Indonesian Journal of Electrical Engineering and Computer Science
Vol. 40, No. 1, October 2025, pp. 530~544
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v40.i1.pp530-544

Predict glucose values with DE algorithm optimized T-LSTM

QingXiang Bian[1], Azizan As'array[1], XiangGuo Cong[2], Khairil Anas bin Md Rezali[1], Raja Mohd Kamil bin Raja Ahmad[3], Mohd Zarhamdy Md. Zain[4]
[1]Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Malaysia
[2]Department of Endocrinology, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Suzhou, China
[3]Department of Electric and Electronics Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Malaysia
[4]Department of Applied Mechanics and Design, Faculty of Engineering, Universiti Teknologi Malaysia, Johor, Malaysia

Article Info

Article history:

Received Dec 30, 2024
Revised Apr 13, 2025
Accepted Jul 4, 2025

Keywords:

Continuous glucose monitoring
Differential evolution
Long short-term memory
Predicting algorithm
Transformer

ABSTRACT

The prevalence of diabetes is rising. According to the International Diabetes Federation (IDF) predictions, the number of diabetic patients worldwide will reach 608 million in 2030, accounting for approximately 11.3% of the total number of people in the world. To monitor and predict the future 1 hour glucose have a great significance meaning for patients. This research utilizes a differential evolution (DE) algorithm, an optimized hybrid model transformer and long short-term memory (T-LSTM) technologies to analyze historical data from continuous blood glucose monitoring (CGM) systems and equipment calibration values. The aim is to predict future blood sugar levels in patients, thereby helping to prevent episodes of hypoglycemia and hyperglycemia. The study tested the model using the CGM data from 8 patients at the Suzhou Municipal Hospital in Jiangsu Province, China. Results show that this DE-optimized T-LSTM model outperforms traditional models. The model's accuracy is evaluated using mean squared error (MSE), with MSE values recorded at 15, 30, and 45 minutes being 0.96, 1.54, and 2.31, respectively.

Corresponding Author:

Azizan As'array
Department of Mechanical and Manufacturing Engineering, Faculty of Engineering
Universiti Putra Malaysia
43400 Serdang, Selangor, Malaysia
Email: zizan@upm.edu.my

1. INTRODUCTION

Diabetes mellitus (DM) poses a significant global health challenge, affecting millions worldwide. Effective management of DM necessitates precise forecasting of blood glucose levels to prevent complications such as hypoglycemia and hyperglycemia. The increasing prevalence of diabetes globally necessitates advancements in monitoring technologies that can provide accurate blood glucose readings. Along with the development, continuous blood glucose monitoring (CGM) systems enable real-time glucose monitoring for patients. However, the market lacks products capable of forecasting glucose values, a feature primarily explored in academic research [1], [2]. Traditional predictive models often fall short in capturing the complex, nonlinear, and time-varying nature of blood glucose dynamics. Recent advancements in machine learning, particularly long short-term memory (LSTM) networks, have shown promise in modeling temporal dependencies in glucose time-series data. However, the performance of LSTM models can be further enhanced by optimizing their parameters. Incorporating differential evolution (DE) algorithms for this optimization has been explored in several studies. To address this, this study introduces a DE algorithm-optimized T-LSTM prediction model. This innovative model merges the capabilities of Transformer and

LSTM, optimizing time series forecasting by leveraging their strengths. These include the ability to model long-term dependencies, execute parallel computations, learn feature representations, and handle multi-scale data. The combination of Transformer and LSTM in this model considerably enhances forecasting accuracy and efficiency, particularly in spotting trends, periodic patterns, and long-term relationships in time series data [3].

Traditional intermittent blood sugar testing does not reflect a patient's genuine blood sugar changes over a 24-hour period [4]. In contrast, CGM systems could show real-time glucose values every 3 or 5 minutes [2], [5]. This frequent monitoring is crucial for evaluating treatment effectiveness and making more accurate adjustments to therapeutic plans [6]. CGM systems are adept at predicting the onset of hyperglycaemia or hypoglycaemia, allowing for a more personalized and targeted approach to enhancing patient quality of life. By alerting patients to irregular glucose levels, CGM-based prediction models empower patients to take timely action [7]. Additionally, these models can more accurately replicate the specific physiological traits of diverse populations and reveal varying responses to treatments among patients [8], [9]. Integrating data from multiple sources enables the creation of customized treatment plans, aiding patients in better controlling their blood sugar levels and minimizing complications [10]. CGM devices provide immediate insights into glucose levels, thereby facilitating informed decisions about insulin dosage adjustments or dietary changes based on blood sugar fluctuations [11]-[13].

Research in the field of diabetes management has increasingly focused on the use of advanced control strategies to enhance the performance of CGM systems. For instance, studies by researchers [14], [15] have explored various algorithmic approaches, including the application of machine learning techniques such as support vector machines (SVM) and Reinforcement Learning, which have shown promise in improving the accuracy of glucose forecasts. These studies underscore the potential of integrating sophisticated computational models with traditional CGM systems to enhance predictive accuracy and patient outcomes.

Recent advancements in artificial intelligence, particularly in machine learning and deep learning, have led to progress in CGM prediction algorithms. These advancements have turned such algorithms into a focal point of research in diabetes management [16]-[21]. The T-LSTM model has exceptional capabilities in parallel processing, handling long-range dependencies, and adaptability across various domains, make it an intriguing option for this application [22], [23]. This study introduces a novel approach, employing the T-LSTM model to forecast future blood glucose levels. The model is designed to be patient-specific, adapting and improving over time to match the individual's metabolic characteristics. The key contribution of this article is the development of this innovative combined T-LSTM model, aiming to predict glucose levels for upcoming intervals of 15, 30, and 45 minutes. Despite these advancements, challenges persist in achieving high-precision glucose predictions across diverse patient populations. Variability in individual responses necessitates personalized modeling approaches. Additionally, the interpretability of machine learning models remains a critical concern, especially when they inform therapeutic decisions. Studies have highlighted the importance of model interpretability in diabetes management, emphasizing that understanding the rationale behind model predictions is crucial for patient safety and trust.

This study introduces a novel DE algorithm optimized transformer and a T-LSTM model that leverages the strengths of both Transformer and LSTM technologies. This hybrid model aims to analyze historical data from CGM systems more effectively, thereby enhancing the prediction of future blood glucose levels. By optimizing the time series forecasting capabilities of the T-LSTM model through the DE algorithm, this research seeks to provide a more reliable tool for diabetes patients to manage their condition proactively. The primary objective of this research is to demonstrate that the DE is optimized. The T-LSTM model can outperform traditional predictive models used in CGM systems. We hypothesize that the integration of Transformer and LSTM models, optimized through a DE algorithm, will provide superior accuracy in predicting short-term glucose fluctuations, thus significantly benefiting diabetes management.

The paper is organized as follows: Section 1 gives the introduction to this thesis; Section 2 details the methodology, including the description of the clinical data used, the theoretical framework of the T-LSTM model, and the optimization process via the DE algorithm. Section 3 presents the results of the model testing, including a comparative analysis of the MSE values obtained against traditional models. Section 4 discusses the implications of these results, potential limitations of the study, and directions for future research. Finally, section 5 concludes with a summary of findings and their implications for real-world applications.

