

• 论 著 •

基于检验数据的机器学习建立 2 型糖尿病患者合并冠心病的风险预测模型

顾智超, 吴昀喆, 杨帆, 陆怡德[△]

上海交通大学医学院附属瑞金医院检验科, 上海 200025

摘要:目的 分析 2 型糖尿病患者的临床指标特征, 利用机器学习筛选风险预测指标, 建立简便、有效的 2 型糖尿病合并冠心病的风险预测模型。方法 采用回顾性研究的方法, 选取 2022 年 1 月至 2023 年 11 月该院心内科住院且诊断为 2 型糖尿病合并冠心病的患者 217 例作为合并冠心病组, 同期在门诊确诊为 2 型糖尿病患者 214 例作为糖尿病组, 记录两组患者实验室常规检测数据。通过最小绝对收缩和选择算法(Lasso)筛选特征, 运用随机森林、决策树、支持向量机、极端梯度提升、Logistic 回归、K 近邻算法、人工神经网络 7 种机器学习算法建立模型, 通过受试者工作特征曲线及曲线下面积(AUC)、校准曲线、特异度、灵敏度、F1 值等指标评价不同模型的诊断效能。**结果** 通过 Lasso 回归共筛选出年龄、性别、收缩压、舒张压、心率、C 反应蛋白、血糖等 20 个关键因素。纳入模型后, 支持向量机模型具有最高的灵敏度(88.37%)、阴性预测值(82.14%)及 AUC(0.845), 随机森林模型具有最高的准确度(76.47%)、阳性预测值(76.74%)、F1 值(0.77), 而极端梯度提升算法具有较好的特异度(80.95%), 引入支持向量机模型及 SHAP 值后, 归纳得出血糖对于糖尿病合并冠心病具有正向影响。**结论** 机器学习可作为 2 型糖尿病患者合并冠心病风险评估的有效工具, 其中支持向量机、随机森林和极端梯度提升模型均有较好的预测效能, 具有一定的临床应用前景。

关键词:机器学习; 2 型糖尿病; 冠心病; 预测模型

DOI:10.3969/j.issn.1673-4130.2025.02.002

中图法分类号:R446.1;R587.2

文章编号:1673-4130(2025)02-0135-06

文献标志码:A

Establishment of a risk prediction model for patients with type 2 diabetes and coronary heart disease based on machine learning of laboratory data

GU Zhichao, WU Yunzhe, YANG Fan, LU Yide[△]

Department of Laboratory Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

Abstract: Objective To analyze the characteristics of clinical indicators in patients with type 2 diabetes, and to establish a simple and effective risk prediction model for type 2 diabetes complicated with coronary heart disease by screening risk prediction indicators with machine learning. **Methods** A retrospective study was conducted, and 217 patients diagnosed with coronary artery disease combined with type 2 diabetes mellitus who were hospitalized in the Hospital from January 2022 to November 2023 were selected. Additionally, 214 patients diagnosed with T2DM during the same period in the outpatient department were selected as the control group. Their routine laboratory test data were recorded. The Least Absolute Shrinkage and Selection Operator (Lasso) algorithm was used to select features, and the models were built by using seven machine learning algorithms: Random Forest, Decision Tree, Support Vector Machine, eXtreme Gradient Boosting, Logistic Regression, K-Nearest Neighbor, and Artificial Neural Network. The diagnostic efficacy of different models through receiver operating characteristic curve (ROC), area under curve (AUC), calibration curve, specificity, sensitivity, F1 value, and other indicators were evaluated. **Results** Twenty key factors, including age, gender, systolic blood pressure, diastolic blood pressure, heart rate, C-reactive protein and blood glucose were selected using Lasso regression. When incorporated into various models, the SVM model exhibited the highest sensitivity (88.37%), negative predictive value (82.14%), and area under curve (0.845). The Random Forest model had the highest accuracy (76.47%), positive predictive value (76.74%), and F1 score (0.77). Meanwhile, the XGBoost algorithm demonstrated relatively good specificity (80.95%). After introducing the SHAP model, it was inferred that blood glucose had a significant positive impact on the occurrence of coronary heart disease in

individuals with type 2 diabetes. **Conclusion** Machine learning can serve as an effective tool for assessing the risk of coronary heart disease in patients with type 2 diabetes. In this study, SVM, Random Forest, and XGBoost models all demonstrate good predictive performance, indicating promising clinical application prospects.

