

基于机器学习的护士职业倦怠风险预测模型

高铭阳¹ 袁晓彤¹ 吴玉兰² 尹娟¹

(¹ 大连大学护理学院 大连 116001 ² 大连市第七人民医院护理部 大连 116023)

〔摘要〕 **目的/意义** 构建我国三级医院护士职业倦怠风险预测模型，探索最优算法。**方法/过程** 采用问卷调查法，面向辽宁省三级医院护士，采集 23 项风险因素数据，并通过哥本哈根职业倦怠量表评估其职业倦怠情况。运用随机森林、极度随机树、梯度提升决策树、极端梯度提升树、K 近邻、支持向量机和逻辑回归 7 种单一算法，以及堆叠泛化集成模型策略，构建风险预测模型，以 AUC、准确率、特异度等指标评估模型性能。**结果/结论** 单一模型中逻辑回归表现最佳，集成模型中随机森林与逻辑回归融合的模式表现最佳。该集成模型可作为护士职业倦怠筛查的高效工具。

〔关键词〕 护士；职业倦怠；风险预测模型；机器学习

〔中图分类号〕 R-058 **〔文献标识码〕** A **〔DOI〕** 10.3969/j.issn.1673-6036.2025.11.008

A Risk Prediction Model for Nurse Burnout Based on Machine Learning

GAO Mingyang¹, YUAN Xiaotong¹, WU Yulan², YIN Juan¹

¹School of Nursing, Dalian University, Dalian 116001, China; ²Nursing Department, Dalian Seventh People's Hospital, Dalian 116023, China

〔Abstract〕 **Purpose/Significance** To construct a risk prediction model for job burnout among nurses in tertiary hospitals in China, and to explore the optimal algorithm. **Method/Process** By using the questionnaire survey method, 23 risk factor data are collected from nurses in tertiary hospitals in Liaoning province, and their job burnout status is evaluated through the Copenhagen burnout inventory. Seven single algorithms, namely random forest, extra-trees, gradient boosting decision tree (GBDT), extreme gradient boosting (XGBoost), K-nearest neighbors (KNN), support vector machine (SVM), and logistic regression, and a stacking ensemble model strategy are used to construct risk prediction models. Model performance is evaluated using indicators such as AUC, accuracy, and specificity. **Result/Conclusion** Among the single models, logistic regression performs the best. Among the stacking ensemble models, the model integrating random forest and logistic regression is the optimal one. The ensemble model can serve as an efficient tool for the screening of nurses' job burnout.

〔Keywords〕 nurse; job burnout; risk prediction model; machine learning

1 引言

护士作为医疗团队的核心成员，长期处于高强

度工作状态。三级医院患者病种复杂，工作强度高、节奏快，且突发事件频发，护士要熟练掌握基础操作，同时付出大量情绪劳动，成为职业倦怠高危人群^[1]。《国际疾病分类》2024 版^[2]将职业倦怠

〔修回日期〕 2025-09-29

〔作者简介〕 高铭阳，硕士研究生；通信作者：吴玉兰，副主任护师。

描述为由长期工作压力未得到成功管理所引起的一种综合征。护士职业倦怠不仅会使其对护理工作丧失动力、对患者缺乏热情,长期情绪失控更易引发焦虑、抑郁,身体上出现眩晕、头痛、恶心等表现,身心失衡严重危害其健康^[3]。

目前测量护士职业倦怠主要依赖各种主观量表^[4]。此类量表虽为评估提供了参考,但存在固有限制,结果易受被试者主观态度、作答偏差影响,难以全面反映职业倦怠客观状态。机器学习^[5]是人工智能技术的重要分支,具有良好的数据处理和前瞻性预测能力。国内关于护士职业倦怠的风险模型构建仍处于探索阶段,现有研究多以统计学方法构建风险预测模型^[3]。探索多元分析方法,构建符合我国护士群体特点的风险预测模型具有重要研究价值。本研究基于文献调研,设计调查问卷,引入客观指标,开发适合护理管理者使用的高效、多维度护士职业倦怠评估工具,并运用机器学习算法构建预测模型,探索最优算法。

2 对象与方法

2.1 研究对象

以辽宁省三级医院临床护士为调查对象。纳入标准:工作地点为三级医院,具有注册护士证书,且实际从事护理工作;工龄(执业注册后在岗工作时间)不低于 1 年,不超过 30 年。排除标准:因病假、事假、产假等原因不在岗,或者处于返聘、进修学习状态;工作地点为三级民营医院。本研究经大连市第七人民医院伦理委员会批准,伦理批件号:2025-04。

