

# A LightGBM and XGBoost Learning Method for Postoperative Critical Illness Key Indicators Analysis

Lei Han<sup>1†</sup>, Yiziting Zhu<sup>2†</sup>, Yuwen Chen<sup>3\*</sup>, Guoqiong Huang<sup>4\*</sup>, and Bin Yi<sup>2\*</sup>

<sup>1</sup> Department of Trauma Medical Center, Daping Hospital, Third Military Medical University, Chongqing 400042, China  
[e-mail: 920774200@qq.com]

<sup>2</sup> Department of Anesthesiology, Southwest Hospital, Third Military Medical University, Chongqing 400038, China  
[e-mail: yibin1974@163.com]

<sup>3</sup> Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, China  
[e-mail: chenyuwen@cigit.a.cn]

<sup>4</sup> Public Administration Center, Third Military Medical University, Chongqing 400038, China  
[e-mail: guoqiongh@sina.com]

\*Corresponding author: Yuwen Chen; Guoqiong Huang; Bin Yi

†These authors contributed equally to this work and share the first authorship

*Received February 8, 2023; revised July 4, 2023; accepted July 18, 2023;  
published August 31, 2023*

---

## Abstract

Accurate prediction of critical illness is significant for ensuring the lives and health of patients. The selection of indicators affects the real-time capability and accuracy of the prediction for critical illness. However, the diversity and complexity of these indicators make it difficult to find potential connections between them and critical illnesses. For the first time, this study proposes an indicator analysis model to extract key indicators from the preoperative and intraoperative clinical indicators and laboratory results of critical illnesses. In this study, preoperative and intraoperative data of heart failure and respiratory failure are used to verify the model. The proposed model processes the datum and extracts key indicators through four parts. To test the effectiveness of the proposed model, the key indicators are used to predict the two critical illnesses. The classifiers used in the prediction are light gradient boosting machine (LightGBM) and eXtreme Gradient Boosting (XGBoost). The predictive performance using key indicators is better than that using all indicators. In the prediction of heart failure, LightGBM and XGBoost have sensitivities of 0.889 and 0.892, and specificities of 0.939 and 0.937, respectively. For respiratory failure, LightGBM and XGBoost have sensitivities of 0.709 and 0.689, and specificity of 0.936 and 0.940, respectively. The proposed model can effectively analyze the correlation between indicators and postoperative critical illness. The analytical results make it possible to find the key indicators for postoperative critical illnesses. This model is meaningful to assist doctors in extracting key indicators in time and improving the reliability and efficiency of prediction.

---

**Keywords:** Critical illness, Light gradient boosting machine (LightGBM), Indicator analysis, Shapley additive explanation (SHAP) value, Machine learning

---

## 1. Introduction

The occurrence of critical illness in perioperative patients will not only increase the medical expenses of patients, but also affect the Rehab Results (FCR) of patients [1, 2], and even lead to the death of patients. The study of Khuri et al. [3] showed that the median survival time of patients would be reduced by 69% if critical adverse events occurred within 30 days after the operation, and the long-term consequences of short-term surgical complications had a profound impact on the life expectancy and quality of life of survivors [4]. Preoperative and intraoperative risk prediction of critical illness, immediate capture of the incidence of critical illness for patients, and targeted measures can greatly reduce the pain and mortality of patients, avoid missing the best time for treatment, and avoid excessive use of drugs, which is conducive to the rational allocation of hospital resources.

For the prediction of critical illness risk, various monitoring indicators of patients are the key points to prediction. The quality and selection of the indicators have an important impact on the accuracy and reliability of the results for prediction. However, the samples in the dataset of critical illness are not complete, and there are various redundant and repetitive phenomena. At the same time, the indicators in the dataset are not necessarily related to critical illness, and the indicators of patients required for different illnesses are inconsistent when making treatment decisions. For doctors and clinicians, excessive patient test indicators will also interfere with real-time diagnosis and treatment decisions and affect the accuracy of diagnosis. Therefore, an efficient method is needed to analyze and sort out these data, remove irrelevant and redundant indicators, and select key indicators, so as to improve the efficiency and accuracy of critical illness prediction and real-time diagnosis and treatment decisions of doctors.

At present, machine learning has been widely used in the medical field by researchers [5-10]. Bendi Venkata Ramana et al. [11] used support vector machine, C4.5 decision tree, BP neural network to complete diagnosis classification based on textual dataset of liver. Patricio Miguel et al. [12] used logistic regression, support vector machine, and random forest classification algorithms to predict the presence of breast cancer based on blood sample data. Aljaaf et al. [13] proposed a multi-level risk assessment of developing heart failure based on C4.5 decision tree. Otoom et al [14] presented a system to analyze and monitor coronary arteries, their data set had 76 features and only 13 features were used. Studies by Morelli V et al. [15] and Ursini F et al. [16] had shown that inflammatory factors, blood lipids and uric acid can be used as indicators for the evaluation and prediction of cardiovascular adverse events in patients with diabetes after surgery. Zhang et al. [17] adopted a classification scheme based on a one-class kernel principle component analysis (KPCA) model ensemble that has been proposed for the classification of medical images, and the accuracy rate of recognition in the UCI breast cancer dataset (diagnostic) was 92%. Spanhol et al. [18] used fusion rules to combine different convolutional neural networks for classification, and the average recognition rate on the BreakHis dataset [19] reached 83.2%. Han et al. [20] designed a structured deep-learning model combined with data enhancement methods, the accuracy of their work on the BreakHis dataset was 96%. These studies demonstrate the positive role of machine learning methods in the prediction of disease. Janez Demsar et al. [21] used the RELIEFF method combined with machine learning methods such as Naive Bayes to prove that a small number of features may carry enough information to build a reasonable and accurate prediction model. Prerna Sharma et al. [22] used the improved gray wolf algorithm to select features and make predictions for Perkins, achieving an estimated accuracy of 94.83%. Filipe Lucini et al. [23] used data mining methods combined with machine learning algorithms

