

# Predicting Blood Glucose using an LSTM Neural Network

Touria El Idriss  
Department of Computer Sciences  
EMI, University Mohamed V  
Rabat, Morocco  
Email: {el.idrissit@gmail.com}

Ali Idri and Ibtissam Abnane  
Software Project Management Research  
Team ENSIAS, University Mohamed V  
Rabat, Morocco  
Email: {ali.idri@um5.ac.ma,  
ibtissam.abnane19@gmail.com}

Zohra Bakkoury  
Department of Computer  
Sciences EMI, University  
Mohamed V Rabat, Morocco  
Email: {bakkoury@gmail.com}

**Abstract**—Diabetes self-management relies on the blood glucose prediction as it allows taking suitable actions to prevent low or high blood glucose level. In this paper, we propose a deep learning neural network (NN) model for blood glucose prediction. It is a sequential one using a Long-Short-Term Memory (LSTM) layer with two fully connected layers. Several experiments were carried out over data of 10 diabetic patients to decide on the model's parameters in order to identify the best variant of it. The performance of the proposed LSTM NN measured in terms of root mean square error (RMSE) was compared with the ones of an existing LSTM and an autoregressive (AR) models. The results show that our LSTM NN is significantly more accurate; in fact, it outperforms the existing LSTM model for all patients and outperforms the AR model in 9 over 10 patients, besides, the performance differences were assessed by the Wilcoxon statistical test. Furthermore, the mean of the RMSE of our model was 12.38 mg/dl while it was 28.84 mg/dl and 50.69 mg/dl for AR and the existing LSTM respectively.

## I. INTRODUCTION

**D**ATA mining (DM) techniques are useful tools for extracting valuable knowledge from (large) databases that helps in decision making [1], [2]. DM has been fruitfully used in different subfields of medical informatics such as diabetes [1], [3], cardiology [1], [4] and cancer [1], [5].

This paper deals with the application of DM for diabetes which is a chronic illness caused by a disorder in the glucose metabolism. There are mainly two types of diabetes): 1) Type 1 Diabetes Mellitus (T1DM) when the pancreas does not produce enough insulin, and Type 2 Diabetes Mellitus (T2DM) which results from an ineffective use of insulin [6]. If not well managed, this disease can lead to serious problems such as heart attacks, kidney damage, blindness, unconsciousness and even death [6]. The prediction of blood glucose level (BGL) is an important task in the diabetes management and self-management as it can help controlling the BGL by taking appropriate actions ahead of time [7]. To

predict the BGL, the previous glucose measurements are required. The BGL can be measured: 1) Manually by self-monitoring of blood glucose (SMBG) using sticks several times a day or 2) Automatically by continuous glucose monitoring (CGM) using sensors [6], [7].

According to El Idrissi et al. [7], considerable work was done for the BGL prediction and various Data Mining approaches including statistical methods and machine learning techniques were investigated for that purpose; the most used ones are Artificial Neural Networks (NNs) and Auto Regression (AR) [7]. Recently, deep learning modeling is gaining more interest, such as LSTM NN [8] and deep NN [9].

This paper proposes a deep learning NN with one LSTM layer and two fully connected layers for the prediction of BGL using CGM data. Predicting glucose using LSTM Nns is promising [8] since LSTM NNs were successfully applied in other domains such as prediction of water quality [10], electricity consumption [11] and stock prices [12].

This work aims at: (1) Setting the parameters of the model to identify the best configuration of our LSTM NN; and (2) Assessing and comparing the accuracy of the proposed model to existing ones. Toward this aim, two research questions are discussed:

(RQ1): Can the proposed LSTM model achieve good performance?

(RQ2): Is the proposed LSTM model significantly more accurate than existing models?

This paper is structured as follows: Section II presents an overview of LSTM NNs. Section III summarizes the related work on predicting blood glucose. Data used and performance measurement are described in Section IV. Section V describes the experimental design. Results are reported and discussed in Section VI. Threats to validity of this study are presented in Section VII and finally conclusion and future work are presented in Section VIII.

## II. LSTM NNS: AN OVERVIEW

Deep learning in NNs is an emerging method that allows the NN to learn automatically the characteristics of data by selecting the relevant features [10], contrary to the classical NNs that require features' selection based on domain knowledge [9].

