

# 基于 H2O 平台自动化机器学习的糖尿病视网膜病变预测模型的建立

王慧霞<sup>1</sup>, 张玉婷<sup>2</sup>, 朱曼辉<sup>2</sup>

(苏州大学理想眼科医院斜视与小儿眼科<sup>1</sup>, 病理科<sup>2</sup>, 江苏 苏州 215000)

**摘要:**目的 利用 H2O 平台推出的自动化机器学习 (AutoML) 算法, 建立预测糖尿病 (DM) 视网膜病变 (DR) 模型。方法 纳入 2019 年 1 月–2021 年 1 月于本院就诊的 606 例 DM 患者, 根据眼底照相分为单纯 DM 组 (DM 组, 303 例) 及 DM 合并 DR 组 (DR 组, 303 例)。采集两组患者基本情况、血生化检测结果及视网膜图像等数据。利用 H2O AutoML 算法建立针对 DR 二分类结局, 进行变量筛选并建立机器学习预测模型, 产生相应预测结果, 据此绘制 ROC 曲线并建立混淆矩阵, 绘制 SHAP 及部分依赖图, 评价模型区分能力。结果 DR 组糖尿病病程长于 DM 组, 吸烟、饮酒、高血压、脂肪肝比例、腰臀比、BMI 及收缩压高于 DM 组, 差异有统计学意义 ( $P < 0.05$ ); DR 组 HDL-C 低于 DM 组, FPG, FINS, HOMA-IR, HbA1c, ALT 和 AST 均高于 DM 组, 差异有统计学意义 ( $P < 0.05$ )。将两组特征数据载入 AutoML 工作环境中, 得到最佳模型为通用梯度回归模型 (GBM), 该模型 Gini 值 0.914,  $R^2$  为 0.679, LogLoss 为 0.260。重要性排名前 3 的变量包括 FPG、糖尿病病程及 FINS。在 Train 数据集中, ROC 曲线下面积为 0.942 (95%CI: 0.921~0.963)。利用混淆矩阵得到特异度为 0.924, 敏感度为 0.959, 准确度为 0.942, 误分类率为 0.058。在 Valid 数据集中, ROC 曲线下面积为 0.831 (95%CI: 0.764~0.897)。利用混淆矩阵得到特异度为 0.828, 敏感度为 0.833, 准确度为 0.831, 误分类率为 0.169。结论 本次利用 AutoML 算法建立的通用梯度回归 DR 患病预测模型可用于 DM 人群中 DR 的筛查。

**关键词:** 糖尿病视网膜病变; 自动化机器学习; 预测模型; 混淆矩阵; SHAP 可视化; 部分依赖图

中图分类号: R587.2; R774.1

文献标识码: A

DOI: 10.3969/j.issn.1006-1959.2023.22.002

文章编号: 1006-1959(2023)22-0008-06

## Establishment of Diabetic Retinopathy Prediction Model Based on H2O Platform Automated Machine Learning

WANG Hui-xia<sup>1</sup>, ZHANG Yu-ting<sup>2</sup>, ZHU Man-hui<sup>2</sup>

(Department of Strabismus and Pediatric Ophthalmology<sup>1</sup>, Department of Pathology<sup>2</sup>, Lixiang Eye Hospital of Soochow University, Suzhou 215000, Jiangsu, China)

