

# 多层 DAE 协同 LSSVM 的瓦斯突出预测模型

付华 梁漪

(辽宁工程技术大学电气与控制工程学院 辽宁 葫芦岛 125105)

**摘要** 为准确预测瓦斯突出,提出多层去噪自编码器(Multi-layer DAE)搭载最小二乘支持向量机(LSSVM)的瓦斯突出预测模型。多层 DAE 网络提取瓦斯突出原始数据的有效特征,并链接 LSSVM 进行突出分类。因突出影响因子边界存在一定的模糊性,从时间角度考虑将其分为动、静态影响因子,并依此对多层 DAE 网络按照交叉熵规则设计新的代价函数。利用收集的 100 组真实样本数据,多次实验确定最稳定的模型结构参数。对比分析 PCA-LSSVM、LLE-LSSVM、BP 神经网络模型,结果表明,该模型有更优越的特征提取能力和预测性能,更适用与瓦斯突出预测问题。

**关键词** 瓦斯突出 突出影响因子 去噪自编码器(DAE) 最小二乘支持向量机(LSSVM) 交叉熵

**中图分类号** TP39 TP183 **文献标识码** A **DOI**:10.3969/j.issn.1000-386x.2019.08.037

## MULTI-LAYER DAE COLLABORATED LSSVM GAS OUTBURST PREDICTION MODEL

Fu Hua Liang Yi

(Faculty of Electrical and Control Engineering, Liaoning Technical University, Huludao 125105, Liaoning, China)

**Abstract** In order to predict the gas outburst accurately, this paper presented a gas outburst prediction model based on multi-layer denoising autoencoder(DAE) with least square support vector machine(LSSVM). The multi-layer DAE network was used to extract the effective features of the gas outburst raw data, and collaborated the LSSVM for gas outburst classification. As there was a certain ambiguity in the boundary of the outburst influence factors, it was divided into dynamic and static influence factors from the perspective of time, and we designed a new cost function for the multi-layer DAE network according to the cross entropy rule. In addition, we collected one hundred real sample data, and used them to test the most stable model structure parameters. The same data was utilized to compare and analyze PCA-LSSVM, LLE-LSSVM, LSSVM and BP neural network models to confirm the validity of our proposed model. The experimental results show that the proposed model has superior feature extraction ability and prediction performance, and it is more suitable for gas outburst prediction problems.

**Keywords** Gas outburst Outburst influence factors Denoising autoencoder(DAE) Least square support vector machine(LSSVM) Cross entropy

## 0 引言

煤岩体高应力、高瓦斯压力采掘空间区配合矿震动载易引发瓦斯突出动力灾害,原空间地质动力系统失稳,存积的弹性能呈爆发式瞬间释放,产生煤岩体破碎并抛出现象,严重威胁井下施工人员生命安全和设备安全。我国浅部煤炭资源接近枯竭,随着深部开采,煤岩体承载更高的应力和能量,瓦斯突出发生的可

能性愈发增大。因此,对其预测成为矿井深部开采迫切解决的科学问题<sup>[1-2]</sup>。

针对瓦斯突出预测的研究有很多成果,经典预测方法从机理角度以单因素作用或多因素多指标综合作用开展预测工作,产生了 D、K 综合指标预测、地质动力区划等方法<sup>[3-5]</sup>。这些方法理论相对简单,需要大量的试验数据支撑,数据获取过程复杂,消耗较高的时间成本和经济成本。为弥补以上经典预测方法的不足,许多学者开发了大量基于智能理论预测方法,提高

突出预测效率和预测准确率,得到了广泛的应用。这类方法通常使用降维手段提取动态监测数据有效的特征信息来加速模型决策,例如:文献[6-7]使用主元分析(Principal Component Analysis, PCA)选取对突出贡献率较高的主元影响因子;文献[8]使用局部线性嵌入(Local Linear Embedding, LLE)将高维突出影响因子数据映射至低维空间;文献[9]使用邻域粗糙集(Neighborhood Rough Set, NRS)提取反映突出的主要敏感指标。其中 PCA 是线性降维方法,而突出影响因子间具有非线性关系;LLE 要求样本集是稠密均匀的且只保留训练样本的局部特征,对于离散的测试样本,不能为突出分类问题提取更有效的特征;NRS 缺乏对模糊数据相应的处理方法。

