

XGBoost-SHAP 机器学习可解释框架用于轻度认知障碍分类研究*

易付良¹ 陈杜荣¹ 杨慧¹ 秦瑶¹ 韩红娟¹ 崔靖¹ 白文琳¹
马艺菲¹ 张荣¹ 余红梅^{1,2Δ}

【摘要】 目的 利用机器学习算法对轻度认知障碍(mild cognitive impairment, MCI)亚型分类有利于患者的个性化治疗,而复杂模型常因分类过程的内部机制不可洞察而饱受诟病,本研究借助可解释技术梳理模型的输出结果,以期对相关领域研究者的决策提供统计支持。**方法** 本研究联合极限梯度提升(eXtreme Gradient Boosting, XGBoost)与沙普利可加性(SHapley Additive exPlanations, SHAP)构建可解释性框架,用于遗忘型 MCI(amnestic MCI, aMCI)和非遗忘型 MCI(non-amnestic naMCI)的分类,并根据联合框架输出结果进行解读。**结果** 联合框架输出 aMCI 和 naMCI 的最佳分界值为 0.51, XGBoost 分类准确率、灵敏度、特异度、F1 值、AUC 分别为 92.81%、94.94%、90.54%、0.93、0.96。SHAP 个性化预测结果,7 号和 31 号示例个体被预测为 aMCI 的概率分别为 0.27 和 0.91;全局性解释结果,不同个体随着 CERALL、ORIENT、CDRSUM、LCMF、RSUPMAR、RMEDORBF、LPOSCENM 等指标的 shapley 值增大,患 aMCI 的风险越大,上述指标可以解释为 aMCI 发生的危险因素,而 RENT、MMSEORDA、CRAFTVRS 等则相反,可以解释为 aMCI 发生的保护因素。**结论** XGBoost-SHAP 联合框架用于 MCI 亚型分类效果较为理想,实现了特定个体不同特征预测效果的比较、不同个体给定特征预测能力的判断,为相关研究者打开了洞察复杂模型内在机制的大门。

【关键词】 机器学习 可解释性 极限梯度提升-沙普利可加性 轻度认知障碍 分类

【中图分类号】 R195 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2024.03.019

轻度认知障碍(mild cognitive impairment, MCI)通常被认为是介于正常认知(normal cognitive, NC)与阿尔茨海默病(Alzheimer's disease, AD)的中间状态,是 AD 早期干预的关键时间窗^[1]。根据 Petersen^[2] 等人的划定标准,可以将轻度认知障碍分为遗忘型 MCI(amnestic mild cognitive impairment, aMCI)和非遗忘型 MCI(non-amnestic mild cognitive impairment, naMCI)。不同亚型 MCI 患者的转归并不相同, aMCI 患者发生 AD 的风险较大,2 年转化率达到 15%~25%,而 naMCI 患者容易进展为原发性失语、额颞叶痴呆和路易体痴呆等^[3-4]。临床上针对不同亚型的 MCI 患者的干预方案不尽相同,因此早期识别 MCI 患者亚型有利于采取个性化治疗与进展风险监测。

机器学习方法凭借其遵循非线性、非参数、无需预先限制模型复杂性以及优异的分类与预测能力而备受医务工作者青睐,在 AD 的分类及 MCI 亚型的识别方面应用也较为广泛^[5-7]。然而随着医疗数据的复杂化,高精度的分类结果往往需要复杂的模型来实现,例如我们可以通过集成学习和深度学习实现数据挖掘^[8-9]。另一方面,更加复杂的机器学习算法在带来更高预测精度的同时也增加了算法的学习难度,其内部机制变得难以洞察,模型高精度度和可解释性之间无法达到平衡,算法本身的“黑箱”由此形成^[10-11]。例如,极限梯度提升(eXtreme Gradient Boosting, XG-

Boost)模型作为集成算法中的佼佼者,在医学^[12]、金融^[13]、智能系统^[14]等诸多领域都取得了不错的预测成绩。然而由于机器学习固有的局限性,算法本身输出的全局性的解释结果如增益(Gain)并不令人信服^[15]。事实上,理解模型做出某种预测的缘由与预测的准确性同等重要:一个可解释的结果对辅助医生做出最终决策具有重大影响。

Doshi-velez^[16] 等人将可解释性定义为:以可理解的术语向人类解释或呈现的能力。模型的可解释性包括内在可解释性和事后可解释性两类^[17]。内在可解释性通过构建自解释模型,将可解释性直接融入到算法中,如线性回归将目标预测为输入特征的加权和,但是模型的复杂性受到限制;事后可解释性则是利用可解释性技术分析模型学习到的信息,如 Lundberg 等人提出沙普利可加性解释(SHapley Additive exPlanations, SHAP)利用可加性归因方法将模型的目标预测解释为每个输入特征的平均边际贡献与平均预测值之和^[18],从而量化各个特征对结局的影响。值得注意的是,事后解释方法因可以融合更复杂的模型且对预测不会造成影响而应用更为广泛。

因此,本研究拟联合 XGBoost 和 SHAP 构建可解释性框架以增强机器学习模型的可解释性,同时利用联合框架对 MCI 亚型进行分类以促进患者的个性化干预,并对联合框架输出的结果进行详细解读,以期对相关领域研究者的决策提供统计支持。

