

医学领域中的机器学习：潜力、挑战与可解释性综述

马亚楠¹, 胡昔奇¹, 宋培培², 唐子颢³

¹中南大学湘雅医学院海口医院, 海南省海口市 570208; ²国立国际医疗研究中心临床研究中心, 日本东京 162-8655; ³东京大学附属医院肝胆胰外科, 日本东京 113-8655

摘要: 机器学习 (machine learning, ML) 在医学领域取得显著进展, 尤其是深度学习在大规模数据分析中的卓越表现, 在疾病预测、诊断及个性化治疗方面展现巨大潜力。然而, 模型复杂性带来的可解释性不足, 限制了其在临床实践中的信任度。本文综述了ML在术后并发症预测、糖尿病与心血管疾病早期预测、术后预后评估及神经退行性疾病研究中的应用, 并介绍了逻辑回归、决策树、随机森林、梯度提升、支持向量机与深度学习等常用算法。此外, 探讨了可解释人工智能的发展及SHAP、LIME等工具如何提高模型透明度, 推动ML在医学中的实际应用, 有助于改善诊断准确性与患者护理质量。

关键词: 特征重要性, 可解释人工智能, 模型透明度, 数据可视化, 预测准确性

Machine learning in the medical field: A review of potential, challenges, and interpretability

Ya-nan Ma¹, Xiqi Hu¹, Peipei Song², Kenji Karako³

¹Department of Neurosurgery, Haikou Affiliated Hospital of Central South University Xiangya School of Medicine, Haikou 570208, China; ²National Center for Global Health and Medicine, Tokyo 162-8655, Japan; ³Hepato-Biliary-Pancreatic Surgery Division, Department of Surgery, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan

Abstract: Machine learning (ML) has made significant progress in the medical field, particularly with the outstanding performance of deep learning in large-scale data analysis, demonstrating great potential in disease prediction, diagnosis, and personalized treatment. However, the complexity of these models has led to limited interpretability, which affects trust in clinical practice. This paper reviews the applications of ML in postoperative complication prediction, early detection of diabetes and cardiovascular diseases, postoperative prognosis assessment, and neurodegenerative disease research. It also introduces commonly used algorithms such as logistic regression, decision trees, random forests, gradient boosting, support vector machines, and deep learning. Furthermore, the development of explainable artificial intelligence (XAI) is discussed, highlighting tools like SHAP and LIME that enhance model transparency, thus promoting the practical application of ML in medicine and contributing to improved diagnostic accuracy and patient care quality.

Keywords: feature importance, explainable AI, model transparency, data visualization, prediction accuracy

1. 引言

近年来, 机器学习技术, 特别是深度学习 (deep learning) 取得了快速的发展。卷积神经网络 (convolutional neural networks, CNN) 和Transformer等技术的出现, 专为图像识别与自然语言处理而设计, 使得即使是复杂问题也能实现高性能预测⁽¹⁻⁴⁾。尽管“深度学习”和

“人工智能 (AI)”等术语近来备受关注, 但它们实际上属于更广泛的机器学习范畴。机器学习是一种通过数据驱动的技术, 算法使用数据发现并学习其中的模式和特征, 并根据学习到的结果进行预测或分类。

机器学习的一个关键特点是: 算法能够自主从数据及其对应的结果中识别模式, 而无需人为定义预测规则 (例如, “如果测量值大于1, 则分类为A, 否则分类为B”)。通过在大型数据集中寻找规律, 机器学习能够实现精确的预测。随着机器学习技术的不断进步, 其应用已扩展至多个领域, 包括医学研究, 在该领域内, 基于机器学习的研究正在积极开展。

在医学领域, 机器学习展现了巨大潜力。它已被用于根据患者的测量数据预测术后结果以及评估疾病风险等⁽⁵⁻⁷⁾。基于机器学习的预测模型有望显著帮助医疗专业人员进行决策并改进患者治疗。然而, 尽管机器学习在医学中

