
Proceedings Paper:
Nemat, H., Khadem, H. orcid.org/0000-0002-6878-875X, Elliott, J. et al. (1 more author) (2020) Data fusion of activity and CGM for predicting blood glucose levels. In: Bach, K., Bunescu, R., Marling, C. and Wiratunga, N., (eds.) Knowledge Discovery in Healthcare Data 2020. 5th International Workshop on Knowledge Discovery in Healthcare Data co-located with 24th European Conference on Artificial Intelligence (ECAI 2020), 29-30 Aug 2020, Santiago de Compostela, Spain (virtual). CEUR Workshop Proceedings , pp. 120-124.

Data Fusion of Activity and CGM for Predicting Blood Glucose Levels

Hoda Nemat[1] and **Heydar Khadem**[1] and **Jackie Elliott**[2] and **Mohammed Benaissa**[1]

Abstract. This work suggests two methods—both relying on stacked regression and data fusion of CGM and activity—to predict the blood glucose level of patients with type 1 diabetes. Method 1 uses histories of CGM data appended with the average of activity data in the same histories to train three base regressions: a multilayer perceptron, a long short- term memory, and a partial least squares regression. In Method 2, histories of CGM and activity data are used separately to train the same base regressions. In both methods, the predictions from the base regressions are used as features to create a combined model. This model is then used to make the final predictions. The results obtained show the effectiveness of both methods. Method 1 provides slightly better results.

1 INTRODUCTION

The literature emphasises the importance of the management of type 1 diabetes mellitus (T1DM) in reducing complications associated with the disease [1], [2]. The key role in T1DM management is to control blood glucose level (BGL) to remain in a normal range [3], [4].

The prediction of BGL from current and past information can be a useful contributor [5]. BGL prediction could provide early warnings concerning inadequate glycaemic control to prevent the occurrence of an adverse glycemic status [6], [7].

BGL prediction models could be classified into three main groups: physiological models, data-driven models, and hybrid models. Data-driven models explain the relationship between the present and past information to BGL prediction. In this regard, machine learning and time series approaches have been widely used [5].

Many studies have proposed data-driven BGL prediction methodologies. Mirshekarian et al. [8], Bertachi et al. [9], Martinsson et al. [10], Zhu et al. [11] and Xie et al. [12] in separate studies, developed prediction models to forecast BGL with a prediction horizon of up to 60 minutes.

Mirshekarian's model was based on a recursive neural network (RNN), which utilised long short- term memory (LSTM) units. CGM, insulin, meal, and activity information were inputs of their model. Bertachi used physiological models of insulin, carbohydrate, and activity on board to train an artificial neural network (ANN). Martinsson proposed an RNN model trained on historical blood glucose information to predict BGL in two horizons of 30 and 60 minutes. Zhu generated a dilated deep convolutional neural network fed by CGM, insulin, and carbohydrate intake as inputs. Xie applied an autoregression with exogenous inputs approach to predict BGL by exploiting current and past information of CGM data.

Physical activity is a critical factor in diabetes management. Therefore, investigation of the activity data in BGL prediction models is encouraged [13]. However, developing models with high accuracy using activity and CGM data is challenging, and limited studies have been done in this area. Data fusion of activity and CGM data normally result in models with a performance not comparable with those using CGM alone.

This paper proposes two novel CGM and activity data fusion methods to generate BGL prediction models with performance comparable with those using CGM data alone.

2 DATASET

To develop BGL prediction algorithms, we used the OhioT1DM dataset [14]. The dataset contains eight weeks' worth data of 12 people with T1DM. The data of six patients was released in 2018 for the first BGL prediction challenge [15] and data for additional six patients (referred by ID 540, 544, 552, 567, 584, and 596) was released for the second BGL prediction challenge in 2020 [14]. In this work, we used the data of the latter six patients.