模型原理

1. XGBoost 原理

XGBoost 是陈天奇^[19] 等人提出的一种梯度提升算法,由多棵 CART 回归树串行化集成,基于前向分

* 基金项目:国家自然科学基金资助项目(81973154;82273742);山西省应用基础研究计划青年科学研究项目(202103021223242);山西省研究生创新项目(2022Y367)

1. 山西医科大学公共卫生学院卫生统计学教研室(030000)

2. 重大疾病风险评估山西省重点实验室

Δ通信作者:余红梅, E-mail:yu@sxmu.edu.cn

布算法,首先生成一棵树拟合原始数据得到一个预测输出,计算真实值 y_i 与第一轮迭代预测输出 $\hat{y}_i^{(1)}$ 的损失函数负梯度,并利用损失函数负梯度矩阵拟合第二棵树,如此反复迭代 t 次,学习到包含 t 棵树的加法模型,最后将 t 棵树对应的预测值求和得到 $\hat{y}_i^{(t)}$ 为最终预测。该学习过程对损失函数进行二阶求导得到负梯度,并引入正则化惩罚项来控制树的复杂度。

2. SHAP 原理

SHAP 是 Lundberg^[18, 20-21] 等人开发的一个可加性解释框架,其思想来源于博弈论。对于任意个体,预测模型都会输出一个预测值,而 SHAP 框架则假定每个特征都是目标预测的“贡献者”,并给它们分配一个重要值(shapley 值),给定个体的所有特征的 shapley 值累计与平均预测值之和即是目标预测。

(1) shapley 值的计算

对于指定的特征 x_j 的 shapley 值是该特征对预测的平均边际贡献——通过计算所有可能的组合情况下,包含和不包含特征 x_j 的预测差值来获得,即通过与所有可能的特征组合进行加权求和得到:

$$\Phi_j(val) = \sum_{S \subseteq \{x_1, \dots, x_p\} \setminus \{x_j\}} \frac{|S|! (p - |S| - 1)!}{p!} (val(S \cup \{x_j\}) - val(S)) \quad (1)$$

$$val(S) = \int \hat{f}(x_1, \dots, x_p) d\mathbb{P}_{x \in S} - E_x(\hat{f}(X)) \quad (2)$$

x 是要解释的个体的特征值, $S \subseteq \{x_1, \dots, x_p\} \setminus \{x_j\}$ 为不包括 $\{x_j\}$ 的所有特征的子集, p 为子集中的特征数量, $val(S \cup \{x_j\})$ 是包含 $\{x_j\}$ 的子集内所有特征的预测, $val(S)$ 是不包含 $\{x_j\}$ 的子集 S 所有特征的预测, $E_x(\hat{f}(X))$ 为所有特征的预测期望, 其中 $\frac{|S|! (p - |S| - 1)!}{p!}$ 为子集 S 的权重。对于子集 S 权重的理解: p 个特征在任意排序下有 $p!$ 种可能,确定子集 S 后, p 个特征在特定排序情况下有 $|S|! (p - |S| - 1)!$ 种可能,故 $\frac{|S|! (p - |S| - 1)!}{p!}$ 为子集 S 的特征组合可能占比。

这样计算得到的 shapley 值满足效率性 (efficiency)、对称性 (symmetry)、虚拟性 (dummy) 和可加性 (additivity)。这些特性保证了 SHAP 将 shapley 值视作一种可加性归因方法,并可以将模型的预测解释为每个输入特征的归因值之和。

(2) Tree SHAP

接口 XGBoost 的是 Tree SHAP,它是 SHAP 的一个变种,适合基于树的机器学习^[22-23]。Tree SHAP 引入条件期望 $E_{x_S|x_C}(\hat{f}(x) | x_S)$ 替代平均边际贡献定义值函数,并将 SHAP 值计算的复杂度从 $O(TL^M)$ 降低到了 $O(TL^D)$, T 是树的数量, L 是任意数的最大叶子数, D 是所有树的最大深度。

联合框架

1. 资料来源

数据来源于 NACC (National Alzheimer's Coordinating Center) 数据库 (<https://naccdata.org/>)。研究对象为 2005 年 11 月-2018 年 4 月首次确诊的 MCI 患者,包括 aMCI 和 naMCI。根据精神疾病统计诊断手册第 5 版 (DSM-5) 对 MCI 的定义:由患者、知情者主诉或临床医生观察到的认知领域改变;和本人过往相比,出现一个或多个认知领域的认知障碍;独立活动能力保留;没有痴呆^[24]。而如果患者认知的情景记忆表现较差则为 aMCI,如果除记忆以外的认知领域(如执行功能、语言或视觉空间能力)表现较差,则为 naMCI^[25]。特征包含患者的基本信息、神经心理测试以及结构核磁共振 (structural magnetic resonance imaging, sMRI)。

2. 数据清洗

由于获取的资料为结构化数据,故不需要对 sMRI 进行图像处理。首先将患者 ID 编号与其基本信息、神经心理测试与 sMRI 进行匹配;然后采用 missForest 算法对匹配后的数据进行缺失填补,得到 510 例样本 (aMCI 270 例,naMCI 240 例);再采用弹性网络算法进行特征选择,最后得到 161 个特征。部分特征中文注释见表 1。

