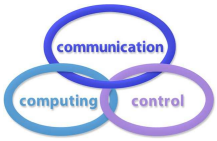
INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 16, Issue: 3, Month: June, Year: 2021
Article Number: 4180, https://doi.org/10.15837/ijccc.2021.3.4180



CCC Publications

Screening Key Indicators for Acute Kidney Injury Prediction Using Machine Learning

Jiaming Wang , Bing Zhu, Pei Liu, Ruiqi Jia, Lijing Jia, Wei Chen, Cong Feng, Jing Li

Jiaming Wang, Bing Zhu, Pei Liu, Ruiqi Jia, Jing Li*
School of Economics and Management
Beijing Jiaotong University, Beijing
*Corresponding author: jingli@bjtu.edu.cn

Lijing Jia, Cong Feng*
Department of Emergency
The First Medical Center to Chinese People's Liberation Army General Hospital, Beijing
*Corresponding author: fc194@163.com

Wei Chen
Department of Emergency
The Third Medical Center to Chinese People's Liberation Army General Hospital, Beijing

Abstract

Acute kidney injury is a common critical disease with a high mortality. The large number of indicators in AKI patients makes it difficult for clinicians to quickly and accurately determine the patient's condition. This study used machine learning methods to filter key indicators and use key indicator data to achieve advance prediction of AKI so that a small number of indicators could be measured to reliably predict AKI and provide auxiliary decision support for clinical staff. Sequential forward selection based on feature importance calculated by XGBoost was used to screen out 17 key indicators. Three machine learning algorithms were used to make predictions, namely, logistic regression (LR), decision tree, and XGBoost. To verify the validity of the method, data were extracted from the MIMIC III database and the eICU-CRD database for 1,009 and 1,327 AKI patients, respectively. The MIMIC III database was used for internal validation, and the eICU-CRD database was used for external validation. For all three machine learning algorithms, the prediction performance from using only the key indicator dataset was very close to that from using the full dataset. The XGBoost algorithm performed the best, and LR was the next best. The decision tree performed the worst. The key indicator screening method proposed in this study can achieve a good predictive performance while streamlining the number of indicators.

Keywords: acute kidney injury, key indicator screening, machine learning, sequential forward selection, XGBoost.

1 Introduction

Acute kidney injury (AKI) is a common critical illness [1, 2, 3]. According to KDIGO (Kidney Disease: Improving Global Outcomes), AKI is defined as the fulfillment of any of the following criteria [4]:

- *An increase in serum creatinine by* $\geqslant$ *0.3 mg/dL (*$\geqslant$ *26.5* μ *mol/L) within 48 h, or*
- *An increase in serum creatinine to* $\geqslant$ *1.5 times baseline within the previous 7 days, or*
- *Urine volume* $\leqslant$ *0.5 ml/kg/h for 6 h.*

AKI has acute morbidity and a high mortality rate [5, 6]. It occurs in at least 5% of hospitalized patients [7] and has many serious consequences, including prolonged ICU and hospital stays, development of chronic kidney disease, and short- and long-term increased risk of death [8]. Therefore, early prediction of AKI and subsequent treatment are essential subjects of research.

Artificial intelligence and the establishment of diagnosis and treatment databases have combined to enable technical and data support for the development and implementation of machine learning methods based on condition prediction algorithms. Scholars have also achieved some success in applying machine learning methods for early warning prediction of AKI. Thottakkara et al. [9] evaluated logistic regression (LR), generalized additive model (GAM), naive Bayes, and support vector machine techniques for predicting the performance of postoperative sepsis and AKI and showed that LR, GAM, and SVM had better performance than naive Bayes, and feature extraction using principal component analysis improved the predictive performance of all models. Kate et al. [10] developed a machine learning approach based on four models—LR, SVM, decision tree, and naive Bayes—and their ensembles for AKI prediction. Their results showed that LR had the best predictive performance (AUC = 0.743).

In a recent study of key indicators and subsequent state prediction in emergency patients, Coster et al. [11], using statistical analysis of six-month patient reports, presented key indicators associated with survival within seven days in emergency scenarios in which the patient had a serious injury. McCoy et al. [12] used machine learning algorithms to assign a risk score for sepsis patients with high mortality rates to provide decision support for early intervention against the condition and constructed a patient index system through manual screening. Levin et al. [13] addressed the shortcomings of the standard of care and risk classification of patients in the U.S. emergency field, which relies heavily on subjective empirical methods, and used a random forest algorithm to predict patient outcomes based on important historical treatment data. They obtained higher accuracy than subjective methods. Hohl et al. [14] adopted a prospective research method and combined it with patient indicators that they collected to establish two sets of pharmacist-led drug treatment decision rules for emergency patients. They reported high diagnostic accuracy in practical use.

Recent studies have shown that machine learning has provided new ideas for early prediction studies of AKI. For outcome prediction and early risk warning studies related to emergency patients, machine learning has been more effective than traditional methods. However, most of the current methods for AKI prediction are based on full datasets. In the field of emergency care, there are often patients with serious conditions that require urgent treatment and quick diagnostic decisions by medical personnel. At the same time, the variety of physiological and laboratory indicators for patients is very large, and it takes a lot of time to collect samples and obtain the values of the indicators. While waiting, the patient's condition may deteriorate. The length of time patients wait for the results of the indicators impacts the timeliness of diagnosis and treatment. When the number of indicators is small, it is easy for physicians to screen key indicators by subjective methods, but when the number of physiological and laboratory indicators is large, physicians relying only on subjective methods to make decisions may make biased conclusions. As such, simultaneously meeting the requirements of speed and accuracy is difficult. This study addresses this problem by proposing a machine learning based method to screen out key indicators and generate predictions by collecting only a small number of key indicators, thus assisting clinicians in making fast and accurate judgments about patients' conditions.

