

基于机器学习与SHAP解释的冠心病风险预测模型研究

杨冰倩^{1,2*}, 张艳敏^{1,2}, 孙凤杰¹, 赵占坤¹

¹河北软件职业技术学院软件工程系, 河北 保定

²保定市人工智能 + 智慧交通领域重点实验室, 河北 保定

收稿日期: 2026年8月17日; 录用日期: 2026年9月30日; 发布日期: 2026年10月13日

摘要

本文基于UCI Cleveland心脏病数据集, 采用机器学习方法构建冠心病风险预测模型。对数据进行预处理后, 分别建立逻辑回归、支持向量机、随机森林、XGBoost和LightGBM五种分类模型, 并结合交叉验证、多指标评价及可靠性分析比较模型性能。同时, 利用SHAP方法对最优模型进行解释, 分析关键临床特征对预测结果的影响。实验结果表明, 五种模型均具有较好的风险识别能力, 其中逻辑回归模型综合性能最佳, 测试集ROC-AUC达到0.9436。Bootstrap、校准曲线和决策曲线分析结果表明, 该模型具有较好的稳定性和应用价值。SHAP分析进一步发现, 年龄、最大心率、运动诱发心绞痛和ST段压低等因素对模型预测具有重要影响。基于机器学习结合可解释分析的方法能够有效实现冠心病风险预测, 并提高模型结果的透明性, 为临床辅助决策提供参考。

关键词

冠心病, 机器学习, 风险预测, 可解释性

Research on Coronary Heart Disease Risk Prediction Model Based on Machine Learning and SHAP Interpretation

Bingqian Yang^{1,2*}, Yanmin Zhang^{1,2}, Fengjie Sun¹, Zhankun Zhao¹

¹Department of Software Engineering, Hebei Software Institute, Baoding Hebei

²Baoding Key Laboratory of Artificial Intelligence and Intelligent Transportation, Baoding Hebei

Received: August 17, 2026; accepted: September 30, 2026; published: October 13, 2026

*通讯作者。

文章引用: 杨冰倩, 张艳敏, 孙凤杰, 赵占坤. 基于机器学习与 SHAP 解释的冠心病风险预测模型研究[J]. 软件工程与应用, 2026, 15(5): 654-663. DOI: 10.12677/sea.2026.155061

Abstract

This paper constructs a risk prediction model for coronary heart disease (CHD) using machine learning methods based on the UCI Cleveland Heart Disease dataset. After data preprocessing, five classification models—Logistic Regression, Support Vector Machine, Random Forest, XGBoost, and LightGBM—were established. Their performances were compared using cross-validation, multi-metric evaluation, and reliability analysis. Meanwhile, the SHAP (SHapley Additive exPlanations) method was applied to interpret the optimal model and to analyze the influence of key clinical features on prediction outcomes. Experimental results show that all five models exhibit good risk discrimination ability, with the Logistic Regression model achieving the best overall performance, attaining a test-set ROC-AUC of 0.9436. Bootstrap resampling, calibration curve, and decision curve analyses indicate that the model has satisfactory stability and clinical utility. SHAP analysis further reveals that age, maximum heart rate, exercise-induced angina, and ST-segment depression are among the most influential factors in model predictions. The proposed approach, combining machine learning with interpretable analysis, can effectively predict CHD risk and enhance the transparency of model outputs, thereby providing a reference for clinical decision support.

Keywords

Coronary Heart Disease, Machine Learning, Risk Prediction, Interpretability

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

冠心病是常见的心血管疾病，早期识别高风险人群对于疾病防治具有重要意义[1]。传统统计模型在处理复杂临床特征关系时存在一定局限[2]，机器学习方法能够从多维临床数据中挖掘潜在规律，为疾病风险预测提供新的技术手段。