Key words: machine learning; type 2 diabetes; coronary heart disease; predictive model

糖尿病在世界范围内影响着约 9.5% 的人口, 大约有 4.5 亿人患有糖尿病^[1]。我国糖尿病患病率呈不断增长趋势, 2011—2021 年我国糖尿病患者人数已从 9 000 万增加至 1.4 亿, 其中 2 型糖尿病患者占总病例数的 90%~95%, 其通常伴随着胰岛素相对不足, 即胰岛素抵抗。糖尿病的主要并发症是包括视网膜病变、神经病变和肾病在内的微血管疾病, 以及冠心病、外周血管疾病和脑血管疾病在内的大血管疾病^[2]。对于 2 型糖尿病患者来说, 动脉粥样硬化性心血管疾病和心力衰竭是患病和死亡的主要危险因素, 动脉粥样硬化性心血管疾病的患病率在糖尿病患者中是非糖尿病患者的 2 倍以上, 因此寻找高效、快速、准确的诊疗工具对 2 型糖尿病患者合并冠心病进行风险评估是极其必要的。人工智能技术当下已渐渐被应用于疾病预测及严重程度分层, 其中机器学习可以检查许多数据并发掘其相关性, 提供解释, 并识别模式, 从而帮助提高许多疾病诊断系统的可靠性、性能、可预测性和准确性^[3]。本文尝试利用 2 型糖尿病患者基本信息和实验室常规检验数据, 探讨机器学习在 2 型糖尿病患者合并冠心病风险评估中的价值, 旨在寻找合适的诊断模型预警 2 型糖尿病合并冠心病, 以便更早期地进行疾病预防, 减少心血管不良事件的发生。

1 资料与方法

1.1 一般资料 选取 2022 年 1 月至 2023 年 11 月在本院心内科住院患者, 根据 2019 年欧洲心脏病学会 (ESC) 颁布的慢性冠状动脉综合征诊断和管理指南^[4] 和《中国 2 型糖尿病防治指南 (2020 年版)》^[5], 将诊断为冠心病合并 2 型糖尿病的患者 217 例作为合并冠心病组, 其中男 138 例, 女 79 例, 住院期间行冠状动脉造影手术且造影记录保存完整。排除标准: (1) 既往已确诊为冠心病; (2) 糖尿病急性并发症; (3) 合并恶性肿瘤; (4) 既往心脏手术史; (5) 严重心血管疾病; (6) 严重肝、肾功能损害; (7) 严重感染性疾病; (8) 孕妇或哺乳期女性。另选取同期在本院门诊确诊的单纯 2 型糖尿病患者 214 例作为糖尿病组, 其中男 131 例, 女 83 例。本研究通过本院伦理委员会批准, 并获得所有患者及家属的知情同意。

1.2 资料收集 记录所有患者性别、年龄、心率 (HR)、收缩压 (SBP)、舒张压 (DBP), 以及血液常规检测项目, 包括葡萄糖 (GLU)、丙氨酸氨基转移酶 (ALT)、天门冬氨酸氨基转移酶 (AST)、尿素氮 (UREA)、肌酐 (Cr)、尿酸 (UA)、胱抑素 C (Cys-C)、甘油三酯 (TG)、总胆固醇 (TC)、低密度脂蛋白胆固醇

(LDL-C)、高密度脂蛋白胆固醇 (HDL-C)、载脂蛋白 B (ApoB)、载脂蛋白 A I (ApoA I)、乳酸脱氢酶 (LDH)、C 反应蛋白 (CRP) 等。检测仪器均为贝克曼-库尔特 AU5821 全自动生化分析仪及其配套试剂。载脂蛋白 E (ApoE) 检测试剂购自宁波瑞源生物公司、脂蛋白 (a) [Lp(a)] 检测试剂购自亚辉龙生物公司, 小而密低密度脂蛋白胆固醇 (sdLDL-C) 检测试剂购自北京九强生物公司, 游离脂肪酸 (FFA) 检测试剂购自德赛诊断公司; 胰岛素 (Ins) 和 C 肽 (C-P) 检测仪器为罗氏 Cobas e602 电化学发光免疫分析仪及其配套试剂; 糖化血红蛋白 (GHb) 检测仪器为伯乐糖化血红蛋白仪 D-100, 糖化血清白蛋白 (GSP) 检测仪器为罗氏 Cobas c701 全自动生化分析仪。应用迈瑞 BC-7500 全自动血液细胞分析仪检测其白细胞计数 (WBC)、血红蛋白 (Hb) 和血小板计数 (PLT)。

