DOI：10.16781/j.CN31-2187/R.20220608

·论　著·

基于机器学习算法的重症缺血性脑卒中早期死亡预测效果评价

罗　枭，程　义，何　倩，涂博祥，吴　骋*，贺　佳*
海军军医大学（第二军医大学）卫生勤务学系军队卫生统计学教研室，上海 200433

［摘要］ **目的**　评价支持向量机（SVM）、随机森林、极限梯度提升（XGBoost）3 种机器学习算法与 logistic 回归模型在重症缺血性脑卒中 30 d 死亡结局预测中的效果。**方法**　使用 2008 年至 2019 年美国重症监护医学信息数据库Ⅳ（MIMIC-Ⅳ）中符合纳入标准的 2 358 例重症缺血性脑卒中患者资料，分别用 SVM、随机森林、XGBoost 3 种机器学习算法与 logistic 回归方法，结合合成少数过采样技术（SMOTE）建立早期死亡预测模型，并使用 ROC 曲线的 AUC 值、准确度、F1 分数、布里尔分数等指标评价模型的预测效果。**结果**　SVM、随机森林、XGBoost 与 logistic 回归模型在原始不平衡数据集中预测早期死亡的 AUC 值分别为 0.78、0.81、0.84、0.83。应用 SMOTE 合成数据集后，SVM、随机森林、XGBoost 与 logistic 回归模型的 AUC 值分别为 0.72、0.84、0.83、0.83。除 SVM 模型外，随机森林、XGBoost 模型与 logistic 回归之间有相似的预测能力，但其准确度、布里尔分数均优于 logistic 回归模型，综合分类性能更优。**结论**　机器学习算法在缺血性脑卒中早期死亡预测中性能较传统 logistic 回归方法更优。

［关键词］ 重症缺血性脑卒中；早期死亡预测；机器学习；合成少数过采样技术
［中图分类号］ R 743.3　［文献标志码］ A　［文章编号］ 2097-1338(2022)12-1365-07

Prediction of early mortality of severe ischemic stroke patients based on machine learning algorithms

LUO Xiao, CHENG Yi, HE Qian, TU Bo-xiang, WU Cheng*, HE Jia*
Department of Military Health Statistics, Faculty of Health Services, Naval Medical University (Second Military Medical University), Shanghai 200433, China

［**Abstract**］ **Objective**　To evaluate the effects of 3 machine learning algorithms (support vector machine［SVM］, random forest, and extreme gradient boosting［XGBoost］) and logistic regression in predicting the 30-d mortality of severe ischemic stroke patients. **Methods**　The data of 2 358 patients with severe ischemic stroke who qualified for the criteria in the Medical Information Mart for Intensive Care Ⅳ (MIMIC-Ⅳ) database from 2008 to 2019 were used. SVM, random forest, XGBoost and logistic regression combined with synthetic minority oversampling technique (SMOTE) were used respectively to build early mortality prediction models. The prediction performance of models was evaluated by the area under curve (AUC) of receiver operating characteristic curve, accuracy, F1-score, and Brier score. **Results**　The AUC values of SVM, random forest, XGBoost and logistic regression models using original unbalance data were 0.78, 0.81, 0.84 and 0.83, respectively. After using SMOTE-based synthetic data, the AUC values of SVM, random forest, XGBoost and logistic regression models were 0.72, 0.84, 0.83 and 0.83, respectively. Except for SVM, random forest and XGBoost had similar predictive ability to logistic regression, but their accuracy and Brier score were better than logistic regression, and their overall classification performance was better. **Conclusion**　Machine learning algorithms have better performance than traditional logistic regression in predicting early mortality of ischemic stroke patients.

［**Key words**］ severe ischemic stroke; early mortality prediction; machine learning; synthetic minority oversampling technique
［Acad J Naval Med Univ, 2022, 43(12): 1365-1371］

脑卒中是一种高死亡率和高发病率的心脑血管疾病。《中国脑卒中防治报告 2019》指出，脑卒中已是我国成人致死、致残的首位病因，其中缺血性脑卒中是最常见的卒中类型，占全部脑卒中

