RESEARCH

A sample article title

Yulong Chen[1*] and John R.S. Smith[1,2]

[*]Correspondence:
jane.e.doe@cambridge.co.uk
[1]Department of Science, University of Cambridge, London, UK
Full list of author information is available at the end of the article

Abstract

Background: Early prediction of sepsis infection can help intensive care unit (ICU) physicians make the best clinical decisions to reduce their patients' risk of sepsis infection. The aim of this study is to develop Stacking-Based Temporal GRU model to predict whether a patient is at risk of sepsis infection within a certain time frame.

Methods: In this study, we used a dataset containing information on approximately 48,000 patients with 36 registrations per patient per hour. The input data for the model was obtained by performing a series of operations on the dataset, such as feature variable filtering, null value filling, feature construction, and sample equalization. Then, we performed fusion experiments using the input data for the stacking ensemble learning method to filter the best combination of base models, and four common machine learning models were chosen for the fusion of base models. Finally, we fused the models obtained by stacking ensemble learning method with the two-layer GRU model by entropy weighting method to obtain our Stacking-Based Temporal GRU model, and compared it with the mainstream machine learning and deep learning models. Three common classifications of evaluation metrics were chosen: area under the curve (AUC), f1 score (F1) and accuracy score (Accuracy).

Results: The experimental results are divided into two parts, the first part is stacking ensemble learning method, we choose the best combination of base model is extreme gradient boosting (XGB), random forest (RF) and categoricalboosting (CatBoost), the metamodel we choose is logistic regression (LR). this combination of model scores 0.898, 0.751 and 0.831 under three evaluation metrics, the best performance among all combinations. Another part is the performance comparison of early sepsis prediction models, comparing the selected mainstream machine learning and deep learning models, our Stacking-Based Temporal model has the best scores in all three evaluation metrics, in order of 0.908, 0.752 and 0.819.

Conclusion: The Stacking-Based Temporal model developed in this study has good early predictive performance for sepsis. It has the potential to help ICU physicians to reduce the risk of patients contracting sepsis by early pharmacological treatment and clinical interventions for patients who are likely to contract sepsis.

Keywords: Intensive care unit; Sepsis; Stacking ensemble learning; Deep learning; Early prediction model

Background

Sepsis is a syndrome of physiological, pathological and biochemical abnormalities caused by infection and its life-threatening organ dysfunction. Due to the heterogeneity of the condition, it poses great difficulties at the level of identification,

treatment and research [1, 2]. It is estimated that sepsis affects more than 30 million people per year worldwide and may cause more than 6 million deaths per year. According to data from intensive care units in the United States and Europe in 2012, the mortality rate of sepsis was approximately 41% in Europe and 28.3% in the United States [3]. Specifically, the most common cause of in-hospital death in the United States is sepsis, with more than 24 billion spent annually on sepsis prevention and treatment [4]. A study conducted in Australia and New Zealand showed that mortality from sepsis decreased from 35% to 20% from 2000-2012, but the incidence continued to increase [5]. Therefore, early prediction, detection and management of sepsis increases the likelihood of survival of patients with sepsis and can effectively reduce the incidence of sepsis [6, 7].

Currently, several hospitals use Sequential Organ Failure Assessment (SOFA) for clinical assessment of sepsis, and this scale is recommended by The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) [2]. In addition, other authors have proposed QuickSOFA as a rapid bedside quick score to identify patients with suspected infection and high risk of death. The sepsis-3 requires monitoring of the patient's vital signs (VS) such as heart rate, oxygen saturation, arterial pressure, respiratory rate, and laboratory test results. These vital signs are considered as characteristic variables to calculate the sepsis score. Therefore, monitoring patient variables is essential for the prevention and treatment of sepsis and improving the research process of sepsis [8, 9].

In recent years, a number of studies have used patient vital sign (VS) data and other characteristic data for early prediction of sepsis. Marcio et al. [10] constructed a kinematic model based on the original vital sign variables and obtained the kinematic characteristics (KF) of the patient by calculating the location of sepsis spatially. When VS and KF were used together as input variables for neural network model prediction, the accuracy was higher compared to using VS data alone as input.

Manaktala et al. [11] constructed a triple combination of change management, electronic monitoring, and algorithms that provided highly sensitive and specific decision support to the point of care, providing recommendations for early goal-directed therapy that led to a significant reduction in sepsis deaths.