1.3 统计学处理 采用 R4.3.2 进行统计学分析。计量资料的正态性采用 K-S 检验, 方差齐性采用 Levene 检验, 呈正态分布的数据以 $\bar{x} \pm s$ 表示, 组间比较采用两独立样本 *t* 检验; 呈非正态分布的数据以 $M(P_{25}, P_{75})$ 表示, 组间比较采用 Mann-Whitney *U* 检验。计数资料以例数或百分率表示, 组间比较采用 χ^2 检验。以 $P < 0.05$ 为差异有统计学意义。

1.4 机器学习模型构建与评价

1.4.1 特征选择 应用最小绝对收缩和选择算法 (Lasso) 进行回归分析, 通过计算选择变量, 在回归系数绝对值 ≤ 1 个常数的前提条件下, 获得参数估计值。通过 10 折交叉验证筛选出预测变量中最重要的特征, 识别最具影响力的预测指标, 提高模型的预测性能和解释能力。

1.4.2 模型构建 采用 7 种机器学习的算法预测 2 型糖尿病患者合并冠心病的风险, 包括随机森林 (RF)、决策树 (DT)、支持向量机 (SVM)、极端梯度提升 (XGBoost)、Logistic 回归 (LR)、K 近邻算法 (KNN) 和人工神经网络 (ANN)。所有研究数据随机按照 4 : 1 比例分为模型的训练集和测试集, 并采用 10 折交叉验证方法评估模型的性能, 以最大曲线下面积 (AUC) 和 F1 值判断模型为最优模型。

1.4.3 模型性能评价 采用计算各个模型受试者工作特征 (ROC) 曲线及 AUC、准确度、灵敏度、特异度、阳性预测值 (PPV)、阴性预测值 (NPV)、F1 值, 绘制标准曲线等对比评估不同机器学习模型的预测效能, 其中, 准确度表示所有标本中被预测正确的样本比例; AUC 揭示了灵敏度与特异度的相互关系, AUC 越大, 诊断准确性越高; F1 值是模型精度和召回率的

调和平均数,是衡量二分类模型精确度的重要指标。

2 结果

2.1 两组一般资料比较 两组年龄、SBP、DBP、

GLU、CRP、TC、Cr、Cys-C、LDL-C、ApoB、FFA、LDH、C-P、Inc 和 GHb 水平比较,差异有统计学意义 ($P < 0.05$)。见表 1。

表 1 合并冠心病组与糖尿病组基本特征比较 [$n(\%)$ 或 $M(P_{25}, P_{75})$ 或 $\bar{x} \pm s$]

