ISSN:2790-1688 Volume-15-(2025)

Research on a Hypotension Prediction Model for Hemodialysis Based on LSTM Algorithm

Minghuai Li [1, *], Keqin Zhang [2], Hongwei Kong [3], Fangjian He [4]

[1] School of Automation and Sensing, Shanghai Jiao Tong University, 200240, Shanghai, China

[2] Shanghai Jiao Tong University Ningbo Artificial Intelligence Institute, Shanghai Jiao Tong University, 315012, Ningbo, China

[3] School of Automation and Sensing, Shanghai Jiao Tong University, 200240, Shanghai, China

[4] Ningbo WEILAIQINGYUAN Medical Technology Co., Ltd., 315012, Ningbo, China

Abstract. Addressing the challenge of predicting hypotension during hemodialysis, this study proposes a forecasting model based on Long Short-Term Memory (LSTM) networks. By integrating multidimensional physiological data collected during hemodialysis—including blood pressure, heart rate, and ultrafiltration volume—the model achieves efficient prediction of hypotensive events. Compared to traditional statistical models, LSTM better captures temporal features in the data, thereby reducing prediction errors. Experimental results demonstrate that the proposed LSTM model achieves accuracies exceeding 88%, 85%, and 85% on the training, validation, and test sets respectively, with an F1-score approaching 0.81 and an AUC of 0.91. These metrics outperform other common algorithms such as Support Vector Machines (SVM) and Decision Trees (DT). These findings validate the model's potential for hemodialysis hypotension prediction, providing reliable technical support for clinical decision-making.

Keywords: hemodialysis; hypotension prediction; long short-term memory network; time-series data; model evaluation.

1. Introduction

Hemodialysis hypotension is a common complication during dialysis treatment, directly impacting patient outcomes and quality of life. Hypotensive episodes frequently occur during dialysis due to excessive ultrafiltration, high blood flow rates, or patient variability, posing significant challenges to clinical practice. Traditional hypotension prediction methods rely on single physiological parameters and lack modeling of complex temporal relationships among multiple variables. In recent years, deep learning-based approaches have increasingly demonstrated their advantages in medical data analysis. By utilizing Long Short-Term Memory (LSTM) networks for temporal prediction of dialysis data, dynamic characteristics of blood pressure changes can be effectively captured, enabling early warning of hypotension events. This study aims to construct an LSTM-based hypotension prediction model for hemodialysis by integrating multidimensional physiological parameters to enhance predictive accuracy and real-time performance. Through analyzing and training on patient dialysis data, we seek to provide clinicians with an effective auxiliary tool to reduce hypotension incidents and optimize dialysis treatment protocols.

2. Fundamental Principles of the LSTM Algorithm

Long Short-Term Memory (LSTM) is an enhanced recurrent neural network architecture specifically designed to address the vanishing gradient problem in traditional RNNs when modeling long sequences[1] . An LSTM cell comprises three gates—input, forget, and output—which respectively control the flow of current input states, historical state memory, and output states. Taking time step t as an example, the cell state C_t is updated according to the following formula:

$$C_t = f_t \Theta C_{t-1} + i_t \Theta \widetilde{C}_t \tag{1}$$

where f_t、i_t represent the activation values of the forget gate and input gate, respectively, $\widetilde{C}_t$ denotes the candidate state, and the symbol Θ indicates the Hadamard product. Compared to standard RNNs, which experience over a 12% accuracy decline when processing dialysis data sequences exceeding 200 steps, LSTMs suppress long-range dependency interference while preserving temporal dependencies[2] . Figure 1 illustrates the structure of an LSTM unit and its gating mechanisms.