LSTM NNs are deep recurrent NNs (RNNs) that were introduced by Hochreiter and Schmidhuber [13] to overcome the problem of exploding or vanishing gradient encountered with traditional RNNs [8]. The LSTM NNs are suitable for sequential data such as speech, video and time series as they can capture long term dependencies [14]. They consist of memory cells with a cell state which is maintained over time, and a gate structure that controls and regulates the information of the cell state.

Fig. 1 illustrates the structure of a memory cell. The index  $t$  refers to time or sequence.  $X_t$ ,  $Y_t$ ,  $h_t$  and  $C_t$  represent respectively the input, the output, the hidden vector and the cell state for  $t$ .

The memory cell contains 3 gates: 1) Input gate selects the information to be retained in the cell; 2) Forget gate decides about the information to be ignored; and 3) Output gate calculates the output and updates the hidden vector. Each of these gates is a NN whose input vector is a concatenation of the hidden vector of the previous cell and the input vector. Let  $W_i$ ,  $W_f$ ,  $W_o$  be the weight matrices corresponding respectively to the input, forget and output gates; and  $b_i$ ,  $b_f$ ,  $b_o$  the corresponding bias vector.  $W_c$  and  $b_c$  are the weight matrix and the bias vector used for updating the cell state.

The result of the input gate is  $i_t$ , which is obtained as follows:

$$i_t = \sigma(W_i * [h_{t-1}, X_t] + b_i) \quad (1)$$

To calculate  $f_t$ , the forget gate uses the following equation:

$$f_t = \sigma(W_f * [h_{t-1}, X_t] + b_f) \quad (2)$$

The output gate uses the equation (3) to obtain  $o_t$ , and the equation (4) to get the hidden vector.

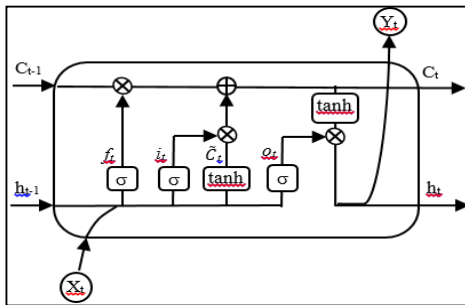


Fig. 1 : Memory cell

$$o_t = \sigma(W_o * [h_{t-1}, X_t] + b_o) \quad (3)$$

$$h_t = o_t * \tanh(C_t) \quad (4)$$

The cell state is updated as follows:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

$$\text{where } \tilde{C}_t = \tanh(W_c * [h_{t-1}, X_t] + b_c) \quad (6)$$

In the Equations (1 to 4 and 6),  $\sigma$  and  $\tanh$  are the activation functions, the former is the sigmoid function defined in Equation (7) and the latter is the hyperbolic tangent function defined in equation (8).

$$\sigma(x) = \frac{e^x}{1+e^x} \quad (7)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (8)$$

$\sigma$  and  $\tanh$  are defined on the set of real numbers.  $\sigma$  ranges between 0 and 1 while  $\tanh$  ranges from -1 to 1.

## III. RELATED WORK

El Idrissi et al. [7] investigated by the mean of a systematic map and review the use of predictive techniques in data mining for the diabetes self-management. The study summarized and reviewed 38 studies published from 2000 to April 2017 with regards to: 1) publication's year and source, 2) diabetes' type, 3) investigated clinical tasks, 4) the used predictive techniques and 5) their performance.

The main findings of [7] are the following:

1. There is a growing interest to the use of DM predictive techniques in the last decade. Conferences and Journals are used for publication, nevertheless, the main publication channel is conferences by 76.32% of the selected studies, while just 23.68% were published in journals.

2. T1DM gained more attention for research than T2DM with 84.21% of the selected studies.

3. Considerable work was done for the BGL prediction comparatively to other clinical tasks, in fact, 57.89% of the considered papers investigated BGL prediction.

4. Various DM predictive techniques were investigated, and the most used ones are NNs and AR.

5. NNs and AR models yield the highest accuracies. However, none of the used DM predictive techniques is dominant over the others.

Table I reports the findings of a set of selected studies from the review of El Idrissi et al. [7], in addition to the two recent studies [8], [9]. From Table I, we note that:

1. Various techniques are investigated to the BGL prediction problem: statistical methods such as AR [15], [16] and Kalman Filter (KF) [17]; machine learning methods such as Artificial NNs [18]-[20] and Support Vector Regression (SVR) [21], [22]; and recently deep learning techniques [8], [9].

