Short-term prediction for dynamic blood glucose trends based on ARIMA-LSSVM-GRU model

Yongjun Zhang1*,2, Yufu Ning1,2, Baotian Li 1, Yanfeng Liu 1 and Yujun Dang1

1 School of Information Engineering, Shandong Youth University of Political Science, Jinan, Shandong, 250103, China
2 Key Laboratory of Intelligent Information Processing and Information Security in Universities of Shandong, Jinan, Shandong, 250103, China
*Corresponding author’s e-mail: Yongjunzhang@sdyu.edu.cn

Abstract. Continuous glucose monitoring (CGM) is an effective strategy to dynamically monitor the patient's blood glucose (BG) levels. The current existing prediction models, such as ARIMA, LSSVM, GRU, LSTM are commonly used for the prediction of BG changes in the process of health monitoring and evaluation. However, these single models cannot obtain the optimal prediction results for the BG series, which possess both linear and nonlinear characteristics. Given the above limitation, a hybrid ARIMA-LSSVM-GRU model based on ARIMA, LSSVM, and GRU technologies is proposed to predict the forthcoming BG trends. This model utilizes an ARIMA model to capture the linear features and predict the BG series. Then, the least-squares support vector machine is used to predict the error series that is generated by ARIMA. Finally, the GRU model is used to combine the prediction results of ARIMA and LSSVM to get the more accurate upcoming BG trends. To test the accuracy of this hybrid model, the BG series is prepared, trained, and tested through real clinical BG monitoring data set. The ARIMA-LSSVM-GRU model is been compared with other prediction models (ARIMA, LSSVM, GRU, LSTM) by the evaluation criteria RMSE, MAPE, TIC. The experimental results show that this model can effectively improve the accuracy of the short-term BG prediction.

1. Introduction
Continuous glucose monitoring (CGM) provides an effective strategy to continuously monitor the blood glucose levels of diabetics, which provides more information support for prediction models to forecast the future BG trends and effectively prevents the occurrence of abnormal events [1-3]. To predict the dynamic BG series will timely facilitate the preventive measures against health risks induced by its abnormal trends. However, due to the BG series usually have both linear and nonlinear characteristics, the prediction models cannot effectively and precisely predict the future BG results [4]. Therefore, it has become one of the most challenging applications to construct the BG series forecasting model and accurately predict future time-varying trends. Jun Yang et al. establishes an ARIMA model with adaptive orders to continuously predict the dynamic BG concentrations and it outperforms the adaptive univariate model and ARIMA model [5]. JaouherBen et al. propose adopted several artificial neural network approaches for the prediction of Type 1 Diabetes [6]. Some adaptive artificial neural networks have been proposed and been proved to be accurate, adaptive, and very encouraging for clinical implementation. Kezhi et al. proposed a CRNN algorithm for glucose level forecasting using a clinical dataset [7]. The proposed method showed superior performance of clinical experiments in BG level forecasting. In addition, the deep neural network that is represented by LSTM has also become a research
hotspot in recent years and the application is conducted to do early hypoglycaemic alarms by the CGM data trends[8]. Ta et al. proposed accurate prediction of continuous blood glucose-based SVR and differential evolution methods to achieve high prediction accuracy [9-10]. However, the above models all assume that the time series is a linear or nonlinear characteristic, but in practical application, there may be both of these characteristics, which affect the performance of the BG prediction models. To solve these problems, this research proposes a new strategy for the combination of linear and nonlinear model prediction results. Firstly, the ARIMA model is used to extract the linear features of the time series, then the LSSVM model is used to predict the error series between the predicted value and the actual value of the ARIMA model. Finally, the prediction results of the former two models are used as the input of the GRU model to predict the time series, and dropout technology is applied to optimize the GRU model for further accurate BG prediction.

2. Methodology

2.1. ARIMA model

The autoregressive integrated moving average process model (ARIMA) is a linear model proposed by Box and Jenkins for time series analysis and prediction. In time series analysis, three parameters need to be set, namely autoregressive order (p), difference order (d), and moving average order (q). The general form \(ARIMA(p,d,q)\) is as the following equations (1) to (3).

\[
\begin{align*}
\phi(B)\nabla^d y_t &= \theta(B)e_t \\
E(e_t) &= 0 \quad \text{var}(e_t) = \sigma^2_e, \quad E(e_t e_s) = 0, s \neq t \\
E(x_t e_s) &= 0 \quad s < t
\end{align*}
\]

\[
\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p
\]

\[
\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q
\]

Where, \(B\) is the backward operator. \(\nabla^d = (1-B)^d\) is the high order difference. \(\phi_i (i = 1, 2, \cdots, p)\) and \(\theta_j (j = 1, 2, \cdots, q)\) are autoregressive parameters and moving average parameters respectively. \(e_t\) is a series of Gaussian random variables with mean 0 and variance \(\sigma^2\). Equations (2) and (3) are polynomial of autoregressive correlation coefficient and polynomial of moving average coefficient respectively. This model is utilized to extract the linear features of the BG series and establish the foundation for BG series prediction.