基于以上分析,文中结合去噪自编码器(Denoising Autoencoder, DAE)抗干扰能力强、鲁棒性强和最小二乘支持向量机(Least Square Support Vector Machine, LSSVM)操作简单的优点,建立瓦斯突出预测模型。以 DAE 获取更高效的隐藏特征为目的,设计多层 DAE 网络,并结合瓦斯突出影响因子在时间轴上的特点和交叉熵规则定义多层 DAE 网络的代价函数,用 BP 算法和梯度下降法更新多层 DAE 网络的权重和阈值,得到的特征作为 LSSVM 的输入,对样本进行瓦斯突出分类。

## 1 预测模型设计

### 1.1 动静态约束下多层 DAE 网络的特征提取

#### 1.1.1 多层 DAE

多层 DAE 网络是由  $L$  个去噪自编码器链接而成, DAE 是 Pascal Vincent<sup>[10]</sup>等在自编码器 AE 的基础上提出的一种神经网络,属于典型的单隐层反馈神经网络。相较于传统的反馈神经网络, DAE 抗干扰能力强,因此选用 DAE 去提取原始样本的隐藏特征。多层 DAE 网络结构如图 1 所示。

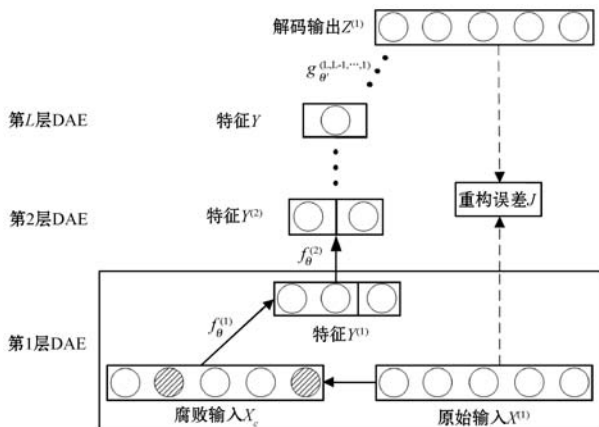


图 1 多层 DAE 网络结构

假设  $N$  组样本构成原始瓦斯突出输入样本集为  $X^{(1)} = \{x_1, x_2, \dots, x_n\}, n \in N$ , 其中包含静态影响因子数据的子样本集为  $X_d^{(1)}$ , 包含静态影响因子数据的子样本集为  $X_s^{(1)}$ 。以概率  $q$  将部分输入层节点置为 0 对输入样本进行噪化, 得到腐败输入  $X_c = X_{dc}^{(1)} + X_{sc}^{(1)}$ 。  $X_c$  通过编码函数  $Y = f_{\theta}(X_c) = s(WX_c + b)$  进行  $L$  次编码得到特征  $Y^{(L)} = Y_d^{(L)} + Y_s^{(L)}$ , 编码器参数  $\theta = \{W, b\}$ ,  $W$  为权重,  $b$  为偏置。特征  $Y^{(L)}$  通过解码函数  $Z = g_{\theta'}(Y) = s(W'Y + b')$  进行  $L$  次解码后得到与输入层相对应的重构输出  $Z^{(1)} = Z_d^{(1)} + Z_s^{(1)}$ , 解码器参数  $\theta' = \{W', b'\}$ , 上标表示网络层数。编码函数  $f_{\theta}$  和解码函数  $g_{\theta'}$  皆为 sigmoid 函数且  $W^T = W'$ 。多层 DAE 网络希望动态的重构输出  $Z_d^{(1)}$  尽可能还原动态输入  $X_d^{(1)}$ , 静态的重构输出  $Z_s^{(1)}$  尽可能还原静态输入  $X_s^{(1)}$ , 此时的动静态影响因子特征  $Y^{(L)}$  能够较完整地表达原始瓦斯突出输入数据。文中设计以交叉熵定义动静态约束下的解码误差用于表达重构输出对输入的还原程度, 同时为避免网络训练过程产生“过拟合”导致网络泛化能力低, 在解码误差中引入  $L_2$  正则化项, 得到用于微调网络参数  $\theta$  的瓦斯突出特征提取代价函数, 表达式为:

