International Journal of Computer Applications (0975 – 8887)
Volume 184– No.39, December 2022

LSTM Variants in Blood Glucose Prediction for Diabetic Patients

Rita Ganguly
Department of Computer Applications
Dr.B.C.Roy Engineering College
Durgapur-713206, West Bengal

ABSTRACT
The prediction and management of blood glucose level is indispensable for precautionary of diabetic management. The precious tool like Continuous Glucose Monitoring (CGM) system is used to minutely monitoring and refluxing the blood current glucose level of patients. The result whatever generated by CGM can be very useful prediction of future blood glucose levels with the help of machine learning models. In this research introducing a CGM based approaches on recurrent deep learning which will forecast the future blood glucose levels from the obtained data. In case of deep learning LSTM (Long Short Term Memory) and derived LSTM are used in various complex fields with complex time series data. In this study a survey report is established to determine blood glucose with LSTM based different variants to generate more accuracy in results. The goal of this survey is to lower the differences between predicted CGM values and the finger stick blood glucose readings. The output might indicate that this survey approach is realizable for more appreciable forecasting of BG that enhance and enrich the diabetes management**.**

Keywords
LSTM, CGM, Deep Learning, Blood Glucose, RNN

1. INTRODUCTION
In modern era one of the fatal and chronic metabolic disease which generate sequence of barrier and alarmingly [1] influencing to deteriorate patient daily life and health condition. The forbearing of diabetes, have to monitoring and maintaining their body glucose level in blood (Blood Glucose Level BGC) within 70-180 mg/dl [2]. Now a day with the help of CGM (Continuous Glucose Monitoring) system is widely accepted to dynamically monitor the blood glucose level of the patient in form of data which is generated through time series.

Prediction of exactly and accurately glucose level in blood is very crucial for proactively monitoring of blood glucose in early stage before it shift to unacceptable and disagreeable result. There is countless number of experimental approaches or based on any physical model to determine the blood glucose levels [6-13]. The data driven experimental models are much more uncomplicated and accessible to estimate dynamically monitoring of blood glucose with time assorted association among the variables [8,15,16]. In comparison the physical models are more difficult to provide dynamic data as it is unfeasible to evaluate metrics of the model [14]. The ensuing glucose levels are instigated by many aspects such as physical activeness, intake of dextrose, historical trends, concentration of different hormone, insulin administration and most of these metrics can be measured accurately and these are mostly variation of time [17,18].

In recent years the deep learning researcher are applying different deep learning techniques and algorithm for prediction of blood glucose . The LSTM (Long Short Time Memory) network with different metrics which provides more accuracy in result to predict bold glucose level compared to SVR based model[23].

A convolutional recurrent neural network(CRNN) proposed by Li.et al to predict blood glucose level. It is the amalgamation of modified recurrent neural network(RNN) layer with multilayered convolutional neural network (CNN) layer with extracted the features of time series data of multi dimensional [24]. Another deep learning techniques developed by Mosquera -Lopez et.al[26] which can store a huge amount of data with superior predictor performance.

Many research studies strive to fabricate deep learning models for better accomplishment considering with the significant differences in patients behavior and physiology [27-29]. The ARX model, with improve prediction accurately Luo and Zhao [30] used incremental learning and transfer learning which resolving the new patients data in sufficiency.

In contrast of traditional machine learning deep learning methods have established more superior performance. This is possible relevant due to their higher complexity features of relevant learning and ability to extract automatic features which instigate to predict accurate level of blood glucose.

In this research work highlighting at: (1) Comparing and estimating the accuracy of existing LSTM variants

(2) To ambient the parameters of the variants to analyze the best variant.

Two research questions are arise in this paper:

R1: Which LSTM variants significantly more accurate?

R2: Which LSTM variants achieve good performance?

The paper is organized as follows: section 2 explains about LSTM network, the dataset are described in section 3 , while the variants of LSTM based model for forecasting of glucose level are described in section 4. In section 5 the advantages of LSTM variants approaches given. In section 6 LSTM variants prediction accuracy in RMSE is shown and finally conclusion section 6.

2. LSTM NETWORK
LSTM (Long Short Term Memory) network is one type RNN [31] which is extensively utilized for numerous sequence labeling and prediction tasks due to higher level performances in Long and Short term data dependencies.