表 1 部分特征中文注释

特征	中文注释	特征	中文注释
CEREALL	颅内总容量	LSUPMAR	左上缘灰质体积(cc)
ORIENT	临床痴呆评分-定向	LPOSCENM	左中央后平均皮质厚度(mm)
CDRSUM	临床痴呆得分总分	MMSEORDA	MMSE 定向分量表得分-时间
LCMF	左尾中额叶灰质体积(cc)	CRAFTVRS	延迟回忆-回忆故事总数
RSUPMAR	右上缘灰质体积(cc)	MEMUNITS	逻辑记忆-回忆故事总数
RENT	右内嗅灰质体积(cc)	RPARCEN	右中央旁灰质体积(cc)
RMEDORBF	右眶额内侧灰质体积(cc)	LPOSCENM	左中央后平均皮质厚度(mm)

3. 联合框架构建

对清洗后数据构建 XGBoost 模型进行预测,输出

目标的预测值,并基于最佳截断值将连续的预测值二值化得到分类结果,模型评价指标包括灵敏度、特异

度、F1 值、ROC 曲线下面积(AUC)。然后将预测模型接口 SHAP 框架,以输出模型解释性结果。

结果

1. 分类结果与可解释性

模型输出 aMCI 和 naMCI 的最佳分界值为 0.51,二值化预测值后模型分类的准确率、灵敏度、特异度、F1 值、AUC 分别为 92.81%、94.94%、90.54%、0.93、0.96,模型的分类效果较好。接口 SHAP 框架后输出个性化解释图(解释力图、局部条形图、瀑布图)以及全局解释图(特征重要度柱状图、摘要图、依赖图)。

2. 个性化解释图

(1)SHAP 解释力图

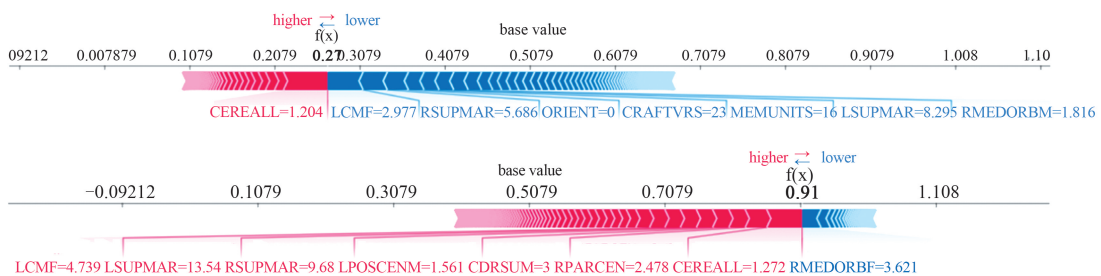


图 1 预测 aMCI 患者 SHAP 解释示例

(2)SHAP 局部条形图

解释力图可以直观地显示特定个体的预测值和影响个体患病的因素,但不能输出每个特征的平均边际贡献(shapley 值),而局部条形图很好地弥补了这一缺陷。图 2 展示了 7 号和 31 号个体的前 10 个特征以及剩余 151 个特征的 shapley 值。

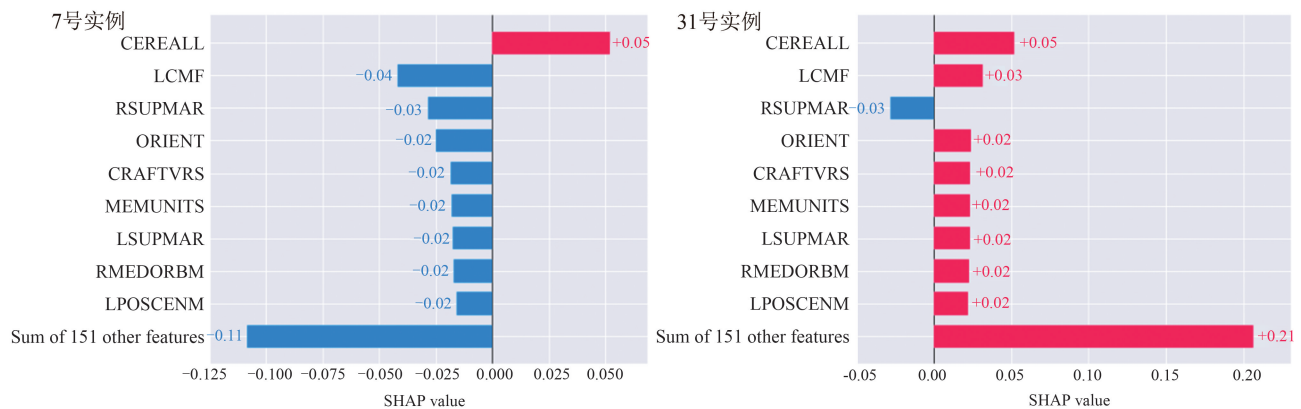


图 2 特定个体不同特征的边际贡献

(3) SHAP 瀑布图

瀑布图可以看作是对解释力图 and 局部条形图的整合,它完整地输出了特定个体的预测值、影响因素、所有特征的平均预测值以及每个特征的平均边际贡献(shapley 值)。

结合 7 号个体,图 3 显示 $E(f(x)) = 0.508$ 为所有特征的平均预测值, $f(x) = 0.27$ 为该个体对应的 aMCI

解释力图可以输出任意个体的预测值(所有特征的累计平均边际贡献(shapley 值)与平均预测值 $E(\hat{f}(X))$ 之和)以及影响个体患病的因素。图 1 解释了两个个体(7 号 naMCI、31 号 aMCI)的 aMCI 预测值,分别为 0.27 和 0.91,平均预测值 $E(\hat{f}(X)) = 0.508$ 。图中红色部分代表增加 aMCI 发生风险的因素,蓝色部分代表降低 aMCI 发生风险的因素(同时也是增加 naMCI 发生风险的因素)。