**Abstract:** **Objective** To establish a model for predicting diabetic retinopathy (DR) by using the automatic machine learning (AutoML) algorithm introduced in the H2O platform. **Methods** A total of 606 patients with DM who were treated in our hospital from January 2019 to January 2021 were divided into simple DM group (DM group, 303 patients) and DM combined with DR group (DR group, 303 patients) according to fundus photography. The basic conditions, blood biochemical test results and retinal images of the two groups were collected. The H2O AutoML algorithm was used to establish a machine learning prediction model for DR binary classification outcomes, and the corresponding prediction results were generated. Based on this, the ROC curve was drawn and the confusion matrix was established. SHAP and partial dependence graph were drawn to evaluate the model's ability to distinguish. **Results** The duration of diabetes in the DR group was longer than that in the DM group, and the proportion of smoking, drinking, hypertension, fatty liver, waist-to-hip ratio, BMI and systolic blood pressure were higher than those in the DM group ( $P < 0.05$ ). HDL-C in DR group was lower than that in DM group, FPG, FINS, HOMA-IR, HbA1c, ALT and AST in DR group were higher than those in DM group, the differences were statistically significant ( $P < 0.05$ ). The two sets of feature data are loaded into the AutoML working environment, and the best model was the general gradient regression model (GBM), the Gini value of the model was 0.914,  $R^2$  was 0.679, and LogLoss was 0.260. The top three important variables included FPG, duration of diabetes and FINS. In the Train dataset, the area under the ROC curve was 0.942 (95%CI: 0.921–0.963). Using the confusion matrix, the specificity was 0.924, the sensitivity was 0.959, the accuracy was 0.942, and the misclassification rate was 0.058. In the Valid dataset, the area under the ROC curve was 0.831 (95%CI: 0.764–0.897). Using the confusion matrix, the specificity was 0.828, the sensitivity was 0.833, the accuracy was 0.831, and the misclassification rate was 0.169. **Conclusion** The gradient boost machine model for DR diagnosis prediction based on H2O AutoML can be used for DR screening in DM population.

**Key words:** Diabetic retinopathy; Automatic machine learning; Prediction model; Confusion matrix; Shapley additive explanations; Partial dependence plots

糖尿病视网膜病变 (diabetic retinopathy, DR) 是最常见且患病率较高的糖尿病 (diabetic mellitus,

DM) 微血管并发症之一, 主要病理改变是视网膜血管闭塞性循环障碍<sup>[1]</sup>。DR 严重威胁着糖尿病患者的

基金项目: 1. 苏州市医疗卫生科技创新-应用基础研究 (编号: SKJYD2021044); 2. 苏州市姑苏卫生人才计划 [编号: (2022)192]

作者简介: 王慧霞 (1978.4–), 女, 江苏苏州人, 本科, 主治医师, 主要从事代谢性眼病的临床诊治研究

通讯作者: 朱曼辉 (1991.1–), 男, 江苏苏州人, 硕士, 主治医师, 主要从事代谢性疾病的临床模型研究

生存质量,同时给社会带来严重经济负担。早期筛查防治 DR,对于改善 DM 患者远期预后至关重要<sup>[2]</sup>。近年来,国内外学者基于医院或社区的横断面或病例对照研究,利用广义线性模型算法,建立了一系列预测 DM 患者发生 DR 的模型及列线图,展现出较好的预判能力和临床运用效果<sup>[3,4]</sup>。广义线性模型属于机器学习中发展较早的统计学模型,而今机器学习飞速发展,算法不断更新,在监督学习中涌现出如支持向量机、决策树、朴素贝叶斯、人工神经网络等算法<sup>[5]</sup>。算法的不断更迭进步,伴随而来是对建模者专业知识不断增长的要求,这大大提高了临床医生利用机器学习算法的门槛。自动化机器学习(AutoML)可通过自动化建模和调参程序,大大降低建模人员的工作门槛和负担<sup>[6]</sup>。本研究采集本院患者一般资料及实验室检查结果,利用 H2O 运算平台推出的 AutoML 算法建立预测模型,旨在为 DM 人群的 DR 筛查提供新的思路。

## 1 资料与方法

1.1 一般资料 选取 2019 年 1 月–2021 年 1 月于苏州大学理想眼科医院就诊患者电子病历数据,采用 ICD-10 疾病分类标准进行编码。纳入 606 例 DM 患者,根据眼底照相分为单纯 DM(DM 组)303 例及 DM 合并 DR(DR 组)303 例。纳入标准:所有患者诊断均符合最新国内临床指南<sup>[1,2]</sup>。排除标准:①其他原因导致的高血糖;②合并 2 型 DM 急性并发症;③妊娠及哺乳期女性;④合并显著肝肾功能异常、严重的心脑血管疾病或恶性肿瘤等;⑤资料不全者。本研究已获苏州大学附属理想眼科医院伦理委员会批准(批准号 SLER2018112),所有患者均签署知情同意书。

