International Journal of Innovative Computing, Information and Control
Volume **16**, Number **6**, December **2020**

ICIC International ©2020 ISSN 1349-4198
pp. **2035–2046**

PREOPERATIVE RISK PREDICTION OF HEART FAILURE WITH NUMERICAL AND TEXTUAL ATTRIBUTES

Yuwen Chen[1,2,3], Lige Zhang[1,3], Ju Zhang[2], Bin Yi[4] and Xiaolin Qin[1,3,*]

[1]Chengdu Institute of Computer Applications
Chinese Academy of Sciences
No. 9, 4th Section of South Renmin Road, Chengdu 610041, P. R. China
chenyuwen@cigit.ac.cn; zhanglige15@mails.ucas.ac.cn; *Corresponding author: keyanche@163.com

[2]Chongqing Institute of Green and Intelligent Technology
Chinese Academy of Sciences
No. 266, Fangzheng Avenue, Beibei District, Chongqing 400714, P. R. China
zhangju@cigit.ac.cn

[3]School of Computer and Control Engineering
University of Chinese Academy of Sciences
No. 19(A), Yuquan Road, Shijingshan District, Beijing 100049, P. R. China

[4]Department of Anesthesia
The First Hospital Affiliated to Army Medical University (Southwest Hospital)
No. 30, Gaotanyanzheng Street, Shapingba District, Chongqing 400038, P. R. China
yibin1974@163.com

Received January 2020; revised August 2020

Abstract. *The occurrence of perioperative heart failure will affect the quality of medical services and threaten the safety of patients. It is of great significance to improve the quality of medical service and ensure the safety of patients with perioperative heart failure by scientific method. A novel method of predicting preoperative critical disease is proposed for the properties of medical heterogeneity data in perioperative patients. This work is the first study that integrates the numerical laboratory data and textual diagnostic data of patient for extracting the predictive feature by XGBOOST and LDA (Latent Dirichlet Allocation) model respectively, and builds a low cost, scalable and effective model of predicting preoperative critical disease by using logistic regression. Firstly, the data of preoperative patients are preprocessed, and the patient data is divided into the structural data of numerical data and the unstructured data of textual data. The numerical data of the patient is constructed by the gradient boosting tree model, and the textual features of the patients are extracted by the topic model of the text-based data. Fuse textual features and numerical features, and finally through a simple logistic regression predict the patient critical illness. To evaluate the performance of the proposed method, we made preoperative prediction of critical events of heart failure in perioperative period based on the real operative data of patients from a hospital, and the results showed that the sensitivity and specificity of the model proposed could reach 90% and 93%, thus verifying the feasibility and effectiveness of the model.*
Keywords: Preoperative risk prediction, LDA model, XGBOOST

1. **Introduction.** The occurrence of perioperative adverse events will increase medical expenses of patients, prolong the recovery time, influence the recovery of patients [1,2], and even lead to the death of patients. Studies have shown that critical adverse events within 30 days of surgery can reduce the median survival time of patients by 69% [3], and the long-term consequences of critical adverse events in short-term surgery have a significant

DOI: 10.24507/ijicic.16.06.2035

impact on the long-term survival and quality of life of patients [4]. The incidence of intraoperative adverse events in patients with heart failure is 2%-17%, which has become an important cause of perioperative complications and mortality. Currently, heart failure is considered as one of the most fatal human diseases in the world. However, early prediction techniques for perioperative adverse cardiac events are still lacking. Simple adverse event warning systems often fail to catch the signs of adverse events in heart failure. The treatment will be difficult and the effect of intervention will be limited when adverse events occur. Therefore, actively predicting the risk of adverse events of heart failure is conducive to the early detection, early warning, intervention clues of adverse events and diagnosis, which has important scientific significance and social value.

Recently, artificial intelligence technology has been widely used in the medical field [5-11]. New techniques have opened up a new way of studying heart failure. At present, the researches on heart failure are mainly based on the patient's medical records, physical characteristics, auxiliary examinations, treatment plans, etc., and algorithms are used to build a predictive model. In addition, most studies mainly analyze the characteristics of ECG data and establish diagnostic models for heart failure [12-17]. Zheng et al. [18] proposed a data analysis method for patients with heart failure based on Support Vector Machine (SVM), including age, medical insurance type, sensitivity assessment (audiovisual and thinking), complications, first aid, drug risk, and final hospital stay, etc., and established a rehospitalization prediction model for patients with heart failure, with a prediction accuracy of 78.4%. Choi et al. [19] used the recursive neural network algorithm to analyze the diagnostic data of patients with heart failure, including the time series of doctor's advice, spatial density and other characteristics. They established a diagnostic model of heart failure, and the Area Under Curve (AUC) of the diagnostic model was verified by experiments to be 0.883. Chen et al. [20] used the Support Vector Machine (SVM) algorithm to analyze the 24-hour dynamic electrocardiogram of patients with heart failure and healthy controls. Shameer et al. [21] used Naive Bayes algorithm to analyze the diagnostic data, treatment data, examination data, doctor's orders records, vital signs data and other data variables of patients with heart failure. They established a read-admission prediction model for patients with heart failure, and the prediction AUC was 0.78.