［收稿日期］ 2022-07-20　［接受日期］ 2022-09-02
［基金项目］ 军队双重学科建设项目-03，上海市公共卫生体系建设三年行动计划学科带头人计划（GWV-10.2-XD05），上海市公共卫生体系建设三年行动计划学科建设项目（GWV-10.1-XK05）. Supported by Project of Dual Key Discipline Construction of PLA-03, Fund for Leading Talents of Shanghai Three-Year Action Plan for Public Health System Construction (GWV-10.2-XD05), and Discipline Construction Program of Shanghai Three-Year Action Plan for Public Health System Construction (GWV-10.1-XK05).
［作者简介］ 罗　枭，硕士生. E-mail: luoxiao930501@163.com
* 通信作者（Corresponding authors）. Tel: 021-81871442, E-mail: wucheng_wu@126.com; Tel: 021-81871441, E-mail: hejia63@yeah.net

的 60%～80%[1]。脑卒中的死亡风险在发病 30 d 内最高，此后每年的平均死亡率为 10%[2-3]。实验室指标和临床数据对于缺血性脑卒的病情评估、临床治疗指导和预后判断有着重要参考价值，为此，如何利用丰富的临床信息构建精准的缺血性脑卒中早期预测模型、制定个体化治疗方案是目前研究的重点之一。

传统缺血性脑卒中预后预测模型通常基于 logistic 回归。作为广义线性模型的一种，当预测变量与因变量存在较为复杂的非线性关系时，logistic 回归的预测效果常不够理想[4]。因此，有必要探索更好的方法用于构建反映变量和结局之间真实关系的预测模型。

机器学习是一种统计学和计算机科学交叉的科学研究方法，它依赖于高效的算法。随着大数据时代的到来，越来越多的机器学习方法应用于医疗领域[5-6]。由于缺血性脑卒中患者的结局涉及复杂的临床指标，且存在较强的非线性关联，适合采用机器学习模型进行分析。

本研究基于美国重症监护医学信息数据库Ⅳ（Medical Information Mart for Intensive Care Ⅳ，MIMIC-Ⅳ）1.0 版[7]开展了一项回顾性队列研究，采用支持向量机（support vector machine，SVM）、随机森林、极限梯度提升（extreme gradient boosting，XGBoost）3 种机器学习算法构建了重症缺血性脑卒中患者 30 d 内死亡的预测模型，并与 logistic 回归模型进行对比，评价不同模型的预测能力，为重症缺血性脑卒中患者早期预后研究提供方法学参考。

1　资料和方法

本研究遵循个体预后或诊断的多变量预测模型透明报告（transparent reporting of a multivariable prediction model for individual prognosis or diagnosis，TRIPOD）[8]。

1.1　研究对象与结局　研究对象数据来源于 MIMIC-Ⅳ 1.0 版[7]。目前，MIMIC-Ⅳ记录了 2008 年至 2019 年波士顿贝斯以色列女执事医疗中心 ICU 收治的患者数据，具有全面、高质量和去隐私化等特征。研究者已获得该数据库的使用授权（编码：48510731）。利用结构化查询语言，在 MIMIC-Ⅳ中提取符合国际疾病分类（ICD-9 代码为 433、434、436、437.0、437.1 或 ICD-10 代码为 I63、I65、I66）缺血性脑卒中诊断的患者信息。共提取 2 831 例重症缺血性脑卒中患者的相关数据，排除年龄＜18 岁、ICU 入院时间＜1 d 或多次入住 ICU 的患者，经过筛选得到 2 358 例患者。根据重症缺血性脑卒中患者 30 d 内是否死亡，将患者分为死亡组与存活组。

1.2　变量选取　综合考虑 MIMIC-Ⅳ中数据的复杂性、数据缺失程度及国内外缺血性脑卒中预后影响因素，从一般资料、合并症、生理指标及实验室指标中筛选出 34 个关键变量：性别、年龄、体重、吸烟史、饮酒史、高血压、高脂血症、糖尿病、慢性阻塞性肺疾病、冠心病、心房颤动、心率、舒张压、收缩压、平均血压、呼吸频率、体温、血氧饱和度、血糖、血清胆固醇、血细胞比容、血红蛋白、白细胞计数、阴离子间隙、碳酸氢盐、血尿素氮、钙离子、氯化物、钠离子、钾离子、国际标准化比值、凝血酶原时间、部分凝血活酶时间、序贯器官衰竭估计（sequential organ failure assess，SOFA）评分。生理指标及实验室指标取值为登记入 ICU 后 24 h 内的平均值。

