

# 2 型糖尿病并发动脉粥样硬化风险预测模型比较\*

王一凡 石超君 马晓洁 冯文佳 安洪庆 高倩倩 井 淇 蔡伟芹 马安宁

(山东第二医科大学 潍坊 261053)

**[摘要]** 目的/意义 探究各种预测模型在 2 型糖尿病并发动脉粥样硬化风险预测中的应用及其准确率。方法/过程 以国家人口健康科学数据中心“糖尿病并发症预警数据集”的生化数据表为基础,运用 MATLAB 软件,基于最近邻、决策树、反向传播神经网络、朴素贝叶斯模型构建 2 型糖尿病并发动脉粥样硬化的风险预测模型,并进行比较分析。结果/结论 从有效性看,朴素贝叶斯算法的预测准确率最高(61.6%);其次为决策树模型算法(58.2%)、最近邻算法(57.7%)、反向传播神经网络算法(55.9%);混淆矩阵结果与受试者操作特征曲线结果显示,朴素贝叶斯模型表现最好。4 种算法构建的预测模型有效性、性能、稳定性等方面均是朴素贝叶斯模型最优。

**[关键词]** 2 型糖尿病并发症;动脉粥样硬化;疾病预测模型

**[中图分类号]** R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2024.07.013

## Comparison of Risk Prediction Models for Atherosclerosis in Type 2 Diabetes Mellitus

WANG Yifan, SHI Chaojun, MA Xiaojie, FENG Wenjia, AN Hongqing, GAO Qianqian, JING Qi, CAI Weiqin, MA Anning  
Shandong Second Medical University, Weifang 261053, China

**[Abstract]** **Purpose/Significance** To explore the application and predictive accuracy of various models in predicting the risk of atherosclerosis in diabetic patients. **Method/Process** Based on the biochemical data table from the “Diabetes Complications Warning Dataset” provided by the National Population Health Science Data Center, MATLAB software is used to construct risk prediction models for diabetes-induced atherosclerosis. The models are built by using k-nearest neighbors (KNN), decision trees, backpropagation (BP) neural networks, and Naive Bayes algorithms, and which are subjected to comparative analysis. **Result/Conclusion** In terms of effectiveness, the predictive accuracy of Naive Bayes algorithm is the highest (61.6%), followed by the decision tree model (58.2%), the KNN model (57.7%), and the BP neural network model (55.9%). The results of the confusion matrix and the receiver operating characteristic (ROC) curve indicate that the Naive Bayes model performs best. When comparing the models in terms of effectiveness, performance and stability, the Naive Bayes model is superior.

**[Keywords]** type 2 diabetes complications; atherosclerosis; disease prediction model

**[修回日期]** 2024-03-25

**[作者简介]** 王一凡, 硕士研究生; 通信作者: 马安宁, 博士。

**[基金项目]** 国家自然科学基金资助项目(项目编号: 72104186、72004165); 山东省自然科学基金项目(项目编号: ZR2021MG019); 山东省重点研发(软科学项目)重大项目(项目编号: 2020RZB14001)。

## 1 引言

2 型糖尿病是慢性代谢性疾病的一种<sup>[1]</sup>。发病的主要缘由是身体无法有效使用胰岛素，其症状与 1 型糖尿病相似，却更难以发现，因此 2 型糖尿病可能是在发病数年之后或并发症出现后才被诊断出。2 型糖尿病可能导致多样化的并发疾病，例如缺血性心肌病、脑卒中、贫血性外围动脉疾病，以及微血管相关疾病，如视网膜病变、神经病变和肾病变等<sup>[2]</sup>。中国 2022 年在糖尿病治疗领域的花费约为 6 210 亿美元，治疗费用较高的原因主要是并发症的治疗<sup>[3]</sup>。因此，预防糖尿病及其并发症是一项重大的公共卫生目标<sup>[4]</sup>。本研究旨在分析 2 型糖尿病患者各项指标与并发症动脉粥样硬化之间的关联程度，通过机器学习方法建立相关指标与动脉粥样硬化的预测模型，以更好地预测并发症。