项目	总体($n=431$)	糖尿病组($n=214$)	合并冠心病组($n=217$)	P
性别				0.681
男	269(62.4)	131(61.2)	138(63.6)	
女	162(37.6)	83(38.8)	79(36.4)	
年龄(岁)	63(54,71)	58(51,66)	68(59,74)	<0.001
心率(次/分)	80(77,90)	81(77,90)	80(77,91)	0.943
SBP(mmHg)	140.1±20.1	135.2±19.4	144.9±19.6	<0.001
DBP(mmHg)	79(72,88)	80(75,89)	78(70,85)	<0.001
WBC($\times 10^9/L$)	6.37(5.39,7.74)	6.27(5.28,7.61)	6.51(5.57,7.85)	0.192
Hb(g/L)	138(126,148)	138(126,148)	136(127,147)	0.514
PLT($\times 10^9/L$)	195(165,243)	203(164,248)	189(167,233)	0.143
CRP(mg/L)	1.74(0.81,3.77)	1.29(0.71,3.12)	2.02(1.14,4.71)	0.003
GLU(mmol/L)	8.24(7.03,9.71)	7.49(6.59,8.98)	8.74(7.86,10.58)	<0.001
ALT(U/L)	21(14,29)	20(13,33)	21(15,27)	0.823
AST(U/L)	20(17,26)	20(16,25)	20(17,26)	0.280
UREA(mmol/L)	5.9(5.0,7.5)	5.9(4.9,7.4)	6.0(5.2,7.5)	0.182
Cr($\mu\text{mol/L}$)	72(62,87)	71(60,85)	75(64,91)	0.015
UA($\mu\text{mol/L}$)	323.3±84.1	321.1±89.5	325.4±78.6	0.594
Cys-C(mg/L)	1.06(0.92,1.28)	1.04(0.91,1.21)	1.10(0.94,1.36)	0.020
TG(mmol/L)	1.51(1.12,2.38)	1.46(1.07,2.35)	1.58(1.15,2.42)	0.158
TC(mmol/L)	4.44(3.71,5.26)	4.56(3.98,5.28)	4.24(3.43,5.21)	<0.001
LDL-C(mmol/L)	2.55(1.89,3.20)	2.64(2.11,3.22)	2.35(1.81,3.10)	0.003
HDL-C(mmol/L)	1.05(0.88,1.24)	1.07(0.89,1.30)	1.04(0.86,1.18)	0.057
ApoA I (g/L)	1.23(1.10,1.37)	1.22(1.12,1.39)	1.24(1.08,1.36)	0.608
ApoB(g/L)	0.95(0.75,1.13)	1.01(0.84,1.15)	0.90(0.71,1.11)	<0.001
Lp(a)(g/L)	0.11(0.05,0.32)	0.10(0.04,0.33)	0.12(0.05,0.32)	0.266
ApoE(mg/dL)	4.0(3.5,4.5)	4.0(3.6,4.5)	3.9(3.5,4.4)	0.107
sdLDL-C(mmol/L)	0.86(0.60,1.23)	0.90(0.64,1.21)	0.83(0.57,1.30)	0.295
FFA(mmol/L)	0.46(0.35,0.59)	0.44(0.34,0.57)	0.50(0.36,0.63)	0.048
LDH(U/L)	176(158,204)	180(163,203)	172(152,204)	0.010
GSP(%)	20.66(17.71,24.59)	20.84(17.43,25.11)	20.57(17.92,24.03)	0.609
C-P($\mu\text{g/L}$)	2.42(1.60,3.47)	2.14(1.34,3.10)	2.71(2.01,3.66)	<0.001
Ins($\mu\text{IU/mL}$)	8.65(5.20,14.30)	7.46(4.01,13.47)	9.57(6.18,14.56)	0.002
GHb(%)	7.9(7.2,9.0)	8(7.3,9.1)	7.9(7.0,9.0)	0.049

2.2 特征选择 利用 Lasso 回归分析了年龄、心率、性别、血压和 26 项检测结果对冠心病的影响。逐步增加正则化参数 λ (图 1),筛选出 20 个最重要的变量(图 2),分别为年龄、性别、SBP、DBP、心率、CRP、GLU、AST、Cr、TG、TC、HDL-C、ApoB、ApoE、sdLDL-C、FFA、C-P、Ins、GHb 和 GSP。单个变量预

测糖尿病合并冠心病的 ROC 曲线见图 3。

2.3 模型构建 随机森林构建了决策树节点处变量个数为 4,决策树数量为 400,并利用多数投票的方式对测试集样本做出预测;XGBoost 迭代次数为 500,学习率为 0.01,树的最大深度为 12,采样比例为 0.8;支持向量机参数 kernel = “radial”, cost = “0.1”,

gamma=0.1, class_weights=c(0.77, 0.23); 决策树从树的顶端开始, 依次按照条件成立与否往下进行, 终点指标为“是/否合并冠心病”的二分类数据; Logistic 回归模型参数为 family=binomial; KNN 中的模型参数为 K=5; ANN 中的模型参数为 hidden=c(3), linear.output=FALSE。

细的性能指标见表 2, 其中 SVM 模型的灵敏度和 NPV 最高, 显示出较高的预测灵敏度, 随机森林模型的准确度、PPV 和 F1 值最高, 显示出较高的准确性, 而 XGBoost 模型则具有较好的特异度。同时, 7 种预测模型的校准曲线中, SVM 模型、XGBoost 模型及随机森林模型都有着较好的校准度。

