

# 人工智能在脑卒中疾病预测研究中的应用

王子敏, 陆 丹

湖北省中医院(湖北中医药大学附属医院), 康复医学中心, 湖北 武汉

收稿日期: 2025年8月19日; 录用日期: 2025年9月13日; 发布日期: 2025年9月25日

## 摘 要

基于人工智能(AI)的决策支持系统可以让医生快速评估脑卒中患者数据,并在初始阶段更准确地预测脑卒中可能引发的危重疾病。为了提高脑卒中疾病预测的准确性,多种基于人工智能的决策支持系统被提出,本文中,我们首先对脑卒中分类、症状及危险因素等情况进行了简单介绍,然后对人工智能在脑卒中疾病预测研究中的应用做了详细的说明,同时通过比较不同机器学习方法的优劣性,并列举国内外脑卒中预测的具体研究案例,对各种基于人工智能的决策支持系统进行了批判性分析,为预测脑卒中提供了相应的研究方法。本文进一步扩展讨论了深度学习方法(CNN、RNN/LSTM、GNN)的应用,并提出了模型可解释性(XAI)和隐私保护计算(联邦学习)在临床转化中的关键意义。

## 关键词

脑卒中预测, 机器学习, 深度学习, 模型可解释性, 联邦学习, 决策支持系统

# Application of Artificial Intelligence in Stroke Prediction Research

Zimin Wang, Dan Lu

Rehabilitation Medicine Center, Hubei Provincial Hospital of Traditional Chinese Medicine (Affiliated Hospital of Hubei University of Chinese Medicine), Wuhan Hubei

Received: Aug. 19<sup>th</sup>, 2025; accepted: Sep. 13<sup>th</sup>, 2025; published: Sep. 25<sup>th</sup>, 2025

## Abstract

AI-based decision support systems enable doctors to quickly evaluate stroke patient data and more accurately predict critical illnesses that may be caused by stroke in the initial stage. To improve the accuracy of stroke prediction, a variety of AI-based decision support systems have been proposed. In this paper, we first briefly introduce the classification, symptoms, and risk factors of stroke, then elaborate on the application of AI in stroke prediction research. Meanwhile, by comparing the advantages

and disadvantages of different machine learning methods and listing specific domestic and international research cases on stroke prediction, we conduct a critical analysis of various AI-based decision support systems and provide corresponding research methods for stroke prediction. This paper further expands the discussion on the application of deep learning methods (CNN, RNN/LSTM, GNN) and points out the key significance of model interpretability (XAI) and privacy-preserving computing (federated learning) in clinical translation.

## Keywords

Stroke Prediction, Machine Learning, Deep Learning, Model Interpretability, Federated Learning, Decision Support Systems

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. AI 在脑卒中预测领域的应用概述

脑卒中是一种医疗疾病, 当为大脑提供氧气和营养的动脉血管堵塞或受损时就会发生脑卒中。由于大脑特定区域的血流受限或中断, 这些区域的细胞缺乏必要的氧气和营养, 从而死亡。脑卒中的后遗症可轻可重, 大多数脑卒中患者永远无法完全康复, 终生带着中度或重度残疾[1]。基于人工智能(AI)的决策支持系统有助于对脑卒中中进行早期预测。决策支持系统(DSS)是一种辅助决策的工具, 决策支持系统对大量非结构化数据进行分析, 以求获得有可能解决问题和帮助决策的洞察力[2] [3]。

DSS 由三个核心部分组成: 数据库管理系统、模型管理系统和用户界面。

- **数据管理系统:** 为协助决策, 该系统收集来自内部和外部的数据。
- **模型管理系统:** 它是一种控制决策支持系统工作方式的软件, 由多个数据处理和分析模型组成。
- **用户界面:** 使系统导航更容易, 它以图表等多种形式展示调查结果。

人工智能、大数据、机器学习、物联网和云计算等技术是第四次工业革命的关键技术, 这些技术在医疗健康领域受到了广泛关注[4]-[6]。由于人工智能在医疗健康领域的广泛应用相对较晚, 有关人工智能在疾病诊断、药物研发、患者护理监测等医疗保健和医学领域的应用研究仍在进行中。

