【电子与信息科学／Electronics and Information Science】

基于Q学习的极限学习机

朱丰君[1]，邹伟东[1]，钟必能[2]

1）北京理工大学(珠海)海洋科技学域，广东珠海 519088；2）广西师范大学广西多源信息挖掘与安全重点实验室，广西桂林 541004

摘　要：针对极限学习机(extreme learning machine，ELM)处理病态矩阵时易引发数值不稳及过拟合的问题，提出一种基于Q学习(Q-learning)的极限学习机(ELM based on Q-learning，Q-ELM)．该方法将输出权重的求解过程建模为马尔可夫决策过程，利用强化学习智能体在离散动作空间中进行迭代寻优．通过引入权重截断机制实现了隐式的硬约束正则化，有效抑制了噪声放大效应．理论上证明了算法在凸目标函数下的收敛性及其抗噪边界．实验结果表明：Q-ELM算法在小样本分类任务(Glass数据集)中的分类准确率比ELM算法提升了5.39个百分点，在高维病态任务(SynapseMNIST3D数据集)中提升了2.19个百分点，显著增强了模型的鲁棒性与分类性能；在9个对比算法的19个数据集上的分类任务中，Q-ELM算法取得了4.16的平均排名，在单隐层前馈神经网络(single-hidden layer feed forward network，SLFN)算法中位居首位，且算法的显著性分析结果优于ELM及其SLFN变体．Q-ELM算法有效提升了ELM的分类性能与抗过拟合能力，为其输出权重的优化提供了新思路．

关键词：人工智能；机器学习；极限学习机；Q学习；智能决策；权重截断机制

中图分类号：TP183　　**文献标志码：**A　　**DOI：**10.3724/SP.J.1249.2026.03357

Extreme learning machine based on Q-learning

ZHU Fengjun[1], ZOU Weidong[1], and ZHONG Bineng[2]

1) Marine Science and Technology Domain, Beijing Institute of Technology, Zhuhai, Zhuhai 519088, Guangdong Province, P. R. China

2) Guangxi Key Laboratory of Multi-source Information Mining and Security, Guangxi Normal University, Guilin 541004, Guangxi Zhuang Autonomous Region, P. R. China

Abstract: Extreme learning machines (ELMs) are prone to numerical instability and overfitting when dealing with ill-conditioned matrices, primarily due to the use of the Moore-Penrose pseudoinverse for output weight estimation. To address these issues, this paper proposes an ELM-based on Q-learning (Q-ELM), in which the determination of output weight is reformulated as a Markov decision process (MDP) by utilizing a reinforcement learning agent to iteratively searches for optimal solution within a discrete action space. A weight clipping mechanism is introduced to impose implicit hard-constrained regularization, effectively suppressing noise amplification. Theoretical analysis establishes the convergence of the proposed algorithm under convex objective functions and derives its noise robustness bounds. Experimental results demonstrate that Q-ELM significantly outperforms the standard ELM, achieving accuracy improvements of up to 5.39 percentage points on the small-sample scenarios (Glass dataset) and 2.19 percentage points on the high-dimensional ill-posed task (SynapseMNIST3D dataset). These gains highlight the enhanced robustness and classification performance of the proposed model. Furthermore, across 19 benchmark

Received: 2025-08-06; **Accepted:** 2025-12-23; **Online** (CNKI)**:** 2026-04-08

Foundation: National Natural Science Foundation of China (61906015); Open Research Fund of Guangxi Key Laboratory of Multi-source Information Mining & Security of Guangxi Normal University (MIMS22-13)

Corresponding author: Lecturer ZOU Weidong (zouweidong1985@163.com)

Citation: ZHU Fengjun, ZOU Weidong, ZHONG Bineng. Extreme learning machine based on Q-learning [J]. Journal of Shenzhen University Science and Engineering, 2026, 43(3): 357-364. (in Chinese)

Open access: This article is licensed under the CC BY 4.0 License (https://creativecommons.org/licenses/by/4.0/).

http://journal.szu.edu.cn

datasets and in comparison with nine competing algorithms, Q-ELM attains an average ranking of 4. 16, securing the top position among all single-hidden layer feed forward network (SLFN) algorithms. Notably, statistical significance tests confirm that Q-ELM yields superior results compared to conventional ELM and its SLFN variants.

Key words: artificial intelligence; machine learning; extreme learning machine; Q-learning; intelligent decision making; weight clipping mechanism

2004年，HUANG等[1]提出极限学习机(extreme learning machine，ELM)，并从数学上严格证明：只要激活函数满足特定条件，ELM可具备通用的函数逼近能力. ELM及其改进算法现已被应用于图像识别和分类[2-3]、工业生产数据检测[4]和医疗诊断[5]等领域. 但是，尽管ELM具有较高计算效率，其基于Moore-Penrose伪逆的解析解在特定工程数据下仍存在鲁棒性问题. 当ELM的隐藏层节点数较多或训练数据存在多重共线性时，隐藏层的输出矩阵可能呈现出病态特征[6]，此时小奇异值会放大噪声从而降低含噪数据上的泛化性能. 因此，探索更加稳定的求解机制具有重要意义.

研究者从多个角度提出ELM改进策略：RELM[7]通过L2正则化缓解过拟合，但对参数敏感；RBF-ELM[8]使用核函数增强局部逼近能力，但计算成本随数据量剧增；I-ELM[9]和QRI-ELM[10]通过自适应调整节点数以平衡精度与复杂度；PSO-ELM[11]和GA-ELM[12]等进化算法通过全局优化权重提升精度，但训练耗时更长. 深层结构方面，ML-ELM[13]通过多层堆叠增强特征提取能力；集成ELM[14]通过组合多个基模型降低方差并提升泛化性能，但模型复杂度和资源开销较大.

