

基于机器学习的镍基单晶高温合金蠕变 寿命预测模型研究

杜晓明¹, 陆瑶¹, 刘纪德²

(1. 沈阳理工大学材料科学与工程学院, 沈阳 110159; 2. 中国科学院金属研究所, 沈阳 110016)

摘要: 构建合适的镍基单晶合金蠕变寿命预测模型, 对于我国航空发动机叶片设计、强度分析和寿命预测具有重要意义。采用多项式回归、最近邻回归、支持向量机回归、决策树回归四种机器学习算法, 建立镍基单晶高温合金蠕变寿命与合金成分、微观组织和蠕变工艺参数的关系模型, 为镍基单晶高温合金的蠕变性能调控提供了新方法。基于蠕变寿命预测模型, 系统地比较了四种算法和特征选择对模型性能的影响。结果表明, 支持向量机回归模型的预测结果最优, 相关性较高的四个特征依次为 γ' 固溶温度、Ta、W、Re。研究结果可为获得更有效的镍基单晶高温合金蠕变性能预测方法提供参考。

关键词: 机器学习; 镍基单晶高温合金; 特征选择; 蠕变寿命

中图分类号: TG132.3; TP391 **文献标志码:** A **DOI:** 10.3969/j.issn.1003-1251.2025.01.007

Research on Prediction Model of Creep Life of Nickel-based Single Crystal Superalloys Based on Machine Learning

DU Xiaoming¹, LU Yao¹, LIU Jide²

(1. Shenyang Ligong University, Shenyang 110159, China;
2. Institute of Metal Research, Chinese Academy of Sciences, Shenyang 110016, China)

Abstract: Building the suitable creep life prediction model of nickel-based single crystal superalloys is of great significance for the design, strength analysis, and life prediction of aircraft engine blades in China. In this paper, four machine learning algorithms are used, including polynomial regression, nearest neighbor regression, support vector machine regression and decision tree regression. The relationship models between the creep life of nickel-based single crystal superalloys and the alloy composition, microstructure and creep process parameters are established, which provide a new method for the control of creep properties of nickel-based single-crystal superalloys. Based on the creep life prediction model, the effects of algorithm selection and feature selection on the comprehensive performance of the model are systematically compared. The results show that the SVR regression model has the best prediction, and the four features with relatively high correlations are in order as follows: γ' solution temperature, Ta, W, and Re. The research results can provide a reference for the prediction method of nickel-based single crystal superalloys with more effective creep properties.

收稿日期: 2024-03-21

基金项目: 国家自然科学基金项目(12375305); 辽宁省应用基础研究计划项目(2023JH2/101300233); 辽宁省教育厅高等学校基本科研项目(JYTZD20230004)

作者简介: 杜晓明(1976—), 男, 教授, 博士。

Key words: machine learning; nickel-based single crystal superalloys; feature selection; creep life

镍基单晶高温合金作为航空发动机涡轮叶片的关键材料,具有优异的承温性能和抗蠕变性能^[1]。但由于其成分体系复杂、元素间相互作用影响显著且涉及大量昂贵的金属元素、单晶制造技术成本高、蠕变试验周期长,同时缺乏成分、组织和性能之间的定量关系,通过传统试错法提高合金的蠕变抗力需要耗费大量的时间和成本。并且,合金设计的突破往往需要对复杂的多组分/多相体系进行全面的理解,这种开发方式步骤繁琐,研发到应用周期冗长,且达不到预期性能,仅靠人力无法挖掘材料特征与性能之间的深层联系。

近年来,随着材料数据库的不断积累,材料信息学得到了飞速发展^[2],其独特的 DIKW (data→information→knowledge→wisdom) 层级研究模型为机器学习和大数据背景下建立非显著参考体系分布奠定了基础。Jiang 等^[3]针对铜合金材料性能的小数据集,将材料成分和材料性能分别作为输入和目标量,通过建立机器学习模型表示两者之间的函数关系。Raccuglia 等^[4]利用失败实验数据结合支持向量机 (SVM) 回归算法成功建立了无机、有机杂化材料结晶过程反应模型并预测了新化合物的形成条件,准确率达 89%。