$$\theta^*, \theta'^* = \arg \min_{\theta, \theta'} J_{\text{Multi-layer DAE}} = \arg \min_{\theta, \theta'} \alpha J_d(x, z) + \beta J_s(x, z) + \frac{\lambda_w}{2N} \sum_{l=1}^L \|W^{(l)}\|_2^2 = \arg \min - \frac{\alpha}{N} \sum_{n=1}^N [x_{d_n} \log z_{d_n} + (1 - x_{d_n}) \cdot \log(1 - z_{d_n})] - \frac{\beta}{N} \sum_{n=1}^N [x_{s_n} \log z_{s_n} + (1 - x_{s_n}) \cdot \log(1 - z_{s_n})] + \frac{\lambda_w}{2N} \sum_{l=1}^L \|W^{(l)}\|_2^2 \quad (1)$$

式中:  $\alpha$  和  $\beta$  分别为动态影响因子和静态影响因子的控制参数,  $\alpha + \beta = 1$ 。  $\lambda_w$  为正则参数, 常用交叉验证法确定。

#### 1.1.2 多层 DAE 网络参数更新

初始化多层 DAE 网络相关参数, 包括网络层数  $L$ , 加噪概率  $q$ , 输入、重构层节点数, 隐层节点数, 动态影响因子控制参数  $\alpha$ , 静态影响因子控制参数  $\beta$ , 随机初始化编码器参数  $\theta$  和解码器参数  $\theta'$ 。

首先进行网络预训练, 利用“最大-最小值”归一化处理后的瓦斯突出数据在多层 DAE 网络中正向传播得到特征和重构输出。

其次进行网络微调, 采用 BP 算法以式(1)为调优标准计算多层 DAE 网络权重  $W$  的变化量, 表达式为:

$$\frac{\partial J_{\text{Multi-layer DAE}}}{\partial w_l} = \alpha \frac{\partial J_d}{\partial w_l} + \beta \frac{\partial J_s}{\partial w_l} + \frac{\lambda_w}{2N} \cdot \frac{\partial}{\partial w_l} \sum_{l=1}^L \|W^{(l)}\|_2^2 \quad (2)$$

式中:

$$\frac{\partial J_d}{\partial w_l} = \frac{1}{N} \sum_{n=1}^N [g(f(x_{d_n})) - z_{d_n}] x_{d_j} \quad (3)$$

$$\frac{\partial J_s}{\partial w_l} = \frac{1}{N} \sum_{n=1}^N [g(f(x_{s_n})) - z_{s_n}] x_{s_j} \quad (4)$$

因此,式(2)可以写为:

$$\frac{\partial J}{\partial w_l} = \frac{\alpha}{N} \sum_{n=1}^N [g(f(x_{d_n})) - z_{d_n}] x_{d_j} + \frac{\beta}{N} \sum_{n=1}^N [g(f(x_{s_n})) - z_{s_n}] x_{s_j} + \frac{\lambda_w}{N} w_l \quad (5)$$

同理,偏置  $b$  的变化量为:

$$\frac{\partial J}{\partial b_l} = \frac{\alpha}{N} \sum_{n=1}^N [g(f(x_{d_n})) - z_{d_n}] + \frac{\beta}{N} \sum_{n=1}^N [g(f(x_{s_n})) - z_{s_n}] \quad (6)$$

最后,按照式(5)和式(6)梯度减小的方向对权重  $W$  和偏置  $b$  进行更新,置学习率  $\eta = 0.01$ ,更新规则如下:

$$w_l = w_l - \eta \frac{\partial J}{\partial w_l} \quad (7)$$

$$b_l = b_l - \eta \frac{\partial J}{\partial b_l}$$

按照式(7)进行  $t$  次迭代,直至多层 DAE 网络收敛,此时的动静态影响因子特征  $Y^{(L)}$  方可用于下一步的分类。

## 1.2 多层 DAE 协同 LSSVM 预测模型

多层 DAE 网络提取的特征  $Y^{(L)}$  结合样本对应的瓦斯突出信息  $\{c_i\}_{i=1}^N$  用于训练最小二乘支持向量机 LSSVM 分类器,其中突出样本的  $c_i$  为 1、非突样本的  $c_i$  为 0。通过非线性映射  $\Psi(y) = (\varphi(y_1), \varphi(y_2), \dots, \varphi(y_m))$  将特征  $Y^{(L)}$  从原空间映射到  $\Psi$  空间,并在此寻找最优突出决策函数  $p(y) = u\varphi(y) + v$ ,即确定最优权重系数  $u$  和偏差  $v$ ,此时突出预测模型优化目标为:

