







# Prediction of Blood Glucose Levels in Patients with Type 1 Diabetes via LSTM Neural Networks

Ciro Rodriguez Leon<sup>1</sup> , Oresti Banos<sup>1</sup> , Oscar Fernandez Mora<sup>2</sup>,  
Alex Martinez Bedmar<sup>2</sup>, Fernando Rufo Jimenez<sup>2</sup>,  
and Claudia Villalonga<sup>1</sup>  

- <sup>1</sup> Research Centre for Information and Communication Technologies (CITIC-UGR),  
University of Granada, Granada, Spain  
{crleon,oresti,cvillalonga}@ugr.es
- <sup>2</sup> School of Engineering and Technology, Universidad Internacional de La Rioja  
(UNIR), Logroño, Spain  
{oscar.fernandez007,alex.martinez559,fernando.rufo448}@comunidadunir.net

**Abstract.** Diabetes is one of the most prevalent diseases of the 21st century, with more than 500 million people affected. Having tools to estimate blood glucose levels is critical for these patients in their management of the disease. In this work, we present a comparison of three neural network architectures based on long short-term memory (LSTM). Their predictive ability has been evaluated against a longitudinal dataset with continuous glucose level measurements of patients with type 1 diabetes. All models, trained for different prediction horizons of 30, 60, 90 and 180 min, have generally yielded good prediction results. These results are further validated using clinical standards resulting in more than 95% of accurate blood glucose level predictions, mostly leading to correct treatments.

**Keywords:** Type 1 Diabetes Mellitus · Deep Learning · Blood glucose level prediction

## 1 Introduction

Diabetes, a metabolic disorder that affects the way the body processes and uses blood sugar, is one of the most common chronic diseases worldwide. In fact, 537 million adults are living with diabetes and according to predictions this number will rise to 643 million by 2030 and 783 million by 2045. The economic impact of diabetes in 2021 was at least \$966 billion in health expenditures, which was a 316% increase over the last 15 years. In Europe, one in eleven adults has some form of diabetes, accounting for 61 million in total, even if it is estimated that one in three adults has undiagnosed diabetes. Furthermore, the number of adults with diabetes in Europe is expected to grow up to 67 million by 2030 and 69

million by 2045 [7]. This situation will place a large burden on healthcare systems and have a significant economic and social impact.

Although there is no cure for diabetes, it can be controlled through medication, insulin and healthy lifestyles, such as having a healthy diet and exercising regularly. In fact, it is important for patients with diabetes to maintain adequate control of their disease, as this can help prevent long-term complications such as heart disease, stroke, kidney damage and vision loss. The ability to predict the evolution of blood glucose levels in the near future can also be useful for patients with diabetes, as it allows them to anticipate how their blood glucose level will evolve in the coming hours or days. This is specially important for patients with type 1 diabetes mellitus, one of the main types of diabetes which occurs when the body's immune system destroys insulin-producing cells in the pancreas, preventing the body from producing enough insulin to regulate blood sugar levels [9]. Patients with type 1 diabetes mellitus are the ones who find more difficult to maintain their glucose levels in range. Therefore, knowing the future evolution of blood glucose levels will allow them to adapt their lifestyle so that they can maintain these levels close to those of a healthy person and prevent possible complications related to this chronic disease.

One of the applications of IoT is real-time continuous glucose monitoring in diabetic patients. In the first systems marketed in 1999 the devices stored glucose level information and afterwards transmitted and analyzed it. Today systems measure not only glucose, but also blood pressure, temperature, physical activity and dietary data via mobile apps that directly transmit the data to a server [5]. Continuous blood glucose level monitoring has been a major advance in the management of diabetes, as it provides an accurate record of the evolution over time of blood glucose levels in patients with diabetes. This has allowed the development of models that try to predict how blood glucose levels will vary over a short period of time from previous blood glucose level measurements, insulin administered and other collected data. These predictive models can be very useful for patients with diabetes, as they allow them to anticipate how their blood glucose level may be affected by each decision they make in their daily lives, such as the amount of carbohydrates they consume, the amount of exercise they do, or the dose of insulin they take.

This work compares several neural networks used in blood glucose level prediction for patients with type 1 diabetes mellitus. The neural networks are trained with data obtained from continuous glucose monitoring systems. These algorithms use deep learning techniques to process large amounts of data and try to predict how blood glucose levels will vary in the near future. The results of this work may help to improve the accuracy of the proposed neural networks and thus improving the effectiveness of short-term blood glucose level prediction for patients with type 1 diabetes mellitus.