1.3　数据预处理　本研究对 MIMIC-Ⅳ中的缺失值及异常值进行了处理，删除了缺失比例＞20% 的变量。异常值定义为各计量数据中＜0.01% 和＞99.99% 的数据，删除异常值后进行填补，尽可能保留原始数据信息。针对本研究使用的高维数据特征，采用线性多重填补法进行缺失值填补。计量资料为了保持原始数值信息且便于计算，使用标准化法对数据进行归一化处理，以 0 为均数、1 为标准差。计数资料采用设置哑变量的方法纳入建模。在变量纳入模型训练前剔除近零方差的变量饮酒史，以避免纳入该变量而导致模型破坏或数据拟合不稳定。

1.4　统计学处理　符合正态分布的计量资料以 $\bar{x}\pm s$ 表示，组间比较采用独立样本 t 检验；不符合正态分布的计量资料以中位数（下四分位数，上四分位数）表示，组间比较采用 Wilcoxon 秩和检验。计数资料以例数和百分数表示，组间比较采用 χ^2 检验。

本研究使用了 4 种模型，分别是 logistic 回归、SVM、随机森林和 XGBoost。数据集（n＝2 358）按 7 ∶ 3 随机分为训练集（n＝1 660）和测试集（n＝698），其中训练集存活患者 1 476 例、死亡患者 184 例，测试集存活患者 619 例、死亡患者 79 例。

训练集用于拟合模型，测试集用于评估模型性能。对训练集采用合成少数过采样技术（synthetic minority oversampling technique，SMOTE）处理结局类别的不平衡问题，新生成死亡患者 1 472 例与原训练集存活患者 1 476 例组合生成含 2 948 例样本的训练集（死亡例数：存活例数≈1 ： 1）。多因素 logistic 回归模型基于赤池信息准则（Akaike information criterion，AIC），使用逐步法筛选变量。为使 SVM、随机森林和 XGBoost 3 种机器学习算法获得最佳的训练模型，各模型训练使用重复 5 次的 5 折交叉验证以减少模型过拟合，并使用网格搜索的方法进行超参数寻优。本研究使用 ROC 曲线的 AUC 值作为主要评价指标，并选取 F1 分数（F1-score）、布里尔分数（Brier score）和准确度评估模型的性能[9]，模型间 AUC 值比较使用 DeLong 检验[10]。根据测试集的模型验证结果，选取每个算法表现最好的模型。根据模型变量重要性分析影响缺血性脑卒中患者早期不良预后的危险因素。

所有统计学分析均采用 R 4.1.2 软件，均为双侧检验，检验水准（α）为 0.05。

2　结　果

2.1　*患者基本特征*　纳入重症缺血性脑卒中患者 2 358 例，男女比为 1 ： 0.95，平均年龄为（72.0±14.3）岁，其中存活组 2 095 例、死亡组 263 例，重症缺血性脑卒中患者 30 d 内死亡率为 11.2%。与存活组相比，死亡组患者年龄较大，体重较轻，较少有高血压、高血脂症，较多伴有心房颤动，心率、呼吸频率、体温、血氧饱和度、血糖、白细胞计数、阴离子间隙、血尿素氮、钠离子、钾离子、国际标准化比值、凝血酶原时间、SOFA 评分较高，碳酸氢盐水平较低，差异均有统计学意义（P 均＜0.05）。见表 1。

表 1　根据 30 d 预后分组的重症缺血性脑卒中患者的基本资料

Tab 1　Basic characteristics of severe ischemic stroke patients with different 30 d prognoses

Variable	Survival group N=2 095	Death group N=263	Statistic	P value
Gender, *n* (%)			χ^2=1.29	0.26
Female	1 009 (48.2)	137 (52.1)		
Male	1 086 (51.8)	126 (47.9)		
Age/year, $\bar{x}\pm s$	71.2±14.4	78.3±11.4	t=9.26	＜0.01
Body weight/kg, $\bar{x}\pm s$	80.2±21.0	74.9±20.3	t=3.98	＜0.01
Smoking history, *n* (%)			χ^2=1.71	0.19
Yes	401 (19.1)	41 (15.6)		
No	1 694 (80.9)	222 (84.4)		
History of alcohol consumption, *n* (%)			χ^2=0.01	0.97
Yes	61 (2.9)	7 (2.7)		
No	2 034 (97.1)	256 (97.3)		
Hypertension, *n* (%)			χ^2=14.68	＜0.01
Yes	1 127 (53.8)	108 (41.1)		
No	968 (46.2)	155 (58.9)		
Hyperlipidemia, *n* (%)			χ^2=23.42	＜0.01
Yes	1 418 (67.7)	138 (52.5)		
No	677 (32.3)	125 (47.5)		
Diabetes mellitus, *n* (%)			χ^2=0.10	0.75
Yes	534 (25.5)	70 (26.6)		
No	1 561 (74.5)	193 (73.4)		
COPD, *n* (%)			χ^2=0.10	0.75
Yes	534 (25.5)	70 (26.6)		
No	1 561 (74.5)	193 (73.4)		
Coronary heart disease, *n* (%)			χ^2=0.49	0.48
Yes	800 (38.2)	94 (35.7)		
No	1 295 (61.8)	169 (64.3)		
Atrial fibrillation, *n* (%)			χ^2=8.13	＜0.01
Yes	732 (34.9)	116 (44.1)		
No	1 363 (65.1)	147 (55.9)		
Heart rate/min^{-1}, $\bar{x}\pm s$	79.2±14.5	86.2±17.1	t=6.31	＜0.01
Systolic blood pressure/mmHg, $\bar{x}\pm s$	130.4±18.6	132.2±0.7	t=1.39	0.16