近年来，逻辑回归、支持向量机和集成学习等方法已广泛应用于心血管疾病预测[3]。然而，在小样本医学数据场景下，不同模型的稳定性存在差异，部分复杂模型同时面临预测结果难以解释的问题。针对上述问题，本文基于 UCI Cleveland 心脏病数据集，构建逻辑回归、支持向量机、随机森林、XGBoost 和 LightGBM 五种预测模型，通过模型比较、性能评价及 SHAP 解释分析，探讨机器学习模型在冠心病风险预测中的应用效果。

2. 相关理论基础

2.1. 机器学习分类算法

本研究选取逻辑回归(LR)、支持向量机(SVM)、随机森林(RF)、XGBoost 和 LightGBM 五种典型分类算法进行冠心病风险预测。其中，逻辑回归具有较好的可解释性，常作为医学预测任务中的基准模型；SVM 适用于复杂分类边界建模；随机森林和梯度提升树模型通过集成学习提高模型的预测能力[4]。五种模型涵盖线性模型、核方法和集成学习方法，可用于比较不同算法在冠心病风险预测中的适用性。

2.2. SHAP 可解释性方法

SHAP (Shapley Additive Explanations)基于合作博弈理论中的 Shapley 值，将模型预测结果分解为各

特征的贡献值,实现对机器学习模型的解释[5]。SHAP 值不仅能够反映单个样本中特征对预测结果的影响方向,还可以通过平均绝对 SHAP 值分析全局特征重要性,并利用依赖图展示特征取值变化与预测结果之间的关系。该方法能够增强模型预测过程的透明性,为本文冠心病风险预测模型的结果解释提供依据。

3. 数据来源与预处理

3.1. 数据来源与数据集描述

本文采用公开的 Heart Disease 数据集中的 Cleveland 子集(processed.cleveland.data)作为实验数据。该数据集来源于克利夫兰诊所冠状动脉造影检查记录[6],共包含 303 例患者样本,每个样本包括 13 项临床特征和 1 项诊断标签。特征主要涵盖人口学信息(年龄、性别)、症状与体征(胸痛类型、静息血压、最大心率、运动诱发心绞痛、ST 段下降幅度、ST 段斜率)、实验室指标(胆固醇、空腹血糖)以及辅助检查结果(静息心电图、主要血管数量、地中海贫血状态)。其中,年龄、静息血压、胆固醇、最大心率和 ST 段下降幅度为连续变量,其余特征为分类变量。原始标签取值范围为 0~4,其中 0 表示未检出冠心病,1~4 表示不同程度冠心病。为满足二分类预测任务需求,本文将标签进行二值化处理,将 0 定义为无冠心病样本,将 1~4 合并为冠心病样本,处理后正类样本 139 例,负类样本 164 例。

3.2. 数据预处理

数据质量分析发现,ca(主要血管数量)与 thal(地中海贫血)分别存在 4 例和 2 例缺失,合计占样本总量的 2.0%。鉴于样本规模有限,本文不采用直接删除策略,而是采用填补的方法。数值变量按 IQR 准则检出离群值共 20 处,其中静息血压 9 处、胆固醇 5 处、ST 段下降 5 处、最大心率 1 处。这些数值多为高血压、高血脂等患者的真实极端测量,剔除反而会损失临床信息,故予以保留,异常值的耐受交由各模型的机制实现。

特征变换严格遵循避免数据泄露的原则,全部封装在 ColumnTransformer 构建的两路 Pipeline 中:数值特征依次执行中位数填补与标准化,类别特征依次执行众数填补与 One-Hot 编码。所有预处理统计量均仅由训练集估计,测试集在模型训练完成前不参与任何参数学习。

4. 模型构建与实验设计

4.1. 实验总体流程

为验证不同机器学习方法在冠心病风险预测任务中的有效性,本文设计了完整的模型构建与评价流程。首先,对原始心脏病数据进行整理,并按照分层抽样原则划分训练集和测试集;随后,在训练集上完成缺失值处理、特征转换以及模型参数优化,并分别训练逻辑回归、支持向量机、随机森林、XGBoost 和 LightGBM 五种分类模型。模型训练完成后,利用测试集数据对各模型进行性能评估,并结合多项评价指标比较不同模型的预测效果。最后,针对综合性能较优的模型,引入 SHAP 方法分析各临床特征对预测结果的影响,以提升模型结果的可解释性。

