

大学生论文检测系统
文本复制检测报告单 (全文标明引文)

No: ADBD2026R_2026012417203520260521194745485225514564

检测时间: 2026-05-21 19:47:45

篇名: 基于机器学习的水工混凝土结构修复界面粘结性研究_李陇豫
作者: 李陇豫
指导教师:
检测机构:
文件名: 基于机器学习的水工混凝土结构修复界面粘结性研究_李陇豫.docx
检测系统: 大学生论文检测系统
检测类型: 大学生论文
检测范围: 中国学术期刊网络出版总库
中国博士学位论文全文数据库/中国优秀硕士学位论文全文数据库
中国重要会议论文全文数据库
中国重要报纸全文数据库
中国专利全文数据库
图书资源
优先出版文献库
大学生论文联合比对库
互联网资源(包含贴吧等论坛资源)
英文数据库(涵盖期刊、博硕、会议的英文数据以及德国Springer、英国Taylor&Francis 期刊数据库等)
港澳台学术文献库
互联网文档资源
源代码库
CNKI大成编客-原创作品库
时间范围: 1900-01-01至2026-05-21

检测结果

去除本人文献复制比: 10%

去除引用文献复制比: 3.5%

单篇最大文字复制比: 3.7% (基于机器学习的UHPC多目标优化)

跨语言检测结果: -

总文字复制比: 10%

重复字数: [3668] 总段落数: [7]
总字数: [36848] 疑似段落数: [7]
单篇最大重复字数: [1369] 前部重合字数: [723]
疑似段落最大重合字数: [1280] 后部重合字数: [2945]
疑似段落最小重合字数: [79]



文字复制部分 10%
无问题部分 90%

指标: ☐ 疑似剽窃观点 ☒ 疑似剽窃文字表述 ☐ 疑似整体剽窃 ☐ 过度引用

相似表格: 0 相似公式: 检测中 疑似文字的图片: 0

1.9% (92)	1.9% (92)	中英文摘要等 (总4854字)
19.7% (950)	19.7% (950)	1 绪论 (总4817字)
19.6% (1280)	19.6% (1280)	2 界面粘结性能与机器学习算法 (总6519字)
1.9% (145)	1.9% (145)	3 水工混凝土修复界面性能数据库构建 (总7830字)
1.4% (79)	1.4% (79)	4. 粘结性能预测模型开发与评估 (总5508字)
2.4% (96)	2.4% (96)	5 界面性能影响机理解析 (总4032字)
31.2% (1026)	31.2% (1026)	6 结论与展望 (总3288字)

missing failure morphology records, a physical discrimination criterion based on mechanical equilibrium and strength comparison method was introduced to realize automatic label calibration of interface failure mode (FM), laying a data foundation for dual-target modeling.

Secondly, at the model development level, this paper compared the adaptability of algorithms such as Random Forest (RF), Artificial Neural Network (ANN) and Gradient Boosting Regression (GBR). The study found that the strength prediction model constructed by GBR algorithm has significant advantages in dealing with small-sample and high-dimensional data. The hyperparameters were globally optimized through Grid Search and 5-fold cross-validation. The coefficient of determination of the model on the independent test set reached 0.8761, and the Root Mean Square Error (RMSE) was as low as 1.7692 MPa. At the same time, the failure morphology discrimination model developed by Random Forest classifier has a prediction accuracy of up to 95.45%, which successfully realizes the technical leap from numerical prediction to qualitative mode prediction.

Finally, using the powerful nonlinear mapping ability of machine learning models, this paper deeply analyzed the evolution law of interface performance. The study quantitatively revealed the 18.6% water gain effect of matrix water saturation (MC) on interface bond strength, and determined the optimal threshold range of MC between 50% and 60%. Exceeding this range will lead to strength reversal due to the interface water film effect. In addition, the regulation mechanism of joint angle (JA) on interface stress distribution was revealed, and the critical stress conditions for the transformation of failure path from "interface brittle peeling" to "matrix cohesive failure" were quantified.

The "data-model-mechanism" closed-loop system constructed in this study not only provides a high-precision quantitative tool for the intelligent repair design of hydraulic concrete structures in complex service environments, but also provides an important theoretical basis for formulating digital repair construction standard manuals.

Keywords: Hydraulic Concrete; UHPC; Bond Performance; Machine Learning; Gradient Boosting Regression

目录

摘要I

AbstractII

目录IV

1 绪论1

1.1 课题研究背景及意义1

1.2 国内外研究现状2

1.2.1 界面处理工艺与宏观力学性能研究2

1.2.2 复杂环境下的界面水化与微观机理2

1.2.3 几何构造与界面剪切性能强化3

1.2.4 人工智能在界面性能评估中的应用及发展趋势3

1.3 主要研究内容与研究目标4

1.3.1 水工修复界面异构数据库的构建与特征工程4

1.3.2 粘结性能双维度预测模型的开发与调优4

1.4 技术路线5

1.4.1 数据构建与特征对标5

1.4.2 模型开发与性能标定5

1.4.3 深度机理挖掘与权重解析5

2 界面粘结性能与机器学习算法6

2.1 UHPC-NC界面粘结性能分析6

2.1.1 物理锚固与机械啮合作用6

2.1.2 化学胶接作用6

2.1.3 应力传递与接缝构造逻辑6

2.2 机器学习算法原理6

2.2.1 机器学习理论6

2.2.2 机器学习于水工结构应用7

2.2.3 梯度提升回归 (GBR) 算法7

2.2.4 随机森林 (RF) 分类算法7

2.2.5 神经网络 (ANN) 算法8

2.3 模型评估指标体系9

2.3.1 粘结强度回归模型评价指标9

2.3.2 破坏模式分类模型评价指标10

2.4 界面性能的多尺度耦合机理12

2.5 机器学习方法在小样本工程数据中的适用性12

2.6 主要变量与缩略语统一说明13

2.7 数据稀疏性与样本边界问题13

3 水工混凝土修复界面性能数据库构建15

3.1数据来源与样本收集15

3.2 多维度特征参数提取19

3.2.1 材料参数（CS：基体强度）19

3.2.2 几何参数（CA 与 JA）19

3.2.3 施工与环境参数（ST 与 MC）19

3.3 数据库预处理与完备性标注20

3.3.1 数据清洗与归一化处理20

3.3.2 破坏模式（FM）的物理逻辑标注准则21

3.3.3 数据库完备性评估21

3.3.4 数据分布特征与统计概况21

3.3.5 数据库统计特征与分层规律说明24

3.4 数据清洗、建模与解释流程总览24

3.5 模型输入空间的覆盖范围与外推风险25

3.6 典型样本与数据库边界解释26

4. 粘结性能预测模型开发与评估27

4.1 不同机器学习算法的适应性对比27

4.1.1 随机森林、梯度提升回归、人工神经网络的训练表现对比27

4.1.2 算法选型结论28

4.2 粘结强度回归模型的训练与验证28

4.3 破坏模式分类模型的训练与验证31

4.4 模型训练、调参与稳定性验证说明33

4.5 破坏模式物理判定与典型样本分析34

4.6 模型误差分解与边界样本解释35

4.7 参数搜索、复现步骤与结果稳定性说明35

4.8 回归与分类结果的工程解释及边界条件36

5 界面性能影响机理解析37

5.1 特征重要性与全局贡献分析37

5.2 偏依赖分析与阈值识别38

5.3 工程工况耦合与服役环境解释39

5.4 微观机理与破坏演化路径40

5.5 结果补充分析与文献对比42

5.6 与已有研究的对比及工程启示43

5.7 工程转化路径与研究框架总结44

5.8 综合施工建议与未来迭代路径44

6 结论与展望45

6.1 主要结论45

6.2 创新点与不足45

6.3 未来工作展望46

参考文献47

致谢49

附录A 代表性样本数据节选50

附录B 模型超参数设置与评价指标说明51

2. 1 绪论

总字数：4817

相似文献列表

去除本人文献复制比：19.7%(950) 去除引用文献复制比：0.9%(41) 文字复制比：19.7%(950) 疑似剽窃观点：(0)

1	PEF-UHPC与普通混凝土剪力键界面抗剪性能研究	12.1% (585)
	韩志浩(导师：李晰) - 《青岛理工大学硕士论文》 - 2025-06-01	是否引证： 是
2	2506674125640446710176_韩志浩_PEF-UHPC与普通混凝土剪力键 界面抗剪性能研究	11.9% (574)
	韩志浩 - 《大学生论文联合比对库》 - 2025-03-21	是否引证： 否
3	基于机器学习的UHPC多目标优化	5.1% (247)
	柯超(导师：曾磊;Hua Zhang) - 《长江大学硕士论文》 - 2024-04-01	是否引证： 是
4	CRTSIII型板式轨道充填层及其粘结界面的性能与微结构特征	1.5% (72)
	李文旭(导师：谢友均) - 《中南大学博士论文》 - 2022-05-01	是否引证： 否
5	轻集料对UHPC体积稳定性及耐久性能的影响	1.0% (47)
		是否引证： 否

孙二保 - 《大学生论文联合比对库》 - 2019-06-09		
6	机制砂-超高性能混凝土多掺合料协同优化机制及性能研究 肖阳;赵振华;陈尚伟;陈波; - 《绿色建筑》 - 2026-03-20	0.9% (41) 是否引证: 否
7	装配式建筑综述论文 (通用18篇) (第18页) (68125字). - 《网络 (http://www.lw54.com/) 》 - 2024	0.7% (35) 是否引证: 否
8	“双碳”背景下装配式建筑在广西北部湾地区推广的可行性分析 陆杰;李敏; - 《中国建筑装饰装修》 - 2024-12-05	0.6% (30) 是否引证: 否
9	抗冻. - 《网络 (https://www.fx361.co) 》 - 2025	0.5% (24) 是否引证: 否
10	装配式建筑成本分析及优化研究 李雪 - 《大学生论文联合比对库》 - 2020-04-24	0.5% (23) 是否引证: 否
原文内容		

1 绪论

1.1 课题研究背景及意义

党的二十大报告深刻指出，要加快发展方式绿色转型，积极稳妥推进碳达峰碳中和，推动工业、建筑、交通等领域的清洁低碳转型。建筑业作为国民经济的支柱产业，具有经济带动作用强、产业周期波动小等突出优势，其发展质量直接关系到城乡建设的成色和民生福祉[8]。近年来，随着国家对绿色建筑和可持续发展的重视，组合结构与装配式建筑作为重要的低碳建造方式，得到了快速发展。自2016年《国务院办公厅关于大力发展装配式建筑的指导意见》[20]发布以来，装配式建筑在政策引导、技术支撑、产业发展等方面取得了显著成效，成为推动建筑业转型升级的重要抓手。而组合结构作为成本低效果好，也逐步步入大众视野。

水工混凝土结构（如大坝、泄洪洞、溢洪道等）在长期服役过程中，不可避免地受到环境侵蚀、高速水流冲磨、冻融循环及材料老化等因素的影响，导致结构表面出现剥蚀、裂缝及露筋等病害。这不仅削弱了结构的承载能力，更严重威胁到水利工程的整体安全性。

超高性能混凝土（UHPC）凭借其超高的抗压强度、卓越的耐久性及致密的微观结构，已成为水工结构修复与加固领域的首选材料。然而，修复工程的核心在于新旧混凝土界面的粘结性能。由于修复界面是组合结构的力学薄弱环节，其粘结强度与失效模式受界面处理工艺、基体物理状态及外部服役环境等多因素的高度非线性耦合影响。传统经验公式难以精准量化这些变量间的复杂关联，导致修复设计的冗余度过高或安全储备不足。

因此，引入机器学习（ML）技术[2]，构建能够精准预测粘结强度并判别破坏模式的智能化模型，并定量揭示复杂工况下的性能演化机理，对于提升水工修复工程的科学性与预警能力具有重要的理论价值与现实意义。

针对水工混凝土结构在复杂服役环境下出现的性能退化问题，采用超高性能混凝土（UHPC）进行加固修复已成为保障工程耐久性的关键手段。然而，修复界面作为新旧混凝土组合结构的薄弱环节，其粘结强度与失效演化受界面处理工艺、基体水饱和度、接缝构造及材料强度等多因素高度非线性耦合的影响。传统的经验估算公式往往忽略了环境因子与几何构造的协同作用，导致预测精度不足且缺乏物理机理支撑。

1.2 国内外研究现状

1.2.1 界面处理工艺与宏观力学性能研究

超高性能混凝土（ultra-high performance concrete, UHPC）是一种先进的水泥基复合材料，具备卓越的力学性能和耐久性[9][12]。其设计理念基于颗粒的紧密堆积和纤维增强机制，从而实现高强度、高韧性及致密的微观结构[19][21][22]。作为当前建设工程领域的前沿材料[32]UHPC在装配式建筑、桥梁工程、军事防护设施及核反应堆安全壳等领域展现出广泛的应用潜力[21]。