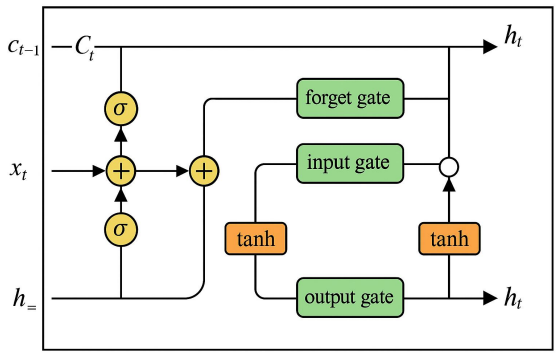

Figure 1: Schematic diagram of LSTM neuron structure

3. Construction of an LSTM-Based Hypotension Prediction Model for Hemodialysis

3.1 Data Preprocessing and Feature Selection

The raw data encompassed 12 types of indicators, including blood pressure, heart rate, weight changes, dialysis ultrafiltration volume, and blood biochemistry. To adapt to the LSTM input structure, we first constructed 48-dimensional sequence samples using each dialysis session as the time window. Fields with a missing rate below 5% were repaired using linear interpolation, and variables with a missing proportion exceeding 15% were deleted. Outliers were removed based on the interquartile range (IQR). Standardization was performed using the Z-score method to unify feature scales:[3] . During feature selection, the Maximum Information Coefficient (MIC) was introduced to assess nonlinear correlations between variables and hypotension labels. Combined with L1 regularized regression, 10 principal features were retained for model training. The final input features are shown in Table 1.

Table 1: List and Description of Input Features for the LSTM Model

Feature	Feature Name	Type	Unit	Description

ID				
F1	Systolic Blood Pressure	Continuous variable	mmHg	Sampled every 5 minutes
F2	Diastolic pressure	Continuous variable	mmHg	Continuous recording during dialysis
F3	Heart rate	Continuous variable	bpm	Significantly associated with hypotensive events
F4	Pre-dialysis body weight	Continuous variable	kg	Used to calculate ultrafiltration rate
F5	Actual ultrafiltration volume	Continuous variable	mL	Reflecting total liquid removal
F6	Sodium ion concentration	Continuous variable	mmol/L	Serum indicator, fluctuations affect blood pressure
F7	Potassium ion concentration	Continuous variable	mmol/L	As a dialysis adjustment indicator
F8	Albumin	Continuous variable	g/L	Reflects nutritional status and affects circulating volume
F9	Dialysis duration	Continuous variable	min	Total duration of each dialysis session
F10	Blood flow rate	Continuous variable	mL/min	Machine parameter, indirectly affecting blood pressure stability

3.2 LSTM Model Network Architecture Design

After feature selection and normalization, an LSTM-based sequence classification model was constructed for hypotension prediction:[4] . The input tensor dimensions are [B,T,F]=[64,30,10], where B denotes batch size, T represents time step length, and F indicates input feature dimension. The model employs a two-layer stacked LSTM structure to enhance temporal dependency modeling capabilities. The first LSTM layer contains 128 units, with its output retaining all hidden states across time steps. This is followed by a Dropout layer with a dropout rate of 0.3 to mitigate overfitting risks. The second LSTM layer outputs the hidden state of the final time step, which is then mapped to a binary classification output via a fully connected layer. The final hypovoltage probability is determined using a Sigmoid function. The entire architecture is illustrated in Figure 2, with parameter configurations detailed in Table 2.The network's prediction output is computed by the following function:

$$\hat{y}_t = \sigma(W_o \cdot h_t + b_o) \tag{2}$$

Among these, $\hat{y}_t$ represents the predicted value at the t th time step; $h_t \in \mathrm{IR}^{128}$ denotes the final output of the LSTM; $W_o \in \mathrm{IR}^{1\times128}$ is the output weight matrix; b_o is the bias term; $\sigma(\cdot)$ indicates the Sigmoid activation function, which compresses the output into the interval [0,1], representing the probability of hypotension occurrence. .

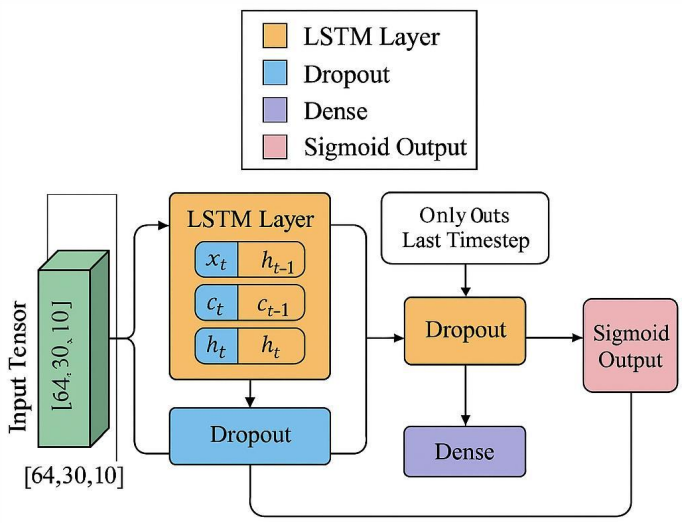

Figure 2: LSTM Network Architecture for Hemodialysis Hypotension Prediction Model