4.2. 数据划分与交叉验证

本研究将 303 例样本按照 7:3 的比例进行分层随机划分,其中 212 例作为训练集,91 例作为测试集。模型训练和参数优化均基于训练集完成,本实验采用分层 5 折交叉验证方法提高模型评价的稳定性,并以 ROC-AUC 作为模型参数选择的主要依据。

4.3. 模型构建与超参数设置

五种分类模型均与数据预处理流程集成为统一 Pipeline，并在交叉验证过程中完成参数优化。各模型
的参数搜索范围及最终确定的参数组合如表 1 所示。参数设置主要结合不同模型的特点进行调整，其中
逻辑回归重点优化正则化参数，支持向量机调整惩罚系数和核函数参数，集成树模型则围绕模型复杂度、
学习率及采样策略等参数进行优化。各模型搜索组合数量控制在 12~24 组之间，在保证参数探索有效性的
同时，兼顾实验效率。

Table 1. Hyperparameter search ranges and optimal parameters for each model
表 1. 各模型超参数搜索范围与最优参数

模型	搜索参数(候选值)	最优参数
逻辑回归	C: 0.01/0.1/1/10	C = 0.1
SVM (RBF)	C: 0.1/1/10/100; γ : scale/0.01/0.1	C = 1.0; γ = 0.01
随机森林	n_estimators: 100/200/300; max_depth: None/5/10; min_samples_split: 2/5	n = 100; depth = None; min_split = 5
XGBoost	n_estimators: 100/200; max_depth: 3/5/7; learning_rate: 0.05/0.1; subsample: 0.8/1.0	n = 100; depth = 5; lr = 0.05; sub = 0.8
LightGBM	n_estimators: 100/200; num_leaves: 15/31/63; learning_rate: 0.05/0.1	n = 100; leaves = 15; lr = 0.1

4.4. 模型评价指标

在测试集上对各模型进行统一评价，计算准确率(Accuracy)、精确率(Precision)、召回率(Recall)、F1
值、ROC-AUC 和 PR-AUC 等指标，并绘制混淆矩阵进行结果分析。其中，Accuracy 用于衡量模型整体
分类效果，Precision 和 Recall 分别反映预测结果的可靠性以及对阳性样本的识别能力，F1 值综合考虑二
者的影响；ROC-AUC 和 PR-AUC 则用于评价模型在不同分类阈值下的整体判别性能。针对综合表现最
优的模型，进一步开展稳定性与临床应用价值分析，包括采用 Bootstrap 重采样方法(1000 次)计算 ROC-
AUC 的 95%置信区间[7]，利用校准曲线和 Brier 分数评价预测概率的一致性，并通过决策曲线分析模型
的潜在临床获益[8]。所有性能指标均基于独立测试集一次性计算，以保证评价结果的客观性。

5. 实验结果与分析

5.1. 不同模型性能比较

五种模型在测试集上的预测结果如表 2 所示，交叉验证结果如表 3 所示。由表 2 可知，各模型 ROC-
AUC 均超过 0.90，说明基于 13 项临床特征建立的模型均具有较好的冠心病风险识别能力。其中，逻辑
回归综合性能最佳，ROC-AUC 和 PR-AUC 分别达到 0.9436 和 0.9429，准确率为 0.8681，F1 值为 0.8571；
支持向量机性能与其接近，在准确率(0.8791)和 F1 值(0.8675)方面略高于逻辑回归。随机森林、XGBoost
和 LightGBM 的 ROC-AUC 分别为 0.9339、0.9096 和 0.9052。

对比表 2 和表 3 可以发现，交叉验证结果与独立测试集结果存在一定差异。例如，随机森林在交叉
验证中的平均 ROC-AUC 高于逻辑回归，但测试集表现略低。考虑到 UCI Cleveland 数据集样本规模较
小，复杂模型在不同数据划分下可能存在一定波动，而逻辑回归凭借结构简单和正则化约束表现出较好
的稳定性。因此，在小样本医学预测任务中，模型性能不仅取决于算法复杂度，还需要结合实际验证结