2. METHOD

2.1. Clinical data

Data were collected from continuous glucose monitoring (CGM) devices. A sliding window technique was applied to segment the time-series data into input-output pairs suitable for modeling. The dataset was then divided into training, validation, and test sets to ensure robust model evaluation. The data

has been authorized by the Department of Endocrinology at The Affiliated Suzhou Hospital of Nanjing Medical University, located in Suzhou Municipal Hospital, Suzhou, China.

The 2023 ADC diabetes diagnosis and treatment standards state that for non-pregnant adults without significant hypoglycemia, the ideal Glycated hemoglobin (A1C) target should be below 7% (53 mmol/mol) [24]-[26]. An A1C level above 7% is considered abnormal. As shown in Table 1, for the purpose of this research, 8 patients from the Department of Endocrinology at Suzhou Municipal Hospital in Jiangsu Province, China, were selected. These patients, all under 18 years old, had A1C levels exceeding 7.0%, were diagnosed with Type 1 Diabetes Mellitus (T1DM) as per ADA guidelines, and were using a CGM system. This study aims to demonstrate the effectiveness of the T-LSTM model in comparison to traditional models.

Patient data confidentiality was maintained by excluding personal identifiers like names, ages, and hospital stay duration, and excluding data from patients with atypical gestational diabetes or those who experienced data transmission loss during the CGM process. Each patient underwent a glucose tolerance test upon hospital admission. The patients' A1C levels and islet function loss statuses varied, which was integral for model validation. The methodology involves algorithmically dividing patients' data into training and testing sets and setting a threshold to optimize the algorithm model.

Table 1. The standards for data choosing

Research standards	2023 ADC diabetes diagnosis and treatment standards
A1C Target	<7% (53 mmol/mol) for non-pregnant adults
Patient Selection Criteria	Location: Department of Endocrinology, Suzhou Municipal Hospital, Jiangsu Province, China.
Age	Below 18 years.
A1C Level	>7.0%.
Condition	Diagnosed with T1DM as per ADA guidelines.
Monitoring System	Using CGM.
Data Confidentiality	Exclusion of personal identifiers and specific patient conditions.
Patient Assessment	Glucose tolerance test upon admission; varying A1C levels and islet function loss statuses.

2.2. Transformer-LSTM theory and formula

Figure 1 depicts a composite machine learning architecture, integrating the functionalities of LSTM networks to create a hybrid T-LSTM model. This innovative model framework is divided into two subclasses derived from the `nn.Module` and `model Base`: the Transformer model subclass and the LSTM model subclass. The Transformer model subclass is intricately designed, featuring a series of components including encoder, decoder, projection head, and linear layer, which collectively contribute to its robust encoding capabilities. Complementing this, the LSTM model subclass extends the functionalities of the Transformer model by incorporating an LSTM layer, which enhances the model's ability to handle sequential data effectively. Both subclasses are equipped with a forward method, enabling forward propagation through the network and for computing outputs. The constructors within these subclasses are parameter-rich, providing versatility in model configuration. These parameters, input embedding, positional encoding layers, and dropout rates, are critical for tailoring the model to the specific characteristics of the dataset being analyzed.

Model architecture: The T-LSTM network is designed to capture temporal dependencies in glucose time-series data. The architecture consists of multiple LSTM layers with a specified number of memory units, optimized to balance model complexity and performance. Previous studies have identified that an optimal number of memory units enhances prediction accuracy without leading to overfitting.

Incorporating Transformer mechanisms allows the model to focus on different time steps in the input sequence, capturing global dependencies and improving prediction accuracy. The self-attention mechanism enables the model to weigh the importance of various input features dynamically.

T-LSTM model uniquely integrates the capabilities of both Transformer and LSTM models: the Transformer is adept at processing long-term dependencies, while the LSTM excels in managing short-term dependencies. This contrasts with other models that typically rely solely on either LSTM or CNN architectures. Despite its innovative approach, the T-LSTM algorithm has not been directly compared to other models in existing research [27], [28]. The model is designed for immediate, short-term (within 1 hour) glucose predictions in Type 1 diabetes patients, while other models are utilized for different tasks or patient groups. The choice of a suitable deep learning model for glucose prediction is contingent upon the specific task requirements and the available data. The T-LSTM model algorithm is particularly effective for short-term predictions, whereas other models may be more appropriate for longer-term forecasts or for use with different demographic groups.

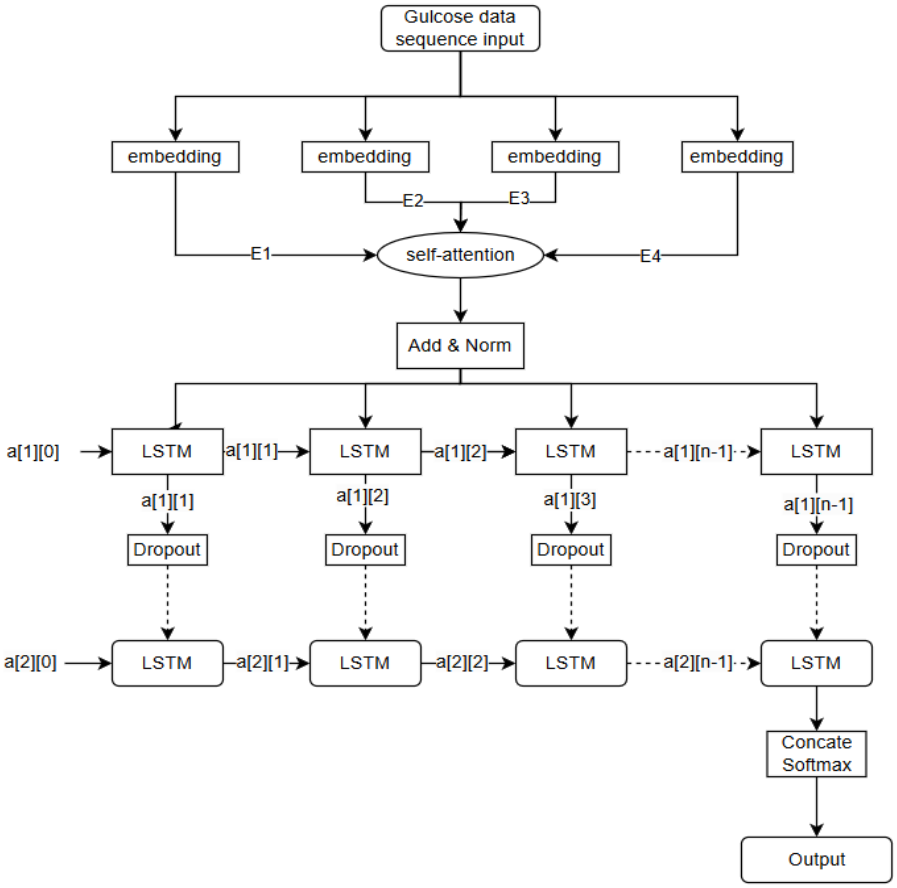


Figure 1. The architecture of T-LSTM algorithm

The underlying concept of the hybrid transformer model is its ability to interpret glucose data as a sequence of attributes, like time and glucose levels. The Transformer component is designed to analyze the interrelationships among these attributes, while the LSTM focuses on extracting localized features from the data. The outputs from both models are then merged to forecast subsequent glucose levels.