本研究提出一种基于Q学习的极限学习机(ELM based on Q-learning，Q-ELM)，将输出权重的求解建模为马尔可夫决策过程(Markov decision process，MDP)，构建“状态-动作-奖励”优化框架，并结合权重截断机制实现有界搜索. 该方法相比传统伪逆解具有更好的数值稳定性和抗噪性能，为病态线性系统提供了新的求解思路.

1 ELM算法

ELM是一种单隐层前馈神经网络(single-hidden layer feed forward network，SLFN)算法，其核心机制在于将网络的训练过程解耦为两个阶段：① 通过随机生成输入权重和偏置，将输入数据非线性映射至高维特征空间；② 采用Moore-Penrose伪逆解析求解输出权重. 表1给出了ELM算法中主要变量的定义.

表1 主要变量定义

Table 1 Definitions of main variables

符号	注释	符号	注释
N	样本总量	j	样本索引
n	样本特征维数	i	隐藏层节点索引
L	隐藏层维数	m	输出层维数
$\boldsymbol{t}_j$	样本j的标签向量	$\boldsymbol{x}_j$	样本j的特征向量
$\boldsymbol{w}_i$	第i个隐藏层节点的输入权重	$\boldsymbol{\beta}_i$	第i个隐藏层节点的输出权重
b_i	第i个隐藏层节点的输入偏置	ϕ	激活函数
S	状态空间	A	动作空间
R	奖励函数	τ	迭代次数

ELM的数学模型为

$$\boldsymbol{t}_j = \sum_{i=1}^{L} \boldsymbol{\beta}_i \phi\left(\boldsymbol{w}_i \boldsymbol{x}_j + b_i\right), \quad j = 1, 2, \cdots, N \tag{1}$$

若令隐藏层的输出权重$\boldsymbol{\beta} = \left[\boldsymbol{\beta}_1\ \boldsymbol{\beta}_2\ \cdots\ \boldsymbol{\beta}_L\right]^{\mathrm{T}}$，样本标签矩阵$\boldsymbol{T} = \left[\boldsymbol{t}_1\ \ \boldsymbol{t}_2\ \ \cdots\ \ \boldsymbol{t}_N\right]^{\mathrm{T}}$，且隐藏层矩阵为$\boldsymbol{H} = \begin{bmatrix} \phi\left(\boldsymbol{w}_1 \boldsymbol{x}_1 + b_1\right) & \cdots & \phi\left(\boldsymbol{w}_L \boldsymbol{x}_1 + b_L\right) \\ \vdots & & \vdots \\ \phi\left(\boldsymbol{w}_1 \boldsymbol{x}_N + b_1\right) & \cdots & \phi\left(\boldsymbol{w}_L \boldsymbol{x}_N + b_L\right) \end{bmatrix}$，则式(1)可写成

$$\boldsymbol{T} = \boldsymbol{H}\boldsymbol{\beta} \tag{2}$$

求解式(2)，可得ELM输出层权重矩阵的最小二乘解为

$$\hat{\boldsymbol{\beta}} = \boldsymbol{H}^{\dagger}\boldsymbol{T} \tag{3}$$

其中，$\boldsymbol{H}^{\dagger}$为$\boldsymbol{H}$的Moore-Penrose广义逆矩阵.

2 Q-ELM算法

2.1 $\hat{\boldsymbol{\beta}}$求解的MDP建模

将ELM的权重求解构建为确定性MDP. 对输出层权重矩阵按列分解，对于其中某列的权重向量$\boldsymbol{\beta}_{\mathrm{col}} \in \mathbf{R}^{L\times 1}$，其MDP四元组分别为：

1）状态空间$\boldsymbol{S}$. 状态$\boldsymbol{s}_\tau(\boldsymbol{s}_\tau \in S)$为第$\tau$次迭代的残差向量，$\boldsymbol{s}_\tau = \boldsymbol{T}_{\mathrm{col}} - \boldsymbol{H}\boldsymbol{\beta}_{\mathrm{col}}^{(\tau)}$. 其中，$\boldsymbol{\beta}_{\mathrm{col}}^{(\tau)}$为第$\tau$次

http://journal.szu.edu.cn

迭代的$\boldsymbol{\beta}_{\text{col}}$. 范数$\|\boldsymbol{s}_\tau\|$表征当前解的逼近误差.

2）动作空间A. 第τ次迭代的动作$\boldsymbol{a}_\tau(\boldsymbol{a}_\tau \in A)$为沿$\boldsymbol{\beta}$的权重空间$\mathbf{R}^{L\times m}$基向量方向的离散更新，即

$$\boldsymbol{a}_\tau = \delta \boldsymbol{e}_d, \quad \delta \in \Delta, d \in \{1, 2, \cdots, L\} \tag{4}$$

其中，δ为步长；Δ为预设步长集合；$\boldsymbol{e}_d$为第d个权重分量的标准基向量. 动作空间的对称性保证了算法可沿任意坐标正负方向调整权重.

3）状态转移P. 执行动作$\boldsymbol{a}_\tau$并采用$\boldsymbol{\beta}^{(\tau+1)} = \boldsymbol{\beta}^{(\tau)} + \boldsymbol{a}_\tau$更新权重，因此，导出状态转移方程为

$$\boldsymbol{s}_{\tau+1} = \boldsymbol{T} - \boldsymbol{H}\left[\boldsymbol{\beta}^{(\tau)} + \boldsymbol{a}_\tau\right] = \boldsymbol{s}_\tau - \boldsymbol{H}\boldsymbol{a}_\tau \tag{5}$$

4）奖励函数R. 为实现最优化目标函数$\min\|\boldsymbol{s}_\tau\|^2$，将单步奖励$r_\tau$定义为残差范数的衰减量，即$r_\tau = R(\boldsymbol{s}_\tau, \boldsymbol{a}_t) = \|\boldsymbol{s}_\tau\| - \|\boldsymbol{s}_{\tau+1}\| \times f_{\text{scale}}$. 其中，$f_{\text{scale}}$为缩放因子. 当残差下降率超过阈值$\rho$时，引入稀疏奖励以加速收敛.

