scientific reports

Check for updates

OPEN

Enhancing healthcare AI stability with edge computing and machine learning for extubation prediction

Kuo-Yang Huang[1,2,3], Ying-Lin Hsu[4,6], Che-Liang Chung[2,3,5], Huang-Chi Chen[5], Ming-Hwarng Horng[5], Ching-Hsiung Lin[1,2], Ching-Sen Liu[6] & Jia-Lang Xu[7✉]

The advancement of the Internet of Medical Things (IoMT) has revolutionized data acquisition and processing in critical care settings. Given the pivotal role of ventilators, accurately predicting extubation outcomes is essential to optimize patient care. This study presents an edge computing-based framework that incorporates machine learning algorithms to predict ventilator extubation success using real-time data collected directly from ventilators. The system was deployed on edge devices to enable on-site inference with minimal latency. Among the evaluated models, Random Forest and XGBoost, the latter demonstrated superior predictive performance under both holdout and tenfold cross-validation schemes. Notably, the edge-based architecture reduced server data transmissions by 83.33%, while improving system stability, resilience, and sustainability. This paper details the model evaluation and demonstrates the feasibility and efficiency of edge intelligence in ventilator weaning decision support.

Abbreviations

IoMT	Internet of medical things
ICU	Intensive care units
MV	Mechanical ventilators
RF	Random forest
Vte	Expiratory tidal volume
RR	Respiration rate
Ppeak	Peak airway pressure
Pmean	Mean airway pressure
PEEP	Positive end-expiratory pressure
FiO2	Fraction of inspired oxygen

Currently, there are ever increasing technologies employed to integrate medical treatment with technology, enabling doctors to integrate technology to enhance medical treatment and reduce human errors, thus improving benefits. In numerous medical articles many scholars apply machine learning methods for both study and prediction purposes. Yin et al.[1] developed a stroke outcome prediction model using a deep convolutional neural network. Huang et al.[2] applied the YOLO model to construct a lung cancer detection system. Additionally, Huang et al.[3] employed a deep learning-based time series approach to predict ventilator weaning outcomes. Lee et al.[4] proposed a prediction model for malaria diagnosis and concluded that RF could achieve better prediction results when comparing multiple machine learning models. Talukder and Ahammed[5] proposed a predictive model for childhood malnutrition in Bangladesh using a Random Forest model. Seo et al.[6] proposed a method to predict postprandial hypoglycemia and concluded that a Random Forest approach could effectively predict postprandial hypoglycemia better than a variety of models. Hsiue et al.[7] found that cancer patients receiving MV

[1]Division of Chest Medicine, Department of Internal Medicine, Changhua Christian Hospital, Changhua, Taiwan. [2]Institute of Genomics and Bioinformatics, National Chung Hsing University, Taichung, Taiwan. [3]Ph.D. Program in Medical Biotechnology, National Chung Hsing University, Taichung, Taiwan. [4]Institute of Statistics, National Chung Hsing University, Taichung, Taiwan. [5]Division of Chest Medicine, Department of Internal Medicine, Yuanlin Christian Hospital, Changhua, Taiwan. [6]Department of Applied Mathematics, National Chung Hsing University, Taichung, Taiwan. [7]Department of Computer Science and Information Engineering, Chaoyang University of Technology, Taichung, Taiwan. ✉email: jlxu@cyut.edu.tw

in the ICU have a lower rate of successful extubation compared to non-cancer patients. Su et al.[8] suggest that age is an important factor for successful extubation and long-term survival. Milbrandt et al.[9] found that half of patients admitted to intensive care units (ICU) are aged over 65 years, thus, aging populations will increase stress in ICUs. Danaga et al.[10] note that nearly half of ICU patients need to use mechanical ventilators (MV) to ensure adequate oxygenation to maintain basic physiological functions. Respiratory failure occurs when patients are unable to maintain proper respiratory function due to various disease conditions. Wu et al.[11] believe that many factors, such as prolonged hospital stay and increased BUN levels, will affect the likelihood of extubation failure. Yang et al.[12] found that the inclusion of MV use under Taiwan's national health insurance scheme has increased their use.

The rapid advancements in the Internet of Medical Things (IoMT) have revolutionized healthcare by enabling real-time medical monitoring and data analysis. Efficient medical monitoring systems require high-speed data processing and minimal latency to ensure timely clinical decision-making. However, traditional cloud-based medical systems often suffer from delays due to data transmission and server processing limitations. Edge computing offers a promising solution by processing and analyzing data close to the source; this reduces transmission time, thereby reducing the burden on central servers, and enhancing system reliability. By integrating edge computing with machine learning techniques, medical data can be processed efficiently at the point of care before being uploaded to the server, thereby improving both system responsiveness and decision-making accuracy. Several studies have demonstrated the potential of edge computing in healthcare applications. Jiang et al.[13] highlighted that combining fuzzy logic, neural networks, and edge computing can enhance disease prediction and diagnosis. Rahman and Hossain[14] proposed an IoMT-based system for COVID-19 symptom detection, showing its effectiveness in home health management. Kong et al.[15] introduced an edge computing-based mask detection framework (ECMask) for real-time public health monitoring, while Hsu et al.[16] applied edge computing in preprocessing signal data for aircraft engine operation monitoring using deep learning techniques.

In recent years, machine learning has been widely applied in the field of medicine, providing doctors with expanded reference for decision-making. Maini and Dhanka[17] proposed an RBF-SVM model for breast cancer prediction and concluded that the method can effectively enhance prediction. Kumar et al.[18] used a hybrid genetic algorithm model for heart disease prediction and concluded that the method could improve early detection. Uddin et al.[19] stated that Random Forest can achieve excellent prediction results compared to SVM models in disease prediction by many scholars.

Despite these advancements, one critical challenge in critical care remains: accurately predicting the success or failure of ventilator extubation in patients with acute respiratory failure. Incorrect extubation decisions can lead to severe complications, prolonged ICU stays, and increased healthcare costs. To address this issue, this study integrates edge computing and machine learning algorithms, specifically Random Forest and XGBoost, to develop a predictive model for ventilator extubation. By leveraging real-time ventilator-generated data, the proposed system aims to provide clinicians with a reliable decision-support tool, enabling faster and more accurate extubation assessments while reducing server processing loads. The findings of this study have the potential to enhance clinical efficiency, improve patient outcomes, and optimize healthcare resource utilization.

