

引用格式: 杨帆, 程学云, 朱鹏程, 等. 基于角度-振幅混合编码的量子神经网络及其应用研究 [J]. 电子科技大学学报, 2025, 54(5): 789-800.
YANG F, CHENG X Y, ZHU P C, et al. Quantum neural networks based on angle-amplitude mixed encoding and its applications[J]. Journal of University of Electronic Science and Technology of China, 2025, 54(5): 789-800.

基于角度-振幅混合编码的量子神经网络 及其应用研究



杨帆¹, 程学云^{1*}, 朱鹏程², 姜一博¹, 顾晖¹, 管致锦¹

(1. 南通大学人工智能与计算机学院, 南通 226019; 2. 泰州学院信息工程学院, 泰州 225300)

摘要: 传统量子神经网络与自注意机制结合模型需消耗较高的量子位资源, 针对其在当前 NISQ 设备上运行效率低和设计复杂度高问题, 提出了一种混合编码方式, 将数据集特征通过特定的方式嵌入量子态中, 从而实现角度编码与振幅编码的有效混合; 基于该编码方法设计出一种结构独特的双环 Ansatz, 借鉴自注意机制中的分而治之思想, 构建出具备更高表现力的量子神经网络。在鸢尾花分类任务中训练损失收敛于 0, 证明模型有效捕捉到鸢尾花特征之间的内在联系; 在文本分类任务中与已有方法相比, 分类精确度平均提升了 8.9%, 且在保证效果良好的前提下, 成功减少了训练参数的数量。基于角度-振幅混合编码的量子神经网络的轻量化和低复杂度特性使其更适用于当前的 NISQ 设备。

关键词: 量子神经网络; 混合编码; 自注意机制; 文本分类

中图分类号: TP301

文献标志码: A

DOI: 10.12178/1001-0548.2024289

Quantum neural networks based on angle-amplitude mixed encoding and its applications

YANG Fan¹, CHENG Xueyun^{1*}, ZHU Pengcheng², JIANG Yibo¹, GU Hui¹, and GUAN Zhijin¹

(1. School of Artificial Intelligence and Computer Science, Nantong University, Nantong 226019, China;

2. School of Information Engineering, Taizhou University, Taizhou 225300, China)

Abstract: Models combining traditional quantum neural networks with the self-attention mechanism require a high consumption of qubit resources. In response to the issues of low operational efficiency and high design complexity on current NISQ (noisy intermediate-scale quantum) devices, this paper proposes a mixed encoding method that embeds dataset features into quantum states in a specific manner, achieving an effective mix of angle encoding and amplitude encoding. Based on this encoding method, a uniquely structured double-ring Ansatz is designed. Drawing on the divide-and-conquer strategy of the self-attention mechanism, a quantum neural network with higher expressiveness is constructed. In the Iris classification task, the training loss converges to 0, proving that the model effectively captures the intrinsic relationships between the Iris features. In the text classification task, the classification accuracy improves by an average of 8.9% compared with existing methods, while also significantly reducing the number of training parameters without compromising performance. The lightweight and low-complexity characteristics of the quantum neural network based on angle-amplitude hybrid encoding make it more suitable for current NISQ devices.

Key words: quantum neural networks; hybrid encoding; self-attention mechanism; text categorization

量子计算是一种前景广阔的计算范式^[1], 利用量子位的叠加态和纠缠态处理信息, 为密码学^[2-3]、优化^[4-5]和机器学习^[6-9]等领域提供了解决方案。目前, 量子计算处于噪声中等规模量子 (noisy

intermediate-scale quantum, NISQ) 时代, 量子设备拥有的物理量子位数量最多为数百个, 但这些设备易受噪声和退相干的影响, 尤其在实现复杂量子算法时面临显著挑战。另一方面, 自注意机制 (self-

收稿日期: 2024-10-22

基金项目: 国家自然科学基金面上项目 (62072259); 江苏省自然科学基金 (BK20221411); 南通大学博士启动基金 (23B03); 南通市民生科技计划 (MS2023070); 南通大学信息科学技术学院研究生科研与实践创新计划 (NTUSISTPR24_04)

作者简介: 杨帆, 主要从事量子机器学习方面的研究。

*通信作者 E-mail: chen.xy@ntu.edu.cn

attention mechanism, SAM) 作为多种深度学习架构的核心, 在自然语言处理和计算机视觉领域取得了显著进展, 特别是在捕捉复杂关系、提高并行计算能力以及处理大规模数据方面发挥了重要作用。这种机制在经典神经网络中的成功经验启发了将其与量子计算结合, 尤其是在量子神经网络 (quantum neural networks, QNN) 中的应用。QNN 通过参数化量子线路 (parameterized quantum circuits, PQC) 构建多层深度网络。然而, 基于经验架构的 QNN 往往忽略了量子数据间的内在联系, 导致对量子存储资源的需求高, 因而在 NISQ 器件上实现困难。因此, 如何实现量子自注意机制成为一个值得深入研究的课题。

目前, 已有研究尝试将 SAM 融入 QNN。如文献 [10] 提出了量子自注意神经网络模型, 该模型能够有效挖掘高维量子特征空间中元素之间的相关性。文献 [11] 通过量子线路制备 Query、Key 和 Value 向量, 构建了量子 Transformer。文献 [12] 提出了一种量子自注意网络框架, 通过延迟和条件测量释放量子资源。文献 [13] 引入量子逻辑相似度 (QSL) 来可视化元素之间的相似度, 该模型在计算注意力分数方面表现优异, 过程中需要借助辅助寄存器来完成某些交换操作。文献 [14] 提出了一种自然的量子自注意模型, 该模型在准确度和线路复杂度方面都表现良好, 其鲁棒性在模拟噪声中得到了验证。

SAM 与 QNN 的结合, 能够通过量子线路实现输入特征间相似性的计算与全局信息的提取。然而, 部分研究通常忽视了量子线路宽度对实际运行的影响, 导致量子位资源需求过大。此外, 网络结构复杂的量子自注意模型运行效率较低, 在面对噪声时相对不稳定, 难以适应当前的 NISQ 设备。

为了解决模型网络结构复杂的问题, 本文提出了角度-振幅混合编码方法, 以降低模型的复杂性, 并设计了能够自然集成 SAM 的 Ansatz。通过将角度编码和振幅编码的优势结合, 基于角度-振幅混合编码的量子神经网络 (quantum neural networks based on angle-amplitude mixed encoding, AAM-QNN) 显著缩减了模型规模。此外, 基于这一编码方式, 设计出了具备纠缠功能的双环架构 Ansatz。与其他量子学习模型相比, 本文模型在不同类型的数据集分类任务中表现出更优的性能, 且更适用于当前的 NISQ 设备。

1 基本概念

1.1 量子基础知识

在量子计算机中, 量子位是携带量子信息的最小元素, 不同于经典位的二进制状态, 通常用 $|\psi\rangle$ 表示两个特征状态 $|0\rangle$ 和 $|1\rangle$ 线性叠加的量子态^[1]:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (1)$$

式中, $|0\rangle$ 和 $|1\rangle$ 分别表示量子位的基态和激发态; α 和 β 称为振幅, 且满足归一化条件:

$$|\alpha|^2 + |\beta|^2 = 1 \quad (2)$$

为了在量子计算中实现信息的处理和传输, 需要对量子态进行操作。量子门提供了这样一种机制, 其作用是将一种量子态线性转换为另一种量子态, 其数学形式可表示为 $2^n \times 2^n$ 的酉矩阵, 其中 n 为量子位的数量。本文使用的量子门主要包括控制旋转门 CRX、受控门 CNOT、非门 X、旋转门 RX, 如表 1 所示。

表 1 量子门表示

符号表示	矩阵表示	图形表示
CRX	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\frac{\theta}{2}) & -i\sin(\frac{\theta}{2}) \\ 0 & 0 & -i\sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{pmatrix}$	
CNOT	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$	
X	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	
RX	$\begin{pmatrix} \cos(\frac{\theta}{2}) & -i\sin(\frac{\theta}{2}) \\ -i\sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{pmatrix}$	

