



稳健的量子重上传分类器

韩思予, 贾林瀚, 李宇峰

引用本文:

韩思予, 贾林瀚, 李宇峰. 稳健的量子重上传分类器[J]. *智能系统学报*, 2026, 21(3): 617-626.

HAN Siyu, JIA Linhan, LI Yufeng. Robust data re-uploading quantum model[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(3): 617-626.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202507029>

您可能感兴趣的其他文章

一种卷积神经网络集成的多样性度量方法

Diversity measuring method of a convolutional neural network ensemble

智能系统学报. 2021, 16(6): 1030-1038 <https://dx.doi.org/10.11992/tis.202011023>

记忆神经网络在机器人导航领域的应用与研究进展

Research progress and application of memory neural network in robot navigation

智能系统学报. 2020, 15(5): 835-846 <https://dx.doi.org/10.11992/tis.202002020>

图像情境下的数字序列逻辑学习

Number sequence logic learning in image context

智能系统学报. 2019, 14(6): 1189-1198 <https://dx.doi.org/10.11992/tis.201905044>

关于深度学习的综述与讨论

Overview on deep learning

智能系统学报. 2019, 14(1): 1-19 <https://dx.doi.org/10.11992/tis.201808019>

BP神经网络和支持向量机相结合的电容器介损角辨识

Capacitor dielectric loss angle identification based on a BP neural network and SVM

智能系统学报. 2019, 14(1): 134-140 <https://dx.doi.org/10.11992/tis.201805034>

多标记学习自编码网络无监督维数约简

Unsupervised dimensionality reduction of multi-label learning via autoencoder networks

智能系统学报. 2018, 13(5): 808-817 <https://dx.doi.org/10.11992/tis.201804051>

DOI: 10.11992/tis.202507029

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20260316.1711.002>

稳健的量子重上传分类器

韩思予^{1,2}, 贾林瀚^{1,2}, 李宇峰^{1,2}

(1. 南京大学 计算机软件新技术全国重点实验室, 江苏 南京 210023; 2. 南京大学 人工智能学院, 江苏 南京 210023)

摘要: 针对数据重上传模型在实际应用中仍面临严重的过拟合与超参数敏感性问题, 本研究提出了稳健重上传分类器。它通过一种创新的损失函数设计, 在传统的量子态保真度基础上引入了正则化项, 从而有效抑制过拟合现象并增强模型的泛化性能与稳健性。实验结果充分证明了稳健重上传分类器在降低过拟合风险、增强模型稳健性方面的显著优势。这种创新不仅为量子机器学习中损失函数的设计开拓了新路径, 也为分类任务的泛化能力研究提供了新的实验依据, 为未来量子机器学习的进一步发展奠定了实践基础。

关键词: 量子机器学习; 数据重上传; 正则化; 过拟合; 稳健性; 量子计算; 量子神经网络; 变分子量子电路

中图分类号: TP18 文献标志码: A 文章编号: 1673-4785(2026)03-0617-10

中文引用格式: 韩思予, 贾林瀚, 李宇峰. 稳健的量子重上传分类器 [J]. 智能系统学报, 2026, 21(3): 617-626.

英文引用格式: HAN Siyu, JIA Linhan, LI Yufeng. Robust data re-uploading quantum model [J]. CAAI transactions on intelligent systems, 2026, 21(3): 617-626.

Robust data re-uploading quantum model

HAN Siyu^{1,2}, JIA Linhan^{1,2}, LI Yufeng^{1,2}

(1. State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China; 2. School of Artificial Intelligence, Nanjing University, Nanjing 210023, China)

Abstract: To address the severe overfitting and hyperparameter sensitivity problem of the data re-uploading model, we introduce the robust quantum re-uploading classifier, featuring an innovative loss function that combines quantum state fidelity with a strategically designed regularization term. This combination effectively reduces overfitting and considerably enhances model stability and generalization. Experimental results on benchmark datasets clearly demonstrate the classifier's superiority over existing methods, particularly in mitigating overfitting and sustaining robust predictions. Consequently, this novel strategy not only opens new avenues for designing loss functions in QML but also provides empirical insights into generalization capabilities. It establishes a strong foundation for future quantum classification research and applications.

Keywords: quantum machine learning; data re-uploading; regularization; overfitting; robustness; quantum computing; quantum neural network; variational quantum circuit

随着量子计算的迅猛发展, 量子机器学习 (quantum machine learning, QML) 逐渐成为一个前沿且极具潜力的交叉研究领域。量子机器学习通过将量子计算的优势与经典机器学习方法相结合, 旨在利用量子叠加、纠缠以及干涉等独特的量子效应, 高效地处理高维度数据, 实现复杂函数的精准逼近, 并在优化问题求解方面展现显著优势。这种融合不仅有望在计算效率和速度上实现突破, 更可能在经典算法遇到计算瓶颈的场景

下提供全新的解决方案。

作为量子机器学习的一个典型代表, 数据重上传模型通过重复编码的方式将数据重载入量子机器学习模型中, 绕开了量子计算中的不可克隆定理的限制, 将经典数据多次上传到同一个量子比特当中。若将每一次的编码视为一个神经元, 则万能逼近定理可以保证其理论上强大的泛化潜力。过去的研究也针对该模型作出了许多改进, 其中针对量子重上传模型, 通过统合多种数据编码设计的多嵌入重上传模型, 克服了单一的数据编码导致的泛化能力的限制, 在分类任务的性能上获得了巨大的提升。

收稿日期: 2025-07-27. 网络出版日期: 2026-03-18.

基金项目: 国家自然科学基金项目 (62576162); 中央高校基本科研业务费专项经费项目 (022114380023).

通信作者: 李宇峰. E-mail: liyf@nju.edu.cn.

然而,在实际应用中,重上传模型依然面临诸多挑战,过拟合问题便是其中之一。过拟合是指机器学习模型因过度拟合训练数据噪声导致泛化能力下降的现象。由于量子机器学习模型通常具有强大的表达能力,能够非常精确地拟合训练数据中的量子态和对应的经典特征,这种高表达能力极易导致过拟合,限制了量子机器学习模型在实际问题中的应用效能。

此外,多嵌入重上传模型尽管具有强大的泛化能力,但未能解决自身对于层数这一超参数的敏感性问题。超参数敏感使得模型调参变得更加困难,往往需要大量的实验和尝试来找到最合适的超参数组合。因此,寻找一个能够有效规避过拟合并提高稳健性的模型结构,具有重要的研究意义。

本研究针对上述问题,提出了稳健重上传量子模型(robust multi-embedding data re-uploading quantum model, ROME)。它提供了一种创新的误差设计方案,以明确优化目标并改善量子机器学习模型的泛化性能。实验结果表明,该方法不仅能够显著降低量子模型的过拟合现象,而且在不同的机器学习任务中均展现出卓越的泛化效果。

