

# 基于模拟数据的机器学习脑动静脉畸形介入栓塞术后再通风险预测模型构建及性能比较

孙乐飞, 逢锦峰\*, 董浩宇

北华大学附属医院神经外科, 吉林 吉林

收稿日期: 2026年1月12日; 录用日期: 2026年2月6日; 发布日期: 2026年2月24日

## 摘要

本研究旨在构建一种基于模拟数据的机器学习模型,用于预测脑动静脉畸形(AVM)患者术后的再通风险。研究参考了10~20篇近期已发表的相关文献,构建了涵盖患者基本信息、病灶特征、血管构筑特征、临床病史以及治疗信息等多维度模拟数据集。采用Logistic Regression、Random Forest、XGBoost、LightGBM以及Balanced Random Forest这五种机器学习模型进行训练与比较,同时运用SHAP、热图等方法开展特征分析与可视化工作。结果表明,一项基于机器学习算法的研究显示,XGBoost模型在分类问题上的性能评估中表现出色,其AUC值达到0.98, F1分数为0.81,优于其他模型。本研究证实,基于模拟数据构建的机器学习模型能够有效识别AVM栓塞术后再通的高危因素,为临床个性化治疗决策提供可靠的辅助预测工具。

## 关键词

脑动静脉畸形, 介入栓塞, 再通风险, 机器学习, 预测模型, XGBoost, 模拟数据

# Construction and Performance Comparison of Machine Learning-Based Prediction Models for Recanalization Risk Following Interventional Embolization of Brain Arteriovenous Malformations Using Simulated Data

Lefei Sun, Jinfeng Pang\*, Haoyu Dong

\*通讯作者。

文章引用: 孙乐飞, 逢锦峰, 董浩宇. 基于模拟数据的机器学习脑动静脉畸形介入栓塞术后再通风险预测模型构建及性能比较[J]. 临床医学进展, 2026, 16(2): 3053-3067. DOI: 10.12677/acm.2026.162718

## Abstract

This study aims to construct a machine learning model based on simulated data to predict the risk of recanalization after interventional embolization in patients with brain arteriovenous malformations (AVMs). The study references 10 to 20 recently published relevant literature to construct a multi-dimensional simulated dataset covering patient basic information, lesion characteristics, vascular architecture features, clinical history, and treatment information. Five machine learning models—Logistic Regression, Random Forest, XGBoost, LightGBM, and Balanced Random Forest—were used for training and comparison, while SHAP, heatmaps, and other methods were employed for feature analysis and visualization. The results indicate that a study based on machine learning algorithms demonstrates that the XGBoost model excels in the performance evaluation of classification problems, with an AUC value of 0.98 and an F1 score of 0.81, outperforming other models. This study confirms that machine learning models constructed based on simulated data can effectively identify high-risk factors for recanalization after AVM embolization, providing a reliable auxiliary predictive tool for personalized clinical treatment decision-making.

## Keywords

Brain Arteriovenous Malformation, Interventional Embolization, Recanalization Risk, Machine Learning, Predictive Model, XGBoost, Simulated Data

Copyright © 2026 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. 引言

脑动静脉畸形(Brain Arteriovenous Malformations, bAVM)作为一种先天性脑血管发育异常病症,是引发颅内出血、癫痫及神经功能障碍的重要因素[1]。介入栓塞术作为 AVM 的主要治疗手段之一,因其微创、可重复的特点被广泛应用[2]。然而,术后脑血管畸形治疗后的复发率是一个不容忽视的关键问题,可能导致再次出血、需二次干预,进而增加医疗负担与患者痛苦。

目前,临床对 AVM 栓塞术后再通风险的预测多依赖于经验判断及有限的单一影像学或临床因素,如 Spetzler-Martin (SM)分级、畸形团大小等[3]。尽管传统方法在医疗领域具有一定的参考价值,但它们在预测精度、综合多维度信息以及量化个体风险方面存在局限。相比之下,机器学习技术的应用已经显著提高了医疗诊断的准确性,能够处理和分析大量多维度数据,并有助于个性化风险评估[4]。随着医疗数据维度的增加,迫切需要一种能够整合多元特征、提供高精度个性化风险预测的工具。

随着信息技术的飞速发展和大数据时代的到来,机器学习(Machine Learning, ML)技术在医疗风险预测领域展现出巨大潜力。例如,利用深度学习算法,科学家们成功地预测了疾病的发展风险,这一突破有望在临床实践中用于提前预警和个性化治疗,为患者提供更好的医疗保健服务。此外,机器学习算法在医疗预测中的应用还包括了对心脏病、糖尿病和肺癌等疾病风险的预测[5],这些研究显示了机器学习在提高早期干预准确性和有效性方面的巨大潜力。例如,随着医疗科技的快速发展,基于机器学习的疾

病风险评估方法在临床实践中得到广泛应用。该方法通过描述性分析、预测分析和个性化推荐,能够有效识别高风险个体,从而采取针对性的预防措施,降低疾病发生的概率。此外,它有助于医生根据风险等级分配医疗资源,提高医疗服务效率,减少医疗资源浪费。同时,通过早期干预,疾病风险评估有助于改善患者的生活质量,减轻疾病带来的痛苦。此外,深度学习算法如循环神经网络(RNN)已被用于预测疾病的发展风险,为临床预警和个性化治疗提供了新的可能性。其能够从复杂且高维的数据中自动学习规律,挖掘潜在的非线性关系,且已成功应用于多种疾病的预后预测。然而,在AVM研究领域,特别是针对术后再通这一特定结局,高质量、大样本的临床数据常常难以获取,数据稀缺已然成为模型构建的主要瓶颈。借助模拟数据(Synthetic Data)生成技术,依据现有研究证据构建具有统计合理性的数据集,为在数据受限的情况下开展方法学探索与模型初步构建提供了可行路径。

因此,本研究旨在基于现有文献证据,生成涵盖人口学、病灶形态学、血管构筑及治疗信息的综合性模拟数据集,系统构建并比较多种机器学习模型,以期建立一种高精度的AVM介入栓塞术后再通风险预测模型。研究结果不仅可为临床医生提供数据驱动的决策支持,助力优化治疗策略,还能为在有限真实数据条件下开展预测模型研究提供方法学参考。

## 2. 材料与方法

### 2.1. 数据来源与模拟数据生成

本研究的数据基础源于对近十年内发表的24篇关于AVM治疗及预后因素的权威文献的系统回顾及相关研究[2][6]-[12]。构建AVM再通风险预测模型时,我们借鉴了类似缺血性脑卒中后认知障碍风险预测模型的构建方法,收集并分析大量数据,提取关键变量及其统计学描述(如均值、标准差、分布比例、优势比等),确定与AVM再通风险可能相关的核心特征集。最终,模拟数据集共包含17个特征,归为以下五类:

- 1、患者基本信息:年龄、性别。
- 2、病灶核心特征:病灶体积、体积分类(如大、中、小)、病灶位置(如幕上/幕下、具体脑叶)、是否位于功能区、SM分级、畸形团结构(如是否合并动脉瘤、瘘口)。
- 3、静脉引流特征:引流静脉数量、是否存在深静脉引流、引流静脉分类。
- 4、供血动脉特征:供血动脉数量、是否存在深部供血动脉、供血动脉分类。
- 5、临床与治疗特征:出血次数、出血史分类(有/无)、手术方式(如单纯栓塞、栓塞联合其他治疗)。

模拟数据的生成采用参数化方法。在医学研究中,对于连续变量(如年龄、病灶体积),可以参考文献报道的分布特征(如正态分布、对数正态分布)来生成随机值,以模拟AVM病灶大小等变量的分布。对于分类变量(如性别、SM分级、是否功能区),依据文献中的发生率进行随机抽样生成。各特征之间存在潜在关联,比如较大的病灶更有可能处于非功能区,深部引流与高SM分级存在相关性,这些关联通过预设的条件概率矩阵加以约束,从而确保生成的数据具备临床合理性与内在一致性[9]。数据集总共生成了5000条样本,其中再通阳性样本占比约14.98%,以此模拟临床中再通事件相对较少的情形。