### 1.2 临床及实验室检测

1.2.1 人体数据测量 测量并记录纳入人群的身高、体重、血压(SBP/DBP)、臀围和腰围,并计算体质量指数(BMI)和腰臀比(WHR)。采集患者既往病史、服药史、烟酒史,并通过计算代谢当量商(metabolic equivalent of task, MET)计算体力活动量。

1.2.2 DR 检查方法 视网膜检查使用免散瞳眼底数码相机(型号:TRC-NW300),由专科技师在摄片暗室进行。患者进入暗室休息 5 min,待视觉适应后,由技师应用免散瞳眼底数码相机进行拍摄以黄斑为中心的视网膜彩色图像,每只眼睛拍摄一张照片。照片由对此项研究盲法的眼科医生进行阅读。

1.2.3 血清学检测 获取所有纳入者空腹静脉血 10 ml。将收集的血样统一离心,立刻上机检测或储存

于-80℃冰箱待检。血清生化分析采用 Mindray 迈瑞 800 全自动生化分析仪。检测项目包括:①糖代谢和胰岛素功能相关指标:空腹血糖(FPG)、空腹胰岛素(FINS)及糖化血红蛋白(HbA1c),并计算稳态模型胰岛素抵抗指数(HOMA-IR);②肝酶指标:谷丙转氨酶(ALT)、谷草转氨酶(AST)及  $\gamma$ -谷氨酰转肽酶(GGT);③脂质代谢相关指标:三酰甘油(TG)、高密度脂蛋白胆固醇(HDL-C)及低密度脂蛋白胆固醇(LDL-C)。

1.3 评价方法 利用 H2O 运算平台推出的 AutoML 算法建立针对 DR 二分类结局的机器学习预测模型,产生相应预测结果,据此绘制受试者工作特征(ROC)曲线并建立混淆矩阵,计算特异度、敏感度、准确度及误分类率,评价模型区分能力。

1.4 统计学方法 本研究建模及绘图软件包括:R(4.0.4 版)、H2O 包(H2O cluster 版本:h2o\_3.32.1.7)、tableone 包(0.13.2 版本)及 lime 包(0.5.3 版本)。计量资料以( $\bar{x} \pm s$ )或 $[M(P_{25}, P_{75})]$ 表示,组间比较采用 Student's *t* 检验或 Mann-Whitney *U* 检验。计数资料采用 $[n(\%)]$ 表示,比较采用  $\chi^2$  检验。为进一步了解变量在模型中的重要性及分布情况,进行可视化分析,包括 Shapley Additive exPlanations (SHAP 分析)、Partial dependence(部分依赖)及 LIME 可视化。双侧  $P < 0.05$  为差异有统计学意义。

## 2 结果

2.1 两组一般及临床资料比较 DR 组糖尿病病程长于 DM 组,吸烟、饮酒、高血压、脂肪肝比例、腰臀比、BMI 及收缩压高于 DM 组,差异有统计学意义( $P < 0.05$ );DR 组 HDL-C 低于 DM 组,FPG、FINS、HOMA-IR、HbA1c、ALT 和 AST 均高于 DM 组,差异有统计学意义( $P < 0.05$ ),见表 1。

2.2 模型建立判断 DR 发病风险 将上述单因素分析中存在差异的变量纳入 AutoML 机器学习工作环境中,利用 H2O 平台进行随机分组、特征选择、建模运算及验证。将 606 例 DM 患者按照 8:2 比例随机分组为 Train 集(482 例)及 Valid 集(124 例)。最佳模型为通用梯度回归模型(generalized boosted regression model, GBM)。这是一种由多棵决策树组成的迭代决策树算法。该模型(基于 Train 数据集的 5 折交叉验证):Gini 值 0.914,  $R^2$  为 0.679, LogLoss 为 0.260。模型中各变量的重要性见表 2。其中在最佳模型 GBM 中,重要性排名前 3 的变量(即对模型贡献的排名)分别为空腹血糖、糖尿病病程及空腹胰岛素,其占比均超过 10%。