续表 1

Variable	Survival group N=2 095	Death group N=263	Statistic	P value
Diastolic blood pressure/mmHg, $\bar{x}\pm s$	68.1±12.8	67.5±12.2	t=0.71	0.49
Mean blood pressure/mmHg, $\bar{x}\pm s$	85.2±12.8	84.9±12.9	t=0.30	0.76
Respiratory rate/min^{-1}, $\bar{x}\pm s$	18.8±3.1	20.8±4.2	t=7.56	<0.01
Temperature/℃, $\bar{x}\pm s$	36.9±0.4	37.0±0.5	t=3.94	<0.01
Oxygen saturation/%, $\bar{x}\pm s$	96.9±1.7	97.5±1.9	t=5.32	<0.01
Glucose/(mmol•L^{-1}), $\bar{x}\pm s$	7.5±2.3	8.6±2.9	t=6.12	<0.01
Cholesterol/(mmol•L^{-1}), $\bar{x}\pm s$	2.7±0.1	2.7±0.2	t=0.76	0.45
Erythrocyte specific volume/%, $\bar{x}\pm s$	35.6±6.0	35.7±6.4	t=0.39	0.70
Hemoglobin/(g•L^{-1}), $\bar{x}\pm s$	118.0±21.0	117.0±22.0	t=0.95	0.34
WBC/(L^{-1}, ×10^{9}) , $\bar{x}\pm s$	10.8±4.5	13.4±6.1	t=6.84	<0.01
Anion gap/(mmol•L^{-1}), $\bar{x}\pm s$	14.4±3.0	16.2±3.8	t=7.66	<0.01
Bicarbonate/(mmol•L^{-1}), $\bar{x}\pm s$	23.5±3.1	22.2±3.6	t=5.34	<0.01
Blood urea nitrogen/(mmol•L^{-1}), $\bar{x}\pm s$	12.2±8.7	16.2±11.9	t=5.31	<0.01
Calcium/(mmol•L^{-1}), $\bar{x}\pm s$	8.6±0.7	8.6±0.7	t=1.78	0.07
Chloride/(mmol•L^{-1}), $\bar{x}\pm s$	104.2±4.8	104.3±5.8	t=0.20	0.84
Sodium/(mmol•L^{-1}), $\bar{x}\pm s$	139.1±4.0	139.9±5.4	t=2.47	0.01
Potassium/(mmol•L^{-1}), $\bar{x}\pm s$	4.2±0.5	4.3±0.6	t=2.31	0.02
International normalized ratio, $\bar{x}\pm s$	1.3±0.5	1.4±0.7	t=2.59	0.01
Prothrombin time/s, $\bar{x}\pm s$	14.2±5.8	15.3±6.8	t=2.57	0.01
Partial thromboplastin time/s, $\bar{x}\pm s$	35.6±16.3	36.4±18.6	t=0.74	0.46
SOFA score, M (Q_L, Q_U)	1 (0, 2)	1 (0, 3)	Z=34.95	<0.01

1 mmHg=0.133 kPa. COPD: Chronic obstructive pulmonary disease; WBC: White blood cell; SOFA: Sequential organ failure assess; M (Q_L, Q_U): Median (lower quartile, upper quartile).

