

doi: 10.7690/bgzd.2026.06.006

基于贝叶斯-支持向量回归的压装药密度预测算法

赵维亮¹, 张雷², 庄存波¹, 袁申³, 黄求安³, 张旭¹(1. 北京理工大学, 北京 100081; 2. 天津商业大学, 天津 300134;
3. 中国兵器装备集团自动化研究所有限公司智能制造事业部, 四川 绵阳 621000)

摘要: 针对压装药成型密度预测问题, 提出一种贝叶斯-支持向量回归算法(Bayesian optimization-support vector regression, BO-SVR)。利用正交实验法采集压装工艺参数及质量数据, 对数据样本进行相关性分析, 在此基础上构建支持向量回归模型, 使用贝叶斯算法搜寻支持向量回归模型惩罚系数以及核函数参数等超参数的最优组合, 对不同参数组合的效果进行评估, 并对分析 BO-SVR 模型与传统支持向量回归(support vector regression, SVR)模型的预测效果。结果表明, BO-SVR 比传统 SVR 模型各项评价指标均提升近一倍。

关键词: 支持向量回归; 贝叶斯优化算法; 交叉验证; 参数优化; 质量预测

中图分类号: TJ410.3⁺4; TP301.6 **文献标志码:** A

Prediction Algorithm of Charge Density Based on Bayesian Optimization-support Vector Regression

Zhao Weiliang¹, Zhang Lei², Zhuang Cunbo¹, Yuan Shen³, Huang Qiu'an³, Zhang Xu¹(1. Beijing Institute of Technology, Beijing 100081, China; 2. Tianjin University of Commerce, Tianjin 300134, China;
3. Department of Intelligent Manufacture, Automation Research Institute Co., Ltd. of China South Industries Group Corporation, Mianyang 621000, China)

Abstract: A Bayesian optimization-support vector regression (BO-SVR) algorithm is proposed for the problem of predicting the molding density of press-loaded drugs. The orthogonal experimental method is used to collect the press-loading process parameters and quality data, and the data samples are analyzed for correlation, on the basis of which the support vector regression model is constructed, and the Bayesian algorithm is used to search for the optimal combinations of the penalty coefficients of the support vector regression model as well as hyperparameters such as the kernel function parameter, to evaluate the effects of the different combinations of the parameters, and to compare and analyze the effect of the BO-SVR model with that of the traditional support vector regression (SVR) model. The results show that BO-SVR nearly doubles all the evaluation indexes than the traditional SVR model.

Keywords: support vector regression; Bayesian optimization algorithm; cross-validation; parameter optimization; quality prediction

0 引言

随着兵工技术的不断发展, 大量现代高性能武器装备涌现出来, 其中影响这些武器性能最重要的因素之一就是装药质量^[1]。目前, 大多数装药工艺研发较为传统, 对装药质量的测量主要采用排水法等较为传统的方法测量, 存在生产工艺繁琐、生产周期长和生产效率较低等问题。

近年来, 机器学习和深度学习被广泛应用于质量预测领域, 因此, 利用机器学习和深度学习实现对装药压制成型质量的预测是一个值得研究的方向。

机器学习和深度学习在压装装药质量预测方面应用较少, 但在其他质量预测领域应用较为成熟。

在机器学习方面, 王伟等^[2]通过建立 DP 钢力学性能梯度提升决策树预测模型, 实现对 DP 钢的力学性能在线质量监控, 决策树算法对于多维度的数据回归效果较好, 但在低维数据以及少样本数据集上, 容易出现过拟合等现象; 刘琳岚等^[3]利用随机森林回归算法对链路质量进行预测, 但随机森林算法本质上是由决策树组成的集成模型, 与决策树算法类似, 都存在类似的问题; 深度学习也被广泛应用于质量预测中, 董海等^[4]提出一种基于优化深度信念网络的完全填充型 FDM 零件质量预测方法, 该方法优点是能够对高维数据集进行处理, 且对数据的非线性建模能力很强, 但缺点是需要大量的数据集进行模型训练; 与此类似的还有陶磊等^[5]构建的基于 B 样条神经网络的温度场预测模型, 该模型用以