## 2 相关研究

英国糖尿病研究是来自英国的糖尿病前瞻性队列研究，构建了首个 2 型糖尿病并发症预测模型，在糖尿病研究领域居标志性地位<sup>[5]</sup>，侧重于评价不同疗法对 2 型糖尿病患者可能带来的影响，还关注 2 型糖尿病并发症，如心血管病、视网膜疾病、肾脏疾病之间的关联性。后由 Stevens R J 等<sup>[6]</sup>构建糖尿病并发症预测模型，揭示了年龄、性别、种族、吸烟历史等因素对预测 2 型糖尿病患者心血管并发症风险的重要性；Clarke P M 等<sup>[7]</sup>采用 Weibull 回归创建模型预测糖尿病相关并发症风险。

虽有众多国外学者研究 2 型糖尿病并发症风险预测模型的方法及应用，但因为归属人群不同，诸多方面存在差异，导致众多模型并不能完全适用于中国人群，且少有针对微血管并发症的研究；研究算法比较局限，一般使用单一算法，缺乏算法的横向对比，应有针对性地比较算法，更好地为知识发现及临床决策提供支持。

本研究从数据挖掘的思路出发，分析 2 型糖尿病并发症发生的关联因素，探讨多种预测模型对 2 型糖

尿病并发症动脉粥样硬化发生与否预测的准确率<sup>[8]</sup>，用测试集检验预测率，结合受试者操作特征曲线（receiver operating characteristic, ROC）的曲线下面积及预测概率图分析模型预测率，对 2 型糖尿病并发症预防以及临床康复和治疗提供科学依据。

## 3 资料与方法

### 3.1 数据来源

数据源自中国人民解放军总医院国家人口健康科学数据中心“糖尿病并发症预警数据集”<sup>[9]</sup>。

### 3.2 数据清洗

筛选出所有 2 型糖尿病患者个例数据。由于数据缺失过多和数据存在过多异常值，为了保证数据处理结果的准确性，将变量缺失值达到 30% 的变量去除，对所有有缺失的样本采用删除法剔除。采用可视化方法初步识别数据异常值，运用 Python 软件将数据异常值替换为数据均值，评估数据集的分布情况以及替换操作对分析结果的影响，发现并无明显影响。将数据无量纲化处理、Z-score 标准化处理，将数据压缩到 [0, 1]，使数据服从正态分布，以便于不同单位或量级的指标能够比较和加权。对数据库信息清洗整理后，得到有效个案 1 345 例，其中患动脉粥样硬化的 759 例，占 56.4%；无动脉粥样硬化的 586 例，占 43.6%。运用 SPSS 25.0 进行 Pearson 相关分析，得到 13 个具有明显影响的变量：年龄、体重、体重指数、高血压、高脂血、空腹血糖、糖化血红蛋白、糖化血清蛋白、总蛋白、血清白蛋白、谷丙转氨酶、 $\gamma$  谷氨酰胺转移酶、碱性磷酸酶。其他变量均排除，建立预测模型时不予考虑。

### 3.3 模型构建

3.3.1 KNN 算法 最近邻算法（K-nearest neighbor, KNN）主要解决分类问题，应用十分广泛，是一种使用所有训练样本分类的方法，也会增加计算复杂性。采用常规 KNN 方法聚类时，为提高识别准确率，通常改进 KNN 的距离度量、K 值的最优和采样的预处理等方面<sup>[10]</sup>。

3.3.2 决策树算法 决策树算法是直观且容易理解的机器学习方法，其总结和分类训练数据集的特征<sup>[11]</sup>。该算法建立在已标记的数据集基础上，通过自下而上的递归手段形成树状结构<sup>[12]</sup>。在决策树里，顶端节点被称为根节点，底端节点被名为叶节点，每个叶节点都象征着一种样本种类<sup>[13]</sup>。

3.3.3 朴素贝叶斯算法 朴素贝叶斯的核心思想就是在给定的条件下，通过预先估计原始数据，结合采样点知识，求取目标的后验概率，以概率推理的方式构建图状网络以解决更多的复杂问题<sup>[14]</sup>。

3.3.4 BP 神经网络算法 利用反向传播 (back-propagation, BP) 算法训练并优化多层前馈神经网络的非线性权重，构成 BP 神经网络。有针对 BP 神经网络预测 2 型糖尿病并发症的研究，为糖尿病肾病、视网膜病变、神经病变等构建模型，并运用思维进化算法优化神经网络。