大多数机器学习的任务是做出预测,系统在给定的训练示例上进行泛化。泛化的方式主要有两种:一种是基于实例的学习,另一种是基于模型的泛化。Han 等^[5]基于从文献中收集的小数据集,通过数据驱动的机器学习方法,开发了一种基

于双模型联动的方法预测合金的蠕变寿命。Liu 等^[6]基于设计的适应度函数,提出分而治之自适应 (DCSA) 学习方法,该方法基于随机森林 (RF)、支持矢量回归 (SVR)、高斯过程回归 (GPR)、套索回归 (LR) 和脊回归 (RR) 共五个最优机器学习模型,将上述机器学习方法对应不同聚类结果,经训练后,结果显示聚类模型可以很好地将数据区分,且蠕变寿命预测模型的预测值与实验值的误差在合理范围内,大大提高了机器学习模型的预测能力。

本文通过四种机器学习算法分别建立镍基单晶高温合金成分、工艺、微观组织参数和蠕变寿命的机器学习模型,优化出镍基高温合金蠕变寿命预测模型的参数,以便应用于合金成分与工艺的逆向设计。

1 数据和机器学习模型

1.1 数据集与描述符

本实验数据集部分来源于中科院金属研究所的研究结果,其余来自于文献[7-12]。数据处理后共计 550 条镍基单晶高温合金蠕变寿命数据,在机器学习模型中输入特征标签所对应的描述符,如表 1 所示。数据集样本中共包含 24 个输入特征 (14 种合金元素、8 个微观参数、2 个工艺参数)。

表 2 为镍基单晶高温合金蠕变寿命数据的统计分析结果。

表 1 特征标签所对应的描述符

Table 1 Descriptors corresponding to the feature labels

编号	特征标签	描述符	编号	特征标签	描述符
1	Ni 的质量分数/%	Ni	13	Nb 的质量分数/%	Nb
2	Re 的质量分数/%	Re	14	Hf 的质量分数/%	Hf
3	Co 的质量分数/%	Co	15	γ/γ' 错配度 δ /%	LM
4	Al 的质量分数/%	Al	16	有效扩散系数 $\lg(x)/(m^2 \cdot s^{-1})$	EDC
5	Ti 的质量分数/%	Ti	17	γ' 体积分/%	$V_{\gamma'}$
6	W 的质量分数/%	W	18	γ' 固溶温度/ $^{\circ}C$	$T_{\gamma'}$
7	Mo 的质量分数/%	Mo	19	反相畴界能/ $(mJ \cdot m^{-2})$	APDB
8	Cr 的质量分数/%	Cr	20	剪切模量/GPa	SM
9	Ta 的质量分数/%	Ta	21	层错能/ $(mJ \cdot m^{-2})$	SFE
10	C 的质量分数/%	C	22	γ 基体通道宽度/nm	D_{γ}
11	B 的质量分数/%	B	23	蠕变测试温度/ $^{\circ}C$	Test. T
12	Y 的质量分数/%	Y	24	蠕变测试压力/MPa	Test. S

表2 镍基单晶高温合金蠕变寿命数据描述符的统计分析结果

Table 2 Statistical analysis of creep life data descriptors for nickel-based single crystal superalloys

