<u>Original Paper</u>

Early Triage of Critically Ill Adult Patients With Mushroom Poisoning: Machine Learning Approach

Yuxuan Liu[1*], MSc; Xiaoguang Lyu[2*], MD; Bo Yang[3*], MSc; Zhixiang Fang[1], PhD; Dejun Hu[3], BSc; Lei Shi[4], BSc; Bisheng Wu[5], BSc; Yong Tian[6], BSc; Enli Zhang[7], BSc; YuanChao Yang[8], BSc

[1]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China

[2]Department of Gastroenterology, Renmin Hospital of Wuhan University, Wuhan, China

[3]Department of Internal Medicine, Renmin Hospital of Xianfeng, Enshi, China

[4]Department of Nephrology, Minda Hospital of Hubei Minzu University, Enshi, China

[5]Department of General Surgery, Renmin Hospital of Xianfeng, Enshi, China

[6]Department of Internal Medicine, Renmin Hospital of Laifeng, Enshi, China

[7]Department of General Surgery, Central Hospital of Hefeng, Enshi, China

[8]Department of Gastroenterology, Renmin Hospital of Xuanen, Enshi, China

[*]these authors contributed equally

Corresponding Author:
Zhixiang Fang, PhD
State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing
Wuhan University
No129, Luoyu Road
Hongshan District
Wuhan, 430072
China
Phone: 86 1 592 703 5199
Email: zxfang@whu.edu.cn

Abstract

Background: Early triage of patients with mushroom poisoning is essential for administering precise treatment and reducing mortality. To our knowledge, there has been no established method to triage patients with mushroom poisoning based on clinical data.

Objective: The purpose of this work was to construct a triage system to identify patients with mushroom poisoning based on clinical indicators using several machine learning approaches and to assess the prediction accuracy of these strategies.

Methods: In all, 567 patients were collected from 5 primary care hospitals and facilities in Enshi, Hubei Province, China, and divided into 2 groups; 322 patients from 2 hospitals were used as the training cohort, and 245 patients from 3 hospitals were used as the test cohort. Four machine learning algorithms were used to construct the triage model for patients with mushroom poisoning. Performance was assessed using the area under the receiver operating characteristic curve (AUC), decision curve, sensitivity, specificity, and other representative statistics. Feature contributions were evaluated using Shapley additive explanations.

Results: Among several machine learning algorithms, extreme gradient boosting (XGBoost) showed the best discriminative ability in 5-fold cross-validation (AUC=0.83, 95% CI 0.77-0.90) and the test set (AUC=0.90, 95% CI 0.83-0.96). In the test set, the XGBoost model had a sensitivity of 0.93 (95% CI 0.81-0.99) and a specificity of 0.79 (95% CI 0.73-0.85), whereas the physicians' assessment had a sensitivity of 0.86 (95% CI 0.72-0.95) and a specificity of 0.66 (95% CI 0.59-0.73).

Conclusions: The 14-factor XGBoost model for the early triage of mushroom poisoning can rapidly and accurately identify critically ill patients and will possibly serve as an important basis for the selection of treatment options and referral of patients, potentially reducing patient mortality and improving clinical outcomes.

XSL•FO

RenderX

Introduction

Approximately 5-10 per 100,000 people die annually from accidental wild mushroom poisoning worldwide, mainly in European countries, the United States, Japan, China, and Iran [1,2]. The American Association of Poison Control Centers reported 86,462 (10,808 cases/year) cases of mushroom exposure from 2012 to 2019. In Japan, there were 1920 cases of mushroom poisoning from 2001-2010, with a morbidity and mortality rate of 0.52% [3]. China's foodborne disease outbreak surveillance system recorded 10,036 outbreaks of mushroom poisoning between 2010 and 2020, resulting in 38,676 illnesses, 21,967 hospitalizations, and 779 deaths [4]. According to the National Health and Family Planning Commission and the Chinese Center for Disease Control and Prevention, mushroom poisoning is the leading cause of death from food poisoning in China [5,6]. Rapid and effective triage is essential for the early treatment of patients with mushroom poisoning and the effective allocation of hospital resources.

The HOPE6 and TALK scores [7] are recommended by the Emergency Physicians Branch of the Chinese Physicians Association, the Chinese Emergency Medical Specialists Consortium, the Emergency Resuscitation and Disaster Medicine Specialty Committee of the Chinese Physicians Association, and the Beijing Emergency Medicine Society, and they serve as an important basis for patient treatment plan determination and are the most widely used condition assessment models in clinical work in China. However, there are some limitations of the abovementioned condition assessment methods: (1) Medical technology is limited in areas with a high incidence of mushroom poisoning, and toxin detection tools and some treatment tools are not available in primary hospitals. (2) Some of the assessment items are poorly clinically operable in practice; for example, patients basically cannot provide the exact information of the poisonous mushroom consumed to determine the type of mushroom. (3) The model may be overly simplistic for complex clinical events such as the development of a critical illness because it assumes risk is a linear mix of numerous factors. (4) The scoring model is proposed based on the literature rather than clinical data, the existing mushroom poisoning literature is mostly summarized based on severe cases, and the patient's condition may be greatly overestimated [8].

In recent years, with the emergence of biomedical big data, machine learning has attracted great attention for developing clinical informatics tools for disease diagnosis, staging, and prognosis [9-11] and has been used in personalized medicine [12-15]; therapeutics [16,17]; surgery [18,19]; radiology [20-25]; and hematology, oncology, and pathology [14,26-29]. It has been demonstrated that machine learning may predict clinical outcomes more accurately than traditional statistical models, particularly when applied to huge data sets [30]. Machine learning algorithms, in contrast to regression-based methods, are capable of capturing higher-order nonlinear interactions among predictors [31]. In this paper, we present the first machine learning–based early condition assessment model for mushroom poisoning assessment, which aims to improve the efficiency and accuracy of condition assessment.

Methods

Our study was divided into 4 steps, and the flowchart of the whole work is shown in Figure 1.