在实现量子算法时, 这些量子门被系统地组合为量子线路, 以执行特定的计算任务。量子线路的宽度, 即同时操作的量子位数量, 成为量子线路设计的关键。如图 1 所示, 量子线路包含 3 个量子位 $|q_0\rangle$ 、 $|q_1\rangle$ 和 $|q_2\rangle$, 因此线路的宽度为 3。

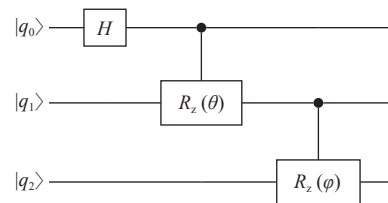


图 1 量子线路

宽度表示空间的维度, 不仅直接影响到系统的计算能力, 也决定了线路的复杂度和所需的物理资

源, 因此是优化量子线路设计的重要因素。

1.2 量子神经网络

QNN 是一种结合了量子计算和神经网络概念的前沿技术^[15], 由 4 部分组成: 数据编码、Ansatz、量子测量以及参数优化。QNN 架构如图 2 所示。

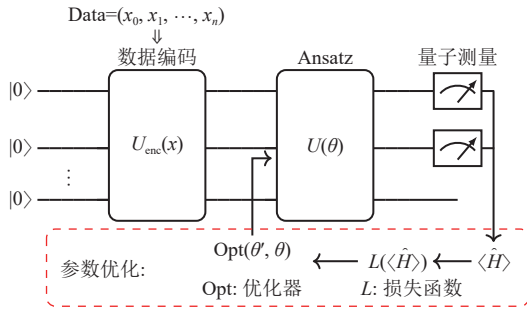


图 2 QNN 模型架构图

1) 数据编码

数据编码将经典数据转化为量子态, 通过量子门执行酉变换。首先, 将量子线路初始化为态 $|0\rangle$, 然后设计编码线路 (一个酉矩阵)。经过编码线路后, 初始态通过线性矩阵运算转换为量子数据。这个过程表示为 $|x\rangle = U_{enc}|0\rangle^{\otimes n}$, 其中 x 是经典数据向量, U_{enc} 是编码线路。

2) Ansatz

Ansatz 是量子计算与经典计算交互的关键。它接收编码后的量子态, 并通过 PQC 进行处理。Ansatz 中的可调参数通过经典优化算法调整, 以优化某个目标。通过不断优化这些参数, 使 QNN 逐渐逼近最优解。设计合适的 Ansatz 线路对于在 NISQ 设备上有效运行至关重要, 因此需要在提出的编码方式的基础上, 设计出适合混合编码的 Ansatz 架构。

1.3 自注意机制

SAM 在自然语言处理中的核心思想是通过将数据集样本中的每个特征相互关联, 以方便机器学习模型捕获特征之间的关键信息和内在联系^[16]。具体步骤如下。

1) 给定一个数据样本 $x = (a_1, a_2, \dots, a_n)$, 其中 $a_i \in R^d, i = 1, 2, \dots, n$ 。计算查询向量、键向量、值向量。计算过程为:

$$Q_i = W_Q \cdot a_i, K_i = W_K \cdot a_i, V_i = W_V \cdot a_i \quad (3)$$

式中, $Q_i、K_i、V_i$ 分别表示第 i 个元素的查询、键、值向量; $W_Q、W_K、W_V$ 为权重矩阵。

2) 计算注意力分数。对样本里每一对向量 a_i 和 a_j , 通过点积计算注意力分数 $a_{i,j}$, 计算过程为:

$$a_{i,j} = \frac{Q_i \cdot K_j^T}{\sqrt{d_k}} \quad (4)$$

式中, d_k 是键向量 K_i 的维度。为了匹配查询向量 Q 和键向量 K 之间的维度, 使得点积运算可以正确执行, 将 K 转置变为 K^T 。

3) 应用 softmax 函数。将注意力分数 $a_{i,j}$ 转化为注意力权重 $w_{i,j}$, 计算过程为:

$$w_{i,j} = \text{softmax}(a_{i,j}) \quad (5)$$

4) 计算最终输出。将样本中每个元素的值向量与注意力权重 $w_{i,j}$ 相乘并求和, 得到最终的输出:

$$A_i(Q_i, K_j, V_j) = \sum_{j=1}^n w_{i,j} V_j \quad (6)$$

2 量子神经网络的数据编码设计

在使用 QNN 解决问题的过程中, 数据编码是重要的步骤之一。将 SAM 融入 QNN, 首先需要获取机制所需要的量子数据向量, 而数据编码能够将经典数据转化为适合在量子计算中进行处理和分析的量子态表示, 在获取经典数据的量子态表示之后, 通过借鉴 SAM 对数据分而治之的思想设计合适的 Ansatz。现有的数据编码方法包括角度编码和振幅编码^[17-18]。

2.1 现有编码方法

2.1.1 角度编码

角度编码是把数据集里样本的每个特征逐个输入到旋转量子门的角度中, 以制备相应的量子态。图 3 为通过角度编码将经典数据嵌入编码线路的过程。

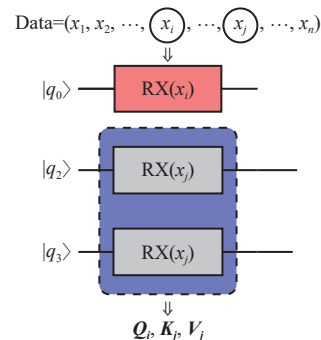


图 3 经典数据利用角度编码转换为量子态

经典数据经过编码转化为量子态, 以此表示经典数据对应的量子数据。按照式 (3) 获取查询向量 Q_i , 键向量 K_j , 值向量 V_j 。在图 3 中, 量子门 RX 实现式 (3) 中的权重矩阵 $W_{Q,K,V}$; 数据样本 Data 的

特征数量，即数据集维度，决定了整个编码层的量子位数量，其中 x_i 和 x_j 为样本中任意两个特征值。

在 1.3 节中提到 SAM 需要将不同的特征按照特定的组合方式放在一起，主要是为了捕捉输入特征之间的依赖关系和内在联系，让特征之间产生关联。特征之间重复组合，用数学符号描述为每次以 (x_i, x_j, x_j) 的形式重新组合，即 1 个样本 (x_1, x_2, \dots, x_n) 在按照上述形式重新组合之后变为 $((x_1, x_1, x_1, x_1, x_2, x_2, \dots, x_1, x_n, x_n), \dots, (x_n, x_1, x_1, x_n, x_2, x_2, \dots, x_n, x_n, x_n))$ ，其中 $1 \leq i \leq j \leq n$ ；利用角度编码对重新组合之后的新样本进行编码，然而按照 (x_i, x_j, x_j) 的形式逐个编码，需要编码 $3n^2$ 个特征，每个特征对应一个量子位，这使得整个量子线路的宽度为指数级别。文献 [14] 给出了减少冗余特征嵌入的方法，这种方法可以优化编码线路宽度，如图 4 所示。

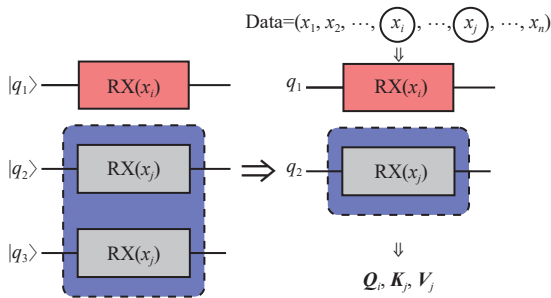


图 4 角度编码优化过程

由图 4 可知优化后的角度编码从一开始按照每组 (x_i, x_j, x_j) 的形式改变为 (x_i, x_j) 的组合编码到量子线路中，不仅使得量子位数减少，量子门也相应减少。虽然优化后的编码让量子线路宽度从原来的 $3n^2$ 降低至 $2n^2$ ，但仍然是指数级别的宽度。如果利用角度编码搭建 QNN，在处理高维数据时易造成量子位资源浪费。