结合 7 号个体,我们可以得出: CEREAALL = 1204ml、LCMF = 2.977ml、RSUPMAR = 5.686ml、CRAFTVRS = 23、MEMUNITS = 16 等情况下,预测该个体为 aMCI 的概率为 0.27。31 号个体预测值与危险因素解释类似。

7 号个体,CEREAALL 对 aMCI 预测的平均边际贡献为 0.05,LCMF 的平均边际贡献为 -0.04,其他特征解释类似。所有特征的累计边际贡献为 -0.25,平均预测值为 0.508,故该个体 aMCI 的预测值为 0.27。30 号个体解释类似。

预测值,纵坐标展示了前 10 个特征及其对应的特征值,横坐标为特征的 shapley 值。31 号个体解释类似。

3. 全局解释图

(1)特征重要度排序图

XGBoost 模型本身可以根据增益输出特征重要度排序,增益是对给定特征的所有分割所贡献的损失或熵(不纯度)的总减少量。根据贪心算法,树根节点附

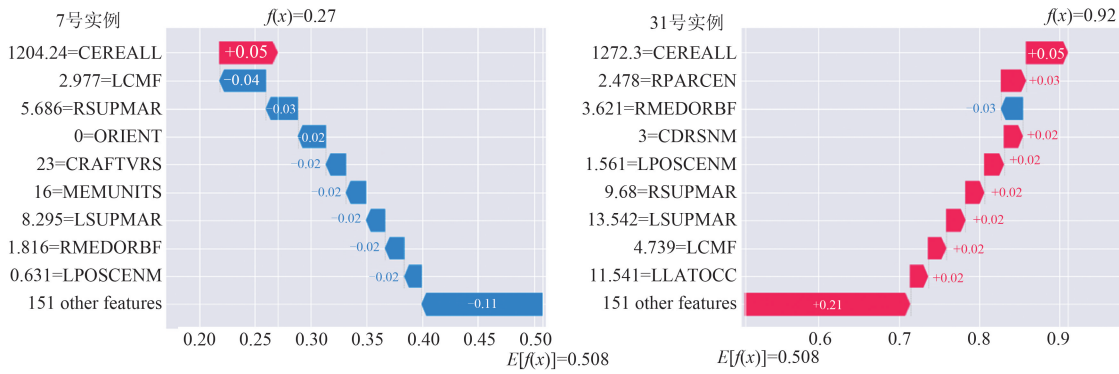


图3 特定个体个性化解释的 SHAP 瀑布图

近的分裂的特征比叶子节点附近分裂的特征更重要,然而增益更偏重于叶节点附近的分裂,这种偏差会导致通过增益获得的特征重要度可靠性降低^[22]。SHAP 框架也可以输出类似 XGBoost 模型的特征重要度排序,但不同的是:后者的给定特征重要度基于该特征所有个体 shapley 的平均绝对值 ($I_j = \sum_{i=1}^n$

$|\Phi_{ij}|$),其中 j 代表第 j 个特征, i 代表第 i 个个体, n 代表样本量。

图 4 显示两种方法输出前 10 个特征的重要度,可以看到二者差别较大。根据 SHAP 框架输出的结果:CEREALL 是预测 aMCI 发生最重要的特征,平均改变了 aMCI 发生绝对概率的 3.5%。

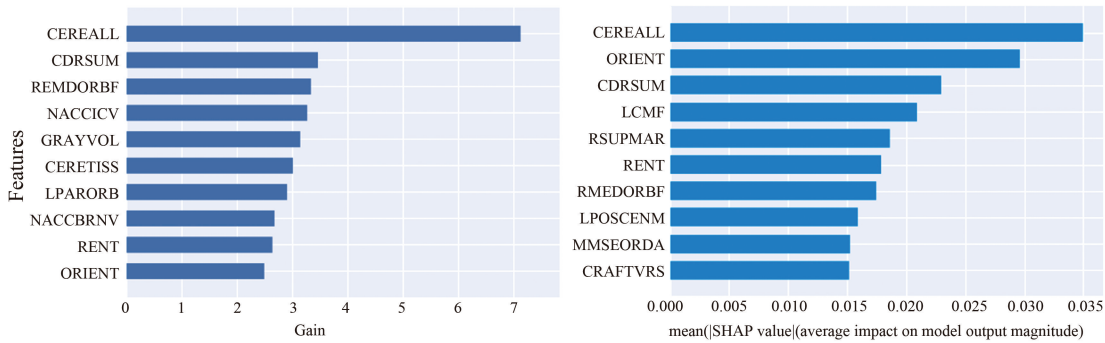


图4 基于增益和 SHAP 框架输出的特征重要度比较

(2) SHAP 摘要图

特征重要性排序图给出了特征相对重要性的概念,但它们并不能反映给定特征在个体间的影响范围分布,SHAP 摘要图则较好地弥补了这一点。图 5 纵坐标反映特征重要度;横坐标表示给定特征在不同个体间的平均边际贡献 (shapley 值);颜色反映 shapley 值的大小,从低(蓝色)到高(红色)。