Sidney et al. [12] used electronic health record (EHR) data from a cohort of inpatient and emergency department visits of children aged 2-17 years at the University of California, San Francisco Medical Center to predict the severity of pediatric patients using machine learning algorithms with significant effects due to the Moderate Logical Organ Dysfunction Score (PELOD-2) and the Moderate Systemic Inflammatory Response Syndrome (SiRS), suggesting that machine learning algorithms can be useful for the detection and prevention of sepsis severity. Guilan et.al [13] used machine learning algorithms of least absolute shrinkage and selection operator (LASSO), random forest (RF), gradient boosting machine (GBM) and traditional logistic regression (LR) methods to develop prediction models and compared the prediction performance of the developed prediction models with simplified acute physiology score (SAPS) and found that the machine learning based models have good sepsis prediction performance, and thus may help improve the prognoses of sepsis patients in the ICU.

Xiang et al. [14] proposed a Time-phased machine learning model for sepsis prediction, which was able to predict sepsis conditions in intensive care units in real time. Akram et al. [15] used a support vector machine (SVM) to classify 29,552 adult patients from five regional hospitals and found that physiological indicators of sepsis patients were significantly differentially expressed over time, and artificial intelligence models can be developed to predict sepsis patients using time series.

Lucas et al. [16] and Michael et al. [17] used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) and PRISMA guidelines, respectively, to compile a large number of models from the literature. They found that machine learning is increasingly used for early prediction of sepsis, and that machine learning models can accurately predict sepsis episodes in advance.

Hye et al. [18] further improved the performance of sepsis prediction models by learning temporal patterns of time series data through LSTM deep learning models. Dongdong et al. [19] proposed a deep learning model with temporal encoding to predict sepsis using patients' electronic health records (EHR), confirming that the deep learning model provided high accuracy and transparency and clinical interpretability.

Tunc et al. [20] first used a regression-based analysis to screen data on seven vital signs of patients in ICU beds and then used an algorithm called the Deep SOFA Sepsis Prediction Algorithm (DSPA) to predict SOFA scores in septic patients to measure the severity of organ failure in septic patients.

Gavin et al. [21] provided a statistically significant improvement to machine learning modeling applications by proposing a new deep learning application for effective prediction of sepsis six hours prior to onset, including a boosted cascading training methodology and adjustable margin hinge loss fuction.

All of the above prediction models show that both machine learning models and deep learning models can achieve better performance in sepsis prediction compared to the traditional sepsis scoring criteria. In addition, processing of time series variables and patient vital sign variables can further improve model performance. Although these studies have yielded good results, they have limitations, such as deep learning models do not perform as well as machine learning models on small data sets, and machine learning models do not learn the variation of variables over time series as well as deep learning models. Therefore, we developed A Stacking-Based Temporal GRU for sepsis prediction model. The first level of the whole prediction model consists of random forest (RF), categoricalboosting (CatBoost) and extreme gradient boosting (XGB) fused by an integrated learning method, the second level model is composed of a deep learning model double-layer GRU, and finally these two levels are finally fused by an entropy weighting method to obtain our sepsis prediction model. The innovation is that we fuse the machine learning model using the integrated learning approach as the first level model, and then fuse the first level with the deep learning model in the second depth, which combines the advantages of machine learning model and deep learning model and enables the model to handle different types of data sets, and also considers the VS behavior of patient time series. We validate the higher accuracy compared to the unilateral use of machine learning models and deep learning models. We also compared with a sepsis prediction neural network model that introduces kinematic features [10] and found

higher performance of our model, further confirming the accuracy and completeness of our model.

Methods

Data source

As a data source, we used the clinical multivariate timeseries database published in the Early Prediction of Sepsis From Clinical Data: The PhysioNet Computing in Cardiology Challenge 2019 [22–24]. The dataset source data for 40,336 patients, saved separately in individual text files, with each line in the individual text file representing the status of each variable for the patient at the current hour. Available patient variables include demographics, vital signs (VS), and laboratory values, for a total of more than 40 characteristic variables, with each patient having a different time series length.

Data preprocessing

Since each patient in the original dataset may or may not be ill in the entire time series, we first defined the SepsisLabel variable (For sepsis patients: 1, For non-sepsis patients: 0) to be non-sepsis patients if all of them are 0 in the entire time series, and sepsis patients whenever 1 occurs, and redefined a Label variable (For sepsis patients: 1, For non-sepsis patients: 0) indicates whether the patient is a patient in the whole sequence. The number of patients was obtained as 2932 and the number of non-patients as 37404. Secondly, we only kept the data of patients with more than 36 hours of recording and removed the characteristic variables with null values greater than 82% in these patient data. The final data showed a total number of patients of 24,071, of which the number of patients was 15,08 and the number of non-patients was 22,563. Considering that even after removing the feature variables with missing values greater than 82%, there may still exist variables that are not very helpful for model prediction, we used the CatBoost model [25] to filter the remaining variables and took the variables with the top 9 scores as predictor variables, and the variable names and their values are introduced as shown in Table 1.

