Prediction-Coherent LSTM-Based Recurrent Neural Network for Safer Glucose Predictions in Diabetic People

Maxime De Bois[1(✉)], Mounîm A. El Yacoubi[2], and Mehdi Ammi[3]

[1] CNRS-LIMSI and Université Paris Saclay, Orsay, France
`maxime.debois@limsi.fr`
[2] Samovar, CNRS, Télécom SudParis, Institut Polytechnique de Paris, Évry, France
`mounim.el_yacoubi@telecom-sudparis.eu`
[3] Université Paris 8, Saint-Denis, France
`ammi@ai.univ-paris8.fr`

Abstract. In the context of time-series forecasting, we propose a LSTM-based recurrent neural network architecture and loss function that enhance the stability of the predictions. In particular, the loss function penalizes the model, not only on the prediction error (mean-squared error), but also on the predicted variation error.

We apply this idea to the prediction of future glucose values in diabetes, which is a delicate task as unstable predictions can leave the patient in doubt and make him/her take the wrong action, threatening his/her life. The study is conducted on type 1 and type 2 diabetic people, with a focus on predictions made 30-min ahead of time.

First, we confirm the superiority, in the context of glucose prediction, of the LSTM model by comparing it to other state-of-the-art models (Extreme Learning Machine, Gaussian Process regressor, Support Vector Regressor).

Then, we show the importance of making stable predictions by smoothing the predictions made by the models, resulting in an overall improvement of the clinical acceptability of the models at the cost in a slight loss in prediction accuracy.

Finally, we show that the proposed approach, outperforms all baseline results. More precisely, it trades a loss of 4.3% in the prediction accuracy for an improvement of the clinical acceptability of 27.1%. When compared to the moving average post-processing method, we show that the trade-off is more efficient with our approach.

Keywords: Glucose prediction · Recurrent neural network · Loss function · Stability · Clinical acceptability

1 Introduction

With 1.5 milion inputed deaths in 2012, diabetes is one of the leading diseases in the modern world [26]. Diabetic people, due to the non-production of insulin

(type 1) or an increased resistance to its action (type 2), have a lot of trouble managing their blood glucose. In one hand, when their glycemia falls too low (state of hypoglycemia), they are at risk of short-term complications (e.g., coma, death). In the other hand, if their glycemia is too high (hyperglycemia), the complications are long-term (e.g., cardiovascular diseases, blindness).

A lot of efforts are focused towards helping diabetic people in their daily life, with, for instance, continuous glucose monitoring (CGM) devices (e.g., FreeStyle Libre [18]), artificial pancreas (e.g., MiniMed 670G [16]), or coaching smartphone applications for diabetes (e.g., mySugr [20]). Thanks to the advances in the field of machine learning and the increased availability of data, a lot of researchers are following the lead of the prediction of future glucose values. The goal is to build data-driven models that, using the patient's past information (e.g., glucose values, carbohydrate intakes, insulin boluses), predict glucose values multiple minutes ahead of time (we call those models multi-step predictive models).

While a lot of the early work in the glucose prediction field were focused on the use of autoregressive (AR) models [22], the models that are used nowadays are more complex. Georga *et al.* explored the use of Support Vector Regression (SVR) in predicting glucose up to 120 min ahead of time in type 1 diabetes [9]. Valletta *et al.* proposed the use of Gaussian Process regressor (GP) to include a measure of the physical activity of type 1 diabetic patients into the predictive models [25]. In their work, Daskalaki *et al.* demonstrated the superiority of feed-forward neural networks compared to AR models [2]. As for them, Georga *et al.* studied the use of Extreme Learning Machine models (ELM) in short-term (PH of 30 min) type 1 diabetes glucose prediction [10]. Finally, recurrent neural networks (RNN) have shown a lot of interest in the field [27], and in particular those with long short-term memory (LSTM) units [5,15,17,24].

However, neural-network-based models, while exhibiting very promising results, often show instability in the predictions. This comes from the training of the models that, most of the time, aims at minimizing the mean-squared error (MSE) loss function. It makes the model focus on getting a good point-accuracy, without questioning the coherence of consecutive predictions.