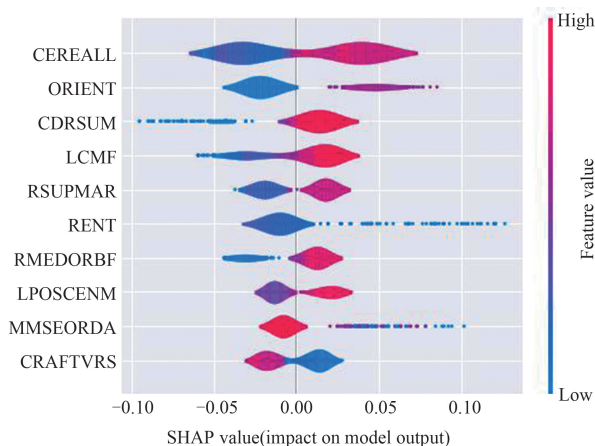


图5 SHAP 摘要图

就 CEREALL 而言,首先它是影响 aMCI 分类和预测的最重要指标,不同个体间 CEREALL 对应的 shapley 值越高,其患 aMCI 的风险越大。ORIENT、CDRSUM、LCMF、RSUPMAR 等解释类似,MMSEORDA、CRAFTVRS 的平均边际贡献恰好相反。

(3) SHAP 依赖图

SHAP 依赖图可以展示给定特征在不同个体间的 shapley 值的分布。随着特征的 shapley 值在不同个体间变化,该特征映射在个体上的预期输出也会改变,这有助于解释模型如何依赖该特征。图 6 展示了特征重要度排序靠前的 6 个特征的依赖图,横坐标为特征值,纵坐标为特征的 shapley 值。

就 CEREALL 而言,CEREALL = 1500 的个体相较于 CEREALL = 900 的个体 shapley 值更高,故前者被预测为 aMCI 的概率也更大。RENT 则相反,更大的体积意味着更低的 shapley 值,对应个体被预测为 aMCI 的概率也更小。

此外,SHAP 框架还提供了一种考虑特征交互作用的依赖图(图 7B),它捕捉了模型中由于交互作用而导致的垂直色散,这些效果可以通过交互特征的值

着色每个点实现可视化。

我们可以看到图 7A、B 中散点 (CEREALL 在不同个体间的 shapley 值) 的分布一致, 而 B 图通过 LC-

MF 的特征值对各散点进行着色。结果显示: 随着 LC-MF 特征值的增加, CEREALL 平均边际贡献 (shapley 值) 越大的个体被预测为 aMCI 的概率越大。