However, the above research still focuses on the manual calculation and analysis of scoring and testing indices. Its prediction has poor time-validity, low accuracy, and the not fused heterogeneity data of patients. Specifically, the heterogeneity of patients is embodied in many aspects, such as age, occupation, gender, various physiological indices [22] as well as various types of numerical laboratory data, textual diagnostic information. The predictive results of perioperative critical illness using these indices are more practical and targeted. To make the results of research and analysis more accurate and more convenient to apply to practice, we use machine learning method to model the risk prediction of heart failure during and after operation based on the preoperative medical data of patients, so as to construct a scalable, low-cost and effective prediction solution.

The remainder of this paper is organized as follows. The method of this paper is presented in Section 2. Section 3 describes the process and results of experiment. Section 4 discusses the advantages, disadvantages and future improvement of this method. The conclusions are given in Section 5.

2. Methods.

This section mainly describes the preoperative risk prediction model of critical disease based on XGBOOST and Latent Dirichlet Allocation (LDA) models. Figure 1 shows an abstract framework for predicting critical disease using the gradient boosted decision tree [23], the topic model [24] and logistic regression [25]. From the medical

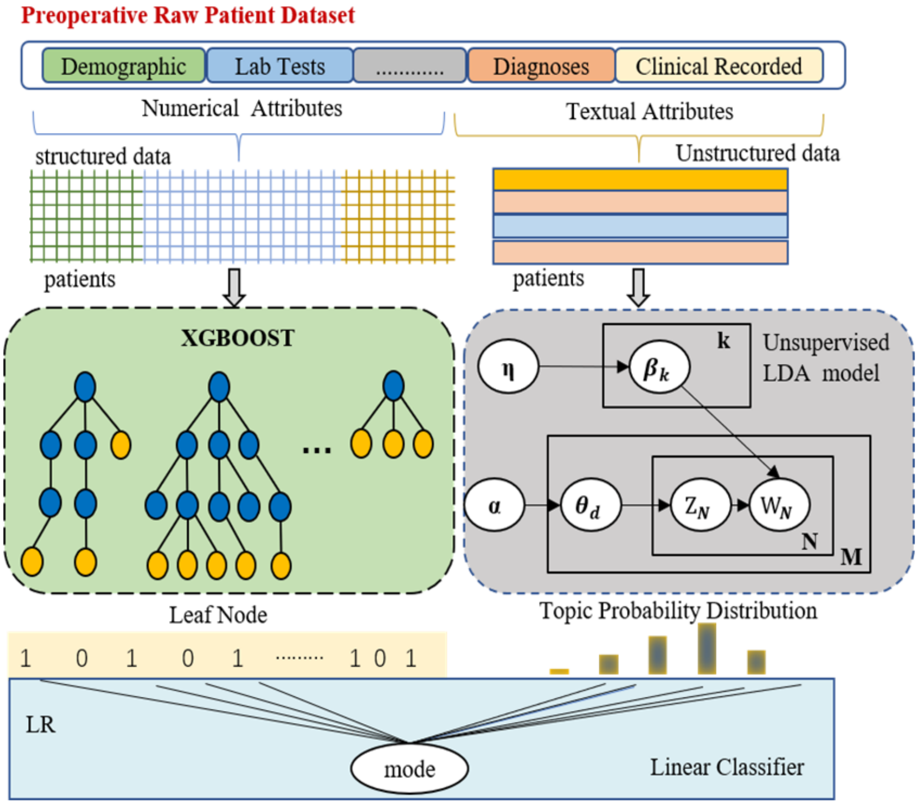

FIGURE 1. Conceptual framework for predicting critical illness models

information system (HIS, PACK, LIS) to extract the raw electronic medical record data of preoperative patients, each patient data can be preprocessed into structural numerical data (Age, Weight, BMI, etc.) and non-structural textual data (preoperative diagnosis, medical history, etc.). Use the patient numerical data and textual data as input, and the nonlinear features of a high level are extracted by the feature learning algorithm. Each patient then is represented by the combined numerical features and textual features, and the critical event prediction task can be executed through a simple and efficient logistic regression.

The numerical data of patients are extracted by gradient boosting algorithm, although neural networks [26-29] have been revived and popular in recent years. Gradient boosting algorithm still has its indispensable advantages in the scene of limited quantity of training sample and short training time, and is especially suitable for heterogeneous data. In the gradient boost tree, a weak learning regression tree is first initialized with the patient data, and then a regression tree is fitted according to the residual learned by the weak learner. Repeatedly in the direction of the gradient to learn fitting, a series of weak classifiers (called basic classifiers) are obtained, and finally these weak classifiers are combined to form a strong classifier for classification and judgment. The process can be likened to the diagnosis of the same patient by multiple doctors at the same time, and finally, a conclusion can be drawn by combining the results of all doctor visits. In this way, many regression trees are used to judge and learn more information of patients, which is not easy to have an overfitting problem, and the results are interpretable. The topic model is used to extract textual features from patients. LDA is unsupervised learning that can be used to identify topic information that lurks in large-scale document collection or corpus. The relevant text information of preoperative diagnosis of patients is formed into a text of patient, and the information of each patient is equivalent to a text description.