果进行选择。

逻辑回归模型的混淆矩阵如图 1 所示，测试集中正确分类样本共 79 例，其中真阴性 43 例、真阳性 36 例，假阳性和假阴性均为 6 例，对应敏感度 85.7%、特异度 87.8%。模型在两类样本上的识别能力较为均衡，具有较低的分类误差，可作为后续可解释性分析的基础模型。

Table 2. Performance comparison of different models on the test set
表 2. 不同模型在测试集上的性能比较

模型	Accuracy	Precision	Recall	F1	ROC-AUC	PR-AUC
逻辑回归	0.8891	0.8871	0.8571	0.8721	0.9052	0.9035
SVM (RBF)	0.8791	0.8780	0.8571	0.8675	0.9427	0.9425
随机森林	0.8242	0.7826	0.8571	0.8182	0.9339	0.9313
XGBoost	0.8242	0.8250	0.7857	0.8049	0.9096	0.9077
LightGBM	0.7912	0.7805	0.7619	0.7711	0.9052	0.9035

Table 3. 5-fold cross-validation ROC-AUC for the five models
表 3. 五种模型 5 折交叉验证 ROC-AUC

模型	均值	标准差
随机森林	0.9067	0.0232
逻辑回归	0.9030	0.0301
SVM (RBF)	0.8968	0.0276
XGBoost	0.8930	0.0369
LightGBM	0.8885	0.0377

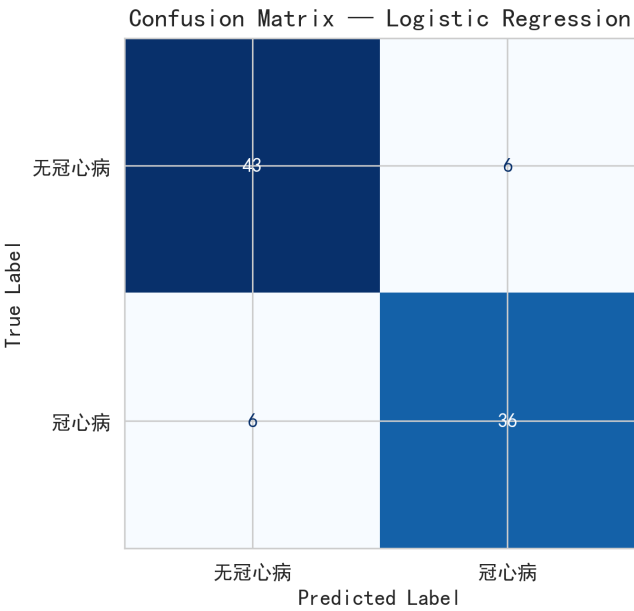


Figure 1. Confusion matrix of the optimal model (Logistic Regression)
图 1. 最优模型(逻辑回归)混淆矩阵

5.2. ROC-AUC 与模型可靠性分析

ROC-AUC 结果表明, 逻辑回归模型在测试集上取得最佳判别性能($AUC = 0.9436$), 支持向量机表现接近($AUC = 0.9427$), 而随机森林、XGBoost 和 LightGBM 略低。该结果说明, 基于 13 项临床特征建立的模型均具有较好的风险识别能力。与复杂集成模型相比, 逻辑回归在小样本条件下表现出更好的稳定性, 可能与其模型结构简单、正则化约束较强有关。

Bootstrap 重采样分析进一步验证了模型性能的可靠性。通过多次随机重采样计算 ROC-AUC 置信区间, 可以降低单次数据划分带来的偶然影响, 结果表明逻辑回归模型具有较稳定的预测能力。

在模型校准和临床应用价值方面, 校准曲线显示预测风险与实际结果具有较好一致性, Brier 分数进一步反映模型概率预测的可靠程度; 决策曲线分析表明, 该模型在一定风险阈值范围内能够获得净收益, 具有辅助冠心病风险评估的应用潜力(图 2~4)。