收稿日期: 2024-12-08; 修回日期: 2025-01-09

第一作者: 赵维亮(1999—), 男, 江西人, 硕士。

预测熔铸装药过程中模具内部温度场分布及其变化规律，而该模型也需要大量的仿真数据作为训练数据集以保障其准确率。

质量预测模型参数很大程度上会影响模型的预测准确度，通常是根据经验测试来确定这些参数的最优值，随着参数优化研究的不断深入，智能优化算法开始应用于预测模型参数寻优中。Bazi 等^[6]利用遗传算法对基于高斯核的非线性支持向量机分类器参数优化；侯红玲等^[7]利用遗传算法对 BP 神经网络权值和阈值进行优化，结果表明预测结果更加准确；Zhang 等^[8]开发了一种新的蚁群算法解决 SVM 模型参数寻优问题；石权^[9]基于贝叶斯优化算法对随机森林、梯度提升树的融合模型参数进行调优，得到了更好的预测精度；鲁娟等^[10]提出一种基于差分进化算法优化的支持向量机回归模型，该模型用来预测蠕墨铸铁加工过程中的表面质量，指导加工参数调整。

由于已有的压装药成型质量数据集样本较小，不足以支撑深度神经网络等含有大量参数的模型进行训练，故采用以 SVM 为预测模型的机器学习算法进行压装药成型质量预测。同时，采用贝叶斯优化算法对 SVM 模型进行参数寻优，针对样本较少的问题，采用交叉验证的方法对模型进行训练与评估。

1 支持向量回归模型

支持向量回归是一种基于 SVM 的回归方法，其目标是在样本空间中寻找一个最优超平面实现对回归过程的建模。

假设有训练样本集 $T = \{(\mathbf{x}_i, y_i), i = 1, 2, \dots, n\}$ ， \mathbf{x}_i 指第 i 个样本的输入向量， y_i 指第 i 个样本的输出值， n 为样本总个数。在一般的回归问题中，回归效果的好坏即损失函数用回归模型 $f(x) = wx + b$ 与真实值 y 的差别来衡量，当 $f(x)$ 与 y 一致时，损失为 0；但在支持向量回归中，存在一个容忍度 ε ，当 $|f(x) - y| > \varepsilon$ 时，才计算损失。故支持向量回归 (SVR) 问题可以形式化为：

$$\begin{aligned} \min_{w,b} & \frac{1}{2} \|w\|^2 \\ \text{s.t.} & |y_i - f(\mathbf{x}_i)| \leq \varepsilon, \quad i = 1, 2, \dots, n. \end{aligned} \quad (1)$$

在此基础上，对每个样本点 (\mathbf{x}_i, y_i) 引入一个松弛变量 $\xi_i \geq 0$ ，当样本点 (\mathbf{x}_i, y_i) 满足 $|y_i - f(\mathbf{x}_i)| \leq \varepsilon$ 时， $\xi_i = 0$ ，当样本点 (\mathbf{x}_i, y_i) 不满足 $|y_i - f(\mathbf{x}_i)| \leq \varepsilon$ 时，

则存在一个松弛变量 ξ_i ，使得式(2)成立：

$$|y_i - f(\mathbf{x}_i)| \leq \varepsilon + \xi_i. \quad (2)$$

去掉绝对值，则可以表示为：

$$\left. \begin{aligned} y_i - f(\mathbf{x}_i) & \leq \varepsilon + \xi_i \\ f(\mathbf{x}_i) - y_i & \leq \varepsilon + \xi_i^* \end{aligned} \right\}. \quad (3)$$

由上面的约束条件可知，除了要使权值 w 的二范数平方尽可能小之外，还需要使各个样本点的松弛变量 ξ_i 尽可能小，所以可得到 SVR 的约束最优化原始问题：

$$\left. \begin{aligned} \min_{w,b} & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{s.t.} & f(\mathbf{x}_i) - y_i \leq \varepsilon + \xi_i \\ & y_i - f(\mathbf{x}_i) \leq \varepsilon + \xi_i^* \\ & \xi_i \geq 0, \quad \xi_i^* \geq 0, \quad i = 1, 2, \dots, n \end{aligned} \right\}. \quad (4)$$

式中 $C > 0$ 为惩罚系数，表示回归的约束程度，通过调整惩罚系数 C 的大小来控制 SVR 的适应能力。

通过拉格朗日乘子法可得到拉格朗日无约束优化函数：

$$\begin{aligned} L(w, b, \xi_i, \xi_i^*, \alpha, \alpha^*, \mu, \mu^*) & = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) + \\ & \sum_{i=1}^n \alpha_i (w\mathbf{x}_i + b - y_i - \varepsilon - \xi_i) + \sum_{i=1}^n \alpha_i^* (y_i - w\mathbf{x}_i + \\ & b - \varepsilon - \xi_i^*) - \sum_{i=1}^n \mu_i \xi_i - \sum_{i=1}^n \mu_i^* \xi_i^*, \quad i = 1, 2, \dots, n. \end{aligned} \quad (5)$$