## 2 State of the Art

Maybe the first approach to glucose prediction in patients with type 1 diabetes using deep learning is presented in 1999 by Tresp, Briegel and Moody [19]

Several Recurrent Neural Network (RNN) models are evaluated and compared with other linear and nonlinear models. Insulin levels, meals, exercise level, and current and previous estimates of blood glucose are used to train the models. An RNN is also trained in [2], but in this case signals from a continuous monitoring devices are used as input. Prediction horizons of 15, 30, 45 and 60 min are compared with the results of a standard feed-forward network and it is found that long-term estimates are more accurate than ones obtained with the RNN.

Long short-term memory (LSTM) networks have been widely used to predict blood glucose levels in patients with type 1 diabetes mellitus. For example, [17] describes a sequential model in which an LSTM and a bidirectional LSTM (BiLSTM) of four units each are combined with three fully-connected layers. 26 datasets from 20 different patients, real and in silico, are used in the evaluation, which shows that the LSTM model improves the predictions of the classical models. [6] presents an LSTM network to characterize the temporal dimensions of the data and two dense layers to extract features. This work tests different combinations of hyperparameters for 10-patient data, obtaining the best results with 50 units in the LSTM and 30 for each dense layer. [12] proposes an LSTM architecture based on a physiological model from which the dependencies between the parameters are extracted. The three-layer architecture, with an LSTM layer and a dense layer for the results, is trained using glucose, insulin, sleep and exercise levels data from real patients in a total of 1600 days. [13], a posterior version of the previous work, proposes an LSTM coupled with a neural attention model. [1] proposes two LSTM networks working in parallel and then connected in a fully connected layer. The first network works with observed data and the second one with estimated data. To improve the model, the weights of the LSTM are adjusted for each patient, obtaining good results both in real patients and in silico at different prediction horizons. [14] proposes four models consisting on an LSTM layer followed by a dense layer, one for each of the inputs: glucose, carbohydrates and fast and slow insulin units. Once the inputs are processed separately, the networks for insulin and carbohydrate concatenate, returning a prediction and then, the glucose information is concatenated to evaluate the final glucose values. [11] also presents an architecture with one LSTM layer that alternates with two fully connected layers, but treats glucose predictions as a classification problem, rather than a classical time series problem. Hypo- and hyperglycemia ranges are normalized and divided into 100 bins, which will be the different classes returned by the model.

LSTM networks are also predominant in the models presented to the second Blood Glucose Level Prediction (BGLP) Challenge, which took place in 2020. In this challenge, the OhioT1DM dataset [10] was used by several researchers to train their own models and to compare the efficacy of their different prediction approaches. The results of BGLP are presented in [3], where eight systems that conformed to the challenge rules are ranked based on their errors for 30 and 60 min prediction horizons. The best prediction model is [16], a neural network architecture based on Neural Basis Expansion for Interpretable Time-Series Forecasting (N-BEATS) but replacing the fully connected block structure of

N-BEATS with LSTMs. This winning work presents an architecture which learns to forecast gradually in stages or blocks. Each residual block contains a BiLSTM with a single output layer that produces the forecast and back projection, and additional variables are added as input channels to each block. In fact, this and other ensemble models have been recently used to estimate blood glucose levels in patients with type 1 diabetes mellitus. These approaches train multiple models and combine their independent outcomes into a unified prediction. For example, [8] proposes a system that combines six models called base-learners: two LSTM networks, two Multilayer perceptrons (MLP) and two Partial Least Square Regression (PLSR) models. These base-learners converge into a PLSR layer, the meta-learner, which provides the output prediction of the blood glucose level. Two ensembles based on Bayesian voting to predict the blood glucose level are presented in [18]. These ensembles use three and four LSTM models, respectively, which are selected as the best from a set of ten different neural network architectures. The two proposed ensembles are compared with many of the previously described models. The OhioT1DM dataset is also used to evaluate them under the same conditions at prediction horizons of 30, 60 and 120 min and using the variables glucose levels, basal insulin, insulin dose and carbohydrate intake. The work concludes that there is little difference in predictive capacity since the values of the performance metrics are very close, and the confidence intervals overlap. In fact, although differences have been found statistically between the worst and the best models, from a medical perspective they are irrelevant.

