

Machine Learning Methods to Predict Mortality of Patients with Intracerebral hemorrhage in the ICU

dongfeng liu¹, Siru Liu¹, Qinwen Shi¹, Ke Li¹, and Jialin Liu¹

¹Affiliation not available

May 5, 2020

Abstract

Rationale, aims and objectives: Intracerebral hemorrhage (ICH), the second most common cause of stroke, has a high fatality rate. The establishment of mortality prediction models based on ICH patients and disease characteristics is very useful for clinical decision-making and corresponding treatment methods. Therefore, we used five machine learning methods to establish models for predicting in-hospital mortality in ICH patients and compared models' performance. **Methods:** Model development and performance comparisons were performed using the medical information mart for intensive care (MIMIC-III) database. We took the maximum and minimum values of each index of 1143 ICH patients in the first, second and third days after admission as the input variables of the model, and established five machine learning models including random forest (RF), Gradient Boosting Decision Tree (GBDT), decision tree, Naïve Bayes and KNN. The most important feature variables were selected by the RF model and Least Absolute Shrinkage and Selection Operator (LASSO) method. The area under the receiver operating characteristic curve (AUROC), accuracy, precision, recall, and F1 score were used as the assessment criteria of the model prediction effect. **Results:** After 5-fold cross-validation, the AUROC of RF, GBDT, Naïve Bayes, Decision Tree and KNN models were 0.92, 0.93, 0.9, 0.89, 0.89, respectively. The performance of GBDT was better than other prediction models. The accuracy, precision, recall, and F1 score of the GBDT model were respectively 0.87, 0.84, 0.76, and 0.79. **Conclusions:** There is great potential for machine learning in mortality prediction for ICH patients in ICU. Considering the above five models, we believe that GBDT is an appropriate tool for clinicians to predict ICH patient mortality.

1. Introduction

Intracerebral hemorrhage (ICH) is a serious disease caused by rupture of blood vessels in the brain and is common in the intensive care unit (ICU)^[1]. Unlike ischemic strokes that are often preceded by a transient ischemic attack, ICH appears suddenly without any warning^[2]. Consequently, ICH is the most fatal form of stroke^[3,4]. Two million people per year are affected by ICH, accounting for 10-15% of the world's cerebral stroke patients^[1,5]. ICH patients' mortality approaches 40-50%, and disability in survivors is common^[3,4,6,7]. Although much progress has been made in ICU research, clinical outcomes after ICH have not improved significantly in the last few decades^[6,8,9]. Therefore, identifying potential patients to provide them early treatment would be an effective approach to control the ICH disease.

Several prognostic tools have been proposed for mortality and functional outcome prediction in ICH. These tools can help clinicians select the best treatment for ICH patients, facilitate communication between clinicians and patients, and serve as indicators for optimal allocation of medical resources in the ICU^[10,11,12]. Peng et al.^[13] established models such as RF, support vector machine (SVM), and logistic regression to predict the 30-day mortality of ICH patients. Eighteen indicators including demographic information, physiological characteristics, and laboratory parameters were used and the results showed that RF had the best predictive performance. The majority of the patients included in this study are Asian, and the number of cases (423 ICU patients) and model parameters is small, so the universality is limited.

Considering the limitations mentioned above, most of the existing prognostic tools for ICH patients have the problems of insufficient cases or few parameters, and less research has added the temporal information that can observe the changes of diseases and make the model prediction more accurate to the indicators of ICH patients, which has questioned the accuracy and applicability of the model prediction^[13]. In this study, we aim to use machine learning to develop and validate an ICH in-hospital mortality prediction model. The publicly accessible ICU database MIMIC III^[14] was used for data selection and model development. We added temporal information to ICH patients and analyzed broader variables that might affect mortality in ICH patients. Then we compared random forest (RF) model with Gradient Boosting Decision Tree (GBDT), decision tree, K-Nearest Neighbor (KNN), and Naïve Bayes models. Considering the explainability, we also used the feature importance of RF and LASSO regression to select some important features.

2. Method

2.1. Data source

In this study, we used the Medical Information Mart for Intensive Care III (MIMIC-III) 2001-2012^[14] for the data extraction and model development. There were 53,423 patients with different diseases and the database was included in the vital signs, drugs and laboratory measurement, observation and care providers of record, liquid balance, program code, diagnosis, imaging report, length of hospital stays, survival data, etc. MIMIC-III included almost comprehensive clinical data of patients admitted to the Beth Israel Deaconess Medical Center in Boston, Massachusetts.

2.2. Patient selection

Patients admitted to ICU who were diagnosed with ICH (ICD-9 = 431) were eligible for inclusion. In the MIMIC-III database, each patient has its ID (SUBJECT_ID), and each admission is reassigned with an ID (HADM_ID). We screened out the variables of each ICH patient at the time of the last admission in the database record period (that is, each SUBJECT_ID corresponds to a HADM_ID). After extracting, the data of 1143 ICH patients older than 18 years were included and divided into the dead (383 people) and the alive group (760 people).