This study uses eXtreme Gradient Boosting (XGBoost) and sequential forward selection (SFS) to extract key indicators. The extracted key indicators are used in three algorithms: LR, decision

tree, and XGBoost to assess their relative performance. They are also validated on the full dataset to compare and analyze the effectiveness of the key indicators' extraction.

2 Methods

2.1 Data collection and processing

The data sources of this study are the MIMIC III database (Medical Information Mart for Intensive Care III) and eICU-CRD database (eICU Collaborative Research Database). MIMIC III is a publicly available database developed by the Massachusetts Institute of Technology Computational Physiology Laboratory. The database contains data related to >6,000 hospitalizations between 2001 and 2012, including demographics, vital signs, laboratory tests, drugs, and other data. The eICU-CRD database is a large public database created by the Philips Group and the Massachusetts Institute of Technology Computational Physiology Experiment. It contains a large amount of high-quality clinical information related to admissions to the intensive care unit in 2014 and 2015.

Data were extracted from the MIMIC III database and the eICU-CRD database according to the inclusion and exclusion criteria given by doctors. As the number of people in the non-AKI group was higher than that in the AKI group, we used equalization, that is, the number of individuals in the selected control group was similar to the number in the experimental group [15, 16] . From the MIMIC III database, 1,009 people each in the non-AKI and AKI group were extracted, with a total of 47 indicators. From the eICU-CRD database, 1,327 people each in the non-AKI and AKI group were extracted, with a total of 41 indicators. The sample set extraction process is shown in Fig. 1.

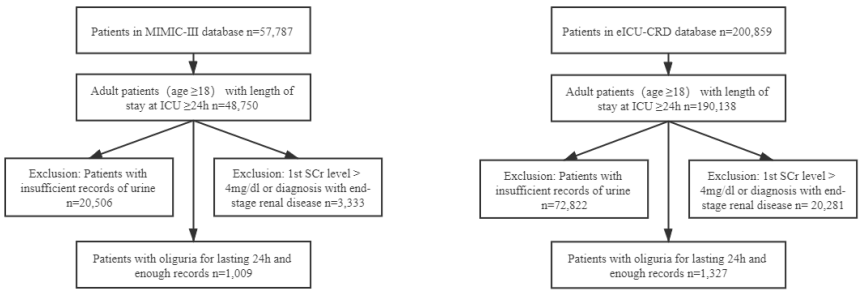


Figure 1: The sample set extraction process of (a) MIMIC III database and (b) eICU-CRD database.

Owing to the differences in the time of data collection and the specificity of each patient's condition, there were a large number of missing values in the data. The data parts wherein the proportion of missing values was too high were eliminated. Finally, 39 indicators were selected from the MIMIC III database, and 25 indicators were selected from the eICU-CRD database.

Taking the data extracted from the MIMIC III database as an example, the percentage of missing data is visualized in Appendix Fig. 1.

With the development of the patient's state and condition, various indicators also change dynamically. In this case, if the mean, median, and mode are used to fill in missing values, a large number of data entries will be the same, which will not accurately reflect the dynamic development of the patient's physical state. In this study, we assumed that patient indicators varied linearly. Therefore, we first performed linear interpolation for the same patient. For the remaining missing data, we used a clustering approach to find similar patients and then populated the missing data with values of the similar patients.

2.2 Model evaluation metrics

In this study, 10-fold cross validation was adopted to evaluate model performance. The dataset was randomly partitioned into 10 equal-sized subsamples, nine of which were used to train the model

prior to validation using the remaining subsample. To verify the generalizability of the methods and models, we used the data from the MIMIC III database as internal validation data and the data from the eICU-CRD database as external validation data. Accuracy, recall, and precision scores were used to assess the prediction results. The F-score combines the two indicators and provides a more realistic measure of model performance using both precision and recall as follows:

$$F = \frac{(\beta^2 + 1) * \text{ Precision } * \text{ Recall}}{\beta^2 * \text{ Precision } + \text{ Recall}}$$

The F-score takes into account both precision and recall and can be considered as a harmonic average of precision and recall. In the setting of this study, potential AKI patients should be identified to the extent possible. Therefore, recall is more important. This article sets $\beta = 1.5$, that is, F1.5 score. The performance evaluation of the model is mainly based on the AUC (area under the receiver operating characteristic curve) value and reference to the F1.5 value. The ROC (receiver operating characteristic) curve was used to evaluate algorithm performance. It is a comprehensive indicator representing the continuous variables of sensitivity and specificity. The AUC represents the probability that a predicted positive sample is ranked above a negative sample, thus reflecting a classifier's ability to sort samples. Values closer to 1 indicate better model performance.

2.3 Machine learning algorithms

This study uses three machine learning algorithms: LR, decision tree, and XGBoost. LR is a prediction tool that is often used in medical research [17] and does not require much calculation. Moreover, the calculation speed is efficient. Instead of fitting a straight line or hyperplane, we used the logistic function to squeeze the output of a linear equation between 0 and 1.

The interpretability of this algorithm is that we can see the change of prediction when one of the features x_j is changed by 1 unit, as the algorithm is an extension of the linear regression model.

The essence of a decision tree is that many decision structures are combined in a tree shape, and the leaf nodes represent the final category. This algorithm is interpretable, and the logic is easy to understand. In this study, the CART decision tree algorithm was applied.