Figure 1. Study flowchart. The phases of data processing, model selection, model comparison, and analysis are illustrated. DCA: decision curve analysis; LR: logistic regression; ML: machine learning; RF: random forest; ROC: receiver operating characteristic; SHAP: Shapley additive explanations; SVM: support vector machine; XGBoost: extreme gradient boosting.
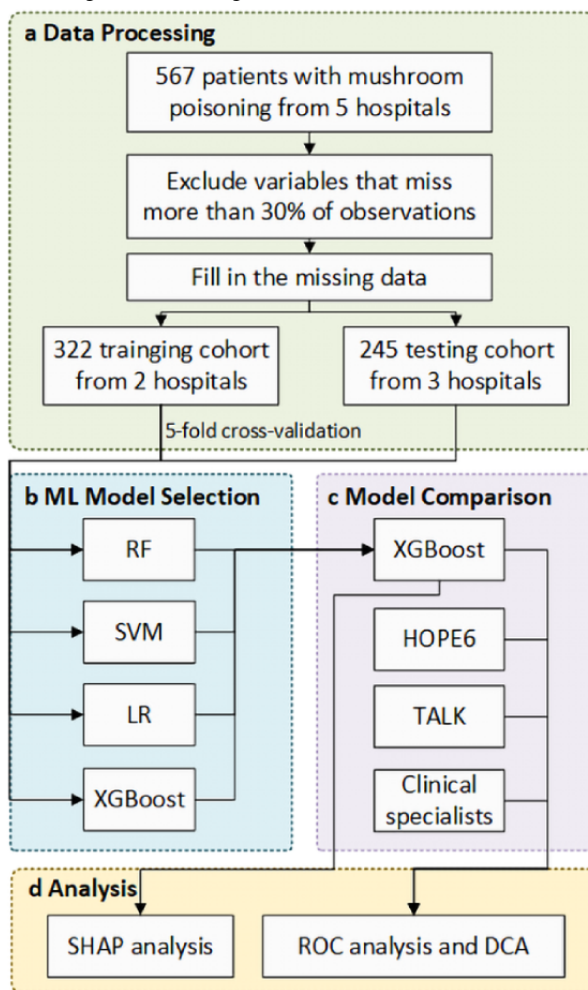
Patient Population

We collected data on patients with wild mushroom poisoning admitted to 5 county hospitals in Enshi, Hubei Province, China, between January 2010 and May 2022. Critical illness was defined as the occurrence of an admission to an intensive care unit, hemodialysis therapy, referral to a higher-level hospital, or death. We collected 24-hour data from 567 patients from 5 hospitals. We used data from 322 patients from 2 of the hospitals as the training set, of which 56 were critically ill and 245 were noncritically ill. Data from 245 patients from 3 other hospitals were used as the test set to assess the performance of our model, of which 43 were critically ill and 202 were noncritically ill.

We used the following inclusion criteria to develop the condition assessment model: (1) patients older than 14 years of age and (2) definite consumption of wild mushrooms before the onset of the disease. Data used in the diagnostic model were excluded when the following conditions were met: (1) age younger than 14 years; (2) possible concurrent consumption of other foods causing acute poisoning; and (3) previous cardiac, hepatic, and renal disorders, as well as mental disorders.

Data Preprocessing

To ensure the reliability of the results and that model use characteristics were readily available, we excluded variables that omitted more than 30% of the observations. The missing values of continuous attributes were filled with the mean value of each attribute, and the missing values of discrete attributes were filled with the mode of each attribute. The values of each feature were normalized for the support vector machine and logistic regression.

Development of Machine Learning Model

Four popular machine learning classification algorithms, including extreme gradient boosting (XGBoost) [32], random forest, support vector machine, and logistic regression, were applied in this study to build the classification models. We implemented machine learning algorithms using Python (version 3.9; Python Software Foundation) and several Python modules (*panda*, *numpy*, *scipy*, *sklearn*, *xgboost*, *shap*, and *matplotlib*). Hyperparameter tuning was performed by a grid search based on 5-fold cross-validation to select the best area under the receiver operating characteristic curve (AUC) value for the classification models.

A proper model interpretation must be supplied for the machine learning model. Model predictions were interpreted using Shapley additive explanations (SHAP) [33,34]. SHAP is a model-independent interpretation technique that helps to interpret the results of prediction models. The interpretation is

based on the SHA*P* value for each feature, which indicates the feature's contribution to the risk of being predicted as critically ill. Having a positive SHA*P* value indicates that the corresponding feature contributes to a higher risk of the patient being critically ill and is a risk factor. On the other hand, having a negative SHA*P* value indicates that the corresponding feature contributes to a lower risk of the patient being critically ill and is a protective factor.

Statistical Analysis

To validate the performance of the model, we compared the best machine learning model with the HOPE6 and TALK scoring models [7]. In addition, we included the results of each patient's primary care physician's assessment of the patient's condition to explore the diagnostic performance of our model compared with that of the physician's judgment.

We evaluated model performance in the test set by calculating (1) the AUC; (2) decision curve analysis (DCA); and (3) sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio. The AUC is often used to assess the performance of various prediction models and is robust to category imbalance [35]. Based on the receiver operating characteristic curve, we chose the best predictive value (i.e., the value that is closest to the perfect model) for cases to fix the category imbalance for whether the condition is critical or not [36]. By calculating the AUC of different models, the discriminatory ability of different models can be compared. However, the AUC only focuses on the overall accuracy of the models and does not focus on the relationship between benefit and risk associated with different cutoff values in different models. DCA, on the other hand, permits the assessment of the range of threshold probabilities for a model to have value, the magnitude of the benefit, and the best model among numerous candidates. DCA figures out the "clinical net benefit" of one or more predictive models over a range of threshold probabilities. A threshold probability is a minimum chance that a disease needs further intervention, and the "clinical net benefit" takes into account the relative harms of false positives and false negatives [37].

Ethics Approval

This study was approved by the Ethical Committee of Renmin Hospital of Xianfeng, and informed consent was waived because this study was retrospective and used deidentified data (XFRY2021-12). The privacy and confidentiality of all individuals included in this study were strictly protected, and their data were used only for the purposes of this research.

Results

Patient Characteristics

The patient cohort participating in this study included data from 567 cases of mushroom poisoning of patients admitted to 5 county hospitals in the Enshi area. The case data included the following types of mushroom poisoning: gastroenteritis, neuropsychiatric symptoms, acute liver damage, acute renal failure, myocardial injury, and combined types [7]. For all data, the length of stay was 4.45 (95% CI, 3.45-5.44) and 3.52 (95% CI 3.21-3.82) days for critically ill and noncritically ill patients, respectively, and the cost of hospitalization was ¥11,113.09 (95% CI ¥8059.92-¥14,166.27; ¥1=US $0.14) and ¥2392.08 (95% CI ¥2188.34-¥2595.82), respectively.

Table 1 shows the baseline data of patients. The majority of the indicators did not differ significantly between the training set and the test set, and the remaining indicators (*P*<.01) were excluded from the machine learning model.

Table 1. Patient baseline characteristics.