### 2.2. 特征预处理与数据集构建

针对生成的脑血管畸形(AVM)再通风险预测模拟数据集,开展系统性的特征预处理工作。预处理流程涵盖数据清洗、特征编码、标准化处理以及数据集划分四个主要环节,以此确保数据质量,并为后续的机器学习建模奠定基础。

#### 2.2.1. 数据清洗与缺失值处理

首先,对数据集进行完整性检查,识别并处理缺失值。数值型特征(如年龄、病灶体积、引流静脉数量等)采用中位数插补法,以此保持数据分布的稳定性;分类变量(如病灶位置、手术方式等)采用众数插

补法。针对数据集中可能存在的异常值,运用 Z-score 检测法,加以识别,同时采用 Winsorization 方法(截尾处理)进行修正,把极端值控制在合理范围内。

### 2.2.2. 特征编码策略

依据特征类型与临床意义,采用差异化的编码策略:

1) 连续变量编码:年龄、病灶体积等连续变量保留其数值特性,为提升模型对非线性的捕捉能力,同时构建对数变换( $\log(1+x)$ )和平方根变换特征。

2) 二分类变量编码:性别(男/女)、深静脉引流(是/否)等变量运用二进制编码(0/1)。

3) 有序分类变量编码:出血史分类(无出血/单次出血/多次出血)、体积分类(小/中/大)等具备明确等级顺序的变量,采用整数序列编码(0, 1, 2...),保留其顺序信息。

4) 名义分类变量编码:病灶位置(基底节区、额叶、颞叶等)采用基于临床风险评估的加权编码方式,把临床专业知识融入特征表示。例如,基底节区位置因再通风险较高编码为 4,而枕叶位置风险较低编码为 1 [13] [14]。

5) 手术方式编码:根据文献报道的再通率,将不同治疗方式转换为风险评分,例如单纯栓塞编码为 4 (高再通风险),而手术切除编码为 1 (低再通风险),这一评估方法在急性脑动脉闭塞治疗中得到了应用比较,其中机械取栓相较于动脉溶栓具有更高的再通率和更低的出血率[15]。

### 2.2.3. 特征工程与交互特征创建

为了进一步深入挖掘特征之间的复杂关系,我们创建了多种衍生特征:

1) 非线性变换特征:针对病灶体积等呈现右偏分布的变量,我们采用对数变换的方法,以改善其分布形态。

2) 交互特征:基于临床病理生理机制,我们创建了年龄与病灶体积的乘积交互项,以及引流静脉数量与供血动脉数量的比值特征等,以此捕捉变量间的协同效应。

3) 复合风险评分:我们结合多个临床风险因素(如 SM 分级、深静脉引流、病灶位置等),创建了加权复合风险评分,从而为模型提供综合风险评估特征。

### 2.2.4. 数据标准化处理

为了消除特征量纲差异对模型训练的影响,我们采用 Z-score 标准化方法对所有连续变量进行处理。在标准化过程中,我们严格遵循数据隔离原则:仅使用训练集来计算均值和标准差,然后将相同的变换应用于测试集,以此避免数据泄露问题。对于可能包含异常值的特征,我们同时备选了稳健标准化方法(RobustScaler),该方法基于中位数和四分位距,对异常值具有更强的鲁棒性。

### 2.2.5. 数据集划分与验证策略

为了确保模型评估的可靠性和泛化能力,我们采用分层抽样方法,将数据集按照 8:2 的比例划分为训练集和测试集。通过分层抽样,我们确保了两个数据子集中的正负样本比例与原始数据集一致,从而避免了随机划分可能引发的类别分布偏差。具体而言,训练集包含 4000 个样本(80%),用于模型训练和超参数调优;测试集包含 1000 个样本(20%),作为最终评估模型泛化性能的独立数据集。

在整个模型开发过程中,我们严格执行数据隔离原则,即测试集仅用于最终性能评估,不参与特征工程、参数调优或模型选择等任何环节。此外,我们在训练集内部采用 5 折分层交叉验证方法进行模型选择和超参数优化,以进一步提升模型评估的稳定性和可靠性。

## 2.3. 机器学习模型构建与调优

本研究构建了五种机器学习模型,以期找到预测 AVM 栓塞术后再通风险的最优方法。模型的选择

兼顾了经典算法与现代高性能算法，并特别考虑了医疗数据中常见的不平衡问题。

### 2.3.1. 逻辑回归

逻辑回归模型结构简单，结果易于解释，常作为医疗预测模型的基线。为防止模型因过于复杂而“死记硬背”训练数据(即过拟合)，我们为其引入了 L2 正则化惩罚。同时，我们设定模型在训练时自动调整类别权重，以应对“再通”患者数量显著少于“未再通”患者的数据不平衡问题。

### 2.3.2. 随机森林

随机森林通过构建大量决策树，并综合这些决策树的结果进行预测，能够有效捕捉复杂的特征关系。为避免单棵决策树过于复杂，我们将其最大深度限制为 10 层。为增强模型的泛化能力，每棵决策树仅随机选取部分特征进行学习。该模型同样内置了处理类别不平衡的机制。

### 2.3.3. XGBoost

XGBoost 是一种高效的梯度提升算法，它通过顺序构建决策树来不断修正前序树的错误，并通过提前停止等技术防止过拟合，从而保持强劲预测性能。我们对其进行了严格设置以防止过拟合：限制树深度为 6 层，在每次构建新树时只随机使用 80% 的数据和 80% 的特征，并增加了额外的正则化项。针对正负样本不平衡的情况，我们按照样本数量的比例为其设置了专门的权重参数。

### 2.3.4. LightGBM

LightGBM 是另一种高效的梯度提升框架，训练速度较快。我们为其设置了与 XGBoost 相似的正则化约束，涵盖树深度、L1/L2 正则化等方面，并启用了其处理不平衡数据的功能。

### 2.3.5. 平衡随机森林

该模型是专门为不平衡数据设计的随机森林变体。其核心思想是在构建每棵决策树时，对原始数据进行有方向的抽样，并有意使抽样后的数据中正负样本数量达到平衡，从而让每棵树都能更好地学习到少数类(即“再通”类)的特征。

所有模型的训练数据均经过标准化处理，具体而言，是先减去均值，再除以标准差，以此确保所有特征处于同一量纲，进而保障模型训练的稳定性。

## 2.4. 模型评估指标

为全面、客观地衡量模型性能，本研究采用了一套综合评估体系，所有评估均在模型从未见过的独立测试集上进行。主要指标包括：

**AUC 值：**作为本研究的核心评价指标，它能够综合衡量模型在不同判定阈值下，对“再通”与“未再通”患者进行区分的能力。AUC 值越接近 1，表明模型的判别能力越强；而 0.5 则相当于随机猜测。

**准确率：**指模型做出正确预测的病例数占总病例数的比例。然而，在样本极不平衡的情况下，该指标可能会出现失真现象(例如，若全部预测为“未再通”，准确率也可能很高)。

**精确率：**在所有被模型预测为“再通”的病例中，真正发生了再通的比例，即精确率，是衡量模型预测准确性的重要指标。高精确率表明模型在预测“再通”病例时的误报较少，这对于医疗诊断等应用尤为重要。

**召回率(灵敏度)：**指在所有实际发生再通的病例中，被模型成功预测出来的比例。高召回率意味着模型的“漏报”情况较少。

**F1 分数：**是精确率和召回率的调和平均数。当精确率和召回率出现矛盾，即一个高一个低时，F1 分数能够更好地反映模型的综合性能，尤其适用于不平衡数据的评估。

平衡准确率：即分别计算正例和负例的准确率，随后取其平均值。该指标可有效规避样本数量不均对总体准确率评估产生的偏差。

在临床实践中，我们既希望模型能尽可能找出所有高风险患者(高召回率)，又希望避免对低风险患者进行不必要的干预(高精确率)。因此，AUC 值和 F1 分数是本研究中权衡模型性能的关键依据。