2.2. LSSVM regression model

The LSSVM has played an important role in many fields including data regression, pattern recognition, time series prediction, etc. Figure 1 illustrates the structure diagram of LSSVM.

![Figure 1. Structure of LSSVM regression model.](image)
For the given N training data pairs \((x_i, y_i)\), \(i = 1, 2, ..., N\), where \(x_i = (x_{i1}, x_{i2}, ..., x_{id})^T \in \mathbb{R}^d\) is the input vector, and \(y_i \in \mathbb{R}\) is the corresponding output value, the input data can be mapped to a higher dimensional space from the original feature space \(\mathbb{R}^d\) by adopting a nonlinear function \(\phi(x)\). \(K(x_i, x)\) is the kernel function, satisfying Mercer's conditions. The kernel function plays an important role in constructing high-performance LSSVM, with the ability to reduce the computational complexity in high-dimensional space. Then the LSSVM model can be constructed as follows:

\[
y(x) = \sum_{i=1}^{N} \alpha_i K(x_i, x) + b
\]

(4)

2.3. Gated Recurrent Unit (GRU) model
Compared with Long Short-Term Memory (LSTM) model, GRU is a kind of Recurrent Neural Network (RNN) and it has the same data mining ability with the improved gate control structure. It combines the input gate and forgetting gate into the gate \(tz\), and replaces the output gate of LSTM with the gate \(tr\). Among them, \(z_i\) determines the integration of historical information of new input information, \(r_i\) determines the proportion of the previous state information into the model. The training parameters are reduced in the GRU structure due to its reduced door numbers. Therefore, the training speed is greatly improved. The mathematical expression of the GRU is described as follows:

\[
\tilde{s}_i = \phi_{\text{tanh}} \left( W'_r (r_i \odot s_{i-1}) + U_x x_i + b_z \right)
\]

(5)

\[
s_i = (1 - z_i) \odot s_{i-1} + z_i \odot \tilde{s}_i
\]

(6)

\[
z_i = \sigma_{\text{sig}} \left( W'_z s_{i-1} + U_x x_i + b_z \right)
\]

(7)

\[
r_i = \sigma_{\text{sig}} \left( W'_r s_{i-1} + U_x x_i + b_r \right)
\]

(8)

Where \(\odot\) stands for the product formula of elements. \(W'_r\) and \(W'_z\) are the weight matrix of the gate \(r_i\) and \(z_i\). \(W_x\) is the weight matrix of the output state. \(x_i\) is the input data at time t. \(\tilde{s}_i\) and \(s_i\) are candidate states and output states at time t, respectively. \(b_z\), \(b_r\) and \(b_r\) are the constants. \(\sigma_{\text{sig}}\) and \(\phi_{\text{tanh}}\) are respectively sigmoid and tanh activation functions, which are used to activate gates. The expressions of sigmoid and tanh functions are illustrated as formulas (9) and (10).

\[
tanh(x) = 2 \cdot \text{sig}(x) + 1
\]

(9)

\[
\text{sig}(x) = \left(1 + e^{-x}\right)^{-1}
\]

(10)

Since the BG series are prepared already, they are utilized to train the GRU model. In the training process, the parameter matrices \(W_r, W'_r, W_z, U_r, U_z, U_x\) are modified in each iteration, which follows the backpropagation through the time procedure of typical RNN. After the learning process, a GRU is trained to be constructed and competent for time series forecasting tasks.

2.4. Dropout technology
The increasing number of layers, neurons, and parameters in the deep neural network will improve its performance. But too large parameters or the less training data set can easily lead to the problem of overfitting. At present, the method to prevent overfitting is to add the inactivation regularization weight
constraint technology to the training data set. In the training process, dropout technology randomly discards the hidden neurons in the network according to a certain probability, so that these neurons do not participate in forward and backpropagation in the training process, but their weights are still retained, as shown in figure 2. It shows the results of a standard neural network processed by dropout technology, in which the solid circle represents the neurons randomly discarded by dropout technology. The optimal dropout probability in this research is set to 0.3.