Table 1 Predictor variables remaining after feature variable screening

Varible	Unit	Positive	Negative
Number	#	1508	3016
MAP	mm Hg	80.64 ± 59.46	82.23 ± 62.23
Temp	°C	36.98 ± 4.41	36.97 ± 16.07
O2Sat	%	97.46 ± 66.46	97.25 ± 77.25
SBP	mm Hg	121.79 ± 147.20	123.35 ± 174.64
HR	beats/min	86.84 ± 124.15	84.62 ± 195.37
DBP	mm Hg	62.08 ± 185.91	63.67 ± 234.32
Resp	beats/min	19.13 ± 51.86	18.60 ± 81.39
HospAdmTime	h	-75.48 ± 76.27	-53.52 ± 5313.33
Age	years	62.35 ± 37.64	62.39 ± 37.60

After the feature filtering was performed, the null values under each variable obtained from the filtering were filled by means of mean filling. The mean value was selected as the mean value of all patient data, which is more effective compared to using the mean value of individual patient data, and the data were normalized after mean filling, and the sample data obtained are shown in Table 2.

Next, we used feature construction to expand the data set by first obtaining the mean (mean), standard deviation (std), 25th percentile, 50th percentile, 75th percentile, and maximum (max) corresponding to each patient's nine feature variables,

Table 2 Mean-filled sample data (selected from a single patient for three consecutive hours)

MAP	Temp	O2Sat	SBP	HR	DBP	Resp	HospAdmTime	Age
0.02	0.70	1.00	0.44	0.27	0.23	0.30	0.99	0.51
0.22	0.70	0.97	0.32	0.24	0.26	0.23	0.99	0.51
0.24	0.70	1.00	0.36	0.28	0.28	0.29	0.99	0.51

which were added to the data set as new features, and the sample data obtained after feature construction are shown in Table 3.

Table 3 Mean-filled sample data (selected from a single patient for three consecutive hours)

MAP	MAP_mean	MAP_std	MAP_25%	MAP_50%	MAP_75%	MAP_max
0.02	0.36	0.36	0.33	0.37	0.38	0.17
0.22	0.34	0.24	0.35	0.34	0.30	0.15
0.24	0.58	0.38	0.52	0.60	0.57	0.28

Finally, to ensure a uniform distribution of positive and negative samples (positive samples are patients and negative samples are non-patients), we randomly selected 4524 from 22,563 non-sepsis patients as negative samples to achieve a patient to non-patient ratio of 1:3, and the entire data preprocessing process is shown in Figure 1.

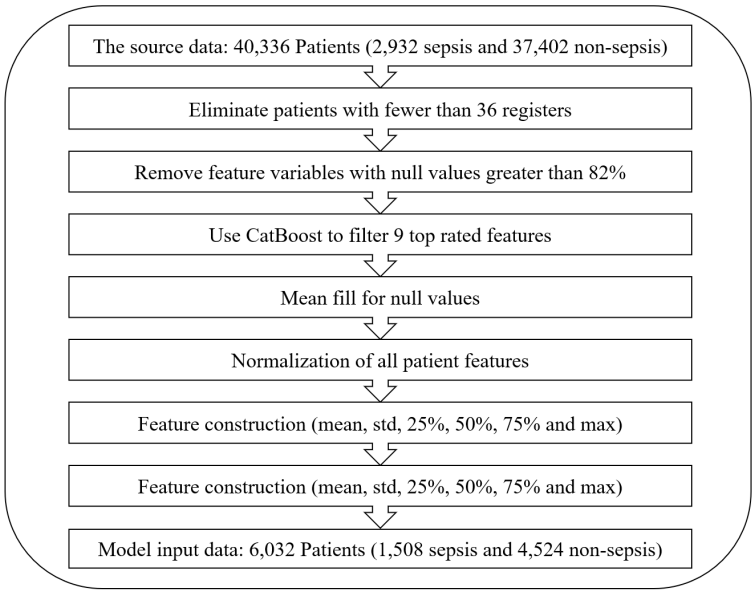

Figure 1 Preprocessing flowchart

Stacking ensemble learning method

The stacking ensemble learning method mixes many machine learning model types through some sort of fusion, utilizing the benefits of various machine learning models to analyze data from various angles. As a result, the first layer's base learner must choose a variety of models in addition to high-performing ones. To repair the faults of the base learner in the first layer and get the best predicting, the meta-learner in the second layer should choose models with strong generalization potential. CatBoost is an innovative algorithm for processing classification features

