Evolution in Electrical and Electronic Engineering Vol. 4 No. 2 (2023) 474-483

EEEE

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/eeee
e-ISSN : 2756-8458

Assessment of Blood Glucose Level Prediction Using LSTM Deep Learning Method

Nur Iffah Najiha Abdullah[1], Farhanahani Mahmud[1,2]*

[1]Faculty of Electrical and Electronic Engineering,
Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, MALAYSIA

[2]Microelectronics and Nanotechnology Shamsuddin Research Centre (MiNT-SRC),
Institute of Integrated Engineering (I2E),
Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, MALAYSIA

*Corresponding Author Designation

DOI: https://doi.org/10.30880/eeee.2023.04.02.058
Received 14 July 2023; Accepted 05 September 2023; Available online 30 October 2023

Abstract: Continuous glucose monitoring has significantly improved the situations of patients with T1DM in this day and age of technological advancement as new procedures and technologies in clinical medicine have been developed. An artificial pancreas was recently developed to help these people manage their glucose levels. An artificial pancreas is a three-part device that mimics how the body's functional pancreas maintains blood glucose, also known as blood sugar. A synthetic pancreas is mostly used to help people with type 1 diabetes mellitus (T1DM). However, to create a successful artificial pancreas, the first step is to predict the future blood glucose level of diabetes patients to enable proactive and precise control of insulin delivery to them, with the application of artificial intelligence (AI) to accelerate and simplify this process. Thus, a study on deep learning which is a subset of the AI was conducted for the blood glucose level time-series prediction based on the Cobelli model of T1DM and is presented in this paper. In this study, Long Short-Term Memory (LSTM) deep learning model was built to predict the blood glucose levels for 3- and 6-minute prediction horizon (PH) using Python Programming. Besides, the time-series data of the blood glucose level used to train the model was generated from the T1DM-typed Cobelli model-based open-loop insulin delivery system developed in MATLAB Simulink software. From the results, the 3-minute PH generally outperformed the 6-minute PH in terms of mean absolute error (MAE), mean squared error (MSE) and root mean squared error (RMSE), indicating better prediction performance. In conclusion, to some extent, the finding contributes to the ongoing efforts in developing advanced technologies for diabetes management.

Keywords: LSTM, Cobelli Type 1 Diabetic Model, Blood Glucose Level

Corresponding author: farhanah@uthm.edu.my

1. Introduction

Based on the World Health Organization, more than 422 million persons globally have diabetes by 2020, with the figure anticipated to rise to 700 million by 2045. For example, in 2019, the United States healthcare system spent 25% of its expenditure on diabetes [1]. Diabetes mellitus is distinguished by elevated blood glucose levels and decreased insulin synthesis. Diabetes is classified into two types: type 1 and type 2 diabetes mellitus. Type 1 diabetes mellitus (T1DM) is a progressive disorder caused by the gradual degradation of beta cells in the pancreas [2]. It is also known as insulin-dependent diabetes since persons with T1DM must take insulin every day to survive. While the illness primarily affects youngsters, it can also impact adults. Adult-onset diabetes is another name for type 2 diabetes mellitus (T2DM). T2DM is caused by lifestyle, has high blood sugar levels, and has the same risks as type 1 diabetes. The natural effect of insulin on insulin resistance pulls glucose from the circulation into the cells, resulting in T2DM. As a result, glucose accumulates in the blood. However, not all people with T2DM require insulin.

People with T1DM must constantly test their blood glucose (glycemia) and self-inject appropriate quantities of insulin. Patients must carefully manage their schedules around their care activities to receive the best T1DM care. They must also make challenging judgments and calculations about their treatment regimen regularly [3], particularly when food intake and lifestyle variables are taken into consideration. This behavioral load of T1DM therapy can cause patients discomfort and dissatisfaction, especially among those who fail to achieve their goal glycemic levels [3] In this day and age of technological growth, continuous glucose monitoring has substantially improved the situations of patients with T1DM as new procedures and technologies in clinical medicine have been developed [4]. An artificial pancreas has recently been created to improve glucose management in these individuals. An artificial pancreas is a three-part device that mimics how a functional pancreas maintains blood glucose, often known as blood sugar, in the body. An artificial pancreas is mostly used to aid persons with type 1 diabetes [5]. Researchers have studied numerous ways for developing the artificial pancreas to overcome the problem of T1DM patients. Several studies have focused on utilizing deep learning techniques to create an effective artificial pancreas [6]-[8]. However, to develop a successful artificial pancreas, the early phase must be done by predicting the blood glucose level that is suitable for the human body. To further evaluate the potential of the artificial pancreas for better glucose management, a study on deep learning prediction of blood glucose levels based on the Cobelli model of type 1 diabetes will be undertaken and presented in this work [9]. Deep learning is a subfield of machine learning that is inspired by the structure and function of the brain, specifically neural networks. It involves training artificial neural networks on a large dataset, allowing the network to learn and make intelligent decisions on its own. These neural networks are called deep because they have a large number of layers, allowing them to learn and represent complex patterns and relationships in data [10].