The structure of the LSTM represents a memory cell which stores all information through time to time and the flow of information controls through a structured gate[32]. The internal memory vector is represented like C_t where t is each step cell in LSTM which is an RNN. In internal memory the controlling part of three gates are f_t that is forget gate, the other one is input gate i_t where all the inputs are stored and the final one is output gate O_t which includes all the outputs. The two activation functions are σ which represents sigmoid function and tanh heperbolic tangent function. Through gates all information are controlled and protected in LSTM. The input vectors of the input gate is a_t and the hidden state previous output $h_{t-1.}$ The new internal memory C_t and h_t cell output calculation might follow these following expressions

$$it = \sigma(Wiat + Uiht - 1 + bi) \quad (1)$$

$$ft = \sigma(wfat + Ufht - 1 + bf) \quad (2)$$

$$ot = \sigma(woat + Uoht - 1 + bo) \quad (3)$$

$$Ct = ft.ct - 1 + it.\tanh(wcat + ucht - 1 + bc) \quad (4)$$

$$ht = ot.\tanh(ct) \quad (5)$$

where b is the bias vector

U is the previous weight metric

W is the weight matric of LSTM cell

The cell output ht produce the result of current state with nonlinear tanh function . the forgate gate f_t gives the information of previous output of the LSTM cell. Input gate depends on previous state h_{t-1} and also determine how much new candidate updated. Lastly the output gate determines the ultimate output. The current external input a_t and output previous cell h_{t-1} in which every gate is dependable.

Recurrent Neural Network (RNN) is every efficient time series applications or sequential data [33, 34]. The major challenge of RNN is gradient exploding or vanishing problem [35]. The problem is sorted out with recent deep RNN which introduces the memory cell and forget gate into classical RNN [36].

The different variants of LSTM are –convolution LSTM (C-LSTM) [37], Convolution Neural Network (CNN-LSTM)[38], Vanilla LSTM(V-LSTM) [39] ,bidirectional LSTM (Bi-LSTM)[40],Stacked –LSTM [41] have shown excellencies in prediction of result with complex time series data.

In this research, introduce a best LSTM based variants for prediction of blood glucose level with more clinical accuracy and improved prediction.

3. DATASET

The dataset used in this study collected from open source. The data consist of more than 750 adult cases. There are 3 to 4 meals each day for each patient. The data field mainly relating of pregnancy, plasma glucose, BP, skin thickness, insulin, BMI, DPF, age and the last field diabetic or non diabetic. Here mainly considerable is CGM.

4. VARIANTS OF LSTM BASED MODEL FOR FORECASTING OF GLUCOSE LEVEL IN BLOOD

4.1 Convolutional LSTM(C-LSTM)

C-LSTM make over the convolutional operation and internal computation logic. C-LSTM take the input first with convolutional part and after that feed the output of each LSTM. The convolutional LSTM is represented like:

$$ft = \sigma(wf * at + Uf * ht - 1 + bf)$$

Where "*" denotes convolution.

4.2 CNN-LSTM

CNN-LSTM is the amalgamation of CNN layers in order to take the cream of both CNN and LSTM.Some spatial inputs like prediction of image and video sequence prediction , it was designed. The CNN-LSTM architecture follows a CNN layers on derivation of max pooling layer then a pile of LSTM layers to fully connected output.

4.3 Vanilla LSTM

The architecture of the vanilla LSTM for determining of glucose where a sequence of glucose values are consider as input from RNN-LSTM and at the end the targeted output is generated.

4.4 Bidirectional LSTM(Bi-LSTM)

The procedure of making any neural network who have information sequence in both directions backward (future to past) or forward(past to future). In Bi-LSTM make the input flow in both directions to amalgamate future and past information together.

4.5 Stacked LSTM

A stacked LSTM architecture can be defined the LSTM model engrossed of multipleLSTM layers. In stacked LSTM rather than a single value output a sequence of output provided . Specially one output per input time step, rather than one input time step for all input time time steps.

It is an extension of the vanilla LSTM network by stacking a sequence of lSTM layers. In operation LSTM layer output a sequence of vectors that will be used as the input of the subsequent LSTM layer.