式中 $\alpha_i > 0$ 、 $\alpha_i^* > 0$ 、 $\mu_i > 0$ 、 $\mu_i^* > 0$ 为拉格朗日乘子。分别对 $L(w, b, \xi_i, \xi_i^*, \alpha, \alpha^*, \mu, \mu^*)$ 中的 w 、 b 、 ξ_i 、 ξ_i^* 求偏导，得到：

$$w^* = \sum_{i=1}^n (\alpha_i^* - \alpha_i) \mathbf{x}_i; \quad (6)$$

$$\sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0; \quad (7)$$

$$C - \alpha_i - \mu_i = 0; \quad (8)$$

$$C - \alpha_i^* - \mu_i^* = 0. \quad (9)$$

将式(6)–(9)代入式(5)后，可以得到其对偶问题：

$$\left. \begin{aligned} \min_{\alpha, \alpha^*} & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) (\mathbf{x}_i * \mathbf{x}_j) - \\ & \sum_{i=1}^n y_i (\alpha_i^* - \alpha_i) + \varepsilon (\alpha_i^* + \alpha_i) \\ \text{s.t.} & \sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0 \\ & 0 \leq \alpha_i \leq C \\ & 0 \leq \alpha_i^* \leq C, \quad i = 1, 2, \dots, n \end{aligned} \right\}. \quad (10)$$

从而求出原始最优化问题的 w^* 和 b^* :

$$w^* = \sum_{i=1}^n (\alpha_i^* - \alpha_i) \mathbf{x}_i; \tag{11}$$

$$b^* = y_j + \varepsilon - \sum_{i=1}^n (\alpha_i^* - \alpha_i) (\mathbf{x}_i \cdot \mathbf{x}_j). \tag{12}$$

最终的回归函数为:

$$f(x) = w^* x + b^* = \sum_{i=1}^n (\alpha_i^* - \alpha_i) (\mathbf{x}_i \cdot x) + b^*. \tag{13}$$

式中 $(\mathbf{x}_i \cdot x)$ 为输入 x 和训练样本 \mathbf{x}_i 的内积, 构造一个新函数 $k(\mathbf{x}_i, x) = \mathcal{O}(\mathbf{x}_i) \cdot \mathcal{O}(x)$, 则回归函数为:

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \alpha_i) k(\mathbf{x}_i, x) + b^*. \tag{14}$$

式中 $k(\mathbf{x}_i, x)$ 为支持向量回归模型的核函数。

2 贝叶斯优化算法

2.1 贝叶斯理论

贝叶斯优化算法是由美国学者 Pelikan 等^[11]提出的。其主要逻辑是通过不同的超参数组合, 利用贝叶斯公式不断更新高斯过程建立的目标函数先验概率分布。其中贝叶斯公式的表达式为:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^m P(A_j)P(B|A_j)}. \tag{15}$$

式中: A 和 B 分别为 2 个随机事件; $P(A_i|B)$ 为已知事件 B 发生后 A_i 发生的条件概率, 称 $P(A_i|B)$ 为 A_i 的后验概率; $P(B|A_i)$ 为 B 的似然函数; $P(A_i)$ 和 $P(A_j)$ 不考虑事件 B 的影响, 故称为 A 的先验概率。

2.2 高斯过程

贝叶斯优化算法常采用的一种概率模型就是高斯过程 (Gaussian process, GP) 模型, 高斯过程是指观测值分布在一个连续的空间或时间中的统计模型。

假设预测的目标函数 $g(x)$ 服从高斯过程, $g(x) \sim GP(E(x), K(x, x'))$, 设其均值为 0, 即有一组样本点 $A = \{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)\}$, $i = 1, 2, \dots$ 满足 $g(x) \sim GP(0, K(x, x'))$, 其中核函数矩阵定义如下:

$$\mathbf{K}(X, X) = \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_i) \\ \vdots & & \vdots \\ k(x_i, x_1) & \dots & k(x_i, x_i) \end{bmatrix}. \tag{16}$$

经过高斯过程后, 通过输入采集函数确定新的采样点 x^* , 并生成一个新的高斯分布, 可得到联合高斯分布:

$$\begin{bmatrix} g \\ g^* \end{bmatrix} \sim N\left(0, \begin{bmatrix} \mathbf{K}(X, X) & K(X, x^*) \\ \mathbf{K}(x^*, X) & K(x^*, x^*) \end{bmatrix}\right). \tag{17}$$

