# Preoperative Risk Prediction of heart failure with Numerical and Textual Attributes

**Yuwen Chen** ( ✉ chenyuwen@cigit.ac.cn )

Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences; https://orcid.org/0000-0003-4032-5937

**Lige Zhang**

Chinese Academy of sciences,chengdu

**Xiaolin Qin**

Chinese Academy of Sciences,chengdu

**Bin Yi**

Army medical university

# Preoperative Risk Prediction of heart failure with Numerical and Textual Attributes

Yuwen Chen[123*], Lige Zhang[13], Xiaolin Qin[13] ,Bin Yi[4]

*1Chengdu Institute of Computer Applications, Chinese Academy of Sciences, Chengdu, China*

*2Chongqing Institute of Green and Intelligent Technology,*

*Chinese Academy of Sciences, Chongqing, China*

*3University of Chinese Academy of Sciences, Beijing, China*

*4 Department of Anesthesia, Southwest Hospital*

*Army Medical University, Chongqing 400038, P. R. China*

*\*Correspondence: chenyuwen@cigit.ac.cn*

## Abstract

Background: 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.

Methods: 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. Fusion of textual features and numerical features, and finally through a simple logistic regression to predict the patient critical illness.

Results: 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 94% and 97%, thus verifying the feasibility and effectiveness of the model.

Conclusions: 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 model respectively, and builds a low cost, scalable and effective model of predicting preoperative critical disease by using logistic regression. The experimental demonstrate show that it is better to predict whether the patients have heart failure with Numerical and Textual Attributes of patients

Keywords: Critical illness, LDA model, preoperative risk prediction, XGOOST

## Background

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. At present, the incidence of adverse events of various critical diseases during the perioperative period in China is as high as 12%, resulting in inpatients with a maximum fatality rate of 1.1%. The "aura ratings" critical adverse events early warning system often cannot catch signs of critical adverse events in time. Therefore, once the critical adverse events occur, the disease is heavier or terminal, resulting in the difficult treatment and the limited effect of intervention. 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 impact on the long-term survival and quality of life of patients[4]. Therefore, it is great scientific significance and social value to actively carry out the risk prediction of critical adverse events, which is helpful to the early detection, early warning, diagnosis and intervention cues of critical adverse events.

Recently, artificial intelligence technology has been widely used in the medical field [5-11]. The

medical expert system has experienced the problem of modeling uncertainty to the wide application of Bayesian network [12].The diabetes database was analyzed by using CART analysis method, and the prediction model was established by using the simple Bayesian method, IBI method and C4.5 method, and the accuracy can reach 79%[13].In addition, for the prediction of critical events, the study reported that 5 vital signs of the patient bedside data were collected and analyzed by Visensia physiological monitoring software, and the deterioration of the index could be an early warning of cardiac respiratory arrest [14].The early warning system score of early signs of clinical deterioration of diseases can be used to warn against cardiovascular events such as vascular pressurization, heart failure and cardiac arrest within 48 hours [15]。Thrombolytic Therapy in Acute Myocardial Infarction (TIMI) score and the Global Registry of Acute Coronary Events (GRACE) score can predict postoperative cardiovascular adverse events in patients with emergency chest pain[16]；inflammatory factors, blood lipids and serum uric acid can be used as indices for the evaluation and prediction of postoperative cardiovascular adverse events in diabetic patients [17, 18].

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[19] as well as various types of numerical laboratory data, textual diagnostic information. These indices are used to predict perioperative critical illness, the results are more practical and targeted. Therefore, in order to make the results of research and analysis more accurate and more convenient to apply to practice, in this paper, we use machine learning method to model the risk prediction of critical illness during and after operation based on the preoperative medical data of patients, so as to construct a scalable, low-cost and effective prediction solution for critical events.