Literature review
Random forest
Random Forest is a decision tree generated by multiple Gini algorithms. Random Forest adds training data in a random manner to achieve the final calculation results, as shown in Eq. (1).

$$\text{Gini}\,(\text{D}) = 1 - \sum_{i=1}^{n} P(i|t)^2 \tag{1}$$

where D is defined as the patient extubation data containing n samples, and $P(i|t)$ is the probability generated by the attribute value. Sharma et al.[20] evaluated six ML models for predicting arrhythmia and found that Random Forest generated better prediction results. Dhanka and Maini[21] compared four machine learning models for predicting heart disease and found that random forests had good predictive power. Huang et al.[22] proposed a model using Random Forest to effectively predict successful extubation in mechanically ventilated patients. Dhanka and Maini[23] used a Random Forest model to predict heart disease. Menon[24] proposed a model using logistic regression and Random Forest to effectively predict the high-cost and non-high-cost categories of patients' medical expenditures. Hanko et al.[25] proposed using Random Forest to generate a model to evaluate the mortality and postoperative results of patients with traumatic brain injury, positing that such an approach could produce good predictive effects. Chandana and Krishna[26] proposed using random forests to generate an algorithm to achieve more accurate predictions for breast cancer. Mursalin et al.[27] proposed applying correlation-based feature selection and Random Forest models to EEG signals to effectively detect epileptic seizures.

XGBoost
XGBoost proposed by Chen and Guestrin[28] provides a fast and scalable way to solve prediction and classification problems in many application fields. It provides early prediction values for the tree root and calculates the residual value of the data set to produce a tree diagram. Dhanka and Maini[29] proposed a novel hybrid XGBoost framework to predict heart disease results. Dhanka and Maini[30] proposed two models named HyOPTRF and HyOPTXGBoost to predict heart disease outcomes. Dhanka et al.[31] compared logistic regression and XGBoost for monitoring coronary artery heart disease, achieving an accuracy of 91.85% after parameter optimization in XGBoost. Kuo et al.[32] proposed combining XGBoost analysis and rule-based methods to achieve a good

monitoring effect for fetal heart rate monitor signals. Prabha et al.[33] proposed the use of photovolume signals and basic physiological data to detect diabetes and suggested that XGBoost can achieve good prediction results with fewer features and lower workload. Półchłopek et al.[34] proposed using time mode for data mining in electronic medical records, finding that XGBoost achieves the highest predictive effect in selected age groups. Tseng et al.[35] suggested the RF + XGBoost model can achieve better preset results than by random forest models alone to predict cardiac surgery–associated acute kidney injury (CSA-AKI). The XGBoost model is mainly generated through addition, as shown in Eqs. (2)–(3).

$$\widehat{y_i} = \sum_{k=1}^{t} f_k\left(x_i\right) \tag{2}$$

$$\widehat{y_i}^{(t)} = \widehat{y_i}^{(t-1)} + f_t(x_i) \tag{3}$$

where $\widehat{y_i}$ is the final predicted value; $f_k\left(x_i\right)$ is the prediction of the k-th tree for the i-th data point, and $\widehat{y_i}^{(t-1)}$ is the prediction form of the previous iteration.

Model evaluation
This research uses Sensitivity, Specificity and Accuracy for model evaluation, as shown in Eqs. (4)–(6).

$$Sensitivity = \frac{TP}{TP + FN} \tag{4}$$

$$Specificity = \frac{TN}{TN + FP} \tag{5}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

where TP is True positive; FP is False positive; TN is True negative; and FN is False negative.

Methods
The hardware setup for this study includes a Raspberry Pi 4 Model B with 4 GB RAM, operating on Ubuntu 22.04. The software implementation is based on Python, incorporating key libraries such as Pandas, NumPy, and Scikit-learn. The research process of this research is mainly divided into four steps: Dataset, Data Preprocessing, Feature Extraction, and Model Training and Evaluation as shown in Fig. 1. The complete architecture of the proposed system is illustrated in Fig. 2, which outlines the integration of edge computing with machine learning for real-time ventilator extubation prediction.

Data collection
This study analyzed demographic and clinical data from a total of 233 patients collected at a hospital in Taiwan. Of these, 28 patients experienced extubation failure, while 205 achieved successful extubation. The dataset includes 3.5 h of continuous physiological recordings per patient, capturing key respiratory parameters such as tidal volume (Vte), respiratory rate (RR), peak airway pressure (Ppeak), mean airway pressure (Pmean), positive end-expiratory pressure (PEEP), and fraction of inspired oxygen (FiO$_2$). A detailed breakdown of the demographic variables is presented in Table 1, while the structure and types of physiological data are described in Table 2.

Data preprocessing
In data preprocessing, this research employs an approach which averages values over specific time intervals to reduce noise and enhance data interpretability. The averaging process is conducted at intervals of 1, 30, 60, 120, 180, and 300 s. This step is crucial for summarizing raw data into more manageable and representative forms, enabling a clearer analysis of trends and patterns.

Feature selection
In feature selection, the Entropy of the original and expanded fields on the success or failure of extraction was calculated by using the Information Gain method proposed by Kullback and Leibler (1951) for feature selection. The entropy value is expressed between 0 and 1, where 0 means the field cannot be partitioned for extubation success or failure, while 1 means that the field can be effectively divided, shown as Eqs. (7)–(9), and the analysis results are shown in Table 3. The original field can predict the extubation data without any difference in the squared data, with good predictive effect for the degree of difference.