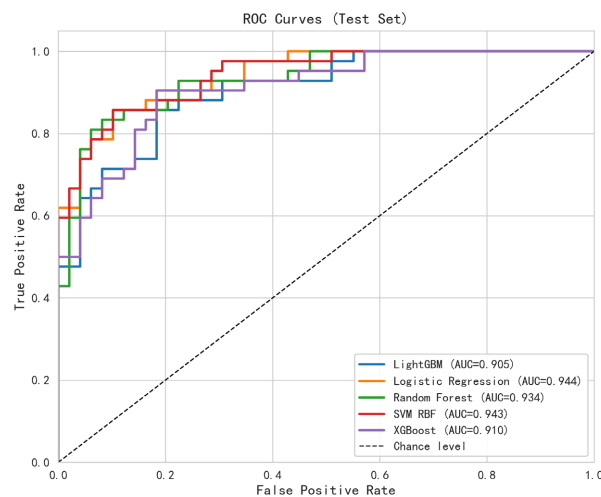


Figure 2. ROC curves of the five models
图 2. 五种模型 ROC 曲线

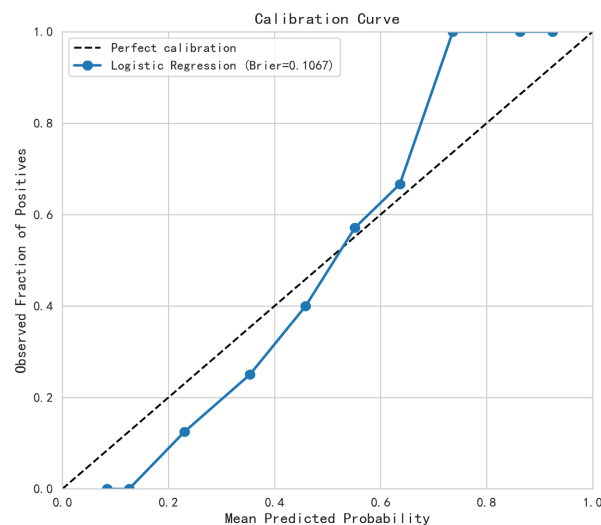


Figure 3. Calibration curve of the optimal model
图 3. 最优模型校准曲线

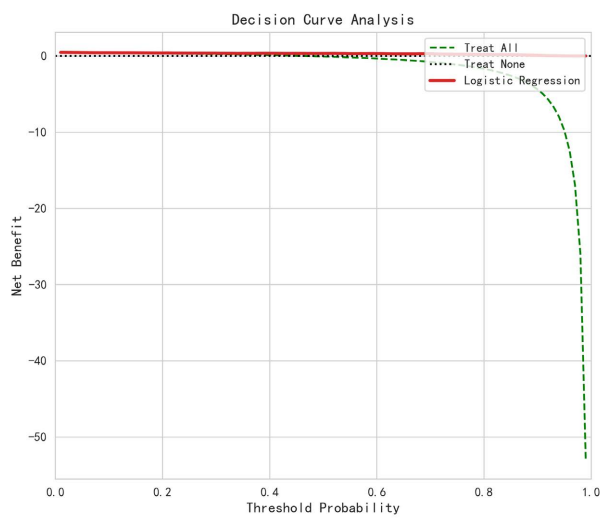


Figure 4. Decision curve analysis (DCA) of the optimal model
图 4. 最优模型决策曲线(DCA)

5.3. 特征重要性分析

基于平均绝对 SHAP 值计算得到的特征重要性排序如表 4 和图 5 所示。结果显示，影响模型预测的主要特征依次为胸痛类型(0.4523)、地中海贫血(0.4226)、主要血管数量(0.4078)、性别(0.3134)和最大心率(0.2770)，前五项特征贡献占全部特征贡献的约 64.7%。相比之下，静息血压、胆固醇和空腹血糖等指标的 SHAP 贡献相对较低。