## 4 预测模型评估

### 4.1 数据集划分

以 7:3 随机分割 1 345 例患者的资料，得到 942 个样本的训练集。通过对测试集 403 个样本进行建模评价，得到最终结果。

### 4.2 模型构建算法

使用 scikit - learn 库监督学习算法，通过算法构建、训练和评估机器学习模型，得到最优构建模型。KNN 模型近邻数设置为 5，使用 KDTree 算法计算最近临近点，度量距离使用 Minkowski 距离；决策树模型使用基尼不纯度度量树的切分质量，每个节点上的拆分策略为最佳拆分策略；朴素贝叶斯模型假设特征的可能性是高斯分布，使用极大似然去估计类变量的均值和方差；BP 神经网络算法模型隐藏层神经元设置为 100 个，激活函数为 relu 函数，使用随机梯度下降优化对数损失函数，学习率设置为 0.001，优化公差为 0.000 1。

### 4.3 参数调整方法

在该模型训练的基础上，利用 10 折交叉验证

方法优化模型。取 10 次迭代的性能指标平均值，作为模型的最终性能评估指标，以更充分利用有限的数据集，减少由于数据划分不合理或者偶然性造成的评估误差，提高模型评估的准确性和可信度。通过多次迭代训练和验证，有助于更全面地了解模型的稳定性和泛化能力，更好地选择模型和调优参数。

## 4.4 模型评估指标

4.4.1 模型有效性评估 采用准确率 (A)、精确率 (P)、召回率 (R)、F1 值作为评价指标，对预测模型的效能做出准确评估<sup>[15-16]</sup>。准确率指模型在测试集中正确预测动脉粥样硬化样本的数量占总样本数的比例。精确率指模型对动脉粥样硬化样本的正确预测数量占预测为该类型样本总量的比例。召回率指模型正确预测为动脉粥样硬化样本的数量占实际动脉粥样硬化样本总数的比例。F1 值为精确率和召回率的调和平均数，能够有效评估模型的全方位效益。TP 表示正确预测的动脉粥样硬化样本数量，FP 表示错误预测为动脉粥样硬化的非动脉粥样硬化样本数量，FN 表示被错误预测为非动脉粥样硬化的动脉粥样硬化样本数量，TN 表示正确预测为非动脉粥样硬化的样本数量。

$$A = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2PR}{P + R} \quad (4)$$

4.4.2 模型性能评估 以混淆矩阵为评估工具，确定模型预测的能力。混淆矩阵是一种用于评估分类模型性能的表格，以真实类别与模型预测类别之间的对应关系为基础进行构建。能将样本的真实类别和预测类别结合起来，生成 4 个基本指标 (TP、FP、FN、TN) 来展示分类结果的分布情况。通过使用混淆矩阵，可以有效确定分类模型的预测精确程度<sup>[17]</sup>。模型构建后，通过模糊矩阵的可视化评价其表现特征，实现对样本的识别。

4.4.3 模型稳定性评估 ROC 曲线是用于衡量模型识别负样本和正样本能力的视图，识别正样本的概率为纵轴、错误识别负样本的可能性为横轴。提升灵敏度等同于提高了正确识别正样本的概率，而

提升特异度则等同于增加了正确识别负样本的概率<sup>[18]</sup>。ROC 曲线越接近左上角，预测模型的能力就越出色。为了更有效地度量模型的预测性能，选取 ROC 曲线下面积作为直观评估模型预测精准度的方法，其数值越大意味着预测精度越高。

## 5 结果

### 5.1 预测模型有效性比较

运用上述 4 种算法分别构建预测模型，得到 4 个模型的预测表现，见表 1。其中朴素贝叶斯在测试集上的综合效果最好。

表 1 4 个模型有效性评价指标 (%)

模型名称	准确率	精确率	召回率	F1
KNN	57.7	64.7	63.3	64.0
决策树	58.2	65.5	62.5	64.0
朴素贝叶斯	61.6	62.4	89.2	73.4
BP 神经网络	55.9	64.0	59.2	61.5

### 5.2 预测模型混淆矩阵比较

混淆矩阵的行表示真实类别，列表示预测类别，因此每个单元格包含了一个类别的样本数量。在混淆矩阵中，通常将正例标记为 1，负例标记为 0。因此，“0”和“1”分别代表模型预测的两个类别，为患病和未患病。4 个模型的混淆矩阵，见图 1—4。其中朴素贝叶斯模型正确预测为非动脉粥样硬化的样本个数为 214 例，在 4 个模型中是最优的，其次为 KNN 模型（152 例）、决策树模型（150 例），BP 神经网络模型最低（142 例）。