such as support vector machines, to predict the future hospitalization and discharge of patients. Since a single classifier cannot make the diagnosis of all diseases, MadhuSudana Rao Nalluri et al. [24] proposed a hybrid classifier parameter optimization diagnosis system based on three evolutionary algorithms, support vector machines, and multilayer perceptron to achieve the diagnosis for mixture of diseases. The above researches show that artificial intelligence is a kind of effective method to predict the risk of illness and the features in the dataset are not need to be fully used in the prediction. However, most of the data features used in the prediction of the above studies were extracted using different methods for specific illnesses or relying on the experience of the physician. In 2018, L. Nelson Sanchez-Pinto et al. [25] analyzed 8 different feature selection methods for the variable selection methods of machine learning models currently used in clinical diagnosis. Their results showed that the regression-based feature selection method seems to achieve better parsimony in clinical prediction on smaller data sets, while tree-based methods perform better on larger data sets. L. Nelson Sanchez-Pinto et al. studied the effectiveness of two types of feature extraction algorithms in different-size data sets, and their work is instructive. For a range of critical illnesses, an effective method is still needed to analyze their association with preoperative and intraoperative indicators of patients.

In view of the characteristics of various medical indicators and different formats of critical illness patients, this paper proposed a critical illness indicator analysis model based on a supervised feature selection method. This model used a light gradient boosting machine (LightGBM) [26] and Shapley additive explanation (SHAP) values [27] to analyze the correlation between the indicators and critical illness and select the corresponding key indicators, so as to facilitate real-time diagnosis and treatment by doctors and improve the accuracy of prediction for critical illness.

## 2. The Proposed Model

This section mainly describes the key indicators analysis model for the prediction of critical illness. There are four parts in the proposed model and the overall flow of this model is shown in Fig. 1. Data extraction and merging, missing values processing and single unique values processing make up the first part. Collinear indicators processing is the second part. Analysis of the importance of indicators is the third part. Key indicators selection is the last part.

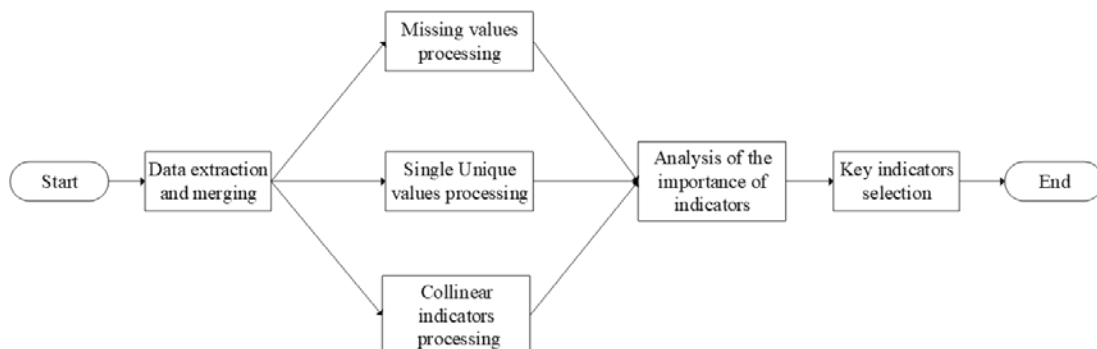


Fig. 1. Flowchart of the proposed method.

## 2.1 Data preprocessing

### 2.1.1 Data extraction and merging

To analyze the correlation between indicators and critical illness, the preoperative and intraoperative test data should be extracted and combined correctly. Since patients were not tested for all indicators during each preoperative examination, there were some missing values in various preoperative test tables. In the preoperative samples, the proposed model took 14 days as the threshold and completes the examination information of each patient through (1), where  $e_{i,j}, e_{k,j}$  represent the examination values of the indicator  $j$  detected of a patient at time  $i$  and  $k$ . Threshold set to 14 days was the result after a discussion with doctors. Then, the preoperative information of patients in each table was combined with reference to the medical record number and the operation time of patients in the surgical information table.

$$e_{i,j} = \begin{cases} \text{null} & i - k > 14 \text{ days} \\ e_{k,j} & i - k \leq 14 \text{ days} \end{cases} \quad (1)$$

The intraoperative monitoring data of patients were time series. To combine these data with preoperative test data, the statistical values such as mean values, variance, standard deviation, max values, min values, kurtosis and skewness of indicators in the intraoperative monitoring data were calculated. Then the intraoperative monitoring data were represented by the statistic values.

After extracting the preoperative and intraoperative test data, the integrated examination data were generated by combining the two kinds of data with reference to the medical record number and the operation time of patients. Then, the test dataset has been extracted.

