Glucose Prediction with Long Short-Term Memory (LSTM) Models on Three Distinct Populations

Ubiquitous and Personal Computing Lab

Cleber Carvalho [1], Zilu Liang [2]

1. cleberfc23@gmail.com, Faculty of Engineering, Kyoto University of Advanced Science (KUAS), Kyoto, Japan
2. Faculty of Engineering, Kyoto University of Advanced Science (KUAS), Kyoto, Japan

1 Introduction

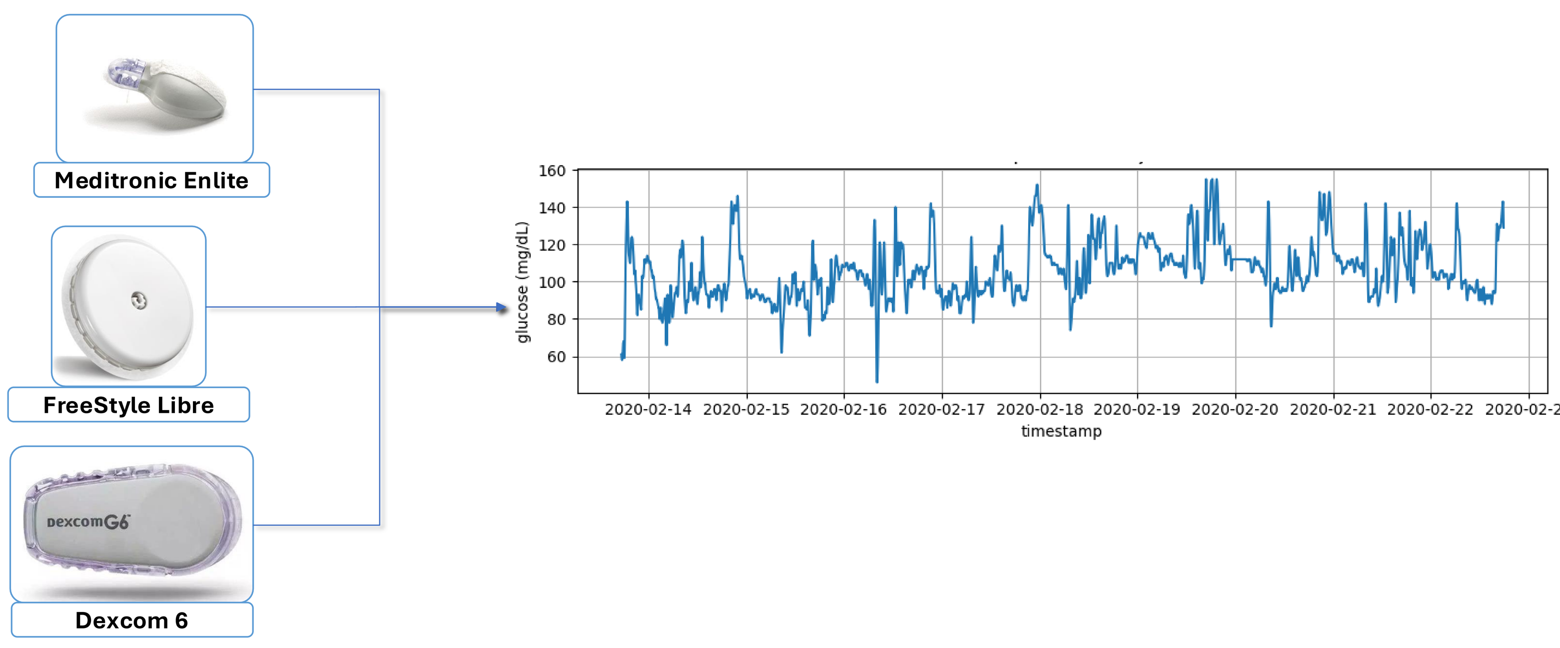
1.1 Objective

The goal of this study is to develop Long Short-Term Memory models and predict the glucose levels on three distinct diabetic populations: Type 1 Diabetes (T1D), Type 2 Diabetes (T2D) and Prediabetics (PRED). Glucose prediction is important both for healthy people, helping to improve their quality of life, and especially for diabetics, preventing cases of hyperglycemia, hypoglycemia and reducing long-term health complications[1].