早期脑卒中诊断可以减少残疾和死亡率, 医疗决策支持系统可以让医生快速评估患者数据并做出诊断。在初步医疗数据和测试的基础上, 决策支持系统可以帮助早期发现脑卒中。决策支持系统中使用的机器学习算法可提高预测准确性, 同时减少人为误差[7] [8]。

## 2. 脑卒中分类、症状及危险因素

### 2.1. 脑卒中分类

缺血性脑卒中、出血性脑卒中和短暂性脑缺血发作是脑卒中的三种不同形式。

- **缺血性脑卒中:** 缺血是一种由血流减少引起的疾病。将血液和营养物质从心脏输送到大脑的动脉可能会硬化或堵塞, 从而影响血液和营养物质到达大脑的数量。这种情况会导致缺血性脑卒中。脂质、胆固醇和其他物质会在动脉壁内部和表面堆积, 这种情况被称为斑块。斑块堆积导致动脉狭窄, 阻碍血液流动。当斑块破裂时, 就会形成血栓[9]。

- **出血性脑卒中:** 脑血管破裂或渗漏引起的出血会导致出血性脑卒中。出血会增加脑部压力, 从而损伤特定脑区。高血压是导致出血性脑卒中的主要原因之一。与年龄有关的动脉血管衰弱是导致出血性

脑卒中的另一个因素[10]。

● **短暂性脑缺血发作(TIA):** 短暂性脑缺血发作有时被称为脑卒中或小脑卒中。它是即将发生脑卒中的征兆。它只会持续很短的时间。它与脑卒中一样, 是一种危及生命的紧急医疗状况。短暂性脑缺血发作通常由血栓引起。当大脑的一部分血流暂时中断时就会发生这种情况。短暂性脑缺血发作的症状与脑卒中相似, 但持续时间较短[11]。

## 2.2. 脑卒中症状

脑卒中的几个征兆如下[12]。

- 没有感觉或不适, 尤其是身体一侧的手臂、腿部或面部。
- 单眼或双眼视力受损或模糊。
- 走路焦躁不安。
- 说话困难。
- 协调问题。
- 无明确原因的剧烈头痛。

## 2.3. 导致脑卒中的危险因素

- **年龄:** 年龄是脑卒中的首要风险因素。超过 55 岁, 患脑卒中的几率会增加一倍[13]。
- **性别:** 脑卒中对男性和女性都有影响。女性平均寿命较长, 高龄脑卒中发病率较高, 因此女性脑卒中发病率较高[13]。
- **高体重指数:** 高体重指数是 1990 年至 2019 年间增长最快的脑卒中危险因素[13]。
- **高血压:** 高血压会影响脑血流量。由于高血压导致脑血管增厚和改变, 血管外径减小。高血压患者发生脑卒中的风险增加了三至四倍[14]。
- **吸烟:** 吸烟会影响心率、血压和血栓形成等生理过程, 从而增加脑卒中的风险。根据不同来源的研究, 与终生吸烟者或十多年前戒烟的人相比, 经常吸烟者发生脑卒中的几率至少高出 2 到 4 倍[13]。
- **胆固醇:** 血液中胆固醇水平过高会导致脂肪沉积在动脉中, 动脉可能会收紧和变硬, 使血液流动更加困难。此外, 它还会增加血液凝结的风险[13]。
- **血压:** 血压超过约 115/75 mm Hg 时, 脑卒中风险会逐渐增加。血压与脑卒中密切相关, 而且大多数成年人的血压值过高, 这也是造成全球约三分之二脑卒中负担的根本原因[13]。
- **糖尿病:** 在脑卒中治疗期间和治疗后, 糖尿病患者哮喘和尿路感染的发病率更高。糖尿病引起脑卒中的原因有很多, 如血糖控制不足、缺乏活动、血浆中炎症敏感蛋白水平升高、2 型糖尿病患者的药物不良反应等[13]。