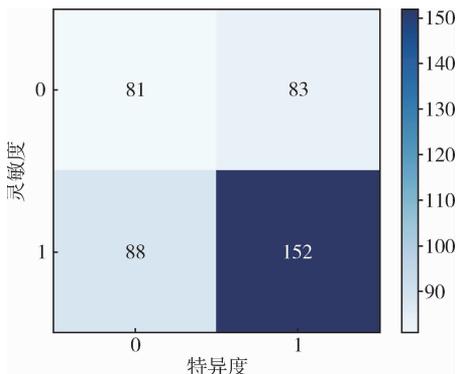


图 1 KNN 模型算法预测混淆矩阵

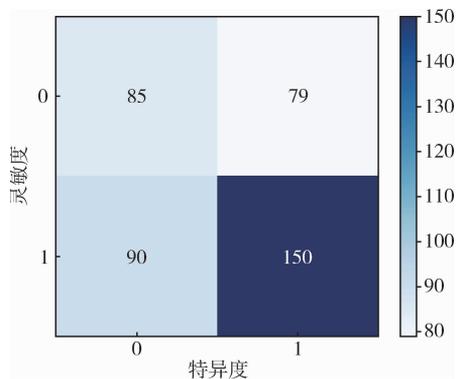


图 2 决策树算法预测混淆矩阵

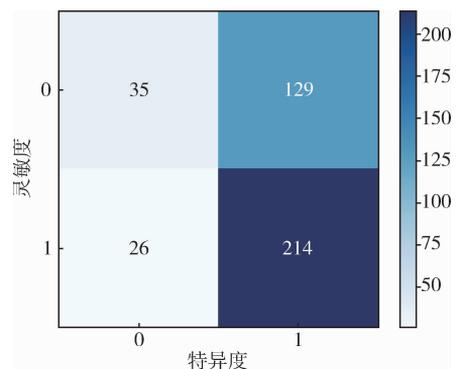


图 3 朴素贝叶斯算法预测混淆矩阵

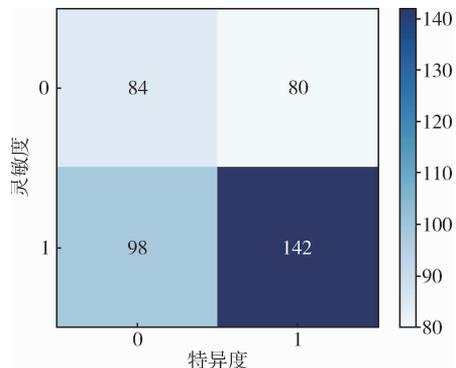


图 4 BP 神经网络算法预测混淆矩阵

### 5.3 预测模型稳定性比较

运用 ROC 曲线评价模型的稳定性，分别得出 4 个预测模型的 ROC 曲线下面积：其中朴素贝叶斯模型最高（67%），其次是 KNN 模型（59%），决策树模型和 BP 神经网络模型均为 57%，见图 5—8。说明基于朴素贝叶斯算法构建的 2 型糖尿病并发症预测模型稳定性最好。