### 3 Methodology

The objective of this work is to evaluate the performance of three popular recurrent neural network architectures in the field of glucose prediction: long short-term memory (LSTM), bidirectional LSTM (BiLSTM) and convolutional LSTM (ConvLSTM). The evaluation will be performed for different prediction horizons when training the models with a longitudinal dataset of continuous glucose measurements from patients with type 1 diabetes mellitus.

#### 3.1 *T1DiabetesGranada* Dataset

*T1DiabetesGranada*: a longitudinal multi-modal dataset of type 1 diabetes mellitus [15] is a public dataset which comprises continuous blood glucose levels, demographic and clinical information of 736 patients with type 1 diabetes mellitus. The dataset contains over four years of data collected from patients at the Clinical Unit of Endocrinology and Nutrition of the San Cecilio University Hospital of Granada, Spain. Blood glucose levels are measured every 15 min using *FreeStyle Libre 2*, a flash glucose meter manufactured by *Abbott Diabetes Care, Inc.* The dataset provides more than 22.6 million records that constitute the time series of continuous blood glucose level measurements of the patients during the duration of the study.

### 3.2 Data Analysis and Preparation

An exploratory analysis of the *T1DiabetesGranada* dataset has been performed and it has been decided to only use the continuous blood glucose level measurements of the patients to train the prediction models. This data has been processed by eliminating possible outliers. The blood glucose level measurements outside the range from 40 to 400 mg/dl have been removed as done in previous works like [18]. Due to the functioning of the flash glucose meter and the interaction of the patient, the time series of blood glucose levels can contain measurements in intervals of less than 15 min. This happens because each time a patient scans the device, the current blood glucose level is measured and it is added as an extra measurement point to the time series. Therefore, the data is processed by eliminating the smaller of the two intervals in the time series, thus obtaining an interval that is closer to 15 min. Furthermore, the time series might also contain data gaps without blood glucose level measurements. This situation occurs in two situations. First, if the patient does not scan the device in less than 8 h, which is the maximum storage time, and the flash glucose meter overwrites the previous measurements. Second, if the patient, does not activate the replacement device early enough after its 14-days life span. To solve this problem, the data is interpolated using the cubic spline method which provides smooth and continuous data, characteristics of blood glucose levels, and generates values adjusted to different data forms. For each patient, the longest sequence of continuous blood glucose level measurements is selected. In order to do so, a tolerance window of 90 min is defined, which is the maximum time allowed in the sequence without data and represents a gap of up to six missing measurements. Cubic spline interpolation is used to obtain the complete time series over the window. The five patients with the longest data sequences are used in this work and the information about their data is presented in Table 1.

### 3.3 Training, Validation and Test

After the exploratory analysis and data preparation, the data has been separated into the training, validation and test sets, with a split of 70%, 20% and 10%. It is not possible to perform random distributions of the time series, since their temporal correlation must be maintained. Therefore, data windows are implemented to provide the neural network with a set of historical data that can be used to predict future blood glucose levels. The prediction horizon is the time frame within the model is expected to make accurate predictions when trained on a data history of a given size. Since blood glucose levels can vary significantly in a short time due to diet, physical activity and other factors, state-of-the-art prediction horizons of 30 and 60 min are commonly used when the models are trained on a history of 120 min. Considering that the *T1DiabetesGranada* dataset used in this work provides blood glucose level measurements every 15 min, whereas the OhioT1DM dataset provides them every 5 min, it might be necessary to increase the prediction horizon to obtain more accurate predictions. Therefore,

**Table 1.** Information about the blood glucose level measurements of the patients used to train the prediction models.

	LIB193385	LIB193327	LIB193367	LIB193313	LIB193316
Start date	22/08/2021 23:41:00	13/01/2021 2:28:00	05/04/2021 22:49:00	27/06/2021 22:02:00	27/01/2022 9:47:00
End date	08/11/2021 22:22:00	27/03/2021 1:52:00	17/05/2021 22:31:00	08/08/2021 2:39:00	10/03/2022 9:47:00
Records	746	6989	4025	4008	5059
Interpolated records	34	45	16	18	514
Interpolation percentage	0.46%	0.64%	0.40%	0.45%	10.16%
Mean	128.81	179.35	148.67	132.27	104.58
Deviation	29.90	60.35	46.31	44.91	36.00
Min	52	53	53	52	51
25%	108	135	113	101	76
50%	126	174	141	124	99
75%	146	217	180	156	124
Max	259	400	314	337	268