that solves the problem of prediction bias that can result from gradient enhancement algorithms [25]. The random forest (RF) model, which can successfully avoid overfitting and is immune to interference and parallel processing, is a model integrated by bagging [26,27]; Extreme gradient boosting (XGB) integrated by boosting with a regularization term to prevent overfitting, with high computational efficiency and multi-threaded operation [28]; The lightgbm model (LGB) is a Histogram-based decision number algorithm that allows the information gain from large gradient samples to be amplified while retaining large gradient samples [29]; Logistic regression (LR) is a widely used classification algorithm that is widely accepted in the field of statistics [30]. In summary, we consider the combination of CatBoost, RF, XGB, and LGB in different combinations as the base learner of the Stacking ensemble learning method, and LR models as a meta-learning model, using a five-fold cross-validation approach for Stacking integration, and the integrated model is called the first-level model of the overall prediction model. We created trials to individually forecast the various combinations and assessed the prediction outcomes (the evaluation metrics we used were Accuracy, F1 and AUC in Evaluation). The finally selected optimal base model combination is RF, CatBoost and XGB, the model framework is shown in Figure 2.

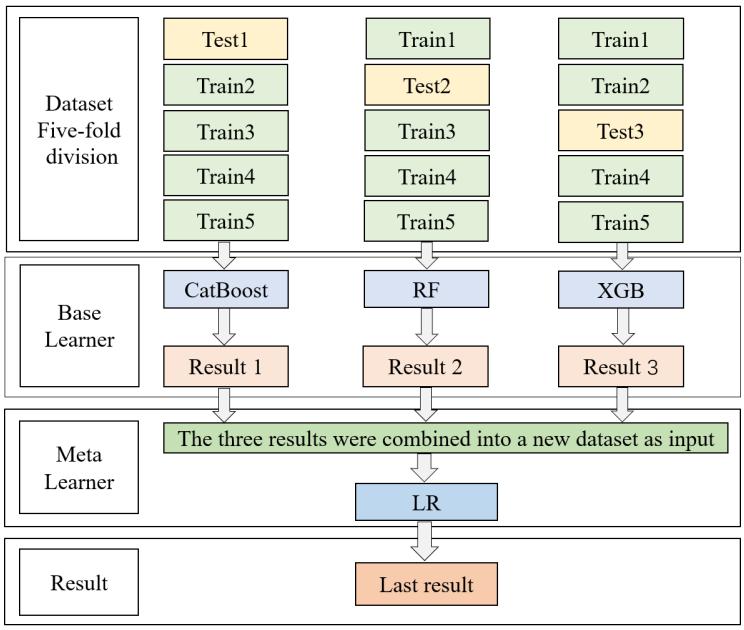

Figure 2 The model framework obtained after Stacking ensemble learning method

Two-layer GRU model

The Gated Recurrent Unit (GRU) is a variant of recurrent neural networks that Junyoung et al. first proposed in 2014 and used in sequence studies [31]. It shares the same gating unit as the LSTM to minimize information loss during the propagation of long sequences (long-term) through the gating unit, but it lacks a recurrent memory unit, making it less complex than the LSTM. In a different research, Cho

demonstrated that GRU outperforms LSTM with a sufficient dataset and is quicker than LSTM [32]. The mathematical expressions can be denoted as

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \tag{1}$$

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \tag{2}$$

$$\tilde{h} = tanh(W[r_t \odot h_{t-1}, x_t]) \tag{3}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{4}$$

In this paper, we have used the two-layer GRU model as a second level for the overall prediction model and we can see in Figure 3 how the two-layer GRU model works when we input x_t.

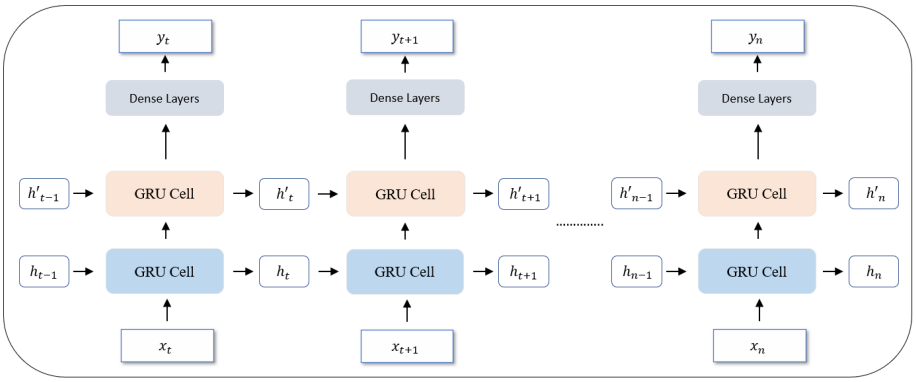

Figure 3 Two-tier GRU model framework (the second level of the overall prediction model)