Table 4 shows whether using the first n selected feature attributes can achieve a good prediction effect due to feature extraction. Using the first three hours of data for the 233 patients with the Random Forest algorithm, this study finds that the model can obtain good prediction results without deleting any of the six features. Therefore, the feature selection in this research uses the original Vte, RR, Ppeak, Pmean, PEEP, and FiO2 data as model input fields.

$$Entropy\ (Q) = \sum_{i=1}^{j} p_i log_j\left(\frac{1}{p_i}\right) = -\sum_{i=1}^{j} p_i log_j\left(\frac{1}{p_i}\right) \tag{7}$$
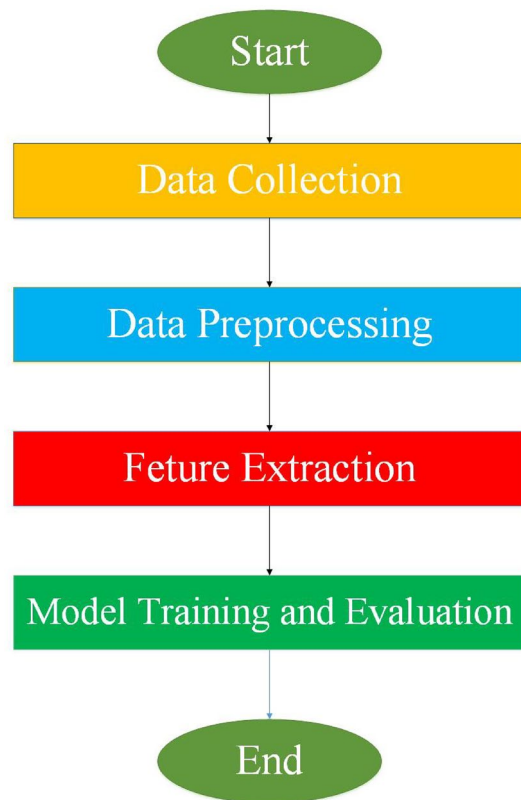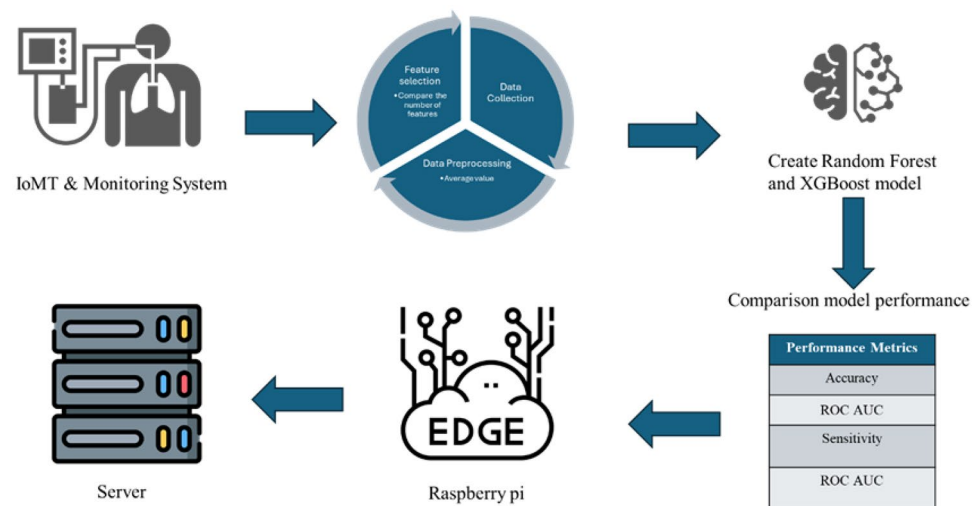
Fig. 1. Research flow.

Fig. 2. Research Structure.

Characteristic	Success (n = 205)	Failed (n = 25)	P-value
Age (year)	73.0 (61.0–82.0)	72.5 (64.5–80.5)	0.980
Gender male (%)	132 (64.4)	15 (53.6)	0.267
APACHEII score	21.0 (16.0–26.0)	20.0 (16.0–25.5)	0.982
Mechanical ventilation duration (days)	3.8 (2.2–7.4)	5.4 (3.0–7.0)	0.311

Table 1. Demographic characteristics table.

Parameter	Data types	Data range
Vte	Numeric	420.6–612.9
RR	Numeric	13.0–20.8
Ppeak	Numeric	15.9–19.0
Pmean	Numeric	7.8–9.4
PEEP	Numeric	4.4–5.8
FiO2	Numeric	30.1–32.0

Table 2. Dataset description.

Parameter	Per 1 s	Per 30 s	Per 60 s	Per 120 s	Per 180 s	Per 300 s
FiO2	0.0608	0.0708	0.0735	0.0734	0.0742	0.0784
Ppeak	0.0340	0.0581	0.0640	0.0722	0.0716	0.0621
PEEP	0.0225	0.0420	0.0481	0.0563	0.0609	0.0606
Pmean	0.0147	0.0248	0.0291	0.0329	0.0312	0.0283
Vte	0.0121	0.0153	0.0166	0.0191	0.0210	0.0215
RR	0.0100	0.0115	0.0118	0.0117	0.0106	0.0105

Table 3. Information gain results.

	3 Feature	4 Feature	5 Feature	6 Feature
Per 1 s	91.32%	92.45%	94.51%	98.80%
Per 30 s	95.35%	98.72%	99.77%	99.86%
Per 60 s	97.10%	99.57%	99.86%	99.88%
Per 120 s	98.59%	99.80%	99.90%	99.91%
Per 180 s	99.22%	99.74%	99.94%	99.92%
Per 300 s	99.53%	99.80%	99.79%	99.93%

Table 4. Model accuracy with n features.

$$\text{Entropy}\left(\text{N}_{di}\right) = \sum_{i=1}^{|N_d|} \frac{N_{di}}{Q} * \text{Entropy}\left(\text{N}_{di}\right) \tag{8}$$

$$\text{IG}\left(\text{N}_d\right) = \text{Entropy}\left(\text{Q}\right) - \text{Entropy}\left(\text{N}_{di}\right) \tag{9}$$

Entropy(Q) is the calculation of the total information content in the entire classification, where Q is the number of data, j is the number of categories, and p_i is the probability of the i-th category appearing in the Q items of data. $\text{Entropy}\left(\text{N}_{di}\right)$ is the information content before the i-th value, category, and data quantity of the d features, and the obtained information content is summed and expressed as the information content of a single feature. $\text{IG}\left(\text{N}_d\right)$ is the information gain obtained for the d-th feature, which is the total information content of the entire data set, which will be generated by the difference between the d-th feature contents to obtain the information gain of each feature.

Model training and evaluation
This research uses four verification methods for model training on the edge device, where M1, M2, M3 Holdout Cross-Validation respectively splits the data from the first 3 h of the 233 data sets into ratios of 6:4, 7:3 and 8:2 for training and Validation sets (respectively 1h48m/1h12m, 2h6m, 54 m, and 2h24m/36 m), while M4 uses a tenfold method to split the data for the first three hours of the 233 sets into one set used for validation, while the rest are used for training. This research splits the data from half an hour of the 233 data sets before extubation for testing set. This research uses two different machine learning methods: Random Forest and XGBoost. Comparison is made with four validation methods, and the model evaluation method should be compared with validation Accuracy. In this study, the specific parameter settings of Random Forest and XGBoost are shown in Table 5.