2.3. Predictors and outcome variables

For each ICH patient, we collected variables including, patient physiological characteristics, laboratory parameters, and CT findings. For variables with temporal information, the maximum and minimum values of daily variables within three days after admission were extracted, and we took them as the parallel input of the model, because we believed that these variables with time information could more directly reflect the limiting condition of the patient's body and had a more obvious influence on whether the patient died or not. After consulting with experts in this area, we selected variables: Gender, marital status, ethnicity, age, complication (high blood pressure, diabetes, ischemic heart disease, heart failure, pneumonia), and ICH location (basal ganglia, lobe, infratentorial). In addition, variables with temporal information (the maximum and minimum values of the first, second and third days of admission) included: systolic pressure, diastolic pressure, mean blood pressure, heart rate, body temperature, GCS score, white blood cells, lymphocytes, neutrophils, eosinophils, basophils, red blood cells, hemoglobin, platelet count, hematocrit, glucose, potassium, sodium, creatinine, and urea nitrogen. The outcome of this study was in-hospital mortality of ICH patients. If the variable's missing values were more than 20% of the total, we removed it.

2.4. Preprocessing

Missing values were common in data extraction. If a patient died or was discharged on the first day (or second day) of admission, there was no data entry on the remaining days. In that case, we assumed that the data on the first day (or second day) could represent the data on the remaining days. Patients whose data lacked more than 20% of the variables were removed. Other missing values were supplemented by the median of the dead and alive group. Descriptive data are expressed as actual numbers and percentages or mean \pm standard deviations. Five-fold cross-validation is used in the model. The whole group of data was randomly divided into five pieces. One of the subsamples was retained as the test set, and the remaining

four subsamples were used as the training data. Then the cross-validation process was repeated five times. The average of the five times generated was used to represent the performance of each model.

In order to better illustrate the relationship between variables, One-Hot Encoding is adopted to deal with text-based variables. Since Naïve Bayes is not good at dealing with data sets with different dimensions, we standardized the data before the training to make the data conform to the standard normal distribution, that is, the mean value is 0 and the standard deviation is 1.

2.5. Modeling

In this study, five prediction models of RF, GBDT, decision tree, KNN, and Naïve Bayes models were established and compared.

Naïve Bayes is the Bayesian theory based on probability theory and mathematical statistics. It is a supervised algorithm that directly measures the probability relationship between labels and features^[15]. Naïve Bayes' simplicity makes the model run quickly. However, the condition of its validity is that the features of samples are independent^[16]. That is difficult in practical application.

K-Nearest Neighbor (KNN) is a simple non-parametric classification method. Assuming the data set d to be classified, k of its nearest neighbors was retrieved and calculated as the *neighborhood of d* . Whether or not weight based on distance is considered, the category of d is usually determined by most data records in the neighborhood^[17].

Gradient Boosting Decision Tree (GBDT) was first proposed by Jerome H. Friedman^[18]. The trees in GBDT are regression trees, which can be used for regression prediction and classification. The core of GBDT is that each tree learns the residual (negative gradient) of the sum of all previous tree conclusions, which is the sum of the true value after adding the predicted value.

The essence of the decision tree algorithm is the graph structure, which, like KNN, is a non-parametric supervised learning algorithm. It can summarize decision rules from a series of data with features and labels and present them in a tree structure. A decision tree contains root nodes, intermediate nodes, and leaf nodes. The decision tree follows the principle of top-down segmentation^[19], which means from the root node, the division is carried out according to the principle of minimum impurity, and the growth stops when the number of records in the node falls below the preset threshold. The commonly used measurement methods of impurities are Gini index and information entropy^[20]:

$$\text{Gini}(t) = 1 - \sum_{i=0}^{c-1} p(i|t)^2 \quad (1)$$

$$\text{Entropy}(t) = - \sum_{i=0}^{c-1} p(i|t) \log_2 p(i|t) \quad (2)$$

Where t represents a given node, i represents any classification of labels, and $p(i|t)$ represents the proportion of label classification i in node t . Since the effect of information entropy and Gini coefficient is the same in practical application, information entropy is selected to calculate impurities in the decision tree constructed in this study.

Random forest (RF) is a very representative bagged ensemble algorithm. All its basic evaluators are decision trees, and the forest composed of classification trees is the random forest classifier. Each decision tree provides a different solution to the problem. The solutions of all the decision trees are eventually combined (usually by voting or averaging) into a single final model output^[21], which is usually a more stable and accurate prediction. In this study, the model predicted whether a patient would die in hospital by providing the probability of death for each patient. The probability is determined by the ratio of the decision tree with positive results to the total number of decision trees.

For the selection of feature variables, we adopted the feature importance attribute of RF and LASSO regression. The feature importance of RF indicates the contribution of each predictor to the model, so the most relevant predictor can be selected to represent most of the performance of the model. Least Absolute Shrinkage and Selection Operator (LASSO) imposes a constraint on the sum of the absolute values of the model parameters, applying a regularization process in which it penalizes the coefficients of the regression variables and sets some of them precisely to zero^[22]. Variables with non-zero results were selected as important variables.