	描述符	平均值	标准差	最小值	25%分位数	50%分位数	75%分位数	最大值
输入	Ni/%	63.42	3.42	52.71	60.99	63.73	65.70	77.26
	Re/%	2.85	2.24	0.00	0.00	3.00	4.50	7.60
	Co/%	6.76	3.51	0.00	4.50	7.50	9.83	16.86
	Al/%	5.37	0.82	1.05	5.07	5.50	5.80	7.50
	Ti/%	1.04	1.37	0.00	0.00	0.60	1.50	9.00
	W/%	5.72	2.65	0.00	4.00	5.50	6.79	18.60
	Mo/%	1.03	1.10	0.00	0.00	0.83	1.50	10.50
	Cr/%	6.78	3.12	0.00	5.00	7.00	8.00	20.50
	Ta/%	6.55	2.78	0.00	5.00	7.00	8.51	13.94
	C/%	0.01	0.02	0.00	0.00	0.00	0.00	0.10
	B/%	0.00	0.01	0.00	0.00	0.00	0.00	0.05
	Y/%	0.08	0.45	0.00	0.00	0.00	0.00	5.40
	Nb/%	0.22	0.44	0.00	0.00	0.00	0.30	2.59
	Hf/%	0.15	0.30	0.00	0.00	0.10	0.15	2.00
	LM/%	0.12	0.20	-0.38	-0.03	0.09	0.23	0.74
	EDC/(m ² ·s ⁻¹)	19.30	1.26	15.89	18.27	19.35	20.16	22.06
	V _γ /%	66.31	8.85	16.45	63.57	67.29	70.68	95.89
	T _γ /°C	1 267.67	37.09	1 121.45	1 248.31	1 269.46	1 296.10	1 333.92
	APDB/(mJ·m ⁻²)	211.91	4.78	204.04	209.10	211.02	213.13	230.08
	SM/GPa	59.06	1.99	53.43	57.39	59.37	60.54	63.64
SFE/(mJ·m ⁻²)	34.85	9.33	17.07	28.34	32.89	39.85	71.56	
D _γ /nm	34.37	47.04	1.75	19.45	23.73	30.69	437.56	
Test. T/°C	1 028.14	70.13	927.00	982.00	991.50	1 093.00	1 200.00	
Test. S/MPa	184.35	72.44	70.00	135.25	163.50	248.20	345.00	
输出	蠕变寿命/h	2.35	0.54	1.48	1.91	2.28	2.69	4.18

1.2 数据预处理

由于数据集的质量直接影响机器学习模型的预测能力,且大部分机器学习算法无法在缺失的特征上工作,故对数据进行预处理、缺失值处理、异常值处理和去重。在若干行训练数据集中,每行数据皆涉及到14种合金元素,但并非所有行都出现某一个或几个合金元素数据为空的情况,故引入Pandas数据处理库将合金元素描述符转化为DataFrame二维标签化数据结构,用select_types.columns找出所有的描述性列中缺失值,再用fillna将缺失值用0填充。对于数据值完全相同的多条数据记录则用drop_duplicates方法直接去重。数据集中将所有输入列分为输入特征和输出目标变量,对于数据集中输入特征完全相同,但输出目标变量不同的情况,先用groupby将输入特征分组,再用idmax在每个分组中选取目标变量

较优的记录。

1.3 机器学习模型的建立和评价

基于机器学习回归算法是在PyCharm集成开发环境中建立不同蠕变寿命的最优预测模型,基本流程如图1所示。图中:PR为多项式回归;KNN为最近邻回归;SVR(RBF)为支持向量机回归;DT为决策树回归。

输入特征引入数据库后,选择Filter过滤法中的Pearson、Spearman和Kendall三种系数对数据特征进行相关性分析。通过减掉均值并将数据缩放到单位方差来标准化特征,即均值-方差标准化,Z-Score标准化转换为

