

Explainable Machine Learning Model for Predicting Early Neurological Deterioration in Patients with Acute Ischemic Stroke

Tingting Huang

huangtingting1022@126.com

The First Affiliated Hospital of Anhui Medical University

Shoucai Zhao

First Affiliated Hospital of Wannan Medical College

Kai Wang

The First Affiliated Hospital of Anhui Medical University

Research Article

Keywords:

Posted Date: April 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9044703/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Explainable Machine Learning Model for Predicting Early Neurological Deterioration in Patients with Acute Ischemic Stroke

Tingting Huang^{1*}, Shoucai Zhao², Kai Wang^{1#}

1.The First Affiliated Hospital of Anhui Medical University,Hefei City,Anhui Province,China.

2.The First Affiliated Hospital of Wannan Medical College,Wuhu City,Anhui Province,China.

*First author:Tingting Huang

#Corresponding author:Kai Wang;The First Affiliated Hospital of Anhui Medical University,218 Jixi Road, Shushan District, Hefei City,Anhui Province,230022,China
E-mail:wangkai1964@126.com

Objective

This study aims to develop and validate predictive models for early neurological deterioration in patients with acute ischemic stroke using multiple machine learning methods.

Methods

A total of 1,285 patients with ischemic stroke admitted to Yijishan Hospital of Anhui Province from November 2020 to November 2024 were enrolled. The patients were randomly divided into a training set (70%) and a validation set (30%). Potential predictors were selected using a combination of Lasso regression and the Boruta algorithm. Seven machine learning algorithms—logistic regression, decision tree, random forest, XGBoost, k-nearest neighbor, light gradient boosting machine, and naïve Bayes—were used to build predictive models. Model performance was evaluated using AUC, sensitivity, specificity, and other indicators.

Results

Lasso regression and the Boruta algorithm jointly identified nine potential predictors: history of hypertension, TACI(Total Anterior Circulation Infarct), LACI(lacunar anterior circulation infarct), SII(Systemic Inflammatory Response Index), ARC(Acute-to-chronic glycemic ratio), HDL-C, LDL-C, ALB, and NIHSS . All seven

machine learning models demonstrated good performance in both the training and validation sets. Among them, the XGBoost model performed best in the validation set, with an AUC of 0.881 (95% CI: 0.834–0.928), sensitivity of 0.746, and specificity of 0.874, showing superior overall predictive ability compared to other models. Decision curve analysis (DCA) and calibration plots indicated excellent clinical benefit and discrimination ability. Finally, the SHAP summary plot was used to visualize and interpret the XGBoost model.

Conclusion

This study successfully developed the first machine learning-based predictive model for progressive ischemic stroke. Through model comparison and explainability analysis, the XGBoost model demonstrated superior predictive accuracy and clinical applicability, providing a reliable tool for early intervention.

Introduction

Stroke is one of the leading causes of death and disability in developing countries. Among them, the incidence of ischemic stroke has been rising year by year, with approximately 10%-40% of patients experiencing progressive ischemic stroke (PIS) within a short period after the initial onset, which is closely related to poor patient prognosis[1]. Better identification of high-risk groups for PIS is one of the important measures to improve patient outcomes. Currently, some studies have proposed predictive and identification methods such as the SITS, SIHS, and NIHSS scores[2-3]. However, these models have poor overall predictive efficiency, specificity, and applicability to different populations, and they do not account for varying racial, ethnic, and risk factors in a unified study.

As an important branch of artificial intelligence, machine learning algorithms can digitize relevant patient information and effectively screen and assess high-risk pathogenic factors. Some disciplines have integrated these algorithms with predictive models, validating the significant value of machine learning in disease prediction[4-5]. Based on the above considerations, this study aims to collect demographic data,

laboratory results, and non-traditional risk factor indicators of stroke patients. Through a combination of classical statistical analysis and novel machine learning algorithms, multiple predictive models for the risk of PIS will be developed, along with the creation of a visual risk prediction tool to better predict the occurrence of PIS, thereby providing strong support for clinical decision-making.

Materials and Methods

1.1 Study Population

This study was based on the inpatient electronic medical record system and included patients diagnosed with ischemic stroke who were admitted to Yijishan Hospital of Anhui Province between November 2020 and November 2024. All patients underwent neuroimaging examinations and were diagnosed with ischemic stroke, receiving standardized treatment according to relevant clinical guidelines. The NIHSS score was assessed at admission and again within 7 days after admission[6]. Progressive stroke was defined as an increase of ≥ 2 points in the NIHSS score within 7 days (including 1 point each for consciousness and limb motor function) or an increase of ≥ 4 points in total.

Inclusion criteria:

- (1) Age > 18 years, diagnosis of acute ischemic stroke confirmed by emergency clinical symptoms, neurological signs, and imaging findings;
- (2) Time from onset to admission < 24 hours;
- (3) All included patients received standardized guideline-based treatment, and their clinical and laboratory data were complete.