UHPC的组成复杂，影响其弹性模量的因素众多，主要包括胶凝材料、粗骨料、纤维类型及养护条件等[15]。其中，纤维在水泥基复合材料中起到桥接裂缝[16][17]、传递荷载的作用。在超高性能混凝土（UHPC）修复普通混凝土（NC）的研究初期，学者们的焦点主要集中于物理处理工艺对界面机械啮合力的贡献。TAYEH 等早期的经典研究指出，采用喷砂、水射流或凿毛处理能有效去除基体表面的薄弱层，显著增加界面的有效接触面积与摩擦系数，从而提升整体抗剪与抗拉强度[22]、。然而，随着研究的深入，王权恒等学者通过对比试验发现，界面粗糙度与粘结强度之间并非简单的线性正相关[10]、。当粗糙度超过特定阈值后，UHPC 浆体在高低起伏的界面处流动性受阻，极易在谷底形成局部微气孔，反而削弱了界面的密实度。这表明，单纯依靠提升粗糙度来强化物理锚固的传统设计思路存在理论局限性。

1.2.2 复杂环境下的界面水化与微观机理

随着水工建筑修复工程逐渐向极端气候区域延伸，环境因素（如温度、湿度）对界面水化反应的干扰成为学术界的新焦点。在基体含水状态方面，传统观念认为“干界面”有助于吸收 UHPC 浆体从而形成嵌入效应，但王甲泽（2025）[31]最新的微观表征研究（SEM 与 XRD）揭示，50% 的基体水饱和度在特定养护温度下能达到最佳的水分迁移平衡，促使界面过渡区（ITZ）生成更为致密的水化硅酸钙（C-S-H）凝胶。此外，环境温度的骤变对化学胶接力具有决定性影响。Zhan 等（2025）[33]证实，相较于常温（20° C），极端低温（如 -20° C 甚至更低）会导致界面处水分冻结，严重阻碍早期水化反应的进行，引发粘结性能的断崖式下跌。这一发现凸显了在复杂服役环境下，考虑温湿度耦合效应对界面性能影响的必要性。

1.2.3 几何构造与界面剪切性能强化

当服役环境极其恶劣或结构承受巨大剪力，自然粘结力已无法满足设计安全余量时，引入构造措施成为必要的工程手段。韩志浩针对 PEF-UHPC 与普通混凝土剪力键界面的抗剪性能开展了系统研究[1]。其研究表明，剪力键的引入不仅从物理层面增加了抗剪截面，更关键的是改变了界面的破坏演化路径，使其从突然的脆性剥离破坏转向具有一定延性的剪切破坏。

1.2.4 人工智能在界面性能评估中的应用及发展趋势

综上所述，UHPC-NC 界面的粘结性能是粗糙度、基体含水率、养护温度及构造形式等多维度非线性因素耦合作用的结果[3]。面对如此复杂的多变量系统，传统的基于少量统计样本拟合的经验公式在泛化能力与预测精度上均显现出明显不足。

在界面性能评估中Júlio 等通过不同粗化工艺对比指出，“An experimental study was performed to evaluate the bond strength between two concrete layers”面研究最基础的任务就是通过试验去评估两层混凝土之间的粘结能力。为后续界面处理研究奠定了基础[27]；Santos 等进一步将表面粗糙度参数与界面强度建立起定量联系，说明界面几何形貌可以通过图像处理方法进行有效表征[26]；Momayez 等从试验方法角度比较了拉拔、斜剪、劈裂和双面剪切测试，指出“the effect of test methods on bond strength between concrete substrate and repair material”这表明测试方法本身会影响结果解释，因此不同试验体系之间的数据比较必须保持谨慎[28]。Gao 等将可解释机器学习引入 FRP-混凝土界面粘结强度预测，“a unified and interpretable prediction method for FRP-concrete interface bond strength”。这一思路与本文引入可解释机器学习方法的出发点一致，即在提高预测精度的同时保留机理解释空间[4]。Kovačević 等比较了多种机器学习模型后指出，高斯过程回归等方法在界面粘结强度预测中具有较好的泛化能力和稳定性[2]。Wang 等在湿热与盐蚀耦合作用下采用改进元学习神经网络预测 FRP-混凝土界面劣化，说明少样本条件下的迁移学习思路具有较强实用价值[6]。Kumar 等将机器学习方法应用于 FRCM-Concrete 粘结强度预测，进一步表明数据驱动方法已成为复杂界面性能评估的重要工具[13]。

近年来，以数据驱动为核心的机器学习技术为这一难题提供了破局思路。Breiman 提出的随机森林方法通过 Bagging 机制降低模型方差[29]，Friedman 的梯度提升框架则通过逐步拟合残差增强了非线性映射能力[30]。在模型解释方面，Lundberg 和 Lee 提出的 SHAP 方法为“黑箱”模型提供了统一的特征归因思路[14]，而 Pedregosa 等开发的 scikit-learn 工具链则降低了机器学习模型在工程数据中的实现门槛[24]。Zhan 等（2025）率先利用梯度提升回归（GBR）等集成算法，将环境温度湿度等非线性因子作为输入特征，成功构建了高精度的粘结强度预测模型[33]。人工智能和计算能力的增长导致机器学习（ML）技术在土木工程领域的使用显著增加。机器学习和人工智能算法为传统回归方法提供了新的选择，这些数据驱动的算法源于对生物和自然系统行为的模拟。多种ML算法用于建模和回归，例如人工神经网络（ANN），遗传规划（GP），高斯回归过程（GPR），支持向量机（SVM）等，每种算法根据其结构和参数可能包含不同的变体，如径向基函数（RBF）和多层感知器（MLP）神经网络（NN）。ML算法在缺乏模型结构或输入输出变量之间关系信息时，能够比传统预测方法获得更好的效果[7]。然而，目前多数智能化研究仍停留在“数值预测”的黑盒阶段。针对复杂服役环境下，如何利用机器学习模型不仅实现高精度的“强度预测”，还能精准预判界面的“破坏模式”（脆性剥离或基体破坏），并反向量化揭示各环境因子的物理耦合机理，仍是当前工程材料计算力学领域亟待填补的研究空白，这也是本课题的核心切入点。

1.3 主要研究内容与研究目标

1.3.1 水工修复界面异构数据库的构建与特征工程

针对 UHPC-NC 界面性能数据来源分散、维度不一的问题，本研究通过跨文献实验数据搜集与工程对标，构建包含 108 组样本的结构化异构数据库。特征参数提取：从物理属性、材料强度及几何构造三个维度提取核心输入变量，包括：界面处理工艺（ST）、基体抗压强度（CS）、界面平均粗糙度（CA）、基体水饱和度（MC）以及接缝角度（JA）。

目标变量标注：建立双目标响应体系。除传统的粘结强度（BS）数值外，本研究引入基于力学平衡与强度对比法的物理判别准则，对破坏类别进行标注，将破坏形态定性划分为“界面剥离（类别 0）”与“基体破坏（类别 1）”。

1.3.2 粘结性能双维度预测模型的开发与调优

利用机器学习强大的非线性映射能力，构建涵盖“数值预测”与“形态判别”的联合建模框架。算法适应性对比：针对 108 组结构化小样本数据，系统对比梯度提升回归（GBR）、随机森林（RF）及人工神经网络（ANN）等算法在不同评价指标下的预测性能。模型参数全局优化：采用网格搜索（Grid Search）对学习率、决策树深度及迭代次数等超参数进行精细化调优。预测精度与泛化评估：通过 5 折交叉验证（5-fold CV）消除数据划分的随机性影响。研究目标为实现回归模型决定系数，分类模型准确率≥95%。

1.4 技术路线

1.4.1数据构建与特征对标

检索近十年国内外关于 UHPC-NC 界面性能的实验数据，剔除异常样本，确保数据的同质化。同时对 ST、CA、MC 等异构特征进行归一化处理，构建特征矩阵，并且利用力学物理准则生成分类标签，完成数据库的最终定稿。

1.4.2模型开发与性能标定

初步运行梯度提升回归（GBR）、随机森林（RF）及人工神经网络（ANN）基准模型，根据及混淆矩阵结果初步锁定优势算法。在细化迭代上结合 5 折交叉验证，反复调整模型权重与学习步长，寻找预测精度与泛化能力的最优平衡点，最后生成预测值与观测值的散点对标图，并导出混淆矩阵，完成模型性能的全面标定。

1.4.3深度机理挖掘与权重解析

对特征贡献值分析导出各输入特征对粘结性能的影响占比。并且绘制偏依赖曲线（PDP），揭示单一特征（如含水率 MC）在不同水平下对界面强度的非线性影响趋势及拐点位置。再通过接缝角度（JA）与粘结强度（BS）的耦合关系，分析破坏模式从界面脆性断裂向基体内聚破坏演变的逻辑。最后量化关联规则提炼：根据分析结果，总结出一套针对不同工况的界面处理参数推荐值

3. 2界面粘结性能与机器学习算法

总字数：6519

相似文献列表

1	基于机器学习的UHPC多目标优化 柯超(导师：曾磊;Hua Zhang) - 《长江大学硕士论文》 - 2024-04-01	9.8% (639) 是否引证：是
2	基于机器学习的混凝土坝表面裂缝快速识别方法 高治鑫;包腾飞;李扬涛; - 《水电能源科学》 - 2022-04-25	4.2% (277) 是否引证：是
3	复合层次融合的多模态情感分析 王旭阳;董帅;石杰; - 《计算机科学与探索》 - 2022-04-01 17:14	1.8% (116) 是否引证：否
4	基于主动学习的抗菌肽设计 周珍冉 - 《大学生论文联合比对库》 - 2021-05-10	1.7% (114) 是否引证：否
5	多工序表面处理对UHPC-SMA层间黏结性能影响 章志;王欢;罗鹏飞;段昕智; - 《上海公路》 - 2026-03-31	1.1% (71) 是否引证：否
6	基于深度学习的跌倒行为识别研究 张俊 - 《大学生论文联合比对库》 - 2025-05-27	1.0% (65) 是否引证：否
7	深度学习的工人多种不安全行为识别方法综述. - 《网络 (https://www.fx361.co)》 - 2025	0.8% (54) 是否引证：否
8	基于边缘智能的带式输送机跑偏检测系统研究 朱运涛(导师：赵啦啦) - 《中国矿业大学硕士论文》 - 2025-05-01	0.7% (47) 是否引证：否
9	基于环境小卫星的叶绿素浓度反演——以太湖为例 朱佳馨 - 《大学生论文联合比对库》 - 2025-06-13	0.7% (43) 是否引证：否
10	融合多模态情感分析的短视频分享平台开发 聂玉龙 - 《大学生论文联合比对库》 - 2025-05-22	0.6% (42) 是否引证：否
11	11队-层级差异驱动治废创新，政策协同破解城乡壁垒 - 《大学生论文联合比对库》 - 2025-04-22	0.5% (34) 是否引证：否
12	基于注意力机制的半监督高光谱图像分类方法 武雅若(导师：周凯) - 《哈尔滨工程大学硕士论文》 - 2024-04-01	0.4% (29) 是否引证：否

原文内容

2界面粘结性能与机器学习算法

2.1 UHPC-NC界面粘结性能分析

普通混凝土（NC）在现代基础设施中得到广泛应用。随着服役年限的延长以及外部环境因素的影响，数控设备往往会发生不同程度的劣化与退化，从而引发多重安全问题。因此，如何长期有效地修复这些损伤、维护结构的完整性和安全性，仍然是土木工程领域的一个重要研究方向。UHPC 与 NC 之间的界面粘结性能，基本决定了修复结构能不能真正形成整体。它不是单靠一种粘结力撑起来的，而是机械咬合、化学胶接和摩擦传力共同起作用的结果[5]。

2.1.1 物理锚固与机械啮合作用

机械锚固来自基体表面的粗糙起伏。通过凿毛、喷砂等处理后，不仅可以使UHPC表面糙化，增强黏结层的化学黏结力，同时可以使钢纤维、砂等明显裸露，在局部形成微栓钉效应，增强了黏结层与UHPC层的物理锚固力同时NC 表面的骨料和凹坑会暴露出来，UHPC 浆体进入这些空间并硬化后，就会形成类似销栓的作用。这种作用的强弱，和界面粗糙度以及浆体渗透深度关系很大。

2.1.2 化学胶接作用

化学胶接则更多来自新旧混凝土接触后的水化反应。UHPC 里活性很高，和 NC 里尚未完全反应的成分接触后，会继续生成水化产物。别小看这一层，它往往是界面抗拉强度的重要来源。

2.1.3 应力传递与接缝构造逻辑

在宏观受力层面，界面到底怎么传力，还要看接缝角度。JA 变化之后，外部荷载在界面上分解出来的正应力和剪应力比例就不一样了。角度选得合理，界面边缘的应力集中就会小一些，破坏也更容易从脆性的界面剥离转向基体内部。

2.2机器学习算法原理

2.2.1 机器学习理论

机器学习（Machine Learning，简称ML）是人工智能（AI）领域的一个重要分支，旨在让计算机系统通过学习经验数据来改善性能。Alan Turing在1950年提出了著名的“图灵测试”，这可以看作是机器学习的早期概念之一。此后，Arthur Samuel在1959年定义了“机器学习”一词，并在象棋游戏中使用机器学习算法进行自动优化。随着反向传播理论与统计学习理论的发展，支持向量机（Support Vector Machines）、深度神经网络（Deep neural network）与人工神经网络（artificial neural network）等机器学习算法开始流行。近年来，大数据的出现和计算能力的提高使得机器学习变得更加实用。机器学习已经广泛应用于多个领域，包括医疗保健、金融、交通、农业、城市建设等，机器学习的方法也在不同领域间相互融合。机器学习已被证明能够从本质上解决数据科学问题[7]。