## 2.5. 可解释性分析

一个“黑箱”模型即便预测精准，也难以获得临床医生的接纳与信任。因此，我们通过以下方法对模型的预测逻辑进行解读，使其决策过程透明化。

### 2.5.1. 分析特征重要性

利用随机森林、XGBoost 等模型自身提供的功能，计算每个特征(如“年龄”、“病灶体积”)对最终预测结果的贡献度，并排序。通过图 1 展示最重要的前 20 个特征(包括原始数据以及编码后的变量)，可以直观地回答“模型主要依据哪些信息做出判断”这个问题，帮助医生聚焦关键临床指标。

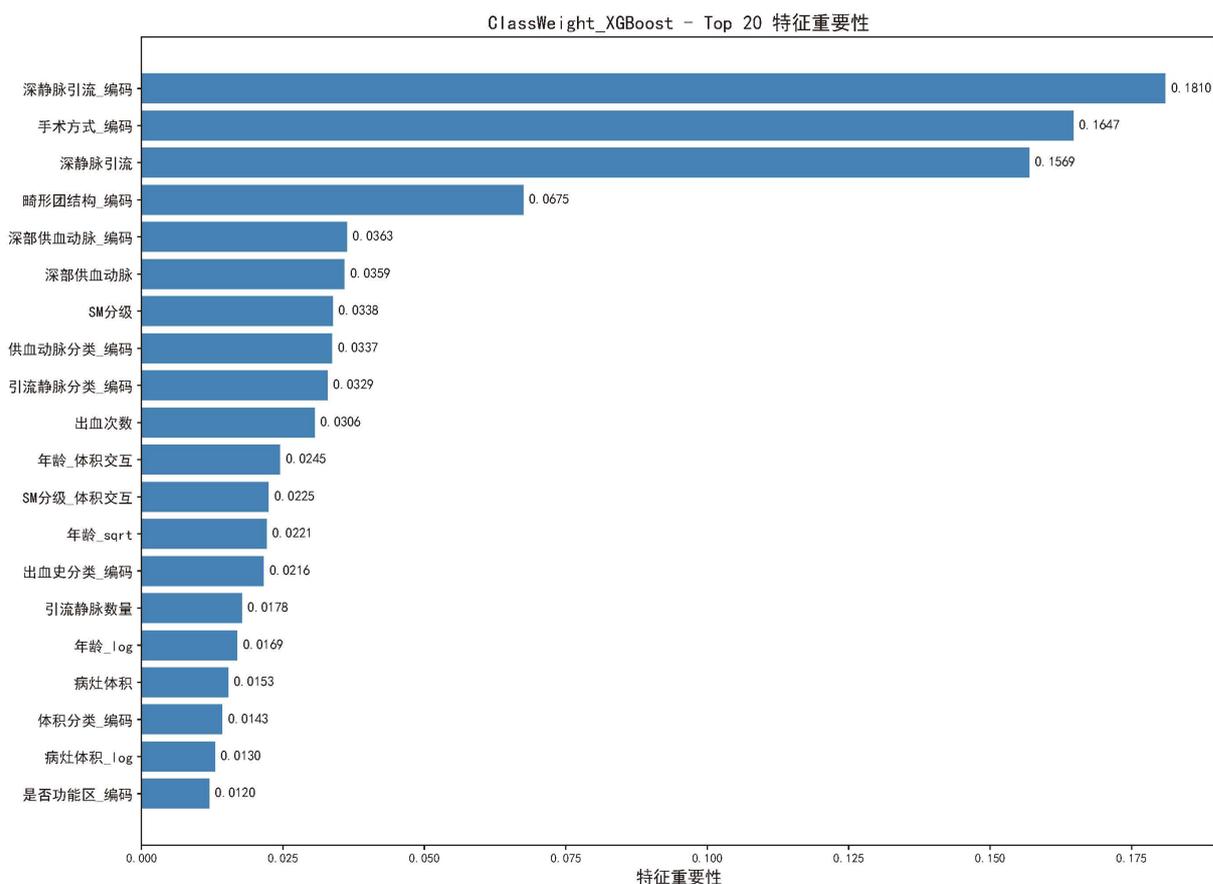


Figure 1. Top 20 feature importance of the XGBoost model

图 1. XGBoost 模型 top20 特征重要性

### 2.5.2. 可视化特征如何影响预测(SHAP 分析)

我们运用 SHAP 框架展开深入剖析。该框架能够针对每个患者的每个预测结果，计算出每个特征的具体贡献值。汇总所有患者的分析结果后，我们获得了两个关键洞察：

全局解释：通过计算所有样本中各特征的 SHAP 值平均绝对值，我们可得出特征的全局重要性排序，

进而识别出影响预测的普遍性关键因素。例如，我们可以确定在整个数据集中，“手术方式”、“畸形团结构”和“深静脉引流”是预测再通风险的最重要特征。

局部解释：针对单个患者，SHAP 值能够展示各特征如何协同作用，从而得出最终的预测结果。例如，对于一个具体患者，我们可以用瀑布图展示：特征“病灶体积大”将其预测风险得分推高了 0.3 分，而特征“无深静脉引流”将风险得分降低了 0.2 分，最终模型给出的预测风险得分为 0.65 (即 65% 的再通风险)。这种解释有助于临床医生理解模型对特定患者的决策依据，进而增强信任。

### 2.5.3. 绘制特征间相关性热图

我们计算了所有特征间的相关系数，并通过热图进行可视化展示。这有助于发现高度相关的特征组，理解特征间的关系可辅助评估模型的稳定性，并在必要时进行特征筛选。

### 2.5.4. 可视化模型性能比较

使用多种图表对所有训练好的模型进行横向比较。此举不仅旨在遴选最优模型，更在于揭示不同算法与数据处理策略的优劣，为方法学选择提供有力支撑。基于上述分析，我们致力于构建一个既精准又可靠、实用的预测工具，确保其预测结果能为临床医生所理解与验证，进而切实辅助临床决策。

## 3. 结果

### 3.1. 模拟数据集基本特征描述

本研究共纳入模拟的 AVM 患者数据，通过系统参数设置生成 5000 例患者信息。如表 1 所示，数据集包含年龄、性别、病灶体积、引流静脉数量、SM 分级等关键临床特征。目标变量“再通”的分布显示，再通患者占 14.98%，未再通患者占 85.02%，存在一定的不平衡性。

Table 1. Baseline characteristics of the simulated AVM patient dataset (n = 5000)

表 1. 模拟 AVM 患者数据集基本特征(n = 5000)

总数	5000 例	
年龄	45.3 ± 14.3 岁	
性别	男	2820 例(56.4%)
	女	2180 例(43.6%)
病灶体积	平均体积 10.8 ± 14.3 岁	
	小(≤10 cm <sup>3</sup> )	2839 例(56.8%)
	中(10~30 cm <sup>3</sup> )	2057 例(41.1%)
	大(30~50 cm <sup>3</sup> )	93 例(1.9%)
	巨大(>50 cm <sup>3</sup> )	11 例(0.2%)
SM 分级分布	1 级	1312 例(26.2%)
	2 级	2253 例(45.1%)
	3 级	1212 例(24.2%)
	4 级	207 例(4.1%)
	5 级	16 例(0.3%)

续表

手术方式分布	保守治疗	1646 例(32.9%)
	单纯栓塞	1563 例(31.3%)
	手术切除	775 例(15.5%)
	栓塞 + 放疗	585 例(11.7%)
	栓塞 + 手术	431 例(8.6%)
引流静脉类型分布	多支浅静脉	2159 例(43.2%)
	深 + 浅静脉	530 例(10.6%)
	单一浅静脉	1882 例(37.6%)
	深静脉引流	429 例(8.6%)

### 3.2. 模型性能比较

在测试集上, 五种机器学习模型的性能表现如表 2 所示。其中, XGBoost 模型表现最为出色, 其 AUC 值达到 0.98, F1 分数为 0.81, 平衡准确率也达到 0.81, 这些指标表明 XGBoost 在分类任务中具有较高的准确性和良好的泛化能力。