Exclusion criteria:

- (1) Baseline modified Rankin Scale (mRS) score > 2 ;
- (2) Contraindications to intravenous thrombolysis, endovascular therapy, or multimodal CT scanning;
- (3) Presence of malignancy, coagulation disorders, severe hepatic or renal insufficiency, or severe infections;
- (4) Missing clinical data.

1.2 Data Collection and Definitions

Clinical data during hospitalization were collected from the electronic medical record system, including demographic characteristics and medical history. Laboratory test indicators included routine blood tests, blood lipids, blood glucose, and liver and kidney function. All laboratory data were obtained from the first fasting blood sample collected on the morning after admission.

Definitions:

1.ARC (Acute-to-chronic glycemic ratio) = Admission blood glucose / $[28.7 \times \text{HbA1c} (\%) - 46.7]$

2.SII (Systemic Immune-Inflammation Index) = Platelet count (P) $\times$ Neutrophil count (N) / Lymphocyte count (L)

1.3 Statistical Analysis

All statistical analyses were performed using SPSS version 25.0 and R software version 3.6.3. The Kolmogorov–Smirnov test was used to assess normality of continuous variables. Normally distributed variables were expressed as mean $\pm$ standard deviation ($\bar{x} \pm s$) and compared between groups using the independent samples t-test. Non-normally distributed variables were presented as median (M) and interquartile range (P25–P75) and compared using the Mann–Whitney U (Z) test. Categorical variables were expressed as frequencies and percentages and compared using the chi-square (χ^2) test.

The variable selection process consisted of three steps: Preliminary screening using Lasso regression analysis; Identification of independent predictors using logistic regression analysis; Further optimization of feature selection using the Boruta algorithm.

Seven machine learning algorithms were used to construct predictive models: logistic regression (LR), light gradient boosting machine (LGBM), random forest (RF), decision tree (DT), XGBoost, k-nearest neighbor (KNN), and naïve Bayes (NB). Model training and validation were performed using tenfold cross-validation. Model performance was evaluated using multiple metrics, including area under the receiver

operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), recall, F1 score, and accuracy. Decision curve analysis (DCA) was used to assess clinical net benefit, calibration curves were used to evaluate the agreement between predicted and actual probabilities, and clinical impact curves were used to analyze the clinical decision-making value of the models. To enhance model interpretability, SHAP (SHapley Additive exPlanations) was applied for visualization of feature importance and to analyze each variable's contribution to the prediction outcome. All statistical tests were two-sided, and a P value <0.05 was considered statistically significant.

3. Results

3.1 Comparison of Baseline Characteristics

According to imaging diagnosis and the inclusion and exclusion criteria, a total of 1,285 patients were enrolled in the study. The detailed selection process is shown in Figure 1. Among the enrolled patients, 208 experienced disease progression within a short period, while 1,071 showed no progression. The characteristics of all included patients, including demographic information, medical history, and laboratory results, are summarized in Table 1.

Subsequently, the entire cohort was randomly divided into a training set (70%) and a validation set (30%). There were no significant differences in baseline characteristics between the two groups ($P > 0.05$). Detailed information is provided in Supplementary Table S1.

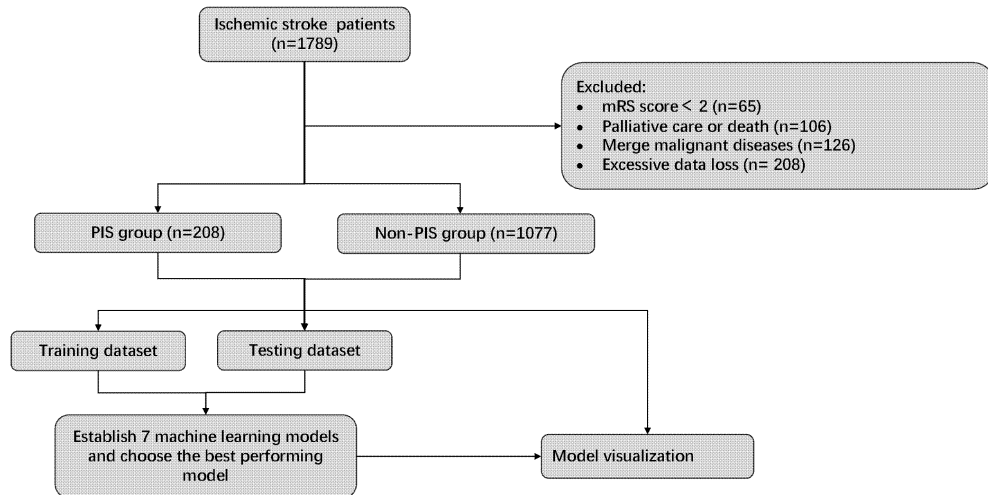


Figure 1 Flowchart and Framework of Model Construction

Table 1 Comparison of the Characteristics data Between the Training and Validation Groups for Model