2.6. Data analysis

We used MySQL and Tableau (version 2019.1.0) to extract data from MIMIC-III v1.4, and Python3.7 to build models and process data. The working environment was PyCharm 2018.3.2.

Each of machine learning models involves a parameter-tuning process, so we used the learning curve and grid search to select the parameters to get the best model (Table A.1 in Appendices). In this study, accuracy rate, precision rate (P), recall rate (R), F1 score, and area under ROC curve (AUROC) were used to measure the model performance.

Results:

A total of 1143 ICH patients were obtained from the MIMIC-III database, including 760 survivors and 383 died. Table 1 shows the clinical characteristics of patients who survived and died during hospitalization. Table A.2 in Appendices shows the changes of physiological characteristics and laboratory parameters of dead and alive patients over time.

First, we used all 122 variables to construct five models, and used the learning curve and grid search to determine the optimal parameters. Prediction performance comparison results after 5-fold cross validation are shown in Table 2.

We can find that the GBDT model have the best accuracy (0.87) and the best F1 score (0.80). In terms of accuracy and AUROC, GBDT model had better values than other models, 0.87 and 0.93, respectively. Naïve Bayes had the best recall rate (0.85), but its accuracy and precision were the lowest. KNN model had the lowest recall rate (0.60), F1 score (0.70) and AUROC (0.87). Then we used the feature importance of RF and LASSO regression to select the most important feature variables. The first 39 most important variables were selected by two methods respectively (Table A.3 in Appendices), and the intersection of the two methods was taken as the screened variables, a total of 18. The importance order of the intersection variables is shown in Fig. 1. The importance score is normalized value, distributed between 0 and 1, and the closer to 1, the more important the variable is.

We reconstructed and trained five models with 18 variables obtained, and observed the changes of each indicator as shown in Table 3. The ROC curves of these predictive models are presented in Fig. 2.

Compared with the model constructed with all variables, it was found that although GBDT model has a small decline in precision, recall and F1 value, it can be seen from the AUROC index that GBDT model was the best among the five models. From the results, we found that the prediction effect of all the five models had not decreased significantly. Therefore, the input variables of our models were reduced from 122 to 18 successfully, greatly improving the practicability.

4. Discussion

This study established a prediction model based on RF, GBDT, decision tree, Naïve Bayes, and KNN algorithms to predict the in-patient mortality of ICH patients, and compared the effects of the five models. The results showed that the AUROC and precision of the RF model were better than other models. RF feature importance and LASSO method were used to screen out fewer variables, so that the model was simplified and the performance is not affected. This had important implications for practical applications of the model, because researchers can predict mortality rates from just 18 data sets of patients without using more than 100 variables.

This study adds temporal information into the prediction of mortality model of ICH patients. We took the maximum and minimum values of each indicator in the first, second and third days after admission of ICH patients as the model input variables, which can make the model learn more information automatically and make the prediction effect better. Compared with previous studies^[23-26], the AUROC obtained by the GBDT model in this study reached 0.93, which was higher than the previous results, although the data sources and methods used were different.

From the results, the indicators of the first, second and third days all had an important impact on the death of ICH patients. This has not been considered in previous studies. The most important influencing factor is GCS score, which is consistent with literature reports^[27]. The lower the GCS score, the higher the mortality rate (Fig. 3). Another important factor affecting ICH mortality was hematoma volume^[27], but this index was not included in this study, because only 405 patients (35.43% of the total number) had a record of this index, which was seriously missing. It has been reported that the larger the size of hematoma, the higher the mortality rate of patients with ICH.

This study is also the first time to use the MIMIC-III database as the data source to establish ICH patient's mortality prediction model. The dataset spanned more than a decade and details of patient care. The integrity and normality of the data were assured, and it was the basis for our models to incorporate temporal information. Therefore, there was no problem in our data that the population age was relatively small [13] or the number of participants was small^[23,26,28]. But the MIMIC-III data was based on a single medical center, which meant the generality of our results deserves consideration.

Machine learning has great potential in the field of health and even in the field of critical care^[29]. Specifically, the ICH mortality prediction model constructed by using machine learning methods (such as RF, GBDT, etc.) is far better than the traditional scoring system ICH score^[30]. Since ICH score involved hematoma size, ICH score was not compared with other machine learning models in this paper. However, from the existing research results, there is no doubt that the machine learning model is better than the ICH score.

5. Conclusions

The model used MIMIC-III database to extract variables with temporal information, which greatly improved the prediction effect of the model. Compared with the other four machine learning models, the GBDT model had a better recognition rate, regardless of whether multiple variables were used or fewer variables were obtained after screening.

Authors' contributions

Jialin Liu, Ke Li, and Siru Liu conceived the study. Dongfeng, Liu, Jialin Liu, Ke Li, Siru Liu, and Qinwen Shi performed the analysis, interpreted the results and drafted the manuscript. All authors revised the manuscript. All authors read and approved the final manuscript.

Statement on conflicts of interest

The authors have no conflicts of interest to declare.