Results
The results in Fig. 3 show the verification for M1, indicating that Random Forest can achieve the best results by averaging every 120 s, while XGBoost can achieve the best results by averaging Per 180 s; the validation results of XGBoost are slightly lower than Random Forest except for Per 1 s and Per180 seconds. The results in Fig. 4 show

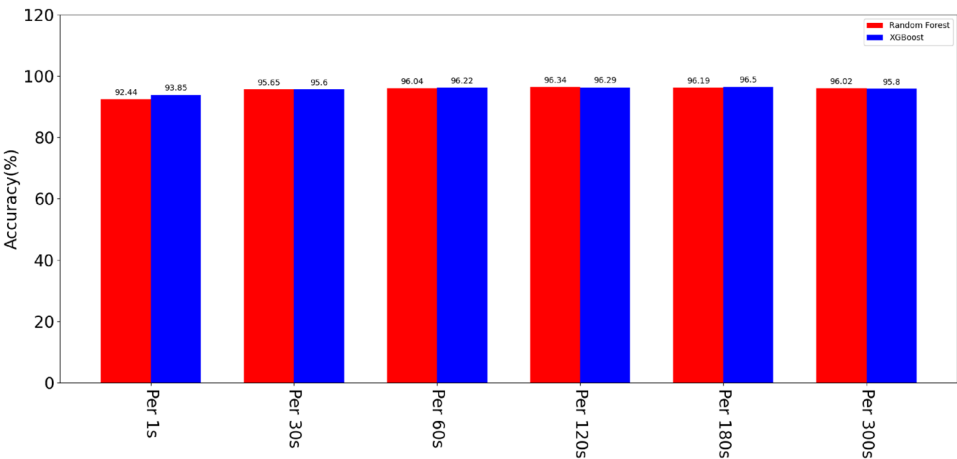| Model name	Parameter name	Parameter value
Random forest	n_estimators	100
	max_depth	None
	min_samples_split	2
Xgboost	Gamma	0
	Eta	0.3
	max_depth	6

Table 5. Model reference settings.
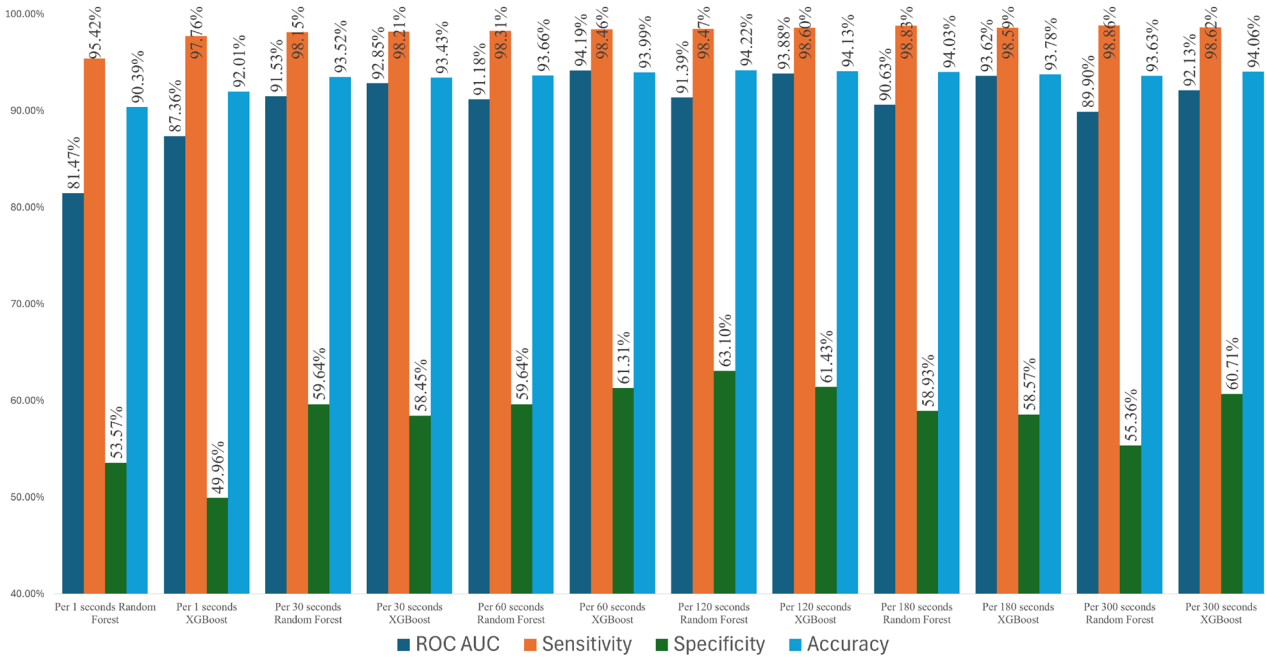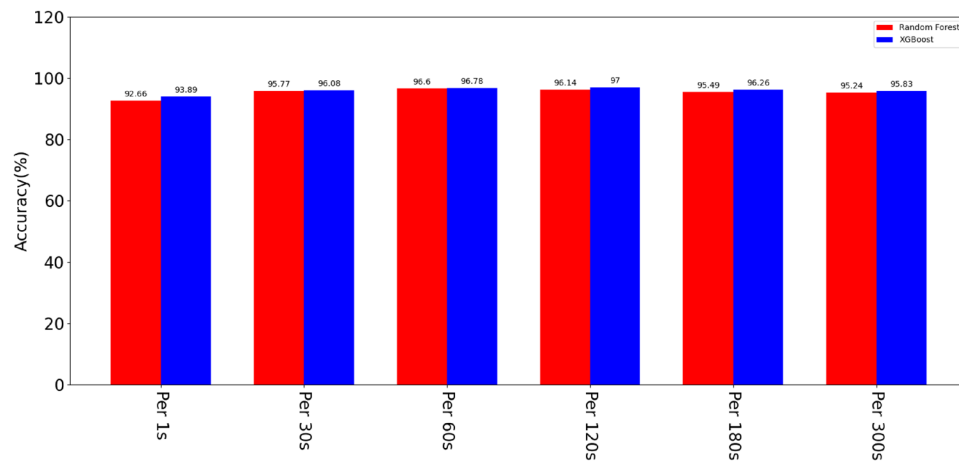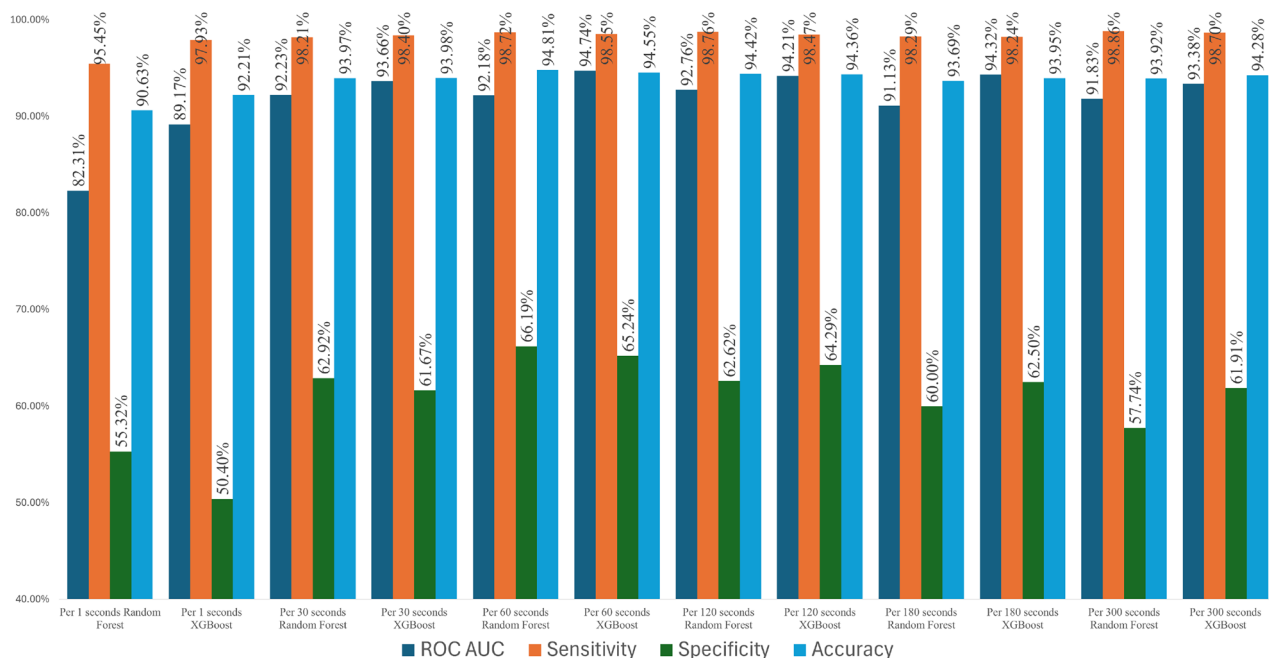