the prediction models have been trained in four different scenarios: (1) prediction horizon of 30 min with a history of 120 min; (2) prediction horizon of 60 min with a history of 120 min; (3) prediction horizon of 90 min with a history of 360 min; and (4) prediction horizon of 180 min with a history of 360 min. For each scenario, the Mean Absolute Error (MAE) of the trained models has been calculated. In the prediction of blood glucose levels, the MAE is preferred to the Mean Square Error (MSE) and the Root Mean Squared Error (RMSE) because it is considered more robust as it gives less weight to outliers. Although the MAE measures the error between the predictions and the actual values, it does not take into account the clinical context in which the model is used. Therefore, the Clarke Error Grid analysis [4] has been used to represent the expected and estimated values of blood glucose levels and quantify their clinical accuracy. The grid is divided into five zones: zone A represents values clinically accurate thus leading to correct treatments, zone B those leading to a benign or no treatment, zone C to unnecessary treatment, zone D to a failure to detect and treat, and zone E to an erroneous treatment.

### 3.4 Neural Network Architectures

Three recurrent network architectures based on models presented in the literature have been implemented using Tensorflow. The first model is a 128-unit LSTM recurrent neural network (see Fig. 1a). After the LSTM layer, there are four dense layers with 150, 100, 50 and 20 units and connected to the previous and next layers. Before the second and fourth dense layers, there is a dropout layer with a rate of 0.20 and 0.15, respectively, used to reduce overfitting. ReLu

activation function is used in all the layers and the output layer has a single neuron, which returns the predicted value of the blood glucose level. The second model, the BiLSTM network (see Fig. 1b) is a variant of the LSTM network, replacing the recurrent network layer with a 128-unit BiLSTM but leaving the rest of the network unchanged. The third model, the ConvLSTM network (see Fig. 1c) consists of a convolutional layer with 32 filters of kernel size 1. The result of this layer is connected to the original 128-unit LSTM network architecture, with a slight variation in the last dense layer, which has only 16 neurons instead of 20. All three models have been implemented using the same settings. Adam has been used as optimizer and the loss function has been calculated in MSE. The models have been trained for 100 epochs with a batch size of 32, and early stopping is included in some runs to avoid overtraining the model.

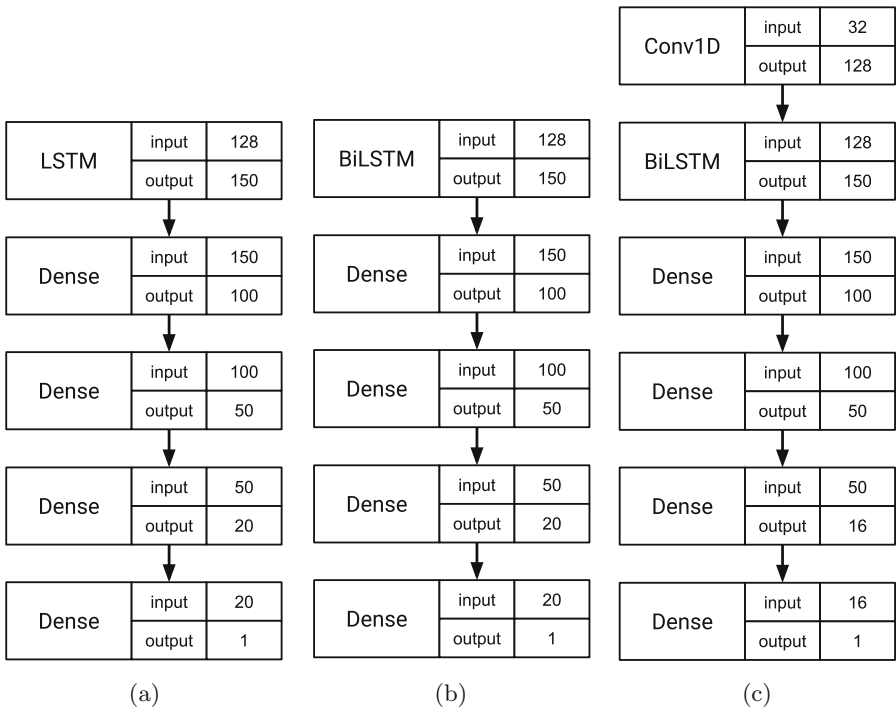


Fig. 1. Neural network architectures: (a) LSTM. (b) BiLSTM. (c) ConvLSTM.