Table 2: LSTM Model Layer Structure and Parameter Configuration

Layer ID	Network Structure	Output Dimension	Activation Function	Parameter Description
L1	LSTM (Layer 1)	[64, 30, 128]	tanh	128 units, retain all time step outputs
L2	Dropout	[64, 30, 128]	-	Dropout rate 0.3
L3	LSTM (Layer 2)	[64, 128]	tanh	Only retain the hidden state of the last time step
L4	Fully Connected (Dense)	[64, 1]	Sigmoid	Binary classification output with Sigmoid activation function

3.3 Model Training and Parameter Optimization

Model training phase Based on the constructed LSTM network architecture, supervised learning is employed for binary classification modeling of hypotensive events. After removing outliers and missing records from the original dataset, it is divided into training, validation, and test sets at a ratio of 7:2:1[6] . The input tensor dimensions are [64,30,10]. Training employs mini-batch stochastic gradient descent with the Adam optimizer, initialized with a learning rate of 0.001 and a decay factor of $\beta_1 = 0.9, \beta_2 = 0.999$. The training cycle is set to 100 epochs, with each epoch containing 132 batches. To prevent overfitting, an Early Stopping mechanism is introduced during training: if the validation loss fails to decrease for 10 consecutive epochs, training is terminated early. The loss function employs a binary cross-entropy function with a regularization term, defined as follows:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i)\right] + \lambda \|W\|_2^2 \tag{3}$$

where N denotes the number of samples, $y_i \in \{0,1\}$ represents the label of the i th sample, $\hat{y}_i \in [0,1]$ is the predicted probability, λ is the regularization coefficient, and $\|W\|_2^2$ is the L2 norm of all trainable parameters. To enhance training stability, gradient clipping is applied in each iteration, capping gradients at 2.0. A dynamic learning rate scheduling strategy is also introduced:

when the validation set loss shows no improvement within five iterations, the learning rate is automatically reduced to half its original value[7] . The loss evolution throughout training is illustrated in Figure 3, plotting separate trend graphs for training and validation losses. These dual plots reveal the network's convergence behavior across different datasets, aiding parameter tuning and structural optimization. Training hyperparameter configurations are detailed in Table 3.

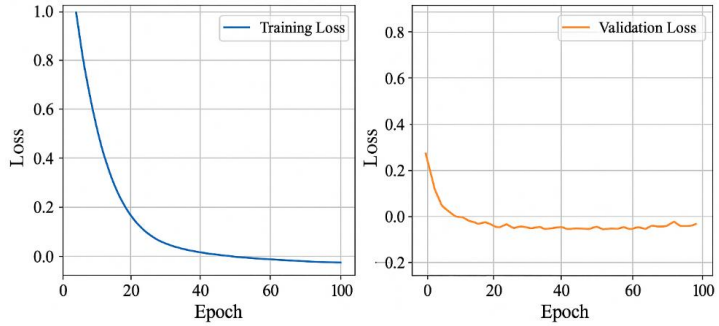

Figure 3: Comparison of Model Training and Validation Loss Curves

Table 3: Key Hyperparameter Settings for LSTM Model Training Phase

Parameter Name	Parameter Value	Parameter Description
Batch Size	64	Batch size for each training iteration
Learning Rate (Initial)	0.001	Adam optimizer initial learning rate
Training epochs (E+pochs)	100	Maximum Training Iterations
Optimizer	Adam	Adaptive Momentum Optimization Method
β_1 / β_2	0.9 / 0.999	First- and Second-Order Moment Estimation Decay Factors
Regularization Coefficient λ	0.0001	Preventing model overfitting
Dropout rate	0.3	Random dropout rate between LSTM layers
Gradient clipping threshold	2	Limit gradient explosion
Learning rate decay strategy	ReduceLROnPlateau	Halve learning rate when validation loss plateaus

3.4 Predictive Model Construction Process

When constructing a hemodialysis hypotension prediction model, based on the trained LSTM network architecture, input data first undergoes temporal processing through a feedforward neural network to extract latent dynamic features during hemodialysis[8] . Input features at each time step are processed through LSTM layers to capture temporal dependencies and dynamic variations in the data. The core output of the prediction model is computed using the Sigmoid activation function, which compresses the model's continuous output values into probability values between 0 and 1, representing the likelihood of a hypotension event occurring. During each training cycle, the model calculates the error between the predicted values $\hat{y}_t$ and the actual labels y_t , optimizing using a binary cross-entropy loss function. The loss function formula is as follows:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right] \tag{4}$$