The stability of the predictions is very important in predicting future glucose values. Predicting towards the wrong direction or with consecutive inconsistent directions can make the diabetic patient take the wrong action, potentially threatening his/her life. This is why the accuracy of the predicted glucose variations is taken into account when assessing the clinical acceptability of glucose predictive models, with, for instance, the widely-used Continuous Glucose-Error Grid Analysis (CG-EGA) [19]. We identified that this issue is not specific to the field of glucose prediction and can be extended to other multi-step forecasting applications, such as stock market prediction [6] or flood levels forecasting [1].

In this paper, to enhance the stability of the predictions, we propose a new LSTM-based RNN architecture and loss function. We demonstrate the usefulness of the idea by applying it to the challenging task of predicting future glucose values of diabetic patients which directly benefits from an increased stability.

We can summarize our contributions as follows:

1. We propose a new loss function that penalizes the model simultaneously during its training, not only on the classical MSE, but also on the predicted variation error. To be able to compute the penalty, we propose to use the loss function in a two-output LSTM-based RNN architecture. We validate the proposed approach by comparing it to four other state-of-the-art models.
2. We demonstrate the importance of making stable predictions in the context of glucose predictions as accurate but unstable predictions lead the models to have a bad clinical acceptability.
3. We confirm the overall usefullness of using LSTM-based RNN in predicting future glucose values by comparing it to other state-of-the-art models. In particular, the LSTM model shows more clinical acceptable results.
4. We have conducted the study on two different datasets, one with type 1 and one with type 2 diabetic patients. This is worth mentioning as glucose prediction studies are very rarely done on type 2 diabetes (although it represents around 90% of the whole diabetic population).
5. Finally, we have made all the source code and a standalone implementation of the CG-EGA available in Github.

The rest of the paper is organized as follows. First, we introduce the proposed architecture and loss function. Then, we present its application to the prediction of future glucose values. Finally, we provide the reader with the results and takeaways from the experiments.

2 Prediction-Coherent LSTM-Based Recurrent Neural Network

2.1 Presentation of the Model

In multi-step time-series forecasting, at time t, the model takes a set of features $\boldsymbol{X}$ to predict the future value of the time-series $\boldsymbol{y}$ at a prediction horizon PH: $\hat{y}_{t+PH}$. Most of the time, the input features $\boldsymbol{X}$ comprises the past H known values of the time-series $\boldsymbol{y}$ as well as other time-related features.

RNN, and in particular those based on LSTM cells, are neural networks that are particularly suited for time-series forecasting as they include the temporal component of the features and the predictions into their architecture [13]. Such models are usually trained with the MSE loss function (see Eq. 1) which estimates the mean accuracy of the predictions.

$$MSE(\boldsymbol{y}, \hat{\boldsymbol{y}}) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{1}$$

However, using the MSE does not incentivize the model to make successive predictions that are coherent with their respective true values. More formally, we can call two consecutive predictions, $\hat{y}_{t+PH-1}$ and $\hat{y}_{t+PH}$, coherent with the true values when the predicted variation from one to the other, $\Delta\hat{y}_{t+PH}$, reflect the true variation of the time-series Δy_{t+PH}.

To enhance to coherence of consecutive predictions, we propose the idea of using a two-output LSTM that takes advantage of its architecture to penalize incoherent successive predictions during its training. We call this neural network a Prediction-Coherent LSTM-based recurrent neural network (pcLSTM).

Two-Output LSTM. The two-output LSTM is a standard LSTM unrolled H times and that outputs the predictions of the last two steps (see Fig. 1).
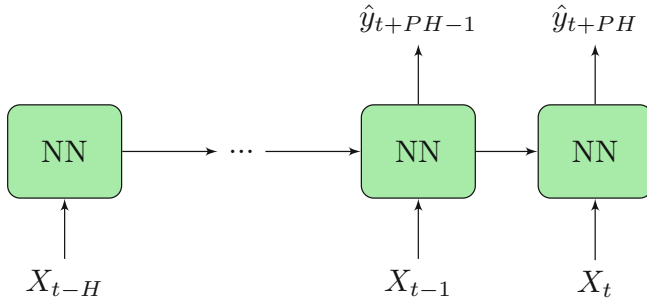

Fig. 1. Two-output LSTM which has been unrolled H times. X_t are the input features at time t and $\hat{y}_{t+PH}$ is the forecast of the time-series y at a time $t + PH$.