## 4 Results and Discussion

The three neural network models trained for each patient under the prediction horizons of 30, 60, 90 and 180 min, with and without early stopping, have been evaluated. The MAE of the prediction models are shown in Table 2. The models yielding best results for each scenario have been highlighted. The performance of

the models deteriorates as the prediction horizon increases. This is expected since the further out the predicted value is in time, the more complicated is to predict it and the less accurate the prediction will be. The prediction performance of the models varies depending on whether or not early stopping is used during the training phase. With early stopping, the training has been completed in a few epochs, in most cases after six complete training cycles, and in the case of BiLSTM in as few as three training cycles. Without early stopping, the models are trained up to 100 epochs, which can lead to overfitting. For prediction horizons of 30 and 60 min, the ConvLSTM model provides the best results when using early stopping. Without early stopping, the LSTM obtains the best results for three of the patients. For the other two patients, the ConvLSTM performs best for the prediction horizon of 30 min and the BiLSTM for the prediction horizon of 60 min. In view of the results, the performance of the models may vary from one patient to another, and therefore each patient could have a different optimal model.

**Table 2.** MAE (mg/dl) of the prediction models trained for each patient under different prediction horizons with and without early stopping.

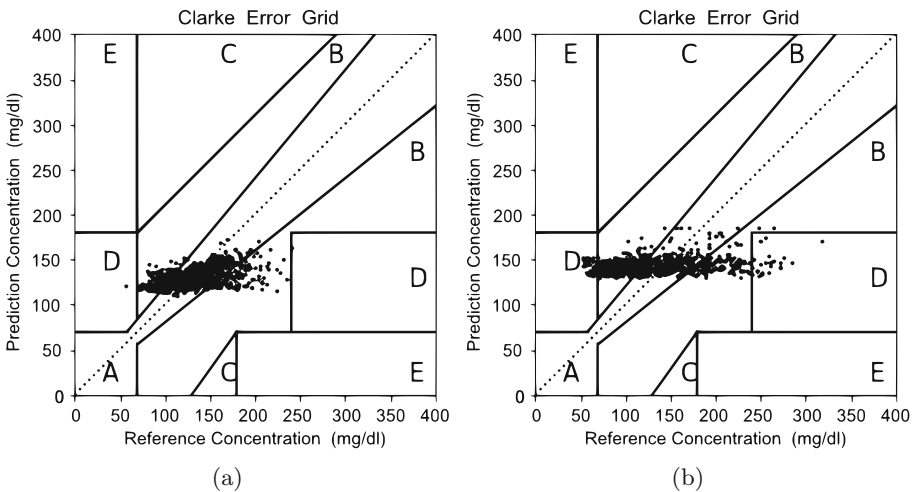
Prediction horizon	Model	LIB193385	LIB193327	LIB193367	LIB193313	LIB193316
With early stopping						
30 min	LSTM	20.67	26.22	25.27	26.74	19.80
	BiLSTM	18.08	27.43	30.44	18.48	15.60
	ConvLSTM	<b>14.83</b>	<b>22.10</b>	<b>24.69</b>	<b>14.04</b>	<b>11.24</b>
60 min	LSTM	27.31	31.21	35.22	30.01	25.29
	BiLSTM	21.04	30.64	35.74	22.61	19.43
	ConvLSTM	<b>20.67</b>	<b>26.36</b>	<b>32.58</b>	<b>20.72</b>	<b>17.11</b>
90 min	LSTM	28.31	38.30	44.11	35.66	24.88
	BiLSTM	<b>22.98</b>	33.30	45.39	26.66	22.86
	ConvLSTM	23.04	<b>31.98</b>	<b>41.34</b>	<b>24.67</b>	<b>19.81</b>
180 min	LSTM	31.89	43.62	49.47	40.07	30.84
	BiLSTM	<b>25.58</b>	<b>39.41</b>	50.79	<b>30.78</b>	31.57
	ConvLSTM	26.01	40.17	<b>49.30</b>	31.32	<b>29.25</b>
Without early stopping						
30 min	LSTM	<b>16.46</b>	<b>21.23</b>	<b>20.33</b>	20.09	16.80
	BiLSTM	25.14	30.25	26.47	16.94	12.41
	ConvLSTM	22.32	30.40	21.35	<b>13.79</b>	<b>11.40</b>
60 min	LSTM	<b>15.62</b>	<b>26.06</b>	<b>27.64</b>	24.61	23.45
	BiLSTM	24.48	34.33	28.50	<b>22.41</b>	<b>17.01</b>
	ConvLSTM	22.56	32.64	29.10	22.65	17.44
90 min	LSTM	<b>17.92</b>	<b>31.09</b>	<b>31.58</b>	33.25	26.50
	BiLSTM	19.50	37.79	35.62	<b>27.90</b>	<b>20.70</b>
	ConvLSTM	18.29	39.78	33.61	31.62	23.52
180 min	LSTM	21.24	<b>41.71</b>	40.73	41.43	31.95
	BiLSTM	<b>20.26</b>	43.61	<b>37.51</b>	<b>34.83</b>	<b>27.59</b>
	ConvLSTM	20.85	48.13	38.17	38.94	29.56