TABLE I.  
LITERATURE OVERVIEW OF THE BGL PREDICTION

Reference	Technique	Diabetes type	Input Data	Findings
Lu et al., 2010 [15]	AR	T1DM	CGM	AR models yield accurate BGL prediction with short length signals; it is not required to consider exogenous inputs explicitly nor all frequency bands of the glucose signals.
Novara et al., 2016 [16]	AR	T1DM	CGM	A blind identification using AR technique was proposed to predict the BGL and recover the unmeasured inputs.
Wang et al., 2013 [17]	KF	T1DM	CGM	The model based on an extended KF achieves mostly reliable BGL predictions, and made significant improvement compared to zero-order hold.
Zarkogianni et al. 2011 [18]	RNNs	T1DM	CGM	The proposed RNN model to simulate the blood glucose–insulin metabolism makes it possible to personalize the system and to handle the environment’s variations with efficiency.
Allam et al., 2011 [19]	Feedforward NNs	T1DM	CGM	For short prediction horizon, feedforward NN based model gets accurate BGL prediction without time lagging.
Mathiyazhagan & Schechter, 2014 [20]	Fuzzy NNs	T1DM	CGM	This study proposes a soft computing approach which tolerates imprecision by using a fuzzy NN to predict the BGL.
Bunescu et al., 2013 [21]	SVR	T1DM	CGM	The incorporation of physiological features into an existing SVR model for BGL prediction made a significant performance enhancement .
Georga et al., 2010 [22]	SVR	T1DM	CGM	Predicting BGL is possible by compartmental models and SVR with a satisfactory accuracy and a clinical acceptability.
Sun et al, 2018 [8]	LSTM NN	T1DM	CGM	An LSTM network with one LSTM layer followed by one bi-directional LSTM layer and several fully connected layers were proposed for BGL prediction. This LSTM model outperformed the baseline methods ARIMA and SVR.
Mhaskar et al, 2017 [9]	Deep NN	T1DM	CGM	The proposed deep NN for BGL prediction outperforms shallow NN.

2. Different types of Artificial NNs were explored: RNNs [18], feedforward NNs [19], fuzzy NNs [20], and deep NNs [8], [9].

3. Studies are using data coming from CGM for T1DM patients.

#### IV. DATA DESCRIPTION AND PERFORMANCE MEASUREMENT

The data set used in this study is the historical data set DirecNetInpatientAccuracyStudy provided by Diabetes Research in Children Network (DirecNet) [23]. It holds collected data of 110 T1DM patients particularly the recorded data from CGM devices which give the BGLs at intervals of 5 minutes. As we did not specify any prerequisite for our model, we considered randomly a subset of 10 patients. However, a pre-processing of the data was done to eliminate redundancy and outliers between consecutive measurements. Table II summarizes information about the considered patients.

To evaluate the performance of our model, we use the root mean square error (RMSE) which is a common performance metric used to assess BGL prediction [7]. It evaluates the difference between the actual value and the predicted one by means of the Equation (9):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{e}_i - e_i)^2} \quad (9)$$

where  $e_i$  and  $\hat{e}_i$  are respectively the actual and the predicted values while  $n$  is the sample’s size. The value of RMSE goes

from 0 to  $+\infty$  and the prediction is better when the RMSE is low.

#### V. EXPERIMENTAL DESIGN

This paper proposes an LSTM NN to predict the BGL of diabetic patients. This NN uses one LSTM layer and two dense layers. To identify the best configuration to adopt, we performed by means of a Grid Search (GS) described in Table III and inspired from [24], the following steps:

Step 1: Train and test the model for each patient by varying the number of LSTM units (LU) according to the GS values of Table III. The chosen value is the one that gives the best RMSE over the patients’ datasets.

TABLE II.  
DATA SET DESCRIPTION. THE BGL IS IN MG/DL

Patient ID	Number of measurements	Min BGL	Max BGL	Mean BGL
PT01	766	40	339	114.78
PT02	278	57	283	120.96
PT03	283	103	322	185.89
PT04	923	40	400	188.44
PT05	562	50	270	179.71
PT06	771	62	400	187.45
PT07	897	42	400	210.26
PT08	546	43	310	152.88
PT09	831	40	400	157.50
PT10	246	72	189	116.51

Step 2: Using the LU obtained in Step 1, we train and evaluate the model for each patient by varying the number of dense units (DU) respecting the GS values of Table III. The chosen value is the one that gives the best RMSE over the patients' datasets.

