A Multiheaded Convolution Neural Network for Blood Glucose Time Series Forecasting

Sofia Goel*, Sudhansh Sharma**

*School of Computer and Information Sciences, Indira Gandhi National Open University, India

Abstract- Diabetes is a common chronic condition occurs due to imbalance of blood glucose levels in the body.In this paper, a forecasting model based on deep learning algorithm is proposed for accurate blood glucose level prediction in both short (15 minute) and long (30 minute) forecasting scenarios. The proposed model is based on a Multiheaded Convolution Neural Network (MHCNN) and multiple convolutional layers for extracting useful features providing meaningful information for pattern formation. The model is tested and trained on 30 subjects that includes 10 subjects of three different categories namely adult, adolescent and child on UVA/Padova dataset. The proposed model is tested against three types of Long Short-Term Memory (LSTM) networks namely Vanilla LSTM, Layered LSTM and Bidirectional LSTM. In addition, the performance of MHCNN is compared with CNN exhibiting the role of multiple heads in architecture.The MHCNN outperformed existing LSTM models and CNN in terms of accuracy and time of execution, according to preliminary experimental results. MHCNN is much faster than other models.

Index Terms-Diabetes, Convolution Neural Network, Multiheaded, LSTM, Deep Learning, Forecasting

I. Introduction

Type 2 diabetes is a life-threatening condition that happens when the body generates insulin but does not use it properly. When the amount of glucose produced isn't properly utilized, blood sugar levels fluctuate, either too low or too high hypoglycemia or hyperglycemia arises [1][2].Therefore, blood glucose forecasting with high accuracy plays a vital role in diabetes management. Several predictive algorithms, most of which built on machine learning, were created to identify the glycemic trends of persons with diabetes based on the readings acquired by the CGM[3][4][5]. These algorithms, in particular, employ deep learning models such as CNN and RNN to recognise the structure of CGM and its temporal correlation, and then treat it using time series analysis approaches [6][7][25].In addition, research circulated developing hybrid of CNN and LSTM for forecasting BG levels [8][[9]. The main objective around all research for using neural networks was featuring longer sequences and feature extraction. For which CNN was found to be highly competitive and even better than LSTM[[10]] [11][24].The above review proves that deep learning models made a remarkable performance in all types of prediction problems. One of the most noteworthy work was proposed by Kezhi Li et al. [12][26]who significantly performed glucose forecasting by introducing a deep learning framework named Glunet. A different approach was performed by by Maxime De Bois et al. [13] and by Eleni I. Georga et al.[14] where they treat the machine learning algorithms from two different perspectives, one for short term prediction of hypoglycemic events and another for long term for prediction of hyperglycemic events.Despite their great results, one of the primary issues remains the prediction of sudden changes in blood glucose readings caused by insufficient feature extraction in time series data[27][28]. Another major challenge is in deep learning models is overfitting. Simply adding multiple layers to the neural network makes the model complicated. Another noticeable factor in diabetes data is there exists a temporal relationship between all samples of different features. Thus to capture all features optimization of hyperparameters such as kernel size is a matter of concern [29][30]. In this research, a real-time glucose forecasting model centered on time series algorithms to aid in the prevention of diabetic issues is proposed. The proposed work is based on a deep learning architecture of CNN with multiple kernel sizes, multiple convolution layers with multiple head sizes. In this study, we are performing long and short-term forecasting of blood glucose using various deep learning models. The dataset used in silico data UVA/PADOVA of 15 days for 30patients.The subjects fall into three categories namely adults, adolescents and children.

Blood glucose data is highly nonlinear and non-stationary, thus to forecast the data multiheaded CNN model is employed for feature learning and is compared with various deep learning LSTM models [15]. The reason behind the selection of LSTM and CNN for forecasting is LSTM perform sequence to sequence time series forecasting and has proven to perform long term trend analysis and CNN performs auto feature selection, extraction and forecasting all in one model [15] [16]. In this study, we have compared vanilla, layered and bidirectional LSTM models of deep learning. LSTM models used for forecasting problems of blood glucose time series data are well suited to discover the long-term dependencies in sequential data due to its potential of internal memory [10].CNN is a sort of deep neural network applied for two purposes: feature extraction component and classification part. Unlike the traditional model, in this model independent CNNs are used, which are better known as convolution heads, to deal with the prediction of blood glucose levels. Here, data is addressed individually thus avoiding the need for preprocessing of data and delivering a more customised architecture for each type of observation. This architecture is referred to as Multiheaded CNN. It is implemented as a multiheaded model to capture the correlations of blood glucose levels of the past and future. The aim of the paper is to provide the best suited model forecasting model for sequential non stationary data in two forms short term and long term forecasting.