Data accessibility

The MIMIC-III database (<https://mimic.physionet.org/>) is derived from the de-private medical record of the ICU of the Boston Medical Center in the United States. All patient data were anonymized prior to extraction and data analysis. The creation, maintenance, and use of the MIMIC-III database was approved by the institutional review boards of the Massachusetts Institute of Technology and Beth Israel Deaconess Medical Center. The database contains detailed information on patient vital signs, laboratory tests and disease diagnosis codes for research by global scholars.

Acknowledgements

This project was funded by Sichuan Science and Technology Program Grant No. 18PTDJ0117 & 18ZDYF3402.

Reference

1. J.M. Mackenzie. Intracerebral hemorrhage. *Lancet* . 1992; 373(9675): 1632-1644.
2. M.J. Ariesen, S. P. Claus, G.J.E. Rinkel, et al.. Risk Factors for Intracerebral Hemorrhage in the General Population: A Systematic Review. *Stroke* . 2003; 34(8): 2060-2065.
3. I.C. Hostettler, D.J. Seiffge, D.J. Werring. Intracerebral hemorrhage: an update on diagnosis and treatment. *Expert Review of Neurotherapeutics* . 2019; 19(7): 679-694..
4. W.M.T. Jolink, C.J.M. Klijn, P.J.A.M. Brouwers, et al.. Time trends in incidence, case fatality, and mortality of intracerebral hemorrhage. *Neurology* . 2015; 85(15): 1318-24.
5. A.M. Thabet, M. Kottapally, J.C. Hemphill. Management of intracerebral hemorrhage. *Critical Care Neurology Part I* . 2017; 140(3): 177-193.
6. Z.T. Brodrick, R. Freeze-Ramsey, R.A. Seupaul. Among Patients With Intracerebral Hemorrhage, Is Intensive Blood Pressure Decreasing Associated With Improved Outcome? *Annals of Emergency Medicine* . 2018; 72(5): 611-612.
7. N.I.W. Participants. Priorities for clinical research in intracerebral hemorrhage: report from a National Institute of Neurological Disorders and Stroke workshop. *Stroke; a journal of cerebral circulation* . 2005; 36(3): 23-41.
8. J. Emelia, S. Salim, W. Clifton, et al.. Heart Disease and Stroke Statistics—2018 Update: A Report From the American Heart Association. *Circulation* . 2018; 137: e67–e492.
9. A.R. Parry-Jones, L. Paley, B.D. Bray, et al.. Care-limiting decisions in acute stroke and association with survival: analyses of UK national quality register data. *International Journal of Stroke* . 2016; 11(3): 321-331.
10. L.B. Morgenstern, J.C. Hemphill Rd, C. Anderson, et al.. Guidelines for the management of spontaneous intracerebral hemorrhage: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke; a journal of cerebral circulation* . 2015; 46(7): 2032.
11. S. Thorsten, A.S. Rustam, B. Ronnie, et al.. European Stroke Organisation (ESO) guidelines for the management of spontaneous intracerebral hemorrhage. *International Journal of Stroke* . 2014; 9(7): 840-855.
12. G.S. Power, D.A. Harrison. Why try to predict ICU outcomes? *Current Opinion in Critical Care* . 2014; 20(5): 544-549.
13. S.Y. Peng, Y.C. Chuang, T.W. Kang, K.H. Tseng. Random forest can predict 30-day mortality of spontaneous intracerebral hemorrhage with remarkable discrimination. *European Journal of Neurology* . 2010; 17(7): 945-950.
14. A.E.W. Johnson, T.J. Pollard, L. Shen, et al.. MIMIC-III, a freely accessible critical care database. *Scientific Data* . 2016; 3: 160035.
15. E. Mendes. Introduction to Bayesian Networks. *Springer Berlin Heidelberg* . 2014: 61-71.
16. S.B. Kim, K.S. Han, H.C. Rim, et al.. Some Effective Techniques for Naive Bayes Text Classification. *IEEE Transactions on Knowledge and Data Engineering* . 2006; 18(11):1457-1466.
17. G. Guo, H. Wang, D. A. Bell, et al. KNN Model-Based Approach in Classification. *Springer-Verlag Berlin Heidelberg* . 2003: 986-996.
18. Friedman J H . Greedy Function Approximation: A Gradient Boosting Machine[J]. *The Annals of Statistics* . 2001; 29(5):1189-1232.
19. J.R. Quinlan. Induction of decision trees. *Machine Learning* . 1986; 1: 81-106.
20. L. Rokach, O. Maimon. Top-Down Induction of Decision Trees Classifiers-A Survey. *IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS PART C APPLICATIONS AND REVIEWS* . 2005; 35(4): 476-487.
21. L. Breiman. Random Forests. *Machine Learning* . 2001; 45(1): 5-32.
22. R. Tibshirani. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society Series B (Methodological)* . 1996; 58(1): 267-288.