表 1 两组一般及临床资料比较 $[\bar{x}\pm s, M(P_{25}, P_{75})]$ 

项目	DM 组( $n=303$ )	DR 组( $n=303$ )	统计值	$P$
年龄(岁)	64.51 $\pm$ 11.77	64.6 $\pm$ 8.02	$t=0.171$	0.867
糖尿病病程(年)	6.88 $\pm$ 2.26	9.72 $\pm$ 3.51	$t=11.840$	0.000
男性 $[n(\%)]$	225(74.25)	225(74.25)	$\chi^2=0.000$	0.999
吸烟史 $[n(\%)]$	75(24.75)	115(37.95)	$\chi^2=12.270$	0.001
饮酒史 $[n(\%)]$	104(34.32)	143(47.19)	$\chi^2=10.390$	0.002
高血压 $[n(\%)]$	74(24.42)	121(39.93)	$\chi^2=16.702$	0.000
脂肪肝 $[n(\%)]$	112(36.96)	222(73.26)	$\chi^2=80.711$	0.000
WHR $[n(\%)]$	74(24.42)	121(39.93)	$\chi^2=17.87$	0.205
BMI( $\text{kg}/\text{m}^2$ )	23.71 $\pm$ 3.45	24.63 $\pm$ 3.52	$t=3.214$	0.001
SBP(mmHg)	126.97 $\pm$ 18.65	130.62 $\pm$ 20.72	$t=2.588$	0.023
DBP(mmHg)	76.96 $\pm$ 11.57	78.36 $\pm$ 13.12	$t=1.398$	0.164
MET(h/周)	115.00 $\pm$ 38.00	75.00 $\pm$ 24.80	$t=15.49$	0.000
UA( $\mu\text{mol}/\text{L}$ )	317.20 $\pm$ 89.02	302.37 $\pm$ 95.04	$t=2.006$	0.048
TG( $\text{mmol}/\text{L}$ )	1.97 $\pm$ 0.87	1.81 $\pm$ 0.77	$t=1.875$	0.061
HDL-C( $\text{mmol}/\text{L}$ )	1.20 $\pm$ 0.32	1.07 $\pm$ 0.34	$t=4.847$	0.000
LDL-C( $\text{mmol}/\text{L}$ )	2.54 $\pm$ 0.85	2.51 $\pm$ 0.80	$t=0.447$	0.643
FPG( $\text{mmol}/\text{L}$ )	5.14(4.71, 5.78)	7.37(5.80, 10.36)	$U=-2.580$	0.000
FINS( $\text{U}/\text{ml}$ )	10.30(7.20, 14.80)	17.30(11.90, 24.10)	$U=-3.531$	0.000
HOMA-IR	3.02 $\pm$ 1.93	6.23 $\pm$ 5.17	$t=10.130$	0.000
HbA1c(%)	6.10(5.40, 7.40)	8.60(6.90, 10.55)	$U=-2.044$	0.000
ALT( $\text{U}/\text{L}$ )	26.00(18.50, 36.00)	26.00(21.00, 41.00)	$U=-1.901$	0.008
AST( $\text{U}/\text{L}$ )	26.00(20.00, 34.00)	31.00(27.00, 37.00)	$U=-3.017$	0.000
GGT( $\text{U}/\text{L}$ )	25.00(17.00, 42.00)	29.00(17.00, 39.00)	$U=-0.875$	0.311

表 2 最佳模型中各变量重要性排名

项目	相对重要性	重要性百分比(%)	项目	相对重要性	重要性百分比(%)
空腹血糖	116.6	22.41	LDL-C	13.8	2.76
糖尿病病程	113.5	21.82	HDL-C	13.5	2.62
空腹胰岛素	89.2	17.20	AST	11.4	2.24
HOMA-IR	35.5	6.85	BMI	11.1	2.13
HbA1c	25.3	4.98	SBP	8.8	1.73
TG	22.7	4.42	吸烟	5.1	1.02
体力活动	19.6	3.85	高血压	1.5	<0.10
UA	16.0	3.16	饮酒	1.3	<0.10
ALT	14.4	2.80	脂肪肝	0.5	<0.10