1 相关工作

量子机器学习近年来成为量子计算与经典机器学习相结合的一个重要研究领域^[1-5]。许多传统的机器学习方法的量子形式被开发出来^[6-9]。分为以量子支持向量机为代表的隐式核方法^[10],以及由变分量子线路构成的显式量子模型^[11-12]。显式量子模型一般由以下几个部分组成:数据编码^[13]、量子电路^[14]以及测量部分^[15]。针对以下这几个组成部分的优化一直受到研究者们广泛关注^[16]。

量子数据编码(quantum data encoding)是量子机器学习中的关键技术之一。早期的研究主要集中在如何将经典数据映射到量子比特上,以实现量子信息处理的高效性。例如, Kitaev 提出的量子支持向量机(quantum support vector machine, QSVM)^[17]和 Rebentrost 等^[18]提出的量子分类方法,均是通过量子态的叠加与干涉对经典数据进行映射和处理,从而加速数据分类和回归任务的计算。常见的方法有振幅编码、相位编码、量子近似优化算法编码^[19]等。

随着量子硬件的不断发展,量子神经网络(quantum neural networks, QNNs)成为量子机器学习量子电路部分的一个研究热点。量子神经网络结构采用量子比特作为节点,通过量子门实现神

经元间的复杂非线性关系。Harrow 等^[20]首次提出了量子神经网络的基本框架,其中量子比特通过参数化量子电路进行计算,并展示了量子神经网络在解决某些类型的问题时具有比经典神经网络更优的计算效率。此外,量子卷积神经网络(quantum convolutional neural networks, QCNNs)^[21]被提议用于图像识别等任务,利用量子叠加态的并行性提高计算能力。

在测量方面,有一些工作^[22-23]希望通过对量子模型的测量建立阴影模型,从而希望利用经典机器学习模型在特定任务上实现量子优势。除此以外,也有一些量子机器学习和经典机器学习之间关系的探究^[24-27]。

此外,数据重上传(data re-uploading)方法在量子机器学习中也得到了广泛的关注。该方法通过多次上传经典数据到量子电路中,从而增强量子模型的表达能力,尤其在量子计算资源有限的情况下,能够有效提高模型的表现力和泛化能力。该策略已被应用于多个量子神经网络架构中,并取得了显著的效果^[28-29]。例如,先前的工作中提出的一种基于多重数据上传的量子机器学习框架——多重编码数据重上传^[30]。其通过结合多种数据嵌入方法,充分利用数据编码的多样性,从而提升了模型的泛化能力。通过实验验证,多重编码数据重上传相较于现有的最先进方法在多个数据集上均表现出显著的性能提升,尤其在面对复杂数据集时,展现了更强的适应能力和更高的准确性。

2 预备知识

2.1 数据重上传模型

数据重载模型,由 Pérez-Salinas 等^[26]提出,目的是实现使用最小量子资源的通用量子分类器。其核心理念在于,针对单个量子比特的计算空间有限(只有两个自由度)的问题,模型在量子线路中多次“重上传”数据,从而近似复杂的分类函数。数据重上传绕过了量子计算中的不可克隆定理(不允许直接复制量子数据),通过在多层量子线路中重新引入经典数据实现这一目标。

该方法的理论基础可以与经典神经网络中的多层感知机(multilayer perceptron, MLP)和通用逼近定理(universal approximation theorem)相类比,其中单一的量子比特通过多次数据重上传实现对连续函数的近似,从而增强模型的表达能力。

数据重上传模型的量子线路包含以下几个关键组件。

1) 数据编码: 将经典输入数据 $\mathbf{x} \in \mathbf{R}^d$ 编码到量子态中, 使用参数化的单比特旋转门 $U(\mathbf{x})$ 。

2) 门控网络层: 量子线路由多层网络组成, 每一层包括两个量子门部分:

$$L(i) = U(\mathbf{x})U(\theta_i)$$

式中: θ_i 是可调参数, 数据 \mathbf{x} 在每一层都会被重新上传。

尽管数据重上传模型在理论上展现出强大的表达潜力, 但在实际应用中, 其性能却未能完全达到预期, 尤其是在处理特定类型的数据集 (如线性可分数据) 时表现出明显的不足。具体而言, 当模型的网络层数较少时, 数据编码对决策边界的影响将愈发显著, 导致模型难以有效捕捉数据中的关键特征, 从而无法准确地拟合决策边界。此外, 随着模型结构变得更浅, 数据重上传模型对输入特征的有效表达受到进一步的限制, 这在一定程度上阻碍了其对更复杂数据模式的有效特征提取和泛化。这种现象表明, 数据重上传方法尚未充分挖掘数据嵌入策略的潜力, 因此有必要进一步优化数据编码机制, 以克服上述局限性。

2.2 多重编码数据重上传量子模型

为了解决传统数据重上传模型泛化能力不足的问题, 多重编码数据重上传量子模型通过将多个数据嵌入方法结合在一起, 以提升模型的泛化能力。具体而言, 多重编码数据重上传量子模型 (multi-embedding data re-uploading quantum model, MEDQ) 在数据重上传过程中引入了多个数据编码策略, 使得模型能够充分利用数据编码的多样性, 从而在处理不同类型的任务时获得更好的表现。

在多重编码数据重上传量子模型中, 经典数据首先通过多种不同的量子数据编码方式进行嵌入, 例如旋转编码 (Rotation Encoding)、量子近似优化算法 (quantum approximate optimization algorithm, QAOA) 编码以及角度编码 (Angle Encoding) 等。这些编码策略能够从不同角度提取数据特征, 从而避免了单一数据编码方式所带来的局限性。每种嵌入方式的输出被传递到量子电路中进行处理, 并通过量子旋转门进行量子态的更新。由于每种数据嵌入策略的参数均为可训练的, 模型能够通过训练过程优化这些参数, 从而提高对复杂数据集的拟合能力。

多重编码数据重上传量子模型通过多层数据上传的方式进一步增强其泛化能力。每一层的量子电路都采用不同的嵌入策略, 这使得每一层都能提取输入数据的不同特征。通过这种多层次的

嵌入和上传, MEDQ 能够处理更为复杂的非线性数据关系, 提高了模型在实际应用中的表现。与传统的数据重上传模型相比, MEDQ 在多个基准数据集上的实验结果显示出显著的性能提升。

MEDQ 在实验中表现出优于现有方法的泛化能力, 特别是在处理线性可分数据集时性能显著提高。然而, 尽管 MEDQ 结构在实验层面表现出色, 但仍存在过拟合风险, 尤其当量子线路层数增加时该问题尤为明显。因此, 针对 MEDQ 模型提出有效的正则化方法具有重要的研究价值, 能够进一步提升量子机器学习模型的稳健性。基于上述观察, 我们提出了一种新颖的正则化优化策略, 以改善 MEDQ 模型的过拟合问题, 在保证模型泛化能力的同时增强模型的稳健性。