23. D.F. Edwards, H. Hollingsworth, A.R. Zazulia, et al.. Artificial neural networks improve the prediction of mortality in intracerebral hemorrhage. *Neurology* . 1999; 53(2): 351.
24. S. Lukic, Zarko Cojbasic, Z. Peric, et al.. Artificial neural networks based early clinical prediction of mortality after spontaneous intracerebral hemorrhage. *Acta Neurologica Belgica* . 2012; 112(4): 375-382.
25. G. Celik, O.K. Baykan, Y. Kara, et al. Predicting 10-day Mortality in Patients with Strokes Using Neural Networks and Multivariate Statistical Methods. *Journal of Stroke and Cerebrovascular Diseases* . 2014; 23(6): 1506-1512.
26. D.M.S. Boon, H.H.D.M.V. Vliet, R. Zietse, et al.. To Explore Intracerebral Hematoma with a Hybrid Approach and Combination of Discriminative Factors. *Methods of Information in Medicine* . 2016; 55(05): 450-454.
27. R.A. Hanel, A.R. Xavier, Y. Mohammad, et al.. Outcome following intracerebral hemorrhage and subarachnoid hemorrhage. *Neurological Research* . 2002; 24(Supplement-1): 58-62.
28. C. Weimar, M. Roth, V. Willig, et al.. Development and validation of a prognostic model to predict recovery following intracerebral hemorrhage. *Journal of Neurology* . 2006; 253(6): 788-793.
29. A. Awad, M. Bader-El-Den, J. McNicholas, et al.. Early Hospital Mortality Prediction of Intensive Care Unit Patients Using an Ensemble Learning Approach. *International Journal of Medical Informatics* . 2017; 108: 185-195.
30. J. C. Hemphill, D.C. Bonovich, L. Besmertis, et al.. The ICH Score. *Stroke* . 2001; 32(4): 891.

Table:

Table 1. The clinical characteristics of ICH patients

Variables	Survival (n=760)	Death (n=383)	p-value
Age	67.4±15.81	72.12±13.93	<0.01
Male	417 (54.89%)	196 (51.17%)	0.237
Ethnicity			
White	580 (76.32%)	295 (77.02%)	0.790
Black	55 (7.24%)	21 (5.48%)	0.261
Asian	29 (3.42%)	12 (3.13%)	0.558
Other	30 (3.95%)	13 (3.39%)	0.643
Unknown	66 (8.68%)	42 (10.97%)	0.213
Marital Status			
Married	383 (50.39%)	187 (48.83%)	0.616
Single	163 (21.45%)	54 (14.10%)	<0.01
Widowed	121 (15.92%)	63 (16.45%)	0.819
Divorced	36 (4.74%)	16 (4.18%)	0.668
Separated	4 (0.53%)	7 (1.83%)	0.033
Unknown	53 (6.97%)	56 (14.62%)	<0.01
Complication			
Hypertension	562 (56.31%)	286 (53.76%)	0.339
Diabetes	156 (15.63%)	90 (16.92%)	0.514
Ischemic Heart Disease	121 (12.12%)	75 (14.10%)	0.271
Heart failure	69 (6.91%)	48 (9.02%)	0.139
Pneumonia	90 (9.02%)	33 (6.20%)	0.054
Location of ICH	Location of ICH		
Basal ganglia	193 (25.39%)	81 (21.15%)	0.112
Uncharted	108 (14.21%)	58 (15.14%)	
Thalamus	67 (8.82%)	23 (6.01%)	
Caudate	6 (0.79%)	0 (0%)	
Putamen	12 (1.58%)	0 (0%)	
Lobe	345 (45.39%)	159 (41.51%)	0.212

Variables	Survival (n=760)	Death (n=383)	p-value
<i>Infratentorial</i>	78 (10.26%)	25 (6.53%)	0.037
<i>Unknown</i>	144 (18.95%)	118 (30.81)	<0.01

Table 2. Performance comparison of the five models with all variables for mortality prediction of ICH

	Accuracy	Precision	Recall	F1 score	AUROC	p-value
RF	0.87	0.84	0.75	0.79	0.92	0.276 vs. GBDT 0.112 vs. Naïve Bayes <0.05 vs. Decision Tree 0.082 vs. KNN
GBDT	0.87	0.84	0.76	0.79	0.93	<0.01 vs. Naïve Bayes <0.05 vs. Decision Tree <0.05 vs KNN
Naïve Bayes	0.82	0.68	0.87	0.77	0.9	0.116 vs. Decision Tree 0.356 vs. KNN
Decision Tree	0.85	0.8	0.72	0.76	0.89	0.716 vs. KNN
KNN	0.84	0.82	0.69	0.75	0.89	

Table 3. Performance comparison of the 5 models with 18 variables for mortality prediction of ICH

	Accuracy	Precision	Recall	F1 score	AUROC	p-value
RF	0.87	0.84	0.75	0.79	0.92	0.276 vs. GBDT 0.112 vs. Naïve Bayes <0.05 vs. Decision Tree 0.082 vs. KNN
GBDT	0.87	0.84	0.76	0.79	0.93	<0.01 vs. Naïve Bayes <0.05 vs. Decision Tree <0.05 vs KNN
Naïve Bayes	0.82	0.68	0.87	0.77	0.9	0.116 vs. Decision Tree 0.356 vs. KNN
Decision Tree	0.85	0.8	0.72	0.76	0.89	0.716 vs. KNN
KNN	0.84	0.82	0.69	0.75	0.89	