2.3 变量在模型中的作用 在最佳模型 GBM 中,各变量的 SHAP 特征图绘制在图 1 中。重要性排名前 3 的变量是空腹血糖、糖尿病病程及空腹胰岛素。三者结局二分类中的分布,体现其标准化数值与发病呈正相关趋势。图 2 为 LIME 可视化,显示是随机抽取的 8 个样本(DM 组 5 例,DR 组 3 例),3 个重要变量对预测结果的重要性贡献。图 3 显示的 3 个变量在模型中的部分依赖图,可以看出三者与结果基本上呈现的是单调的上升趋势。

2.4 预测模型的区分能力 采用上述获得的预测模型 GBM 绘制预判 DR 发病的 ROC 曲线并建立混淆矩阵。在 Train 数据集中,ROC 曲线下面积为 0.942(95%CI:0.921~0.963)。利用混淆矩阵得到特异度为 0.924,敏感度为 0.959,准确度为 0.942,误分类率为 0.058。在 Valid 数据集中,ROC 曲线下面积为 0.831(95%CI:0.764~0.897)。利用混淆矩阵得到特异度为 0.828,敏感度为 0.833,准确度为 0.831,误分类率为 0.169,见表 3。

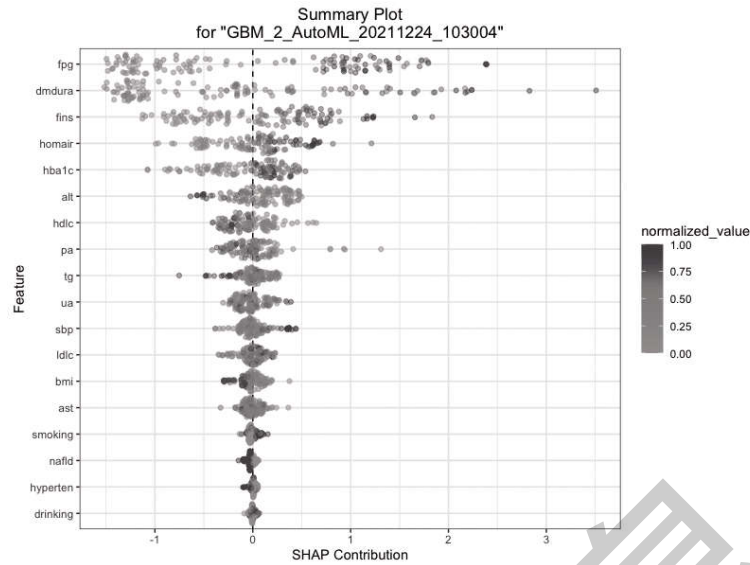


图 1 最佳模型中各变量 SHAP 特征

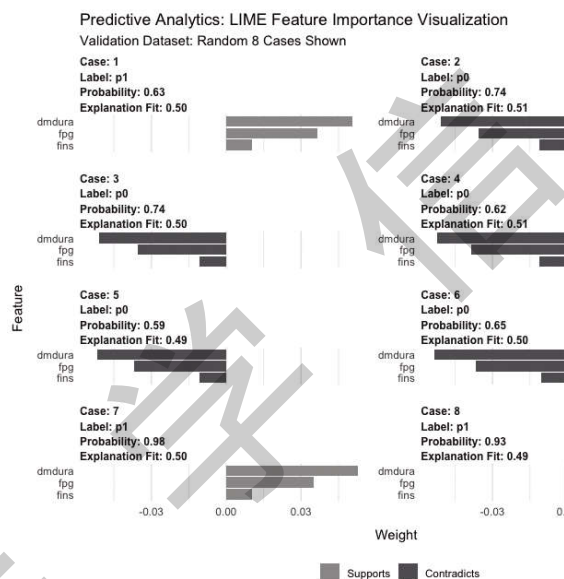
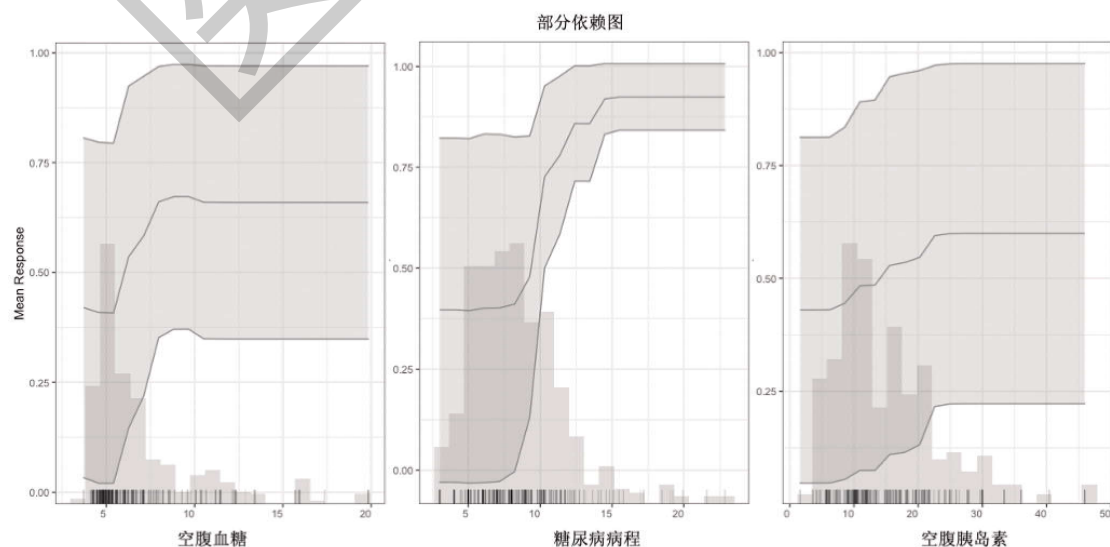


图 2 随机样本中变量重要性 LIME 可视化



注:部分依赖显示的是 3 个变量对最佳模型的预测结果的边际效应

图 3 最佳模型中变量部分依赖图

表 3 最佳预测模型 GBM 在各个数据集中的区分能力

数据集	ROC 曲线		混淆矩阵				