$$\min J(u, v, \xi) = \frac{1}{2} (u^T u + C \sum_{i=1}^m \xi_i^2) \quad (8)$$

$$\text{s. t. } c_i = u^T \varphi(y_i) + v + \xi_i$$

式中:  $J(u, v, \xi)$  为突出预测模型目标函数,  $\xi_i$  为容许误差,  $C$  为惩罚参数。

同 SVM 求解方式一样,使用拉格朗日算子将上述优化问题转变为求解对偶无约束问题,将约束条件整合到式(8)的目标函数。

$$L(u, v, \xi, \lambda) = \frac{1}{2} (u^T u + C \sum_{i=1}^m \xi_i^2) + \sum_{i=1}^m \lambda_i [c_i - (u^T \varphi(y_i) + v + \xi_i)] \quad (9)$$

式中:  $L(u, v, \xi, \lambda)$  为重新定义的突出预测模型目标函数,  $\lambda_i$  为拉格朗日算子。

利用式(9)对变量  $u, v, \xi_i, \lambda_i$  分别求偏导并令偏导数为 0, 有  $u = \sum_{i=1}^m \lambda_i \varphi(y_i)$ 、 $\sum_{i=1}^m \lambda_i = 0$ 、 $\lambda_i = C\xi_i$ 。将上述关系代入式(8),其中  $\varphi(y_i)^T \varphi(y_j)$  项用高斯核函数  $\kappa(y_i, y_j) = \varphi(y_i)^T \varphi(y_j)$  代替。由此可得模型分类函数:

$$p(y) = \text{Sign} \left( \sum_{i=1}^m \lambda_i c_i \kappa(y, y_i) + v \right) \quad (10)$$

一般地,惩罚参数  $C$  根据经验选择。

## 2 瓦斯突出影响因子选取

瓦斯突出是在特殊的地质赋存条件下,煤岩体系系统能量先以稳态积聚再以非稳态释放的非线性动力过程,是外部荷载环境、内部结构、构造及其物理力学性质的多物理场耦合作用的结果。通常瓦斯突出影响因子之间呈现复杂的非线性关系,致使某一因子是否为突出必要条件没有明确的划分。因此,结合相关文献<sup>[12-15]</sup>,综合考虑瓦斯、地质构造、煤的物理性质因素,选取 13 个突出影响因子,并从时间角度出发,将其划分为动态和静态两种类别。

动态影响因子  $X_d = \{ \text{瓦斯含量 } x'_1, \text{瓦斯压力 } x'_2, \text{瓦斯放散初速度 } x'_3, \text{采动应力 } x'_4 \}$ 。

静态影响因子  $X_s = \{ \text{煤层地质构造 } x_1, \text{地应力 } x_2, \text{煤的破坏类型 } x_3, \text{煤层厚度 } x_4, \text{煤层埋深 } x_5, \text{煤的坚固性系数 } x_6, \text{煤系地层的含水性 } x_7, \text{煤岩渗透率 } x_8, \text{岩浆侵入 } x_9 \}$ 。

部分影响因子为定性因子,使用前需量化处理,量化标准见表 1。

表 1 定性因子量化说明

定性因子	赋值				
	1	2	3	4	5
煤层地质构造 $x_1$	不集中	一般集中	集中	较集中	十分集中
煤的破坏类型 $x_3$	I(非破坏煤)	II(破坏煤)	III(强烈破坏煤)	IV(粉碎煤)	V(全粉碎煤)
煤系地层含水性 $x_7$	丰富	一般贫乏	贫乏	较贫乏	十分贫乏
火成岩侵入 $x_9$	无	轻微	严重	较严重	十分严重

### 3 确定模型结构参数

经整理,研究选取开滦矿业集团钱家营煤矿 5 号煤层 14 组典型的瓦斯突出矿井监测数据和 86 组无突出矿井监测数据形成样本空间,根据表 1 对定性因子进行量化,同时将所有样本按照“最大-最小值”进行归一化处理形成样本集。置测试样本与检验样本的比例为 85:15。按照表 2 将样本空间分为 5 种分配类型,组内抽样为无放回随机抽取,所有分配类型下的样本均用于模型的各项检验。置动态影响因子控制参数  $\alpha = 0.5$ ,静态影响因子控制参数  $\beta = 0.5$ ,高斯核参数为 1 000,惩罚参数  $C = 50$ ,迭代次数  $t = 200$ 。

