arXiv:2302.02980v1 [quant-ph] 6 Feb 2023

量子特征映射的混合遗传优化

Rowan Pellow-Jarman[1,2*]、Anban Pillay[1,3], Ilya Sinayskiy[1,2]和 Francesco Petruccione[4,2]

[1*]夸祖鲁-纳塔尔大学。[2]国家理论与计算科学研究所 (NITheCS)。[3]人工智能研究中心 (CAIR)。[4]斯泰伦博斯大学。

*通讯作者。电子邮箱：rowanpellow@gmail.com；撰稿作者：pillayw4@ukzn.ac.za；sinayskiy@ukzn.ac.za；petruccione@sun.ac.za；

致谢

我们谨向国家理论与计算科学研究所 (NITheCS) 和人工智能研究中心 (CAIR) 致谢，感谢其为本研究提供经费支持。

摘要

核方法是机器学习中的一类重要技术。要取得良好效果，优秀的特征映射对于将无法线性分离的输入数据映射到更高维的（特征）空间至关重要，从而使数据在特征空间中可线性分离。以往研究表明，对于给定数据集，可以使用 NSGA-II（一种遗传算法）自动设计量子特征映射，同时最小化量子线路规模并最大化分类准确率。然而，评估候选特征映射所达到的准确率代价高昂。本研究展示了在基于遗传算法的量子特征映射设计中，核目标对齐可用作准确率的替代指标。核目标对齐的评估速度快于准确率，且无需预留部分数据点用于评估。为进一步加快遗传适应度的评估，我们提出了一种近似核目标对齐的方法。为了提高核目标对齐

和均方根误差，我们进一步使用 COBYLA 对生成电路的最终可训练参数进行训练，以确定采用传统电路参数训练的混合方法能否轻松补充遗传结构优化方法。我们在 UCI机器学习库中的九个不同二元分类问题上，将八种新方法与原方法进行了比较。结果表明，核目标对齐及其近似方法生成的特征映射电路，其准确率与以往研究相当，但在训练数据上的间隔更大（超过 20%），并且通过电路参数训练可进一步扩大间隔。

关键词： 量子计算、机器学习、支持向量机、量子支持向量机、遗传算法、NSGA-II、核目标对齐

1 引言

量子计算机利用纠缠等量子性质，在应用于特定问题时，有望比经典算法更具速度优势。理论研究已表明，对于相同任务，某些算法的性能优于所有已知的经典算法，例如用于因数分解的Shor算法（Shor，1997年）和Grover搜索算法（Grover，1996）。量子算法设计颇具挑战，因为量子纠缠的特性不符合直觉，必须有效利用它，才能获得相对于经典算法的优势。量子机器学习旨在将量子计算应用于机器学习任务，从而实现相对于经典机器学习的量子优势。

量子机器学习和经典机器学习有望实现许多实际任务的自动化；否则，这些任务将需要人类智能才能完成，包括疾病诊断（Myszczynskaet al，2020）、自然语言处理（Khurana etal，2022年）和图像分类（Horak and Sab-latnig，2019年）。机器学习算法旨在从数据中学习函数（Goodfellow etal，2016）。这些算法可分为三类：监督学习、无监督学习和强化学习。监督学习算法利用预先标注的训练数据，而无监督学习算法则不使用此类数据（Alloghani et al，2020）。在强化学习中，智能体通过与环境交互来学习行为，环境提供奖励和惩罚以引导其学习。

核方法是经典机器学习和量子机器学习中的一类重要技术。支持向量机（SVM）（Boser 等人，1996）是一种重要的经典监督式核方法，这得益于它与其他学习模型的理论联系（Jacot 等人，2018；Schuld，2021），以及关于其泛化能力的研究结果（Vapnik，1998；Vapnik 和 Chervonenkis，2015）。它的核心是一个优化算法，用于寻找能够将数据点划分为两个类别的最优线性超平面。该超平面经选择后，可使训练数据集中任意点的最小间隔达到最大，

其中，点的间隔定义为该点到分离超平面的距离。从理论上看，较大的间隔大小与更好的泛化性能相关（Vapnik，1998；Vapnik 和 Chervonenkis，2015）。将非线性数据映射到更高维的特征空间，即可实现非线性决策边界。所用的映射函数称为特征映射，而特征映射的值域称为特征空间。借助一种称为“核技巧”的技术，可以将决策边界优化问题重新表述为核函数的形式，该核函数用于计算特征空间中一对数据点的相似度（Boser 等人，1996）。只要能够计算相应的核函数，就无需显式计算数据点的特征映射输出。

必须谨慎选择特征映射，才能有效地区分数据。对于任意核函数与带标签训练集的组合，都可以计算一个称为核函数的核目标对齐度的量。它表示核函数与假想预言机核之间的一致程度；该预言机核由训练标签诱导生成，且适用于训练数据（Cristianini 等人，2001）。其他研究表明，较高的核目标对齐度与分类性能的提升相关（Cristianini 等人，2001），并且有人提出将其用作分类问题中为数据集选择合适核函数的度量（Cristianini 等人，2001；Hubregtsen 等人，2022年）。

量子支持向量机算法通过将特征映射函数实现为量子电路来增强支持向量机算法（见第 2.1.2 节）。量子特征映射电路以单个数据点的特征值为参数，但也可以包含额外的可训练参数，这些参数可通过优化来提升核电路对特定数据集的适用性（Hubregtsen 等人，2022年）。包含可训练参数的量子电路称为拟设。以往研究使用经典优化器优化带参数的量子核，以最大化其核目标对齐度，从而提升最终支持向量机模型的分类准确率（Hubregtsen 等人，2022）。

Hubregtsen等人报告了固定结构的量子特征映射电路，这些电路使用可训练参数值Hubregtsen等人(2022年）。他们使用随机梯度上升优化可训练参数值，以最大化相应核函数的核目标对齐。此举旨在测试核目标对齐优化能否提升固定结构量子特征映射在给定数据集上的表现。此前，Cristianini 等人(2001) 已证明，核目标对齐度提高与分类能力增强相关。所述技术既可用于针对某一数据集调整现有特征映射，也可根据具有预定电路结构和参数布局的特征映射拟设，为该数据集完整生成特征映射。这项工作同时采用了量子计算机的经典无噪声模拟和真实

NISQ计算机来运行量子电路，并报告称，在最大化核目标对齐后，分类准确率有所提高（Hubregtsen等人，2022年）。

量子核分类器可采用的第二种模型训练指标是分类器的均方根误差（RMSE）。为分类任务训练参数化模型时，通常会对该指标进行优化。与直接评估准确率相比，它可能更适合训练，因为它考虑的是误分类错误的大小，而非仅仅统计错误的数量。如果数据集不够大，准确率对电路参数变化的敏感度可能过低，无法显示量子线路是否略有改进。

优化数据集所用核电路选择的另一种方法，是在优化可训练电路参数值的同时，优化特征映射中所用电路门的选择。这种方法已用于优化应用于其他问题的量子线路（Ostaszewski 等人，2021年）。遗传启发算法也已用于量子线路结构优化，因为它们能够进行组合优化（Lukac 和 Perkowski，2002；Bautu 和 Bautu，2007；Rasconi 和 Oddi，2019年）。

Altares-L´opez 等人（2021年）详细介绍了用于自动化特征映射电路设计的遗传算法实现，该算法供量子支持向量机分类器使用，同时最大化分类准确率并最小化电路规模。该优化采用了一种名为 NSGA-II（Deb 等人，2002）的遗传算法变体。该算法定制了遗传算法中常规的遗传选择和适应度评估操作（如第 2.3节所述），以评估多个适应度函数，并优先选择非支配解进行交叉。

在具有两个适应度函数的最小化问题中，如果某个解 s 的适应度值为（a，b），则相对于另一组具有适应度值 $\{$（f_i，g_i）的 n 个解$|i \in \{1,2,...,n\}\}$，当且仅当 $\forall i \in \{1,2,...,n\}$，（$f_i < a$）$\implies$（$g_i > b$）时，该解被视为非支配解；也就是说，当且仅当该集合中不存在所有适应度值都优于 s对应适应度值的解。NSGA-II 还采用精英保留策略，以确保保留至少在某一个单独适应度函数上实现最佳优化的解。

为了将遗传算法应用于量子特征映射设计，本文还定义了一种用于编码特征映射电路的二进制串表示。该编码策略见 Altares-L´opez 等人 (2021年)，并在此作了总结。

每个电路门由一个 5 比特序列编码，其中前三比特用于编码所应用的量子门类型，后两比特用于编码比例参数；当前三比特选定参数化旋转门时，该比例参数将用于该门。比特与量子门及比例参数之间的映射如表 1所示。

超参数 M 和 N 分别表示量子比特的最大数量和门层的最大数量，必须在应用遗传算法之前选定，并在整个遗传优化过程中保持不变。单个解用长度为 $5MN$的比特串表示，

Bits	Gate
000	H
001	CNOT
010	I
011	Rx
100	Rz
101	I
110	I
111	Ry

Bits	参数
00	π
01	$\pi/2$
10	$\pi/4$
11	$\pi/8$

表1：对每个连续的5比特序列进行映射。该序列中的比特用于编码特征映射量子门，以确定量子门类型以及在使用参数化门时所采用的比例参数值。编码可选用的量子门包括哈达玛门（H）、受控非门、恒等门（I），以及绕布洛赫球X、Y和Z轴的参数化旋转门，这些门用于将数据点特征值编码到量子线路中。我们将绕x轴旋转门 Rx(θ) 定义为 cos($\theta/2$)I$-i$sin($\theta/2$)X，将绕y轴旋转门 Ry(θ) 定义为 cos($\theta/2$)I$-i$sin($\theta/2$)Y，将绕z轴旋转门 Rz(θ) 定义为 cos($\theta/2$)I$-i$sin($\theta/2$)Z。如果选择了绕 a 轴的参数化旋转门 R_a，则将使用五个量子门比特中最后两位所选的参数来选择比例参数 p 。将量子门 R_a应用于编码特征值 x_i时，所应用的量子门为 $R_a(px_i)$。如果选择受控非门并将其应用于量子比特i，则量子比特 i 将作为控制量子比特，量子比特 $(i+1) \bmod M$ 将作为目标量子比特。

其中包含特征映射中各个量子门编码的串接结果。

编码中的量子门依次应用，其中目标量子比特和可能被编码的特征值以轮询方式选取。具体而言，对于每个连续的 5 个量子门比特组，第 i 个量子门描述将应用于量子比特 $i \bmod M$ ；如果该量子门执行参数化旋转，则会编码特征 $i \bmod N$ 。选择这种编码策略是为了简便，不过也可以尝试其他可行的策略。

本研究对两个适应度函数进行了优化：最大化测试集上的准确率，同时最小化加权尺寸指标。非加权规模指标 SM 根据量子比特数量M、单量子比特门数量 N_{local}和纠缠门数量N_{CNOT}计算，表达式为

$$\text{SM} = \frac{N_{\text{local}} + 2N_{\text{CNOT}}}{M}.$$

提供给遗传算法的加权尺寸指标 WS 定义为以下表达式

$$\text{WS} = \text{SM} + \text{SM} \cdot \text{accuracy}^2.$$

这项工作展示了使用 NSGA-II，结合所设计的特征映射二进制字符串编码策略、准确率最大化和加权规模最小化，能够自动生成量子

特征映射电路，用于量子支持向量机分类，其输入仅需一个数据集和少量超参数。实验还表明，生成的电路能够泛化到未见过的数据。除了使用较少的量子比特和量子门外，研究还发现，该方法生成的电路几乎不使用纠缠，甚至完全不使用纠缠；这意味着它们可以通过经典方式高效模拟，而且该方法有望成为一种量子启发式经典机器学习算法。

已有一些研究尝试改进遗传优化，并将其与其他方法进行比较。Chen 和 Chern (2022 年的工作基于 Altares-L´opez 等人 (2021) 提出的算法。他们采用了修改后的编码方案，用三个比特而不是两个比特对比例参数值进行编码，使可编码的参数值数量增加了一倍。参数值选择受限是Altares-L´opez 等人 (2021)设计的算法可能存在的局限之一。此外，该算法还经过进一步修改，将量子门成本和分类准确率纳入同一个目标表达式进行优化，采用单目标遗传算法，而不是 NSGA-II 之类的多目标遗传算法。值得注意的是，这引入了一个新的超参数，用于权衡优化对电路规模和准确率的侧重；而 NSGA-II 则无需此参数。

将遗传算法生成的特征映射与另外两种拟设进行了比较：一种是 Kandala 等人(2017)提出的硬件高效拟设，另一种是Shende 等人(2006年)提出的酉分解拟设。每种拟设均使用 COBYLA（Powell，1994年）进行训练，并直接将分类准确率作为代价函数进行评估。研究发现，该工作所用遗传算法变体生成的特征映射电路，其表现与硬件高效拟设相近，具体取决于生成硬件高效拟设时所选的电路深度超参数。然而，在准确率方面，两者均不及酉分解拟设；后者取得了最高准确率，但代价是规模庞大且固定。

我们的研究探讨核目标对齐能否作为一种度量，用于自动化量子增强支持向量机（QSVM）算法的量子特征映射设计（Schuld 和 Killoran，2019年；Rebentrost 等人，2013年）。我们的研究有两个目标：一是探究在 Altares-L´opez 等人（2021年）所述的遗传优化过程中，使用准确率以外的代价函数是否合适；二是探究能否通过遗传训练与电路参数训练相结合的混合流程，解决遗传算法中量子线路参数选择有限的问题。

我们通过评估中准确率指标的两种替代代价函数来实现第一个目标：Altares-L´opez 等人(2021年提出的遗传优化步骤采用核目标对齐；其次，采用启发式方法估计核目标对齐，仅执行一部分核评估。为实现第二个目标，我们评估了一种混合方法：对于上述每种方法，在遗传算法结束后，进一步优化最终选定的可训练电路参数值。此最终优化