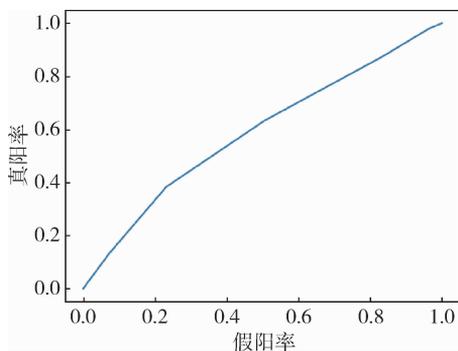


图5 KNN 模型 ROC 曲线预测指标

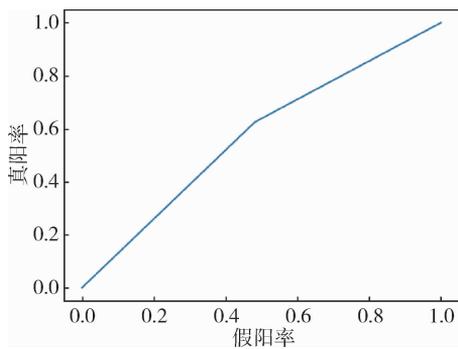


图6 决策树模型 ROC 曲线预测指标

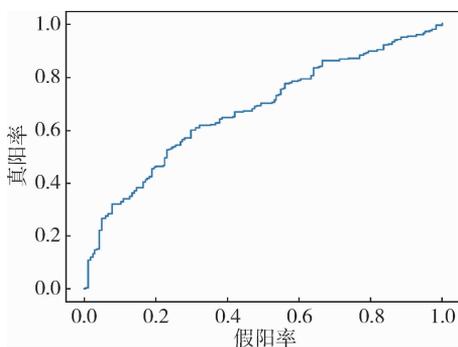


图7 朴素贝叶斯模型 ROC 曲线预测指标

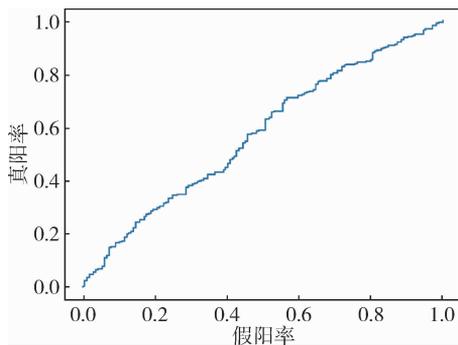


图8 BP 神经网络模型 ROC 曲线预测指标

比较 4 种模型的 ROC 曲线，ROC 曲线下面积越大，预测模型的稳定性能越好。朴素贝叶斯模型 ROC 曲线下面积最大，为 67%，说明基于朴素贝叶斯算法构建的 2 型糖尿病并发症预测模型稳定性最好。

## 6 讨论

### 6.1 模型预测性能比较

朴素贝叶斯模型在预测 2 型糖尿病患者的动脉粥样硬化方面效果最佳，因此在临床使用中，推荐采用此模型。准确性方面采用 4 个指标，朴素贝叶斯算法的预测准确率最高（61.6%）， $F1$  分数也最大，充分说明该模型的有效性在 4 个模型中最好。模型性能方面采用混淆矩阵比较，朴素贝叶斯预测模型整体上优于其他模型。高的召回率意味着可能会有更多的误检，但是会尽力找到每一个应该找到的样本。4 个模型精确率差距较小。稳定性方面，4 个模型中，朴素贝叶斯模型 ROC 曲线下面积较大，模型分类效果较好，其余 3 个预测模型近似于线性直线，因此朴素贝叶斯模型预测稳定程度更高，疾病检测比较准确。

### 6.2 模型性能原因分析

分析朴素贝叶斯预测模型性能整体优于其他预测模型的原因，主要是由于数据特点适合朴素贝叶斯模型；朴素贝叶斯模型假设特征之间相互独立，该数据集特征之间的关联性较低，且特征的独立性假设成立，所以朴素贝叶斯模型可以更好地拟合数据。朴素贝叶斯模型在小样本数据集上通常表现良好，因为其对数据的估计参数较少，对少量的训练数据也能有较好的分类效果。本研究中样本量较小，朴素贝叶斯模型比其他模型更稳定和可靠。朴素贝叶斯模型具有简单高效的特点，训练速度快，计算成本低。如果临床应用中需要处理大规模数据集或频繁地进行模型训练和调整，朴素贝叶斯模型也是一个较合适的选择。

决策树模型在本研究中表现一般，可能的原因是其对数据特征的处理方式较特殊，即通过递归将

数据集划分为不同的子集构建树形结构。如果数据特征之间存在复杂的非线性关系,决策树模型可能无法很好地拟合数据。

KNN 算法可能由于数据集中的样本分布不均匀,即不同类别的样本数量差异很大,导致在分类时偏向于样本量较大的类别,影响模型性能。

BP 神经网络需要大量的训练数据训练模型,由于本研究数据量较小,训练数据不足,无法满足 BP 神经网络模型的训练要求,影响了模型的泛化能力和性能。

### 6.3 风险因素分析

年龄在 2 型糖尿病预测模型中扮演至关重要的角色,且对动脉粥样硬化有显著影响。Wang X J 等<sup>[18]</sup>的一项全国性横向研究表明,动脉粥样硬化在中老年人中的发病率非常高,这与年龄有直接的正相关关系。年龄增大导致 2 型糖尿病并发症的风险增大,年龄增大为 2 型糖尿病并发症的非保护性因素。