II.Proposed Multiheaded Convolution Neural Network(MHCNN)

CNN is generally used for image processing and has become a cutting edge in this arena. We know in case of image processing 2-D convolutions are used where the kernel travels in all directions, while times series data is one dimensional thus need 1-D convolution with a single channel. Multiheaded convolution is a one-dimensional convolution neural network in which each input sequence is processed by a completely independent convolution referred to as convolution heads. Here, the role of multiple heads is to extract the features of each input sequence independently. As a result, it generates a feature map for each head consist of time series data. This generates a sequence of feature maps for each time series independently. These sequences are then concatenated together because they were obtained independently of one another.

For time-series predictions, deep learning is progressively being employed in the healthcare field. The creation of multi-headed neural network architectures for multivariate time-series forecasting is gaining popularity in research. This is due to the unique topology of multiheaded neural network where each input series comprises of independent variables can be handled by a distinct head. In multi-headed neural network topologies, the output of each of these heads can be merged before a prediction is produced. In this paper, two multi-headed ML architectures is used to predict patient's blood glucose on a quarterly basis.Another important reason to use MHCNN is when RNN gradient vanishing is a significant issue [20] . In RNN, the weights shift and eventually become so little that they have no effect on the output. As a result, the network's ability to learn from the past deteriorates, as does its operational competence for analysing extended data sequences for predictions [23]. MHCNN is used for forecasting problems because it has convolution layers for feature extraction and pattern recognition, which results in quick prediction.The MHCNN architecture used in this paper has multiple convolutional layers, each followed by a pooling layer. Each convolutional layer has a headsize of two with an independent input sequence and filter size of three. To increase the accuracy of recognition of blood glucose levels recurrent sequential processing of input features is done. These input sequences are passed through two CNN models independently and finally fused to predict the output. The suggested method uses two CNNs, each with a different configuration. One has a kernel size of 3 and the other has a kernel size of 5.

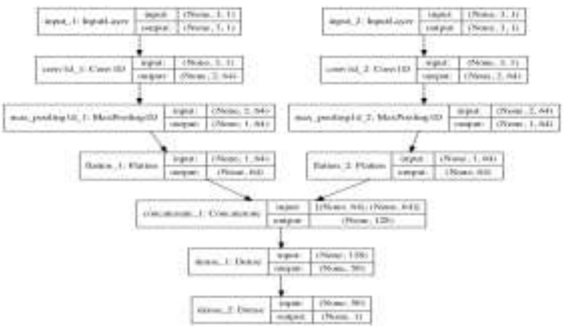

Figure 1: Structure of MHCNN

The above figure-1 shows the structure of the proposed MHCNN model.

III. Experiment

In this research simulated UVA-PADOVA silico dataset is used for training and testing. It includes a population of 30 silico subjects of three different age groups namely adults (>13 years), adolescents (age group 13-18 years) and children (2-12 years). Data is collected to forecast the blood glucose level of 5 minutes and 10 minutes based on continuous glucose monitoring done for 15 minutes and 30 minutes respectively. The experiment is conducted into two stages: Input stage and Forecasting stage.

A. Input Stage

The following steps are employed to generate the readings of blood glucose on variable meals and different calories.Simulation time is accepted in hours for 360 hrs i.e.15 days.Among two scenarios random and custom, a custom scenario is selected.The simulation started for three meals covering the amount of calorie intake for breakfast, lunch and dinner.The process of the simulation was performed for three categories of patients namely adolescent, adult and child.To collect appropriate blood glucose readings CGM sensors of three different types were taken: Dexcom, Guardian RT and Navigator which measured interstitial glucose levels for every few minutes. In addition, the Basal Bolus controller was used for keeping these blood glucose levels under control.Insulin pump glucose of two types cozmo and insulet were used to read glucose level trends over time and were visible on a built-in device screen.Finally, readings were saved in csv format to forecast the blood glucose data using machine learning algorithms.