3 稳健重上传量子模型

在量子机器学习领域, 特别是在涉及数据重上传架构的模型中, 过拟合现象已逐渐显现为一个普遍存在却鲜有深入探讨的关键问题。尽管该现象严重制约了量子机器学习模型在实际应用中的性能与可靠性, 但当前学术界尚未构建出针对数据重上传模型过拟合现象的全面分析框架, 亦未提供有效的解决方法。这一重要且尚未填补的研究空白不仅极大地限制了量子机器学习模型泛化性能的提升, 更本质性地制约了此类模型在复杂实际环境中的应用效能与推广价值。如图 1 所示, 其中蓝色为传统重载模型, 红色为 MEDQ, 实线为训练准确率, 虚线为测试准确率。该图展现了过拟合现象普遍存在于重载模型族中, 并且随着层数的增加过拟合现象越明显; 同时 MEDQ 也出现了训练不稳定的问题。

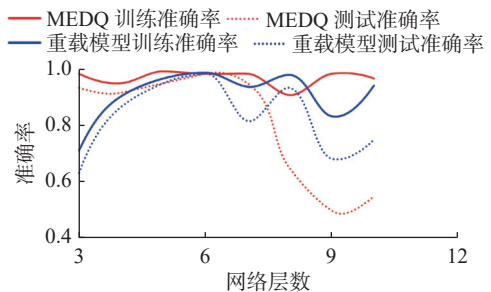


图 1 重上传模型的过拟合现象
Fig. 1 Overfitting phenomenon in re-uploading models

针对这一关键性问题, 本研究尝试从理论与实验两个维度对数据重上传量子机器学习模型的过拟合机制展开探究。具体而言, 我们创新性地提出了一种加权量子态保真度的混合损失函数, 该函数通过引入可调节的正则化约束项, 在优化

模型训练过程的同时有效抑制了过度拟合。

传统机器学习分类任务通常采用交叉熵损失 (cross-entropy loss) 或负对数似然损失 (negative log-likelihood loss) 作为优化目标, 这些损失函数的本质是对预测概率分布与真实标签分布之间的差异性进行度量。具体而言, 它们通过 Kullback-Leibler 散度或似然函数的形式, 量化了模型预测的概率分布与目标分布之间的信息差异。然而, 这类基于概率差异的损失函数在量子机器学习框架下存在局限性, 因为它们无法直接利用量子态空间的几何特性进行优化。

量子态保真度 (quantum state fidelity) 是量子信息领域中用来衡量两个量子态“相似程度”的指标。它可以理解为两个态之间的“距离”或“重合程度”。对于两个量子比特, 其保真度可以定义为

$$F(|\psi\rangle, |\varphi\rangle) = |\langle\psi|\varphi\rangle|^2$$

为此, 权重量子态保真度损失函数被设计出来用于训练量子机器学习模型:

$$\chi_{wf}^2(\theta, \omega, \alpha) = \frac{1}{MC} \sum_{\mu=1}^M \sum_{c=1}^C (\alpha_c F_c(\theta, \omega, \mathbf{x}_\mu) - Y_c(\mathbf{x}_\mu))^2$$

式中: $F_c(\theta, \omega, \mathbf{x}) = |\langle\psi_c|\psi(\theta, \omega, \mathbf{x})\rangle|^2$ 为预测的量子状态和标签的量子状态的保真度, C 代表分类任务的标签数量, M 代表样本数量。 $|\psi_c\rangle$ 为类 c 标签的量子状态。 $Y_c(\mathbf{x})$ 代表了一个成功的分类情况下期望的量子态保真度。其量子信息论中的物理意义通过 Bloch 球面中的正多面体的形式表示出来, 如图 2 所示。

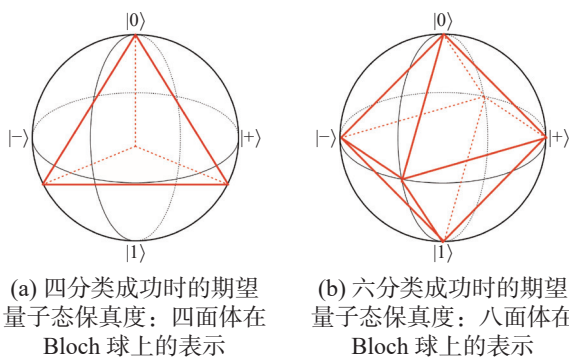


图 2 分类成功情况下的期望量子态保真度的 Bloch 球面表示

Fig. 2 Representation in the Bloch sphere of the expected quantum state fidelity in the case of successful classification

具体而言, 每个类别标签被映射为特定的量子态向量, 预测输出也以量子态向量的形式进行表示, 两者之间的相似性可通过保真度进行量化。将离散的类别标签嵌入到连续希尔伯特空间中, 每个类别被唯一地映射为一个特定的量子基

态。而权重量子态保真度损失函数根据量子信息论中的 Bures 度量被定义为真实量子态与预测量子态之间距离。因此, 这种损失函数的优化过程本质上等价于在连续特征空间内最小化预测值与目标值之间的欧氏距离。通过将分类误差重新表达为希尔伯特空间中的回归误差, 我们能够利用回归任务的优化方法 (如梯度下降法) 直接进行参数更新。

从机器学习的视角重新审视权重量子态保真度损失函数的结构与性质, 不仅加深了我们对其数学本质的理解, 也为解决数据重上传模型中的过拟合问题提供了新的切入点。受传统机器学习中应对过拟合策略的启发, 我们可以在损失函数中引入正则化项。这一策略为量子模型训练提供了与经典机器学习接轨的思路, 也拓展了现有量子损失函数设计的研究范式。

我们设计的稳健量子态保真度损失函数为

$$\chi_{wf}^2(\theta, \omega, \alpha) = \frac{1}{MC} \sum_{\mu=1}^M \sum_{c=1}^C (\alpha_c F_c(\theta, \omega, \mathbf{x}_\mu) - Y_c(\mathbf{x}_\mu))^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

式中: $\mathbf{w} = (\theta, \omega, \alpha)$ 代表量子模型的参数。

这一损失函数明显分成了两个部分: 保真度损失和正则化损失。一方面通过最小化模型输出与真实标签之间的差距, 提高训练精度; 另一方面通过正则项限制参数幅度, 提升模型在测试数据上的泛化能力, 同时也保证了模型能够在一定程度上抵抗量子模型本身的噪声影响。