使用 COBYLA（Powell，1994年）来最大化核目标对齐或均方根误差。我们在若干难度各异的二元分类问题上，将新方法与原方法进行了比较。结果表明，尽管核-目标对齐度指标在量子核评估方面的计算成本较低，且无需训练支持向量机分类器，但所构建分类器的性能与原方法相当，并且在训练数据上的间隔分布通常更优。理论研究已表明，间隔大小增加意味着泛化能力更强（Vapnik，1998年；Vapnik 和 Chervonenkis，2015年）。结果显示，核目标对齐近似启发式方法的表现略逊于精确的核目标对齐优化，但计算成本仅为其一小部分。混合方法能够提高间隔大小。此外，结果还表明，原方法有时会对用于评估准确率指标的测试数据过拟合，尤其是在较难的问题上。

在接下来的章节中，我们将更详细地介绍理解本研究及相关研究所需的背景知识，并说明我们的实验设置。随后一节将介绍我们的研究发现及其解读。最后一节将概述我们所作的贡献，并提出进一步研究的方向。

2 方法

2.1 使用量子核的二元分类器

2.1.1 支持向量机（SVM）

支持向量机算法是一种用于二元分类问题的经典监督式机器学习算法，其通过在两类数据点之间寻找最优分离超平面来实现分类。当数据点可以表示为实值特征向量时，支持向量机算法便适用。为简化定义，类别标签通常替换为正一和负一。

单个数据点的间隔定义为该数据点到 SVM 所选超平面的距离。SVM 分类器的间隔是指各数据点间隔中的最小值。具有最小间隔的数据点称为支持向量。SVM 通过求解二次规划优化问题来选择超平面，使其最大化训练集数据点间隔中的最小值，因此该超平面是最优的。

支持向量机算法也能够对类别不可线性分离的数据集进行分类。为此，可以先将数据点映射到高维空间，使其在高维空间中变得线性可分（示意见图1 ）。用于执行此映射的函数称为特征映射，该函数的值域称为特征空间。特征映射的选择必须

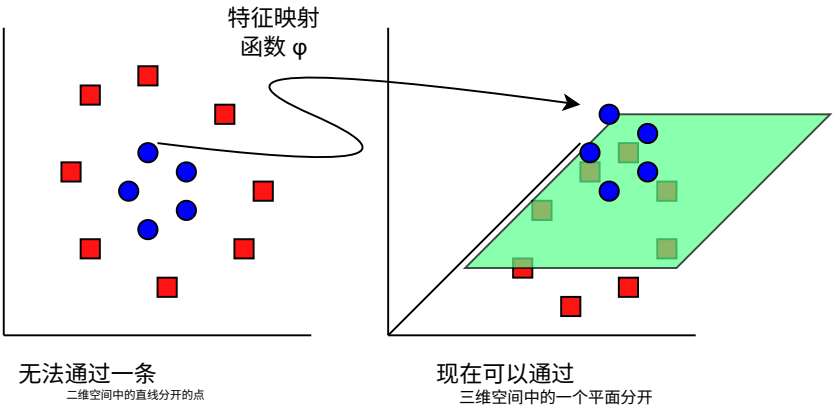


图 1：示例说明如何使用特征映射函数，将非线性可分的数据点映射为高维空间中的线性可分数据点。在本例中，特征映射可以通过一个函数实现：为数据点增加第三个维度，并且该维度的值会随着数据点到中心区域的距离增加而减小。

适合数据集，才能很好地对其进行分类，因为它决定了数据经过变换后是否会变得线性可分。

将数据点映射到特征空间的特征映射 $\phi(x)$具有相应的核函数 $\kappa(x_i, x_j) = \langle \phi(x_i)\ \phi(x_j) \rangle$，该核函数计算特征空间中一对数据点的内积。间隔优化问题可以使用核函数等价地重新表述为对偶问题（Boser 等人，1996），从而有时可以避免显式计算特征映射。这一优势通常被称为“核技巧”。鉴于在许多情况下，特征映射无法高效计算，但相应的核函数却可以高效计算，因此该问题的对偶形式增加了可应用于数据集的潜在特征映射数量。

在支持向量机的对偶形式中，给定类别的分类器输出由系数序列 $\alpha = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ 和偏移量序列 $b = \{b_1, b_2, ..., b_n\}$ 决定，这些序列由支持向量机算法在超平面优化过程中选定。决策函数 df 输出输入点映射到特征空间后与超平面之间距离的指示值。它由核函数 κ、训练样本 $\{x_1, x_2, ..., x_n\}$、α 系数和 b 偏移量定义如下

$$\mathrm{df}(x) = \sum_{i=1}^{n} (\alpha_i \kappa(x, x_i) + b_i).$$

df(x) 的符号用于确定参数点 x 的预测类别：

$$\text{Class}(x) = \text{sgn}(\text{df}(x)).$$

2.1.2 量子增强支持向量机

量子增强支持向量机（QSVM）算法通过在量子计算机（Schuld 和 Killoran，2019年；Rebentrost 等，2013年）上执行核计算来扩展支持向量机算法。将单个数据点的值作为参数的量子电路被用作特征映射，将数据点映射到量子希尔伯特空间中的高维量子态。

对于将数据点编码到 q 个量子比特中的量子特征映射，特征希尔伯特空间的维数为 2^q。尽管量子计算机能够高效地计算数据点的量子态特征空间表示，但通常情况下，由于特征空间的维数呈指数增长，量子态无法在经典计算机上高效表示。为克服这一经典局限性，可直接在量子计算机上计算核函数，并使用支持向量机的基于核的公式。一对数据点的核计算可以通过测量其在量子特征空间中对应量子态的重叠来高效完成。

要训练量子支持向量机模型，必须计算训练点的格拉姆矩阵 $K_{n\times n}$。对于 n 个训练点 $\{x_1, x_2, ..., x_n\}$ 和量子核函数 κ，格拉姆矩阵定义为 $K_{ij} = \kappa(x_i, x_j)$，其中$i, j \in \{1, 2, ..., n\}$。如果使用无噪声量子计算机或模拟器来执行 κ，则可以利用对称性 $\kappa(x_i, x_j) = \kappa(x_j, x_i)$ 以及性质 $\kappa(x_i, x_i) = 1$，以减少所需的核评估次数。

假设上述性质成立，则 K (K_{11},K_{22}, ..., K_{nn}) 的 n 主对角线元素无需进行核评估，因为 $K_{ii} = \kappa(x_i, x_i) = 1$。对于其余 $n^2 - n$ 个非主对角线元素 K_{ij}，每个元素都有一个值相同的对称元素 K_{ji}，因为 $K_{ij} = \kappa(x_i, x_j) = \kappa(x_j, x_i) = K_{ji}$。这意味着只需通过核评估显式计算一半的元素。实际上，构建 K 只需进行 $\frac{n^2-n}{2}$ 次核评估。对于含噪声中等规模量子计算机，如果采取措施缓解噪声并校正核矩阵（如 Hubregtsen 等人（2022年）所述），则有可能应用这一技术。

支持向量机算法与量子支持向量机算法之间唯一的区别在于核计算的执行方式，因此量子支持向量机算法的潜在优势在于能够计算经典方法难以估计的核函数（Liu 等人，2021年）。虽然已经为人工数据集发现了此类核函数的实例（Liu 等人，2021年），但如何设计量子特征映射，以获得兼具实用性和量子速度优势的核函数，仍是一个未解的问题。

2.2 核质量指标

2.2.1 核目标对齐

核目标对齐是一种评估核函数质量的启发式方法，用于衡量两个核函数之间的相似程度，或核函数与数据集之间的一致程度（Cristianini 等人，2001）。其计算方式是对由训练样本构造的格拉姆矩阵与由训练标签构造的预言机矩阵求矩阵内积。预言机矩阵充当假想核函数的格拉姆矩阵的替代物，而该假想核函数与数据高度契合。

对于一组训练点 $\{x_1, x_2, ..., x_n\}$ 及其对应的标签 $\{y_1, y_2,..., y_n\}$ 其中 $\forall i$, $y_i \in \{-1, 1\}$，核函数 $\kappa(x_i, x_j)$，以及矩阵的 $Frobenius$内积定义为 $\langle A, B\rangle_F = \sum_{i,j} A_{ij}B_{ij}$，核目标对齐可按如下方式计算（Cristianini等人，2001）

1. 使用核函数和训练点，按照规则计算格拉姆矩阵 $K_{n\times n}$

$$K_{ij} = \kappa(x_i, x_j).$$

2. 使用训练标签，按照规则计算预言机矩阵 $O_{n\times n}$

$$O_{ij} = y_i y_j.$$

3. 使用弗罗贝尼乌斯内积计算核目标对齐 核目标对齐，公式如下：

$$\mathrm{KTA} = \frac{\langle K, O\rangle_F}{\sqrt{\langle K, K\rangle_F \langle O, O\rangle_F}}.$$

其他研究表明，较高的核目标对齐与更好的分类性能相关（Cristianini 等人，2001年），并且有人提出将其用作分类问题中为数据集选择适用核函数的指标（Cristianini 等人，2001年；Hubregtsen 等人，2022年）。

2.2.2 均方根误差

均方根误差（RMSE）是衡量模型误差的常用指标。其计算方法是：对分类器在训练集中的每个数据点上的预测误差求平方，取这些平方误差的平均值，再开平方根。在本研究中，分类器的均方根误差根据决策函数在训练数据上的误差计算，并对误差计算方式进行了调整。这一调整旨在考虑到，对于给定的样本和标签对，SVM 决策函数并不存在一个明确正确的输出。误差是相对于正的目标决策函数输出 m 进行衡量的，在本研究中我们将其设为一。

我们使用以下规则计算决策函数输出 a 与训练标签b 之间的误差:

$$\text{error}(a,b) = \begin{cases} (m-a) & \text{if } b = 1 \text{ and } a < m \\ (a-m) & \text{if } b = -1 \text{ and } a > -m \\ 0 & \text{otherwise} \end{cases}$$

这种误差函数的选择意味着，只有未达到所需置信度 m 要求而被分类的点才会计入误差计算；所考虑点的误差会随着其与目标输出之间距离的增加而增大。

对于一组训练点 $\{x_1, x_2, ..., x_n\}$ 及其对应的标签 $\{y_1, y_2,.. y_n\}$，其中 $\forall i$, $y_i \in \{-1, 1\}$，均方根误差根据这一调整后的误差函数和决策函数 df 按以下规则计算

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \text{error}(\text{df}(x_i), y_i)^2}{n}}$$

2.3 遗传算法概述

遗传算法是一类灵活的元启发式算法，其灵感源于现实世界中的进化原理，包括自然选择、遗传和随机突变。在已知能够产生全局最优解的算法不存在或不可行时，遗传算法是优化复杂目标函数的常用选择。

实现遗传算法，首先需要设计一种可执行遗传操作的解表示。解表示通常（但并非总是）采用二进制串的形式，并通过特定问题的解码函数映射为可用的解。遗传算法管理一组这样的解表示，称为种群。初始种群可以由随机生成的解组成，也可以根据特定问题的启发式方法选取。

优化过程是迭代式的，通常会持续进行，直到找到合适的解、达到预定的迭代次数，或解的改进速度变慢。每一代都会对种群应用受遗传机制启发的操作，以创建新的替代种群。种群的创建过程包括适应度评估、选择、交叉和变异操作。

适应度评估通过将解表示解码为一个解，然后评估其对问题的适用性数值评分来完成。对整个种群进行评估后，会应用选择操作，选出一些解进行交叉和变异。让较优的解更有可能被选中进行交叉非常重要，因为这是推动不同代之间不断改进的主要机制。被选中进行交叉的解称为“父代”。

交叉操作在两个父代解之间进行，以生成一个或多个称为子代解的新解。通常

通过简单组合父代的解表示来完成。对于二进制字符串解表示，可以通过组合父代中两个互不重叠的子序列来进行简单交叉：从第一个父代中随机选取若干比特，其余部分则取自第二个父代对应位置上的比特。可以对一个子代解应用变异操作，随机对其表示进行少量修改。这模拟了现实生活中发生的随机突变，并会影响种群中可用遗传物质的多样性。

在确定哪些个体组成下一代时，通常采用的一种策略是保留当前代和新生成子代中表现最佳的解。这被称为精英保留策略，可确保解能够在多代中持续存在，甚至无限期保留，只要它们继续优于更新的解。这有助于防止随着代数增加，已获得的适应度因偶然而出现退化。

遗传算法的总体思路具有足够的灵活性，因此，针对上述组件的许多变体和扩展也得到了研究（Chahar 等人，2021年）。

2.4 实验

文献Altares-L´opez 等人（2021年）中所述的自动化特征映射设计算法，使用 Julia 编程语言（Bezan- son 等人，2017）、Yao 量子模拟器框架（Luo 等人，2020）以及 NSGA-II 的 pymoo（Blank 和 Deb，2020）实现进行了重新实现。所有实验均将最大量子比特数和特征映射深度超参数设置为 6。遗传算法的种群规模设为 100，每代生成 15 个新个体。新个体中有 30% 通过交叉产生，其余个体则从父代中随机选取。每一代中，种群的 70% 会进行变异。发生变异时，变异解中 20% 的比特会发生翻转。所有实验均使用 Yao 提供的无噪声量子模拟器运行。

在九个难度各异的数据集上运行了原始算法的三种配置（见表 2），以比较它们的效果。第一种配置最大化测试集上的准确率，并最小化加权规模，与原始研究相同（Altares-L´opez 等人，2021年）。第二种配置最大化训练数据上的核-目标对齐，而不考虑测试数据；同时最小化 Altares-L´opez 等人（2021年）中定义的非加权规模指标。第三种配置最大化训练数据上核-目标对齐的近似值，并最小化相同的非加权规模指标。

在进行遗传优化之前，将数据集划分为三个互不相交的子集，即训练数据、测试数据和验证数据。训练数据用于评估核目标对齐及其近似值，并针对给定的特征映射电路训练量子支持向量机模型。

数据集	类别 -1	类别 1	特征（主成分分析）	训练集	Test	验证集
月牙形数据集	左上	右下	2（不适用）	210	90	500
癌症	良性	恶性	30（10）	210	90	124
Iris	变色鸢尾	维吉尼亚鸢尾	4（不适用）	42	18	40
Digits	八	Nine	64（10）	140	60	148
Circles	外部	内部	2（不适用）	210	90	500
Random	Red	Blue	2（不适用）	210	90	N/A
Voice	可接受	不可接受	309 (10)	28	12	44
SUSY	背景	信号	18 (10)	210	90	500
SUSY 简化	背景	信号	8（不适用）	210	90	500