式中 $\mathbf{K}(X, x^*)$ 为 X 和 x^* 的协方差矩阵。由于多维正态分布的条件分布仍为正态分布, 故可以计算出在已知 $g(x_{1:i})$ 的条件下 (其中 $x_{1:i}$ 表示 x_1, x_2, \dots, x_i), $g(x^*)$ 所服从的条件分布, 即 $g(x^*)$ 服从 1 维正态分布:

$$g(x^*) | g(x_{1:i}) \sim N(\mu, \sigma^2). \tag{18}$$

根据多维正态分布条件分布的计算公式, 可以计算出此条件分布的均值和方差:

$$\left. \begin{aligned} \mu &= \mathbf{K}(X, x^*)^T \mathbf{K}(x^*, X)^{-1} g \\ \sigma^2 &= K(x^*, x^*) - \mathbf{K}(X, x^*)^T \mathbf{K}(x^*, X)^{-1} \mathbf{K}(X, x^*) \end{aligned} \right\}. \tag{19}$$

根据此条件分布的均值和方差选择采集函数, 求得该采集函数的极大值所对应的采样点和该采样点的观测值。目前, 常用的采集函数主要有概率提升函数 (probability of improvement, POI)^[12] 及置信边界函数 (upper confidence bound, UCB)^[13]。笔者选用 UCB 采集函数作为概率分布的采集函数, 其函数公式为:

$$GP - UCB(X) = \mu(x) + \sqrt{\beta} \sigma(x). \tag{20}$$

式中 β 用于平衡均值与均方差。

2.3 贝叶斯优化算法流程

贝叶斯优化算法流程为:

- 1) 随机选取一个采样点 A_n , 计算采样点对应的解;
- 2) 根据已有的采样点和对应解的分布形式, 构造贝叶斯网络概率分布函数;
- 3) 根据分布函数的均值和方差计算采集函数;
- 4) 根据采集函数的极大值确定下一个采样点 A_{n+1} ;
- 5) 将 A_{n+1} 放入采样点种群 A 中, 更新种群 A ;
- 6) 判断 A_{n+1} 是否满足条件, 否则重复步骤 2)。

3 k-fold 交叉验证

模型训练一般需要用到训练数据集和测试数据集, 传统的数据集制作方法是按一定的比例进行拆分, 如 70% 用于训练, 30% 用于测试; 但这种静态的拆分方式会导致模型在测试数据集上容易出现过拟合现象, 尤其针对现有压装药成型工艺参数与质量数据集样本数较小的情况下, 为了充分利用全部数据, 采用 k -fold 交叉验证方法划分数据集, 以泛化模型的预测能力。

以 5 折交叉验证为例, 描述交叉验证的基本流

程，如图 1 所示。将原始数据集分成 5 个相似大小的子集，将这些子集中的一个作为验证集，其余 4 个子集作为训练集。首先用训练集上的数据来训练模型，然后在验证集上评估模型性能。重复上述过程 4 次，每次选择不同的验证集。计算 5 次验证的性能指标的平均值，以评估模型的性能。

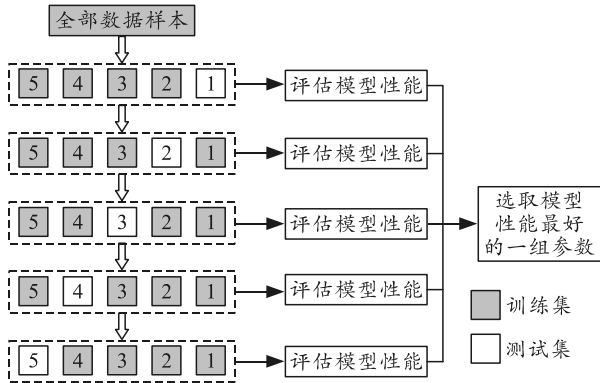


图 1 5 折交叉验证流程

评估预测模型性能有多种指标，如均方误差 (MSE)、 R^2 、平均绝对误差 (MAE) 等。均方误差计算公式如下：

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (21)$$

式中： y_i 为样本目标变量的真实值； \hat{y}_i 为模型预测值；MSE 为回归模型的损失函数，表示真实值与预测值之间的差异程度。 R^2 的计算公式为：