## 3. 预测脑卒中的机器学习算法及面临的问题

### 3.1. 几类预测脑卒中的机器学习算法

脑卒中病因较为复杂, 多种危险因素并存, 并且脑卒中病人的身高、年龄、体重等信息也各不相同, 同时表现出的症状也有所差异。面对这么多的信息标签, 如何对脑卒中进行准确预测, 就需要借助基于人工智能(AI)的决策支持系统。通常运用在脑卒中预测中的机器学习算法属于监督类学习算法, 常见的包括支持向量机(SVM)、朴素贝叶斯(Naive Bayesian)、K 均值算法(K-Means)、AdaBoost 算法(Adaptive Boosting)、逻辑回归(Logistic Regression)、决策树算法(Decision Tree)和随机森林法(Random Forests)等[14]-[16]。

### 3.2. 深度学习算法的应用

随着医疗数据体量和复杂度的增加,深度学习在脑卒中预测中的应用日益广泛,主要包括以下几类:

#### 卷积神经网络(CNN)

CNN 广泛用于医学影像分析,如 CT、MRI。其自动特征提取能力可识别缺血性或出血性卒中的早期征象。例如,研究显示基于 CNN 的影像模型在梗死灶分割和出血检测中准确性优于人工判读,并可通过热力图直观显示模型关注区域。

#### 循环神经网络(RNN)及长短期记忆网络(LSTM)

RNN/LSTM 适合处理时序数据,在脑卒中预测中主要用于电子健康记录(EHR)、动态血压、血糖监测数据的建模。LSTM 能捕捉长期依赖特征,实现对卒中发生或复发风险的动态预测。例如,有学者开发了基于 EHR 的 LSTM 模型,对卒中复发风险的预测表现优于传统 Cox 回归模型。

#### 图神经网络(GNN)

脑卒中涉及多种危险因素和患者间复杂的关联关系。GNN 可以建模“患者-疾病-药物”的图结构,揭示多因素间潜在交互关系。部分研究利用 GNN 对患者关系网络进行建模,实现了卒中分型和并发症预测的性能提升。

这些深度学习方法在处理不同模态数据(影像、EHR、关系网络)时各具优势,为脑卒中预测提供了新的研究路径。

### 3.3. 机器学习预测脑卒中面临的问题

上述机器学习算法的准确性及有效性虽然在机器视觉、计算机安全、计算生物学等诸多人工智能领域得到了充分的验证,但是如何将其应用在对脑卒中疾病的预测时,仍然需要解决以下几个问题。

- 1) 数据不平衡是研究人员面临的问题。当一个类的实例多于另一个类时,就会出现类不平衡。如果不用不平衡的数据集来训练模型,它的表现就不会好。
- 2) 训练数据集包含各种不同的特征类型和数值范围,现有数据可能存在样本选择偏差,例如某些人群或疾病亚型的数据不足,导致模型预测结果存在偏差。
- 3) 医疗数据存在缺失、不一致、标注错误等问题,不同机构的数据标准不统一,影响模型训练和泛化能力,同时对预测准确性也会造成影响。
- 4) 数据集中可能存在不重要的属性,从而降低系统性能。因此,在训练模型之前,必须使用特定的特征选择策略来降低维度。由于每种算法都可能为同一数据集创建一组不同的特征,因此合适的特征选择方法是非常困难的。
- 5) 要阻止这些疾病恶化,早期检测和预后评估至关重要。机器学习技术可用于生成具有不同准确度的模型,选择最佳的机器学习技术非常困难。
- 6) 脑卒中的症状和危害与其他疾病类似,很难确定某种症状是由脑卒中还是其他疾病引起的。
- 7) 在某个数据集上,会对灵敏度、特异性和准确性等性能指标进行评估,每个数据集的性能指标可以有不同的取值范围。

### 4. 各种脑卒中预测实例与比较

研究人员利用不同的机器学习技术提出了许多预测脑卒中的决策支持系统。

Kim 等[17]采用基于自然语言处理(NLP)的算法从核磁共振成像文本报告中识别急性缺血性脑卒中患者。在数据分析中,作者使用了 3024 份 MRI 文档。采用 SVM、单一决策树(SDT)、二元逻辑回归(BLR)和 NB 分类器进行分类。