2.1.2 振幅编码

振幅编码是另一种常见的编码形式。图 5 展示了实现振幅编码的量子线路。图中振幅编码线路的旋转角度 $\alpha_1, \alpha_2, \dots, \alpha_m$ 获取方式与角度编码不同，角度编码是直接数据作为量子门中的旋转角度设置，而振幅编码线路中量子门的旋转角度是通过算法计算得到，文献 [18] 详细描述了量子门旋转角度的计算过程。在编码相同数量的数据特征时，使用振幅编码所需的量子位从指数级别变成对数级别。

振幅编码虽然可以以少量量子位编码更多的数据，但相应使用了更多的多量子位门（占用量子位

数 ≥ 3 ），这导致振幅编码线路复杂且分解后的线路深度大，因而对噪声敏感，特别是在当前 NISQ 设备上，所以需要尽可能地减少多量子位门的数量。

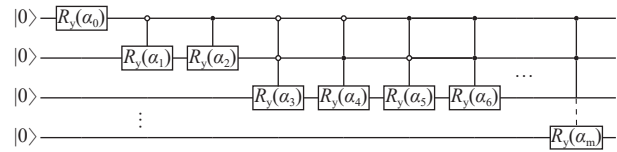


图 5 振幅编码线路

2.2 量子态振幅单向驱动的特征嵌入二叉树

给定一个数据集样本 $y = [y_1, y_2, \dots, y_n]$ ，其中 y_i 为第 i 个特征， n 为样本特征数量。从 2.1.1 节角度编码可知，样本 y 的每个特征在优化后需要以 $((y_i, y_1), (y_i, y_2), \dots, (y_i, y_n))$ 的形式去嵌入，其中 $1 \leq i \leq n$ 。

定义 1 样本的每个特征作为量子门角度嵌入量子门后转换为量子态，该量子态的振幅状态可以影响其他样本特征的嵌入，这个过程称为量子态振幅单向驱动的特征嵌入（简称振幅驱动嵌入）。

具体而言，每个样本特征 y_i 作为量子门角度嵌入量子门后转化为量子态 $|\psi_i\rangle = \alpha_i|0\rangle + \beta_i|1\rangle$ ，其中振幅 α_i 和 β_i 的取值取决于 y_i 和量子门的类型；假设以 y_i 为角度的量子门所在量子位为 q_i ，而以 y_j 为角度的控制量子门的控制位为 q_i 时， y_j 的嵌入由振幅 α_i 和 β_i 状态所决定，这是因为测量 q_i 时，量子态会以 $|\alpha_i|^2$ 的概率塌缩成 $|0\rangle$ ，以 $|\beta_i|^2$ 的概率塌缩成 $|1\rangle$ ，即 q_i 为 $|0\rangle$ 时， y_j 不会被嵌入； q_i 为 $|1\rangle$ 时， y_j 被嵌入。所以在嵌入 y_i 的同时会影响 y_j 的嵌入，产生振幅驱动嵌入的机制。

现有的角度编码容易造成大量量子位浪费，而振幅编码则会导致线路深度过高并引入更多噪声，致使编码效率低下。因此，为了优化量子位的使用和线路深度，融入振幅驱动嵌入思想，在构建编码线路之前，构建出一种新的数据结构——量子态振幅单向驱动的特征嵌入二叉树（简称振幅驱动嵌入树），以优化量子编码过程中对特征嵌入关系的表征和处理效率。

振幅驱动嵌入树是一种有方向性的二叉树，每个节点表示一个待编码的特征，特征之间具有依赖关系。通过量子门角度嵌入，每个节点的特征转化为相应的量子态。树的每条边则反映了不同特征之间的依赖关系，尤其是在量子门作用下的特征传递和控制机制。振幅驱动嵌入树的构建过程如下。

1) 根节点选取。选取重复出现的特征作为二叉树的根节点 v_0 ，通过量子门角度嵌入将其转化为

初始量子态。接下来根据特征间的依赖关系, 逐步构建子节点。

2) 子节点与边的建立。每个子节点 v_i 通过量子门与父节点相连, 边 $e_{i,j}$ 表示父节点对子节点的控制关系。假设 v_i 是控制位, 则它会影响子节点 v_j 的嵌入状态。若父节点的量子态为 0, 则子节点不会被嵌入; 若为 1, 则子节点被嵌入。振幅驱动嵌入树如图 6 所示。

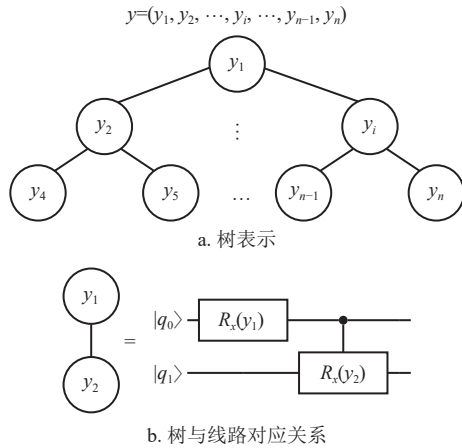


图 6 振幅驱动嵌入树

在振幅驱动嵌入树中, 结点中的值直接表示需要嵌入的数据集样本特征值; 每条边表示一个双量子位门, 其中控制位为父节点所在的量子位; 如图 6a 中 y_1 和 y_2 子树, y_1 所在结点为控制结点, q_0 为控制位, y_2 所在结点为目标结点, q_1 为目标位。 y_1 在嵌入后, q_0 量子位的振幅状态会发生变化, 而同时 q_0 作为双量子门控制位, 又由振幅驱动嵌入定义可知 y_2 的嵌入由 q_0 的振幅状态所决定。这种特性为后面提出的混合编码可以优化量子线路宽度且进一步减少特征嵌入奠定了基础。

引理 1 若通过构建振幅驱动嵌入树进行数据编码, 则原始数据集的维度不能超过 4, 即 $1 \leq n \leq 4$ 。

证明: 假设给定一个数据集样本特征为 5, 基于图 4 得到 1 组简化特征 $((y_1, y_1), (y_1, y_2), (y_1, y_3), (y_1, y_4), (y_1, y_5))$, 如图 7 所示。图 7a 中如果按照两两振幅驱动嵌入的方式, 结点中只能出现两次 y_1 , 因为只有这样才不会重复编码; 然而根据图 7b 可知, 只有两个 y_1 无法控制其他特征完成全部的嵌入, 这与假设的特征嵌入方式相矛盾, 因此数据集的维度不可以超过 4。证毕。

定理 1 在 $((y_i, y_1), (y_i, y_2), \dots, (y_i, y_n))$ 中, y_i 需要

嵌入 n 次, 其中 $n > 2$; 在振幅驱动嵌入特征后, 每个特征 y_i 只需嵌入 m 次, 其中 $1 \leq m \leq 2$ 。

证明: 在引理 1 给出的特征维数上界内, y_i 在振幅驱动嵌入树中, 最多出现在两个结点里, 即 y_i 的嵌入次数最多为 2。证毕。

定理 2 在样本中存在 n 个重复特征, 经过振幅驱动嵌入后, 重复特征的数量减少至 m 个, 量子门的数量相应减少 $(n-m)$ 个。

证明: 在振幅驱动嵌入树中, 每个节点表示一个待编码的特征, 而每编码一个特征就需要一个量子门, 又由定理 1 可知, 在振幅驱动嵌入树中, 重复出现 n 次的特征减少到 m 次, 因此所需要的量子门也会减少 $(n-m)$ 个。证毕。

以 $(x_1, x_1, x_1, x_2, x_1, x_3)$ 数据样本为例, x_1 在这个样本中出现多次, 如果利用振幅驱动嵌入树, x_1 能够控制其他结点的特征嵌入, 因此可以进一步减少特征的输入, 过程示意图如图 8 所示。