Step 3: Using the LU and DU obtained in Step 1 and Step 2 respectively, we train and evaluate the model for each patient by varying the length of the input sequences (SL) according to the GS values of Table III. The chosen value is the one that gives the best RMSE over the patients' datasets.

Step 4: We compare in terms of the RMSE criterion our best LSTM model with the LSTM of [8] and an AR model.

Step 5: We evaluate the statistical significance of performance differences between the three models by means of the Wilcoxon statistical test [25].

## VI. RESULTS AND DISCUSSION

This section presents and discusses the obtained results: (1) We present the empirical results of the Steps 1 to 3 of the experimental process related to the parameters' setting. (2) The results of the performance comparison of our model to the two other models are presented. (3) We report the results of the statistical test. And (4), We discuss all the empirical results.

Our LSTM model is developed using Python 3.6 with the framework Keras 2.2.4 and Tensorflow 1.12.0 as backend and under the operating system Windows 10.

### A. Parameters selection

The main objective of this research is to propose an LSTM model that achieves good performance (RQ1). To do that, a phase of parameters selection was carried out. This phase consists of the three first steps of the experimental design of Section V. The results of the first step are shown in Fig. 2 which represents the RMSE for each patient according to the search space of LU defined in Table III. We observe that the RMSE is in general better for LU=50. Therefore, at that step the LU was chosen to be 50 units.

TABLE III.  
PARAMETERS FOR SEARCH GRID

Parameter	Signification	Search space
LSTM Units	Number of neurons in the gates	{5, 10, 20, 30, 40, 50, 60, 70}
Dense Unit	Number of the neurons in the dense layer	{10, 20, 30, 40, 50}
Sequence length	Dimension of the input vector	{5, 10, 15, 20}

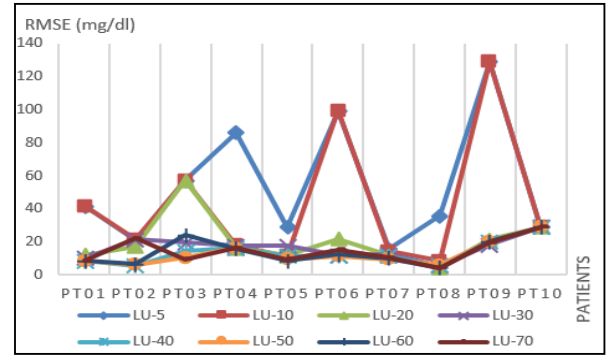


Fig. 2 : RMSE with LU variations

In step 2, we did the same experimentations with LU=50 and we vary DU according to the GS values of Table III. Fig. 3 shows the results obtained: we observe that for DU=40 and DU=50, the RMSE is worse than the others. Furthermore, the values 10, 20 and 30 of DU gave comparable results. However, by considering the times that a configuration (DU is equal to 30, 20 or 10) is better than the others, the minimum RMSE was reached 4 times, 3 times and once respectively. Therefore, the chosen DU value is 30.

The Step 3 uses LU=50, DU=30, and varies the SL according to the GS values of Table III. Fig. 4 shows that SL=10 achieves in general better RMSE values.

To sum up, the best configuration of our LSTM uses LU=50, DU=30 and SL=10. This LSTM variant will be used to compare it with existing BGL predictors.

### B. Models Comparison

To answer RQ2, we applied the LSTM model of [8] referred to as Sun\_LSTM and AR on the same dataset. We carried out the comparison with Sun\_LSTM since it is, according to the best of our knowledge, the only study that proposed an LSTM model for BGL prediction. Regarding the comparison with an AR model, it is motivated by the fact that AR models are suitable for time series prediction and according to [7] they yield along with NNs highest accuracy rates.

The Sun\_LSTM contains one LSTM layer with 4 LU followed by one bi-directional LSTM layer with 4 LU and 3 fully connected dense layers with respectively 8, 64 and 8 DU and finally the output Dense Layer. The SL is set to 4. Regarding the AR model, we used the default model developed in Weka with a lag of 10 similarly to the SL proposed in our model which is equal to 10.

Fig. 5 presents the RMSE got for the three models: we observe that our LSTM model outperforms Sun\_LSTM for all patients and outperforms AR in 9 cases over 10 patients.