Fig. 3. M1 validation results.

Fig. 4. M1 prediction sample test results.

the ROC values of XGBoost are all better than those of Random Forest, except for the accuracy of Per 30 s, Per 120 s, and Per 180 s, which is slightly lower than that of Random Forest.

The results in Fig. 5 show the verification results for M2, indicating that those of XGBoost are better than Random Forest's, and the results in Fig. 6 show the ROC values of XGBoost are all better than Random Forest,

Fig. 5. M2 validation results.

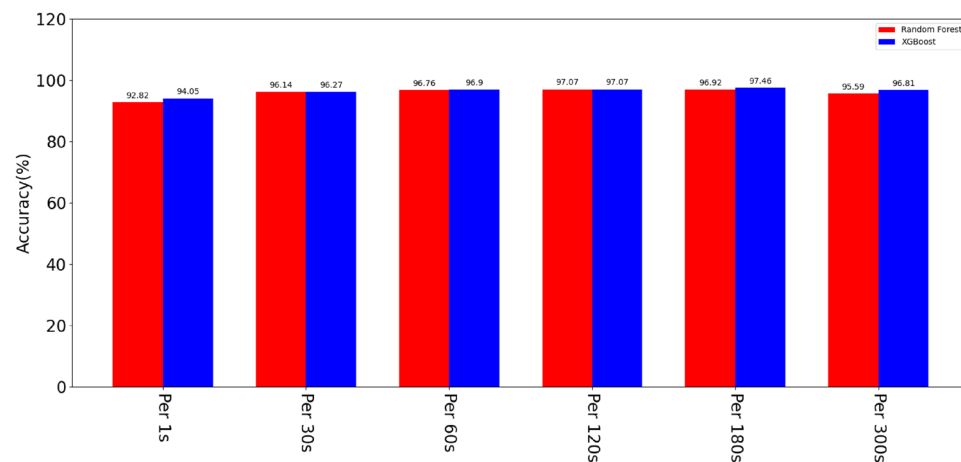**Fig. 6.** M2 prediction sample test results.

and XGBoost is better than Random Forest's except for the accuracy of Per 60 s and Per 120 s, which is slightly lower than Random Forest.

The results in Fig. 7 show the verification for M3, indicating that XGBoost is better than Random Forest and can achieve the best results; the results in Fig. 8 show the ROC values of XGBoost are all from Random Forest, and XGBoost accuracy is better than Random Forest's except that per 120 s is slightly lower than Random Forest every 120 s.

The results in Fig. 9 show the verification for M4, indicating that XGBoost is better than that of Random Forest in Per 1 s, Per 120 s, and Per 300 s, while the others are better than that of Random Forest; the results in Fig. 10 show the ROC values of XGBoost are better than Random Forest's, and XGBoost accuracy is better than Random Forest except that Per 60 s, Per 120 s, and Per 300 s are slightly lower than Random Forest.

Table 6 shows the time required for M1, M2 and M3 to make predictions at different data acquisition frequencies. The results show that when the model is trained using data acquired per second, although a larger amount of data is acquired, this also leads to a longer computation time for the model to perform the prediction. In contrast, averaging the data over a longer time interval effectively reduces the amount of input data, which in turn reduces the computational burden on the model, resulting in a more efficient prediction process and shorter inference time. This suggests that appropriate adjustment of the data acquisition method can help improve the predictive efficiency of the model.

