

## ◆特邀栏目◆

## 基于机器学习的静脉溶栓时间窗外急性缺血性脑卒中患者短期预后模型建立与验证\*

周辉<sup>1</sup>, 冯兵<sup>2\*\*</sup>, 刘旭日<sup>1</sup>, 邹勇<sup>1</sup>, 唐慧<sup>1</sup>, 黄钦彬<sup>2</sup>, 陈庭军<sup>2</sup>, 邹东华<sup>3</sup>

(1. 桂平市中医医院, 广西桂平 537200; 2. 桂平市人民医院, 广西桂平 537200; 3. 广西医科大学第二附属医院, 广西南宁 530007)

**摘要:**本研究旨在探讨静脉溶栓时间窗外, 急性缺血性脑卒中(Acute Ischemic Stroke, AIS)患者入院时常规临床资料对其短期不良结局的影响。该研究通过收集桂平市中医医院神经内科住院治疗的溶栓时间窗外 AIS 患者(774 例)临床资料, 比较短期不同预后患者入院时的临床资料, 并利用机器学习分类(Classification)算法建模分析患者 90 d 内预后不良的影响因素, 同时采用受试者工作特征曲线(ROC)和校准图验证模型的预测效能及准确度。研究结果显示:(1)3 个月随访期内预后不良患者占比 13.95%(108 例)。(2)预后不良患者的年龄、血白细胞(WBC)水平、血 C 反应蛋白(C-Reactive Protein, CRP)水平、血甘油三酯(TG)水平、入院美国国立卫生研究院卒中量表(National Institutes of Health Stroke Scale, NIHSS)评分、梗死灶体积和发病-治疗时间大于或高于预后良好患者;预后不良患者的入院格拉斯哥昏迷(Glasgow Coma Scale, GCS)评分低于预后良好组, 均差异显著( $P < 0.05$ )。机器学习分类算法中, 极限梯度提升(Extreme Gradient Boosting, XGB)算法的效果最佳[受试者工作特征曲线下面积(Area Under Curve, AUC) = 0.81], 且校准曲线显示模型评估预测风险与实际发生风险的一致程度高。(3)XGB 算法筛选的预测变量影响权重居前 6 位的依次为发病-治疗时间、血 CRP 水平、年龄、梗死灶体积、血 TG 水平与 NIHSS 评分。上述结果说明 XGB 算法可用于预测静脉溶栓时间窗外 AIS 患者短期不良预后的影响因素。

**关键词:**急性缺血性脑卒中; 超溶栓时间窗; 预后; 机器学习; 临床预测模型

中图分类号: R743.33 文献标识码: A 文章编号: 1005-9164(2024)04-0781-07

DOI: 10.13656/j.cnki.gxkx.20240919.003

脑卒中是全球排名第二的致死性疾病, 我国每年每 10 万人中有 340 例以上患者因其死亡<sup>[1]</sup>。急性缺

血性脑卒中(Acute Ischemic Stroke, AIS)是脑卒中最主要的临床亚型, 我国 AIS 住院患者病死率或致

收稿日期: 2023-10-17

修回日期: 2023-11-13

\* 广西壮族自治区卫生健康委员会科研课题(Z20200212)资助。

【第一作者简介】

周辉(1980—), 男, 副主任医师, 主要从事脑血管病的临床和基础研究, E-mail: zhouhui7739@163.com。

【\*\*通信作者简介】

冯兵(1979—), 男, 副主任医师, 主要从事脑血管病的临床和基础研究, E-mail: 51658923@qq.com。

【引用本文】

周辉, 冯兵, 刘旭日, 等. 基于机器学习的静脉溶栓时间窗外急性缺血性脑卒中患者短期预后模型建立与验证[J]. 广西科学, 2024, 31(4): 781-787.  
ZHOU H, FENG B, LIU X R, et al. Establishment and Validation of a Short-term Prognostic Model for Acute Ischemic Stroke Patients Beyond the Time Window for Intravenous Thrombolysis Based on Machine Learning [J]. Guangxi Sciences, 2024, 31(4): 781-787.