B. Forecasting Stage

Once the results are obtained of blood glucose levels of patients then it was used to forecast the data on different deep learning models. The different LSTM deep learning models are applied for testing and training like Vanilla, Layered and Bidirectional along with CNN and Multiheaded (multiheaded CNN with varied head size). The parameters like filters, pool_size,and kernel_size are associated with the Convolutional Neural Networks models only. Multiheaded CNN model with headsize=2 is used in the experiment. The models can be used for both types of forecasting (short-term and long-term). Short-term forecasting means predicting the blood glucose level for the next 15 mins, while the long-term forecasting means predicting the blood glucose level for the next 30 minutes.In the dictionary pf parameters there were two variables named input_sequence and output_sequence. The time interval taken between the two samples is 3 minutes. Therefore, if the input_sequence is 50 and the output sequence is 10, it means 150 minutes (1.5 hours) of input data is used to forecast the next 15 minutes of blood glucose level.The performance of each model is evaluated in terms of MSE, RMSE, MAE and MAPE.The quality of prediction is also evaluated by calculating the R_square value. R_square is the coefficient of determination that determines the quality of fitness among the actual value and the forecasted value.It is observed in the experiment performed that accuracy achieved in CNN is many times faster than LSTM and performance is also very high. In case of multiheaded kernel size is taken as 3 or 5 or 7.

IV. Results and Discussion

In this section, we evaluate the performance of the proposed MHCNN model for forecasting blood glucose levels on UVA/PADOVA silico dataset. The 15 days of data is taken which comprises of three categories of instances (10 adults,10 adolescents and 10 children). The entire data is divided into two parts 70% for testing and 30% for training. The performance of the model could be analyzed from two perspectives. Firstly, based on error rate and secondly on the execution time. Five types of errors RMSE, MSE, MAE, MAPE, Square error are examined. The results of the experiment are presented in six tables. The first three tables show performance measures for predicting 15 min of data and the next three tables shows 30 min of forecasting. All the models were trained for 200 epochs with optimizer ADAM with a kernel size of 3 and 5. Moreover, to avoid dropping of features during convolution operations, the recursive sequential pattern extraction is applied.

Tables 1,2 and 3depicts the performance of the model for short term forecasting of blood glucose levels (15 min) in the case of adults, adolescents and children respectively. Here the length of the input sequence is taken as 50 and the length of the output sequence is 5. The time interval of BGM is three minutes. If we compare the error rate of MHCNN with LSTM models (Vanilla, Layered and Bidirectional) and CNN, results depicts that the proposed model is outperforming in all three cases. More specifically the value of errors exhibited in case of adults are: MSE 5.14±2.85, RMSE = 2.26±2.55, MAE =1.76±2.37, MAPE =1.25±2.38 and R_Square =0.99±0.14. Similarly, the value of error is found to be least for the MHCNN model in the case of adolescents as well as for the child. When we look at the execution time of all the models it is pertinent that the MHCNN model is significantly many times faster than all other models.

Table 1: Performance measure of Adults class of UVA/PADOVA sample dataset.For Ph=15 min (Input sequence =50 and output sequence =5)

	Ph=15min Adults					
Approach	**MSE**	**RMSE**	**MAE**	**MAPE**	**R_square**	**ExecutionTime**
LayeredLSTM	573.52 ±2.35	23.95 ±2.02	7.51±2.89	5.23± 2.98	0.92±1.33	46m14s
Vanilla LSTM	329.42 ±2.56	18.15 ±2.78	6.57±2.23	4.23± 2.10	0.98±1.23	39m19s
Bi-LSTM	118.96 ±2.23	10.90 ±3.32	9.89±2.27	7.43± 2.15	0.91±1.34	40m10s
CNN Headsize=1	11.83± 2.19	3.44± 2.15	3.05±2.35	2.15± 2.39	0.98±0.98	5m26s
MHCNN Headsize=2	**5.14±2.85**	**2.26± 2.55**	**1.76±2.37**	**1.25± 2.38**	**0.99±0.14**	**4m19s**

