHAHRAMANI Z. arXiv: 1506.02142
- [39] PEARCE T, LEIBFRIED F, BRINTRUP A, et al. arXiv: 1810.05546
- [40] MÖLLER P, SIERK A J, ICHIKAWA T, et al. *Atom Data Nucl Data Tabl*, 2016, 109-110: 1.
- [41] PEDREGOSA F, VAROQUAUX G, GRAMFORT A, et al. *Journal of Machine Learning Research*, 2011, 12: 2825.
- [42] ABADI M, AGARWAL A, BARHAM P, et al. arXiv: 1603.04467.

Deep Learning for Nuclear Physics

PANG Longgang^{1,2,1)}, ZHOU Kai³, WANG Xinnian^{1,2,4}

(1. Key Laboratory of Quark and Lepton Physics of Ministry of Education, Central China Normal University, WuHan 430079, China;

2. Institute of Particle Physics (IOPP), Central China Normal University, WuHan 430079, China;

3. Frankfurt Institute For Advanced Studies, Frankfurt 60438, Germany;

4. Lawrence Berkeley National Laboratory, Berkeley 94706, USA)

Abstract: Deep learning is the state-of-the-art pattern recognition method. It is expected to help scientists to discover most relevant features from big amount of complex data. Different categories of deep learning, the best deep neural network architectures for different data structures, the interpretability of black-box models and the uncertainties of model predictions are reviewed in this article. The applications of deep learning in nuclear equation of state, nuclear structure, mass, decay and fissions are also introduced. In the end, a simple neural network is trained to predict the mass of nucleus. We found that the artificial neural network trained on experimental data has low prediction error for experimental data that are held back. Trained with experimental data, the network predictions for light neutron-rich nuclei deviate from Macro-Micro Liquid model, which indicate that there might be new physics missing in the theoretical model and more data are needed to verify this.

Key words: machine learning; deep learning; nuclear structure; nuclear equation of state; nuclear fission

Received date: 31 Dec. 2019; Revised date: 03 Apr. 2020

Foundation item: National Natural Science Foundation of China(11861131009)

1) E-mail: lgpang@mail.ccnu.edu.cn.