$$Z = \frac{X_i - \mu}{\sigma} \quad (1)$$

式中:Z为标准分数; μ 为总体数据的均值; σ 为总体数据的标准差; X_i 为数据集中的观测值。

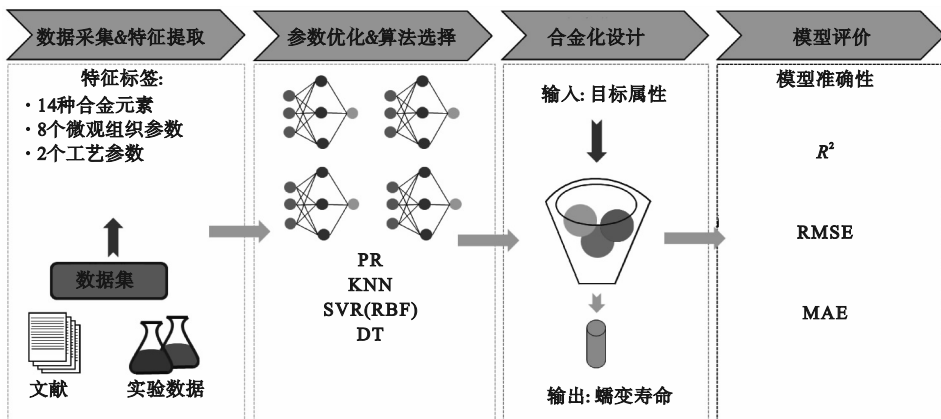


图 1 最优预测模型建立流程

Fig. 1 The establishing process of the optimal prediction model

在 Scikit-learn 中使用 StandardScaler 实现特征标准化,此时特征服从标准正态分布,方差为 1,均值为 0。标准化完成后对数据进行拆分,采用

train_test_split 方法,按照 80% 训练集、10% 测试集、10% 验证集进行划分,采用 K 折交叉验证法, K 取 10,十折交叉验证示意如图 2 所示。

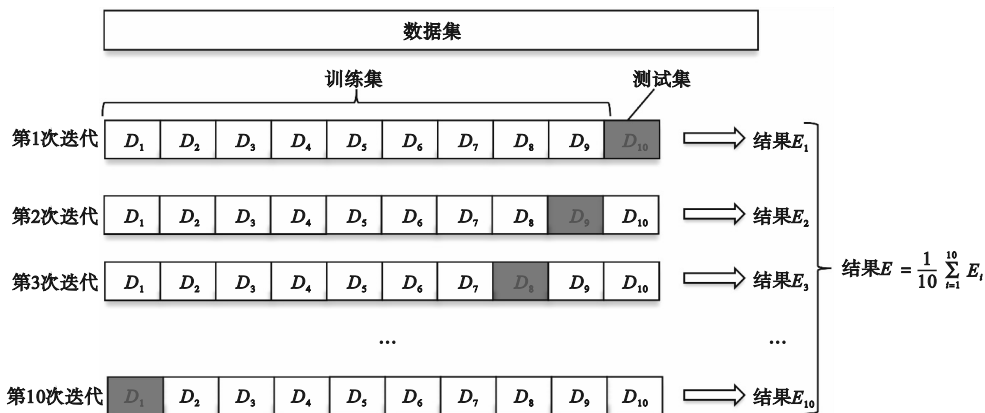


图 2 十折交叉验证示意图

Fig. 2 Ten-fold cross validation diagram

使用 PR、KNN、SVR(RBF)、DT 四种机器学习回归算法,分别对训练模型进行参数调优,得到对应的四种最佳模型。最后引入 metrics 库,使用决定系数 (R^2)、预测误差均方根值 (RMSE)、预测误差平均绝对值 (MAE) 进行模型性能评估,计算式分别为