The dataset includes data of CGM sensor, physical activity band, physiological sensor, and self-reported life-event. Among the different collected data, we explored CGM and activity data which were collected every 5 and 1 minutes, respectively. Detailed information about the sensors and devices as well as characteristics of the patients has been published [14], [15].

In the dataset, there are three types of activity data consisting of galvanic skin response, skin temperature, and magnitude of acceleration. In this work, we only used the data of the magnitude of acceleration. Hereafter, for simplicity, 'magnitude of acceleration' is referred to as 'activity'.

3 METHODOLOGY

This section presents the information about data preprocessing and the methodologies developed for the prediction of BGL.

This paper is submitted to the second Blood Glucose Level Prediction Challenge 2020, the 5th International Workshop on Knowledge Discovery in Healthcare Data.

[1] Department of Electronic and Electrical Engineering, University of Sheffield, UK, email addresses: hoda.nemat@sheffield.ac.uk, h.khadem@sheffield.ac.uk, m.benaissa@sheffiels.ac.uk

[2] Department of Oncology and Metabolism, University of Sheffield, UK, email address: j.elliott@sheffield.ac.uk

3.1 Preprocessing

Missing data in the training set is imputed using linear interpolation. For the testing set, on the other hand, linear extrapolation is used. This is to assure that future data is not seen by the model, and that the model can be used for a real-time application. Thus, we convert CGM and activity data to regular time series without any missing data in 5-minute and 1-minute intervals, respectively.

The next step was to unify the resolution of CGM and activity data. To do so, we downsampled the activity time series data to 5-minute intervals by capturing the nearest activity data to each CGM data and discarding the rest.

There were a considerable number of unavailable activity data at the beginning and/or end of training and/or test set. This was due to the difference in wear time of CGM and activity sensors. For these points average of activity data in the training set is used rather than linear interpolation or extrapolation. Table 1 shows the number of unavailable activity data for each patient ID.

Table 1. The number of non-existent activity data points in training and testing sets per data contributor.

Patient ID	testing set	training set
540	547	31
544	0	125
552	622	505
567	0	108
584	3	123
596	80	18

Another data preprocessing step was to reframe a time series problem to a supervised learning task. To this end, time series data were transformed into samples with lag observations as input and future observations as output. We use a rolling window with a history length of 6 or 12 data points for the input, which has the information of 30- or 60- minute history, respectively. Also, the output of each sample is a vector with 6 or 12 data points corresponding to prediction horizons of 30- and 60- minute, respectively.

3.2 Regression tools

Three base regressions and a stacked regression technique are used as tools to develop the final prediction models.

3.2.1 Base regressions

- Multilayer perceptron (MLP)

MLP [16] is an ANN that can be used for time series forecasting. In this work, a single-hidden-layer MLP model was used. The model comprised a dense layer of 100 nodes with an activation function of rectified linear unit (ReLU) followed by an output layer. Adam and mean absolute error were used as an optimiser and a loss function, respectively. The learning rate was 0.01, and the model was fitted with 100 epochs.

- Long short-term memory (LSTM)

RNN is also an artificial neural network suitable for working with sequential data. We used a vanilla LSTM recurrent network [17] with vector output which is used for multi-step ahead forecast. The model was composed of a hidden layer with 200 units followed by a fully-connected layer with 100 nodes and an output layer. Both hidden layers used ReLU as the activation function. Mean squared error was the loss function, Adam was the optimiser. The model trained with 100 epochs with a learning rate of 0.01.

- Partial least squares regression (PLSR)

PLSR carries considerable popularity in different applications, such as glucose sensing [18]. In this work, PLSR was applied as a regression tool. Different values were considered for the number of components—ranging from one to the length of the input window. Each time, the predicted residual sum of squares ($PRESS$) was calculated as follows.

$$PRESS = \sum_{i=1}^{N} (y - \hat{y}_i)^2 \quad (1)$$

Where, N is the size of the evaluation set, and y_i is the reference value, and $\hat{y}_i$ is the predicted value.