从特征含义来看，胸痛类型能够反映患者心肌缺血相关症状，主要血管数量与冠状动脉病变程度密切相关，最大心率则反映运动状态下的心功能储备。需要指出的是，模型中的特征重要性反映的是变量对预测结果的贡献大小，并不等同于医学上的因果关系。部分传统危险因素虽然与冠心病发生密切相关，但在多特征联合建模过程中，其提供的额外判别信息可能被其他临床指标部分替代，因此在 SHAP 排序中未表现出较高权重。

Table 4. Feature importance ranking based on SHAP values
表 4. 基于 SHAP 值的特征重要性排序

排名	特征	平均 SHAP	排名	特征	平均 SHAP
1	cp (胸痛类型)	0.4523	8	slope (ST 段斜率)	0.1788
2	thal (地中海贫血)	0.4226	9	restecg (静息心电图)	0.1352
3	ca (主要血管数量)	0.4078	10	age (年龄)	0.1129
4	sex (性别)	0.3134	11	trestbps (静息血压)	0.0847
5	thalach (最大心率)	0.2770	12	chol (胆固醇)	0.0512
6	exang (运动心绞痛)	0.2137	13	fbs (空腹血糖)	0.0397
7	oldpeak (ST 段下降)	0.2043			

5.4. SHAP 可解释性分析

图 6 展示了 SHAP Summary 图中各特征对模型预测结果的影响方向。结果显示，胸痛类型、主要血管数量、地中海贫血、最大心率以及性别等特征对模型输出具有较明显影响。其中，胸痛类型取值升高时，模型风险预测贡献呈增加趋势；主要血管数量增加以及 thal 取值较高时，对预测风险具有正向影响；最大

心率表现为负向贡献，即较低的最大心率对应更高的风险预测值；男性样本(sex = 1)整体呈现正向贡献。

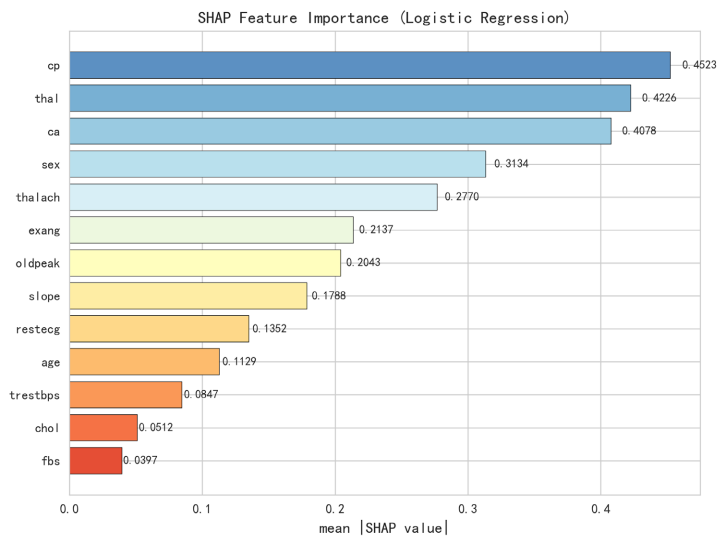


Figure 5. SHAP feature importance ranking
图 5. SHAP 特征重要性排序

从解释结果来看，模型关注的主要特征与冠心病相关临床指标具有较好的一致性。例如，胸痛类型反映患者缺血相关症状，主要血管数量与冠状动脉病变程度相关，最大心率能够体现运动状态下的心功能水平。需要注意的是，SHAP 分析反映的是特征对模型预测结果的贡献，而非变量之间的因果关系。因此，该结果主要用于增强模型预测过程的可解释性，为理解模型决策依据提供参考。

