JOURNAL OF SCIENCE AND TECHNOLOGY
ISSN: 2229-8460 e-ISSN: 2600-7924

JST

Vol. 16 No. 1 (2024) 65-75
https://publisher.uthm.edu.my/ojs/index.php/jst

Assessment of Deep Learning Model System for Blood Glucose Time-Series Prediction

Ade Anggian Hakim Ade Anggian Hakim[1], Farhanahani Mahmud[1,2]*, Marlia Morsin[1,2]

[1] *Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia (UTHM), 86400 Parit Raja, Batu Pahat, Johor, MALAYSIA*
[2] *Microelectronics and Nanotechnology Shamsuddin Research Centre (MiNT-SRC), Institute of Integrated Engineering (I2E), Universiti Tun Hussein Onn Malaysia (UTHM), 86400 Parit Raja, Batu Pahat, Johor, MALAYSIA*

*Corresponding Author: farhanah@uthm.edu.my
DOI: https://doi.org/10.30880/jst.2024.16.01.007

Article Info

Received: 14 March 2024
Accepted: 26 May 2024
Available online: 23 June 2024

Keywords

Deep learning, temporal fusion transformer (TFT), long short-term memory (LSTM), time series prediction, blood glucose prediction, type 1 diabetes mellitus (T1DM)

Abstract

Diabetes has become one of the most severe and prevalent chronic diseases, leading to life-threatening, costly, and disabling consequences and reduced life expectancy. Uncontrolled blood glucose (BG) conditions become a factor in diabetes mellitus sufferers, which then causes BG levels that are too high (hyperglycemia) and too low (hypoglycemia). People with Type 1 Diabetes Mellitus (T1DM) require long-term BG management to keep BG levels. Deep learning models using Continuous Glucose Monitoring (CGM) data to monitor and regulate BG concentrations in diabetic patients with prediction values to prevent hypoglycemia and hyperglycemia is very important. Based on some of the latest research, the deep learning Temporal Fusion Transformer (TFT) model is considered an approach method with superior performance in time-series prediction. Therefore, in this study, two TFT models, the TFT and AutoTFT univariate models, were proposed for the time-series BG prediction for T1DM patients. In this study, the two proposed TFT models with two baseline models, were trained and tested on the ShanghaiT1DM dataset. The proposed and baseline models were trained using manual and auto-tuning hyperparameters with Optuna on cross-validation for prediction horizons (PHs) of 30 and 60 minutes, respectively. The performance metrics used to evaluate the models were mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE). As a result, the TFT model is superior to the baseline LSTM model, also the proposed AutoTFT models achieved the smallest MAE, MAPE, and RMSE for both 30 and 60-minute PHs, respectively of all models used. Besides, the BG prediction results with 30-minute PHs are better than those with 60-minute PHs for all the models. This shows that the AutoTFT model stands as a promising tool for the accurate prediction of adverse glycemic events.

1. Introduction

Diabetes is now recognized as one of the most severe and widespread chronic illnesses today, resulting in life-threatening, expensive, and disabling outcomes and a decrease in average life span. Over 500 million people

globally are affected by diabetes, imposing an increasing financial burden on numerous countries[1]. Individuals experiencing unregulated blood glucose (BG) levels due to diabetes are in danger of elevated (hyperglycemia) or low (hypoglycemia) BG levels. In 2019, approximately 463 million individuals were estimated to have diabetes (with a 95% confidence interval ranging from 369 to 601 million). If swift and efficient measures are not taken, it is projected that the number of diabetes cases will increase to 578 million by 2030 and escalate to 700 million by 2045[2].

To date, there is no cure for diabetes. Most diabetes cases stem from the body's inability or failure to produce or release insulin, known as Type 1 Diabetes Mellitus (T1DM). Individuals diagnosed with T1DM must manage BG levels to maintain safe BG levels. It involves administering insulin externally and closely monitoring BG levels due to the lack of natural insulin secretion. Otherwise, there is a heightened risk of experiencing hypoglycemia and hyperglycemia, leading to various immediate and long-term consequences. Conditions such as nephropathy, retinopathy, and coronary heart disease are primarily attributed to elevated BG levels[3].

Simultaneously, severe hypoglycemia poses a greater risk and can lead to unconsciousness, seizures, or potentially fatal outcomes[4]. This issue will elevate healthcare expenses related to diabetes in the concerned nations. Presently, there is no viable approach to prevent T1DM. Individuals with T1DM must maintain continuous and consistent management of their BG levels, effectively managing these significant risks necessitates precise controls[5]. Nevertheless, maintaining the optimal BG levels in individuals with diabetes is a complex task because glycemic control, involving insulin bolus injections, can pose an elevated risk of unexpected glycemic episodes, which can be highly detrimental to the individual. Precise prediction of BG levels utilizing patient time series data for those with T1DM is a valuable resource for enabling proactive intervention and timely administration of medications to enhance the management of T1DM patients.

Recent research has centered on predicting BG levels by utilizing data collected from Continuous Glucose Monitoring (CGM) devices employed to oversee and manage BG levels in individuals with diabetes. Most of these studies have primarily involved patients diagnosed with T1DM[6,7,8]. Research has shown that previous BG readings can be leveraged to foresee upcoming BG levels, as demonstrated in initial efforts to predict BG levels using historical BG data acquired from CGM records[3]. Since then, numerous investigations have been carried out to anticipate BG levels by implementing statistical learning methods and conventional machine approaches[9,10].

There are two types of time series models based on input variables, namely univariate and multivariate models. Univariate time series models focus solely on analyzing and forecasting blood glucose levels over time without considering other variables[11]. Conversely, multivariate time series models take into account the interdependencies between multiple variables, such as insulin and dietary intake, to enhance the accuracy of blood glucose level predictions[12]. Univariate models have the advantage of simplicity and ease of interpretation. They are particularly useful when the relationships between different factors influencing blood glucose levels are unclear [13]. On the other hand, multivariate models consider the interactions between multiple variables, providing a more comprehensive understanding. However, multivariate models are more complex to develop and interpret compared to univariate models, requiring careful consideration of the interplay between various input variables[14].