5. ADVANTAGES OF LSTM VARIANTS APPROACHES

Name of the LSTM Variants	Advantages
Convolutional LSTM(C-LSTM)	• Explicit maximization of the concordance correlation coefficient(ρ_c)[16] is used and show that this improve performance in terms of emotion prediction compared to optimizing the mean square error objective which is used traditionally.
CNN-LSTM	• CNN,LSTM and DNN block captures the information for future

		improvements by combining information at different scales. CLDNN architecture provides a 4% relative improvement in WER over the LSTM and features of multi scale provides 1% relative improvement. CLDNN architecture provides the robust larger data sets with different environmental constraints.
Vanilla LSTM	•	It is a combination of a set of recurrently connected sub networks, known as memory blocks. The idea behind the memory block is to maintain its state over time and regulate information flow through nonlinear gating units.
Bi-Directional LSTM	•	Bi-Directional LSTM (Bi LSTM) is a recurrent neural network used primarily on NLP unlike enable the additional training by traversing the input data twice.
Stacked LSTM	•	In stacked LSTM , each LSTM layer outputs a sequence of vectors which will be used as an input to a subsequent LSTM layer. The hierarchy of hidden layers enables more complex representation of time series data, capturing information at different scales.

6. LSTM VARIANTS PREDICTION ACCURACY IN RMSE

Among comparing of performance regarding the research [42,43,44] the Table I describe the prediction of blood glucose accuracy in C-LSTM,CNN LSTM, Vanilla LSTM and Stacked LSTM regarding the prediction horizon 30,45 and 60 mins. In respect of the observations the stacked LSTM provides the best performance of RMSE of 30 min(Short term) and 45 min(Mid term). In 60 min prediction horizon of Vanilla LSTM providing the lowest result 19.01 where as Stacked LSTM providing 19.24. Considering the overall prediction horizon performance where Stacked LSTM providing the best performance regarding the other variants of LSTM.

Table 1: Comparison of Blood Glucose Prediction Accuracy

Methods	**PH 30 min**	**PH 45 min**	**PH 60 min**
CLSTM	12.20	15.82	19.60
CNN-LSTM	13.05	16.72	19.80
Vanilla LSTM	12.33	15.86	19.01
Stacked LSTM	11.96	15.81	19.24

7. CONCLUSIONS

In this research work a comparative study result is evaluated for prediction of best LSTM variants based on prediction of blood glucose model. A comparative study among LSTM variants of widely used Algorithm and established learning algorithm to determine real time glucose using CGM data. The stacked LSTM providing the superior result regarding all variants of LSTM.

The limitations of this work is the longest term prediction horizon is showing 60 mins but it should be more than 2 Hrs. In future work more clinical data required regarding the information from Electronic Health Record for detailing of patient phenotype and personal prediction of data model and also the patients co morbidities data are very much required.

8. REFERENCES

[1] American Diabetes A (2014) Diagnosis and classifcation of diabetes mellitus. Diabetes Care 37(Suppl 1):S81-90. https://doi. org/10.2337/dc14-S081 (Epub 2013/12/21, PubMed PMID: 24357215)

[2] Kirchsteiger H, Jørgensen JB, Renard E, Del Re L (2015) Prediction Methods for Blood Glucose Concentration: Design, Use and Evaluation. Springer, Berlin

[3] Klonof DC (2005) Continuous glucose monitoring roadmap for 21st century diabetes therapy. Diabetes Care 28(5):1231–1239

[4] Deiss D, Bolinder J, Riveline J-P, Battelino T, Bosi E, TubianaRuf N et al (2006) Improved glycemic control in poorly controlled patients with type 1 diabetes using real-time continuous glucose monitoring. Diabetes Care 29(12):2730–2732

[5] The Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group (2008) Continuous glucose monitoring and intensive treatment of type 1 diabetes. N Engl J Med 2008(359):1464–1476

[6] Ozogur HN, Ozogur G, Orman Z (2020) Blood glucose level prediction for diabetes based on modifed fuzzy time series and particle swarm optimization. Comput Intell-US. https://doi.org/ 10.1111/coin.12396 (PubMedPMID:WOS:000566826000001)

[7] Xia Y, Rashid M, Feng J, Hobbs N, Cinar A (2020) Online glucose prediction using computationally efcient sparse kernel fltering algorithms in type-1 diabetes. IEEE Trans Control Syst Technol 28(1):3–15

[8] Woldaregay AZ, Arsand E, Walderhaug S, Albers D, Mamykina L, Botsis T et al (2019) Data-driven modeling and prediction of blood glucose dynamics: machine learning applications in type 1 diabetes. Artif Intell Med 98:109–134. https://doi.org/10.1016/j. artmed.2019.07.007 (PubMedPMID:WOS:000488323400010)

[9] Yu X, Turksoy K, Rashid M, Feng JY, Hobbs N, Hajizadeh I et al (2018) Model-fusion-based online glucose concentration predictions in people with type 1 diabetes. Control Eng Pract 71:129– 141. https://doi.org/10.1016/j.conengprac.2017.10.013 (PubMe dPMID:WOS:000424175400013)