Thus, this study aims to explore the power of deep learning long short-term memory (LSTM) model for predicting blood glucose levels and intends to contribute to the field by evaluating the effectiveness of LSTM models in predicting blood glucose levels in T1DM patients, using Python as the primary programming language, where the time-series blood glucose datasets used to train the LSTM model were generated from the simulations of an open-loop insulin delivery based on Cobelli type 1 diabetic model. Thus, the objectives of this study are to generate 24-hour samples of the time-series blood glucose dataset of a virtual T1DM patient based on the Cobelli model using MATLAB Simulink software by simulating the open-loop insulin delivery system, and to design and evaluate the deep learning model based on LSTM to predict the blood glucose level based on the time-series blood glucose dataset using Python programming. The performance of the model in predicting the blood glucose level was evaluated based on the performance metrics of MAE, MSE and RMSE.

2. Materials and Methods

Figure 1 shows the flowchart of the overall steps taken in this work, starting from the generation of 10 datasets of 24-hour samples of the time-series blood glucose values of a virtual T1DM patient based on Cobelli model using MATLAB Simulink software by simulating an open-loop insulin delivery system. Utilizing these datasets, an LSTM deep learning model was developed to predict the blood glucose levels by implementing deep learning techniques in Python programming language with Google Colab. Lastly, the performance of the deep learning model was analysed based on the performance metrics of MAE, MSE and RMSE.

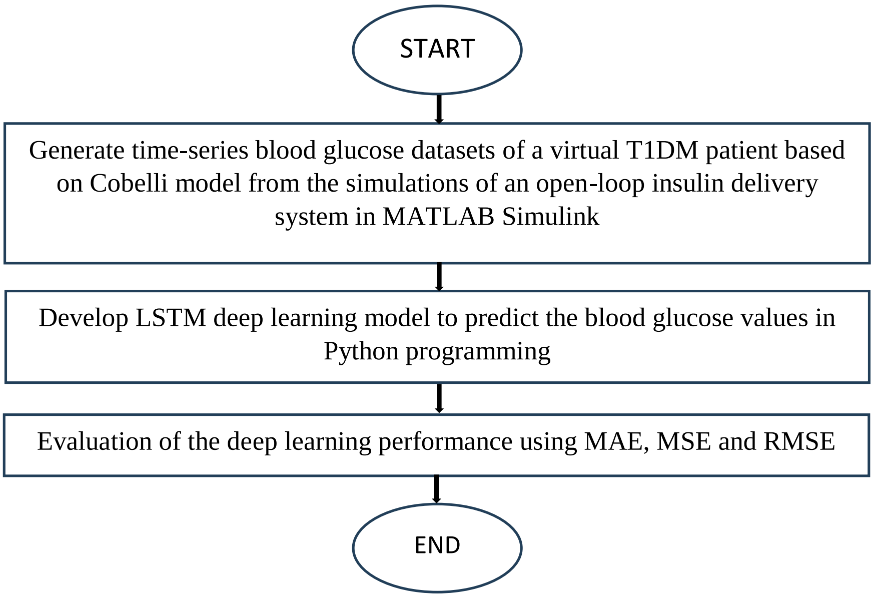


Figure 1: The flowchart of the overall steps involve in the methodology

2.1 Generation of the time-series blood glucose datasets

Cobelli model was used to generate the dataset of blood glucose level of a virtual T1DM patient as it more physiologically-based, taking into account a greater number of factors that impact glucose metabolism, including insulin secretion, glucose absorption, and utilization [11] makes it a more accurate model for simulating blood glucose levels in response to different inputs. Additionally, it is the first computer simulator to be approved by the Food and Drug Administration (FDA), and can be easily implemented in MATLAB software [12].

The Simulink system design for the insulin-glucose open loop for the Cobelli Type 1 diabetic model dataset was developed using MATLAB Simulink. The simulation was conducted in a 24 hours. Figure 2 shows the block diagram of the Simulink system design for the Cobelli model dataset. The system generated input signals for meal disturbance and insulin, which were used to generate 10 datasets. These datasets have different values of insulin and meal disturbance, resulting in varied graph outcomes.