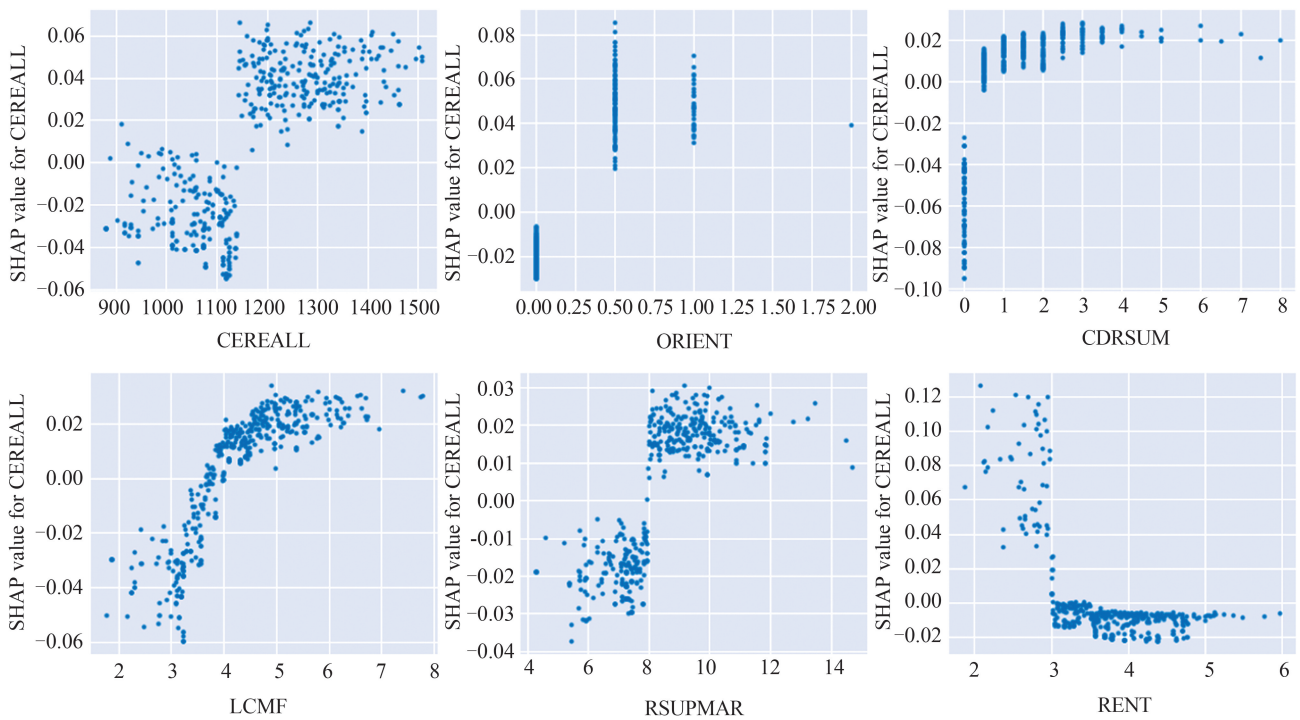


图 6 给定特征的 SHAP 依赖图

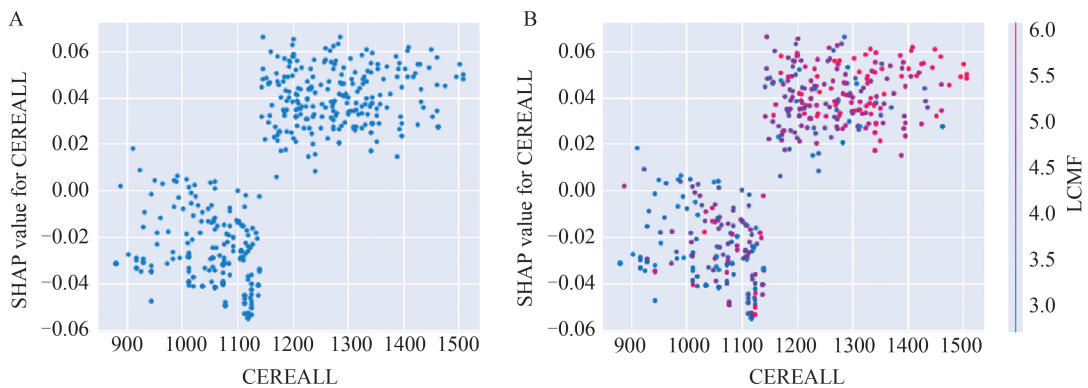


图 7 特征交互作用的 SHAP 依赖图

讨论

本研究通过联合 XGBoost 机器学习模型与 SHAP 构建可解释性框架, 展示个性化以及全局化的解释图, 分别实现了特定个体情况下不同特征的预测效果比较、不同个体间给定特征预测能力的判断。这在模型预测的准确性和解释性日益紧张的情况下, 破除了传统机器学习算法的黑箱, 并以用户友好的形式展示解释结果, 促进了方法学与决策工作的衔接。无论是基于个性化的还是全局性的解释图, 输出的不同特征的 shapley 值均可为 MCI 患者干预策略提供可靠依据。

本研究基于二分类目标结局构建的联合框架取得了较好的预测和解释效果。无独有偶, Chun^[26] 等人将

aMCI 患者 3 年随访资料 (包括神经心理测试和 APOE 基因) 作为特征, 利用逻辑回归、随机森林 (random forest, RF)、支持向量机 (support vector machine, SVM)、人工神经网络以及 XGBoost 等算法对 aMCI 和 AD 转归患者进行分类, 结果显示: XGBoost 的分类效果最好, 准确率和 AUC 分别达到 80.70% 和 0.852; XGBoost-SHAP 联合框架提示 CDRSB、MMSE 以及 APOE 基因被认为是识别 aMCI 和 AD 的重要特征。同样地, Martínez-Florez^[27] 等人将成套神经心理测试作为特征, 利用 AdaBoost、SVM、RF、决策树、梯度提升树、K-近邻以及 XGBoost 等算法对 NC 和 aMCI 进行识别, 结果显示 XGBoost 的准确率最佳 (81.20%); 根据可解释性框架, 听觉语言学习测试、短期记忆回忆、

以及语言流利性等测试项被认为是识别 NC 和 aMCI 的重要特征。然而上述研究准确率略低于本研究,可能的原因是前者未纳入 sMRI 资料,而 sMRI 被证实存在 AD 的分类中存在潜在价值^[28]。综上,我们有理由相信 XGBoost-SHAP 联合框架在 MCI 有关领域研究中的实用性和可靠性。

值得一提的是,临床工作中 AD 的多分类任务并不罕见,而 SHAP 并不会受到目标类别个数的约束;模型无关的归因性解释方法可以接口任意输出类别的预测^[29];并且 SHAP 还可以与其他的主流算法,如 RF、lightGBM、SVM、深度学习等联合^[30-31]。例如,El-sappagh^[32]等人联合 RF-SHAP 框架实现了对 NC、MCI 以及 AD 患者的多分类工作,并取得不错的预测效果:准确率和 F1 值分别为 93.95% 和 0.939;又如 Yongxing Lai^[33]等人利用 SVM-SHAP 探讨内质网应激相关基因在 AD 患者进展中的作用,结果显示:SVM 识别 NC 和 AD 效果理想(AUC>0.8),并根据 SVM-SHAP 框架得到 6 个可以精确预测 AD 进展的特征基因。

He^[23]等学者对 MCI 亚型的识别评估研究证实:aMCI 的颅内总容量(CEREALL)与 naMCI 患者存在差异,后者的脑萎缩表现更为明显。Kim^[34]等学者通过 2 年随访观察发现,进展为 AD 的 MCI 患者 CDR 总分(CDRSUM)、MMSE 总分均高于没有进展为 AD 者,且结合 CDR 定向力子条目(ORIENT)后预测 MCI 进展 AD 的准确率进一步提高。Csukly^[35]等学者对正常对照、aMCI 以及 naMCI 脑磁共振检查发现,与 naMCI 和正常对照相比,aMCI 组的海马、内嗅皮质(REN)和杏仁核体积均有所减小。Rusconi^[36]学者的研究报道,额叶(LCMF)、颞叶萎缩更多见于 naMCI 患者。这是因为 aMCI 主要是记忆领域受损,而记忆受损又与海马、内嗅皮质厚度下降和杏仁核萎缩息息相关;naMCI 则在记忆以外的其他领域(如语言或注意力、执行能力等)出现损伤,这些领域的损伤又与额叶、颞叶萎缩联系紧密,且研究认为 aMCI 的表现提示阿尔茨海默病病理,而 naMCI 的表现提示额颞叶痴呆、路易体痴呆病理^[37-38]。