2.2 研究设计

2.2.1 样本量计算 根据既往研究^[6],护士职业倦怠发生率约 41%,设允许误差为 5%, α 值为 0.05,通过 PASS 15.0 软件中 Proportions 模块计算样本量 $N=372$,按科室与工龄进行分层整群抽样。考虑每层脱落率为 20%,至少应纳入 500 人。

2.2.2 调查工具 一是护士工作情况调查问卷。通过系统综述,归纳护士职业倦怠风险因素主要包

括个体、工作特征和组织环境 3 类。结合我国国情与护士工作特点设计问卷,共 23 个客观题项。经检验,问卷的取样适当性(Kaiser-Meyer-Olkin, KMO)值为 0.786,表明其具有良好的结构效度,适合因子分析。二是哥本哈根职业倦怠量表(the Copenhagen burnout inventory, CBI),由 Kristensen T S 等^[4]于 2005 年编制,含个体、工作、患者或同事倦怠 3 个维度,19 个条目,采用李克特 5 级评分(1—5 分),总分 19~95 分。原量表 Cronbach's α 系数为 0.86,各维度 Cronbach's α 系数为 0.85~0.87。2020 年吴军等^[7]对其进行修订,修订后量表 Cronbach's α 系数为 0.936,各维度 Cronbach's α 系数为 0.853~0.932, KMO=0.945,信效度良好,适用于我国医务人员。因此采用修订版 CBI 测量护士职业倦怠程度,50~95 分为职业倦怠,19~49 分为非职业倦怠。本次调查 Cronbach's α 系数为 0.975,各维度分 Cronbach's α 系数为 0.834~0.960,信度良好。

2.2.3 调查方法 2025 年 3—5 月,先以便利抽样对符合纳排标准的 25 名护士开展预调查,完善调查工具。之后采用分层整群抽样,通过问卷星发放问卷,同 IP 限答 1 次,要求填写完整。剔除作答低于 2 分钟、答案规律重复、不符合纳排标准、未通过身份验证以及注意力题回答错误的问卷。共回收 998 份问卷,有效问卷 908 份,有效率为 90.98%。

2.2.4 统计学分析 使用 SPSS 27.0 分析数据,分类变量以 $n(\%)$ 表示,组间比较应用卡方检验或 Fisher 确切概率法;连续变量以中位数和四分位数间距表示,组间比较应用 Mann-Whitney U 检验。单因素分析非职业倦怠组与职业倦怠组差异, $P<0.05$ 表示差异有统计学意义。问卷信度评价中,采用 Cronbach's α 系数检验内部一致性信度,采用探索性因子分析检验结构效度。

2.2.5 数据预处理与分析 为构建非职业倦怠(标签“0”,441 例)与职业倦怠(标签“1”,467 例)二分类预测模型,对变量进行提取并编码。数据集类别基本平衡(比例接近 1:1),避免了后续模型训练因类别不平衡产生偏差。对数据进行平滑与

归一化预处理, 然后进行基线分析、共线性分析、特征重要性排序。使用 Python 3.6 及其 sklearn.ensemble、sklearn.model_selection、sklearn.metrics 等库进行模型搭建和相关计算分析。

2.2.6 预测模型构建 采用集成学习中的堆叠泛化 (stacking) 集成模型策略^[8], 以提升预测精度与泛化能力。(1) 算法选择。选取随机森林 (random forest, RF)、极度随机树 (extra trees, ET)、梯度提升决策树 (gradient boosting decision tree, GBDT)、极端梯度提升树 (extreme gradient boosting, XGBoost)、K 近邻 (K-nearest neighbors, KNN)、支持向量机 (support vector machine, SVM) 和逻辑回归 (logistic regression, LR) 7 种算法。(2) 分层交叉验证。采用 10 折分层交叉验证, 确保每折中训练集和测试集的类别分布 (非职业倦怠/职业倦怠) 与原始数据一致。每折并行训练多个基础模型并计算评估指标。(3) 最优折选取。分析每种模型组合在 10 折交叉验证中所有测试集上的受试者工作特征曲线下面积 (area under the receiver operating characteristic curve, AUC) 表现, 选取 AUC 值最高的第 8 折, 作为后续模型构建基础。(4) 模型训练与测试。基于选定的第 8 折数据划分方案, 使用其中的训练集独立训练 7 种不同的学习器, 并在对应测试集上获取各模型的预测结果。(5) 模型性能评估。采用 AUC、F1 值、敏感性、特异性、准确率、平均精确率、阳性预测值、阴性预测值评估模型性能。(6) Stacking 集成模型构建。以最优模型作为元学习器, 同时选取排名前 3 位的模型作为基学习器 (包含元学习器), 构建 4 种 Stacking 集成模型 (包括 3 种基学习器的两两组合模型, 以及 3 种基学习器的整体组合模型), 对最终选定的最优 Stacking 集成模型进行全面评价, 包括受试者工作特征曲线、决策曲线、校准曲线。使用 SHAP 值^[9]量化分解各输入特征对预测结果的边际贡献, 以明确并展示预测模型中每个特征的贡献程度。筛选性能最优者为最终预测模型。模型构建流程, 见图 1。