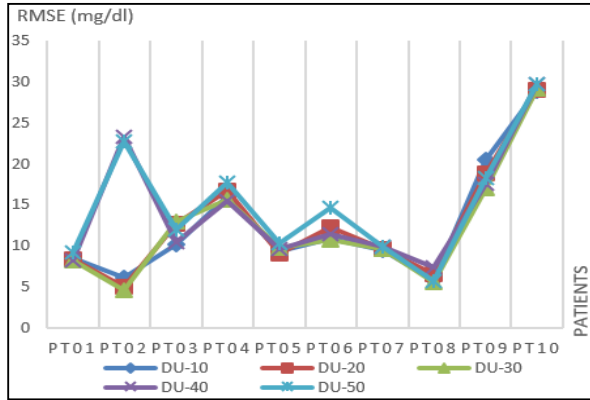


Fig. 3: RMSE with DU variations

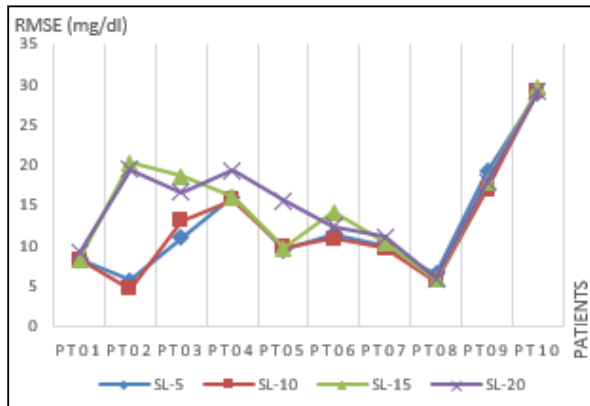


Fig. 4: RMSE with SL variations

Furthermore, we have compared the training time for the two LSTM models presented in Fig. 6. Our model needs overall less time to be trained comparatively to Sun\_LSTM.

Finally, we recorded the number of epochs required for each patient to train the model. Results are presented in Table IV.

The number of epochs is low and ranges between 46 and 260, the mean is equal to 132.8.

TABLE IV.  
NUMBER OF EPOCHS

Patient ID	Number of epochs	Patient ID	Number of epochs
PT01	91	PT06	260
PT02	94	PT07	46
PT03	259	PT08	49
PT04	71	PT09	168
PT05	88	PT10	81

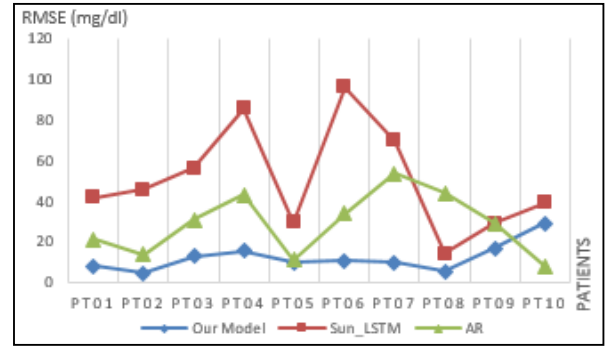


Fig. 5: RMSE for Our model, Sun\_LSTM and AR

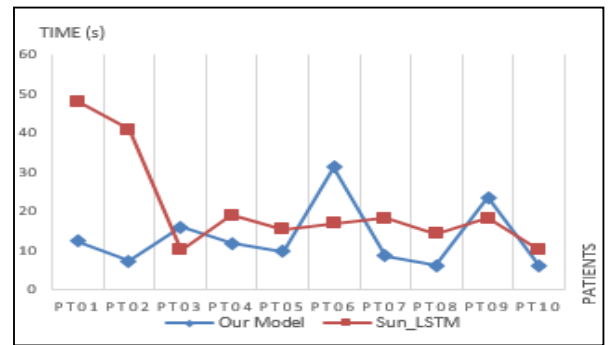


Fig. 6: Training time for Our model and Sun\_LSTM

C. Statistical test

We statistically assessed the results obtained based on the Wilcoxon statistical test. It is a non-parametric test that validates if the differences between the compared models are statistically significant by the means of a statistical hypothesis [25]. In our case, two null hypotheses (NH) were considered to test if our LSTM model is better than Sun\_LSTM and AR. These hypotheses are the following:

- NH1 : Our LSTM model does not outperform Sun\_LSTM.
- NH2 : Our LSTM model does not outperform AR model.