XGBoost is a strong classifier that integrates with decision trees, which are weak classifiers [19]. The importance of each candidate predictor is ranked by selection frequency. The sum of all importance scores is scaled to 100, which means each importance can be interpreted as the share of the overall model importance. Additionally, individual predictions of XGBoost can be explained by decomposing the decision path into one component per feature. Therefore, a decision can be tracked through the tree and explained as a prediction by the contributions added at each decision node. The principle is as follows. Let $\mathrm{D} = \{(x_i, y_i)\}\ (|D| = n, x_i \in \mathbb{R}^m, y_i \in \mathbb{R}^n)$ represent a database with n examples and m features. A tree boosting model output $\widehat{y_i}$ with K trees is defined as follows:

$$\widehat{y_i} = \sum_{k=1}^{K} f_k\left(x_i\right), f_k \in \mathrm{F}$$

where $\mathrm{F} = \{f(x) = \omega_q(x)\}\left(q : \mathbb{R}^m \to T, \omega \in \mathbb{R}^T\right)$ is the space of the regression or classification trees (i.e., CART). The tree is divided by f_k into the structure part q and the leaf weights part ω. Here, T indicates the number of leaves. XGBoost adds second derivative information and regularization in the cost function to control the complexity of the model and prevent overfitting. For samples with missing feature values, XGBoost can automatically learn the splitting direction [20].

2.4 Key indicators extraction and visualization method

The key indicator screening method in this study is SFS based on XGBoost. SFS is a greedy algorithm that avoids a simple exhaustive search. The set of key indicators starts with an empty set. One indicator at a time is selected to be added to the set of key indicators for iteration, and the results of each iteration of the model are recorded. During the search process, features are continuously added

or removed from the current key indicator set according to a certain order so as to obtain an optimized key indicator set [21].

For interpretability, this study uses Shapley additive explanations (SHAPs). SHAP is an additive explanatory model inspired by cooperative game theory where all features are considered as "contributors." For each prediction sample, the model generates a prediction value, and the SHAP value is the value assigned to each feature in that sample. The specific formula is as follows [22, 23]:

$$g\left(z'\right) = \phi_0 + \sum_{j=1}^{M} \phi_j z'_j,$$

where g is the explanation model, $z' \in \{0,1\}^M$ are the simplified features, M is the maximum features size, and $\phi_j \in \mathbb{R}$ is the feature attribution for a feature j. The advantages of the SHAP value is that it can substantially reflect the characteristics in each sample and also show positive and negative effects. We used SHAP summary plots to reflect the impacts of all key indicators on the risk of developing AKI in patients as well as partial SHAP dependent plots to illustrate the impact of individual indicators.

3 Results

3.1 Patient characteristics

The demographic data, vital signs, laboratory indicators, and other data were obtained from the MIMIC III database and eICU-CRD database. The differences in age, gender, hospital length of stay, and mortality between the AKI patients group and the non-AKI patients group are shown in Appendix Tables 2 and 3. For continuous variables, the Mann-Whitney U-test was used to compare the differences between the two groups, and for the categorical variables, the chi-square test was used. The MIMIC III database and eICU-CRD database had significant differences in age, gender, hospital length of stay, and mortality in the hospital. The AKI group had a higher mortality (MIMIC III: 49.7% vs. 7.9%; eICU-CRD: 40.7% vs. 6.9%).

3.2 Key indicators extraction results

All indicator weights for XGBoost were calculated. The dataset was added in descending order of the indicator weights, and predictions were made using XGBoost. A total of 39 iterations were performed, and the results of each iteration were recorded. As can be seen in Table 3, the F1.5 score of the iterative process changed. As shown in Fig. 2, the F1.5 score reached a high value when the number of indicators reached 17 (F1.5 = 0.8853). So, we selected the key indicator set as the top 17 indicators for input into the model. They are serum creatinine, urea nitrogen, INR, PT, pH, urine pH, leukocytes, SBP, SpO2, potassium, chloride, total CO2, lactate, temperature, AST, DBP, and ALP. The weight of all key indicators is 0.6861.

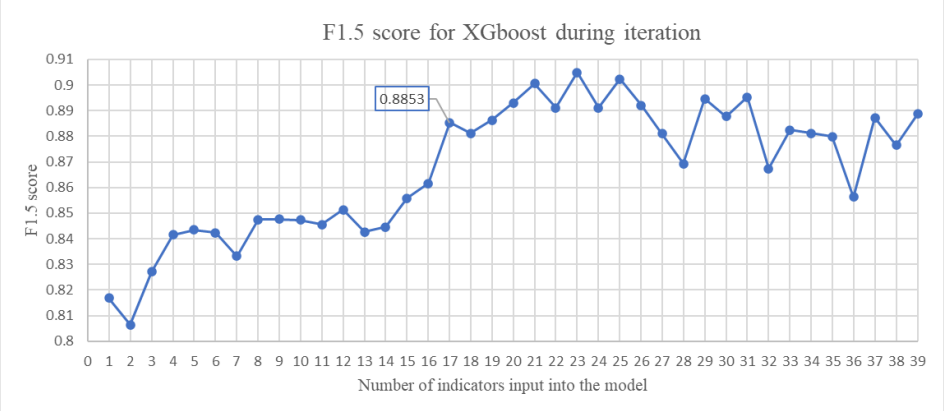
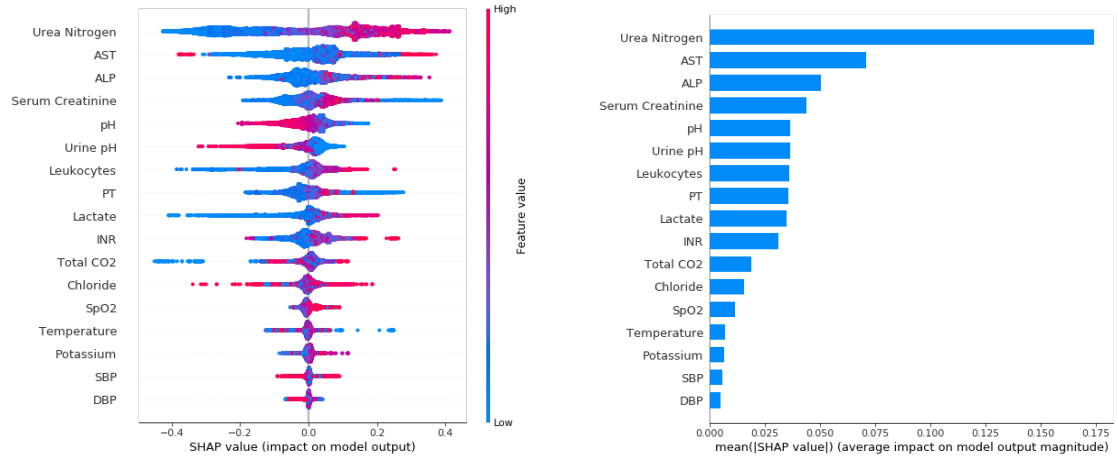