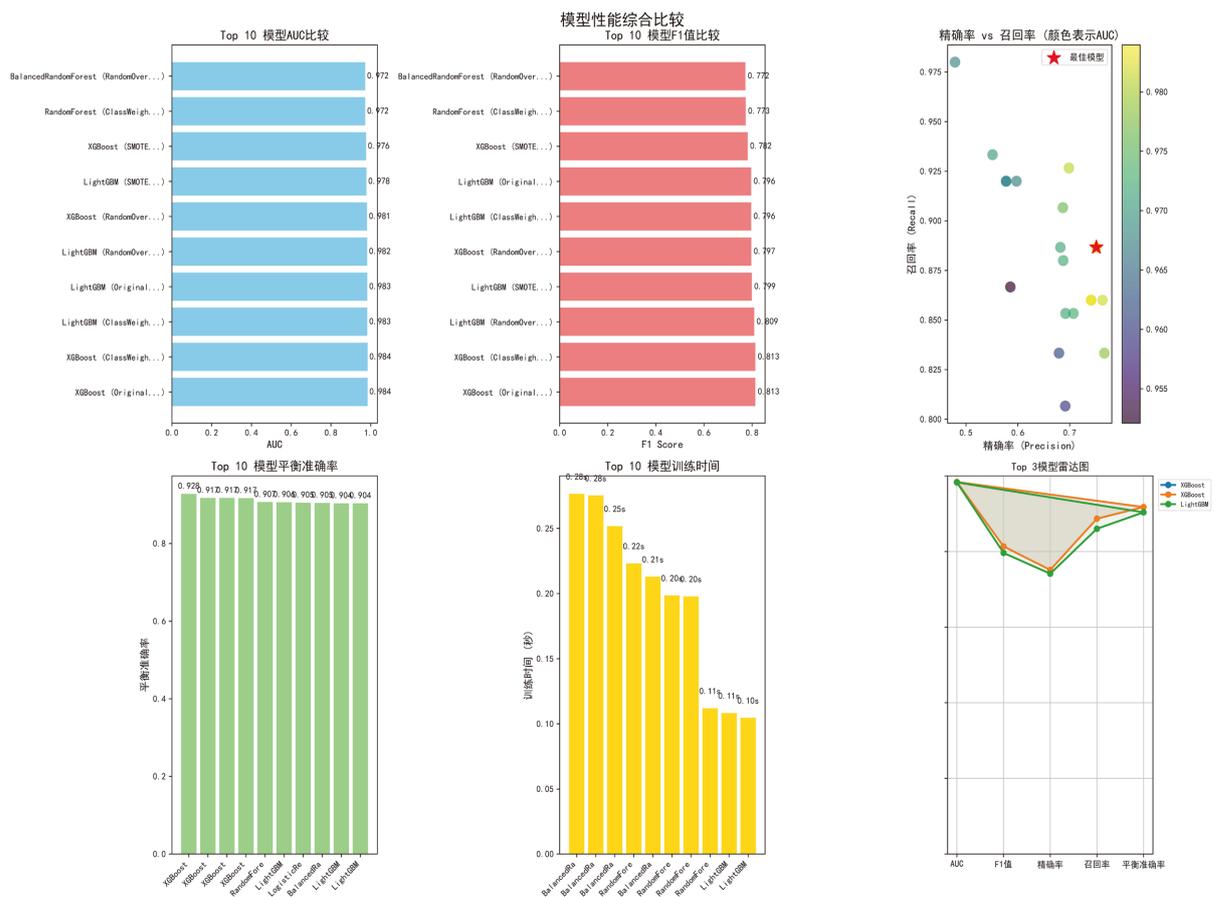


Figure 2. Bar chart of model performance comparison  
图 2. 模型性能对比条形图

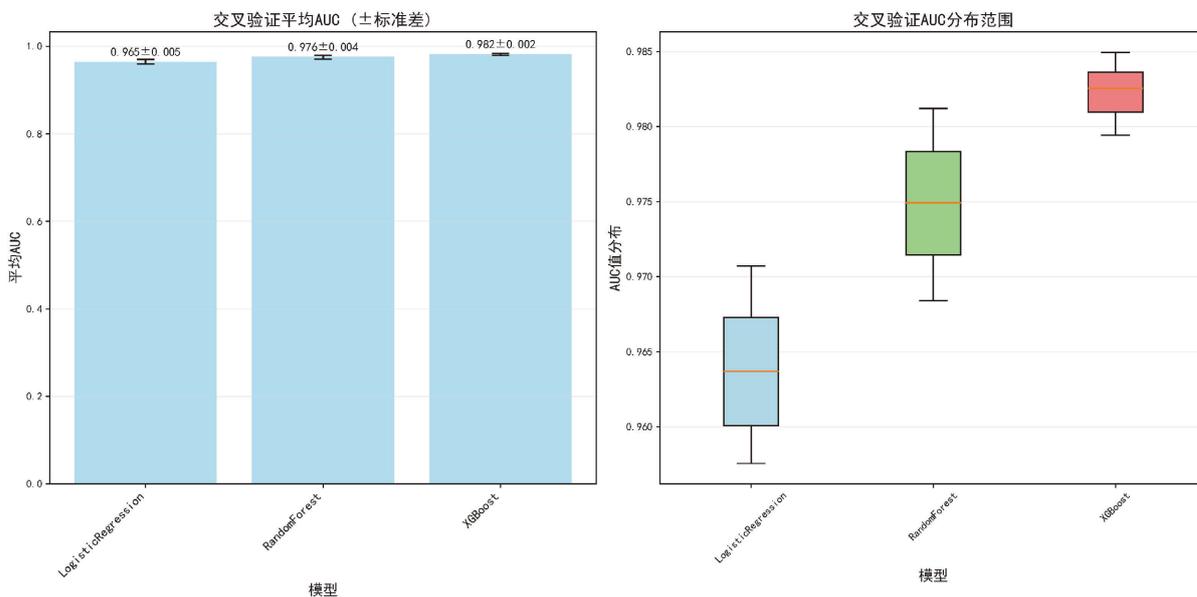


Figure 3. Comparison chart of cross-validation AUC results  
图 3. 交叉验证 AUC 结果对比图

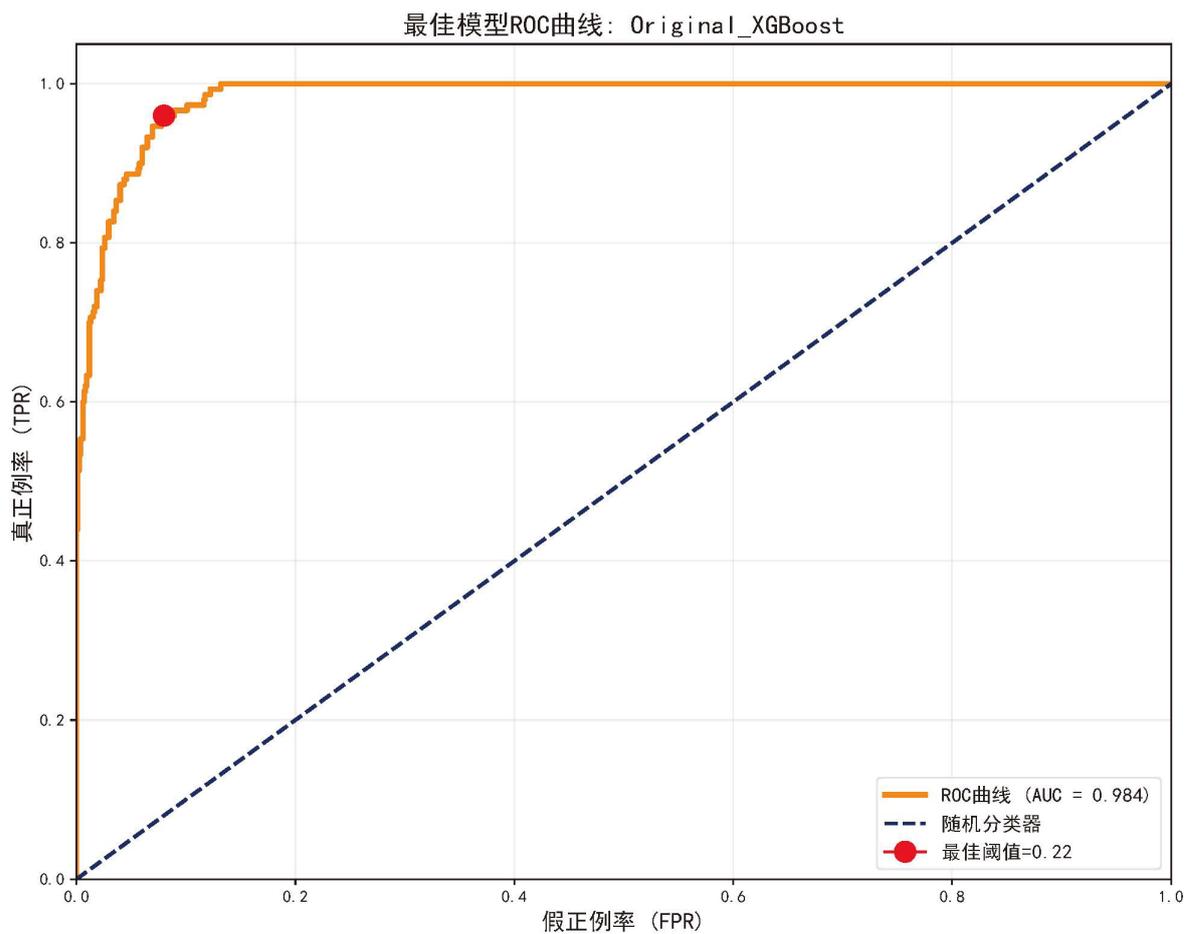


Figure 4. Comparison chart of ROC curves  
图 4. ROC 曲线对比图

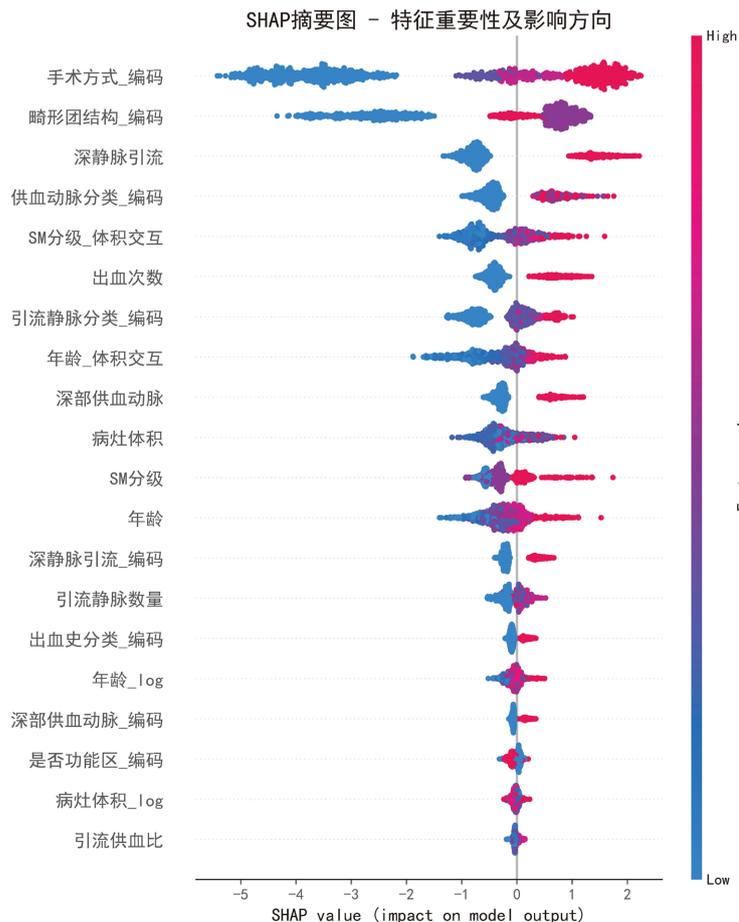
**Table 2.** Detailed performance comparison of different machine learning models on the test set  
**表 2.** 不同机器学习模型在测试集上的性能对比详情

模型	最佳采样方法	AUC	准确率	精确率	召回率	F1 值	平衡准确率	AP
XGBoost	Original	0.98	0.93	0.75	0.88	0.81	0.91	0.92