Table 2: Performance measure of Adolescent class of UVA/PADOVA sample dataset.For Ph=15 min (Input sequence =50 and output sequence =5)

	Ph=15min Adolescents					
Approach	**MSE**	**RMSE**	**MAE**	**MAPE**	**R_square**	**Time**
Layered LSTM	792.42 ±2.35	28.15 ±2.45	8.51±2.87	6.23± 2.12	0.92±1.98	29m67s
Vanilla LSTM	535.42 ±2.98	23.15 ±2.78	6.57±2.23	4.23± 2.10	0.98±1.09	34m11s
Bi-LSTM	240.56 ±2.35	**15.51 ±3.32**	8.77±2.27	6.56± 2.15	0.93±1.28	30m18s
CNN Headsize=1	19.73± 2.79	4.44± 2.15	6.05±2.78	4.15± 2.31	0.98±1.17	5m2s
MHCNN Headsize=2	11.76± 2.83	3.43± 2.58	2.74±2.49	1.78± 2.38	0.91±0.19	4m2s

Table 3 : Performance measure of Child class of UVA/PADOVA sample dataset.For Ph=15 min (Input sequence =50 and output sequence =5)

Ph=15min Child						
Approach	**MSE**	**RMSE**	**MAE**	**MAPE**	**R_square**	**Time**
LayeredLSTM	473.06 ±2.89	21.75 ±2.67	11.23± 2.79	5.23± 2.98	0.92±1.33	32m24s
Vanilla LSTM	329.42 ±2.98	18.15 ±2.78	6.57±2.23	4.23± 2.10	0.98±1.09	27m18s
Bi-LSTM	166.66 ±2.35	**12.91 ±3.32**	8.37±2.21	4.39± 2.15	0.93±1.28	29m13s
CNN Headsize=1	10.83± 2.34	3.14± 2.15	3.75±2.15	2.15± 2.69	0.98±1.11	4m56s
MHCNN Headsize=2	8.35±2.56	2.89± 2.23	1.46±2.38	1.55± 2.18	0.95±0.34	4m8s

Tables 4,5and 6 present the error rate for long term forecasting (30 min of prediction). The length of the input sequence remains the same as 50 while the output sequence is taken as 10. In other words, 150 min of data is predicting the BG level of the next 30min. From these tables, it can be seen that the proposed model produced lower prediction error values for all three subjects as compared to the results predicted by the LSTM models. The MSE difference in the adults using the MHCNN method was only 3.87 while the value of MSE in the case of CNN was 5.03 followed by Bidirectional LSTM as 15.91 then Vanilla LSTM with 22.5 and layered LSTM with 39.5. In addition, execution time was also 3m48s in the case of MHCNN which was gradually increased many times in other models. This proves that the MHCNN demonstrates better predictions shown by the lesser RMSE values and faster execution time.

Table 4: Performance measure of Adult class of UVA/PADOVA sample dataset.For Ph=30 min (Input sequence =50 and output sequence =10)

Ph=30min Adult						
Approach	**MSE**	**RMSE**	**MAE**	**MAPE**	**R_square**	**Time**
LayeredLSTM	1562.06±2.19	39.53± 2.19	27.51 ±2.19	15.23 ±2.19	0.82±2.19	1hr14m
Vanilla LSTM	507.15 ±2.19	22.52± 2.19	9.55± 2.19	8.16± 2.19	0.99±0.59	1hr23m
Bi-LSTM	238.57 ±2.19	15.91± 2.19	10.97 ±2.19	6.69± 2.19	0.83±0.39	56m10s
CNN Headsize=1	25.31± 2.19	5.03±2.19	2.89± 2.19	1.92± 2.19	0.98±0.45	4m26s

MHCNN Headsize=2	15.00± 2.12	3.87±2. 19	2.35± 2.19	1.65± 2.19	0.9893± 0.19	3m48 s

Table 5: Performance measure of Adolescent class of UVA/PADOVA sample dataset.For Ph=30 min (Input sequence =50 and output sequence =10)