2.2　机器学习算法及 logistic 回归模型在原始数据集上的表现　表 2 显示了 4 种模型应用于原始训练集数据的结果。XGBoost 模型的分类效能最优，AUC 值为 0.84；SVM 模型的分类效能较差，AUC 值为 0.78；XGBoost 与 logistic 回归模型的 AUC 值比较差异无统计学意义（P=0.78）。F1 分数从高到低依次为 logistic 回归、XGBoost、SVM、随机森林模型，4 种模型的布里尔分数及准确度相近。

表 2　4 种模型在原始数据集上的性能表现

Tab 2　Performance of 4 models on original dataset

Indicator	Logistic regression	XGBoost	Random forest	SVM
AUC (95% *CI*)	0.83 (0.77, 0.89)	0.84 (0.78, 0.90)	0.81 (0.77, 0.89)	0.78 (0.72, 0.84)
F1-score	0.38	0.24	0.15	0.26
Brier score	0.08	0.08	0.09	0.09
Accuracy (95% *CI*)	0.90 (0.87, 0.92)	0.89 (0.87, 0.91)	0.89 (0.87, 0.91)	0.89 (0.86, 0.91)
Sensitivity (95% *CI*)	0.28 (0.19, 0.39)	0.11 (0.06, 0.20)	0.09 (0.04, 0.17)	0.18 (0.11, 0.28)
Specificity (95% *CI*)	0.98 (0.96, 0.99)	0.99 (0.98, 1.00)	0.99 (0.98, 0.99)	0.98 (0.96, 0.99)
Positive predictive value (95% *CI*)	0.61 (0.45, 0.75)	0.64 (0.39, 0.84)	0.50 (0.27, 0.73)	0.52 (0.34, 0.69)
Negative predictive value (95% *CI*)	0.91 (0.89, 0.93)	0.90 (0.87, 0.92)	0.90 (0.87, 0.92)	0.90 (0.88, 0.92)

XGBoost: Extreme gradient boosting; SVM: Support vector machine; AUC: Area under curve; *CI*: Confidence interval.

2.3　机器学习算法及 logistic 回归模型在合成数据集上的表现　表 3 显示了 4 种模型应用于 SMOTE 合成后训练集数据在测试集上的分类表现。随机森林模型的分类效能最优，AUC 值为 0.84；SVM 模型的分类效能较差，AUC 值为 0.72。与使用原始训练集的分类结果相比，随机森林模型的分类表现有所提升，虽然随机森林模型的 AUC 值与 logistic 回归模型比较差异无统计学意义（P=0.55），但准确度及布里尔分数均优于 logistic 回归模型，分类性能综合表现更优。

2.4　机器学习算法及 logistic 回归模型变量的重要性　经过 5 折交叉验证后，合成训练集中相对重要性排名前 10 位的变量见图 1。对于 logistic 回归模型，血糖是影响模型性能的最重要变量，其次是呼吸频率和白细胞计数。XGBoost 模型相对重要性排名前 3 位的变量为高血压、高血脂症和年龄。随机森林模型相对重要性排名前 3 位的变量为年龄、血糖和血氧饱和度。上述 2 种机器学习算法和 logistic 回归模型相对重要性排名前 10 位的变量基本相同。

表 3　4 种模型在合成数据集上的性能表现

Tab 3　Performance of 4 models on synthetic dataset

Indicator	Logistic regression	XGBoost	Random forest	SVM
AUC (95% *CI*)	0.83 (0.77, 0.89)	0.83 (0.77, 0.89)	0.84 (0.78, 0.90)	0.72 (0.65, 0.79)
F1-score	0.41	0.31	0.24	0.12
Brier score	0.16	0.09	0.10	0.11
Accuracy (95% *CI*)	0.77 (0.73, 0.80)	0.88 (0.86, 0.91)	0.89 (0.86, 0.91)	0.87 (0.84, 0.90)
Sensitivity (95% *CI*)	0.72 (0.61, 0.81)	0.23 (0.15, 0.33)	0.17 (0.10, 0.26)	0.08 (0.04, 0.16)
Specificity (95% *CI*)	0.77 (0.74, 0.80)	0.97 (0.95, 0.98)	0.97 (0.96, 0.98)	0.97 (0.96, 0.98)
Positive predictive value (95% *CI*)	0.29 (0.23, 0.35)	0.46 (0.32, 0.61)	0.43 (0.27, 0.61)	0.26 (0.13, 0.46)
Negative predictive value (95% *CI*)	0.96 (0.93, 0.97)	0.91 (0.88, 0.93)	0.90 (0.88, 0.92)	0.89 (0.87, 0.91)

XGBoost: Extreme gradient boosting; SVM: Support vector machine; AUC: Area under curve; *CI*: Confidence interval.