Predicting BG levels is crucial for managing T1DM. Fortunately, with the expansion of CGM data, there has been increased interest in machine-learning methods, resulting in the testing of numerous techniques to predict BG readings[15]. The standard strategy in this scenario is to treat BG prediction as supervised learning assignments, utilizing continuous sets of CGM data as input for the models and future BG levels as the intended outcomes (univariate model). Many researchers have researched for this task such as conventional machine learning choices encompass employing the ARIMA method with Random Forest for regressive integration[16], Artificial Neural Networks (ANN)[17], and Support Vector Machines (SVM)[18].

Deep learning-based regression techniques have gained significant attention due to their ability to effectively handle complex patterns and large datasets. These techniques leverage deep neural networks to automatically learn representations from data, enabling them to capture intricate relationships and achieve state-of-the-art performance across various domains[19]. In the context of regression, the gradient computation using backpropagation in deep neural networks facilitates the implementation of these techniques[20]. Applied to time series data deep learning-based regression techniques are well-suited for capturing temporal dependencies in sequential data[21]. Deep learning models for predicting BG time series have exhibited strong prediction capabilities. The research focus involves leveraging various Deep Neural Network (DNN) architectures. A recent study demonstrated that model-based deep learning surpassed the conventional approach of training traditional machines and yielded superior outcomes in BG prediction. Furthermore, challenges and limitations such as the vanishing and exploding gradient of vanilla recurrent neural network (vanilla RNNs) problems were addressed through the utilization of long short-term memory (LSTM) and recurrent gate units (GRU)[22]. This approach has seen extensive application in prior research to predict BG levels. LSTM-based models were employed to analyze the physiological patterns of BG dynamics utilizing solely CGM input[23].

J. of Science and Technology Vol. 16 No. 1 (2024) 65-75

The Bidirectional LSTM model accurately predicted BG concentrations and demonstrated superior performance compared to the standard ARIMA baseline[24]. LSTM is an effective deep-learning model utilizing gradient-based techniques. It efficiently addresses the challenge of recurrent backpropagation, which typically takes considerable time to learn to retain information over extended periods due to diminishing error backflow. LSTM, by default, can maintain data for extended durations and finds application in processing, predicting, and classifying time-series data[25]. Temporal Fusion Transformers (TFT) are known for their interpretability, setting them apart from black-box machine learning models [1]. This interpretability is crucial in healthcare applications such as blood glucose prediction, where understanding the reasoning behind predictions is essential for medical decision-making. Additionally, TFT is designed for multi-horizon time series forecasting, allowing it to predict blood glucose values accurately over varying time intervals [2]. Compared to single regression prediction models, TFT leverages model fusion techniques, such as incorporating LSTM in its architecture, to address the high volatility of blood glucose levels in diabetic patients [3]. This fusion approach enhances the model's ability to capture complex patterns and dependencies in the data, leading to more robust and accurate predictions. Furthermore, TFT transformer-based architecture enables it to efficiently capture long-range dependencies in the data, which is crucial for accurate forecasting of blood glucose values [4]. Moreover, TFT has been successfully applied in various domains beyond healthcare, such as electric demand forecasting and PV power forecasting, showcasing its versatility and effectiveness across different time series prediction tasks [5][6]. This broad applicability highlights TFT robustness and scalability in handling diverse forecasting challenges. In summary, TFT's interpretability, multi-horizon forecasting capabilities, fusion techniques, transformer-based architecture, and proven success in various forecasting tasks make it a robust and reliable model for predicting blood glucose values with the specific approach applied.

Therefore, in this work, the future BG values were predicted from the BG time series based on the TFT model in 30 and 60-minute prediction horizons (PHs). The reasons for selecting those PHs settings are because the PH of 30 to 60 minutes is considered a practical timeframe for anticipating and managing glucose level changes [26], and they are well supported by previous research works [27][28][29] for comparison purposes. The TFT model was trained, tested, and compared to the LSTM baseline models using the established ShanghaiT1DM dataset [30], the latest open-source dataset of type 1 diabetics. Furthermore, the prediction performance of each model was assessed using mean absolute error (MAE), mean absolute percentage error (MAPE), and mean squared error (RMSE) performance matrices to produce error scores.

2. Methodology

This study proposes a deep learning BG prediction with TFT and AutoTFT univariate models using past BG values from CGM (mg/dL) to time-series values as input to the univariate model. The difference between TFT and AutoTFT is that TFT is only trained based on fixed set values of hyperparameters model during training that refers to previous work that has been produced. Meanwhile, AutoTFT is trained with several hyperparameter models based on an auto-tuning function, with the same approach for LSTM and AutoLSTM as baseline models.