The mathematical formulation for the T-LSTM model is presented as (1):

$$f_t = \sigma\left(w_f \left[h_{t-1,} x_t\right] + b_f\ \right) \quad (1)$$

The components of the equation are defined as follows:
f_t represents the forget gate,
h_{t-1} denotes the output status of the previous node,
x_t is the input at the current moment,
σ is the sigmoid activation function,
w_f and b_f is the learn rate parameters.

$$i_t = \sigma\left(w_i \left[h_{t-1,} x_t\right] + b_i\ \right) \quad (2)$$

$$\tilde{C}_t = tanh\left(w_c \left[h_{t-1,} x_t\right] + b_c\ \right) \quad (3)$$

The architecture features an input gate that determines the relevance of the current glucose input within the broader context. When the input gate is activated, it enables the model to disregard the current glucose input, thereby preventing the transmission of this input information to subsequent nodes or time steps. The input gate is composed of two segments, each activated by distinct activation functions: the sigmoid function and the tanh function, respectively.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

The C_t variable in the model serves a crucial role in retaining both previous and current related information, preserving it over extended sequences. This capability effectively addresses the challenge of long-distance dependencies within the data. Once C_t is updated, the information is subsequently propagated to the next node or time step.

$$o_t = \sigma\left(w_o \left[h_{t-1,} x_t\right] + b_o\right) \quad (5)$$

$$h_t = o_t * tanh(C_t) \quad (6)$$

The function of the output gate, as delineated by formulas 5 and 6, is to deliver the final content of the model. This output is derived from variables C_t, h_{t-1} and x_t. Additionally, formula (6) incorporates a filtering mechanism o_t ; this mechanism evaluates the utility of the information contained in C_t , allowing the gate to selectively output information deemed useful while discarding what is considered non-essential.

$$y = f(x, W_T * O_{Transformer}, W_L * O_{LSTM}) \quad (7)$$

In the specified model, y represents the forecasting glucose values, while x denotes the input glucose sets. Furthermore, W_T is identified as the weight matrix for the Transformer model, and W_L , which serves as the weight matrix for the LSTM model.

By utilizing past data, glucose prediction attempts to forecast future glucose levels. The T-LSTM model emerges as a potent solution for this task, synthesizing the strengths of both architectural frameworks. The Transformer model excels in capturing long-range dependencies due to its self-attention mechanisms, which allows it to examine different time stages within the input sequence. This ability is helpful in identifying intricate relationships and patterns found in glucose data. On the other hand, the LSTM is an architecture of RNN that is well-known for its effectiveness at modelling sequential data. The memory cell that allows for the storage and retrieval of information across extended sequences. This feature is especially useful for capturing the temporal relationships typical of glucose time series.

2.3. DE algorithm optimizes T-LSTM

A. Optimization with DE algorithm

The DE algorithm is employed to optimize the hyperparameters of the T-LSTM network, including the number of memory units, learning rate, and batch size. DE's population-based search strategy effectively explores the hyperparameter space, aiming to find the optimal configuration that minimizes prediction error. This optimization process enhances the model's ability to generalize across diverse patient data.

B. Model training and evaluation

The optimized T-LSTM model is trained using the Adam optimizer with a mean squared error (MSE) loss function. Performance is evaluated using metrics such as root mean square error (RMSE) and mean absolute percentage error (MAPE). Cross-validation techniques are employed to assess the model's robustness and prevent overfitting.

C. Comparison with existing models

The proposed hybrid model's performance is compared with traditional models like autoregressive integrated moving average (ARIMA) and other machine learning models, including support vector regression (SVR) and recurrent neural networks (RNNs). Studies have demonstrated that integrating DE with LSTM networks improves prediction accuracy over standard modeling approaches.

The DE algorithm generates population individuals by encoding with floating point vectors. In the process of DE algorithm optimization, first, select two individuals from the parent individuals and perform vector difference to generate a difference vector; secondly, select another individual and sum the difference vector to generate an experimental individual; then, compare the parent individuals with the corresponding experimental individuals undergo a crossover operation to generate new offspring individuals; finally, a selection operation is performed between parent individuals and offspring individuals, and individuals that meet the requirements are saved to the next generation group. The formulation can be expressed as:

$$Min\ (x_1, x_2, ..., x_d) \quad (8)$$

$$s.t\ \ X_j^L \leq x_j \leq X_j^U,\ j = 1,2, ..., D \quad (9)$$

Where, D represents the dimension of the solution space, X_j^L and X_j^U respectively represent the Upper and Lower limitations on the value range of the j-th component x_j.

As shown in Figure 2, the flowchart illustrates the optimization process of a T-LSTM model for glucose prediction using the DE algorithm, beginning with Initialization where a diverse population of 1 to 50 solutions is generated, each defined by specific hyperparameters such as a learning rate ranging from 0.001 to 0.01, 1 to 3 LSTM layers, hidden layer sizes between 50 and 200 units, a dropout rate of 0.2 to 0.5, and a batch size from 32 to 128. In the Mutation step, mutant vectors are created by adjusting these hyperparameters using a scaling factor between 0.5 and 0.8, applied to differences between randomly selected solution pairs, introducing variability. The Crossover phase mixes these mutants with original vectors at a crossover rate of 0.7 to 0.9, effectively blending traits to potentially enhance solution efficacy. Selection evaluates these new configurations via the MSE fitness function, choosing those with lower MSE for progression, thereby optimizing predictive accuracy. If the Check Convergence criteria of a maximum of 100 to 200 generations or an MSE threshold are not met, the process cycles again from mutation, else it concludes with Finalize Model, where the T-LSTM's hyperparameters are fine-tuned and ready for deployment, ensuring the model's capability to provide precise glucose level forecasts essential for effective diabetes management. The flowchart details the optimization of proposed T-LSTM model using the DE algorithm for glucose prediction. It starts by initializing a population with specific hyperparameters, which are then refined through mutation and crossover steps based on performance measured by MSE. This iterative process continues until convergence criteria are met, resulting in a finely tuned model optimized for accurate glucose forecasting, enhancing diabetes management.

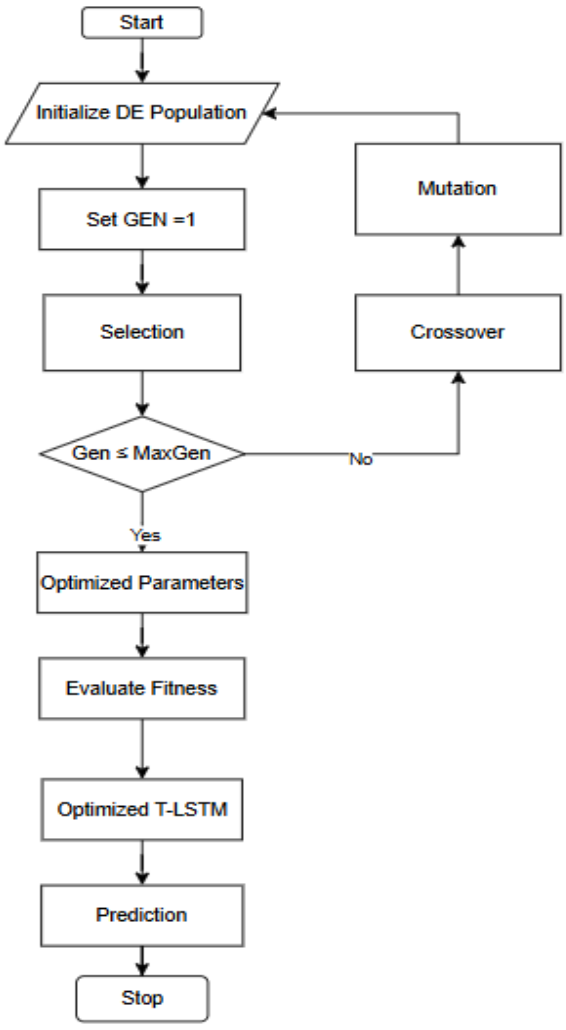


Figure 2. DE optimization diagram

3. RESULTS AND DISCUSSION

3.1. Results of MSE comparison

To evaluate the efficacy of the T-LSTM prediction model, refer to Table 1, which compares six models utilized in this study based on their MSE. The formula for calculating the MSE is provided below:

$$MSE = \frac{1}{N} \sum_{l=1}^{N} (\overset{\Lambda}{y}(t|t - PH) - y(t))^2 \quad (30)$$