残率高,分别达 34.5%和 37.1%<sup>[2]</sup>。AIS 患者最佳治疗方案是在溶栓时间窗内(发病至溶栓时间 $<6$  h)行静脉溶栓,但据我国国家卒中登记数据库数据显示,我国 AIS 患者静脉重组组织型纤溶酶原激活剂(Recombinant Tissue Plasminogen Activator, rt-PA)溶栓率仅为 1.23%<sup>[3]</sup>。可见,绝大部分 AIS 患者未能静脉溶栓。对于溶栓的 AIS 患者,其预后的影响因素较多,如血管再通情况、基础疾病等<sup>[4]</sup>;另外,人工神经网络(ANN)可有效地预测 AIS 患者预后,并识别出影响预后的关键因素<sup>[5]</sup>。但对于占比更高的未能溶栓的 AIS 患者,其预后影响因素的相关报道,尤其是使用机器学习或人工智能方法对患者预后进行预测的研究较少。因此,本研究通过回顾性研究 AIS 患者入院时的临床资料并结合机器学习分类(Classification)算法,对 AIS 患者短期预后进行研究,以尽早识别不良预后的高风险患者,为及早干预并改善预后提供科学依据。

## 1 材料与方法

### 1.1 研究对象

选择 2019 年 1 月至 2023 年 3 月在我院神经内科住院治疗的 AIS 患者 774 例。纳入标准:(1)符合 AIS 诊断标准(采用中华医学会制定的《中国急性缺血性脑卒中诊治指南 2018》<sup>[6]</sup>);(2)发病时间 $>6$  h;(3)未行静脉溶栓治疗患者。排除标准:(1)伴精神疾病、依从性差无法配合患者;(2)多发梗死昏迷患者;(3)伴遗传代谢性疾病、血液系统疾病患者;(4)伴重度心、肝、肾、肺功能障碍,多器官功能损伤及凝血功能障碍患者;(5)伴恶性肿瘤、颅脑外伤、颅内感染患者;(6)临床资料不全患者;(7)发病至入院时间(患者出现 AIS 症状至入院时间) $>2$  周;(8)患者年龄 $>90$  岁。

### 1.2 资料收集

收集患者入院时(治疗前)的临床资料,包括性别、年龄、身体质量指数(Body Mass Index, BMI)、美国国立卫生研究院卒中量表(National Institutes of Health Stroke Scale, NIHSS)评分、格拉斯哥昏迷(Glasgow Coma Scale, GCS)评分、发病-治疗时间(h)、梗死灶体积(CT 诊断,  $\text{mm}^3$ )、2 型糖尿病史(TMD2)、高血压史、卒中史、心绞痛史、心梗史、房颤史、现病史 $\geq 3$  种、血白细胞(WBC)计数、血中性粒细胞计数(NEU)、血淋巴细胞计数(LYM)、血中性粒细胞/淋巴细胞比值(NLR)、血单核细胞计数

(MON)、血小板计数(PLT)、凝血酶原时间(PT)、低密度脂蛋白(LDL)水平、高密度脂蛋白(HDL)水平、HDL-LDL 比值、血总胆固醇(TC)水平、血甘油三酯(TG)水平、血清总胆红素(T-BIL)水平、血 C 反应蛋白(C-Reactive Protein, CRP)水平、血半胱氨酸(Cysteine, CYS)水平。

### 1.3 预后评价

随访至出院 90 d,根据改良 Rankin 量表(modified Rankin Scale, mRS)界定短期预后是否良好:残疾( $\text{mRS} \geq 3$  分)或死亡为预后不良; $\text{mRS} < 3$  分为预后良好。

### 1.4 统计学方法

计量资料以均数 $\pm$ 标准差表示,非正态分布计量资料则以中位数[四分位数间距,  $M(P_{25}, P_{75})$ ]表示,计数资料以例数(百分比)表示。组间连续性资料比较采用  $t$  检验(正态分布,方差齐)或非参数检验(Mann-Whitney 检验),分类资料采用卡方检验或 Fisher's 精确检验。 $P < 0.05$  为差异显著,表示有统计学意义。

### 1.5 机器学习建模

使用 Python 编程语言(3.10),使用的库主要包括 NumPy(1.23.5)、Pandas(2.1.1),机器学习方法使用 PyCaret<sup>[7]</sup>(PyCaret 3.0,使用方法详见网址 <https://pycaret.org/>)。将 50%数据划分为训练样本(用于模型构建),余下 50%数据为验证样本(用于模型评估)。建模前对数据进行预处理,对连续型资料(数值型变量)进行归一化(Normalization)处理,对计数资料(分类变量)进行哑变量处理。利用受试者工作特征曲线(ROC)计算曲线下面积(Area Under Curve, AUC)来判断模型的诊断准确性,使用校准图(Calibration plot)评估预测风险与实际发生风险的一致性。