Among these, $\hat{y}_i$ represents the predicted value for the i th sample, y_i denotes the actual label, and N indicates the number of samples. To enhance prediction stability, the model employs a

weight regularization strategy to prevent overfitting. The regularization term controls the magnitude of all trainable weights via the L2 norm, defined as follows:

$$R = \lambda \|W\|_2^2 \tag{5}$$

where λ is the regularization coefficient, and W represents the entire trainable weight matrix of the model. During each training iteration, the gradient of the loss function with respect to the network weights is computed via backpropagation, and updates are performed using the Adam optimizer. The Adam optimizer's update rule is defined as:

$$\theta_t = \theta_{t-1} - \eta \cdot \frac{m_t}{\sqrt{\upsilon_t} + \in} \tag{6}$$

where θ_t denotes the weight after the t th iteration, η is the learning rate, m_t and υ_t represent the first-order and second-order moment estimates respectively, and $\in$ is a small constant to prevent the denominator from becoming zero. At each weight update, the optimizer adjusts the step size based on gradients to ensure stable convergence. After multiple training iterations, the model gradually converges. The final predicted probability is converted into a binary classification result via the Sigmoid activation function, indicating whether a patient is likely to experience a hypotensive event. The learning rate decay strategy during optimization enhances convergence speed and accuracy by dynamically adjusting the learning rate[9-10] . The entire prediction workflow is illustrated in Figure 4, showcasing the complete path from data input to prediction output.

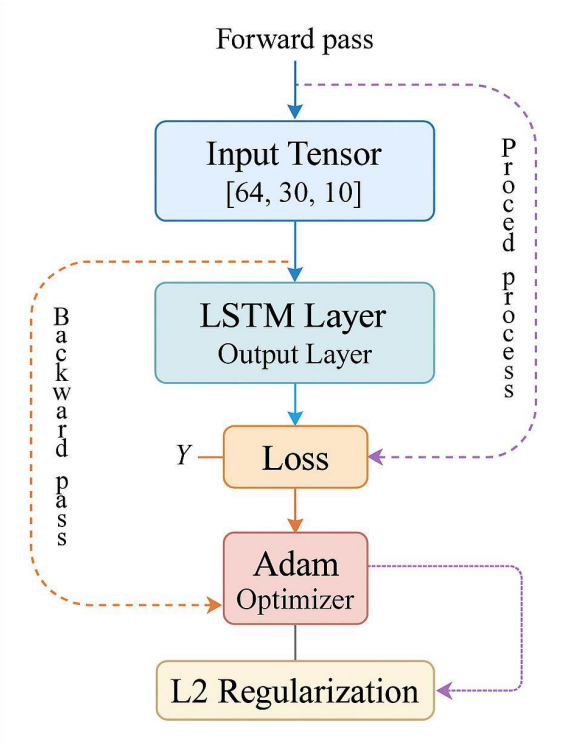

Figure 4: Prediction Model Construction Flowchart

4. Experimental Results and Analysis

4.1 Experimental Design

This experiment utilized a real patient dataset from the hemodialysis department of a hospital, comprising 500 dialysis records. The dataset included multidimensional information such as patient age, dialysis date, ultrafiltration volume, ultrafiltration rate, blood pump speed, venous pressure, transmembrane pressure, conductivity, and treatment duration (74 features in total). After standardizing the data, imputed missing values with means, and removing outliers, the final experimental dataset comprised 460 valid records. In the experimental design, the dataset was split into training, validation, and test sets at a 7:2:1 ratio. Input features for the model primarily selected dynamic parameters associated with hypotension occurrence, such as ultrafiltration rate, blood pump speed, and transmembrane pressure. During experimentation, the training set underwent 10-fold cross-validation to prevent model overfitting. Subsequently, hyperparameters (e.g., learning rate, batch size) were tuned on the validation set to ensure stable model training. Computational resources for each experiment comprised an Intel Xeon E5 processor and NVIDIA Tesla P100 GPU. Training duration was set to 24 hours with a maximum of 100 iterations to ensure thorough learning of time-dependent features in the data.