The number of components (A) resulting in the minimum value for $PRESS/(N - A - 1)$ is then selected [19].

3.2.2 Stacked regression

Stacked regression is applied to enhance the performance of BGL prediction [20]. This technique uses predictions from a number of models—first-level models—as features to train a new model—second-level model. In this work, a stacked regression structure was employed where the three base regressions mentioned in 2.3.1 were set as its first-level models and a PLSR as the second-level model (Figure 1).

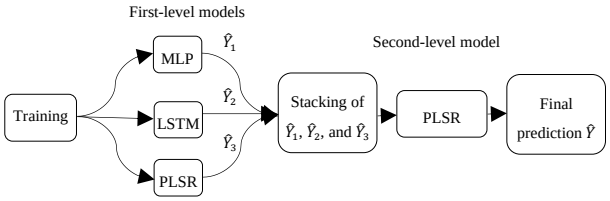

Figure 1. Diagram of the developed stacked regression.

3.3 Prediction methods

We developed two different methods using the stacked regression structure mentioned above to fuse CGM and activity data. Using these methods, models were then created to predict BGL of each patient for both horizons of 30 and 60 minutes. For each prediction horizon, two histories of 30 and 60 minutes were tried for training purposes.

3.3.1 Method 1

This method used the average value of activity data added to the window of CGM data to train the first-level models.

3.3.2 Method 2

In this method, the first-level models were trained twice. Once using a history of CGM data, and once using a history of activity data, thus producing six first-level models rather than three.

Table 3. Evaluation results of the first-level models of Method 1 using a history of 30 minutes.

Patient ID	Model	PH: 30 min		PH: 60 min	
		RMSE	MAE	RMSE	MAE
540	PLSR	22.13	16.60	41.09	31.74
	MLP	21.96 ± 0.29	16.46 ± 0.21	40.53 ± 0.38	30.95 ± 0.33
	LSTM	**21.22 ± 0.12**	**15.82 ± 0.08**	**39.65 ± 0.28**	**30.38 ± 0.28**
544	PLSR	18.08	13.33	31.80	24.71
	MLP	17.95 ± 0.07	12.87 ± 0.13	31.61 ± 0.32	24.27 ± 0.71
	LSTM	**17.62 ± 0.20**	**12.60 ± 0.32**	**30.79 ± 0.29**	**23.02 ± 0.67**
552	PLSR	16.76	12.77	30.23	23.67
	MLP	16.96 ± 0.19	12.69 ± 0.21	30.38 ± 0.36	23.42 ± 0.61
	LSTM	**16.44 ± 0.17**	**12.18 ± 0.22**	**29.89 ± 0.47**	**22.53 ± 0.40**
567	PLSR	20.97	15.04	37.41	28.15
	MLP	21.44 ± 0.63	15.60 ± 0.76	37.96 ± 1.45	29.01 ± 1.35
	LSTM	**20.61 ± 0.20**	**14.64 ± 0.32**	**36.36 ± 0.31**	**27.08 ± 0.43**
584	PLSR	22.07	16.21	36.85	27.85
	MLP	21.60 ± 0.12	15.61 ± 0.14	**36.54 ± 0.74**	**27.27 ± 0.89**
	LSTM	**21.55 ± 0.26**	**15.58 ± 0.27**	36.75 ± 1.69	27.62 ± 2.08
596	PLSR	17.79	12.76	29.63	22.05
	MLP	18.01 ± 0.16	12.99 ± 0.17	29.75 ± 0.69	21.93 ± 0.38
	LSTM	**17.23 ± 0.17**	**12.25 ± 0.29**	**29.17 ± 0.22**	**21.29 ± 0.32**
Average	PLSR	19.63	14.45	34.50	26.36
	MLP	19.65 ± 0.24	14.37 ± 0.27	34.46 ± 0.66	26.14 ± 0.71
	LSTM	**19.11 ± 0.19**	**13.85 ± 0.25**	**33.77 ± 0.54**	**25.32 ± 0.70**

3.4 Evaluation

In the Ohio dataset, the last 10 days' worth of data for each contributor was allocated as the testing set and the rest as training [14]. To train and evaluation purposes, we used the training and testing sets, respectively. Extrapolated data and, the first 60 minutes of the test set was excluded when calculating the evaluation metrics. The latter is because the testing set starts immediately after the training set, and they are chronologically close to each other. Summarised statistics of the testing set for each patient is given in Table 2.