表 2 样本分配类型

样本数	样本组别				
	1	2	3	4	5
训练样本中突出样本数	12	11	10	9	8
检验样本中突出样本数	2	3	4	5	6

#### 3.1 确定模型网络层数

固定加噪概率  $q = 0.4$ ,正则参数  $\lambda_w = 0.01$ ,分别检验带有  $L = 2 \sim 5$  层 DAE 网络的模型预测性能,每个网络的第  $l = 2, 3, 4, 5$  隐层节点数分别设置为 11、9、7、5,以平均分类准确率和训练时间为检验标准,试验统计结果如图 2 所示。

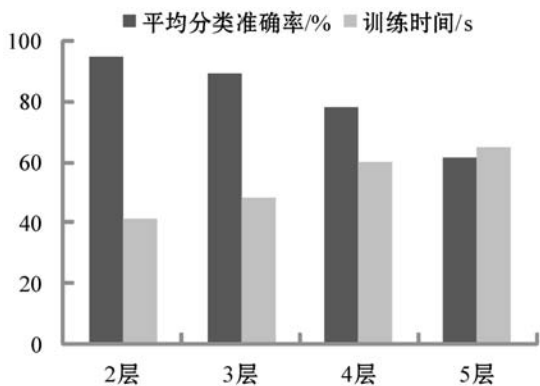


图 2 多层 DAE 网络平均分类准确率和训练时间统计

由图 2 可得,不是网络层数越多模型的效果越好。训练时间随网络层数增加而迅速递增。根据试验结果,从模型准确性和节约成本的角度出发,置模型中多层 DAE 网络层数  $L = 2$ 。

#### 3.2 确定模型最优隐层节点数

同 BP 神经网络类似,隐层节点数对模型性能有重大影响,数量过少,网络无法训练或压缩性能很差;数量过多,训练容易陷入局部最优,所以最优隐层节点数需要人为多次实验确定。结合“控制变量法”和“穷

举法”的思想,反复试验并计算模型平均均方根误差 RMSE,选择最小误差所对应的隐层节点数确定多层 DAE 网络拓扑结构。具体做法如下:

$h^{(1)}$  表示第一隐层节点数, $h^{(2)}$  表示第二隐层节点数。因多层 DAE 网络在预测模型中起维度压缩作用,因此有  $1 < h^{(2)} < h^{(1)} < d = 13$ , $d$  为样本维度,即影响因子个数。试验过程先固定其他变量,拟定加噪概率  $q = 0.4$ ,正则参数  $\lambda_w = 0.01$ 。首先置  $h^{(2)} = 1$ ,对  $h^{(1)}$  在  $[2, 12]$  区间内的候选整数值进行逐一实验, RMSE 统计结果如图 3 实线所示。

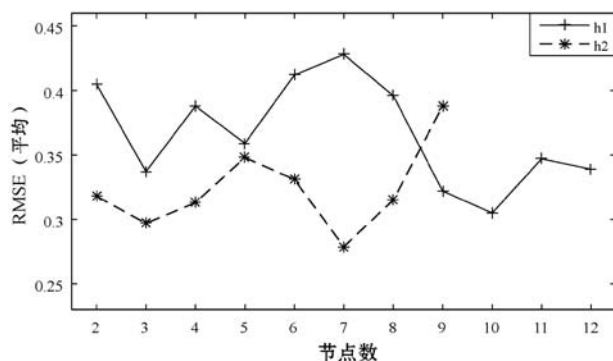


图 3 模型采用不同隐层节点数的预测效果

显然  $h^{(1)} = 10$  时, RMSE 达到最小值 0.305。其次置  $h^{(1)} = 10$ ,对  $h^{(2)}$  在  $[2, 9]$  区间内的候选整数值进行逐一实验, RMSE 统计结果如图 3 虚线所示。可见  $h^{(2)} = 7$  时, RMSE 达到最小值 0.279。因此,瓦斯突出预测模型中多层 DAE 网络拓扑结构为 13(输入层节点数)→10(第一隐层节点数)→7(第二隐层节点数)→13(重构层点数)。

