



基于改进花朵授粉算法的极限学习机模型

邵良杉, 兰亭洋, 李臣浩

(辽宁工程技术大学 系统工程研究所, 辽宁 葫芦岛 125105)

摘 要: 为提高瓦斯突出风险预测的准确率和效率, 在极限学习机(ELM)模型的基础上构建预测模型 ACFPA-ELM。采用核线性鉴别分析(KLDA)对瓦斯突出样本数据进行特征抽取, 利用代价敏感思想修正 ELM 适应度函数, 同时将 Tent 混沌搜索和自适应算子引入花朵授粉算法(FPA)中, 优化 ELM 的初始输入权值和阈值, 从而提高对瓦斯突出风险的预测能力。实验结果表明, 相较于经典的 SVM、BP 和 ELM 单一预测模型以及改进的 FPA-ELM 和 PSO-ELM 复合预测模型, ACFPA-ELM 模型在瓦斯突出风险预测的准确率、预测一致性以及运行效率方面均具有明显的优势。

关键词: 瓦斯突出; 花朵授粉算法; 极限学习机; 核线性鉴别分析; 混沌映射

开放科学(资源服务)标志码(OSID):



中文引用格式: 邵良杉, 兰亭洋, 李臣浩. 基于改进花朵授粉算法的极限学习机模型[J]. 计算机工程, 2019, 45(12): 281-288.

英文引用格式: SHAO Liangshan, LAN Tingyang, LI Chenhao. Extreme learning machine model based on improved flower pollination algorithm[J]. Computer Engineering, 2019, 45(12): 281-288.

Extreme Learning Machine Model Based on Improved Flower Pollination Algorithm

SHAO Liangshan, LAN Tingyang, LI Chenhao

(System Engineering Institute, Liaoning Technical University, Huludao, Liaoning 125105, China)

[Abstract] In order to improve the accuracy and efficiency of gas outburst risk prediction, this paper proposes a prediction model ACFPA-ELM based on Extreme Learning Machine (ELM). First, this paper adopts Kernel Linear Discriminant Analysis (KLDA) to extract the features of gas outburst sample data. Then, this paper utilizes the cost sensitive ideas to modify ELM fitness function. At the same time, the Tent chaotic search and adaptive operator are introduced into the Flower Pollination Algorithm (FPA) to optimize the initial input weight and threshold of the ELM, thus improving the prediction ability for gas outburst risk. Experimental results show that, compared with the classic SVM, BP and ELM single prediction models, as well as the improved FPA-ELM and PSO-ELM composite prediction models, the proposed model is superior in the accuracy, consistency and efficiency of gas outburst risk prediction.

[Key words] gas outburst; Flower Pollination Algorithm (FPA); Extreme Learning Machine (ELM); Kernel Linear Discriminant Analysis (KLDA); chaos mapping

DOI:10.19678/j.issn.1000-3428.0053285

0 概述

瓦斯突出作为煤矿作业中的一种地质灾害, 是多种复杂因素共同作用的结果, 其具有极大的破坏性, 严重威胁煤矿的安全生产。目前对瓦斯突出危险性预测的主要方法有支持向量机^[1]、综合指标法^[2]、模糊模式识别^[3]、灰色理论^[4]和神经网络等。

神经网络作为一种智能学习算法, 能够解决非线性、高维变换的问题。文献[5]提出 PCA-BP 神经网络预测模型, 但由于 BP 网络训练参数过多, 因此模型精度不高、泛化性差。文献[6]提出 KPCA-ELM 预测模型, 用极限学习机(Extreme Learning Machine, ELM)替换 BP 网络, 但由于其随机选取 ELM 的初始参数, 会导致模型预测结果不唯一。针对此问题, 文

基金项目: 国家自然科学基金(71371091)。

作者简介: 邵良杉(1961—), 男, 教授、博士、博士生导师, 主研方向为数据挖掘、矿业系统工程; 兰亭洋、李臣浩, 硕士。

收稿日期: 2018-12-03 修回日期: 2019-01-15 E-mail: abc551218@126.com

献[7-8]分别采用人工蜂群算法和粒子群优化(Particle Swarm Optimization, PSO)算法优化 ELM,但是这 2 种算法都存在易陷入局部最优解、参数过多、执行过程复杂的问题。此外,无监督 KPCA 算法在特征提取时未能充分利用样本的标签项,导致特征提取后同类样本的距离变大。

花朵授粉算法(Flower Pollination Algorithm, FPA)^[9]作为一种新型元启发式优化算法,具有参数少、易实现的特点。FPA 算法利用概率 $p(0 < p < 1)$ 按一定比例进行异花授粉和自花授粉,体现了 Levy 飞行策略,在一定程度上解决了输出结果容易陷入局部最优解的问题,相比 PSO 和遗传算法有更强的寻优能力^[9]。

本文构建 ACFPA-ELM 模型对瓦斯突出风险进行预测。通过引入代价敏感的思想修正 ELM 模型的适应度函数,并采用改进的 FPA 算法对其输入权值和阈值进行优化,从而得到最优的 ELM 瓦斯风险预测模型。此外,利用 KLDA 算法抽取瓦斯突出数据的特征,以解决 PCA 和 KPCA 算法在特征抽取时忽略标签项的问题,提高了特征提取的可分性。

1 ACFPA-ELM 预测模型

ELM^[10-11]是一种单隐层前馈神经网络学习算法。与其他神经网络相比,ELM 优势在于不需要预设大量的训练参数以及训练中间层偏置和输入权值^[12],并且输出的最优解唯一,解决了神经网络训练过程复杂、易陷入局部最优解和过分拟合等问题^[13]。但输入权值和阈值的随机选取在一定程度上限制了 ELM 的精度,其结果会产生一定的波动性。特别地,当输入权值和阈值出现零值时,部分节点失去作用,从而影响 ELM 的预测准确率。