Figure 2: F1.5 score for XGBoost during iteration.

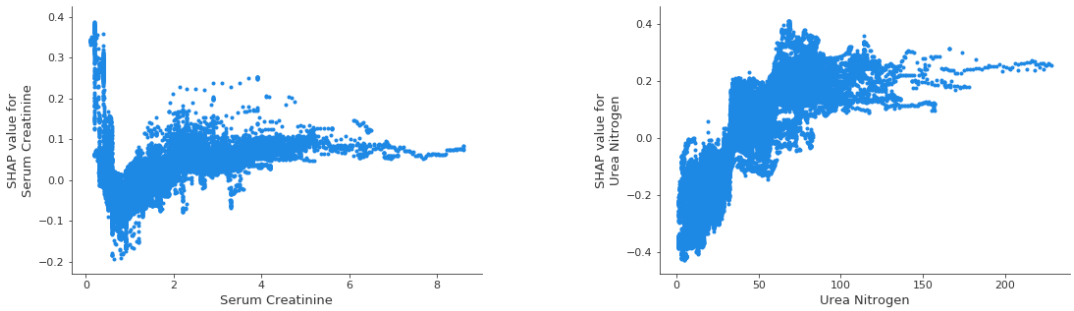
3.3 Visualization of feature importance

To enhance clinicians' intuitive understanding of the importance of indicators, this study used SHAP to illustrate how key indicators affect AKI onset (Fig. 3). Among the key indicators, higher urea nitrogen, AST, ALT, lactate, leukocytes, INR, total CO2, chloride, potassium, SpO2, SBP, lower creatinine, pH, urine pH, PT, temperature, and DBP were noted. SHAP can also be used to illustrate the influence of individual indicators on the development of AKI in patients. As shown in Figs. 4, patients are at higher risk of developing AKI when urea nitrogen is >40 mg/dL and when serum creatinine is <0.7 mg/dL.



(a) (b)

Figure 3: SHAP summary plots of key indicators for AKI patients. (a) The importance ranking of the key indicators. (b) The importance ranking of the key indicators according to the mean (|SHAP value|). INR: international normalized ratio; PT: prothrombin time; SBP: systolic blood pressure; AST: aspartate amino transferase; DBP: diastolic blood pressure; ALP: alkaline phosphatase.



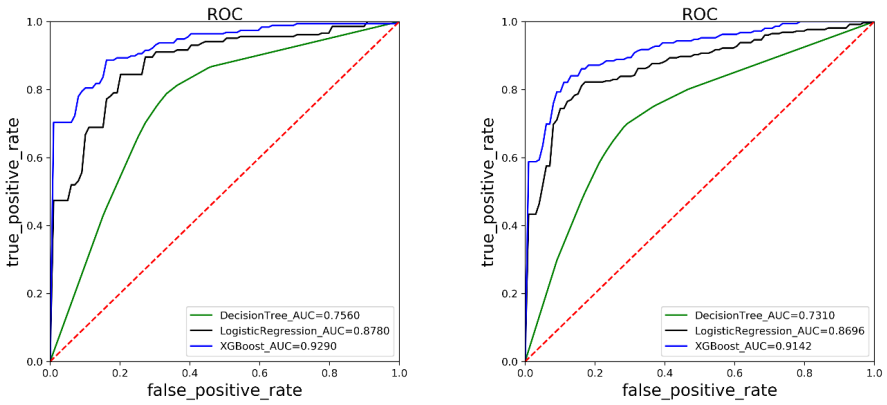
(a) (b)

Figure 4: Partial SHAP dependence plots. (a) Serum creatinine. (b) Urea nitrogen

3.4 Key indicators and full dataset prediction results

The key indicator data were input into the algorithm model. The ROC curves and AUC values of the prediction results are shown in Fig. 5. From the results, it can be seen that the prediction performance of the key indicator set is comparable to that of the full dataset in internal and external validation. Among the three algorithms, the XGBoost algorithm had the best prediction performance (MIMIC III AUC: 0.9290, 95% CI: 0.8971–0.9680; eICU-CRD AUC: 0.9142, 95% CI: 0.8745–0.9601). LR was second best (MIMIC III AUC: 0.8780, 95% CI: 0.8210–0.9379; eICU-CRD AUC: 0.8696, 95%