Variables	Training set (n=322)	Test set (n=245)	*P* value
Age (years), mean (SD)	50.89 (17.37)	55.62 (15.9)	<.001
Sex, n (%)			.21
Male	155 (48.1)	131 (53.5)	
Female	167 (51.9)	114 (46.5)	
Critical illness, n (%)			.96
Yes	56 (17.4)	43 (17.6)	
No	266 (82.6)	202 (82.4)	
Vital signs, mean (SD)			
Systolic pressure	120.9 (20.1)	122.43 (21.45)	.39
Diastolic pressure	76.07 (11.83)	77.82 (13.09)	.10
Respiratory rate	18.89 (2.4)	19.74 (1.86)	<.001
Temperature	36.5 (0.4)	36.4 (0.43)	<.001
Heart rate	81.39 (13.87)	79.58 (14.29)	.13
Blood routine test, mean (SD)			
WBC[a]	11.98 (39.69)	9.39 (4.18)	.31
RBC[b]	4.63 (0.79)	4.62 (4.18)	.89
Hemoglobin	137.57 (26.6)	137.74 (24.52)	.94
Hematocrit	42.88 (8.74)	37.87 (13.71)	<.001
Heart rate	81.39 (13.87)	79.58 (14.29)	.13
Liver function, mean (SD)			
TBiL[c]	16.89 (11.65)	17.7 (15.98)	.49
DBiL[d]	6.05 (5.73)	6.09 (8.27)	.94
IBiL[e]	10.4 (5.64)	12.23 (12.14)	.02
ALT[f]	106.23 (380.7)	142.23 (689.61)	.43
AST[g]	152.04 (713.41)	166.09 (824.13)	.83
Kidney function, mean (SD)			
Urea	7.03 (3.17)	7.69 (4.5)	.08
Creatinine	91.27 (54.42)	100.15 (94.5)	.16
Coagulation function, mean (SD)			
PT[h]	13.14 (3.52)	12.88 (4.86)	.48
TT[i]	13.82 (2.57)	15.21 (16.76)	.17
PPT[j]	31.65 (6.24)	33.88 (15.91)	.03
Fibrinogen	2.77 (0.76)	3.01 (0.83)	<.001
INR[k]	1.08 (0.34)	1.06 (0.3)	.39
Serum electrolytes, mean (SD)			
Serum potassium	4.94 (17.39)	4.97 (14.5)	.99
Serum sodium	138.24 (11.62)	136.21 (16.4)	.09
Other blood biochemistry and enzymatic parameters, mean (SD)			
CO2CP[l]	22.6 (9.81)	23.59 (13.71)	.15

Variables	Training set (n=322)	Test set (n=245)	*P* value
CK[m]	394 (1981.36)	49.29 (527.65)	.16
CK-MB[n]	38.22 (138.74)	35.89 (118.64)	.84
LDH[o]	298.95 (570.85)	342.67 (997.28)	.52

[a]WBC: white blood cell.

[b]RBC: red blood cell.

[c]TBiL: total bilirubin.

[d]DBiL: direct bilirubin.

[e]IBiL: indirect bilirubin.

[f]ALT: alanine transaminase.

[g]AST: aspartate aminotransferase.

[h]PT: prothrombin time.

[i]TT: thrombin time.

[j]PPT: partial thromboplastin time.

[k]INR: international normalized ratio.

[l]CO2CP: carbon dioxide combining power.

[m]CK: creatine kinase.

[n]CK-MB: creatine kinase-MB.

[o]LDH: lactate dehydrogenase.

Comparison Among the Machine Learning Algorithms

Among the algorithms, XGBoost achieved the highest AUC value, with AUCs of 0.83 (95% CI 0.77-0.90) and 0.90 (95% CI 0.83-0.96) in the internal and external validation sets (Figures 2 and 3). In the DCA (Figure 4), the XGBoost model had a greater net benefit compared to other methods over a wide range of threshold probabilities. As a result, we selected XGBoost as a suitable algorithm for developing the prediction model and conducted additional analyses to determine its predictive validity.

The feature ranking interpretation of the XGBoost model based on the SHAP algorithm (Figure 5) shows that lactate dehydrogenase (LDH), aspartate aminotransferase (AST), international normalized ratio (INR), serum sodium, alanine transaminase (ALT), hemoglobin, white blood cell, urea, total bilirubin (TBiL), creatine kinase-MB (CK-MB), creatinine, heart rate, indirect bilirubin (IBiL), and prothrombin time (PT) were important features of the XGBoost model. LDH and AST were the most influential factors, and their contribution was considerably greater than that of other indicators (Figure 5). Overall, the characteristics of LDH, AST, INR, serum sodium, ALT, hemoglobin, white blood cell, urea, CK-MB, creatinine, heart rate, and PT were positively correlated with the results and were risk factors; meanwhile, TBiL and IBiL were negatively correlated with the results and were protective factors (Figure 6).

Figure 2. ROC curves for each machine learning algorithm in the 5-fold stratified cross-validation. Diagonal dotted lines represent the random classifier. AUC: area under the receiver operating characteristic curve; ROC: receiver operating characteristic; SVM: support vector machine; XGBoost: extreme gradient boosting.
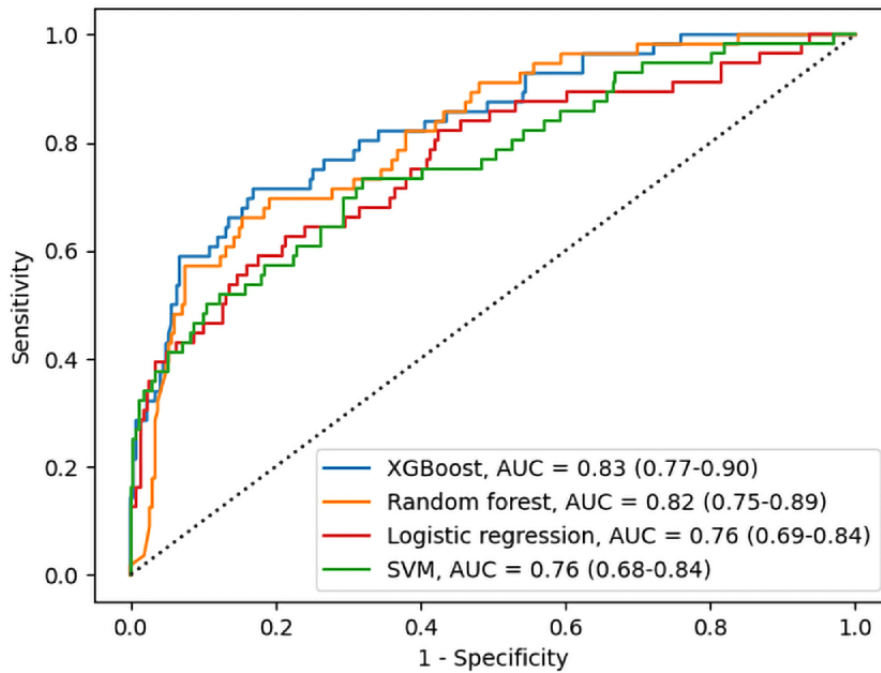