Stacking-Based Temporal GRU model

The model obtained by the Stacking integration method after filtering the base model is used as the first level model of the sepsis early prediction model, and the two-layer GRU model is used as the second level model of the sepsis prediction model, and the two levels of models are combined to form our overall prediction model. The overall model framework is shown in Figure 3. The weights for the entropy weighting fusion are determined according to the entropy weighting method [33]. The mathematical expressions can be denoted as

$$x'_{ij} = \frac{|x_{ij} - min(x_j)|}{max(x_j) - min(x_j)} \tag{5}$$

$$e_j = -\frac{1}{ln(n)} \sum_{i=1}^{n} \frac{x'_{ij}}{\sum_{i=1}^{n} x'_{ij}} ln(\frac{x'_{ij}}{\sum_{i=1}^{m} x'_{ij}}), j = 1, ..., m \tag{6}$$

$$\omega_j = \frac{1 - e_j}{\sum_{j=1}^{m}(1 - e_j)}, j = 1, ..., m \tag{7}$$

$$s_j = \frac{\sum_{j=1}^{m} \omega_j x_{ij}}{\sum_{j=1}^{m} \omega_j}, i = 1, ...n \tag{8}$$

x_{ij} refers to the ith prediction result of the jth model, and x'_{ij} is the result obtained after normalizing x_{ij}. i ranges from 1 n and j ranges from 1 m.e_j represents the information entropy of the jth model, ω_j represents the weight of the jth model, and s_j represents the result obtained after entropy weighting fusion.

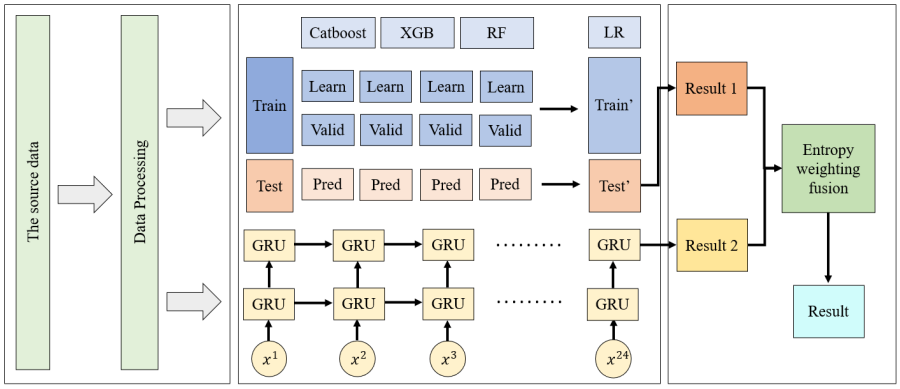

Figure 4 Stacking-Based Temporal GRU model framework (Train: original training data, Test: original test data, Learn: training set divided in five-fold cross-validation, Valid: validation set divided in five-fold cross-validation, Pred: prediction result of the training set, Train': new training data obtained by combining the prediction result of the base model and the original training data)

Evaluation

This article selects Accuracy,F1 and AUC to measure the performance of the different models. They are denoted by

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{10}$$

Where: TP means predicted 1, actual 1, predicted correctly. FP means predicted 1, actual 0, predicted incorrectly. FN means predicted 0, actual 1, predicted incorrectly. TN means predicted 0, actual 0, predicted correctly. $precision = TP/(TP + FP)$, $recall = TP/(TP + FN)$. $FPR = TP/(TP + FN)$, $TPR = TP/(FP + TN)$. AUC refers to the area below the ROC curve consisting of FPR as the $x - axis$ and TPR as the $y - axis$.

Results

Filter the base model in Stacking ensemble method

We use the four machine learning models XGB, LGB, RF, and CatBoost as the base models for the Stacking ensemble method, and then take LR as the meta model.

To determine the ideal hyperparameter values for each of the four machine learning models, we employed the grid parameter search technique offered by Scikit-Learn. We can see from the findings in Table1 that the base model of XGB, CatBoost and RF combined with LR as the meta-model beats the other base and meta model combinations. As a result, we decided to employ the model created utilizing the aforementioned integration techniques as the first level model for our overall prediction model.