Ph=30min Adolescent						
Approach	**MSE**	**RMSE**	**MAE**	**MAPE**	**R_square**	**Time**
LayeredLSTM	683.82 ±2.68	26.15 ±2.78	11.12± 2.34	9.86± 2.19	0.96±1.19	1hr24 m
Vanilla LSTM	240.87 ±3.23	15.52 ±2.19	9.58±2.78	7.98± 2.23	0.99±1.12	1hr13 m
Bi-LSTM	166.66 ±2.56	12.91 ±2.19	10.17± 2.89	6.62± 1.23	0.93±1.26	56m40s
CNN Headsize=1	121.66 ±2.49	11.03 ±2.19	2.89±2.39	1.92± 1.45	0.98±0.19	4m27 s
MHCNN Headsize=2	59.14 ±2.19	7.69 ±2.12	2.53±2.10	1.35± 1.67	0.99±0.15	3m44 s

Table 6: Performance measure of Child class of UVA/PADOVA sample dataset.For Ph=30 min (Input sequence =50 and output sequence =10)

Ph=30min Child						
Approach	**MSE**	**RMSE**	**MAE**	**MAPE**	**R_square**	**Time**
LayeredLSTM	270.60 ±2.19	16.45 ±2.19	9.82 ±2.67	8.86± 2.19	0.96±2.19	1hr29 m
Vanilla LSTM	156.25 ±2.19	12.50 ±2.19	9.55±3.45	6.16± 2.89	0.99±2.19	1hr13 m
Bi-LSTM	253.12 ±2.10	15.91 ±2.19	9.37±2.67	4.79± 2.23	0.99±2.02	1hr10 m
CNN Headsize=1	81.54± 1.79	9.03± 2.19	2.29±2.19	2.92± 2.34	0.98±1.34	4m16 s
MHCNN Headsize=2	31.36 ±2.19	5.60 ±2.19	3.09± 2.19	3.06± 2.19	0.99±0.89	4m24 s

Examining the results of all tables it is also noted that all the models tested in this work achieved noticeably improved performance for all the age groups as compared to other models. Furthermore, unlike previous work done, the proposed model predicts the blood glucose level with high accuracy and in less time making it a promising model for time series forecasting problems.

III. Conclusion

MHCNN is developed as a valuable solution for prediction of blood glucose level. A multi-headed 1D CNN is followed by a multilayered structure, with the CNN capturing the characteristics or patterns of the multi-dimensional time series. The improved CNN can process previous sequential data and predict the blood glucose level. Using that time series data, the proposed model performs pattern formation for each diabetes subject. The pattern obtained exhibits into a trained deep neural network which could be used wearable devices in future. In the silico dataset for short and long term forecasting of blood glucose levels, the suggested MHCNN approach outperformed the existing neural networks. It significantly proved that MHCNN is better in terms of accuracy and execution time as compared to LSTM models.Future improvements could be addressed regarding the optimization of hyper parameters such as kernel size, head sizes, pooling layers with more combinations. The model could be improved concerning the different meals intake in various conditions. It would be interesting to gather data from patients for more days. In addition, the model could be implemented on smartphone devices and wearable sensors.

IV. REFERENCES

[1] G.paracino, F.anderigo and S.Corazza, "Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series," IEEE Transactions, pp. 931-937, 2007.

[2] T.Zhu, L. Kuang, K. Li, J. Zeng and P.Herrero, "Blood Glucose Prediction in Type 1 Diabetes Using Deep Learning on the Edge," IEEE International Symposium on Circuits and Systems (ISCAS),, pp. 1-5, 2021.

[3] M.Canizoa, "Multi-head CNN–RNN for multi-time series anomaly detection: An industrial case study," Neurocomputing, pp. 246-260, 2019.

[4] A. Z.Woldaregay, E.Årsand and S.Walderhaug, "Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes.," Artificial intelligence in medicine, no. 3, pp. 109-134, 2019.

[5] H.I.Fawaz, G.Forestier and J.Weber, "Deep learning for time series classification: a review," Data mining and knowledge discovery, pp. 917-963, 2019.