收稿日期: 2024-11-15; 修回日期: 2024-12-20

基金项目: 海南省博士后科研项目/日本科学研究费助成 (24K14216)

通讯作者/Corresponding author: 唐子颢/Kenji Karako, E-mail: tri.leafs@gmail.com

的潜力巨大，却面临一个重要的挑战，即“黑箱”问题。所谓“黑箱”问题，指的是机器学习模型生成的预测结果及其过程对人类而言难以理解。随着机器学习算法的复杂性增加，其行为在宏观层面上变得更加难以解释，尽管在微观层面上某些方面仍然可以被理解。这种复杂性导致了预测结果的产生原因或其背后的逻辑难以明确，从而限制了机器学习在医学领域的应用接受度。由于预测过程不透明，医生和患者可能不愿意信任这些预测结果。

近年来，为了解决“黑箱”问题，一种名为可解释人工智能（Explainable AI, XAI）的技术逐渐受到关注^(8,9)。XAI旨在通过分析机器学习模型，解释预测结果的生成过程，识别预测趋势，并提供预测结果的解释。XAI能够以人类可理解的方式展示不同特征的重要性，从而揭示影响算法结果的因素。这有望使机器学习在医学领域更容易被接受。

本文首先展示了机器学习在医学中的具体应用实例。接下来，描述了常用的机器学习算法的原理与机制。最后，本文将深入探讨可解释的机器学习技术，作为解决“黑箱”问题的方案。通过本次探讨，我们旨在分享机器学习在医疗保健领域的潜在应用见解。

2. 医学领域中的机器学习应用

随着近年来人工智能和深度学习技术的快速发展，利用机器学习进行疾病预测、诊断和预后预测的研究在医学领域得到了广泛开展。这些模型被期望能够分析复杂的患者数据，用于预测并发症、恢复结果，并在治疗策略决策中提供支持。以下将介绍几种典型的医学领域机器学习应用，这些应用有望为医学决策提供重要帮助。

2.1. 术后并发症预测

机器学习在预测术后并发症风险方面非常有效⁽⁵⁾。例如，已有模型被提出用于评估严重术后并发症的风险，如肺炎、急性肾损伤、深静脉血栓形成和肺栓塞。研究基于111,888例手术数据（包括患者特征和临床信息），使用了五种不同的机器学习算法进行术后并发症风险预测的比较，包括逻辑回归（logistic regression）、支持向量机（Support Vector Machine, SVM）、随机森林（random forest）、梯度提升（gradient boosting）和深度神经网络（deep neural networks）⁽¹⁰⁻¹³⁾。结果表明，结合术前和术中数据的模型能提供最高的预测准确性，突出了机器学习在术后风险管理中的有效性。

2.2. 糖尿病和心血管疾病的早期预测

机器学习模型也被用于有效预测糖尿病和心血管疾病风险⁽⁷⁾。这些模型整合了多种数据，如家族病史、年龄、体重、血压、胆固醇水平以及生活习惯（如吸烟与运动情况），以评估疾病风险。研究采用了适用于线性关系的算法（如线性回归和支持向量机），以及能够处理非线性关系的算法（如随机森林和梯度提升），从而实现了高精度的风险预测。

2.3. 阿尔茨海默病的诊疗

结合了机器学习技术的Cordax回归模型能够深入解析假定的淀粉样蛋白原纤维结构，尤其是在易聚集区域相关的结构方面，提供了丰富且互补的结构洞察⁽¹⁴⁾。这些洞察对于理解与淀粉样蛋白相关的疾病及功能性疾病具有重要意义。

通过ML方法，研究人员已鉴定出*BDHI*、*TRAP1*、*OPAI*、*DLD*和*SPG7*等基因，作为阿尔茨海默病线粒体基因的潜在分子治疗靶点。这些基因在免疫微环境的相互作用中起着关键作用，为探索阿尔茨海默病的分子机制和治疗策略提供了新的研究方向⁽¹⁵⁾。