Fig. 7. M3 validation results.
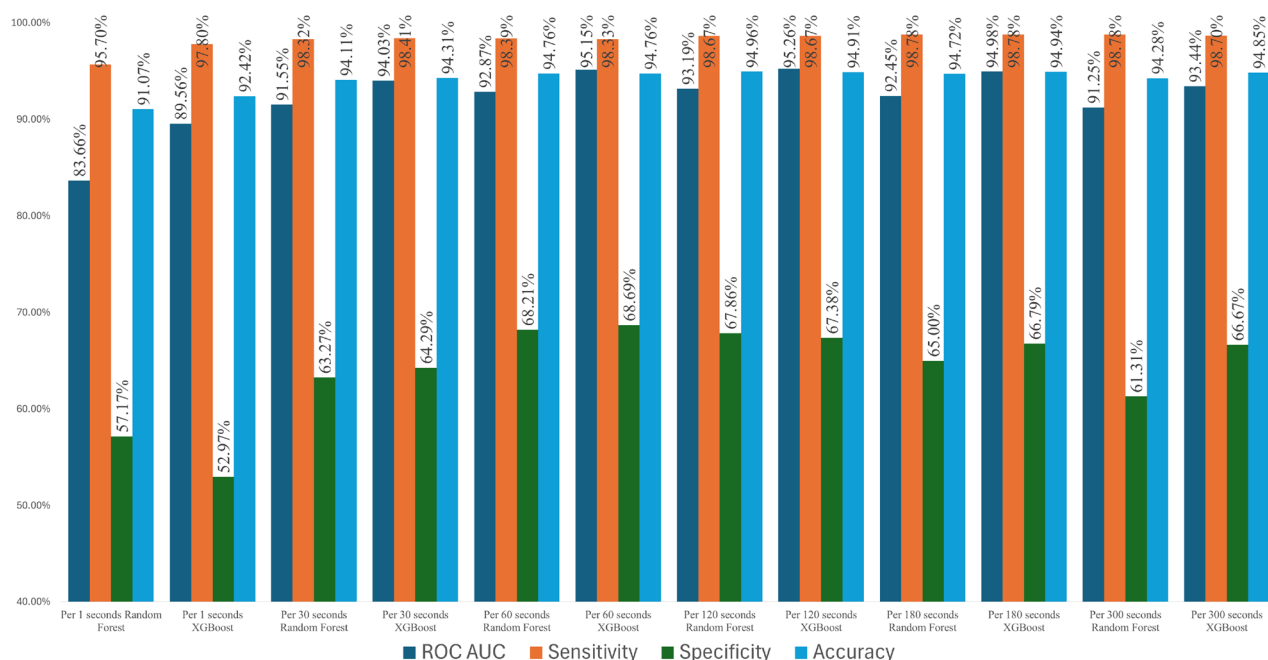
Fig. 8. M3 prediction sample test results.

Discussion

In this study, it is found that XGBoost outperforms Random Forest in different cross validation methods such as 6:4, 7:3, 8:2, and tenfold cross validation, which indicates that Random Forest is stable but not flexible enough to handle low entropy data, whereas XGBoost belongs to the Boosting technique domain. XGBoost is a Boosting technique, which can summarize the data efficiently and accurately derive better prediction results. There is a significant difference between Random Forest and XGBoost in terms of prediction time as XGBoost has faster prediction time.

Since the data used in this study are processed on a per second basis, edge computing can effectively reduce the delay in data transmission because for a patient, if the prediction is performed every three minutes, there will be 1080 pieces of data that need to be uploaded to the server and then averaged and computed, which will lead to the possibility of delay in data transmission. If the prediction is performed on an edge device, it is possible that the data will be delayed. If the computation is performed on the edge device, only the prediction results need to be uploaded to the server for clinical staff to make decisions, which will reduce the data transmission by 83.33%, making the healthcare system more stable and thus reducing the burden on the server. Edge computing can minimize the data leakage problem.

This study found a way to improve the success rate of extubation: real-time AI prediction can not only help clinicians to more accurately determine whether or not the patient is suitable for extubation, as well as reduce
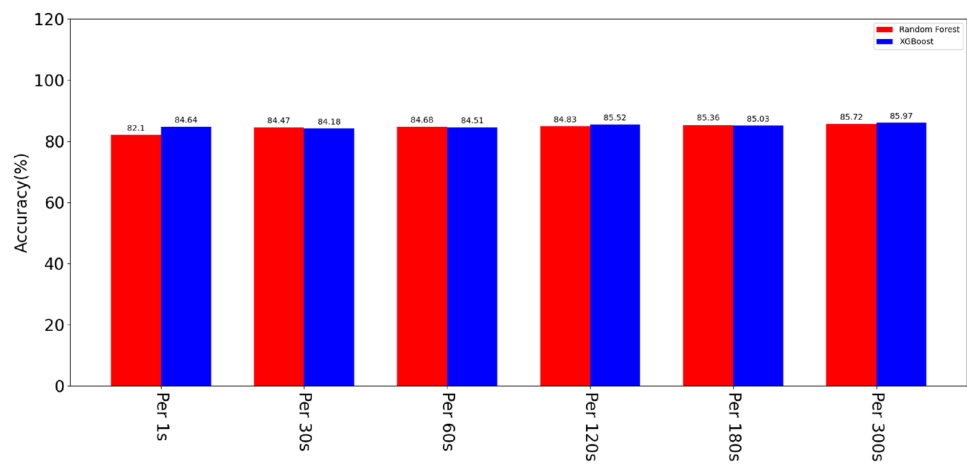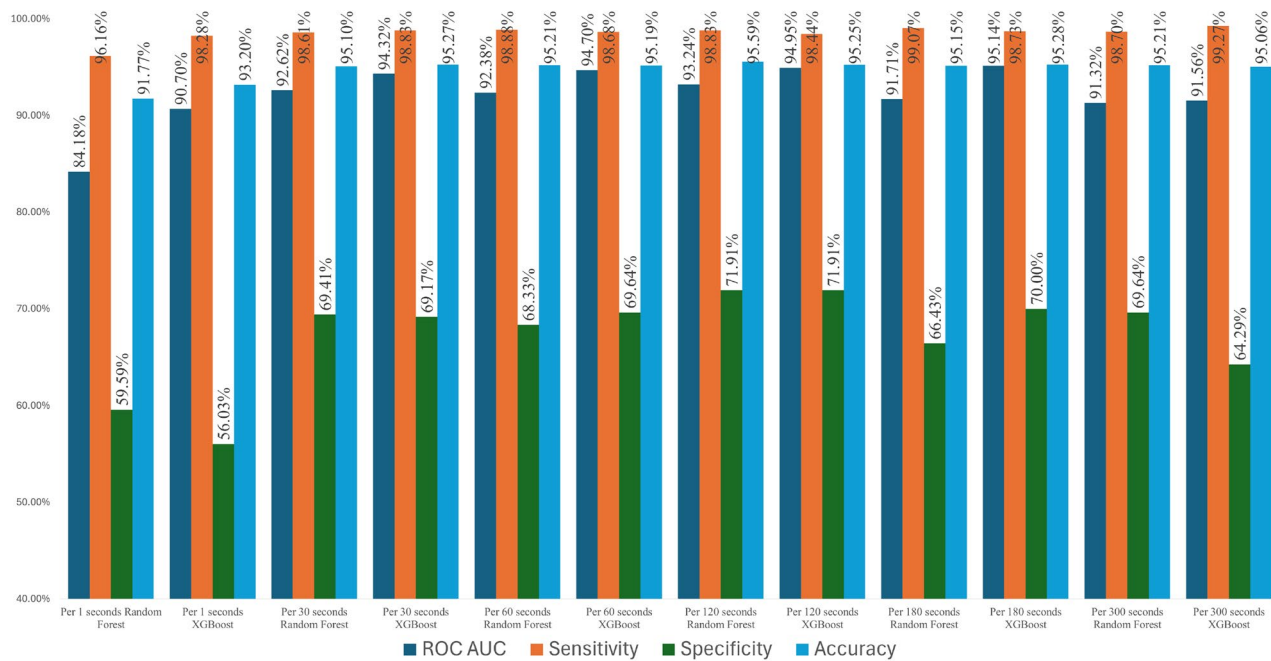
Fig. 9. M4 validation results.