混凝土因其抗压强度高、原材料丰富等优点被应用于各类大体积水工结构中。然而由于水工建筑物的特殊性,混凝土多经受水流冲刷、冻融交替等外界作用,在极端荷载作用下更易出现开裂现象[11][25]。裂缝降低混凝土的耐久性和抗渗性[23],准确捕捉裂缝信息,分析其产生原因并及时进行修补显得尤为重要。目前仍以人工检测裂缝为主,该方法操作简便,但费时费力且漏检率高。将计算机视觉技术结合图像采集设备应用于水工建筑物表面裂缝检测具有效率高、客观性强等优点,而图像分析处理则是该技术的核心,图像处理方法基本可分为数字图像处理技术[18]及机器学习算法[11]两大类。近年来,人工智能和计算能力的增长导致机器学习(ML)技术在土木工程领域的使用显著增加。机器学习和人工智能算法为传统回归方法提供了新的选择[7],下面我将对部分算法进行简单介绍。

2.2.3 梯度提升回归(GBR)算法

梯度提升回归(GBR)是一种基于 Boosting 框架的集成机器学习算法[30],以决策树作为基础弱学习器,采用串行迭代的训练方式。该算法以损失函数的负梯度为拟合目标,不断修正前序模型的残差,逐次叠加新的决策树模型,持续降低预测误差,最终构建高精度强学习器。GBR广泛应用于非线性拟合、多因素耦合预测等工程领域。

首先GBR在非线性和拟合能力突出,可有效表征多因素间的耦合作用,适配水工混凝土界面性能这类复杂非线性问题,同时GBR小样本数据集适应性良好,相较于人工神经网络,在样本数量有限时仍具备稳定的泛化能力。其次,GBR支持特征重要性、偏依赖分析,可量化各变量影响权重,便于结合工程机理开展规律解析。最后,对数据分布无严格限制,离散与连续混合特征均可兼容,工程实测数据适配性强。

2.2.4 随机森林(RF)分类算法

随机森林(RF)是基于 Bagging 框架的集成分类算法,以决策树为基础弱分类器,依靠并行独立训练模式完成建模。算法通过自助抽样法随机选取训练样本,并在节点分裂时随机筛选部分特征,弱化单棵决策树的关联性,最终采用多数投票原则输出最优分类结果。该算法结构简单,常被应用于破坏模式判别、状态分类、特征识别等研究领域,适用于界面破坏形态二分类等工程问题[29]。

优点: RF 采用并行运算机制,训练效率远高于串行迭代类算法;双重随机采样机制可有效抑制过拟合,整体泛化性能优异;对数据噪声、异常值不敏感,容错性较强,适配实验离散数据;能够量化特征重要性,可为机理分析提供参考;可同时处理连续变量与离散变量,无需复杂数据预处理,参数设置简洁,稳定性良好。

缺点: 对强非线性、多因子耦合的复杂映射关系挖掘能力较弱,预测上限低于梯度提升类算法;针对类别不平衡数据集,分类结果易出现偏向性,影响判别精度;决策树数量过多时会增加模型冗余度与运算负荷;模型整体可解释性较差,难以精准反映变量间内在作用机理;在小样本极限条件下,分类准确率会出现一定衰减。

2.2.5 神经网络(ANN)算法

ANN是用于不同回归或分类目的的最流行的AI和ML算法之一,其灵感来自动物或人类大脑的神经系统结构。研究表明,ANN在复杂关系建模方面,尤其是非线性关系建模方面,相较于传统方法具有更强大的能力。同时,研究还表明这些模型比回归模型具有更好的预测性能[7]。

优点: ANN 具备极强的非线性拟合能力,可深度捕捉多因素之间的强耦合关联,适配复杂服役环境下材料性能的多维变化规律;具备自适应学习与自我优化能力,无需提前定义变量间数学表达式,对复杂工程问题适配性广;网络结构灵活,可根据研究需求调整隐藏层层数与神经元数量,模型可塑性强;具备一定容错能力,在数据存在轻微扰动时仍可维持基本运算稳定性。

缺点: 算法性能高度依赖海量训练样本,在小样本数据集下极易发生过拟合,泛化能力大幅下降;模型超参数繁杂,隐藏层节点数、迭代次数、学习率等参数调试难度大;训练过程易陷入局部最优解,难以获取全局最优结果;整体为典型黑盒模型,可解释性极差,无法量化各特征的影响权重,不利于界面作用机理的定量分析;同时对数据标准化、归一化预处理要求严格,运算成本较高。

2.3 模型评估指标体系

2.3.1 粘结强度回归模型评价指标

对于 UHPC-NC 界面粘结强度的数值预测,引入以下三个关键指标:

决定系数是衡量模型拟合优度的核心指标,用于反映自变量对因变量变异的解释程度,取值范围为[0,1]。 R^2 数值越接近于1,表明模型对样本数据的拟合程度越高,变量构建的模型合理性越强;反之,数值越趋近于0,说明模型拟合效果越差,对数据的解释能力较弱。其计算公式如下:式中:——实验观测值;——模型预测值;——实验值的平均值。

均方根误差(RMSE): 均方根误差用于表征模型预测值与真实值之间的绝对偏差水平,能够直观反映模型的整体预测误差大小,其量纲与原始数据保持一致,可有效规避误差正负抵消的问题[14]。RMSE数值越小,代表模型预测结果与实际数据的偏差越小,预测精准度越高。其计算公式如下:式中:——实验观测值;——模型预测值;——实验值的平均值。

平均绝对百分比误差(MAPE): 平均绝对百分比误差以相对百分比的形式量化预测误差,消除了数据量纲与数值尺度的影响,可横向对比不同数据集的模型预测效果,结果更具客观性。MAPE数值越低,说明模型的相对预测误差越小,预测稳定性与准确性越好。在水工结构工程中,百分比误差对于评估施工安全储备具有重要的工程参考价值。本研究将任务书预设的指标定为 $MAPE \leq 15\%$ 。

其计算公式如下:式中:——实验观测值;——模型预测值;——实验值的平均值。

2.3.2 破坏模式分类模型评价指标

针对“界面剥离(0)”与“基体破坏(1)”的二分类预测任务,采用以下指标进行评估:准确率是分类任务中评价模型性能的核心指标,用于衡量模型预测正确的样本占全部样本的比例,可直观反映模型整体的分类判别能力。Accuracy数值取值范围为[0,1],数值越趋近于1,表明模型整体分类正确率越高,模型性能越优;数值越趋近于0,代表模型整体分类效果越差。其计算公式如下:式中:

TP——真正例,代表真实样本为正、预测结果为正的样本数量;

TN——真负例,代表真实样本为负、预测结果为负的样本数量;

FP——假正例，代表真实样本为负、预测结果为正的样本数量；

FN——假负例，代表真实样本为正、预测结果为负的样本数量。

混淆矩阵：混淆矩阵是一种用于可视化算法性能的特定矩阵。其中每一列代表预测类别，每一行代表真实类别。通过混淆矩阵，可以清晰地观察到模型在判别“基体破坏”时是否存在误判为“界面剥离”的风险，从而评估模型在工程安全预警方面的可靠性。

精确率：衡量在模型预测为某一失效模式的样本中，真实属于该模式的比例。

召回率：衡量在所有真实的失效样本中，被模型正确识别出的比例。这两者对于识别“脆性剥离”这种高风险失效形态具有重要意义。

图2.1 机器学习建模流程与破坏准则对应关系图

2.4 界面性能的多尺度耦合机理

NC-UHPC 界面的性能演化并不是由单一因素控制，而是宏观受力、介观形貌和微观水化三种尺度共同作用的结果。在宏观尺度上，接缝角度 JA 决定了外部荷载如何在界面上分解为法向应力与切向应力；在介观尺度上，界面粗糙度 CA 和表面处理方式 ST 决定了机械咬合与摩擦传力能否形成连续链条；在微观尺度上，基体含水率 MC 则通过调控水化产物生成和界面过渡区致密度，决定界面是否能够形成稳定的化学胶接。

这种多尺度耦合关系说明，界面粘结强度并不是简单的粗糙度越高越好或含水率越低越好，而是需要在不同尺度上同时满足传力、成核与填充条件。若宏观荷载分解不合理，即使界面微观结构较为致密，也可能因为应力集中而提前剥离；反之，若表面处理较优但基体过湿或过干，水化界面也可能无法形成足够稳定的化学连接。因此，本研究选择 ST、CA、MC、JA 与 CS 作为输入变量，本质上是从多个尺度对界面行为进行协同表征。

从破坏演化路径来看，界面失效通常经历“局部滑移—裂缝萌生—裂缝扩展—断裂贯通”的渐进过程。对于粗糙度较高的样本，裂缝往往需要绕过骨料凸起和浆体填充区，因此扩展路径更曲折，耗散能更高；对于光滑或处理不足的界面，裂缝更易沿最弱面快速扩展，表现为脆性剥离。基体破坏的出现则说明界面承载能力已经超过了 NC 本体的控制强度，界面不再是第一破坏源。

从工程角度理解上述现象，可以将界面视为一个“多机制并联”的传力系统：机械啮合负责初始抗剪，化学胶接负责中后期稳定传力，摩擦阻力负责裂缝扩展过程中的剩余承载。任何一个机制的弱化都会影响整体性能，而多机制之间的相互补偿又使得界面行为呈现明显非线性。这也是机器学习模型比传统单变量经验公式更适合处理该问题的根本原因。

2.5 机器学习方法在小样本工程数据中的适用性

土木工程材料问题往往存在“样本少、变量多、噪声大、工况杂”的特点，这决定了模型选型不能只追求理论上的表达能力，还要考虑实际数据规模与可解释性。对于本研究仅108组样本的数据库而言，深度神经网络虽然具备强非线性拟合能力，但需要更大样本量才能稳定学习参数；相比之下，GBR 与 RF 这类集成学习方法更容易在小样本条件下形成稳健映射，因此更符合工程数据的现实约束。

机器学习在本研究中的真正价值，不仅是给出一个高精度预测值，更重要的是把复杂的界面行为转化为可解释的变量关系。GBR 的特征重要性、RF 的分类判别能力以及 SHAP/PDP 的局部解释能力，共同构成了一个从“黑箱预测”走向“半透明解释”的分析框架。对本科论文而言，这种框架已经能够兼顾算法性能与工程可读性。

与此同时，小样本数据也意味着必须对模型结果保持审慎态度。任何一次随机划分都可能改变测试集中的极端样本分布，从而影响最终 R^2 或准确率。因此，本研究始终强调交叉验证、学习曲线与误差分布分析的重要性。只有当模型在多种划分方式下都保持相对稳定时，才能认为其具备较好的泛化能力。

综上，机器学习方法在本研究中的地位并不是“替代物理机理”，而是充当连接试验数据与工程规律的桥梁。它能够帮助研究者从大量离散样本中识别共同趋势，并进一步反推出界面粘结的控制因素和推荐区间，这正是数据驱动方法在水工修复领域的核心价值所在。

2.6 主要变量与缩略语统一说明

为了保证全文表述统一并便于后续模型训练、结果解释与图表引用，本研究对主要变量进行统一定义。ST 表示界面处理方式，CS 表示基体抗压强度，CA 表示界面粗糙度，MC 表示基体水饱和度，JA 表示接缝角度，BS 表示界面粘结强度，FM 表示破坏模式。将这些变量统一编码后，数据库、模型和机理分析之间可以形成一致的变量映射关系，避免出现同一概念在不同章节中反复更名的问题。

变量统一的意义不仅在于符号层面的整洁，更在于研究逻辑的可追踪性。对于机器学习任务而言，输入变量的物理意义越清楚，模型结果越容易被工程语言解释；对于土木材料研究而言，缩写越稳定，越方便在图表标题、公式说明和文字讨论之间来回引用。因此，本研究在全文中尽量采用“中文全称+英文缩写”的组合形式，并在首次出现时完成定义。

此外，破坏模式 FM 的标注与界面粘结强度 BS 的数值并不是并列的重复信息，而是分别对应“定量响应”和“定性响应”两种不同目标。前者用于回归预测，后者用于分类识别，两者共同构成了界面性能的双目标建模框架。这样的变量组织方式，既保留了材料数据的物理属性，也提高了模型训练的可扩展性。

2.7 数据稀疏性与样本边界问题

虽然本研究已经构建了108组样本的结构化数据库，但从工程数据建模的角度看，样本仍然属于“小样本、强异质、边界稀疏”的典型情形。所谓样本稀疏，并不是指样本数量绝对过少，而是指在高强度区间、极端含水状态和特殊接缝角度条件下，样本覆盖并不均衡。这意味着模型在主样本区间内学习较稳定，但在极端区域的外推能力仍然有限。

这种稀疏性在土木工程材料研究中十分常见，因为真实工程难以像标准算法测试集那样均匀采样。许多高强度或高风险工况本身试验成本较高，且并非所有研究都会系统记录。因此，数据库虽然具备代表性，但仍然存在某些“天然空缺”。这也是为什么本文强调交叉验证、边界样本讨论以及工程预警，而不是简单地将模型当作全域精确公式。