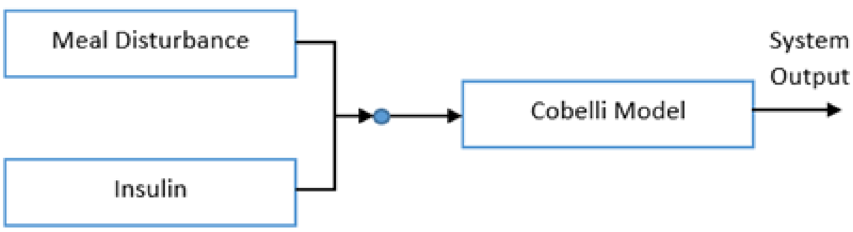


Figure 2: Block diagram of the Simulink system design for Cobelli model dataset

The S-Function block was a computer language description written in MATLAB®, C, C++, or Fortran, and its function was to simulate the Cobelli model coding in the .m file format. The To File block was used to write all input signal data into a MAT-file, including the time interval, glucose set point, glucose highest limit, hyperglycemia and lowest limit, hypoglycemia, insulin, and meal disturbance input. The output of the data file was the glucose level from the insulin and meal disturbance input.

The Simulink block design for the Cobelli model dataset provided a dataset of blood glucose dynamics for LSTM deep learning modeling. Figure 3 shows the MATLAB Simulink design for the Cobelli model dataset. The Signal Builder block generated input signals for meal disturbance and insulin, while the S-Function block simulated the Cobelli model coding in the .m file format.

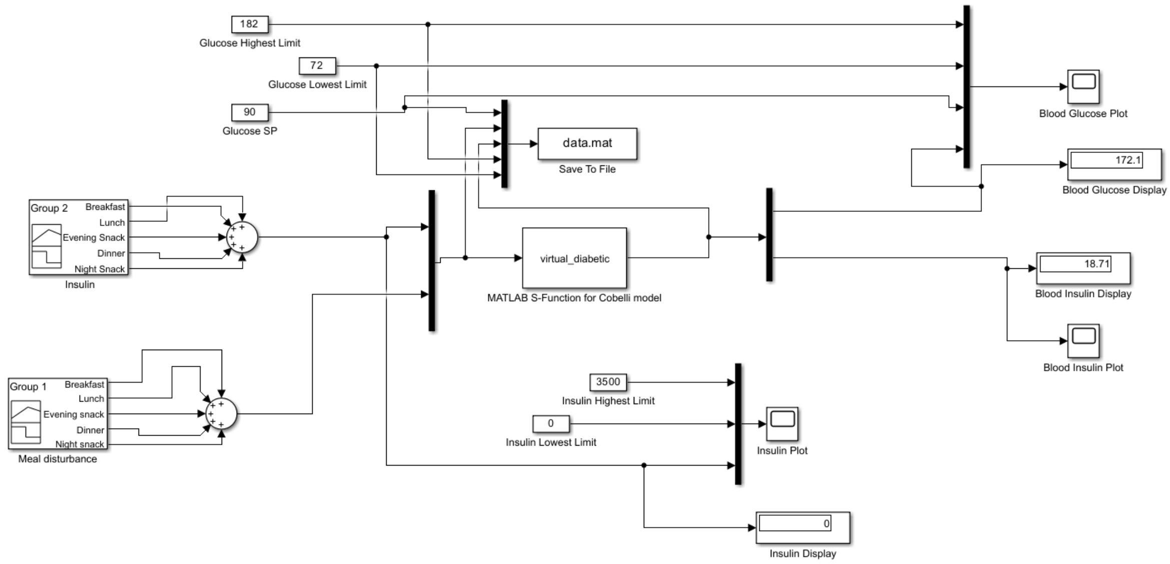


Figure 3: The MATLAB Simulink design for Cobelli model dataset

2.2 Development of LSTM deep learning model for the blood glucose prediction

The study utilized Long Short-Term Memory (LSTM) techniques to forecast blood glucose levels using the Cobelli model dataset. LSTM networks were ideal for time series prediction due to their ability to handle sequential data and learn long-term dependencies. Python was used as the main programming language, along with Anaconda and Jupyter notebook, to generate the deep learning model using ten datasets from MATLAB Simulink. The model was built using Keras' Sequential API, consisting of an LSTM layer and a dense output layer. The Adam optimizer and mean squared error loss function were used to build the model.

The code imported libraries for data manipulation and visualization, such as matplotlib.pyplot, pandas, numpy, torch, torch.nn, torch.optim, and torch.utils.data [13]. The create_dataset function was defined from the pandas library, transforming time series data into input-output pairs suitable for training an LSTM model.