LightGBM	ClassWeight	0.98	0.93	0.74	0.86	0.79	0.90	0.91
RandomForest	ClassWeight	0.97	0.92	0.70	0.8	0.77	0.89	0.86
BalancedRandomForest	RandomOverSample	0.97	0.92	0.68	0.88	0.77	0.90	0.86
LogisticRegression	RandomOverSample	0.96	0.89	0.59	0.92	0.72	0.90	0.83

图 2 直观展示了各模型的 AUC 和 F1 值对比，XGBoost 模型在两个指标上均表现最优。通过 5 折交叉验证(图 3)验证，XGBoost 模型表现出良好的稳定性，平均 AUC 为  $0.982 \pm 0.002$ 。图 4 展示了最佳模型(XGBoost)的 ROC 曲线，其 AUC 值为 0.98，明显高于随机分类器的 AUC 值。

### 3.3. 特征重要性分析

借助 SHAP 框架对 XGBoost 模型的决策依据展开分析，我们发现，影响 AVM 再通风险预测的关键特征涵盖手术方式、病灶体积、引流静脉数量、SM 分级以及是否存在深静脉引流等方面。在图 5 的 SHAP 摘要图中，清晰展示了各特征对预测的贡献方向及幅度，其中，红色点代表高特征值，蓝色点代表低特征值。