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

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}| \quad (4)$$

式中: n 为测试集中样本的数量; y_i 为样本真实值; \hat{y} 为模型预测值; \bar{y} 为测试集中样本真实值的

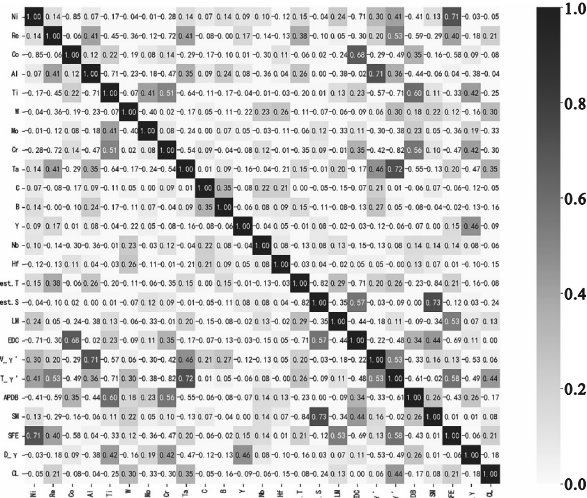
平均值。

2 结果与讨论

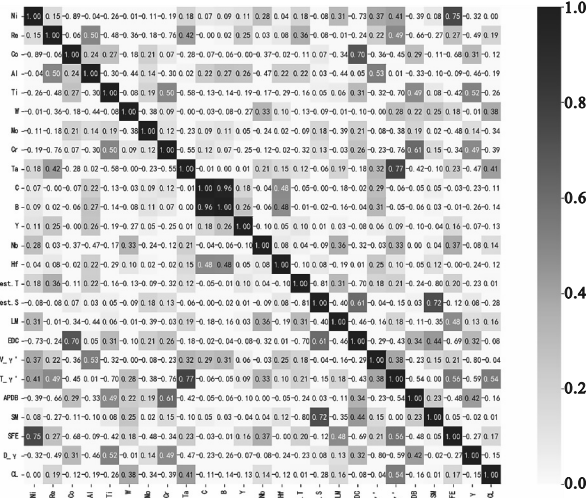
2.1 特征选择对机器学习模型性能的影响

采用 Pearson、Kendall、Spearman 三种相关系数对特征进行选择和分析,结果如图 3 所示。

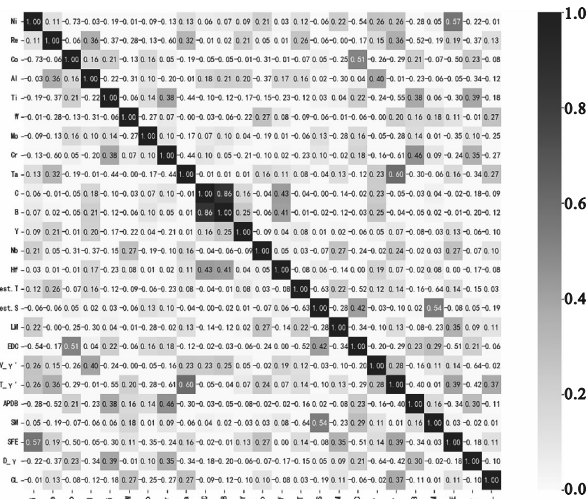
由图 3 可知: Pearson 相关系数中,输入特征 SM 与 Test. S 相关性最高,为 0.73,其值小于 0.8,呈中度相关; Kendall 相关系数中,输入特征 C 与 B 相关性最高,为 0.86,大于 0.8,呈高度相关; Spearman 相关系数中,输入特征 C 与 B 相关性最高,为 0.96,接近 1,呈显著性相关。由此可见, C 元素和 B 元素对蠕变寿命会产生重要影响,且存在多重共线性的情况。这是因为 C 和 B 两



(a) Pearson



(b) Spearman



(c) Kendall

图3 三种相关性热力图

Fig. 3 Three correlation heat maps

种元素可能共存形成复合碳化物,有助于提高合金的高温硬度和抗蠕变性。然而在训练机器学习模型时,输入特征之间相关性过高可能会导致拟合训练数据,导致模型过分关注 C 和 B,忽视其他可能对泛化蠕变寿命性能有帮助的特征,从而导致模型参数的大幅变化。