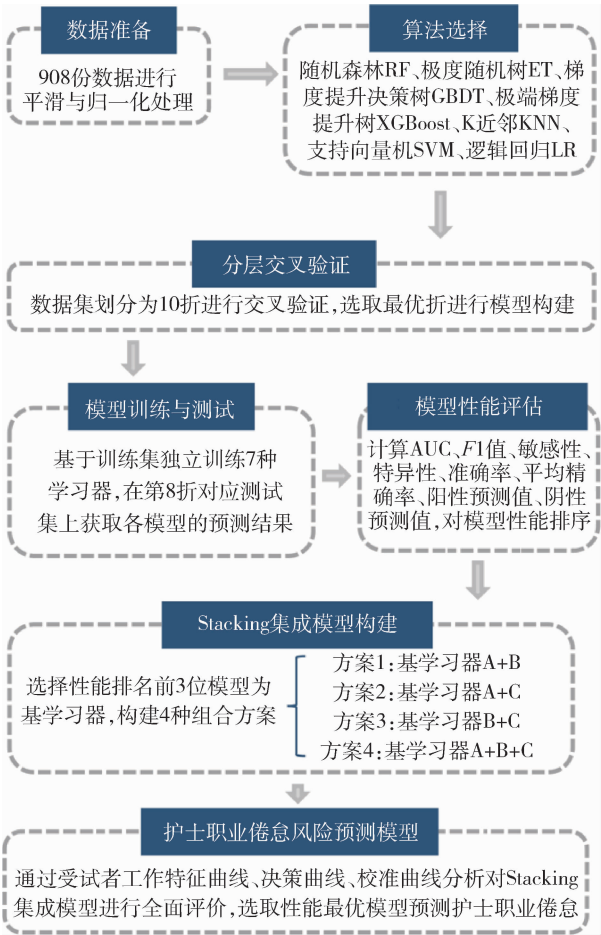


图 1 模型构建流程

3 结果

3.1 风险因素特征

本研究共纳入 908 例数据, 职业倦怠组 467 例, 非职业倦怠组 441 例。两组在科室、聘用形式、每周工作时间、遭受暴力次数、被投诉次数、被批评次数的组间差异均有统计学意义 ($P < 0.05$), 数据集基线特征, 见表 1。共线性分析结果, 见图 2。年龄与工龄相关系数为 0.97, 呈强相关; 婚姻状况与子女数量相关系数为 0.62, 呈显著正相关; 性别与科研时间占比, 考试或培训次数与患者沟通时间占比, 相关系数接近 0, 线性关联弱。