Variations Penalized Loss Function. To enhance the coherence between two consecutive predictions, we propose to penalize the network on the error of the predicted variation. We define the cMSE (see Eq. 2), which is the weighted sum of the MSE of the predictions and the MSE of the predicted variations. We call the parameter c the *coherence factor*. It represents the relative importance of the variation-based penalty compared to the accuracy of the predictions.

$$cMSE(\boldsymbol{y}, \hat{\boldsymbol{y}}) = MSE(\boldsymbol{y}, \hat{\boldsymbol{y}}) + c \cdot MSE(\boldsymbol{\Delta y}, \boldsymbol{\Delta \hat{y}})$$
$$= \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + c \cdot (\Delta y_i - \Delta \hat{y}_i)^2 \qquad (2)$$

The coherence factor c is a problem-dependent parameter that has to be optimized depending on the relative importance of having coherent or stable predictions versus having accurate predictions.

We note that, if the coherence factor, c, is set to 0, the $cMSE$ becomes the MSE and the model then behaves like a standard one-output LSTM model.

3   Methods

In this section, we go through the experimental details of the study, and, in particular, the data we used, the preprocessing steps we followed, the models we implemented, and the evaluation metrics we used.

We made the source code used in this study available in the pcLSTM Github repository [4].

3.1   Experimental Data

Our data come from two distinct datasets: the Ohio T1DM dataset and the IDIAB dataset accounting for 6 type 1 and 5 type 2 diabetic patients respectively.

Ohio Dataset. First published for the Blood Glucose Level Prediction Challenge in 2018, the OhioT1DM Dataset comprises data from six type 1 diabetic people who were monitored during 8 weeks [14]. For the sake of simplicity and the uniformity with the IDIAB dataset, we restrict the dataset to the glucose readings (in mg/dL), the daily insulin boluses (in units) and the meal information (in g of CHO).

IDIAB Dataset. For this study, we conducted a data collection on the type 2 diabetic population. The data collection and the use of the data in this study has been approved by the french ethical committee "Comités de protection des personnes" (ID RCB 2018-A00312-53).

Five people with type 2 diabetes (4F/1M, age 58.8 ± 8.28 years old, BMI 30.76 ± 5.14 kg/m^2, HbA1c 6.8 ± 0.71 %), have been monitored for 31.8 ± 1.17 days in free-living conditions. The patients were equipped with FreeStyle Libre (FSL) CGM devices (Abbott Diabetes Care) [18], which were recording their glucose levels (in mg/dL), and with the mySugr (mySugr GmbH) coaching app for diabetes [20], in which the patient logged his/her food intakes (in g of CHO) and insulin boluses (in units).

3.2   Preprocessing

The goal of the preprocessing part is to uniformize the two datasets and prepare them for the training and testing of the models.

Data Cleaning. To balance the training and the testing sets regarding the distribution of the samples on the daily timeline, we have chosen to remove incomplete days from the datasets. As a result, for every patient, we ended up with an average of 38.5 (± 4.82) and 29.4 (± 1.62) days worth of data for the Ohio and IDIAB datasets respectively.

We noticed that several glucose readings in the IDIAB dataset were erroneous (characterized by high amplitude spikes). As this is not particularly surprising (a study by Fokkert *et al.* reported that only 85.5% of the FSL readings were within $\pm 20\%$ of the reference sensor values [7]), we removed them to prevent them from disturbing the training of the model.

Resampling and Interpolation. To synchronize the data between them, we have resampled both datasets to get a sample every 5 min. During the resampling process, glucose values have been averaged, insulin boluses and CHO intakes have been summed up.

To make up for the introduced missing glucose values in the IDIAB dataset (which has one reading every 15 min, instead of 5), we interpolated the glucose signals as it has already been done in the context of glucose prediction [23]. In particular, we used a piecewise cubic hermite interpolating polynomial (PCHIP) [8] to avoid oscillations in the interpolated signal (which occurred with a single polynomial interpolation) and to preserve the monotonicity of the fitted signal (which was an issue with a spline interpolation) [12].

Datasets Splitting. To ready up the datasets for the training and testing of the models, we have to create the training, validation and testing sets. The splitting of the data has been done on full days of data to ensure an uniform distribution of the daily sequences across the datasets. We split the data into training, validation and testing sets following a 50%/25%/25% distribution.

Input Scaling. Lastly, the training sets data have been standardized (zero-mean and unit-variance). The same transformation has then been applied to the validation and testing sets.