## 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[20],the topic model[21]and logistic regression[22]. From the medical 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.). Using 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 [23-26] 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 patient text, 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 below.
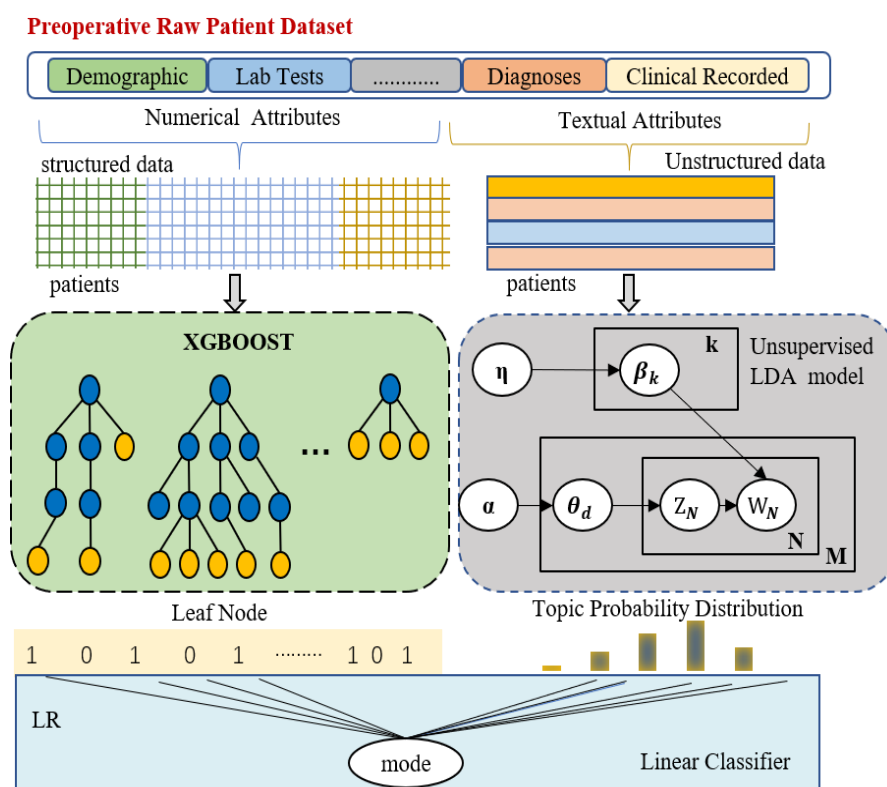


Fig. 1 Conceptual framework for predicting critical illness models.

**Numerical feature extraction of perioperative patients**

XGBOOST[27] is used to extract the numerical features of patients, which is an integrated learning method proposed by Tianqi Chen based on GBDT[20]. 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, 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. 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].
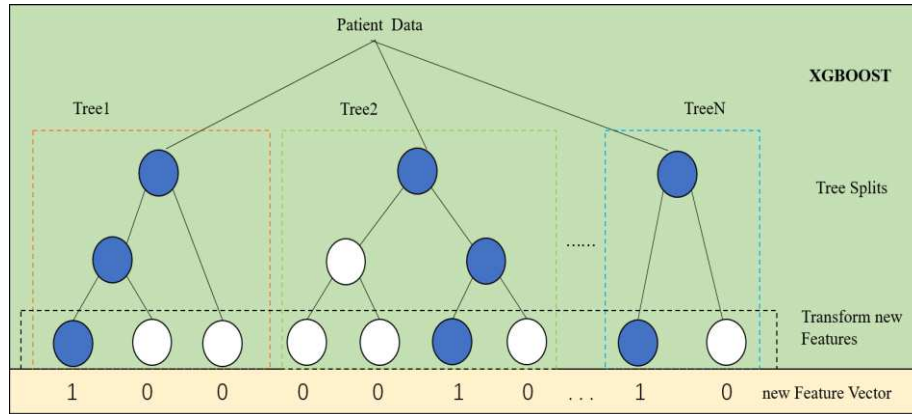
Fig.2 Patient numerical feature extraction