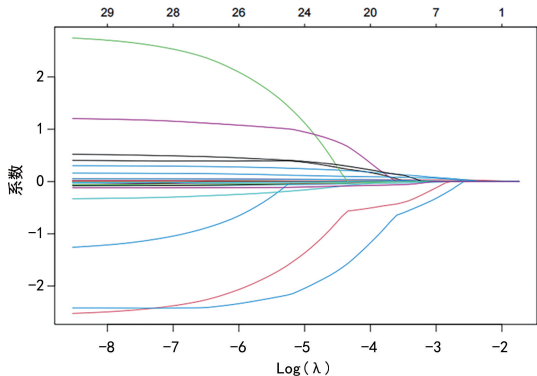


图 1 log(λ) 值与 Lasso 回归系数

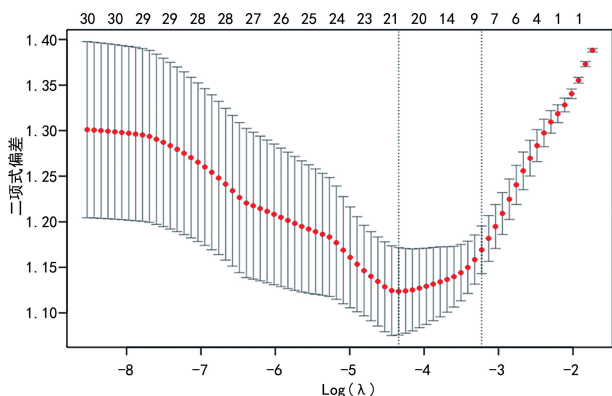


图 2 Lasso 回归算法函数筛选的关键变量

2.4 不同模型的效能比较 比较 7 个模型对于糖尿病患者合并冠心病的预测效能(图 4), 其中有 4 个预测模型的 ROC 曲线 AUC 大于 0.75, 表明具有较好的预测性能, 由高到低分别为 SVM(AUC 为 0.845)、随机森林(AUC 为 0.833)、XGBoost(AUC 为 0.802)、Logistic 回归(AUC 为 0.769)。7 个模型详

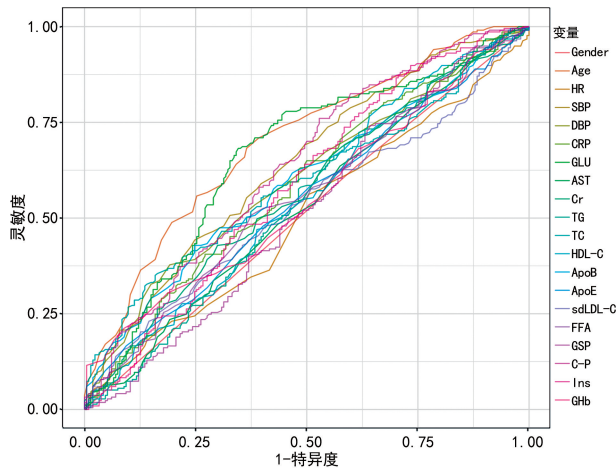
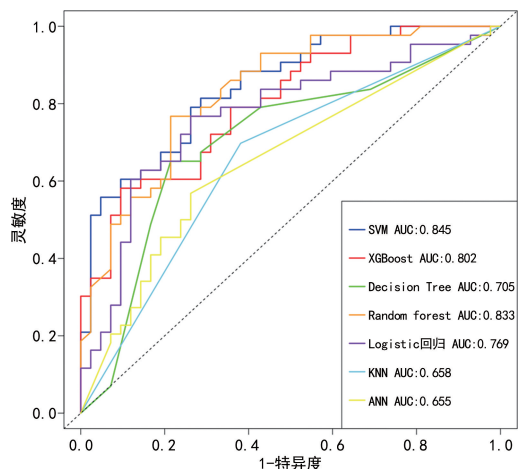


图 3 单个变量预测糖尿病合并冠心病的 ROC 曲线



注: SVM 为支持向量机, XGBoost 为极端梯度提升, Decision Tree 为决策树, Random Forest 为随机森林, KNN 为 K 近邻算法, ANN 为人工神经网络。