Appendix:

Table A.1. Parameter setting of 5 ICH mortality prediction models

Models	Parameters	Explanation	Values
RF	criterion	measurement methods of impurities	gini
	n_estimators	number of trees in the forest	110
	max_depth	the maximum depth of the tree	5
	max_features	maximum number of retained features when branching	10
GBDT	n_estimators	number of trees in the forest	45
	max_depth	the maximum depth of the tree	2
	min_samples_split	the number of training samples at least included in the sub-node	2
	learning_rate	Learning rate	0.1
Decision Tree	criterion	measurement methods of impurities	entropy
	random_state	parameters of random patterns in branches	13
	splitter	way of the tree branches	random
	max_depth	the maximum depth of the tree	5
KNN	n_neighbors	number of nearest neighbor samples with voting rights	5
	weights	voting proportion method	uniform
Naïve Bayes	alpha	Laplacian smoothing coefficient	20

Table A.2. Variables with temporal information

Variables	Variables	Survival (n=760)	Death (n=383)	p-value
<i>SBP (1_Max)</i>	166.64±22.95	166.64±22.95	171.38±27.76	<0.01
<i>SBP (1_Min)</i>	100.59±26.54	100.59±26.54	88.31±35.66	<0.01
<i>SBP (2_Max)</i>	161.77±21.03	161.77±21.03	162.28±27.28	0.727
<i>SBP (2_Min)</i>	108.21±21.68	108.21±21.68	91.82±37.84	<0.01
<i>SBP (3_Max)</i>	161.46±21.49	161.46±21.49	160.78±30.24	0.662
<i>SBP (3_Min)</i>	110.5±23.03	110.5±23.03	87.04±41.39	<0.01
<i>DBP (1_Max)</i>	91.61±18.68	91.61±18.68	88.8±20.9	0.021
<i>DBP (1_Min)</i>	46.68±14.52	46.68±14.52	41.24±16.95	<0.01
<i>DBP (2_Max)</i>	89.92±31.01	89.92±31.01	82.01±20.58	<0.01
<i>DBP (2_Min)</i>	50.49±12.47	50.49±12.47	42.98±17.57	<0.01
<i>DBP (3_Max)</i>	88.68±21.93	88.68±21.93	81.59±22.91	<0.01
<i>DBP (3_Min)</i>	51.52±12.77	51.52±12.77	41.16±19.12	<0.01
<i>MBP (1_Max)</i>	115.69±32.3	115.69±32.3	115.67±32.82	0.990
<i>MBP (1_Min)</i>	63±19.11	63±19.11	58.39±20.17	<0.01
<i>MBP (2_Max)</i>	111.14±26.18	111.14±26.18	107.25±30.97	0.026
<i>MBP (2_Min)</i>	70.21±13.1	70.21±13.1	60.22±21.4	<0.01
<i>MBP (3_Max)</i>	110.51±29.44	110.51±29.44	105.58±29.68	<0.01
<i>MBP (3_Min)</i>	70.28±15.07	70.28±15.07	58.18±22.93	<0.01
<i>HR (1_Max)</i>	95.26±17.29	95.26±17.29	104.28±23.29	<0.01
<i>HR (1_Min)</i>	65.14±12.08	65.14±12.08	59.86±24.81	<0.01
<i>HR (2_Max)</i>	95.15±17.87	95.15±17.87	105.36±23.99	<0.01
<i>HR (2_Min)</i>	67.15±11.9	67.15±11.9	57.23±31.05	<0.01
<i>HR (3_Max)</i>	94.4±18.18	94.4±18.18	104.09±24.41	<0.01
<i>HR (3_Min)</i>	67.96±12.21	67.96±12.21	53.08±33.71	<0.01
<i>T (1_Max)</i>	37.47±0.66	37.47±0.66	37.8±1.04	<0.01
<i>T (1_Min)</i>	36.18±2.02	36.18±2.02	36.28±0.97	0.365