Fig. 10. M4 prediction sample test results.

	M1		M2		M3	
	Random forest	**XGBoost**	**Random forest**	**XGBoost**	**Random forest**	**XGBoost**
Per 1 s	407 s	113 s	477 s	132 s	558 s	153 s
Per 30 s	14 s	3 s	16 s	4 s	19 s	5 s
Per 60 s	7 s	2 s	8 s	2 s	9 s	3 s
Per 120 s	3 s	1 s	4 s	1 s	4 s	1 s
Per 180 s	2 s	< 1 s	3 s	< 1 s	3 s	1 s
Per 300 s	1 s	< 1 s	2 s	< 1 s	2 s	< 1 s

Table 6. Time complexity of model predict on Raspberry Pi.

the failure rate of extubation and the incidence of complications. It can also reduce the failure rate of extubation and decrease the incidence of complications, which in turn can shorten the ICU hospitalization time, reduce the waste of medical resources and increase the recovery rate of patients. In this study, only the parameters of the ventilator can be used to make effective predictions without the need for additional vital signs, which means that it can be more immediate, and the type of the ventilator only needs to be able to allow it to be carried out in the way of the Internet of Things (IoT) in order to use the system of this study.

Currently, only one hospital's data is used in this study to verify the applicability of the inter-organizational testing model, which is an important limitation of the study, and other inter-organizational collaborations can ensure that it can be run stably in different healthcare scenarios. The feasibility of this study can also be demonstrated in different hospital data environments.

Conclusion

In this study, an edge computing-based machine learning approach was implemented to predict ventilator extubation outcomes using real-time ventilator-generated data. The model was trained and validated on an edge device, leveraging different data processing and validation techniques, including Holdout cross-validation and tenfold cross-validation. Among the evaluated algorithms, while Random Forest showed competitive performance in some cases, XGBoost demonstrated superior overall accuracy and efficiency in processing medical data on edge devices. The proposed system effectively reduces server workload by processing data locally before transmission, leading to an 83.33% reduction in data uploads. This architecture enhances system stability, robustness, and sustainability, making it well-suited for real-time clinical applications. Furthermore, this research highlights the potential of integrating patient-specific variables and additional algorithms to refine prediction accuracy in future studies. By enabling physicians to use machine learning-based predictions as a reference for extubation decisions, this approach aims to improve patient management and outcomes in critical care settings.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