The novelty of this work is related to the generalizability of the models among different populations, investigating internal and external validation. For internal validation, each model was tested on the rest of the subjects in the correspondent dataset. For external validation, each model was tested on the entirely different datasets.

1.2 Continuous Glucose Monitoring (CGM)

Continuous glucose monitoring (CGM) is a method for tracking blood glucose levels in real-time, providing a view of glucose trends and fluctuations throughout the day and generating a large amount of data. This data can be utilized to uncover insights into glycemic dynamics and their relationship to other aspects of human physiology and behavior [2,3].



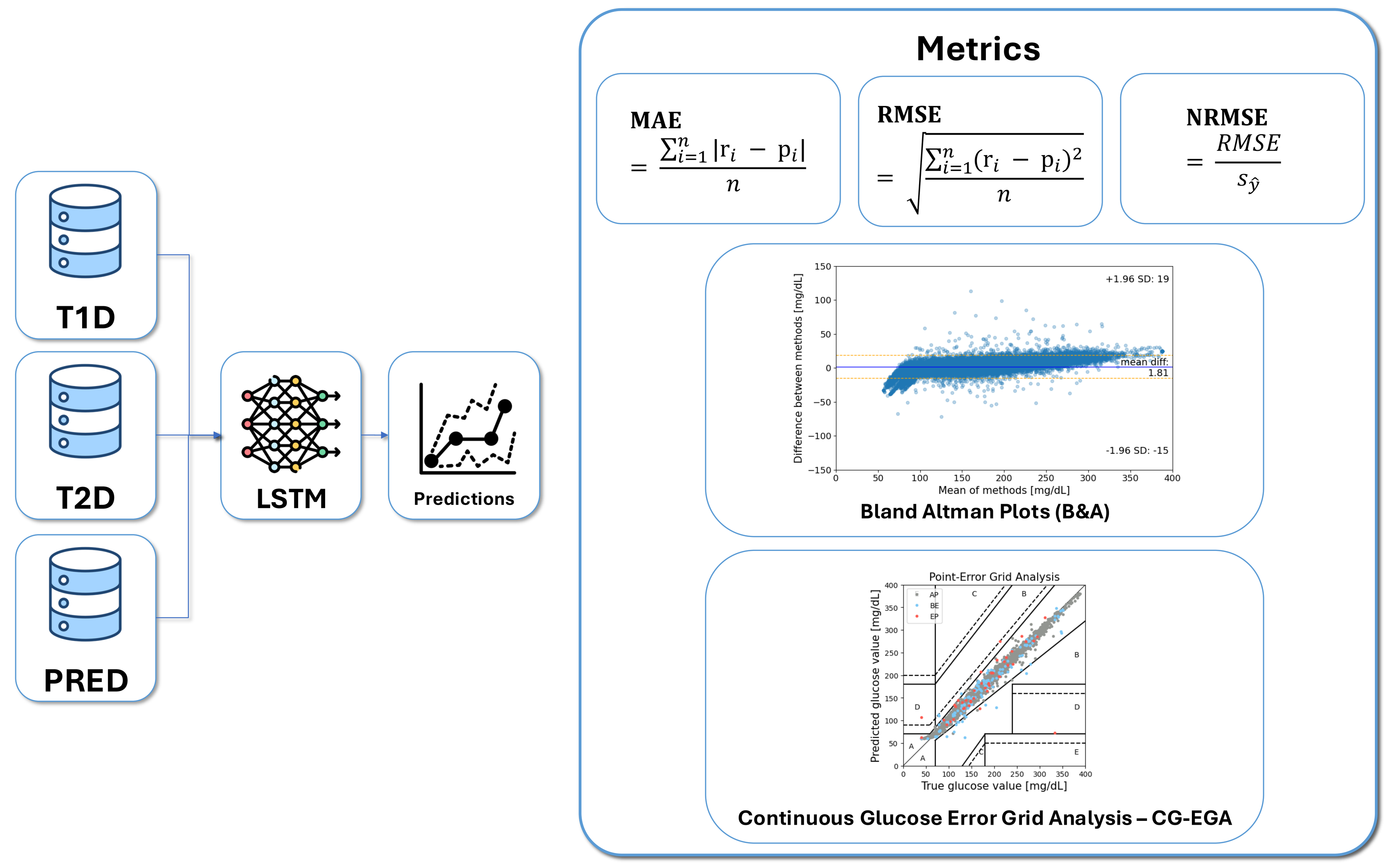
1.3 Long Short-Term Memory (LSTM)