[6] K.Shankar, Y. Zhang, Y. Liu, L. Wu and C.-H. Chen, "Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification," IEEE Access, vol. 8, pp. 118164-118173, 2020.

[7] M. Schuster and K. Paliwal, "Bidirectional recurrent neural networks," IEEE transactions on Signal processing, pp. 2673-2681, 1997.

[8] Y. Heryadi and H.Warnars, "Learning temporal representation of transaction amount for fraudulent transaction recognition using CNN," IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom), 2017.

[9] K.Kaya and S.Gündüz, "Deep Flexible Sequential (DFS) Model for Air Pollution Forecasting," NatureResearch, pp. 1-12, 2020.

[10] N.Xue, I.Triguero and G.P.Figueredo, "Evolving deep CNN-LSTMs for inventory time series prediction," IEEE Congress on Evolutionary Computation (CEC), pp. 1517-1524, 2019.

[11] X.Zheng and Q. Huang, "Deep Sentiment Representation Based on CNN and LSTM," International Conference on Green Informatics, 2017.

[12] K. Li, J. Daniels, C. Liu, P. Herrero and P. Georgiou, "Convolutional Recurrent Neural Networksfor Glucose Prediction," IEEE journal of biomedical and health informatics,, vol. 24, pp. 603-613, 2019.

[13] M. D. Bois, A. Mounîm, E.Yacoubi and M.Ammi, "Study of Short-Term Personalized Glucose Predictive Models on Type-1 Diabetic Children," in International Joint Conference on Neural Networks (IJCNN), 2019.

[14] E. Georga, V. Protopappas and D.Ardigò, "Multivariate Prediction of Subcutaneous Glucose Concentration in Type 1 Diabetes Patients Based on Support Vector Regression," IEEE Journal of Biomedical and Health Informatics, vol. 17, no. 1, Jan 2013.

[15] P.Chandana, G. S. P. Ghantasala, J. Rethna, K.Sekaran and Y.Nam, "An effective identification of crop diseases using faster region based convolutional neural network and expert systems.," ScholarWorks, 2020.

[16] P.Rizwan, G. Ghantasala, R.Sekaran and a. M. D.Gupta, "Smart healthcare and quality of service in IoT using grey filter

convolutional based cyber physical system.," Sustainable Cities and Society, 2020.

[17] A.Graves and J.Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," Neural networks, pp. 602-610, 2005.

[18] S. Qummar, F. G. Khan, S. Shah, A. Khan and S. Shamshirband, "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection," IEEE Access, vol. 7, pp. 150530 - 150539, 2019.

[19] E.Sreehari and G. Ghantasala, "Climate Changes Prediction Using Simple Linear Regression," ournal of Computational and Theoretical Nanoscience , vol. 16, 2019.

[20] Y. Bengio, P. Simard and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," IEEE transactions on neural networks, pp. 157-166, 1994.

[21] M. Goyal, N. D. Reeves, A. K. Davison, S. Rajbhandari, J. Spragg and M. H. Yap, "DFUNet: Convolutional Neural Networks forDiabetic Foot Ulcer Classification," IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE, vol. 4(5), pp. 728-739, 2018.

[22] L. S. and S. M., "A Novel 1-D Convolution Neural Network With SVM Architecture for Real-Time Detection Applications," IEEE sensors journal, vol. 18(2), pp. 724-731, 2017.

[23] X. Wang, Y. Lu, Y. Wang and W.-B. Chen, "Diabetic retinopathy stage classification using convolutional neural networks," in IEEE International Conference on Information Reuse and Integration (IRI), 2018.

[24] S. Oviedo, J. Veh, R. Calm, and J. Armengol, "A review of personalized blood glucose prediction strategies for t1dm patients," InternationalJournal for Numerical Methods in Biomedical Engineering, vol. 33,no. 6, p. 2833, 2017.

[25] P. Pesl, P. Herrero, M. Reddy, M. Xenou, N. Oliver, D. Johnston,C. Toumazou, and P. Georgiou, "An advanced bolus calculator for typediabetes: System architecture and usability results," IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 1, pp. 11–17, Jan 2016.