To evaluate the performance of the prediction models from a clinical perspective a Clarke Error Grid analysis is performed. Table 3 reports on the percentage of predictions falling in the zones A and B, which lead to clinically correct treatments and those leading to a benign treatment respectively. All percentages are above 94% irrespective of the model, prediction horizon and patient. The best result (99.99%) is obtained for the patient LIB193327 when training the LSTM network under a prediction horizon of 60 min. The worst result (94.02%) is achieved for the patient LIB193313 when training the BiLSTM under a prediction horizon of 180 min. The Clarke Error Grid analysis for these two cases are shown in Fig. 2. Clearly, most of the predictions fall in zones A and B, thus confirming the clinical validity of the developed models even for the worst ones.

**Table 3.** Percentage of predictions falling in Clarke Error Grid zones A and B.

Prediction horizon	Model	LIB193385	LIB193327	LIB193367	LIB193313	LIB193316
30 min	LSTM	99,95%	99,89%	99,78%	98,96%	95,35%
	BiLSTM	99,89%	99,98%	99,89%	98,01%	97,70%
	ConvLSTM	99,90%	99,95%	99,90%	98,80%	97,79%
60 min	LSTM	99,48%	<b>99,99%</b>	99,48%	97,96%	95,36%
	BiLSTM	99,95%	99,90%	99,95%	94,36%	96,65%
	ConvLSTM	99,90%	99,96%	99,90%	96,50%	94,65%
90 min	LSTM	99,95%	99,93%	97,95%	96,35%	94,36%
	BiLSTM	99,95%	99,95%	98,01%	94,65%	95,58%
	ConvLSTM	99,93%	99,89%	98,50%	96,50%	96,66%
180 min	LSTM	99,93%	99,60%	96,54%	95,65%	95,65%
	BiLSTM	99,95%	99,71%	98,88%	<b>94,02%</b>	94,36%
	ConvLSTM	99,98%	99,68%	97,01%	96,58%	95,02%



**Fig. 2.** Clarke Error Grid analysis for the best and worst prediction models: (a) LIB193327 - LSTM - 60' (99.99%). (b) LIB193313 - BiLSTM - 180' (94.02%).

## 5 Conclusions

This work has compared the ability of three neural network models (LSTM, BiLSTM, and ConvLSTM) for predicting blood glucose level measurements in type 1 diabetes patients. The models have been evaluated on four different scenarios with varying prediction horizons (30, 60, 90, and 180 min) and history (120 and 360 min). Few differences are found with respect to the performance of the models, yielding similar prediction errors. Regarding the neural network training strategy, the ConvLSTM stands out as the best model when using early stopping while the LSTM network is found to prevail without early stopping for a majority of patients. According to the experiments, there is no one-fits-all model but rather some models work best for some patients. From a medical point of view, practically all the predictions made by the learned models are in zone A and zone B of Clarke error grid. These results are considered clinically accurate and therefore demonstrate that these models could be used in practice.

**Acknowledgements.** This research has been funded by the Andalusian Ministry of Economic Transformation, Industry, Knowledge and Universities under grant P20-00163.