Table 2. The statistics of the patients' testing set.

Patient ID	Original data point	Imputed data point	Evaluation data point
540	2896	3066	2884
544	2716	3136	2704
552	2364	3950	2352
567	2389	2871	2377
584	2665	2995	2653
596	2743	3003	2731

Root mean square error (RMSE) and mean absolute error (MAE) were calculated as follows and considered as evaluation metrics.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \quad (2)$$

$$MAE = \frac{\sum_{i=1}^{N}|y_i - \hat{y}_i|}{N} \quad (3)$$

Where y_i, $\hat{y}_i$, and N have the same meaning as in (1).

4 RESULTS AND DISCUSSION

In this section, the results of RMSE and MAE for prediction models are provided for both prediction horizons of 30 and 60 minutes. Models with a performance dependent on random initialisation ran five times, and the mean and standard deviation of results are reported. We have used the acronym PH for the prediction horizon in the tables.

4.1 Method 1

Table 3 displays the evaluation results of the first-level models of Method 1 when a history of 30 minutes is used for training. Based on the RMSE and MAE values, in both prediction horizons, LSTM had the best prediction performance for all patients except 584. For this patient, MLP had the best result. PLSR, as a simple linear regressor, produced results comparable to the non-linear neural network models.

Table 4. Evaluation results of the second-level model of Method 1 using a history of 30 minutes.

Patient ID	PH: 30 min		PH: 60 min	
	RMSE	MAE	RMSE	MAE
540	21.19 ± 0.07	15.73 ± 0.09	39.41 ± 0.09	30.04 ± 0.15
544	17.40 ± 0.08	12.45 ± 0.08	30.48 ± 0.07	22.90 ± 0.08
552	16.25 ± 0.07	12.02 ± 0.05	29.32 ± 0.09	22.21 ± 0.02
567	20.40 ± 0.07	14.44 ± 0.07	36.12 ± 0.02	27.12 ± 0.07
584	21.54 ± 0.06	15.62 ± 0.06	36.27 ± 0.15	27.17 ± 0.16
596	17.17 ± 0.10	12.13 ± 0.09	28.77 ± 0.26	20.80 ± 0.17
Average	18.99 ± 0.08	13.73 ± 0.07	33.39 ± 0.12	25.04 ± 0.11

Table 5. Evaluation results of the first-level models of Method 1 using a history of 60 minutes.