Figure 2. Schematic diagram of dropout technique.

3. ARIMA-LSSVM-GRU model
This research presents a hybrid ARIMA-LSSVM-GRU prediction model that is based on ARIMA, LSSVM, and GRU technologies. The hybrid model consists of three parts. Figure 3 shows the prediction process of the combined prediction algorithm and the specific components of this model are described as follows.

Figure 3. The BG forecasting procedures of the ARIMA-LSSVM-GRU model.
The first part is the linear ARIMA model, which realizes the one-step prediction of blood glucose time series \( tBG \), and the prediction result is expressed as \( \text{Result}_t \). The second part is the LSSVM prediction for error \( \Delta_t \) that is forecasted by the ARIMA model. The prediction result is expressed as \( \hat{\Delta}_t \). In the last part of the ARIMA-LSSVM-GRU model, the GRU model is used to non-linearly combine the prediction results \( \text{Result}_t \) and \( \hat{\Delta}_t \), and the prediction results are expressed as \( \hat{tBG}_t \).

4. Results and discussion

The configuration of our experimental environment is as follows. We take Ubuntu 16.04 as the operating system to establish the BG forecasting experiments. Python 3.6 is taken as the development language and the testing framework is TensorFlow2.2. The data in this study are acquired by the CGM system (RGMS-III, MeiQi) in a municipal hospital (Jinan, Shandong province of China). CGM device measures the blood glucose data about every 3 minutes to obtain the continuous BG data during three days for training and testing of our proposed prediction model.

4.1. Data preprocessing

In neural network training, the dimensional difference between data plays a key role in whether the network training can converge and the prediction accuracy. Therefore, it is necessary to pre-process the input time series data before the prediction modeling. In this paper, formula (11) is used to normalize the sample data to the interval \([0, 1]\) for the needs of modeling and prediction.

\[
x(t)_{\text{normal}} = \frac{x(t) - \text{Min}[x(t)]}{\text{Max}[x(t)] - \text{Min}[x(t)]}
\]

(11)

Where, \( t \) is time series. \( x(t) \) is the original BG series. \( x(t)_{\text{normal}} \) is the normalized BG series. \( \text{Min}[x(t)] \) and \( \text{Max}[x(t)] \) are the minimum and maximum values of the training data set respectively. The output BG series after training and modeling are denormalized to obtain the predicted value.

4.2. Evaluation criteria

To evaluate the prediction performance of the ARIMA-LSSVM-GRU model, this research uses the root means square error (RMSE), average absolute percentage error (MAPE), as well as Theil inequality coefficient (TIC) as evaluation indicators to measure the prediction accuracy. The expressions of evaluation criteria are as follows.

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{x}_i)^2}
\]

(12)

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{x_i - \hat{x}_i}{x_i} \right| \times 100
\]

(13)

\[
TIC = \frac{\sqrt{\sum_{i=1}^{n} (x_i - \hat{x}_i)^2} / n}{\sqrt{\sum_{i=1}^{n} x_i^2} / n + \sqrt{\sum_{i=1}^{n} \hat{x}_i^2} / n}
\]

(14)
In equations, \( x_i \) is the actual value of the \( i \)th sample. \( \hat{x}_i \) is the forecasted value. \( n \) is the number of samples in the BG data series. Theil inequality coefficient is between 0 and 1. The closer the TIC value is to 0, the higher the fitting degree between the predicted value and the actual value.

### 4.3. Experiment and result analysis

This paper constructs and compares a batch of short-term BG prediction models. It is illustrated in figure 4. The prediction results of a patient's blood glucose test data set, in which the step length in the prediction process is set to 30-min and 60-min respectively. From the comparison of the prediction results with different forecasting steps, it is easily seen from the forecasted results that the ARIMA-LSSVM-GRU model is better than the other traditional models (GRU, LSTM, ARIMA, LSSVM).

![Figure 4. The prediction results in the BG series by using the ARIMA-LSSVM-GRU.](image)

The dynamic BG prediction results of different models with different prediction steps were evaluated from the prediction results. It shows the ARIMA-LSSVM-GRU model has an advanced performance than the other models. The prediction evaluation has illustrated as follows in detail.