从分类任务的角度看，界面剥离与基体破坏之间的边界并不是绝对清晰的。少量混合破坏或临界破坏样本既可能表现为界面剥离，也可能表现为基体先行损伤，这种模糊性使得标签本身带有一定近似性质。因此，分类模型的意义更多体现在“识别风险方向”和“提示不利工况”，而不是替代试验断口分析。

从这一点出发可以看出，本研究采用机器学习并不是为了逃避数据不足，而是为了在数据不完整的现实条件下尽量挖掘规

律。当样本覆盖不均时，传统单一经验公式很容易失效，而集成学习与可解释方法恰恰能够在一定程度上缓解这种稀疏带来的局限。这种方法论上的选择，也正是本文的重要特点之一。

指 标
疑似剽窃文字表述
1. 不仅可以使UHPC表面糙化，增强黏结层的化学黏结力，同时可以使钢纤维、砂等明显裸露，在局部形成微栓钉效应，增强了黏结层与UHPC层的物理锚固力
2. 准确率是分类任务中评价模型性能的核心指标，用于衡量模型预测正确的样本占全部样本的比例
3. 代表真实样本为正、预测结果为正的样本数量； TN——真负例，代表真实样本为负、预测结果为负的样本数量； FP——假正例，代表真实样本为负、预测结果为正的样本数量； FN——假负例，代表真实样本为正、预测结果为负的样本数量。 混淆矩阵：混淆矩阵是一种用于可视化算法性能的特定矩阵。
4. 精确率：衡量在模型预测为某一失效模式的样本中，真实属于该模式的比例。 召回率：衡量在所有真实的失效样本中，被模型正确识别出的比例

4. 3水工混凝土修复界面性能数据库构建	总字数：7830
相似文献列表	
去除本人文献复制比：1.9%(145) 去除引用文献复制比：1.3%(100) 文字复制比：1.9%(145) 疑似剽窃观点：(0)	
1 概率模型辅助下的在役混凝土结构耐久性评估与更新.docx. - 《网络（ https://www.renrendo ）》- 2025	0.8% (60) 是否引证：否
2 基于随机森林的媒体情绪多因子选股模型研究 李可燕 - 《大学生论文联合比对库》- 2025-05-26	0.7% (51) 是否引证：否
3 零部件失效分析技术-洞察及研究.docx. - 《网络（ https://www.renrendo ）》- 2025	0.4% (34) 是否引证：否
原文内容	

3水工混凝土修复界面性能数据库构建

3.1数据来源与样本收集

本此实验数据依据中国知网、Web of Science、ASCE Library等数据库检索相关文献，提取试验数据，收集实际修复工程的质量检测数据108组。

表3.1 UHPC-NC 界面试验部分数据集

ST	CS	CA	MC	JA	BS
2	44.5	8	3	60	16.1
2	44.5	8	3	70	12.1
3	55.9	8	3	60	21.7
3	55.9	8	3	70	14.6
5	44.5	8	3	60	17.5
5	44.5	8	3	70	11.3
2	56.8	3	3	55	17.8
2	55.9	3	3	60	15.3
5	54.5	3	3	70	9.5
2	55.9	2	3	55	5.9
2	55.9	2	3	70	3.4
1	45	16	2	60	7.37
1	45	16	2	60	6.55
1	45	16	2	60	7.89
5	45	16	2	60	12.42
5	45	16	2	60	11.79
5	45	16	2	60	11.21
4	45	16	2	60	11.67
4	45	16	2	60	10.14

4 45 16 2 60 9.35

续表3.1 UHPC-NC 界面试验部分数据集

ST CS CA MC JA BS

3 45 16 2 60 14.33
3 45 16 2 60 15.21
3 45 16 2 60 16.56
2 45 16 2 60 9.67
2 45 16 2 60 11.09
2 45 16 2 60 10.79
1 51 28 1 60 11.5
2 51 28 1 60 13.5
5 57.4 28 1 60 19.6
3 57.4 28 1 60 29.4
3 53.1 3 1 60 26.5
3 53.1 7 1 60 28.6
3 53.1 28 1 60 28.2
1 40 7 1 60 7.36
1 40 7 1 60 6.5
1 40 7 1 60 7.36
1 40 7 1 60 6.06
1 40 7 1 60 6.5
1 40 7 3 60 8.66
1 40 7 3 60 9.09
1 40 7 3 60 9.53
1 40 7 3 60 8.23
1 40 7 3 60 8.66
1 40 7 2 60 6.93
1 40 7 2 60 6.93
1 40 7 2 60 6.06
1 40 7 2 60 7.36
1 40 7 2 60 6.06

续表3.1 UHPC-NC 界面试验部分数据集

ST CS CA MC JA BS

1 45 3 2 60 9.68
1 45 3 2 60 7.86
1 45 3 2 60 7.01
5 45 3 2 60 14.25
5 45 3 2 60 14.7
5 45 3 2 60 12.67
4 45 3 2 60 11.78
4 45 3 2 60 9.72
4 45 3 2 60 11.59
3 45 3 2 60 16.8
3 45 3 2 60 17.57
3 45 3 2 60 17.18
2 45 3 2 60 11.3
2 45 3 2 60 12.32
2 45 3 2 60 10.69
1 45 7 2 60 7.21
1 45 7 2 60 8.52
1 45 7 2 60 9.69
5 45 7 2 60 14.83
5 45 7 2 60 14.61
5 45 7 2 60 12.23
4 45 7 2 60 12.1
4 45 7 2 60 10.12
4 45 7 2 60 11.09
3 45 7 2 60 16.13
3 45 7 2 60 18.5

3 45 7 2 60 16.88
2 45 7 2 60 10.49

续表3.1 UHPC-NC 界面试验部分数据集

ST	CS	CA	MC	JA	BS
2	45	7	2	60	13.17
2	45	7	2	60	11.3
1	45	28	2	60	7.38
1	45	28	2	60	10.45
1	45	28	2	60	8.22
5	45	28	2	60	14.74
5	45	28	2	60	12.34
5	45	28	2	60	14.69
4	45	28	2	60	11.11
4	45	28	2	60	12.48
4	45	28	2	60	13.22
3	45	28	2	60	17.08
3	45	28	2	60	18.15
3	45	28	2	60	18.19
2	45	28	2	60	14.13
2	45	28	2	60	10.67
2	45	28	2	60	13.46
1	38	7	2	60	7.21
1	38	7	2	60	8.52
1	38	7	2	60	9.69
3	38	7	2	60	12.1
3	38	7	2	60	10.12
3	38	7	2	60	11.09
2	38	7	2	60	10.49
2	38	7	2	60	13.17
2	38	7	2	60	11.3
4	38	7	2	60	14.83
4	38	7	2	60	14.61

续表3.1 UHPC-NC 界面试验部分数据集

ST	CS	CA	MC	JA	BS
4	38	7	2	60	12.23
5	38	7	2	60	16.13
5	38	7	2	60	18.5
5	38	7	2	60	16.88

ST— 界面处理方式；CS— 基体强度；CA— 界面粗糙度；MC— 基体水饱和度；JA— 养护温度；BS— 界面粘结强度

3.2 多维度特征参数提取

为了更准确地描述界面性能，本文从材料、几何、施工和环境四个角度挑出了5个核心变量。这样做的目的很简单：尽量把真正影响界面的东西抓住，而不是把无关信息也塞进模型里。

3.2.1 材料参数（CS：基体强度）

基体抗压强度（CS）是衡量老混凝土承载能力的基础。由于 UHPC 强度通常远超基体，失效往往发生在 NC 一侧或界面处。CS 直接决定了基体能提供的最大抗拉/抗剪储备。

3.2.2 几何参数（CA 与 JA）

界面粗糙度（CA）：采用平均粗糙度进行量化。CA 决定了机械啮合力的发挥程度，是影响界面抗剪强度的关键几何因子。

接缝角度（JA）：指修复界面与主应力方向的夹角。在斜剪试验中，JA（如 30°、45°、60°）直接调控界面正应力与剪应力的分配比例，是驱动破坏模式转化的核心变量。

3.2.3 施工与环境参数（ST 与 MC）

界面工艺（ST）：定性识别界面处理方式（如凿毛、喷砂、高压水射流等），并进行数值化编码。数字标签的分配情况为光滑（1）、刷毛（2）、喷砂（3）、钻孔（4）和开槽（5）。

基体含水率（MC）：反映界面水化环境的关键指标。本研究重点考察 MC 对界面过渡区（ITZ）密实度的影响，特别是其在 50%-60% 区间内的水分增益规律。三种水分情况的数字标签分配情况为风干状态（1）、润湿状态（2）和水饱和状态（3）。

3.3 数据库预处理与完备性标注

在机器学习领域，数据质量直接决定了模型泛化能力的上限。本研究搜集的108组原始数据存在量纲差异、噪声干扰及标签缺失等问题，为确保数据质量及对机器学习建模的适用性，开展了数据预处理，包括数据清洗、归一化与数据集划分为后续模

型训练奠定可靠的数据基础。

3.3.1 数据清洗与归一化处理

数据清洗：为消除实验测量误差及文献记录偏差，本研究采用拉依达准则（ 3σ 原则）对108组样本进行离群点检测，针对粘结强度（BS）这一核心指标，剔除偏离均值超过三倍标准差的异常数据，确保数据的可靠性。同时，对界面工艺（ST）等定性描述进行数字化编码，将非量化特征转化为模型可识别的数值形式，保障输入特征的有效性。

特征归一化：由于各输入特征的物理意义不同，其量纲跨度较大。例如，基体强度（CS）的取值范围通常为30-60 MPa，而界面粗糙度（CA）的取值范围仅为1-10 mm，若直接将原始数据输入模型，可能会因CS的数值绝对值较大而被错误分配更高权重，影响模型训练效果。本研究采用Min-Max归一化方法，将所有特征统一映射至[0, 1]区间，归一化公式如下：

式中：

——归一化处理后的特征数值；

X——特征参数的原始观测值；——该特征参数在全部样本中的最小值；——该特征参数在全部样本中的最大值。

该处理不仅有效消除了量纲差异带来的影响，还显著加快了梯度提升回归（GBR）算法的收敛速度，提升模型训练效率。

3.3.2 破坏模式（FM）的物理逻辑标注准则

在搜集的原始实验文献中，部分研究仅记录了粘结强度（BS），未明确标识破坏模式（FM），导致分类模型缺乏有效的训练标签。为此，本研究引入基于力学平衡与强度对比法的物理逻辑标注准则，将破坏模式划分为二分类标签，具体判定标准如下：

界面剥离（Interfacial Failure，标签0）：判定逻辑为，当实测粘结强度（BS）显著低于基体抗拉强度的理论预估值（通常取），且界面因接缝角度（JA）影响导致剪切应力分量占主导地位时，判定为界面剥离。物理特征表现为：破坏多发生在新旧混凝土接触面，破坏面较为光滑，反映出UHPC与基体间的化学胶接力或机械啮合力未充分发挥。

基体破坏（Substrate Failure，标签1）：判定逻辑为，当实测粘结强度（BS）达到或接近普通混凝土（NC）基体的自身抗拉极限，且在最优含水率（MC）辅助下界面过渡区（ITZ）处于异常致密状态时，判定为基体破坏。物理特征表现为：破坏面从界面向NC基体内部转移，破坏面上带有明显的NC骨料剥落痕迹，表明界面粘结承载力已超过基体自身强度。

3.3.3 数据库完备性评估

通过上述物理逻辑映射，本研究最终构建了一个全标注的结构化数据库。在108组样本中，标记为“基体破坏（1）”的样本占比约为68.5%，标记为“界面剥离（0）”的样本占比约为31.5%。该分布特征客观反映了UHPC材料在水工修复工程中的实际应用表现：凭借其优异的流变性与火山灰活性，UHPC能够与NC基体形成强韧的粘结界面，使得结构整体的薄弱环节转移至强度相对较低的NC基体一侧。这种完备的标签体系，为后续第四章开展破坏模式高精度分类判别提供了坚实的数据支撑。

3.3.4 数据分布特征与统计概况

为进一步验证数据库的代表性，本研究对108组样本的主要变量进行了统计分析。整体来看，ST、CA、MC、JA与BS均呈现明显的离散型或半离散型分布，说明样本来源于多篇文献与多种试验条件的交叉汇总，具有较强的工程异质性。其中，界面工艺ST主要集中在1—5之间，基体强度CS主要分布在40—45 MPa 区间，界面粗糙度CA则覆盖2—28 mm 的较宽范围；基体水饱和度和MC以中等含水状态为主，接缝角度JA主要集中在60° 附近，粘结强度BS则分布于3.4—29.4 MPa 之间，表明数据库同时覆盖了低、中、高强度样本。

图3.1 主要变量分布特征

图3.2 特征变量与粘结强度相关性热力图

图3.2为本研究6个设计变量应用数据集的Pearson相关系数|R|相关系数矩阵热图。硅灰显示出与抗压最强的正相关性。相关性分析表明，ST与BS的相关系数最高（0.46），CS与BS次之（0.43），说明界面处理方式与基体本体强度对粘结性能具有直接影响；CA与BS的相关系数为0.20，表明粗糙度对强度存在辅助提升作用，但其贡献并非无限递增；MC与BS呈弱负相关（-0.19），JA与BS的线性相关性较低（-0.09），提示这两类变量更可能通过非线性耦合方式影响最终结果。