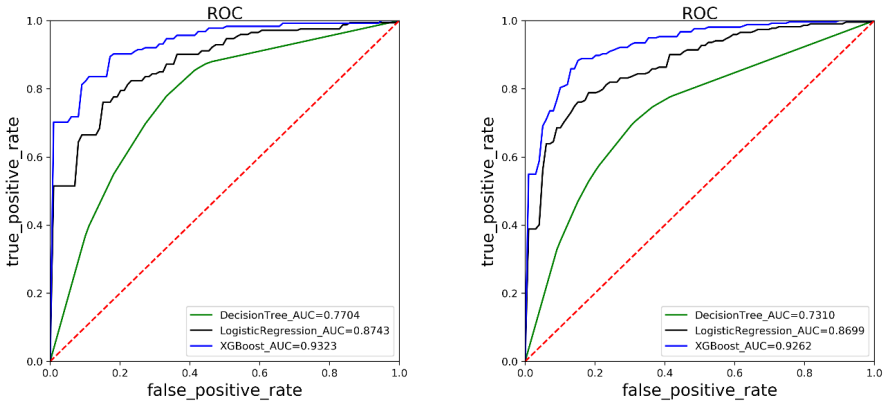
CI: 0.8469–0.8955). Decision tree had the worst performance (MIMIC III AUC: 0.7560, 95% CI: 0.7230–0.7890; eICU-CRD AUC: 0.7310, 95% CI: 0.6957–0.7664).



(a) (b)

Figure 5: ROC curves of the key indicators showing the performance of the models in predicting AKI. (a) Internal verification results (MIMIC III database). (b) External verification results (eICU-CRD database).

To validate the prediction of the key indicators, we input all data into the three machine learning models. The ROC curves and AUC values predicted by the models are shown in Fig. 6. From the internal verification and external verification results, we can see that the prediction performances of LR (MIMIC III AUC: 0.8743, 95% CI: 0.8375–0.9170; eICU-CRD AUC: 0.8699, 95% CI: 0.8459–0.8972), and XGBoost (MIMIC III AUC: 0.9323, 95% CI: 0.9041–0.9669; eICU-CRD AUC: 0.9262, 95% CI: 0.9140–0.9433) were high, and the prediction performance of the decision tree was the worst (MIMIC III AUC: 0.7704, 95% CI: 0.7144–0.8265; eICU-CRD AUC: 0.7310, 95% CI: 0.6928–0.7691).



(a) (b)

Figure 6: ROC curves of the full dataset showing the performance of the models in predicting AKI. (a) Internal verification results (MIMIC III database). (b) External verification results (eICU-CRD database).

The prediction results for all indicator sets and key indicator sets by the three algorithms are shown in Table 1. From the results, it can be seen that the three algorithms perform consistently with respect to the key indicator dataset and all the indicators dataset. XGBoost performs first best; LR was second best; and decision tree was the worst. When using the same algorithm, the prediction of the key indicator dataset is very close to that of the full indicators dataset, which verifies that using the screening key indicators method proposed in this study for prediction is effective.

Table 1: Summary of the internal and external verification results of the prediction of the full dataset and key indicator dataset of the three algorithms

		MIMIC III					eICU-CRD				
		AUC	F1.5	Accuracy	Precision	Recall	AUC	F1.5	Accuracy	Precision	Recall
The full dataset	Decision tree	0.7704	0.8010	0.7584	0.8060	0.8040	0.7310	0.7360	0.7363	0.7441	0.7394
	LR	0.8743	0.8214	0.7851	0.8281	0.8227	0.8699	0.8132	0.8126	0.7911	0.8336
	XGBoost	0.9323	0.8796	0.8566	0.8962	0.8763	0.9262	0.8542	0.8538	0.8446	0.8613
The key indicator dataset	Decision tree	0.7793	0.8215	0.7786	0.7949	0.8413	0.7310	0.7629	0.7567	0.7221	0.7872
	LR	0.8811	0.8444	0.8133	0.8621	0.8418	0.8696	0.8421	0.8311	0.8004	0.8730
	XGBoost	0.9302	0.8764	0.8567	0.8714	0.8819	0.9142	0.8556	0.8468	0.8267	0.8747

4 Discussion

AKI is a common critical illness with a rapid onset and high mortality rate. Early detection of the condition is important for saving the patient's life and reducing the burden of medical costs. This research proposed a machine learning technique for screening key indicators for AKI. Three machine learning methods, namely, LR, decision tree, and XGBoost, were compared. LR required a short training time, and it predicted well. Its AUC was higher than 0.85. Decision tree construction is simple and interpretable, but its prediction performance was average. XGBoost had the highest prediction accuracy. It also allows the existence of null values in the features and effectively prevents overfitting.

For making predictions on key indicator datasets, the results showed that all three algorithms can achieve the prediction purpose. To verify the effectiveness of the key indicator index screening method, this research used the above three machine learning methods to compare the prediction effects on the full datasets. It was found that using the key indicator screening method proposed in this study, it was possible to achieve similar predictive performance using only the key indicator dataset to that from using the full dataset. To increase the interpretability of the results and give clinicians an intuitive understanding of the importance of the indicators, this study used SHAP to illustrate the impact of the key indicators on the risk of AKI.