虽然 FPA 算法相比 PSO 和遗传算法有更好的寻优能力,但与其他种群优化算法一样,其也存在易陷入局部最优、收敛速度过慢的问题。许多研究人员对其进行改进,以进一步提高算法性能。文献[14]在算法本身的 Levy 飞行中引入万有引力的思想,提高配子更新的效率。文献[15]将量子系统的态叠加特性融入算法,用波函数描绘种群个体的位置,增强异花授粉效率。文献[16]利用高斯变异对配子更新过程进行扰动从而增加了种群的多样性。但上述方法大多存在改进角度单一、时间空间复杂度过大的问题。因此,本文从多角度对 FPA 进行改进,提出 ACFPA 算法优化 ELM 的输入权值和阈值以提高其精度。

1.1 ACFPA 算法设计

1.1.1 FPA 算法

设花粉配子个数为 N ,目标函数维度为 d ,配子的当前位置为 $\mathbf{X}_i = [x_{i1}, x_{i2}, \dots, x_{id}] (i = 1, 2, \dots, N)$,配子局部搜索(自花授粉)公式为:

$$\mathbf{X}_i^{t+1} = \mathbf{X}_i^t + \varepsilon(\mathbf{X}_i^t - \mathbf{X}_k^t) \quad (1)$$

其中, \mathbf{X}_i^t 表示第 i 个配子在 t 时刻的位置, \mathbf{X}_i^{t+1} 表示第 i 个配子在 $t+1$ 时刻的位置, $\mathbf{X}_j^t, \mathbf{X}_k^t$ 为所有配子中

除第 i 个配子以外的任意 2 个配子在 t 时刻的位置, ε 是服从均匀分布的比例系数。配子的全局搜索(异花授粉)公式为:

$$\mathbf{X}_i^{t+1} = \mathbf{X}_i^t + L(\mathbf{X}_i^t - \mathbf{g}^*) \quad (2)$$

其中, L 为服从 Levy 分布的随机变量, \mathbf{g}^* 为全局最优解。 L 近似表示为:

$$L \sim \frac{\gamma \cdot \text{gam}(\gamma) \cdot \sin(\pi\gamma/2)}{\pi s^{1+\gamma}} \quad (3)$$

其中, $s \ll s_0 \ll 0$, $\text{gam}(\gamma)$ 为 gamma 函数。通常, $\gamma = 1.5$ 。

由 FPA 的授粉策略可知,在异花授粉阶段全程受到当前全局最优解的引导,而随着迭代次数的增加,大量粒子过分聚集,导致算法出现重叠现象,并陷入局部最优。相反,在自花授粉阶段,配子更新不受任何因素引导,导致配子更新出现明显的盲目性和随机性,从而大幅限制了算法收敛的速度。同时传统的 FPA 算法全局以固定比例进行异花授粉和自花授粉,不能根据不同阶段的实际情况动态分配两者的比重,降低了算法寻优的效率。此外, FPA 初始花粉配子位置的随机选取,在一定程度上也降低了 FPA 搜索的效率。上述缺陷限制了 FPA 的寻优精度和收敛速度,为此,本文借助 Tent 混沌搜索和自适应算子对其进行改进。

1.1.2 Tent 混沌映射

Tent 映射公式为:

$$x_{i+1} = \begin{cases} 2x_i, & 0 \leq x_i \leq \frac{1}{2} \\ 2(1-x_i), & \frac{1}{2} \leq x_i \leq 1 \end{cases} \quad (4)$$

Tent 映射的移位变换公式为:

$$x_{i+1} = (2x_i) \bmod 1 \quad (5)$$

反向解求解公式为:

$$ox_i = k(Du_i - Dl_i) - x_i \quad (6)$$

基于反向学习的混沌序列生成步骤为:

步骤 1 随机产生 $(0, 1)$ 的初值 $x_0, z_1 = x_0$ 。

步骤 2 根据式(5)生成一个 x 序列。

步骤 3 若 $x_i = \{0, 0.25, 0.5, 0.75\}$ 或 $x_i = x_i - s, s = \{0, 1, 2, 3, 4\}$, 则带入 $x_i = z_{j+1} = z_j + e$ (e 为 $0 \sim 1$ 的随机变量);否则执行步骤 2。

步骤 4 若满足迭代次数,则停止迭代返回 x 序列;否则转向步骤 2。

步骤 5 根据式(6)生成 x 序列的反向解,并计算全部解向量的适应度值,返回前 N 个最优的解向量。

1.1.3 Tent 混沌搜索

本文采用混沌搜索对 FPA 算法的异花授粉进行优化,其混沌搜索的步骤如下:

步骤 1 采用式(7)将 x_i 投影到 $(0, 1)$ 区间。

步骤 2 将式(7)代入式(4)进行迭代,产生混沌序列 $Z_k^d(0), (k = 1, 2, \dots, C_{\max}), C_{\max}$ 为 Tent 搜索的最大迭代次数。

步骤 3 采用式(8)将 $z_i(k)$ 载波到解空间。

步骤4 计算当前解的适应度, 并与历史最优适应度进行比较, 保留最优解。

步骤5 若达到最大迭代次数, 则返回最优解, 否则执行步骤2。

$$Z_k^d(0) = \frac{(X_k^d - X_{\min}^d)}{(X_{\max}^d - X_{\min}^d)} \quad (7)$$

$$V_k^d = X_k^d + \frac{(X_{\max}^d - X_{\min}^d)}{2} \times (2Z_k^d(t) - 1) \quad (8)$$

Tent 混沌映射在点 0.2、0.4、0.6、0.8 上会出现混沌吸引子, 并呈现小周期现象, 因此, 当 $x_i = \{0, 0.25, 0.5, 0.75\}$ 时, 利用公式 $x_i = z_{j+1} = z_j + e$ (e 为 0~1 的随机变量) 替换迭代初值, 增加混沌搜索的内随机性, 从而增加 Tent 混沌的混沌强度, 同时通过反向学习公式增加了 Tent 混沌的遍历性, 有助于提高解的质量和求解效率^[17]。