[26] A. G. Karegowda, M. Jayaram, and A. Manjunath, "Cascading kmeans clustering and k-nearest neighbor classifier for categorization of diabetic patients," International Journal of Engineering and AdvancedTechnology, vol. 1, no. 3, pp. 2249 – 8958, 2012.

[27] A. G. Karegowda, M. Jayaram, and A. Manjunath, "Cascading kmeans clustering and k-nearest neighbor classifier for categorization of diabetic patients," International Journal of Engineering and AdvancedTechnology, vol. 1, no. 3, pp. 2249 – 8958, 2012.

[28] E. I. Georga, V. C. Protopappas, D. Ardig, M. Marina, I. Zavaroni,D. Polyzos, and D. I. Fotiadis, "Multivariate prediction of subcutaneousglucose concentration in type 1 diabetes patients based on supportvector regression," IEEE Journal of Biomedical and Health Informatics,vol. 17, no. 1, pp. 71–81, Jan 2013.

[29] K. Yan and D. Zhang, "Blood glucose prediction by breath analysis system with feature selection and model fusion," in 36th Annual International Conference of the IEEE Engineering in Medicine and BiologySociety, Aug. 2014, pp. 6406–6409.

[30] H. Abdi and L. J. Williams, "Principal component analysis," WileyInterdisciplinary Reviews: Computational Statistics, vol. 2, no. 4, pp.433–459, 2010.

[31] K. Polat, S. Gne, and A. Arslan, "A cascade learning system for classification of diabetes disease: Generalized discriminant analysis andleast square support vector machine," Expert Systems with Applications,vol. 34, no. 1, pp. 482 – 487, 2008.

[32] C. Perez-Gand ́ ́ıa, A. Facchinetti, G. Sparacino, C. Cobelli, E. Gomez, ́M. Rigla, A. de Leiva, and M. Hernando, "Artificial neural networkalgorithm for online glucose prediction from continuous glucose monitoring," Diabetes technology & therapeutics, vol. 12, no. 1, pp. 81–88, 2010.

[33] C. Zecchin, A. Facchinetti, G. Sparacino, G. D. Nicolao, and C. Cobelli,"Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration," IEEE Transactions onBiomedical Engineering, vol. 59, no. 6, pp. 1550–1560, Jun. 2012.

[34] K. Plis, R. Bunescu, C. Marling, J. Shubrook, and F. Schwartz, "A machine learning approach to predicting blood glucose levels for diabetes management," in Modern Artificial Intelligence for Health AnalyticsPapers from the AAAI-14.

[35] H. Mhaskar, S. Pereverzyev, and M. van der Walt, "Adeep learning approach to diabetic blood glucose prediction,"https://arxiv.org/abs/1707.05828, 2017.

[36] C. Marling and R. Bunescu, "The OhioT1DM dataset for blood glucoselevel prediction," in The 3rd International Workshop on Knowledge Discovery in Healthcare Data, Stockholm, Sweden, July 2018.

[37] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick,S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in Proceedings of the 22Nd ACM InternationalConference on Multimedia, ser. MM '14, 2014, pp. 675–678.

[38] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: review, opportunities and challenges," Briefings in Bioinformatics, vol. 19, no. 6, pp. 1236–1246, 05 2017. [Online].Available: https://dx.doi.org/10.1093/bib/bbx044

[39] Q. Zhang, Y. N. Wu, and S.-C. Zhu, "Interpretable convolutional neuralnetworks," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2018, pp. 8827–8836.

[40] Y. Bengio, "Deep learning of representations: Looking forward," inStatistical Language and Speech Processing. Berlin, Heidelberg:Springer Berlin Heidelberg, 2013, pp. 1–37.

[41] J. Schmidhuber, "Deep learning in neural networks: An overview,"Neural Networks, vol. 61, pp. 85 – 117, 2015.

First Author – Sofia Goel, Pursing Ph.D., Indira Gandhi National Open University, (sofiagoel@gmail.com)

.

Second Author – Dr.Sudhansh Sharma,Ph.D,
School of Computer and Information Sciences (SOCIS),
Indira Gandhi National Open University,
New Delhi, India
(sudhansh@ignou.ac.in)