Figure 3. ROC curves for machine learning algorithms, HOPE6, TALK, and physicians' assessment in the test set. AUC: area under the receiver operating characteristic curve; ROC: receiver operating characteristic; SVM: support vector machine; XGBoost: extreme gradient boosting.
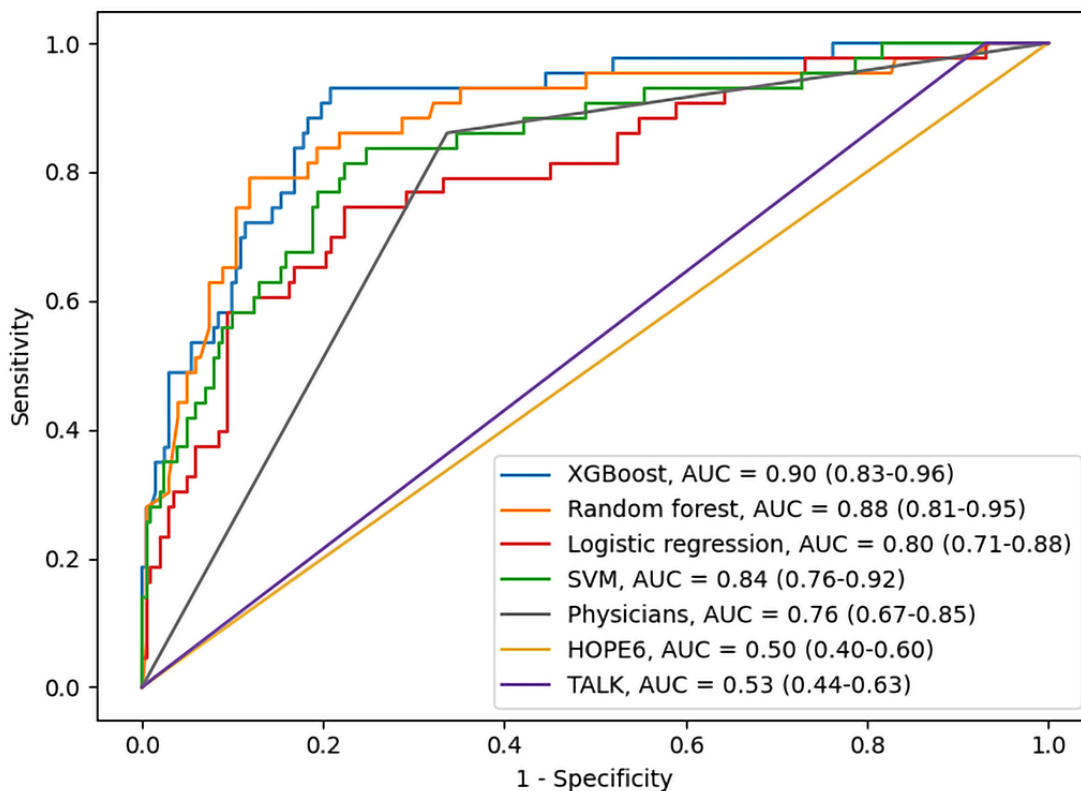
Figure 4. Decision curve analysis on the test set. The x-axis represents the threshold probability of the hospitalization outcome. The y-axis represents the net benefit. The curves (decision curves) represent the threshold probabilities of net benefit outcomes for the 6 models (4 machine learning models, HOPE6 model, and TALK model), physician classification, and the 2 clinical alternatives (classifying no patients as critical vs classifying all patients as critical) within the specified range. SVM: support vector machine; XGBoost: extreme gradient boosting.
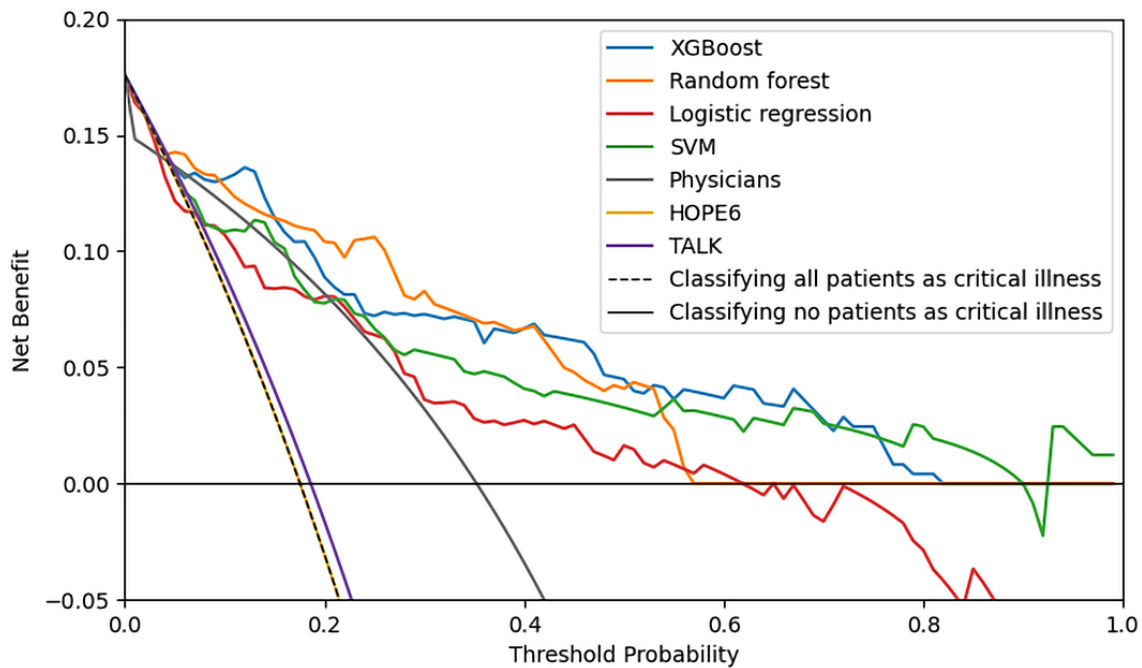