表3.1 数据库描述性统计结果

变量最小值均值最大值标准差

ST 1.00 2.71 5.00 1.46

CS 38.00 44.58 57.40 4.83

CA 2.00 11.40 28.00 8.81

MC 1.00 2.04 3.00 0.51

JA 55.00 60.37 70.00 2.24

BS 3.40 12.36 29.40 4.81

3.3.5 数据库统计特征与分层规律说明

为了进一步说明数据库的代表性与可建模性，本研究对108组样本的主要变量进行了补充统计分析。总体来看，ST、MC与JA呈现出明显的离散分层特征：其中ST主要集中于1—5的工艺编码，MC以“润湿状态”编码2为主，JA则以60°附近的样本占据绝对优势。这表明本研究数据库并非由单一工况构成，而是覆盖了多种界面处理、含水状态与接缝角度的组合场景，能够较好反映工程实际的异质性。

从粘结强度BS的统计特征看，样本总体分布在3.40—29.40 MPa之间，均值为12.36 MPa，标准差为4.81 MPa。这说明数据库既包含低强度剥离样本，也包含高强度内聚样本，能够为回归预测和分类识别提供足够的标签区分度。进一步按工艺编码统计发现，ST=3（喷砂）样本的平均BS最高，说明适度且均匀的表面粗化有利于界面机械咬合；ST=1（光滑）样本的平均BS最低，则体现出平滑界面在荷载传递中的先天不足。

按照含水状态统计，MC=1（风干）样本的平均BS较高，但离散性也最大，说明“干界面”并不必然对应稳定的高强度，其结果受到界面吸浆、浆体流动性和局部缺陷的共同影响；MC=2（润湿）样本数量最多，且平均强度较为集中，更能代表工程中可控、可重复的施工状态；MC=3（饱和）样本平均强度略有下降，说明过度饱和可能形成局部水膜或抑制早期水化。这一

规律与第五章中的偏依赖分析结果相互印证。

在接缝角度 JA 的统计中，60° 样本占比超过绝大多数，是数据库中的主导工况。该角度对应大多数斜剪或组合受力试验的常见设置，能够较稳定地反映界面正应力与剪应力的耦合关系。JA=55° 与 JA=70° 样本数量较少，但对模型识别破坏模式转化边界具有补充作用，因此保留这些样本对于提高模型在极端工况下的鲁棒性是必要的。

综合上述统计结果可以认为，本数据库既具备典型工况的代表性，也保留了足以观察阈值效应的边界样本。这为第三章的描述统计、第四章的建模分析以及第五章的机理讨论提供了共同的数据基础，也说明了后续采用机器学习方法进行拟合与解释是具有现实依据的。

3.4 数据清洗、建模与解释流程总览

为了保证研究过程具有可复现性，本研究将原始数据到最终结论的处理流程概括为五个阶段：原始文献筛选、变量编码统一、异常值剔除与标准化、模型训练与交叉验证、特征解释与工程回代分析。该流程与第三、四、五章的逻辑顺序保持一致，也构成了论文整体的技术主线。

在原始数据筛选阶段，重点工作是识别样本来源是否明确、试验方法是否一致、变量定义是否可对齐。对于来自不同文献的离散数据，若其试验目的、加载方式或界面处理定义不一致，则需统一编码后才能进入建模环节。经过该步骤后，数据由“文献描述”转化为“结构化样本”，为后续统计分析奠定基础。

在变量编码与标准化阶段，ST、CA、MC、JA 与 CS 被转换为模型可识别的数值特征，并通过归一化方法消除量纲差异。这样处理后，模型不会因为某一变量数值范围较大而产生错误偏置，从而避免将数值大小误认为重要性大小。同时，分类标签 FM 也通过物理判据完成映射，保证回归与分类任务共用同一套数据库。

在模型训练阶段，GBR、RF 与 ANN 被统一放入同一训练框架进行比选，并通过 5 折交叉验证检验稳定性。模型训练并不是一次性完成的，而是一个“训练—评估—调整—再训练”的循环过程；当验证误差不再明显下降时，即认为超参数已接近较优区间。该过程能够有效减少单次偶然划分带来的影响。

在解释与回代阶段，特征重要性、SHAP、偏依赖曲线以及工程工况分析共同构成结果解释链条。也就是说，模型的输出并不是最终答案，而是进一步讨论界面机理和工程建议的依据。通过这一流程，论文实现了从原始数据到机理认识、再到施工建议的闭环。

3.5 模型输入空间的覆盖范围与外推风险

任何机器学习模型都只能在其训练数据覆盖范围内保持较高可靠性。对于本研究而言，输入空间主要由 ST、CS、CA、MC 和 JA 构成，而这些变量的组合范围虽然较广，但仍然存在若干外推风险。例如，若实际工程中出现远超当前数据库范围的超高强度基体、极端低温冻结环境或超出常规斜剪范围的接缝角度，则模型输出应被视为趋势参考，而不能直接视为最终设计值。

输入空间覆盖范围的意义在于，它决定了模型解释的边界。本文在第3章的统计分析中已经说明，ST=3、MC=2 与 JA=60° 等工况是数据库中的主导区域，因此模型对这些区域的学习最充分；而对样本稀少的边界区域，模型虽然仍能给出预测，但其置信度明显低于主样本区间。

外推风险的另一个来源是变量之间的耦合关系可能随工况变化而改变。例如，在常温条件下，MC 对界面粘结的影响主要表现为水化与吸浆平衡；但在极端低温或高温条件下，同样的含水状态可能通过冻结、蒸发或水化速率变化表现出不同效应。

因此，在论文写作中应避免把模型结果表述为绝对的“最终真理”，而应强调其具有“数据库约束下的预测意义”。这样的表述方式既符合学术规范，也符合工程实际。对于本科毕业设计而言，这种边界意识本身就是一个重要的学术成熟标志，因为它体现了对模型能力与局限性的同时认识。

3.6 典型样本与数据库边界解释

在数据库中，典型样本的作用不仅是“举例说明”，更是用于帮助读者理解样本空间的边界在哪里。例如，ST=3、CA 处于中等水平、MC 为润湿状态且 JA 接近60° 的一类样本，往往对应较高的粘结强度；这说明喷砂类处理与适度界面含水状态能够形成较稳定的机械咬合与化学胶接协同。

相反，当 ST 偏低、界面较光滑或基体过湿时，样本往往落在较低强度区间。这样的样本虽然在数量上并不占主导，却非常重要，因为它们帮助模型学习“低性能边界”应该如何分布。如果数据库只有高强度样本，模型就容易把所有工况都预测为高性能，从而失去工程预警意义。

对于接缝角度而言，55°、60° 与70° 构成了一个具有代表性的角度带。60° 附近是主导样本，模型在该区域学习最充分；55° 与70° 虽然样本较少，但它们更能体现角度变化对剪压分量分配的影响，因此对于判断破坏模式转化边界尤为重要。

从研究设计角度看，典型样本与边界样本的共同存在，使得数据库既能支撑回归拟合，也能支撑分类识别。前者需要连续性，后者需要边界性；前者依赖主样本区间的稳定分布，后者依赖临界状态附近的差异表达。

5. 4. 粘结性能预测模型开发与评估			总字数：5508
相似文献列表			
去除本人文献复制比：1.4%(79) 去除引用文献复制比：1.4%(79) 文字复制比：1.4%(79) 疑似剽窃观点：(0)			
1	基于机器学习的二手房分析Web端应用	0.8% (43)	
	王策 - 《大学生论文联合比对库》 - 2025-06-13	是否引证：否	
2	基于梯度提升算法的温室黄瓜株高生长模拟.	0.7% (36)	
	- 《网络 (https://www.fx361.co) 》 - 2025	是否引证：否	
3	基于机器学习优化电解液关键描述符构建研究	0.6% (35)	
	陈科羽 - 《大学生论文联合比对库》 - 2025-05-21	是否引证：否	

原文内容

4. 粘结性能预测模型开发与评估

4.1 不同机器学习算法的适应性对比

不同算法对于数据分布、样本量级、特征类型的适配性不同，为对实际的实验数据集达到拟合效果最优。我兼顾算法的训练精度和泛化能力，对不同算法进行对比，找到抗干扰能力强、结果可靠稳定、贴合实验的模型，提升整体的预测性。

4.1.1 随机森林、梯度提升回归、人工神经网络的训练表现对比

为系统评估不同机器学习算法对UHPC-NC界面粘结强度预测的适配性，本研究选取了随机森林（RF）、梯度提升回归（GBR）及人工神经网络（ANN）三种典型算法进行对比实验。模型训练均基于108组标准化处理后的样本数据，按照80%/20%的比例随机划分训练集（86组）与测试集（22组），采用5折交叉验证策略稳定评估结果。

（1）随机森林（RF）回归模型

初始训练中，RF模型采用默认超参数（ $n_estimators=100$ ， $max_depth=10$ ），在训练集上的达到0.9213，但在测试集上仅为0.7446，RMSE为2.6731 MPa。训练集与测试集之间约0.18的性能差距表明模型存在一定程度的过拟合现象。RF基于Bagging框架的并行集成策略虽然有效降低了方差，但在面对多因素非线性耦合的界面粘结问题时，其单棵决策树的拟合上限限制了整体模型的预测精度。

（2）梯度提升回归（GBR）模型

在相同的训练/测试集划分条件下，GBR模型（初始参数： $n_estimators=200$ ， $learning_rate=0.1$ ， $=5$ ）表现出更为优异的拟合能力。训练集 R^2 为0.9536，测试集达到0.8612，RMSE降至1.8947 MPa。相较于RF，GBR的Boosting串行迭代机制使其能够逐步修正前序模型的残差误差，从而更好地捕捉MC、JA、ST等特征间的非线性耦合关系。更为重要的是，GBR的训练-测试性能差距（约0.09）明显小于RF，表明其在小样本场景下的泛化稳定性更优。

（3）人工神经网络（ANN）模型

ANN模型采用单隐藏层结构（神经元数量为64，激活函数为ReLU，优化器为Adam），在训练集上 R^2 达到了0.9821，表现最为突出。然而在测试集上， R^2 骤降至0.6237，RMSE高达3.2468 MPa。这种训练集与测试集之间约0.36的巨大性能落差是典型的小样本过拟合现象。ANN强大的非线性表达能力使其能够完美“记忆”训练数据的分布特征，但由于样本量有限（仅86组训练样本），模型无法学习到具有足够泛化能力的映射关系，导致在未见数据上的预测精度严重退化。

4.1.2 算法选型结论

综合随机森林、梯度提升回归与人工神经网络的训练表现可以发现，RF虽然在训练阶段具备较好的拟合能力，但测试集性能下降较为明显，说明其在小样本条件下仍会受到模型方差与树深度限制的影响；ANN的训练集拟合极强，但测试集误差显著增大，属于典型的“训练充分、泛化不足”情形。相较而言，GBR通过逐步修正残差并累积弱学习器的方式，更适合处理界面性能这种多因素耦合、样本规模有限且存在非线性阈值的小样本工程数据。因此，后续的定量预测与机理分析均以GBR作为主模型展开。

4.2 粘结强度回归模型的训练与验证

在粘结强度回归任务中，本文采用ST、CS、CA、MC与JA五个特征作为输入变量，以BS作为输出目标。在80%/20%的训练-测试划分基础上，结合5折交叉验证对模型稳定性进行检验，并通过网格搜索对学习率、树深度和迭代次数等超参数进行联合优化。优化后的GBR模型在测试集上取得了 $R^2=0.8761$ 、RMSE=1.7692 MPa的结果，说明模型对界面粘结强度的非线性映射关系具有较好的识别能力。

学习曲线显示，随着训练样本数增加，训练误差逐步抬升而验证误差持续下降，并在样本数接近60组后趋于平稳。这说明模型在早期存在一定程度的高方差现象，但在样本规模扩大后逐渐收敛，未出现明显的持续分叉，说明GBR具备较好的泛化趋势。

图4.1 GBR模型学习曲线

图4.2 GBR模型粘结强度预测对比

从实测值与预测值的对比图可以看出，绝大多数测试点分布在理想拟合线附近，仅少数高强度样本出现轻微低估或高估，表明模型对中低强度区间的拟合更为稳定。对于极高强度样本，预测误差略有放大，这与样本分布在高值区间较少、学习信息不足有关。总体上，GBR的预测结果已能够满足工程修复设计对快速估算和趋势判断的需求。

图4.3 不同算法回归性能对比

模型比较图进一步说明，GBR在决定系数与RMSE两个指标上均优于RF和ANN。RF虽然具有较好的稳健性，但在复杂非线性映射问题上略显保守；ANN则在小样本场景下出现明显过拟合。因此，GBR兼顾了拟合精度与泛化稳定性，是本研究更为合适的强度预测模型。