**Figure 5.** Swarm plot of SHAP value distributions  
**图 5.** SHAP 值分布蜂群图

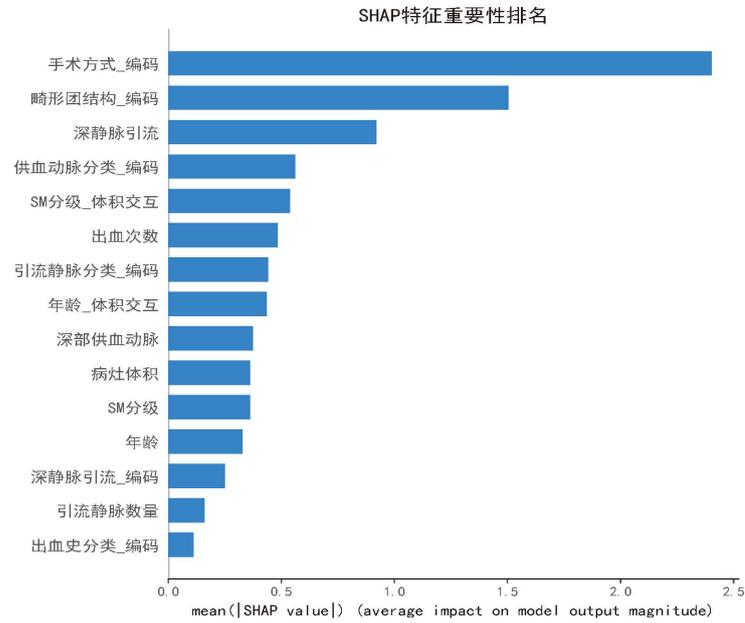


Figure 6. Bar chart of SHAP feature importance  
图 6. SHAP 特征重要性条形图

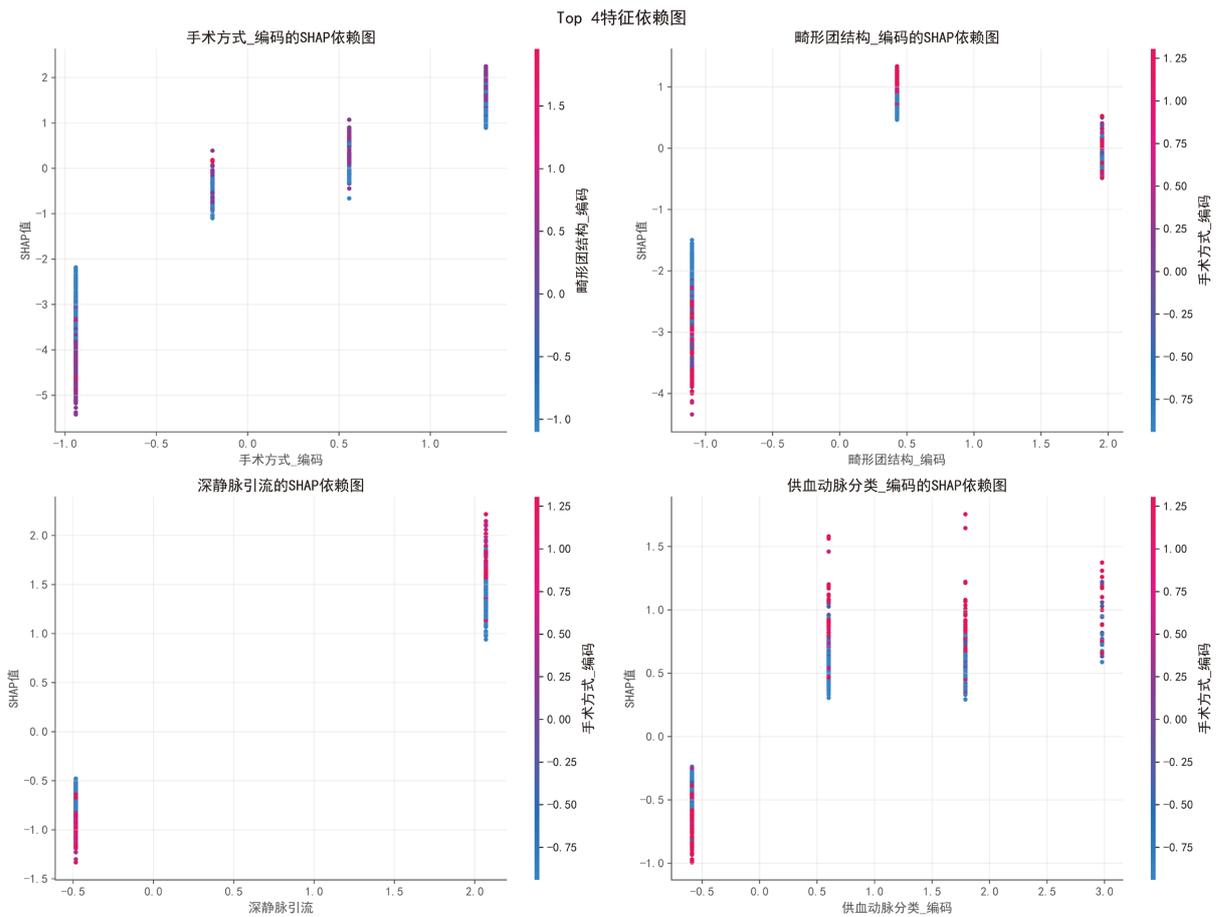


Figure 7. SHAP dependence plot  
图 7. SHAP 依赖图

图 6 的特征重要性条形图量化了各特征的贡献度，其中手术方式对预测的影响最大(平均|SHAP 值| = 2.5)，这一点与 AVM 病灶大小、血流动力学、位置、与重要脑组织的关系、手术医生的经验和技能、患者的年龄和整体健康状况、术前医疗条件、术后护理等因素密切相关，这些因素共同决定了手术风险的高低。其次为病灶体积(1.5)，病灶体积越大，手术风险越大，手术治疗后遗症的发生率和严重程度也越高。引流静脉数量(1.0)也是影响手术操作复杂度和出血风险的重要因素。图 7 的依赖图进一步揭示了关键特征与预测风险之间的非线性关系，例如病灶体积越大，SHAP 值越高，表明再通风险随病灶体积增大而增加。这一点在脑动静脉畸形(AVM)的研究中得到了证实，其中 AVM 的大小被发现与出血危险性以及手术风险密切相关。

### 3.4. 特征相关性分析

图 8 呈现了主要临床特征之间的 Spearman 相关性情况。结果显示，病灶体积与 SM 分级呈中等正相关( $r = 0.56, p < 0.01$ )；与此同时，SM 分级与引流静脉数量也呈正相关( $r = 0.41, p < 0.05$ )。而年龄与其他特征的相关性较弱( $|r| < 0.3$ )，这表明其独立影响再通风险。