The CSV file, which consisted of the time-series blood glucose data, was loaded and plotted using matplotlib.pyplot, allowing for a visual representation of blood glucose measurements over time. A 67% - 33% split was used, with 67% allocated for training and 33% for testing. The lookback parameter was set to 60 and 120 time steps, for 3- and 6-minute PH, respectively, to capture patterns and dependencies in the data. The PH was set to 3 and 6 minutes as they are in the range of continuous glucose monitoring's (CGM's) sampling period of 1 to 15 minutes.

The create_dataset function transformed a time series into a prediction dataset for supervised learning by selecting lookback and horizon consecutive time steps from the dataset. The ten datasets were split into training and testing sets, and the input data was reshaped to fit the LSTM model's input shape. The function returned two arrays: X and y, which contained input sequences and target sequences, respectively.

The AirModel class was used as the architecture for the LSTM model in blood glucose prediction, designed as a subclass of nn.Module. The model consisted of two main components: an LSTM layer and a linear layer. The input_size parameter was set to 1, representing the number of features at each time step. The hidden_size parameter was set to 50, determining the number of LSTM cells in the hidden state. The num_layers parameter was set to 1, indicating a single recurrent layer. The batch_first parameter was set to True to ensure the input tensor shape aligned with PyTorch data conventions. A linear layer (self.linear) was added after the LSTM layer to map the output to the desired prediction horizon, transforming the hidden state of the LSTM to match the output size required for predicting future time steps.

2.3 Performance evaluation of the blood glucose prediction

The evaluation of blood glucose prediction performance relied on errors such as MAE, MSE and RMSE. These errors were obtained from the Cobelli diabetic model's deep learning system. The MAE measured the difference between the original and predicted values and could be interpreted as the mean of the total differences in the dataset. The formulas for these error metrics are as shown in Eq.1 to Eq.3, where $\hat{y}_k$ represented the actual value, y_k represented the predicted value and n is the total number of the samples.

$$MAE = \frac{1}{n}\sum_{k=1}^{n}|\widehat{y_k} - y_k| \qquad Eq.1$$

While MSE identified the difference between the original and predicted values, it could be interpreted as the mean of the total differences in the dataset.

$$MSE = \frac{1}{n}\sum_{k=1}^{n}(\widehat{y_k} - y_k)^2 \qquad Eq.2$$

RMSE evaluated the model's prediction ability by calculating the square root of the average squared difference between predicted and actual values. A lower value indicated better average prediction performance[14].

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(\widehat{y_k} - y_k)^2} \qquad Eq.3$$

3. Results and Discussion

Figure 4 shows the blood glucose data from Dataset 1, one of the ten datasets generated using the designed MATLAB Simulink blocks of the open-loop insulin delivery system based on the Cobelli type 1 diabetic model. The y-axis and x-axis of Figure 4 represented the blood glucose value and the number of data points, respectively. Each of the ten datasets contained 28,800 data points representing 24-hour time-series blood glucose simulated data of a Cobelli model-based virtual type 1 diabetic patient with a time sample of 0.05 minutes.

Table 1 shows the graphs of the RMSE versus epoch for the 3- and 6-minute PHs. Small PHs were chosen due to the small time sample of the used data. Based on the graphs, the results on how the model's performance improved or fluctuated over time could be observed. Moreover, the lower the error that intersected between the train and test values, the better the model's performance.

According to the graphs in Table 1, the RMSE was plotted against the epochs during the model training process for each dataset, with a fixed number of 10 epochs. The graphs showed a decreasing trend in the errors as the number of epochs increased, indicating that the model's performance improved over time as it learned from the training data and adjusted its weights and biases. Furthermore, the errors reached a point of stability or showed minimal changes over subsequent epochs of 8 to 10, suggesting that the model likely converged and achieved its optimal performance.