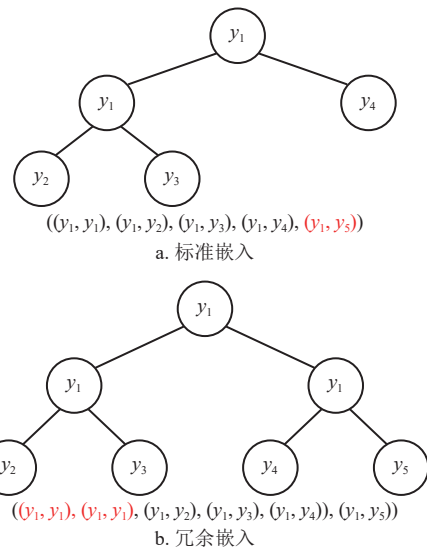


图 7 5 特征振幅驱动嵌入树

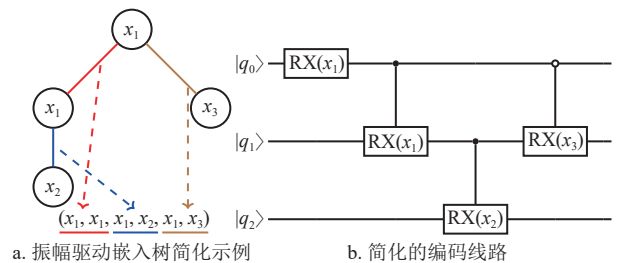


图 8 编码优化

图 8b 中, 当 x_1 以角度的形式编码到线路后, q_0 量子位就变成了如式 (1) 所示的形式, 因此对于 x_2 和 x_3 特征的嵌入取决于 q_0 和 q_1 量子位的振幅状态, 由于最后需要进行大量测量获取量子态的概率

分布, 因此 x_2 和 x_3 特征都能够以一定的概率嵌入量子线路中, 从而实现对所有特征的嵌入。编码相同数量的特征, 不仅融合角度编码的简单嵌入方式, 而且相比于图 5 中的振幅编码线路, 图 8b 所示的线路只有双量子位门。

2.3 角度-振幅混合编码

选择合适的数据编码方法须在计算效率与硬件限制间权衡。本文为开发高效的 QNN 综合考虑角度编码和振幅编码的优劣, 以优化模型性能, 提出角度-振幅混合编码。角度-振幅混合编码是一种结合了角度编码和振幅编码的量子态表示方法, 用于在量子线路中高效嵌入经典数据。

2.3.1 混合编码原理

在混合编码方式中, 经典数据的一部分被嵌入到量子位的旋转角度, 通过量子门控制量子态的相位。角度参数决定了每个量子位的相对旋转, 从而影响量子态的整体相位分布。另一部分经典数据则通过振幅参数来影响量子态的振幅分布, 直接决定了量子态各分量的概率幅度。

假设有一个经典数据向量 $\mathbf{x} = [x_1, x_2, \dots, x_n]$, 可以使用以下的混合编码方法:

$$|\psi\rangle = \otimes_{i=1}^n \left(\cos\left(\frac{\theta(x_i)}{2}\right)|0\rangle + \sin\left(\frac{\theta(x_i)}{2}\right)\alpha(x_i)|1\rangle \right) \quad (7)$$

式中, \otimes 为张量积符号; $\theta(x_i)$ 表示与特征 x_i 相关的角度参数; $\alpha(x_i)$ 为与 x_i 相关的振幅参数。若 x_i 的量子态测量为 $|1\rangle$, 则该特征控制其他特征的嵌入; 若为 $|0\rangle$, 则控制不同的特征嵌入。这种振幅驱动嵌入机制形成了完整的量子系统, 利用量子态间的相互依赖性优化了经典数据的量子表示。

角度-振幅混合编码为量子线路设计提供了一种高效且资源节约的方案。

1) 与角度编码相比, 该方法有效减少了量子位需求。通过振幅驱动嵌入树, 所需的量子位数量等于树的深度。编码 n 个特征时, 角度编码需要 n 个量子位, 而混合编码仅需 $\log_2 n$ 个量子位, 优化了线路宽度。

2) 振幅驱动嵌入树中的每条边仅涉及两量子位门, 减少了对多量子位门的依赖, 线路深度进一步降低; 根据定理 2, 可知量子门资源也得到减少。这一优化不仅简化了量子态制备过程, 还通过降低线路宽度和深度, 降低了噪声的影响, 进而减小了量子计算过程中误差积累的风险, 提升了量子线路计算的稳定性。

2.3.2 编码线路构建

基于角度-振幅混合编码原理, 构建 QNN 的编码层。2.1 节中提到, 在角度编码中可通过减少冗余特征输入以减少量子资源浪费, 但仍会有一些特征会多次被利用, 因此到这一步也并不是特征输入的最简办法。

基于 2.2 节振幅驱动嵌入树来实现混合编码, 即当一个数据特征嵌入的同时还能控制另一个数据特征的嵌入, 可以再一次减少数据特征的重复编码。如图 9 所示, 对 $(x_i, x_i, x_i, x_j, x_i, x_{j+1}, x_i, x_{j+2})$ 这组数据特征, 使用图 9a 的二叉树只需要嵌入 x_i 两次, 而使用图 9b 的角度编码需要重复 5 次 x_i 的嵌入。图 9a 中二叉树的结点代表一个量子门, 边代表门控制位数。

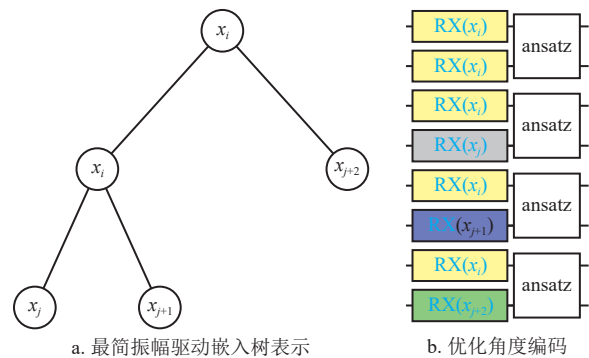


图 9 二叉树与编码线路

二叉树的建立目的是构建相应的编码层, 按照图 9a 中节点和边与量子门的对应关系, 画出相应的编码线路。这种编码线路是基于提出的新编码方式而构建, 如图 10 所示。图中编码层可以通过将数据集特征直接嵌入旋转门角度中来制备相应的振幅, 其中 4 条红色虚线代表同时嵌入的 4 组数据。因为量子态的叠加特性, 其中一个特征嵌入后得到相应的量子叠加态可以控制另一特征的嵌入, 因此在量子线路的多次运行下, 各组数据的嵌入都可以实现。这里需要注意, 所有数据集在预处理时都需要降维至 4 维及以下, 这种编码方式才适用。

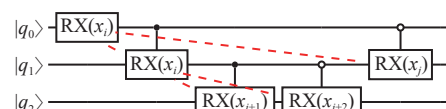


图 10 角度-振幅混合编码线路

2.3.3 基于混合编码线路的数据预处理

在将数据集嵌入图 10 所示的编码线路之前, 需要对嵌入的数据特征进行简化并重新排列, 在和

编码线路中的量子门形成一一对应的关系之后, 将数据集直接嵌入到线路中。数据集处理的过程总结如下。

设 X 表示一个二分类数据集 $X = (x_1, x_2, \dots, x_K)$, x_i 称为 X 中的第 i 个样本, $i = 1, 2, \dots, K$; 每个样本都包含 L 个特征, $1 \leq L \leq 4$, 即 $x_i = (a_1, a_2, \dots, a_L)$, a_j 为 x_i 的第 j 个特征, $j = 1, 2, \dots, L$ 。对数据集进行预处理, 由引理 1 可知, 需要对数据集特征数量大于 4 维的特征降维, 将数据集维度降至 4 维或以下。

1) 设 x'_i 为 x_i 将特征重排之后的样本, 此时的 x'_i 特征数量从 L 个变成 $3L^2$ 个, 有:

$$x'_i = ((a_1, a_1, a_1, a_1, a_2, a_2, \dots, a_1, a_L, a_L), \dots, (a_L, a_1, a_1, a_L, a_2, a_2, \dots, a_L, a_L, a_L)) \quad (8)$$