#### 3.3 确定模型加噪概率和正则参数

针对动静态类型输入数据,在合适的区间内设定加噪概率  $q$  和匹配恰当的正则参数  $\lambda_w$  对模型的预测表现极为关键。置正则参数  $\lambda_w = 0.01$ ,其他参数均按常规选取。以  $q = 0.10$  为起点,按照 0.1 步长逐渐增加  $q$  值至  $q = 0.90$ ,重复 5 次试验获取平均 RMSE,并用 MATLAB R2014a 中的 cftool 工具箱拟合加噪概率  $q$  与 RMSE 之间的关系,拟合曲线如图 4 所示。

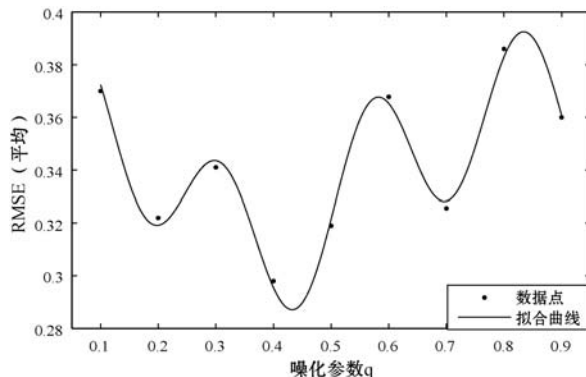


图 4 噪化参数  $q$  对模型预测效果的拟合关系

由图 4 可得,所尝试的数据点中  $q=0.40$  时 RMSE 最小为 0.298。实际中不需精确计算使 RMSE 达到最小值的加噪概率  $q$ ,根据拟合结果认定在以 0.4 为左边界的闭区间  $[0.4, 0.47]$  内选择  $q$  值即为合理,后续试验皆规定  $q=0.40$ 。

此外,采用交叉验证法确定合适的正则参数,设预选范围  $\lambda_w \in \{0.001, 0.01, 0.1, 1, 10, 100, 1\ 000\}$ , 改后  $\lambda_w \in \{0.001, 0.01, 0.1, 1, 10, 100\}$  将该范围内  $\lambda_w$  的取值逐次代入模型中,得到相应的 RMSE 分别为 0.315、0.303、0.338、0.376、0.402、0.457,因此预测模型的最优  $\lambda_w$  为 0.01。

## 4 模型预测对比

### 4.1 DAE 与多层 DAE 特征提取能力对比

本节进行 DAE-10-LSSVM 模型、DAE-7-LSSVM 模型和多层 DAE-LSSVM 模型预测效果对比,用于观察单层 DAE 与多层 DAE 在瓦斯突出数据集上特征提取能力的差别。其中 DAE-10-LSSVM 相当选用与多层 DAE-LSSVM 中第一隐层输出的同等维度特征作用于分类器进行预测,其网络拓扑结构为  $13 \rightarrow 10 \rightarrow 13$ 。DAE-7-LSSVM 相当选用与多层 DAE-LSSVM 中第二隐层输出的同等维度特征作用于分类器进行预测,其网络拓扑结构为  $13 \rightarrow 7 \rightarrow 13$ ,两模型其他参数设置均与多层 DAE-LSSVM 保持一致。采用  $F_1$  度量和模型训练时间  $T_p$  描述模型预测性能和收敛性能,对比结果见表 3。

表 3 5 种样本分配方式下 3 个模型预测准确性和有效性对比

模型		样本组别					平均值
		1	2	3	4	5	
DAE-10-LSSVM	$F_1$ 度量	0.400	0.571	0.500	0.667	0.750	0.577 6
	$T_p/s$	37.384 8	39.834 5	36.830 2	39.077 0	35.774 9	37.780 3
DAE-7-LSSVM	$F_1$ 度量	0.000	0.545	0.600	0.545	0.769	0.491 8
	$T_p/s$	33.047 5	35.298 4	32.647 4	34.684 8	36.787 4	34.493 1
多层 DAE-LSSVM	$F_1$ 度量	0.800	0.750	0.727	0.889	0.909	0.771 8
	$T_p/s$	40.878 3	43.757 8	44.242 4	42.886 0	44.582 4	43.269 3