算法目标为学习最优策略$\boldsymbol{\pi}^*$：$S \to A$，以最大化累计折扣回报[15]($G_\tau = r_{\tau-1} + \gamma r_{\tau-2} + \gamma^2 r_{\tau-3} + \cdots$. 其中，$\gamma \in [0, 1)$为折扣因子). 该最优策略为

$$\boldsymbol{\pi}^*(\boldsymbol{s}) = \arg\max_{\boldsymbol{a}\in A} \boldsymbol{Q}^*(\boldsymbol{s}, \boldsymbol{a}) \tag{6}$$

其中，$\boldsymbol{s}$为当前状态；$\boldsymbol{a}$为智能体在当前状态下可选择的具体“动作”，$\boldsymbol{Q}^*(\boldsymbol{s}, \boldsymbol{a})$为最优状态-动作价值函数.

2.2　Q-ELM算法实现

对于输出权重$\boldsymbol{\beta} \in \mathbf{R}^{L\times m}$的每列$\boldsymbol{\beta}_{\text{col}}$，智能体(agent)负责搜索其最优解. Q-ELM算法实现的核心流程分为两个阶段：① 随机生成输入层参数，并计算隐藏层矩阵$\boldsymbol{H}$；② 通过Q-Learning迭代更新$\boldsymbol{\beta}$，其间引入ε-greedy探索策略与权重截断机制以保证算法的鲁棒性. 基于Q-ELM算法的伪代码见图1.

由图1可知，Q-ELM算法两阶段的时间复杂度为$O(NL + mE_{\max}S_{\max}NL)$. 其中，$E_{\max}$为最大训练轮数；$S_{\max}$为每轮最大迭代步数. 又由文献[1]可知，ELM通过伪逆求解的时间复杂度为$O(NL^2 + L^3)$. 可见在一般情况下，Q-ELM训练算法的时间较ELM算法略高，但该额外计算开销避免了矩阵求逆运算，以适度的计算效率换取了在病态矩阵和高维特征场景下更为可靠的泛化表现.

2.3　收敛性分析

ELM算法通过求解$\min J(\boldsymbol{\beta})$来最小化残差平方和，则算法的目标函数为

输入：数据集$\{(x_j, t_j)\}_{j=1}^N$，N为样本总量，隐藏层节点个数L，超参数包括：学习率η，折扣因子γ，探索率ε，最大训练轮数$E_{\max}$，单轮步数$S_{\max}$，步长集合$\Delta = \{\pm\delta_1, \pm\delta_2, \cdots, \delta_{\max}\}$，收敛容差tol

输出：权重矩阵$\boldsymbol{\beta}$

```
// 阶段1：ELM初始化
生成输入权重 W ~ U(−1, 1) 和偏置 b ~ U(−1, 1)
计算隐藏层输出矩阵 H ∈ R^{N×L}
初始化 L × m 维输出权重矩阵 β，所有元素为0
生成动作空间 A = {δe_d | δ ∈ Δ, d ∈ {1, 2, ···, L}}
// 阶段2：权重迭代求解
for col = 1 to m do
    初始化Q-Table Q_col(s, a) ← 0
    for episode = 1 to E_max do
        for step τ = 1 to S_max do
            if 随机数 < ε then 随机选择 a_τ ∈ A
            else a_τ = arg max_{a∈A} Q_col(s_{τ−1}, a)
            β_col^imp ← β_col + a_τ
            β_col ← clip(β_col^imp, β_min, β_max)
            s_τ = T_col − Hβ_col^(τ)
            //计算奖励与更新Q值
            r_τ = (‖s_{τ−1}‖ − ‖s_τ‖) × f_scale      //f_scale为缩放因子
            Q_col(s_{τ−1}, a_τ) = (1 − α)Q_col(s_{τ−1}, a_t) +
                α[r_τ + max_{a'} Q_col(s_τ, a')]
            if ‖s_τ‖ < tol then break
        end for
    end for
end for
return β
```

图1　Q-ELM算法伪代码

Fig. 1　Pseudocode of the Q-ELM algorithm.

$$J(\boldsymbol{\beta}) = \frac{1}{2}\|\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{T}\|^2 \tag{7}$$

$J(\boldsymbol{\beta})$关于$\boldsymbol{\beta}$的海森(Hessian)矩阵为

$$\frac{\partial^2 J(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2} = \boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} \tag{8}$$

对于任意非零向量$\boldsymbol{v} \in \mathbf{R}^{L\times 1}$，式(8)的2次型满足$\boldsymbol{v}^{\mathrm{T}}(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})\boldsymbol{v} = \|\boldsymbol{H}\boldsymbol{v}\| \geqslant 0$. 因此，$\frac{\partial^2 J(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2}$恒为半正定矩阵. 根据凸优化理论，$J(\boldsymbol{\beta})$是定义域内的全局凸函数，所有局部极小值点即全局极小值点. 基于此前提，以下将进一步证明算法的收敛性.

根据强化学习理论，最优动作价值函数$\boldsymbol{Q}^*(\boldsymbol{s}, \boldsymbol{a})$满足贝尔曼最优方程[15]，即

http://journal.szu.edu.cn

$$\boldsymbol{Q}^*(\boldsymbol{s},\boldsymbol{a})=R(\boldsymbol{s},\boldsymbol{a})+\gamma\max_{\boldsymbol{a}'\in A}\boldsymbol{Q}^*(\boldsymbol{s}',\boldsymbol{a}') \tag{9}$$

其中，$\boldsymbol{s}'=\boldsymbol{s}-\boldsymbol{H}\boldsymbol{a}$为执行动作后的次态；$\boldsymbol{a}'$表示在状态$\boldsymbol{s}'$时，在动作空间$A$中可能执行的所有动作；$R(\boldsymbol{s},\boldsymbol{a})$为奖励函数.

