

极限学习机的预测精度受到隐藏层节点数目、隐藏层的任意生成的输入参数和数据噪声的影响. 这种不更新隐藏层参数, 通过最小二乘调整的输出权重使极限学习机的抗错能力较差, 容易夸大离群点和噪声的影响得到不准确的结果. 在一些应用中, 针对极限学习机隐藏层节点过多的缺陷, 在隐藏层中增加了一类分类神经元^[2]. 或者通过粒子群算法优化选择极限学习机的隐藏层偏置, 验证了粒子群极限学习机算法在隐含层节点数目选择上具有优势^[3].

现实生活中存在许多与机器学习方法的应用条件不一致的情况, 因此对大多数传统算法进行改进以适应当前情况是正常的. 王莉等^[4]在代价敏感的理论基础上, 提出一种新的基于代价敏感集成学习的非平衡数据分类方法. 郑仙花等^[5]通过进化学习改进了克隆选择算法实现了多类监督分类, 避免了只能针对某一类样本数据进行监督学习. 沈宋衍等^[6]基于在线回归学习提出一种轮廓跟踪算法, 解决了目标快速运动以及严重形变导致跟踪失败的问题. 王英博等^[7]提出采用修正型果蝇优化算法优化广义回归神经网络进行参数优化. 蒙凯等^[8]基于集成问题的离散特征, 提出面向多目标优化的改进灰狼算法. 赵燕伟等^[9]以关联函数为基础, 重新定义神经网络中的误差计算方法, 构建了一种基于改进 BP 神经网络的可拓分类器.

由于极限学习机的学习效果依赖于初始的隐藏层输入权值和偏置. 本研究认为极限学习机可以利用模拟退火算法不断尝试隐藏层输入权值和偏置的选择, 提升预测能力. 首先, 传统的极限学习机对训练集的学习过程, 可以得到一组隐藏层的输入权值、偏置、输出权值和均方根误差. 然后把得到的隐藏层输入权值和偏置作为初始解, 均方根误差视为目标函数, 通过模拟退火过程, 找到训练过程均方根误差最小的极限学习机的隐藏层输入权值和偏置, 再通过传统的极限学习机计算得到隐藏层输出权值. 最后文本为了测试改进后的极限学习机的预测能力, 选取了鸢尾花分类数据和波士顿房价预测数据分别进行了分类和回归实验, 实验结果表明基于模拟退火改进的极限学习机在分类和回归的预测能力上优于传统的极限学习机.

1 基于 SA 改进的 ELM 预测模型

1.1 极限学习机

极限学习机作为单隐层神经网络, 与传统的神经

网络相比, 优势在于收敛速度快、泛化能力强, 并且避免了反向传播神经网络易陷入局部最优, 由于迭代, 训练过程十分耗时等特点^[10]. 极限学习机任意初始化输入权重和偏置, 通过计算隐藏层神经元的输出权值, 加快了极限学习机的学习速度. 根据线性方程组的求解方法可知, 当样本隐藏层神经元输出值矩阵是满秩时, 只需要矩阵求逆这一次性的操作, 就可以得到隐藏层神经元权重. 这一过程恰好可以学习不同的观测样本. 极限学习机的学习算法如下所示:

对于 N 个随机样本 (x_i, y_i) , 其中 $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$. 一个具有 L 个隐藏层神经元, 激活函数为 $g(x)$ 的极限学习机可以表示为 $\sum_{j=1}^L \beta_j g(W_j * x_i + b_j) = o_i$, $i = 1, \dots, N$, $W_j = [w_{j1}, w_{j2}, \dots, w_{jn}]^T \in R^n$, β_j 为输出权重, b_j 是第 j 个隐藏层神经元的偏置, $W_j * x_i$ 表示 W_j 和 x_i 的内积. 极限学习机训练样本的目的是使输出的误差小, 即 $\sum_{i=1}^N \|o_i - y_i\| = 0$, 也就是存在 β_j , W_j 和 b_j , 使 $\sum_{j=1}^L \beta_j g(W_j * x_i + b_j) = y_i$, $i = 1, \dots, N$, 矩阵形式表示为: $H\beta = Y$, 其中 H 是隐藏层神经元输出, β 是输出权重, Y 为样本观测值.