2) 对于重复的特征, 只取重复特征的其中一个进行简化, 设 x''_i 为 x'_i 简化后的样本, x''_i 的特征数量由 $3L^2$ 变成了 $2L^2$ 个, 简化后的数据特征变成:

$$x''_i = ((a_1, a_1, a_1, a_2, \dots, a_1, a_L), \dots, (a_L, a_1, a_L, a_2, \dots, a_L, a_L)) \quad (9)$$

3) 根据图 9a 二叉树, x''_i 还可以进一步简化, x'''_i 是基于角度-振幅混合编码简化至能够直接嵌入线路中的样本, 因为 $2 \leq L \leq 4$, 则 x'''_i 分情况讨论, 表示为:

$$x'''_i = \begin{cases} ((a_1, a_1, a_2), (a_2, a_2, a_1)) & L = 2 \\ ((a_1, a_1, a_2, a_3), (a_2, a_2, a_1, a_3), (a_3, a_3, a_1, a_2)) & L = 3 \\ ((a_1, a_1, a_4, a_2, a_3), (a_2, a_2, a_4, a_1, a_3), (a_3, a_3, a_4, a_1, a_2), (a_4, a_4, a_1, a_2, a_3)) & L = 4 \end{cases} \quad (10)$$

最终, 初始数据集 X 经过一番简化后, 获得新的数据集 $X_s = (x'''_1, x'''_2, \dots, x'''_K)$ 。而新的数据集能够直接嵌入到图 10 中的编码层。

3 基于角度-振幅混合编码的 Ansatz 设计

通过 2.2 节提出的编码方式和数据预处理方法, 量子线路宽度有效减少并得到最简数据集, 将新的数据集直接嵌入到量子态中, 得到相应的量子数据。虽然模型结构复杂的问题得到了有力的解决, 而针对提出的使用混合编码得到的经典数据的量子态, 设计出符合角度-振幅混合编码理念的 Ansatz, 建立不同量子态的纠缠, 实现数据之间的内在联系同样重要。

3.1 双环 Ansatz

本文提出一种新型的量子线路结构, 称为“双环” Ansatz。该 Ansatz 由两个嵌套的旋转操作环路组成, 旨在优化量子位间的相互作用并提升量子线路的表达能力。根据提出的角度混合编码的原理, 构建 Ansatz, 如图 11 所示。

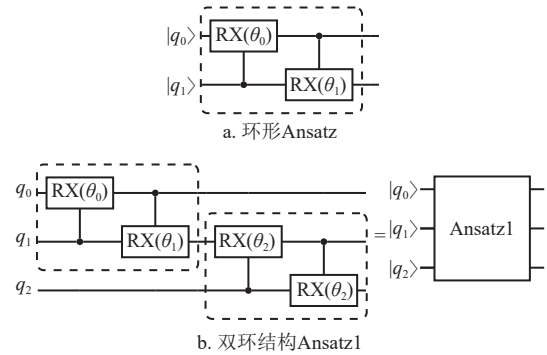


图 11 QNN 训练层

首先通过对需要编码的数据集进行预处理, 将其按照角度-振幅混合编码方式进行量子编码; 接着将生成的量子编码系统进行合理划分, 将整个编码数据的过程分组, 确保不同子系统内的量子信息能够独立处理; 然后将每组编码视为一个子问题, 最终把各组编码的量子态通过构建的训练层汇总在一起, 如图 12 所示。其中 3 量子位的量子系统切割为 $A: (q_0, q_1)$ 与 $B: (q_1, q_2)$ 两个量子子系统, 把 (x_i, x_i, x_i, x_j) 这组数据编码到 A 量子系统中, 而 $(x_i, x_{j+1}, x_i, x_{j+2})$ 这组数据编码到 B 量子系统中。在 A 量子系统中, 数据经过编码到量子态进行演化后, 子系统所产生的量子态位串表示为:

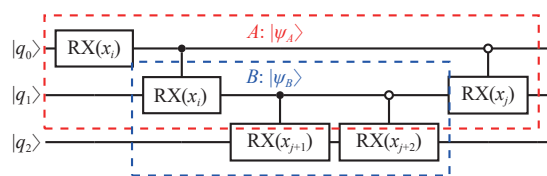


图 12 编码线路的划分

$$|\psi_A\rangle = \alpha_0 |00\rangle + \alpha_1 |01\rangle + \alpha_2 |10\rangle + \alpha_3 |11\rangle \quad (11)$$

式中, $\alpha_0 |00\rangle, \alpha_1 |01\rangle, \alpha_2 |10\rangle, \alpha_3 |11\rangle$ 可视为将一组经典数据 (x_i, x_i, x_i, x_j) 嵌入量子门角度后所得到的量子数据。同理 B 量子系统表示为:

$$|\psi_B\rangle = \beta_0 |00\rangle + \beta_1 |01\rangle + \beta_2 |10\rangle + \beta_3 |11\rangle \quad (12)$$

式中, $\beta_0 |00\rangle, \beta_1 |01\rangle, \beta_2 |10\rangle, \beta_3 |11\rangle$ 可视为将另一组经典数据 $(x_i, x_{j+1}, x_i, x_{j+2})$ 嵌入量子门角度后所得到的

量子数据。将编码线路所代表的量子系统划分之后,利用量子线路的纠缠特性,针对每一步划分,构建相同架构的图 11a 环形 Ansatz,然后将这些所有的 Ansatz 嵌套在一起,嵌套后的 Ansatz 如图 11b 所示。

在量子线路设计中,图 11b 中构建的 Ansatz1 通过纠缠操作将编码后的量子数据进行处理,从而近似实现式 (4)。具体而言, Ansatz1 的纠缠结构使得不同量子态之间能够进行复杂的相互作用,这种相互作用对应于式 (4) 中的查询向量 Q_i 和键向量 K_i 的点积计算。通过这种量子纠缠操作,量子线路能够有效地模拟经典 SAM 中向量的相似度计算过程,并通过量子态的叠加和纠缠特性实现 SAM 的核心功能。这一方法自然地将 SAM 嵌入量子线路的设计中,充分利用了量子计算的并行性和纠缠性,实现对经典数据的高效处理。

图 11b 中的 Ansatz1 完全基于角度-振幅混合编码而设计,它是一个两层环形线路组合而成的结构,这种结构相对简单,线路深度较浅,从而减少量子退相干和噪声影响,有利于在实际硬件上实现稳定的量子计算。

3.2 基于双环 Ansatz 的线路宽度选择

经过第 1 层 Ansatz1 后,得到式 (4) 中的 $a_{i,j}$, $a_{i,j}$ 在量子线路中以振幅的形式存在。由于在量子系统中所有振幅的平方和为 1,因此注意力权重 $w_{i,j}$ 可以自然地从业式 (5) 中获得。训练过程中使用的其他 Ansatz 结构如图 13 所示。

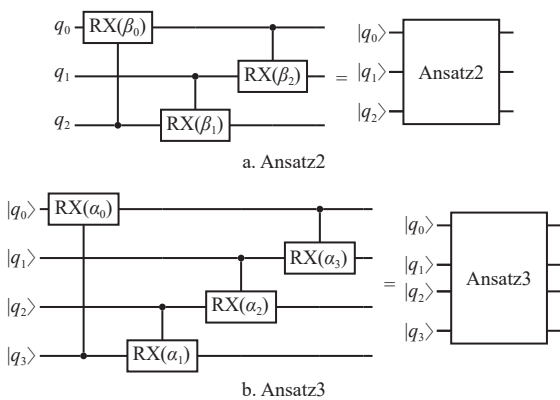


图 13 QNN 训练层

具体而言,1.3 节提到的注意力分数的计算采用点积运算,因此式 (6) 可以用矩阵运算形式在式 (13) 中直观表示。图 13 中的 Ansatz 结构展示了实现这些计算的量子线路。给定一个数据集,通过 2.3.2 节中构建的编码线路将经典数据嵌入量子态