The statistical test was two-tailed considering the significance level  $\alpha$  equals to 0.05. Note that the test is considered significant if the result which is the p-value is less than  $\alpha$ .

For our tests, the p-value of NH1 and NH2 were 0.00512 and 0.0285 which means that the performance differences between our models and respectively the two other models were statistically significant.

D. Discussion

In this study, we propose an LSTM NN to predict BGL. Our first concern was to determine the best configuration by tuning the three following parameters: LU, DU and SL. As we observe from the results of Fig. 2, Fig. 3 and Fig. 4, these parameters have an impact on the model's performance. Thus, when building an LSTM model, one has to determine

the optimal values of these parameters. The final model achieved good accuracy comparatively to what was found in literature by [7]. In fact, for the RMSE, the minimum, maximum and mean values were equal to 4.67, 29.12 and 12.38 respectively. On the other hand, our model outperforms a previous LSTM [8] and AR models. Furthermore, the performance of our model over the two others was assessed by the Wilcoxon statistical test which showed that our model significantly outperformed Sun\_LSTM and AR.

These results show that the use of LSTM NNs for BGL prediction is promising. In fact, LSTM NNs and Deep NNs in general are gaining more interest due to their performances in different fields [26]. Actually, the deep NNs took advantage from their training algorithms that are computationally efficient [27], and from the large hidden neurons' number which results on huge number of free parameters [26]. On another hand, the performance of our model over Sun\_LSTM can be explained by the fact that our LSTM is deeper than Sun\_LSTM: in fact, the former uses 50 LU while the latter uses 4 LU.

However, we believe that the accuracy can be improved by considering: 1) other parameters to tune such as learning rate, the loss function and the optimizer, and 2) more information as input such as medication and exercise.

From another point of view, the training time and the epoch numbers were low, which means that the model can be used on a mobile platform or wearable device [7] which require rapidity [8]. In fact, knowing the importance of the mobile personal health records in health management [28], [29], the deployment of our model on a mobile platform is another direction to explore. Toward that aim, we used TensorFlow as it provides an end-to-end support which goes till the deployment of the models on mobile apps, cloud server and others [30].

#### VII. THREADS TO VALIDITY

Four of threads to validity are identified which are:

**Internal validity:** This thread concerns the evaluation methodology which can be inappropriate. To overcome this limitation, we used 10 data sets, each of them was divided in training and test sets representing respectively 66% and 34% of the whole data set. The model was trained on the former and evaluated on the latter.

**External validity:** This thread is related to the perimeter of validity. This was settled by considering 10 diabetic patients randomly taken from a public dataset. Note that some studies used only one dataset as reported in [7]. Furthermore, the datasets are with different sizes, it goes from 246 to 923 instances.

**Construct validity:** It concerns the measurement validity. To limit this thread, we used the RMSE which is a common performance criterion used to assess BGL prediction [7].

Moreover, it was used by many studies such as [8], [15], [18], [19], [21], [22]. Note that it was used lonely by [15], [21].

**Statistical conclusion validity:** It affects the conclusion related to the comparison performed. To avoid this thread, we used the Wilcoxon statistical test to assess the significance of the performance differences.

#### VIII. CONCLUSION AND FUTURE WORK

This paper considered a deep NN for BGL prediction which is a sequential model with one LSTM layer and two fully connected layers. Multiple runs were done by varying the three parameters: LU, DU and SL to determine the best configuration. The model achieves good accuracy and significantly outperforms, based on RMSE, an existing LSTM model and an AR model.

These promising results encourage to carry out further investigations using the LSTM NNs. Ongoing work aims at building an LSTM model for multi-step prediction by exploring different strategies from literature. Considering more input data such as medication and exercise may improve the accuracy. Furthermore, a problem that can be encountered is the missing data when for example the patient takes off the CGM device, investigating how to handle automatically the missing data is an interesting direction to consider.