	AUC	95% CI	特异度	敏感度	查准率	准确度	误分类率
Train 数据集	0.942	0.921~0.963	0.925	0.959	0.928	0.942	0.058
Valid 数据集	0.831	0.764~0.897	0.828	0.833	0.820	0.831	0.169

### 3 讨论

随着机器学习及深度学习的快速发展,选择合适的模型并根据超参数不断调整优化模型,这一过程变得十分繁琐且耗时。此外,它还对建模者的计算机知识提出了较高要求<sup>[5]</sup>。近几年,各大科技公司陆续推出自动化学习平台,如谷歌公司推出的 Cloud AutoML、H2O 平台的 AutoML 等<sup>[6]</sup>。AutoML 可以自动化完成机器学习的前期工作:包括数据准备、编码、功能选择/提取以及工程化环境。同时,在模型生成过程中,如模型算法选择、优化、迭代以及验证,AutoML 均可在少量代码基础上实现。

H2O AutoML 是 H2O 平台提供一种针对自动化工作流程的算法,主要功能涵盖:轻数据准备、环境建立、模型选择及优化<sup>[6]</sup>。其中亮点包括:通过堆叠一组集成学习模型,自动化迭代模型。当前版本 H2O AutoML 可建立并交叉验证以下模型:广义线性模型、随机森林、极随机森林、梯度提升机随机网络、XGBoosts、深神经网络随机网络以及相关的集成学习。H2O AutoML 的一大特点就是组合堆叠多个集成学习算法,以获得比从单个成分学习算法更好的预测性能,广泛运用在监督学习中。许多流行的现代机器学习算法实际上都是集成学习。例如,随机森林和梯度提升机都是通过组合弱学习算法(例如决策树)并形成单一、强学习的集成学习法。