![Figure 5. Error comparisons of different blood glucose prediction models.](image)
From experimental results, it can be concluded from above figure 5 that the prediction deviations generally are increased with the extension of prediction steps. The prediction error that is measured by RMSE is reduced by 22.7%, 25.1%, 25.2%, and 23.0% than those prediction effects of GRU, LSTM, ARIMA, and LSSVM algorithm respectively in the 60-minute predicting step. Meanwhile, the MAPE is decreased by 14.1%, 18.2%, 27.3%, and 25.1% according to the above comparison orders. The result of TIC indicates the prediction error is reduced by about 14.3%, 18.1%, 26.4%, and 25.7%. The BG forecasting precision is improved greatly when the prediction step length is shortening to 30-minute. The RMSE is decreased by 24.2%, 25.7%, 34.4% and 35.4%. The MAPE is also reduced by 14.2%, 17.9%, 19.52%, 19.49%, and TIC is diminished by 16.9%, 20.3%, 31.6%, 30.2% respectively. From the above prediction result comparison and analysis, the evaluation results show that our fusion algorithm greatly improves the accuracy of dynamic BG trends prediction.

5. Conclusions
This paper proposes a hybrid ARIMA-LSSVM-GRU model for BG series prediction. Firstly, the combination of ARIMA and LSSVM models is jointly used to extract the linear features of the BG series and the nonlinear features of the error series separately. After that, the GRU model with dropout optimization is utilized to combine the linear and nonlinear prediction results for the accurate prediction of the future BG trends. This study makes an empirical analysis of the BG series of different experimental samples. The experimental results show that our proposed model improves about 24.0% and 21.2% by RMSE and MAPE respectively when compares with other models (GRU, LSTM, ARIMA, LSSVM) in 60-min predication. At the same time, TIC evaluation also shows that the prediction result in this region is increased by 21.1% than the other methods. The prediction accuracy is also largely improved in 30-min than it in 60-min. In conclusion, the ARIMA-LSSVM-GRU prediction model can effectively improve prediction accuracy. Therefore, it is a better short-term BG prediction method for a certain practical application. In the future, our proposed hybrid model should be further optimized to make it more suitable for time-varying BG prediction.

Acknowledgments
This research is supported by the 2019 Key R&D Program of the Shandong province public welfare project (2019GNC106079) and the second batch of an Industry-University Cooperative Education Project of China's Ministry of Education in 2020 (Title: Research on teaching and curriculum reform of introduction to IoT- based on CDIO mode).

References
[1] J Li, T Igbe, Y Liu, A Kandwal, Z Nie (2021) Non-invasive monitoring of three glucose ranges based on ECG by using DBSCAN-CNN. IEEE Journal of biomedical and health informatics. J.7:1-12.
[2] A Zaitcev, MR Eissa, H Zheng, T Good, M Benaisa. (2020) A Deep Neural Network Application for Improved Prediction of in Type 1 Diabetes. IEEE J Biomed Health Inform. J.24(10): 2932-2941.
[3] Shuai Gao, Yuefei Huang, Shuo Zhang. (2020) Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation. Journal of Hydrology. J. 589:125188.
[4] K Li, C Liu, T Zhu, P Herrero, P Georgiou. (2020) GluNet: A Deep Learning Framework for Accurate Glucose Forecasting. IEEE Journal of biomedical and health informatics. J.24(2): 414-423.
[5] Yang, J., et al. (2019) An ARIMA Model with Adaptive Orders for Predicting Blood Glucose Concentrations and Hypoglycaemia. IEEE J Biomed Health Inform. J. 23(3):1251-1260.
[6] Jaouher Ben, Ali Takoua Hamdi, Nader Fnaiech. (2018) Continuous blood glucose level prediction of Type 1 Diabetes based on Artificial Neural Network. Biocybernetics and Biomedical Engineering. J. 38(4): 828-840
[7] Li, K., et al. (2020) Convolutional Recurrent Neural Networks for Glucose Prediction. IEEE J Biomed Health Inform. J.24(2): 603-613.
[8] Wang, W., M. Tong and M. Yu. (2020) Blood Glucose Prediction with VMD and LSTM Optimized by Improved Particle Swarm Optimization. IEEE ACCESS. J.8: 217908-217916.
[9] Costanzo, T.H. J.A. (2018) Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm. Biocybernetics and Biomedical Engineering. J. 38:362-372.
[10] X Yu, K Turksoy, M Rashid, J Feng, N Hobbs. (2018) Model-fusion-based online glucose concentration predictions. Control Engineering Practice. J. 7:129-141.