3.3 Models

In this study, we compare the proposed approach (pcLSTM) to four other state-of-the-art models, namely an Extreme Learning Machine neural network (ELM), a Gaussian Process regressor (GP), a LSTM recurrent neural network (LSTM), and a Support Vector Regression model (SVR).

Every model is personalized to the patient. To be able to model long-term dependencies, every model takes the past 3 h of glucose, insulin, and CHO values as input. The hyperparameters of every model have been tuned on the validation sets by grid search.

ELM. The ELM architecture has 10^5 neurons in its single hidden layer. To reduce the impact of overfitting, we applied a L2 penalty (500) to the weights.

GP. The GP model has been implemented with a dot-product kernel. The dot-product has been chosen instead of a traditional radial basis function kernel as it has been shown to perform better in the context of glucose prediction [5]. The inhomogeneity parameter of the kernel has been set to 10^{-8}. To ease the fitting of the model, white noise (value of 10^{-2}) has been added to the observations.

LSTM. The LSTM model is made of a single hidden layer of 128 LSTM units. It has been trained to minimize the MSE loss function using the Adam optimizer with batches of 10 samples and a learning rate of 5×10^{-3}. To prevent the overfitting of the network to the training data, we added a L2 penalty (10^{-4}) and used the early stopping methodology.

pcLSTM. The pcLSTM recurrent neural network shares the same characteristics with the LSTM model. The only difference is its two-output architecture and its associated cMSE loss function (see Sect. 2). In particular, the coherence factor has been optimized through grid search to ensure a good trade-off between the accuracy of the predictions and the accuracy of the predicted variations. We settled down with a coherence factor of 2.

SVR. The SVR model has been implemented with a radial basis function (RBF) kernel. The coefficient of the kernel has been set to 5×10^{-4}. The wideness of the no-penalty tube has been set to 0.1 and the penalty itself has been set to 50.

3.4   Post-processing

By using the cMSE loss function, we incentivize the model to make consecutive predictions reflecting the actual glucose rate of change. In a way, it can be viewed as a smoothing effect integrated to the training of the model.

Some post-processing time-series smoothing techniques exist, such as the exponential smoothing or the moving average smoothing [21]. The latter, yielding a better trade-off between the accuracy of the predictions and the accuracy of the predicted variations, has been used with a window of the last 3 predictions.

3.5   Evaluation Metrics

In this study, three evaluation metrics have been used: the Root-Mean-Squared prediction Error (RMSE), the Root-Mean-Squared predicted variation Error (dRMSE), and the Continuous Glucose-Error Grid Analysis (CG-EGA) measuring the clinical acceptability of the predictions.

RMSE. The RMSE is the most used metric in the world of glucose prediction as it measures the overall accuracy of the predictions [19].

dRMSE. We call the dRMSE the RMSE applied to the difference between two consecutive predictions. Therefore, it measures the accuracy of the predicted variations and can be used to estimate the impact of the variation-based penalty in the cMSE loss function.

CG-EGA. The CG-EGA provides a measure of the clinical acceptability of the predictions [19]. Indeed, predictions, depending on the current state of the patient's glycemia (hypoglycemia, euglycemia[1], or hyperglycemia), can be more or less dangerous, which is not taken into account in metrics such as the RMSE.

Technically, the CG-EGA is made of two grids: the Point-Error Grid Analysis (P-EGA) and the Rate-Error Grid Analysis (R-EGA). Whereas the P-EGA provides an acceptability score (from A to E) to the glucose predictions, the R-EGA gives each prediction a score (also from A to E) based on the variation from the previous prediction to the current one [11]. The CG-EGA combines both grids and gives, for every prediction, in its simplified representation, a clinical acceptability category: accurate prediction (AP), benign error (BE), or erroneous prediction (EP). For a prediction to be categorized as an AP, it needs to have a score of A or B in both the P-EGA and the R-EGA.

We published the source code of the CG-EGA implementation in Github [3].