Through LDA, the inter-class variance of patients becomes larger and the intra-class variance becomes smaller, and the model learns the text features of different critical patients through unsupervised cluster. Details of the model are described as follows. The XGBOOST and LDA are used to extract the features of numerical data and textual data for patients. Then, these features are used as the input of a linear classifier to achieve the prediction for the critical illnesses.

2.1. **Numerical feature extraction of perioperative patients.** XGBOOST [30] is used to extract the numerical features of patients, which is an integrated learning method proposed by Tianqi Chen based on GBDT [23]. The improvement of XGBOOST algorithm to GBDT algorithm lies in that the second derivative is used to calculate the objective function in the process of model optimization, besides, the regularization term is added to the objective function to prevent the algorithm from over-fitting in the training process, and moreover, XGBOOST algorithm uses the idea of random forest for reference in the training process, and does not use all samples in the iteration process, and does not use every iteration. The generalization ability of the model is effectively improved by sampling all the features of the samples and training some of the features of the samples. As shown in Figure 2, the XGBOOST model was first trained with the data of the perioperative patients, then the new features were constructed by the leaf nodes of n trees which value is zero or one. The trees in Figure 2 represent the classification and regression trees which are used as base classifiers in XGBOOST. Patients are predicted by n trees respectively. If the prediction result falls on the leaves, the value is assigned to 1; otherwise it is 0. Therefore, the length of the patient's numerical feature is equal to the number of tree leaf nodes. The algorithm is described in detail in [27].

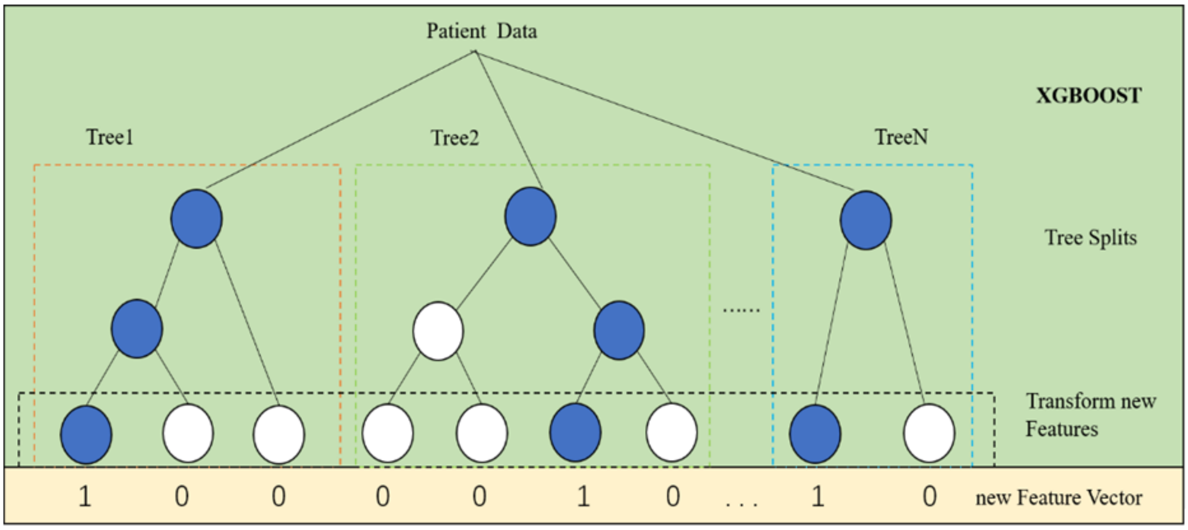

FIGURE 2. Patient numerical feature extraction

2.2. **Patient text feature extraction.** The text data of diagnostic of patients contain a large number of medical information, which has a great value in judging whether or not the patient develops critically illness events during or after surgery. For example: the conclusion of electrocardiogram and echocardiogram is: sinus arrhythmia, counterclockwise rotation, moderate tricuspid regurgitation, mild increased pulmonary arterial systolic pressure, mild pulmonary valve regurgitation, mild mitral valve regurgitation, mild aortic valve regurgitation, early decreased right as well as left ventricular function and the conclusion of another patient: normal electrocardiogram. The risk of heart failure during and after surgery is not the same. In addition, the patient previous illness and previous history of the intraoperative and postoperative critical adverse events also have the same judgment value, and patients with no previous history of the risk of critical events during and after surgery should be lower than patients with a previous medical history. The important textual information should not be overlooked when establishing a risk prediction model. Therefore, we pre-processed the diagnostic data of each patient into patient documentation. The topic model was used to cluster the patient documentation to get the topic probability distribution of the document. Then, the dimension of each patient medical diagnostic information can be reduced through the topic probability distribution of the document, as shown in Figure 3. Topic model is based on Latent Dirichlet Allocation (LDA) [24] which was proposed by Blei et al. in 2003 to estimate the topic distribution of the document. It can give the topic of each document in the document set as a probability distribution. The algorithm is described in detail in [21].