$$H(W_1, \dots, W_L, b_1, \dots, b_L, x_1, \dots, x_N) = \begin{bmatrix} g(W_1 * x_1 + b_1) & \dots & g(W_L * x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(W_1 * x_N + b_1) & \dots & g(W_L * x_N + b_L) \end{bmatrix}_{N \times L}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m}, Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N \times m}$$

通过解方程得到 $\hat{\beta}$, $\hat{\beta} = H^+ Y$, H^+ 是矩阵 H 的 Moore-penrose 广义逆. 为解决隐藏层输入权值和偏置的任意性, 周华平等^[11]结合改进鱼群算法对极限学习机的隐藏层偏置进行优化, 提出了乳腺肿瘤诊断的方法. 赵建堂^[12]采用有序加权平均算子融合单一极限学习的输出信息, 使分割式极限学习机的输出结果更为稳定.

1.2 模拟退火算法

模拟退火算法作为一种迭代自适应启发式概率性搜索算法, 模拟了一个高温固体的退火过程, 将优化过程分成加温、等温、冷却等 3 个部分, 利用 Metropolis 算法适当的控制温度下降过程. Metropolis 准则是模拟退火算法收敛于全局最优解的关键所在, 它以一定的

概率接受恶化解,这就使算法跳离局部最优的陷阱^[13]. 具体来说,模拟退火算法通过迭代的方式尝试改进目标函数的最优解,改进的新解将被接受为最优解,当新解劣于当前最优解,由波尔茨曼概率确定一个劣解的概率,使目标函数避免局部最优,最终获得全局最优解^[14]. 模拟退火算法的具体实现步骤算法 1 所示.

算法 1. 模拟退火算法

- 1) 初始化,取足够大的初始温度 T_0 ,令 $T=T_0$,任意选取初始解 S_1 ,确定每个温度下的迭代次数 L 和结束温度 T_{end} .
- 2) 对当前温度 T 和 $k=1,2,\dots,L$,重复第 3) 步至第 5) 步.
- 3) 随机产生一个新解 S_2 .
- 4) 计算 $\Delta f=f(S_2)-f(S_1)$,其中 $f(S_1)$, $f(S_2)$ 分别是 S_1 , S_2 的代价函数.
- 5) 若 $\Delta f<0$,则接受 S_2 作为新的当前解,令 $S_1=S_2$;否则计算 S_2 的接受概率 $\exp(-\Delta f/T)$,若产生 $(0,1)$ 区间的随机数 $rand$, $\exp(-\Delta f/T)>rand$,也接受 S_2 作为新的当前解, $S_1=S_2$;否则当前解 S_1 不变.
- 6) 降温速度为 q , $T=q*T$,若 T 小于 T_{end} ,则停止迭代,输出当前解,否则继续迭代.

在应用中模拟退火算法发挥了重要的作用,高鹰等^[15]提出一种基于模拟退火的粒子群优化算法,改善了粒子群优化算法摆脱局部极值点的能力,提高了算法的收敛速度和精度.杨若黎等^[16]提出一种确定模拟退火算法温度更新函数的启发式准则,数值计算结果表明采用新的温度更新函数以及相应的概率密度函数的模拟退火算法可以显著地提高求解全局优化问题的计算效率.张世睿等^[17]提出一种基于模拟退火算法的单隐藏层 BP 神经网络隐藏层节点估算算法,通过模拟退火不断增加隐藏层节点个数直至算法结束,得到最优解.凌静等^[18]用模拟退火算法改进遗传算法的变异操作,改善了遗传算法的早熟现象.黄联标等^[19]基于模拟退火算法对多工程系统维护时刻和维护方案进行寻优,确定各个工段最佳的预防性维护策略.

1.3 改进的极限学习机

由于极限学习机是任意给定隐藏层神经元的权值和偏置,这导致一些隐藏层神经元在训练过程中无效,使得极限学习机的泛化能力降低.由于 ELM 学习算法随机选择隐藏层神经元的输入权值和偏置,但是这些输入权值和偏置相对与输入数据来说,不是最优的选择,这使得极限学习机的泛化能力降低.在实际应用中,为了使神经网络有较好的泛化能力,需要较多的隐含层神经元,因而增加了网络的复杂度.罗庚合^[20]为了减少隐含层神经元个数、提高网络的泛化性能,引入可拓聚类算法,动态调整隐藏层节点数目.针对以上问题

本文提出基于模拟退火算法的极限学习机,利用模拟退火算法选择极限学习机的输入权值和偏置,从而得到一个最优的训练网络.