Patient ID	Model	PH: 30 min		PH: 60 min	
		RMSE	MAE	RMSE	MAE
540	PLSR	22.10	16.58	41.10	31.76
	MLP	21.58 ± 0.28	16.12 ± 0.22	40.53 ± 1.23	31.12 ± 0.91
	LSTM	**21.11 ± 0.18**	**15.56 ± 0.11**	**39.18 ± 0.37**	**30.00 ± 0.33**
544	PLSR	18.09	13.33	31.83	24.71
	MLP	18.09 ± 0.03	**13.05 ± 0.08**	32.34 ± 1.00	24.80 ± 1.76
	LSTM	**18.04 ± 0.35**	13.06 ± 0.48	**30.79 ± 0.39**	**23.15 ± 0.68**
552	PLSR	**16.79**	12.78	30.25	23.67
	MLP	17.58 ± 0.46	13.39 ± 0.70	**30.16 ± 0.43**	**22.89 ± 0.14**
	LSTM	16.97 ± 0.78	**12.59 ± 0.55**	30.69 ± 0.70	23.19 ± 0.55
567	PLSR	20.99	15.03	37.51	28.21
	MLP	21.71 ± 0.92	15.80 ± 1.06	37.34 ± 0.78	28.02 ± 0.76
	LSTM	**20.74 ± 0.50**	**14.75 ± 0.59**	**36.67 ± 0.98**	**27.52 ± 1.06**
584	PLSR	22.04	16.19	37.04	27.97
	MLP	22.10 ± 0.25	15.98 ± 0.23	37.13 ± 0.74	27.68 ± 0.89
	LSTM	**21.66 ± 0.10**	**15.63 ± 0.12**	**36.76 ± 0.46**	**27.18 ± 0.44**
596	PLSR	17.62	12.66	**29.48**	21.97
	MLP	18.05 ± 0.29	12.71 ± 0.27	29.71 ± 0.35	21.83 ± 0.21
	LSTM	**17.58 ± 0.19**	**12.55 ± 0.34**	29.55 ± 0.52	**21.63 ± 0.34**
Average	PLSR	19.60	14.43	34.53	26.38
	MLP	19.85 ± 0.37	14.51 ± 0.43	34.54 ± 0.75	26.06 ± 0.78
	LSTM	**19.35 ± 0.35**	**14.02 ± 0.36**	**33.94 ± 0.57**	**25.44 ± 0.57**

Table 4 shows the evaluation results of the second-level model of Method 1 when a history of 30 minutes was used for training. Comparing these results with those in Table 3, the second-level model resulted in better prediction performance than all the first-level models for all patients and both prediction horizons. This means that the stacked regression technique helped improve prediction performance.

Table 5 displays the evaluation results of the first-level models of Method 1, when a history of 60 minutes was used for training. As results show, for both prediction horizons, LSTM had the best performance for a majority of the patients. In overall, PLSR provided the second-best results.

The evaluation results of the second-level model of Method 1 using 60-minute history are shown in Table 6. In comparison with Table 5, it can be observed that the stacked regression technique advanced the prediction performance for all patients for this history, too. Also, in comparison with Table 4, Method 1 had a better overall performance when it used a history of 30 minutes than a history of 60 minutes.

Table 6. Evaluation results of the second-level model of Method 1 using a history of 60 minutes.

Patient ID	PH: 30 min		PH: 60 min	
	RMSE	MAE	RMSE	MAE
540	20.98 ± 0.13	15.50 ± 0.14	39.05 ± 0.17	29.68 ± 0.18
544	17.66 ± 0.09	12.66 ± 0.08	30.42 ± 0.36	22.82 ± 0.42
552	16.30 ± 0.09	12.04 ± 0.06	29.38 ± 0.24	22.26 ± 0.21
567	20.52 ± 0.17	14.54 ± 0.10	36.52 ± 0.10	27.31 ± 0.14
584	21.62 ± 0.17	15.63 ± 0.08	37.01 ± 0.28	27.64 ± 0.20
596	17.45 ± 0.08	12.27 ± 0.09	28.92 ± 0.27	20.92 ± 0.19
Average	19.09 ± 0.12	13.77 ± 0.09	33.55 ± 0.24	25.11 ± 0.23

4.2 Method 2

In this section, the evaluation result of Method 2 is presented. To be concise, the results of the second-level model only are reported, which are the final predictions of the method.

Table 7 shows the evaluation results of Method 2 using a 30-minute history. Comparing these results with those in Table 4, the prediction performance of Method 2 was comparable with that of Method 1 for all patients, except patient 552. This may be due to the existence of a large number of missing activity data points in this patient's data (as can be seen in Table 1).

Table 7. Evaluation results of Method 2 using a history of 30 minutes.