### 2.1.2 Missing value processing

Although the preoperative test dataset has been filled in the previous part, a large number of missing values still exist in the test dataset, which is caused by incomplete data records and different test indicators for different patients. The amount of missing value affects the contribution of this indicator for the prediction of critical illness and the indicators with a large number of missing values are invalid for classification. The proposed model statistically analyzed the missing values of all indicators. Firstly, the proportion of missing values in each indicator was calculated through (2), where  $m_i$  represents the proportion of missing values for indicator  $i$ ,  $MissNum$  is the number of missing values in indicator  $i$ ,  $TotalNum$  is the total number of values in indicator  $i$ , namely the number of samples in the test dataset. Then, a threshold  $M_T$  of missing value was set (as shown in **Table 2**,  $M_T=0.90$ ). If the proportion of an indicator was higher than the threshold, the indicator was removed.

$$m_i = \frac{MissNum_i}{TotalNum} \quad (2)$$

### 2.1.3 Single unique value processing

In addition to missing values, some indicators only have one kind of values in the test dataset. The cause of this problem is that the values of some test indicators for the patients are the same. For example, all patients in the test dataset may not have used a certain drug in the past. These indicators are the same for all patients, regardless of whether the patients are sick or not. Therefore, these indicators have no contribution to the prediction of critical illness.

To overcome this problem, the proposed model calculated the number of different values for all indicators in the datasets. Then the indicators which only have a single unique value were removed.

## 2.2 Correlation analysis

Due to the large number of indicators in the test dataset, there will inevitably be some correlation between different indicators. However, two indicators with high correlation have the same contribution for predicting critical illness. Therefore, it is necessary to remove indicators with higher correlation to improve the predictive efficiency of critical illness.

In addition to continuous test indicators, there are also non-continuous indicators such as gender. Therefore, the proposed method analyzed the correlation between all the indicators in the datasets by Spearman correlation coefficient. First, the values of the two indicators that need to calculate the correlation were sorted, and the order of each value in the sorting was recorded. Then the recorded sequences of the two indicators were subtracted to obtain the vector  $d$ . Finally, the correlation coefficient of the two indicators is calculated according to (3), where  $d_i$  is the value of  $d$  at position  $i$ , and  $n$  is the number of elements in the indicator, namely the number of samples in the dataset. To select highly correlated pairs of indicators, a threshold was set and the pairs of indicators with a correlation coefficient greater than the threshold was selected. Afterward, one of the indicators in each pair of the selected indicators pairs was removed.

$$\rho = 1 - \frac{6 \times \sum_{i=1}^n d_i^2}{n \times (n^2 - 1)} \quad (3)$$

## 2.3 Importance analysis

To analyze the correlation between the key indicators and critical illness, the crucial part was estimating the importance of each indicator for the prediction of critical illness. The importance of the indicator represents how much this indicator contributes to the prediction.

In this part, the proposed model used LightGBM [26] as the classifier to predict critical illness. LightGBM was proposed by Guolin Ke et al. at 2017. It was an implementation of Gradient Boosting Decision Tree (GBDT) [28]. This algorithm used two novel techniques: Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB), which make LightGBM outperform XGBoost [29] and SGB [30] in terms of computational speed and memory consumption.

The objective function of LightGBM in the proposed model is shown as (4), where  $t$  is the current iteration,  $n$  is the total number of samples,  $y_i$  is the label of  $x_i$ ,  $f_t$  represents the CART decision tree in the  $t$  iteration,  $f_t(x_i)$  is the prediction of  $x_i$  in the  $t$  iteration,  $L$  is the loss function and  $\Omega$  is the regularization term.

$$Obj^t = \sum_{i=1}^n L(y_i, f_t(x_i)) + \Omega(f_t) \quad (4)$$

Since the model was trained to predict critical illness, the contribution of each feature to the model was equivalent to the contribution of each indicator to preoperative critical illness prediction. Therefore, the importance of each feature in the model can be taken as the importance of each indicator for preoperative prediction of critical illness after training the model. In the proposed model, SHAP values was used with an explainer for LightGBM and analyze the contribution of indicators to the prediction. Through the explainer, each indicator

was assigned an important value.

## 2.4 Key indicators selection

The importance of indicators was used to select the key indicators in this part. Firstly, the indicators were sorted in descending order of their importance. Secondly, the importance of all indicators was normalized by (5), where  $v_i$  is the importance of indicator  $i$ ,  $N_i$  is the normalized importance of indicator  $i$ ,  $m$  is the number of indicators in the test dataset. Thirdly, the cumulative sum (cumsum) of the normalized indicators was calculated by (6), where  $N_k$  is the cumsum of indicator  $k$ .

Finally, a threshold  $S_r$  was set and the indicators with cumsum values lower than the threshold was selected as the rough key indicators. The parameters are shown in [Table 2](#). If the number of the rough key indicators was more than 25, the rough key indicators were set as the key indicator set. If the number of the rough key indicators was less than 25, the first 25 indicators in the indicator series were set as the key indicators.

$$N_i = \frac{v_i}{\sum_{i=1}^m v_i} \quad (5)$$

$$C_k = \sum_{i=1}^k N_i \quad (6)$$

## 3. Experiment and Results

### 3.1 Description of the experiment

The data of two critical illnesses were collected from preoperative and intraoperative clinical features and laboratory results of patients in the southwest hospital in China. Diagnoses of heart failure and respiratory failure were derived from database entries. After data preprocessing, the number of samples for heart failure and respiratory failure is shown in [Table 1](#). It is obvious that the number of negative samples is greater than the number of positive samples in each type of critical illness. In response to this problem, the downsampling method which called random undersampling was used to reduce the number of negative samples to the same as the number of positive samples. The parameters of the proposed model are shown in [Table 2](#).

The test dataset was split into training samples and test samples in a ratio of 4:1. The indicators of each patient were used as the features of each sample. The labels of patients with critical illness are set to 1 and the labels of patients without critical illness are set to 0. In the experiment, all the importance of indicators and the results of prediction are taken from the average value of the classifiers after 10 runs.