基于模拟退火算法改进的极限学习机算法算法 2 所示.

算法 2. 基于模拟退火算法改进的极限学习机算法

- 1) 任意生成输入权值 $W_{L \times N}^0$ 和偏置 b_L^0 ,极限学习机学习训练集,得到输出权值 β^0 ,均方根误差为预测评价标准,计算训练误差 $rmse^0$.
- 2) 设置初始温度 $T=T_0$,结束温度 T_{end} ,每个温度下的迭代次数 L ,降温速度 q ,令 $k=0$.
- 3) 任意生成输入权值 $W_{L \times N}$ 和偏置 b_L ,并且 $W_{L \times N} \neq W_{L \times N}^0$, $b_L \neq b_L^0$.极限学习机学习训练集,得到输出权值 β .设置代价函数即均方根误差 $rmse$, $k=k+1$.
- 4) 若 $\Delta f=rmse-rmse^0$, $\Delta f<0$ 成立,则 $W_{L \times N}^0=W_{L \times N}$, $b_L^0=b_L$;若不成立,则计算输入权值 $W_{L \times N}$ 和偏置 b_L 的接受概率 $\exp(-\Delta f/T)$,并且生成一个 $(0,1)$ 区间的随机数 $rand$,若 $\exp(-\Delta f/T)>rand$ 成立,则 $W_{L \times N}^0=W_{L \times N}$, $b_L^0=b_L$;否则返回第 3) 步.
- 5) 若 $k>L$ 成立,则 $T=T*q$, $k=0$,否则返回第 3) 步.
- 6) 若 $T=T_{end}$ 不成立,则返回第 3) 步,若成立,则算法结束,得到最优的输入权值 $W_{L \times N}^0$ 和偏置 b_L^0 .

基于模拟退火算法改进的极限学习机,结构复杂不便理论分析,若要了解算法的收敛性,可采用数值实验的方法.计算该算法的目标值与问题已有最优值之比,利用概率统计的方法考察所得比值与 1 的接近程度,比值越接近于 1,说明算法性能越好^[21].

2 实验分析

为了说明基于模拟退火算法改进的极限学习机的收敛性和泛化效果,文本选取了鸢尾花分类样本和波士顿房价预测两个数据集分别进行定性预测和定量预测两组实验.将基于模拟退火算法改进的极限学习机的预测结果与极限学习机等其他方法的预测结果进行对比.两组实验参数如表 1 所示.

表 1 参数设置

| 初始温度 | 降温速度 | 结束温度 | 每个温度下的迭代次数 |
|------|------|------|------------|
| 1 | 0.9 | 0.1 | 10 |

2.1 鸢尾花分类

本研究采用的鸢尾花数据包含 4 个解释变量分别是萼片的长度、萼片的宽度、花瓣的长度、花瓣的宽度,被解释变量即鸢尾花的种类.在这 150 条数据中,包含了 3 种鸢尾花,分别为 setosa、versicolor、virginica,每种花各有 50 条数据.实验按 2:1 的比例将数据随机地划分成训练数据集和测试数据集.研究选取精度为

衡量预测准确性的标准,对实验结果进行分析,精度计算公式如表2所示.

表2 精度计算公式

| 真实值 | 估计值 | |
|-----|---------|---------|
| | 正例 | 反例 |
| 正例 | TP(真正例) | FN(假反例) |
| 反例 | FP(假正例) | TN(真反例) |

精度=(TP+TN)/(TP+TN+FP+FN)

任意选取一次实验进行观测,结果如图1所示,基于模拟退火改进的极限学习机在降温的过程中,分类预测误差得到优化,未改进的极限学习机的预测精度是88%,改进后的极限学习机的预测精度将近98%,预测精度提高了10%.

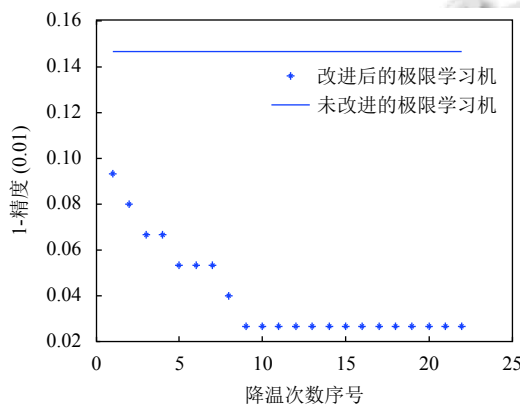


图1 改进的极限学习机分类的优化过程

实验对每种分类方法的50次分类精度取平均值,再进行比较.实验发现极限学习机的分类精度是6种方法里最低的,只有83.2%,不足90%,其余的分类方法的预测精度都高于90%.其中,本文方法的预测精度最高,达到99.0%,可见本文方法可以极大的提高极限学习机的分类预测精度.如表3所示