表 1 护士职业倦怠风险预测模型数据集基线特征

| 变量 | | 非职业倦怠组（n = 441） | 职业倦怠组（n = 467） | P |
|--|-------------|---------------------|---------------------|---------|
| 性别 [n（%）] | 女 | 393(89.1) | 428(91.6) | 0.236 7 |
| | 男 | 48(10.9) | 39(8.4) | |
| 年龄 [岁, M（P ₂₅ , P ₇₅ ）] | | 32(27, 41) | 33(28, 42) | 0.086 5 |
| 工龄 [年, M（P ₂₅ , P ₇₅ ）] | | 10(5, 17) | 10(5, 20) | 0.308 6 |
| 科室 [n（%）] | 内科病房 | 118(26.7) | 104(22.3) | 0.006 1 |
| | 外科病房 | 118(26.7) | 107(22.9) | |
| | 手术室 | 35(7.9) | 56(11.9) | |
| | 重症监护室 | 45(10.2) | 49(10.4) | |
| | 急诊 | 36(8.2) | 53(11.4) | |
| | 儿科 | 40(9.2) | 55(11.8) | |
| | 精神科和心理科 | 49(11.1) | 43(9.3) | |
| 教育程度 [n（%）] | 中专 | 8(1.8) | 10(2.1) | 0.320 0 |
| | 大专 | 47(10.7) | 50(10.7) | |
| | 大学本科（非全日制） | 225(51.0) | 218(46.7) | |
| | 大学本科（全日制） | 152(34.5) | 175(37.5) | |
| | 硕士研究生（非全日制） | 8(1.8) | 7(1.5) | |
| | 硕士研究生（全日制） | 1(0.2) | 7(1.5) | |
| 被聘用职称 [n（%）] | 护士 | 126(28.6) | 100(21.4) | 0.624 6 |
| | 护师 | 137(31) | 152(32.6) | |
| | 主管护师 | 127(28.8) | 166(35.6) | |
| | 副主任护师 | 43(9.7) | 47(10.0) | |
| | 主任护师 | 8(1.9) | 2(0.4) | |
| 婚姻状况 [n（%）] | 未婚 | 163(36.9) | 165(35.4) | 0.951 2 |
| | 已婚 | 260(58.9) | 293(62.7) | |
| | 离异 | 15(3.5) | 8(1.7) | |
| | 丧偶 | 3(0.7) | 1(0.2) | |
| 聘用形式 [n（%）] | 合同制（院聘） | 258(58.5) | 225(48.2) | 0.010 4 |
| | 编制内（事业编制） | 173(39.2) | 227(48.6) | |
| | 劳务派遣（第三方聘用） | 10(2.3) | 15(3.2) | |
| 子女数量 [个, M（P ₂₅ , P ₇₅ ）] | | 2(1, 2) | 2(1, 2) | 0.691 2 |
| 一个月内夜班数 [次, M（P ₂₅ , P ₇₅ ）] | | 3(0, 6) | 3(0, 7) | 0.849 4 |
| 工资区间 [元/月, M（P ₂₅ , P ₇₅ ）] | | 6 000(4 000, 7 000) | 5 000(4 000, 7 000) | 0.429 9 |
| 每周工作时间 [小时, M（P ₂₅ , P ₇₅ ）] | | 43(40, 48) | 48(40, 48) | 0.000 0 |
| 执行治疗时间占比 [% , M（P ₂₅ , P ₇₅ ）] | | 50(0, 75) | 50(25, 75) | 0.480 6 |
| 文书工作时间占比 [% , M（P ₂₅ , P ₇₅ ）] | | 50(0, 75) | 50(0, 75) | 0.497 0 |
| 患者沟通时间占比 [% , M（P ₂₅ , P ₇₅ ）] | | 25(0, 50) | 25(0, 50) | 0.334 3 |
| 管理工作时间占比 [% , M（P ₂₅ , P ₇₅ ）] | | 50(25, 50) | 50(25, 50) | 0.934 6 |
| 科研时间占比 [% , M（P ₂₅ , P ₇₅ ）] | | 0(0, 50) | 0(0, 50) | 0.135 8 |
| 遭受暴力次数 [次, M（P ₂₅ , P ₇₅ ）] | | 0(0, 1) | 0(0, 2) | 0.000 0 |
| 被投诉次数 [次, M（P ₂₅ , P ₇₅ ）] | | 0(0, 0) | 0(0, 0) | 0.000 0 |
| 被批评次数 [次, M（P ₂₅ , P ₇₅ ）] | | 0(0, 0) | 0(0, 2) | 0.000 0 |
| 单位表扬次数 [次, M（P ₂₅ , P ₇₅ ）] | | 0(0, 1) | 0(0, 1) | 0.242 7 |
| 患者感谢次数 [次, M（P ₂₅ , P ₇₅ ）] | | 0(0, 1) | 0(0, 1) | 0.380 2 |
| 考试或培训次数 [次, M（P ₂₅ , P ₇₅ ）] | | 0(0, 4) | 0(0, 4) | 0.893 4 |