4 实验验证

本文所提出的损失函数的改进本质上可视为数据重上传 (data reuploading) 模型的一种扩展形式。研究者可依据具体任务需求灵活选取正则项。为便于与传统数据重上传模型进行对比分析, 本文的网络结构如下:

整体模型由 $3n$ 层组成, 其中前 n 层采用标准的数据重上传嵌入方式, 中间 n 层引入 QAOA 嵌入, 最后 n 层则应用角度嵌入 (angle embedding)。该架构在遵循量子电路通用近似定理的基础上, 于参数数量一致的前提下实现更具针对性的编码优化, 从而确保对比实验的公平性与合理性。

为全面评估所提方法的有效性, 选取线性可分数据集与 MNIST 数据集作为评估平台。比较对象包括: 标准数据重上传模型和多嵌入数据重上传模型。此外, 本次实验中选择的批大小为 32, 学习率为 0.001, 测量算子的类型为全测量量

子 (full)。实验的设备条件为 CPU: Intel(R) Xeon(R) Gold 6226R CPU @ 2.90 GHz, GPU: NVIDIA A800 80 GB PCIe, 操作系统为 Ubuntu 20.04.6 LTS. Python 版本 3.10.15, torch 版本 2.5.1+cu121。

实验证实, 相较于传统训练方法, 本文的新型损失函数能够在保持性能优势的同时, 使模型展现出更为稳健的预测能力。这一创新性方法为解决量子机器学习中的过拟合问题提供了新的研究思路和技术路径。

4.1 线性可分数据实验

该数据集通过在 d 维超立方体中随机采样生成, 并利用一个正交于向量 $(1, 1, \dots, 1)^T$ 的超平面进行划分, 保留了有限但明确的数据无关边界, 具备良好的可解释性与分析价值。

本文在 10、12 和 14 维数据集上开展实验, 此处选择的正则化惩罚参数 $\lambda = 0.01$, 重点比较稳健重上传模型与 MEDQ 模型之间的性能差异。在每一组层数配置下, 均通过网格搜索确定训练误差最小的超参数组合, 并进行 5 次重复实验以报告平均性能, 从而更准确地反映模型在实际应用场景中的稳健性。

图 3~5 分别给出了在 10、12 和 14 维数据集上, 重载模型、MEDQ 模型和稳健重载模型的拟合情况。从图中可以明显看出, 随着层数的增加, 重载模型和 MEDQ 模型均表现出测试准确率下降幅度远远大于训练准确率的情况, 表现出明显的过拟合现象, 并伴随着训练过程中不稳定性的出现, 导致训练失败。而稳健重载模型在保持优异拟合能力的同时, 成功缓解了过拟合, 避免了训练失败的情况, 表现出更强的稳健性。因此, 稳健重载模型在此实验设置下展现出了显著的优势。

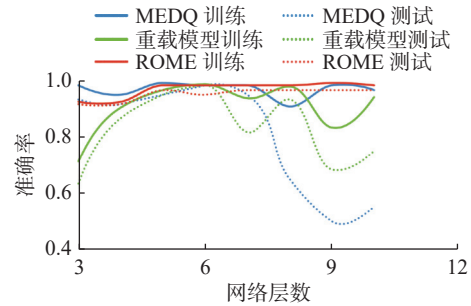


图 3 重载模型族在线性可分数据集-10d 上的性能表现
Fig. 3 Performance of data re-uploading models on Linear Separable-10d

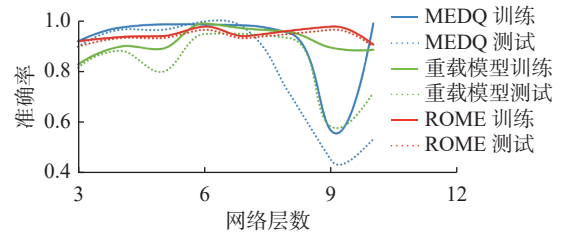


图 4 重载模型族在线性可分数据集-12d 上的性能表现
Fig. 4 Performance of data re-uploading models on Linear Separable-12d

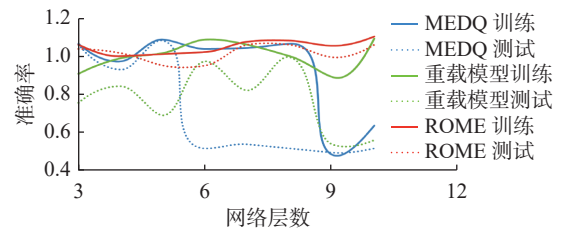


图 5 重载模型族在线性可分数据集-14d 上的性能表现
Fig. 5 Performance of data re-uploading models on Linear Separable-14d

实验结果表明, 在相同实验条件下, 稳健重上传模型可在 MEDQ 过拟合的场景中保持自身的性能, 具体准确率结果如表 1~3 所示。

表 1 在线性可分数据集-10d 上的实验结果

Table 1 Experimental results on Linear Separable-10d dataset

模型	网络层数 L							
	3	4	5	6	7	8	9	10
重上传	0.7508	0.8825	0.9591	0.9341	0.9691	0.9683	0.8658	0.9650
MEDQ	0.9533	0.9633	0.9633	0.9566	0.9800	0.7633	0.8200	0.6933
ROME	0.9066	0.9100	0.9433	0.9266	0.9533	0.9566	0.9800	0.9633

注: 加粗表示最优结果。

表 2 在线性可分数据集-12d 上的实验结果

Table 2 Experimental results on Linear Separable-12d dataset

模型	网络层数 L							
	3	4	5	6	7	8	9	10
重上传	0.8033	0.9033	0.9133	0.9533	0.9166	0.8566	0.7760	0.7500
MEDQ	0.8966	0.9630	0.9600	0.9766	0.9833	0.6433	0.4633	0.5433
ROME	0.9100	0.9533	0.9266	0.9566	0.9366	0.9433	0.9400	0.9100

注: 加粗表示最优结果。

表 3 在线性可分数据集-14d 上的实验结果
Table 3 Experimental results on Linear Separable-14d dataset

模型	网络层数 L							
	3	4	5	6	7	8	9	10
重上传	0.7333	0.79	0.7333	0.9266	0.8300	0.8266	0.7166	0.6966
MEDQ	0.8766	0.8633	0.9266	0.7433	0.5833	0.6700	0.4766	0.4966
ROME	0.9160	0.8933	0.9166	0.9166	0.9300	0.9200	0.9200	0.9166

注: 加粗表示最优结果。

4.2 MNIST 数据集实验

MNIST 是计算机视觉与深度学习领域中的经典数据集, 广泛用于图像分类与模式识别算法的性能评估。考虑到当前量子模型对高维数据处理的局限性, 本文采用主成分分析对数据