本研究也存在一定局限,为了保持联合框架输出可解释性结果的可读性,我们在纳入患者基本信息、神经心理测试以及 sMRI 时未对数据进行融合,数据之间的异构性可能导致建模噪声,我们期待后续研究有其他方法可以完美地解决这一问题。

参 考 文 献

- [1] Anderson ND. State of the science on mild cognitive impairment (MCI). *CNS Spectr*, 2019, 24(1): 78-87.
- [2] Petersen RC. Mild Cognitive Impairment. *Continuum (Minneapolis)*, 2016, 22(2 Dementia): 404-418.
- [3] Li X, Zhang ZJ. Neuropsychological and neuroimaging characteristics of amnesic mild cognitive impairment subtypes: a selective overview. *CNS Neurosci Ther*, 2015, 21(10): 776-783.
- [4] Emmert NA, Reiter KE, Butts A, et al. Hippocampal Volumes in Amnesic and Non-Amnesic Mild Cognitive Impairment Types Using Two Common Methods of MCI Classification. *J Int Neuropsychol Soc*, 2022, 28(4): 391-400.
- [5] Chang CH, Lin CH, Lane HY. Machine Learning and Novel Biomarkers for the Diagnosis of Alzheimer's Disease. *Int J Mol Sci*, 2021, 22(5): 2761.
- [6] Jitsuishi T, Yamaguchi A. Searching for optimal machine learning model to classify mild cognitive impairment(MCI) subtypes using multimodal MRI data. *Sci Rep*, 2022, 12(1): 1-14.
- [7] Binaco R, Calzaretto N, Epifano J, et al. Machine Learning Analysis of Digital Clock Drawing Test Performance for Differential Classification of Mild Cognitive Impairment Subtypes Versus Alzheimer's Disease. *J Int Neuropsychol Soc*, 2020, 26(7): 690-700.
- [8] Zhang Y, Feng T, Wang S, et al. A Novel XGBoost Method to Identify Cancer Tissue-of-Origin Based on Copy Number Variations. *Front Genet*, 2020, 11: 585029.
- [9] Nagaraj S, Duong TQ. Deep Learning and Risk Score Classification of Mild Cognitive Impairment and Alzheimer's Disease. *J Alzheimers Dis*, 2021, 80(3): 1079-1090.
- [10] Yu H, Zhao Z, Cheng F. Predicting and investigating cytotoxicity of nanoparticles by translucent machine learning. *Chemosphere*, 2021, 276:130164.
- [11] Ribeiro MT, Singh S, Guestrin C. "Why should i trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016: 1135-1144.
- [12] Pang L, Wang J, Zhao L, et al. A Novel Protein Subcellular Localization Method With CNN-XGBoost Model for Alzheimer's Disease. *Front Genet*, 2019, 9: 751.
- [13] Loli ı P, Sori ı P, Logaruši ı M. Economic Policy Uncertainty Index Meets Ensemble Learning. *Comput Econ*, 2021, 15: 1-37.
- [14] Luckner M, Topolski B, Mazurek M. Application of XGBoost Algorithm in Fingerprinting Localisation Task. *Computer Information Systems and Industrial Management CISIM*, 2017.
- [15] Peng J, Zou K, Zhou M, et al. An Explainable Artificial Intelligence Framework for the Deterioration Risk Prediction of Hepatitis Patients. *J Med Syst*, 2021, 45(5): 61.
- [16] Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv*, 2017, 1702: 08608.
- [17] Du M, Liu N, Hu X. Techniques for interpretable machine learning. *Communications of the ACM*, 2019, 63(1): 68-77.
- [18] Lundberg SM, Lee SI. A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 2017:4768-4777.
- [19] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016:785-794.
- [20] Lundberg SM, Nair B, Vavilala MS, et al. Explainable machine-learning predictions for the prevention of hypoxemia during surgery. *Nat Biomed Eng*, 2018, 2(10): 749-760.
- [21] Wang J, Wiens J, Lundberg SM. Shapley Flow: A Graph-based Ap-