表2：展示数据集特征及所用样本划分的表格。基础数据集中的所有点并未全部使用，以确保每个样本集中的样本划分保持平衡。在确定数据划分时，还需考虑实验运行时间，同时维持足够大的测试点与训练点之比，以及足够多的验证点。所有特征值超过10个的数据集均使用主成分分析（PCA）降至10个特征。月牙形数据集、Circles数据集和Random数据集均为人工生成，其中月牙形数据集和Circles数据集由 Scikit-learn 生成（Pedregosa 等人，2011）。其余数据集均来源于 UCI机器学习库（Dua 和 Graff，2017），或直接获取，或通过 Scikit-learn 间接获取（Pedregosa 等人，2011）。

测试数据仅用于在第一种方法中评估准确率指标。验证数据用于确定生成模型的泛化能力，并且必须与测试数据分开，因为第一种方法可以通过准确率指标间接访问测试数据，从而可能对其过拟合。

为了计算 n 个训练点的核-目标对齐近似，将 n 个点划分为 a 个互不相交的互补子集，每个子集的大小约为 n/a。子集数量 a 可根据训练点的数量进行调整，以平衡速度和精度。依次在每个子集上计算核目标对齐，然后取其平均值。

假设不利用核函数的性质来加速格拉姆矩阵的计算，则计算精确的核目标对齐需要进行 n^2 次核评估，而计算近似值只需进行 $a\,(n/a)^2 = (n^2)/a$ 次评估，速度可提升 a 倍。如果利用核函数的性质，则计算精确的核目标对齐需要进行 $(n^2 - n)/2 = n^2/2 - n/2$ 次核评估。因此，评估核-目标对齐近似所需的核评估次数可推导如下：

$$a\left(\frac{(\frac{n}{a})^2 - \frac{n}{a}}{2}\right) = \frac{n^2}{2a} - \frac{n}{2},$$

这意味着，通过计算近似所获得的加速因子大于 a，但随着 n 增大，应趋近于 a。在本研究中，所有涉及核-目标对齐近似的实验均采用 $a = 5$。

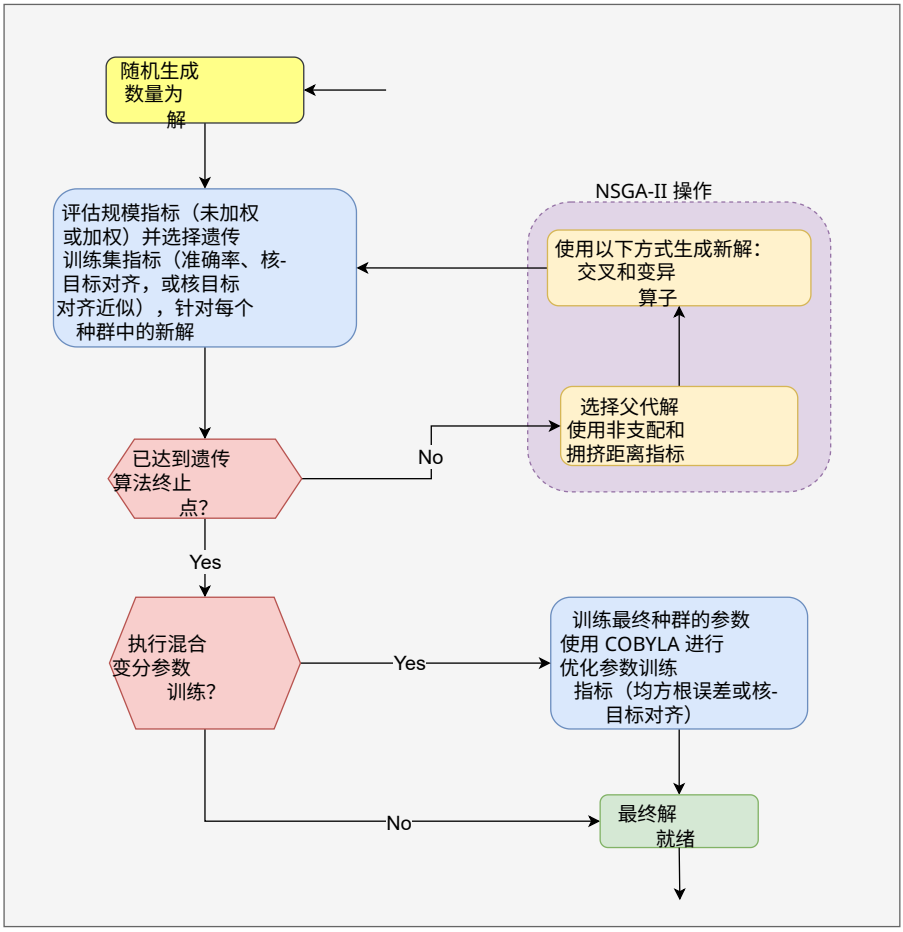


图 2：概述通过遗传方式训练量子特征映射电路所采用算法的流程图。该图还展示了如何在遗传优化后，采用涉及电路参数训练的混合方法进行训练。

每种配置下的遗传优化完成后，我们会尝试进一步优化门表示中最后两位所编码的比例参数，以进一步改进结果；为此，我们使用 NLopt 优化库（Powell，1994年）提供的 COBYLA 实现（Johnson，2011）。这样，参数值就不再局限于仅有的四种可能性之一。此优化以训练集计算各项指标，目标是最小化均方根误差或最大化核目标对齐。COBYLA 优化器最多可评估代价函数一百次以完成优化。图 2展示了算法流程图。

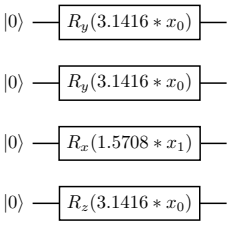
用于均方根误差（RMSE）和核目标对齐的 COBYLA 代价函数在每次评估时都需要计算一个格拉姆矩阵，这意味着所执行的核评估次数为 $^{100}\,_{(n^2-n)}{}_2$。相比之下，在对准确率或核目标对齐进行遗传优化时，仅评估遗传算法中首代 100 个解的适应度，所执行的核评估次数就至少与此相同。在随后的 1199 代中，$15 \times 1199 = 17985$ 会执行更多格拉姆矩阵评估，总数达到 18085 次；这意味着，在对准确率或精确核目标对齐进行遗传优化的情况下，对整个输出种群进行最终参数训练所需的核评估次数，约为遗传优化过程中所执行次数的 55%。

我们分别将三种基础方法命名为方法 1、方法 2 和方法 3。每种方法还有两个额外的子方法，分别用于进一步训练均方根误差或核目标对齐，共计九种方法。均方根误差和核目标对齐的变体分别以 .1 和 .2 为后缀命名。我们绘制每种方法所生成的最佳模型的分类准确率、平均间隔、ROC曲线、特征映射电路和混淆矩阵，其中，种群中的最佳模型被定义为验证集准确率最高的模型。对于二维数据集，我们还会绘制决策边界。

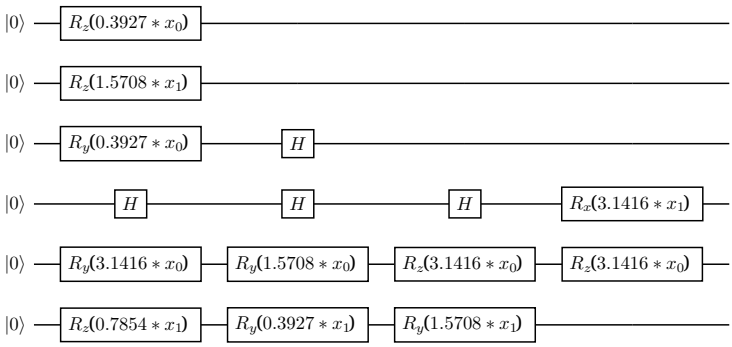
实现实验和绘制结果图表的代码可在 GitHub 上找到（Pellow-Jarman，2022年）。

3 结果

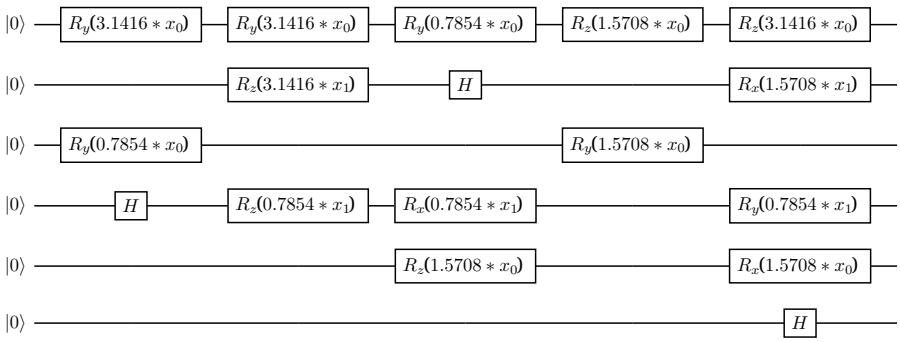
与 Altares-L´opez 等人 (2021年) 的原始研究一致，各种方法生成的特征映射电路几乎不使用纠缠门（见图 3）。然而，基于核目标对齐指标进行优化所生成的电路，整体规模往往显著更大（见图 3）。这可能是因为遗传优化中的加权规模指标被替换成了非加权规模指标，因为加权规模指标取决于测试集准确率，而该准确率并未进行评估。方法 2 和 3 在分配的 1200 代内未能实现电路规模优化的收敛，这一点可从存在冗余门看出。或许可以通过增加迭代代数，或像原始方法那样使用与核目标对齐而非准确率加权的规模指标来解决这一问题。电路规模较大的另一种可能解释是，即使电路已经达到完美准确率，其核目标对齐仍可能继续提升；在准确率最大化的情况下，遗传算法在达到 100% 准确率后可以转而专注于缩小电路规模，但在最大化核目标对齐时则难以做到这一点，因为其极限值 1 更难达到。此外，可以降低 70% 的高变异率，尝试在分配的 1200 代内实现收敛。



（a）方法 1 - 准确率、加权规模



(b) 方法2 - 核目标对齐，未加权大小



(c) 方法 3 - 核-目标对齐近似，未加权大小

图 3：使用三种基础遗传方法为 Moons数据集创建量子特征映射时，生成的验证集准确率最高的量子线路。(a) 显示的是按照原始研究中的方法，在训练时最大化准确率并最小化加权规模所生成的最佳量子线路；(b) 显示的是在训练时最大化精确核目标对齐并最小化未加权大小所生成的最佳量子线路；(c) 显示的是在训练时最大化核目标对齐近似并最小化未加权大小所生成的最佳量子线路。(b) 和 (c) 中的量子线路明显更大。图中省略了未使用的门层和量子比特。

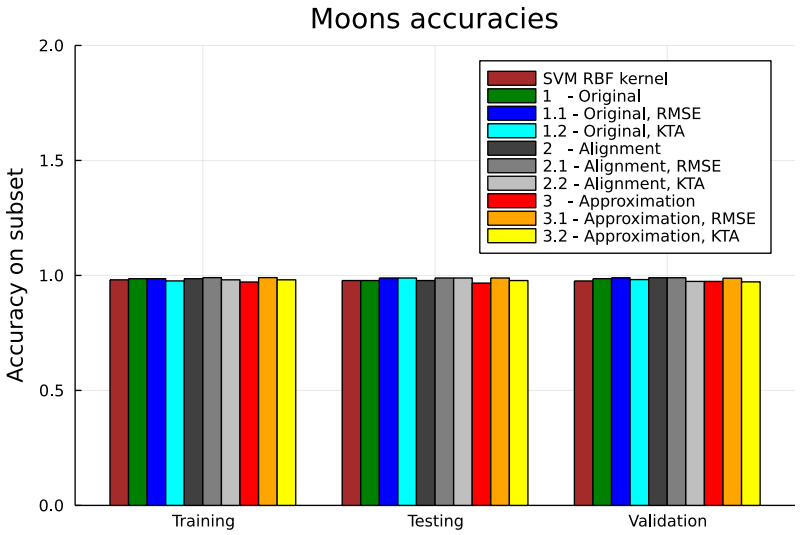


图 4：该图展示了采用量子特征映射设计的各种方法所生成的最佳模型在月牙形数据集上的分类准确率，并以经典 RBF 核作为参考进行比较。可以看出，所有方法在不同子集上均达到了相当的准确率。

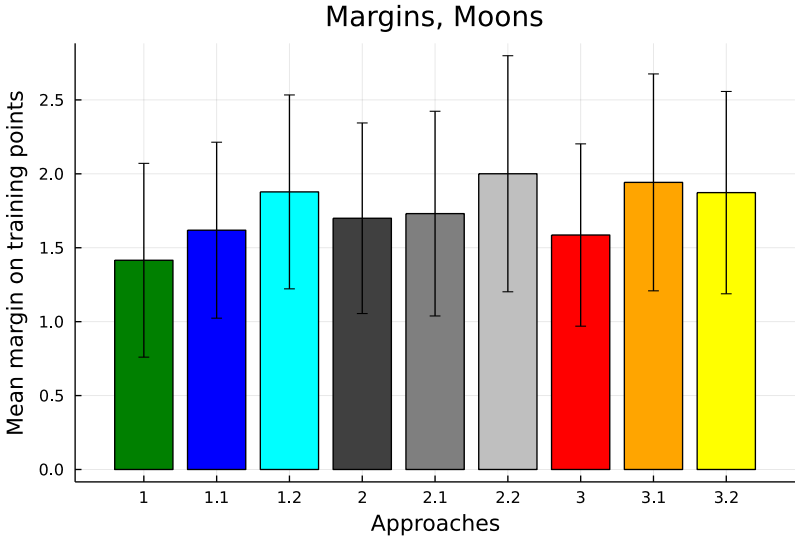


图 5：该图展示了各方法生成的最佳分类器对 Moons 训练集样本的平均间隔，误差条表示标准差。电路参数训练和核目标对齐的遗传训练均显示出能够增大平均间隔大小。方法编号与图 4 中使用的编号相对应。

.