图 4 7 种模型预测效能的 ROC 曲线

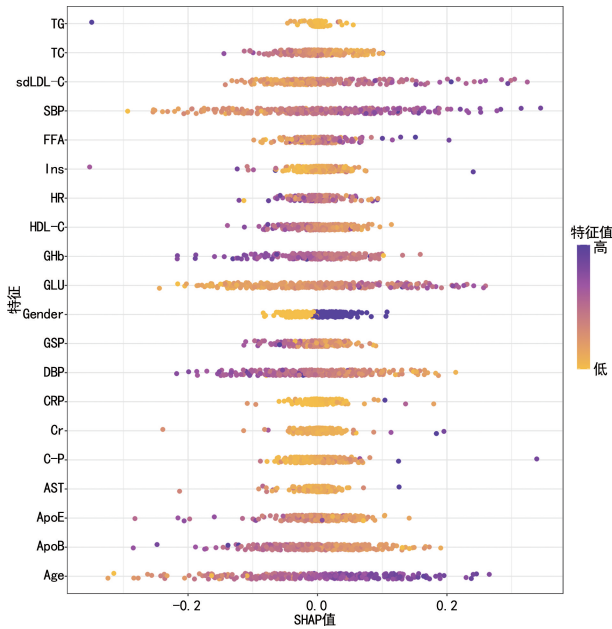
表 2 7 种模型预测效能评价

项目	准确度(%)	灵敏度(%)	特异度(%)	PPV(%)	NPV(%)	F1 值
Logistic 回归	75.29	76.74	73.81	75.00	75.61	0.76
极端梯度提升	70.59	60.47	80.95	76.47	66.67	0.68
支持向量机	71.76	88.37	54.76	66.67	82.14	0.76
决策树	68.24	65.12	71.43	70.00	66.67	0.67
随机森林	76.47	76.74	76.19	76.74	76.19	0.77
K 近邻算法	65.88	69.77	61.90	65.22	66.67	0.67
人工神经网络	65.12	56.82	73.81	69.44	62.00	0.63

2.5 不同模型的可解释化模型 采用 SHAP 值对 SVM 模型中的特征性因素进行重要性分析见图 5, 纵

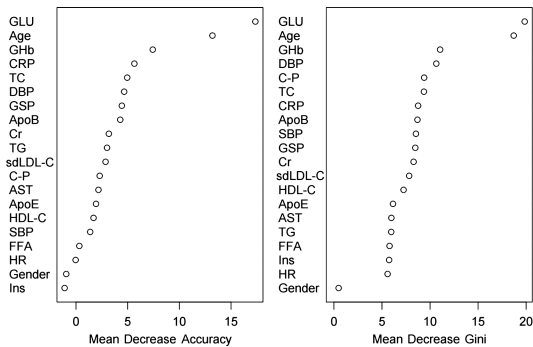
轴为各个因变量, 横轴为 SHAP 值, 正值代表对模型预测具有正向贡献。可见血糖、年龄、性别对 SVM 模

型的预测结果影响最大,表明随着血糖值的增高,预测合并冠心病的概率亦增大;将各项参数纳入随机森林模型,其重要性评分见图 6,血糖、年龄和 GHb 的重要性排名靠前。



注: Age 表示年龄; Gender 表示性别。

图 5 SVM 模型重要特征的 SHAP 值分布图



注: Mean Decrease Accuracy 表示随机森林模型预测准确性的降低程度,即值越大越重要; Mean Decrease Gini 表示通过基尼指数计算每个变量对分类树每个节点观测值的异质性影响; Age 表示年龄; Gender 表示性别。

图 6 随机森林模型变量重要性评分

3 讨论

与传统的统计学方法相比,机器学习更多的是依据实际数据特征建立模型,并在建模过程中自动学习改进。目前,机器学习在医疗领域的应用越来越广泛,主要集中在疾病的预测、辅助诊断、预后评估、健康管理以及医学图像识别等^[6-7]。刘巧红等^[8]提出基于 XGBoost 的糖尿病风险预测及特征模型; HE 等^[9]开发出针对糖尿病微血管并发症的机器学习模型; CHEN 等^[10]则通过使用 Lasso 回归和 Boruta 特征选择进行变量筛选,最终开发出以神经网络模型评估 2 型糖尿病患者罹患高尿酸血症风险的预测模型。传统的 2 型糖尿病患者合并冠心病的确诊依赖于侵入

性检查,如冠状动脉造影。这些检查方法不仅成本高,而且存在一定的风险,本研究旨在通过机器学习算法,利用常规实验室检测指标建立一种经济又高效的 2 型糖尿病患者合并冠心病的风险预测模型。