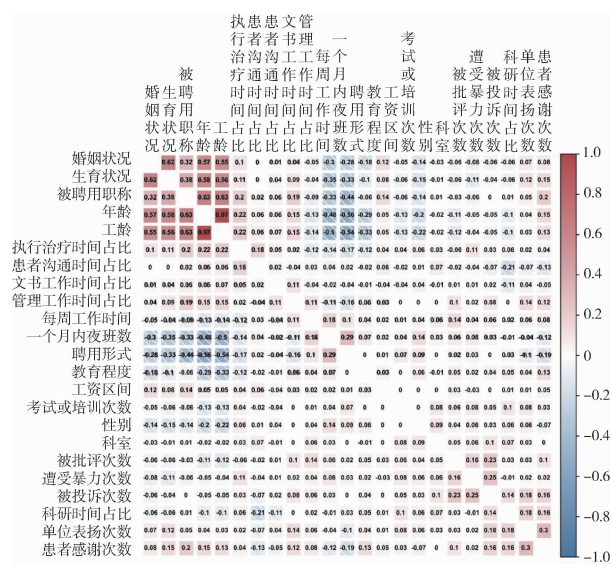


图 2 共线性分析

特征重要性排序，见图 3。被批评次数的重要性最高，其次是遭受暴力次数、每周工作时间、年龄、科室；婚姻状况最低。

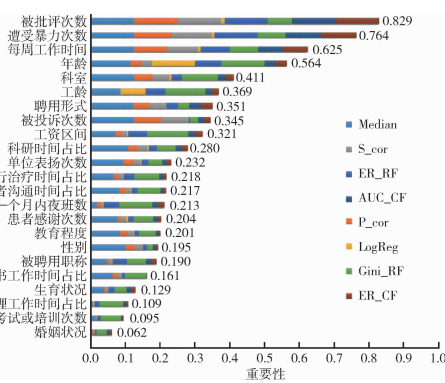


图 3 风险因素特征重要性排序

表 2 7 种机器学习算法的预测性能

| 算法 | AUC | 准确率 | 敏感性 | 特异性 | F1 | 阳性预测值 | 阴性预测值 | 平均精确率 |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|
| RF | 0.718 | 0.667 | 0.617 | 0.717 | 0.652 | 0.691 | 0.647 | 0.620 |
| ET | 0.695 | 0.624 | 0.489 | 0.761 | 0.568 | 0.677 | 0.593 | 0.589 |
| KNN | 0.667 | 0.581 | 0.426 | 0.739 | 0.506 | 0.625 | 0.557 | 0.556 |
| GBDT | 0.692 | 0.634 | 0.596 | 0.674 | 0.622 | 0.651 | 0.620 | 0.592 |
| XGBoost | 0.717 | 0.645 | 0.575 | 0.717 | 0.621 | 0.675 | 0.623 | 0.603 |
| SVM | 0.687 | 0.645 | 0.575 | 0.717 | 0.621 | 0.675 | 0.623 | 0.603 |
| LR | 0.730 | 0.720 | 0.638 | 0.804 | 0.698 | 0.769 | 0.685 | 0.674 |

综合评估显示，LR 表现最佳，AUC 和准确率显著高于其他模型，且特异性和阳性预测值表现突出，在正确识别非职业倦怠护士以及预测可靠性方面具有显著优势。RF 和 XGBoost 的 AUC 值表现次之，且 RF 显示出较高的敏感性。ET 虽然特异性较高，但敏感性过低，导致其漏诊率较高。GBDT 和 SVM 表现中等。KNN 在所有指标上均表现最差，不适于本预测任务。最终 LR 模型被确定为性能最优的单一模型。后续 Stacking 集成学习将以 LR 为元学习器，LR、RF 和 XGBoost 模型作为基学习器。

3.3 Stacking 集成模型结果

RF_LR Stacking 集成模型在 AUC (0.764)、准

注：Median 为中位数，S_cor 为 Spearman 相关系数，ER_RF 为随机森林的错误率，AUC_CF 为分类器的曲线下面积，P_cor 为 Pearson 相关系数，LogReg 为逻辑回归系数，Gini_RF 为随机森林的基尼系数，ER_CF 分类器的错误率。

3.2 单一预测模型结果

7 种算法在测试集的结果，见表 2。

确率 (0.720)、特异性 (0.804) 上表现最优，见表 3。融合 3 种算法的 RF_XGBoost_LR Stacking 集成模型 AUC (0.763)、准确率 (0.699)、特异性 (0.761) 表现次之。RF_XGBoost 和 XGBoost_LR Stacking 集成模型表现稍弱，两者特异性均为 0.761，平均精确率分别为 0.632、0.641，整体在减少非职业倦怠群体误判与提升预测结果可信度上表现不及前两个模型。4 种模型的 AUC 均集中在 0.755 ~ 0.764 之间，核心指标差异较小，但结合临床筛查需求，要优先保证减少漏诊与准确识别非倦怠群体，RF_LR Stacking 集成模型在特异性与准确率等关键指标上的优势，成为综合性能最优的预测模型。各模型受试者工作特征曲线、决策曲线和校准曲线，见图 4。

表 3 4 种 stacking 集成模型的预测性能

| 算法 | AUC | 准确率 | 敏感性 | 特异性 | F1 | 阳性预测值 | 阴性预测值 | 平均精确率 |
|---------------|-------|-------|-------|-------|-------|-------|-------|-------|
| RF_XGBoost | 0.759 | 0.677 | 0.596 | 0.761 | 0.651 | 0.718 | 0.648 | 0.632 |
| RF_LR | 0.764 | 0.720 | 0.638 | 0.804 | 0.698 | 0.769 | 0.685 | 0.674 |
| XGBoost_LR | 0.755 | 0.688 | 0.617 | 0.761 | 0.667 | 0.725 | 0.660 | 0.641 |
| RF_XGBoost_LR | 0.763 | 0.699 | 0.638 | 0.761 | 0.682 | 0.732 | 0.673 | 0.650 |