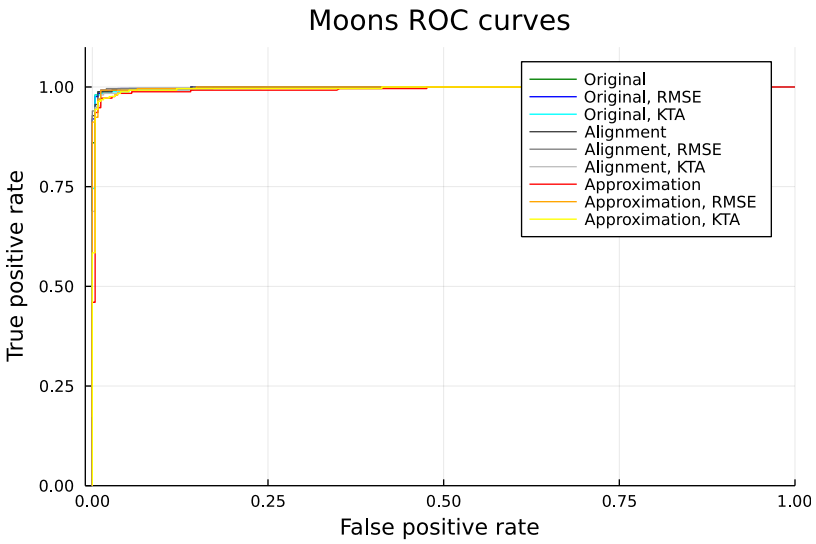


图 6：图中展示了在 Moons数据集上，采用不同量子特征映射设计方法得到的最佳模型的 ROC曲线。所有生成的模型在该数据集上的表现相近。

实验表明，在遗传优化过程中，以核目标对齐或近似核目标对齐替代准确率，在所有数据集上都能得到准确率与原方法相当的特征映射电路（例如，参见图 4 和6 ）。使用 COBYLA 进一步优化最终种群的可训练参数，往往能够增大分类器的平均间隔（参见图 5），并且有时能提高验证集上的分类准确率（参见图 4 和 9）。这表明，尽管需要付出计算成本，但如果提高准确率很重要，仍值得尝试采用混合方法，进一步优化最终种群的参数值。对单个解的参数进行训练，进行 100 次代价评估所需的核评估次数仅为遗传优化过程的一半百分比，因此，以低得多的成本训练一小部分最终解也是可行的。此外，如果进一步进行训练，解的二进制字符串中编码的未训练参数并不会丢失；如果它们的表现优于训练后的参数，仍可使用。

结果表明，在 SUSY、SUSY精简特征、Voice 和 Random 数据集等难度较高的数据集上，原始方法的模型可能会对用于评估准确率指标的测试数据过拟合（见图7）。这很可能是因为遗传算法直接优化测试集准确率，而未考虑训练集准确率。由于核目标对齐方法在遗传优化过程中仅使用训练数据，因此不会遇到同样的问题；不过，在这些难度较高的问题上，这些方法并未提升验证数据上的表现。

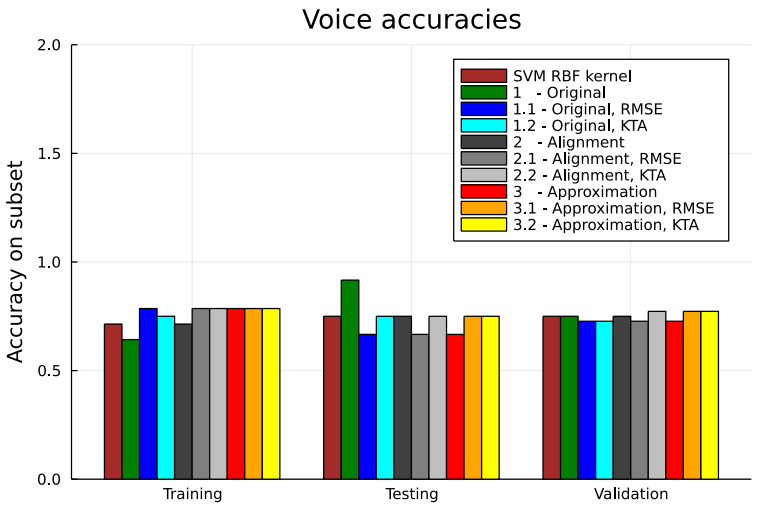


图7：该图展示了在 Voice 数据集上，采用不同量子特征映射设计方法得到的最佳模型的分类准确率，并以经典 RBF 核作为参照。结果表明，遗传算法最大化准确率会对用于评估准确率指标的测试数据过拟合，这证明了单独设置验证集的必要性。

一种可能的解决办法是每一代都打乱训练数据和测试数据，不过这样一来，准确率指标将取决于评估准确率时所处的代数，并且可能导致遗传算法无法缓存解适应度。另一种可能的解决办法是对数据子集上的准确率取平均值。对于包含 n 个点的数据集，最坏情况下最多需要进行 $\frac{n^2-n}{2}$ 次核评估即可完成此操作，因为可以一次性计算整个数据集的格拉姆矩阵，并将其用作缓存；在使用任意训练子集和测试子集创建模型时，可据此查找任意点对的核输出。

采用第二种和第三种方法训练的分类器，其间隔往往大于采用第一种方法训练的分类器（见图 5 和8，以及表 3）。这可能是因为，核-目标对齐度量及其近似值是在训练子集上评估的，而准确率是在测试子集上评估的，因此前两种方法在训练子集上的置信度更高。无论如何，根据相关理论研究（Vapnik，1998年；Vapnik 和 Chervonenkis，2015年），间隔大小的增大表明泛化能力有所提升；这些研究表明，间隔大小可以界定 VC 维，而 VC 维可以界定预期泛化误差。在最终一代上进一步进行参数训练，也往往能使间隔大小有所改善，即使是在月牙形数据集和 Circles数据集这类较简单的数据集上也如此，而这些数据集的总体

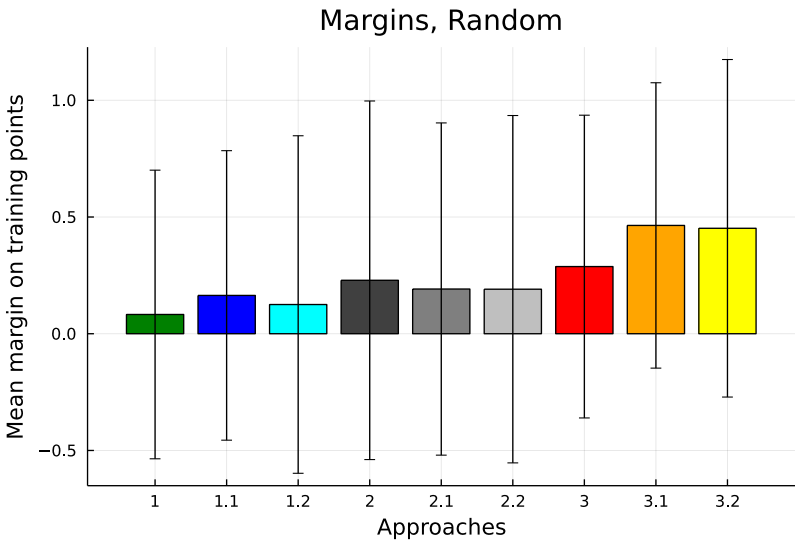


图 8：该图展示了各方法生成的最佳分类器在随机训练集样本点上的平均间隔，误差条表示标准差。方法编号与图4中的编号对应。

分类准确率并未受到太大影响。分类器的决策边界图直观地展现了间隔的这一改善（见图 9）。

4 结论

本文将我们对Altares-L´opez 等人 (2021年) 所提出方法的实现，与对遗传算法代价函数的调整进行了比较。这些调整旨在探究核目标对齐作为测试集准确率替代指标的适用性，并减少该方法所需的核评估次数。结果表明，尽管电路规模有所增加，新方法仍能以更少的核评估次数有效地设计出准确的分类器；而且，它们往往还能生成在训练数据上间隔更大的分类器。我们还提出了一种混合方法，在原有工作的基础上，于遗传算法终止后应用 COBYLA（Powell，1994年）进一步优化生成的量子特征映射电路中的可训练参数，以期在遗传算法基础成本之外仅增加较低的计算成本，进一步提升性能。结果还表明，这种参数训练能够增大间隔，有时也能提高准确率，同时不增加电路门成本。

在保持量子门成本较低的同时加速遗传算法，仍有更多工作需要完成。实现这一目标的一种潜在途径是采用多阶段遗传算法：其代价函数最初易于评估，但在经过一定数量的代后提高精度。

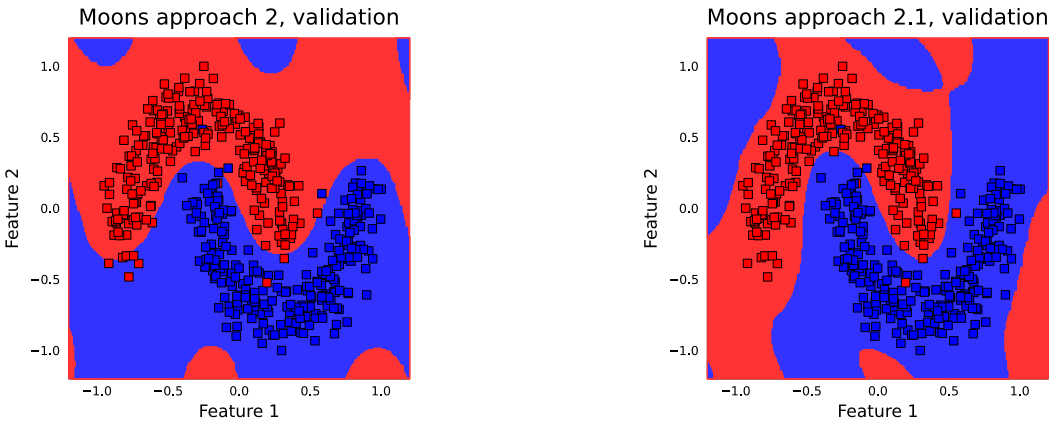


图 9：在 Moons 验证数据上，方法 2 和 2.1 的最佳分类器的决策边界。先通过遗传算法设计特征映射以最大化核目标对齐，再在训练数据上进一步进行参数训练以最小化均方根误差，结果表明这能提升模型在验证数据上的分类能力。

方法	平均间隔	绝对变化	百分比变化
1-准确率（原始工作）	0.838	N/A	N/A
准确率，均方根误差训练	0.993	+0.154	+18.42%
准确率，核目标对齐训练	0.971	+0.133	+15.87%
2-对齐	1.043	+0.205	+24.46%
对齐，RMSE 训练	1.016	-0.028	-2.66%
对齐，KTA 训练	1.090	+0.047	+4.49%
3-近似	1.065	0.226	+26.99%
近似，RMSE 训练	1.145	+0.080	+7.54%
近似，KTA 训练	1.124	+0.059	+5.59%

表3：该表展示了每种方法生成的最佳分类器的平均间隔大小，该数值是在九个数据集上取平均得到的，并对每个数据集赋予相同权重。在此目的下，最佳分类器定义为在目标数据集上取得最高验证集准确率的分类器。基础方法（2和3）的改进列展示了对核目标对齐及其近似进行遗传优化后，相较于原始工作中所达到的间隔有何改进。对于混合方法（包括额外的均方根误差和核目标对齐参数训练），各列展示的是相对于基础方法的变化。

例如，可以让核-目标对齐近似的 a 参数随着代数增加而降低，使近似结果更加准确，但代价是需要进行更多核评估；或者，在达到预先设定的核目标对齐后，将代价函数从核目标对齐切换为分类准确率，以降低量子门成本。

原始方法还可以进一步扩展，让参数化门的门编码选择一种经典数据编码函数，例如 Suzuki 等人(2020) 所使用的函数，从而在编码中引入经典非线性，有望生成量子门成本更低或精度更高的量子线路。

参考文献

Alloghani M, Al-Jumeily Obe D, Mustafina J 等（2020）《数据科学的监督式与无监督式机器学习算法系统综述》，第3–21页。
https://doi.org/10.1007/978-3-030-22475-2 1 -

Altares-L´opez S, RibeiroA,Garc´ıa-Ripoll JJ（2021）量子特征映射的自动设计。《Quantum Science and Technology》6(4):045,015。
https://doi.org/10.1088/2058-9565/ac1ab1，URL https://doi.org/10.1088%2F2058-9565%2Fac1ab1

Bautu A, Bautu E（2007）通过遗传编程进行量子电路设计

Bezanson J, Edelman A, Karpinski S 等（2017）Julia：一种数值计算的新方法。《SIAM Review》59(1):65–98。 https://doi.org/10.1137/141000671，URL https://epubs.siam.org/doi/10.1137/141000671

Blank J, Deb K（2020）pymoo：Python 中的多目标优化。《IEEE Access》8：89,497–89,509

Boser BE, Guyon IM, Vapnik VN（1996）最优间隔分类器的训练算法。载于：第 5 届 ACM 计算学习理论研讨会论文集，第 144–152 页

Chahar V, Katoch S, Chauhan S（2021）遗传算法综述：过去、现在与未来。《Multimedia Tools and Applications》80。
https://doi.org/10.1007/s11042-020-10139-6

Chen BS, Chern JL（2022）为支持向量机分类器生成量子特征映射。
https://doi.org/10.48550/ARXIV.2207.11449，网址https://arxiv.org/abs/2207.11449

Cristianini N, Shawe-Taylor J, Elisseeff A 等（2001）关于核目标对齐。载于：Dietterich T, Becker S, Ghahramani Z（编）《神经信息处理系统进展》，第 14 卷。麻省理工学院出版社，网址https://proceedings.neurips.cc/paper/2001/file/1f71e393b3809197ed66df836fe833e5-Paper.pdf

Deb K、Pratap A、Agarwal S 等（2002）一种快速且具精英保留策略的多目标遗传算法：NSGA-II。IEEE Evolutionary Transactions

《Computation》6(2):182–197. https://doi.org/10.1109/4235.996017

Dua D、Graff C（2017）UCI机器学习库。网址 http://archive.ics.uci.edu/ml

Goodfellow I、Bengio Y、Courville A（2016）深度学习。麻省理工学院出版社

Grover LK（1996）一种用于数据库搜索的快速量子力学算法。载于：第二十八届ACM计算理论研讨会论文集。计算机协会，纽约州纽约市，STOC'96，第212–219页，https://doi.org/10.1145/237814.237866，网址 https://doi.org/10.1145/237814.237866