4   Results and Discussion

The results of the models, presented with and without the moving average smoothing technique discussed in Sect. 3.4, are reported in Table 1. Figure 2 gives a graphical representation of the effect of the proposed approach on the predictions. A detailed graphical clinical acceptability classification of the predictions is given by Fig. 3.
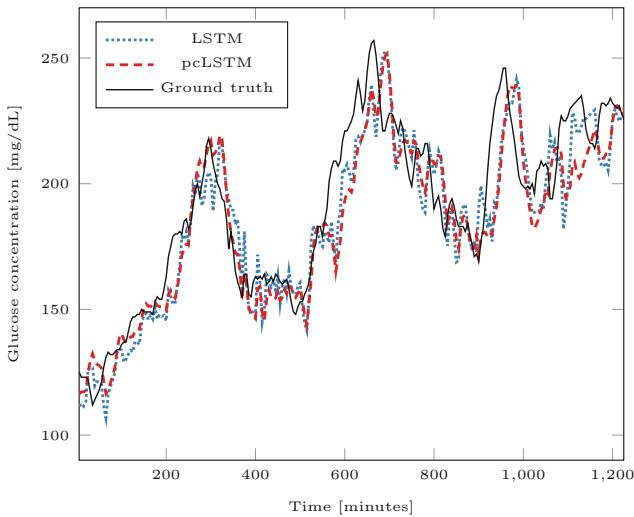

Fig. 2. Glucose predictions of the unsmoothed LSTM and pcLSTM against the ground truth, for a given day of one of the patients.

---

[1] The euglycemia region is the region between hypoglycemia and hyperglycemia.

Table 1. Performances of the ELM, GP, LSTM, pcLSTM, and SVR models, evaluated at a prediction horizon of 30 min with and without the smoothing of the predictions (mean ± standard deviation, averaged on the subjects from both datasets).

Model	RMSE	dRMSE	CG-EGA		
			AP	BE	EP
Without smoothing					
ELM	25.54 ± 5.02	1.90 ± 0.45	79.34 ± 7.53	14.92 ± 5.50	5.74 ± 2.33
GP	**18.92 ± 4.56**	2.21 ± 0.44	81.70 ± 6.21	13.88 ± 4.09	4.41 ± 2.28
LSTM	19.48 ± 4.42	1.95 ± 0.40	82.98 ± 5.65	12.42 ± 3.73	4.60 ± 2.06
pcLSTM	20.32 ± 4.56	**1.47 ± 0.31**	**87.60 ± 4.74**	**8.76 ± 3.23**	**4.01 ± 1.84**
SVR	20.08 ± 4.24	1.74 ± 0.44	83.92 ± 6.10	11.75 ± 4.24	4.32 ± 2.01
With smoothing					
ELM	26.64 ± 5.17	1.42 ± 0.31	86.13 ± 5.26	9.31 ± 3.89	4.57 ± 1.79
GP	**20.42 ± 4.70**	1.48 ± 0.31	87.17 ± 4.38	9.08 ± 3.03	**3.74 ± 1.65**
LSTM	21.21 ± 4.63	1.41 ± 0.30	87.51 ± 4.45	8.60 ± 3.15	3.89 ± 1.60
pcLSTM	22.42 ± 4.85	**1.29 ± 0.28**	**88.82 ± 4.43**	**7.36 ± 3.18**	3.81 ± 1.68
SVR	21.81 ± 4.43	1.42 ± 0.32	87.38 ± 4.78	8.79 ± 3.37	3.83 ± 1.63

First, when looking at the unsmoothed baseline results, apart from the ELM model that has overall the worse performances (excluding it from the following analysis), we can see that the models have different strengths and weaknesses. Whereas the GP model stands out as being the most point-accurate model (RMSE), it is also the most unstable model (dRMSE). This makes it the least clinically acceptable model of the remaining three, having the lowest AP and the highest EP rates. On the other hand, the SVR model has the worse RMSE, the best dRMSE, and the best AP and EP rates, making it the most clinically acceptable baseline model. Finally, the LSTM model displays competitive results with respect to the GP and SVR models, which validates the use of the LSTM model in the context of glucose prediction.

When looking at the unsmoothed performances of the pcLSTM model, we can see that, compared to the LSTM model, its RMSE is slightly worse (+4.3%), its dRMSE drastically improved (−24.6%) and so is its clinical acceptability (+27.1% and −12.8% for the room for improvement in the AP and EP rates respectively). This shows the importance of focusing on the coherence of successive predictions as the increased accuracy in predicted variations (dRMSE) is the main contributor to the increased clinical acceptability.