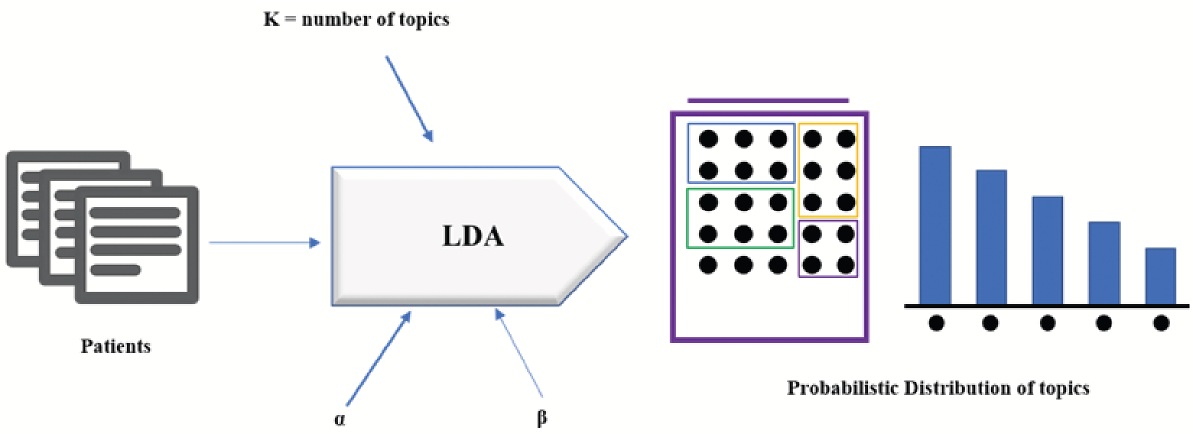

FIGURE 3. Feature extraction of diagnostic text information of patients

2.3. **Efficient and rapid preoperative diagnosis of critical illness.** The purpose of this paper is to construct a low cost, scalable and effective prediction model of critical diseases. Logistic regression is simple, fast, robust and interpretive, and it is an optimization algorithm which is widely used at present. Mainly used in epidemiology, more commonly used cases are to explore the risk factors of a disease, according to risk factors to predict the probability of the occurrence of a disease. The text is based on the textual features and numerical features of the patients with logistic regression fusion to recognize the final preoperative critical disease. The feature linear fusion formula is as follows:

$$f(x) = \sum_{i=0}^{n} \alpha_i x_i + \beta_i x_i \tag{1}$$

where $\alpha_i x_i$ is a linear combination of the numerical types of patients, and $\beta_i x_i$ is a linear combination of the topic probability distribution of the patient textual document.

We put the fusion feature into the logistic function, the higher the calculated score, the greater the critical risk, and the lower the score, the smaller the critical risk. The prediction formula is as follows:

$$h(x) = \frac{1}{1 + e^{-f(x)}} \tag{2}$$

The maximum likelihood is used to train sample dataset for model solving, and the cost function of the prediction model is shown in Formula (3).

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\left(y^{(i)}\log h_\theta\left(x^{(i)}\right) + \left(1-y^{(i)}\right)\log\left(1-h_\theta\left(x^{(i)}\right)\right)\right)\right] \tag{3}$$

Optimizing the cost function by using gradient descent method for model solving, the risk of patients with critical disease can be predicted. Here m is the number of samples, $h_\theta(x)$ is predicted by taking the parameter θ and the sample x, it is the probability that x is positive, y is the label of the sample x.

3. **Experiment and Results.** Experiment was based on the preoperative demographic information, examination information, preoperative diagnosis information of surgical patients from a hospital in China. Risk prediction analysis was carried out based on the data of surgical patients in the hospital for the period from June 2018 to October 2018. A total of 8797 data was collected for surgical patients, including 86 patients with heart failure during and after surgery, and the other without critical disease during and after surgery. The number of patients with heart failure and without complications is 86 and 8711 respectively, which means the distribution of the data is extremely unbalanced. Therefore, we used the resampling method to solve the problem of data imbalance. The resampling method consists of the up-sampling method [31] which increases number of small samples and the down-sampling method [32] which decreases the number of large samples. The sample distribution becomes more balanced so as to improve the recognition accuracy of rare classes. Due to the extreme imbalance of medical data, we considered up-sampling and down-sampling [33] to increase the number of patients with heart failure and reduce the data of patients without complications.