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (22)$$

式中： \bar{y} 为样本目标变量真实值的平均值； R^2 值的范围为 0~1，其越接近于 1，表示模型拟合效果越好。

平均绝对误差计算公式为：

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (23)$$

式中 MAE 为预测值与真实值之间绝对误差的平均值。对于模型性能的评估，使用以上几种评估标准，进而筛选出模型性能最好的参数。

4 贝叶斯-支持向量回归算法应用验证

为了实现小数据样本的压装药质量成型预测，使用压装药工艺参数数据以及其对应的质量数据对贝叶斯-支持向量回归算法模型进行训练。由于药柱原材料获得较困难，便可进行压装实验的次数有限。为了最大程度采集不同工艺参数下药柱的质量情况，根据正交实验法确定的 36 组不同的工艺参数组

合采集压装药数据集，每组数据样本由输入变量 x 及目标变量 y 组成， x 表示压装药成型的工艺参数，即压制温度、保温时间和压力， y 表示压制成型密度。

4.1 数据预处理及相关分析

首先，对这 36 组数据进行数据预处理，将输入变量进行归一化处理。观察数据样本输出变量的分布，发现大部分数据的密度值都分布在 1.7 g/cm^3 左右。接着，对输入变量进行相关性分析，得出每 2 个变量之间的相关性系数，其计算公式为：

$$\rho_{x,y} = \frac{E(x,y) - E(x)E(y)}{\sqrt{E(x^2) - E^2(x)}\sqrt{E(y^2) - E^2(y)}} \quad (24)$$

式中： x 、 y 为任意 2 个变量； E 为期望的计算。将每个变量之间的相关性按式(24)进行计算，并以热力图的形式表示出来，如图 2 所示。

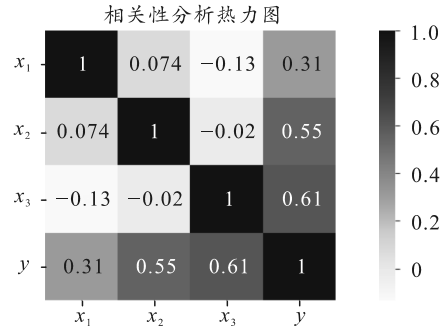


图 2 压装药数据样本相关性分析

图 2 中 x_1 、 x_2 、 x_3 分别表示工艺参数压制温度、保温时间和压力， y 表示压制成型密度。可以发现：压制成型密度与压制压力相关性最大，为 0.61，与压制温度相关性最小，为 0.31。

4.2 贝叶斯-支持向量回归算法基本流程

在对数据进行预处理及相关性分析之后，将数据作为贝叶斯-支持向量回归算法的输入，贝叶斯-支持向量回归算法是指用贝叶斯算法优化支持向量回归的参数。该算法流程如图 3 所示，其主要步骤为：

- 1) 根据压装药过程采集的数据，选取数据样本的 80% 作为训练样本集，20% 作为测试样本集；
- 2) 初始化模型，设置 C 和 γ 参数的取值范围和模型交叉验证的迭代次数；
- 3) 随机挑取一组参数作为初始解；
- 4) 通过高斯过程构造贝叶斯概率分布函数，并更新目标函数模型；
- 5) 根据采集函数确定下一组参数的位置；

- 6) 计算该组参数的概率值, 并更新先验分布;
- 7) 判断新的参数组合是否满足条件, 如果满足条件, 则输出结果 C 和 γ 参数的值, 否则返回 4) 重复迭代, 直到满足最大迭代次数为止;
- 8) 利用优化后的模型对输入数据进行预测, 并输出预测结果。

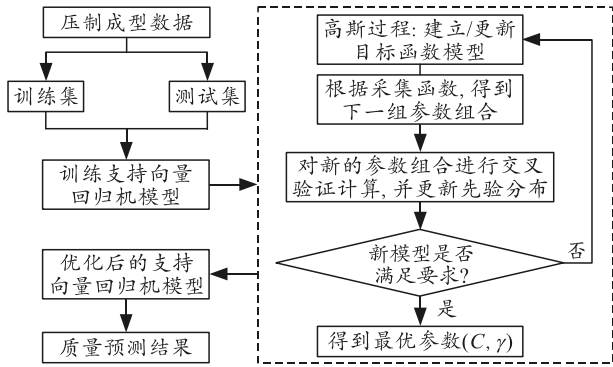


图 3 基于贝叶斯优化支持向量回归模型参数流程

4.3 预测结果对比

根据图 3, 设置参数 C 的取值范围为 $[0.01, 2]$, 参数 γ 的取值范围为 $[0.000 1, 0.01]$, 设置交叉验证的次数为 5, 即 5 折交叉验证。经过不断的迭代, 使目标函数值达到最小, 得到该数据样本下的模型最优参数值。输出优化后模型的预测曲线如图 4 所示。