此外，AI和ML技术在阿尔茨海默病治疗化合物的筛选方面也展现出巨大潜力。基于计算分析，研究发现了一类通过ML鉴定的新型组蛋白去乙酰化酶抑制剂，为阿尔茨海默病的治疗研究提供了新的候选药物⁽¹⁶⁾。

采用无监督机器学习（包括筛选与验证方法），研究人员鉴定出两种可诱导神经元线粒体自噬的化合物，有助于清除阿尔茨海默病相关的淀粉样蛋白- β （A β ）和tau病变。这一发现不仅加深了对阿尔茨海默病病因的理解，还为探索潜在的生物标志物及治疗策略提供了新视角⁽¹⁷⁾。

2.4. 术后预后预测

机器学习还用于预测术后预后情况。一项研究旨在预测颈前路椎间盘切除融合术后的四种短期不良事件：住院时间延长、出院至非家庭环境、术后30天内再入院及严重并发症⁽⁶⁾。研究使用了五种机器学习算法构建模型：TabPFN、TabNET、XGBoost、LightGBM和随机森林⁽¹⁸⁻²¹⁾。结果表明，在这五种算法中，随机森林表现最佳，其受试者工作特征曲线下面积范围为0.776到0.846。预测术后不良事件风险有助于为每位患者早期实施个性化干预，从而更好地控制潜在病情恶化。如上所述，已有大量基于临床数据的预测性研究被开展。传统的线性回归和逻辑回归等技术仅限于处理线性问题。然而，随着机器学习的进步及各种算法学习与预测性能的提升，这些模型如今能够应用于更加复杂的问题。

3. 近年来常用的代表性机器学习算法

在机器学习领域，已开发出多种算法并被广泛应用。其中，一些具有代表性的基础方法可用于分类与预测任务，包括逻辑回归、决策树（decision trees）、随机森林、梯度提升、支持向量机和深度学习⁽²²⁾。逻辑回归和决策树结构简单且易于解释，但其预测准确性相对较低。相比之下，梯度提升和深度学习等算法尽管预测性能优越，但更难以解释其内部工作原理。

3.1. 逻辑回归

逻辑回归是一种在医学领域常用的基础机器学习算法之一，特别适用于二分类任务⁽¹⁰⁾。其工作原理类似于线性回归，但专门用于处理分类问题。逻辑回归通过对输入解释变量进行加权线性组合，并将结果传递到一个Sigmoid（S型）函数，从而输出一个介于0和1之间的概率值。这些加权系数是根据训练数据计算得到的参数，有效地确定这些参数能够实现预测任务。这一过程被称为“学习”。

简而言之，“学习”指的是找到一个能够准确描述解释变量与目标变量之间关系的函数参数集。该学习过程通常通过优化技术实现。

在优化过程中，会定义一个损失函数（Loss function），用于表示需要最小化的目标，算法通过调整参数以使该损失函数最小化。在逻辑回归中，目标是最大化对数似然函数，但通常会将其转换成最小化损失函数的形式。通过这种优化过程，逻辑回归能够找到最符合训练数据特征的参数，从而实现二分类任务的预测。

然而，逻辑回归假设问题是线性可分的，因此当解释变量与目标变量之间存在非线性关系时，其表现可能不佳。尽管如此，由于逻辑回归的简单性及其通过每个解释变量的权重提供的可解释性，使其在医学领域被广泛使用。

3.2. 决策树

与逻辑回归类似，决策树也是一种直观且易于解释的机器学习算法⁽²²⁾。决策树通过递归地根据特定规则对输入数据进行划分，从而对数据进行分类。这些划分形成的结构类似于树，因此被称为“决策树”。

在该树形结构中，决策路径是从树的顶部根据输入数据的值依次向下进行的。例如，如果天气是晴天，模型会沿着树的右侧路径继续，而如果天气为多云或下雨，则沿着左侧路径进行判断。这一过程会持续进行，直到模型最终预测当天温度是否会超过22°C。