- proach to Interpreting Model Predictions. International Conference on Artificial Intelligence and Statistics, 2021; 721-729.
- [22] Lundberg SM, Erion G, Chen H, et al. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat Mach Intell*, 2020, 2(1): 56-67.
- [23] He J, Farias S, Martinez O, et al. Differences in brain volume, hippocampal volume, cerebrovascular risk factors, and apolipoprotein E4 among mild cognitive impairment subtypes. *Arch Neurol*, 2009, 66(11): 1393-1399.
- [24] Sachdev PS, Blacker D, Blazer DG, et al. Classifying neurocognitive disorders: the DSM-5 approach. *Nat Rev Neurol*, 2014, 10(11): 634-642.
- [25] Petersen RC, Caracciolo B, Brayne C, et al. Mild cognitive impairment: a concept in evolution. *J Intern Med*, 2014, 275(3): 214-228.
- [26] Chun MY, Park CJ, Kim J, et al. Prediction of conversion to dementia using interpretable machine learning in patients with amnesic mild cognitive impairment. *Front Aging Neurosci*, 2022, 14: 898940.
- [27] Martínez-Florez JF, Osorio JD, Cediel JC, et al. Short-Term Memory Binding Distinguishing Amnesic Mild Cognitive Impairment from Healthy Aging: A Machine Learning Study. *J Alzheimers Dis*, 2021, 81(2): 729-742.
- [28] Csukly G, Sirály E, Fodor Z, et al. The Differentiation of Amnesic Type MCI from the Non-Amnesic Types by Structural MRI. *Front Aging Neurosci*, 2016, 8: 52.
- [29] Lundberg SM, Erion GG, Lee SI. Consistent Individualized Feature Attribution for Tree Ensembles. *arXiv preprint arXiv*, 2018, 1802: 03888.
- [30] Molnar C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Independently published, 2022.
- [31] 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(3): e212240.
- [32] El-Sappagh S, Alonso JM, Islam SM, et al. A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer's disease. *Sci Rep*, 2021, 11(1): 1-26.
- [33] Lai Y, Lin X, Lin C, et al. Identification of endoplasmic reticulum stress-associated genes and subtypes for prediction of Alzheimer's disease based on interpretable machine learning. *Front Pharmacol*, 2022, 13: 975774.
- [34] Kim JW, Byun MS, Sohn BK, et al. Clinical dementia rating orientation score as an excellent predictor of the progression to Alzheimer's disease in mild cognitive impairment. *Psychiatry investigation*, 2017, 14(4): 420.
- [35] Csukly G, Sirály E, Fodor Z, et al. The Differentiation of Amnesic Type MCI from the Non-Amnesic Types by Structural MRI. *Front Aging Neurosci*, 2016, 8: 52.
- [36] Rusconi ML, Suardi A, Zanetti M, et al. Spatial navigation in elderly healthy subjects, amnesic and non amnesic MCI patients. *Journal of the neurological sciences*, 2015, 359(1-2): 430-437.
- [37] Iaria G, Chen JK, Guariglia C, et al. Retrosplenial and hippocampal brain regions in human navigation: complementary functional contributions to the formation and use of cognitive maps. *European Journal of Neuroscience*, 2007, 25(3): 890-899.
- [38] Weakley A, Schmitter-Edgecombe M, Anderson J. Analysis of verbal fluency ability in amnesic and non-amnesic mild cognitive impairment. *Archives of Clinical Neuropsychology*, 2013, 28(7): 721-731.

(责任编辑:邓妍)

(上接第 422 页)

参 考 文 献

- [1] 中国政府网. 中共中央、国务院印发《乡村振兴战略规划(2018-2022年)》. http://www.moa.gov.cn/nybgb/2018/201810/201812/t20181218_6165130.htm?phdjecimglgnoh.
- [2] 中共江苏省委新闻网. 江苏印发《关于进一步深化改革促进乡村医疗卫生体系健康发展实施意见》. http://www.zgjssw.gov.cn/fabuting/shengweiwenjian/202310/t20231019_8124924.shtml.
- [3] 刘思敏,李巧梅,唐佳宜,等.基于数据包络分析方法的我国村卫生室运行效率评价. *社区医学杂志*, 2023, 21(16): 853-859.
- [4] 林谋贵,兰韦娟,倪杰文,等.我国农村村级卫生资源配置公平与效率分析. *中国卫生经济*, 2023, 42(2): 53-57.
- [5] 高点,史卢少博,林锦慧,等.基于 DEA-GIS 方法的我国农村医疗卫生资源配置效率及公平性研究. *中国全科医学*, 2024, 27(7): 849-856.
- [6] 唐立健,王长青,钱东福.江苏省农村医疗卫生资源配置现状及效率分析. *中国卫生事业管理*, 2021, 38(8): 610-614.
- [7] 施雯慧,周健,张敏,等.江苏省村卫生室卫生资源配置现状及公平性研究. *中国农村卫生事业管理*, 2022, 42(10): 726-731.
- [8] 苗鑫,于欣波,王舒阳.乡村振兴战略下的“合村并居”模式发展策略研究. *农村经济与科技*, 2022, 33(3): 250-252.
- [9] 龚庆成.推进乡村医疗卫生人才建设筑牢医疗预防保健网底. *中国农村卫生*, 2023, 15(7): 52-53+67.
- [10] 杭苒枫,王丹晖,石武祥,等.基于数据包络分析法的江苏省卫生资源配置效率评价. *江苏预防医学*, 2022, 33(4): 417-421.
- [11] 张涛,孙立奇,刘肖肖,等.我国村卫生室卫生人力资源配置公平性研究. *中国卫生资源*, 2017, 20(1): 36-39.
- [12] 梅子鸿,刘婵娟.2012—2020 年我国基层医疗卫生资源配置效率分析. *中国卫生经济*, 2022, 41(10): 54-58.
- [13] 江恬雨,林小军,喻文茜.县域医疗共同体模式下乡镇卫生院效率影响因素实证研究. *中国卫生统计*, 2021, 38(3): 446-448.
- [14] 曹萌,薛佳.关于农村基层医疗卫生服务机构存在的问题分析及对策探讨. *农村实用技术*, 2021(3): 10-11.
- [15] 陈琦.关于农村基层医疗卫生服务机构存在的问题及对策探讨. *中国农村卫生*, 2021, 13(20): 13-14.
- [16] 何孟芹.农村基层医疗卫生服务机构存在的问题及对策研究:以贵州省为例. *农村经济与科技*, 2018, 29(13): 258-259.
- [17] 张端亮.乡镇卫生院医疗服务能力提升路径研究. *湖南中医药大学*.

(责任编辑:郭海强)