In the equation, $\overset{\Lambda}{y}(t|t - PH)$ represents the forecasted glucose level at time t, $y(t)$ denotes the real time measured value of glucose at the corresponding time t.

The RNN maintains memory via hidden states within each node. These hidden states function as memory units and are updated with each new input, integrating information from the current input as well as the prior hidden state. This mechanism allows the RNN to preserve information from previous sequence steps. In this context, the MSE value over 15 minutes is recorded at 1.41. Additionally, there are advanced versions of RNNs, such as GRU and LSTM networks, which enhance the foundational model. These versions solve concerns such as the vanishing gradient problem, which can inhibit learning over long sequences. For these advanced networks, the MSE values over 15 minutes are 1.50 and 1.75, respectively. The stacked LSTM represents a further development of the standard LSTM model. This method stacks multiple LSTM layers atop each other, resulting in a 15-minute MSE value of 1.69. The bidirectional LSTM (Bi-LSTM) uses context from both the past and the future by processing the input sequence forwards and backwards at the same time. It comprises two distinct LSTM units: one processes the sequence from start to finish, while the other does so from finish to start. This allows the Bi-LSTM to detect dependencies in both directions, which is useful for applications like named entity recognition and sentiment analysis. However, when applied to lengthy time series data, this model achieves an MSE value of 1.72, indicating suboptimal performance. Based on the analysis presented in Table 2, it is evident that the six models' prediction algorithms exhibited varying behaviors when applied to the glucose data of the same patient. Notably, the Transformer-LSTM model, optimized using a Genetic algorithm, demonstrated superior performance on average. This was evident when comparing the MSEs calculated for each patient individually, consistently outperforming the traditional models. Consequently, the DE optimised T-LSTM model is more capable of forecasting glucose values which suggesting its enhanced suitability for applications in this domain.

When benchmarked against traditional models such as ARIMA and SVR, as well as other machine learning models like convolutional recurrent neural networks (CRNN), the proposed hybrid model demonstrated superior performance. For instance, studies have reported that CNN-LSTM architectures can achieve up to 94.71% accuracy in predicting glucose levels at a 90-minute horizon [29]. Similarly, models integrating Transformer components have shown improved prediction accuracy by effectively capturing global dependencies in time-series data [30].

Figure 3 displays a 15-minute forecast result of glucose. The black dashed line represents the actual changes in glucose levels, while the red dashed line indicates the forecasted glucose values. The Transformer-LSTM algorithm, as demonstrated here, can predict future glucose changes with considerable accuracy, closely aligning with the real glucose fluctuations.

The T-LSTM model's predictions for the patient's blood glucose in 45 minutes, based on the CGM data of sample 1, are illustrated in Figures 3-5. The red dashed line depicts the expected values, whereas the black dashed line represents the actual glucose levels. When the three numbers are compared, the forecasts' accuracy decreases as the prediction interval is larger. In other words, the association between accuracy and forecast period duration has decreased significantly.

Figure 6 illustrates the process of K-Fold cross-validation, a method used to assess a model's performance on a dataset by dividing it into multiple folds, as implemented in the provided code. Cross-validation ensures that each data point has an opportunity to be part of the training and testing phases, providing a robust measure of model accuracy and reliability across different subsets of data.

Table 2. MSE values

	115min	330min	445min
RNN	11.41	22.34	22.71
GRU	11.50	22.16	22.51
LSTM	11.75	22.43	22.75
Stacked LSTM	11.69	22.49	22.74
Bidirectional LSTM	11.72	22.50	22.90
DE- T- LSTM	00.96	11.54	22.31