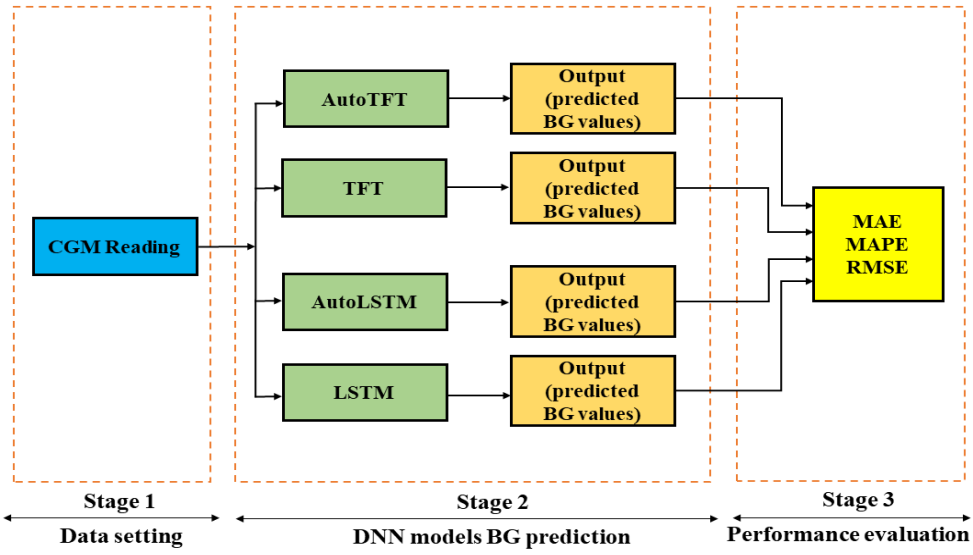

Fig. 1 *Block diagram of the BG prediction framework*

The overall research block diagram of this study is shown in Fig 1. which consists of three main stages which are the data setting, BG prediction based on the DNN models, and performance evaluation of the models. In Stage 1, the data were set for preprocessing before it was used in Stage 2 for the DNN modeling. In Stage 2, the models of four univariate models were trained with the BG time-series data and according to a time-series cross-validation method for selecting and optimizing hyperparameters[31][32][33], in 30 and 60 minutes prediction horizons (PHs). The four univariate models are the proposed models (AutoTFT and TFT) and baseline models (AutoLSTM and LSTM) as a comparison of the best model performance. Finally, Stage 3 is the performance evaluation of the models for BG time-series prediction using MAE, MAPE, and RMSE. This step is also greatly assisted by the implementation of cross-validation to provide a more robust and accurate evaluation of model performance by avoiding overfitting on a particular data set. This is done by dividing the data into chunks and ensuring that the model is tested on data that has never been seen during the training process[34][35]. Here, the performance accuracy of the model to produce BG prediction is important to maintain glucose for Type 1 diabetes patients in acceptable range glucose values which are 90-130 mg/dL, with the minimum threshold to avoid hypoglycemia being less than 80 mg/dL, and to effectively manage blood glucose levels and reduce the risk of diabetes-related complications[36][37].

2.1 Stage 1: Data Setting

We conducted an 80/20 data split for each individual in the ShanghaiT1DM dataset, using the first 80% of the data for training and the remaining 20% for testing. A validation set was created from the last 25% of each training set such as Figure 2. In the current work on machine learning-based BG prediction, this two-step data split is frequently employed[38][39] [40][41].

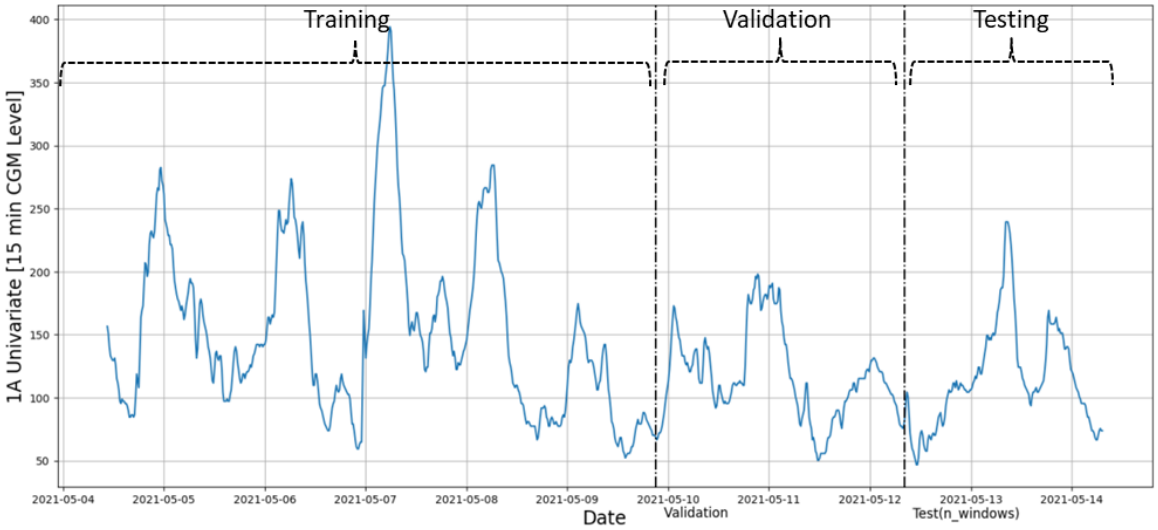

Fig. 2 *Graf of individual ShanghaiT1DM dataset*

The summary of the ShanghaiT1DM dataset used in this study which consists of 16 samples of CGM data is shown in Table 1. The dataset used is the ShanghaiT1DM dataset, and it is publicly available for research [30].

Table 1 *Summary of the ShanghaiT1DM dataset (T1DM)[30]*

Patient Number	Record Number	Age	Gender (Female=1, Male=2)	Total Number of CGM Samples data
1001_0_20210730	1A	66	1	658
1002_0_20210504	**2A**	**68**	**2**	**948**
1002_1_20210521	**2B**	**68**	**2**	**933**
1002_2_20210909	**2C**	**68**	**2**	**357**
1003_0_20210831	3A	37	2	1339
1004_0_20210425	4A	67	1	917
1005_0_20210522	5A	58	2	1256
1006_0_20210114	**6A**	**57**	**2**	**1236**
1006_1_20210209	**6B**	**57**	**2**	**1339**
1006_2_20210303	**6C**	**57**	**2**	**1135**

1007_0_20210726	7A	40	2	1337
1008_0_20210713	8A	73	1	766
1009_0_20210803	9A	59	1	681
1010_0_20210915	10A	65	1	918
1011_0_20210622	11A	51	1	536
1012_0_20210923	12A	53	1	1339

The dataset is from 12 adult T1DM patients which were acquired in actual life conditions. Of the 12 T1DM patients, there are 10 patients with one CGM recording period and the other two with three recording periods (written in bold in Table 1) which makes 16 samples in total. The dataset contains CGM readings (mg/dL) for 3 to 14 days with a sample interval of 5 minutes per sample, insulin doses, and daily dietary information. In addition, clinical characteristics, laboratory measurements, and patient care are provided. For the deep learning modeling of the BG time-series prediction, the data was indexed based on the date and time using the datetime function in Python, and a series of BG values were set as input to predict the future BG value sequentially[30]. The CGM readings have missing values for a variety of realistic causes, including sensor calibration and signal loss. The missing values were handled by applying clip values and linear extrapolation with the 40–400 mg/dL sensor working range[42].

2.2 Stage 2: Prediction Algorithm

This research proposed model design architecture of TFT and AutoTFT models as the time-series prediction model to predict BG values with the PHs of 30 and 60 minutes, respectively[43]. MAE was chosen as a loss function of the proposed model architecture. In the meantime, hyperparameter configuration is important to achieve better prediction value performance and hyperparameter tuning is a crucial aspect of optimizing models for time series forecasting by reducing the error scores in predicting blood glucose levels. Meanwhile, the baseline models of LSTM and AutoLSTM were also developed in the same way as the proposed models for a comparative purpose on the model performance.