Among the existing studies on the interpretation of key indicators of AKI, several authors have pointed out that serum creatinine and urea nitrogen are key indicators that affect the development of AKI, but the existing studies mostly focus on AKI patients with comorbidities. For example Maiwall [24] studied the effect of AKI on the prognosis of patients with acute chronic liver failure (ACLF) for 48 hours. It was found that when the serum creatinine value is 1.14 mg/dL and the value increase is small, risk stratification of ACLF patients should be considered while adopting intervention strategies. Dinna [25] examined the association between increased creatinine and AKI risk in the ICU adult population. The study found that based on the combined effect of clinical factors and early creatinine elevation, the possibility of severe AKI in critically ill patients is significantly increased. Vanmassenhove [26] found an association with mortality for AKI patients with septic sepsis only if the evolution of serum creatinine over the first 24 h after ICU admission was taken into account. Smith [27] established a latent variable mixture model for patients to evaluate the relationship between AKI risk factors and changes in serum creatinine and predict the risk of AKI. Dewitte [28] evaluated the performance of fractional excretion urea in identifying persistent and transient AKI in patients admitted to the intensive care unit. The author found that in patients in the intensive care unit, urea excretion rate of less than 40% was a sensitive and specific indicator for the identification of transient and persistent AKI, especially in the case of diuretics. Corte [29] assessed whether serum urea concentration or a different recommended serum urea concentration was related to hospital mortality when AKI patients received RRT. This retrospective study showed that serum urea concentration and the cut-off value of serum urea concentration at the time of initiation of RRT had no predictive value for the hospital mortality of ICU AKI patients.

However, studies have not indicated the warning range for serum creatinine and urea nitrogen in patients with AKI. This study predicted the risk of AKI in the adult population of ICU, screened 17 key indicators, and used SHAP plots to find the early warning range of serum creatinine and urea nitrogen. It was found that patients with urea nitrogen >40 mg/dL and serum creatinine <0.7 mg/dL had a higher risk of AKI.

To test the reliability of the model, this study conducted external verification. Damen suggested that the future research direction of disease prediction should be in the external verification of the model [30]. This study validated the results using the MIMIC III database and eICU database. From the external validation results, it can be seen that the model developed in this study performs equally on both databases, indicating its reliability. The cross-validation results also confirm that the model developed in this paper is generalizable.

5 Conclusion

This study provides a reference for other applications of machine learning methods and similar clinical decision support studies and improves the feasibility and timeliness of early prediction and warning of AKI. In the case of urgency or limited equipment, using only the key indicator data for prediction can not only achieve similar results as using the full data but also reduce the frequency of collection of data, shorten the length of the wait for indicator results, and enable early detection and treatment. There are some limitations of this study. For example, the number of cases studied was limited. The study sample size needs to be expanded to improve prediction accuracy. The ethnicity of the data is predominantly Western. More people from other ethnic groups could be included in the future. Owing to the different collection times of the samples, there are long intervals between the collection of certain data, which made it difficult to fill in some data. This factor may have an impact on the prediction accuracy of the models. This study is a retrospective study based on existing data. Although it has achieved an effective prediction performance, there may still be a gap in clinical use. These issues will be the focus of our future research.

Acknowledgment

This work was partly supported by the National Key Research and Development Plan for Science and Technology Winter Olympics of the Ministry of Science and Technology of China (2019YFF0302301), the Hainan Provincial Natural Science Foundation of China (818MS156) and the Special Fund for Health Care of the Ministry of Logistics(16BJZ19).

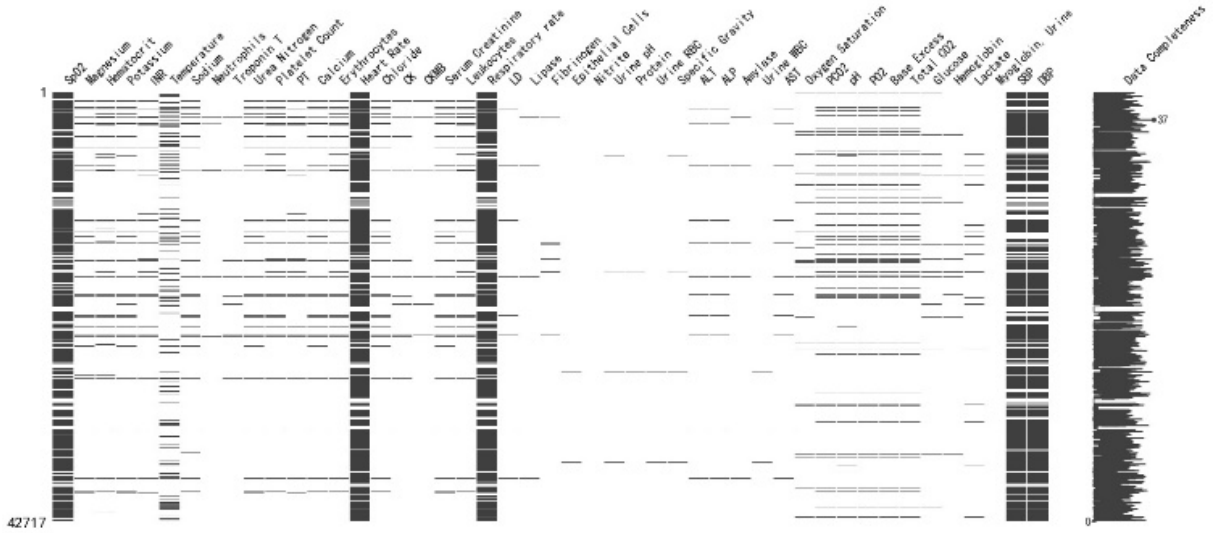
Appendix

Appendix Table. 1: MIMIC III database research population baseline statistics

Characteristics	AKI (n = 1009)	Non-AKI (n = 1009)	P value
Age (years)	72.15 (58.58,82.19)	64.23 (50.78,77.16)	< 0.01
Gender, n (%)			<0.01
Female	533 (48.5%)	486 (44.2%)	
Male	476 (43.3%)	614 (55.8%)	
Hospital length of stay (days)	13.78 (6.44,25.80)	7.02 (4.31, 12.25)	<0.01
Mortality (%)	547 (49.7%)	87 (7.9%)	<0.01