**Table 1.** Number of samples

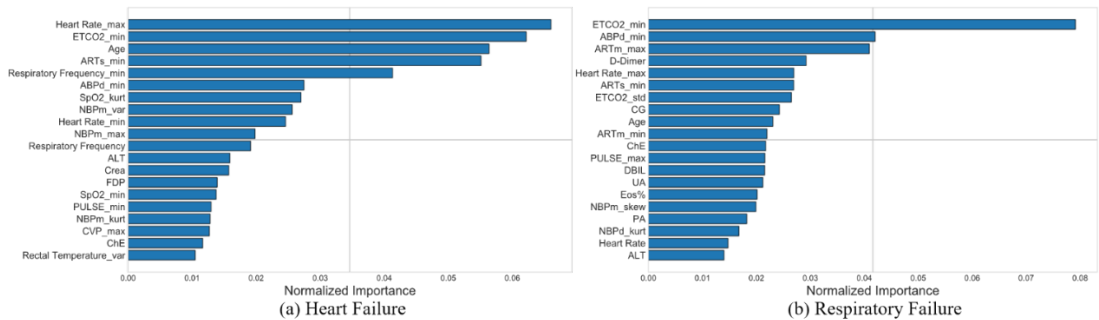
Samples	Heart failure	Respiratory failure
Positive samples	609	194
Negative samples	4223	4223

**Table 2.** Experiment parameters

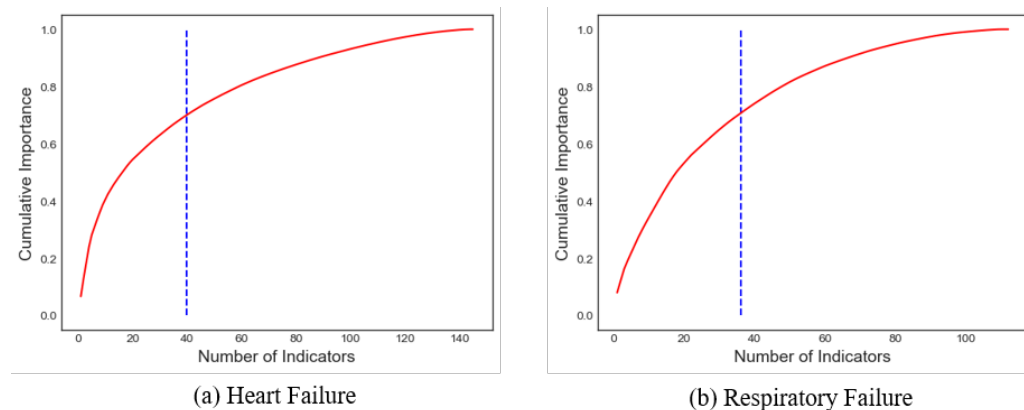
Parameters	Values
$M_T$	0.90
$C_T$	0.98
$I$	10
$S_T$	0.70

### 3.2 Results of the experiment

The number of indicators in the samples of heart failure and respiratory failure are 197 and 189 respectively. After data preprocessing, the number of indicators removed from the samples of heart failure and respiratory failure are 52 and 77 respectively. Fig. 2 shows the normalized importance for top 20 indicators of the two critical illnesses. It can be seen that the contribution of the top 20 indicators for the prediction of heart failure is lower than the contribution of the top 20 indicators for the prediction of respiratory failure. Fig. 3 shows the graph for cumulative indicators importance of heart failure and respiratory failure. The number of indicators of the heart failure with cumsum values lower than  $S_T$  is 39. Thus, the top 39 indicators are used as the key indicators of heart failure. The number of the indicators of the respiratory failure with cumsum values lower than  $S_T$  is 35. Hence, the top 35 indicators are used as the key indicators of respiratory failure. The two kinds of key indicators are shown in Table 3.



**Fig. 2.** Normalized importance for top20 indicators



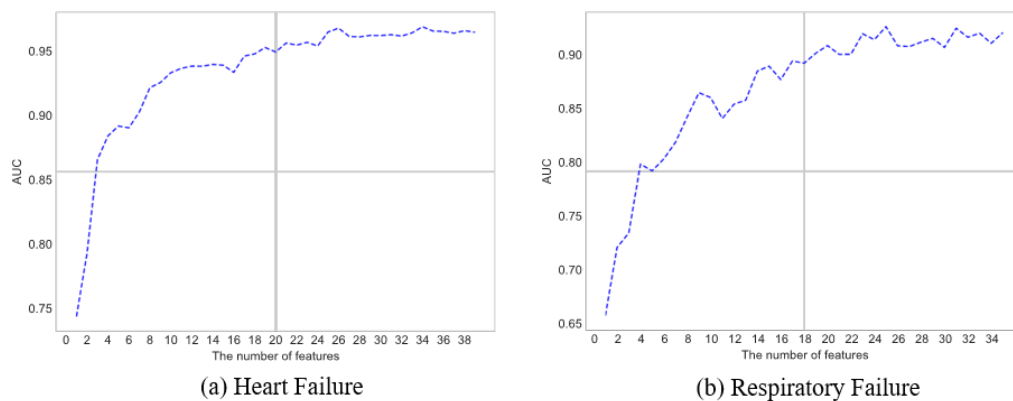
**Fig. 3.** Cumulative indicators importance of heart failure. The value of the ordinate for the intersection of the vertical dotted line and the curve in the three figures is  $S_T$ . The vertical dotted line divides the indicators into two parts, and the left part represents the number of indicators with cumsum values lower than  $S_T$ .