References
1. Yin, H. C., Lei, R. L., Xu, J. L., Lin, C. M., & Hsu, Y. L. Enhancing stroke prognosis prediction using deep convolution neural networks. *J. Mech. Med. Biol.* (2025).
2. Huang, K. Y., Chung, C. L. & Xu, J. L. Deep learning object detection-based early detection of lung cancer. *Front. Med.* **12**, 1567119 (2025).
3. Huang, K. Y., Lin, C. H., Chi, S. H., Hsu, Y. L. & Xu, J. L. Optimizing extubation success: a comparative analysis of time series algorithms and activation functions. *Front. Comput. Neurosci.* **18**, 1456771 (2024).
4. Lee, Y. W., Choi, J. W. & Shin, E. H. Machine learning model for predicting malaria using clinical information. *Comput. Biol. Med.* **129**, 104151 (2021).
5. Talukder, A. & Ahammed, B. Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh. *Nutrition* **78**, 110861 (2020).
6. Seo, W., Lee, Y. B., Lee, S., Jin, S. M. & Park, S. M. A machine-learning approach to predict postprandial hypoglycemia. *BMC Med. Inform. Decis. Mak.* **19**(1), 1–13 (2019).
7. Hsiue, E. H. C., Lee, P. L., Chen, Y. H., Wu, T. H., Cheng, C. F., Cheng, K. M., et al. Weaning outcome of solid cancer patients requiring mechanical ventilation in the intensive care unit. *J. Formosan Med. Assoc.* **118**(6), 995–1004 (2019).
8. Su, J., Lin, C. Y., Chen, S. K., Peng, M. J. & Wu, C. L. Characteristics and outcome for very elderly patients (≥ 80 years) admitted to a respiratory care center in Taiwan. *Int. J. Gerontol.* **6**(4), 262–266 (2012).
9. Milbrandt, E. B., Eldadah, B., Nayfield, S., Hadley, E. & Angus, D. C. Toward an integrated research agenda for critical illness in aging. *Am. J. Respir. Crit. Care Med.* **182**(8), 995–1003 (2010).
10. Danaga, A. R. et al. Evaluation of the diagnostic performance and cut-off value for the rapid shallow breathing index in predicting extubation failure. *J. Bras. Pneumol.* **35**, 541–547 (2009).
11. Wu, Y. K., Kao, K. C., Hsu, K. H., Hsieh, M. J. & Tsai, Y. H. Predictors of successful weaning from prolonged mechanical ventilation in Taiwan. *Respir. Med.* **103**(8), 1189–1195 (2009).
12. Yang, P. H. et al. Successful weaning predictors in a respiratory care center in Taiwan. *Kaohsiung J. Med. Sci.* **24**(2), 85–91 (2008).
13. Jiang, Q., Zhou, X., Wang, R., Ding, W., Chu, Y., Tang, S., et al. Intelligent monitoring for infectious diseases with fuzzy systems and edge computing: A survey. *Appl. Soft Comput.* 108835 (2022).
14. Rahman, M. A. & Hossain, M. S. An internet-of-medical-things-enabled edge computing framework for tackling COVID-19. *IEEE Internet Things J.* **8**(21), 15847–15854 (2021).
15. Kong, X. et al. Real-time mask identification for COVID-19: An edge-computing-based deep learning framework. *IEEE Internet Things J.* **8**(21), 15929–15938 (2021).
16. Hsu, H. Y., Srivastava, G., Wu, H. T. & Chen, M. Y. Remaining useful life prediction based on state assessment using edge computing on deep learning. *Comput. Commun.* **160**, 91–100 (2020).
17. Maini, S., & Dhanka, S. Hyper tuned RBF SVM: A new approach for the prediction of breast cancer. In *2024 1st International Conference on Smart Energy Systems and Artificial Intelligence (SESAI)*, 1–4 (IEEE, 2024).
18. Kumar, A., Dhanka, S., Singh, J., Ali Khan, A. & Maini, S. Hybrid machine learning techniques based on genetic algorithm for heart disease detection. *Innov. Emerg. Technol.* **11**, 2450008 (2024).
19. Uddin, S., Khan, A., Hossain, M. E. & Moni, M. A. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med. Inform. Decis. Mak.* **19**(1), 1–16 (2019).
20. Sharma, A., Dhanka, S., Kumar, A. & Maini, S. A comparative study of heterogeneous machine learning algorithms for arrhythmia classification using feature selection technique and multi-dimensional datasets. *Eng. Res. Express* **6**(3), 035209 (2024).
21. Dhanka, S., & Maini, S. Multiple machine learning intelligent approaches for the heart disease diagnosis. In *IEEE EUROCON 2023–20th International Conference on Smart Technologies*, 147–152. (IEEE, 2023).

22. Huang, K. Y. et al. Developing a machine-learning model for real-time prediction of successful extubation in mechanically ventilated patients using time-series ventilator-derived parameters. *Front. Med.* **10**, 1167445 (2023).
23. Dhanka, S., & Maini, S. Random forest for heart disease detection: a classification approach. In *2021 IEEE 2nd International Conference on Electrical Power and Energy Systems (ICEPES)*, 1–3 (IEEE, 2021).
24. Menon, J. PNS88 classifying high medical expenditure patients using logistic regression and random forest methods. *Value Health.* **24**, S188–S189 (2021).
25. Hanko, M. et al. Random forest-based prediction of outcome and mortality in patients with traumatic brain injury undergoing primary decompressive craniectomy. *World Neurosurg.* **148**, e450–e458 (2021).
26. Chandana, C. H., & Krishna, G. B. Breast cancer detection using random forest classifier. *Mater. Today Proc.* (2021).
27. Mursalin, M., Zhang, Y., Chen, Y. & Chawla, N. V. Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier. *Neurocomputing* **241**, 204–214 (2017).
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*, 785–794 (2016).
29. Dhanka, S. & Maini, S. A hybridization of XGBoost machine learning model by Optuna hyperparameter tuning suite for cardiovascular disease classification with significant effect of outliers and heterogeneous training datasets. *Int. J. Cardiol.* **420**, 132757 (2025).
30. Dhanka, S., & Maini, S. HyOPTXGBoost and HyOPTRF: Hybridized intelligent systems using optuna optimization framework for heart disease prediction with clinical interpretations. *Multimed. Tools Appl.* 1–49 (2024).
31. Dhanka, S., Bhardwaj, V. K. & Maini, S. Comprehensive analysis of supervised algorithms for coronary artery heart disease detection. *Expert. Syst.* **40**(7), e13300 (2023).
32. Kuo, P. L., Lim, B. Y., Du, Y. C., Chen, P. F., & Tsai, P. Y. Combination of XGBoost analysis and rule-based method for intrapartum cardiotocograph classification. (2021).
33. Prabha, A., Yadav, J., Rani, A., & Singh, V. Design of intelligent diabetes mellitus detection system using hybrid feature selection based XGBoost classifier. *Comput. Biol. Med.* 104664 (2021).
34. Półchłopek, O. et al. Quantitative and temporal approach to utilising electronic medical records from general practices in mental health prediction. *Comput. Biol. Med.* **125**, 103973 (2020).
35. Tseng, P. Y. et al. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. *Crit. Care* **24**(1), 1–13 (2020).

Author contributions
K.-Y.H.: Writing—original draft. Y.-L.H.: Writing—review and editing. H.-C.C.: Writing—review and editing. M.-H.H.: Writing—review and editing. C.-L.C.: Writing—review and editing. C.-H.L.: Writing—review and editing. C.-S.L.: Writing—review and editing. J.-L.X.: Writing—original draft.

Funding

Declarations

Competing interests
The authors declare no competing interests.

Ethics statement
The study was approved by the Institutional Review Board of Changhua Christian Hospital (approval no.: 210716). The Institutional Review Board waived the need for informed consent considering the retrospective nature of data collected. All methods were performed in accordance with the relevant guidelines and regulations or declaration of Helsinki.

Additional information
Correspondence and requests for materials should be addressed to J.-L.X.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.