Saleh 等[18]开发了一个脑卒中预测系统。预测系统架构的前五个步骤包括数据集加载、数据预处理、交叉验证和超参数调整、分类和性能评估。脑卒中预测使用了 Kaggle 的医疗数据集。SVM、决策树(DT)、逻辑回归(LR)和 RF 分类器被用于分类。

Govindarajan 等利用人工神经网络(ANN)、SVM、RF、LR、Boosting 和 Bagging 技术进行脑卒中预测。利用标记和最大熵技术从原始数据中检索出有意义的信息。为了剔除重复信息,对收集到的数据进行了相关性分析。

Bandi 等[19]提出了一种基于简易射频的脑卒中预测模型。在研究中,作者使用了包含 4799 条记录的数据集。他们采用了 LR、自适应提升、线性 SVM、DT、Poly SVM、高斯直觉贝叶斯、RBG SVM、RF 和简易 RF 机器学习分类器来确定脑卒中风险。

Tazin 等[1]使用四种机器学习技术 DT、LR、RF 和投票分类器来估计脑卒中的可能性。作者对准确率、F1 分数、召回率和精确度等性能指标进行了评估。

Sailasya 等[20]利用 Kaggle 5110 人的数据集开发了一个脑卒中预测模型。为了预测脑卒中,他们使用了 LR、RF、DT、KNN、SVM 和 NB 机器学习算法进行分类。

Kaur 等[21]创建了一个脑卒中早期诊断框架。建议的框架使用了以下四种深度学习方法:前馈神经网络(FFNN)、双向 LSTM (biLSTM)、LSTM 和 GRU。评估深度学习方法时使用了多个性能指标。

Dev 等[22]使用了 29,072 名患者的电子健康记录数据集。作者使用特征相关性分析来选择最佳特征集。在脑卒中预测方面,作者采用了神经网络(NN)、DT 和 RF 等机器学习技术。就准确性而言,神经网络优于其他机器学习分类器。

## 5. 讨论

近年来,人工智能技术在脑卒中预测中展现出巨大潜力,但仍存在诸多挑战。基于前文分析,未来研究可从以下几个方向深化。

### 5.1. 针对特定脑卒中亚型的多模态融合预测模型

隐源性卒中(cryptogenic stroke)因病因复杂和诊断困难,长期以来是临床预测的难点。未来研究可整合医学影像、基因组学、血液生物标志物及临床检查指标,构建多模态融合预测模型。通过深度学习实现不同数据源的协同分析,有望提高隐源性卒中的早期预测能力,推动精准医学的发展。

### 5.2. 基于动态 EHR 数据的卒中复发风险实时预警系统

卒中复发是影响患者长期预后的核心问题。依托电子健康记录(EHR)中的动态随访数据,利用 LSTM、Transformer 等时序模型,可以实时更新患者的风险评估。通过构建智能预警系统,临床医生能够在复发高危时段提前干预,制定个性化的随访与康复方案。

### 5.3. 利用可穿戴设备数据的社区卒中初筛轻量化模型

随着智能手环、血压计、心电贴片等可穿戴设备的普及,大量社区人群健康数据得以实时采集。未来可基于这些低维度但连续性强的数据,构建轻量化的深度学习模型,实现大规模卒中风险初筛。此类模型在基层医疗和公共卫生管理中具有广阔应用前景。

### 5.4. 跨地域联邦学习框架促进多中心协作

脑卒中预测需要大规模、多样化的数据支持。然而数据孤岛和隐私保护问题限制了模型的训练与推广。未来可构建跨地域、多中心的联邦学习框架,实现数据不出院所的前提下进行联合建模。结合差分

隐私和安全多方计算,能够在确保隐私安全的同时提升模型泛化能力。

## 5.5. 模型可解释性与临床整合度的提升

AI 预测模型若缺乏可解释性,将难以获得临床医生与患者的信任。未来应加强 XAI 技术的应用,通过可视化热力图、特征权重排名、因果推断方法等手段,帮助医生理解预测逻辑。同时应开发临床可交互的平台,使 AI 预测结果与医生经验形成互补,提高 AI 系统在真实临床中的可用性和整合度。