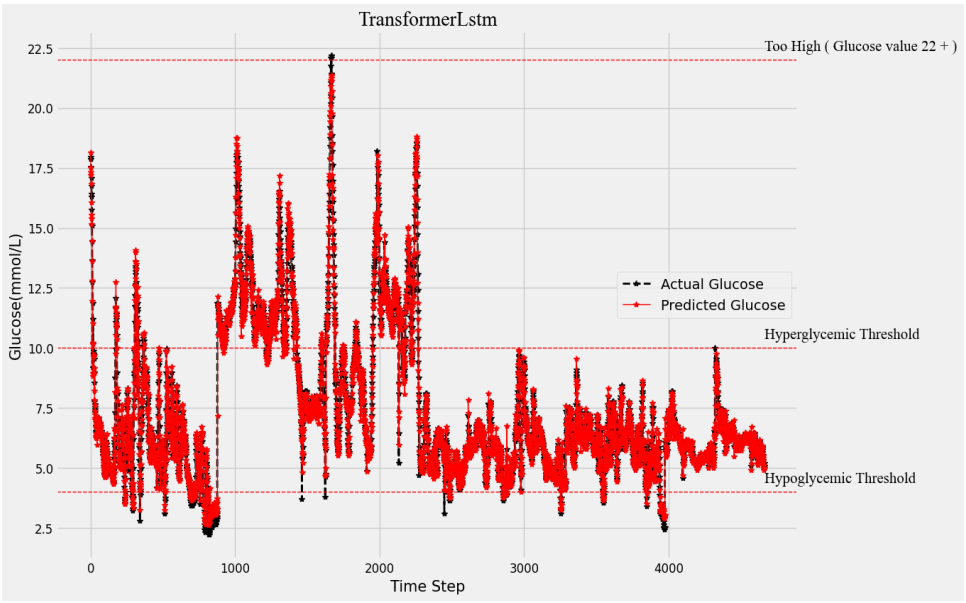


Figure 3. 15 minutes forecast

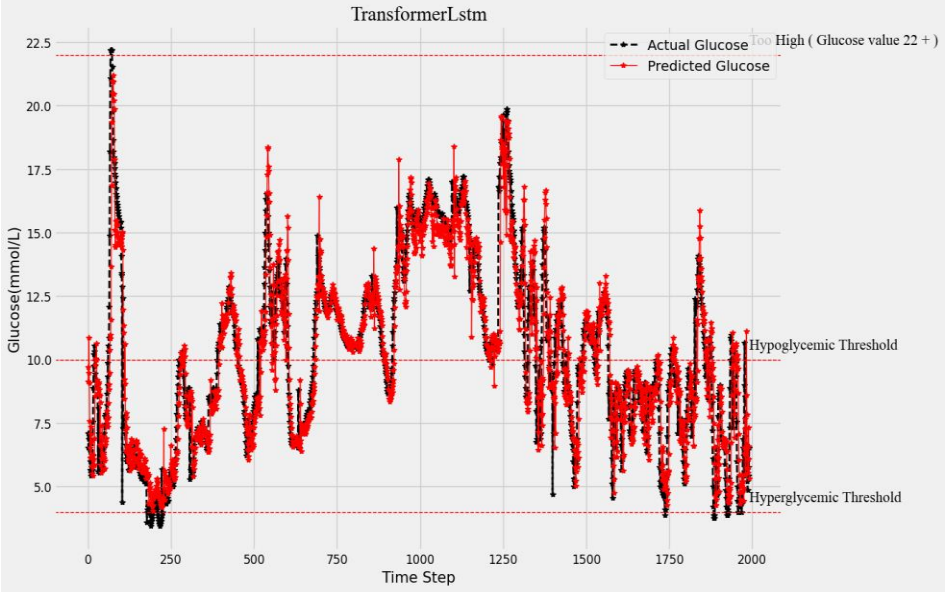


Figure 4. 30 minutes forecast

In this process, the dataset is first shuffled and split into two primary sets: training and test. The training set is then further divided into K folds (here, 14 folds in the code), which allows the model to be trained on (K-1) folds and validated on the remaining one. This rotation continues until each fold has served as the validation set once, and the model's performance is averaged across all folds to obtain a final score.

The code applies this cross-validation technique to compare the performance of two models: Transformer-LSTM and ARIMA. For each fold, a new instance of the Transformer-LSTM model is initialized, trained on the training subset, and validated on the test subset within the fold. The mean loss for each fold is recorded in Figure 7, and results are visualized across folds to understand the models' performance consistency and accuracy. By using K-fold cross-validation, the code ensures that the model is tested on all data points, providing a comprehensive assessment of its generalizability and highlighting any data-dependent performance variations. This image complements the code by visually summarizing the K-fold process, from initial dataset shuffling to sequential fold evaluation, highlighting how this approach provides a reliable measure of model accuracy and generalization across various subsets of data.

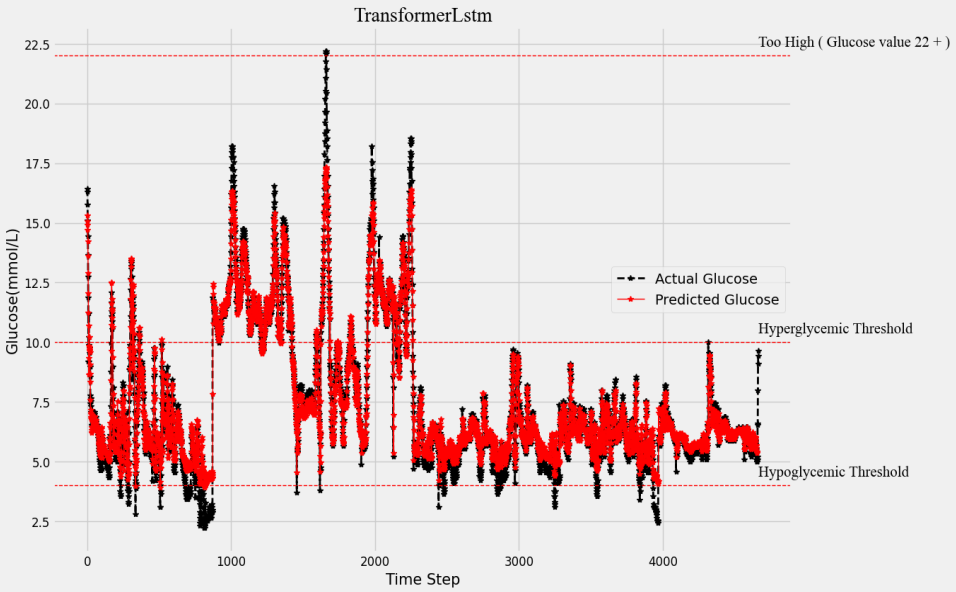


Figure 5. 45 minutes forecast

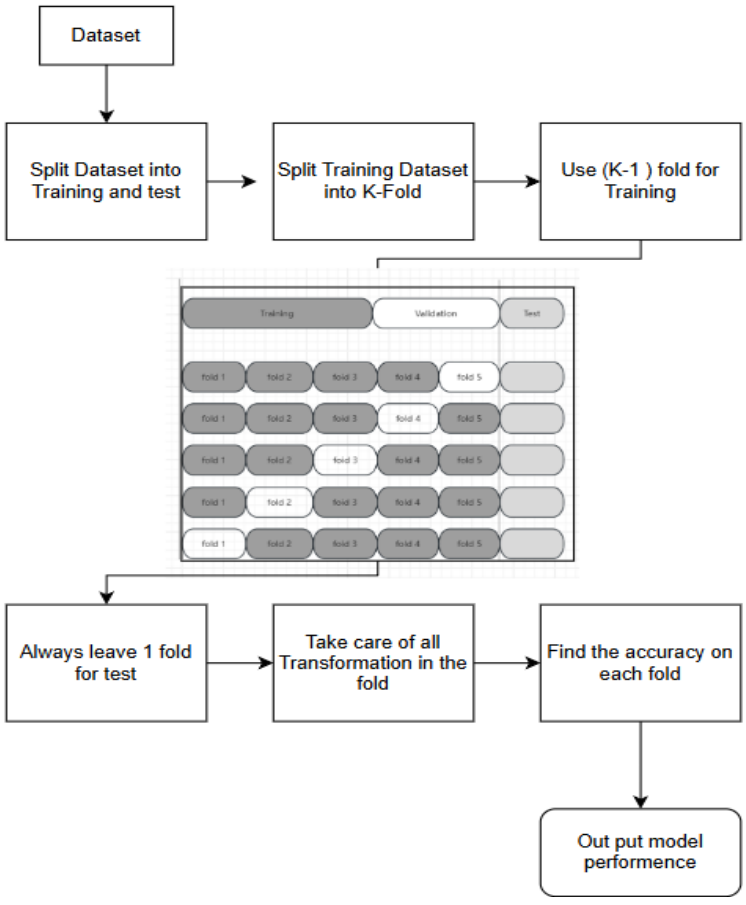
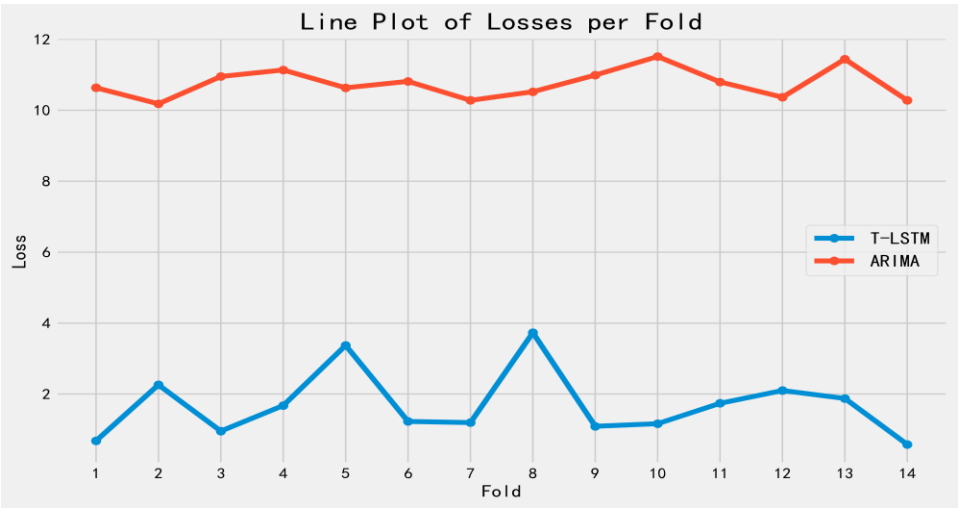


Figure 6. K-folder cross validation

The boxplot and line plot in Figure 7(a) indicate that the Transformer-LSTM model may provide slightly more reliable and consistent predictions of blood glucose levels due to its lower median loss and smaller IQR. However, the fold-specific variations in Figure 7(b) observed in the line plot suggest that

ARIMA can occasionally achieve lower losses depending on the data subset, though it is less consistent overall. This comparative analysis highlights that Transformer-LSTM may be a better choice for general applications in blood glucose prediction due to its balance of accuracy and stability.