1.1.4 ACFPA 算法

本文利用基于反向学习的 Tent 混沌映射对配子群进行初始化。Tent 混沌映射能够运用其特有的非线性规律在目标空间中遍历所有可行解, 同时反向学习策略能够增加花粉配子的多样性, 采用基于反向学习混沌映射对 FPA 算法的配子进行初始化, 能使配子较为均匀地分布在解空间并增加配子的多样性, 提高初始配子的分布质量。具体实现步骤为:

- 1) 随机产生初始向量, 并根据式(1)和式(2)生成混沌解序列。
- 2) 根据式(3)生成对应的反向解。
- 3) 计算全部解序列的适应度值, 并将前 N 个最优的解序列作为 N 个花粉配子的初始位置。

在 FPA 的异花授粉策略上, 为避免大量配子过分聚集, 增强 FPA 算法跳出局部最优的能力, 本文在原有异花授粉的基础上, 通过小概率变异将混沌搜索引入 FPA 的异花授粉中。虽然混沌搜索能够提高异花授粉的搜索能力并且防止陷入局部最优, 但大范围的混沌搜索会增加每次搜索位置重叠的概率, 从而产生冗余搜索。因此, 本文设置变异阈值为 0.2, 使其以小概率在整个解空间上进行混沌搜索。改进的异花授粉步骤为:

- 1) 生成随机数 e 。
- 2) 若 e 大于阈值则跳转步骤3, 否则跳转步骤5。
- 3) 根据式(7)求得 x_i 在 (0, 1) 区间的对应的解 z_{i0} , 然后对 z_{i0} 进行混沌搜索并返回 z_{ik} 。
- 4) 将 z_{ik} 投影到源空间, 生成对应的新解 v_i , 令 $x_i = v_i$, 跳转步骤6。
- 5) 令 $X_i = X_i + L(X_i - g^*)$
- 6) 返回 X_i

本文在自花授粉策略中引入自适应算子, 使自花授粉的范围根据算法的不同阶段动态调整, 增加算法的寻优效率。改进的自花授粉策略为:

$$X_i^{t+1} = X_{\text{best}}^t + \alpha(X_r^t - X_i^t) \quad (9)$$

$$\alpha = \beta + (1 - \beta) \left(\frac{t_{\max} - t}{t_{\max}} \right) \quad (10)$$

其中, X_r^t 为 t 次迭代不为 X_i^t 的解, $\beta \in (0, 1)$, 一般为 0.1, α 为自适应算子。通过式(10)构造自适应算子 α , 使 α 在算法寻优的过程中逐步减小, 进而减小自花授粉中围绕全局最优解搜索的范围。此外算法后期 $t_{\max} - t$ 和 t_{\max} 的比值趋近于 0, 导致 α 同时趋于 0, 进而使自花授粉范围趋近于 0, $X_i^{t+1} = X_{\text{best}}^t$ 。为此, 利用参数 β 将 α 限制在 $(\beta, 1)$ 之间, 避免在算法后期出现上述问题。

为使算法在不同阶段合理分配自花授粉和异花授粉的比例, 本文通过二次函数关系式联系切换概率 p 和迭代次数 t , 使 p 随 t 的增加而指数增长, 增加自花授粉比例在算法后期的增长速度, 使算法的寻优在后期更具针对性。同时定义系数 ε ($\varepsilon \in [0, 1]$), 使 p 在算法运行过程中不会超过 ε , 限制区间上限, 防止搜索后期异花授粉趋于 0。改进的切换概率 p 为:

$$p = \varepsilon \left(\frac{t}{t_{\max}} \right)^2 \quad (11)$$

其中, t 为算法当前迭代次数, t_{\max} 为算法最大迭代次数。

通过改进切换概率 p , 使算法在初期以较大概率执行异花授粉, 在整个解空间上搜寻全局最优解, 在后期逐渐缩小并锁定解的目标范围, p 值逐渐变大, 此时算法以较大的概率进行自花授粉, 在较小的范围内进行针对性寻优, 从而加快算法的收敛速度。

1.2 ACFPA-ELM 模型构建

1.2.1 基于代价敏感的适应度函数

传统的复合 ELM 模型将 0-1 函数作为适应度函数。然而, 将多节点的连续值转换成 0、1 二进制离散值, 无疑会丢失部分信息, 降低适应度函数对 ELM 模型的评级能力, 进而使 ELM 预测模型的收敛速度下降。为此, 本文参考 ELM 输出结构, 在原有适应度函数的基础上构造一种基于代价敏感的适应度函数, 进一步参考每个输出节点值的大小, 提高算法对 ELM 初始输入权值和阈值的优化速度。

ELM 在处理多分类问题时将输出节点中最大值所对应的节点序号作为模型的预测结果, 因此, 优化 ELM 的目标就是要尽可能地增大正确类别节点对应的输出值。传统的 0-1 适应度函数仅将所有样本中被正确分类的比例作为评价标准。为增强适应度函数对 ELM 模型的评价能力, 本文根据每个样本下正确节点对应输出值占总输出值的比例计算样本被分到正确类别的概率, 并使适应度函数和此概率成正相关。因此, 对于相同的分类结果, 改进的适应度函数根据每个样本被正确分类的概率给予每个分类结果不同的分类代价, 这将大幅加快 PSO、FPA 等优化算

法对神经网络参数的优化速度,并且在一定程度上降低分类模型的误差。改进的适应度函数为:

$$loss_p = \exp\left(-\frac{1}{N} \sum_{i=1}^N \frac{out_{right}}{\sum_{j=1}^k out_j}\right) \quad (12)$$