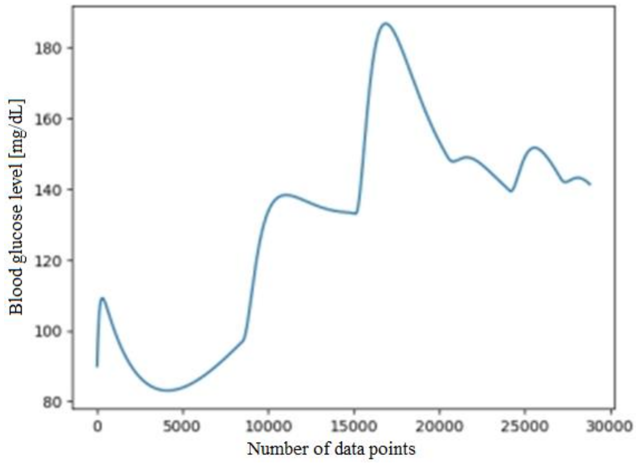
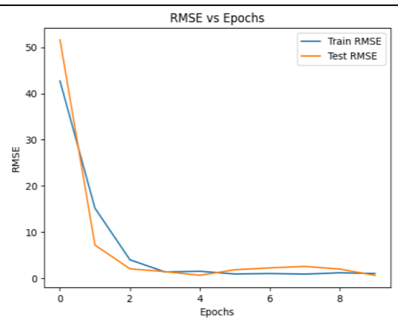
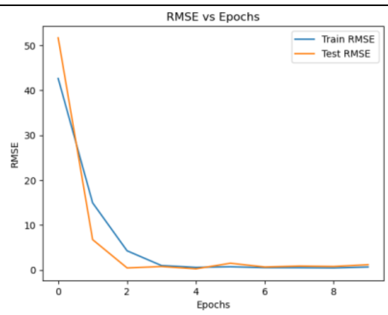
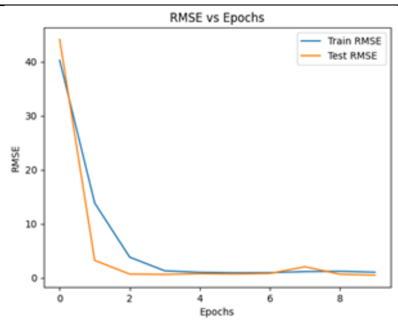
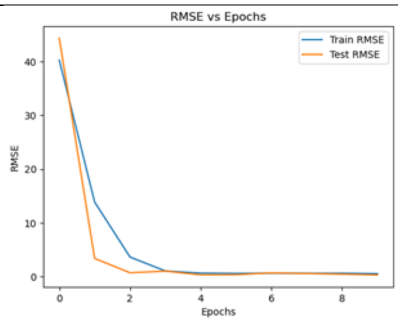
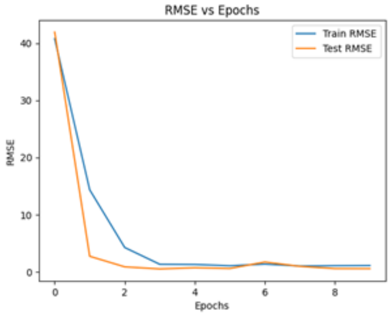
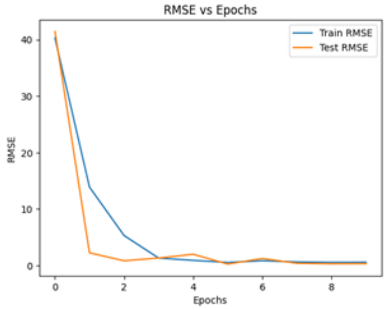
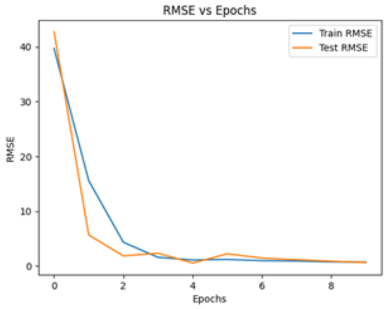
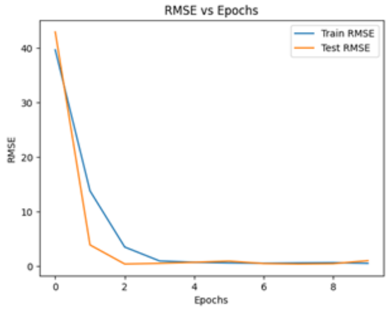
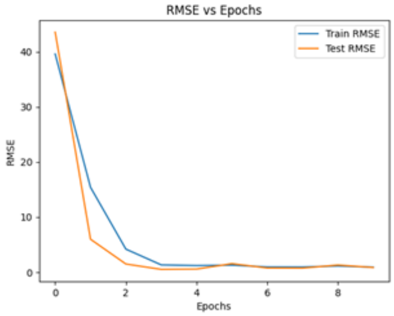
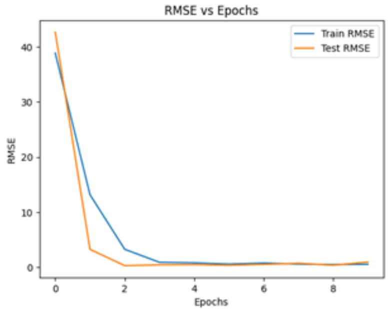