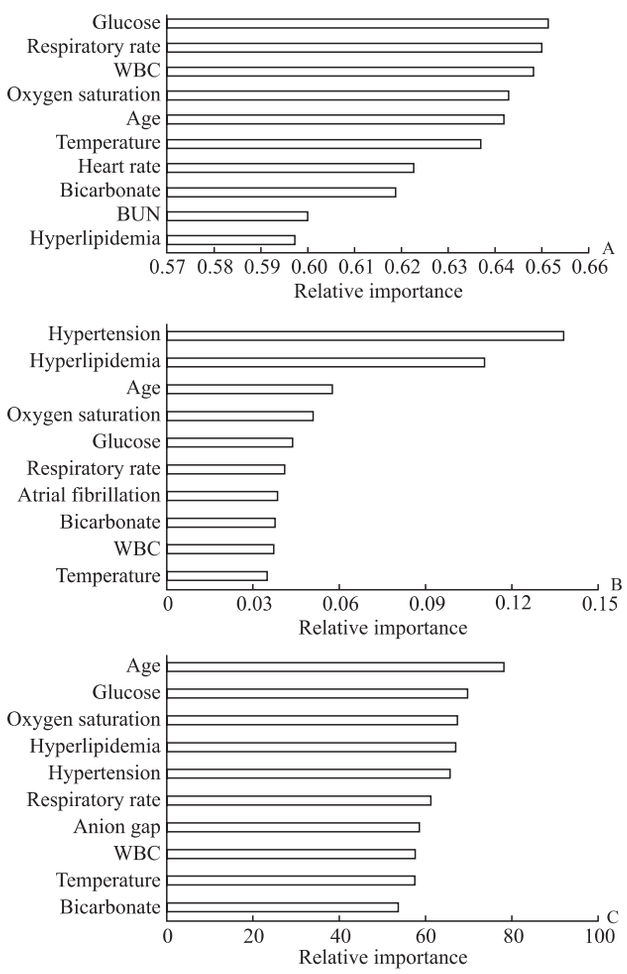

图 1　2 种机器学习算法和 logistic 回归模型变量的相对重要性排序

Fig 1　Relative importance ranking of variables of 2 machine learning algorithms and logistic regression model

A: Logistic regression model; B: XGBoost (extreme gradient boosting) model; C: Random forest model. WBC: White blood cell; BUN: Blood urea nitrogen.

3　讨　论

本研究将机器学习算法与 logistic 回归应用于重症缺血性脑卒中患者的早期死亡预测，利用多个评估指标对模型的预测及泛化能力进行了对比，并对变量的相对重要性进行排序。4 种模型算法都得到了初步验证，即使在缺乏脑卒中特异性评价指标如美国国立卫生研究院卒中量表（National Institutes of Health stroke scale，NIHSS）[11]的情况下，本研究算法也有良好的表现，AUC 值均能达到 0.8 左右。值得注意的是，模型纳入的指标是患者入住 ICU 时的初步评估和检查结果，其优点是能在短时间内快速预测患者的早期预后结局，帮助医师做出临床决定。

作为传统的临床预测模型，logistic 回归模型在缺血性脑卒中的预后预测中应用较多，且有不错的预测效果[12]，因其具有易于实施、便于解释等特点常被用于对比分析机器学习算法的预测效果[13-15]。本研究使用原始数据集并没有得出机器学习模型在重症缺血性脑卒中的临床预测中表现优于 logistic 回归模型的结论，这与既往系统综述结果[16]相似。分析原因可能有以下几点：（1）本研究用于训练机器学习和 logistic 回归模型的样本量较为充足，不足以发挥机器学习应用于更大样本和更多变量的优势。（2）本研究纳入的自变量数目对于机器学习算法和 logistic 回归模型都可能是足够的，且可能存在线性相关的连续变量，这可能使机器学习算法和 logistic 回归模型之间的差异难以区分。（3）由于分类结局中死亡发生率相对较低，低信噪比可能是难以区分机器学习算法和 logistic 回归模型的一个原因[17]。考虑到死亡结局发生率较低，且机器学习算法在应用于极不平衡数据时通常不能提供明显的分类改进[18]，本研究在训练集中使用了 SMOTE 对原始数据进行了合成以平衡结局比例，其优点是能够通过人工合成样本取代随机复制少数类的方法减弱过拟合，并且没有丢失有用的信息[19]。

对于测试集而言，当训练集使用 SMOTE 合成数据集时，机器学习算法在综合分类效能方面有所提升且优于 logistic 回归模型，说明其减弱了结局分类失衡对机器学习算法的影响。但同时本研究也发现 SVM 的分类效能较低，推测可能是因为合成后训练集增大了类别间的重叠，使模型的泛化能力降低。因此，在机器学习算法的应用中是否使用 SMOTE，尚需结合算法本身的特点及其稳健性的要求判断，不能一概而论。