(a)



(b)

Figure 7. Comparison of model losses using two visualization methods; (a) Boxplot showing the distribution of loss values for T-LSTM and ARIMA models (b) Line plot illustrating the loss values across 14 validation folds for both models

Figure 8 displays a line graph representing the change in fitness across 30 iterations. The x-axis labelled "Iteration," ranges from 1 to 30, indicating the sequence of iterations. The y-axis, labelled "Fitness," is scaled from 0.96 to 1.10, which measures the fitness value. The line graph shows a pronounced initial decrease in fitness from iteration 1 to iteration 8, after which the reduction in fitness value becomes less steep and more gradual from iteration 9 onwards, eventually plateauing near iteration 22 through 30. The visual presentation suggests a quick initial optimization followed by a slower, more incremental improvement or stabilization in the fitness value as iterations progress. This pattern could indicate this T-LSTM algorithm approaching an optimal solution or a limitation in further improvement after significant initial gains.

As shown in Table 3, the improvement after 30 generations of optimization, all the results had been optimized. On average, there's a significant reduction with 28% improvement in MSE across all patients" post-optimization, indicating a strong overall efficacy of the optimization process. When a patient's (Patient 6) blood glucose levels stabilize and remain within a relatively normal range over a long period, using the DE optimized model yields better optimization results. This indicates that the DE optimization is more effective

in stable blood glucose scenarios, because the model can more accurately predict and adapt to the consistent patterns observed in glucose levels when they are within normal ranges.

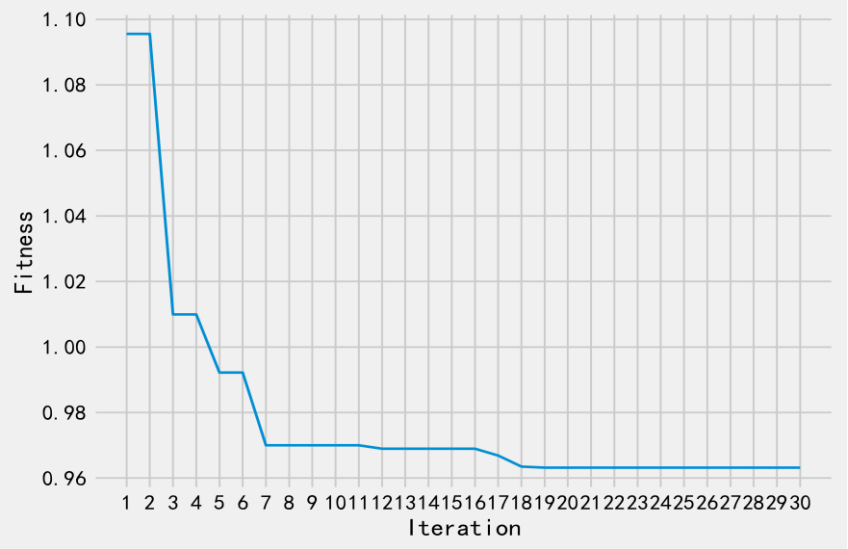


Figure 8. Iteration-fitness curve

Table 3. DE algorithm optimization results

Patient ID	MSE before optimization	MSE after optimization	Improvement (%)	Comments
Patient 1	1.244	0.987	21%	Notable improvement
Patient 2	0.646	0.477	26%	Notable improvement
Patient 3	0.657	0.585	11%	Improvement observed
Patient 4	0.635	0.432	32%	Responds well to changes
Patient 5	1.355	1.242	8%	Minor improvement
Patient 6	1.177	0.521	56%	Significant improvement
Patient 7	0.858	0.507	41%	Good improvement
Patient 8	0.927	0.674	27%	Notable improvement
Average	0.849	0.613	28%	NA

In conclusion, the results for Patient 1 using the DE optimized T-LSTM model, with predictions at 15, 30, and 45 minutes, shows a high consistency with the actual glucose values. The MSE results for T-LSTM about patient are 0.96, 1.54, and 2.31, outperforming those of the standard Transformer model. The DE optimized T-LSTM model demonstrates superior predictive performance in estimating blood glucose levels. These cases indicate that the T-LSTM model's performance varies significantly with changes in patient lifestyle and health status, underscoring the need for personalized diabetes management tools. The model’s superior handling of temporal data points, facilitated by the LSTM component, makes it adept at adapting to patient-specific fluctuations in glucose levels.

The high accuracy rates observed have significant meaning for Clinical Implications, particularly within the clinically acceptable zones, underscore the potential of the proposed model for real-world diabetes management. Accurate prediction of blood glucose levels can significantly aid in timely insulin administration and dietary planning, thereby reducing the risks of hypo- and hyperglycemia. However, it is essential to acknowledge that individual variability exists, and personalized adjustments may be necessary to optimize the model's performance for different patients.

3.2. Discussion

The results of this study highlight the effectiveness of the DE algorithm-optimized T-LSTM model in predicting glucose levels with improved accuracy. Compared to traditional methods such as RNN, GRU, and standard LSTM models, the proposed model achieved lower MSE values across 15, 30, and 45-minute prediction intervals. The integration of Transformer components enabled the model to capture global dependencies in glucose fluctuations, while the LSTM layers preserved short-term temporal patterns. This combination resulted in a more robust and adaptive model capable of providing reliable short-term glucose

forecasts. The findings confirm that optimizing the T-LSTM architecture with the DE algorithm significantly enhances predictive performance, making it a promising tool for real-time diabetes management.

When compared with existing literature, the proposed model demonstrates notable improvements in predictive accuracy. Previous studies have explored CNN-LSTM and ARIMA models for glucose forecasting, yet these approaches often struggled with delayed responses to rapid changes in glucose levels. Our findings suggest that DE-optimized T-LSTM outperforms these models by dynamically adjusting weights and learning rates to optimize prediction accuracy. However, the study also revealed certain limitations, such as reduced accuracy in longer prediction horizons and increased variability among different patients. Unexpected results were observed in some cases where glucose fluctuations were heavily influenced by external factors like meal intake, physical activity, and insulin response—variables that were not explicitly included in this model. This suggests that incorporating additional physiological and behavioral parameters could further enhance prediction reliability.