The results of the models with smoothed predictions show us the general benefit of improving the stability of the predictions to make them more clinically acceptable. Even though all the models see their clinical acceptability improved, the improvement varies from model to model: the models with the highest instability benefit from the smoothing the most. In average, the improvement due to the smoothing applied on the baseline models (still excluding the ELM model) is
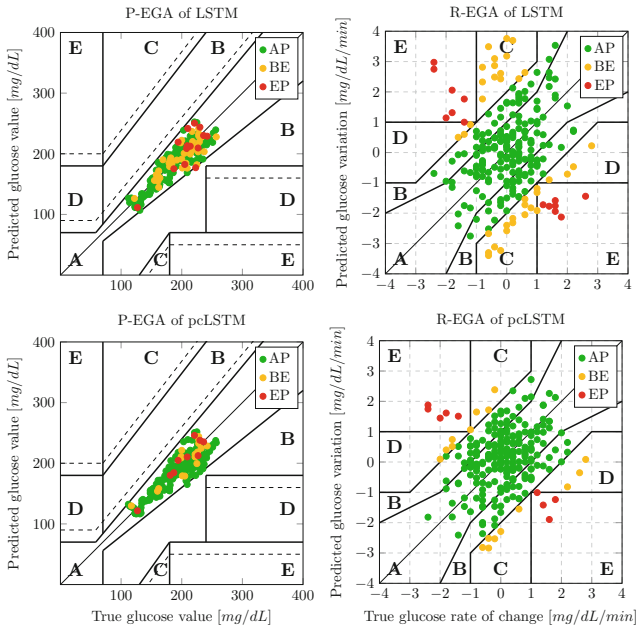
Fig. 3. P-EGA (left) and R-EGA (right) for LSTM (top) and pcLSTM (bottom) models for a patient during a given day. The CG-EGA classification (AP, BE, or EP) is computed by combining both P and R-EGA ranks.

of $+8.5\%$, -24.3%, $+26.0\%$, and -14.14% in RMSE, dRMSE, AP and EP rates respectively. Those results show us that the trade-off made by the pcLSTM is much more efficient ($+8.5\%$ against $+4.3\%$ in RMSE for overall the same improvement in the other metrics).

5 Conclusion

In this paper, we have presented a new loss function for recurrent neural networks which, by penalizing the model on the predicted variation errors in addition to the prediction errors, helps the network making more stable predictions.

We apply the proposed model to the prediction of future glucose values in diabetes. First, we validate the use of recurrent neural networks (in particular with LSTM units) by showing that our baseline LSTM model is competitive when compared to other state-of-the-art models. Then, we demonstrate the importance of the proposed approach as it greatly improves the clinical acceptability of the predictions. Lastly, we compare the proposed approach to another smoothing technique. While the effect on the clinical acceptability is the same, the loss in the accuracy of the prediction is higher, making our proposed approach more efficient.

The tuning of the coherence factor in the cMSE loss function is of paramount importance for the proposed approach. The desired stability is application dependant and must, in the case of glucose prediction, be assessed by practitioners. In the future we plan on improving the loss function further by adding penalties directly tied to the CG-EGA (e.g., penalizing the model when the prediction is an EP).

References

1. Chang, F.J., Chiang, Y.M., Chang, L.C.: Multi-step-ahead neural networks for flood forecasting. Hydrol. Sci. J. **52**(1), 114–130 (2007)
2. Daskalaki, E., Prountzou, A., Diem, P., Mougiakakou, S.G.: Real-time adaptive models for the personalized prediction of glycemic profile in type 1 diabetes patients. Diab. Technol. Ther. **14**(2), 168–174 (2012)
3. De Bois, M.: CG-EGA python implementation (2019). https://github.com/dotXem/CG-EGA, https://doi.org/10.5281/zenodo.3459590
4. De Bois, M.: PCLSTM (2019). https://github.com/dotXem/pcLSTM, https://doi.org/10.5281/zenodo.3459566
5. De Bois, M., El Yacoubi, M., Ammi, M.: Study of short-term personalized glucose predictive models on type-1 diabetic children. In: IJCNN (2019, accepted)
6. Dong, G., Fataliyev, K., Wang, L.: One-step and multi-step ahead stock prediction using backpropagation neural networks. In: 2013 9th International Conference on Information, Communications & Signal Processing, pp. 1–5. IEEE (2013)
7. Fokkert, M., et al.: Performance of the freestyle libre flash glucose monitoring system in patients with type 1 and 2 diabetes mellitus. BMJ Open Diabetes Res. Care **5**(1), e000320 (2017)
8. Fritsch, F.N., Carlson, R.E.: Monotone piecewise cubic interpolation. SIAM J. Numer. Anal. **17**(2), 238–246 (1980)
9. Georga, E.I., et al.: Multivariate prediction of subcutaneous glucose concentration in type 1 diabetes patients based on support vector regression. IEEE J. Biomed. Health Inf. **17**(1), 71–81 (2013)
10. Georga, E.I., Protopappas, V.C., Polyzos, D., Fotiadis, D.I.: Online prediction of glucose concentration in type 1 diabetes using extreme learning machines. In: Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, pp. 3262–3265. IEEE (2015)
11. Kovatchev, B.P., Gonder-Frederick, L.A., Cox, D.J., Clarke, W.L.: Evaluating the accuracy of continuous glucose-monitoring sensors: continuous glucose-error grid analysis illustrated by therasense freestyle navigator data. Diabetes Care **27**(8), 1922–1928 (2004)
12. Li, C.S.R., Yan, P., Bergquist, K.L., Sinha, R.: Greater activation of the "default" brain regions predicts stop signal errors. Neuroimage **38**(3), 640–648 (2007)
13. Mandic, D.P., Chambers, J.: Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability. John Wiley & Sons, Inc. (2001)