其中, N 表示样本个数, k 表示样本类别数, p 表示迭代次数。 out_{right} 表示每个样本分类结果中正确类别所对应输出的输出值, out_j 表示模型第 j 个输出的输出值。

1.2.2 ACFPA-ELM 实现步骤

ACFPA 算法优化 ELM 模型的具体步骤如下:

步骤 1 定义 FPA 相关参数,随机生成一个 d 维的花粉配子,然后利用基于反向学习的混沌映射对花粉配子进行初始化。

步骤 2 计算切换概率 p ,对于每一个配子而言,产生一个随机数 $e \in (0,1)$ 。如果 $e > p$,则进行自花授粉;否则异花授粉,更新配子位置信息。

步骤 3 若 $X_{id} > Upper_d$ (d 维空间上边界),则 $X_{id} = Upper_d$;若 $X_{id} < Low_d$ (d 维空间下边界),则 $X_{id} = Low_d$ 。

步骤 4 将当前配子位置作为 ELM 模型参数对瓦斯突出样本进行预测并将预测结果转化成适应度函数值,比较其与历史最优位置的适应度函数值,若优于历史最优值,则将历史最优值更新为当前位置,否则直接进行下一步。

步骤 5 将当前配子位置对应的适应度函数值和全局最优位置的适应度函数值进行比较,若优于全局最优值,则将全局最优值的位置更新为当前配子位置,否则直接进行下一步。

步骤 6 对所有配子执行以上步骤。

步骤 7 迭代以上步骤指定次数,将全局最优位置返回并作为 ELM 预测模型的初始输入权值和阈值。

2 KLDA 特征提取

线性鉴别分析(Linear Discriminant Analysis, LDA)^[18]是一种有监督的线性降维算法。LDA 特征抽取的 2 个基本原则是同类别的样本在样本空间中间隔尽可能地增加和不同类别的样本间距尽可能地减小,其本质是将样本空间映射到分类性能好的空间上。与其他谱方法类似, LDA 最终也归结为求矩阵特征值、特征向量的问题。相比于无监督的 PCA 降维算法, LDA 能充分利用瓦斯样本的标签项对样本空间进行转化。

核线性鉴别分析(Kernel Linear Discriminant Analysis, KLDA)结合核方法和 LDA 方法,首先将样本数据非线性地映射到高维空间,然后在高维空间使用 LDA 方法进行特征提取,从而间接地实现了空间非线性变换,提高了特征提取的有效性和可靠性。

对于一组总数为 N 、类别数为 C 的样本 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, KLDA 实现过程为:

步骤 1 将样本空间通过核函数映射到高维空间 Z ,映射为 $\Phi(x_j) = \Phi_j$,具体形式为:

$$\phi: x \in \mathbb{R}^d \rightarrow \phi(x) \in Z \quad (13)$$

步骤 2 在高维空间找到最佳投影方向,使样本数据的类间散度和类内散度的比值 $J(w)$ 最大化。

$$J(w) = \arg\max_w \frac{|w^T s_{bw}|}{|w^T s_{ww}|} \quad (14)$$

步骤 3 计算同类别的散度矩阵 D_s 和不同类别的散度矩阵 D_o 。

$$D_s = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{N_i} (y^{(i,j)} - y^{(i)}) (y^{(i,j)} - y^{(i)})^T \quad (15)$$

$$D_o = \frac{1}{N} \sum_{i=1}^C N_i \left(y^{(i)} - \frac{\sum_{j=1}^C y^{(j)}}{N} \right) \left(y^{(i)} - \frac{\sum_{j=1}^C y^{(j)}}{N} \right)^T \quad (16)$$

其中, $y^{(i,j)}$ 表示第 i 类中的第 j 个样本, $y^{(i)}$ 表示第 i 类样本的均值。

步骤 4 计算 $D_s^{-1} D_o$ 的前 k 个特征值和对应的特征向量 $w_i (i=1, 2, \dots, k)$, 得到投影矩阵 $W = \{w_1, w_2, \dots, w_k\}$ 。

步骤 5 将每个样本数据 x_i 转化为新的样本数据 $z_i = W^{-1} x_i$ 。

步骤 6 输出矩阵 $D' = \{(x_1, z_1) (x_2, z_2) \dots, (x_N, z_N)\}$ 。

3 实验与结果分析

由于瓦斯突出样本数量有限,为提高模型稳定性和可靠性,本文采用 10 组 5 折交叉验证实验对模型进行训练。

3.1 瓦斯突出影响因素分析与选取

本文收集 20 个典型煤矿的 130 条历史数据,并将瓦斯突出的影响因素划分为煤与瓦斯、矿区设备、管理和人为 3 大类,具体分析如下:

1) 煤与瓦斯是影响瓦斯突出风险的主体因素。煤与瓦斯因素包含瓦斯体积分数 h 、瓦斯抽采 d 、煤层瓦斯含量 c 以及瓦斯涌出量 o 。

2) 矿区设备是指采矿区的各项硬件设施,是影响瓦斯突出风险的辅助因素。矿区设备因素具体包含通风设施设备 w 、监测监控系统 s 、机电设备 p 、防护设施 g 、通讯设施 n 以及风量供需比 r 。

3) 管理和人为因素同为影响瓦斯突出风险的辅助因素。管理和人为因素包含技术人员配备 a 、安全教育与培训 sa 、安全制度与执行 ex 以及安全科技与投入 de 。