进行降维, 分别构建了 3 维与 4 维的子集用于训练与测试。实验结果显示, 在这两种不同维度下, 稳健重载模型能够保持较好的性能, 此处选择的正则化惩罚参数 $\lambda = 0.01$, 具体结果如表 4~5 所示。

表 4 在 MNIST 数据集-3d 上的实验结果
Table 4 Experimental results on MNIST-3d dataset

模型	网络层数 L							
	3	4	5	6	7	8	9	10
重载	0.7132	0.7239	0.7227	0.7277	0.7358	0.7297	0.7217	0.7232
MEDQ	0.7135	0.7159	0.7299	0.7240	0.7272	0.7338	0.7383	0.7304
ROME	0.7224	0.7255	0.7255	0.7315	0.7287	0.7271	0.7313	0.7345

注: 加粗表示最优结果。

表 5 在 MNIST 数据集-4d 上的实验结果
Table 5 Experimental results on MNIST-4d dataset

模型	网络层数 L							
	3	4	5	6	7	8	9	10
重载	0.7455	0.7873	0.7931	0.7858	0.7942	0.7974	0.7727	0.7873
MEDQ	0.7447	0.7946	0.7787	0.7796	0.7999	0.7960	0.8086	0.8034
ROME	0.7459	0.7429	0.7953	0.7962	0.8091	0.7992	0.8031	0.8047

注: 加粗表示最优结果。

4.3 正则化参数的选取实验

此外, 为了更系统地评估正则化强度对模型性能的影响, 这里进一步在线性可分性更强、效果提升更为显著的数据集上开展了正则化参数敏感性实验。该实验旨在通过控制正则化系数的取值, 观察模型在训练收敛行为与泛化表现 (如验证/测试指标变化) 上的差异, 从而为后续参数设

定提供更具解释性的依据。需要说明的是, 受限于现有实验设备的计算资源与可承受的训练开销, 我们无法在过宽的区间内对正则化参数进行高密度网格搜索; 因此, 本工作将正则化参数的候选取值限制在一个可行的范围内, 其取值范围设定为: $[0.001, 0.005, 0.01]$, 具体结果如表 6~8 所示。

表 6 在线性可分数据集-10d 上的参数选取实验结果
Table 6 Experimental results on Linear Separable-10d dataset

正则化参数 λ	网络层数 L							
	3	4	5	6	7	8	9	10
0.001	0.9467	0.9033	0.8967	0.8967	0.9367	0.8167	0.7600	0.7000
0.005	0.9233	0.9400	0.9333	0.9267	0.9400	0.9533	0.9400	0.8900
0.01	0.9067	0.9300	0.9367	0.9400	0.9533	0.9600	0.9767	0.9367

注: 加粗表示最优结果。

表 7 在线性可分数据集-12d 上的参数选取实验结果
Table 7 Experimental results on Linear Separable-12d dataset

正则化参数 λ	网络层数 L							
	3	4	5	6	7	8	9	10
0.001	0.9133	0.9333	0.9000	0.7967	0.7300	0.6533	0.5433	0.7500
0.005	0.9300	0.8967	0.9367	0.9433	0.9033	0.8533	0.9600	0.6600
0.01	0.9167	0.9267	0.9100	0.9467	0.9467	0.9467	0.9533	0.9300

注: 加粗表示最优结果。

表 8 在线性可分数据集-14d 上的参数选取实验结果
Table 8 Experimental results on Linear Separable-14d dataset

正则化参数 λ	网络层数 L							
	3	4	5	6	7	8	9	10
0.001	0.8867	0.9067	0.9067	0.7667	0.7900	0.5600	0.6633	0.5667
0.005	0.9000	0.9533	0.9300	0.8800	0.9233	0.9200	0.8867	0.8800
0.01	0.9233	0.9367	0.9233	0.9200	0.9433	0.9267	0.9400	0.9300

注: 加粗表示最优结果。

5 理论验证

从实验的验证中, 不难发现该正则化损失函数对过拟合的抑制作用, 接下来将从理论的角度尝试对该损失函数的实验效果做出一定的证明。

5.1 损失函数基本性质

对该成果的证明, 首先从对损失函数的性质开始分析, 即

$$\chi_{w_f}^2(\boldsymbol{\theta}, \boldsymbol{\omega}, \boldsymbol{\alpha}) = \frac{1}{MC} \sum_{\mu=1}^M \sum_{c=1}^C (\alpha_c F_c(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{x}_\mu) - Y_c(\mathbf{x}_\mu))^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

中的权重量子保真度损失函数:

这里假设输入数据 $\mathbf{x} \in \mathbf{R}^d$, 所有的参数为 $\mathbf{w} = (\boldsymbol{\theta}, \boldsymbol{\omega}, \boldsymbol{\alpha})$, 有以下结论成立。

引理 1 (保真度梯度上界) 对于损失函数中的保真度函数 $F_c(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{x}) = |\langle \psi_c | \psi(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{x}) \rangle|^2$, 有:

$$\|\nabla F_c\|^2 \leq \sqrt{d}$$

证明: 由量子参数移位规则, 有:

$$\frac{\delta F_c}{\delta \theta_j} = \frac{1}{2} \left[F\left(\theta_j + \frac{\pi}{2} \mathbf{e}_j\right) - F\left(\theta_j - \frac{\pi}{2} \mathbf{e}_j\right) \right]$$

而由于 $F_c \in [0, 1]$, 因此 $\left| \frac{\delta F_c}{\delta \theta_j} \right| \leq 1$, 进而:

$$\|\nabla F_c\|^2 = \left\| \sum_{i=1}^d \frac{\delta F_c}{\delta \theta_j} \right\|^2 \leq d$$

证毕。

由此能得到损失函数的 Lipschitz 连续性:

定理 1 (损失函数 Lipschitz 连续性) 权重量子保真度损失函数是 Lipschitz 连续的。

证明: 对于 $\forall \mathbf{w}_1, \mathbf{w}_2, \exists \xi$, 由拉格朗日中值定理有:

$$\|\hat{\chi}_{w_f}^2(\mathbf{w}_1) - \hat{\chi}_{w_f}^2(\mathbf{w}_2)\| = \|\nabla \hat{\chi}_{w_f}^2(\xi)\| \cdot \|\mathbf{w}_1 - \mathbf{w}_2\|$$

而

$$\|\nabla \hat{\chi}_{w_f}^2(\xi)\| = \left\| \sum_{c=1}^C (F_c(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{x}_\mu) + \alpha_c \nabla F_c) \right\| \leq \sum_{c=1}^C |F_c(\boldsymbol{\theta}, \boldsymbol{\omega}, \mathbf{x}_\mu)| + \sqrt{d} \sum_{c=1}^C |\alpha_c| \leq C(\sqrt{d} + 1)$$