Horak K、Sablatnig R（2019年）用于图像识别的深度学习概念与数据集：综述2019年。第100页，https://doi.org/10.1117/12.2539806

Hubregtsen T、Wierichs D、Gil-FusterE 等（2022年）在近期量子计算机上训练量子嵌入核函数。《Physical Review A》106：042,431。https://doi.org/10.1103/PhysRevA.106.042431，网址 https://link.aps.org/doi/10.1103/PhysRevA.106.042431

Jacot A, Gabriel F, Hongler C（2018）神经切线核：神经网络中的收敛与泛化。载于：第32届神经信息处理系统国际会议论文集。Curran Associates公司，美国纽约州雷德胡克，NIPS'18，第8580–8589页

Johnson SG（2011）NLopt非线性优化软件包。网址 http://ab-initio.mit.edu/nlopt

Kandala A, Mezzacapo A, Temme K 等（2017）适用于小分子和量子磁体的高效硬件变分量子本征求解器。《Nature》549(7671)：242–246。https://doi.org/10.1038/nature23879，网址https://doi.org/10.1038%2Fnature23879

Khurana D, Koli A, Khatter K 等（2022年）自然语言处理：最新进展、当前趋势与挑战。Multimedia Tools and Applicationshttps://doi.org/10.1007/s11042-022-13428-4

Liu Y、Arunachalam S、Temme K（2021年）监督式机器学习中严谨且稳健的量子加速。《Nature Physics》17(9)：1013–1017

Lukac M、Perkowski M（2002）使用遗传算法演化量子电路。载于：2002年NASA/美国国防部可进化硬件会议论文集，第177–185页，https://doi.org/10.1109/EH.2002.1029883

Luo XZ、Liu JG、Zhang P 等（2020）Yao.jl：用于量子算法设计的可扩展、高效框架。Quantum 4：341。https://doi.org/10.22331/

q-2020-10-11-341，网址https://doi.org/10.22331/q-2020-10-11-341

Myszczynska MA、Ojamies PN、Lacoste AMB 等（2020）机器学习在神经退行性疾病诊断和治疗中的应用。《Nature Reviews Neurology》16(8):440–456。https://doi.org/10.1038/s41582-020-0377-8，URLhttps://doi.org/10.1038/s41582-020-0377-8

Ostaszewski M、Grant E、Benedetti M（2021）参数化量子电路的结构优化。Quantum 5:391。 https://doi.org/10.22331/q-2021-01-28-391，URL https://doi.org/10.22331/q-2021-01-28-391

Pedregosa F、Varoquaux G、Gramfort A 等（2011）Scikit-learn：Python 中的机器学习。《Journal of Machine Learning Research》12:2825–2830

Pellow-Jarman R（2022年）用于量子特征映射设计的混合遗传优化。https://github.com/RowPJ/hybrid-genetic-optimisation-for-quantum-feature-map-design

Powell MJD（1994年）一种通过线性插值对目标函数和约束函数建模的直接搜索优化方法，Springer Netherlands，Dordrecht，第51–67页。https://doi.org/10.1007/978-94-015-8330-5_4，URL https://doi.org/10.1007/978-94-015-8330-5_4

Rasconi R、Oddi A（2019年）一种用于量子线路编译问题的创新遗传算法。载于：第三十三届AAAI人工智能会议、第三十一届创新人工智能应用会议和第九届AAAI人工智能教育进展研讨会论文集。AAAI Press，AAAI'19/IAAI'19/EAAI'19，https://doi.org/10.1609/aaai.v33i01.33017707，URL https://doi.org/10.1609/aaai.v33i01.33017707

Rebentrost P、Mohseni M、Lloyd S（2013年）用于大数据分类的量子支持向量机。《物理评论快报》113。https://doi.org/10.1103/PhysRevLett.113.130503

Schuld M（2021年）量子机器学习模型是核方法

Schuld M、Killoran N（2019年）量子机器学习特征希尔伯特空间。《物理评论快报》122:040,504。https://doi.org/10.1103/PhysRevLett.122.040504，网址 https://link.aps.org/doi/10.1103/PhysRevLett.122.040504

Shende V, Bullock S, Markov I（2006年）量子逻辑电路综合。《IEEE集成电路与系统计算机辅助设计汇刊》25(6):1000–1010。https://doi.org/10.1109/tcad.2005.855930，网址 https://doi.org/10.1109%2Ftcad.2005.855930

Shor PW（1997）量子计算机上用于质因数分解和离散对数的多项式时间算法。《SIAM计算杂志》26(5):1484–1509。https://doi.org/10.1137/s0097539795293172，网址 http://dx.doi.org/10.1137/S0097539795293172

Suzuki Y、Yano H、Gao Q 等（2020）基于核的量子分类器中特征映射的分析与合成。《量子机器智能》2(1)。https://doi.org/10.1007/s42484-020-00020-y，网址 http://dx.doi.org/10.1007/s42484-020-00020-y

Vapnik V、Chervonenkis A（2015）关于事件相对频率一致收敛至其概率，页码 11–30。https://doi.org/10.1007/978-3-319-21852-6_3

Vapnik VN（1998年）《统计学习理论》。Wiley-Interscience

附录A 图

对于图 A1 - A9，我们展示了每个数据集上效果最佳的核函数的准确率，并同时展示经典 RBF 核以供比较。这些图表表明，尽管新提出的方法在训练期间并未直接评估分类准确率，但其分类准确率与原方法相当。对于较难的问题（Random、Voice、SUSY、SUSY精简特征），图表显示，原方法倾向于对用于评估准确率指标的测试数据过拟合，因此必须留出单独的验证集，以估计其模型的泛化能力。在这些情况下，与其他方法相比，基于最大化核目标对齐近似的方法也更容易对训练数据过拟合。

图A10 - A18 展示了各数据集训练数据上表现最佳的相同核函数的间隔。基于最大化核目标对齐及其近似的方法往往具有更大的平均间隔大小，而针对均方根误差和核目标对齐进行最终参数训练，也往往会增大间隔。鸢尾花数据集是这一趋势的例外，它也是第二小的数据集，训练集仅有42个数据点。

图 A19 - A27 展示了各数据集上表现最佳模型的ROC曲线。比较不同方法时，各数据集的曲线大多相似。

图 A28 - A30 展示了针对具有二维特征向量的数据集，各方法生成的最佳特征映射所得到的决策边界。通过观察这些边界，可以了解参数训练带来的改进，以及各方法生成的决策规则的复杂程度。

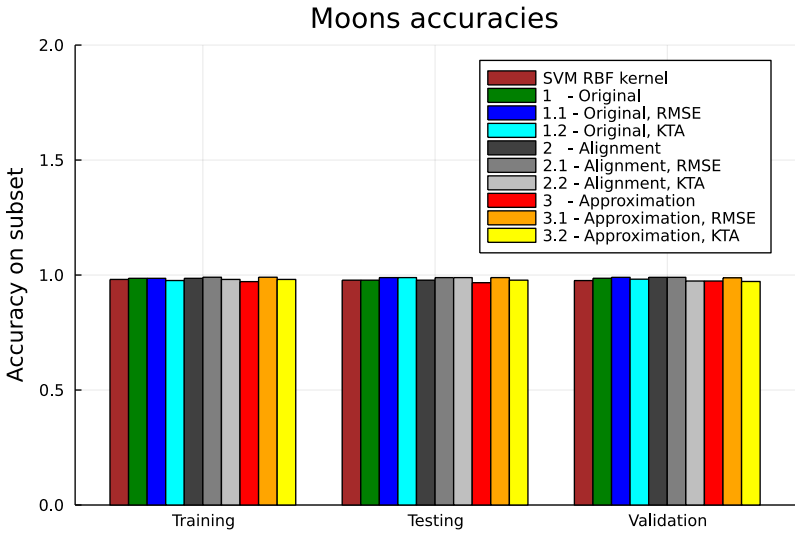


图 A1：图中展示了量子特征映射设计的不同方法所生成的最佳模型在Moons数据集上的分类准确率，并以经典 RBF 核作为参考进行比较。

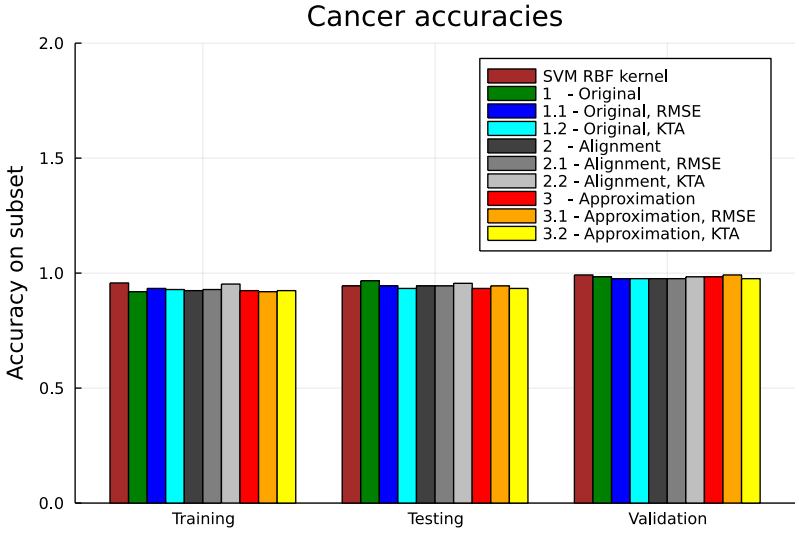


图 A2：图中展示了量子特征映射设计的不同方法所生成的最佳模型在Cancer数据集上的分类准确率，并以经典 RBF 核作为参考进行比较。

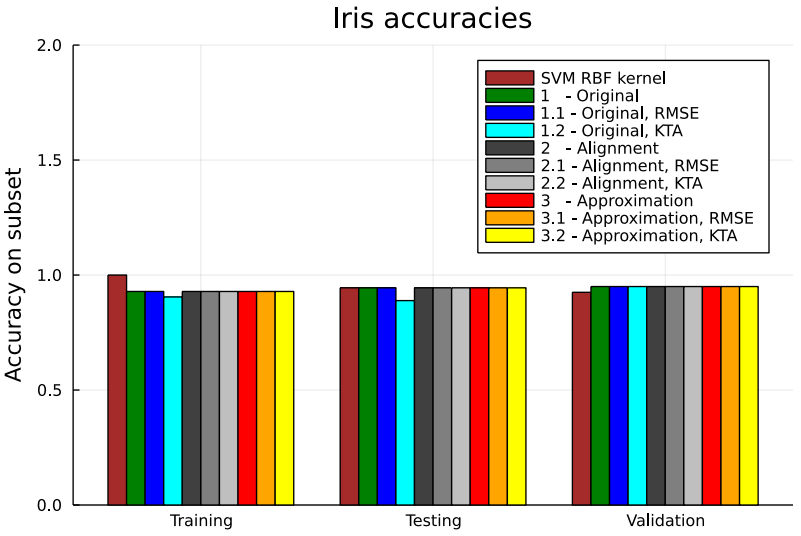


图 A3：图表展示了在鸢尾花数据集上，采用多种量子特征映射设计方法得到的最佳模型的分类准确率，并以经典 RBF 核作为参照。

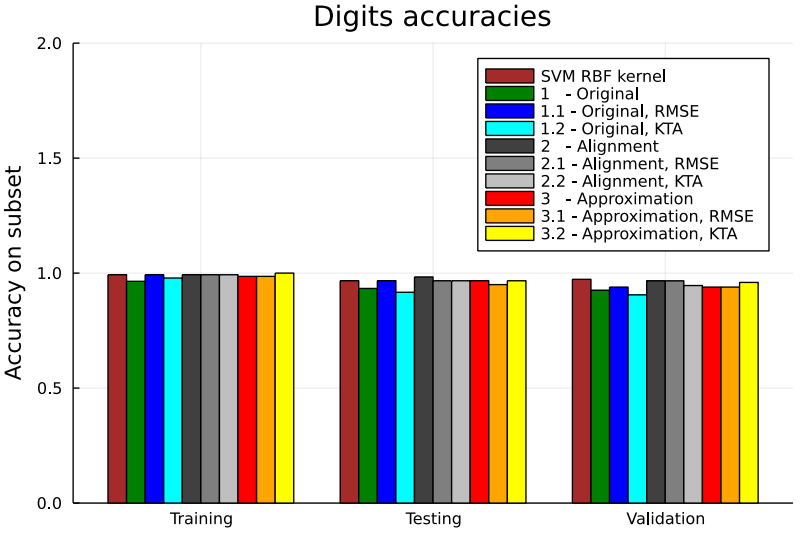


图 A4：图表展示了在手写数字数据集上，采用多种量子特征映射设计方法得到的最佳模型的分类准确率，并以经典 RBF 核作为参照。

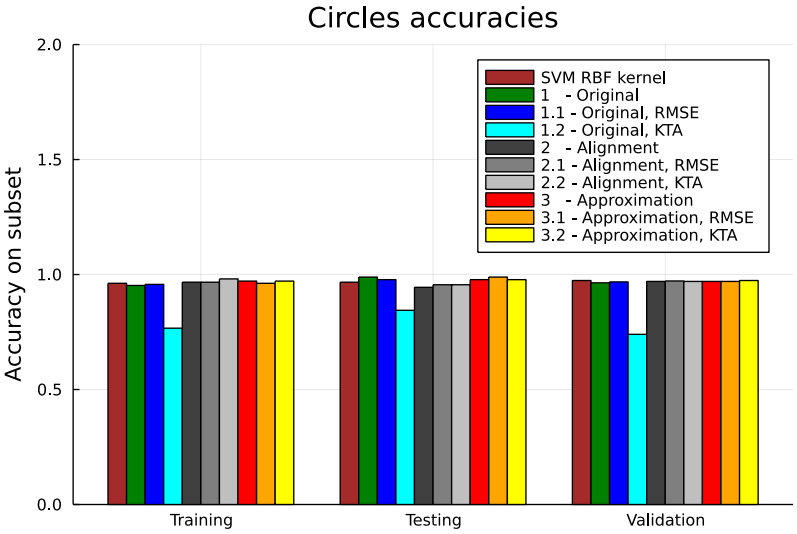


图 A5：图表展示了在 Circles数据集上，通过多种量子特征映射设计方法得到的最佳模型的分类准确率，并以经典 RBF 核作为参考进行比较。

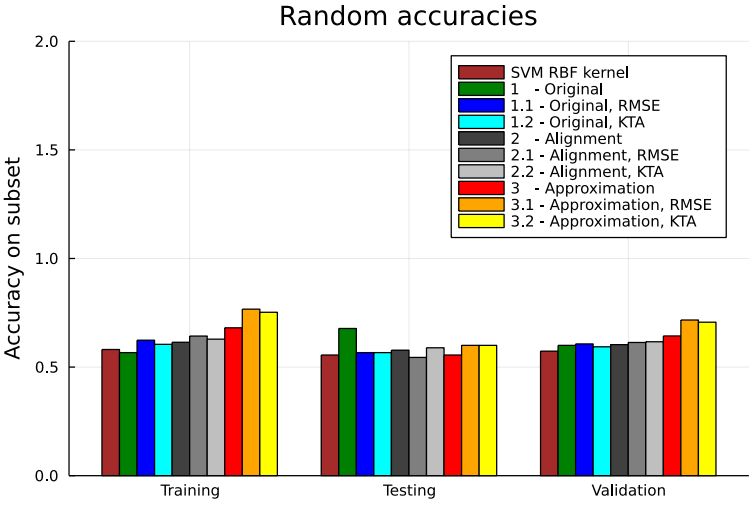


图 A6：图表展示了在 Random 数据集上，通过多种量子特征映射设计方法得到的最佳模型的分类准确率，并以经典 RBF 核作为参考进行比较。对于此数据集，验证集由训练集和测试点的并集构成。这可以反映该方法记住所有点的程度。