本研究利用 Lasso 算法对数据进行压缩筛选,筛选出的 20 个关键因素与传统独立样本 *t* 检验相比,得到了相近的结果,表明 Lasso 算法能够有效去除无关、弱相关且冗余的特征,提取到最优特征集^[11]。将经过筛选的关键变量输入不同的机器学习方法进行模型预测,不同的机器学习模型对于糖尿病并发症风险的预测结果不尽相同。本次研究中 SVM 模型对预测糖尿病合并冠心病的 AUC 最大,显示出了最好的诊断效能。SVM 算法采用结构风险最小化原则进行模型选择,能够有效地避免过拟合现象的发生,使得 SVM 算法具有较强的泛化能力,使得此模型具有高灵敏度的特点。因此,使用 SVM 算法能够在众多的 2 型糖尿病患者中有效识别高冠心病风险人群,提高诊断效率。除此之外,本研究对样本特征进行了筛选,质量较高,虽样本量较小,但也证明了二分类问题的确更有利于 SVM 的矩阵计算^[12];其余 6 个模型中,随机森林模型对于 2 型糖尿病合并冠心病也体现出了较高的分类效果和识别能力。随机森林模型是以建立 N 个决策树为基分类器进行集成学习后得到一个组合分类器^[13],通过大量随机选择样本的方法平衡了样本误差的影响,因此随机森林往往有较高的单个决策树的准确度,也不易出现过拟合现象^[12]。相较于决策树、KNN、ANN 等单一模型由于不确定因素的增加和特征维度的增加导致预测精度的降低,随机森林模型的结果更为可靠。XGBoost 是计算机并行化领域中的最优算法之一,有多篇文献显示其广泛应用于疾病预测、死亡风险评估等,具有较好的识别能力和拟合优度^[13-16],在此次研究中也同样显示出了较高的特异度。若联合应用高灵敏度的 SVM 模型和高特异度的 XGBoost 模型,更准确地从 2 型糖尿病患者中筛选出高冠心病风险的人群、排除低风险人群,为提高诊断效能提供了新的方法和思路。

随机森林模型作为单个学习器具有较好的可解释性,通过重要性分析中识别出血糖、年龄、GHb 是影响病情发展的最重要因素。而 SVM 模型可以通过 LUNDBERG 等^[17]提出的 SHAP 值进行解释,也得到了与随机森林模型一致的结果。2 个模型的分析结果都证明了积极进行科学有效的血糖干预,保证血糖的长期稳定,对于 2 型糖尿病患者降低冠心病的合并发生具有积极的意义^[18]。近 30 年来,我国心血管疾病谱发生重要变化,冠心病负担显著增加^[19],除了控制血脂以外,监测血糖亦可作为临床实验室早期预防、早期干预的重点关注内容。

机器学习将检验大数据结合,纳入更多的预测因子并识别其重要性,探索了综合不同检测项目数据的实用性,提供了改善传统风险因素之外的检测思路,并最终为防止疾病发生、发展的临床决策提供有效信息^[20]。本文仍有不足之处,研究中采用的样本量共 431 例,略显不足,后续会增加样本量或采用其他公共数据库信息,提升机器学习模型的预测效能。