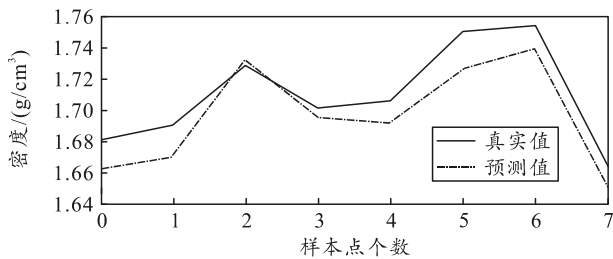


图 4 BO-SVR 模型预测

发现模型的预测值与真实值之间的平均绝对误差为 0.014 7, 预测效果良好。为了验证贝叶斯优化算法对支持向量回归模型是否有优化作用, 将在默认参数下的支持向量回归模型对同样的数据进行预测, 预测结果如图 5 所示。

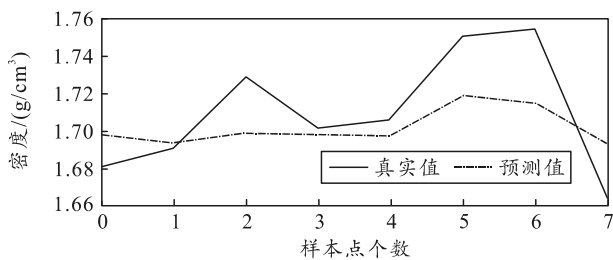


图 5 默认参数的 SVR 模型预测

其预测值与真实值之间的平均绝对误差为

0.020 4, 比 BO-SVR 模型大了 0.005 7。经过 2 个模型的质量预测曲线对比, 发现基于贝叶斯优化的支持向量回归模型拟合效果比原模型有很大提高。

为了说明 2 种模型的预测效果, 依据回归模型常用评价指标对 2 个模型进行对比分析。BO-SVR 算法模型和默认参数下的 SVR 模型对同样的数据样本预测指标如表 1 所示。

表 1 不同模型训练指标

模型名称	R^2	均方误差	方差	平均绝对误差
BO-SVR	0.721 4	0.000 3	0.926 4	0.014 7
传统 SVR	0.359 0	0.000 6	0.431 2	0.020 4

其中回归模型评价指标采用 R^2 、均方误差、方差以及平均绝对误差等 4 种指标, 对 2 种模型的预测效果进行对比。数据表明: BO-SVR 模型的预测效果比传统默认参数下的 SVR 模型训练指标提高了近一倍, 说明贝叶斯优化算法对模型参数具有较大的优化作用。

5 结论

针对压装药成型质量预测问题, 提出一种贝叶斯-支持向量回归算法, 并利用实际压装药数据集对该模型进行验证。实验结果表明: 贝叶斯-支持向量回归比传统的支持向量回归算法预测指标提高了近一倍, 证明了该模型的有效性。

参考文献:

- [1] 马增祥, 卢连军, 张均法, 等. 弹药装药技术的数值研究[J]. 兵工自动化, 2015, 34(4): 92-93, 96.
- [2] 王伟, 马乾伦, 白振华, 等. 基于梯度提升决策树的冷轧高强钢卷力学性能预测[J]. 中国机械工程, 2023, 34(18): 2222-2229.
- [3] 刘琳岚, 高声荣, 舒坚. 基于随机森林的链路质量预测[J]. 通信学报, 2019, 40(4): 202-211.
- [4] 董海, 高秀秀, 魏铭琦. 基于深度学习的完全填充型熔沉积成型零件质量预测方法[J]. 计算机集成制造系统, 2023, 29(1): 200-211.
- [5] 陶磊, 刘检华, 夏焕雄, 等. 基于 B 样条神经网络的熔铸装药温度场预测[J]. 兵工学报, 2023, 44(5): 1339-1349.
- [6] BAZI Y, MELGANI F. Toward an Optimal SVM Classification System for Hyperspectral Remote Sensing Images[J]. Geoscience and Remote Sensing, 2006, 44(11): 3374-3385.
- [7] 侯红玲, 陈鑫, 常向龙, 等. 基于遗传算法优化 BP 神经网络的内螺纹冷挤压质量预测[J]. 塑性工程学报, 2022, 29(1): 102-109.