**Table 3.** Key indicators of two critical illnesses

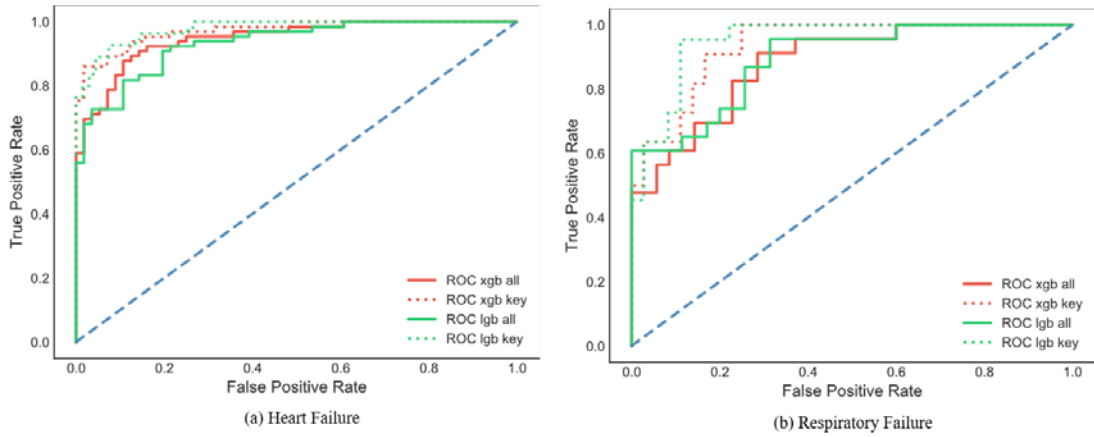
Critical illness	Indicators name
Heart failure	Heart Rate_max, ETCO2_min, Age, ARTs_min, Respiratory Frequency_min, ABPd_min, SpO2_kurt, NBPm_var, Heart Rate_min, NBPm_max, Respiratory Frequency, ALT, Crea, FDP, SpO2_min, PULSE_min, NBPm_kurt, CVP_max, ChE, Rectal Temperature_var, Heart Rate, CK-MB, Respiratory Frequency_kurt, UA, SpO2, NBPm_max, ARTm_min, Respiratory Frequency_max, DBIL, APTT, NBPm_mean, Rectal Temperature_kurt, ARTs, Cys-C, HBDH, ALB/GLO, Tempmure_var, NBPm_kurt, Na
Respiratory	ETCO2_min, ABPd_min, ARTm_max, D-Dimer, Heart Rate_max, ARTs_min, ETCO2_std, CG, Age, ARTm_min, ChE, PULSE_max, DBIL, UA, Eos%, NBPm_skew, PA, NBPd_kurt, Heart Rate, ALT, Crea, ETCO2_max, SpO2, NBPm_min, CK-MB, AST, ETCO2_skew, CVP_max, CK, PULSE_var, INR, Eos#, ABPd_max, ETCO2_kurt, Hb

To verify the effectiveness of the key indicators, it was used to generate multiple indicators of subsets. The number of subsets is the number of indicators in the key indicators. The first subset contains only the most important indicators of the key indicators. The second set contains the top two most important metrics from the key indicators. And so on, the last set contains all indicators in the key indicators. These indicators subsets were used to generate data subsets from the test dataset. Then, the data subsets were sorted in ascending order by the number of indicators included. The AUC (Area Under Curve) of these subsets in predicting critical illness was analyzed in turn by this order. The classifier used in the analysis was LightGBM.

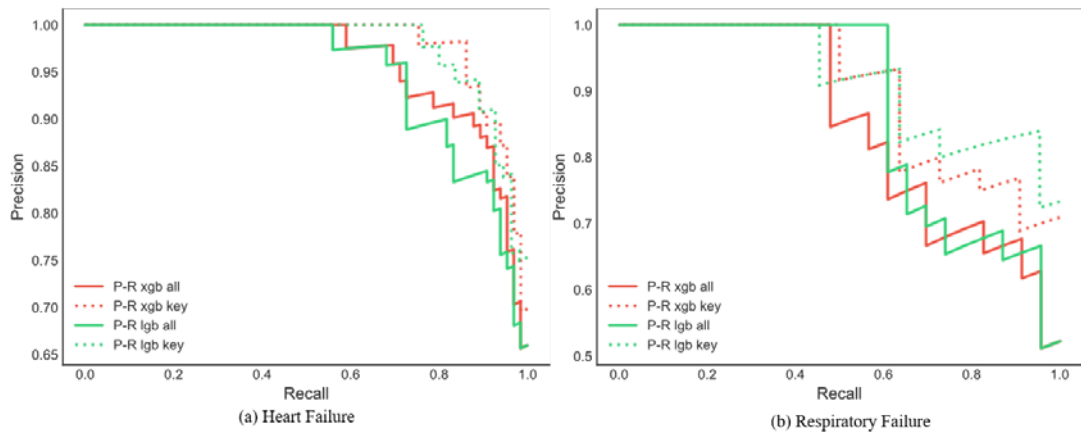
The AUC of prediction using data subsets generated by the key indicators of heart failure and respiratory failure are shown in Fig. 4. It can be seen that, with the increase of indicators number, the AUC for the prediction of heart failure and respiratory failure keep rising and have reached the convergence around 0.96 and 0.92, respectively. Fig. 5 and Fig. 6 show the ROC curve and P-R curve for the prediction of heart failure and respiratory failure. For the two kinds of critical illnesses, the area under the dotted line is both slightly greater than that under the solid line. This indicates that the performance of the prediction using key indicators is better than the performance of the prediction using all indicators.