Initially, the hyperparameter values for all of the models were referred to as the default configuration value of Neural_forecast library[44], and previous work[45][46][42], and then the hyperparameters were tuned to optimize the models. For the TFT dan LSTM (Non-Automodel) models, the best hyperparameters were determined through trial-and-error(manual tuning)[47]. On the other hand, an auto-tuning function, Optuna was used in AutoTFT and AutoLSTM to get the best hyperparameter values. This auto-tuning function is a search algorithm from the Ray Tune library[44], which is designed for hyperparameter optimization. Optuna offers advantages in automating hyperparameter tuning, optimizing complex hyperparameter spaces, and supporting various optimization algorithms[43]. By leveraging Optuna in time series modeling, researchers can enhance the accuracy and efficiency of forecasting models[48]. This automation saves time and effort compared to manual tuning, allowing for a more systematic and comprehensive exploration of hyperparameter configurations[49].

In this research, developing the deep learning model for the BG prediction involved cross-validation with time-series data makes it possible to train and test data on a prediction model by defining a transition window across the data and sliding the window in time for better estimation of the model's predictive abilities. The time-based cross-validation was used in this study to maximize the use of the data in the model evaluation. The cross-validation method was taken from the NeuralForecast class cross-validation employs a sliding window approach to predict future periods based on past observations, maintaining chronological order. This method enhances the estimation of the model's predictive performance by considering multiple periods. It resembles a standard train-test split when only one window is used, with the test data being the last set of observations and the training set comprising earlier data[44].

Then, the prediction and test output were analyzed using MAE, MAPE, and RMSE performance metrics. The prediction models were developed using Python 3.10.12, PyTorch 2.0.1, and Nvidia GeForce GTX 1650 SUPER as the computing device on the (Graphics Processing Unit) GPU.

2.3 Stage 3: Evaluation

Stage 3 describes assessing the mode performance based on BG prediction value compared to BG actual value using statistical metrics. In BG prediction, the most common metrics error scores are the RMSE and MAE[50], as well as the MAPE, a percentage metric that offers insights into relative prediction errors[46]. Performance assessment of univariate models on 16 records of time-series data measuring BG values of T1DM patients provided by the ShanghaiT1DM dataset[30]. The deep learning models were trained in an offline environment using the Graphics Processing Unit (GPU) in Python programming language for the training data and evaluated using the testing data for each of the 16 records mentioned in Table 1. MAE, MAPE, and RMSE were used to evaluate model performance and can be calculated based on equations(1), (2), and (3), respectively.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|a_i - b_i| \quad (1)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|a_i - b_i|}{a_i} * 100\% \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(a_i - b_i)^2}{n}} \quad (3)$$

In this study, the variables are defined as follows: n represents the amount of data in the variables a_i (actual value) and b_i (predicted value). Here, a_i denotes the actual data value at the i-th observation, and b_i indicates the predicted data value at the i -th observation. The index i is the value index of each observation.

3. Results and Discussion

This section is based on the BG time-series prediction conducted by univariate input models, the proposed TFT, and AutoTFT models, which used CGM time-series data from the ShanghaiT1DM dataset. The models predict the BG value based on PHs of 30 and 60 minutes. The results from the proposed models were also compared with those from the baseline models (LSTM and AutoLSTM). Figure 3 to Figure 6 show the comparison graphs of the BG time-series data between the actual (measurement) data from the test data and its prediction data from each model. There are three zones of blood glucose levels hypoglycemia, euglycemia, and hyperglycemia. Hypoglycemia occurs when blood glucose levels drop below 70 mg/dL. Euglycemia refers to normal blood glucose levels, which range from 70 to 100 mg/dL when fasting and remain below 140 mg/dL after eating. Meanwhile, hyperglycemia occurs when blood glucose levels are above normal, with fasting levels exceeding 126 mg/dL and postprandial levels above 180 mg/dL [36][37].

Due to the limited space, the comparison graphs shown in this paper are only from the ShanghaiT1DM dataset with record numbers of 7A and 11A. The graphs show the BG values in the 30 and 60-minute PHs based on historical BG values (CGM). From the observation of the graphs between the actual and the prediction data, the two proposed models of TFT and AutoTFT especially the AutoTFT outperformed the LSTM and Auto-LSTM, and the difference is more clearly visible when observing the graph in Figure 6. Moreover, from the proposed two models, at PHs of 30 minutes, the graphs show that the actual and prediction BG values are closer than those at PHs of 60 minutes.

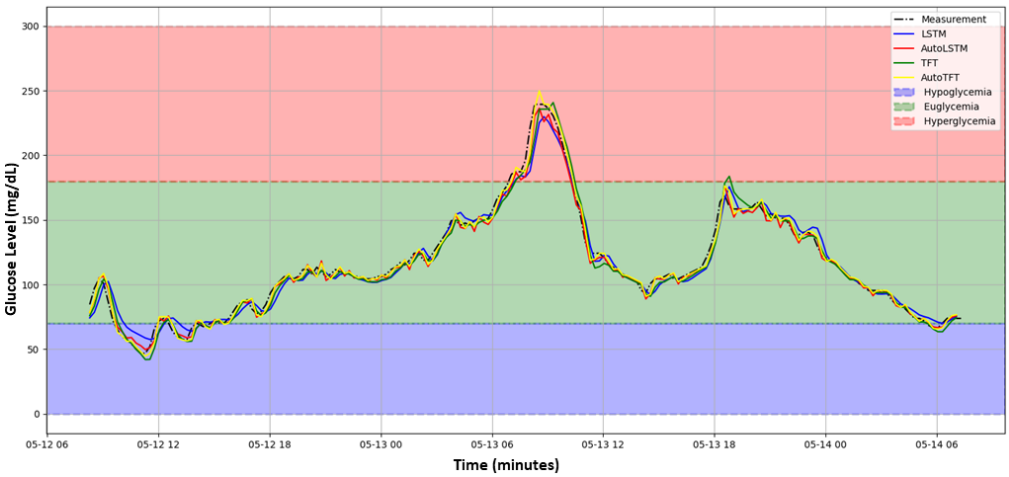

Fig. 3 *A T1D individual in the Shanghai (T1DM) dataset on record number 7A (PH=30 minutes)*