构建能够准确表示数据的树形结构至关重要。为此，已提出了多种算法，包括ID3、CART和C4.5等⁽²²⁻²⁴⁾。决策树的一大优点是其树形结构清晰地展示了用于做出预测的标准及使用的特征，从而使模型非常容易解释。与逻辑回归不同，决策树可用于解决非线性问题。然而，决策树容易发生过度拟合，即模型在训练数据上的表现过于优秀，但在未见过的数据上表现不佳。

3.3. 随机森林

随机森林是一种通过集成多个决策树以提升预测准确性的算法⁽¹²⁾。尽管单个决策树容易过拟合且预测性能有限，但随机森林通过生成多个决策树并对其预测结果进行集成，从而提高了整体的预测性能。术语“森林”指的是由多个决策树组成的集合。这种将多个弱预测器（即多个决策树）组合起来以提升整体性能的技术被称为集成学习（ensemble learning）。

随机森林的运行主要包括三个步骤：自助采样（bootstrap sampling）、决策树构建以及预测结果集成。

①自助采样：训练数据被随机分成多个子数据集，每个数据集用于训练不同的决策树；

②决策树构建：对每个子数据集，使用随机选择的输入特征构建决策树。通过随机选择特征，确保了每棵决策树学习到不同的特征组合，从而增加了模型的多样性，减少了过拟合风险；

③预测结果集成：对于分类任务，通过多数投票，而对于回归任务，则通过平均值整合所有决策树的预测结果，从而生成最终的预测结果。

与单独的决策树相比，随机森林通常具有更高的预测

精度，并且较少出现过拟合现象。然而，这种方法的一个缺点是模型的可解释性降低。虽然单个决策树可以很容易理解，但当多个树组合在一起时，变量之间如何共同作用生成最终预测结果会变得更加难以解释。

3.4. 梯度提升

梯度提升是一种集成学习方法，它通过组合多个弱预测器（通常是决策树）来构建一个强大的模型⁽¹³⁾。与随机森林通过多个决策树的独立预测进行集成不同，梯度提升采用了一种逐步改进的策略，即通过序列化地创建决策树，每棵新树的目标是修正前一棵树的预测误差。

梯度提升的过程从创建一个初始决策树开始，这个初始树通常会生成较大的预测误差，即预测值与真实数据之间的偏差。为了改进模型，计算预测结果与真实值之间的误差，并使用这些误差作为训练目标，训练一个新的决策树来预测这些误差。随后，将初始树的预测结果与新训练的修正误差的决策树结果进行组合，从而提高整体模型的性能。

这种“误差修正”的过程会反复进行，使得模型能够不断自我优化，并通过每次迭代逐步减少预测误差，最终实现较高的预测准确性。常见的梯度提升实现算法包括XGBoost和LightGBM，这些算法在性能和计算效率方面经过了高度优化，适用于大规模数据集的训练^(20,21)。

与单独的决策树相比，梯度提升通常能够生成更为精确的模型，并且相比于随机森林，更不容易发生过拟合。然而，与随机森林类似，由于组合了多个决策树，该方法也存在可解释性较差的问题，难以清楚地解释模型生成预测结果的具体原因。

3.5. 支持向量机

支持向量机是一种用于分类和回归问题的强大机器学习算法⁽¹¹⁾。它通过寻找一个最优的边界（超平面），将训练数据中的不同类别进行分隔，并利用该边界进行预测。最初，SVM被设计用于解决线性可分问题，但随后得到了改进，可以处理非线性问题。对于线性问题，SVM的解释性较强，但在处理非线性问题时，由于引入了核函数（Kernel），模型的可解释性变得更加复杂。