**Fig. 4.** AUC of the data subsets





**Fig. 5.** ROC curve for the prediction of the two critical illnesses. The red curve and green curve are the results of prediction using XGBoost and LightGBM respectively. The dotted line and solid line represent the results of prediction using all indicators and the key indicators respectively.



**Fig. 6.** P-R curve for the prediction of the two critical illnesses. The red curve and green curve are the results of prediction using XGBoost and LightGBM respectively. The dotted line and solid line represent the results of prediction using all indicators and the key indicators respectively.

In order to further verify the effect of the key indicators, LightGBM and XGBoost were used as classifiers to analyze the prediction using all indicators and the key indicator respectively. The predicted results using all indicators and the key indicators of the two kinds of critical illnesses were shown in [Table 4](#) and [Table 5](#). The results in [Table 4](#) and [Table 5](#) were taken from the 10-fold cross validation. HFA, HFK, RFA and RFK represent all indicators of heart failure, key indicators of heart failure, all indicators of respiratory failure and key indicators of respiratory failure respectively. The classifiers used in [Table 4](#) and [Table 5](#) are LightGBM and XGBoost respectively. These results include the accuracy, AUC, f1\_score, sensitivity and specificity of the prediction for the heart failure and respiratory failure. For the heart failure, the experimental results showed that the prediction results using key indicators are similar to those using all indicators in accuracy and specificity, and slightly higher in AUC, f1\_score and sensitive than those using all indicators. For the respiratory, the experimental results show that the prediction results using the key indicators are higher than the prediction results using all indicators, especially in terms of AUC, f1\_score and sensitive.

These results suggest that the key indicators extracted from the model can effectively predict the critical illness.

**Table 4.** Predicting results using LightGBM

Category	Accuracy	AUC	F1_score	Sensitive	Specificity
HFA	0.901(0.884, 0.918)	0.963(0.953, 0.973)	0.897(0.880, 0.915)	0.869(0.839, 0.899)	0.934(0.914, 0.953)
HFK	0.913(0.891, 0.936)	0.969(0.959, 0.980)	0.911(0.889, 0.933)	0.889(0.864, 0.914)	0.939(0.914, 0.964)
RFA	0.84(0.79, 0.89)	0.91(0.87, 0.95)	0.73(0.64, 0.81)	0.66(0.56, 0.75)	0.93(0.89, 0.97)
RFK	0.86(0.82, 0.90)	0.92(0.89, 0.95)	0.77(0.70, 0.83)	0.71(0.63, 0.79)	0.94(0.90, 0.97)

**Table 5.** Predicting results using XGBoost

Category	Accuracy	AUC	F1_score	Sensitive	Specificity
HFA	0.91(0.89, 0.93)	0.97(0.96, 0.97)	0.91(0.89, 0.93)	0.89(0.86, 0.91)	0.93(0.91, 0.96)
HFK	0.91(0.89, 0.94)	0.97(0.96, 0.98)	0.91(0.89, 0.94)	0.89(0.86, 0.92)	0.94(0.91, 0.96)
RFA	0.83(0.80, 0.86)	0.91(0.88, 0.94)	0.72(0.67, 0.77)	0.66(0.59, 0.74)	0.92(0.88, 0.96)
RFK	0.85(0.81, 0.90)	0.93(0.90, 0.95)	0.76(0.68, 0.83)	0.69(0.61, 0.77)	0.94(0.90, 0.98)

## 4. Discussion

Machine learning algorithms which use distance measurement are sensitive to data with missing values. Therefore, the proposed model in this paper chooses LightGBM to analyze the importance of indicators.

The prediction performance of these critical illnesses using the key indicators extracted by the proposed model from the two kinds of critical illnesses are convergent. This shows that the key indicators extracted by the model are effective. At the same time, for the samples of heart failure and respiratory, the performance of key indicators extracted from the model in the prediction of critical illness was slightly higher than the performance of using all indicators. These experimental results show that the key indicators can replace all indicators to make an effective prediction of critical illness. This indicates that the model successfully removes redundant indicators from the dataset and avoids the influence of these redundant indicators on the prediction performance of critical illness.

However, the key indicators extracted from the model still have some shortcomings. To ensure the validity of patient samples, the model does not fill in the missing values in the

examined data of patients. Although the model has removed some indicators with a lot of missing values, there still exist indicators with missing values in the dataset. Due to the influence of missing values and other factors of the data, the importance of some indicators that should be significant for the critical illness according to doctors' experience may not be ranked in the first few in the important analysis of the model. However, some common indicators have no missing values, so the importance obtained in the model is higher than the actual importance of them. For the model, the future work needs to analyze the impact of missing values and other factors on the importance of indicators, combined with the doctor's prior knowledge of the importance of the indicators, so that the analysis results obtained by the model are closer to their real correlation with critical illness.

For the model's generalizability, it is possible to obtain different results when using the same modeling approach on data from different centers. This can occur due to several factors: differences in data distribution, sample size and quality, variations across centers. In the future study, the model's generalization and versatility need to be improved through proper data and model handling techniques, leading to more consistent predictions across various data centers.