中,从而获得式 (13) 中的值向量 $V = \{V_j, 1 \leq j \leq n\}$ 。

图 13a 中的 Ansatz2 用于模拟式 (13) 中的权重系数 $W = \{w_{i,j}, 1 \leq i \leq n, 1 \leq j \leq n\}$ 与值向量 V_j 的乘积。

$$\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,n} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & \cdots & w_{n,n} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} = \begin{bmatrix} w_{1,1}V_1 + w_{1,2}V_2 + \cdots + w_{1,n}V_n \\ w_{2,1}V_1 + w_{2,2}V_2 + \cdots + w_{2,n}V_n \\ \vdots \\ w_{n,1}V_1 + w_{n,2}V_2 + \cdots + w_{n,n}V_n \end{bmatrix} \quad (13)$$

经过多组 Ansatz2 的训练后,需要对所有的 $w_{i,j}V_j$ 进行加和计算,即式 (13) 中 $A = \{w_{i,1}V_1 + w_{i,2}V_2 + \cdots + w_{i,n}V_n, 1 \leq i \leq n\}$ 。图 13b 中的 Ansatz3 通过构建一个维环形量子线路,实现了 Attention_i 计算。

综上所述, Ansatz1、 Ansatz2 和 Ansatz3 的结构设计借鉴了 SAM “分而治之”的计算流程: Ansatz1 用于生成初始权重 $w_{i,j}$; Ansatz2 完成权重与值向量的逐元素乘积 $w_{i,j}V_j$; Ansatz3 进一步将所有结果求和,形成最终注意力向量 A_i 。由于这些 Ansatz 用于 QNN 的参数训练,因此需要在训练过程中不断优化参数,以逼近式 (13) 的结果。

在双环 Ansatz 结构确定后,后续 Ansatz 量子线路的宽度由双环 Ansatz 的宽度决定。

定理 3 设数据集特征维度为 N , 则双环 Ansatz1 的线路宽度为 $\lceil \log_2(N+1) \rceil$, Ansatz2 线路宽度为 $\lceil \log_2(N+1) \rceil$, Ansatz3 线路宽度为 N 。

证明:特征维度重排与分组嵌入:已知数据集特征维度为 $d = N$,通过式 (8) 和式 (9) 可得重排后的数据集特征维度 $d' = 2N^2 - N(N-1)$;调整后的特征需分组嵌入,这是基于 SAM 的分而治之策略,特征维度为对应划分的组数 $G = N$,每组特征数量为 $k = N+1$ 。

双环 Ansatz1 的线路宽度 W_1 和数量 N_1 :由于特征以振幅形式编码至量子线路,编码每组特征所需的量子位数量为 $q = \lceil \log_2(k) \rceil$,由此得出双环 Ansatz1 的线路宽度为 $W_1 = q$;由于需对 G 组特征进行独立处理, Ansatz1 的数量为 $N_1 = G$ 。

Ansatz2 的线路宽度 W_2 : Ansatz2 连接双环 Ansatz1 的全部量子位,因此其线路宽度为 $W_2 = W_1$ 。

Ansatz3 的线路宽度 W_3 : Ansatz3 只选择 $N_1/2$ 个 Ansatz2 的两个量子位进行计算,因此, Ansatz3 的线路宽度 $W_3 = 2N_1/2 = G$ 。证毕。

4 基于角度-振幅混合编码的量子神经网络

4.1 模型构建

基于角度-振幅混合编码量子神经网络完整模型由 4 个部分组成, 如图 14 所示。

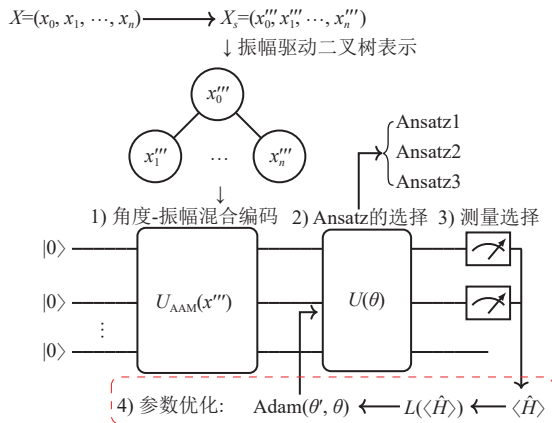


图 14 AAM-QNN 模型示意图

1) 角度-振幅混合编码

根据式 (10) 对数据集样本进行预处理, 通过振幅驱动二叉树表示数据集样本的每个特征, 最终根据二叉树结构确定编码线路结构。

2) Ansatz 的选择

基于图 11 中的环形 Ansatz 构建双环结构 Ansatz1, 并基于双环 Ansatz 选择图 13 中的 Ansatz2 和 Ansatz3 的线路宽度, 最终确立训练层的线路结构。

3) 测量选择

量子测量用于输出分类或回归任务的预测。在量子测量过程中, 可以仅对与目标可观测量或代价函数相关性最高的部分量子位进行测量^[14,19]。本文中因为量子神经网络的每个量子位与最终分类精度有关联, 有些量子位得到的测量结果有助于提高分类精度, 而有些量子位对分类精度并不敏感。本文选取部分量子位进行测量, 与选取全部量子位测量结果的预测误差不超过 1%。采用 Pauli-Z 测量, 首先测量每个待测量子 Z 方向的期望值, 然后将该期望值与任务标签进行对比, 测量过程中期望值接近+1 时, 判定为标签 1, 期望值接近-1 时, 判定为标签 0。通过这种方式, 量子测量结果被映射为具体的分类标签或回归值, 进而实现预测功能。

4) 参数优化

在经典计算机上利用 Adam 优化器更新 Ansatz

中的量子门角度参数。数据编码线路以及所有的 Ansatz 线路都是基于 PQC 实现的。对于提出的新编码线路, 其宽度和结构不同, 采用 Ansatz 的宽度和结构也有所不同。图 11 中 Ansatz1 是为双环形架构, 图 13 中的 Ansatz2 和 Ansatz3 都是采用的环形架构^[20]。

4.2 算法实现

基于角度-振幅混合编码方式的量子神经网络生成算法如算法 1 所示。

算法 1 AAM-QNN 生成算法

Input: the feature matrix //输入特征矩阵

Output: AAM-QNN //输出 AAM-QNN 模型

feature_re_sort feature_matrix //按式 (3) 排序特征

dim \leftarrow feature_matrix//获取特征维度

If dim > 4 then

reduce_dim = PCA (dimension = 4) //降维

feature_simplified \leftarrow feature_re_sort //简化特征

N = feature_simplified_num//获取简化后特征数量

feature_simplified_group \leftarrow feature_simplified//分

组处理简化特征

M = feature_simplified_group.num

Gate_list = Amplitude_encode_modified_circuit//获取编码层所有量子门

While i <= dim do //嵌入所有特征

i \leftarrow 1

j \leftarrow 1

While j < N/dim do

Gate_list_angle[j] = feature_simplified_group [i][j]//将角度嵌入门集合

j \leftarrow j+1

i \leftarrow i+1

End

Return circuit_based_on_Angle-Amplitude Mixed

End

AAM-QNN \leftarrow circuit_based_on_Angle-Amplitude

Mixed //嵌入编码层到 QNN

Divide the circuit_based_on_Angle-Amplitude Mixed

//划分线路

Count= Number of copies divided //统计划分份数

Get (circular_Ansatz) //获取环形 Ansatz

Ansatz1_based_on_AAM circular_Ansatz//嵌入 Ansatz1

AAM-QNN Ansatz1_based_on_AAM, Ansatz2, Ansatz3//获得 AAM-QNN

算法 1 通过分组的方式对特征排序与简化进行并行处理, 然后将处理后的特征嵌入到构建的角度-振幅混合编码线路中, 最后设计出双环架构的量子线路, 如图 14 所示的 AAM-QNN 模型。角度-振幅混合编码方式对量子位的降低是数量级的。

对于 N 维数据集, 使用角度编码所需要的量子位数近似为 $O(N^2)$, 而使用角度-振幅混合编码所需要的量子位数近似为 $O(N \log_2(2N-3))$ 个量子位。再看参数部分, 对于 N 维数据集, 使用角度编码需要 $2N^2+8N+8$ 个参数, 而 AAM-QNN 模型只需要 $(N+2)\log_2(2N-3)+N-2$ 个参数, 研发的模型在计算资源上更加高效。