**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, 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. Through the topic model to cluster the patient documentation, each patient medical diagnostic information can be through the topic probability distribution of the document to represent with Dimensionality reduction. As shown in Figure 3. Topic model based on Latent Dirichlet Allocation(LDA) [21] which was proposed by Blei David 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].
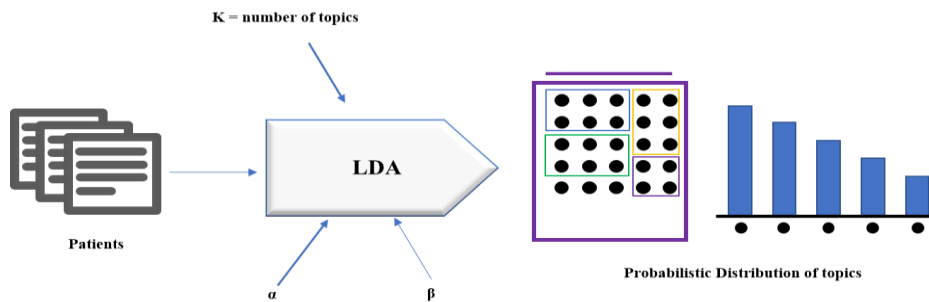


Fig.3 Feature extraction of diagnostic text information of patients

**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, $\beta_i x_i$ 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 9.

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

Optimizing the cost function by using gradient descent method for model solving, the risk of patients with critical disease can be predicted. Where m is the number of samples, $h_\theta(x)$ is y predicted by taking the parameter $\theta$ and x, y is y in the original training sample.

**Experiment**

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 distribution is shown in Figure 4, it is extremely unbalanced. Therefore, we used the method of resampling to solve the problem of data imbalance. The resampling is the up-sampling[28] that increases number of small samples and the down-sampling[[29] that 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 [30] to increase the number of patients with heart failure and reduce the data of patients without complications.
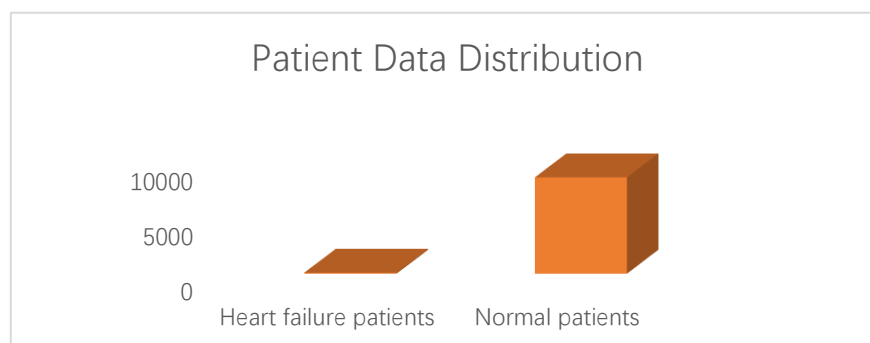


Fig.4 Patient data distribution (The number of patients with heart failure is 86, and the number of patients without complications is 8711)