从图3(a)~3(c)中最后一列或最后一行可以看出,每个特征对于蠕变寿命列的系数的绝对值都大于0,表明24个特征对于蠕变寿命都有各自贡献度。Pearson、Spearman 和 Kendall 三种特征选择方法中,相关性最高的四个特征相同,且顺序由高到低皆为: γ' 固溶温度(T_γ)、Ta、W、Re。 T_γ 的选择对合金的性能具有重要影响, T_γ 升高可能导致 γ' 相的溶解,从而可以提高合金塑性,但同时也降低了合金的强度和抗蠕变性。 T_γ 的选择还影响晶粒尺寸,较高的 T_γ 可能导致较大的晶粒。在高温环境下,较大的晶粒可能更易发生晶粒滑移和变形,影响合金的蠕变行为。Ta、W、Re 三种元素都是高熔点金属元素,其含量会显著影响合金的蠕变寿命。Ta 的适量加入能有效降低层错能,提高形成不全位错剪切 γ' 相能力和形成显微孪晶能力,增加蠕变抗力,有效改善合金蠕变性能^[9]。W 可以提高高温合金的硬度和强度,较高的硬度和强度有助于防止蠕变变形,从而延长蠕变寿命。Re 是一种具有良好高温强度的元素,其加入可以提高合金在高温下的抗变形能力,提高蠕变寿命,促进位错网络的形成及抑制拓扑密堆相(TCP)的析出^[10],从而改善合金的高温稳定性。综上, T_γ 、Ta、W、Re 四种特征在蠕变寿命预测模型中都有重要意义,表明数据集具有一定的可靠性,使机器学习模型稳定性更好,提高模型对未知数据预测的可信度。

2.2 算法选择对机器学习模型性能的影响

本文比较了 PR、KNN、SVR (RBF)、DT 四种算法最佳模型的真实值与预测值,结果如图4所示。由图4可知,测试集数据预测的四种最优模型预测结果都比较好,其中 SVR (RBF) 算法的预测结果最集中,拟合效果最好。

图5给出了四种最优预测模型的 R^2 、RMSE 和 MAE 评价指标结果,其中 SVR (RBF) 算法的预测模型 R^2 最高。

由前述可知,输入特征中存在较多非线性情况,PR 算法可以尝试捕捉并一定程度上处理这些非线性关系,故在四种算法中其 R^2 值相对较低,但这也容易导致模型过拟合。KNN 算法适用于