表3 鸢尾花分类结果

| 方法 | 平均分类精度 (%) |
|---------------|------------|
| 本文方法 | 99.0 |
| 极限学习机 | 83.2 |
| 基于遗传算法的BP神经网络 | 96.0 |
| BP神经网络 | 94.6 |
| 分类树 | 94.2 |
| 支持向量机 | 96.2 |

通过对50次实验结果取平均值,由表4发现随着鸢尾花样本数据规模越大,比值越接近1,说明样本数据越大,改进后的极限学习机分类性能越好.

表4 鸢尾花实验收敛趋势

| 训练样本/ 测试样本 | 改进的极限学习机分类精度/极限学习机分类精度 | | |
|---------------|------------------------|------|------|
| | L=10 | L=20 | L=30 |
| 25/12 | 1.25 | 1.25 | 1.25 |
| 50/25 | 1.22 | 1.22 | 1.23 |
| 75/37 | 1.21 | 1.22 | 1.23 |
| 100/50 | 1.19 | 1.19 | 1.19 |

2.2 波士顿房价预测

本研究选取的波士顿房价数据集共有506条数据,每一条数据包括13个解释变量和1个被解释变量.解释变量包括城镇人均犯罪率、住宅用地所占面积、城镇中非商业用地的所占面积、是否邻近查尔斯河、环保指标、每栋住宅的房间数、1940年以前建成的自住建筑的比例、距离波士顿就业中心的加权距离、距离高速公路的便利指数、每万美元的不动产税率、城镇中教师学生的比例、城镇中黑人比例、房东是低收入阶层的比例.被解释变量即自住房屋房价的中位数.实验按2:1的比例将数据随机地划分成训练数据集和测试数据集,选择均方根误差为预测误差的评价标准,

$$rmse = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}$$

其中, m 是预测样本数, y_i 是实际值, \hat{y}_i 是预测值.

任意选取一次实验进行观测,结果如图2~图4所示.其中图2说明了未改进的极限学习机在测试集上的估计值和真实值的对比.图3说明了改进后的极限学习机的在测试集上的估计值和真实值的对比.

从图2中可知,未改进的极限学习机在测试集上的一些预测值比实际值偏大.从图3中可知,改进后的极限学习机在测试集上的预测值相对平稳,波动较为平缓.图4表明基于模拟退火改进的极限学习机的优化过程,均方根误差在优化过程中减小.

实验对每种分类方法的50次预测误差取平均值,再进行比较,实验结果如表5所示.实验发现极限学习机的预测误差相比于其他传统机器学习方法偏高.本文方法可以提高极限学习机的回归预测能力.由表6发现波士顿房价样本数据规模越大,比值在1附近有微小的变动,说明改进后的极限学习机回归性能稳定.

3 结论与展望

为了提高极限学习机的分类和回归的预测能力,提出一种基于模拟退火改进的极限学习机.本文利用

模拟退火算法的降温过程对隐藏层的输入权值和偏置进行优化, 避免了任意选择的输入权值和偏置使训练的模型无效的情况, 使极限学习机的表现更加稳定. 实

验结果表明通过模拟退火算法改进的极限学习机分类预测能力极好, 回归预测能力劣于 BP 神经网络和回归树的预测能力.