3.6. 深度学习

深度学习是一种模拟大脑神经元行为的机器学习模型，其使用人工神经元作为数学模型进行计算⁽¹⁾。一个人工神经元接收解释变量的输入，对输入数据进行加权线性组合，然后将结果传递给一个激活函数（activation function），激活函数的输出即为神经元的响应值。

如果使用Sigmoid函数作为激活函数，那么该过程与逻辑回归几乎相同。然而，在深度学习中，多个人工神经元使用相同的输入变量进行计算，并被组合成神经网络的层。一个神经网络层的输出会作为下一层的输入，通过多层连接形成完整的神经网络。这种多层结构使深度学习能够学习到更加复杂的特征关系。

每个人工神经元在计算中都包含权重参数，通过调整这些权重，神经网络能够实现更强的预测能力。深度学习

的训练过程涉及以下几个关键步骤:

- ①参数初始化: 初始时, 权重参数被随机分配;
- ②前向传播: 根据输入数据计算神经网络的输出;
- ③计算误差: 计算预测结果与真实标签之间的误差;
- ④反向传播: 根据误差对权重进行调整, 使用梯度下降算法逐步减少误差;
- ⑤迭代训练: 重复进行多轮训练, 直到模型能够在新输入数据上进行准确预测。

从概念上讲, 深度学习可以被视为由多个连接的逻辑回归模型组成的复杂网络。

在过去十年中, 深度学习被广泛研究, 并取得了诸多改进和创新的网络结构, 与传统的机器学习模型相比, 深度学习展现了更高的性能。特别是用于图像识别的卷积神经网络和用于自然语言处理及时间序列分析的Transformer模型已成为研究的热点, 并在多个领域取得了显著的成果⁽³⁾。

4. 可解释的机器学习

到目前为止, 我们已经介绍了几种代表性的机器学习算法。尽管每种算法都有其优势和劣势, 但它们都展现了较高的性能。然而, 在医学领域应用这些算法时, 面临着一个显著的挑战: 可解释性。诸如逻辑回归和决策树等算法较为简单, 使其预测结果相对容易解释。然而, 近年来发展出的更为先进且性能卓越的算法, 如梯度提升和深度学习, 由于模型更加复杂, 难以解释其预测结果, 从而引发了“黑箱”问题。

尽管可以在微观层面理解模型的个别计算过程, 但从整体上解释整个模型的决策过程却十分困难。这种“黑箱”问题在医学等领域尤为突出, 因为医学诊断需要清晰的推理过程和明确的解释, 以确保医生和患者对预测结果的信任。

为了解决“黑箱”问题, 当前的研究正在致力于开发能够以人类可理解的方式解释模型内部结构和决策过程的技术。这些技术被统称为可解释人工智能, 并且在这一领域, 已提出了多种方法^(8,9)。XAI主要从两个角度尝试解释机器学习模型的行为: 全局解释和局部解释。

4.1. 全局解释

全局解释旨在描述机器学习模型的整体特征和行为。一个机器学习模型通过学习训练数据, 获得用于执行预测任务的参数和结构。通过分析模型在做出预测时强调的解释变量, 可以使用一种称为特征重要性的技术来计算并评估哪些变量对模型最为重要。

另一种全局解释方法是构建一个更容易解释的简化模型, 用于理解更复杂模型的行为。例如, 可以使用决策树或逻辑回归这样的简单可解释模型来近似深度学习模型的行为。从外部观察, 深度学习模型的作用就像一个黑箱预测器, 对输入数据生成输出结果。通过收集深度学习模型在不同输入下的输出数据, 并使用这些数据训练一个决策树, 该决策树可以近似地反映深度学习模型的行为。随后, 可以可视化该决策树, 从而解释特定输入变量如何影响模型的预测结果。