Variables	Variables	Survival (n=760)	Death (n=383)	p-value
<i>T (2_Max)</i>	37.43±0.69	37.43±0.69	37.89±3.3	<0.01
<i>T (2_Min)</i>	36.42±1.45	36.42±1.45	36.34±2.7	0.531
<i>T (3_Max)</i>	37.43±0.7	37.43±0.7	37.9±3.31	<0.01
<i>T (3_Min)</i>	36.39±2.38	36.39±2.38	36.43±2.72	0.816
<i>GCS (1_Max)</i>	12.96±3.04	12.96±3.04	7.57±3.63	<0.01
<i>GCS (1_Min)</i>	10±4.16	10±4.16	5.11±2.93	<0.01
<i>GCS (2_Max)</i>	13.21±2.75	13.21±2.75	7.27±3.62	<0.01
<i>GCS (2_Min)</i>	11.09±3.62	11.09±3.62	5.31±2.89	<0.01
<i>GCS (3_Max)</i>	13.25±2.63	13.25±2.63	6.96±3.52	<0.01
<i>GCS (3_Min)</i>	11.41±3.61	11.41±3.61	5.27±2.88	<0.01
<i>WBC (1_Max)</i>	10.93±6.95	10.93±6.95	14.6±13.15	<0.01
<i>WBC (1_Min)</i>	10.1±5.57	10.1±5.57	12.73±6.22	<0.01
<i>WBC (2_Max)</i>	11.16±5.65	11.16±5.65	14.51±8.57	<0.01
<i>WBC (2_Min)</i>	10.82±5.53	10.82±5.53	13.63±6.65	<0.01
<i>WBC (3_Max)</i>	10.96±6.31	10.96±6.31	14.27±8.34	<0.01
<i>WBC (3_Min)</i>	10.74±6.27	10.74±6.27	13.4±6.24	<0.01
<i>RBC (1_Max)</i>	4.38±0.64	4.38±0.64	4.16±0.72	<0.01
<i>RBC (1_Min)</i>	3.98±0.64	3.98±0.64	3.75±0.72	<0.01
<i>RBC (2_Max)</i>	4±0.63	4±0.63	3.88±0.71	<0.01
<i>RBC (2_Min)</i>	3.94±0.64	3.94±0.64	3.72±0.69	<0.01
<i>RBC (3_Max)</i>	3.97±0.65	3.97±0.65	3.84±0.7	<0.01
<i>RBC (3_Min)</i>	3.93±0.67	3.93±0.67	3.7±0.68	<0.01
<i>HB (1_Max)</i>	13.22±1.82	13.22±1.82	12.77±2.04	<0.01
<i>HB (1_Min)</i>	12.02±1.85	12.02±1.85	11.5±2.11	<0.01
<i>HB (2_Max)</i>	12.08±1.83	12.08±1.83	11.89±2.02	0.111
<i>HB (2_Min)</i>	11.9±1.85	11.9±1.85	11.42±2.01	<0.01
<i>HB (3_Max)</i>	12±1.88	12±1.88	11.81±2.05	0.120
<i>HB (3_Min)</i>	11.86±1.93	11.86±1.93	11.37±2.02	0.000
<i>PLT (1_Max)</i>	251.76±92.43	251.76±92.43	240.43±99.57	0.057
<i>PLT (1_Min)</i>	221.43±81.01	221.43±81.01	205.7±90.79	<0.01
<i>PLT (2_Max)</i>	224.2±81.97	224.2±81.97	215.09±91.3	0.088
<i>PLT (2_Min)</i>	218.75±80.25	218.75±80.25	201.17±86.28	<0.01
<i>PLT (3_Max)</i>	225.06±83.76	225.06±83.76	211.35±90.07	0.011
<i>PLT (3_Min)</i>	220.57±83.52	220.57±83.52	197.82±84.24	<0.01
<i>HCT (1_Max)</i>	38.69±5.12	38.69±5.12	37.63±5.71	<0.01
<i>HCT (1_Min)</i>	35.02±5.25	35.02±5.25	33.7±6.14	<0.01
<i>HCT (2_Max)</i>	35.46±5.14	35.46±5.14	35.17±5.77	0.397
<i>HCT (2_Min)</i>	34.86±5.32	34.86±5.32	33.6±5.83	<0.01
<i>HCT (3_Max)</i>	35.28±5.28	35.28±5.28	34.89±5.79	0.255
<i>HCT (3_Min)</i>	34.78±5.53	34.78±5.53	33.43±5.84	<0.01
<i>GLU (1_Max)</i>	160.1±74.32	160.1±74.32	191.78±65.09	<0.01
<i>GLU (1_Min)</i>	124.05±37.63	124.05±37.63	144.23±49.8	<0.01
<i>GLU (2_Max)</i>	133.85±41.99	133.85±41.99	172.69±68.97	<0.01
<i>GLU (2_Min)</i>	124.77±35.83	124.77±35.83	146.29±49.94	<0.01
<i>GLU (3_Max)</i>	131.93±45.2	131.93±45.2	173.05±63.65	<0.01
<i>GLU (3_Min)</i>	124.75±38.07	124.75±38.07	148.61±52.47	<0.01
<i>Na (1_Max)</i>	140.48±3.73	140.48±3.73	142.3±6.21	<0.01
<i>Na (1_Min)</i>	137.95±4.03	137.95±4.03	138.28±4.94	0.215
<i>Na (2_Max)</i>	140.16±3.9	140.16±3.9	142.83±6.37	<0.01
<i>Na (2_Min)</i>	139.22±4.17	139.22±4.17	140.59±5.55	<0.01