Table 4 Mean-filled sample data (Scoring results of Stacking fusion with different base models and LR combinations

Base model	AUC	F1	Accuracy
RF+Catboost	0.884	0.791	0.793
RF+XGB	0.894	0.761	0.737
RF+LGB	0.861	0.681	0.818
catboost+XGB	0.897	0.681	0.731
catboost+LGB	0.863	0.729	0.787
XGB+LGB	0.881	0.742	0.815
RF+Catboost+LGB	0.875	0.701	0.762
RF+XGB+LGB	0.886	0.721	0.782
CatBoost+XGB+LGB	0.888	0.736	0.795
XGB+LGB+CatBoost+RF	0.891	0.733	0.796
RF+CatBoost_XGB	0.898	0.751	0.831

Performance of different models on the sepsis early prediction problem

We compare the evaluation results of the Stacking-Based Temporal GRU model constructed with XGB, LGB, the first-level model constructed in this paper through the Stacking ensemble method, Bi-GRU and Bi-LSTM for the sepsis early prediction problem. It is important to note that the two-layer LSTM and two-layer GRU models we selected are both two-layer unidirectional models since, according to our tests, they outperform all other two-layer models. In addition, the parameter values for the XGB model, the LGB model, and the crucial parameter settings for the GRU and LSTM are displayed in Table5.

Table 5 Mean-filled sample data (Important hyperparameter settings for GRU and LSTM

Hyperparameter	LSTM	GRU
Sequence Length	36	36
Epochs	100	100
Batch_size	64	64

The final evaluation findings for the various models on the sepsis prediction are provided in Table 6. In terms of AUC, Accuracy and F1 scores, our Stacking-Based Temporal GRU model performs better than the other models, notably in the F1 and Accuracy assessment metrics, which are noticeably higher than the other models, reaching 0.752 and 0.819 respectively. This shows that our model performs well when used to predict sepsis.

Discussion

In this study, sepsis patient data were first extracted from the Early Prediction of Sepsis From Clinical Data: The PhysioNet Computing in Cardiology Challenge 2019 dataset. Then, We apply a series of pre-processing operations to the source data to obtain the input data for the early prediction model. Finally, we propose the Stacking-Based Temporal GRU model and compare the performance with other mainstream models under three evaluation metrics.

Table 6 Comparison of the evaluation results of various models applied in the early prediction of sepsis

Models	AUC	F1	Accuracy
RF	0.871	0.713	0.773
XGB	0.887	0.744	0.813
LGB	0.833	0.663	0.701
CatBoost	0.874	0.694	0.796
Stacking(best)	0.898	0.742	0.815
Bi-GRU [34]	0.888	0.719	0.771
Bi-LSTM [35]	0.884	0.705	0.742
KANNEDS [10]	0.894	0.731	0.798
STGEDS	0.908	0.752	0.819

In the data preprocessing section, firstly, in consideration of the fact that different time series measured for each patient may have an effect on the effectiveness of the deep learning model, we selected the first 36 hours of data for each patient and removed those patients with less than 36 hours of time series length. Secondly, since too many null values in the feature variables of some patients would affect the classification effect of the model, we removed the feature variables with null values accounting for more than 82%. Thirdly, we performed feature variable screening by the Catboost model, and finally kept the top 9 feature variables with the highest model scores to avoid the interference of unnecessary variables on the model performance, and the null values are filled using the mean value. Besides, to improve the model effect further, we performed feature construction and data normalization operations. Specifically, feature construction is variable expansion using the mean, standard deviation, 25% quantile, 50% quantile, 75% quantile, and maximum value of each feature variable. Finally, in order to make the positive and negative samples more balanced, we randomly selected some samples from the negative samples to make the proportion of positive and negative samples 1:3.

The stacking ensemble learning method [36] tends to perform better by using different machine learning models to learn different features of the overall data, and after fusion. We used different combinations of LGB, XGB, RF and CatBoost models as the base model for the stacking ensemble learning method. After a series of experiments we finally selected a combination of base models suitable for early prediction of sepsis, which are RF, XGB and CatBoost. Stacking fusion of RF, XGB and CatBoost as the base model and LR as the meta-model constitutes the first level model of our early sepsis prediction model. Considering that the prediction performance of the GRU model may be higher than that of the LSTM model under the larger data size, we used the two-layer GRU model as the second-level model of our early sepsis prediction model, and the final experimental results also confirmed the suitability of the two-layer GRU model as the second-level model. In the end, the two-level model was fused by entropy weighting method to form our early sepsis prediction model. After constructing the Stacking-Based Temporal GRU model, we conducted extensive comparison experiments on our processed dataset using RF, XGB, CatBoost, Bi-GRU, Bi-LSTM model and KANNEDS methods. Our early sepsis prediction model was found to perform significantly better than other models in three metrics, Accuracy, F1 and AUC, and proved the accuracy and completeness of our model.