Figure 5. Ranking of feature importance indicated by SHAP. ALT: alanine transaminase; AST: aspartate aminotransferase; CK-MB: creatine kinase-MB; IBiL: indirect bilirubin; INR: international normalized ratio; LDH: lactate dehydrogenase; PT: prothrombin time; SHAP: Shapley additive explanations; TBiL: total bilirubin; WBC: white blood cell.
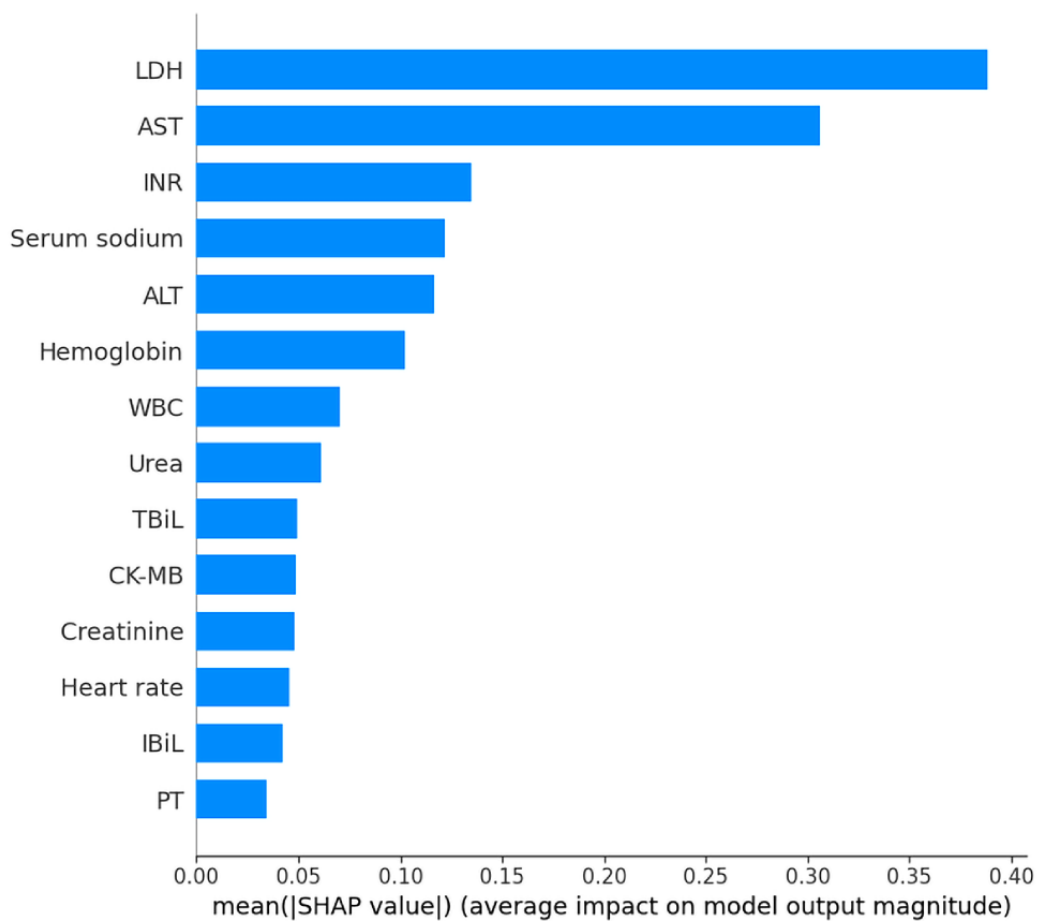
Figure 6. Attributes of the model's features. Each line is a feature, and the horizontal coordinate is the SHAP value, which shows how much that feature affected the outcome. Every point on the graph is a sample. The value of a feature goes up as the color gets redder, and it goes down as the color gets bluer. ALT: alanine transaminase; AST: aspartate aminotransferase; CK-MB: creatine kinase-MB; IBiL: indirect bilirubin; INR: international normalized ratio; LDH: lactate dehydrogenase; PT: prothrombin time; SHAP: Shapley additive explanations; TBiL: total bilirubin; WBC: white blood cell.
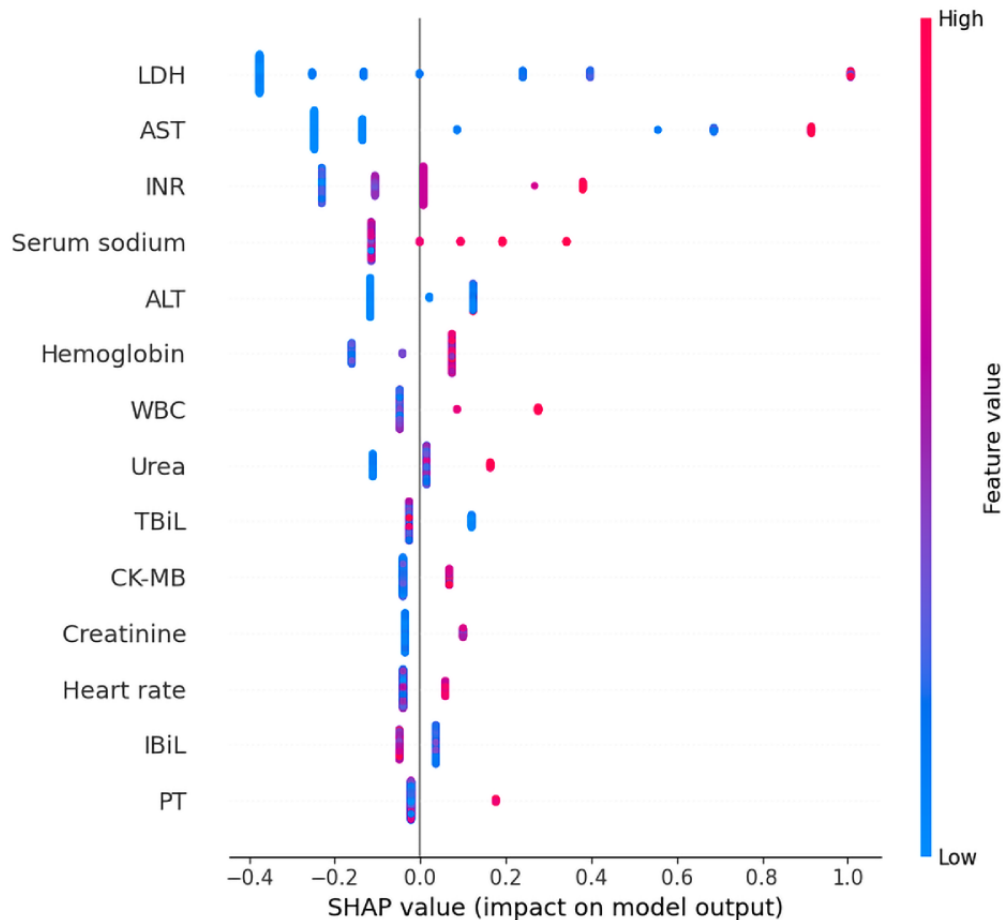
Comparison of the Performances of the XGBoost, HOPE6, TALK, and Physicians' Assessment