Figure 4 presented the overall flowchart of the experiment. First, we resampled the data of patients. This work increases 2 times number of patients with heart failure, and down-sampling 3% of normal patients. After resampling, the ratio of the number of patients with heart failure and normal patients is approximately 1.5 : 1. Then, 10-fold cross validation was used in the experiment. We randomly shuffled the order of the samples and divided them into 10 equal subsets. Each subset contains both positive and negative samples. The proportion among the number of positive and negative samples was slightly different in each subset. Each subset is then used once as a testing set while the 9 remaining subsets form the training set. The ratio of the number of samples between the training dataset and the testing dataset was about 9 : 1, as shown in step 1. Then, the classification model of numerical heart failure risk was trained based on XGBOOST, and the model validation and numerical feature extraction were performed, as shown in step 2. In the textual model processing, this paper uses the diagnostic textual data of

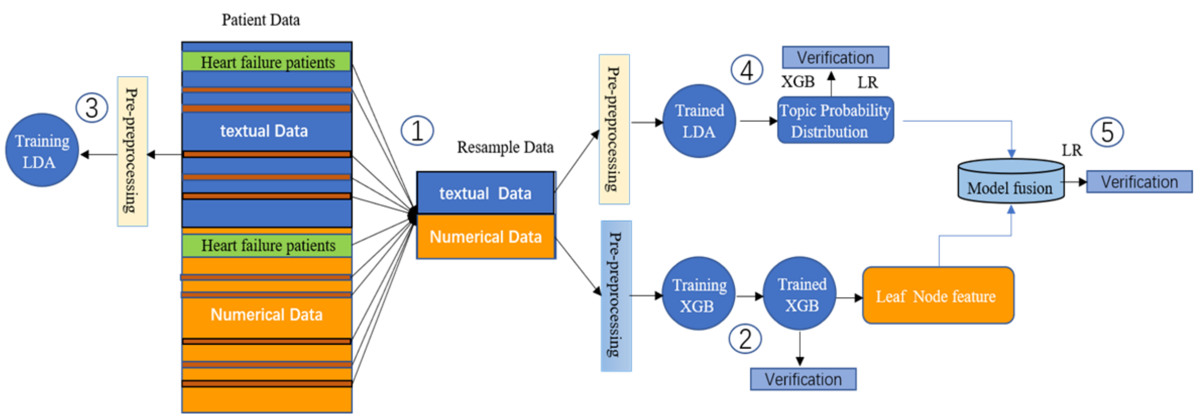

FIGURE 4. Flowchart of experiment

all patients (8797) to firstly carry out topic training, as shown in step 3. Then, based on the diagnostic textual data of the sampled patients, the topic distribution was obtained by the trained LDA model, and the XGBOOST and LR models were used to verify the textual feature heart failure risk prediction, as shown in step 4. Finally, numerical and textual feature models were fused and LR models are used to verify and compare the proposed models.

Because different patients have different numerical test items, in order to reduce the complexity of the model and improve the generalization ability of the model, this paper removes the attribute columns with a missing rate of more than 70%, as well as the numerical feature columns with large correlation and low importance, and finally extracts 56 feature indexes as the input of the numerical model. Table 1 shows the attribute results of preprocessing of patient. The numerical test items consist of physical indices and venous blood test indices. Physical indices include the age, anesthesia mode, weight

Table 1. Numeric and textual attribute results of preprocessing for patient

Index categories		Index name
Numeric	Physical indices	Age, Anesthesia Mode, Body Mass Index (BMI), Diastolic Blood Pressure (DBP), Duration, Systolic Blood Pressure (SBP), Weight
	Venous blood test indices	C-Reactive Protein (CRP), D-Dimer, Gamma-Glutamyl Transferase (GGT), Alanine, Aminotransferase (ALT), Neutrophilic, Granulocyte, Percentage (Neut%), Hepatitis B Virus-e Antibody (HBV-eAb), Hepatitis B Virus-e Antigen (HBV-eAg), Hepatitis B Virus-Core Antibody (HBV-cAb), Hepatitis B Virus-Core Antigen Immunoglobulin M (HBV-IgM), Hepatitis B Virus-Surface Antibody (HBV-sAb), Hepatitis B Virus-Surface Antigen (HBV-sAg), Lactic Dehydrogenase (LDH), Low Density Lipoprotein Cholesterol (LDL-C), International Normalized Ratio (INR), Prothrombin Time (PT), Thrombin Time (TT), Percentage of Monocytes (Mono%), Aspartate Aminotransferase (AST), Urea, Total Cholesterol (TC), Total Bile Acid (TBA), Total Protein (TP), Activated Partial Thromboplastin Time (APTT), Triglyceride (TG), White Blood Cell Count (WBC#), Albumin (ALB), Albumin/Globulin Ratio (ALB/GLO), Direct Bilirubin (DBIL), Alkaline Phosphatase (ALP), Hematocrit Value (Hct), Fibrinogen (Fbg), Fibrinogen Degradation Product (FDP), Creatinine (Crea), Bilirubin (BIL), Cystatin (Cys-C), Adenosine Deaminase (ADA), Glucose (Glu), Platelet Count (PLT#), Hemoglobin (Hb), Retinol Binding Protein (RBP), Superoxide Dismutase (SOD), Ca, Na, K, Indirect Bilirubin (IBIL)
Text		Previous Disease, Past Medical Mistory, Electrocardiogram-Examination, Echocardiography-Examination, Preoperative Clinical Diagnosis