Patient ID	PH: 30 min		PH: 60 min	
	RMSE	MAE	RMSE	MAE
540	21.26 ± 0.09	15.89 ± 0.07	39.48 ± 0.16	30.26 ± 0.19
544	17.59 ± 0.11	12.62 ± 0.12	30.68 ± 0.15	23.14 ± 0.20
552	19.85 ± 4.51	12.65 ± 0.46	35.70 ± 3.32	23.76 ± 0.40
567	20.52 ± 0.12	14.49 ± 0.12	36.39 ± 0.20	27.14 ± 0.19
584	21.72 ± 0.17	15.78 ± 0.10	36.53 ± 0.13	27.45 ± 0.08
596	17.24 ± 0.11	12.19 ± 0.07	28.83 ± 0.11	21.03 ± 0.13
Average	19.70 ± 0.85	13.94 ± 0.16	34.60 ± 0.68	25.46 ± 0.20

Table 8 lists the evaluation result of Method 2 using a history of 60 minutes. Comparing these results with those in Table 6, the evaluation results for both methods were close to each other. Also, comparing these results with those in Table 7, Method 2 made better predictions using a history of 60 minutes than a history of 30 minutes.

Table 8. Evaluation results of Method 2 using a history of 60 minutes.

Patient ID	PH: 30 min		PH: 60 min	
	RMSE	MAE	RMSE	MAE
540	20.89 ± 0.05	15.49 ± 0.11	39.30 ± 0.35	29.80 ± 0.21
544	17.70 ± 0.14	12.68 ± 0.13	30.71 ± 0.22	23.25 ± 0.29
552	16.73 ± 0.51	12.33 ± 0.18	34.67 ± 3.51	23.47 ± 0.58
567	20.57 ± 0.14	14.63 ± 0.11	36.70 ± 0.30	27.48 ± 0.18
584	21.72 ± 0.06	15.71 ± 0.05	36.85 ± 0.09	27.69 ± 0.13
596	17.53 ± 0.21	12.26 ± 0.18	28.88 ± 0.21	21.02 ± 0.17
Average	19.19 ± 0.18	13.85 ± 0.13	34.52 ± 0.78	25.45 ± 0.26

5 SUMMARY AND CONCLUSION

This work contributes to the prediction of BGL by proposing two methodologies for data fusion of CGM and activity using stacked regression.

In the first method, the average value of activity data added to a window of CGM data was used as input to train prediction models. Initially, three base regression models consist of MLP, LSTM, and PLSR were trained. Subsequently, predictions from these base models were used as features to train a new PLSR model which then made final predictions.

In the second method, the same base regressions were trained once using windows of activity data and once using CGM data. The predictions of all trained base models were then fed as features to a new PLSR model for its training process. The new PLSR was used to make refined predictions.

The results obtained show that Method 1 (average value of activity data added to the window of CGM data) had a slightly better performance than Method 2 (first-level models trained twice, once with a history of CGM data, once using a history of activity data). In overall, Method 1 using a history of 30 minutes had the best results by providing a RMSE of 18.99 and 33.39 for the prediction horizon of 30 minutes and 60 minutes, respectively.

6 SOFTWARE AND CODE

To implement the models, we used Python 3.6, TensorFlow 1.15.0 and Keras 2.2.5. Also, Pandas, NumPy and Sklearn packages of python were used. The codes were run on a commodity laptop. The codes of our implementation are available at: `https://gitlab.com/Hoda-Nemat/data-fusion-stacking.git`

REFERENCES

[1] G. S. Jeha et al., "Continuous glucose monitoring and the reality of metabolic control in preschool children with type 1 diabetes," Diabetes Care, vol. 27, no. 12, pp. 2881–2886, 2004.

[2] L. S. Schilling et al., "A new self-report measure of self-management of type 1 diabetes for adolescents," Nurs. Res., vol. 58, no. 4, p. 228, 2009.

[3] E. R. Seaquist et al., "Hypoglycemia and diabetes: A report of a workgroup of the american diabetes association and the endocrine society," J. Clin. Endocrinol. Metab., vol. 98, no. 5, pp. 1845–1859, 2013.