Appendix Table. 2: eICU-CRD database research population baseline statistics

Characteristics	AKI (n = 1327)	Non-AKI (n = 1327)	P value
Age (years)	69.00 (57.00, 79.00)	65.00 (53.00, 76.00)	<0.01
Gender, n (%)			<0.01
Female	707 (47.1%)	667 (44.5%)	
Male	618 (41.2%)	832 (55.5%)	
Hospital length of stay (days)	5.70 (2.82, 10.85)	4.78 (2.71,7.95)	<0.01
Mortality (%)	610 (40.7%)	104 (6.9%)	<0.01



Appendix Fig. 1: The visualization of the proportion of missing data (white indicates missing data). INR: international normalized ratio; PT: prothrombin time; CK: creatine kinase; CKMB: creatine kinase-MB; LD: lactate dehydrogenase; Urine RBC: urine red blood cells; ALT: alanine amino transferase; ALP: alkaline phosphatase; AST: aspartate amino transferase; SBP: systolic blood pressure; DBP: diastolic blood pressure; Urine WBC: urine white cells.

References

[1] S. Uchino, J. A. Kellum, R. Bellomo, G. S. Doig, H. Morimatsu, S. Morgera, M. Schetz, I. Tan, C. S. C. Bouman, and E. Macedo, *"Acute renal failure in critically ill patients: A multinational, multicenter study,"* JAMA, vol. 294, pp. 813-818, 2005.

[2] N. Pannu, M. T. James, B. R. Hemmelgarn, and S. Klarenbach, "Association between AKI, Recovery of Renal Function, and Long-Term Outcomes after Hospital Discharge," *Clinical Journal of The American Society of Nephrology*, vol. 8, pp. 194-202, 2013.

[3] X. Xu, S. Nie, Z. Liu, C. Chen, G. Xu, Y. Zha, J. Qian, B. Liu, S. Han, and A. Xu, "Epidemiology and Clinical Correlates of AKI in Chinese Hospitalized Adults," *Clinical Journal of The American Society of Nephrology*, vol. 10, pp. 1510-1518, 2015.

[4] J. A. Kellum, N. Lameire, P. Aspelin, R. S. Barsoum, E. A. Burdmann, S. L. Goldstein, C. A. Herzog, M. Joannidis, A. Kribben, and A. S. Levey, "Kidney disease: Improving global outcomes (KDIGO) acute kidney injury work group.KDIGO clinical practice guideline for acute kidney injury," *Kidney International*, vol. 2, pp. 1-138, 2012.

[5] Y. Fang, X. Ding, Y. Zhong, J. Zou, J. Teng, Y. Tang, J. Lin, and P. Lin, "Acute Kidney Injury in a Chinese Hospitalized Population," *Blood Purification*, vol. 30, pp. 120-126, 2010.

[6] J. Lafrance and D. R. Miller, "Acute Kidney Injury Associates with Increased Long-Term Mortality," *Journal of The American Society of Nephrology*, vol. 21, pp. 345-352, 2010.

[7] R. J. Kate, R. M. Perez, D. Mazumdar, K. S. Pasupathy, and V. Nilakantan, "Prediction and detection models for acute kidney injury in hospitalized older adults," *BMC Medical Informatics and Decision Making*, vol. 16, p. 39-39, 2016.

[8] E. Hoste, S. M. Bagshaw, R. Bellomo, C. M. Cely, R. Colman, D. N. Cruz, K. Edipidis, L. G. Forni, C. D. Gomersall, and D. Govil, "Epidemiology of acute kidney injury in critically ill patients: the multinational AKI-EPI study," *Intensive Care Medicine*, vol. 41, pp. 1411-1423, 2015.

[9] P. Thottakkara, T. Ozrazgatbaslanti, B. B. Hupf, P. Rashidi, P. M. Pardalos, P. Momcilovic, and A. Bihorac, "Application of Machine Learning Techniques to High-Dimensional Clinical Data to Forecast Postoperative Complications," *PLOS ONE*, vol. 11, 2016.

[10] R. J. Kate, R. M. Perez, D. Mazumdar, K. S. Pasupathy, and V. Nilakantan, "Prediction and detection models for acute kidney injury in hospitalized older adults," *BMC Medical Informatics and Decision Making*, vol. 16, p. 39-39, 2016.

[11] J. Coster, R. Jacques, J. Turner, A. Crum, J. Nicholl, and A. N. Siriwardena, "PP12 New indicators for measuring patient survival following ambulance service care," *Emergency Medicine Journal*, vol. 34, 2017.

[12] McCoy A, Das R., Reducing patient mortality, length of stay and readmissions through machine learningbased sepsis prediction in the emergency department, intensive care unit and hospital floor units.BMJ Open Quality 2017;6:e000158. doi:10.1136/bmjoq-2017-000158

[13] S. Levin, M. Toerper, E. Hamrock, J. S. Hinson, S. Barnes, H. Gardner, A. S. F. Dugas, B. Linton, T. D. Kirsch, and G. D. Kelen, "Machine-Learning-Based Electronic Triage More Accurately Differentiates Patients With Respect to Clinical Outcomes Compared With the Emergency Severity Index," *Annals of Emergency Medicine*, vol. 71, p. 565, 2017.