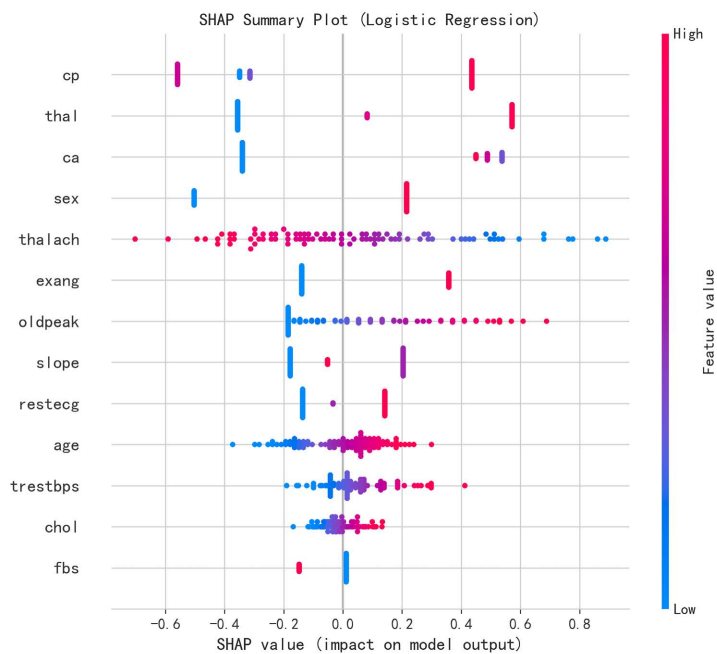


Figure 6. SHAP summary plot for the optimal model
图 6. SHAP summary 图(最优模型)

SHAP 依赖图进一步展示了关键特征取值变化对模型预测结果的影响趋势。如图 7 所示,胸痛类型随等级升高,其 SHAP 值整体呈增加趋势,表明典型胸痛特征会提高模型对冠心病风险的预测倾向,这与逻辑回归模型的线性特征表达方式相一致。图 8 展示了最大心率对预测结果的影响,随着最大心率升高,SHAP 值逐渐降低,说明较低最大心率样本在模型中具有更高的风险贡献。

依赖图分析不仅反映了特征的重要程度,也揭示了不同特征取值对模型输出的影响方向,使模型预测结果更加直观。需要注意的是,由于本文最优模型为逻辑回归,其特征贡献主要表现为线性叠加关系;若采用树集成模型,特征之间可能存在更复杂的非线性关系和交互作用,这将在后续研究中进一步探索。