[10] Oviedo S, Vehi J, Calm R, Armengol J (2017) A review of personalized blood glucose prediction strategies for T1DM patients. Int J Numer Method Biomed Eng. https://doi.org/10.1002/cnm. 2833 (Epub 2016/10/30, PubMed PMID: 27644067)

[11] Georga EI, Protopappas VC, Polyzos D, Fotiadis DI (2015) Evaluation of short-term predictors of glucose concentration in type 1 diabetes combining feature ranking with regression models. Med Biol Eng Comput 53(12):1305–1318. https://doi.org/10.1007/ s11517-015-1263-1 (PubMedPMID:WOS:000365753000006)

[12] Munoz-Organero M (2020) Deep physiological model for blood glucose prediction in T1DM patients. Sensors. https://doi.org/10. 3390/s20143896 (PubMed PMID: WOS:000557994000001)

[13] Montaser E, Diez JL, Rossetti P, Rashid M, Cinar A, Bondia J (2020) Seasonal Local Models for Glucose Prediction in Type 1 Diabetes. IEEE J Biomed Health Inform 24(7):2064–2072. https://doi.org/10.1109/Jbhi.2019.2956704 (PubMedPMID: WOS:000545429400022)

[14] Reifman J, Rajaraman S, Gribok A, Ward WK (2007) Predictive monitoring for improved management of glucose levels. J Diabetes Sci Technol 1(4):478–486

[15] Cherkassky V, Mulier FM (2007) Learning from data: concepts, theory, and methods. Wiley, New YOrk

[16] Araghinejad S (2013) Data-driven modeling: using MATLAB® in water resources and environmental engineering. Springer Science & Business Media, Berlin

[17] Diabetes Research in Children Network Study Group (2005) Impact of exercise on overnight glycemic control in children with type 1 diabetes mellitus. J Pediatr 147(4):528–534

[18] Nomura M, Fujimoto K, Higashino A, Denzumi M, Miyagawa M, Miyajima H et al (2000) Stress and coping behavior in patients with diabetes mellitus. Acta Diabetol 37(2):61–64

[19] Brazeau A-S, Rabasa-Lhoret R, Strychar I, Mircescu H (2008) Barriers to physical activity among patients with type 1 diabetes. Diabetes Care 31(11):2108–2109

[20] Sparacino G, Zanderigo F, Corazza S, Maran A, Facchinetti A, Cobelli C (2007) Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series. IEEE Trans Biomed Eng 54(5):931–937. https://doi.org/10. 1109/TBME.2006.889774 (Epub 2007/05/24, PubMed PMID: 17518291)

[21] Reymann MP, Dorschky E, Groh BH, Martindale C, Blank P, Eskofer BM (2016) Blood glucose level prediction based on support vector regression using mobile platforms. Conf Proc IEEE Eng Med Biol Soc 2016:2990–2993. https://doi.org/10.1109/ EMBC.2016.7591358 (Epub 2017/03/09, PubMed PMID: 28268941)

[22] Pérez-Gandía C, Facchinetti A, Sparacino G, Cobelli C, Gómez E, Rigla M et al (2010) Artifcial neural network algorithm for online glucose prediction from continuous glucose monitoring. Diabetes Technol Ther 12(1):81–88

[23] Mirshekarian S, Bunescu R, Marling C, Schwartz F (2017) Using LSTMs to learn physiological models of blood glucose behavior. Conf Proc IEEE Eng Med Biol Soc 2017:2887–2891. https://doi. org/10.1109/EMBC.2017.8037460 (Epub 2017/10/25, PubMed PMID: 29060501)

[24] Li K, Daniels J, Liu C, Herrero P, Georgiou P (2020) Convolutional recurrent neural networks for glucose prediction. IEEE J Biomed Health Inform 24(2):603–613. https://doi.org/10. 1109/JBHI.2019.2908488 (Epub 2019/04/05, PubMed PMID: 30946685)

[25] Aiello EM, Lisanti G, Magni L, Musci M, Tofanin C (2020) Therapy-driven deep glucose forecasting. Eng Appl Artif Intel. https://doi.org/10.1016/j.engappai.2019.103255

[26] Mosquera-Lopez C, Dodier R, Tyler N, Resalat N, Jacobs P (2019) Leveraging a big dataset to develop a recurrent neural network to predict adverse glycemic events in type 1 diabetes. IEEE J Biomed Health Inform. https://doi.org/10.1109/JBHI.2019.29117 01 (Epub 2019/04/19, PubMed PMID: 30998484) Complex & Intelligent Systems 1 3