Development				
	ALL (N=1285)	Non-PIS (N=1077)	PIS (N=208)	P
Sex:				0.141
Female	494(38.44%)	424(39.37%)	70(33.65%)	
Male	791(61.56%)	653(60.63%)	138(66.35%)	
HTN:				<0.001
NO	717(55.80%)	627(58.22%)	90(43.27%)	
YES	568(44.20%)	450(41.78%)	118(56.73%)	
DM:				0.799
NO	1039(80.86%)	869(80.69%)	170(81.73%)	
YES	246(19.14%)	208(19.31%)	38(18.27%)	
AF:				0.087
NO	929(72.30%)	768(71.31%)	161(77.40%)	
YES	356(27.70%)	309(28.69%)	47(22.60%)	

Smoke:				0.722
NO	771(60.00%)	649(60.26%)	122(58.65%)	
YES	514(40.00%)	428(39.74%)	86(41.35%)	
Drink:				1.000
NO	1157(90.04%)	970(90.06%)	187(89.90%)	
YES	128(9.96%)	107(9.94%)	21(10.10%)	
TACI:				<0.001
NO	1066(82.96%)	940(87.28%)	126(60.58%)	
YES	219(17.04%)	137(12.72%)	82(39.42%)	
PACI:				0.121
NO	1081(84.12%)	914(84.87%)	167(80.29%)	
YES	204(15.88%)	163(15.13%)	41(19.71%)	
POCI:				0.892
NO	947(73.70%)	795(73.82%)	152(73.08%)	
YES	338(26.30%)	282(26.18%)	56(26.92%)	
LACI:				<0.001
NO	762(59.30%)	583(54.13%)	179(86.06%)	
YES	523(40.70%)	494(45.87%)	29(13.94%)	
Age	63.00[54.00;72.00]	62.00[54.00;72.00]	64.00[55.00;73.00]	0.171
SBP	133.00[123.00;146.00]	133.00[123.00;146.00]	136.00[123.00;148.00]	0.190
DBP	79.00[71.00;86.00]	78.00[71.00;86.00]	79.50[70.00;89.00]	0.926
Hb	136.00[124.00;148.00]	136.00[125.00;148.00]	135.00[122.75;148.00]	0.529
WBC	7.56[6.01;9.42]	7.55[5.95;9.34]	7.72[6.38;9.81]	0.171
Neut	5.89[4.16;9.03]	5.74[4.00;9.10]	6.60[5.20;8.85]	0.004
Lym	1.55[1.23;1.96]	1.56[1.23;1.97]	1.52[1.19;1.90]	0.364
Mono	0.55[0.41;0.74]	0.54[0.41;0.73]	0.60[0.43;0.77]	0.126
PLT	181.00[150.00;222.00]	181.00[150.00;222.00]	178.50[152.00;221.50]	0.921
SII_100	6.91[4.45;10.66]	6.73[4.33;10.47]	8.06[5.14;11.92]	<0.001
Glu	5.82[5.05;6.91]	5.63[4.97;6.82]	6.50[5.70;7.77]	<0.001

HbA1c	5.70[5.30;6.40]	5.60[5.30;6.30]	5.90[5.50;6.54]	0.002
ARC	0.89[0.76;1.04]	0.88[0.75;1.02]	0.97[0.85;1.17]	<0.001
TG	1.80[1.24;2.85]	1.75[1.21;2.76]	2.08[1.39;3.15]	0.003
TC	4.34[3.56;5.05]	4.34[3.55;5.04]	4.37[3.64;5.11]	0.187
HDL	1.07[0.89;1.33]	1.09[0.90;1.37]	0.99[0.84;1.17]	<0.001
LDL	2.81[2.14;3.61]	2.73[2.03;3.56]	3.32[2.74;3.73]	<0.001
LPA	183.00[105.30;315.20]	172.70[99.00;294.80]	228.10[145.55;387.65]	<0.001
apoA1	1.25[1.11;1.43]	1.26[1.11;1.44]	1.23[1.10;1.39]	0.237
apoB	0.93[0.75;1.12]	0.92[0.74;1.11]	0.97[0.78;1.20]	0.002
ALB	41.10[37.80;44.30]	41.50[38.40;45.10]	38.80[35.85;41.10]	<0.001
ALT	29.00[20.00;43.00]	27.00[19.00;41.10]	36.50[23.75;45.00]	<0.001
AST	29.00[24.00;47.00]	29.00[23.00;47.00]	38.00[26.00;47.00]	<0.001
K	3.97[3.70;4.24]	3.99[3.71;4.26]	3.87[3.62;4.12]	0.001
UA	341.00[278.50;403.50]	340.30[280.00;399.40]	343.10[268.30;421.25]	0.900
Cr	74.95[61.00;87.00]	74.90[61.00;86.00]	75.10[63.75;90.25]	0.166
NHIS	10.00[8.00;12.00]	9.00[8.00;12.00]	12.50[9.00;16.25]	<0.001