#### REFERENCES

- [1] N. Esfandiari, M. R. Babavalian, A. M. E. Moghadam & V. K. Tabar, "Knowledge discovery in medicine: Current issue and future trend", *In Expert Systems with Applications*, vol. 41, no. 9, pp. 4434-4463, 2014, <https://doi.org/10.1016/j.eswa.2014.01.011>
- [2] H. Benhar, A. Idri and J.-L. Fernández-Alemán, "Data preprocessing for decision making in medical informatics: potential and analysis." *World Conference on Information Systems and Technologies*. Springer, Cham, 2018, pp. 1208-1218, [https://doi.org/10.1007/978-3-319-77712-2\\_116](https://doi.org/10.1007/978-3-319-77712-2_116).
- [3] T. El Idrissi, A. Idri, and Z. Bakkoury, "Data Mining Techniques in Diabetes Self-management: A Systematic Map", *In World Conference on Information Systems and Technologies*, Springer, Cham, 2018, pp. 1142-1152, [https://doi.org/10.1007/978-3-319-77712-2\\_109](https://doi.org/10.1007/978-3-319-77712-2_109).
- [4] I. Kadi, A. Idri and J.-L. Fernandez-Aleman. "Knowledge discovery in cardiology: A systematic literature review". *International Journal of Medical Informatics*, vol. 97, pp. 12-32, 2017, <https://doi.org/10.1016/j.ijmedinf.2016.09.005>.
- [5] A. Idri, I. Chlioui and B. EL Ouassif, "A systematic map of data analytics in breast cancer", *In Proceedings of the Australasian Computer Science Week Multiconference*, ACM, 2018, p. 26, <https://doi.org/10.1145/3167918.3167930>.
- [6] R. Billous, R. Donnally, "Handbook of Diabetes", Blackwell, 2010.
- [7] T. EL Idrissi, A. Idri and Z. Bakkoury, "Systematic map and review of predictive techniques in diabetes self-management", *International Journal of Information Management*, vol. 46, pp. 263-277, 2019, <https://doi.org/10.1016/j.ijinfomgt.2018.09.011>.
- [8] Q. Sun, M. V. Jankovic, L. Bally and S. G. Mougiakakou, "Predicting Blood Glucose with an LSTM and Bi-LSTM Based Deep Neural Network," *In Symposium on Neural Networks and Applications (NEUREL)*, Belgrade, 2018, pp. 1-5, <https://doi.org/10.1109/NEUREL.2018.8586990>
- [9] H. N. Mhaskar, S. V. Pereverzyev, and M. D. van der Walt, "A Deep Learning Approach to Diabetic Blood Glucose Prediction," *Front. Appl. Math. Stat.*, vol. 3, no. July, pp. 1-11, 2017, <https://doi.org/10.3389/fams.2017.00014>