结合上述分析,选取 14 个瓦斯突出影响因素对瓦斯突出风险进行预测,部分样本数据如表 1 所示。

表 1 瓦斯突出危险等级样本数据示例

样本编号	<i>h</i>	<i>d</i>	<i>c</i>	<i>o</i>	<i>w</i>	<i>s</i>	<i>p</i>	<i>g</i>	<i>n</i>	<i>r</i>	<i>a</i>	<i>sa</i>	<i>ex</i>	<i>de</i>	等级
1	5.00	4.1	1.00	5.0	5.0	5.0	5.0	5.0	5.0	2.18	5.0	5.0	4.0	4.0	4
2	5.00	4.6	1.00	4.0	4.0	4.0	5.0	4.0	4.0	3.10	5.0	4.0	5.0	5.0	4
3	5.00	5.0	5.00	5.0	5.0	5.0	5.0	5.0	5.0	5.00	5.0	5.0	5.0	5.0	5
4	2.67	1.1	3.25	1.9	3.6	4.4	3.2	1.7	2.3	2.77	2.2	2.3	4.4	1.4	3
5	4.00	3.8	2.56	3.0	4.0	5.0	5.0	5.0	5.0	4.01	5.0	4.0	5.0	4.0	4
6	4.00	1.4	1.00	3.0	1.7	5.0	3.0	4.0	4.0	1.90	2.3	1.7	5.0	3.0	3
7	1.00	2.0	2.00	1.0	1.0	1.0	1.0	1.0	1.0	2.00	1.0	1.0	1.0	1.0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
124	2.00	1.3	1.77	1.8	1.9	1.8	1.2	1.3	1.4	2.50	2.9	1.3	1.3	1.8	2
125	5.00	5.0	4.51	5.0	5.0	5.0	5.0	5.0	5.0	4.93	5.0	5.0	5.0	5.0	5
126	5.00	2.5	1.00	2.5	3.8	5.0	5.0	3.0	5.0	2.98	1.7	3.8	5.0	5.0	4
127	5.00	3.9	4.55	4.0	3.0	4.0	5.0	4.0	4.0	1.00	1.3	1.7	5.0	3.0	4
128	3.30	3.2	1.74	4.8	2.2	3.8	3.6	3.4	3.8	5.00	4.4	2.5	2.6	2.2	3
129	1.00	3.0	3.00	2.0	2.0	2.0	2.0	2.0	2.0	3.00	2.0	2.0	2.0	2.0	2
130	1.00	4.0	4.00	3.0	3.0	3.0	3.0	3.0	3.0	4.00	3.0	3.0	3.0	3.0	3

3.2 特征提取

在实验中对 130 条数据进行归一化处理, 将瓦斯突出危险等级数目作为 ELM 的输出节点数, 为匹配 ELM 模型的输出, 将 5 个瓦斯突出危险等级标签项分别转化成一维二进制向量。转化后的危险等级标签项如表 2 所示。

表 2 危险等级标签转化对应表

原危险等级	转化后的危险等级
1	10000
2	01000
3	00100
4	00010
5	00001

根据 KLDA 算法原理可知, 样本经处理后的特征维度可以压缩到 1~4 的整数范围内。统一使用 KLDA 算法进行特征提取, 去除瓦斯突出风险数据的相关性, 结果如表 3 所示。可以看出, 在 KLDA 降维后的 4 个主成分中, 前 3 个主成分贡献率(方差百分比)为 98.4% (大于 85%), 可解释 91.97% 的原始数据信息。因此, 本文选取降维后的前 3 个主成分对瓦斯突出危险等级进行预测。

表 3 特征提取后不同特征值的贡献率

特征序号	特征值	贡献率/%	累计贡献率/%
1	7.264	59.8	59.8
2	2.822	23.2	83.0
3	1.873	15.4	98.4
4	0.175	1.4	100.0

3.3 模型参数确定

3.3.1 ELM 的隐含层节点数和激活函数

如表 4 所示, 激活函数分别为 sigmoid、softmax、tanh 函数, 节点数为 2~100, 将空间划分为 99×3 的

网格空间, 对于每组参数组合迭代 200 次, 利用网格搜索法选择 ELM 最优的节点数和激活函数, 结果如图 1 所示。

表 4 ELM 网格搜索实验结果

节点数	sigmoid	softmax	tanh
2	0.678	0.595	0.820
3	0.861	0.586	0.886
4	0.836	0.586	0.905
5	0.903	0.778	0.860
6	0.861	0.820	0.886
7	0.861	0.778	0.886
8	0.883	0.803	0.082
9	0.861	0.853	0.087
10	0.861	0.836	0.087
11	0.853	0.820	0.820
12	0.861	0.820	0.836
13	0.870	0.845	0.836
14	0.861	0.820	0.803
15	0.082	0.836	0.836
⋮	⋮	⋮	⋮

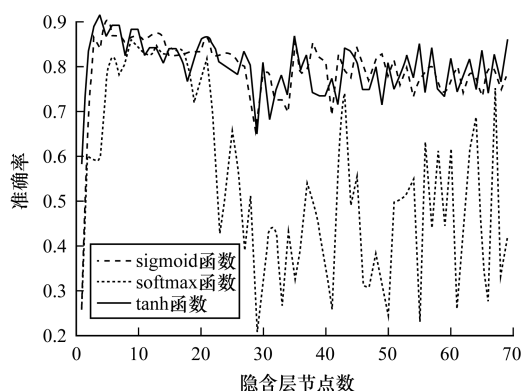


图 1 不同激活函数对 ELM 模型准确率的影响

从表 4 和图 1 可以看出:

1) 使用 sigmoid 和 tanh 函数的 ELM 在不同隐含层节点数的情况下,其预测准确率和稳定性均高于使用 softmax 函数的 ELM。特别是在节点数较高的情况下,使用 softmax 函数的 ELM 准确率出现大幅波动。因此,softmax 激活函数不适用于此模型。

2) 对比 sigmoid 和 tanh 函数在 ELM 上的效果,发现当采用 tanh 激活函数、节点数为 4 时,模型达到最高的准确率 0.905。因此,本文选择 tanh 函数作为模型的激活函数,模型的节点数设定为 4。

3.3.2 最优花粉配子个数

在实验中固定 ELM 模型中参数,将 10~100 的整数值作为配子个数的测试值,采用穷举法确定 ACFPA 算法的最优花粉配子个数。ACFPA-ELM 模型的预测准确率和 ACFPA 花粉配子个数的关系如图 2 所示。