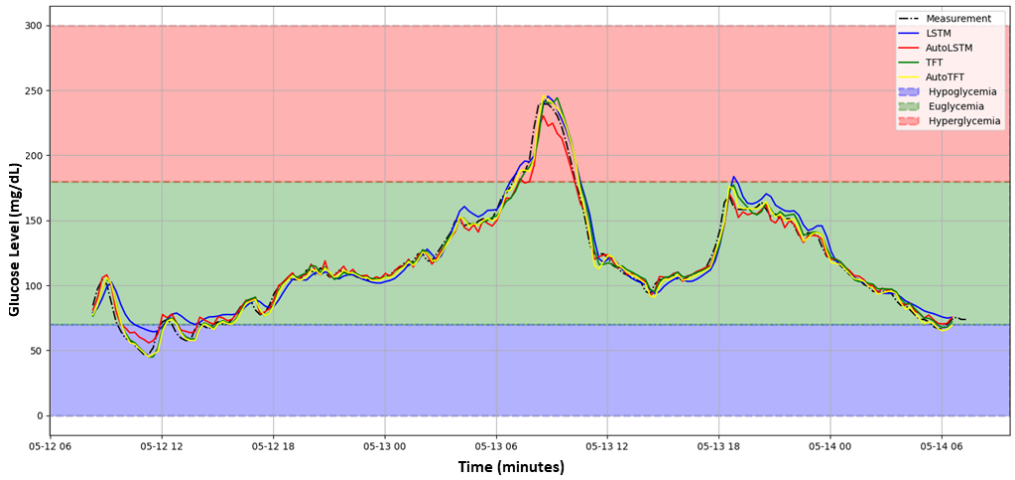

Fig. 4 *T1D individual in the ShanghaiT1DM dataset on record number 7A (PH=60 minutes)*

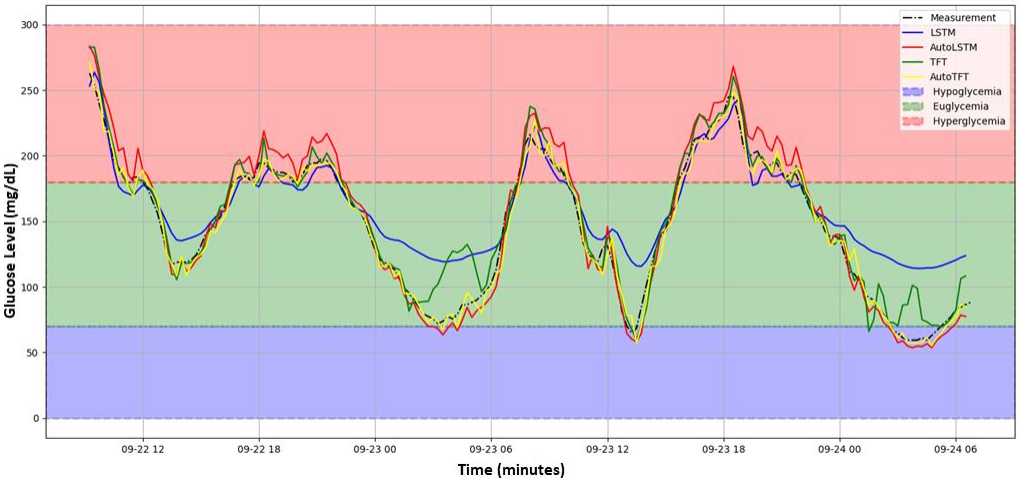

Fig. 5 *T1D individual in the ShanghaiT1DM dataset on record number 11A (PH=60 minutes)*

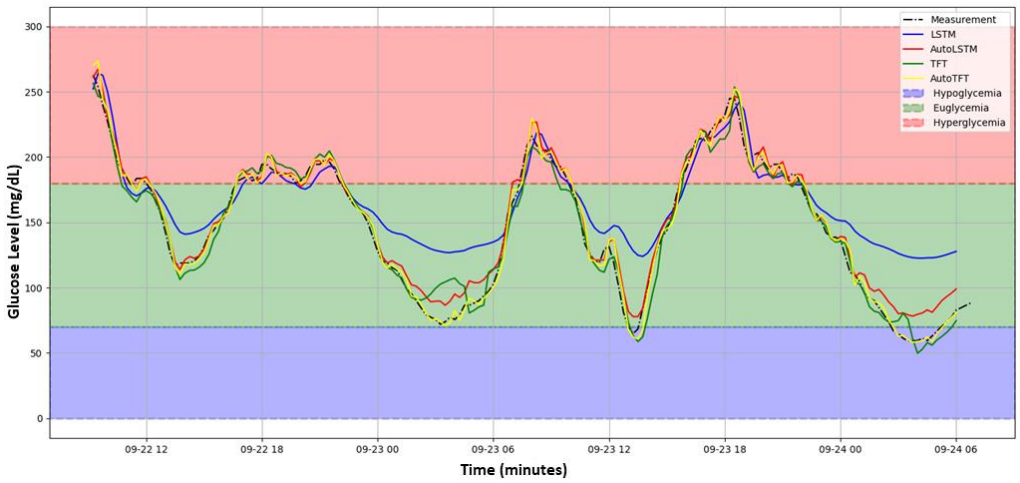

Fig. 6 *T1D individual in the ShanghaiT1DM dataset in record number 11A (PH=60 minutes)*

The performance evaluation results based on the testing of the four models, which are according to the performance metrics of error scores shown in Table 2, are for the mean error scores from a total of 16 records. According to the results, the mean error scores of AutoTFT were the smallest in MAE, MAPE, and RMSE

compared with the TFT, LSTM, and AutoLSTM. This proves that our approach to the proposed AutoTFT model is the best model compared to the other three models. Besides, by comparing with the previous research work that used the AutoTFT with a different hyperparameter auto-tuning method and with a similar dataset [45], this work has shown smaller error scores. Meanwhile, although the AutoLSTM outperformed the TFT, it is interesting to note that the proposed TFT model has smaller error scores than the LSTM baseline model. This proves that the TFT model has a better performance, compared to the LSTM model.