由表 3 可知,5 次试验中多层 DAE-LSSVM 模型  $F_1$  度量值都高于其他两模型,平均度量更是远超其他两模型,表现了较高的准确性。DAE-7-LSSVM 模型对于不同组合方式的样本,预测准确性浮动较大,模型稳定性较差,DAE-10-LSSVM 模型较其稍有改观。在模型

训练时间方面,多层 DAE-LSSVM 模型收敛速度略慢于另外两模型,因为该模型复杂度相较而言更为复杂,势必多花费一些时间训练模型,鉴于其优秀的预测性能,并且收敛时间也在可接受的范围内,总体上认为多层 DAE-LSSVM 模型较另外两模型有更大的优势。

### 4.2 不同模型预测效果比较

为验证所提模型相较于其他智能预测模型是否具有优越性,与 PCA-LSSVM 模型、LLE-LSSVM 模型、LSSVM 分类器、BP 神经网络进行预测精度对比。BP 神经网络采用  $13 \rightarrow 7 \rightarrow 1$  拓扑结构,激励函数为 Sigmoid 函数,训练网络时设置学习率为 0.01,迭代次数为 500 次。采用表 2 中 5 种分配方式下的样本组合,从均方根误差 RMSE、 $F_1$  度量两方面评价模型预测性能。对比结果见表 4。

表 4 5 个模型预测效果对比

模型	RMSE(平均)	$F_1$ 度量(平均)
多层 DAE-LSSVM	0.258	0.889
PCA-LSSVM	0.356	0.727
LLE-LSSVM	0.356	0.727
LSSVM	0.516	0.615
BP 神经网络	0.577	0.571

由表 4 可知,多层 DAE-LSSVM 模型的 RMSE 和  $F_1$  度量值在 5 个智能预测模型中均为最小。相较 PCA-LSSVM 模型、LLE-LSSVM 模型,所提模型的 RMSE 和  $F_1$  度量值分别提高 27.5% 和 22.3%,亦证明所提模型能够获得更有效的特征表示。相较 LSSVM 分类器,所提模型的 RMSE 和  $F_1$  度量值分别提高 50.0% 和 44.5%,亦证明在瓦斯突出预测过程中,减少样本维度能构使得预测准确率大大提升。相较 BP 神经网络,所提模型的 RMSE 和  $F_1$  度量值分别提高 55.3% 和 55.7%,亦证明 LSSVM 分类器比 Sigmoid 函数在分类问题上获得更优越的性能。综上,多层 DAE-LSSVM 模型能更适用于瓦斯突出预测,具有良好的理论价值和实用价值。

## 5 结 语

以瓦斯突出影响因子为基础数据,根据瓦斯突出影响因子在时间域上的分布特点,将其分为动、静态影响因子,并基于交叉熵准则设计适用于突出样本特征提取的多层 DAE 网络代价函数,以此作为网络的调优标准。由该网络提取的样本特征用于训练 LSSVM 分类器,建立多层 DAE 网络协同 LSSVM 模型用于试验

区的瓦斯突出预测。

应用实际的瓦斯突出影响因子数据进行多次重复试验,确定多层 DAE-LSSVM 预测模型的结构参数,与不同网络结构的单层 DAE-LSSVM 相比,证明多层 DAE-LSSVM 模型有较好的降维效果。同 PCA-LSSVM、LLE-LSSVM、LSSVM、BP 神经网络瓦斯突出预测模型对比,证明所提模型预测准确率高,鲁棒性强。