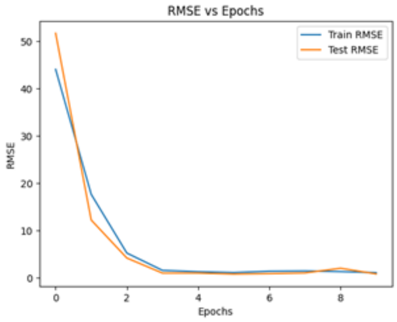
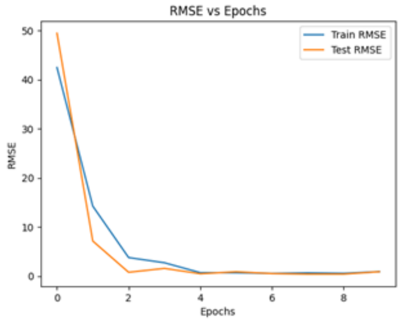
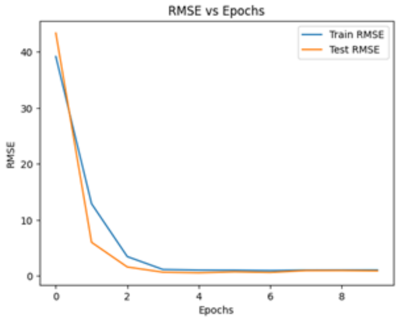
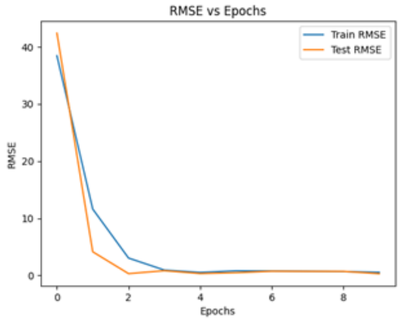
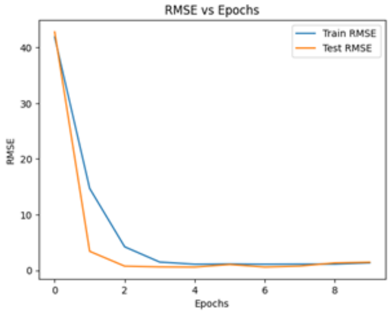
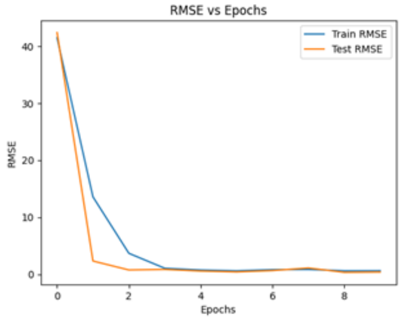
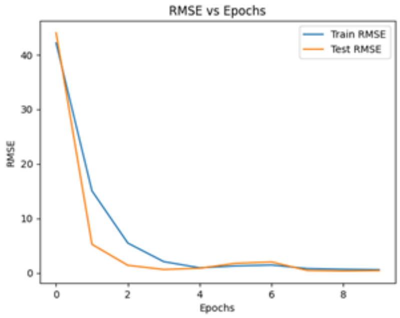
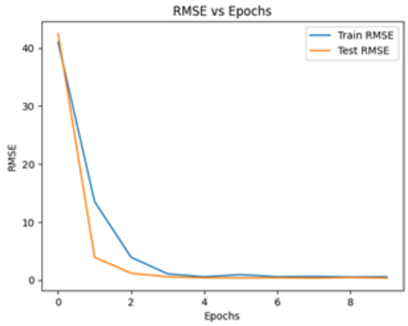
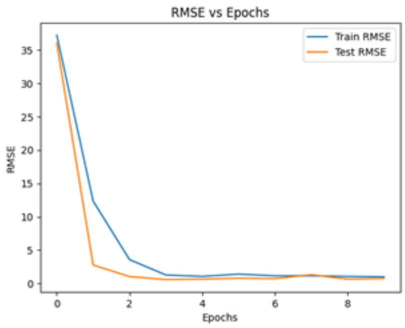
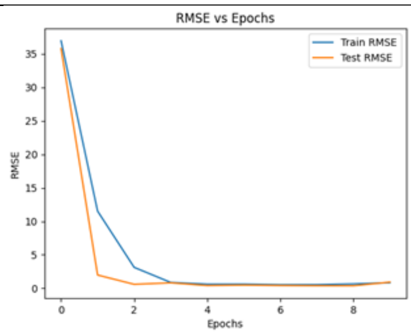