定义贝尔曼算子$\mathcal{T}$为$(\mathcal{T}\boldsymbol{Q})(\boldsymbol{s},\boldsymbol{a})=R(\boldsymbol{s},\boldsymbol{a})+\gamma\max_{\boldsymbol{a}'}\boldsymbol{Q}(\boldsymbol{s}',\boldsymbol{a}')$. 其中，$\boldsymbol{Q}(\boldsymbol{s},\ \boldsymbol{a})$为状态-动作价值函数. 对于折扣因子$\gamma(0\leqslant\gamma<1)$，算子$\mathcal{T}$在$L_\infty$范数下构成压缩映射[16]，即对于第$\tau$次迭代有

$$\left\|\mathcal{T}\boldsymbol{Q}_\tau-\mathcal{T}\boldsymbol{Q}^*\right\|_\infty\leqslant\gamma\left\|\boldsymbol{Q}_\tau-\boldsymbol{Q}^*\right\|_\infty \tag{10}$$

依据Banach不动点定理，迭代序列必然收敛于最优不动点$\boldsymbol{Q}^*$.

在Q-ELM中采用时序差分更新规则

$$\boldsymbol{Q}_{\tau+1}(\boldsymbol{s},\boldsymbol{a})=(1-\alpha_\tau)\boldsymbol{Q}_\tau(\boldsymbol{s},\boldsymbol{a})+\alpha_\tau\left[r_\tau+\gamma\max_{\boldsymbol{a}'}\boldsymbol{Q}_\tau(\boldsymbol{s}_{\tau+1},\boldsymbol{a}')\right] \tag{11}$$

其中，α_τ为第τ次迭代的学习率. 当序列满足Robbins-Monro条件$\left(\sum_\tau^\infty\alpha_\tau=\infty且\sum_\tau^\infty\alpha_\tau^2<\infty\right)$[17]时，迭代值$\boldsymbol{Q}_\tau$将依概率收敛至最优值函数$\boldsymbol{Q}^*$[18].

综上，由于函数$J(\boldsymbol{\beta})$是凸性的，因此，对于任意非最优状态$\boldsymbol{s}_\tau\neq0$，梯度$\nabla J(\boldsymbol{\beta})=-\boldsymbol{H}^{\mathrm{T}}\boldsymbol{s}_\tau\neq0$. 由于动作空间包含沿基向量的正负步长，必然存在$\boldsymbol{a}\in A$使得内积$\langle\nabla J(\boldsymbol{\beta}),\ \boldsymbol{a}\rangle<0$，证得算法在有限步长与$\varepsilon$-greedy探索下，权重更新轨迹能够以高概率进入并停留在以最优权重$\boldsymbol{\beta}^*$为中心，边长为$2\delta_{\min}$的超立方体区域

$$\Omega_\delta=\left\{\boldsymbol{\beta}\in\mathrm{R}^{L\times m}\left\|\boldsymbol{\beta}-\boldsymbol{\beta}^*\right\|_\infty\leqslant\delta_{\min}\right\} \tag{12}$$

其中，$\delta_{\min}$为最小非零搜索步长.

2.4 鲁棒性与抗噪性分析

设训练数据含加性噪声$\boldsymbol{E}$，则含噪声的隐藏层输出矩阵为$\tilde{\boldsymbol{H}}=\boldsymbol{H}+\boldsymbol{E}$. 采用ELM解出此时的输出权重为$\tilde{\boldsymbol{\beta}}=\tilde{\boldsymbol{H}}^{\dagger}\boldsymbol{T}$. 又由奇异值分解可知，$\tilde{\boldsymbol{H}}$的Moore-Penrose伪逆为$\tilde{\boldsymbol{H}}^{\dagger}=\sum\sigma_q\boldsymbol{v}_q\boldsymbol{u}_q^{\mathrm{T}}$. 其中，$\sigma_q$、$\boldsymbol{v}_q$和$\boldsymbol{u}_q$分别为矩阵的第$q$个奇异值、左奇异向量和右奇异向量. 当矩阵病态，小奇异值$(\sigma_q\to0)$时，将噪声$\boldsymbol{E}$放大σ_q^{-1}倍，会导致模型误差增大.

Q-ELM算法通过权重截断机制进行硬约束：

$$\boldsymbol{\beta}_j\leftarrow\mathrm{clip}\left(\boldsymbol{\beta}_j+\boldsymbol{a}_\tau,\boldsymbol{\beta}_{\min},\boldsymbol{\beta}_{\max}\right) \tag{13}$$

其中，clip为截断函数，可将权重限制在预设阈值$[-C,\ C]$(C为大于0的常数)范围内；$\boldsymbol{\beta}_{\min}$和$\boldsymbol{\beta}_{\max}$分别为预设的隐藏节点的最大和最小权重阈值. 式(13)在优化理论上等价于引入L_∞范数正则化，即

$$\min_{\boldsymbol{\beta}}\left\|\boldsymbol{H}\boldsymbol{\beta}-\boldsymbol{T}\right\|^2,\quad\text{s.t. }\left\|\boldsymbol{\beta}\right\|_\infty\leqslant C \tag{14}$$

统计学习理论指出，权重有界性限制了假设空间的Rademacher复杂度[19-20]，因此，设由输入扰动$\Delta\boldsymbol{x}$引发的隐藏层偏差为$\Delta\boldsymbol{h}$，则Q-ELM输出的偏差$(|\Delta\boldsymbol{y}|)$的上界为

$$|\Delta\boldsymbol{y}|=\left\|\Delta\boldsymbol{h}\boldsymbol{\beta}\right\|\leqslant\left\|\Delta\boldsymbol{h}\right\|_1\left\|\boldsymbol{\beta}\right\|_\infty\leqslant\left\|\Delta\boldsymbol{h}_1\right\|C \tag{15}$$

式(15)表明，即便是面对病态矩阵或含噪声的输入，Q-ELM算法仍能通过参数截断抑制噪声的放大，从理论上保证了输出偏差$|\Delta y|$存在一个不随矩阵条件数恶化而发散的常数上界，显著提升了模型的抗噪性能.