In terms of the AUC, all the machine learning models performed significantly better than HOPE6, TALK, and physicians' assessment (Figure 3). Moreover, across a wide range of threshold probabilities (or clinical preferences), all machine learning models performed better than HOPE6, TALK, and physicians' assessment (Figure 4). In terms of AUC and DCA, XGBoost was the top model (Figures 3 and 4).

The sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative

likelihood ratio of each model are shown in Table 2. Almost all patients were classified as critically ill in the HOPE6 and TALK scores (specificity: 0.00, 95% CI 0.00-0.02 for HOPE6 and 0.07, 95% CI 0.04-0.11 for TALK). In the test set, the XGBoost model had a sensitivity of 0.93 (95% CI 0.81-0.99) and a specificity of 0.79 (95% CI 0.73-0.85), whereas the physicians' assessment had a sensitivity of 0.86 (95% CI 0.72-0.95) and a specificity of 0.66 (95% CI 0.59-0.73). The results indicate that the current diagnostic model can more precisely evaluate a patient's condition, assist physicians in formulating more precise treatment plans, lessen the harm caused by incorrect treatment, and reduce patients' treatment expenses.

Table 2. Comparison of the prediction ability of the extreme gradient boost (XGBoost model), HOPE6, TALK, and physicians' assessment on the test set.

	Sensitivity (95% CI)	Specificity (95% CI)	PPV[a] (95% CI)	NPV[b] (95% CI)	PLR[c] (95% CI)	NLR[d] (95% CI)
XGBoost	0.93 (0.81-0.99)	0.79 (0.73-0.85)	0.49 (0.38-0.60)	0.98 (0.95-1.00)	4.47 (3.38-5.93)	0.09 (0.03-0.26)
HOPE6	1.00 (0.92-1.00)	0.00 (0.00-0.02)	0.18 (0.13-0.23)	—[e]	1.00 (1.00-1.00)	—
TALK	1.00 (0.92-1.00)	0.07 (0.04-0.11)	0.19 (0.14-0.24)	1.00 (0.77-1.00)	1.07 (1.03-1.12)	—
Physicians' assessment	0.86 (0.72-0.95)	0.66 (0.59-0.73)	0.35 (0.26-0.45)	0.96 (0.91-0.98)	2.56 (2.04-3.21)	0.21 (0.10-0.44)

[a]PPV: positive predictive value.

[b]NPV: negative predictive value.

[c]PLR: positive likelihood ratio.

[d]NLR: negative likelihood ratio.

[e]Not available.

Discussion

Principal Findings

Mushroom poisoning is a global food safety event. Early triage of patients with mushroom poisoning is essential for the formulation of treatment options and reduction of mortality. The objective of this study was to develop a machine learning–based triage model that could assess whether a patient with mushroom poisoning was critically ill within 24 hours of admission to support clinical decision-making. To our knowledge, this is the first time a machine learning algorithm has been used for the early triage of mushroom poisoning. The research demonstrates that the model developed using the XGBoost algorithm is superior to previous methods for triaging critically ill patients with mushroom poisoning. In addition, we discovered that liver dysfunction had the greatest impact on the model (more than 50%), with LDH and AST being the 2 most influential factors.

Comparison With Prior Work

In this paper, we compared 4 machine learning models, 2 scoring models (HOPE6 and TALK), and clinical experts' assessment results and found that machine learning models outperformed conventional methods. First, machine learning algorithms are entirely data driven, whereas scoring models and physicians' evaluations are based on expert knowledge. Second, machine learning algorithms can learn and infer nonlinear higher-order connections between clinical factors and patient outcomes. Scoring models have the advantage of being simple to calculate and interpret; however, for complex clinical episodes (e.g., progression to critical illness), they may be overly simplistic in assuming that the severity of a patient's condition is a linear combination of multiple factors.

XGBoost, a cutting-edge tree-based gradient boosting method, allowed us to create more accurate predictive models than other machine learning models; consequently, it was selected as our final model. The XGBoost model had greater sensitivity (0.93, 95% CI 0.81-0.99) and specificity (0.79, 95% CI 0.73-0.85) than physicians' assessment (sensitivity: 0.86, 95% CI 0.72-0.95; and specificity: 0.66, 95% CI 0.59-0.73). As expert consensus, the HOPE6 and TALK models have not been validated by clinical data, and because the identification and treatment of critically ill patients have important clinical implications, scoring models based on expert experience will overestimate the condition of every patient. In this study, the HOPE6 and TALK models have extremely high sensitivity (1.00, 95% CI 0.92-1.00) and extremely low specificity (HOPE6: 0.00, 95% CI 0.00-0.02; and TALK: 0.07, 95% CI 0.04-0.11). Consequently, they are able to identify all critical cases and play an important role in clinical practice, allowing critical patients to be treated and the mortality rate to be as low as possible. However, an excessive number of noncritically ill patients will be misidentified as critically ill patients, which may result in the waste of medical resources and physical harm to patients as a result of overtreatment. Therefore, our model has greater advantages for the identification of critically ill patients and may aid physicians in making assessments. Moreover, the model's ability to adjust cutoff values provides greater flexibility and more nuanced insights into the patient's condition, making it a useful tool for physicians. The application software created by our current model is available on GitHub [38].

The study identified 14 clinical variables, of which AST, ALT, LDH, INR, PT, TBiL, IBiL, urea, creatinine, and CK-MB were consistent with current clinical evidence. These factors were classified as hepatic (AST, ALT [39-41], LDH [42], TBiL [39], and IBiL [43]), coagulation (INR [39,41] and PT [43]), renal (urea [43] and creatinine[40,44]), and cardiac impairment (CK-MB [42]). Using the SHAP method, we found that the indicators characterizing liver function impairment contributed 57.5% to the prediction results of the model and had the greatest impact. Interestingly, the study also identified several clinical indicators that lack relevant clinical evidence, such as serum sodium, hemoglobin, and heart rate, highlighting the potential for machine learning to identify novel relationships between clinical factors and patient outcomes.