这种全局解释方法有助于识别重要的解释变量, 并用于改

进模型。此外, 如果模型提供的解释与现有研究一致, 还可以增强模型的科学可信度与有效性。

4.2. 局部解释

与全局解释不同, 局部解释关注的是针对特定输入数据模型的单次预测结果进行解释。例如, 如果一个机器学习模型根据电子健康记录预测患者是否患有某种疾病, 医生可能难以理解模型为何做出“患者患病”或“患者未患病”的预测。局部解释可为这些个别案例提供具体解释。

目前, 有几种常用的局部解释方法, 其中两种常用的方法介绍如下:

识别主要影响因素: 该方法分析哪些特征在单次预测中贡献最大。例如, 如果一个模型预测某位患者患有糖尿病, XAI可能会指出患者的血糖水平和血红蛋白数值特别高, 从而解释这些特征是模型判断的关键因素。常用的技术包括:

Shapley加性解释 (Shapley Additive Explanations, SHAP): 通过计算各特征对预测结果的贡献程度, 以可视化的方式展示特征重要性⁽²⁵⁾。局部可解释模型无关解释 (Local Interpretable Model-agnostic Explanations, LIME): 通过对输入数据进行轻微扰动, 观察预测结果的变化, 以识别最关键的特征⁽²⁶⁾。

相似案例对比: 该方法通过查找训练数据中与当前样本最相似的历史案例, 解释模型的预测。例如, 如果模型预测患者患有糖尿病, XAI可能会在训练数据中查找相似情况的患者, 并提供解释: “该患者的特征与其他糖尿病确诊患者相似。”

影响函数 (influence functions): 用于计算训练数据中每个样本对特定预测的贡献程度, 从而识别对模型影响最大的训练样本⁽²⁷⁾。通过回顾这些关键样本, 可以更好地理解模型为何做出当前预测。此外, 该方法还能识别异常数据点 (如离群点), 以便改进模型的稳定性。

因此, XAI技术提供了全局解释和局部解释两种方法, 以提高机器学习模型的可解释性。尽管XAI仍处于不断发展的阶段, 但它正逐步奠定基础, 以提供合理的解释, 从而解决“黑箱”问题, 并促进机器学习在医学领域的实际应用。

5. 结论

特别是随着深度学习的出现, 机器学习在包括医学在内的多个领域展示了巨大的潜力。其能够分析大规模且复杂的数据集, 并做出高准确性的预测, 为疾病预测、诊断及辅助治疗方案规划提供了显著优势。然而, 机器学习在医学领域的应用仍面临着重要挑战, 尤其是在模型可解释性方面。

传统的模型, 如逻辑回归和决策树, 由于其结构相对简单, 因此其预测结果更容易解释。相比之下, 梯度提升、随机森林和深度学习等更先进的模型, 尽管预测精度更高, 但往往表现为“黑箱”, 即缺乏透明性。这种透明度的缺失在医学领域是一个重大障碍, 因为临床医生和患者需要理解预测背后的推理过程, 以便信任和接受这些结果。

为了解决这一问题, XAI技术的开发至关重要。XAI旨

在弥合现代机器学习模型的高性能与可理解性之间的差距。诸如SHAP、LIME和影响函数等XAI工具的出现,使得在医学领域使用机器学习能够获得更高的信任度。这些工具不仅提供了更高的透明性,还增强了模型的可靠性和科学有效性,有助于使预测结果与已建立的医学知识保持一致。

随着机器学习的不断发展,将这些可解释性技术集成到模型中将成为确保其在医疗保健中实现实际且符合伦理使用的关键。未来的医学可能越来越依赖机器学习,而可解释模型将成为提升诊断准确性和优化临床决策过程的重要工具。通过这些技术进步,机器学习有望显著改善患者护理,推动个性化治疗方案,并帮助医护人员做出更为科学的决策。利益冲突:所有作者均声明不存在利益冲突。