[4] K. Makris and L. Spanou, "Is there a relationship between mean blood glucose and glycated hemoglobin?," J. Diabetes Sci. Technol., vol. 5, no. 6, pp. 1572–1583, 2011.

[5] A. Z. Woldaregay, E. Årsand, T. Botsis, D. Albers, L. Mamykina, and G. Hartvigsen, "Data-driven blood glucose pattern classification and anomalies detection: machine-learning applications in type 1 diabetes," J. Med. Internet Res., vol. 21, no. 5, p. e11030, 2019.

[6] J. Vehí, I. Contreras, S. Oviedo, L. Biagi, and A. Bertachi, "Prediction and prevention of hypoglycaemic events in type-1 diabetic patients using machine learning," Health Informatics J., p. 1460458219850682, 2019.

[7] C. Berra et al., "Hypoglycemia and hyperglycemia are risk factors for falls in the hospital population," Acta Diabetol., vol. 56, no. 8, pp. 931–938, 2019.

[8] S. Mirshekarian, R. Bunescu, C. Marling, and F. Schwartz, "Using LSTMs to learn physiological models of blood glucose behavior," Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, pp. 2887–2891, 2017.

[9] A. Bertachi, L. Biagi, I. Contreras, N. Luo, and J. Vehí, "Prediction of Blood Glucose Levels And Nocturnal Hypoglycemia Using Physiological Models and Artificial Neural Networks.," in 3rd International Workshop on Knowledge Discovery in Healthcare Data, 2018, pp. 85–90.

[10] J. Martinsson, A. Schliep, B. Eliasson, C. Meijner, S. Persson, and O. Mogren, "Automatic blood glucose prediction with confidence using recurrent neural networks," 3rd Int. Work. Knowl. Discov. Healthc. Data, vol. 2148, pp. 64–68, 2018.

[11] T. Zhu, K. Li, P. Herrero, J. Chen, and P. Georgiou, "A Deep Learning Algorithm for Personalized Blood Glucose Prediction.," in 3rd International Workshop on Knowledge Discovery in Healthcare Data, 2018, pp. 64–78.

[12] J. Xie and Q. Wang, "Benchmark Machine Learning Approaches with Classical Time Series Approaches on the Blood Glucose Level Prediction Challenge.," in 3rd International Workshop on Knowledge Discovery in Healthcare Data, 2018, pp. 97–102.

[13] M. H. Jensen, C. Dethlefsen, P. Vestergaard, and O. Hejlesen, "Prediction of Nocturnal Hypoglycemia From Continuous Glucose Monitoring Data in People With Type 1 Diabetes: A Proof-of-Concept Study," J. Diabetes Sci. Technol., vol. 14, no. 2, pp. 250–256, 2020.

[14] C. Marling and R. Bunescu, "The OhioT1DM Dataset for Blood Glucose Level Prediction: Update 2020," in 5th International Workshop on Knowledge Discovery in Healthcare Data, 2020.

[15] C. Marling and R. C. Bunescu, "The OhioT1DM Dataset For Blood Glucose Level Prediction.," in 3rd International Workshop on Knowledge Discovery in Healthcare Data, 2018, pp. 60–63.

[16] F. Murtagh, "Multilayer perceptrons for classification and regression," Neurocomputing, vol. 2, no. 5–6, pp. 183–197, 1991.

[17] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[18] H. Khadem, M. R. Eissa, H. Nemat, O. Alrezj, and M. Benaissa, "Classification before regression for improving the accuracy of glucose quantification using absorption spectroscopy," Talanta, vol. 211, 2020.

[19] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: a basic tool of chemometrics," Chemom. Intell. Lab. Syst., vol. 58, no. 2, pp. 109–130, 2001.

[20] L. Breiman, "Stacked regressions," Mach. Learn., vol. 24, no. 1, pp. 49–64, 1996.