Variables	Variables	Survival (n=760)	Death (n=383)	p-value
<i>Na (3_Max)</i>	140.04±4.16	140.04±4.16	142.88±6.49	<0.01
<i>Na (3_Min)</i>	139.37±4	139.37±4	140.81±5.8	<0.01
<i>K (1_Max)</i>	4.26±0.81	4.26±0.81	4.34±0.85	0.145
<i>K (1_Min)</i>	3.71±0.47	3.71±0.47	3.68±0.61	0.422
<i>K (2_Max)</i>	3.96±0.52	3.96±0.52	4.1±0.79	<0.01
<i>K (2_Min)</i>	3.79±0.43	3.79±0.43	3.71±0.56	<0.01
<i>K (3_Max)</i>	3.94±0.53	3.94±0.53	4.07±0.83	<0.01
<i>K (3_Min)</i>	3.8±0.41	3.8±0.41	3.72±0.56	<0.01
<i>CR (1_Max)</i>	1.16±1.74	1.16±1.74	1.38±1.3	0.027
<i>CR (1_Min)</i>	1±1.09	1±1.09	1.18±1.11	<0.01
<i>CR (2_Max)</i>	1.04±1.2	1.04±1.2	1.33±1.26	<0.01
<i>CR (2_Min)</i>	0.99±0.99	0.99±0.99	1.21±1.09	<0.01
<i>CR (3_Max)</i>	1.01±1.05	1.01±1.05	1.33±1.26	<0.01
<i>CR (3_Min)</i>	0.98±1.02	0.98±1.02	1.24±1.18	<0.01
<i>BUN (1_Max)</i>	21.09±14.63	21.09±14.63	25.5±18.57	<0.01
<i>BUN (1_Min)</i>	17.54±11.86	17.54±11.86	21.68±16.67	<0.01
<i>BUN (2_Max)</i>	18.51±12.65	18.51±12.65	24.07±16.68	<0.01
<i>BUN (2_Min)</i>	17.56±11.52	17.56±11.52	22.2±14.96	<0.01
<i>BUN (3_Max)</i>	18.94±11.66	18.94±11.66	24.87±16.56	<0.01
<i>BUN (3_Min)</i>	18.44±11.43	18.44±11.43	23.36±15.76	<0.01

SBP : Systolic Blood Pressure. **DBP** : Diastolic Blood Pressure. **MBP** : Mean Blood Pressure. **HR** : Heart Rate. **T** : Temperature. **GCS** : GCS score. **WBC** : White Blood Cell. **RBC** : Red Blood Cell. **HB** : Hemoglobin. **PLT** : Platelet. **HCT** : Haematocrit. **GLU** : Glucose. **Na** : Serum sodium. **K** : Serum potassium. **CR** : Creatinine. **BUN** : Blood Urea Nitrogen.

***_1(or 2, or 3) _max (or min)** : the maximum (or minimum) value of * on the first (or second, or third) day of admission

Table A.3. RF model and LASSO selected 39 most important variables

RF model selected variables:	LASSO selected variables:
GCS_2_max	GCS_3_max
GCS_3_max	HR_3_min
GCS_1_max	GCS_3_min
GCS_3_min	Pne
GCS_2_min	BUN_3_max
GCS_1_min	AGE
HR_3_min	WBC_1_min
GLU_3_max	GLU_3_max
HR_2_min	RBC_1_max
WBC_1_max	HR_1_max
SBP_3_max	DBP_3_min
SBP_2_min	GCS_2_min
WBC_3_max	HB_3_max
GLU_2_max	GCS_1_max
DBP_3_min	SBP_1_min
MBP_3_min	x3_no
MBP_2_min	Na_2_min
WBC_2_max	x3_putamen

RF model selected variables:	LASSO selected variables:
WBC_1_min	MBP_2_max
T_1_max	MBP_2_min
DBP_3_max	GLU_1_max
T_2_min	WBC_2_min
AGE	x2_ASIAN
PLT_3_min	PLT_1_min
WBC_3_min	HR_2_max
GLU_1_max	T_2_min
T_2_max	IHD
BUN_2_min	x1_SINGLE
HR_1_min	MBP_1_max
GLU_3_min	x1_SEPARATED
HB_3_max	K_3_min
BUN_2_max	WBC_2_max
T_1_min	HF
PLT_2_max	SBP_3_max
PLT_1_min	GCS_2_max
BUN_3_max	x3_caудate
DBP_2_min	K_1_max
PLT_3_max	CR_3_max
RBC_1_min	Dia

GCS : GCS score. **HR** : Heart Rate. **GLU** : Glucose. **WBC** : White Blood Cell. **SBP** : Systolic Blood Pressure. **DBP** : Diastolic Blood Pressure. **MBP** : Mean Blood Pressure. **T** : Temperature. **PLT** : Platelet. **BUN** : Blood Urea Nitrogen. **HB** : Hemoglobin. **RBC** : Red Blood Cell. **Pne** : Patient with pneumonia. **x3_no** : The location of the patient's cerebral hemorrhage is unknown. **Na** : Serum sodium. **x3_putamen** : The location of the patient's cerebral hemorrhage is putamen. **x2_ASIAN** : The patient's ethnicity is Asian. **IHD** : Patient with ischemic heart disease. **x1_SINGLE** : Marital status of the patient is single. **x1_SEPARATED** : Marital status of the patient is separated. **HF** : Patient with ischemic heart failure. **x3_caудate** : The location of the patient's cerebral hemorrhage is caudate. **K** : Serum potassium. **CR** : Creatinine. **Dia** : Patient with diabetes.

***_1(or 2, or 3) _max (or min)** : the maximum (or minimum) value of * on the first (or second, or third) day of admission