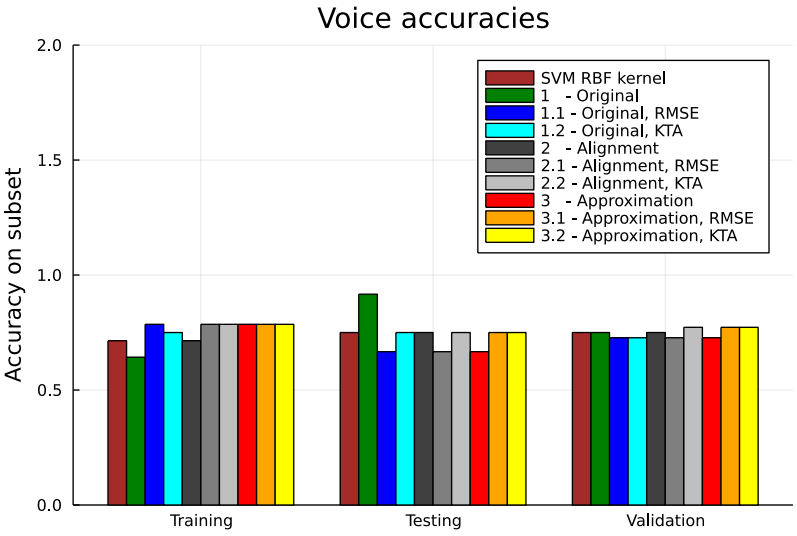


图 A7：图中展示了在 Voice 数据集上，采用不同量子特征映射设计方法得到的最佳模型的分类准确率，并以经典 RBF 核作为参照。

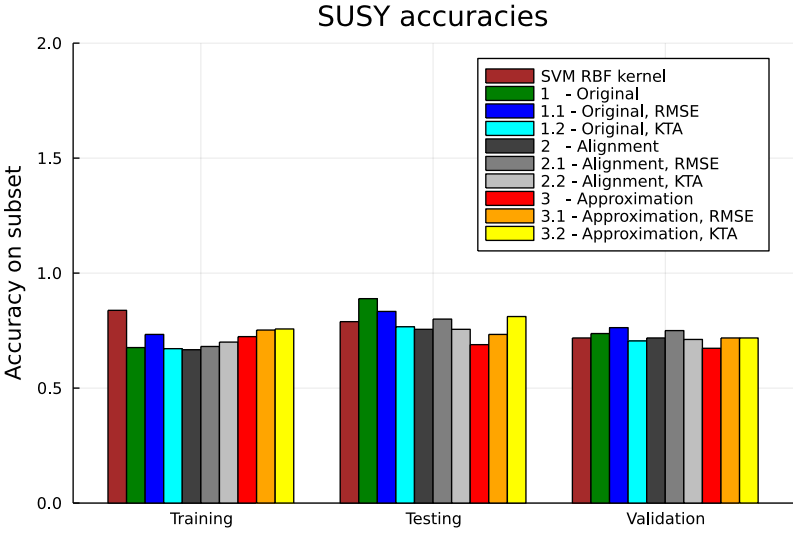


图 A8：图中展示了在 SUSY 数据集上，采用不同量子特征映射设计方法得到的最佳模型的分类准确率，并以经典 RBF 核作为参照。

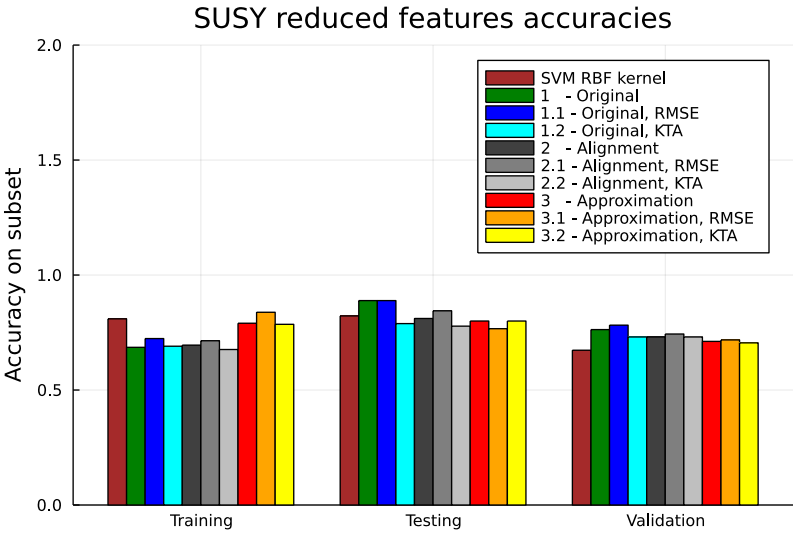


图 A9：图中展示了各种量子特征映射设计方法在 SUSY 降维特征数据集上得到的最佳模型的分类准确率，并以经典 RBF 核作为参照。

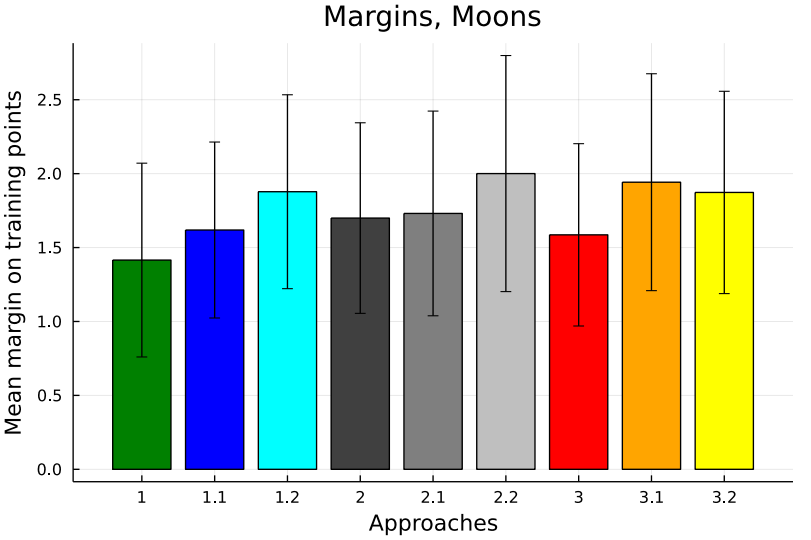


图 A10：图中展示了各方法得到的最佳分类器对 Moons 训练集样本点的平均间隔，误差条表示标准差。

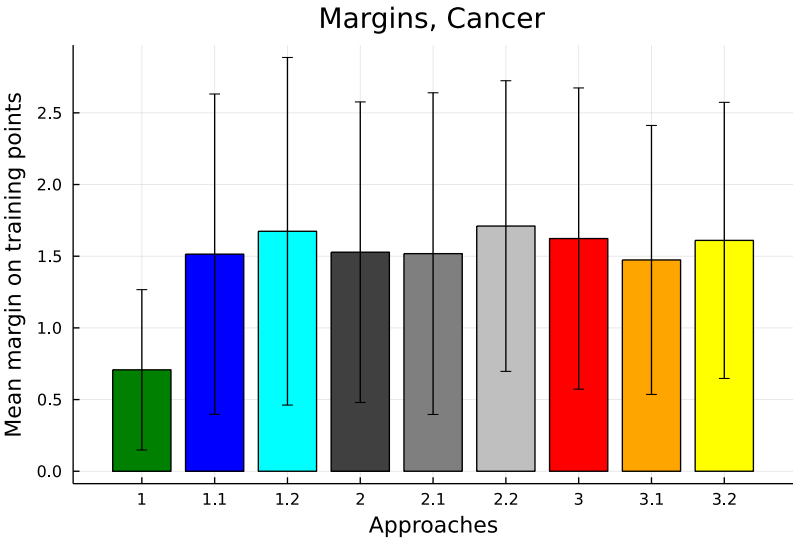


图 A11：该图展示了各方法生成的最佳分类器对癌症训练集样本点的平均间隔，误差条表示标准差。

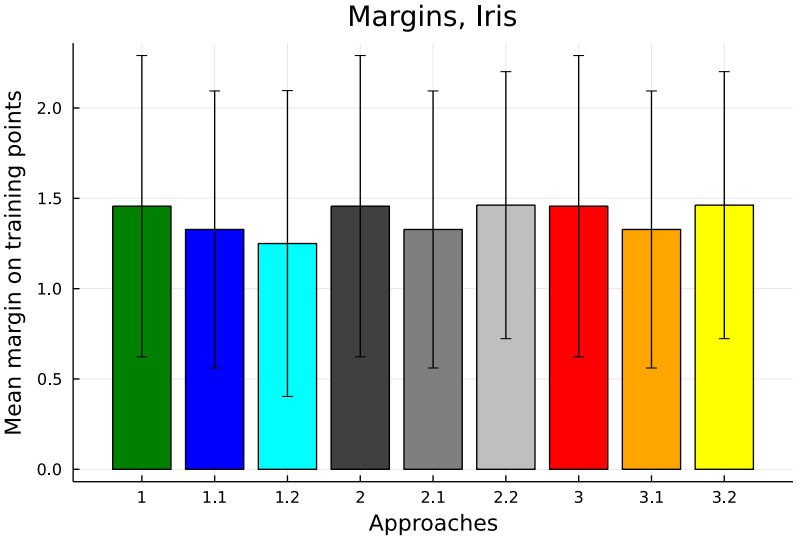


图 A12：该图展示了各方法生成的最佳分类器对鸢尾花训练集样本点的平均间隔，误差条表示标准差。

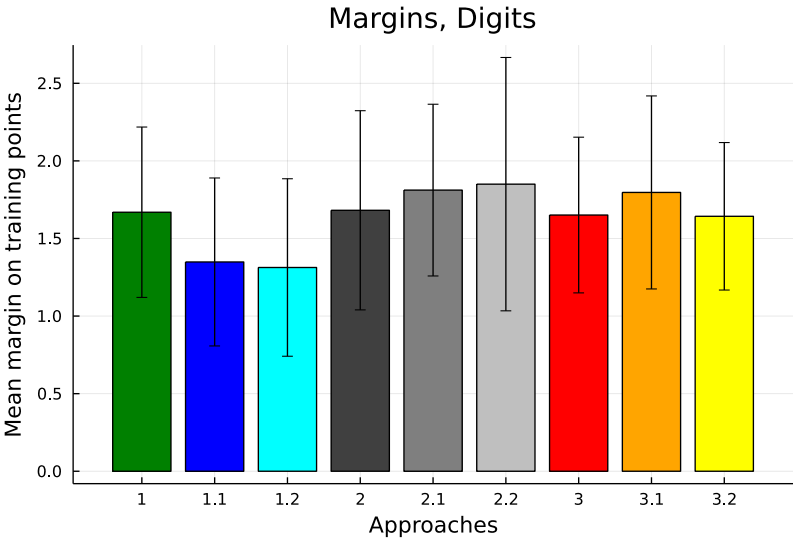


图 A13：该图展示了各方法生成的最佳分类器对 Digits 训练集样本点的平均间隔，误差棒表示标准差。

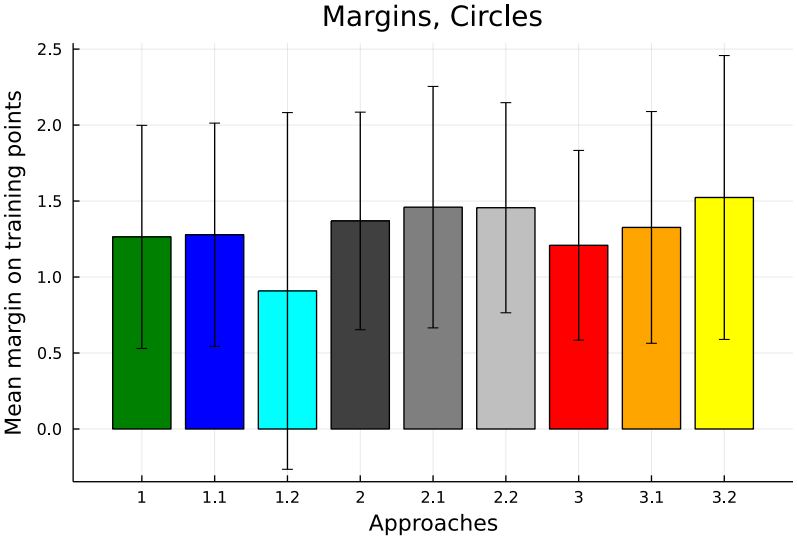


图 A14：该图展示了各方法生成的最佳分类器对 Circles 训练集样本点的平均间隔，误差棒表示标准差。