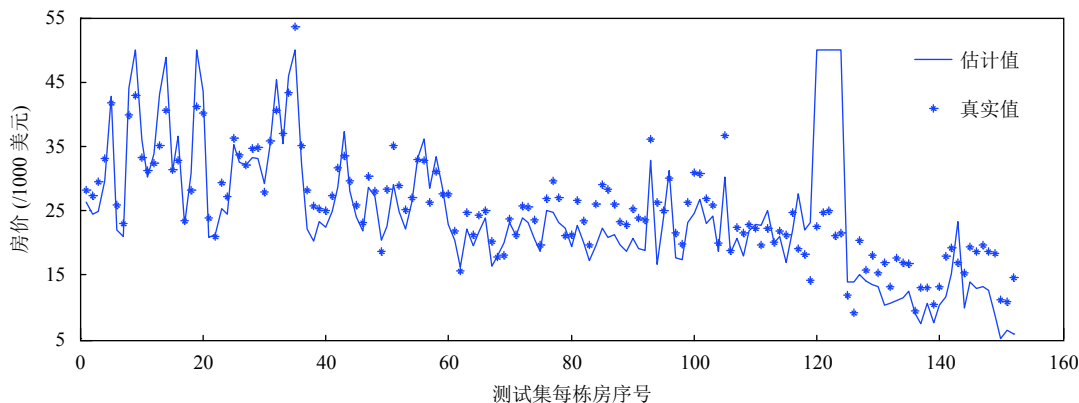


图2 极限学习机的波士顿房价预测结果

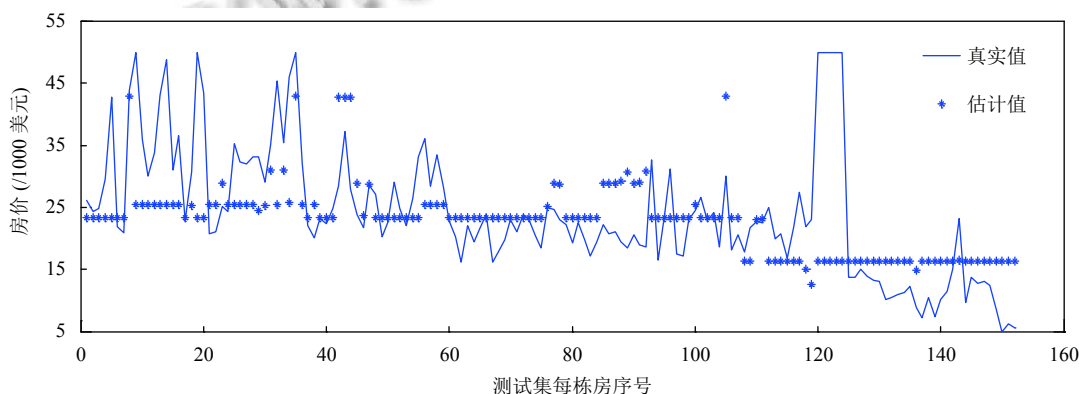


图3 改进的极限学习机的波士顿房价预测结果

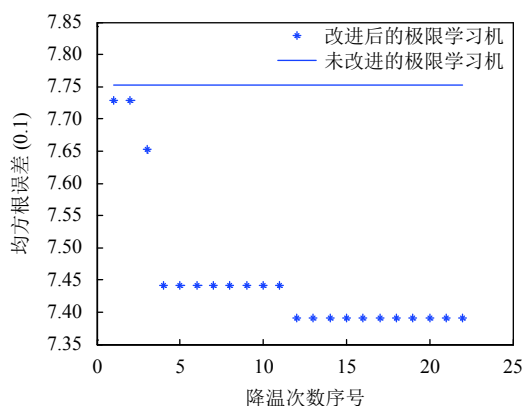


图4 改进的极限学习机回归的优化过程

下一步工作考虑改进的极限学习机在降温优化过程中如何选择最优的迭代次数、在优化时选择最优的隐藏层神经元个数以及如何进一步提高极限学习机的回归预测能力.

表5 波士顿房价预测结果

| 方法 | 平均预测误差 (/1000 美元) |
|-----------------|-------------------|
| 本文方法 | 7.47 |
| 极限学习机 | 8.14 |
| 基于遗传算法的 BP 神经网络 | 8.38 |
| BP 神经网络 | 5.50 |
| 回归树 | 4.75 |
| 支持向量机 | 12.32 |

表6 波士顿房价实验收敛趋势

| 训练样本/ 测试样本 | 改进的极限学习机回归精度/极限学习机回归精度 | | |
|---------------|------------------------|------|------|
| | L=10 | L=20 | L=30 |
| 88/32 | 0.97 | 0.97 | 0.97 |
| 177/63 | 1.02 | 0.97 | 0.98 |
| 265/94 | 0.99 | 0.96 | 1.00 |
| 354/152 | 0.92 | 0.88 | 0.91 |