- [10] Y. Wang, J. Zhou, K. Chen, Y. Wang and L. Liu, "Water quality prediction method based on LSTM neural network," 2017 12<sup>th</sup> International Conference on Intelligent Systems and Knowledge Engineering (ISKE), Nanjing, 2017, pp. 1-5, <https://doi.org/10.1109/ISKE.2017.8258814>
- [11] N. Kim, M. Kim and J. K. Choi, "LSTM Based Short-term Electricity Consumption Forecast with Daily Load Profile Sequences," 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE), Nara, 2018, pp. 136-137, <https://doi.org/10.1109/GCCE.2018.8574484>
- [12] M. Roondiwala, H. Patel, and S. Varma, "Predicting Stock Prices Using LSTM," International Journal of Science and Research (IJSR), vol. 6, no. 4, pp.1754-1756, 2017.
- [13] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput., vol. 9, no. 8, pp. 1735-1780, 1997, <https://doi.org/10.1162/neco.1997.9.8.1735>
- [14] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink and J. Schmidhuber, "LSTM: A Search Space Odyssey," in IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 10, pp. 2222-2232, Oct. 2017, <https://doi.org/10.1109/TNNLS.2016.2582924>.
- [15] Y. Lu, A. V. Gribok, W. K. Ward, and J. Reifman.. "The Importance of Different Frequency Bands in Predicting Subcutaneous Glucose Concentration in Type 1 Diabetic Patients," in IEEE Transactions on Biomedical Engineering, vol. 57, no. 8, pp. 1839-1846, Aug. 2010, <https://doi.org/10.1109/TBME.2010.2047504>
- [16] C. Novara, N. M. Pour, T. Vincent, G. Grassi, "A Nonlinear Blind Identification Approach to Modeling of Diabetic Patients," in IEEE Transactions on Control Systems Technology, vol. 24, no. 3, pp. 1092-1100, May 2016, <https://doi.org/10.1109/TCST.2015.2462734>
- [17] Q. Wang S. Harsh, P. Molenaar, and K. Freeman, "Developing personalized empirical models for Type-I diabetes: An extended Kalman filter approach," American Control Conference, IEEE, 2013, pp. 2923-2928, <https://doi.org/10.1109/ACC.2013.6580278>.
- [18] K. Zarkogianni, A.Vazeou, S. G. Mougiakakou, A. Prountzou and K. S. Nikita, "An Insulin Infusion Advisory System Based on Autotuning Nonlinear Model-Predictive Control," in IEEE Transactions on Biomedical Engineering, vol. 58, no. 9, pp. 2467-2477, Sept. 2011, <https://doi.org/10.1109/TBME.2011.2157823>.
- [19] F. Allam, Z. Nossair, H. Gomma, I. Ibrahim, and M. Abd-el Salam. "Prediction of subcutaneous glucose concentration for type-1 diabetic patients using a feed forward neural network," The Int. Conf. On Computer Engineering & Systems, Cairo, 2011, pp. 129-133, <https://doi.org/10.1109/CCES.2011.6141026>.
- [20] N. Mathiyazhagan, H. B. Schechter, "Soft computing approach for predictive blood glucose management using a fuzzy neural network," IEEE Conf. on Norbert Wiener in the 21st Century (21CW), Boston, MA, 2014, pp. 1-3, <https://doi.org/10.1109/NORBERT.2014.6893906>.
- [21] R. Bunescu, N. Struble, C. Marling, J. Shubrook, and F. Schwartz, "Blood Glucose Level Prediction Using Physiological Models and Support Vector Regression," In 2013 12th International Conference on Machine Learning and Applications, vol. 1, pp. 135-140. IEEE, 2013, <https://doi.org/10.1109/ICMLA.2013.30>.
- [22] E. I. Georga, V. C. Protopappas, D. Polyzos, "Prediction of glucose concentration in type 1 diabetic patients using support vector regression," Proceedings of the 10th IEEE Int. Conf. on Information Technology and Applications in Biomedicine, Corfu, 2010, pp. 1-4, <https://doi.org/10.1109/ITAB.2010.5687764>
- [23] Diabetes Research in Children Network (DirecNet). Available online at: <http://direcnet.jaeb.org/Studies.aspx> [Ap. 1,2019]
- [24] M. Hosni, A. Idri and A. Abran, "Investigating heterogeneous ensembles with filter feature selection for software effort estimation." In Proceedings of the 27th International Workshop on Software Measurement and 12th International Conference on Software Process and Product Measurement, ACM, 2017, pp. 207-220, <https://doi.org/10.1145/3143434.3143456>
- [25] A. Idri, I. Abnane and A. Abran. "Missing data techniques in analogy-based software development effort estimation." Journal of Systems and Software, vol. 117, pp. 595-611, 2016, <https://doi.org/10.1016/j.jss.2016.04.058>.
- [26] X. Chen and X. Lin, "Big Data Deep Learning: Challenges and Perspectives," in IEEE Access, vol. 2, pp. 514-525, 2014, <https://doi.org/10.1109/ACCESS.2014.2325029>.
- [27] N. Papernot, P. McDaniel, S. Jha, M. Fredrikson, Z. B. Celik and A. Swami, "The Limitations of Deep Learning in Adversarial Settings," 2016 IEEE European Symposium on Security and Privacy (EuroS&P), Saarbrucken, 2016, pp. 372-387, <https://doi.org/10.1109/EuroSP.2016.36>.
- [28] S. Ouhbi, A. Idri, J. L. Fernández-Alemán and A. Toval, "Mobile personal health records for cardiovascular patients," 2015 Third World Conference on Complex Systems (WCCS), Marrakech, 2015, pp. 1-6, <https://doi.org/10.1109/ICoCS.2015.7483226>.
- [29] M. Bachiri, A. Idri, J.-L. Fernández-Alemán, A. Toval. "Mobile personal health records for pregnancy monitoring functionalities: Analysis and potential." Computer methods and programs in biomedicine, vol. 134, pp. 121-135, 2016, <https://doi.org/10.1016/j.cmpb.2016.06.008>.
- [30] Sergeev, A., & Del Balso, M. (2018). Horovod: fast and easy distributed deep learning in TensorFlow. arXiv preprint arXiv:1802.05799.