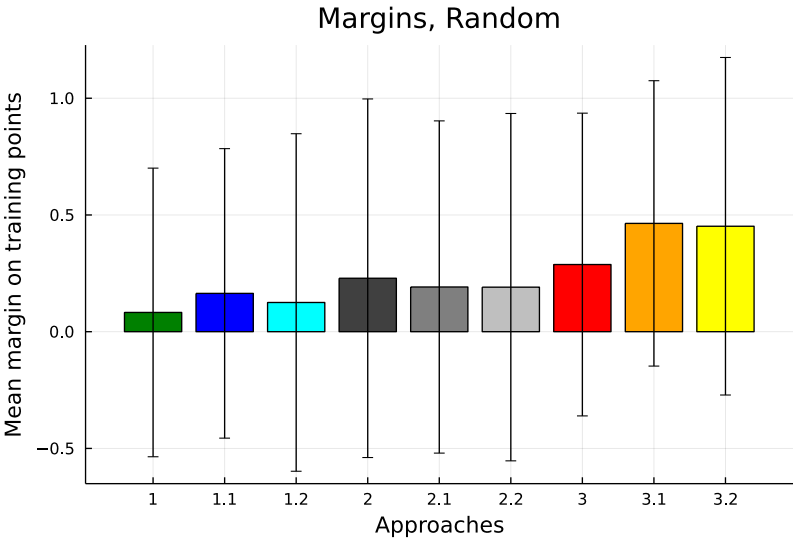


图 A15：图中显示了各方法生成的最佳分类器对随机训练集样本的平均间隔，误差线表示标准差。

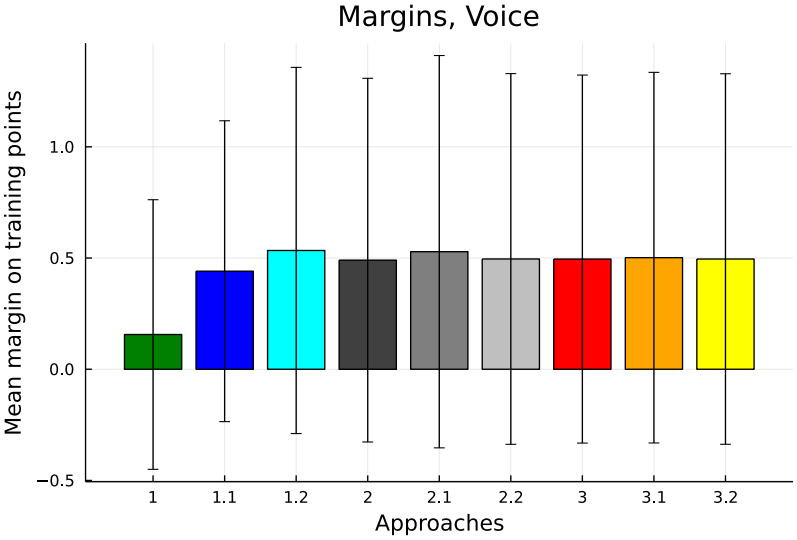


图 A16：图中显示了各方法生成的最佳分类器对语音训练集样本的平均间隔，误差线表示标准差。

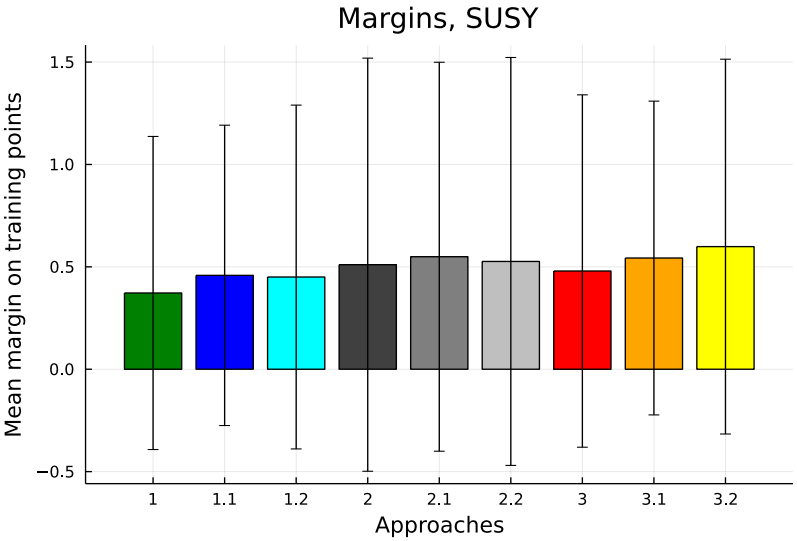


图 A17：图中展示了各方法生成的最佳分类器在 SUSY 训练集样本上的平均间隔，误差线表示标准差。

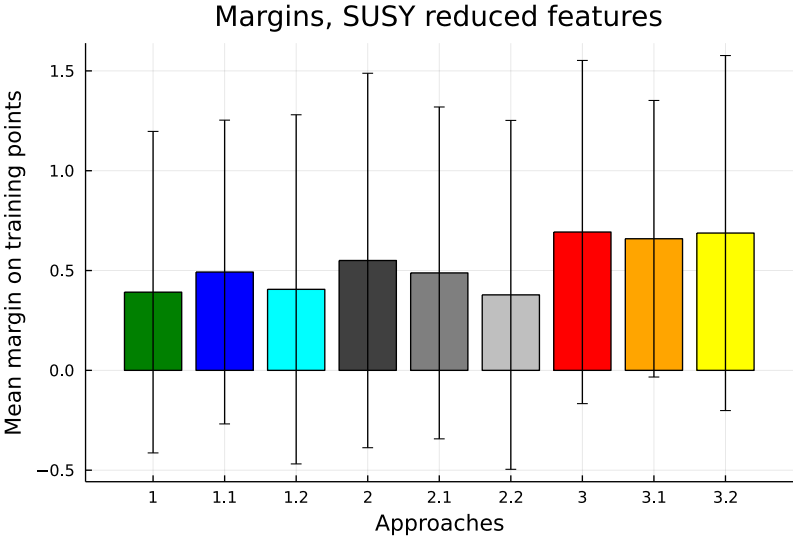


图 A18：图中展示了各方法生成的最佳分类器在 SUSY精简特征训练集样本上的平均间隔，误差线表示标准差。

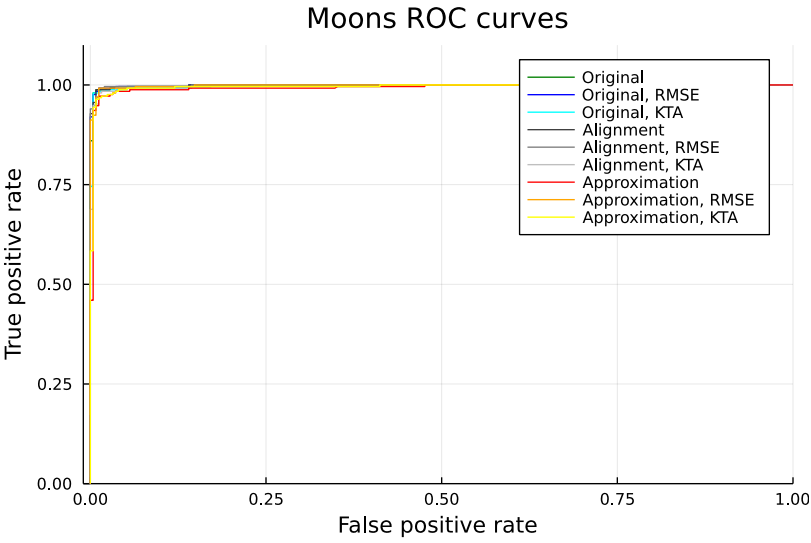


图 A19：展示在Moons数据集上采用量子特征映射设计的各种方法所生成的最佳模型ROC曲线的图。

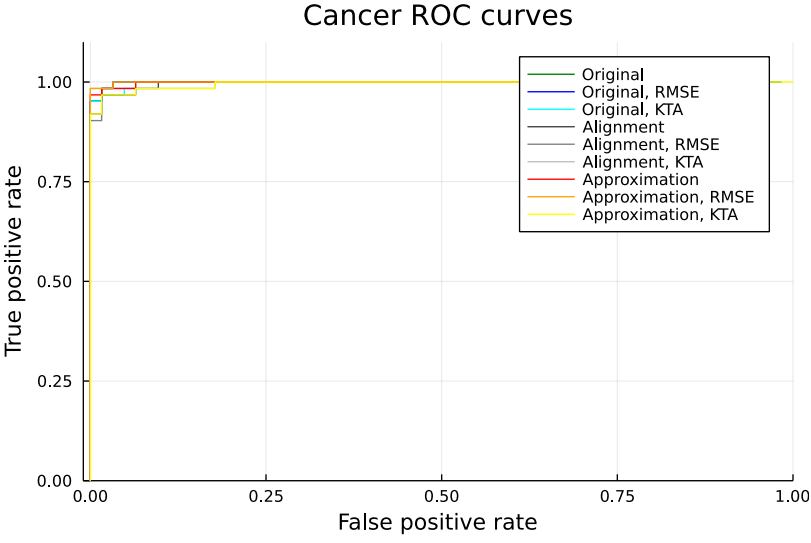


图 A20：展示在Cancer数据集上采用量子特征映射设计的各种方法所生成的最佳模型ROC曲线的图。

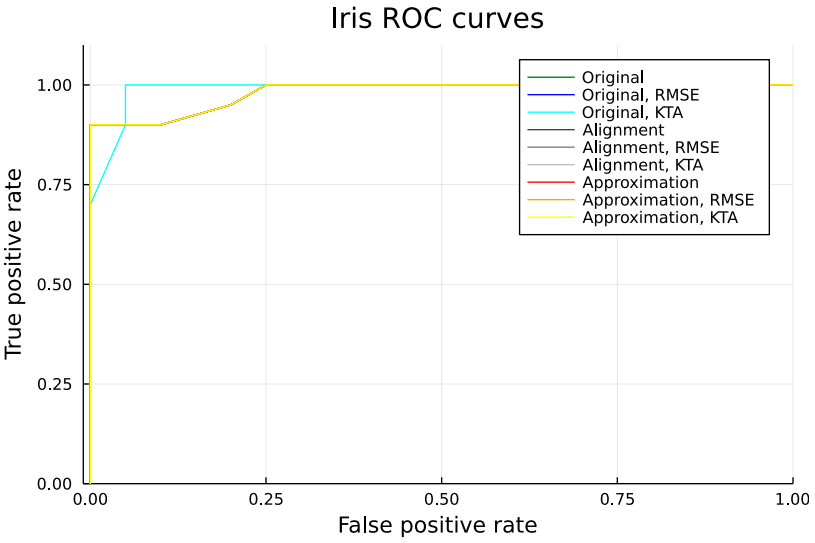


图 A21：展示了在鸢尾花数据集上，通过各种量子特征映射设计方法生成的最佳模型的 ROC曲线。

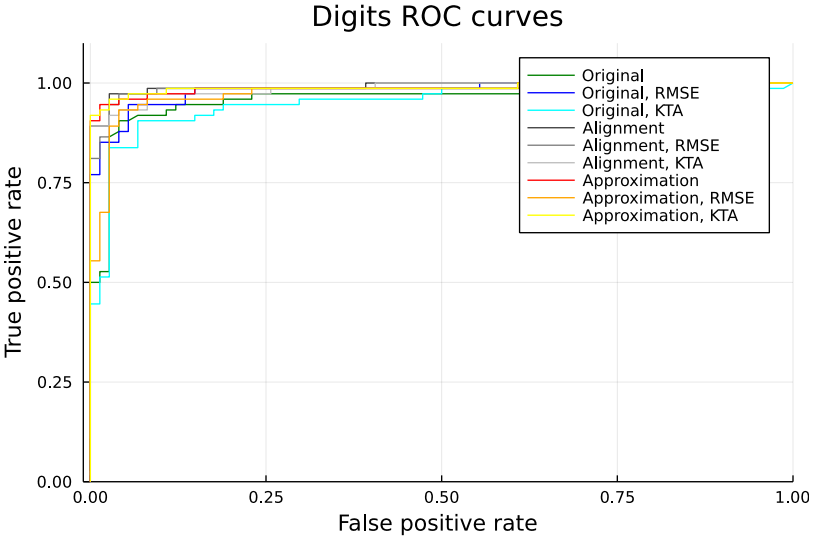


图 A22：展示了在手写数字数据集上，通过各种量子特征映射设计方法生成的最佳模型的 ROC曲线。