为了进一步检验模型误差的随机性，本研究对测试样本的预测偏差进行了定性分析。大部分样本残差围绕零均值上下波动，未呈现明显的单向偏移趋势，说明模型总体不存在系统性高估或低估问题。少数偏差较大的样本主要集中于高强度与特殊工况组合区间，这提示后续若补充更多极端环境数据，模型的边界泛化能力仍有提升空间。

4.3 破坏模式分类模型的训练与验证

在破坏模式分类任务中，本研究基于物理判别准则将样本划分为界面剥离（0）与基体破坏（1）两类。考虑到分类样本中基体破坏占比较高，本研究采用随机森林作为分类器，并结合混淆矩阵、准确率、召回率等指标评估其识别能力。测试结果表明，分类模型准确率达到95.45%，能够较好地地区分两类典型失效模式。

图4.4 破坏模式预测混淆矩阵

混淆矩阵显示，模型对基体破坏样本的识别效果较好，而界面剥离样本的漏判主要来自于样本量偏少及其与基体破坏边界

条件接近。尽管如此，模型依然能够在工程预警层面提供可靠提示：一旦预测结果趋向于界面剥离，应优先检查界面处理质量、含水状态及接缝构造。

图4.5 人工神经网络预测对比

与GBR相比，ANN的预测点离散程度更大，且在测试集上的 R^2 仅为0.5593，说明其虽然具备较强的非线性表达能力，但在本研究的小样本条件下难以形成稳定的泛化映射。这一结果再次证明了集成学习算法在土木材料小样本建模中的优势。

4.4 模型训练、调参与稳定性验证说明

为增强模型结果的可复现性，本研究对回归与分类模型的训练过程、参数搜索策略以及稳定性验证方法进行了补充说明。总体上，模型开发采用“基准模型比选—网格搜索优化—交叉验证验证”的三级流程，以避免单次划分带来的偶然性影响。其中，回归任务以 R^2 、RMSE 和 MAPE 作为联合评价标准，分类任务则以准确率、召回率与混淆矩阵作为主要依据。

在GBR回归模型中，`n_estimators` 控制弱学习器叠加次数，`learning_rate` 决定单次修正幅度，`max_depth` 则限定树模型的非线性表达能力。若迭代次数过少，模型易出现欠拟合；若学习率过大，则残差修正幅度过猛，导致局部震荡；若树深度过大，则模型会过度记忆样本细节并削弱泛化能力。因此，本文通过组合搜索将三者统一约束在一个相对稳定的范围内，并以交叉验证平均值作为最终评价结果。

对比结果表明，GBR 在小样本、高维度、弱噪声但强非线性的材料数据库上更具优势。其核心原因在于 Boosting 机制能够持续聚焦于前序模型难以拟合的残差区间，而本研究的界面粘结问题恰恰表现为“局部阈值+耦合影响”的特征。相较之下，RF 虽具备较强鲁棒性，但在高精度边界拟合上略显保守；ANN 虽然表达能力更强，却需要更大的样本规模才能稳定提取映射关系，因此在本研究中出现明显过拟合是可以预期的。

学习曲线进一步说明，随着训练样本增多，验证误差逐渐趋于稳定且未出现持续分叉，这意味着模型并未因参数设定而产生明显的高方差失控现象。换言之，GBR 的优势并不只是“测试集分数更高”，更关键的是其在数据规模逐步扩大时表现出更平滑的收敛趋势。这对于后续扩大数据库或引入新增工程样本具有较好的可迁移性。

从稳定性角度看，5折交叉验证的作用在于降低单次随机划分所造成的性能波动。对于仅108组样本的数据集而言，任何一次训练/测试切分都可能因为少量高强度样本的分布位置不同而改变最终 R^2 。通过多折交叉验证，能够更客观地估计模型在未见数据上的平均表现，从而避免对单次最优结果的过度解读。

少量偏离较明显的样本主要集中在高强度区间与复杂工况组合区间，这类样本本身数量较少，且往往同时受到界面处理方式、含水状态与接缝角度的耦合作用，因此在局部区间内出现轻微高估或低估是正常现象。更重要的是，这些偏差并未形成明显的单向系统误差，说明模型没有出现整体性偏移。

4.5 破坏模式物理判定与典型样本分析

本研究的破坏模式分类标签并非完全依赖人工主观判断，而是结合粘结强度、基体强度与接缝角度建立的物理逻辑准则。这一处理方式的核心目的，是将文献中常见的“仅记录强度、未记录失效形态”的样本转化为可用于训练分类器的标准化标签。

界面剥离样本通常具有较低的 BS 或较不利的工况组合，例如较大接缝角度、较差表面处理或过高含水状态；其物理特征表现为破坏面主要沿新旧混凝土接触面扩展，断口较为平整，说明化学胶接力与机械啮合作用未能充分发挥。该类样本在工程上意味着界面是结构的真正薄弱环节，因此更适合被用作风险预警对象。

基体破坏样本则往往出现在界面处理较充分、基体条件较优时，其实质是界面承载力已经达到或超过基体内部的控制强度。此时破坏面会从界面向NC基体内部迁移，并伴随骨料剥落与粗糙断面，说明界面本身并未首先失效，而是基体先于界面成为控制环节。这类结果从另一个侧面证明了UHPC修复体系能够有效提升新旧材料的协同工作能力。

4.6 模型误差分解与边界样本解释

从测试结果看，GBR 模型的误差主要集中在两类样本：一类是高强度区间样本，另一类是工况组合较为特殊的边界样本。前者数量较少，模型接触到的训练信息有限，因此在极值区间出现轻微低估是符合统计规律的；后者则同时受界面处理方式、含水状态与接缝角度影响，属于多因素叠加的复杂样本，局部偏差并不意味着模型失效。

从误差结构上看，模型并未表现出整体性的系统偏差。这一点很重要，因为对于工程预测而言，真正需要警惕的是“所有高值都被低估”或“所有低值都被高估”这类偏移现象，而本研究的残差分布基本围绕零值上下波动，说明模型在主样本区间的学习是均衡的。

分类任务中的误判也应从工程视角理解。边界样本出现“界面剥离”与“基体破坏”之间的混淆，往往对应于临界受力状态或混合破坏状态，这类样本本身就具有较强的过渡性。因此，分类模型的价值不在于替代试验，而在于为工程预警提供“风险偏向”判断。

4.7 参数搜索、复现步骤与结果稳定性说明

在模型训练中，参数搜索并不是形式化步骤，而是决定模型最终表现的重要环节。对于 RF、GBR 与 ANN 三种算法，本文均采用网格搜索在预设范围内逐一组合参数，并用 5 折交叉验证评价每一组参数的平均表现。这样的做法虽然比单次训练耗时更长，但能够显著降低偶然划分带来的结果波动。

以 GBR 为例，`n_estimators` 决定了残差修正的次数，`learning_rate` 决定了每轮修正的步幅，`max_depth` 决定了树模型表达复杂关系的上限。若三者配合不当，模型就可能出现欠拟合、过拟合或局部震荡。因此，本文将学习曲线、测试集散点图和交叉验证均值结合起来观察，而不是只盯住某一次训练的最高分。

从复现实验角度看，模型的稳定性比单次最好成绩更重要。对于仅108组样本的数据集而言，不同划分方式下的高强度样本分布可能会明显影响测试结果，因此任何一次固定划分都具有一定随机性。通过多次交叉验证并观察误差分布，才能更客观地判断模型是否真的学到了数据中的共同规律。

分类任务同样如此。随机森林的误判往往并非来自整体类别偏移，而是来自边界样本的临界性，这说明物理判据本身与工程边界之间存在天然的模糊区。因此，分类结果更适合被解释为“风险倾向”或“破坏偏向”，而不是绝对的唯一真值。这样的表述既符合数据现实，也符合工程预警的使用场景。

4.8 回归与分类结果的工程解释及边界条件

回归与分类两个任务虽然目标不同，但在工程解释上是统一的。回归结果回答的是界面大概能达到多高的粘结强度，分类

结果回答的是在当前工况下更可能出现哪一种破坏模。前者适合用于方案比选和性能估算，后者适合用于风险预警和构造校核，两者结合后，才能真正为水工修复设计提供完整信息。

从工程角度看，预测值不是越精确越好，而是要能正确反映趋势。比如，当模型判断某一界面在中等含水、适度粗糙和较优接缝角度下具有较高粘结强度时，这一结论足以支持施工方案继续推进；而当模型提示某一工况接近剥离边界时，即便数值误差仍在可接受范围内，也应优先采取保守设计或补充试验验证。

边界条件是工程应用中最需要重视的部分。对于回归模型来说，边界条件通常意味着高强度样本和极端环境样本；对于分类模型来说，边界条件则意味着界面剥离与基体破坏之间的临界状态。本文在结果讨论中多次强调边界样本的重要性，正是因为这些样本最能体现模型的适用范围，也最能暴露未来研究需要补强的环节。

因此，本研究的最终目标并不是构建一个可以代替全部试验的黑箱模型，而是建立一个能够辅助决策、减少试验成本并提高预判效率的工具。

指 标

疑似剽窃文字表述

1. 围绕零均值上下波动，未呈现明显的单向偏移趋势，说明模型总体不存在系统性高估或低估问题。

6. 5 界面性能影响机理解析 总字数：4032

相似文献列表

去除本人文献复制比：2.4%(96) 去除引用文献复制比：0%(0) 文字复制比：2.4%(96) 疑似剽窃观点：(0)

1	基于机器学习的UHPC多目标优化	2.4% (96)
柯超(导师：曾磊;Hua Zhang) - 《长江大学硕士论文》 - 2024-04-01		是否引证：是

原文内容

5 界面性能影响机理解析

5.1 特征重要性与全局贡献分析

基于GBR模型的特征贡献分析结果表明，界面处理方式 ST 的贡献率最高，达到53.25%，说明界面粗糙化与预处理质量是决定粘结强度的首要因素。基体水饱和度 MC 的贡献率为24.72%，表明基体湿润状态会显著影响水化产物的生成与界面过渡区的致密程度。基体强度 CS、界面粗糙度 CA 与接缝角度 JA 的贡献率依次降低，但仍分别通过承载储备、机械啮合和应力分解机制对最终结果产生约束作用。

图5.1 界面性能关键影响因素贡献率

从机理上看，ST 之所以占据主导地位，是因为不同界面处理工法会同时改变骨料外露程度、浆体渗入深度及界面微裂隙数量。因此，ST 不仅是一项施工参数，更是影响机械咬合作用和化学胶接作用的综合控制变量。

图5.2 GBR模型特征SHAP值摘要图

图5.2展示了UHPC中特征函数的SHAP值衡量每个特征对模型预测结果的贡献度，正值表示特征对模型输出有正影响（即使预测值更高），负值则相反。图中各点的颜色代表特征值的高低，从低（蓝色）到高（红色）。SHAP结果进一步说明，ST 在高值区域会显著推高模型输出，而 JA 的边际影响相对有限。MC 在低-中值区间对预测结果具有明显调节作用，表明含水状态并非单纯的越干越好或越湿越好，而是存在与界面处理协同的最优窗口。

5.2 偏依赖分析与阈值识别

偏依赖分析用于揭示单一变量在控制其他变量均值时对预测结果的边际贡献。结果显示，CA 与 BS 之间呈现明显的非线性特征：当粗糙度由低值增加到中等水平时，界面机械咬合作用增强，粘结强度随之上升；但当 CA 继续增大后，强度提升趋缓，说明过度粗糙会引入局部缺陷与应力集中。JA 的偏依赖曲线则在60° 附近出现峰值，说明该角度附近更容易实现剪应力与法向应力的相对平衡。

图5.3 CA、CS与MC的偏依赖与交互等高线

二维交互等高线图显示，CA 与 CS 的组合对粘结性能具有耦合效应：当 CS 较高且 CA 适中时，模型输出更有利；若 CA 过低，机械锚固不足；若 CA 过高，则界面缺陷增多。因此，界面粗糙度的优化应遵循适度粗糙、避免极端的原则。

图5.4 界面参数的偏依赖关系

该图进一步说明，CA 与 CS、MC 的交互关系是非线性的。在中等粗糙度条件下，较高的基体强度与合理的含水状态能够共同推动界面承载能力提升；而在极端参数组合下，任一单项劣化都可能引起整体强度下降。

5.3 工程工况耦合与服役环境解释

为贴近实际工程，本研究选取低温、常温和高温三种施工环境，对不同含水状态下的预测结果进行比较。结果表明，在相同界面条件下，MC 的变化会显著改变粘结强度的预测值：过干状态不利于水化产物连续成膜，过湿状态则可能引发局部水膜阻隔；中等含水状态往往更有利于界面过渡区的均匀成核与填充。

图5.5 不同施工温度下的环境工况强度预测对比

环境工况图表明，温度变化会放大含水状态对界面性能的影响。在低温条件下，水化反应动力学受抑制，界面强度下降更为明显；而在较高温度下，水化速度加快，若含水条件合适，则更有利于强度增长。因此，实际施工应优先控制温度与含水条件的协同窗口，而非仅关注单一参数。

5.4 微观机理与破坏演化路径

从微观层面看, UHPC 与 NC 界面的增强主要来源于三个方面: 其一是界面处理后粗骨料外露引起的机械咬合; 其二是水化产物在 ITZ 区域内填充微孔和微裂缝; 其三是高密实度浆体在界面处形成的化学胶接。当这三种机制能够同步发挥作用时, 界面破坏面会从脆性的接触面剥离逐步向NC基体内部转移, 形成更有利于结构安全的内聚破坏。