On the other hand, the auto-tuning function for optimal hyperparameters is important as it enhances the model's performance. This improvement is evident in the error scores shown in Table 2, indicating the significance of the hyperparameter tuning in improving not only the performance of the AutoTFT model when it is compared to the TFT model but also the performance of the AutoLSTM baseline model compared to the LSTM baseline model.

Table 2 *Mean error scores for univariate time-series model from the 16 records of the ShanghaiT1DM dataset*

Models	PH = 30 min			PH = 60 min		
	MAE (mg/dL)	MAPE (%)	RMSE (mg/dL)	MAE (mg/dL)	MAPE (%)	RMSE (mg/dL)
LSTM	11.92±1.59	10.10±1.60	16.23±2.31	13.55±2.03	11.24±2.02	18.40±2.79
TFT	8.73±0.81	6.73±0.65	13.36±1.42	9.49±0.84	7.32±0.71	14.06±1.43
AutoLSTM	7.27±0.72	5.56±0.51	9.10±0.88	8.40±1.23	6.97±1.52	11.13±1.60
AutoTFT	5.88±0.56	4.48±0.51	7.71±0.74	6.44±0.45	4.86±0.32	8.44±0.58

Regarding the prediction results between the PH of 30 and 60 minutes, from Table 2, it is found that the error scores from the PH of 30 minutes are smaller compared to those from PH of 60 minutes. The larger error scores of PH of 60 minutes could be due to the dynamic nature of the BG time-series data that makes the BG values more complex to predict in longer PHs. Furthermore, there is a tradeoff between the computation time and the performance of the prediction system when determining the length of PH, as a low PH would call for less calculation time.

4. Conclusion

The TFT model, a new high-performance model for the prediction, is presented in this research as the proposed model. Two models were proposed in this work, the TFT and AutoTFT models. They were trained and tested on the ShanghaiT1DM dataset by a specific approach using Optuna auto-tuning hyperparameters setting at the Auto model and using cross-validation as a strategy to evaluate the model performance. The TFT model is superior to the baseline LSTM model, although TFT with manual tuning has no better accuracy than AutoLSTM, the proposed AutoTFT models achieved the smallest MAE, MAPE, and RMSE for both 30 and 60-minute PHs, which stands as a promising tool in the accurate prediction of adverse glycemic events. For all the models, the BG prediction results with 30 minutes PH are better than those with 60 minutes PH. Soon, the proposed TFT model will be trained and tested on several other datasets to further evaluate the model's performance in predicting the BG values before it can potentially be implemented on edge computing devices for particular use in BG treatment and management.

Acknowledgement

The authors would like to thank Universiti Tun Hussein Onn Malaysia through TIER 1 Grant No. Q359. Extended gratitude is expressed to Microelectronics & Nanotechnology - Shamsuddin Research Centre (MiNT-SRC) for the laboratory facilities.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Ade Anggian Hakim, Farhanahani Mahmud;* ***data collection:*** *Ade Anggian Hakim, Farhanahani Mahmud;* ***analysis and interpretation of results:*** *Ade Anggian Hakim, Farhanahani Mahmud, Marlia Morsin;* ***draft manuscript preparation:*** *Ade Anggian Hakim, Farhanahani Mahmud, Marlia Morsin reviewed the results and approved the final version of the manuscript.*

References

[1] Sun, H., Saeedi, P., Karuranga, S., Pinkepank, M., Ogurtsova, K., Duncan, B. B., Stein, C., Basit, A., Chan, J. C. N., Mbanya, J. C., Pavkov, M. E., Ramachandaran, A., Wild, S. H., James, S., Herman, W. H., Zhang, P., Bommer, C., Kuo, S., Boyko, E. J., & Magliano, D. J. (2022). IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Research and Clinical Practice*, *183*, 109119. https://doi.org/10.1016/j.diabres.2021.109119

[2] Saeedi, P., Petersohn, I., Salpea, P., Malanda, B., Karuranga, S., Unwin, N., Colagiuri, S., Guariguata, L., Motala, A. A., Ogurtsova, K., Shaw, J. E., Bright, D., & Williams, R. (2019). Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Research and Clinical Practice*, *157*. https://doi.org/10.1016/j.diabres.2019.107843

[3] HENRY, W. L. (1962). Perspectives in diabetes. *Journal of the National Medical Association*, *54*(March), 476–478. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2642293/pdf/jnma00686-0078.pdf

[4] Yale, J. F., Paty, B., & Senior, P. A. (2018). Hypoglycemia. *Canadian Journal of Diabetes*, *42*, S104–S108. https://doi.org/10.1016/j.jcjd.2017.10.010

[5] Zhu, T., Li, K., Chen, J., Herrero, P., & Georgiou, P. (2020). Dilated Recurrent Neural Networks for Glucose Forecasting in Type 1 Diabetes. *Journal of Healthcare Informatics Research*, *4*(3), 308–324. https://doi.org/10.1007/s41666-020-00068-2

[6] Georga, E. I., Protopappas, V. C., & Polyzos, D. (2010). Prediction of glucose concentration in type 1 diabetic patients using Support Vector Regression. *Proceedings of the IEEE/EMBS Region 8 International Conference on Information Technology Applications in Biomedicine, ITAB*, *26500*, 1–4. https://doi.org/10.1109/ITAB.2010.5687764

[7] Ly, T. T., Breton, M. D., Keith-Hynes, P., De Salvo, D., Clinton, P., Benassi, K., Mize, B., Chernavvsky, D., Place, J., Wilson, D. M., Kovatchev, B. P., & Buckingham, B. A. (2014). Overnight glucose control with an automated, unified safety system in children and adolescents with type 1 diabetes at diabetes camp. *Diabetes Care*, *37*(8), 2310–2316. https://doi.org/10.2337/dc14-0147

[8] Lee, J. J., Gondhalekar, R., & Doyle, F. J. (2014). Design of an artificial pancreas using zone model predictive control with a Moving Horizon State Estimator. *Proceedings of the IEEE Conference on Decision and Control*, *2015-Febru*(February), 6975–6980. https://doi.org/10.1109/CDC.2014.7040485