近年来,随着人民群众饮食结构改变、人口老龄化等因素的影响,我国糖尿病患病率大幅增加。流行病学调查显示<sup>[7]</sup>,我国大陆糖尿病患者中合并视网膜病变约占 23%。由于 DR 的早期隐蔽性、慢性进展性、不可逆性等特点,目前早期筛查工作仍然面临诸多困难,选择基于临床资料、实验室检查的无创性筛查工具是 DR 的防治工作重点<sup>[8]</sup>。DR 的发生发展是一个较复杂的病理生理过程,具体的机制尚不明确<sup>[9]</sup>。目前观点认为<sup>[10]</sup>,DM 患者机体能量代谢障碍,特别是胰岛素抵抗和糖脂代谢紊乱可诱发视网膜病变。DR 在病程 10 年以上的 DM 患者中患病率高达

80%,是全球中老年人视力丧失的首要病因<sup>[11]</sup>。患者随着 DM 病程的延长,一系列机体能量代谢功能紊乱逐渐发展并加重,其中如高血糖或血糖波动大、脂类代谢紊乱等都被证实是 DR 的独立危险因素。

及时的 DR 筛查对于高危病例至关重要,通过全面的眼科检查与干预,来避免永久性的视力丧失。在过去几年中,各个国家通过大型流行病学研究提出了综合各类风险因素的个性化筛查方案,体现了较好的成本-效益比。DCCT 是北美地区一项开展了近 40 年的大型队列研究,该研究提示长病程、高 HbA1c、高血压等是 DR 发生发展的独立危险因素,为基于社区的筛查提供标记工具<sup>[12]</sup>。有研究<sup>[4]</sup>报道了一项跨我国多省份的横断面研究,发现 DM 发病早、病程久、高血压、高血糖及高 HbA1c 是 DR 的独立危险因素;该团队构建了列线图模型,为 DR 早筛早诊提供帮助。基于人工智能的新技术,包括移动设备检测系统、数字图像算法等,将在未来改变筛查模式,再次改善成本-效益比<sup>[13,14]</sup>。

本研究收集单纯 DM 患者和合并 DR 的 DM 患者临床资料及实验室结果,利用 H2O 平台的 AutoML 算法进行自动化的变量筛选,发现高血糖、DM 病程、高空腹胰岛素等是该地区 DR 发生的重要独立危险因素,这与国内外研究相似<sup>[3,4,12]</sup>。采用上述因素建立的 GBM 预测模型判断 DR 发病,显示出较好区分能力以及均衡的敏感度和特异度,优于基于线性模型算法的相关报道<sup>[3,15]</sup>。本研究中的机器学习模型在训练集和测试集中都展现了良好的特异性,作为一个初筛工具,可以用于社区高危人群的筛选,大大降低眼科卫生机构的工作压力,避免医疗资源的浪费<sup>[16-18]</sup>。

本研究报道了利用 AutoML 算法处理 DM 患者数据,为今后的 DR 筛查提供参考和思路。该方法优势在于:一方面相比较传统的逻辑回归等统计模型,提高了模型准确率;另外,使用自动化算法,大大降低了医务人员利用人工智能技术的门槛,为今后

AutoML 在临床科研领域的应用提供参考<sup>[19,20]</sup>。本研究样本是基于我院的单中心数据,利用随机分组和交叉验证,具有较高的参考价值。但仍需多中心的外部验证进一步评估该模型在 DR 诊断中的作用。