SBP:systolic blood pressure; DBP:diastolic blood pressure; HB:Hemoglobin; PLT:platelet count; WBC:white blood cell; NE:neutrophilicgranulocyte; Lym:lymphocyte; Mono:monocyte; SII: Systemic Inflammatory Response Index; FBG:fasting blood glucose; TG:Triglyceride; TC:total cholesterol; HDL-C:High density lipoprotein cholesterol; LDL-C:Low-Density Lipoprotein Cholesterol; Apo A1:apolipoprotein A-1; Apo B: Apolipoprotein B; ALT:alanine aminotransferase; AST:Aspartate aminotransferase; ALP :alkaline phosphatase; UA:urine acid; Cr:Creatinine;ARC (Acute-to-chronic glycemic ratio) ; SII (Systemic Immune-Inflammation Index) ; TACI: Total Anterior Circulation Infarct; PACI:Partial Anterior Circulation Infarct; POCI:Posterior Circulation Infarct; LACI:lacunar anterior circulation infarct

3.2 Screening of Relevant Features

We first performed Lasso regression analysis on the features with significant differences in the training set (Supplementary Table S1). When the λ_{min} in the Lasso regression was set to 0.032, 9 predictor variables showed non-zero coefficients, as

shown in Figures 2A and 2B. In the Boruta algorithm, a total of 13 relevant variables were identified (Figure 2C). We used a Venn diagram (Figure 2D) to identify the predictor factors included in both algorithms, ultimately determining 9 predictor variables: history of hypertension, TACI, LACI, SII100, ARC, HDL, LDL, ALB, and NHISSscore. To avoid potential effects of collinearity on the ranking of variable importance, we confirmed through collinearity analysis that there is no potential multicollinearity among the aforementioned variables (Supplementary Table S2). These variables will be incorporated into the subsequent machine learning models.

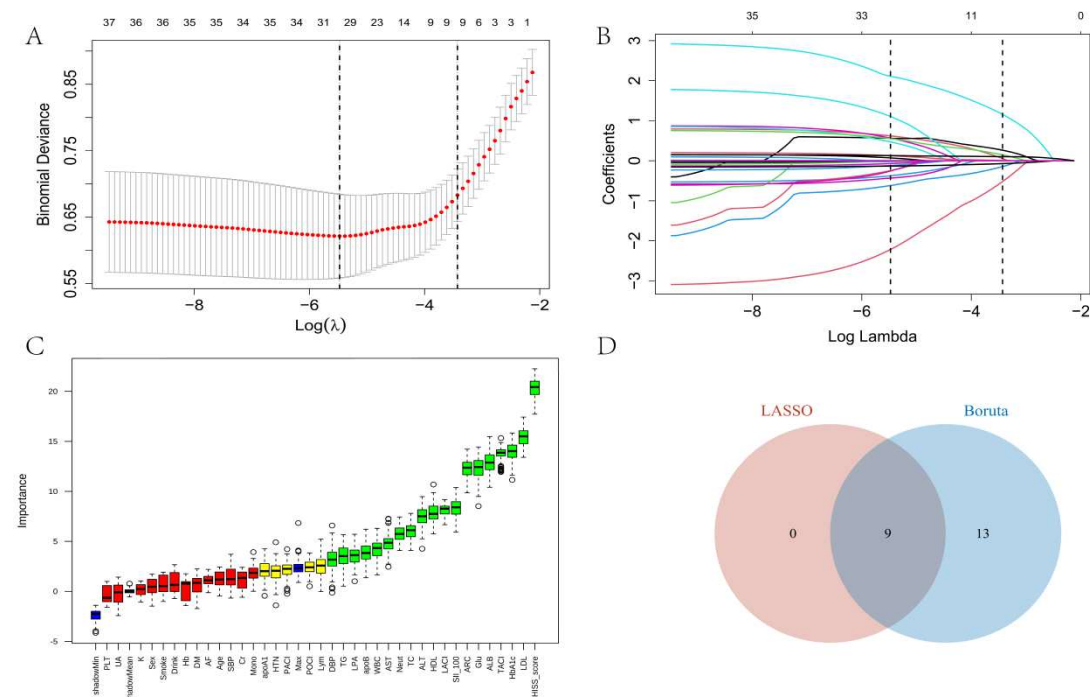


Figure 2. Screening of Relevant Influencing Factors

- (A) Factor selection based on the LASSO regression model;
- (B) Variable trajectory of the LASSO regression model;
- (C) Boruta algorithm;
- (D) Common predictive factors identified between Boruta and LASSO analyses.

3.3 Construction and Performance Evaluation of Multiple Models

All seven models demonstrated good predictive performance in both the training and validation sets (Table 2). Among them, the XGBoost model achieved the highest AUC value, with superior sensitivity, specificity, and other evaluation metrics.

Through comparison using decision curve analysis (DCA), calibration curves, and confusion matrices, we found that within a reasonable threshold range, the XGBoost model provided greater net clinical benefit and showed better discriminative ability than other models.

Based on the comprehensive analysis of all parameters from both the training and validation sets, we concluded that the XGBoost model exhibited the best overall predictive performance (Figure 3).