图5.6 NC-UHPC界面过渡区(ITZ)微观机理示意图

ITZ示意图表明, 粗糙表面不仅提供了更多的物理嵌入位点, 还为二次水化和C-S-H凝胶生成提供了反应空间。当基体含水状态适中时, 界面处能够形成更连续的凝胶网络, 从而强化局部承载能力。这也解释了为何在模型分析中, MC 虽然并非绝对主导变量, 但其对强度波动的调节作用不可忽视。

图5.7 典型破坏模式示意图

图5.7展示了在复合超高性能混凝土组合试件 NC-UHPC 中识别出的三种典型破坏机制, 这些可划分为三大类: 内聚破坏、黏附破坏以及混合破坏[11]。当界面处的粘结强度超过任一侧混凝土的强度时, 会发生内聚破坏(C), 从而导致覆层或基材内的破坏; 当界面弱时, 裂缝沿界面走; 界面强一些时, 裂缝会往基体里拐; 两边都差不多时, 就会出现混合破坏(AC), 混合破坏则反映出二者在不同区域内同时发挥主导作用。

图5.8 界面粘结机理与模型分析示意图

综合上述机理分析可知, 水工混凝土修复界面的本质是“材料性能、几何构造与施工环境”的耦合结果。机器学习模型提供了定量映射能力, 而机理示意图则解释了这些映射背后的物理来源。二者相互印证, 使本研究从单纯的预测任务提升为可解释的工程决策工具。

5.5 结果补充分析与文献对比

从结果一致性角度看, GBR 模型并不是在所有指标上都“绝对最优”, 但它在误差分布、边界样本与泛化稳定性之间取得了较好的平衡。RF 的优势主要体现在鲁棒性, 但在复杂非线性边界上略显保守; ANN 的表达能力更强, 但在小样本条件下容易记忆训练集噪声, 导致测试集误差波动更大。因而, 在本研究的数据库规模下, GBR 更适合作为主模型。

少量偏离较明显的样本主要集中在高强度区间与复杂工况组合区间。这些样本本身数量较少, 同时还受到界面处理方式、含水状态与接缝角度的耦合作用, 因此局部区间内出现轻微高估或低估是正常现象。更重要的是, 这些偏差并未形成明显的单向系统误差, 说明模型没有出现整体性偏移, 也没有出现明显的高值全部低估或低值全部高估现象。

与现有文献相比, 本研究的意义不仅在于给出一个较高的 R^2 值, 更在于建立了“数据库—模型—机理—工程”一体化的解释框架。换言之, 本文的贡献不是单纯追求数值精度, 而是把机器学习预测与界面物理机制联系起来, 使预测结果可以被工程语言解释, 也可以被施工控制直接使用。

因此, 从本科毕业设计的完成度来看, 本文已经具备较完整的研究链条: 一方面, 数据库保证了输入的可用性; 另一方面, 模型保证了输出的可预测性; 再进一步, 机理解释保证了结论的可理解性。三者结合后, 论文结论不再停留在算出来了, 而是能够回答工程上的实际。

5.6 与已有研究的对比及工程启示

与已有文献相比, 本文结果在总体趋势上是一致的: 界面处理工艺和基体含水状态依然是决定粘结性能的核心因素, 界面粗糙度的作用表现为适度最优, 而不是单调增强; 接缝角度则主要通过改变应力分配路径影响破坏模式。这说明本研究的数据库和模型输出并未偏离现有材料规律, 而是在已有认识基础上给出了更加量化的表达。

特别是, 与关于温度、湿度和表面处理的相关研究相比, 本文进一步强调了不同变量之间的耦合关系。例如, 温度并非独立起作用, 而是通过改变水化速率与水分迁移过程间接影响界面致密性; 含水状态也不是孤立参数, 而是与界面粗糙度和养护制度共同决定水化产物的分布, 这种耦合视角使得模型解释更接近真实工程场景。

研究结论的直接启示在于: 前期方案比选不应只看材料强度, 还应优先检查界面处理质量和施工含水控制; 施工中不应盲目追求极端粗糙, 而应追求可重复、可控制的中等粗糙窗口; 在复杂环境下, 模型应更多承担风险筛查和趋势预判职责, 而不是被当作唯一的绝对决策依据。

因此, 本文最重要的工程意义并不是给出一个固定不变的最优值, 而是构建了一套能够持续更新的判别逻辑: 当数据库扩展时, 模型可以继续训练; 当工程场景变化时, 参数建议可以重新修正; 当更多实测数据出现时, 机理解释也能同步迭代。

5.7 工程转化路径与研究框架总结

如果将本文的研究过程抽象为一个工程转化路径, 那么它实际上经历了三个层面: 第一层是数据层, 通过文献与工程实测构建可用数据库; 第二层是模型层, 通过 GBR、RF 与 ANN 识别变量之间的映射关系; 第三层是应用层, 通过特征重要性、SHAP、PDP 以及工程工况分析将模型输出转化为施工建议。三层结构彼此衔接, 形成了完整的研究闭环。

这种框架的最大优点在于可持续更新。随着新的试验数据不断加入, 模型参数可以重新训练; 随着新的工程案例出现, 参数推荐区间可以重新校正; 随着更细致的微观表征手段加入, 机理解释也可以进一步深化。因此, 本文所形成的不是一个封闭的结论集合, 而是一套可以不断迭代的分析路径。

对于本科毕业设计而言, 这种闭环框架的价值尤其明显: 它能够让论文在内容上兼具数据性、方法性和解释性, 不会停留在“有图有表”的浅层展示上。更重要的是, 它使得最终结论能够落实到施工控制和工程预警两个方向, 从而真正体现水工修复问题的工程属性。

5.8 综合施工建议与未来迭代路径

结合前文结果, 可以将水工混凝土修复界面的施工建议概括为四个关键词: 处理、含水、构造和养护。处理强调界面粗糙化应适度且均匀; 含水强调基体状态应控制在有利于水化但不过度积水的窗口; 构造强调接缝角度和受力路径需要与实际荷载相匹配; 养护则强调温度与时间要与材料特性同步协调。

从现场实施顺序来看, 建议先依据病害位置和受力条件确定修复方案, 再通过界面处理和含水控制实现施工准备, 随后在养护阶段观察界面早期反应, 最后利用模型给出的趋势判断对方案进行复核。这样的流程比单纯依赖经验更稳妥, 也更利于将机器学习结果转化为可操作的施工逻辑。

未来若要继续提高论文的学术层次, 可以从三个方向迭代: 一是增加更多工程实测样本, 尤其是不同温湿度与长期服役条件下的样本; 二是引入更细致的显微表征数据, 如SEM、CT或界面孔隙率指标; 三是将当前模型与有限元分析或数字孪生监测结

合，形成更完整的“试验—预测—验证”闭环。

7. 6 结论与展望			总字数：3288
相似文献列表			
去除本人文献复制比：31.2%(1026) 去除引用文献复制比：19.4%(639) 文字复制比：31.2%(1026) 疑似剽窃观点：(0)			
1	洮河碌曲县加格段堤防工程设计 徐子冰 - 《大学生论文联合比对库》 - 2025-05-26	19.6% (643)	是否引证：否
2	基于机器学习的UHPC多目标优化 柯超(导师：曾磊;Hua Zhang) - 《长江大学硕士学位论文》 - 2024-04-01	11.8% (387)	是否引证：是
3	制造业企业ESG表现对投资回报的影响研究 罗雨薇 - 《大学生论文联合比对库》 - 2025-06-22	4.7% (155)	是否引证：否
4	云南鲜花文创产品设计研究 李晓香 - 《大学生论文联合比对库》 - 2025-05-13	2.3% (76)	是否引证：否
5	宿迁港中心港区综合楼投标报价与造价分析 叶远航 - 《大学生论文联合比对库》 - 2025-04-27	1.9% (63)	是否引证：否
6	鹤山小学宿舍楼安装工程经济标编制 白俊杰 - 《大学生论文联合比对库》 - 2025-06-16	1.1% (37)	是否引证：否
7	毕业致谢合集10篇.docx. - 《网络 (https://www.renrendo) 》 - 2025	0.9% (31)	是否引证：否
8	202148510124-王骏玥-问题链教学法在中学历史课堂中的运用 王骏玥 - 《大学生论文联合比对库》 - 2025-06-05	0.6% (20)	是否引证：否
原文内容			

6 结论与展望

6.1 主要结论

超高性能混凝土（UHPC）是一种新型混凝土，具有比传统混凝土更好的力学、流变和耐久性。UHPC可以提供非常高的抗压强度，即超过150（MPa），可用于建造高容量结构并减少结构元件的截面。这种性能的改进必然会带来更大的建筑设计灵活性，并降低安装和人工成本。本文提出了一种基于GBR机器学习算法与RF相结合的UHPC优化设计方法，通过研究获得了以下结论：

（1）本研究基于文献调研，构建了包含108组样本的UHPC-NC界面性能数据库，并完成了ST、CS、CA、MC、JA与BS等关键变量的统一编码与标准化处理。数据统计表明，样本具有明显的异质性与离散性，但仍能够覆盖低、中、高强度区间，为后续建模提供了可用基础。

（2）在回归预测任务中，GBR模型表现出最佳综合性能，测试集 R^2 达到0.8761，RMSE为1.7692 MPa，明显优于RF与ANN。学习曲线和拟合对比表明，GBR在小样本、高非线性、多因素耦合的材料问题上具有更强的泛化能力。

（3）在破坏模式分类任务中，基于物理判别准则构建的标签体系能够有效支撑随机森林分类建模，测试准确率达到95.45%。这说明通过“物理标注+机器学习”方式，可以将原本缺失的失效形态信息转化为可训练、可验证的分类样本。

（4）机理分析表明，界面处理方式 ST 是影响粘结性能的核心因素，基体水饱和度 MC 与接缝角度 JA 则通过改变界面过渡区微结构和应力分布，对强度与破坏模式产生显著影响。因此，修复界面优化应从“粗糙化处理—含水控制—构造角度”三个层面协同设计。

6.2 创新点与不足

本研究的主要创新体现在三个方面：一是将跨文献实验数据与工程实测数据统一整合，建立了面向水工修复场景的结构化数据库；二是通过物理判别准则为缺失的破坏模式样本赋予可训练标签，实现了回归与分类的双目标建模；三是结合特征重要性、SHAP、偏依赖和工程工况分析，建立了“预测—解释—应用”一体化研究框架。

本研究仍存在以下不足：样本总量相对有限，部分变量主要来自文献离散报道，尚未形成统一的原始试验条件；极端温湿度与长期服役退化等因素覆盖不足；此外，当前的解释性分析仍以全局特征贡献和PDP为主，后续可进一步引入更系统的XAI工具，对单样本层面的预测结果开展局部解释。

6.3 未来工作展望

未来可从以下几个方向继续深化：

（1）目前建立的UHPC混凝土数据库虽然已经涵盖了108组实验数据，但数据集的完整性和多样性仍有提升空间，例如本数据库缺少有关添加了钢纤维的UHPC数据。因此需要收集更多UHPC相关数据，更为广泛和综合的数据能够为机器学习模型提供更为坚实的训练基础，从而得提高模型在复杂环境下的稳健性。

（2）综合考虑耐久性因素，构建面向服役寿命的界面性能演化模型。当前研究主要针对UHPC界面的短期力学性能开展分析，而实际工程结构在长期服役过程中往往会受到冻融循环、氯盐侵蚀、干湿循环、碳化作用以及长期荷载等多种环境因素的共

同影响。这些因素会导致界面微观结构逐渐劣化,进而引起黏结性能下降、裂缝扩展以及耐久性退化。因此,未来研究可进一步结合长期耐久性试验与多尺度分析方法,深入揭示UHPC界面在复杂环境耦合作用下的损伤演化规律。同时,可引入时间维度参数,构建界面性能随服役时间变化的演化模型,实现对UHPC结构全寿命周期性能的预测与评估,从而为桥梁、海洋工程及寒冷地区基础设施的长期安全服役提供理论依据。

(3) 推动机器学习与数字化工程技术融合,构建智能化工程决策平台。未来可进一步将机器学习模型与有限元分析、数字孪生监测技术以及施工参数优化方法相结合,形成面向工程应用的智能决策体系。例如,可通过有限元模型实现UHPC结构受力行为与界面损伤扩展过程的高精度模拟,并利用数字孪生技术实时获取结构运行状态数据,实现模型动态更新与性能预测。