5 实验及结果分析

本文所提的方法基于 Python 语言实现。实验在 Qiskit (<https://qiskit-community.github.io/qiskit-machine-learning/tutorials/>) 平台实现, 实验环境为 Windows11 操作系统, AMD Ryzen 7 7735H with Radeon Graphics@3.20 GHz, 32 GB 内存。利用鸢尾花数据和 MC、RP 文本数据对本文提出的 AAM-QNN 模型进行具体的实验验证。本文将二元交叉熵作为损失函数, 其中在优化迭代过程中选择 Adam 优化器对参数进行更新。

5.1 鸢尾花分类

第 1 个任务是鸢尾花分类 (<https://archive.ics.uci.edu/dataset/53/iris>), 使用包含 100 个样本的数据集, 其中每个类别 (setosa 和 versicolour) 各有 50 个样本。每个样本具有 4 个属性: 萼片长度、萼片宽度、花瓣长度和花瓣宽度。数据集被划分为训练集 (80 个样本) 和测试集 (20 个样本)。角度-振幅混合编码是一种通用框架, 能够在数据预处理并按式 (10) 排列后直接嵌入编码线路。根据式 (10), 4 维的鸢尾花数据集可以整理为 20 维。图 15 展示了基于图 10 编码线路构建的编码层。其中虚红线把每个相同编码架构的子线路分割开, 表示编码过程中将数据集分成每组并行编码, 提高编码效率。

AAM-QNN 的整个训练层如图 16 所示。其中图 13 选取 3 种 Ansatz 构建训练层, 选择图 13b 中 Ansatz3 时, 对于虚红线划分的 4 个编码层子线路, 每层线路选择其中的两个量子位作为 Ansatz3 的输入。

在实现对鸢尾花数据集的分类时, 模型的训练损失曲线如图 17 所示。AAM-QNN 模型达到 100%

的分类准确率, 结果表明本模型也可以有效捕捉鸢尾花样本 4 个属性之间的内在联系, 因为机器学习更关注模型在未见数据上的表现, 即从图 17 中的测试集损失曲线可以看出, 从第 150 次迭代训练后, 模型开始收敛, 且损失值逐渐趋近于 0.064, 这说明模型在训练过程中学到了数据中的模式并且能够很好地预测输出。

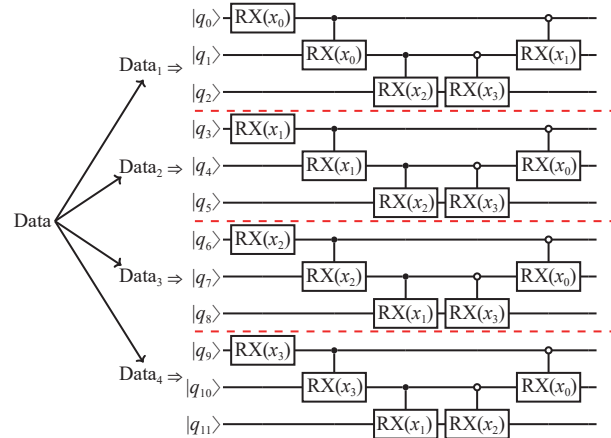


图 15 AAM-QNN 的编码层

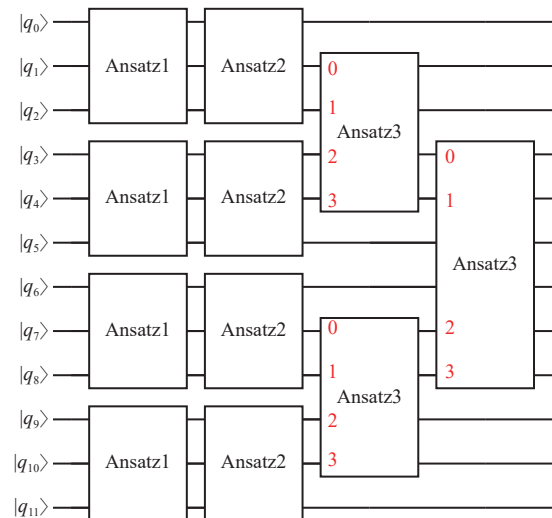


图 16 AAM-QNN 的训练层

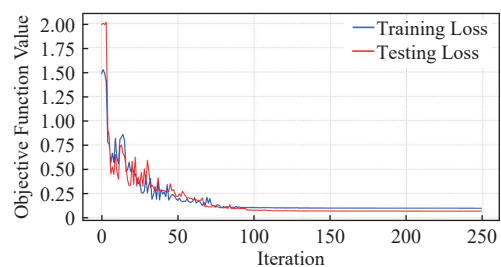


图 17 在鸢尾花数据上的训练损失曲线图

5.2 文本分类

本文采用的合成数据集来自文献 [21], 分别命

名为 MC 和 RP。MC 包含 17 个单词和 130 个句子 (70 个训练+ 30 个开发+ 30 个测试), 每个句子包含 3 或 4 个单词, MC 的任务是将一个句子分类为食物或 IT; RP 有 115 个单词和 105 个句子 (74 个训练+ 31 个测试), 每个句子有 4 个单词, 其目标是预测某个名词短语是否包含主语关系从句或宾语关系从句, 此外, 词汇量的大小和词汇的稀疏性使得该任务比 MC 任务更具挑战性。

对于文本数据集, 首先用词向量来表示词, 这个词向量采用词袋模型对文本型的文件进行一个特征提取, 然后将提取后的文本特征转化为词频矩阵。经过这些处理, 可以得到数值化的文本数据集。

其编码层设计以及训练层设计与图 15 和图 16 一致, 因为该模型是通用模型, 无须对不同类型数据集专门设计不同的编码线路, 实验结果如表 2 所示。Complexity 表示量子位数/两量子位门数, 且分类精度展示的是对测试集上的分类精度。

表 2 显示, 使用角度-振幅混合编码后的模型 AAM-QNN 可以轻松实现 MC 数据集 100% 分类精度, 优于其他用于处理自然语言数据集的量子学习模型; 同时只需要 15 个训练参数以及较低的线路复杂度。而在 RP 任务中, 本文模型的分​​类准确率为 80.65%, 高于其他 3 种模型。

表 2 在 MC 和 RP 测试集上的对比结果

Model	Complexity		Trainable Parameters		Test Accuracy/%	
	MC	RP	MC	RP	MC	RP
DisCoCat ^[21]	—	—	40	168	79.80	72.30
QSANN ^[10]	24/24	48/240	25	109	100	67.74
QSAMB ^[14]	96/136	96/136	22	22	100	72.58
QSAMo ^[14]	64/104	64/104	20	20	100	74.19
AAM-QNN	12/56	12/56	15	15	100	80.65

在实现对文本数据集的分类时, 模型在 MC 数据集上的训练损失曲线如图 18 所示。AAM-QNN 模型的训练曲线在 100 次迭代的训练损失值就已大幅降低, 最终收敛于一个较小的损失值, 这表明本文提出的模型在训练数据上表现良好, 且在测试数据上也有良好的泛化能力。

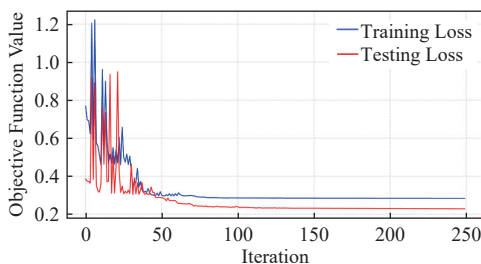


图 18 在 MC 数据集上的训练损失曲线图

模型在 RP 数据集上的训练损失曲线如图 19 所示。RP 数据集上曲线收敛状态相比之下更加稳定。并且在 200 次迭代后, 仍有下降趋势, 随着迭代次数增加, 有望获取更好的收敛效果。此外, 除了参数优化过程是在经典计算机上处理, 其他所有计算处理都是由完整的量子线路来实现, 这也让整个模型更加发挥量子计算的优势。