目前通用观点认为,体重指数作为衡量肥胖程度的关键指数,是加速动脉粥样硬化的危险因素<sup>[19]</sup>。肥胖者通常伴随着脂质代谢紊乱,如高胆固醇、高甘油三酯、低高密度脂蛋白、胆固醇水平等。这些异常脂质代谢会增加动脉内脂质的沉积,促进动脉粥样硬化的形成<sup>[20]</sup>。

糖化血红蛋白 A1c 能够反映过去 3 个月血糖水平,长期高血糖水平已被证明会带来心血管疾病的风险。糖化血红蛋白 A1c 可能是确定 2 型糖尿病患者心血管疾病严重程度的标志物,其水平是 2 型糖尿病微血管并发症的重要危险因素指标,较高的糖化血红蛋白 A1c 也是动脉粥样硬化患者的潜在指标。高血糖时,氧自由基产生增多,活性增强,诱导血管内皮损伤,进而造成动脉粥样硬化<sup>[21]</sup>。

高血压是动脉粥样硬化的已知危险因素之一,高血压增加了动脉壁上的压力和剪切力,使低密度脂蛋白更容易渗透到动脉壁内,在动脉壁内沉积并形成斑块,加速动脉粥样硬化发展。

$\gamma$  谷氨酰基转移酶异常是 2 型糖尿病并发症保护性因素。因为  $\gamma$  谷氨酰基转移酶升高是微血管损

伤的早期表现,体内代谢产生的氧自由基活性对细胞损伤后会引发  $\gamma$  谷氨酰基转移酶的升高,可促进超氧化物的生成。较高水平的  $\gamma$  谷氨酰基转移酶可能意味着更强的抗氧化能力,有助于减轻炎症反应和血管内皮功能损伤,降低 2 型糖尿病并发症的发生风险。

碱性磷酸酶是一种催化血管钙化抑制剂无机焦磷酸盐水解的酶,广泛表达于人体各组织中。血清中碱性磷酸酶水平升高会催化焦磷酸盐的水解,降低焦磷酸水平,加速血管钙化,导致动脉粥样硬化<sup>[22]</sup>。

## 7 结语

糖尿病目前为全球常见且无法完全治愈的疾病之一。2 型糖尿病预测模型可以利用患者的生理、生化和临床数据,帮助医疗专业人员早期发现糖尿病风险,提供及时的干预和治疗,以减缓病情发展。

本研究通过比较分析得出朴素贝叶斯模型比较适合用于临床应用研究,其具有简单高效的特点,训练速度快,计算成本低。若需要处理大规模数据集或频繁地模型训练和调整,朴素贝叶斯模型是一个较为合适的选择。如用于糖尿病视网膜病变预测,基于患者临床数据、生化指标和医疗影像等信息,预测患者是否存在糖尿病视网膜病变的风险;还可以利用患者的年龄、糖尿病病史、血糖水平、血压、眼科检查结果等特征进行训练并预测患者的视网膜病变风险,以及糖尿病肾病、糖尿病足病变、心血管疾病。

本研究存在一定的局限性。一是所提出的 2 型糖尿病并发动脉粥样硬化风险预测模型可以对不同类型的并发症提出预测,但并不意味其适用于所有糖尿病并发症。二是构建预测模型时,过拟合和欠拟合是常见的问题。可能导致模型在实际应用中的性能下降,预测方法未实际应用于不同测试集,对实际应用效果有一定影响;数据量较小,对模型比较结果有一定影响。三是大多数 2 型糖尿病预测模型是基于人群统计数据构建的,可能无法针对个体

的特定情况个性化预测,在一定程度上限制了模型的实用性和准确性。在未来研究中,可以考虑加大研究样本,建立更多数量的测试集测试模型的效能,并进行多中心研究,充分考虑模型应用中的差异性和复杂性,确保研究的科学性和可信度。糖尿病相关课题研究是全球性问题,将全世界糖尿病患者作为数据集预测是未来的研究方向,数据量的增加和并发症影响指标的扩大,有利于临床诊断中的实际需求,可更好地为人类糖尿病预防提供支持。