Table 2. Comparison of predictive performance among different models

Model	AUC(95%CI)	Sensitivity	Specificity	PPV	NPV	Recall	F1	Prevalence	Accuracy
Trainingset									
LR	0.857(0.822-0.893)	0.816	0.75	0.377	0.956	0.816	0.516	0.157	0.816
DT	0.927(0.902-0.952)	0.61	0.988	0.905	0.932	0.61	0.729	0.157	0.61
RF	1(1-1)	1	1	1	1	1	1	0.157	1
XGB	0.941(0.921-0.961)	0.858	0.896	0.605	0.971	0.858	0.71	0.157	0.858
KNN	1(1-1)	1	1	1	1	1	1	0.157	1
LGBM	0.927(0.904-0.95)	0.858	0.881	0.573	0.971	0.858	0.688	0.157	0.858
NBM	0.855(0.823-0.887)	0.794	0.765	0.386	0.952	0.794	0.52	0.157	0.794
Validationset									
LR	0.837(0.785-0.889)	0.791	0.714	0.368	0.942	0.791	0.502	0.174	0.791
DT	0.764(0.699-0.83)	0.343	0.931	0.511	0.871	0.343	0.411	0.174	0.343
RF	0.856(0.809-0.902)	0.373	0.953	0.625	0.878	0.373	0.467	0.174	0.373
XGB	0.881(0.834-0.928)	0.746	0.874	0.556	0.942	0.746	0.637	0.174	0.746
KNN	0.793(0.732-0.854)	0.478	0.903	0.508	0.891	0.478	0.492	0.174	0.478
LGBM	0.831(0.776-0.885)	0.672	0.843	0.474	0.924	0.672	0.556	0.174	0.672
NBM	0.836(0.79-0.882)	0.821	0.758	0.417	0.953	0.821	0.553	0.174	0.821

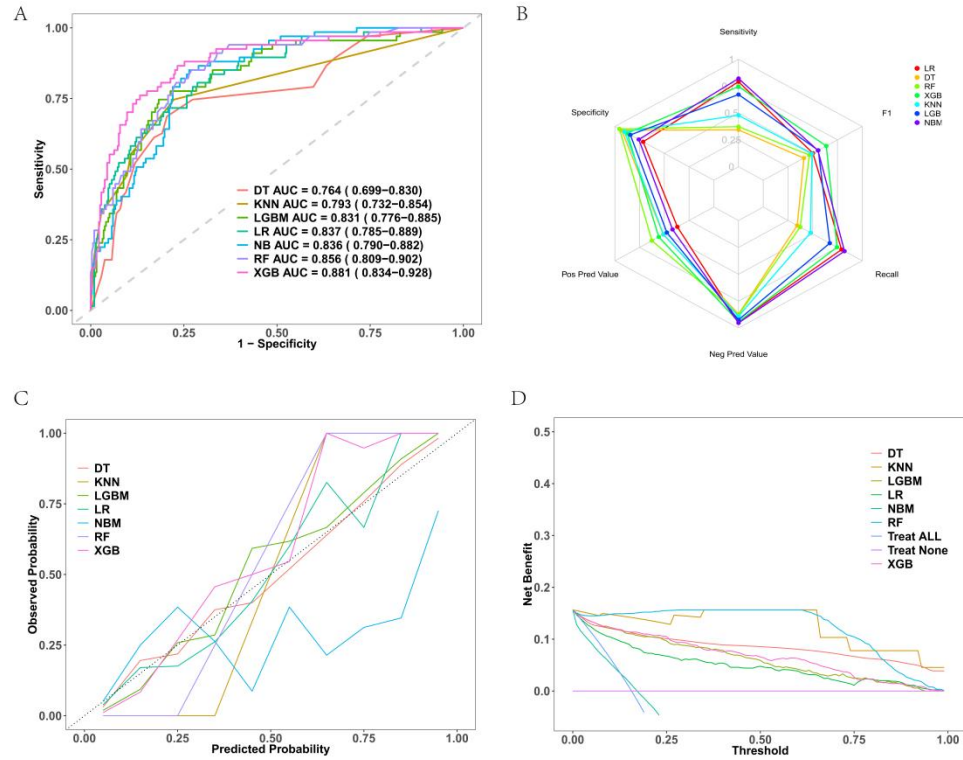


Figure 3. Comparison of predictive performance of various models in the validation set.

Panels A, B, C, and D correspond to the ROC curves, radar plots of predictive capability, DCA curves, and decision curves of different machine learning models in the validation set, respectively.

3.4 Model Interpretation

The SHAP (Shapley Additive Explanations) method was used to interpret and visualize the XGBoost model, ranking the importance of factors influencing disease progression (Figure 4A). Each point on the horizontal axis represents an individual sample, with the color gradient indicating the magnitude of the corresponding feature value, while the vertical axis displays the ranking and distribution of related features. To enhance understanding of the model's decision-making process at the individual level, we conducted detailed interpretability analyses on two representative samples (Figure 4B). By visualizing the SHAP values of these samples, we were able to identify the contribution of each feature to the predictions for these specific cases

(Figures 4C and 4D).