Future research should focus on refining the model by integrating real-time patient data, including dietary intake, insulin dosage, and physical activity levels. Additionally, implementing adaptive learning mechanisms that allow the model to self-adjust based on incoming data could improve long-term forecasting capabilities. Expanding the dataset to include a larger and more diverse patient population will be critical in validating the model's generalizability. Ultimately, this study paves the way for the integration of AI-driven glucose prediction models into clinical practice, enabling personalized diabetes management solutions that can enhance patient outcomes and reduce the risks associated with glucose level fluctuations.

4. CONCLUSION

This study successfully demonstrated the efficacy of a DE algorithm-optimized T-LSTM model for predicting future glucose levels using data from CGM systems. The model consistently outperformed traditional predictive models in terms of MSE across multiple time intervals, highlighting its superior accuracy and reliability in real-time glucose forecasting. The findings from this research provide compelling evidence for the potential integration of advanced machine learning techniques like the T-LSTM model into CGM systems. Such integration can significantly enhance the predictive accuracy of these systems, thereby allowing for better management of diabetes through timely interventions. This can lead to a substantial improvement in patient quality of life by minimizing the risk of episodes of hypo- and hyperglycemia.

While the results are promising, the study was conducted on a relatively small dataset and over short predictive intervals. Future research should focus on validating the model over larger and more diverse populations, as well as extending the predictive horizon to assess the long-term effectiveness of the model. Additionally, integrating more personalized patient data, such as dietary habits, physical activity, and other health metrics, could further refine the model's accuracy and applicability. Looking forward, we plan to expand our research to include multi-modal data integration, where factors such as insulin doses, meal intake, and physical activity are considered in conjunction with CGM data to enhance the model's predictive capability. We also aim to develop a real-time adaptive system that can update its predictions based on live data feeds, thereby providing even more precise glucose level forecasts. Collaborations with biomedical companies and clinical trials will be sought to facilitate the practical application and commercialization of the improved CGM systems.

ACKNOWLEDGEMENTS

The authors would like to thank Universiti Putra Malaysia (UPM) and the Malaysian Ministry of Higher Education (MOHE) for their continuous support in the research work. This work was supported in part by the fundamental research grant scheme (FRGS) FRGS/1/2015/TK03/UPM/02/7 under the grant number Vot. 5524740.

FUNDING INFORMATION

This research was supported by the Fundamental Research Grant Scheme (FRGS) under the Ministry of Higher Education Malaysia (MOHE), project code FRGS/1/2015/TK03/UPM/02/7 and grant number Vot. 5524740. No additional external funding was received for this study.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
QingXiang Bian	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Azizan As'arry		✓				✓		✓	✓	✓	✓	✓		✓
XiangGuo Cong					✓		✓	✓						
Khairil Anas bin Md Rezali	✓		✓	✓		✓	✓				✓	✓	✓	
Raja Mohd Kamil bin Raja Ahmad	✓					✓	✓				✓		✓	
Mohd Zarhamdy Md. Zain					✓		✓			✓		✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study. All participants and/or their legal guardians were fully informed of the objectives, procedures, potential risks, and benefits, and voluntarily agreed to participate. Consent documentation was obtained in writing and is available upon request.

ETHICAL APPROVAL

The research involving human subjects was conducted in accordance with all relevant national regulations and institutional policies, and it followed the principles outlined in the Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board of Suzhou Municipal Hospital, Nanjing Medical University, China. All procedures involving human participants were reviewed and approved under the hospital's ethics protocol.

DATA AVAILABILITY

The data that support the findings of this study were obtained from Suzhou Municipal Hospital and are subject to privacy and ethical restrictions. Therefore, the data are not publicly available. However, de-identified data may be made available from the corresponding author, Azizan As'arry, upon reasonable request and with permission from the hospital's ethics committee.

REFERENCES

[1] M. E. Lean *et al.*, "5-year follow-up of the randomised diabetes remission clinical trial (DiRECT) of continued support for weight loss maintenance in the UK: an extension study," *The Lancet Diabetes and Endocrinology*, vol. 12, no. 4, pp. 233–246, Apr. 2024, doi: 10.1016/S2213-8587(23)00385-6.
[2] M. Rickson, E. E. Wright, A. Bindal, and L. Ghonim, "Advancements in diabetes technology are outpacing the evidence," *Diabetes Technology and Therapeutics*, vol. 25, no. S3, pp. S35–S41, Jun. 2023, doi: 10.1089/dia.2023.0145.
[3] A. Vaswani *et al.*, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.
[4] Y. Tang, G. Li, T. Zhongwei, L. Peng, and Y. Liping, "Review of continuous blood glucose monitoring techniques," 2022.
[5] E. Susana, K. Ramli, P. D. Purnamasari, and N. H. Apriantoro, "Non-invasive classification of blood glucose level based on photoplethysmography using time–frequency analysis," *Information (Switzerland)*, vol. 14, no. 3, p. 145, Feb. 2023, doi: 10.3390/info14030145.
[6] M. Krakauer, J. F. Botero, F. J. Lavalle-González, A. Proietti, and D. E. Barbieri, "A review of flash glucose monitoring in type 2 diabetes," *Diabetology and Metabolic Syndrome*, vol. 13, no. 1, Apr. 2021, doi: 10.1186/s13098-021-00654-3.
[7] W. Wang, M. Tong, and M. Yu, "Blood glucose prediction with VMD and LSTM optimized by improved particle swarm optimization," *IEEE Access*, vol. 8, pp. 217908–217916, 2020, doi: 10.1109/ACCESS.2020.3041355.
[8] D. Zhu, D. Zhou, N. Li, and B. Han, "Predicting diabetes and estimating its economic burden in China using autoregressive integrated moving average model," *International Journal of Public Health*, vol. 66, Jan. 2022, doi: 10.3389/ijph.2021.1604449.