[9] Rob J. Hyndman, & Yeasmin Khandakar. (2008). Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software*, *27*(3), 22. https://doi.org/10.18637/jss.v027.i03

[10] Laptev, N., Yosinski, J., Erran Li, L., & Smyl, S. (2017). [concept drift | LSTM | Auto encoder decoder] Time-series Extreme Event Forecasting with Neural Networks at Uber. *International Conference on Machine Learning - Time Series Workshop*, 1–5. https://www.semanticscholar.org/paper/Time-series-Extreme-Event-Forecasting-with-Neural-Laptev-Yosinski/a9b077367a8ca4cfa8425db31dd339673ddf1579

[11] Freeman, J. R., Hitt, M. P., & Pevehouse, J. C. (2014). *Univariate Time Series Models*. https://doi.org/10.1017/cbo9781139025287.004

[12] Moosa, I. A., & Moosa, I. A. (2000). Multivariate Time Series Models. *Exchange Rate Forecasting: Techniques and Applications*, 98-133. https://doi.org/10.1057/9780230379008_4

[13] Eren-Oruklu, M., Çınar, A., Rollins, D. K., & Quinn, L. (2012). Adaptive System Identification for Estimating Future Glucose Concentrations and Hypoglycemia Alarms. *Automatica*. https://doi.org/10.1016/j.automatica.2012.05.076

[14] Schmidtke, D. W., Freeland, A. C., Heller, A., & Bonnecaze, R. T. (1998). Measurement and Modeling of the Transient Difference Between Blood and Subcutaneous Glucose Concentrations in the Rat After Injection Of insulin. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.95.1.294

[15] Woldaregay, A. Z., Årsand, E., Walderhaug, S., Albers, D., Mamykina, L., Botsis, T., & Hartvigsen, G. (2019). Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes. *Artificial Intelligence in Medicine*, *98*(July), 109–134. https://doi.org/10.1016/j.artmed.2019.07.007

[16] Georga, E. I., Protopappas, V. C., Polyzos, D., & Fotiadis, D. I. (2012). A predictive model of subcutaneous glucose concentration in type 1 diabetes based on Random Forests. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2889–2892. https://doi.org/10.1109/EMBC.2012.6346567

[17] Pérez-Gandía, C., Facchinetti, A., Sparacino, G., Cobelli, C., Gómez, E. J., Rigla, M., ... & Herna19ndo, M. E. (2010). Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring. *Diabetes technology & therapeutics*, *12*(1), 81-88. https://doi.org/10.1089/dia.2009.0076

[18] Georga, E. I., Protopappas, V. C., Ardigò, D., Marina, M., Zavaroni, I., Polyzos, D., & Fotiadis, D. I. (2013). Multivariate prediction of subcutaneous glucose concentration in type 1 diabetes patients based on support vector regression. *IEEE Journal of Biomedical and Health Informatics*, *17*(1), 71–81. https://doi.org/10.1109/TITB.2012.2219876
[19] Rodrigues, F. A. (2023). Machine Learning in Physics: A Short Guide. *Epl (Europhysics Letters)*. https://doi.org/10.1209/0295-5075/ad0575
[20] Montavon, G., Samek, W., & Müller, K. R. (2018). Methods for Interpreting and Understanding Deep Neural Networks. *Digital Signal Processing*. https://doi.org/10.1016/j.dsp.2017.10.011
[21] Ullah, S., Li, S., Khan, K., Khan, S., Khan, I., & Eldin, S. M. (2023). An Investigation of Exhaust Gas Temperature of Aircraft Engine Using LSTM. *Ieee Access*. https://doi.org/10.1109/access.2023.3235619
[22] Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*. 1–9. http://arxiv.org/abs/1412.3555
[23] Martinsson, J., Schliep, A., Eliasson, B., & Mogren, O. (2020). Blood glucose prediction with variance estimation using recurrent neural networks. *Journal of Healthcare Informatics Research*, *4*, 1-18. https://doi.org/10.1007/s41666-019-00059-y
[24] Mohebbi, A., Johansen, A. R., Hansen, N., Christensen, P. E., Tarp, J. M., Jensen, M. L., Bengtsson, H., & Morup, M. (2020). Short Term Blood Glucose Prediction based on Continuous Glucose Monitoring Data. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, *2020-July*, 5140–5145. https://doi.org/10.1109/EMBC44109.2020.9176695
[25] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
[26] Battelino, T., Danne, T., Bergenstal, R. M., Amiel, S. A., Beck, R., Biester, T., Bosi, E., Buckingham, B. A., Cefalu, W. T., Close, K. L., Cobelli, C., Dassau, E., Hans DeVries, J., Donaghue, K. C., Dovc, K., Doyle, F. J., Garg, S., Grunberger, G., Heller, S., … Phillip, M. (2019). Clinical targets for continuous glucose monitoring data interpretation: Recommendations from the international consensus on time in range. *Diabetes Care*, *42*(8), 1593–1603. https://doi.org/10.2337/dci19-0028
[27] Munoz-Organero, M. (2020). Deep Physiological Model for Blood Glucose Prediction in T1DM Patients. *Sensors*. https://doi.org/10.3390/s20143896
[28] Li, K., Daniels, J., Liu, C., Herrero, P., & Georgiou, P. (2020). Convolutional Recurrent Neural Networks for Glucose Prediction. *IEEE Journal of Biomedical and Health Informatics*, *24*(2), 603–613. https://doi.org/10.1109/JBHI.2019.2908488
[29] Deng, Y., Lu, L., Becerra, L. A., Angelidi, A. M., Novak, V., Karniadakis, G. E., & Mantzoros, C. S. (2021). Deep Transfer Learning and Data Augmentation Improve Glucose Levels Prediction in Type 2 Diabetes Patients. *NPJ Digital Medicine*. https://doi.org/10.1038/s41746-021-00480-x
[30] Zhao, Q., Zhu, J., Shen, X., Lin, C., Zhang, Y., Liang, Y., Cao, B., Li, J., Liu, X., Rao, W., & Wang, C. (2023). Chinese diabetes datasets for data-driven machine learning. *Scientific Data*, *10*(1), 1–8. https://doi.org/10.1038/s41597-023-01940-7
[31] Tahkola, M., & Zou, G. (2022). ATSC-NEX: Automated Time Series Classification With Sequential Model-Based Optimization and Nested Cross-Validation. *Ieee Access*, *10*, 39299–39312. https://doi.org/10.1109/access.2022.3166525
[32] Matsui, T., Suzuki, K., Ando, K., Kitai, Y., Haga, C., Masuhara, N., & Kawakubo, S. (2022). A Natural Language Processing Model for Supporting Sustainable Development Goals: Translating Semantics, Visualizing Nexus, and Connecting Stakeholders. *Sustainability Science*, *17*(3), 969–985. https://doi.org/10.1007/s11625-022-01093-3
[33] Dong, H., Gao, X., & Wei, M. (2021). Quality Prediction of Fused Deposition Molding Parts Based on Improved Deep Belief Network. *Computational Intelligence and Neuroscience*, *2021*, 1–14. https://doi.org/10.1155/2021/8100371
[34] Zhang, F., Petersen, M., Johnson, L., Hall, J., & O'Bryant, S. (2022). Hyperparameter Tuning With High Performance Computing Machine Learning for Imbalanced Alzheimer's Disease Data. *Applied Sciences*, *12*(13), 6670. https://doi.org/10.3390/app12136670
[35] Zhang, F., Petersen, M., Johnson, L., Hall, J., & O'Bryant, S. (2021). Accelerating Hyperparameter Tuning in Machine Learning for Alzheimer's Disease With High Performance Computing. *Frontiers in Artificial Intelligence*, *4*. https://doi.org/10.3389/frai.2021.798962
[36] Bode, B. W., Schwartz, S., Stubbs, H. A., & Block, J. E. (2005). Glycemic Characteristics in Continuously Monitored Patients With Type 1 and Type 2 Diabetes. *Diabetes Care*. https://doi.org/10.2337/diacare.28.10.2361
[37] Oba, K., Inuzuka, Y., Yamashita-Onodera, N., Watanabe, K., Matsumura, N., Igari, Y., Suzuki, T., & Nakano, H. (2007). Lower Occurrence of Low Blood Glucose Level in Elderly Type 2 Diabetic Patients Treated With Gliclazide Versus Glibenclamide. *Geriatrics and Gerontology International*.