**作者贡献:** 王一凡负责数据分析、论文初稿撰写;石超君、马晓洁、冯文佳负责研究设计;安洪庆、高倩倩、井淇、蔡伟芹、马安宁负责论文修订。

**利益声明:** 所有作者均声明不存在利益冲突。

## 参考文献

- BRANIMIR L, AMEEN A H, MARIJA S, et al. Predicting complications of diabetes mellitus using advanced machine learning algorithms [J]. *Journal of the American medical informatics association*, 2020, 27 (9): 1343 - 1351.
- DUH E J, SUN J K, STITT A W, et al. Diabetic retinopathy: current understanding, mechanisms, and treatment strategies [J]. *JCI insight*, 2017, 2 (14): e93751.
- 胡玲, 余芳雪, 贾曙光, 等. 成都市 2 型糖尿病及其并发症住院费用影响因素分析 [J]. *现代预防医学*, 2022, 49 (11): 2012 - 2018, 2030.
- 徐冰冰, 李肖静, 张亚伟. 2 型糖尿病患者糖尿病慢性并发症现状调查及影响因素 [J]. *华南预防医学*, 2021, 47 (1): 74 - 76.
- BILOUS R. Microvascular disease: what does the UKPDS tell us about diabetic nephropathy [J]. *Diabetic medicine: a journal of the British diabetic association*, 2008, 25 (S1): 25 - 29.
- STEVENS R J, KOTHARI V, ADLER A I, et al. The UKPD Srisk engine: a model for the risk of coronary heart disease in type II diabetes (UKPDS56) [J]. *Clinical science*, 2001, 101 (6): 671 - 679.
- CLARKE P M, GRAY A M, BRIGGS A, et al. A model to estimate the lifetime health outcomes of patients with type2 diabetes: the United Kingdom prospective diabetes study (UKPDS) outcomes model (UKPDS no. 68) [J]. *Diabe-*

- tologia, 2004, 47 (10): 1747 - 1759.
- 车宏鑫, 王桐, 王伟, 等. 前列腺癌预测模型对比研究 [J]. *数据分析与知识发现*, 2021, 5 (9): 107 - 114.
- 中国人民解放军总医院. 糖尿病并发症预警数据集 [EB/OL]. [2023 - 04 - 30]. <https://www.ncmi.cn/phda/dataDetails.do?id=CSTR:A0006.11.A0005.201905.000282-V1.0>.
- 路敦利, 宁芊, 臧军. 基于 BP 神经网络决策的 KNN 改进算法 [J]. *计算机应用*, 2017, 37 (S2): 65 - 67, 88.
- 周志华. 机器学习 [M]. 北京: 清华大学出版社, 2016.
- 赖春廷. 决策树分类算法研究 [J]. *信息与电脑 (理论版)*, 2020, 32 (14): 59 - 62.
- 陈志泊. 数据仓库与数据挖掘 [M]. 北京: 清华大学出版社, 2009.
- 韦哲, 于金玉, 曹彤, 等. 基于机器学习的糖尿病并发症预测模型研究进展 [J]. *中国医学装备*, 2022, 19 (2): 14 - 17.
- 赵存秀. 基于混淆矩阵的分类器性能评价指标比较 [J]. *电子技术与软件工程*, 2020 (13): 146 - 147.
- 谢永康, 丁梦清, 徐啸, 等. 基于 MLP 神经网络算法的中医肥胖体质分类模型研究 [J]. *无线互联科技*, 2021, 18 (7): 37 - 40.
- 张涵, 薛质, 施勇, 等. 基于多层神经网络的 Webshell 改进检测方法研究 [J]. *通信技术*, 2019, 52 (1): 179 - 183.
- WANG X J, LI W Z, SONG F J, et al. Carotid atherosclerosis detected by ultrasonography: a national cross-sectional study [J]. *Journal of the American heart association*, 2018, 7 (8): e008701.
- 芦晶晶. 2 型糖尿病患者冠状动脉粥样硬化性病变的危险因素分析 [D]. 石家庄: 河北医科大学, 2020.
- LAWRENCE A, LEITE R, STEPHEN C, et al. Cardiovascular risk reduction with once - weekly semaglutide in subjects with type 2 diabetes: a post hoc analysis of gender, age, and baseline CV risk profile in the SUSTAIN 6 trial [J]. *Cardiovascular diabetology*, 2019, (18): 1 - 12.
- 杨柳. 基于医疗大数据及人工智能的 2 型糖尿病患者颈动脉粥样硬化的预测模型研究 [D]. 大连: 大连医科大学, 2022.
- ZONH L X, WANG X W, LI Z X, et al. Alkaline phosphatase and outcomes in patients with preserved renal function: results from China national stroke registry [J]. *Stroke*, 2018, 49 (5): 1176 - 1182.