3 实验设计及结果分析

3.1 Sinc函数拟合性能实验

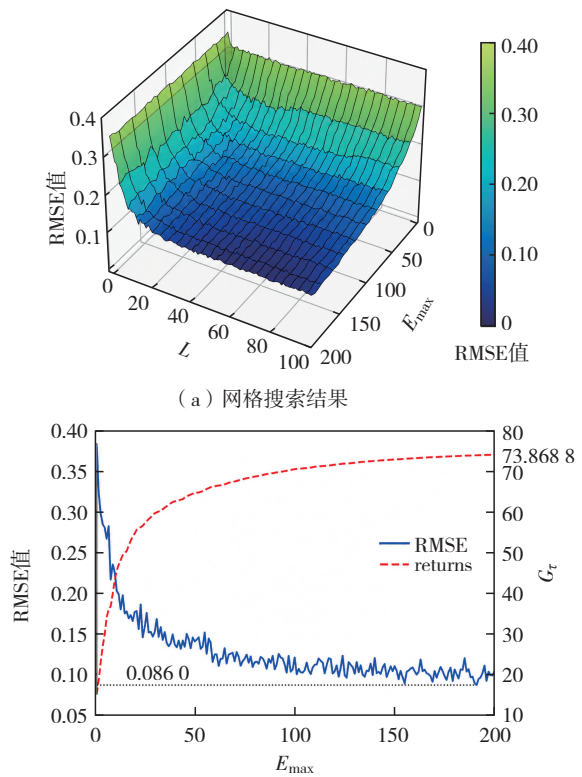
为验证Q-ELM算法的收敛性能，对sinc函数进行拟合实验. 在[−10，10]内均匀选取3 000个数据点并生成对应的函数值构成原始数据集. 将该数据集按7∶3的比例分别划分为训练集与测试集. 选择sigmoid函数作为激活函数，隐藏层节点数设置为50，最大训练轮数($E_{\max}$)以步长1在[1，200]变化. 在训练集的标签中随机加入[−0.02，0.02]均匀分布的噪声，并记录不同$E_{\max}$条件下的均方根误差(root mean squared error，RMSE)值.

图2(a)为Q-ELM算法的网格搜索结果. 由图2(a)可见，在搜索范围内，随着$E_{\max}$和L的增大，模型的拟合精度提高，随后误差降势变缓. 当$E_{\max}>100$且$L>40$时，模型能够保持稳定的拟合精度. 图2(b)为$L=50$，$E_{\max}$从1增至200时的RMSE值和强化学习的累计折扣回报(G_τ)随$E_{\max}$变化的曲线. 由图2(b)可见，EMSE值初期迅速下降然后趋于平缓，累积折扣回报同样在初期迅速上升后趋于平缓，最终算法收敛并有效拟合sinc函数.

3.2 UCI数据集上的分类实验

为验证Q-ELM算法的综合性能，在15个UCI数据集[21]上对不同分类算法的性能进行对比. 这些对比算法分为4类，包括：① 基础ELM、引入L_2正则化的RELM与在线学习算法OS-ELM；② 引入特定优化策略的改进算法，如基于加权网络的WELM、基于交替方向乘子法的ADMM-ELM[22]以及基于梯度下降的迭代网络ADAM-ELM；③ 基于

http://journal.szu.edu.cn



（a）网格搜索结果

（b）RMSE和累计折扣回报随E_{max}变化

图2　Q-ELM算法的收敛性能

Fig. 2　Convergence performance of the Q-ELM algorithm.

动态结构的增量网络 I-ELM；④ 基于核映射的 RBF-ELM 和基于集成学习的 Ensemble-ELM 作为强基准．前 3 类均为 SLFN 框架算法．除 RBF-ELM 外，其他算法的激活函数均为 sigmoid，且随机数相同．不同算法的 UCI 数据集训练结果见表 3．

由表 3 可见，Q-ELM 算法在标准 SLFN 架构下展现出良好的稳定性，而对比算法在各数据集上的泛化表现都存在不稳定的波动，如 I-ELM 算法在面临 Glass 与 Wall-Robot 等特征高度重叠且决策边界非线性的数据集时，分类准确率分别下探至 61.54% 与 64.32%．相比之下，Q-ELM 算法凭借全局 ε-greedy 探索策略，在这两个任务上都取得了稳定表现，在 Glass 数据集上的分类准确率（69.54%）较 ELM 算法（64.15%）高 5.39 个百分点，甚至略胜于 RBF-ELM（68.92%）；而在 Wall-Robot 数据集上，Q-ELM 算法的分类准确率（73.46%）也比 I-ELM 算法（64.32%）高．总体而言，对比其他算法在部分数据集上的性能大幅下降，Q-ELM 在 15 个 UCI 数据集上不存在严重的性能短板，在保持整体竞争力的同时实现了稳定的性能．

3.3　MedMNIST 数据集上的分类实验

为验证算法在高维特征空间的参数搜索能力与鲁棒性，选用 MedMNIST 数据集[23]中模态差异显著的 BreastMNIST、RetinaMNIST、Synapsemnist3d 和 PneumoniaMNIST 数据集（表 4）进行分类实验，结果如表 5．为应对 784 ~ 21 952 维的特征跨度挑战，所有图像均先后经过展平与归一化预处理．由表 5