参考文献

1 Tang JX, Deng CW, Huang GB. Extreme learning machine

- for multilayer perceptron. *IEEE Transactions on Neural Networks and Learning Systems*, 2016, 27(4): 809–821. [doi: 10.1109/TNNLS.2015.2424995]
- 2 常玉清, 李玉朝, 王福利, 等. 基于极限学习机的生化过程软测量建模. *系统仿真学报*, 2007, 19(23): 5587–5590. [doi: 10.3969/j.issn.1004-731X.2007.23.057]
 - 3 王杰, 毕浩洋. 一种基于粒子群优化的极限学习机. *郑州大学学报(理学版)*, 2013, 45(1): 100–104.
 - 4 王莉, 陈红梅, 王生武. 新的基于代价敏感集成学习的非平衡数据集分类方法 NIBoost. *计算机应用*, 2019, 39(3): 629–633. [doi: 10.11772/j.issn.1001-9081.2018071598]
 - 5 郑仙花, 骆炎民. 基于多类数据分类的改进克隆选择算法. *计算机应用*, 2012, 32(11): 3201–3205.
 - 6 沈宋衍, 陈莹. 基于在线回归学习的轮廓跟踪算法. *计算机工程*, 2016, 42(5): 230–234. [doi: 10.3969/j.issn.1000-3428.2016.05.039]
 - 7 王英博, 聂娜娜, 王铭泽, 等. 修正型果蝇算法优化 GRNN 网络的尾矿库安全预测. *计算机工程*, 2015, 41(4): 267–272. [doi: 10.3969/j.issn.1000-3428.2015.04.051]
 - 8 蒙凯, 唐秋华, 张子凯, 等. 基于改进多目标灰狼算法的装配线平衡与预防维护集成优化. *计算机集成制造系统*. <https://www.cnki.net/KCMS/detail/11.5946.tp.20190812.1025.016.html>, [2019-08-18].
 - 9 赵燕伟, 任设东, 陈尉刚, 等. 基于改进 BP 神经网络的可拓分类器构建. *计算机集成制造系统*, 2015, 21(10): 2807–2815.
 - 10 陈于思, 孙林夫. 基于差分进化和极限学习机的并发查询性能预测. *计算机集成制造系统*, 2019, 25(9): 2291–2304.
 - 11 周华平, 袁月. 改进鱼群算法优化的 ELM 在乳腺肿瘤辅助诊断中的应用研究. *计算机工程与科学*, 2017, 39(11): 2145–2152. [doi: 10.3969/j.issn.1007-130X.2017.11.025]
 - 12 赵建堂. 大数据分割式极限学习机算法. *计算机工程与应用*, 2019, 55(10): 73–76. [doi: 10.3778/j.issn.1002-8331.1801-0217]
 - 13 郁磊, 史峰, 王辉, 等. MATLAB 智能算法 30 个案例分析. 2 版. 北京: 北京航空航天大学出版社, 2015. 178–187.
 - 14 Mafarja MM, Mirjalili S. Hybrid Whale Optimization Algorithm with simulated annealing for feature selection. *Neurocomputing*, 2017, 260: 302–312. [doi: 10.1016/j.neucom.2017.04.053]
 - 15 高鹰, 谢胜利. 基于模拟退火的粒子群优化算法. *计算机工程与应用*, 2004, 40(1): 47–50. [doi: 10.3321/j.issn:1002-8331.2004.01.014]
 - 16 杨若黎, 顾基发. 一种高效的模拟退火全局优化算法. *系统工程理论与实践*, 1997, 17(5): 30–36.
 - 17 张世睿, 李心科. 基于模拟退火的 BP 网络隐藏层节点估算算法. *合肥工业大学学报(自然科学版)*, 2017, 40(11): 1489–1491, 1506.
 - 18 凌静, 江凌云, 赵迎. 结合模拟退火算法的遗传 K-Means 聚类方法. *计算机技术与发展*, 2019, 29(9): 61–65.
 - 19 黄联标, 敖银辉. 基于模拟退火算法的多工段系统预防维护策略. *计算机集成制造系统*. <https://www.cnki.net/KCMS/detail/11.5946.tp.20190426.1318.012.html>. [2019-04-29].
 - 20 罗庚合. 基于可拓聚类的极限学习机神经网络. *计算机应用*, 2013, 33(7): 1942–1945. [doi: 10.11772/j.issn.1001-9081.2013.07.1942]
 - 21 白丹宇. 流水车间与开放车间调度算法渐近分析. 北京: 清华大学出版社, 2015. 6–7.