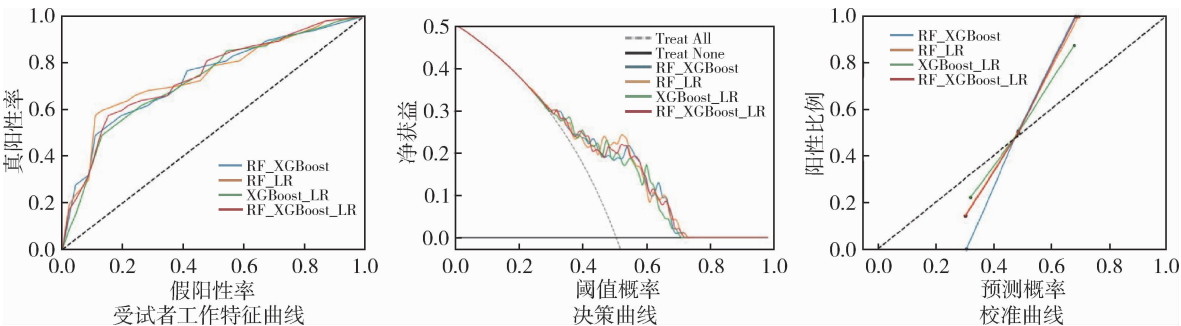


图 4 4 种 Stacking 集成模型性能评价

3. 4 模型特征贡献（图 5）

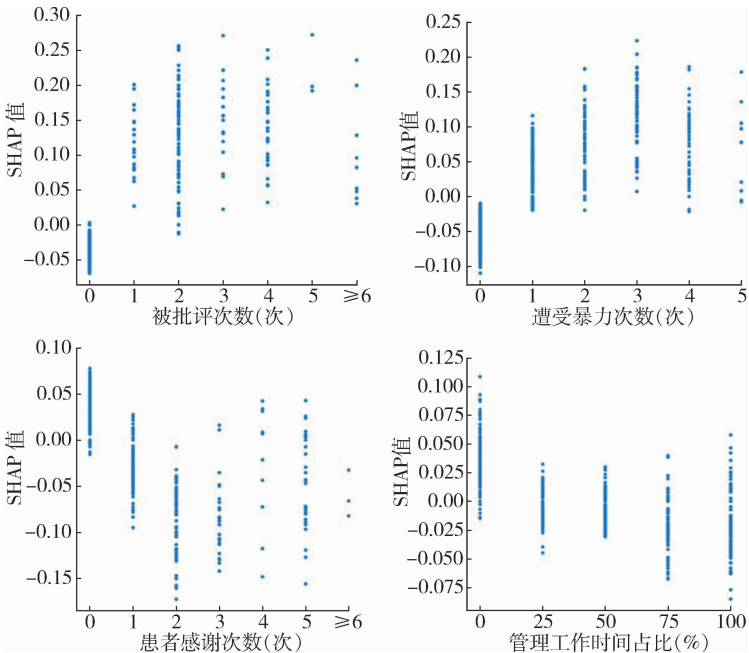


图 5 模型特征贡献 SHAP 值

模型纳入 23 个风险特征，核心风险特征为被批评次数、遭受暴力次数、被投诉次数、一个月内夜班数。当每年被批评 4 次及以上时，SHAP 值急剧上升（0.05 ~ 0.25），每增加 1 次批评，职业倦怠风险提升 8.2%。每遭受 1 次暴力事件，SHAP 值增加 0.03，职业倦怠风险提升 6.5%，当每年遭受

暴力事件 3 次及以上时，SHAP 值突破 0.1（高风险区）。当每年被投诉 3 次及以上时，SHAP 值突破 0.1（风险陡增），当每年被投诉 5 次时，SHAP 值达峰值 0.125。当一个月内夜班不超过 4 次时，影响微弱；当一个月内夜班超过 8 次时，职业倦怠风险急剧提升。关键保护性特征为患者感谢次数、管

理工作时间占比。当护士每年收到患者感谢 4 次及以上时,每增加一次感谢,职业倦怠风险降低 4.3%。每周参与管理工作时间占比维持在 20% ~ 30% 时,不会引发职业倦怠风险,如果过低(低于 10%)或过高(高于 40%),职业倦怠风险会上升 5%。