参考文献

- [1] SAEEDI P, PETERSOHN I, SALPEA P, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045; results from the International Diabetes Federation Diabetes Atlas, 9 (th) edition [J]. *Diabetes Res Clin Pr*, 2019, 157: 107843.
- [2] KALYANI R R. Glucose-lowering drugs to reduce cardiovascular risk in type 2 diabetes [J]. *New Engl J Med*, 2021, 384(13): 1248-1260.
- [3] OBERMEYER Z, EMANUEL E J. Predicting the future: big data, machine learning, and clinical medicine [J]. *New Engl J Med*, 2016, 375(13): 1216-1219.
- [4] KNUUTI J, WIJNS W, SARASTE A, et al. 2019 ESC guidelines for the diagnosis and management of chronic coronary syndromes [J]. *Europ Heart J*, 2020, 41(3): 407-477.
- [5] 中华医学会糖尿病学分会. 中国 2 型糖尿病防治指南 (2020 年版) [J]. *中国实用内科杂志*, 2021, 41(8): 668-695.
- [6] HAUG C J, DRAZEN J M. Artificial intelligence and machine learning in clinical medicine, 2023 [J]. *New Engl J Med*, 2023, 388(13): 1201-1208.
- [7] SHEHAB M, ABUALIGAH L, SHAMBOUR Q, et al. Machine learning in medical applications: a review of state-of-the-art methods [J]. *Comput Biol Med*, 2022, 145: 105458.
- [8] 刘巧红, 马雨生, 蔡雨晨. 基于 XGBoost 算法的糖尿病分类预测模型及应用 [J]. *现代仪器与医疗*, 2023, 29(4): 1-6.
- [9] HE F, NG YIN LING C, NUSINOVICI S, et al. Development and external validation of machine learning models for diabetic microvascular complications: cross-sectional study with metabolites [J]. *J Med Internet Res*, 2024, 26: e41065.
- [10] CHEN Q, HU H, HE Q, et al. Evaluating the risk of developing hyperuricemia in patients with type 2 diabetes mellitus using least absolute shrinkage and selection operator regression and machine learning algorithm [J]. *Digit Health*, 2024, 10: 20552076241241381.
- [11] TIBSHIRANI R, SERIES B. Regression shrinkage and selection via the Lasso [J]. *J Roy Stat Soc B*, 1996, 58(1): 267-288.
- [12] 谭静仪, 蔡灿, 林爱华, 等. 基于 SVM 算法的冠心病分类预测案例研究 [J]. *医学信息*, 2023, 36(1): 37-41.
- [13] 桑祎莹, 黄仕鑫, 易静, 等. Logistic 回归模型和随机森林模型诊断糖尿病周围神经病变效能的比较 [J]. *广西医学*, 2022, 44(5): 511-515.
- [14] 张占林, 孙勇, 妥小青, 等. 随机森林算法对体检人群糖尿病患病风险的预测价值研究 [J]. *中国全科医学*, 2019, 22(9): 1021-1026.
- [15] QUAN K R, LIN W R, HONG J B, et al. A machine learning approach for predicting radiation-induced hypothyroidism in patients with nasopharyngeal carcinoma undergoing tomotherapy [J]. *Sci Rep*, 2024, 14(1): 8436.
- [16] 张富, 颜玉云. 基于 XGBoost 算法构建颈动脉粥样硬化患病风险初筛模型 [J]. *中国医疗器械信息*, 2023, 29(13): 10-15.
- [17] LUNDBERG S M, NAIR B, VAVILALA M S, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery [J]. *Nat Biomed Eng*, 2018, 2(10): 749-760.
- [18] CHEN Y, XU L, CHENG Z, et al. Progression from different blood glucose states to cardiovascular diseases: a prospective study based on multi-state model [J]. *Europ J Prev Cardiol*, 2023, 30(14): 1482-91.
- [19] 施仲伟. 回眸过去 30 年全球和中国的心血管疾病负担及其危险因素: 1990 年至 2019 年全球心血管疾病负担及其危险因素报告解读 [J]. *诊断学理论与实践*, 2021, 20(4): 349-355.
- [20] DE BRUYNE S, SPEECKAERT M M, VAN BIESEN W, et al. Recent evolutions of machine learning applications in clinical laboratory medicine [J]. *Crit Rev Clin Lab Sci*, 2021, 58(2): 131-152.

(收稿日期: 2024-07-19 修回日期: 2024-11-09)

(上接第 134 页)

- [8] OLZMANN J A, CARVALHO P. Dynamics and functions of lipid droplets [J]. *Nat Rev Mol Cell Biol*, 2019, 20(3): 137-155.
- [9] DIAMANTIS S, LONGUET P, LESPRIT P, et al. Terms of use of outpatient parenteral antibiotic therapy [J]. *Infect Dis Now*, 2021, 51(1): 14-38.
- [10] CURTARELLO M, TOGNON M, VENTUROLI C, et al.

Rewiring of lipid metabolism and storage in ovarian cancer cells after anti-VEGF therapy [J]. *Cells*, 2019, 8(12): 1601.

- [11] BENSAAK K, FAVARO E, LEWIS C A, et al. Fatty acid uptake and lipid storage induced by HIF-1 α contribute to cell growth and survival after hypoxia-reoxygenation [J]. *Cell Rep*, 2014, 9(1): 349-365.

(收稿日期: 2024-08-19 修回日期: 2024-09-27)