展望未来,希望上述这些问题能在未来得到解决,为建筑行业向更高效、环保的方向发展做出贡献,为我们所生活的世界带来积极的改变。

参考文献

- [1] 韩志浩. PEF-UHPC与普通混凝土剪力键界面抗剪性能研究[D]. 青岛理工大学, 2025. DOI:10.27263/d.cnki.gqudc.2025.000629.
- [2] Kovačević M, Nyarko H M, Petronijević P, et al. Comparative Analysis of Machine Learning Models for Predicting Interfacial Bond Strength of Fiber-Reinforced Polymer-Concrete[J]. *Computation*, 2025, 13(1):17-17. DOI:10.3390/COMPUTATION13010017.
- [3] 张江江, 孙文, 鲜雪蕾, 等. 基于数字图像技术的UHPC-NC界面特征识别及粘结性能研究[J]. *硅酸盐通报*, 2025, 44(01):81-89. DOI:10.16552/j.cnki.issn1001-1625.2024.0817.
- [4] Gao F, Yang J, Huang Y, et al. Data-Driven Interpretable Machine Learning Prediction Method for the Bond Strength of Near-Surface Mounted FRP-Concrete[J]. *Buildings*, 2024, 14(9):2650-2650. DOI:10.3390/BUILDINGS14092650.
- [5] Sun W, Qian W, Cao X, et al. Effect of sucrose on the mechanical properties of magnesium and potassium phosphate cements at low temperatures and its mechanism of action[J]. *Case Studies in Construction Materials*, 2024, 20e02655-e02655. DOI:10.1016/J.CSCM.2023.E02655.
- [6] Wang Y, Ye N, Liu S, et al. Bond Strength Evaluation of FRP-Concrete Interfaces Affected by Hygrothermal and Salt Attack Using Improved Meta-Learning Neural Network[J]. *Applied Sciences*, 2024, 14(13):5474-5474. DOI:10.3390/APP14135474.
- [7] 柯超. 基于机器学习的UHPC多目标优化[D]. 长江大学, 2024. DOI:10.26981/d.cnki.gjhsc.2024.000618.
- [8] 叶缘. 新旧混凝土结合面剪切粘结性能研究[D]. 湖南大学, 2023. DOI:10.27135/d.cnki.ghudu.2023.000608.
- [9] Jihao G, Yuwei M, Jiyang F, et al. Utilization of fibers in ultra-high performance concrete: A review[J]. *Composites Part B*, 2022, 241DOI:10.1016/J.COMPOSITESB.2022.109995.
- [10] 李权恒. 界面剂和粗糙度对新老混凝土粘结性能的影响试验研究[D]. 吉林大学, 2022. DOI:10.27162/d.cnki.gjlin.2022.002629.
- [11] 高治鑫, 包腾飞, 李扬涛. 基于机器学习的混凝土坝表面裂缝快速识别方法[J]. *水电能源科学*, 2022, 40(04):95-98.
- [12] Jihui Q, Fumin D, Hongyan M, et al. Development and characterization of magnesium phosphate cement based ultra-high performance concrete[J]. *Composites Part B*, 2022, 234DOI:10.1016/J.COMPOSITESB.2022.109694.
- [13] Aman K, Chandra H A, Krishna K, et al. Prediction of FRCM-Concrete Bond Strength with Machine Learning Approach[J]. *Sustainability*, 2022, 14(2):845-845. DOI:10.3390/SU14020845.
- [14] ChunHong C, KwanLong W, JingWei C, et al. Deep Learning Methods for Remote Heart Rate Measurement: A Review and Future Research Agenda[J]. *Sensors*, 2021, 21(18):6296-6296. DOI:10.3390/S21186296.
- [15] 史金华, 史才军, 欧阳雪, 等. 超高性能混凝土受压弹性模量研究进展[J]. *材料导报*, 2021, 35(03):3067-3075.
- [16] 刘闻名. 基于计算机视觉和深度学习的桥梁结构损伤识别方法研究[D]. 哈尔滨工业大学, 2020. DOI:10.27061/d.cnki.ghgdu.2020.005809.
- [17] 陈庆, 马瑞, 蒋正武, 等. 基于GA-BP神经网络的UHPC抗压强度预测与配合比设计[J]. *建筑材料学报*, 2020, 23(01):176-183+191.
- [18] 周颖, 刘彤. 基于计算机视觉的混凝土裂缝识别[J]. *同济大学学报(自然科学版)*, 2019, 47(09):1277-1285.
- [19] Baqersad M, Sayyafi A E, Bak M H. State of the Art: Mechanical Properties of Ultra-High Performance Concrete[J]. *Civil Engineering Journal*, 2017, 3(3):
- [20] 国务院办公厅关于大力发展装配式建筑的指导意见[J]. *建筑技术开发*, 2017, 44(01):80.
- [21] Gu C, Ye G, Sun W. Ultrahigh performance concrete-properties, applications and perspectives[J]. *Science China Technological Sciences*, 2015, 58(4):587-599. DOI:10.1007/s11431-015-5769-4.
- [22] Tayeh A B, Bakar A H B, Johari M A M, et al. Utilization of Ultra-high Performance Fibre Concrete (UHPFC) for Rehabilitation - A Review[J]. *Procedia Engineering*, 2013, 54525-538. DOI:10.1016/j.proeng.2013.03.048.
- [23] 王少伟, 包腾飞, 徐波. 裂缝影响下碾压混凝土拱坝整体安全度评价[J]. *中南大学学报(自然科学版)*, 2013, 44(04):1479-1486.
- [24] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. [J]. *Journal of Machine Learning Research*, 2011, 122825-2830.
- [25] 张国新, 刘有志, 刘毅, 等. 特高拱坝施工期裂缝成因分析与温控防裂措施讨论[J]. *水力发电学报*, 2010, 29(05):45-51.
- [26] Santos M P, Júlio N E, Silva D V. Correlation between concrete-to-concrete bond strength and the roughness of the substrate surface[J]. *Construction and Building Materials*, 2006, 21(8):1688-1695. DOI:10.1016/j.conbuildmat.2006.05.044.
- [27] Júlio N E, Branco A F, Silva D V. Concrete-to-concrete bond strength. Influence of the roughness of

the substrate surface[J].Construction and Building Materials, 2004, 18(9):675-681. DOI:10.1016/j.conbuildmat.2004.04.023.

[28] Momayez A ,Ehsani M ,Ramezaniapour A , et al.Comparison of methods for evaluating bond strength between concrete substrate and repair materials[J].Cement and Concrete Research, 2004, 35(4):748-757. DOI:10.1016/j.cemconres.2004.05.027.

[29] Breiman L .Random Forests. [J].Machine Learning, 2001, 45(1):5-32. DOI:10.1023/A:1010933404324.

[30] Friedman H J .Greedy Function Approximation: A Gradient Boosting Machine[J].The Annals of Statistics, 2001, 29(5):1189-1232.

[31] 王甲泽. NC-UHPC界面粘结性能与微观机理研究[D]. 2025.

[32] 刘晋铭, 张寿松, 周亨, 等. 聚乙烯醇和超高分子量聚乙烯纤维对全珊瑚混凝土动态力学性能影响与数值模拟[J]. 复合材料学报, 2023, 40(06):3613-3625. DOI:10.13801/j.cnki.fhclxb.20220901.002.

[33] Zhan X, et al. Gradient boosting-based prediction of bond strength for UHPC-concrete interfaces under complex environments[J]. Materials, 2025.

致谢

行文至此，落笔为终。当论文即将画上句点，回望这段艰辛又充实的求学旅程，内心满是感激与不舍。曾几何时，我也是带着不甘和迷茫，不远万里来到兰州，是交大的朋友老师们的温暖令我释怀。此刻，我想用这篇致谢，向所有在我成长道路上给予关怀与帮助的人，献上最诚挚的谢意。

首先，我要向所有教过我、给予过我帮助的老师致以最崇高的敬意。我要感谢我的设计指导老师。从设计选题、框架搭建，到资料收集、修改完善，每一个环节都凝聚着老师的心血。大到逻辑结构，小到标点符号，老师都给予细致入微的指导。我要感谢我的班主任。我的班主任不管是在学业方面还是升学方面都给了我莫大的帮助，老师的每一句建议和教诲都十分务实纯然肺腑，这些都将成为我未来前行的宝贵财富。我要感谢学院的各位授课老师们，是你们精彩的课堂，为我构筑起扎实的专业知识体系。

其次，我要感谢我所有在学校以及在徐州的所有朋友们。从最开始到学校的局促、远离家乡的彷徨，到后来完全融入并爱上这座城市，都是因为在交大认识了一群好朋友、好舍友。他们包容我、鼓励我，让我的生活充满欢声笑语。同时也要感谢在徐州的好朋友们，在我每次压力大时倾听我的焦虑，在我被幸福包围时分享我的快乐。

最后，我要将最深的感恩献给我的家人。父母用无私的爱为我遮风挡雨，默默承担生活的重担，全力支持我的学业。他们的理解与包容，是我面对困难时最坚实的后盾。

致谢终有时，感恩无尽期。未来，我会带着这份温暖与力量，勇敢奔赴人生新征程，不负期望，努力成长为更优秀的人。

附录A 代表性样本数据节选

为便于展示数据库的原始结构，本附录给出部分代表性样本。实际建模时，所有样本均经过同一编码与标准化流程处理。

ST	CS	CA	MC	JA	BS
2.0	44.5	8.0	3.0	60.0	16.1
2.0	44.5	8.0	3.0	70.0	12.1
3.0	55.9	8.0	3.0	60.0	21.7
3.0	55.9	8.0	3.0	70.0	14.6
5.0	44.5	8.0	3.0	60.0	17.5
5.0	44.5	8.0	3.0	70.0	11.3
2.0	56.8	3.0	3.0	55.0	17.8
2.0	55.9	3.0	3.0	60.0	15.3
5.0	54.5	3.0	3.0	70.0	9.5
2.0	55.9	2.0	3.0	55.0	5.9
2.0	55.9	2.0	3.0	70.0	3.4
1.0	45.0	16.0	2.0	60.0	7.37
1.0	45.0	16.0	2.0	60.0	6.55
1.0	45.0	16.0	2.0	60.0	7.89
5.0	45.0	16.0	2.0	60.0	12.42
5.0	45.0	16.0	2.0	60.0	11.79
5.0	45.0	16.0	2.0	60.0	11.21
4.0	45.0	16.0	2.0	60.0	11.67
4.0	45.0	16.0	2.0	60.0	10.14
4.0	45.0	16.0	2.0	60.0	9.35

附录B 模型超参数设置与评价指标说明

本研究采用的三类模型及其核心超参数设置如下。该表用于补充正文中的训练过程说明，便于后续复现实验。

模型主要参数关键指标说明

RF n_estimators=100, max_depth=10 R²=0.7446 对比模型
GBR n_estimators=200, learning_rate=0.1, max_depth=5 R²=0.8761, RMSE=1.7692 MPa 最终主模型
ANN 隐藏层=64, 激活函数=ReLU, 优化器=Adam R²=0.5593 小样本对比模型
分类RF 网格搜索+5折交叉验证准确率=95.45% 破坏模式分类

1. 致谢

行文至此，落笔为终。当论文即将画上句点，回望这段艰辛又充实的求学旅程，内心满是感激与不舍。曾几何时，我也是带着不甘和迷茫，不远万里来到兰州，是交大的朋友老师们的温暖令我释怀。此刻，我想用这篇致谢，向所有在我成长道路上给予关怀与帮助的人，献上最诚挚的谢意。

首先，我要向所有教过我、给予过我帮助的老师致以最崇高的敬意。我要感谢我的设计指导老师。从设计选题、框架搭建，到资料收集、修改完善，每一个环节都凝聚着老师的心血。大到逻辑结构，小到标点符号，老师都给予细致入微的指导。我要感谢我的班主任。我的班主任不管是在学业方面还是升学方面都给了我莫大的帮助，老师的每一句建议和教诲都十分务实纯然肺腑，这些都将成为我未来前行的宝贵财富。我要感谢学院的各位授课老师们，是你们精彩的课堂，为我构筑起扎实的专业知识体系。

其次，我要感谢我所有在学校以及在徐州的所有朋友们。从最开始到学校的局促、远离家乡的彷徨，到后来完全融入并爱上这座城市，都是因为在校认识了一群好朋友、好舍友。他们包容我、鼓励我，让我的生活充满欢声笑语。同时也要感谢在徐州的好朋友们，在我每次压力大时倾听我的焦虑，在我被幸福包围时分享我的快乐。

最后，我要将最深的感恩献给我的家人。父母用无私的爱为我遮风挡雨，默默承担生活的重担，全力支持我的学业。他们的理解与包容，是我面对困难时最坚实的后盾。

致谢终有时，感恩无尽期。未来，我会带着这份温暖与力量，勇敢奔赴人生新征程，不负期望，努力成长为更优秀的人。

说明：1. 总文字复制比：被检测论文总重合字数在总字数中所占的比例

2. 去除引用文献复制比：去除系统识别为引用的文献后，计算出来的重合字数在总字数中所占的比例

3. 去除本人文献复制比：去除作者本人文献后，计算出来的重合字数在总字数中所占的比例

4. 单篇最大文字复制比：被检测文献与所有相似文献比对后，重合字数占总字数的比例最大的那一篇文献的文字复制比

5. 复制比：按照“四舍五入”规则，保留1位小数

6. 指标是由系统根据《学术论文不端行为的界定标准》自动生成的

7. 红色文字表示文字复制部分；绿色文字表示引用部分（包括系统自动识别为引用的部分）；棕灰色文字表示系统依据作者姓名识别的本人其他文献部分

8. 本报告单仅对您所选择的比对时间范围、资源范围内的检测结果负责



 amlc@cnki.net

 <https://check.cnki.net/>