4 讨论

4.1 单一模型预测性能整体较高,逻辑回归模型表现最优

7 种单一机器学习算法性能差异显著,其中 LR 展现最佳综合性能,显著优于 RF、XGBoost 等树类算法。该结果与周闯等^[10]提出的线性模型在高特异度需求场景中更具优势的结论高度一致,印证机器学习处理多维度特征时,结合应用场景选择算法的重要性。LR 的高特异度可有效识别非倦怠群体,适合作为临床初筛工具;其阳性预测值较高,有助于护理管理者快速定位高危个体。虽然树类算法在捕捉非线性特征关联方面具有潜力,但其敏感度低于 LR,与张维等^[11]指出的树类算法在样本量有限时易过拟合导致漏诊的结论相符,该缺陷在筛查场景中可能遗漏部分早期职业倦怠筛查对象,提示大规模筛查应优先选择 LR 等兼具精度与稳定性的算法。

4.2 随机森林与逻辑回归 Stacking 集成模型实现性能突破

采用 Stacking 集成模型策略融合基学习器后,RF_LR Stacking 集成模型预测性能显著提升,AUC 较单一 LR 提升 4.7%,且其准确率、F1 值等核心指标表现最佳。印证集成学习融合不同算法优势可提升模型泛化能力的理论^[12]。融合 3 种算法的 RF_XGBoost_LR Stacking 集成模型与 RF_LR Stacking 集成模型性能接近,但计算复杂度更高,提示应权衡模型性能与部署成本。Kong G 等^[13]也应用 Stacking 集成模型提升了预测性能,但本研究中 RF_LR Stacking 集成模型 AUC 更高,能更精准地区分两类群体,尤其适用于须动态监测的高风险科室

(如急诊科、手术室)。结合受试者工作特征曲线分析,该模型在假阳性率与真阳性率的平衡上表现优异,可为护理管理者提供可靠决策依据。

4.3 核心特征贡献揭示干预靶点,为政策落地提供精准路径

《全国护理事业发展规划(2021—2025)》^[14]中提出医疗机构为护士提供必要的卫生防护和医疗保健措施,有效预防和减少护士在职业环境中可能受到的危害,切实维护和保障护士合法权益,以推动护理行业高质量发展。SHAP 值分析揭示了模型预测核心逻辑。被批评次数、遭受暴力次数、被投诉次数是权重最高的风险特征,年被批评次数每增加 1 次,职业倦怠风险提高 8.2%,年暴力事件达到 3 次,护士发生职业倦怠进入高风险区间(SHAP 值超过 0.1),与基线分析及 Haight B L 等^[15]负性事件累积效应的结论一致。据此医疗机构可建立负性事件分级响应机制;优化投诉处理流程,设立多部门投诉调解小组;建立暴力零容忍机制,配备安防设备并开展应急培训,并纳入医院安全考核^[16]。患者感谢次数与管理工作时间占比(20% ~ 30%)具有显著保护效应,与谷良丹等^[17]的职业价值感可缓冲工作压力研究结果相印证。每年收到患者感谢 4 次及以上,可使职业倦怠风险降低 4.3%,管理工作占比低于 10% 或高于 40% 则风险上升 5%。医院可搭建患者反馈平台(如电子感谢系统)、设计分层级管理培训路径,避免管理职责失衡致压力过载。针对暴力事件心理创伤,配备专项支持方案:由专业咨询师提供一对一匿名疏导;对严重创伤者,启动持续干预(阶段性辅导、团体支持),以修复其心理状态^[18]。上述措施可实现风险防控与正向激励双重发力,为政策落地提供精准路径。

4.4 模型部署,从技术工具到实践方案

本研究落地方案包括以下几方面。一是模型工程化,选取 RF_LR Stacking 集成模型进行轻量化改造,确保集成至医院系统后单批次响应时间在 3 秒之内。二是多源数据采集,通过系统接口自动获取工龄、科室等结构化数据,开发端口供护理管理者录入动态指标,内置逻辑校验保障质量。三是平台搭建,建立可

视化预测平台, 输出结果并借 SHAP 值揭示关键特征, 使用仪表盘按个体、工作特征等维度展示风险分布。四是落地保障, 遵循数据隐私保护规范, 形成“数据采集-风险预测”闭环管理, 为护士职业倦怠动态监测提供持续精准保障^[19]。

5 结语

本研究构建的 LR 及 RF_LR Stacking 集成模型为护士职业倦怠筛查提供了高效工具, 后者融合线性与非线性算法优势, 保持可解释性与预测精度。SHAP 值分析揭示的核心特征, 为政策落地提供靶向路径。模型部署能有效推动模型从技术研发阶段转化为临床护理管理的实践应用工具, 助力护理管理者高效开展护士职业倦怠的日常筛查工作, 实现技术转实践的有效落地。本研究仍存在局限, 虽然模型在辽宁省三级医院样本中表现良好, 但其普适性有待不同地区、不同等级医疗机构的多中心数据验证。未来将扩大样本, 覆盖基层医疗机构, 通过动态更新模型参数提升普适性, 为我国护理事业进步及护理队伍稳定发展提供支撑。