使用合成训练集数据时，logistic 回归、XGBoost 及随机森林模型相对重要性排名前 10 位的变量基本相同，主要为年龄、血糖、体温、白细胞计数、血氧饱和度、高脂血症、高血压等，这与既往研究结果[13,20-21]相似。近年来，使用机器学习算法预测脑卒中预后的研究不断涌现。Monteiro 等[22]应用机器学习算法（随机森林、XGBoost、SVM 和决策树）预测缺血性脑卒中患者在初次脑卒中 3 个月后接受重组组织型纤溶酶原激活物治疗的功能结局，初始研究仅纳入入院时的特征，后续通过增加入院后不同时间点的特征来改善预测效果。在使用初入院特征时，机器学习算法的 AUC 值为 0.81，随着新特征的逐步加入 AUC 值增加到 0.90。Bacchi 等[23]使用澳大利亚患者入院临床数据结合机器学习算法来预测脑卒中患者的住院时间和院内结局，研究结果提示 logistic 回归在预测院内死亡率方面表现最好，AUC 值为 0.90。Fernandez-Lozano 等[24]使用随机森林模型预测脑卒中 3 个月的死亡率和功能结局，发现该模型有良好的预测能力，AUC 值最高为 0.91。以上研究除使用本研究涉及的基本资料及临床数据外，还包括脑卒中的特异性评价指标，如 NIHSS 评分或神经影像学结果，这可能是导致模型预测效能高于本研究的原因之一。

尽管本研究中模型的预测效果与现有的缺血性脑卒中风险评分工具[25]相似或略胜一筹，但也有其局限性：（1）未纳入与缺血性脑卒中相关的 NIHSS 评分、神经影像学或药物治疗等数据；（2）未对模型进行外部数据验证；（3）未对出院后的功能结局进行预测，而功能结局也是脑卒中患者及其家属最关心的问题之一。

综上所述，本研究利用 MIMIC-Ⅳ资料，采用机器学习算法及 logistic 回归在重症缺血性脑卒中患者中建立 30 d 内死亡预测模型，所构建的模型分类性能良好且表现稳定。虽然本研究机器学习与 logistic 回归模型的预测效能并无明显差异，但机器学习算法在捕捉非线性关系、处理大规模样本及高维度变量方面存在优势[26]。在今后的研究中可以考虑纳入临床轨迹、神经影像学信息或使用特征优化的方法提升变量数据维度，也可考虑使用不同采样技术或深度学习等方法提高模型的准确性与适用性。

［参 考 文 献］

［1］《中国脑卒中防治报告 2019》编写组.《中国脑卒中防治报告 2019》概要［J］. 中国脑血管病杂志，2020，17：272-281.

［2］BRØNNUM-HANSEN H, DAVIDSEN M, THORVALDSEN P, GROUP D M S. Long-term survival and causes of death after stroke［J］. Stroke, 2001, 32: 2131-2136.

［3］SINGH R J, CHEN S, GANESH A, HILL M D. Long-term neurological, vascular, and mortality outcomes after stroke［J］. Int J Stroke, 2018, 13: 787-796.

［4］TU J V. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes［J］. J Clin Epidemiol, 1996, 49: 1225-1231.

［5］LIEW B X W, PEOLSSON A, RUGAMER D, WIBAULT J, LÖFGREN H, DEDERING A, et al. Clinical predictive modelling of post-surgical recovery in individuals with cervical radiculopathy: a machine learning approach［J/OL］. Sci Rep, 2020, 10: 16782. DOI: 10.1038/s41598-020-73740-7.

［6］KRITTANAWONG C, VIRK H U H, BANGALORE S, WANG Z, JOHNSON K W, PINOTTI R, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis［J/OL］. Sci Rep, 2020, 10: 16057. DOI: 10.1038/s41598-020-72685-1.

［7］JOHNSON A, BULGARELLI L, POLLARD T, HORNG S, CELI L A, MARK R. MIMIC-Ⅳ (version 1.0)［J/OL］. PhysioNet, 2021. (2021-03-16)［2022-07-20］. *https://doi.org/10.13026/s6n6-xd98*.

［8］COLLINS G S, REITSMA J B, ALTMAN D G, MOONS K G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement［J］. Br J Surg, 2015, 102: 148-158.