and other indices of patients. Venous blood test indices were the indices tested form the venous blood of patients, such as percentage of monocytes and count of white blood cell. The experiment for text attribute of patients with heart failure used 5 text data to construct the topic distribution of patient documentation training, including the patient previous disease, past medical history, electrocardiogram-examination, echocardiography-examination, preoperative-clinical diagnosis.

In order to build the best performance of numerical feature extraction of patient and critical disease risk prediction model, we used grid search to tune parameters of XGBOOST. Table 2 shows parameters. The risk prediction of heart failure was verified by the training model, and as shown in Table 3, the sensitivity, specificity and AUC of the model are 0.82, 0.92 and 0.91, respectively. It can be observed that the model can accurately identify patients with heart failure, but the sensitivity of the model is not high, which makes it easy to miss diagnosis. In other words, it is easy to distinguish patients with heart failure as normal patients only by preoperative numerical attributes alone. Therefore, in this paper, the sensitivity was improved without reducing the accuracy by integrating the diagnostic text information of patients (see below for details).

Table 2. XGBOOST parameters and its values

Serial number	Parameter name	Description	Values
1	n_estimators	Number of trees to fit	50
2	learning_rate	Boosting learning	0.01
3	max_depth	Maximum tree depth for base learners	5
4	max_delta_step	Maximum delta step we allow each trees weight estimation to be	0.8
5	subsample	Subsample ratio of the training instance	0.7
6	colsample_bytree	Subsample ratio of columns when constructing each tree	0.9
7	colsample_bylevel	Subsample ratio of columns for each level	0.8
8	reg_alpha	L1 regularization term on weights	0.1
9	reg_lambda	L2 regularization term on weights	0.8
10	scale_pos_weight	Balancing of positive and negative weights	0.6
11	min_child_weight	Minimum sum of instance weight (hessian) needed in a child	0

Table 3. Numerical features of patients with heart failure risk prediction results

Method	Sensitivity	Specificity	F1-score	AUC
XGBOOST	0.82	0.92	0.92	0.91

The number of topics is a hyperparameter of LDA model. The number of topics affects the effectiveness of features extracted by LDA model. In the experiment of topic distribution extraction of patient document, in order to determine the optimal number of topics, we used XGBOOST and LR models to evaluate classification performance based on the extracted features of the patient topic probability distribution, the experimental results as shown in Table 4. It is observed that the risk prediction of postoperative heart failure based solely on the patient diagnostic text is not good, but the predictive performance of XGBOOST model is obviously better than LR model. The dataset used is based on the text conclusion of the first 5 examinations of 8797 patients, so it is limited. Additionally, we do not have a professional participle for the medical field, and the noise in text preprocessing is one of the leading causes of poor prediction. The purpose of this paper is to verify whether the fusion of numerical and textual features can improve the accuracy of prediction, so this part does not have too much optimization work. When the number of topics in the experimental results was 30, the classification evaluation indices of XGBOOST and LR did not improve, but decreased when the number of topics increased, so we chose 30 topics as eigenvectors of the patient diagnostic text. See Table 4.

Table 4. Risk prediction results of heart failure for patients with different topics with textual features

Topic num	Method	Sensitivity	Specificity	F1-score	AUC
5	LR	0.26	0.89	0.64	0.78
	XGBOOST	0.65	0.83	0.76	0.81
10	LR	0.58	0.89	0.77	0.85
	XGBOOST	0.55	0.89	0.76	0.88
30	LR	0.55	0.93	0.79	0.91
	XGBOOST	0.71	0.87	0.81	0.91
50	LR	0.52	0.94	0.77	0.88
	XGBOOST	0.71	0.89	0.82	0.88
100	LR	0.23	0.92	0.68	0.8
	XGBOOST	0.61	0.83	0.74	0.83

We combined the features of 606-dimensional leaf nodes extracted by the trained XGBOOST model and the probability distribution of patients with textual topics extracted from the LDA model to predict the risk of heart failure, and the results were as shown in Table 5. The sensitivity, specificity and AUC of the model are 0.90, 0.92 and 0.93, respectively, which have greatly improvement compared with the single-dimensional features of the prediction.