Limitations

Despite the promising results, the study has several limitations. The machine learning method is data driven, and the model's performance is dependent on the quality and completeness of the data. Additionally, the study cohort was from a single location in China, limiting the model's applicability to other regions. Future studies should collect data from more locations to enhance the model's robustness and generalizability.

Conclusions

Using routinely accessible data at the time of triage, we found that machine learning has good predictive performance in the triage of patients with mushroom poisoning. Compared with other machine learning algorithms, clinical guidelines, and clinician assessment, the XGBoost model has better diagnostic performance and can be used to select core indicators for a triage model. Machine learning may be a powerful prognostic indicator for early warning in critically ill patients, which has a significant impact on the triage of patients, formulation of treatment options, and the allocation of medical resources.

Acknowledgments

Data Availability

The data that support the findings of this study are available from the authors upon reasonable request but cannot be shared with the public due to privacy concerns and restrictions imposed by the institutional review board. The data are available by contacting correspondent authors for collaboration or other reasonable requests with permission from the project directors.

Conflicts of Interest

None declared.

References

1. Chen Z, Zhang P, Zhang Z. Investigation and analysis of 102 mushroom poisoning cases in Southern China from 1994 to 2012. Fungal Diversity 2013 Aug 15;64(1):123-131. [doi: 10.1007/s13225-013-0260-7]
2. Diaz JH. Evolving global epidemiology, syndromic classification, general management, and prevention of unknown mushroom poisonings. Crit Care Med 2005 Feb;33(2):419-426. [doi: 10.1097/01.ccm.0000153530.32162.b7] [Medline: 15699848]
3. Yamaura Y. Recent trends of mushroom poisoning in Japan. Article in Japanese. Chudoku Kenkyu 2013 Mar;26(1):39-43. [Medline: 23600266]
4. Li W, Pires S, Liu Z, Liang J, Wang Y, Chen W, et al. Mushroom poisoning outbreaks - China, 2010-2020. China CDC Wkly 2021 Jun 11;3(24):518-522 [FREE Full text] [doi: 10.46234/ccdcw2021.134] [Medline: 34594925]
5. General Office of the National Health and Family Planning Commission. National food poisoning incidents in 2015. Chinese Journal of Food Hygiene 2016;3:290+391+408 [FREE Full text]
6. Wang X, Ren J, Wang Z, Weng X, Wang R. Epidemiological characteristics of food poisoning events in China, 2017. Disease Surveillance 2018;33(5):359-364 [FREE Full text] [doi: 10.3784/j.issn.1003-9961.2018.05.004]
7. Committee Chinese Society Of Toxicology Poisoning And Treatment Of Specialized, Association Poisoning Group Of Emergency Medicine Branch Of Chinese Medical. Chinese expert consensus on diagnosis and treatment of acute thallium poisoning (2021). Article in Chinese. Zhonghua Wei Zhong Bing Ji Jiu Yi Xue 2021 Apr;33(4):385-392. [doi: 10.3760/cma.j.cn121430-20210305-00338] [Medline: 34053478]
8. Chengye S. Improvement in mushroom poisoning control is in urgent need. Chinese Journal of Emergency Medicine 2016;25(08):981-984. [doi: 10.3760/cma.j.issn.1671-0282.2016.08.001]
9. Xu R, Wei W, Krawczyk M, Wang W, Luo H, Flagg K, et al. Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma. Nat Mater 2017 Nov 09;16(11):1155-1161. [doi: 10.1038/nmat4997] [Medline: 29035356]
10. Zhang JX, Song W, Chen ZH, Wei JH, Liao YJ, Lei J, et al. Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. Lancet Oncol 2013 Dec;14(13):1295-1306. [doi: 10.1016/S1470-2045(13)70491-1] [Medline: 24239208]
11. Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, ACS and GC6-74 cohort study groups. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. Lancet 2016 Jun 04;387(10035):2312-2322 [FREE Full text] [doi: 10.1016/S0140-6736(15)01316-1] [Medline: 27017310]
12. Holzinger A. Trends in interactive knowledge discovery for personalized medicine: cognitive science meets machine learning. The IEEE intelligent informatics bulletin 2014;15(1):6-14 [FREE Full text]
13. Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. Curr Cardiol Rep 2014 Jan 13;16(1):441. [doi: 10.1007/s11886-013-0441-8] [Medline: 24338557]
14. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. Cancer Inform 2007 Feb 11;2:59-77 [FREE Full text] [doi: 10.1177/117693510600200030] [Medline: 19458758]

15. Menden MP, Iorio F, Garnett M, McDermott U, Benes CH, Ballester PJ, et al. Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties. PLoS One 2013 Apr 30;8(4):e61318 [FREE Full text] [doi: 10.1371/journal.pone.0061318] [Medline: 23646105]

16. Kirchmair J, Göller AH, Lang D, Kunze J, Testa B, Wilson ID, et al. Predicting drug metabolism: experiment and/or computation? Nat Rev Drug Discov 2015 Jun 24;14(6):387-404 [FREE Full text] [doi: 10.1038/nrd4581] [Medline: 25907346]

17. Stanfield Z, Coşkun M, Koyutürk M. Drug response prediction as a link prediction problem. Sci Rep 2017 Jan 09;7(1):40321 [FREE Full text] [doi: 10.1038/srep40321] [Medline: 28067293]

18. Kassahun Y, Yu B, Tibebu AT, Stoyanov D, Giannarou S, Metzen JH, et al. Surgical robotics beyond enhanced dexterity instrumentation: a survey of machine learning techniques and their role in intelligent and autonomous surgical actions. Int J Comput Assist Radiol Surg 2016 Apr;11(4):553-568. [doi: 10.1007/s11548-015-1305-z] [Medline: 26450107]

19. Kiranantawat K, Sitpahul N, Taeprasartsit P, Constantinides J, Kruavit A, Srimuninnimit V, et al. The first Smartphone application for microsurgery monitoring: SilpaRamanitor. Plast Reconstr Surg 2014 Jul;134(1):130-139. [doi: 10.1097/PRS.0000000000000276] [Medline: 25028822]

20. García-Lorenzo D, Lecoeur J, Arnold DL, Collins DL, Barillot C. Multiple sclerosis lesion segmentation using an automatic multimodal graph cuts. 2009 Presented at: MICCAI 2009: Medical Image Computing and Computer-Assisted Intervention; September 20-24, 2009; London, UK p. 584-591. [doi: 10.1007/978-3-642-04271-3_71]