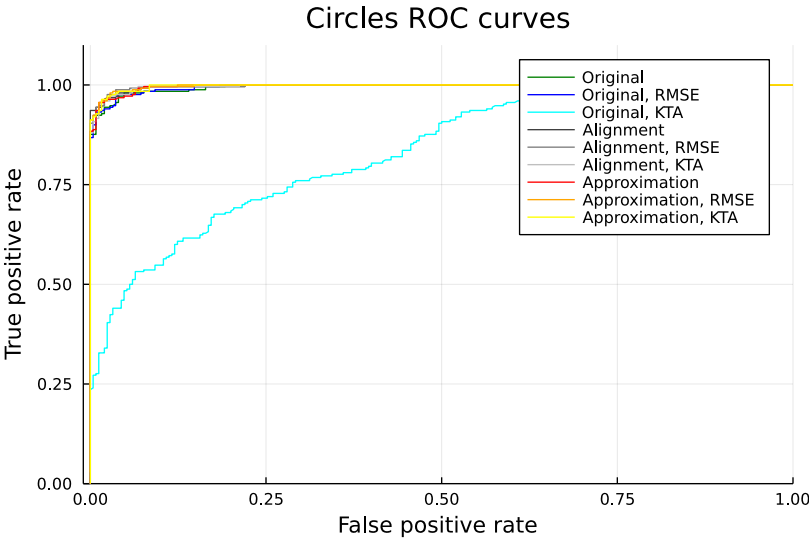


图 A23：展示在Circles数据集上，采用不同量子特征映射设计方法得到的最佳模型的ROC曲线。

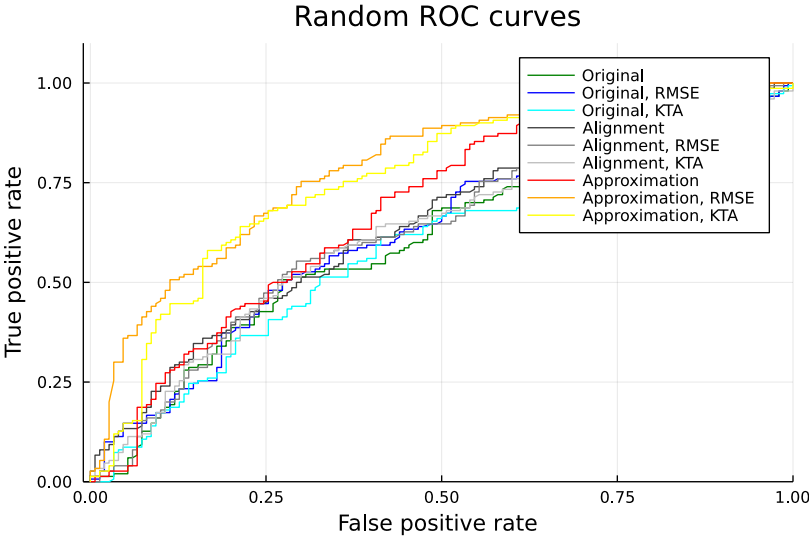


图 A24：展示在Random 数据集上，采用不同量子特征映射设计方法得到的最佳模型的ROC曲线。

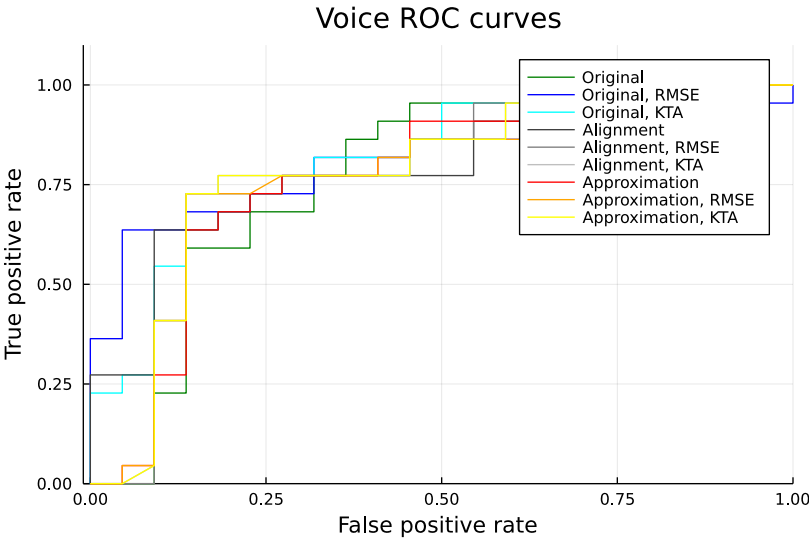


图 A25：展示了在 Voice 数据集上，采用不同量子特征映射设计方法所生成的最佳模型的 ROC曲线。

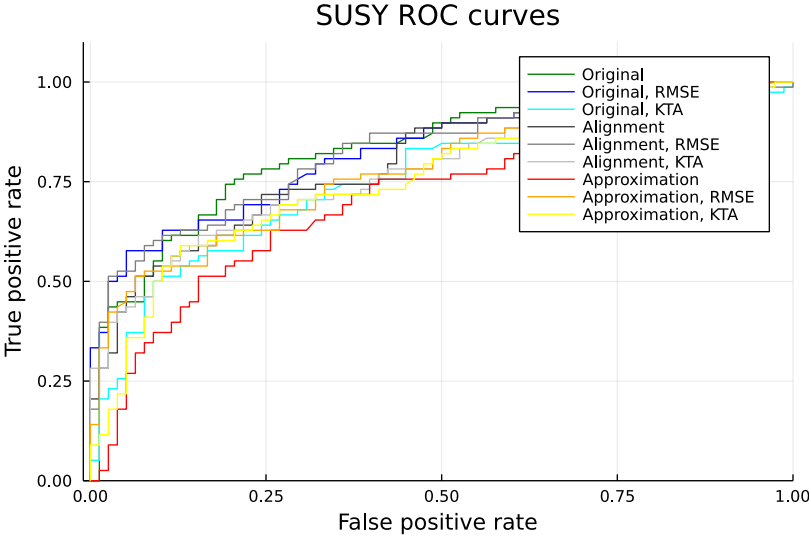


图 A26：展示了在 SUSY 数据集上，采用不同量子特征映射设计方法所生成的最佳模型的 ROC曲线。

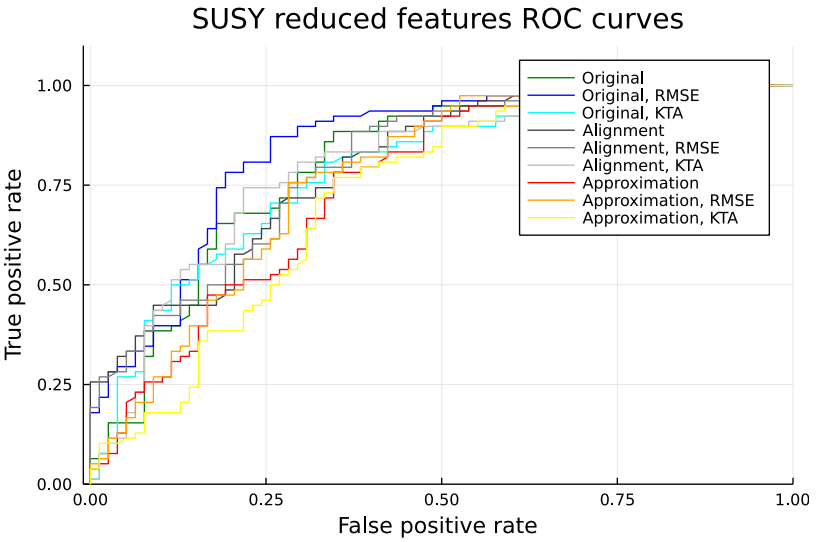


图 A27：展示在 SUSY 降维特征数据集上，通过多种量子特征映射设计方法生成的最佳模型的 ROC曲线。

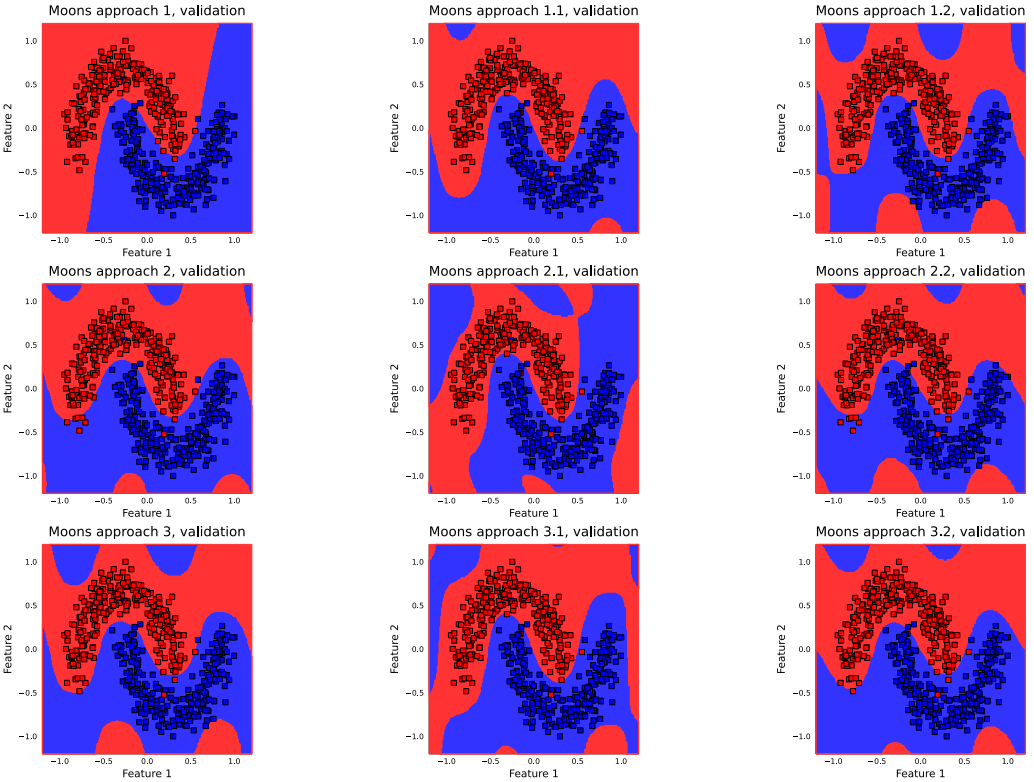


图 A28：各方法在 Moons 验证数据上的决策边界。

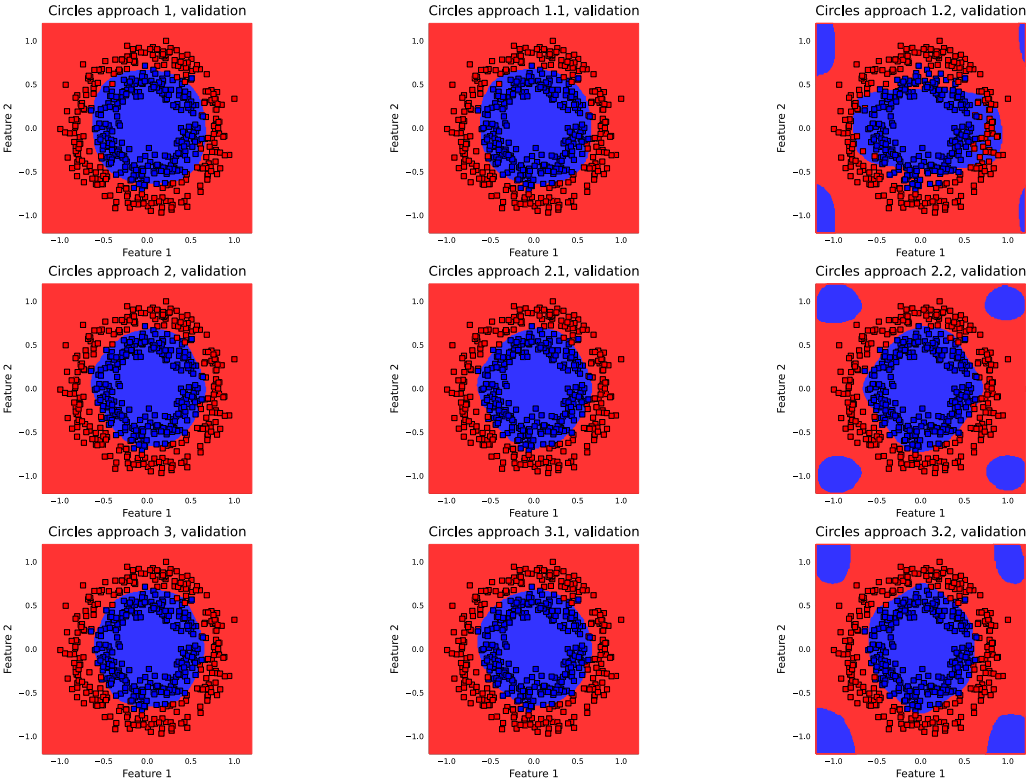


图 A29：各方法在 Circles 验证数据上的决策边界。

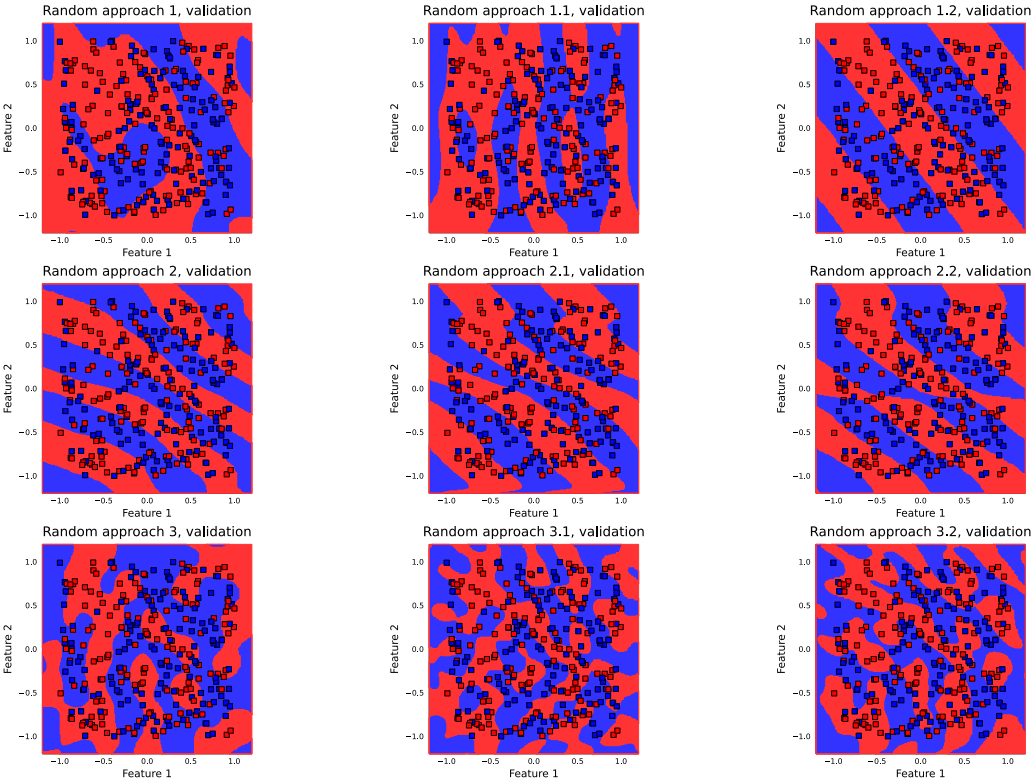


图 A30：随机验证数据上各方法的决策边界。对于 Random 数据集，验证数据并不像其他数据集那样由未见过的随机生成点组成。假设模型无法泛化到新的伪随机数据，验证数据集改为展示训练点和测试点的拼接结果，以此说明不同方法记忆随机点的随机标签分配的能力。