## 2 结果与分析

### 2.1 预后良好与预后不良的 AIS 患者临床资料比较

研究结果显示,774 例患者中,3 个月随访期内预后良好患者占比 86.05%(666 例),预后不良患者占 13.95%(108 例)。组间比较显示,预后良好与预后不良患者在入院时年龄、血 WBC 计数、血 CRP 水平、血 TG 水平、NIHSS 评分、GCS 评分、梗死灶体积和发病-治疗时间上均差异显著( $P < 0.05$ ,表 1)。

表 1 预后良好与预后不良 AIS 患者入院临床资料比较 [(均数±标准差)/n(%)]

Table 1 Comparison of admission clinical indicators between patients with good and poor AIS prognosis [(mean±std)/n(%)]

变量名 Variable name		预后良好(n=666) Good prognosis (n=666)	预后不良(n=108) Poor prognosis (n=108)	P 值 P value
Gender	Female	307(46.10%)	49(45.37%)	0.888
	Male	359(53.90%)	59(54.63%)	
Age/a		60.22 (54.50,65.29)	63.91 (58.01,69.64)	<0.001
BMI/(kg/m <sup>2</sup> )	Normal body weight	257(38.59%)	44(40.74%)	0.469
	Low weight	36(5.41%)	2(1.85%)	
	Overweight	252(37.84%)	42(38.89%)	
	Fat	121(18.17%)	20(18.52%)	
WBC/(×10 <sup>9</sup> /L)		7.73±2.68	7.08±2.38	0.012
NEU/(×10 <sup>9</sup> /L)		4.84±1.52	4.84±1.67	0.997
LYM/(×10 <sup>9</sup> /L)		1.68±0.55	1.65±0.60	0.617
NLR		3.16±5.20	3.34±1.80	0.508
MON/(×10 <sup>9</sup> /L)		0.31±0.10	0.33±0.10	0.147
PLT/(×10 <sup>9</sup> /L)		220.28±7.14	219.51±7.27	0.302
CRP/(mg/L)		9.85±3.28	12.27±3.70	<0.001
PT/s		13.44±5.63	13.31±5.40	0.810
LDL/(mmol/L)		3.52±1.02	3.57±1.15	0.705
HDL/(mmol/L)		0.72±0.23	0.73±0.24	0.636
HDL-LDL ratio		0.23±0.13	0.23±0.14	0.614
TC/(mmol/L)		6.34±2.01	6.36±2.16	0.937
TG/(mmol/L)		1.44±0.33	1.35±0.30	0.008
T-BIL/(mmol/L)		12.60±3.32	12.94±3.52	0.359
CYS/(μmol/L)		14.54±3.87	14.14±3.55	0.294
National Institutes of Health stroke scale score		10.00 (8.00,12.00)	12.00 (10.00,14.00)	<0.001
Glasgow coma scale score		9.00 (7.00,11.00)	8.00 (6.00,10.00)	<0.001
Infarct volume/mm <sup>3</sup>		64.97±20.82	74.40±26.08	<0.001
Onset-treatment time/h		7.82±0.74	8.32±1.20	<0.001
Type 2 diabetes	No	580(87.09%)	96(88.89%)	0.601
	Yes	86(12.91%)	12(11.11%)	
Hypertension	No	584(87.69%)	97(89.81%)	0.528
	Yes	82(12.31%)	11(10.19%)	
Stroke	No	606(90.99%)	97(89.81%)	0.694
	Yes	60(9.01%)	11(10.19%)	
Stenocardia	No	625(93.84%)	100(92.59%)	0.620
	Yes	41(6.16%)	8(7.41%)	
Myocardial infarction	Yes	638(95.80%)	104(96.30%)	>0.999
	No	28(4.20%)	4(3.70%)	
Atrial fibrillation	Yes	603(90.54%)	95(87.96%)	0.404
	No	63(9.46%)	13(12.04%)	
History of present illness≥3 kinds	Yes	651(97.75%)	103(95.37%)	0.181
	No	15(2.25%)	5(4.63%)	

## 2.2 模型构建与评价

本研究采用了目前机器学习中常用的 15 种分类学习算法,结果显示极限梯度提升(Extreme Gradient Boosting,XGB)模型的 AUC 数值为 0.81(表 2,