## 参 考 文 献

- [ 1 ] 窦林名,何学秋, Ren T, 等. 动静载叠加诱发煤岩瓦斯动力灾害原理及防治技术[J]. 中国矿业大学学报, 2018, 47(1): 48 - 59.
- [ 2 ] 荣海,张宏伟,梁冰,等. 煤岩动力系统失稳机理[J]. 煤炭学报, 2017, 42(7): 1663 - 1671.
- [ 3 ] 李冬,彭苏萍,杜文凤,等. 煤层瓦斯突出危险区综合预测方法[J]. 煤炭学报, 2018, 43(2): 466 - 472.
- [ 4 ] 姜福兴,尹永明,朱权洁,等. 基于掘进面应力和瓦斯浓度动态变化的煤与瓦斯突出预警试验研究[J]. 岩石力学与工程学报, 2014, 33(S2): 3581 - 3588.
- [ 5 ] 李楠,王恩元, Ge M C. 微震监测技术及其在煤矿的应用现状与展望[J]. 煤炭学报, 2017, 42(S1): 83 - 96.
- [ 6 ] 付华,李海霞,卢万杰,等. 一种改进的极限学习机煤与瓦斯突出预测模型[J]. 传感技术学报, 2016, 29(1): 69 - 74.
- [ 7 ] 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): 46 - 51.
- [ 8 ] 付华,代巍. 基于 LLE 与 BA-Elman 的瓦斯涌出量动态预测研究[J]. 传感技术学报, 2016, 29(9): 1383 - 1388.
- [ 9 ] 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): 733 - 738.
- [ 10 ] Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders[C]// International Conference on Machine Learning. ACM, 2008: 1096 - 1103.
- [ 11 ] Wang J, Wen Y, Gou Y, et al. Fractional-order gradient descent learning of BP neural networks with Caputo derivative[J]. Neural Networks, 2017, 89: 19 - 30.
- [ 12 ] 闫江伟,张小兵,张子敏. 煤与瓦斯突出地质控制机理探讨[J]. 煤炭学报, 2013, 38(7): 1174 - 1178.
- [ 13 ] 王世超,潘凤龙,申健. 煤与瓦斯突出预测敏感指标确定及应用[J]. 煤炭科学技术, 2013, 41(5): 82 - 85.
- [ 14 ] 唐巨鹏,杨森林,王亚林,等. 地应力和瓦斯压力作用下深部煤与瓦斯突出试验[J]. 岩土力学, 2014, 35(10): 2769 - 2774.
- [ 15 ] 付华,王馨蕊,王志军,等. 基于 PCA 和 PSO-ELM 的煤与瓦斯突出软测量研究[J]. 传感技术学报, 2014, 27(12): 1710 - 1715.

(上接第 180 页)

- [ 2 ] Yu L, Hermann K M, Blunsom P, et al. Deep learning for answer sentence selection[EB]. arXiv preprint arXiv:1412.1632, 2014.
- [ 3 ] Liu P, Qiu X, Chen J, et al. Deep fusion LSTMs for text semantic matching[C]//Proceedings of Annual Meeting of the Association for Computational Linguistics, 2016.
- [ 4 ] Liu P, Qiu X, Huang X. Modelling interaction of sentence pair with coupled-lstms[EB]. arXiv preprint arXiv:1605.05573, 2016.
- [ 5 ] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735 - 1780.
- [ 6 ] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278 - 2324.
- [ 7 ] Santos C, Tan M, Xiang B, et al. Attentive pooling networks[EB]. arXiv preprint arXiv:1602.03609, 2016.
- [ 8 ] LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series[M]//The handbook of brain theory and neural networks. MIT Press, 1998: 255 - 258.
- [ 9 ] Kalchbrenner N, Grefenstette E, Blunsom P. A convolutional neural network for modelling sentences[EB]. arXiv preprint arXiv:1404.2188, 2014.
- [ 10 ] Wang Z, Mi H, Ittycheriah A. Sentence similarity learning by lexical decomposition and composition[EB]. arXiv preprint arXiv:1602.07019, 2016.
- [ 11 ] Hu B, Lu Z, Li H, et al. Convolutional neural network architectures for matching natural language sentences[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems—Volume 2. MIT Press, 2014: 2042 - 2050.
- [ 12 ] Severyn A, Moschitti A. Learning to rank short text pairs with convolutional deep neural networks[C]//Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2015: 373 - 382.
- [ 13 ] Yin W, Schütze H, Xiang B, et al. ABCNN: Attention-Based Convolutional Neural Network for Modeling Sentence Pairs[EB]. arXiv preprint arXiv:1512.05193, 2015.
- [ 14 ] Mikolov T, Sutskever I, Chen K, et al. Distributed Representations of Words and Phrases and their Compositionality[J]. Advances in Neural Information Processing Systems, 2013, 26: 3111 - 3119.
- [ 15 ] Bordes A, Glorot X, Weston J, et al. A semantic matching energy function for learning with multi-relational data[J]. Machine Learning, 2014, 94(2): 233 - 259.
- [ 16 ] Bengio Y. Learning Deep Architectures for AI[J]. Foundations and trends in Machine Learning, 2009, 2(1): 1 - 127.
- [ 17 ] Bordes A, Weston J, Usunier N. Open question answering with weakly supervised embedding models[C]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, 2014: 165 - 180.