Long Short-Term Memory is a popular Deep Learning (DL) algorithm capable of learning long-term dependencies. This quality of LSTMs helps memorize useful parts of the sequence and the model learns parameters more efficiently, making it useful for time series models.

2 Methods

Different databases were used in this study. For individuals with T1D contains data from 12 individuals. Data collection spanned over 8 weeks using Medtronic Enlite CGM sensors. Blood glucose levels were measured every 5 minutes. The second dataset has time series of blood glucose readings, after cleaning data, from 92 individuals, with T2D, who wore a FreeStyle Libre sensor for 3 to 14 days. Glucose data was automatically stored in the sensor every 15 minutes. The last dataset contains information collected from 16 prediabetic subjects. Dexcom 6 for 10 days. Glucose readings were recorded every 5 minutes.

Three LSTM models were trained with data from the subjects in each dataset separated. Then tested with subjects of the same and different test sets. The LSTM models were developed to predict the glucose level at time t + 1 based on the glucose level at time t.



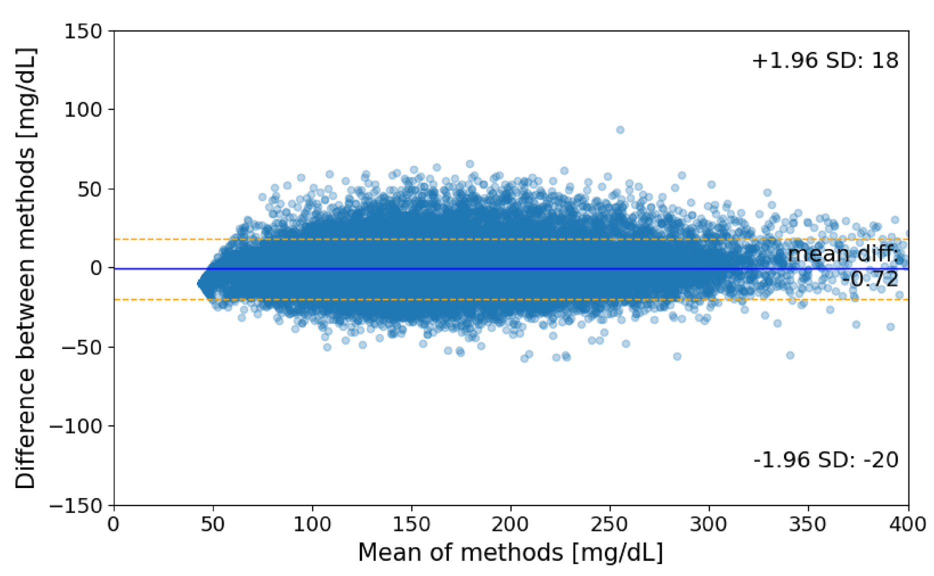
Mean absolute error (MAE), root mean squared error (RMSE) and normalized RMSE (NRMSE) were used to evaluation. The Bland-Altman (B&A) analysis compares the differences between the true values and predictions. The Continuous Glucose-Error Grid Analysis (CG-EGA) evaluates the accuracy of continuous glucose-monitoring to assess the clinical accuracy of blood glucose.

3 Results

The model **LSTM_pred** obtained better performances when compared to the other models.