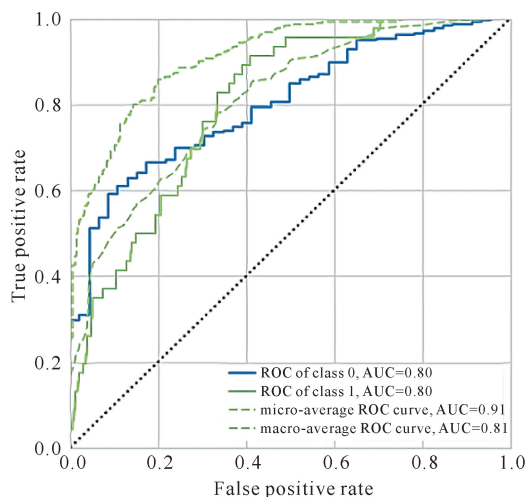
图 1),排在第一,比其他分类算法的 AUC 值高(表 2)。XGB 模型是一种基于决策树的集成学习算法,能有效处理数据中的异常值和缺失值,并通过正则化项和早停策略等技术有效地防止模型过拟合,提高了

模型的泛化能力。根据 XGB 模型输出预测变量影响权重(图 2), 筛选的预测变量影响权重居前 6 位的依次为发病-治疗时间、血 CRP 水平、年龄、梗死灶体积、血 TG 水平与 NIHSS 评分。为评估模型拟合情

况, 通过校准图对预后不良患者组进行校准, 结果显示图中校准线距离理想参考线(图中虚线)较近, 表明模型拟合性能稳健, 模型评估预测风险与实际发生风险的一致程度高(图 3)。

表 2 15 个机器学习分类算法的预测能力比较(验证集)

模型名称 Model name	准确率 Accuracy	曲线下面积 AUC	召回率 Recall	精确率 Precision	F1 分数 F1 score	卡帕系数 Kappa
Extreme Gradient Boosting (XGB)	0.87	0.81	0.28	0.78	0.41	0.35
Light Gradient Boosting (LGB)	0.86	0.80	0.32	0.63	0.40	0.34
CatBoost classifier	0.86	0.79	0.35	0.68	0.43	0.37
Random forest	0.84	0.78	0.11	0.35	0.16	0.13
Extra trees classifier	0.83	0.78	0.02	0.10	0.03	0.02
Gradient boosting decision tree	0.85	0.77	0.32	0.55	0.39	0.31
Logistic regression	0.82	0.76	0.27	0.43	0.30	0.22
Linear discriminant analysis	0.83	0.75	0.27	0.39	0.31	0.22
AdaBoost classifier	0.82	0.72	0.30	0.42	0.34	0.25
Naive bayes model	0.67	0.70	0.39	0.31	0.29	0.15
Decision tree	0.81	0.68	0.48	0.45	0.44	0.33
k-Nearest neighbor	0.82	0.57	0.07	0.20	0.10	0.06
Quadratic discriminant analysis	0.82	0.45	0.00	0.00	0.00	-0.02
SVM-Linear kernel	0.79	0.00	0.39	0.30	0.33	0.22
Ridge regression	0.83	0.00	0.21	0.47	0.28	0.21



Class 0 means favourable prognosis, class 1 means unfavourable prognosis.

图 1 XGB 模型预测溶栓时间窗外 AIS 患者短期不良预后的 ROC 曲线和 AUC(验证集)

Fig. 1 ROC curve and AUC (test dataset) of XGB model predicting short-term poor prognosis of AIS patients beyond the time window for thrombolysis

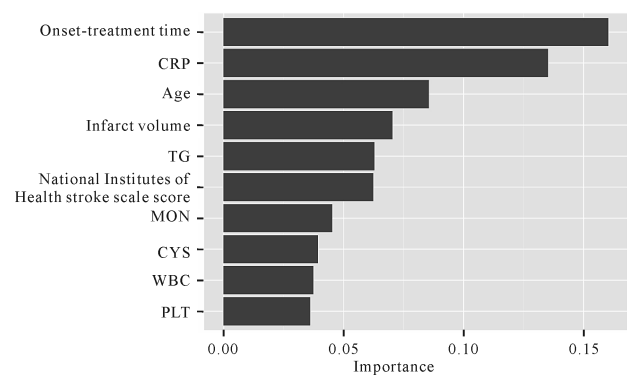


图 2 入院时临床指标影响权重(XGB 模型, 训练集)