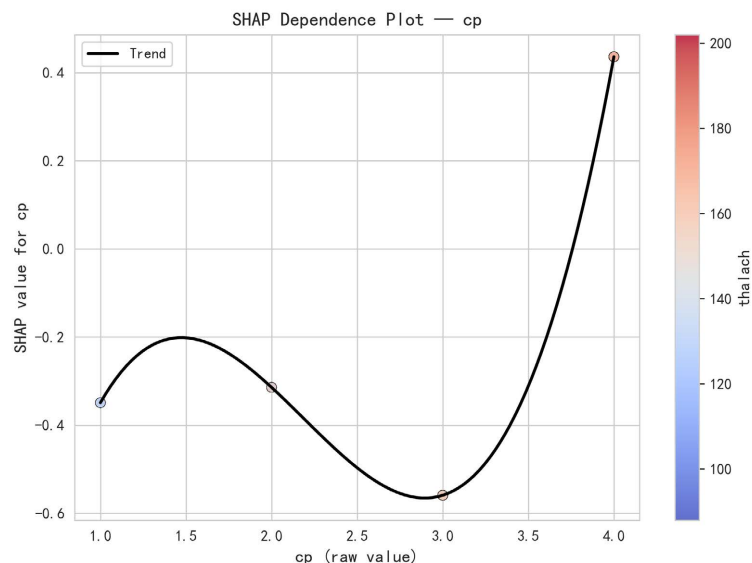


Figure 7. SHAP dependence plot for chest pain type (cp)
图 7. 胸痛类型(cp) SHAP 依赖图

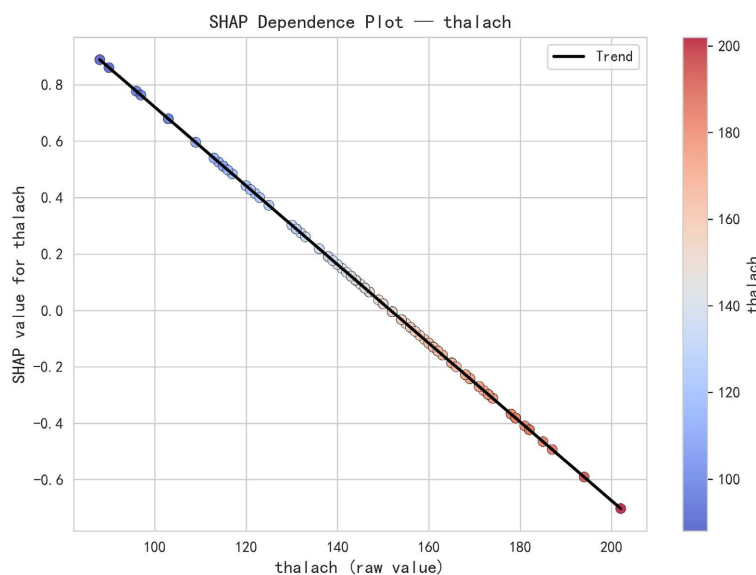


Figure 8. SHAP dependence plot for maximum heart rate (thalach)
图 8. 最大心率(thalach) SHAP 依赖图

6. 结论

本文基于 UCI Cleveland 心脏病数据集，构建了逻辑回归、支持向量机、随机森林、XGBoost 和 LightGBM 五种机器学习模型，并结合性能评价与 SHAP 方法分析模型预测结果及关键影响因素。实验结果表明，五种模型均具备较好的冠心病风险识别能力，其中逻辑回归模型综合性能最佳。可靠性分析和 SHAP 解释结果显示，该模型具有较好的稳定性，年龄、最大心率、运动诱发心绞痛、ST 段压低等特征对预测结果具有重要影响。研究表明，在小样本医学预测场景下，结合机器学习与可解释分析能够实现有效风险评估，并提升模型结果的透明性。由于本文数据来源于公开小样本数据集，模型泛化能力仍需通过更大规模、多中心临床数据进一步验证。未来可结合更多临床指标和外部验证方法，对模型进行优化。

基金项目

2025 年保定市科技计划项目；项目名称：医疗大数据背景下人工智能在疾病预测中的应用研究；项目编号：2511ZG018。

参考文献

- [1] 国家心血管病中心. 中国心血管健康与疾病报告 2022 [M]. 北京: 中国协和医科大学出版社, 2023.
- [2] D'Agostino, R.B., Vasan, R.S., Pencina, M.J., Wolf, P.A., Cobain, M., Massaro, J.M., *et al.* (2008) General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study. *Circulation*, **117**, 743-753. <https://doi.org/10.1161/circulationaha.107.699579>
- [3] Genders, T.S.S., Steyerberg, E.W., Hunink, M.G.M., *et al.* (2012) Prediction Model to Estimate Presence of Coronary Artery Disease: Retrospective Pooled Analysis of Existing Cohorts. *British Medical Journal*, **344**, e3485.
- [4] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. <https://doi.org/10.1023/a:1010933404324>
- [5] Lundberg, S.M. and Lee, S.I. (2017) A Unified Approach to Interpreting Model Predictions. 2017 *Advances in Neural Information Processing Systems* 30, Long Beach, 4-9 December 2017, 4765-4774.
- [6] Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., *et al.* (2020) From Local Explanations to Global Understanding with Explainable AI for Trees. *Nature Machine Intelligence*, **2**, 56-67. <https://doi.org/10.1038/s42256-019-0138-9>
- [7] Steyerberg, E.W., Vickers, A.J., Cook, N.R., Gerds, T., Gonen, M., Obuchowski, N., *et al.* (2010) Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measure. *Epidemiology*, **21**, 128-138. <https://doi.org/10.1097/ede.0b013e3181c30fb2>
- [8] Vickers, A.J. and Elkin, E.B. (2006) Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Medical Decision Making*, **26**, 565-574. <https://doi.org/10.1177/0272989x06295361>