Figure 5 presented the overall flowchart of the experiment. First, we resampled patient data to build the training set and validation set of the model. This work increases 2 times number of patients with heart failure, and down-sampling 0.03% of normal patients. 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 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 are fused and LR models are used to verify and compare the proposed models.
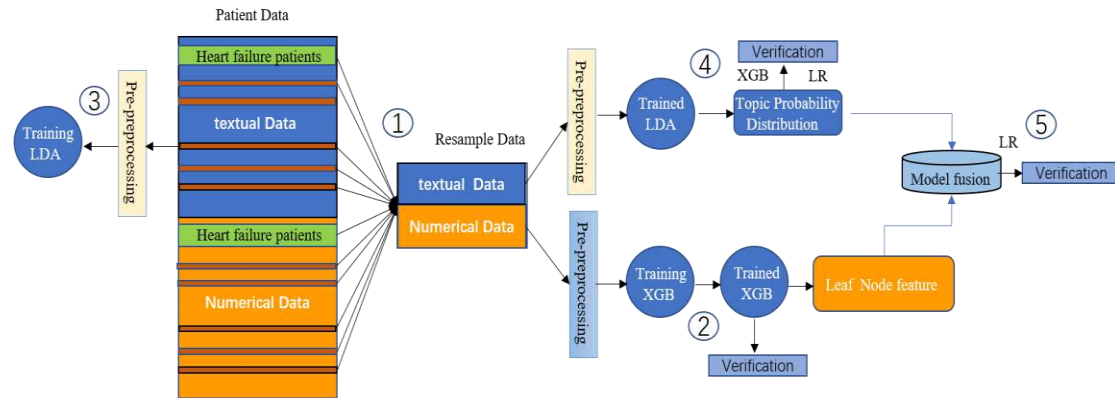


Fig.5 Flow chart of experiment

## Results

Because the 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. 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. Table 1 shows the attribute results of preprocessed of patient

Table .1 Numeric and Textual attribute results of preprocessed

| Index Categories | Index name |
| --- | --- |
| Numeric | Age, Anesthesia Mode, Body Mass Index(BMI), Diastolic Blood Pressure(DBP), Duration, Systolic Blood Pressure(SBP), Weight<br><br>C-Reactive Protein(CRP), D-Dimer, Gamma-Glutamyltransferase (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, 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 |
|------|-----------------------------------------------------------------------------------------------------------------------------------------|

In order to build the best performance of numerical feature extraction of patient and critical disease risk prediction model, we used gridsearch to tune parameters of XGBOOST. Table 2 is shown 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, 0.98 and 0.96, 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 | 0.98 | 0.92 | 0.96 |

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.

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 | 0.96 | 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 | 0.98 | 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 0.94,0.95 and 0.99, respectively, which have greatly improvement comparing 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 | 0.94 | 0.97 | 0.96 | 0.99 |

## Discussion

It can be observed that the combination of numerical and textual data can improve the accuracy and generalization of the model as a whole, and is effective to predict the risk of patient 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, 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 to better and more complete prediction of critical adverse events in patients. 2) 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.

## 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. Fusion of textual features and numerical features, and finally through a simple logistic regression to predict the patient critical illness.

**Abbreviations**

Latent Dirichlet Allocation                    LDA
Extreme Gradient Boosting                   XGBOOST
Logistic Regression                             LR


## Declarations

## Ethics approval and consent to participate
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.

## Consent for publication
Not applicable.

## Availability of data and materials
Not applicable.

## Reference

1. Lawson, E.H., et al., *Association between occurrence of a postoperative complication and readmission: implications for quality improvement and cost savings.* 2013. **258**(1): p. 10-18.

2. Pathak, R., et al., *Mortality, length of stay, and health care costs of febrile neutropenia-related hospitalizations among patients with breast cancer in the United States.* 2015. **23**(3): p. 615-617.

3. Khuri, S.F., et al., *Determinants of long-term survival after major surgery and the adverse effect of postoperative complications.* 2005. **242**(3): p. 326-341.

4. Hamilton, M.A., et al., *A systematic review and meta-analysis on the use of preemptive hemodynamic intervention to improve postoperative outcomes in moderate and high-risk surgical patients.* 2011. **112**(6): p. 1392-1402.

5. Attia, Z.I., et al., *Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram.* 2019. **25**(1): p. 70-74.

6. Chen, Y.W., et al., *Semi-Supervised Surgical Workflow Recognition Based on Convolution Neural Network.* Basic & Clinical Pharmacology & Toxicology, 2019. **124**: p. 52-52.