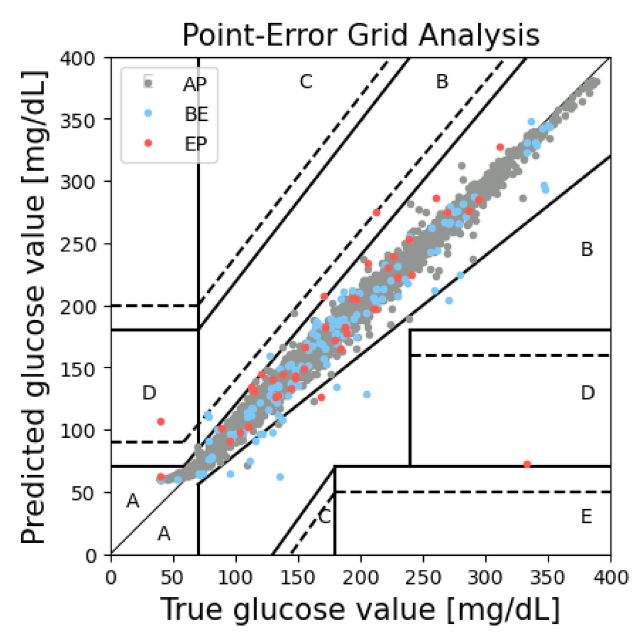
MAE and RMSE are lower metrics compared to the others. When tested on the T2D, the MAE and RMSE metrics are higher. The NRMSE metric demonstrates lower values for the T1D test set.

Model	Dataset	MAE	RMSE	NRMSE
	PRED	2.66 ± 0.54	4.00 ± 0.83	0.21 ± 0.04
LSTM_pred	T1D	4.07 ± 0.43	6.39 ± 1.25	0.11 ± 0.02
	T2D	6.83 ± 1.48	9.55 ± 2.06	0.25 ± 0.06
	PRED	2.70 ± 0.60	4.07 ± 0.88	0.22 ± 0.05
LSTM_t1d	T1D	4.24 ± 0.38	6.59 ± 1.17	0.12 ± 0.02
	T2D	6.70 ± 1.44	9.54 ± 2.04	0.25 ± 0.06
	PRED	3.11 ± 0.65	4.55 ± 0.96	0.26 ± 0.05
LSTM_t2d	T1D	5.97 ± 0.69	8.77 ± 1.18	0.17 ± 0.01
	T2D	7.45 ± 1.74	10.42 ± 2.27	0.29 ± 0.05



In all B&A plots, it was observed that most data points are within the upper and lower limits of agreements and are centered around the average. No systematic bias was detected in any of the models.

The CG-EGA results showed that the **LSTM_pred** achieved the highest values in the AP region when validated on the prediabetic test set and reasonably good AP when validated on the T1D and T2D test sets.



1. American Diabetes A. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2019. Diabetes Care. 2019; 42(Suppl 1):S61–S70. Epub 2018/12/19. https://doi.org/10.2337/dc19-S006 PMID: 30559232.
2. Liang, Z. Mining associations between glycemic variability in awake-time and in-sleep among non-diabetic adults. *Front. Med. Technol.* **2022**, *4*, 1026830.
3. Bertrand, L.; Cleyet-Marrel, N.; Liang, Z. The Role of Continuous Glucose Monitoring in Automatic Detection of Eating Activities. *In Proceedings of 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)*, Nara, Japan, **2021**, pp. 313-314.

4 Discussion

The current work has the following limitations. First, the sizes of the datasets, as the number of subjects with T2D is significantly greater than the individuals with T1D and PRED.

Second, the granularity of demographic information also varies across the datasets used, with the T1D and PRED datasets only contain age range rather than specific age of the participants.

Third, the sampling rate of the glucose levels. While the T2D data were recorded every 15 minutes, the T1D and PRED groups had their records every 5 minutes.

In our future work, we will attempt other types of deep learning algorithms, as well as utilizing larger datasets with diverse demographics. We also plan to incorporate other types of signals in addition to historical glucose data into model construction.