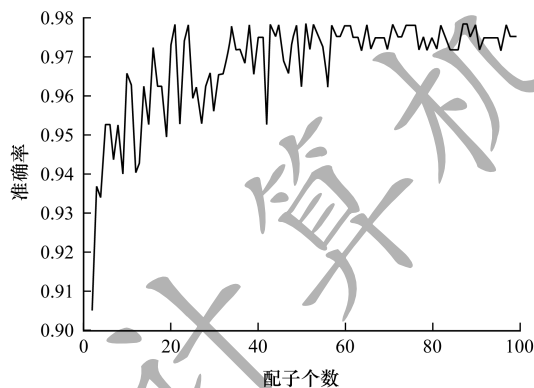


图 2 花粉配子个数对 ACFPA-ELM 模型准确率的影响

从图 2 可知,ACFPA 算法优化效果最佳时的 ELM 模型预测准确率为 0.978,其优化效果最差时的模型预测准确率为 0.905,且 ACFPA 算法达到最佳优化效果时最低的花粉配子个数为 21。由于 ACFPA 算法的运行效率和花粉配子的个数呈负相关,因此为保证优化效果的同时提高算法的运行效率,将 ACFPA 算法的配子个数设置为 21。

3.4 ELM 模型实验结果

本文实验对比 ACFPA-ELM、FPA-ELM、PSO-ELM、ELM 4 个模型的预测精度以及运行效率。首先将瓦斯突出样本数据通过 KLDA 算法进行特征提取,然后依次带入上述模型进行预测,结果如图 3 所示。为保证实验的可靠性和结果的公平性,将 4 个模型中 ELM 的隐含层节点数、激活函数设定为 4 和 tanh 函数。其中,ACFPA-ELM、FPA-ELM、PSO-ELM 3 个模型的种群个数固设定为 20,并根据文献[19-20],将 ACFPA 的切换概率 p 和参数 γ 分别设定为 0.2 和 0.1。

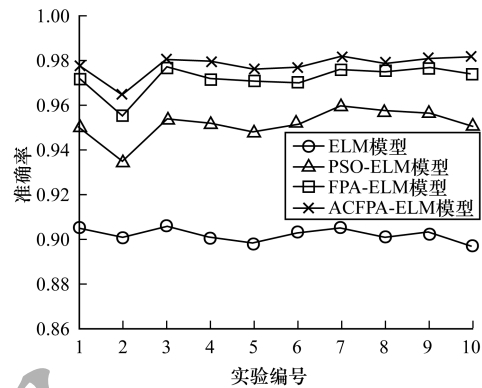


图 3 4 个模型在 10 组实验上的最高准确率

从图 3 可以看出,ACFPA-ELM、FPA-ELM、PSO-ELM、ELM 模型在瓦斯突出危险等级预测上的平均准确率分别为 0.978、0.972、0.952、0.902。由此可知:

1) FPA 和 PSO 算法都能通过优化 ELM 的输入权值和阈值提高预测的准确率。

2) FPA 和 ACFPA 算法通过其自身的 Levy 飞行机制,改善了 PSO 算法易陷入局部最优解的缺陷。

3) ACFPA 算法通过调整初始花粉配子的位置分布,能够进一步提高 FPA-ELM 模型的预测准确率,进而提高瓦斯突出风险模式识别的效果。

表 5 列出了 ELM、PSO-ELM、FPA-ELM 和 ACFPA-ELM 4 个模型的运行时间。由于传统的 ELM 不存在迭代机制,因此本文只对比 PSO-ELM、FPA-ELM、ACFPA-ELM 模型的运行时间。

表 5 4 个模型运行时间对比结果

实验编号	ELM/ 10^{-3}	PSO-ELM	FPA-ELM	ACFPA-ELM
1	7.77	56.82	40.13	32.15
2	7.62	58.27	41.57	34.29
3	7.26	55.11	38.72	27.62
4	6.97	49.15	37.15	26.01
5	7.06	51.39	37.82	25.92
6	7.29	52.24	39.65	28.03
7	7.54	56.38	43.37	32.19
8	6.84	48.15	35.26	25.33
9	6.63	46.68	34.93	22.65
10	6.48	43.94	32.17	21.59
均值	7.15	51.81	38.08	27.58

可以看出,相比于 PSO 算法,FPA 算法运行时间降低了 41.1%~51.5%,表明 FPA 能够更早地跳出局部最优解,从而缩短了模型运行的时间。同时因为 ACFPA 算法优化了花粉配子的初始位置,所以能够更早接近全局最优解。因此,在瓦斯突出风险模式的预测上,ACFPA 算法相比于 PSO 算法和传统的 FPA 算法在时间效率上都有着明显的优势。

3.5 特征提取算法对比

遵循交叉验证的思想, 每次将 104 条数据作为训练集, 26 条数据作为测试集。通过不同隐含层节点数的连续测试, 对比 KPCA 算法、KLDA 算法和不进行特征提取三者对 ELM 模型预测准确率的影响, 结果如图 4 所示。10 次试验中模型预测准确率的均值如表 6 所示。ELM 隐含层节点数和运行时间的关系如图 5 所示。