作者贡献: 高铭阳负责问卷设计、预测模型构建、论文撰写; 袁晓彤负责问卷调查实施、数据收集与整理; 吴玉兰、尹娟负责论文审核与修订。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 吴淑君, 冯伟杰. 护士职业倦怠测评工具及心理干预措施的研究进展 [J]. 职业与健康, 2024, 40 (6): 842-846.
- 2 World Health Organization. ICD-11 for mortality and morbidity statistics 2024-01 [EB/OL]. [2024-09-30]. <https://icd.who.int/browse/2024-01/mms/zh#129180281>.
- 3 沈月, 周炜, 肖雅, 等. 临床护士重度职业倦怠预测模型的构建 [J]. 护理管理杂志, 2023, 23 (10): 797-802.
- 4 KRISTENSEN T S, BORRITZ M, VILLADSEN E, et al. The Copenhagen burnout inventory: a new tool for the assessment of burnout [J]. Work & stress, 2005, 19 (3): 192-207.
- 5 张丽达, 梁晟, 农世相, 等. 基于机器学习算法构建晚期直肠癌病人疼痛危象预测模型 [J]. 护理研究, 2025, 39 (17): 2900-2907.
- 6 CHAKRABORTY R, CHATTERJEE A, CHAUDHURY S.

Internal predictors of burnout in psychiatric nurses: an Indian study [J]. Industrial psychiatry journal, 2012, 21 (2): 119-124.

- 7 吴军, 刘文斌, 冯婷, 等. 哥本哈根倦怠问卷的修订及信效度检验 [J]. 护理研究, 2020, 34 (23): 4149-4153.
- 8 MAO Z, XIA M, JIANG B, et al. Incipient fault diagnosis for high-speed train traction systems via stacked generalization [J]. IEEE transactions on cybernetics, 2020, 52 (8): 7624-7633.
- 9 WU J, CHAN Y T, LU Y, et al. The tumor microenvironment in the postsurgical liver: mechanisms and potential targets of postoperative recurrence in human hepatocellular carcinoma [J]. Medicinal research reviews, 2023, 43 (6): 1946-1973.
- 10 周闯, 金学勤, 郭正丽, 等. 不同机器学习算法的社区老年人认知衰弱风险预测模型比较 [J]. 护理学杂志, 2023, 38 (19): 1-5, 11.
- 11 张维, 张浩晨. 一种基于最优集成随机森林的小样本数据特征提取方法 [J]. 西北工业大学学报, 2022, 40 (6): 1261-1268.
- 12 MAHAJAN P, UDDIN S, HAJATI F, et al. Ensemble learning for disease prediction: a review [J]. Healthcare, 2023, 11 (12): 1808.
- 13 KONG G, WU J, CHU H, et al. Predicting prolonged length of hospital stay for peritoneal dialysis-treated patients using stacked generalization: model development and validation study [J]. JMIR medical informatics, 2021, 9 (5): e17886.
- 14 国家卫生健康委员会. 全国护理事业发展规划 (2021—2025 年) [EB/OL]. [2025-09-24]. https://www.fj.gov.cn/zwgk/ztlz/tjzfznzb/zcwj/gj/202205/t20220510_5906448.htm.
- 15 HAIGHT B L, PEDDIE L, CROSSWELL A D, et al. Combined effects of cumulative stress and daily stressors on daily health [J]. Health psychology, 2023, 42 (5): 325.
- 16 MONTGOMERY A P, PATRICIAN P A, AZUERO A. Nurse burnout syndrome and work environment impact patient safety grade [J]. Journal of nursing care quality, 2022, 37 (1): 87-93.
- 17 谷良丹, 林金燕, 陈瑶偏, 等. 精神科一线临床护士职业获益感与职业价值观和知觉压力的相关性 [J]. 健康研究, 2024, 44 (3): 259-264.
- 18 廖俪雯, 王琳. 遭受工作场所暴力急诊护士情绪调节策略在应对方式与抗逆力之间的调节效应 [J]. 护理学杂志, 2023, 38 (13): 1-6.
- 19 周丽娟, 温贤秀, 蒋蓉, 等. 机器学习在护理领域中的应用研究进展 [J]. 护士进修杂志, 2022, 37 (15): 1388-1392.