https://doi.org/10.1111/j.1447-0594.2007.00369.x
[38] He, M., Gu, W., Kong, Y., Zhang, L., Spanos, C. J., & Mosalam, K. M. (2020). CausalBG: Causal Recurrent Neural Network for the Blood Glucose Inference with IoT Platform. *IEEE Internet of Things Journal*, *7*(1), 598–610. https://doi.org/10.1109/JIOT.2019.2946693
[39] Zhu, T., Uduku, C., Li, K., Herrero, P., Oliver, N., & Georgiou, P. (2022). Enhancing self-management in type 1 diabetes with wearables and deep learning. *Npj Digital Medicine*, *5*(1), 1–11. https://doi.org/10.1038/s41746-022-00626-5
[40] Mirshekarian, S., Shen, H., Bunescu, R., & Marling, C. (2019). LSTMs and Neural Attention Models for Blood Glucose Prediction: Comparative Experiments on Real and Synthetic Data. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 706–712. https://doi.org/10.1109/EMBC.2019.8856940
[41] Zhu, T., Li, K., Herrero, P., & Georgiou, P. (2023). Personalized Blood Glucose Prediction for Type 1 Diabetes Using Evidential Deep Learning and Meta-Learning. *IEEE Transactions on Biomedical Engineering*, *70*(1), 193–204. https://doi.org/10.1109/TBME.2022.3187703
[42] Zhu, T., Chen, T., Kuang, L., Zeng, J., Li, K., & Georgiou, P. (n.d.). *Edge-Based Temporal Fusion Transformer for Multi-Horizon Blood Glucose Prediction*. https://doi.org/10.1109/ISCAS46773.2023.10181448
[43] Elshewey, A. M. (2023). *hyOPTGB: An Efficient OPTUNA Hyperparameter Optimization Framework for Hepatitis C Virus (HCV) Disease Prediction in Egypt*. https://doi.org/10.21203/rs.3.rs-2768795/v1
[44] Open Source Time Series Ecosystem (2022, December 7). *Evaluate the model's performance.* Retrieved December 7, 2022, from https://nixtlaverse.nixtla.io/neuralforecast/index.html
[45] Lim, B., Arık, S., Loeff, N., & Pfister, T. (2021). Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, *37*(4), 1748–1764. https://doi.org/10.1016/j.ijforecast.2021.03.012
[46] Zhu, T., Kuang, L., Piao, C., Zeng, J., Li, K., & Georgiou, P. (2024). Population-Specific Glucose Prediction in Diabetes Care With Transformer-Based Deep Learning on the Edge. *IEEE Transactions on Biomedical Circuits and Systems*, *18*(2), 236–246. https://doi.org/10.1109/TBCAS.2023.3348844
[47] Kim, D. Y., Choi, D. S., Kang, A. R., Woo, J., Han, Y., Chun, S. W., & Kim, J. (2022). Intelligent Ensemble Deep Learning System for Blood Glucose Prediction Using Genetic Algorithms. *Complexity*, *2022*. https://doi.org/10.1155/2022/7902418
[48] Lai, J.-P., Lin, Y.-L., Lin, H.-C., Shih, C.-Y., Wang, Y.-P., & Pai, P.-F. (2023). Tree-Based Machine Learning Models With Optuna in Predicting Impedance Values for Circuit Analysis. *Micromachines*, *14*(2), 265. https://doi.org/10.3390/mi14020265
[49] Yu, J., Zhao, Y., Pan, R., Zhou, X., & Wei, Z. (2023). Prediction of the Critical Temperature of Superconductors Based on Two-Layer Feature Selection and the Optuna-Stacking Ensemble Learning Model. *Acs Omega*, *8*(3), 3078–3090. https://doi.org/10.1021/acsomega.2c06324
[50] Zhu, T., Li, K., Herrero, P., & Georgiou, P. (2021). Deep Learning for Diabetes: A Systematic Review. *IEEE Journal of Biomedical and Health Informatics*, *25*(7), 2744–2757. https://doi.org/10.1109/JBHI.2020.3040225