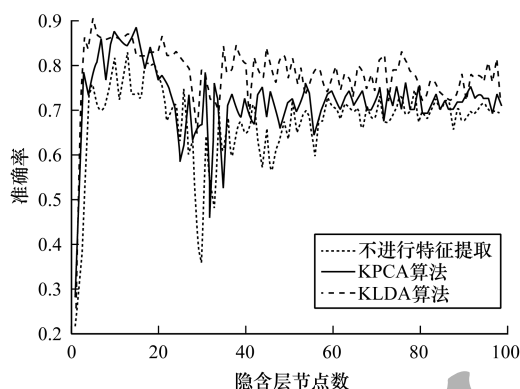


图 4 特征提取算法对 ELM 模型的影响

表 6 2 种特征提取算法效果对比

算法	模型准确率	算法运行时间/ 10^{-3} s	最优节点数
KPCA	0.885	5.072	13
KLDA	0.903	2.519	4

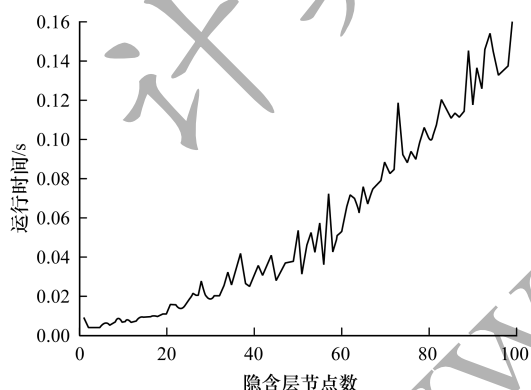


图 5 ELM 算法隐含层节点数和模型准确率的关系

特征提取实验结果表明:

1) KPCA 和 KLDA 降维算法在多数情况下能够提高 ELM 模型在瓦斯突出预测上的准确率。

2) 未经降维后的瓦斯突出数据, 在隐含层节点数为 15 时模型预测准确率达到最优值 0.831; 用 KPCA 降维后的数据对模型进行训练, 当 ELM 的节点数为 13 时, 模型预测准确率达到最优值 0.885; 用 KLDA 降维后的数据对模型训练, 当 ELM 的隐含层节点数为 4 时, 模型预测准确率达到最优值 0.903。由此可知, KLDA 对瓦斯突出数据的降维效果优于 KPCA。

3) 将分别用 3 种方式处理的瓦斯突出数据在 ELM 上进行训练, 对比各自最优的隐含层个数, 其中 KPCA 为 13, KLDA 为 4, 未作特征提取的为 15。由于 ELM 中的大量矩阵的维数取决于设定的隐含层节点个数, 且矩阵乘法的时间和空间复杂度和矩阵维数成正比, 如图 5 所示。当 ELM 的隐含层节点数为 5、13、15 时, 算法中矩阵乘法的时间复杂度分别为 $O(5^3)$ 、 $O(13^3)$ 、 $O(15^3)$, 同理, 可计算出空间复杂度也有同样的关系。因此采用 KLDA 算法对瓦斯突出数据进行特征提取, 在提高算法准确率的同时, 也间接地降低了 ELM 的时间和空间复杂度, 提高了运行效率。

4) KLDA 算法的所用的时间明显低于 KPCA 算法所用的时间, 因此, 在保证特征提取效果的同时, KLDA 的速度更快。

3.6 不同模型对比

实验对比 ACFPA-ELM、SVM、BP 3 个模型的预测准确率和卡帕系数, 结果如表 7 所示。其中, ACFPA-ELM 模型的参数设置同 3.3 节; BP 模型的学习率设为 0.2, 隐含层个数为 1, 隐层节点个数为 4; SVM 模型的步长为 0.2, 核函数采用 RBF 核函数。从表 7 可以看出, ACFPA-ELM 模型的平均预测准确率比 SVM 和 BP 模型分别提高了 2.8% 和 15.1%, 其平均卡帕系数分别提高了 1.8% 和 10.4%。由此表明, 在瓦斯突出风险模型的识别上, 相比于上述经典预测模型, 本文模型预测准确率和结果一致性较高。

表 7 3 个模型的预测准确率和卡帕系数对比结果

实验编号	ACFPA-ELM 模型		SVM 模型		BP 模型	
	准确率	卡帕系数	准确率	卡帕系数	准确率	卡帕系数
1	0.978	0.953	0.953	0.933	0.852	0.852
2	0.965	0.942	0.959	0.928	0.863	0.847
3	0.981	0.958	0.951	0.931	0.859	0.845
4	0.980	0.949	0.944	0.933	0.864	0.846
5	0.976	0.954	0.951	0.925	0.855	0.839
6	0.977	0.956	0.948	0.928	0.858	0.840
7	0.982	0.944	0.950	0.932	0.862	0.847
8	0.979	0.951	0.942	0.934	0.856	0.845
9	0.981	0.950	0.948	0.941	0.849	0.853
10	0.982	0.953	0.951	0.940	0.853	0.857
均值	0.978	0.951	0.950	0.933	0.827	0.847

4 结束语

为提高 ELM 模型在瓦斯突出风险预测上的准确率和速度,本文构建 ACFPA-ELM 模型,通过 KLDA 算法提取样本数据的特征,并利用 ACFPA 算法对 ELM 模型进行优化。KLDA 算法在 KPCA 算法的基础上参考了瓦斯突出样本数据的标签项,提高了样本质量,而采用 ACFPA 算法优化 ELM 预测模型的初始参数,提升了模型识别的准确率和运行效率,并改善了粒子群算法容易陷入局部最优解和算法收敛速度慢的现象。对瓦斯突出风险预测的实验结果表明,与 SVM 和 BP 模型相比,该模型的预测准确率、预测一致性以及运行效率均较优,其通过将 KLDA 降维算法和 ACFPA 算法相结合,克服了 ELM 模型对高维样本预测性能较差的局限性,对其他领域不同维度的预测样本也有较好的适用性。后续将对 KLDA 算法进行改进,进一步提高本文模型对瓦斯突出风险的预测能力。