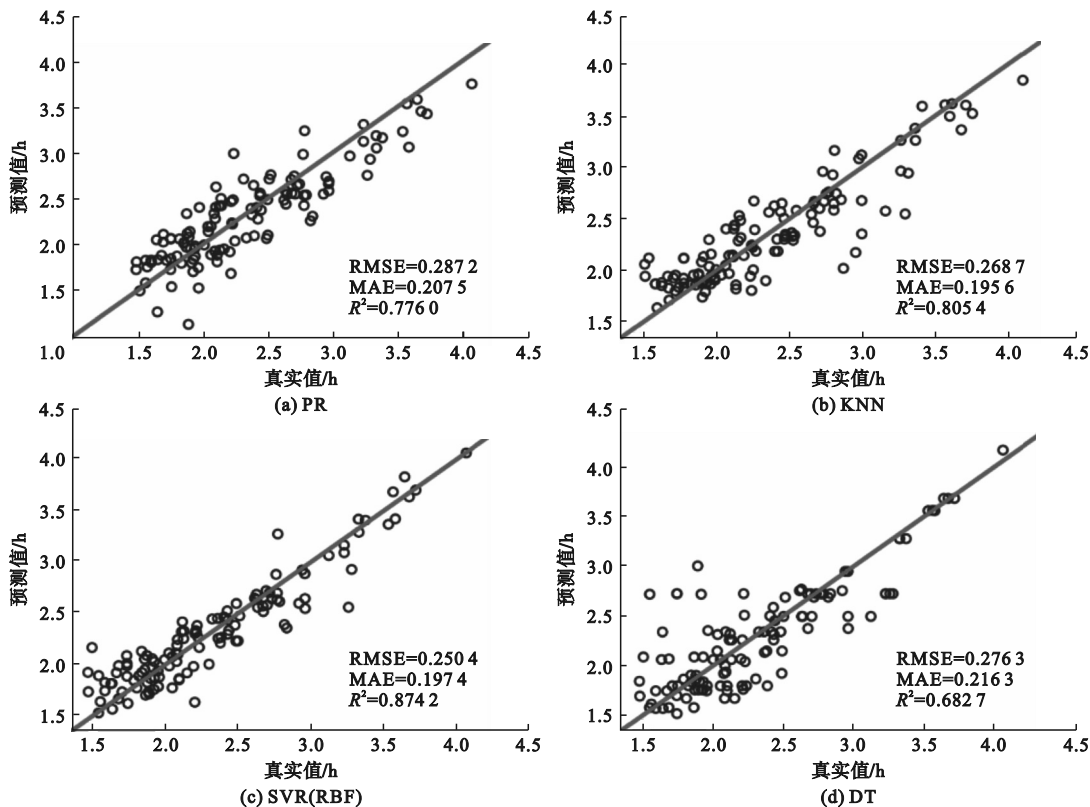


图 4 四种最优回归模型计算蠕变寿命真实值与预测值对比

Fig. 4 Comparisons between actual and predicted values of creep life calculated by four optimal regression models

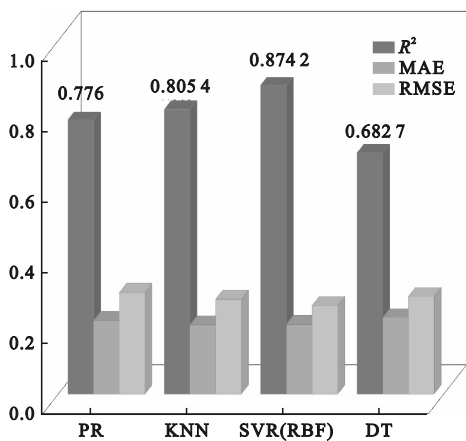


图 5 四种预测模型评价指标

Fig. 5 Evaluation indicators of four prediction models

挖掘蠕变寿命时间的影响因素,但由于输入特征多导致计算成本高。SVR(RBF)算法在处理高维空间中表现较好,RBF核能够将数据映射到无穷维空间,适应数据的复杂结构,且在空间中产生的“径向”特性使其对数据分布的适应性较强,即使在高维数据中也能有效地捕捉数据中不同维度之间的关系。由图 5 可知,SVR(RBF)算法所建立的模型为最优算法模型。经过重要参数(例如核函数、错误项的惩罚因子等)调优后, R^2 达到

0.8742。DT 算法可以处理复杂的非线性关系,但其超参数(例如树的深度、分裂节点的最小样本数等)对模型性能有一定影响,训练后最优参数(见表 3)模型 R^2 值仅为 0.6827,在四种算法中最差。

表 3 四种最优回归模型参数

Table 3 Parameters of four optimal regression models

回归算法	模型内部设置的超参数
PR	degree = 18
KNN	n_neighbors = 4
SVR(RBF)	kernel = rbf, C = 3.42, epsilon = 0.1
DT	max_depth = 7, min_samples_leaf = 1, min_samples_split = 2

3 结论

本文基于国内外镍基单晶高温合金蠕变寿命数据,建立了包含 24 个输入特征的蠕变寿命预测机器学习模型,采用 Filter 过滤法中的 Pearson、Kendall、Spearman 三种特征选择方法,筛选和分析不同输入特征对蠕变寿命预测模型的影响,通过 R^2 、RMSE、MAE 进行模型预测性能评估,分析

了不同回归算法对蠕变寿命预测模型的影响,主要结论如下。

1) 基于 Pearson、Kendall、Spearman 相关性分析可知,每一个输入特征对蠕变寿命预测模型都有贡献,且相关性最高的四个特征相同,按相关性由高到低排序为 T_{γ} 、Ta、W、Re。

2) 基于四种回归算法建立的蠕变寿命预测模型中,SVR(RBF) 回归算法的模型预测能力最优,在测试集上的决定系数 R^2 为 0.874 2, RMSE 为 0.250 4, MAE 为 0.197 4。