综上所述,本次利用 AutoML 算法建立的通用梯度回归 DR 患病预测模型可用于 DM 人群中 DR 的筛查。

#### 参考文献:

- [1]中华医学会糖尿病学分会视网膜病变学组.糖尿病视网膜病变防治专家共识[J].中华糖尿病杂志,2018,10(4):241-247.
- [2]中华医学会糖尿病学分会.中国2型糖尿病防治指南(2017年版)[J].中国实用内科杂志,2018,38(4):292-344.
- [3]Semeraro F,Parrinello G,Cancarini A,et al.Predicting the risk of diabetic retinopathy in type 2 diabetic patients [J].J Diabetes Complications,2011,25(5):292-297.
- [4]Liu Y,Yang J,Tao L,et al.Risk factors of diabetic retinopathy and sight-threatening diabetic retinopathy: a cross-sectional study of 13 473 patients with type 2 diabetes mellitus in mainland China[J].BMJ Open,2017,7(9):e016280.
- [5]Handelman GS,Kok HK,Chandra RV,et al.eDoctor: machine learning and the future of medicine[J].J Intern Med,2018,284(6):603-619.
- [6]LeDell E,Poirier S.H2O AutoML:Scalable Automatic Machine Learning [C]//7th ICML Workshop on Automated Machine Learning(AutoML).2020.
- [7]Pan CW,Wang S,Qian DJ,et al.Prevalence, Awareness, and Risk Factors of Diabetic Retinopathy among Adults with Known Type 2 Diabetes Mellitus in an Urban Community in China[J].Ophthalmic Epidemiol,2017,24(3):188-194.
- [8]Saeidian J,Mahmoudi T,Riazi-Esfahani H,et al.Automated assessment of the smoothness of retinal layers in optical coherence tomography images using a machine learning algorithm[J].BMC Med Imaging,2023,23(1):21.
- [9]Han R,Jin M,Xu G,et al.Progressively Decreased HCN1 Channels Results in Cone Morphological Defects in Diabetic Retinopathy[J].J Neurosci,2022,42(43):8200-8212.
- [10]Petersen MC,Shulman GI.Mechanisms of Insulin Action and Insulin Resistance[J].Physiol Rev,2018,98(4):2133-2223.
- [11]Cheung N,Mitchell P,Wong TY.Diabetic retinopathy [J].Lancet,2010,376(9735):124-136.
- [12]Hainsworth DP,Bebu I,Aiello LP,et al.Risk Factors for Retinopathy in Type 1 Diabetes:The DCCT/EDIC Study [J].Diabetes Care,2019,42(5):875-882.
- [13]Hayes AJ,Leal J,Gray AM,et al.UKPDS outcomes model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study:UKPDS 82[J].Diabetologia,2013,56(9):1925-1933.
- [14]Vujosevic S,Aldington SJ,Silva P,et al.Screening for diabetic retinopathy: new perspectives and challenges [J].Lancet Diabetes Endocrinol,2020,8(4):337-347.
- [15]Mo R,Shi R,Hu Y,et al.Nomogram-Based Prediction of the Risk of Diabetic Retinopathy:A Retrospective Study [J].J Diabetes Res,2020,2020:7261047.
- [16]明帅,姚溪,谢坤鹏,等.糖尿病视网膜病变人工智能自动诊断系统在社区和医院老年糖尿病患者中的应用效果分析[J].中华眼底病杂志,2022,38(2):120-125.
- [17]Ji Y,Chen N,Liu S,et al.Research Progress of Artificial Intelligence Image Analysis in Systemic Disease-Related Ophthalmopathy[J].Dis Markers,2022,2022:3406890.
- [18]Dong L,He W,Zhang R,et al.Artificial Intelligence for Screening of Multiple Retinal and Optic Nerve Disease [J].JAMA Netw Open,2022,5(5):e229960.
- [19]Ma J,Jiang S,Liu Z,et al.Machine Learning Models for Slope Stability Classification of Circular Mode Failure: An Updated Database and Automated Machine Learning (AutoML) Approach[J].Sensors (Basel),2022,22(23):9166.
- [20]Wang Y,Hong Y,Wang Y,et al.Automated Multimodal Machine Learning for Esophageal Variceal Bleeding Prediction Based on Endoscopy and Structured Data [J].J Digit Imaging,2023,36(1):326-338.

收稿日期:2023-01-04;修回日期:2023-02-07

编辑/成森