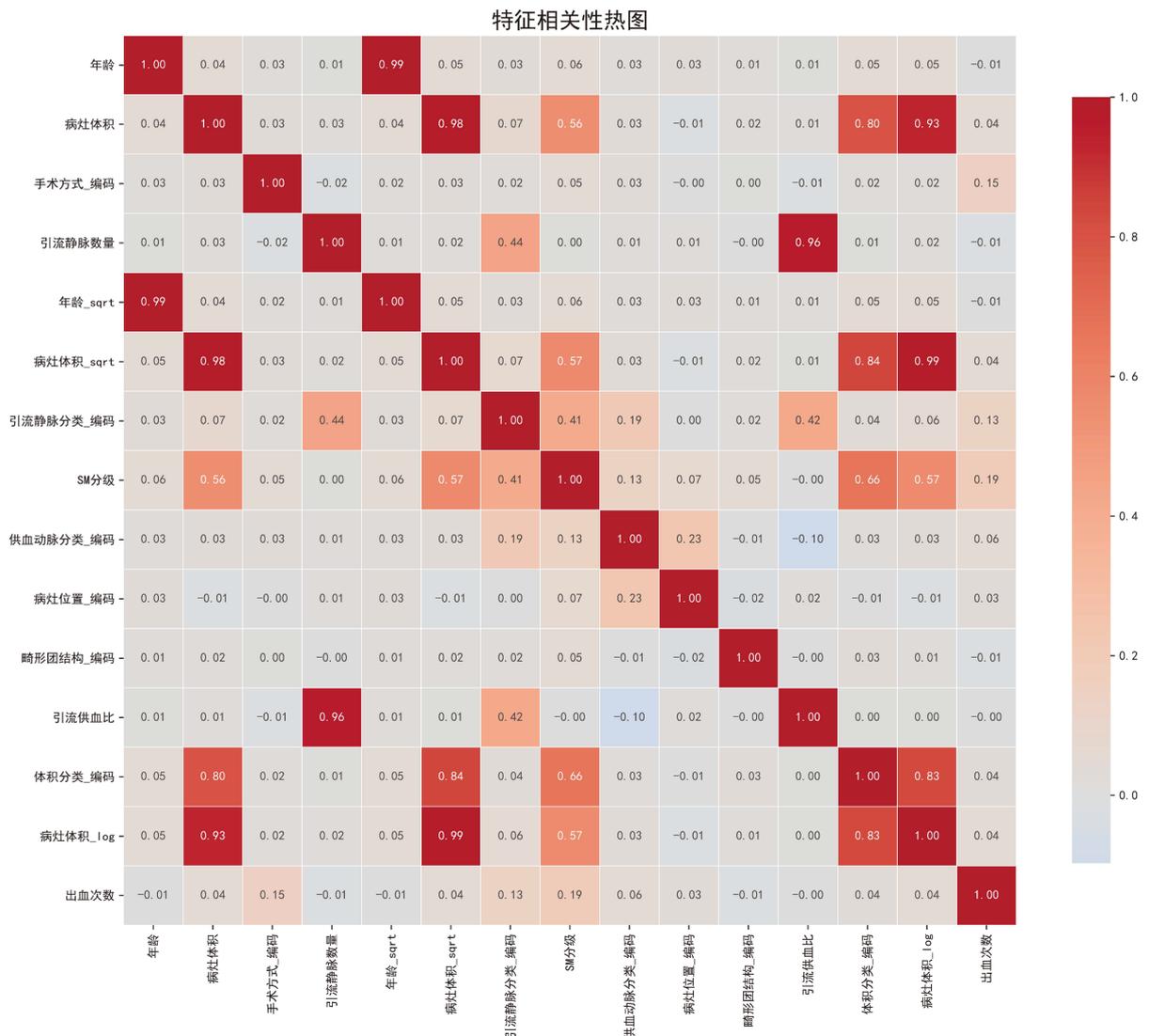


Figure 8. Heatmap of feature correlations  
图 8. 特征相关性 heatmap

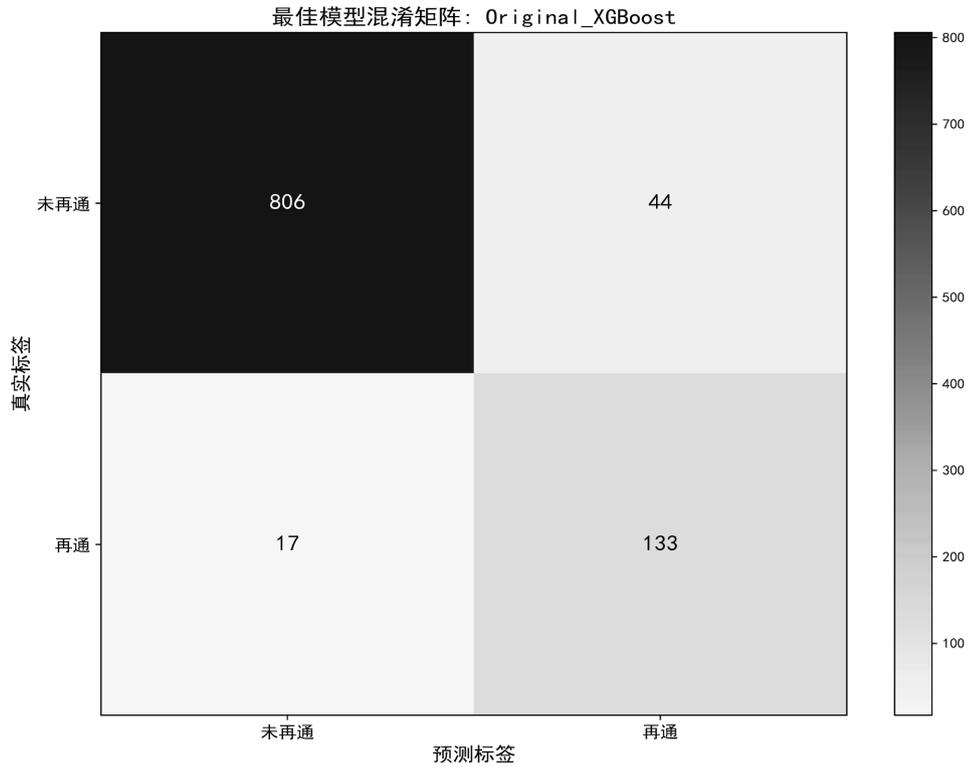


Figure 9. Confusion matrix  
图 9. 混淆矩阵

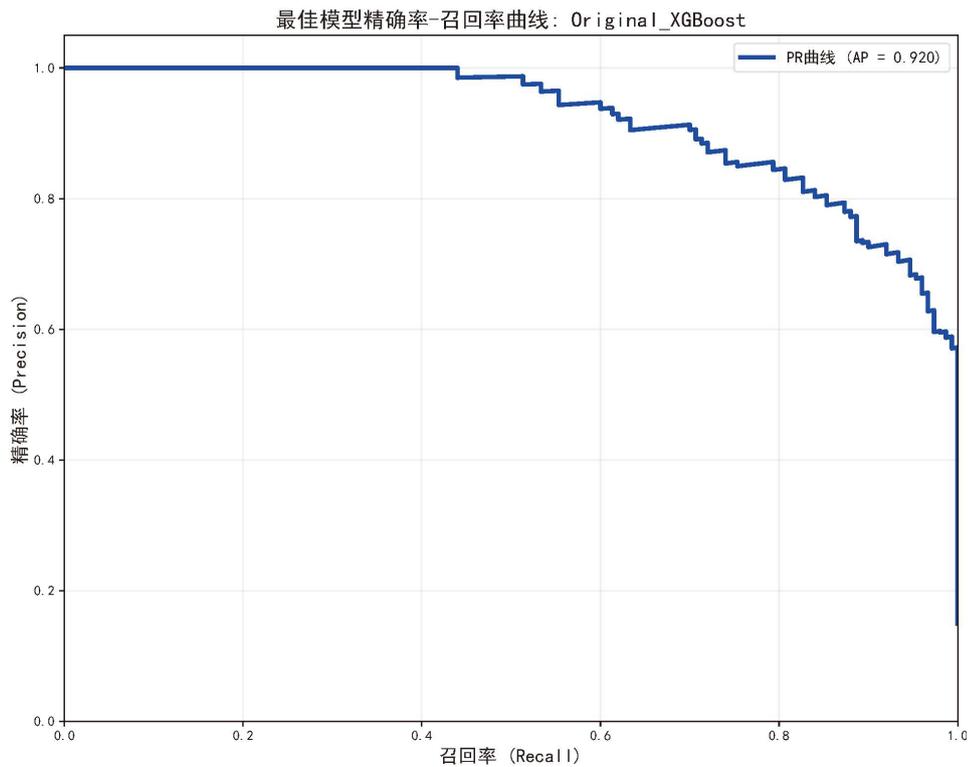


Figure 10. Precision-recall (PR) curve  
图 10. PR 曲线

### 3.5. 最优模型的详细预测表现

XGBoost 模型在测试集上的详细预测表现如图 9 所示。模型正确识别了 850 例未再通中的 806 例(特异度 94.8%), 以及 150 例再通中的 133 例(灵敏度 88.7%)。总体准确率达到 93.9%。图 10 的精确率 - 召回率曲线显示, 在召回率为 60%时, 模型的精确率可达 92%, 表明模型在识别高风险患者方面具有较好的平衡性。

## 4. 讨论

本研究依托系统生成的模拟数据, 成功构建并对比了五种机器学习模型, 以预测 AVM 介入栓塞术后的再通风险。结果表明, 集成学习算法, 尤其是 XGBoost, 在预测任务中表现出色, 其 AUC 值达到 0.98, F1 分数超过 0.81, 在多项性能指标上均优于传统的逻辑回归模型(表 2, 图 1~3)。这证实了机器学习方法, 特别是能够处理复杂特征交互的非线性模型, 在整合多维度临床信息进行精准风险预测方面的巨大优势。

通过 SHAP 分析, 我们识别出了一系列对预测再通风险至关重要的特征。手术方式、病灶体积、引流静脉数量、SM 分级以及是否存在深静脉引流等特征, 在模型中的贡献度最高, 这与既往临床研究对再通高危因素的认知基本相符。例如, 体积较大且引流静脉复杂的畸形团, 其血管构筑通常更为复杂, 栓塞的完全性难以保证, 理论上再通风险也更高。模型不仅对这些已知因素的影响权重进行了量化, 还可能揭示了某些特征间交互作用的潜在规律, 为深入理解再通机制提供了数据支撑。

本研究以模拟数据作为模型构建的起点, 这是针对高质量临床队列数据获取周期长、难度大这一现实挑战所做出的积极尝试。在可控条件下进行系统评估, 我们能够深入理解不同预测变量和算法的效能。例如, XGBoost 算法已被成功应用于重症急性胰腺炎的早期预测, 其预测准确度达到 0.830, ROC 曲线下面积(AUC)为 0.927, 显著高于传统评分系统。类似地, XGBoost 在复杂性阑尾炎风险预测和糖尿病风险预测中也展现了卓越的预测效能, 这为后续基于真实世界数据的研究提供了坚实的方法学基础, 并帮助我们优先筛选出最具潜力的预测模型。与单纯使用不平衡的真实小样本数据相比, 基于文献证据的模拟数据可以在一定程度上保证数据的“临床合理性”和统计效力, 降低了初始探索阶段因数据问题导致的偏误。