This study also has limitations. Firstly, the Early Prediction of Sepsis From Clinical Data: The PhysioNet Computing in Cardiology Challenge 2019 dataset is a

dataset containing only some of the medical center patients. The application of the developed Stacking-Based Temporal GRU model to other datasets or domains requires further clinical evaluation. Secondly, this study discarded some patients' characteristic variables in the data preprocessing stage, which may have some impact on the experimental results. Thirdly, the developed model is based on baseline data from the 36 hours in the ICU and cannot be used to dynamically predict the prevalence of sepsis.

Conclusion

This study contributes to clinical areas with a Stacking-Based Temporal GRU model for early detection of sepsis in ICU. The prediction model has the ability to help ICU physicians predict which patients are likely to contract sepsis so that they can take medication and clinical interventions to reduce their patients' risk of infection. Since all variables used for model construction in this study are regularly collected by the information systems of three separate hospital ICU, it is feasible to ensure that we use the Stacking-Based Temporal GRU model to help ICU physicians with clinical interventions. We plan to ensemble the Stacking-Based Temporal GRU model into existing ICU medical systems for real-time prediction of whether a patient will contract sepsis to provide clinical decision support. Our future research will focus on evaluating the general applicability of our early sepsis prediction model, conducting experiments on data from additional medical centers to confirm the generalizability and validity of the model.

Additional Files

Abbreviations

ICU: Intensive Care Unit; XGB: Extreme gradient boosting; RF: Random forest; CatBoost: categoricalboosting; LR: logistic regression; LSTM: Long short-term memory; GRU: Gate Recurrent Unit;Null: Not a Nnumber. KF: Kinematic feature; VS: Vital sign.

Acknowledgements

The authors would like to thank the reviewers of this manuscript for their valuable comments, which have helped the authors improve the paper significantly.

Declarations of interest

None

Competing interest

None

Authors' contributions

Conception and design: YL; Collection and assembly of data: YL and XY; Data analysis and interpretation: YL, XY and QL. All authors provided valuable inputs and comments on manuscript revision. The final manuscript was written by YL, and all authors have read and approved the final manuscript.

Funding
None

Ethics approval and consent to participate
None

Consent for publication
Not applicable

Competing interests
All authors have declared that they do not have any potential conficts of interest.

Author details
[1]*School* of Computer and Information Engineering, Hubei Normal University, Huangshi, China

Author details
[1]Department of Science, University of Cambridge, London, UK. [2]Institute of Biology, National University of Sciences, Kiel, Germany.