[27] Eberle C, Ament C (2012) Real-time state estimation and longterm model adaptation: a two-sided approach toward personalized diagnosis of glucose and insulin levels. J Diabetes Sci Technol 6(5):1148–1158. https://doi.org/10.1177/193229681200600520 (Epub 2012/10/16, PubMed PMID: 23063042; PubMed Central PMCID: PMCPMC3570850)

[28] Balakrishnan NP, Rangaiah GP, Samavedham L (2012) Personalized blood glucose models for exercise, meal and insulin interventions in type 1 diabetic children. Conf Proc IEEE Eng Med Biol Soc 2012:1250–1253. https://doi.org/10.1109/EMBC.2012.63461 64 (Epub 2013/02/01, PubMed PMID: 23366125)

[29] Capel I, Rigla M, Garcia-Saez G, Rodriguez-Herrero A, Pons B, Subias D et al (2014) Artifcial pancreas using a personalized rule-based controller achieves overnight normoglycemia in patients with type 1 diabetes. Diabetes Technol Ther 16(3):172– 179. https://doi.org/10.1089/dia.2013.0229 (Epub 2013/10/25, PubMed PMID: 24152323; PubMed Central PMCID: PMCPMC3934437)

[30] Luo S, Zhao C (2019) Transfer and incremental learning method for blood glucose prediction of new subjects with type 1 diabetes. In: 2019 12th Asian control conference (ASCC); Kitakyushu-shi, Japan, pp 73–78

[31] S. Hochreiter, J. Schmidhuber: Long short-term memory. Neural computation, (1997), 1735–1780.

[32] Ashour, A. S., El-Attar, A., Dey, N., Abd El-Kader, H., & Abd El-Naby, M. M. (2020). Long short term memory based patient-dependent model for FOG detection in Parkinson's disease. *Pattern recognition letters*, *131*, 23-29.

[33] Wang, Y., Wu, Q., Dey, N., Fong, S., & Ashour, A. S. (2020). Deep back propagation–long short-term memory network based upper-limb sEMG signal classification for automated rehabilitation. *Biocybernetics and Biomedical Engineering*, *40*(3), 987-1001.

[34] Lan, K., Wang, D. T., Fong, S., Liu, L. S., Wong, K. K., & Dey, N. (2018). A survey of data mining and deep learning in bioinformatics. *Journal of medical systems*, *42*(8), 1-20.

[35] S. Park, S. Min, H.-S. Choi, and S. Yoon, "Deep Recurrent Neural Network-Based Identification of Precursor microRNAs," Nips, no. Nips, 2017.

[36] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[37] G. Trigeorgis, F. Ringeval, R. Brueckner, E. Marchi, M. A. Nicolaou, B. Schuller, and S. Zafeiriou, "Adieu features? end-to-end speech emotion recognition using a deep convolutional recurrent network," in Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on. IEEE, 2016, pp. 5200–5204.

[38] Sainath TN, Vinyals O, Senior A, Sak H. Convolutional, long short-term memory, fully connected deep neural networks. In2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2015 Apr 19 (pp. 4580-4584). IEEE.

[39] Wu Y, Yuan M, Dong S, Lin L, Liu Y. Remaining useful life estimation of engineered systems using vanilla LSTM neural networks. Neurocomputing. 2018 Jan 31;275:167-79.

[40] P. Su, X. Ding, Y. Zhang, F. Miao, and N. Zhao, "Learning to Predict Blood Pressure with Deep Bidirectional LSTM Network," pp. 1–19, 2017.

[41] A. Graves, "Generating sequences with recurrent neural networks," arXiv preprint arXiv:1308.0850, 2013.

[42] Li, Z., Wang, D., Dey, N., Ashour, A. S., Sherratt, R. S., & Shi, F. (2019). Plantar pressure image fusion for comfort fusion in diabetes mellitus using an improved fuzzy hidden Markov model. *Biocybernetics and Biomedical Engineering*, *39*(3), 742-752.

[43] Elhayatmy, G., Dey, N., & Ashour, A. S. (2018). Internet of Things based wireless body area network in healthcare. In Internet of things and big data analytics toward next-generation intelligence (pp. 3-20). Springer, Cham.

[44] Wang,T and Li.W(2020).Blood Glucose Forecasting Using LSTM Variants Under the Context of Open Source Artiicial Pancreas System. Proceedings of the 53rd Hawaii International Conference on System Science,pp-3257-3263.