3) 四种回归算法中超参数的调优极为重要,其中学习率、决策树深度、迭代次数、正则化参数和核函数对模型预测性能影响最为重要。

参考文献 (References):

- [1] REED R C. The superalloys: fundamentals and applications [M]. Cambridge, UK: Cambridge University Press, 2006.
- [2] 李云琦, 刘伦洋, 陈文多, 等. 材料基因组学的发展现状、研究思路与建议 [J]. 中国科学: 化学, 2018, 48 (3): 243 - 255.
LI Y Q, LIU L Y, CHEN W D, et al. Materials genome: research progress, challenges and outlook [J]. Scientia Sinica (Chimica), 2018, 48 (3): 243 - 255. (in Chinese)
- [3] JIANG L, WANG C S, FU H D, et al. Discovery of aluminum alloys with ultra-strength and high-toughness via a property-oriented design strategy [J]. Journal of Materials Science & Technology, 2022, 98: 33 - 43.
- [4] RACUGLIA P, ELBERT K C, ADLER P D F, et al. Machine-learning-assisted materials discovery using failed experiments [J]. Nature, 2016, 533: 73 - 76.
- [5] HAN H Y, LI W D, ANTONOV S, et al. Mapping the creep life of nickel-based SX superalloys in a large compositional space by a two-model linkage machine learning method [J]. Computational Materials Science, 2022, 205: 111229.
- [6] LIU Y, WU J M, WANG Z C, et al. Predicting creep rupture life of Ni-based single crystal superalloys using divide-and-conquer approach based machine learning [J]. Acta Materialia, 2020, 195: 454 - 467.
- [7] 祝祥, 杜晓明, 刘纪德. DD419 镍基单晶高温合金 980 °C 下低周疲劳行为研究 [J]. 沈阳理工大学学报, 2023, 42 (4): 69 - 74.
ZHU X, DU X M, LIU J D. Study on low cycle fatigue behavior of DD419 nickel base single crystal superalloy at 980 °C [J]. Journal of Shenyang Ligong University, 2023, 42 (4): 69 - 74. (in Chinese)
- [8] 付佳博, 王晨充, MATEO Carlos Gracia, 等. 物理冶金信息指导机器学习的镍基单晶高温合金蠕变寿命预测 [J]. 中国材料进展, 2023, 42 (9): 722 - 731.
FU J B, WANG C C, GRACIA M C, et al. Creep life prediction of Ni-based single crystal superalloys by physical metallurgy information guided machine learning [J]. Materials China, 2023, 42 (9): 722 - 731. (in Chinese)
- [9] 杨志昆, 王浩, 张义文, 等. Ta 含量对镍基粉末高温合金高温蠕变变形行为和性能的影响 [J]. 金属学报, 2021, 57 (8): 1027 - 1038.
YANG Z K, WANG H, ZHANG Y W, et al. Effect of Ta content on high temperature creep deformation behaviors and properties of PM nickel base superalloys [J]. Acta Metallurgica Sinica, 2021, 57 (8): 1027 - 1038. (in Chinese)
- [10] 叶新玥, 赵云松, 张迈, 等. 镍基单晶高温合金中 Ru 元素作用的研究进展 [J]. 铸造技术, 2023, 44 (2): 114 - 121.
YE X Y, ZHAO Y S, ZHANG M, et al. Research progress on Ru element in nickel-based single crystal superalloys [J]. Foundry Technology, 2023, 44 (2): 114 - 121. (in Chinese)
- [11] KOSTIUCHENKO T, KÖRMANN F, NEUGEBAUER J, et al. Impact of lattice relaxations on phase transitions in a high-entropy alloy studied by machine-learning potentials [J]. NPJ Computational Materials, 2019, 5: 55.
- [12] SONG W, WANG X G, LI J G, et al. Effect of ruthenium on microstructure and high-temperature creep properties of fourth generation Ni-based single-crystal superalloys [J]. Materials Science and Engineering A, 2020, 772: 138646.

(责任编辑: 徐淑姣)