References

1. Gyawali B, Ramakrishna K, Dhamoon A. Sepsis: The evolution in definition, pathophysiology, and management. SAGE Open Medicine. 2019 03;7:205031211983504.
2. Singer M, Deutschman C, Seymour C, Shankar-Hari M, Annane D, Bauer M, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016 02;315:801.
3. Levy M, Artigas A, Phillips G, Rhodes A, Beale R, Osborn T, et al. Outcomes of the Surviving Sepsis Campaign in intensive care units in the USA and Europe: A prospective cohort study. The Lancet infectious diseases. 2012 10;12.
4. Rudd K, Johnson S, Agesa K, Shackelford K, Tsoi D, Kievlan D, et al. Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study. The Lancet. 2020 01;395:200–211.
5. Kaukonen KM, Bailey M, Suzuki S, Pilcher D, Bellomo R. Mortality Related to Severe Sepsis and Septic Shock Among Critically Ill Patients in Australia and New Zealand, 2000-2012. JAMA : the journal of the American Medical Association. 2014 03;311.
6. Zheng R, Zhang Y, Rong Z, Huang W, Fu X. Surviving Sepsis Campaign: international guidelines for management of sepsis and septic shock 2021, interpretation and expectation. Zhonghua wei zhong bing ji jiu yi xue. 2021 10;33:1159–1164.
7. Westphal G, Gonçalves A, Caldeira-Filho M, Silva E, Salomão R, Bernardo W, et al. Guidelines for treatment of severe sepsis/septic shock –tissue perfusion assessment. RBTI. 2011 03;23:6–12.
8. Cruz M, Cavalcante C, Barretto S. Using OPC and HL7 Standards to Incorporate an Industrial Big Data Historian in a Health IT Environment. Journal of Medical Systems. 2018 05;42:122.
9. Rousselot J, Decotignie JD. Wireless Communication Systems for Continuous Multiparameter Health Monitoring; 2009. p. 480 – 484.
10. Cruz M, Ono N, Huang M, Amin A, Kanaya S, Cavalcante C. Kinematics approach with neural networks for early detection of sepsis (KANNEDS). BMC Medical Informatics and Decision Making. 2021 05;21.
11. Manaktala S, Claypool S. Evaluating the impact of a computerized surveillance algorithm and decision support system on sepsis mortality. Journal of the American Medical Informatics Association. 2016 05;24:ocw056.
12. Le S, Hoffman J, Barton C, Fitzgerald J, Allen A, Pellegrini E, et al. Pediatric Severe Sepsis Prediction Using Machine Learning. Frontiers in Pediatrics. 2019 10;7.
13. Kong G, Lin K, Hu Y. Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU. BMC medical informatics and decision making. 2020 10;20:251.
14. Li X, Xu X, Xie F, Xu X, Sun Y, Liu X, et al. A time-phased machine learning model for real-time prediction of sepsis in critical care. Critical Care Medicine. 2020;48(10):e884–e888.
15. Mohammed A, Wyk F, Chinthala L, Khojandi A, Davis R, Coopersmith C, et al. Temporal Differential Expression of Physiomarkers Predicts Sepsis in Critically Ill Adults. Shock (Augusta, Ga). 2020 09;Publish Ahead of Print.
16. Fleuren LM, Klausch TL, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. Intensive care medicine. 2020;46(3):383–400.
17. Moor M, Rieck B, Horn M, Jutzeler CR, Borgwardt K. Early prediction of sepsis in the ICU using machine learning: a systematic review. Frontiers in medicine. 2021;8:607952.
18. Kam HJ, Kim HY. Learning representations for the early detection of sepsis with deep neural networks. Computers in biology and medicine. 2017;89:248–255.
19. Zhang D, Yin C, Hunold Buck K, Jiang X, Caterino J, Zhang P. An Interpretable Deep Learning Model for Early Prediction of Sepsis in the Emergency Department; 2020.

20. Al-Mualemi BY, Lu L. A deep learning-based sepsis estimation scheme. Ieee Access. 2020;9:5442–5452.
21. Tsang G, Xie X. Deep learning based sepsis intervention: the modelling and prediction of severe sepsis onset. In: 2020 25th international conference on pattern recognition (ICPR). IEEE; 2021. p. 8671–8678.
22. Reyna MA, Josef C, Seyedi S, Jeter R, Shashikumar SP, Westover MB, et al. Early prediction of sepsis from clinical data: the PhysioNet/Computing in Cardiology Challenge 2019. In: 2019 Computing in Cardiology (CinC). IEEE; 2019. p. Page–1.
23. Reyna JCJRSSMBWMBSANSCGD M. Early Prediction of Sepsis from Clinical Data: The PhysioNet/Computing in Cardiology Challenge 2019 (version 1.0.0). PhysioNet. 2019;.
24. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. circulation. 2000;101(23):e215–e220.
25. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. Advances in neural information processing systems. 2018;31.
26. Breiman L. Bagging predictors. Machine learning. 1996;24(2):123–140.
27. Breiman L. Random forests. Machine learning. 2001;45(1):5–32.
28. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; 2016. p. 785–794.
29. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems. 2017;30.
30. Cramer JS. The origins of logistic regression. 2002;.
31. Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:14123555. 2014;.
32. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:14061078. 2014;.
33. Zhao J, Ji G, Tian Y, Chen Y, Wang Z. Environmental vulnerability assessment for mainland China based on entropy method. Ecological Indicators. 2018;91:410–422.
34. Zhou R, Hu X, Yuan B, Xu Q. Lithology Classification System for Well Logging Based on Bidirectional Gated Recurrent Unit. In: 2021 4th International Conference on Artificial Intelligence and Big Data (ICAIBD). IEEE; 2021. p. 599–603.
35. Huang Z, Xu W, Yu K. Bidirectional LSTM-CRF models for sequence tagging. arXiv preprint arXiv:150801991. 2015;.
36. Dong X, Yu Z, Cao W, Shi Y, Ma Q. A survey on ensemble learning. Frontiers of Computer Science. 2020;14(2):241–258.