

我系杜江峰、王亚、李兆凯等成功研发新型量子机器学习技术实现数据特征提取

发布时间：2021-09-09 浏览次数：71

我院中国科学院微观磁共振重点实验室杜江峰、王亚、李兆凯等人在量子机器学习研究中取得重要进展，研发出新型量子特征提取算法，实验实现了对未知量子系统矩阵的分析与信息提取。该成果以“Resonant Quantum Principal Component Analysis”为题发表在近期的Science Advances上[Science Advances 7, eabg2589 (2021)]。

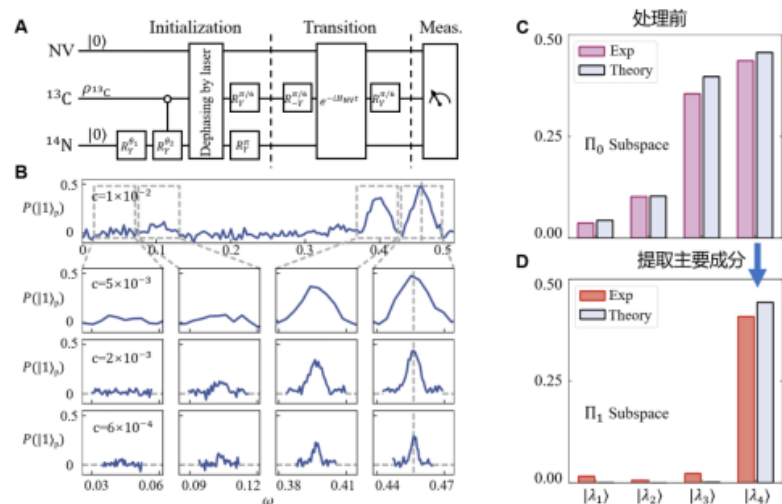
机器学习是指使用计算机从大量历史数据中挖掘隐含规律，并用于后续预测或者分类的过程。机器学习是人工智能的核心，是计算机具有智能的根本途径，其理论和方法已被广泛应用于解决日常应用和科学领域的复杂问题。为了成功完成特定任务，人工智能往往需要大量数据用于总结与分类，这对计算机系统的存储与处理能力提出了很高的要求。量子机器学习可以将量子算法的并行加速特性应用于人工智能领域中，提升人工智能系统的效率与能力，有望在未来实现基于量子系统的人工智能。

杜江峰院士团队自2012年以来率先开展了量子人工智能的实验研究，相关工作如量子手写识别[PRL 114, 140504 (2015)]是量子人工智能应用于实际问题的最早案例，展示了量子技术加速人工智能问题的潜力；特征值检测[PRL 122, 090504 (2019)]、线性方程组求解[PRA 89, 022313 (2014)]等技术为机器学习中的数据运算提供了快速有效的量子方法。此前的工作及国际上的相关实验研究，多集中在如何处理较理想的数据集。但无论是使用经典还是量子计算机进行机器学习，在获得类似数据集之前都需要对原始数据进行分析 and 预处理，提取出其中的核心信息用以学习与总结规律。这一过程被称之为数据特征提取，是量子人工智能运行的关键步骤。

使用量子算法进行特征提取的理论思路最早于2014年提出，但其原始设想基于量子相位估计算法，需要大量量子比特作为辅助寄存器，因此一直未能在真实实验体系中予以实现。为解决这一限制，本文研究团队开发出新型基于共振的量子主成分分析技术，将辅助量子比特的需求降低到1个，大大降低实验难度。同时，为减少实际实验中的噪声干扰，该技术还可以结合量子相干保护手段，有利于在实际量子处理器物理平台上达到高精度与高效率的量子计算。

实验中研究人员使用金刚石氮-空位色心量子处理器，演示了对未知量子数据矩阵进行分析与处理的过程。待研究的数据以量子密度矩阵的形式被输入量子处理器，如图C所示，该数据矩阵包含4种不同成分且占比各不相同（对应数据柱高度）。特征提取任务的目标是将该数据的关键特征，即右侧第一组占比最高的成分提取出来，同时尽量去除其他三组数据或噪声。实验中，研究人员使用一个辅助比特作为探针进行扫

描，精确定位了密度矩阵中不同成分的强度（图B中的不同谱峰）。通过多次迭代逼近，密度矩阵成分的定位误差被降低到小于0.001，相当于原本10个辅助量子比特才能达到的精度。其后，研究人员锁定该数据矩阵的主要成分并将其隔离提取出来（图D），得到的量子态即为输入数据矩阵的关键特征。实验结果显示，这一特征提取过程达到了90%的提取精度与86%的提取效率，展示了该新技术在真实物理平台上的适用性与精确性。



图：(A)金刚石量子处理器上实现共振量子主成分分析的线路图；

(B)通过多次迭代，精确定位量子数据矩阵中不同成分的比例与信息；

(C)处理前的原始量子数据；(D)数据矩阵中的关键部分（右侧第一列）被单独提取并储存。

研究结果显示此次研发的新技术可以实现对数据预处理过程的量子加速，高效率提取出量子数据矩阵中的关键特征，用于后续进一步分类与识别。该技术能够提升机器学习的效率和效果，未来有望在较大规模量子处理器上得到应用。

中国科学院微观磁共振重点实验室副研究员李兆凯、博士生柴梓华为该文共同第一作者，杜江峰院士和王亚教授为该文共同通讯作者。该研究得到了科技部、国家自然科学基金委、中国科学院和安徽省等资助。

论文链接：<https://advances.sciencemag.org/content/7/34/eabg2589>

(中国科学院微观磁共振重点实验室、物理学院、合肥微尺度物质科学国家研究中心、中国科学院量子信息和量子科技创新研究院、科技部)

请使用谷歌Chrome、Edge、双核浏览器极速模式 1366x768以上分辨率访问本站

版权所有：中国科学技术大学物理学院 工程与应用物理系 通讯地址：安徽省合肥市包河区金寨路96号