Figure 4: The blood glucose data from Dataset 1

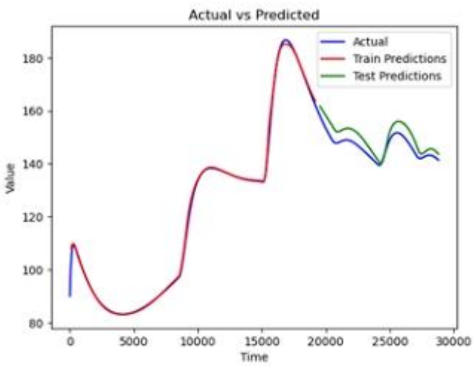
Table 1: The graphs of the RMSE versus epoch for the 3- and 6-minute PHs

Dataset	RMSE	
	6-minute PH	3-minute PH
1		
2		

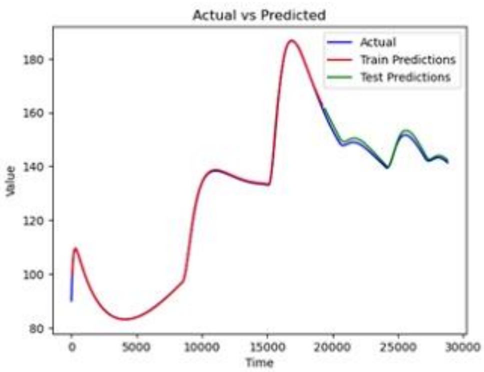
<table>
<tr><td>3</td><td>
</td><td>
</td></tr>
<tr><td>4</td><td>
</td><td>
</td></tr>
<tr><td>5</td><td>
</td><td>
</td></tr>
<tr><td>6</td><td>
</td><td>
</td></tr>
<tr><td>7</td><td>
</td><td>
</td></tr>
</table>

8		
9		
10		

While Figure 5 (a) and (b), respectively show the predictions for the 6- and 3-minute PHs, compared to the actual data from Dataset 1. The training set exhibited strong agreement between the predicted values and the original values, indicating that the model had learned underlying patterns and could replicate them accurately. Besides, the testing set demonstrated close adherence to the original values, suggesting that the model generalized well to unseen data.



(a)



(b)

Figure 5: The graph predictions of the blood glucose from Dataset 1 for (a) the 6-minute PH and (b) 3-minute PH

Table 2 and Table 3 show the summary of results of the LSTM model performance based on the MSE, MAE, and RMSE for the 6- and 3-minute PHs, respectively. The model was trained with the epoch set to 10. When the performance of the 6-minute PH in Table 2 and the 3-minute PH in Table 3 for the "Train" and "Test" sets was compared, it was discovered that the 3-minute PH exhibited lower values of MAE, MSE, and RMSE than those from the 6-minute PH for the training data and testing data. This suggested that the 3 minutes forecasting was generally closer to the actual values compared to the 6 minutes forecasting, indicating that the 3 minutes forecasting was more accurate than the 6 minutes forecasting as the former had a shorter time prediction compared to the latter.

Table 2: Values of RMSE, MSE and MAE for 6-minute PH

Dataset	MAE		MSE		RMSE	
	Train	Test	Train	Test	Train	Test
1	0.559	3.1392	0.6474	8.5895	0.8556	3.3881
2	0.6264	0.3531	0.68	0.5895	1.0224	0.4814
3	0.6747	0.4454	1.0944	0.5433	1.217	0.6986
4	0.4799	0.4766	1.0958	2.8929	0.6963	0.6228
5	0.5835	0.7196	1.1998	0.5094	0.898	0.8394
6	0.6761	0.6174	1.3153	1.1479	1.0659	0.7973
7	0.7029	0.7779	0.8413	0.5418	1.0768	0.9352
8	1.0481	1.3790	0.7251	0.2691	1.3591	1.4616
9	0.3734	0.3388	1.2052	0.4961	0.5712	0.4126
10	0.6363	0.6414	1.0745	0.9298	1.0009	0.7611
MEAN	0.63603	0.88884	0.98788	1.65093	0.97632	1.03981

Table 3: Values of RMSE, MSE and MAE for 3-minute PH

Dataset	MAE		MSE		RMSE	
	Train	Test	Train	Test	Train	Test
1	0.5166	1.1325	0.235	0.9856	0.6785	1.1915
2	0.3528	0.2082	0.2611	0.0837	0.5227	0.2859
3	0.3691	0.2558	0.3005	0.23	0.6039	0.3406
4	0.3911	0.8703	0.4089	0.7286	0.5222	1.0096
5	0.3886	0.7446	0.2689	0.1387	0.5209	0.9354
6	0.7003	0.8340	0.4336	0.8646	0.9168	0.9078
7	0.3905	0.2281	0.4285	0.2317	0.5227	0.3126
8	0.4051	0.3085	0.5467	0.2923	0.6254	0.3886
9	0.3134	0.2748	0.4012	0.1556	0.5407	0.3472
10	0.6763	0.8764	0.3332	0.3308	0.8376	0.9241
MEAN	0.45038	0.57332	0.36176	0.40416	0.62914	0.66433