Fig. 2 Importance of clinical indicators at admission (XGB model, train dataset)



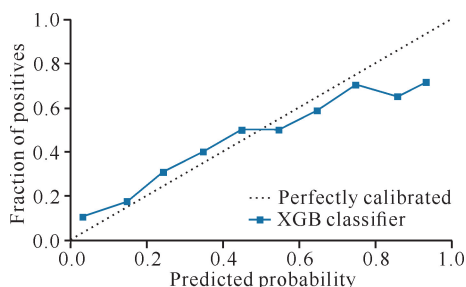


图3 XGB模型预测溶栓时间窗外AIS患者短期不良预后校准图(验证集)

Fig. 3 Calibration plot of short-term poor prognosis of AIS patients beyond the time window for thrombolysis predicted by the XGB model (test dataset)

### 3 讨论

机器学习是目前新颖且热门的人工智能技术,已有不少医学领域的研究将这一方法应用于疾病的诊断以及预后的预测中,并取得了不错的效果<sup>[8-10]</sup>。本研究利用机器学习分类算法,筛选出对AIS患者短期预后具有预测价值的入院临床资料。选择入院时临床资料在于尽早发现患者是否会发生不良预后,当预测不良预后的概率超过50%时,临床医生应提高警惕,并根据患者的具体情况及早干预,以改善预后。

通过比较各种机器学习分类算法发现,XGB模型效果最佳。筛选出的入院临床资料预测变量中,影响权重排名前6位的依次为发病-治疗时间、血CRP水平、年龄、梗死灶体积、血TG水平与NIHSS评分。发病-治疗时间影响权重最大,提示对于未能溶栓的AIS患者,时间窗意义重大,及早治疗可极大地提高病人的存活率,减少致残率,提高病人的生活质量。患者入院时血CRP水平对患者短期预后的影响权重排第二,该指标作为常见的炎症因子,可损伤血管,加重血管局部炎症反应,加重缺血<sup>[11]</sup>,是AIS患者短期预后的独立影响因素<sup>[12]</sup>。Derbisz等<sup>[13]</sup>研究也显示,血CRP $\geq 10$  mg/L是阿替普酶治疗的AIS患者发病后90 d内死亡的独立影响因子,这些结果提示对超出溶栓时间窗的AIS患者,应密切观察血CRP水平,并积极进行抗炎治疗,以改善患者预后。年龄是影响权重排第三的指标,年龄不只是AIS发病的危险因素,也是该病患者预后的重要影响因素,其具体机制可能在于年龄越大,脑组织损伤后修复能力越差,而且老年人往往存在着其他基础疾病,康复难度更大<sup>[14]</sup>。关于梗死灶体积对AIS患者的影响,国内学者报道梗死灶体积越大,神经功能缺损程度越

高<sup>[15]</sup>,这与本研究结果一致。另外,本研究发现血TG水平对患者预后影响较大,预后不良组患者入院血TG水平显著高于预后良好组患者,提示控制AIS患者的血TG水平有助于改善其预后;但血TG水平对AIS患者预后的影响存在争议,崔琴等<sup>[16]</sup>研究认为AIS患者的血TG水平是预后的保护性因素,并认为血TG对患者中枢神经有一定保护作用,只是其作用机制尚不清楚,但该研究未用患者数据验证,因此需要进一步研究。作为公认的影响AIS患者预后的指标,本研究建模结果同样认为入院时NIHSS评分对患者预后存在影响。

对于疾病预后的预测方面,机器学习分类算法相比传统的逻辑回归模型可更好地捕捉非线性关系,具有更好的鲁棒性和泛化能力,可更好地处理噪声和缺失值。但机器学习分类算法需对数据进行预处理(归一化和哑变量),结果的可解释性不佳,无法像逻辑回归模型直接量化影响因素的作用(计算OR值),进而判定因素为独立危险因素抑或是保护因素。这就需通过将机器学习分类算法置于基于浏览器的交互界面中,通过输入患者具体指标的数值,计算出预后不良的概率。此外,本研究为单中心研究,例数不够多,纳入的指标也不够全面和细化(如病因分型、梗死部位、肺部感染情况、治疗用药与时间等),由于纳入资料的时间跨度大,某些指标未能使用更为准确的检查方法测量,如梗死灶体积未能用磁共振扩散加权成像(MR DWI)计算,这些都是本研究的局限所在。因此,还需要在多中心研究中纳入更多、更细化的全面指标,继续深入研究,为超出溶栓时间窗的AIS患者预后判断提供临床依据。