［9］POWERS D M W. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation［J/OL］. 2020: arXiv:2010.16061［cs.LG］(2020-10-11)［2022-07-20］. *https://arxiv.org/abs/*

2010.16061v1.

［10］ DELONG E R, DELONG D M, CLARKE-PEARSON D L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach[J]. Biometrics, 1988, 44: 837-845.

［11］ FAROOQUE U, LOHANO A K, KUMAR A, KARIMI S, YASMIN F, BOLLAMPALLY V C, et al. Validity of national institutes of health stroke scale for severity of stroke to predict mortality among patients presenting with symptoms of stroke[J/OL]. Cureus, 2020, 12: e10255. DOI: 10.7759/cureus.10255.

［12］ 袁水生，马芳杰，韩丽华. 急性缺血性脑卒中患者远期预后的预测因子分析[J]. 中国实用神经疾病杂志，2018，21：2063-2068.

［13］ 饶夫阳，宋艳平，吕芯芮，白旭，覃伟，刘欢，等. 基于机器学习模型缺血性脑卒中 1 年死亡预测效果评价[J]. 中国公共卫生，2019，35：1187-1191.

［14］ HEO J, YOON J G, PARK H, KIM Y D, NAM H S, HEO J H. Machine learning-based model for prediction of outcomes in acute stroke[J]. Stroke, 2019, 50: 1263-1265.

［15］ LI X, PAN X D, JIANG C L, WU M R, LIU Y K, WANG F S, et al. Predicting 6-month unfavorable outcome of acute ischemic stroke using machine learning[J/OL]. Front Neurol, 2020, 11: 539509. DOI: 10.3389/fneur.2020.539509.

［16］ CHRISTODOULOU E, MA J, COLLINS G S, STEYERBERG E W, VERBAKEL J Y, VAN CALSTER B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models[J]. J Clin Epidemiol, 2019, 110: 12-22.

［17］ ENNIS M, HINTON G, NAYLOR D, REVOW M, TIBSHIRANI R. A comparison of statistical learning methods on the Gusto database[J]. Stat Med, 1998, 17: 2501-2508.

［18］ KOIVU A, SAIRANEN M, AIROLA A, PAHIKKALA T. Synthetic minority oversampling of vital statistics data with generative adversarial networks[J]. J Am Med Inform Assoc, 2020, 27: 1667-1674.

［19］ 石洪波，陈雨文，陈鑫. SMOTE 过采样及其改进算法研究综述[J]. 智能系统学报，2019，14：1073-1083.

［20］ 孙勇，王立强，王芬，陈国强，张颖超，刘亚辉，等. 可解释的机器学习模型用于预测远期脑缺血事件[J]. 心脑血管病防治，2022，22：53-56，60.

［21］ REICHE E M V, GELINKSI J R, ALFIERI D F, FLAUZINO T, LEHMANN M F, DE ARAÚJO M C M, et al. Immune-inflammatory, oxidative stress and biochemical biomarkers predict short-term acute ischemic stroke death[J]. Metab Brain Dis, 2019, 34: 789-804.

［22］ MONTEIRO M, FONSECA A C, FREITAS A T, PINHO E MELO T, FRANCISCO A P, FERRO J M, et al. Using machine learning to improve the prediction of functional outcome in ischemic stroke patients[J]. IEEE/ACM Trans Comput Biol Bioinform, 2018, 15: 1953-1959.

［23］ BACCHI S, OAKDEN-RAYNER L, MENON D K, JANNES J, KLEINIG T, KOBLAR S. Stroke prognostication for discharge planning with machine learning: a derivation study[J]. J Clin Neurosci, 2020, 79: 100-103.

［24］ FERNANDEZ-LOZANO C, HERVELLA P, MATO-ABAD V, RODRÍGUEZ-YÁÑEZ M, SUÁREZ-GARABOA S, LÓPEZ-DEQUIDT I, et al. Random forest-based prediction of stroke outcome[J/OL]. Sci Rep, 2021, 11: 10071. DOI: 10.1038/s41598-021-89434-7.

［25］ NTAIOS G, FAOUZI M, FERRARI J, LANG W, VEMMOS K, MICHEL P. An integer-based score to predict functional outcome in acute ischemic stroke: the ASTRAL score[J]. Neurology, 2012, 78: 1916-1922.

［26］ 向超益，吴亚飞，方亚. 数据挖掘技术在心血管疾病预后研究中的应用进展[J]. 中华流行病学杂志，2021，42：2234-2238.

［**本文编辑**］　杨亚红