4. Conclusion

This study presents an effective deep learning algorithm based on Long Short-Term Memory (LSTM) for blood glucose prediction, focusing on T1DM patients. The model was trained and tested using ten datasets generated from the Cobelli type 1 diabetic model, and was evaluated using performance metrics of Mean Squared Error, Mean Absolute Error, and Root Mean Squared Error. Considering the overall performance across both the "Train" and "Test" sets, the 3 minutes forecasting consistently outperformed the 6 minutes in terms of MAE, MSE, and RMSE. The results demonstrate promising results in accurately predicting blood glucose levels, contributing to the development of advanced diabetes management technologies and pave the way for the implementation of artificial pancreas systems in the future.

Acknowledgement

The authors would like to thank the Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia for its support through TIER 1 Grant No. Q359.

References

[1] "Diabetes," World Health Organization. [Online]. Available: https://www.who.int/newsroom/fact-sheets/detail/diabetes. (accessed Nov. 01, 2022).

[2] "Type 1 Diabetes Mellitus: Practice Essentials, Background, Pathophysiology." https://emedicine.medscape.com/article/117739-overview. (accessed Nov. 01, 2022).

[3] Balfe, M., Doyle, F., Smith, D. et al. What's distressing about having type 1 diabetes? A qualitative study of young adults' perspectives. BMC Endocr Disord 13, 25 (2013). https://doi.org/10.1186/1472-6823-13-25 .

[4] Dai, X., Luo, Zc., Zhai, L. et al. Artificial Pancreas as an Effective and Safe Alternative in Patients with Type 1 Diabetes Mellitus: A Systematic Review and Meta-Analysis. Diabetes Ther 9, 1269–1277 (2018). https://doi.org/10.1007/s13300-018-0436-y.

[5] D. and, "Artificial Pancreas," National Institute of Diabetes and Digestive and Kidney Diseases, Nov. 08, 2022. https://www.niddk.nih.gov/health information/diabetes/overview/managing-diabetes/artificial-pancreas (accessed Nov. 07, 2022).

[6] M. F. Rabby et al., "Stacked LSTM based deep recurrent neural network with Kalman smoothing for blood glucose prediction," BMC Medical Informatics and Decision Making, vol. 21, no. 1, 2021. doi:10.1186/s12911-021-01462-5.

[7] H. V. Dudukcu, M. Taskiran, and T. Yildirim, "Blood glucose prediction with deep neural networks using weighted decision level fusion," Biocybernetics and Biomedical Engineering, vol. 41, no. 3, pp. 1208–1223, 2021. doi:10.1016/j.bbe.2021.08.007.

[8] T. Zhu et al., "Blood glucose prediction in type 1 diabetes using deep learning on the edge," 2021 IEEE International Symposium on Circuits and Systems (ISCAS), 2021. doi:10.1109/iscas51556.2021.9401083 .

[9] C. Dalla Man, R. A. Rizza and C. Cobelli, "Meal Simulation Model of the Glucose-Insulin System," in IEEE Transactions on Biomedical Engineering, vol. 54, no. 10, pp. 1740-1749, Oct. 2007, doi: 10.1109/TBME.2007.893506.

[10] "What Is Deep Learning? | How It Works, Techniques & Applications," Mathworks.com, 2022. https://www.mathworks.com/discovery/deep-learning.html (accessed Nov. 09, 2022).

[11] C. Dalla Man et al., "Minimal model estimation of glucose absorption and insulin sensitivity from oral test: Validation with a Tracer Method," American Journal of Physiology-Endocrinology and Metabolism, vol. 287, no. 4, 2004. doi:10.1152/ajpendo.00319.2003.

[12] P. Li, L. Yu, Q. Fang, and S.-Y. Lee, "A simplification of Cobelli's glucose–insulin model for type 1 diabetes mellitus and its FPGA implementation," Medical & Biological Engineering & Computing, vol. 54, no. 10, pp. 1563–1577, 2015.

[13] Parthman Chanda, "Libraries in python," GeeksforGeeks, https://www.geeksforgeeks.org/libraries-in-python/ (accessed June 29, 2023).

[14] H. N. Mhaskar, S. V. Pereverzyev, and M. D. van der Walt, "A deep learning approach to diabetic blood glucose prediction," Frontiers in Applied Mathematics and Statistics, vol. 3, 2017.