[9] J. Yang, L. Li, Y. Shi, and X. Xie, "An ARIMA model with adaptive orders for predicting blood glucose concentrations and hypoglycemia," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 1251–1260, May 2019, doi: 10.1109/JBHI.2018.2840690.
[10] R. Lin, F. Brown, S. James, J. Jones, and E. Ekinci, "Continuous glucose monitoring: A review of the evidence in type 1 and 2 diabetes mellitus," *Diabetic Medicine*, vol. 38, no. 5, Mar. 2021, doi: 10.1111/dme.14528.
[11] R. Ajjan, D. Slattery, and E. Wright, "Continuous glucose monitoring: a brief review for primary care practitioners," *Advances in Therapy*, vol. 36, no. 3, pp. 579–596, Jan. 2019, doi: 10.1007/s12325-019-0870-x.
[12] H. H. Aly, "A proposed hybrid machine learning model based on feature selection technique for tidal power forecasting and its integration," *Electronics (Switzerland)*, vol. 13, no. 11, p. 2155, Jun. 2024, doi: 10.3390/electronics13112155.
[13] T. Zhu, K. Li, J. Chen, P. Herrero, and P. Georgiou, "Dilated recurrent neural networks for glucose forecasting in type 1 diabetes," *Journal of Healthcare Informatics Research*, vol. 4, no. 3, pp. 308–324, Apr. 2020, doi: 10.1007/s41666-020-00068-2.
[14] S. Yamada, "Historical achievements of self-monitoring of blood glucose technology development in japan," *Journal of Diabetes Science and Technology*, vol. 5, no. 5, pp. 1300–1306, Sep. 2011, doi: 10.1177/193229681100500541.
[15] R. Amanipour, H. Nazeran, I. Reyes, M. Franco, and E. Haltiwanger, "The effects of blood glucose changes on frequency-domain measures of HRV signal in type 1 diabetes," in *CONIELECOMP 2012 - 22nd International Conference on Electronics Communications and Computing*, Feb. 2012, pp. 50–54, doi: 10.1109/CONIELECOMP.2012.6189880.
[16] W. P. T. M. van Doorn *et al.*, "Machine learning-based glucose prediction with use of continuous glucose and physical activity monitoring data: The Maastricht Study," *PLoS ONE*, vol. 16, no. 6 June, p. e0253125, Jun. 2021, doi: 10.1371/journal.pone.0253125.
[17] M. F. Rabby, Y. Tu, M. I. Hossen, I. Lee, A. S. Maida, and X. Hei, "Stacked LSTM based deep recurrent neural network with kalman smoothing for blood glucose prediction," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, Mar. 2021, doi: 10.1186/s12911-021-01462-5.
[18] A. E. S. El-Bashbishy and H. M. El-Bakry, "Pediatric diabetes prediction using deep learning," *Scientific Reports*, vol. 14, no. 1, Feb. 2024, doi: 10.1038/s41598-024-51438-4.
[19] X. Jin, A. Cai, T. Xu, and X. Zhang, "Artificial intelligence biosensors for continuous glucose monitoring," *Interdisciplinary Materials*, vol. 2, no. 2, pp. 290–307, Feb. 2023, doi: 10.1002/idm2.12069.
[20] M. Khalifa and M. Albadawy, "Artificial intelligence for diabetes: Enhancing prevention, diagnosis, and effective management," *Computer Methods and Programs in Biomedicine Update*, vol. 5, p. 100141, 2024, doi: 10.1016/j.cmpbup.2024.100141.
[21] S. Kumar Das, K. K. Nayak, P. R. Krishnaswamy, V. Kumar, and N. Bhat, "Review—electrochemistry and other emerging technologies for continuous glucose monitoring devices," *ECS Sensors Plus*, vol. 1, no. 3, p. 31601, Jul. 2022, doi: 10.1149/2754-2726/ac7abb.
[22] D. Rodbard, "Continuous glucose monitoring: a review of successes, challenges, and opportunities," *Diabetes Technology and Therapeutics*, vol. 18, no. S2, pp. S23–S213, Feb. 2016, doi: 10.1089/dia.2015.0417.
[23] A. Z. Woldaregay *et al.*, "Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes," *Artificial Intelligence in Medicine*, vol. 98, pp. 109–134, Jul. 2019, doi: 10.1016/j.artmed.2019.07.007.
[24] F. Munoz and G. Acuna, "Time series forecasting using NARX and NARMAX models with shallow and deep neural networks," in *2021 IEEE Latin American Conference on Computational Intelligence, LA-CCI 2021*, Nov. 2021, pp. 1–6, doi: 10.1109/LA-CCI48322.2021.9769832.
[25] F. Prendin, J. Pavan, G. Cappon, S. Del Favero, G. Sparacino, and A. Facchinetti, "The importance of interpreting machine learning models for blood glucose prediction in diabetes: an analysis using SHAP," *Scientific Reports*, vol. 13, no. 1, Oct. 2023, doi: 10.1038/s41598-023-44155-x.
[26] "G. IEEE Engineering in Medicine and Biology Society. Annual International Conference (41st : 2019 : Berlin, IEEE Engineering in Medicine and Biology Society, and Institute of Electrical and Electronics Engineers."
[27] M. Jaloli and M. Cescon, "Long-term prediction of blood glucose levels in type 1 diabetes using a CNN-LSTM-based deep neural network," *Journal of Diabetes Science and Technology*, vol. 17, no. 6, pp. 1590–1601, Apr. 2023, doi: 10.1177/19322968221092785.
[28] J. B. Disclaimer, "Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in Python," *Machine Learning Mastery*, 2018.
[29] B. Di Martino, A. Esposito, G. J. Pezzullo, and T. H. Weng, "Evaluating machine and deep learning techniques in predicting blood sugar levels within the E-health domain," *Connection Science*, vol. 35, no. 1, Nov. 2023, doi: 10.1080/09540091.2023.2279900.
[30] F. Rustam *et al.*, "Enhanced detection of diabetes mellitus using novel ensemble feature engineering approach and machine learning model," *Scientific Reports*, vol. 14, no. 1, Oct. 2024, doi: 10.1038/s41598-024-74357-w.

BIOGRAPHIES OF AUTHORS

QingXiang Bian is a doctoral candidate in Biomedical Engineering at University Putra Malaysia, expected to complete his studies in 2024, focusing on the development of predictive algorithms for continuous glucose monitoring to enhance diabetes management. Born in Shandong, China, in 1993, he completed his bachelor's degree in mechanical manufacturing and Automation at North Minzu University in 2017, followed by a master's degree in manufacturing system engineering at University Putra Malaysia. His master's research, which involved collaboration with industry partners, provided him with valuable insights into real-time glucose monitoring challenges. He can be contacted at email gs65310@student.upm.edu.my.

Azizan As'arry was born in Kuala Lumpur, Malaysia. He received bachelor's, master's, and Ph.D. degrees in mechanical engineering from Universiti Teknologi Malaysia, Malaysia, in 2007, 2009, and 2013, respectively. His major field of study is control and mechatronics. He is currently a senior lecturer at Universiti Putra Malaysia, Malaysia. His research interests include active force control, system identification, system modeling, and simulation. He can be contacted at email: zizan@upm.edu.my.

XiangGuo Cong He is currently in Department of Endocrinology, Suzhou Municipal Hospital, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou, China. His research interests include metabolic syndrome and diabetes. He can be contacted at email: cxg136999@126.com.

Khairil Anas Md Rezali received his first degree in engineering (mechanical) from Universiti Putra Malaysia in 2010. He then continued to pursue his doctorate study at University of Southampton, United Kingdom and received his Ph.D. in Sound and Vibration in 2015. His research is including the study of mechanical vibration, material for vibration isolation, biodynamics of human body, hand-arm vibration, and vibration in transport. He is now working at Universiti Putra Malaysia, Malaysia. He can be contacted at email: khairilanas@upm.edu.my.

Raja Mohd Kamil Bin Raja Ahmad is a faculty member at the Department of Electrical and Electronic Engineering, Universiti Putra Malaysia (UPM). He holds a Ph.D. in Control Engineering from the University of Sheffield, UK, and a B.Eng. in Electrical Engineering from the University of Southampton, UK. His research interests include control engineering, biomedical engineering, signal processing, and pattern recognition. He is a Professional Engineer registered with the Board of Engineers Malaysia (BEM), a member of IEEE, and formerly served on the Executive Committee of the IEEE CSS Malaysia Section from 2009 to 2020. He can be contacted at email: kamil@upm.edu.my.

Mohd Zarhamdy Bin Md. Zain is an Associate Professor at the Faculty of Mechanical Engineering, Universiti Teknologi Malaysia (UTM), and a member of the Intelligent Control and Automation (iCA) Research Group under the Frontier Materials Research Alliance. His research interests include mechatronics, control systems, fuzzy logic, robotics, automation, and educational assessment. He has secured a total of 14 research grants, including industry, national, and university funds, and has published over 130 scientific papers, including 68 indexed publications. He has also supervised numerous postgraduate students at both Master's and Ph.D. levels. He can be contacted at email: zarhamdy@utm.my.