证毕。

定理 2 (正则化诱导参数约束) 正则化的优化问题的最优解 \mathbf{w}_λ^* 满足:

$$\|\mathbf{w}_\lambda^*\| \leq \frac{C(\sqrt{d} + 1)}{\lambda}$$

证明: 由 KKT 条件, 容易得到:

$$\nabla \chi_{w_f}^2(\mathbf{w}_\lambda^*) = \nabla \hat{R}(\mathbf{w}_\lambda^*) + \lambda \mathbf{w}_\lambda^* = 0$$

由此,

$$\|\nabla \hat{R}(\mathbf{w}_\lambda^*)\| = \lambda \|\mathbf{w}_\lambda^*\|$$

所以由定理 1, 有:

$$\|\mathbf{w}_\lambda^*\| \leq \frac{\|C(\sqrt{d} + 1)\|}{\lambda}$$

证毕。

5.2 广义三角多项式

所有量子神经网络, 其均属于文献 [14] 中提及的参数化量子电路 (parameterized quantum circuit, PQC), 进而输出函数表示为广义三角多项式 (generalized trigonometric polynomial, GTP):

$$f_w(\mathbf{x}) = \sum_{\omega \in \Omega(\mathbf{D})} c_\omega(\mathbf{w}) e^{-i\omega \mathbf{x}}$$

式中: $\Omega(\mathbf{D})$ 是由编码策略决定的频率集, 是由编码哈密顿量的谱结构生成的, 只与编码方式相关:

$$|\Omega(\mathbf{D})| = \sum_{i=1}^d \sum_{j=1}^{N^{(i)}} \{(\lambda_k - \lambda_l) e^{(i)} | \lambda_k, \lambda_l \in \text{spec}(H)\}$$

其中, $N^{(i)}$ 为第 i 维的编码门的数量, $H_j^{(i)}$ 为第 i 维的第 j 个编码门的哈密顿量, $e^{(i)}$ 为第 i 维的标准基。

这里讨论的广义三角多项式, 均在一个频率支持集为 Ω , 范数上界为 B 的广义三角多项式函

数类中:

$$\mathcal{F}_\Omega^B = \left\{ f(\mathbf{x}) = \sum_{\omega \in \Omega} c_\omega e^{-i\omega \mathbf{x}} : \|f\| \leq B \right\}$$

其中 c_ω 可以表示为 $c_\omega(\boldsymbol{\theta}) = \text{Tr}(\rho_\omega(\boldsymbol{\theta})M)$ 。 ω 是频率向量, $\boldsymbol{\theta}$ 是可训练参数。

由于 PQC 模型最后输出的是实数, 所以有 $c_{-\omega} = c_\omega^*$, 神经网络的广义三角多项式可以转换为实函数形式:

$$f(\mathbf{x}) = \frac{a_0}{2} + \sum_{\omega \in \Omega^+} [a_\omega \cos(\omega \mathbf{x}) + b_\omega \sin(\omega \mathbf{x})]$$

借由 Parseval 等式, 引入了新的函数类:

$$\mathcal{H}_\Omega^B = \left\{ f(\mathbf{x}) = \frac{a_0}{2} + \sum_{\omega \in \Omega^+} [a_\omega \cos(\omega \mathbf{x}) + b_\omega \sin(\omega \mathbf{x})] : \|(a_0, (a_\omega), (b_\omega))\|_2 \leq 2(2\pi)^{d/2} B \right\}$$

对于任意一个损失函数 ℓ , 定义如下函数类:

$$\mathcal{G} := \{[0, 2\pi]^d \times \mathbf{R} \ni (\mathbf{x}, y) \mapsto \ell(y, f(\mathbf{x})) : f \in \mathcal{H}_\Omega^B\}$$

根据上述的背景, 在文献 [14] 中, 我们从 Rademacher 复杂度的角度可以得到结论:

定理 3 (广义三角多项式的泛化界) 对 $[0, 2\pi]^d \times \mathbf{R}$ 上的任意概率分布 P , 给定 $\delta > 0$, 从 P 独立同分布抽取大小为 m 的训练集 $S = \{(x_i, y_i)\}_{i=1}^m \sim P$, 损失函数 $\ell(y, f(\mathbf{x}))$ 对于 y 是 L -Lipschitz 连续, 则以概率至少 $1 - \delta$ 有:

$$\begin{aligned} R(f) - \hat{R}_S(f) &\leq \\ 2\mathfrak{R}_S(\mathcal{G}) + 3c \sqrt{\frac{\log(2/\delta)}{2m}}, \quad \forall g \in \mathcal{G} &\leq \\ 2LR_{S_{\text{ix}}}(\mathcal{H}_\Omega^B) + 3c \sqrt{\frac{\log(2/\delta)}{2m}} &\leq \\ 2 \frac{L}{\sqrt{m}} \left(2\pi \max \{K, 2(2\pi)^{d/2} B \sqrt{|\Omega^+|}\} \sqrt{2 \log(2d)} + \right. \\ \left. \max \{\pi, 2(2\pi)^{d/2} B\} \right) + 3c \sqrt{\frac{\log(2/\delta)}{2m}} &\leq \\ 2 \frac{L}{\sqrt{m}} \left((2\pi)^{d/2+1} B \sqrt{|\Omega|} \sqrt{2 \log(2d)} \right) + 3c \sqrt{\frac{\log(2/\delta)}{2m}} \end{aligned}$$

在进一步推进该结论之前, 先从另一个定理出发, 将量子神经网络的参数约束和函数约束通过广义三角多项式结合起来。

引理 2 (量子神经网络的函数约束和参数约束的关系) 对于形如:

$$f(\mathbf{x}) = \text{Tr}[\Phi(|\psi(\boldsymbol{\theta}, \omega, \mathbf{x})\rangle \langle \psi(\boldsymbol{\theta}, \omega, \mathbf{x})|)]$$

形式的量子神经网络, 其广义三角多项式的形式:

$$f(\mathbf{x}) = \frac{a_0}{2} + \sum_{\nu \in \Omega^+} [a_\nu \cos(\nu \mathbf{x}) + b_\nu \sin(\nu \mathbf{x})]$$

式中: Ω 由编码方式和网络参数决定, 且 $|a_\nu|, |b_\nu| \leq 1$ 记:

$$\begin{aligned} \varphi(\mathbf{x}) &:= \left(\frac{1}{2}, (\cos(\nu \mathbf{x}))_{\nu \in \Omega^+}, (\sin(\nu \mathbf{x}))_{\nu \in \Omega^+} \right)^T \\ \mathbf{c}(\mathbf{w}) &:= (a_0(\mathbf{w}), (a_\nu(\mathbf{w}))_{\nu \in \Omega^+}, (b_\nu(\mathbf{w}))_{\nu \in \Omega^+})^T \end{aligned}$$