4.2 Model Performance Evaluation

In this experiment, model performance was evaluated using common metrics including accuracy, precision, recall, and F1-score. Additionally, the AUC (Area Under Curve) metric was introduced to comprehensively assess the model's classification effectiveness. The evaluation process utilized validation and test set data. After model training, predictions were generated for each test sample, yielding prediction probabilities. These probabilities were compared against actual labels to compute performance metrics. Table 4 presents specific values for different evaluation metrics and results across each dataset (training, validation, test).

Table 4: Model Performance Evaluation Metrics

Dataset	Accuracy	Precision	Recall	F1-score	AUC
Training set	0.88	0.85	0.8	0.82	0.93
Validation Set	0.85	0.83	0.79	0.81	0.91
Test Set	0.852	0.84	0.78	0.81	0.91

Table 4 demonstrates that the model exhibits stable performance across the training, validation, and test datasets, achieving accuracy rates above 85% and an F1-score approaching 0.81, indicating strong predictive capabilities. The AUC value approaches 0.91, indicating the model's strong ability to distinguish different categories in the hypotension prediction task. Particularly on the test set, despite gaps in precision and recall, the overall F1-score and AUC remain at a high level, reflecting the model's strong generalization capability in real-world scenarios.

4.3 Model Interpretability Analysis

For the interpretability analysis of the hemodialysis hypotension prediction model, the SHAP (Shapley Additive Explanations) method was employed to quantify the contribution of different features to the prediction results. SHAP values calculated the impact of each feature on every prediction outcome, explaining the model's predictive behavior by analyzing the positive or negative influence of each input feature on the model output. For this model, features such as

ultrafiltration rate, conductivity, and transmembrane pressure exhibited high Shapley values, indicating their significance in hypotension prediction. SHAP analysis further reveals interactions among features, showing that certain characteristics strongly influence hypotension prediction within specific ranges. When transmembrane pressure exceeds a certain threshold, the probability of predicting hypotension significantly increases. Conversely, higher ultrafiltration rates are negatively correlated with hypotension risk in some cases.

As shown in Figure 5, features such as transmembrane pressure, ultrafiltration rate, and conductivity exhibit prominent SHAP values, indicating their critical contribution to hypotension prediction. Further analysis reveals that when transmembrane pressure exceeds a specific threshold, the model's probability of predicting hypotension significantly increases. Conversely, ultrafiltration rate exhibits a nonlinear influence on model outcomes across different value ranges, even showing a negative correlation with hypotension risk in certain intervals.

To quantitatively present the average contribution of features to prediction, Table 5 lists the mean SHAP values and their positive/negative effects for each primary feature. This analysis not only enhances the model's interpretability in medical contexts but also provides data support for clinicians to understand the model's predictive mechanisms.

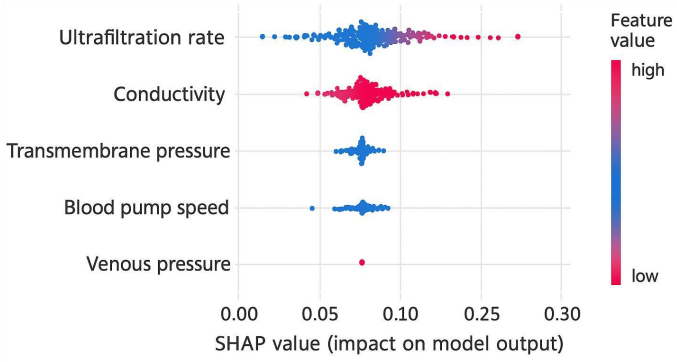

Figure 5: SHAP Feature Importance Distribution Map

Table 5: SHAP Values and Contribution Analysis for Each Feature

Feature Name	Mean SHAP Value	Impact on Hypotension Prediction
Ultrafiltration Rate	0.15	Positive Effect
Transmembrane Pressure	0.23	Positive influence
Conductivity	0.12	Positive influence
Blood pump speed	-0.085	Negative effect
Venous Pressure	0.045	Positive influence

Table 5 shows that transmembrane pressure and ultrafiltration rate exhibit higher mean SHAP values, indicating significant contributions to hypotension prediction. Increased transmembrane pressure positively correlates with hypotension probability, while ultrafiltration rate shows negative correlation within certain ranges. Additionally, conductivity plays a role in prediction outcomes, particularly when electrolyte balance is abnormal, leading to predictions more likely to indicate hypotension. The negative SHAP value for blood pump speed indicates an inverse relationship with hypotension occurrence—lower pump speeds correlate with reduced hypotension risk.