14. Marling, C., Bunescu, R.: The ohiot1dm dataset for blood glucose level prediction. In: The 3rd International Workshop on Knowledge Discovery in Healthcare Data, Stockholm, Sweden (2018)

15. Martinsson, J., Schliep, A., Eliasson, B., Meijner, C., Persson, S., Mogren, O.: Automatic blood glucose prediction with confidence using recurrent neural networks. In: 3rd International Workshop on Knowledge Discovery in Healthcare Data, KDH@ IJCAI-ECAI 2018, 13 July 2018, pp. 64–68 (2018)

16. Messer, L.H., et al.: Optimizing hybrid closed-loop therapy in adolescents and emerging adults using the minimed 670g system. Diabetes Care **41**(4), 789–796 (2018)

17. Mirshekarian, S., Bunescu, R., Marling, C., Schwartz, F.: Using LSTMS to learn physiological models of blood glucose behavior. In: Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE, pp. 2887–2891. IEEE (2017)

18. Ólafsdóttir, A.F., et al.: A clinical trial of the accuracy and treatment experience of the flash glucose monitor freestyle libre in adults with type 1 diabetes. Diabetes Technol. Ther. **19**(3), 164–172 (2017)

19. Oviedo, S., Vehí, J., Calm, R., Armengol, J.: A review of personalized blood glucose prediction strategies for t1dm patients. Int. J. Num. Methods Biomed. Eng. **33**(6), e2833 (2017)

20. Rose, K., Koenig, M., Wiesbauer, F.: Evaluating success for behavioral changein diabetes via mhealth and gamification: mysugr's keys to retention andpatient engagement. Diabetes Technol. Ther. **15**, A114 (2013)

21. Shumway, R.H., Stoffer, D.S.: An approach to time series smoothing and forecasting using the em algorithm. J. Time Ser. Anal. **3**(4), 253–264 (1982)

22. Sparacino, G., Zanderigo, F., Corazza, S., Maran, A., Facchinetti, A., Cobelli, C.: Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series. IEEE Trans. Biomed. Eng. **54**(5), 931–937 (2007)

23. Stahl, F., Johansson, R.: Short-term diabetes blood glucose prediction based on blood glucose measurements. In: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 291–294. IEEE (2008)

24. Sun, Q., Jankovic, M.V., Bally, L., Mougiakakou, S.G.: Predicting blood glucose with an LSTM and Bi-LSTM based deep neural network. In: 2018 14th Symposium on Neural Networks and Applications (NEUREL), pp. 1–5, November 2018. https://doi.org/10.1109/NEUREL.2018.8586990

25. Valletta, J.J., Chipperfield, A.J., Byrne, C.D.: Gaussian process modelling of blood glucose response to free-living physical activity data in people with type 1 diabetes. In: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4913–4916. IEEE (2009)

26. World Health Organization, et al.: Global Report on Diabetes. World Health Organization (2016)

27. Zarkogianni, K., Vazeou, A., Mougiakakou, S.G., Prountzou, A., Nikita, K.S.: An insulin infusion advisory system based on autotuning nonlinear model-predictive control. IEEE Trans. Biomed. Eng. **58**(9), 2467–2477 (2011)