Early Detection of Diabetes Mellitus Type-2 Using Deep Learning Technique

[1]Sai Prakash S and [2]Dr. A.C. Subhajini,
[1]Research Scholar, [2]Associate Professor,
[1]Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India

Abstract: Diabetes mellitus (DM) is aunceasingillness that causesimbalances in glucose level in blood due to the body's reluctance in generating insulin hormone. Because of its high morbidity, it has become a growing worry, and the average age of individuals affected by this disease has now dropped to the mid-twenties. Given its prevalence, it is critical to address this issue effectively. Many academics and doctors have now developed AI-based detection approaches to better tackle problems that are ignored owing to human errors. ML and DL approaches have been utilised to predict diabetes and its consequences in recent years. This research provides a DL strategy for diagnosing DM using CNN-Bi-LSTM.The approach entails retrieving essential elements from a dataset of diabetes clinical records and feeding them into a DNN. The network is then trained to recognise diabetes-related patterns in the data. The model is tested using a distinct dataset. The tests are carried out using the PIDD dataset, which contains 768 record and 8 critical variables connectedwith diabetes, each with a grouptag indicating the result of non-diabetic and diabetic individuals. The primary goal of this study is to maximise the model's accuracy.

Keywords- *Diabetes, Machine learning, deep learning, PIDD dataset, CNN-Bi-LSTM*

I. INTRODUCTION

People's lifestyles are too hectic nowadays, and majority of them do not look after their wellbeing or know how to protect it. It can lead to develop a variety of lifestyle disorders, including DM, which is, as usual, one of the diseases that is strongly tied to our lifestyle. If left untreated, it can be the worst disease [29, 25]. Our bodies require energy to function, and blood glucose is the main source of energy, which is obtained from the foodstuff we consume. The pancreas is a vital organ in our bodies that produces insulin. Insulin is a hormone that is essential for controlling blood sugar levels.Insulin regulates blood sugar levels so that they do not become too low or too high, which are both dangerous.Glucose is derived from carbohydrates in the foods that humans consume and is responsible for the healthy functioning of the body. As shown in Figure 1, normal people's blood sugar levels after meals range from 70 to 80mg/dl, but diabetics' levels of blood sugar ranged from 170mg/dl and higher.

High levels of blood sugar result in symptom such as increased thirst, appetite, and increased urination. Diabetes, if left untreated, can result in a slew of consequences. Diabetic keto-acidosis and non-ketotichyperosmolarcoma are two of the most serious consequences. DM is a toxic condition in which the pancreas in the body is incapable to create sufficient insulin or no insulin, or the body is unable to respond to insulin owing to deficient action. As a result of the anomalies in blood glucose levels, several side effects occur, including eye problems, stroke, renal failure, hyper-tension, and cholesterol.

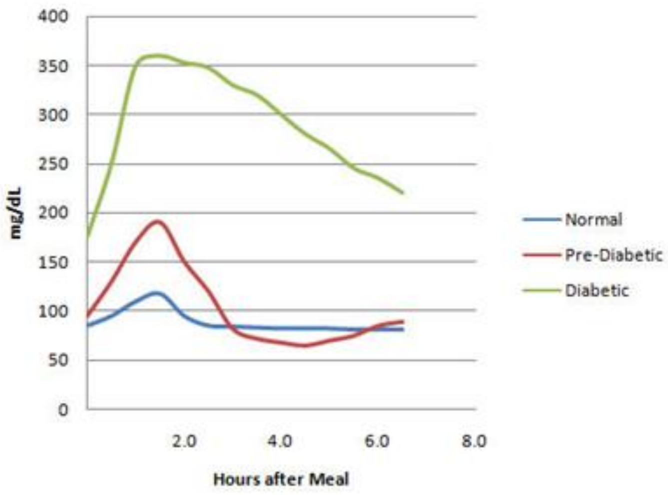

Fig.1: Graph of sugar level in blood

Diabetes [3, 1] is a long-term condition with numerous risk factors and complexities that increase mortality rates. Type-1 [26], type-2 [6], pre-diabetes [19], and gestational diabetes [2] are the four types. Type 1 diabetes (T1D) is a dangerous condition that causes chronic illness in both children and adults. The pancreas completely quits producing insulin at this point. It is completely dependent on insulin from outside sources to control their blood sugar levels. The DCCT (Diabetes Control and Intricacies Trial) supported the individual through the rundown solutions with being followed to stay away from symptom, extreme difficulties on various organs, and live a longer healthier life through the rules and nourishment inclinations [8]. These criteria led to the discovery of a dietary regimen.