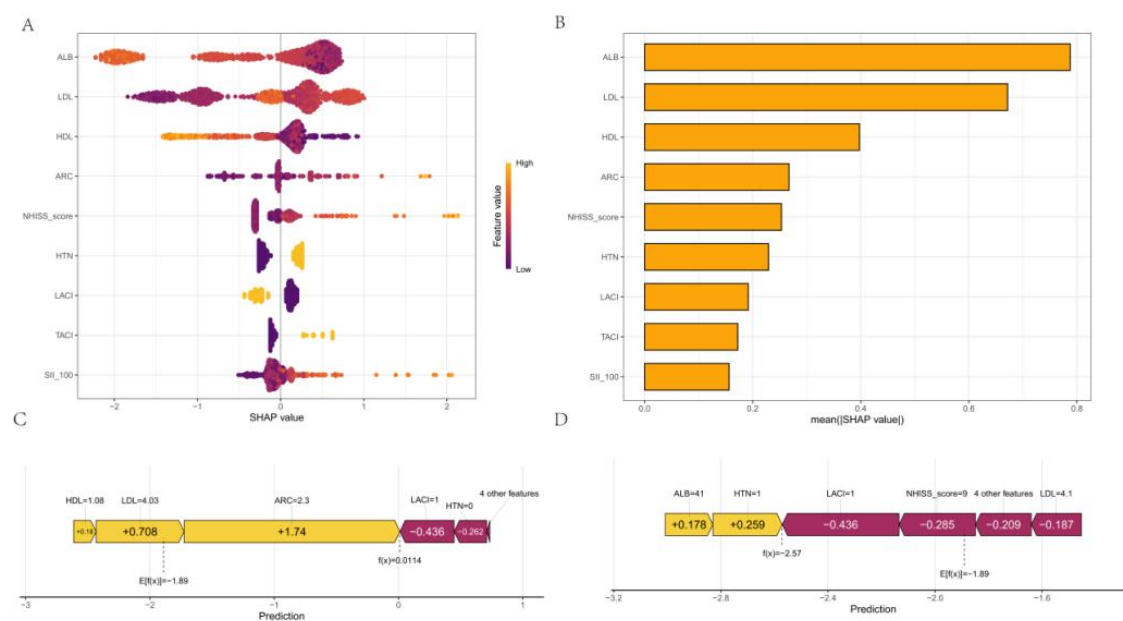


Figure 4.

- (A) SHAP interpretation of the model.
- (B) Feature importance ranking in the SHAP model.
- (C) Example of a patient with PIS.
- (D) Example of a non-PIS patient with no coronary artery progression.

3. Discussion

This study included a total of 37 related influencing factors, including personal medical history, medication use, and laboratory test results. Using a combination of Boruta and LASSO algorithms for feature selection, we identified HTN, TACI, LACI, SII100, ARC, HDL, LDL, ALB, and NIHSSscore as independent predictors of PIS occurrence. Subsequently, we constructed predictive models for PIS using various machine learning algorithms, among which the XGBoost model demonstrated superior predictive efficiency and clinical applicability. Furthermore, we applied the SHAP method for model visualization, enhancing the interpretability and clinical

utility of the model.

Ischemic stroke (AIS) remains one of the leading causes of death and disability worldwide. Short-term neurological progression has significant implications for patient prognosis. PIS is independently associated with in-hospital mortality, increasing the risk of death by more than 2.5 times. Moreover, PIS is closely related to post-stroke depression and recurrent cardiovascular events [7]. Epidemiologically, the occurrence of PIS is associated with multiple factors, including large-artery occlusion, dysglycemia, poor hypertension control, infection, and insufficient collateral circulation compensation [8-9]. In this study, we found that traditional atherosclerotic risk factors such as a history of hypertension (HTN), HDL, and LDL were independent risk factors influencing early progression of AIS. The first International Stroke Trial reported a “U-shaped” relationship between systolic blood pressure (SBP) and outcomes in ischemic stroke patients, with both excessively high and low SBP independently associated with poor outcomes [10]. San Martin et al. found that when SBP was below 150 mm Hg, each 10 mm Hg decrease increased mortality by 17.9%; when SBP exceeded 150 mm Hg, each 10 mm Hg increase raised early mortality by 3.8% ($P = .016$). Additionally, each 10 mm Hg increase in SBP raised the 2-week risk of recurrent stroke by 4.2% ($P = .023$) [11]. The balance between LDL-C and HDL-C also plays an important role in AIS onset and early progression [12]. A recent large randomized controlled trial demonstrated that intensive lipid-lowering therapy with the PCSK9 inhibitor Evolocumab combined with statins effectively reduced LDL-C and inflammatory marker levels within 7 days after non-cardioembolic AIS, thereby reducing the risk of PIS (RR = 0.90; 95% CI -1.59 to -0.22; $p = 0.010$), highlighting the importance of lipid management in improving PIS risk [13]. Similarly, Shan Lv et al. found in a predictive model for AIS patients receiving thrombolytic therapy that HDL-C/LDL-C was a risk factor for PIS (OR 3.29; 95% CI 1.00–10.89; $p = 0.049$) [14].