参考文献

- [1] XIE Guomin, XIE Hong, FU Hua, et al. Prediction model for coal and gas outburst based on NN-SVM [J]. Chinese Journal of Sensors and Actuators, 2016, 29(5): 732-738. (in Chinese)
谢国民, 谢鸿, 付华, 等. 煤与瓦斯突出预测的 NN-SVM 模型[J]. 传感技术学报, 2016, 29(5): 732-738.
- [2] GUAN Weijuan, ZHANG Guoshu, ZHAO Zhigen, et al. Multi-index comprehensive identification and real-time warning of coal and gas outburst [J]. Journal of Mining and Safety Engineering, 2013, 30(6): 922-929. (in Chinese)
关维娟, 张国枢, 赵志根, 等. 煤与瓦斯突出多指标综合辨识与实时预警研究[J]. 采矿与安全工程学报, 2013, 30(6): 922-929.
- [3] ZHU Zhijie, ZHANG Hongwei, LIU Xin. Mine dynamic disaster prediction based on fuzzy pattern recognition [J]. Journal of Natural Disasters, 2014, 23(4): 19-25. (in Chinese)
朱志洁, 张宏伟, 刘鑫. 基于模糊模式识别的矿井动力灾害预测[J]. 自然灾害学报, 2014, 23(4): 19-25.
- [4] GUO Deyong, LI Nianyou, PEI Dawen, et al. Prediction method of coal and gas outburst using the grey theory and neural network [J]. Journal of University of Science and Technology Beijing, 2007, 29(4): 354-357. (in Chinese)
郭德勇, 李念友, 裴大文, 等. 煤与瓦斯突出预测灰色理论——神经网络方法[J]. 北京科技大学学报, 2007, 29(4): 354-357.
- [5] ZHU Zhihao, ZHANG Hongwei, HAN Jun, et al. Prediction of coal and gas outburst based on PCA-BP neural network [J]. China Safety Science Journal, 2013, 23(4): 45-50. (in Chinese)
朱志洁, 张宏伟, 韩军, 等. 基于 PCA-BP 神经网络的煤与瓦斯突出预测研究[J]. 中国安全科学学报, 2013, 23(4): 45-50.
- [6] LI Sheng, HU Haiyong. Risk identification of coal and gas outburst based on KPCA and improved extreme learning machine model [J]. Application Research of Computers, 2018, 35(1): 172-176. (in Chinese)
李胜, 胡海永. 基于 KPCA 和改进极限学习机的煤与瓦斯突出危险性识别[J]. 计算机应用研究, 2018, 35(1): 172-176.
- [7] HUANG Guangbin, CHEN Lei. Enhanced random search based incremental extreme learning machine [J]. Neurocomputing, 2008, 71(16/17/18): 3460-3468.
- [8] WANG Jie, BI Haoyang. A new extreme learning machine optimized by PSO [J]. Journal of Zhengzhou University (Natural Science Edition), 2013, 45(1): 100-104. (in Chinese)
王杰, 毕浩洋. 一种基于粒子群优化的极限学习机[J]. 郑州大学学报(理学版), 2013, 45(1): 100-104.
- [9] YANG Xinshe. Flower pollination algorithm for global optimization [C] // Proceedings of the 11th International Conference on Unconventional Computation and Natural Computation. Berlin, Germany: Springer, 2012: 240-249.
- [10] IGELNIK B, PAO Y H. Stochastic choice of basis functions in adaptive function approximation and the functional-link net [J]. IEEE Transactions on Neural Networks, 1995, 6(6): 1320-1329.
- [11] HUANG Guangbin, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications [J]. Neurocomputing, 2006, 70(1/2/3): 489-501.
- [12] DENG Wanyu, ZHENG Qinghua, CHEN Lin, et al. Research on extreme learning of neural networks [J]. Chinese Journal of Computers, 2010, 33(2): 279-287. (in Chinese)
邓万宇, 郑庆华, 陈琳, 等. 神经网络极速学习方法研究[J]. 计算机学报, 2010, 33(2): 279-287.
- [13] ZHANG Nan, DING Shifei, SHI Zhongzhi. Denoising Laplacian multi layer extreme learning machine [J]. Neurocomputing, 2015, 171: 1066-1074.
- [14] XIAO Huihui, WAN Changxuan, DUAN Yanming, et al. Flower pollination algorithm based on gravity search mechanism [J]. Acta Automatica Sinica, 2017, 43(4): 576-594. (in Chinese)
肖辉辉, 万常选, 段艳明, 等. 基于引力搜索机制的花朵授粉算法[J]. 自动化学报, 2017, 43(4): 576-594.
- [15] DUAN Yanming, XIAO Huihui. Flower pollination algorithm based on quantum behavior [J]. Journal of Lanzhou University of Technology, 2016, 42(3): 88-95. (in Chinese)
段艳明, 肖辉辉. 基于量子行为的花朵授粉算法[J]. 兰州理工大学学报, 2016, 42(3): 88-95.
- [16] XIAO Huihui, WAN Changxuan, DUAN Yanming, et al. Flower pollination algorithm combination with gauss mutation and Powell search method [J]. Journal of Frontiers of Computer Science & Technology, 2017, 11(3): 478-490. (in Chinese)
肖辉辉, 万常选, 段艳明, 等. 融合高斯变异和 Powell 法的花朵授粉优化算法[J]. 计算机科学与探索, 2017, 11(3): 478-490.
- [17] ZHANG Yunpeng, ZUO Fei, ZHAI Zhengjun. A color image encryption algorithm based on chaotic Chebyshev and variable-parameters logistic systems [J]. Journal of North Western Polytechnical University of Science and Technology: Nature Science Edition, 2009, 37(3): 102-105.
- [18] LIU Xiaozhang, YUEN P C, FENG Guocan, et al. Learning kernel in kernel-based LDA for face recognition under illumination variations [J]. IEEE Signal Processing Letters, 2009, 16(12): 1019-1022.
- [19] DRAA A. On the performances of the flower pollination algorithm-qualitative and qualitative analyses [J]. Applied Soft Computing, 2015, 34: 349-371.
- [20] YANG Xinshe, MEHMET K, HE Xingshe. Multi-objective flower algorithm for optimization [J]. Procedia Computer Science, 2013, 18(1): 861-868.

编辑 刘盛龄