Table 5. The heart failure classification of combining textual and numerical features

Method	Sensitivity	Specificity	F1-score	AUC
XGBOOST+LR	0.90	0.92	0.92	0.93

4. **Discussion.** It can be observed that the combination of numerical and textual data can improve the accuracy and generalization of the model, and is effective to predict the risk of patients with critical disease. Through curve comparison, the AUC of the model that combines numerical and textual information can reach 0.99. Additionally, the method that combines XGBOOST and LR has many advantages. Specifically, the LR model cannot achieve the feature combination, but the feature combination in the model is very important, and it is time-consuming and may not have a good effect by the human experience. XGBOOST mainly mines features and feature combinations, which can automatically find effective features. Each iteration of the model creates a new decision tree on the gradient direction of the residual reduction, and each leaf node corresponds to a path, which is a combination method. Since each path of the tree is calculated by methods such as minimizing mean square error, the obtained path (feature combination) is as distinguished as the human experience. However, the LR model is

simple and efficient, which can be quickly and expansively deployed in practical medical applications. On a whole, the proposed model is better than the single numeric and text-based model in predicting the risk of heart failure.

There are some directions that can be further studied in future work. Firstly, the effect of intraoperative monitoring data on critical events. When analyzing the risk prediction of patients with critical illness, this study mainly considered the patients electronic medical record data. However, the intraoperative operation on patients and postoperative critical complications also has a great impact for the risk prediction of patients with critical illness. Therefore, we can introduce from the integration of electronic medical record data and intraoperative monitoring data in future work, in order for better and more complete prediction of critical adverse events in patients. Perioperative critical illness occurs in different stages and has different effects on patients. At present, this work only classifies and predicts whether critical illness will occur in patients during or after surgery according to preoperative medical examination data of patients, but the time when critical illness will occur in patients is not involved, which is also our future research directions.

5. **Conclusion.** This paper discusses an interesting medical problem. The prediction of critical events in perioperative patients is a complex process, and whether or not critical adverse events occur after surgery depends entirely on the doctor experience and judgment, and the prediction accuracy of long-term and experienced doctors is higher, but the prediction accuracy of doctors with short working hours or inexperience is slightly lower. Additionally, the judgment has lag and the evaluation result cannot be applied directly and effectively, which is a serious problem. Therefore, based on the machine learning method, this paper establishes a prediction model of critical adverse events in patients to predict the risk of critical disease with the preoperative index of any patient. Firstly, the data of preoperative patients are preprocessed, and the patient data is divided into the structural data of numerical data and the unstructured data of textual data. The numerical data of the patient is constructed by the gradient boosting tree model, and the textual features of the patients are extracted by the topic model of the text-based data. Fuse textual features and numerical features, and finally through a simple logistic regression predict the patient critical illness.

Acknowledgment. This work is supported by the National Key Research & Development Plan of China (2018YFC0116704) in data collection. In addition, it is supported by Chongqing Technology Innovation and Application Research and Development Project (cstc2019jscx-msxmx0237) in the design of the study. This study was approved by the Ethics Committee of the First Affiliated Hospital of Army Medical University, PLA, and the Approved No. of ethic committee is KY201936.

REFERENCES

[1] E. H. Lawson, B. L. Hall, R. Louie, D. S. Zingmond, L. Han et al., Association between occurrence of a postoperative complication and readmission: Implications for quality improvement and cost savings, *Annals of Surgery*, vol.258, no.1, pp.10-18, 2013.

[2] R. Pathak, S. Giri, M. R. Aryal, P. Karmacharya, V. R. Bhatt et al., 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, 2015.

[3] S. F. Khuri, W. G. Henderson, R. G. DePalma, C. Mosca, N. A. Healey et al., Determinants of long-term survival after major surgery and the adverse effect of postoperative complications, *Annals of Surgery*, vol.242, no.3, pp.326-341, 2005.

[4] M. 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, *Anesthesia & Analgesia*, vol.112, no.6, pp.1392-1402, 2011.

[5] Z. I. Attia, S. Kapa, F. Lopez-Jimenez, P. M. McKie, D. J. Ladewig et al., Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram, *Nature Medicine*, vol.25, no.1, pp.70-74, 2019.

[6] Y. W. Chen, P. Tang, K. H. Zhong, L. Han, B. L. Qi et al., Semi-supervised surgical workflow recognition based on convolution neural network, *Basic & Clinical Pharmacology & Toxicology*, vol.124, p.52, 2019.

[7] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo et al., A guide to deep learning in healthcare, *Nature Medicine*, vol.25, no.1, pp.24-29, 2019.

[8] O. Gottesman, F. Johansson, M. Komorowski, A. Faisal, D. Sontag et al., Guidelines for reinforcement learning in healthcare, *Nature Medicine*, vol.25, no.1, pp.16-18, 2019.

[9] Y. Gurovich, Y. Hanani, O. Bar, G. Nadav, N. Fleischer et al., Identifying facial phenotypes of genetic disorders using deep learning, *Nature Medicine*, vol.25, no.1, pp.60-64, 2019.