当 $\|\nabla c_i(\mathbf{w})\| \leq C, \forall i \in \{1, 2, \dots, 2|\Omega^+| + 1\}$, 则存在常数 G 使得:

$$\|f(\mathbf{x})\| \leq G \|\mathbf{w}\|$$

其中 $\mathbf{w} = (\boldsymbol{\theta}, \omega, \alpha)$ 。

证明: 量子神经网络可以写成:

$$f(\mathbf{x}) = C(\mathbf{w})^T \varphi(\mathbf{x})$$

而

$$\|\varphi(\mathbf{x})\|_2 = \sqrt{\frac{1}{2} + \sum_{\nu \in \Omega^+} \cos^2(\nu \mathbf{x}) + \sum_{\nu \in \Omega^+} \sin^2(\nu \mathbf{x})} \leq \sqrt{1 + 2|\Omega^+|} \leq \sqrt{|\Omega|}$$

由于 $\|\nabla c_i(\mathbf{w})\| \leq C$, 有:

$$\|\nabla c(t \cdot \mathbf{w})\| \leq C, t \in [0, 1]$$

因此:

$$\begin{aligned} |c_i(\mathbf{w})| &= \left| \int_0^1 \nabla c(t \cdot \mathbf{w}) \cdot \mathbf{w} dt \right| \leq \int_0^1 \|\nabla c(t \cdot \mathbf{w})\| \cdot \|\mathbf{w}\| dt \leq \\ &\|\mathbf{w}\| \int_0^1 C dt = C \|\mathbf{w}\| \end{aligned}$$

整理后可以得到:

$$\begin{aligned} \|f(\mathbf{x})\| &\leq \|c(\mathbf{w})\| \cdot \|\varphi(\mathbf{x})\| \leq \\ &C \sqrt{|\Omega|} \|\mathbf{w}\| = G \|\mathbf{w}\| \end{aligned}$$

证毕。

5.3 正则化权重量子保真度损失函数的泛化能力 有了前两节的理论准备, 我们可以讨论正则化量子保真度损失函数相比于未正则化的损失函数对模型的泛化能力的促进作用。

对于正则化的优化问题的最优参数 \mathbf{w}_λ^* , 综合定理 2 和引理 2, 有:

$$\|f_{\mathbf{w}_\lambda^*}(\mathbf{x})\| \leq G \|\mathbf{w}_\lambda^*\| \leq G \frac{C(\sqrt{d}+1)}{\lambda} =: B_\lambda$$

同理, 对于未正则化的优化问题的最优参数 \mathbf{w}^* , 有:

$$\|f_{\mathbf{w}^*}(\mathbf{x})\| \leq G \|\mathbf{w}^*\| \leq GR =: B_0$$

因此根据定理 3:

$$R(f_{\mathbf{w}_\lambda^*}) - \hat{R}_S(f_{\mathbf{w}_\lambda^*}) \leq 2LR_{S_{\text{ix}}}(\mathcal{H}_{\Omega^2}^{B_\lambda}) + 3c \sqrt{\frac{\log(2/\delta)}{2m}} \leq$$

$$2 \frac{L}{\sqrt{m}} \left((2\pi)^{d/2+1} B_\lambda \sqrt{|\Omega|} \sqrt{2 \log(2d)} \right) + 3c \sqrt{\frac{\log(2/\delta)}{2m}}$$

同理:

$$R(f_{w^*}) - \hat{R}_S(f_{w^*}) \leq 2LR_{S|x}(\mathcal{H}_{\Omega^{\theta_0}}) + 3c \sqrt{\frac{\log(2/\delta)}{2m}} \leq$$

$$2 \frac{L}{\sqrt{m}} \left((2\pi)^{d/2+1} B_0 \sqrt{|\Omega|} \sqrt{2 \log(2d)} \right) + 3c \sqrt{\frac{\log(2/\delta)}{2m}}$$

因此, 当 $B_\lambda \leq B_0$, 即 $\lambda \geq \frac{C(\sqrt{d}+1)}{R}$ 时, 可以说

明此时正则化了的优化问题相较于于正则化前的优化问题拥有更好的泛化界。

6 结束语

在量子机器学习 (QML) 领域, 数据重上传模型 (如多重嵌入数据重上传量子模型) 虽然具有强大的泛化潜力, 但在实际应用中面临两个主要问题: 过拟合问题和超参数敏感性。这两个问题严重限制了量子机器学习的应用场景。针对此问题, 本文首次对数据重上传类模型的过拟合现象进行了系统而深入的理论分析与实验研究。本文提出一种新颖的损失函数, 该损失函数基于加权量子态保真度损失, 并引入了正则化项, 以有效抑制模型训练过程中产生的过拟合现象。通过一系列实验研究, 验证了所提出的正则化损失函数在不同数据规模与量子电路结构下的泛化性能。实验结果表明, 本文方法能够显著提升数据重上传类量子机器学习模型的稳健性与泛化表现, 并在后面给出了模型的泛化性能的理论证明。

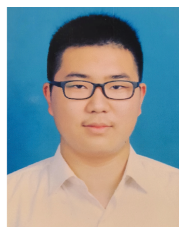
本研究从机器学习的理论视角创新性地重构了权重量子态保真度损失函数的数学框架, 揭示其本质为希尔伯特空间中的量子态回归优化问题。这一理论重构为后续研究提供了重要的启示方向: 一方面, 可深入探索基于量子态统计分布的差异性度量方法, 如量子相对熵或 Wasserstein 距离等更丰富的度量空间; 另一方面, 亟须建立严格的数学理论体系以证明该损失函数在泛化误差界、收敛性等方面的理论完备性。这些理论深化工作将为进一步优化量子分类任务的损失函数设计奠定坚实的理论基础。

参考文献:

- [1] 石金晶, 肖子萌, 王雯萱, 等. 面向含噪中规模量子处理器的量子机器学习[J]. 计算机学报, 2025, 48(3): 602–631.
- SHI Jinjing, XIAO Zimeng, WANG Wenxuan, et al. Quantum machine learning with noisy intermediate-scale quantum processors[J]. Chinese journal of computers, 2025, 48(3): 602–631.
- [2] 侯敏. 量子机器学习算法研究概述[J]. 通讯世界, 2024, 31(8): 139–141.
- HOU Min. Overview of quantum machine learning algorithms[J]. Telecom world, 2024, 31(8): 139–141.
- [3] 王健, 张蕊, 姜楠. 量子机器学习综述[J]. 软件学报, 2024, 35(8): 3843–3877.
- WANG Jian, ZHANG Rui, JIANG Nan. Survey on quantum machine learning[J]. Journal of software, 2024, 35(8): 3843–3877.
- [4] 瞿治国, 陈韦龙, 孙乐, 等. ECG-QGAN: 基于量子生成对抗网络的心电图生成式信息系统[J]. 计算机研究与发展, 2025, 62(7): 1622–1638.
- QU Zhiguo, CHEN Weilong, SUN Le, et al. ECG-QGAN: an ECG generative information system based on quantum generative adversarial networks[J]. Journal of computer research and development, 2025, 62(7): 1622–1638.
- [5] 于瑞祺, 张鑫云, 任爽. 基于变分量子电路的量子机器学习算法综述[J]. 计算机研究与发展, 2025, 62(4): 821–851.
- YU Ruiqi, ZHANG Xinyun, REN Shuang. A review of quantum machine learning algorithms based on variational quantum circuit[J]. Journal of computer research and development, 2025, 62(4): 821–851.
- [6] BIAMONTE J, WITTEK P, PANCOTTI N, et al. Quantum machine learning[J]. Nature, 2017, 549(7671): 195–202.
- [7] CERZO M, VERDON G, HUANG H Y, et al. Challenges and opportunities in quantum machine learning[J]. Nature computational science, 2022, 2(9): 567–576.
- [8] TYCHOLA K A, KALAMPOKAS T, PAPAPOSTAS G A. Quantum machine learning: an overview[J]. Electronics, 2023, 12(11): 2379.
- [9] ZEGUENDRY A, JARIR Z, QUAFAROU M. Quantum machine learning: a review and case studies[J]. Entropy, 2023, 25(2): 287.
- [10] HAVLÍČEK V, CÓRCELES A D, TEMME K, et al. Supervised learning with quantum-enhanced feature spaces[J]. Nature, 2019, 567(7747): 209–212.
- [11] JERBI S, FIDERER L J, POULSEN NAUTRUP H, et al. Quantum machine learning beyond kernel methods[J]. Nature communications, 2023, 14: 517.
- [12] JERBI S, GYURIK C, MARSHALL S C, et al. Shadows of quantum machine learning[J]. Nature communications, 2024, 15: 5676.
- [13] SCHULD M, SWEKE R, MEYER J J. Effect of data encoding on the expressive power of variational quantum-machine-learning models[J]. Physical review A, 2021, 103(3): 032430.
- [14] CARO M C, GIL-FUSTER E, MEYER J J, et al. Encod-

- ing-dependent generalization bounds for parametrized quantum circuits[J]. *Quantum*, 2021, 5: 582.
- [15] CARO M C, HUANG H Y, CERESO M, et al. Generalization in quantum machine learning from few training data[J]. *Nature communications*, 2022, 13: 4919.
- [16] KAVITHA S S, KAULGUD N. Quantum machine learning for support vector machine classification[J]. *Evolutionary intelligence*, 2024, 17(2): 819–828.
- [17] KITAEV A. Quantum computations: algorithms and error correction[J]. *Russian mathematical surveys*, 1997, 52(6): 1191–1249.
- [18] REBENTROST P, MOHSENI M, LLOYD S. Quantum support vector machine for big data classification[J]. *Physical review letters*, 2014, 113(13): 130503.
- [19] LLOYD S, SCHULD M, IJAZ A, et al. Quantum embeddings for machine learning[EB/OL]. (2020–01–10)[2025–01–01]. <https://arxiv.org/abs/2001.03622>.
- [20] HARROW A W, HASSIDIM A, LLOYD S. Quantum algorithm for linear systems of equations[J]. *Physical review letters*, 2009, 103(15): 150502.
- [21] CONG I, CHOI S, LUKIN M D. Quantum convolutional neural networks[J]. *Nature physics*, 2019, 15(12): 1273–1278.
- [22] SCHREIBER F J, EISERT J, MEYER J J. Classical surrogates for quantum learning models[J]. *Physical review letters*, 2023, 131(10): 100803.
- [23] HUANG H Y, KUENG R, PRESKILL J. Predicting many properties of a quantum system from very few measurements[J]. *Nature physics*, 2020, 16(10): 1050–1057.
- [24] GHOBADI R, OBEROI J S, ZAHEDINEJHAD E. The power of one qubit in machine learning[EB/OL]. (2019–03–03)[2025–01–01]. <https://arxiv.org/abs/1905.01390>.
- [25] HUANG H Y, BROUGHTON M, MOHSENI M, et al. Power of data in quantum machine learning[J]. *Nature communications*, 2021, 12: 2631.
- [26] LAROSE R, COYLE B. Robust data encodings for quantum classifiers[J]. *Physical review A*, 2020, 102(3): 032420.
- [27] GUALA D, ZHANG Shaoming, CRUZ E, et al. Practical overview of image classification with tensor-network quantum circuits[J]. *Scientific reports*, 2023, 13: 4427.
- [28] PÉREZ-SALINAS A, CERVERA-LIERTA A, GIL-FUSTER E, et al. Data re-uploading for a universal quantum classifier[J]. *Quantum*, 2020, 4: 226.
- [29] BOWLES J, AHMED S, SCHULD M. Better than classical? The subtle art of benchmarking quantum machine learning models[EB/OL]. (2024–03–11)[2025–01–01]. <https://arxiv.org/abs/2403.07059>.
- [30] HAN Siyu, JIA Lihan, GUO Lanzhe. Multiple embeddings for quantum machine learning[EB/OL]. (2025–03–27)[2025–07–01]. <https://arxiv.org/abs/2503.22758>.

作者简介:



韩思予, 硕士研究生, 主要研究方向为机器学习、数据挖掘。曾担任 ICML、IJCAI 等人工智能领域著名学术会议审稿人。E-mail: hansy@lamda.nju.edu.cn。



贾林瀚, 博士研究生, 主要研究方向为机器学习、数据挖掘。曾担任 ICML、ICLR、NeurIPS 等人工智能领域著名学术会议审稿人, 曾获国家奖学金、南京大学优秀研究生、中国计算机学会优秀大学生等荣誉。发表学术论文 8 篇。E-mail: jialh@lamda.nju.edu.cn。



李宇峰, 教授, 博士生导师, 国家级人才, 中国计算机学会杰出会员, 主要研究方向为机器学习、数据挖掘。担任中国人工智能学会机器学习专委会秘书长, 人工智能领域著名国际期刊《Artificial Intelligence》《Machine Learning》编委, 《Frontiers of Computer Science》《计算机研究与发展》青年编委。入选人工智能旗舰国际会议 IJCAI21 青年成就亮点报告, 获江苏省科学技术一等奖、吴文俊人工智能优博指导教师、华为火花奖等科研奖励。发表学术论文 90 余篇, 获国内外优秀论文奖 4 次。E-mail: liyf@nju.edu.cn。