从临床转化角度看, 本研究所构建的 XGBoost 模型具有成为辅助决策工具的潜力。在术前评估阶段, 医生可输入患者的各项特征信息, 获得一个量化的再通风险概率, 从而与患者及家属进行更为充分的沟通, 并辅助制定更具个体化的治疗方案(例如是否需采取更积极的完全栓塞措施、是否需联合手术或放疗)。在术后随访过程中, 高风险患者可被标识为重点监测对象, 以此优化影像随访策略。

当然, 本研究存在若干局限性。首先, 最为关键的是模拟数据与真实患者群体之间必然存在的差异。尽管我们竭力依据文献设定参数, 但生成的数据终究无法完全替代真实临床数据的复杂性与异质性。其次, 尽管研究纳入的特征已涵盖主要临床维度, 但仍有一些更精细的影像组学特征、血流动力学参数或遗传标记未被纳入, 这或许会限制模型的预测上限。最后, 目前模型缺乏外部验证, 其泛化能力尚需在未来自助多中心真实世界数据进行严格检验。

展望未来, 下一步研究的核心在于将该模型在真实临床队列中开展前瞻性验证与调优。同时, 探索纳入更多维度的数据, 如高分辨率血管影像的纹理特征、血流动力学模拟参数等, 以此进一步提升模型的预测精度和生物学解释能力。最终目标是开发一个用户友好型、可集成至医疗信息系统的临床决策支持模块, 实现从研究到临床实践的有效转化。

## 5. 结论

本研究基于现有文献生成了模拟数据, 系统构建并比较了五种机器学习模型, 用于预测脑动静脉畸形介入栓塞术后的再通风险。结果显示, XGBoost 模型在该任务中表现最佳, 其 AUC 值为 0.98, F1 分

数为 0.81, 能够有效整合多维度临床特征, 实现个体化风险预测。关键影响因素与临床认知相符, 增强了模型的可解释性。尽管模拟数据存在固有局限性, 但本研究为后续在真实临床数据中构建与验证预测模型提供了方法学参考。

## 参考文献

- [1] Pinheiro, L.C.P., Wolak Junior, M., Ferreira, M.Y., Magalhaes, R.B., Fernandes, A.Y., Paiva, W.S., *et al.* (2024) Unruptured Brain Arteriovenous Malformations: A Systematic Review and Meta-Analysis of Mortality and Morbidity in Aruba-Eligible Studies. *World Neurosurgery*, **185**, 381-392.e1. <https://doi.org/10.1016/j.wneu.2024.02.109>
- [2] Bauer, A.M., Bain, M.D. and Rasmussen, P.A. (2015) Onyx Resorbition with AVM Recanalization after Complete AVM Obliteration. *Interventional Neuroradiology*, **21**, 351-356. <https://doi.org/10.1177/1591019915581985>
- [3] Guo, Y., Zhang, S., Sun, S., Zhang, Q., Zhai, Y., Wang, X., *et al.* (2024) Analysis of the Role of Perfusion Parameters in Predicting Rupture of Brain Arteriovenous Malformations: A Multi-Center Study. *Scientific Reports*, **14**, Article No. 25566. <https://doi.org/10.1038/s41598-024-77316-7>
- [4] Rodríguez Mallma, M.J., Vilca-Aguilar, M., Zuloaga-Rotta, L., Borja-Rosales, R., Salas-Ojeda, M. and Mauricio, D. (2023) Machine Learning Approach for Analyzing 3-Year Outcomes of Patients with Brain Arteriovenous Malformation (AVM) after Stereotactic Radiosurgery (SRS). *Diagnostics*, **14**, Article No. 22. <https://doi.org/10.3390/diagnostics14010022>
- [5] Wang, L., Wei, C., Chen, Y., Lu, D., Zhang, H. and Yang, M. (2026) Machine Learning-Based Prediction Model for 30-Day Readmission Risk in Elderly Patients with Type 2 Diabetes Mellitus and Heart Failure: A Retrospective Cohort Study with SHAP Interpretability Analysis. *Frontiers in Cardiovascular Medicine*, **12**, Article ID: 1673159. <https://doi.org/10.3389/fcvm.2025.1673159>
- [6] Hao, Q., Zhang, H., Han, H., *et al.* (2023) Recurrence of Arteriovenous Malformations of the Brain after Complete Surgical Resection. Kuopio University Hospital Experience and Systematic Review of the Literature. *Neurosurgical Review*, **46**, No. 1.
- [7] Jin, H., Huo, X., Jiang, Y., Li, X. and Li, Y. (2017) Safety and Efficacy of Endovascular Therapy and Gamma Knife Surgery for Brain Arteriovenous Malformations in China: Study Protocol for an Observational Clinical Trial. *Contemporary Clinical Trials Communications*, **7**, 103-108. <https://doi.org/10.1016/j.conctc.2017.06.006>
- [8] Khan, H., Sangah, A.B., Nasir, R., Khan, S.A., Shaikh, S.S., Ahmed, I., *et al.* (2024) Efficacy of Radiosurgery with and without Angioembolization: A Subgroup Analysis of Effectiveness in Ruptured versus Unruptured Arteriovenous Malformations—An Updated Systematic Review and Meta-Analysis. *Surgical Neurology International*, **15**, Article No. 467. <https://doi.org/10.25259/sni.737.2024>
- [9] Morgan, M.K., Alsahli, K., Wiedmann, M., Assaad, N.N. and Heller, G.Z. (2016) Factors Associated with Proximal Intracranial Aneurysms to Brain Arteriovenous Malformations. *Neurosurgery*, **78**, 787-792. <https://doi.org/10.1227/neu.0000000000001114>
- [10] Nerva, J.D., Barber, J., Levitt, M.R., Rockhill, J.K., Hallam, D.K., Ghodke, B.V., *et al.* (2017) Onyx Embolization Prior to Stereotactic Radiosurgery for Brain Arteriovenous Malformations: A Single-Center Treatment Algorithm. *Journal of NeuroInterventional Surgery*, **10**, 258-267. <https://doi.org/10.1136/neurintsurg-2017-013084>
- [11] Pabaney, A.H., Reinard, K.A., Kole, M.K., Seyfried, D.M. and Malik, G.M. (2016) Management of Arteriovenous Malformations in the Elderly: A Single-Center Case Series and Analysis of Outcomes. *Journal of Neurosurgery*, **125**, 145-151. <https://doi.org/10.3171/2015.6.jns15293>
- [12] Shotar, E., Debarre, M., Sourour, N., Di Maria, F., Gabrieli, J., Nouet, A., *et al.* (2018) Retrospective Study of Long-Term Outcome after Brain Arteriovenous Malformation Rupture: The RAP Score. *Journal of Neurosurgery*, **128**, 78-85. <https://doi.org/10.3171/2016.9.jns161431>
- [13] Hak, J., Boulouis, G., Kerleroux, B., Benichi, S., Stricker, S., Gariel, F., *et al.* (2021) Pediatric Brain Arteriovenous Malformation Recurrence: A Cohort Study, Systematic Review and Meta-Analysis. *Journal of NeuroInterventional Surgery*, **14**, 611-617. <https://doi.org/10.1136/neurintsurg-2021-017777>
- [14] Hao, Q., Zhang, H., Han, H., Jin, H., Ma, L., Li, R., *et al.* (2023) Recurrence of Cerebral Arteriovenous Malformation Following Complete Obliteration through Endovascular Embolization. *Translational Stroke Research*, **16**, 339-349. <https://doi.org/10.1007/s12975-023-01215-8>
- [15] Sattari, S.A., Shahbandi, A., Yang, W., Feghali, J., Xu, R. and Huang, J. (2022) Microsurgery versus Microsurgery with Preoperative Embolization for Brain Arteriovenous Malformation Treatment: A Systematic Review and Meta-Analysis. *Neurosurgery*, **92**, 27-41. <https://doi.org/10.1227/neu.0000000000002171>