## 5. Conclusion

This paper proposes a key indicator analysis model of postoperative critical illness based on machine learning. The model includes four modules: data preprocessing, analysis of the correlation in indicators, analysis of the importance for indicators and selection of key indicators. The missing value, single value, correlation and importance of indicators are counted and calculated in the first three modules, and the key indicators in the data are selected based on these results in the last module. In this paper, two kinds of samples for critical illnesses from surgical patients were used in the experiment. The prediction accuracy of postoperative critical illness obtained by using the key indicators extracted from the model is slightly higher than those obtained by using all the indicators. The experimental results show that the model presented in this paper can effectively analyze the correlation of the indicators with postoperative critical illnesses and extract the key indicators of postoperative critical illnesses.

## Funding

The work was supported by Chongqing Municipal Natural Science Foundation (NO: CSTB2022NSCQ-MSX0894), Youth Innovation Promotion Association of Chinese Academy of Sciences (No. 2020377), National Natural Science Foundation project(82070630), Chongqing Talent Plan Project (CQYC202103080) and Chongqing Graduate Education Teaching Reform Research Project (Key subject: yjg202039).

## Acknowledgement

We thank all the people who participated in this study.

## References

- [1] E. H. Lawson, B. L. Hall, R. Louie, S. L. Ettner, D. S. Zingmond, L. Han, M. Rapp, and C. Y. Ko, "Association between occurrence of a postoperative complication and readmission: implications for quality improvement and cost savings," *Ann Surg*, vol. 258, no. 1, pp. 10-8, Jul, 2013. [Article \(CrossRef Link\)](#)
- [2] R. Pathak, S. Giri, M. R. Aryal, P. Karmacharya, V. R. Bhatt, and M. G. Martin, "Mortality, length of stay, and health care costs of febrile neutropenia-related hospitalizations among patients with breast cancer in the United States," *Support Care Cancer*, vol. 23, no. 3, pp. 615-617, Mar, 2015. [Article \(CrossRef Link\)](#)
- [3] S. F. Khuri, W. G. Henderson, R. G. DePalma, C. Mosca, N. A. Healey, and D. J. Kumbhani, "Determinants of long-term survival after major surgery and the adverse effect of postoperative complications," *Ann Surg*, vol. 242, no. 3, pp. 326-343, discussion 341-3, Sep, 2005. [Article \(CrossRef Link\)](#)
- [4] M. A. Hamilton, M. Cecconi, and A. Rhodes, "A systematic review and meta-analysis on the use of preemptive hemodynamic intervention to improve postoperative outcomes in moderate and high-risk surgical patients," *Anesth Analg*, vol. 112, no. 6, pp. 1392-1402, Jun, 2011. [Article \(CrossRef Link\)](#)
- [5] Y. Chen, and B. Qi, "Representation learning in intraoperative vital signs for heart failure risk prediction," *BMC Med Inform Decis Mak*, vol. 19, no. 1, pp. 260, Dec 9, 2019. [Article \(CrossRef Link\)](#)
- [6] Y. Chen, J. Zhang, and X. Qin, "Interpretable instance disease prediction based on causal feature selection and effect analysis," *BMC Med Inform Decis Mak*, vol. 22, no. 1, pp. 51, Feb 26, 2022. [Article \(CrossRef Link\)](#)
- [7] Y. Chen, K. Zhong, Y. Zhu, and Q. Sun, "Two-stage hemoglobin prediction based on prior causality," *Front Public Health*, vol. 10, pp. 1079389, 2022. [Article \(CrossRef Link\)](#)
- [8] Y. W. Chen, Y. J. Li, P. Deng, Z. Y. Yang, K. H. Zhong, L. G. Zhang, Y. Chen, H. Y. Zhi, X. Y. Hu, J. T. Gu, J. L. Ning, K. Z. Lu, J. Zhang, Z. Y. Xia, X. L. Qin, and B. Yi, "Learning to predict in-hospital mortality risk in the intensive care unit with attention-based temporal convolution network," *BMC Anesthesiol*, vol. 22, no. 1, pp. 119, Apr 23, 2022. [Article \(CrossRef Link\)](#)
- [9] Y. W. Chen, and J. Liu, "Polynomial dendritic neural networks," *Neural Computing & Applications*, vol. 34, no. 14, pp. 11571-11588, Jul, 2022. [Article \(CrossRef Link\)](#)
- [10] Y. W. Chen, X. L. Qin, L. G. Zhang, and B. Yi, "A Novel Method of Heart Failure Prediction Based on DPCNN-XGBOOST Model," *Cmc-Computers Materials & Continua*, vol. 65, no. 1, pp. 495-510, 2020. [Article \(CrossRef Link\)](#)
- [11] B. V. Ramana, M. S. P. Babu, and N. B. J. I. J. o. D. M. S. Venkateswarlu, "A Critical Study of Selected Classification Algorithms for Liver Disease Diagnosis," *International Journal of Database Management Systems ( IJDMS )*, vol. 3, pp. 101-114, 2011. [Article \(CrossRef Link\)](#)
- [12] M. Patrício, J. Pereira, J. Crisóstomo, P. Matafome, M. Gomes, R. Seiça, and F. Caramelo, "Using Resistin, glucose, age and BMI to predict the presence of breast cancer," *BMC Cancer*, vol. 18, no. 1, pp. 29, Jan 4, 2018. [Article \(CrossRef Link\)](#)
- [13] A. J. Aljaaf, D. Al-Jumeily, A. J. Hussain, T. Dawson, P. Fergus, M. Al-Jumaily, and Ieee, "Predicting the Likelihood of Heart Failure with a Multi Level Risk Assessment Using Decision Tree," in *Proc. of 2015 Third International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAECE)*, pp. 101-106, 2015. [Article \(CrossRef Link\)](#)
- [14] A. Otoom, E. Abdallah, Y. Kilani, A. Kefaye, and M. Ashour, "Effective diagnosis and monitoring of heart disease," *International Journal of Software Engineering and its Applications*, vol. 9, pp. 143-156, 01/01, 2015. [Article \(CrossRef Link\)](#)
- [15] V. Morelli, S. Palmieri, A. Lania, A. Tresoldi, S. Corbetta, E. Cairoli, C. Eller-Vainicher, M. Arosio, M. Copetti, E. Grossi, and I. Chiodini, "Cardiovascular events in patients with mild autonomous cortisol secretion: analysis with artificial neural networks," *Eur J Endocrinol*, vol. 177, no. 1, pp. 73-83, Jul, 2017. [Article \(CrossRef Link\)](#)