7. Esteva, A., et al., *A guide to deep learning in healthcare.* 2019. **25**(1): p. 24-29.

8. Gottesman, O., et al., *Guidelines for reinforcement learning in healthcare.* 2019. **25**(1): p. 16-18.

9. Gurovich, Y., et al., *Identifying facial phenotypes of genetic disorders using deep learning.* 2019. **25**(1): p. 60-64.

10. Hannun, A., et al., *Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network.* 2019. **25**(1): p. 65-69.

11. Topol, E.J.J.N.M., *High-performance medicine: the convergence of human and artificial intelligence.* 2019. **25**(1): p. 44-56.

12. Başçiftçi, F. and E.J.C.M.P.B. Avuçlu, *An expert system design to diagnose cancer by using a new method reduced rule base.* 2018. **157**: p. 113-120.

13. Wang, Q., et al., *Two-Point Discrimination Predicts Pain Relief after Lower Limb Nerve Decompression for Painful Diabetic Peripheral Neuropathy.* 2018. **141**(3): p. 397e.

14. Adrian, F. and E.J.N.M. Naomi, *Early warning scores: a sign of deterioration in patients and systems.* 2015. **22**(1): p. 26-31.

15. Smith, M.E., et al., *Early warning system scores for clinical deterioration in hospitalized patients: a systematic review.* 2014. **11**(9): p. 1454.

16. Takashima, H., et al., *Correlation between T2 relaxation time and intervertebral disk degeneration.* 2012. **41**(2): p. 163-167.

17. Morelli, V., et al., *Cardiovascular events in patients with mild autonomous cortisol secretion: analysis with artificial neural networks.* 2017. **177**(1).

18. Ursini, F., et al., *Prevalence of Undiagnosed Diabetes in Rheumatoid Arthritis: an OGTT Study.* 2016. **95**(7): p. e2552.

19. Koerkamp, B.G., et al., *The Combined Analysis of Uncertainty and Patient Heterogeneity in Medical Decision Models.* 2011. **31**(4): p. 650-661.

20. Ye, J., et al. *Stochastic gradient boosted distributed decision trees.* in *conference on information and knowledge management.* 2009.

21. Blei, D.M., A.Y. Ng, and M.I.J.J.o.M.L.R. Jordan, *Latent dirichlet allocation.* 2003. **3**: p. 993-1022.

22. Friedman, J.H., T. Hastie, and R.J.A.o.S. Tibshirani, *Additive logistic regression : A statistical view of boosting.* 2000. **28**(2): p. 337-407.

23. Ronneberger, O., et al., *U-Net: Convolutional Networks for Biomedical Image Segmentation.* 2015: p. 234-241.

24. Ren, S., et al., *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.* 2017. **39**(6): p. 1137-1149.

25. Litjens, G.J.S., et al., *A survey on deep learning in medical image analysis.* 2017. **42**: p. 60-88.

26. Arik, S.O., et al., *Deep Voice: Real-time Neural Text-to-Speech.* 2017: p. 195-204.

27. Chen, T., C.J.k.d. Guestrin, and d. mining, *XGBoost: A Scalable Tree Boosting System.* 2016: p. 785-794.

28. Bunkhumpornpat, C., K. Sinapiromsaran, and C.J.A.I. Lursinsap, *DBSMOTE: Density-Based Synthetic Minority Over-sampling TEchnique.* 2012. **36**(3): p. 664-684.

29. Kubat, M. and S. Matwin. *Addressing the Curse of Imbalanced Training Sets: One-Sided Selection.* in *international conference on machine learning.* 1997.

30. Estabrooks, A. and N. Japkowicz, *A Mixture-of-Experts Framework for Learning from Imbalanced Data Sets.* 2001.

Fig. 1 Conceptual framework for predicting critical illness models.

Fig.2 Patient numerical feature extraction

Fig.3 Feature extraction of diagnostic text information of patients

Fig.4 Patient data distribution

Fig.5 Flow chart of experiment