表3　不同算法在15个UCI数据集训练的准确率

Table 3　Training results based on UCI datasets of various algorithms　%

数据集	Q-ELM	ELM	RELM	ADMM-ELM	ADAM-ELM	WELM	I-ELM	OS-ELM	RBF-ELM	Ensemble-ELM
Zoo	92.90	92.26	92.58	90.65	90.97	92.58	87.10	92.26	94.52	93.87
Iris	100.00	99.33	99.78	99.56	98.44	99.78	84.44	99.33	100.00	99.78
Wine	98.52	97.41	97.41	97.41	97.96	96.67	94.44	97.41	99.81	97.96
Sonar	76.98	76.98	76.98	76.98	78.10	76.51	79.37	76.98	93.97	79.21
Seeds	93.33	93.02	92.86	92.86	94.13	93.02	92.06	93.02	93.02	93.81
Glass	69.54	64.15	64.31	64.15	66.92	62.00	61.54	64.15	68.92	67.23
Heart-statlog	78.89	78.89	78.89	78.89	81.48	79.63	83.95	78.89	82.35	82.22
Ecoli	89.60	89.70	89.80	89.80	85.05	84.95	84.16	89.80	90.69	89.50
Ionosphere	90.00	89.62	89.62	89.62	90.28	88.30	88.68	89.62	91.79	90.66
Cancer	97.31	97.19	97.19	97.19	97.37	97.19	96.49	97.19	98.01	97.89
Balance-scale	89.41	88.56	88.72	88.56	87.39	79.26	85.64	88.56	89.79	89.41
Banknote-authentication	99.93	100.00	100.00	100.00	98.57	100.00	97.82	100.00	100.00	100.00
Spambase	91.50	91.46	91.46	91.46	89.32	91.75	86.68	91.46	92.57	92.88
Waveform	82.52	82.35	82.35	82.35	80.06	82.39	79.33	82.35	86.45	85.91
Wall-Robot	73.46	73.60	73.62	73.63	67.35	70.13	64.32	73.63	77.65	77.64

注：灰底数据为对应数据集上的最优值；下划线数据表示当前结果为SLFN类算法中的最优值．

http://journal.szu.edu.cn

可见，在BreastMNIST和RetinaMNIST数据集上，Q-ELM与ELM及变体(ADMM-ELM和ADAM-ELM算法)均收敛至一致准确率(78.63%和51.17%)．这是由ELM凸优化本质决定的：固定输入权重下，线性最小二乘问题存在唯一全局最优解，证明Q-ELM达到了与解析解等效的判别能力．

表4 四个MedMNIST数据集的基本信息

Table 4 Basic information of 4 MedMNIST datasets

数据集	N	n	类别数
BreastMNIST	702	784	2
RetinaMNIST	1 480	2 352	5
Synapsemnist3d	1 582	21 952	2
PneumoniaMNIST	5 332	784	2

表5 不同算法在4个MedMNIST数据集上的分类准确率

Table 5 Classification accuracy of various algorithms on the MedMNIST datasets %

算法	Breast-MNIST	Retina-MNIST	Synapsem-nist3d	Pneumonia-MNIST
Q-ELM	78.63	51.17	73.20	83.60
ELM	78.63	51.17	71.01	83.55
RELM	78.63	51.17	73.02	83.55
ADMM-ELM	78.63	51.17	72.96	83.55
ADAM-ELM	77.56	51.17	72.92	80.66
WELM	75.43	44.50	55.97	85.58
I-ELM	75.64	54.25	68.75	80.77
OS-ELM	78.63	51.17	73.17	83.55
RBF-ELM	80.56	54.42	73.01	82.69
Ensemble-ELM	79.91	53.50	73.20	83.44

注：灰底数据为各数据集下的最优值；下划线数据为SLFN类算法中的最优值．

在高维SynapseMNIST3D数据集中，ELM受矩阵病态影响分类准确率仅为71.01%，而Q-ELM达73.20%，提升了2.19个百分点，分类性能几乎与RELM及集成模型持平，证实采用隐式正则化能有效抑制数值误差与过拟合．在PneumoniaMNIST数据集上，各算法表现相近，WELM算法以85.58%准确率居首，表明加权优化策略在此类数据上较固定权重更具优势．

综上，Q-ELM能够在低维任务保证最优收敛，高维任务克服ELM的不稳定性；虽在特定任务略逊于动态网络，但整体展现了无求逆求解器的鲁棒性与实用性．

3.4 分类实验结果统计分析

为客观评估不同算法的性能，对15个UCI基准数据集与4个MedMNIST医学图像数据集的训练结果，参考文献[24]，采用平均排名、Friedman检验和Nemenyi检验进行显著性分析，结果见表6．

表6 不同算法在19个数据集中训练的平均排名

Table 6 Average rankings of various algorithms on 19 datasets

算法	平均排名	算法	平均排名
RBF-ELM	2.16	ELM	6.29
Ensemble-ELM	2.97	ADMM-ELM	6.16
Q-ELM	4.16	ADAM-ELM	6.47
RELM	5.58	WELM	7.00
OS-ELM	5.79	I-ELM	8.42

由表6可见，RBF-ELM和Ensemble-ELM算法凭借核函数和集成学习的结构优势占据了前两名．在SLFN类算法中，Q-ELM的平均排名最高(4.16)，显著优于第4名RELM(5.58)，展现了Q-ELM算法的综合性能优势．

采用Friedman检验验证上述排名差异的统计学意义．假设原假设H_0为所有算法性能相同，根据Friedman统计量计算公式

$$\chi_{\mathrm{F}}^2=\frac{12N_{\mathrm{d}}}{k(k+1)}\left[\sum_l R_l^2-\frac{k(k+1)^2}{4}\right] \tag{16}$$