综上所述,未来研究应从“精准化、实时化、社区化和可解释化”四个维度入手,推动人工智能在脑卒中预测中的真正临床落地,助力早期预防与个体化康复管理。

## 6. 模型可解释性与隐私保护计算在临床转化中的关键作用

### 6.1. 模型可解释性(XAI)

- 1) 临床信任:通过可视化(如 Grad-CAM、SHAP)展示模型关注区域和特征权重,提升医生对预测结果的理解与信任。
- 2) 决策支持:解释性结果帮助医生快速识别高危因素(如房颤、高血压),与医学知识互补。
- 3) 机制探索:揭示数据中潜在的新型病因关联,为卒中发病机制研究提供线索。
- 4) 法规合规:透明的 AI 系统更符合医疗监管要求。

### 6.2. 隐私保护计算与联邦学习(FL)

- 1) 数据本地化:各中心在本地训练模型,仅共享参数更新,不泄露原始数据。
- 2) 泛化能力提升:整合多中心异质性数据,提升模型在真实世界中的适用性。
- 3) 合规与安全:结合差分隐私、多方安全计算,满足 GDPR 与《个人信息保护法》的要求。
- 4) 真实世界研究推动:促进跨区域、跨机构的卒中预测模型研发,加速科研成果的临床落地。

### 6.3. 临床转化意义

XAI 和联邦学习是脑卒中预测模型进入临床实践的两大支撑:前者解决“可信度”问题,后者解决“数据与隐私”问题。二者相辅相成,将共同推动卒中预测模型从实验室研究走向临床应用,实现早期预警、精准干预和个体化康复。

## 7. 结论

人工智能,尤其是深度学习方法(如用于影像分析的卷积神经网络 CNN、用于时序 EHR 建模的 RNN/LSTM、以及用于建模复杂关系网络的图神经网络 GNN),在脑卒中风险预测与诊断领域已显示出显著潜力。本文回顾了这些方法在影像、电子健康记录和患者关系图谱等多模态数据上的典型应用,并指出了当前研究面临的数据不平衡、数据质量、特征冗余与模型泛化能力不足等关键挑战。为促进临床转化,模型可解释性(XAI)与隐私保护计算(如联邦学习)是两项不可或缺的支持技术:前者有助于提升临床信任与决策透明度,后者使得多中心协作在不泄露患者隐私的前提下成为可能。基于此,我们建议未来研究重点围绕:一是开发针对卒中亚型的多模态融合预测模型,二是构建基于动态 EHR 的实时复发预警系统,三是探索可在可穿戴设备和基层医疗中部署的轻量化筛查模型,并在多中心大样本上开展严格的外部验证与伦理合规评估。总体而言,只有在算法性能、可解释性、隐私保护与临床可用性等多方面协同推进时,人工智能才能真正推动脑卒中预测从研究走向临床应用,从而实现早期预警、精准干预与改善患者长期结局。