### 4 结论

本研究通过应用各种机器学习分类算法,分析静脉溶栓时间窗外AIS患者入院时的临床资料。结果发现,XGB算法表现最佳,显示出较高的预测准确性和可靠性;患者短期预后受发病-治疗时间、血CRP水平、年龄、梗死灶体积、血TG水平与NIHSS评分的显著影响。

#### 参考文献

- [1] TU W J, ZHAO Z, YIN P, et al. Estimated burden of stroke in China in 2020 [J]. JAMA Network Open, 2023, 6(3): e231455.
- [2] GBD 2016 Neurology Collaborators. Global, regional, and national burden of neurological disorders, 1990 -

- 2016; a systematic analysis for the global burden of disease study 2016 [J]. *The Lancet Neurology*, 2019, 18(5):459-480.
- [3] 刘琴琴, 杨丽, 赵秋利, 等. 缺血性卒中溶栓就医延迟风险因素研究进展[J]. *中国全科医学*, 2019, 22(16):2000-2007.
- [4] JIN M, PENG Q, WANG Y. Post-thrombolysis early neurological deterioration occurs with or without hemorrhagic transformation in acute cerebral infarction: risk factors, prediction model and prognosis [J]. *Heliyon*, 2023, 9(5):e15620.
- [5] CHUNG C C, CHAN L, BAMODU O A, et al. Artificial neural network based prediction of postthrombolysis intracerebral hemorrhage and death [J]. *Scientific Reports*, 2020, 10(1):20501.
- [6] 中华医学会神经病学分会, 中华医学会神经病学分会脑血管病学组. 中国急性缺血性脑卒中诊治指南 2018[J]. *中华神经科杂志*, 2018, 51(9):666-682.
- [7] ZHAO Q Y, LIU L P, LUO J C, et al. A machine-learning approach for dynamic prediction of sepsis-induced coagulopathy in critically ill patients with sepsis [J]. *Frontiers in Medicine*, 2020, 7:637434.
- [8] LE S, PELLEGRINI E, GREEN-SAXENA A, et al. Supervised machine learning for the early prediction of acute respiratory distress syndrome (ARDS) [J]. *Journal of Critical Care*, 2020, 60:96-102.
- [9] AL'REF S J, SINGH G, VAN ROSENDAEL A R, et al. Determinants of in-hospital mortality after percutaneous coronary intervention: a machine learning approach [J]. *Journal of the American Heart Association*, 2019, 8(5):e011160.
- [10] 李惠萍, 胡安民. 机器学习 DNN 和 XGBoost 算法对危重患者预后预测模型效能评估[J]. *实用医学杂志*, 2020, 36(4):466-469.
- [11] 韩雪, 高燕军, 赵士娇. 血清正五聚蛋白 3、超敏 C 反应蛋白水平与急性脑梗死患者的病情严重程度及预后的关系[J]. *卒中与神经疾病*, 2019, 26(1):22-26.
- [12] 孙娜, 钟丽珍, 赵华, 等. 血清胱抑素 C、C-反应蛋白和同型半胱氨酸联合预测脑梗死患者短期预后的价值[J]. *心脑血管病防治*, 2021, 21(6):567-570.
- [13] DERBISZ J, NOWAK K, WNUK M, et al. Prognostic significance of stroke-associated infection and other readily available parameters in acute ischemic stroke treated by intravenous thrombolysis [J]. *Journal of Stroke and Cerebrovascular Diseases*, 2021, 30(2):105525.
- [14] ZHAO Y, HAN Y, SUN W, et al. Clinical symptoms, etiology and prognosis of acute bilateral posterior circulation cerebral infarction [J]. *International Journal of General Medicine*, 2022, 15:2787-2793.
- [15] 郝泽林, 吴建跃, 滕振飞, 等. 急性脑梗死患者血清学指标与神经功能缺损和脑梗死体积及颈动脉斑块性质的相关性研究[J]. *中华全科医学*, 2020, 18(11):1803-1806.
- [16] 崔琴, 湛彦强, 胡丹, 等. 非空腹甘油三酯与急性脑梗死预后的关系[J]. *神经损伤与功能重建*, 2019, 14(11):547-550.