In this study, we used the Systemic Immune-Inflammation Index (SII) and the acute-to-chronic glycemic ratio (ARC) to assess inflammatory and stress responses, respectively, and confirmed both as independent risk factors for PIS. The SII

combines platelet, neutrophil, and lymphocyte counts to evaluate immune and inflammatory status [15]. Ghulam Jeelani et al. analyzed thrombi from stroke patients and observed abundant activated neutrophils undergoing NETosis in platelet-rich regions and at red blood cell interfaces, releasing citrullinated histone (H3Cit)-modified DNA networks. The close platelet–neutrophil interaction and co-localization of NETs with VWF, platelets, and HMGB1 suggest that these components synergistically enhance thrombus stability. Thus, SII can reflect the relationship between coagulation and inflammation in AIS patients [16]. ARC reflects the balance between stress and chronic glycemia, providing a more accurate representation of pancreatic function under stress [17]. Elisenda Climent’s recent study found that ACR was independently associated with poor 3-month outcomes and mortality in AIS patients: for every 0.1 increase in ACR, the risk of poor outcome increased by 7% and mortality by 10%. Patients in the highest ACR tertile (≥ 1.13) had a 62% higher risk of poor outcome and an 88% higher mortality risk, independent of diabetes status—consistent with our findings [18].

Previous studies have developed multiple models to predict short-term outcomes and PIS risk after AIS, including the N3H3 nomogram (NIHSS, 24h NIHSS change, hypertension, HCY, HDL/LDL) and the START nomogram (NIHSS, age, pre-stroke mRS, onset-to-thrombolysis time) [19–20]. Although these models share some overlapping predictors with ours, their overall predictive performance remains unclear and limited by incomplete consideration of potential risk factors, which may reduce their generalizability. Our study is, to our knowledge, the first to apply machine learning for predicting PIS risk, incorporating demographic, treatment, and laboratory data as well as composite inflammatory and stress-related indices. By jointly applying LASSO and Boruta algorithms for feature selection, and evaluating multiple machine learning algorithms, we constructed, validated, and visualized the prediction models. This approach not only integrated traditional risk factors and clinical characteristics but also introduced nontraditional indicators, improving model robustness and applicability. Therefore, we selected the XGBoost model as the optimal algorithm and applied SHAP to enhance interpretability and precision in risk quantification. All

required predictors are routinely available clinical variables during hospitalization, with no additional testing or computational cost, improving the clinical practicality of our model.

Conclusion

To the best of our knowledge, this study is the most recent machine learning predictive model constructed for the risk of early progression of acute ischemic stroke. We innovatively integrated two feature selection methods (LASSO regression and Boruta) with seven machine learning algorithms (including logistic regression, support vector machine, and XGBoost). Through cross-validation and comprehensive performance evaluation (AUC, precision, recall), we identified XGBoost as the optimal model. To enhance clinical interpretability, we employed SHAP (SHapley Additive exPlanations) for feature importance visualization, revealing the contribution of both traditional and novel risk factors (such as SII and ARC) to PIS occurrence. This model may provide a scientific basis for early identification of high-risk PIS patients and for developing individualized intervention strategies, demonstrating significant translational potential in clinical practice.

Limitations

This study is a single-center retrospective study with a limited sample size, and it only performed internal validation of the model. This this study, as well as the predictive indicators in the model, are limited. Potential influencing factors such as genetics, economics, and culture were not incorporated. Future work will focus on multi-center, large-sample data collection, model improvement, and validation to enhance the predictive capability of the model.

Funding

This study was financially supported by Wuhu Municipal Science and Technology Bureau Project(NO:2024kj048).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

The datasets used and/or analysed during the current study available from the

corresponding author on reasonable request.

Author contributions

Tingting Huang wrote the manuscript and conducted the data analysis. Shoucai Zhao performed the study search and screening. Kai Wang reviewed the article and supervised the entire process. All authors have read and approved the final manuscript.

Ethical statement

This study was conducted in accordance with the ethical standards outlined in the Declaration of Helsinki. The study protocol was approved by the Ethics Committee of The First Affiliated Hospital of Anhui Medical University. Due to the retrospective design of this study, the Ethics Committee agreed to waive the requirement for written informed consent and allowed the use of de-identified patient information for this retrospective study.

Reference

[1]Ryu WS, Hong KS, Jeong SW, et al. Association of ischemic stroke onset time with presenting severity, acute progression, and long-term outcome: A cohort study. *PLoS Med.* 2022;19(2):e1003910. Published 2022 Feb 4. doi:10.1371/journal.pmed.1003910

[2]Sharma D, Aravind S, Joseph S, Fagera N, Rajagopalan G. The Role of Monocyte to High-density Lipoprotein Cholesterol Ratio in Predicting the Severity of Acute Ischemic Stroke and its Association with the NIHSS. *J Assoc Physicians India.* 2023;71(8):11-12. doi:10.59556/japi.71.0321

[3]Montellano FA, Ungethüm K, Ramiro L, et al. Role of Blood-Based Biomarkers in Ischemic Stroke Prognosis: A Systematic Review. *Stroke.* 2021;52(2):543-551. doi:10.1161/STROKEAHA.120.029232

[4]Polson JS, Zhang H, Nael K, et al. Identifying acute ischemic stroke patients within the thrombolytic treatment window using deep learning. *J Neuroimaging.* 2022;32(6):1153-1160. doi:10.1111/jon.13043

[5]Carluccio E, Cameli M, Rossi A, et al. Left Atrial Strain in the Assessment of Diastolic Function in Heart Failure: A Machine Learning Approach. *Circ Cardiovasc Imaging.* 2023;16(2):e014605. doi:10.1161/CIRCIMAGING.122.014605

[6]Weimar C, König IR, Kraywinkel K, Ziegler A, Diener HC; German Stroke Study Collaboration. Age and National Institutes of Health Stroke Scale Score within 6 hours after onset are accurate predictors of outcome after cerebral ischemia: development and external validation of prognostic models. *Stroke*. 2004;35(1):158-162.