## 参考文献

- [1] Tazin, T., Alam, M.N., Dola, N.N., Bari, M.S., Bourouis, S. and Monirujjaman Khan, M. (2021) Stroke Disease Detection and Prediction Using Robust Learning Approaches. *Journal of Healthcare Engineering*, **2021**, Article ID: 7633381. <https://doi.org/10.1155/2021/7633381>
- [2] 鲁君. 人工智能在急性缺血性脑卒中早期诊断及预后预测中的应用研究[D]: [博士学位论文]. 武汉: 华中科技大学, 2023.
- [3] 陈思玓, 刘欢, 黄馨莹, 李皓琳, 谷鸿秋, 姜勇. 基于机器学习算法构建缺血性卒中 3 个月死亡预测模型研究[J]. 中国卒中杂志, 2020, 15(6): 595-599.
- [4] Sachdeva, R.K., Bathla, P., Rani, P., Kukreja, V. and Ahuja, R. (2022) A Systematic Method for Breast Cancer Classification Using RFE Feature Selection. *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Greater Noida, 28-29 April 2022, 1673-1676. <https://doi.org/10.1109/icacite53722.2022.9823464>
- [5] 傅璠, 卢洁. 人工智能在缺血性脑卒中影像的研究进展[J]. 中华老年心脑血管病杂志, 2020, 22(10): 1110-1112.
- [6] 孟钰婷, 徐姣娜, 裴可凡, 牛国忠. 深度学习在缺血性脑卒中的应用[J]. 浙江临床医学, 2022, 24(12): 1892-1894.
- [7] Kakkar, L., Gupta, D., Saxena, S. and Tanwar, S. (2019) An Analysis of Integration of Internet of Things and Cloud Computing. *Journal of Computational and Theoretical Nanoscience*, **16**, 4345-4349. <https://doi.org/10.1166/jctn.2019.8523>
- [8] 顾晓倩, 王云玲. 多参数 MRI 人工智能分析在缺血性脑卒中的研究进展[J]. 新疆医学, 2025, 55(1): 19-21, 32.
- [9] Sirsat, M.S., Fermé, E. and Câmara, J. (2020) Machine Learning for Brain Stroke: A Review. *Journal of Stroke and Cerebrovascular Diseases*, **29**, Article ID: 105162. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.105162>
- [10] Parmet, S. (2004) Hemorrhagic Stroke. *JAMA*, **292**, 1916. <https://doi.org/10.1001/jama.292.15.1916>
- [11] Lioutas, V., Ivan, C.S., Himali, J.J., Aparicio, H.J., Leveille, T., Romero, J.R., et al. (2021) Incidence of Transient Ischemic Attack and Association with Long-Term Risk of Stroke. *JAMA*, **325**, 373-381. <https://doi.org/10.1001/jama.2020.25071>
- [12] 《中国脑卒中防治报告 2019》概要[J]. 中国脑血管病杂志, 2020, 17(5): 272-281.
- [13] Donkor, E.S. (2018) Stroke in the 21st Century: A Snapshot of the Burden, Epidemiology, and Quality of Life. *Stroke Research and Treatment*, **2018**, Article ID: 3238165. <https://doi.org/10.1155/2018/3238165>
- [14] 王宁, 王浩莲, 徐苗. 半监督学习研究综述[J]. 电脑知识与技术, 2024, 20(6): 24-26.
- [15] 张润, 王永滨. 机器学习及其算法和发展研究[J]. 中国传媒大学学报(自然科学版), 2016, 23(2): 10-18, 24.
- [16] 李旭然, 丁晓红. 机器学习的五大类别及其主要算法综述[J]. 软件导刊, 2019, 18(7): 4-9.
- [17] Kim, C., Zhu, V., Obeid, J. and Lenert, L. (2019) Natural Language Processing and Machine Learning Algorithm to Identify Brain MRI Reports with Acute Ischemic Stroke. *PLOS ONE*, **14**, e0212778. <https://doi.org/10.1371/journal.pone.0212778>
- [18] Saleh, H., Abdel Ghany, S.F., Younis, E., Omran, N. and Ali, A. (2019) Stroke Prediction Using Distributed Machine Learning Based on Apache Spark. *International Journal of Advanced Science and Technology*, **28**, 89-97.
- [19] Bandi, V., Bhattacharyya, D. and Midhunchakkravarthy, D. (2020) Prediction of Brain Stroke Severity Using Machine Learning. *Revue d'Intelligence Artificielle*, **34**, 753-761. <https://doi.org/10.18280/ria.340609>
- [20] Sailasya, G. and Kumari, G.L.A. (2021) Analyzing the Performance of Stroke Prediction Using ML Classification Algorithms. *International Journal of Advanced Computer Science and Applications*, **12**, 539-545. <https://doi.org/10.14569/ijacsa.2021.0120662>
- [21] Kaur, M., Sakhare, S.R., Wanjale, K. and Akter, F. (2022) Early Stroke Prediction Methods for Prevention of Strokes. *Behavioural Neurology*, **2022**, Article ID: 7725597. <https://doi.org/10.1155/2022/7725597>
- [22] Dev, S., Wang, H., Nwosu, C.S., Jain, N., Veeravalli, B. and John, D. (2022) A Predictive Analytics Approach for Stroke Prediction Using Machine Learning and Neural Networks. *Healthcare Analytics*, **2**, Article ID: 100032. <https://doi.org/10.1016/j.health.2022.100032>