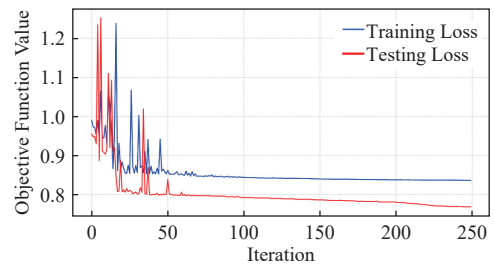


图 19 在 RP 数据集上的训练损失曲线图

5.3 图像分类

为了进一步验证模型在具有更高数据复杂度情况下的性能, 本节选取了 MNIST 数据集的两个具有不同识别难度的二分类任务 (0, 1) 与 (3, 6) 进行测试。该数据集包含大量不同风格的手写数字样本, 具有较高的代表性与挑战性。训练集包括 1 000 个样本, 测试集包括 100 个样本。

在实验之前, 需要按照 2.3.3 节的方法对图像数据集进行预处理, 即需要将图像数据集降维并按照式 (10) 对数据集的特征重排, 然后使用 AAM-QNN 模型进行训练, 如 5.2 节所述, AAM-QNN 是一种通用模型, 因此无须在结构上做任何改变。针对两个二分类任务测试精度结果如表 3 所示。其中, 文献 [22] 中的 QCNN 模型只对 (0, 1) 二分类任务进行测试, 因此在 (3, 6) 的测试任务中没有数据; 此外, 本文提出的模型在测试集上的分类精度采用了平均精度作为评估指标。

通过表 3 可知, 本文提出的模型在 (0, 1) 二分类任务中效果最优, 达到了 100% 的分类精度; 与文献 [22] 提出的张量网络相比, 在 (3, 6) 分类任务上分类精度与张量网络相近, 并且 AAM-QNN 模型只需要训练 15 个参数即可达到很好的训练效果。

表 3 在 MNIST 二分类测试集上的对比结果

Model	Parameters	Accuracy/%	
		(0,1) Test	(3,6) Test
张量网络 ^[22]	1 008	99.6	97.2
MLP ^[23]	630	99	94
QCNN ^[24]	18	98.2	—
AAM-QNN	15	100	97

6 结束语

本文提出了一种角度-振幅混合编码方法, 这种编码方法通过构建量子态振幅单向驱动的特征嵌入二叉树, 允许数据以直接嵌入量子门的方式进行编码, 显著降低了模型的复杂性。通过结合角度编码和振幅编码的优点, 设计了一种能够自然融合SAM的QNN, 称为AAM-QNN。此外, 通过将模型编码层表示的量子系统分割成子问题, 并将子问题的结果整合, 设计出了一种具有强大纠缠能力的双环架构 Ansatz, 使其更适合当前的NISQ设备。AAM-QNN模型在鸢尾花分类任务中的训练损失低于已有模型; 在文本分类任务中优于其他处理自然语言处理的量子模型和简单的QNN, 且在更大数据规模的MNIST不同难度二分类数据集上验证了模型的性能。未来工作将进一步优化模型, 提高其复杂度和准确性, 并采用更先进的技术, 如多头自注意机制, 以提升模型的应用能力。

参考文献

- [1] PRESKILL J. Quantum computing 40 years later[M]// Feynman Lectures on Computation. Boca Raton: CRC Press, 2023: 193-244.
- [2] XU F H, MA X F, ZHANG Q, et al. Secure quantum key distribution with realistic devices[J]. *Reviews of Modern Physics*, 2020, 92(2): 025002.
- [3] 徐雅斌, 陈淑娟, 李艳平. 量子密钥分发网络的多路径密钥传输方法研究[J]. *电子科技大学学报*, 2020, 49(2): 276-282.
XU Y B, CHEN S J, LI Y P. Research on multipath key transmission in quantum key distribution networks[J]. *Journal of University of Electronic Science and Technology of China*, 2020, 49(2): 276-282.
- [4] 吴涵卿, 袁湫木, 陈柄任, 等. 量子近似优化算法在投资组合优化中的应用[J]. *电子科技大学学报*, 2023, 52(5): 642-648.
WU H Q, YUAN H M, CHEN B R, et al. The application of quantum approximation optimization algorithm in portfolio optimization[J]. *Journal of University of Electronic Science and Technology of China*, 2023, 52(5): 642-648.
- [5] BRANDAO F G S L, SVORE K M. Quantum speed-ups for solving semidefinite programs[C]//Proceedings of the IEEE 58th Annual Symposium on Foundations of Computer Science. [S.l.]: IEEE, 2017: 415-426.
- [6] BIAMONTE J, WITTEK P, PANCOTTI N, et al. Quantum machine learning[J]. *Nature*, 2017, 549(7671): 195-202.
- [7] REBENTROST P, MOHSENI M, LLOYD S. Quantum support vector machine for big data classification[J]. *Physical Review Letters*, 2014, 113(13): 130503.
- [8] HUANG H Y, BROUGHTON M, MOHSENI M, et al. Power of data in quantum machine learning[J]. *Nature Communications*, 2021, 12: 2631.
- [9] SCHULD M, PETRUCCIONE F. Machine learning with quantum computers[M]. 2nd ed. Berlin: Springer, 2021.
- [10] LI G X, ZHAO X Q, WANG X. Quantum self-attention neural networks for text classification[J]. *Science China Information Sciences*, 2024, 67(4): 142501.
- [11] DI SIPIO R, HUANG J H, CHEN S Y, et al. The dawn of quantum natural language processing[C]//Proceedings of the ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing. [S.l.]: IEEE, 2022: 8612-8616.
- [12] SHI J J, ZHAO R X, WANG W X, et al. QSAN: A near-term achievable quantum self-attention network[EB/OL]. [2024-10-01]. <https://arxiv.org/abs/2207.07563>.
- [13] ZHAO R X, SHI J J, LI X L. QKSAN: A quantum kernel self-attention network[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(12): 10184-10195.
- [14] SHI S S, WANG Z M, LI J X, et al. A natural NISQ model of quantum self-attention mechanism[EB/OL]. [2024-10-01]. <https://arxiv.org/abs/2305.15680>.
- [15] KWAK Y, YUN W J, JUNG S, et al. Quantum neural networks: Concepts, applications, and challenges[C]//Proceedings of the 12th International Conference on Ubiquitous and Future Networks. [S.l.]: IEEE, 2021: 413-416.
- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 6000-6010.
- [17] SCHULD M. Supervised quantum machine learning models are kernel methods[EB/OL]. [2024-02-21]. <https://arxiv.org/abs/2101.11020>.
- [18] ARAUJO I F, PARK D K, LUDERMIR T B, et al. Configurable sublinear circuits for quantum state preparation[J]. *Quantum Information Processing*, 2023, 22(2): 123.
- [19] SIM S, JOHNSON P D, ASPU - GUZIK A. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum - classical algorithms[J]. *Advanced Quantum Technologies*, 2019, 2(12): 1900070.
- [20] ZHAO A, TRANTER A, KIRBY W M, et al. Measurement reduction in variational quantum algorithms[J]. *Physical Review A*, 2020, 101(6): 062322.
- [21] LORENZ R, PEARSON A, MEICHANETZIDIS K, et al. QNLP in practice: Running compositional models of meaning on a quantum computer[J]. *Journal of Artificial Intelligence Research*, 2023, 76: 1305-1342.
- [22] HUGGINS W, PATIL P, MITCHELL B, et al. Towards quantum machine learning with tensor networks[J]. *Quantum Science and Technology*, 2019, 4(2): 024001.
- [23] TAUD H, MAS J F. Multilayer perceptron (MLP)[M]//Geomatic Approaches for Modeling Land Change Scenarios. Cham: Springer International Publishing, 2017: 451-455.
- [24] HUR T, KIM L, PARK D K. Quantum convolutional neural network for classical data classification[J]. *Quantum Machine Intelligence*, 2022, 4(1): 3.