21. Cocosco CA, Kedenburg G, Niessen WJ, Thoms H. Method a system and a computer program for segmenting a structure associated with a reference structure in an image. Google Patents. 2008. URL: https://patents.google.com/patent/US20080298682A1/en [accessed 2023-03-07]

22. Kedenburg G, Cocosco CA, Köthe U, Niessen WJ, Vonken EJPA, Viergever MA. Automatic cardiac MRI myocardium segmentation using graphcut. 2006 Mar 10 Presented at: Medical Imaging 2006: Image Processing; March 10, 2006; San Diego, CA. [doi: 10.1117/12.653583]

23. Cheng HD, Cai X, Chen X, Hu L, Lou X. Computer-aided detection and classification of microcalcifications in mammograms: a survey. Pattern Recognition 2003 Dec;36(12):2967-2991. [doi: 10.1016/s0031-3203(03)00192-4]

24. Summers RM, Yao J, Pickhardt PJ, Franaszek M, Bitter I, Brickman D, et al. Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population. Gastroenterology 2005 Dec;129(6):1832-1844 [FREE Full text] [doi: 10.1053/j.gastro.2005.08.054] [Medline: 16344052]

25. Chan HP, Hadjiiski L, Zhou C, Sahiner B. Computer-aided diagnosis of lung cancer and pulmonary embolism in computed tomography-a review. Acad Radiol 2008 May;15(5):535-555 [FREE Full text] [doi: 10.1016/j.acra.2008.01.014] [Medline: 18423310]

26. Yu KH, Zhang C, Berry GJ, Altman RB, Ré C, Rubin DL, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun 2016 Aug 16;7:12474 [FREE Full text] [doi: 10.1038/ncomms12474] [Medline: 27527408]

27. Shipp MA, Ross KN, Tamayo P, Weng AP, Kutok JL, Aguiar RC, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. Nat Med 2002 Jan;8(1):68-74. [doi: 10.1038/nm0102-68] [Medline: 11786909]

28. Gupta S, Tran T, Luo W, Phung D, Kennedy RL, Broad A, et al. Machine-learning prediction of cancer survival: a retrospective study using electronic administrative records and a cancer registry. BMJ Open 2014 Mar 17;4(3):e004007 [FREE Full text] [doi: 10.1136/bmjopen-2013-004007] [Medline: 24643167]

29. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017 Feb 02;542(7639):115-118 [FREE Full text] [doi: 10.1038/nature21056] [Medline: 28117445]

30. Zhou ZH. Ensemble Methods: Foundations and Algorithms. Boca Raton, FL: CRC Press; 2012.

31. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized nutrition by prediction of glycemic responses. Cell 2015 Nov 19;163(5):1079-1094 [FREE Full text] [doi: 10.1016/j.cell.2015.11.001] [Medline: 26590418]

32. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. 2016 Aug 13 Presented at: KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, CA p. 785-794. [doi: 10.1145/2939672.2939785]

33. Lundberg S, Lee SI. A unified approach to interpreting model predictions. arXiv. Preprint posted online on May 22, 2017. [doi: 10.48550/arXiv.1705.07874]

34. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2020 Jan;2(1):56-67 [FREE Full text] [doi: 10.1038/s42256-019-0138-9] [Medline: 32607472]

35. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit 1997 Jul;30(7):1145-1159. [doi: 10.1016/S0031-3203(96)00142-2]

36. Kuhn M, Johnson K. Applied Predictive Modeling. New York, NY: Springer; 2013.

37. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26(6):565-574 [FREE Full text] [doi: 10.1177/0272989X06295361] [Medline: 17099194]

38.   BecomeRich-Coder / Early-triage-of-critically-ill-mushroom-poisoning-patients-using-machine-learning. GitHub. URL:
       https://github.com/BecomeRich-Coder/Early-triage-of-critically-ill-mushroom-poisoning-patients-using-machine-learning
       [accessed 2023-02-23]
39.   Trabulus S, Altiparmak MR. Clinical features and outcome of patients with amatoxin-containing mushroom poisoning.
       Clin Toxicol (Phila) 2011 Apr 12;49(4):303-310. [doi: 10.3109/15563650.2011.565772] [Medline: 21563906]
40.   Trakulsrichai S, Sriapha C, Tongpoo A, Udomsubpayakul U, Wongvisavakorn S, Srisuma S, et al. Clinical characteristics
       and outcome of toxicity from Amanita mushroom poisoning. Int J Gen Med 2017 Nov 3;10:395-400 [FREE Full text] [doi:
       10.2147/IJGM.S141111] [Medline: 29138589]
41.   Altuntaş M, Duran L. Predictors of poor outcome in mushroom poisoning: a retrospective cohort study. J Surg Med 2021
       Oct 1;5(10):1029-1032. [doi: 10.28982/josam.979166]
42.   Wen L, Wu W, Li L, Gan L, Ou S. Analysis on clinical features and risk factors of death in 210 patients with acute mushroom
       poisoning. Article in Chinese. Zhonghua Wei Zhong Bing Ji Jiu Yi Xue 2018 Jan;30(1):72-77. [doi:
       10.3760/cma.j.issn.2095-4352.2018.01.014] [Medline: 29308762]
43.   Kim T, Lee D, Lee JH, Lee Y, Oh BJ, Lim KS, et al. Predictors of poor outcomes in patients with wild mushroom-induced
       acute liver injury. World J Gastroenterol 2017 Feb 21;23(7):1262-1267 [FREE Full text] [doi: 10.3748/wjg.v23.i7.1262]
       [Medline: 28275306]
44.   Trakulsrichai S, Jeeratheepatanont P, Sriapha C, Tongpoo A, Wananukul W. Myotoxic mushroom poisoning in Thailand:
       clinical characteristics and outcomes. Int J Gen Med 2020 Nov 17;13:1139-1146 [FREE Full text] [doi:
       10.2147/IJGM.S271914] [Medline: 33235487]

Abbreviations

ALT: alanine transaminase
AST: aspartate aminotransferase
AUC: area under the receiver operating characteristic curve
CK-MB: creatine kinase-MB
DCA: decision curve analysis
IBiL: indirect bilirubin
INR: international normalized ratio
LDH: lactate dehydrogenase
PT: prothrombin time
SHAP: Shapley additive explanations
TBiL: total bilirubin
XGBoost: extreme gradient boosting

XSL•FO
RenderX