致谢:无。

作者贡献声明:无。

参考文献

1. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2016;521:436-444.
2. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proc IEEE*. 1998;86:2278-2324.
3. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM*. 2017;60:84-90.
4. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. In: 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.
5. Xue B, Li D, Lu C, et al. Use of machine learning to develop and evaluate models using preoperative and intraoperative data to identify risks of postoperative complications. *JAMA Netw Open*. 2021;4:e212240.
6. Karabacak M, Bhimani AD, Schupper AJ, et al. Machine learning models on a web application to predict short-term postoperative outcomes following anterior cervical discectomy and fusion. *BMC Musculoskelet Disord*. 2024;25:401.
7. Dinh A, Miertschin S, Young A, et al. A datadriven approach to predicting diabetes and cardiovascular disease with machine learning. *BMC Med Inform Decis Mak*. 2019;19:211.
8. Burkart N, Huber MF. A survey on the explainability of supervised machine learning. *J Artificial Intel Res*. 2021;70:245-317.
9. Guidotti R, Monreale A, Ruggieri S, et al. A survey of methods for explaining black box models. *ACM Comput Surv*. 2018;51:Article 93.
10. Cox D. The regression analysis of binary sequences (with discussion). *J Roy Stat Soc*. 1958;20:215-242.
11. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20:273-279.
12. Breiman L. Random forests. *Mach Learn*. 2001;45:5-32.
13. Hastie T, Tibshirani R, Friedman J. Boosting and additive trees. In: *The elements of statistical learning*. Springer; New York, NY. 2009.
14. Louros N, Orlando G, De Vleeschouwer M, et al. Structure-based machine-guided mapping of amyloid sequence space reveals uncharted sequence clusters with higher solubilities. *Nat Commun*. 2020;11:3314.
15. Zhang Y, Miao Y, Tan J, et al. Identification of mitochondrial related signature associated with immune microenvironment in Alzheimer's disease. *J Transl Med*. 2023;21:458.
16. Gupta R, Ambasta RK, Kumar P. Identification of novel class I and class IIb histone deacetylase inhibitor for Alzheimer's disease therapeutics. *Life Sci*. 2020;256:117912.
17. Xie C, Zhuang XX, Niu Z, et al. Amelioration of Alzheimer's disease pathology by mitophagy inducers identified via machine learning and a cross-species workflow. *Nat Biomed Eng*. 2022;6:76-93.
18. Hollmann N, Müller SG, Eggensperger K, et al. TabPFN: A transformer that solves small tabular classification problems in a second. In: *Proc Internatl Conf Learning Representations*. 2022.
19. Arık S, Pfister T. TabNet: Attentive interpretable tabular learning. In: *Proc 35th AAAI Conf Artificial Intel*. 2021; 35: AAAI-21 Technical Tracks 8.
20. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016;785-794.
21. Ke G, Meng Q, Finley T, et al. LightGBM: A highly efficient gradient boosting decision tree. In: 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.
22. Quinlan JR. Induction of decision trees. *Mach Learn*. 1986; 1:81-106.
23. Breiman L, Friedman JH, Olshen RA, et al. Classification and regression trees. Chapman and Hall; 2017; pages 1-368.
24. Salzberg SL. Book Review: C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Mach Learn*. 1994; 16:235-240.
25. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: 31st conference on neural information processing systems (NIPS 2017), Long Beach, CA, USA.
26. Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?" Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 2016;1135-1144.
27. Koh PW, Liang P. Understanding black-box predictions via influence functions. In: *Proceedings of the 34th International Conference on Machine Learning*, 2017;PMLR 70:1885-1894.

引用本文 / Article Citation:

马亚楠, 胡昔奇, 宋培培, 唐子颢. 医学领域中的机器学习: 潜力、挑战与可解释性综述. *医学新视角*. 2024;1(6):291-295. doi:10.5582/npjm.2024.01089

Ya-nan Ma, Xiqi Hu, Peipei Song, Kenji Karako. Machine learning in the medical field: A review of potential, challenges, and interpretability. *The New Perspectives Journal of Medicine*. 2024;1(4):291-295. doi:10.5582/npjm.2024.01089