## Establishment and Validation of a Short-term Prognostic Model for Acute Ischemic Stroke Patients Beyond the Time Window for Intravenous Thrombolysis Based on Machine Learning

ZHOU Hui<sup>1</sup>, FENG Bing<sup>2\*</sup>, LIU Xuri<sup>1</sup>, ZOU Yong<sup>1</sup>, TANG Hui<sup>1</sup>, HUANG Qinbin<sup>2</sup>, CHEN Tingjun<sup>2</sup>, ZOU Donghua<sup>3</sup>

(1. Traditional Chinese Medicine Hospital of Guiping City, Guiping, Guangxi, 537200, China; 2. Guiping People's Hospital, Guiping, Guangxi, 537200, China; 3. The Second Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, 530007, China)

**Abstract:** The objective of this study is to investigate the effects of clinical data at admission on the short-term adverse outcomes of Acute Ischemic Stroke (AIS) patients beyond the time window for intravenous thrombolysis. The clinical data of 774 AIS patients beyond the time window for thrombolysis treated in the department of neurology in the Traditional Chinese Medical Hospital of Guiping City were collected. The clinical indicators at admission of patients with different short-term prognoses were compared. Machine learning classi-

fication algorithms were used to model and analyze the factors influencing poor prognosis within 90 d. The prediction performance and accuracy of the nomogram model were evaluated based on the Receiver Operating Characteristic (ROC) curve and calibration plot. The results showed: (1) 13.95% (108 cases) of the patients had poor prognosis within the 3-month follow-up period. (2) The patients with poor prognosis had higher age, levels of White Blood Cell (WBC), C-Reactive Protein (CRP), and Triglyceride (TG), National Institutes of Health Stroke Scale (NIHSS) score at admission, infarct volume, and onset-to-treatment time than the patients with good prognosis. The patients with poor prognosis had lower Glasgow Coma Scale (GCS) score at admission than the patients with good prognosis ( $P < 0.05$ ). The Extreme Gradient Boosting (XGB) algorithm performed best, with the area under the ROC curve (AUC) of 0.81, and the calibration curve showed a high degree of consistency between the predicted and actual occurrence of risk. (3) The top six predictive variables selected by the XGB algorithm were onset-to-treatment time, CRP level at admission, age, infarct volume, TG level, and NIHSS score. These results indicate that the XGB algorithm can be used to predict the factors influencing short-term poor prognosis in AIS patients beyond the time window for intravenous thrombolysis.

**Key words:** acute ischemic stroke; beyond the time window for thrombolysis; prognosis; machine learning; clinical prediction model

责任编辑:米慧芝

(上接第 780 页 Continued from page 780)

## Short-term Wind Power Forecasting Based on VMD-GRAU

XIE Jincal<sup>1,2</sup>, YU Qiancheng<sup>1,2\*\*</sup>, WANG Zhici<sup>1,2</sup>, HU Zhiyong<sup>1,2</sup>, YU Xulong<sup>1,2</sup>,  
WANG Aoqiang<sup>1,2</sup>

(1. School of Computer Science and Engineering, North Minzu University, Yinchuan, Ningxia, 750021, China; 2. Laboratory of Graphics and Images of the State Ethnic Affairs Commission, North Minzu University, Yinchuan, Ningxia, 750021, China)

**Abstract:** The existing short-term wind power prediction models are difficult to accurately capture the nonlinear mapping relationship between wind power historical monitoring data and wind power and adapt to the environmental changes in wind speed, which lead to poor prediction results. In view of these problems, a wind power prediction model was designed based on Variational Mode Decomposition (VMD) and Recurrent Neural Network (RNN). The model first performed VMD of the wind power data to avoid the problems of mode mixing and false peaks, and proposed the Gated Recurrent Attention Unit (GRAU) model, which designed an attention gate to enhance the ability of RNN to capture the important sequential features and robustness. An error correction module was designed to reduce the effects of stochasticity and volatility in wind power prediction, and finally a sparse regularization term was added in the loss function to prevent the model from being overfitted. The experimental results showed that the proposed VMD-GRAU model achieved the average values of 0.022, 0.016, and 0.995 in terms of Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and  $R^2$  score, respectively. The model proposed in this study improved the MAE by 26%–31% compared with GRU and Transformer models and had good generalization performance in different environments.

**Key words:** wind power prediction; variational mode decomposition; attention mechanism; regularization

责任编辑:陆雁