其中，χ_{F}^2是Friedman检验的统计量，该统计量在N_{d}较大时近似服从自由度为$k-1$的卡方分布；数据集数N_{d}为19；算法数k为10；R_l为第l个算法的平均排名．由式(16)计得$\chi_{\mathrm{F}}^2=66.80$，从而得到Iman-Davenport统计量为

$$F_{\mathrm{F}}=\frac{(N_{\mathrm{d}}-1)\chi_{\mathrm{F}}^2}{N_{\mathrm{d}}(k-1)-\chi_{\mathrm{F}}^2}=11.54 \tag{17}$$

故在(9，162)的F值分布下计算出p值为6.75×10^{-14}．由于$p\ll0.05$，故拒绝原假设，认为10种算法存在显著的性能差异．

鉴于Friedman检验拒绝了原假设，进一步采取Nemenyi检验，使用式(18)计算临界差值(critical difference，CD)以判断任意两算法之间是否存在显著差异．当两个算法的平均排名差值大于CD时，视为差异显著．

$$D_{\mathrm{C}}=q_{0.05}\sqrt{\frac{k(k+1)}{6N_{\mathrm{d}}}} \tag{18}$$

其中，$q_{0.05}$为Studentized range统计量的临界值．当显著性水平为0.05时，$q_{0.05}\approx3.102$，则有$D_{\mathrm{C}}=$

http://journal.szu.edu.cn

3.047 1，故绘制的临界差值图如图3. 图3上方的横线为统计学上的临界差异阈值，被粗横线覆盖的算法表示它们的性能差异在 0.05 的显著性水平下不显著. 由图3可见，Q-ELM的评价排名处于常规ELM算法类和复杂结构ELM算法之间，Q-ELM凭借基于强化学习的权重寻优策略，取得了单一前馈神经网络中的最佳平均排名.

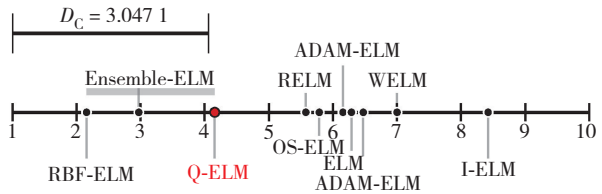


图3　临界差值图

Fig. 3　Critical difference diagram.

3.5　大规模数据集分类实验

为验证在超大规模数据集上Q-ELM算法是否会出现性能瓶颈，并探索改进方向，选择Magic、Adult、Shuttle和Skin-Segmentation四个数据集样本数为1.9万～24.5万的超大规模数据集，使用SLFN类固定隐藏层的算法并保持原参数设置进行实验，结果见表7. 由表7可见，相比基础ELM算法，Q-ELM算法的准确率出现了一定程度的下降，如在Magic数据集上，Q-ELM算法的准确率为84.51%，与ELM(85.71%)的差距为1.20%；在Adult数据集上，两者的准确率差距缩小至约0.11%，而在Shuttle和Skin-Segmentation数据集上准确率的差距则分别为1.3%和1.7%. 在上述4个数据集上，Q-ELM的训练时间明显高于表中的ELM及其他变体. 尽管实验验证了Q-ELM算法在大规模数据上的收敛性，但其在绝对精度和速度上仍有改进空间. 分析实验现象，认为原因主要在于预设的离散动作空间限制了高维海量数据下的寻优分辨率，且逐列求解机制带来了额外的时间开销. 后续研究可在奖励函数优化、连续动作空间优化以及利用图形处理器(graphics processing unit，GPU)多线程技术设计并行智能体等方面深入探索，使算法在该规模数据集上有更好的性能表现.

表7　不同算法在4个超大规模数据集上的分类

Table 7　Classification performance of different algorithms on four large-scale datasets

算法	Magic		Adult		Shuttle		Skin-Segmentation	
	准确率 / %	耗时 / s	准确率 / %	耗时 / s	准确率 / %	耗时 / s	准确率 / %	耗时 / s
Q-ELM	84.51	5.37	85.16	5.51	98.35	7.04	98.04	30.39
ELM	85.71	0.04	85.27	0.15	99.68	0.21	99.77	0.90
RELM	85.48	0.05	85.38	0.06	99.66	0.10	99.65	0.35
ADMM-ELM	85.61	0.05	85.31	0.08	99.70	0.09	99.76	0.33
ADAM-ELM	79.44	0.07	79.65	0.04	91.53	0.10	91.23	1.15
WELM	84.33	0.04	80.04	0.06	98.78	0.09	99.40	0.31

注：灰底数据为对应指标下的最优值.

结　语

提出一种基于强化学习求解输出权重的Q-ELM算法. 该算法不依赖传统Moore-Penrose伪逆解析求解路径，将权重优化建模为马尔可夫决策过程，利用智能体的“探索-利用”机制迭代寻优. 通过引入权重截断机制，Q-ELM算法实现了隐式硬约束正则化，有效抑制噪声放大，提升了模型的数值稳定性与鲁棒性. 在sinc函数拟合、UCI基准及MedMNIST医学图像分类等广泛实验中，Q-ELM算法展现出卓越性能，特别是在小样本及高维共线性任务中，有效克服了梯度算法易陷局部极小值及ELM受数值误差影响的缺陷. 统计检验表明，Q-ELM在单隐层网络算法中平均排名最佳，显著优于ELM、RELM及OS-ELM等主流变体，分类性能逼近复杂结构模型表现.

虽然Q-ELM在常规实验中取得了良好的性能表现，但在超大规模的样本数据集场景下，模型的预测性能和计算速度仍存在一定的瓶颈，后续研究将探索奖励函数优化，并行化处理等方向，进一步提升算法的性能与计算效率.