[7]Saber H, Liebeskind DS. Infarct Progression in the Early and Late Phases of Acute Ischemic Stroke. *Neurology*. 2021;97(20 Suppl 2):S60-S67. doi:10.1212/WNL.0000000000012795

[8]Seners P, Yuen N, Olivot JM, et al. Factors Associated With Fast Early Infarct Growth in Patients With Acute Ischemic Stroke With a Large Vessel Occlusion. *Neurology*. 2023;101(21):e2126-e2137. doi:10.1212/WNL.00000000000207908

[9]Li H, Ke X, Feng B, et al. Research progress on the mechanism and markers of metabolic disorders in the occurrence and development of cognitive dysfunction after ischemic stroke. *Front Endocrinol (Lausanne)*. 2025;16:1500650. Published 2025 Jan 22. doi:10.3389/fendo.2025.1500650

[10]International Stroke Trial Collaborative Group. The International Stroke Trial (IST): a randomised trial of aspirin, subcutaneous heparin, both, or neither among 19435 patients with acute ischaemic stroke. *Lancet* 1997;349:1569–81.

[11]Feldstein CA. Early treatment of hypertension in acute ischemic and intracerebral hemorrhagic stroke: progress achieved, challenges, and perspectives. *J Am Soc Hypertens*. 2014;8(3):192-202. doi:10.1016/j.jash.2013.09.004

[12]Demarin V, Lisak M, Morović S, Cengić T. Low high-density lipoprotein cholesterol as the possible risk factor for stroke. *Acta Clin Croat*. 2010;49(4):429-439.

[13]Tian W, Cao H, Li X, et al. Adjunctive PCSK9 Inhibitor Evolocumab in the Prevention of Early Neurological Deterioration in Non-cardiogenic Acute Ischemic Stroke: A Multicenter, Prospective, Randomized, Open-Label, Clinical Trial. *CNS Drugs*. 2025;39(2):197-208. doi:10.1007/s40263-024-01145-5

[14]Lv S, Song Y, Zhang FL, et al. Early prediction of the 3-month outcome for individual acute ischemic stroke patients who received intravenous thrombolysis using the N2H3 nomogram model. *Ther Adv Neurol Disord*. 2020;13:1756286420953054. Published 2020 Sep 4. doi:10.1177/1756286420953054

[15]Wang RH, Wen WX, Jiang ZP, et al. The clinical value of neutrophil-to-lymphocyte ratio (NLR), systemic immune-inflammation index (SII), platelet-to-lymphocyte ratio (PLR) and systemic inflammation response index (SIRI) for predicting the occurrence and severity of pneumonia in patients with intracerebral hemorrhage. *Front Immunol.* 2023;14:1115031. Published 2023 Feb 13. doi:10.3389/fimmu.2023.1115031

[16]Pir GJ, Parray A, Ayadathil R, et al. Platelet-Neutrophil Association in NETs-Rich Areas in the Retrieved AIS Patient Thrombi. *Int J Mol Sci.* 2022;23(22):14477. Published 2022 Nov 21. doi:10.3390/ijms232214477

[17]Deng Y, Ma Y, Gao J, Sun Q, Zhang J, Zhu L. Association between acute-to-chronic glucose ratio and acute kidney injury after emergency PCI in patients with acute myocardial infarction. *Heart Lung.* 2025;73:48-55. doi:10.1016/j.hrtlng.2025.04.025

[18]Climent E, Rodriguez-Campello A, Jiménez-Balado J, et al. Acute-to-chronic glycemic ratio as an outcome predictor in ischemic stroke in patients with and without diabetes mellitus. *Cardiovasc Diabetol.* 2024;23(1):206. Published 2024 Jun 18. doi:10.1186/s12933-024-02260-9

[19]Cappellari M, Turcato G, Forlivesi S, et al. The START nomogram for individualized prediction of the probability of unfavorable outcome after intravenous thrombolysis for stroke. *Int J Stroke.* 2018;13(7):700-706. doi:10.1177/1747493018765490

[20]Song B, Chen X, Tang D, et al. External Validation of START nomogram to predict 3-Month unfavorable outcome in Chinese acute stroke patients. *J Stroke Cerebrovasc Dis.* 2019;28(6):1618-1622. doi:10.1016/j.jstrokecerebrovasdis.2019.02.032

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplymaterial.docx](#)