[14] C. M. Hohl, K. Badke, A. Zhao, M. E. Wickham, S. A. Woo, M. L. A. Sivilotti, and J. J. Perry, "Prospective Validation of Clinical Criteria to Identify Emergency Department Patients at High Risk for Adverse Drug Events," *Academic Emergency Medicine*, vol. 25, pp. 1015-1026, 2018.

[15] M. A. Tahir, J. Kittler and F. Yan, "Inverse random under sampling for class imbalance problem and its application to multi-label classification," *Pattern Recognition*, vol. 45, pp. 3738-3750, 2012.

[16] Q. A. Rahman, T. Janmohamed, H. Clarke, P. Ritvo, J. M. Heffernan, and J. Katz, "Interpretability and Class Imbalance in Prediction Models for Pain Volatility in Manage My Pain App Users: Analysis Using Feature Selection and Majority Voting Methods.," *JMIR medical informatics*, vol. 7, 2019.

[17] S. Sperandei, "Understanding logistic regression analysis," *Biochemia Medica*, vol. 24, pp. 12-18, 2014.

[18] X. Niuniu and L. Yuxun, *"Notice of Retraction: Review of decision trees,"* vol. 5, pp. 105-109, 2010-01-01 2010.

[19] T. Chen and C. Guestrin, *"XGBoost: A Scalable Tree Boosting System,"* pp. 785-794, 2016.

[20] L. Liu, Y. Yu, Z. Fei, M. Li, F. Wu, H. Li, Y. Pan, and J. Wang, "An interpretable boosting model to predict side effects of analgesics for osteoarthritis," *BMC Systems Biology*, vol. 12, pp. 29-38, 2018.

[21] N. Huang, G. Lu, G. Cai, D. Xu, J. Xu, F. Li, and L. Zhang, "Feature Selection of Power Quality Disturbance Signals with an Entropy-Importance-Based Random Forest," *Entropy*, vol. 18, p. 44, 2016.

[22] S. M. Lundberg and S. Lee, "A unified approach to interpreting model predictions," in *Advances in neural information processing systems*, 2017, pp. 4765-4774.

[23] C. Molnar, "Interpretable Machine Learning: A Guide for Making Black Box Models Explainable.(2019)," URL https://christophm. github. io/interpretable-ml-book, 2019.

[24] R. Maiwall, G. Kumar, A. Bharadwaj, K. Jamwal, A. S. Bhadoria, P. Jain, and S. K. Sarin, "AKI persistence at 48 h predicts mortality in patients with acute on chronic liver failure," *Hepatology International*, vol. 11, pp. 529-539, 2017.

[25] D. N. Cruz, A. Ferrer-Nadal, P. Piccinni, S. L. Goldstein, L. S. Chawla, E. Alessandri, C. Belluomo Anello, W. Bohannon, T. Bove, N. Brienza, M. Carlini, F. Forfori, F. Garzotto, S. Gramaticopolo, M. Iannuzzi, L. Montini, P. Pelaia, C. Ronco, and I. NEFROINT, "Utilization of small changes in serum creatinine with clinical risk factors to assess the risk of AKI in critically lll adults," *Clinical journal of the American Society of Nephrology*, vol. 9, pp. 663-672, 2014-01-01 2014.

[26] J. Vanmassenhove, N. Lameire, A. Dhondt, R. Vanholder, and W. Van Biesen, "Prognostic robustness of serum creatinine based AKI definitions in patients with sepsis: a prospective cohort study.," *BMC Nephrology*, vol. 16, p. 112-112, 2015.

[27] L. E. Smith, D. K. Smith, J. D. Blume, E. D. Siew, and F. T. Billings, "Latent variable modeling improves AKI risk factor identification and AKI prediction compared to traditional methods," *BMC Nephrology*, vol. 18, p. 55, 2017.

[28] A. Dewitte, M. Biais, L. Petit, J. F. Cochard, G. Hilbert, C. Combe, and F. Sztark, "Fractional excretion of urea as a diagnostic index in acute kidney injury in intensive care patients," *Journal of Critical Care*, vol. 27, pp. 505-510, 2012.

[29] W. De Corte, R. Vanholder, A. Dhondt, J. J. De Waele, J. Decruyenaere, C. Danneels, S. Claus, and E. Hoste, "Serum urea concentration is probably not related to outcome in ICU patients with AKI and renal replacement therapy," *Nephrology Dialysis Transplantation*, vol. 26, pp. 3211-3218, 2011.

[30] J. A. A. G. Damen, L. Hooft, E. Schuit, T. P. A. Debray, G. S. Collins, I. Tzoulaki, C. Lassale, G. C. M. Siontis, V. Chiocchia, and C. Roberts, "Prediction models for cardiovascular disease risk in the general population: systematic review," *BMJ*, vol. 353, 2016.

This journal is a member of, and subscribes to the principles of,
the Committee on Publication Ethics (COPE).
https://publicationethics.org/members/international-journal-computers-communications-and-control

Cite this paper as:

Jiaming Wang , Bing Zhu, Pei Liu, Ruiqi Jia, Lijing Jia, Wei Chen, Cong Feng , Jing Li (2021). Screening Key Indicators for Acute Kidney Injury Prediction Using Machine Learning *International Journal of Computers Communications & Control*, 16(3), 4180, 2021.
https://doi.org/10.15837/ijccc.2021.3.4180