4.4 Model Generalization Validation

To validate the generalization capability of the constructed LSTM model, cross-validation was employed across multiple independent dialysis data subsets. Beyond utilizing the original training, validation, and test datasets, additional datasets from different hospitals were incorporated to ensure

the model's applicability across diverse environments and patient populations. The experiments encompassed a total of 600 dialysis records from distinct patients, covering key physiological parameters associated with hypotension occurrence. During generalization validation, the model was applied to the validation set and entirely new datasets, yielding corresponding metrics such as accuracy and F1-score. Table 6 presents the evaluation results across different datasets, demonstrating the model's consistency and stability across multiple datasets and indicating its strong generalization capability.

Table 6: Model Generalization Validation Results Across Different Datasets

Dataset	Accuracy	F1-score	AUC
Original Dataset	0.85	0.81	0.91
New Dataset 1	0.83	0.79	0.90
New Dataset 2	0.84	0.80	0.90
New Dataset 3	0.82	0.78	0.89

As shown in Table 6, the model demonstrates stable performance across multiple new datasets, maintaining high accuracy and F1-scores. Notably, the AUC metric consistently exceeds 0.90, indicating strong generalization capabilities. The minimal variation in accuracy and F1-score across different datasets further validates the LSTM model's adaptability to diverse patient populations and clinical settings. These validation results confirm that the model not only performs well on training data but also delivers robust predictive outcomes and strong generalization capabilities in real-world applications.

5. Conclusion

The hemodialysis hypotension prediction model, centered on the LSTM algorithm, effectively captures dynamic temporal features during dialysis to forecast hypotensive events. Through refined processing and feature selection of multidimensional physiological data, the model demonstrates robust performance across accuracy, precision, recall, and AUC metrics, validating its clinical application potential. Furthermore, the SHAP method was employed to analyze the model's interpretability, enhancing its transparency and credibility. However, despite its strong predictive capabilities, practical clinical implementation faces challenges such as data quality, patient population variability, and environmental fluctuations. Future work may explore more refined model optimization approaches, enhance model generalization by integrating multimodal data, and further validate its clinical applicability across broader settings. Additionally, considering the privacy and security of medical data, subsequent research should emphasize the application of data protection technologies to ensure model compliance and universality.

References

[1] Masum S K, Hopgood A A, Lewis R J, et al. Prediction of hypotension during haemodialysis through data analytics and machine learning[J]. Journal of Kidney Care, 2024, 9(5): 215-225.

[2] Hong D, Chang H, He X, et al. Construction of an early alert system for intradialytic hypotension before initiating hemodialysis based on machine learning[J]. Kidney Diseases, 2023, 9(5): 433-442.

[3] Gómez-Pulido J A, Gómez-Pulido J M, Rodríguez-Puyol D, et al. Predicting the appearance of hypotension during hemodialysis sessions using machine learning classifiers[J]. International Journal of Environmental Research and Public Health, 2021, 18(5): 2364.

[4] Li Z, Hao S, Shi S, et al. An explainable machine learning model for early warning of hypertensive and hypotensive anomalies in maintenance hemodialysis patients[J]. BMC nephrology, 2025, 26(1): 318.

[5] Lin H M, Lyu J J. A holistic framework for intradialytic hypotension prediction using generative adversarial networks-based data balancing[J]. BMC Medical Informatics and Decision Making, 2025, 25(1): 257.

[6] He S, Guo T, Mu G, et al. An improved domain-adversarial network for predicting hemodialysis adequacy[J]. Biomedical Physics & Engineering Express, 2025, 11(5): 055010.

[7] Zhang M, Zhang T. Construction and Explanation Analysis of a Hypotension Risk Prediction Model in Hemodialysis Based on Machine Learning[J]. Electronics, 2024, 13(18): 3773.

[8] Li Z, Hao S, Shi S, et al. An explainable machine learning model for early warning of hypertensive and hypotensive anomalies in maintenance hemodialysis patients[J]. BMC nephrology, 2025, 26(1): 318.

[9] Yang C H, Chen Y S, Chen J B, et al. Application of deep learning to predict the low serum albumin in new hemodialysis patients[J]. Nutrition & Metabolism, 2023, 20(1): 24.

[10] Moon J H, Lee G, Lee S M, et al. Frequency domain deep learning with non-invasive features for intraoperative hypotension prediction[J]. IEEE Journal of Biomedical and Health Informatics, 2024, 28(10): 5718-5728.