## References

1. Aiello, E.M., Lisanti, G., Magni, L., Musci, M., Toffanin, C.: Therapy-driven deep glucose forecasting. *Eng. Appl. Artif. Intell.* **87**, 103255 (2020). <https://doi.org/10.1016/j.engappai.2019.103255>
2. Allam, F., Nossai, Z., Gomma, H., Ibrahim, I., Abdelsalam, M.: A recurrent neural network approach for predicting glucose concentration in type-1 diabetic patients. In: Iliadis, L., Jayne, C. (eds.) *AIAI/EANN -2011. IAICT*, vol. 363, pp. 254–259. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-23957-1\\_29](https://doi.org/10.1007/978-3-642-23957-1_29)
3. Bach, K., Bunescu, R., Marling, C., Wiratunga, N.: Preface the 5th international workshop on knowledge discovery in healthcare data (KDH). In: *CEUR Workshop Proceedings*, vol. 2675, pp. 1–4 (2020)
4. Clarke, W.L., Cox, D., Gonder-Frederick, L.A., Carter, W., Pohl, S.L.: Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care* **10**(5), 622–628 (1987). <https://doi.org/10.2337/diacare.10.5.622>
5. Gia, T.N., et al.: IoT-based continuous glucose monitoring system: a feasibility study. *Procedia Comput. Sci.* **109**, 327–334 (2017). <https://doi.org/10.1016/j.procs.2017.05.359>
6. Idriss, T., Idri, A., Abnane, I., Bakkoury, Z.: Predicting blood glucose using an LSTM neural network. In: *Proceedings of the Federated Conference on Computer Science and Information Systems*, vol. 18, pp. 35–41 (2019)
7. International Diabetes Federation: *IDF Diabetes Atlas*. International Diabetes Federation, Brussels, Belgium, 10 edn (2021). <https://www.diabetesatlas.org>
8. Khadem, H., Nemat, H., Elliott, J., Benaissa, M.: Multi-lag stacking for blood glucose level prediction. In: *CEUR Workshop Proceedings*, vol. 2675, pp. 146–150 (2020)
9. Levy, D.: *Type 1 Diabetes*. Oxford University Press, Oxford (2016)

10. Marling, C., Bunescu, R.: The OhioT1DM dataset for blood glucose level prediction: update 2020. In: CEUR Workshop Proceedings, vol. 2675, pp. 71–74 (2020)
11. Mayo, M., Koutny, T.: Neural multi-class classification approach to blood glucose level forecasting with prediction uncertainty visualisation. In: Proceedings of 5th International Workshop on Knowledge Discovery in Healthcare Data (KDH 2020), vol. 2675, pp. 80–84. Santiago de Compostela, Spain (2020)
12. Mirshekarian, S., Bunescu, R., Marling, C., Schwartz, F.: Using LSTMs to learn physiological models of blood glucose behavior. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2887–2891 (2017). <https://doi.org/10.1109/EMBC.2017.8037460>
13. Mirshekarian, S., Shen, H., Bunescu, R., Marling, C.: LSTMs and neural attention models for blood glucose prediction: comparative experiments on real and synthetic data. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 706–712 (2019). <https://doi.org/10.1109/EMBC.2019.8856940>
14. Munoz-Organero, M.: Deep physiological model for blood glucose prediction in T1DM patients. *Sensors* **20**(14), 3896 (2020). <https://doi.org/10.3390/s20143896>
15. Rodriguez-León, C., et al.: T1DiabetesGranada: a longitudinal multi-modal dataset of type 1 diabetes mellitus (2023). <https://osf.io/vd45b/>. Accessed 31 Mar 2023
16. Rubin-Falcone, H., Fox, I., Wiens, J.: Deep residual time-series forecasting: application to blood glucose prediction. In: CEUR Workshop Proceedings, vol. 2675, pp. 105–109 (2020)
17. Sun, Q., Jankovic, M., Bally, L., Mougiakakou, S.: 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 (2018)
18. Tena, F., Garnica, O., Lanchares, J., Hidalgo, J.I.: Ensemble models of cutting-edge deep neural networks for blood glucose prediction in patients with diabetes. *Sensors* **21**(21), 7090 (2021). <https://doi.org/10.3390/s21217090>
19. Tresp, V., Briegel, T., Moody, J.: Neural-network models for the blood glucose metabolism of a diabetic. *IEEE Trans. Neural Networks* **10**(5), 1204–1213 (1999). <https://doi.org/10.1109/72.788659>