- [16] F. Ursini, E. Russo, S. D'Angelo, F. Arturi, M. L. Hribal, L. D'Antona, C. Bruno, G. Tripepi, S. Naty, G. De Sarro, I. Olivieri, and R. D. Grembiale, "Prevalence of Undiagnosed Diabetes in Rheumatoid Arthritis: an OGTT Study," *Medicine (Baltimore)*, vol. 95, no. 7, pp. e2552, Feb, 2016. [Article \(CrossRef Link\)](#)
- [17] Y. G. Zhang, B. L. Zhang, F. Coenen, J. M. Xiao, and W. J. Lu, "Erratum to: One-class kernel subspace ensemble for medical image classification," *Eurasip Journal on Advances in Signal Processing*, pp. 1, Oct, 2015. [Article \(CrossRef Link\)](#)
- [18] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using Convolutional Neural Networks," in *Proc. of 2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 2560-2567, 2016. [Article \(CrossRef Link\)](#)
- [19] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification," *IEEE Trans Biomed Eng*, vol. 63, no. 7, pp. 1455-1462, Jul, 2016. [Article \(CrossRef Link\)](#)
- [20] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, and S. Li, "Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model," *Sci Rep*, vol. 7, no. 1, pp. 4172, Jun 23, 2017. [Article \(CrossRef Link\)](#)
- [21] J. Demsar, B. Zupan, N. Aoki, M. J. Wall, T. H. Granchi, and J. Robert Beck, "Feature mining and predictive model construction from severe trauma patient's data," *Int J Med Inform*, vol. 63, no. 1-2, pp. 41-50, Sep, 2001. [Article \(CrossRef Link\)](#)
- [22] P. Sharma, S. Sundaram, M. Sharma, A. Sharma, and D. Gupta, "Diagnosis of Parkinson's disease using modified grey wolf optimization," *Cognitive Systems Research*, vol. 54, pp. 100-115, 2019. [Article \(CrossRef Link\)](#)
- [23] F. R. Lucini, F. S. Fogliatto, G. J. C. da Silveira, J. L. Neyeloff, M. J. Anzanello, R. S. Kuchenbecker, and B. D. Schaan, "Text mining approach to predict hospital admissions using early medical records from the emergency department," *Int J Med Inform*, vol. 100, pp. 1-8, Apr, 2017. [Article \(CrossRef Link\)](#)
- [24] M. R. Nalluri, K. K. M. M., and D. S. Roy, "Hybrid Disease Diagnosis Using Multiobjective Optimization with Evolutionary Parameter Optimization," *J Healthc Eng*, vol. 2017, pp. 5907264, 2017. [Article \(CrossRef Link\)](#)
- [25] L. N. Sanchez-Pinto, L. R. Venable, J. Fahrenbach, and M. M. Churpek, "Comparison of variable selection methods for clinical predictive modeling," *Int J Med Inform*, vol. 116, pp. 10-17, Aug, 2018. [Article \(CrossRef Link\)](#)
- [26] M. Qi, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree,".
- [27] S. Lundberg, and S. I. Lee, "A Unified Approach to Interpreting Model Predictions," 2017. [Article \(CrossRef Link\)](#)
- [28] H. F. Jerome, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189-1232, 2001. [Article \(CrossRef Link\)](#)
- [29] T. Chen, and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. of KDD '16*, 2016. [Article \(CrossRef Link\)](#)
- [30] M. Schonlau, "Boosted Regression (Boosting): An Introductory Tutorial and a Stata Plugin," *Stata Journal*, vol. 5, pp. 330-354, 2005. [Article \(CrossRef Link\)](#)



**Lei Han** was born in Sichuan, China, in 1985. He received the MM from the Naval Medical University. He is currently an assistant researcher of Department of Trauma Medical Center, Third Military Medical University. His main research interests include the quality of diagnosis and treatment of critical illness and patient safety management.



**Yiziting Zhu** received her MM degree from Chongqing Medical University. She's now an anesthesiologist of the Department of Anesthesiology at Southwest Hospital. Her main research focus on the application of machine learning in the perioperative medicine.



**Yuwen Chen** received the MA.Sc from the University of Science and Technology of China and Ph.D from the Chengdu Institute of Computer Application, Chinese Academy of Sciences. He is currently an associate researcher and Master Instructor of Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences. His main research interests include machine learning applied to the medical field, providing doctors with decision support tools based on large-scale health data.



**Guoqiong Huang** was born in 1965. She received the MA and Ph.D of Social Medicine and Health service Management from the Army Medical University. She is currently a researcher of Administration Center of Third Military Medical University. Her main research interests include health service management and medical education evaluation.



**Bin Yi** received the Clinical Medicine Doctor degree from the Third Military Medical University and is now the director, professor and chief physician of the Department of Anesthesiology at Southwest Hospital. His main research interest is the application of big data in perioperative critical illness management.