[10] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn et al., Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, *Nature Medicine*, vol.25, no.1, pp.65-69, 2019.

[11] E. J. Topol, High-performance medicine: The convergence of human and artificial intelligence, *Nature Medicine*, vol.25, no.1, pp.44-56, 2019.

[12] R. A. Thuraisingham, A classification system to detect congestive heart failure using second-order difference plot of RR intervals, *Cardiology Research and Practice*, 2009.

[13] Y. Isler and M. Kuntalp, Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure, *Computers in Biology and Medicine*, vol.37, no.10, pp.1502-1510, 2007.

[14] S. N. Yu and M. Y. Lee, Conditional mutual information-based feature selection for congestive heart failure recognition using heart rate variability, *Computer Methods and Programs in Biomedicine*, vol.108, no.1, pp.299-309, 2012.

[15] Z. Masetic and A. Subasi, Congestive heart failure detection using random forest classifier, *Computer Methods and Programs in Biomedicine*, vol.130, pp.54-64, 2016.

[16] P. Melillo, R. Fusco, M. Sansone, M. Bracale and L. Pecchia, Discrimination power of long-term heart rate variability measures for chronic heart failure detection, *Medical & Biological Engineering & Computing*, vol.49, no.1, pp.67-74, 2011.

[17] L. Pecchia, P. Melillo, M. Sansone and M. Bracale, Discrimination power of short-term heart rate variability measures for CHF assessment, *IEEE Trans. Information Technology in Biomedicine*, vol.15, no.1, pp.40-46, 2011.

[18] B. Zheng, J. Zhang, S. W. Yoon, S. S. Lam, M. Khasawneh et al., Predictive modeling of hospital readmissions using metaheuristics and data mining, *Expert Systems with Applications*, vol.42, no.20, pp.7110-7120, 2015.

[19] E. Choi, A. Schuetz, W. F. Stewart and J. Sun, Using recurrent neural network models for early detection of heart failure onset, *Journal of the American Medical Informatics Association*, vol.24, no.2, pp.361-370, 2017.

[20] W. Chen, L. Zheng, K. Li, Q. Wang, G. Liu et al., A novel and effective method for congestive heart failure detection and quantification using dynamic heart rate variability measurement, *PLoS One*, vol.11, no.11, DOI: 10.1371/journal.pone.0165304, 2016.

[21] K. Shameer, K. W. Johnson, A. Yahi, R. Miotto, L. Li et al., Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: A case-study using mount sinai heart failure cohort, *Pacific Symposium on Biocomputing*, Fairmont Orchid, Big Island of Hawaii, pp.276-287, 2016.

[22] B. G. Koerkamp, T. Stijnen, M. C. Weinstein and M. G. M. Hunink, The combined analysis of uncertainty and patient heterogeneity in medical decision models, *Medical Decision Making*, vol.31, no.4, pp.650-661, 2011.

[23] J. Ye, J. Chow, J. Chen and Z. Zheng, Stochastic gradient boosted distributed decision trees, *Proc. of the 18th ACM Conference on Information and Knowledge Management (CIKM'09)*, New York, NY, United States, pp.2061-2064, 2009.

[24] D. M. Blei, A. Y. Ng and M. I. Jordan, Latent Dirichlet allocation, *Journal of Machine Learning Research*, pp.993-1022, 2003.

[25] J. Friedman, T. Hastie and R. Tibshirani, Additive logistic regression: A statistical view of boosting, *The Annals of Statistics*, vol.28, no.2, pp.337-407, 2000.

[26] O. Ronneberger, P. Fischer and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*, Munich, Germany, pp.234-241, 2015.

[27] S. Ren, K. He, R. Girshick and J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.39, no.6, pp.1137-1149, 2017.

[28] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi et al., A survey on deep learning in medical image analysis, *Medical Image Analysis*, vol.42, pp.60-88, 2017.

[29] S. O. Arik, M. Chrzanowski, A. Coates, G. Diamos, A. Gibiansky et al., Deep voice: Real-time neural text-to-speech, *International Conference on Machine Learning*, Sydney, NSW, Australia, pp.195-204, 2017.

[30] T. Chen and C. Guestrin, XGBoost: A scalable tree boosting system, *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, United States, pp.785-794, 2016.

[31] C. Bunkhumpornpat, K. Sinapiromsaran and C. Lursinsap, DBSMOTE: Density-based synthetic minority over-sampling technique, *Applied Intelligence*, vol.36, no.3, pp.664-684, 2012.

[32] M. Kubat and S. Matwin, Addressing the curse of imbalanced training sets: One-sided selection, *International Conference on Machine Learning (ICML)*, Nashville, TN, USA, pp.179-186, 1997.

[33] A. Estabrooks and N. Japkowicz, A mixture-of-experts framework for learning from imbalanced data sets, *Advances in Intelligent Data Analysis*, Cascais, Portugal, pp.34-43, 2001.