基金项目：国家自然科学基金资助项目(61906015)；广西壮族自治区重点实验室开放课题基金资助项目(MIMS22-13)
作者简介：朱丰君(zhufj_0116@163.com)，北京理工大学硕士研究生. 研究方向：机器学习与极限学习机.
引　　文：朱丰君，邹伟东，钟必能. 基于Q学习的极限学习机[J]. 深圳大学学报理工版，2026，43(3)：357-364.

http://journal.szu.edu.cn

参考文献 / References:

[1] HUANG Guangbin, ZHU Qinyu, SIEW C K. Extreme learning machine: theory and applications [J]. Neurocomputing, 2006, 70(1/2/3): 489-501.

[2] 北京理工大学. 一种基于变步长ADMM算法极限学习机的图像分类方法：CN202311033499.1[P]. 2023-10-20.
Beijing Institute of Technology. An image classification method based on variable step-size ADMM-ELM: CN202311033499.1[P]. 2023-10-20. (in Chinese)

[3] XIAO Xiao, LIAO Bolin, LONG Qiuqing, et al. Improved extreme learning machine based on deep learning and its application in handwritten digits recognition [C]// The 11th International Conference on Intelligent Control and Information Processing. Piscataway, USA: IEEE, 2021: 360-366

[4] 邹伟东，张百海，姚分喜，等. 基于改进型极限学习机的日光温室温湿度预测与验证[J]. 农业工程学报，2015，31(24)：194-200.
ZOU Weidong, ZHANG Baihai, YAO Fenxi, et al. Verification and forecasting of temperature and humidity in solar greenhouse based on improved extreme learning machine algorithm [J]. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(24): 194-200. (in Chinese)

[5] 王之琼，刘红艳，肖静，等. 基于ELM的室性早搏检测算法[J]. 计算机研究与发展，2013，50(增刊1)：196-204.
WANG Zhiqiong, LIU Hongyan, XIAO Jing, et al. The ELM-based algorithm on the detection of premature ventricular contraction [J]. Journal of Computer Research and Development, 2013, 50(Suppl. 1): 196-204. (in Chinese)

[6] HORATA P, CHIEWCHANWATTANA S, SUNAT K. Robust extreme learning machine [J]. Neurocomputing, 2013, 102: 31-44.

[7] DENG Wanyu, ZHENG Qinghua, CHEN Lin. Regularized extreme learning machine [C]// IEEE Symposium on Computational Intelligence and Data Mining. Piscataway, USA: IEEE, 2009: 389-395.

[8] LI Qiude, XIONG Qingyu, JI Shengfen, et al. A method for mixed data classification base on RBF-ELM network [J]. Neurocomputing, 2021, 431: 7-22.

[9] HUANG Guangbin, CHEN Lei, SIEW C K. Universal approximation using incremental constructive feedforward networks with random hidden nodes [J]. IEEE Transactions on Neural Networks, 2006, 17(4): 879-892.

[10] YE Yibin, QIN Yang. QR factorization based Incremental Extreme Learning Machine with growth of hidden nodes [J]. Pattern Recognition Letters, 2015, 65: 177-183.

[11] XU You, SHU Yang. Evolutionary extreme learning machine -based on particle swarm optimization[C]// The 3rd International Symposium on Neural Networks. Berlin: Springer, 2006: 644-652.

[12] YAO Lei, XU Shiming, XIAO Yanqiu, et al. Fault identification of lithium-ion battery pack for electric vehicle based on GA optimized ELM neural network [J]. IEEE Access, 2022, 10: 15007-15022.

[13] 席磊，何苗，周博奇，等. 基于改进多隐层极限学习机的电网虚假数据注入攻击检测[J]. 自动化学报，2023，49(4)：881-890.
XI Lei, HE Miao, ZHOU Boqi, et al. Research on false data injection attack detection in power system based on improved multi layer extreme learning machine [J]. Acta Automatica Sinica, 2023, 49(4): 881-890. (in Chinese)

[14] LIU Nan, WANG Han. Ensemble based extreme learning machine [J]. IEEE Signal Processing Letters, 2010, 17(8): 754-757.

[15] SUTTON R S, BARTO A G. Reinforcement learning: an introduction [M]. Cambridge, USA: MIT Press, 1998.

[16] BERTSEKAS D P, TSITSIKLIS J N. Neuro-dynamic programming [M]. Belmont, USA: Athena Scientific, 2011.

[17] OBBINS R, MONRO S. A stochastic approximation method [J]. The Annals of Mathematical Statistics, 1951, 22(3): 400-407.

[18] WATKINS C J C H, DAYAN P. Q-learning [J]. Machine Learning, 1992, 8(3): 279-292.

[19] VAPNIK V N. Statistical learning theory [M]. New York, USA: Wiley, 1999.

[20] BARTLETT P L. The sample complexity of pattern classification with neural networks: the size of the weights is more important than the size of the network [J]. IEEE Transactions on Information Theory, 1998, 44(2): 525-536.

[21] University of California, Irvine, School of Information and Computer Sciences. UCI Machine Learning Repository [DB/OL]. [2025-08-01]. Irvine, USA: University of California. https://archive. ics. uci. edu/.

[22] 刘知音. 基于交替方向乘子算法的L_1正则化极限学习机的算法研究[D]. 北京：北京化工大学. 2018.
LIU Zhiyin. Research on L_1 regularized extreme learning machine based on alternative direction method of multipliers [D]. Beijing: Beijing University of Chemical Technology, 2018. (in Chinese)

[23] YANG Jiancheng, SHI Rui, NI Bingbing. MedMNIST classification decathlon: a lightweight AutoML benchmark for medical image analysis [C]// The 18th International Symposium on Biomedical Imaging. Piscataway, USA: IEEE, 2021: 191-195.

[24] DEMŠAR J. Statistical comparisons of classifiers over multiple data sets [J]. The Journal of Machine Learning Research, 2006, 7: 1-30.

【中文责编：英子；英文责编：木柯】

http://journal.szu.edu.cn