Type-2 diabetes (T2D) is anunceasing, non-insulin-dependent illness that commonly affects adults. Type 2 occurrences are characterised by genetic and metabolic factors, ancestral background, and physiological immobility.Overweight, obesity, poor eating habits, and smoking habits all increase the risk of diabetes [5]. Pre-diabetes is a stage before type 2 diabetes in which the individual's glucose levelis higher than usual but not to the levels of type 2. Under specified settings and methods, a man with pre-diabetes has a higher risk of developing type 2. Gestational is a basic classifying influenced for pregnant women [33, 20]. A multitude of hormones during pregnancy, as well as increased insulin secretion, can cause high level of glucose in blood.Diabetes is a possibility for newly created newborns.

When there is no diabetes, pancreas functions normally and releases adequate insulin. When insulin connects to receptor on the cell's surfaces, the molecule of glucose gains access to the cell. T1D causes the pancreas to gradually stop generating insulin, which impairs the mechanism of glucose transport to cell. Type-2Diabetes is not caused by an inability of the pancreas to generate insulin. Although there is sufficient insulin and glucose reaching cell, the receptors of insulinwhich enable insulin to reach cell have failed their capacity to react to insulin.Figure 2 depicts the methods by which cell absorbs glucose in individual who are normal, have T1D, or have T2D.

The diagram depicts three distinct processes of glucose entry into cells. The first procedure is typical when a person is not diagnosed as having diabetes and his pancreas generates sufficient insulin. When anindividual has T1D, the second procedure is used. The pancreas eventually ceases to produce insulin, affecting the mechanism of glucose transfer to cells.The lastprocedure depicts the effects of T2D, which is not due to the pancreas's inability to make insulin. Both sorts are extremely harmful and induce a slew of problems and illnesses.

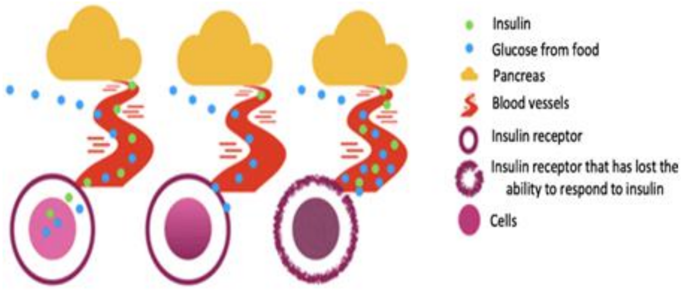

Fig.2:Glucose entry processes in cells

1. DEEP NEURAL NETWORK

Diabetes prediction [12] and early identification [24] are critical for all people who are prone to diabetes. Numerous diseases may now be identified using AI approaches, and deep neural networks [32] have obtained the best classification performance. DNNs have been used to diagnose a variety of ailments in recent years. Deep learning (DL) is a novel study path in the field of machine learning (ML) [34, 23], and it has made significant progress in speech recognition and computer vision applications in recent years.

Deep learning algorithms play acrucialinfluence in the illness prediction system process of forecasting disease level. It can handle enormous amounts of data faster than ML methods [15].When employed with less datasets, the ML algorithm typically outperforms the DL method. At the same time, DL algorithms perform better than ML algorithms when dealing with very huge amounts of data. However, because the dataset is taught repeatedly during the learning process, the DL method is more time consuming. The features of datasets can be thoroughly examined by means of recurrent learning methods.

DNNs are based on the organic nervous system. Synapses are the links between artificial neurons, which are simple functions shown as nodes segregated in layers [18]. It is a data-driven, self-adaptive training method that generates non-linear models suitable of simulating actual issues.Figure 3 depicts a simple deep neural system design. The structure is composed of four layers that simulate neurons and nodes and are all directed in the same direction. Each layer of nodes' output is reliant on the output of the preceding layer.Every node contains a one-way link to the next nodeand it includes two hidden layers in which each node uses local input to train a replica of global modeling variables. Furthermore, it processes the model across numerous threads and applies averaging to contribute to model access across the entire network.

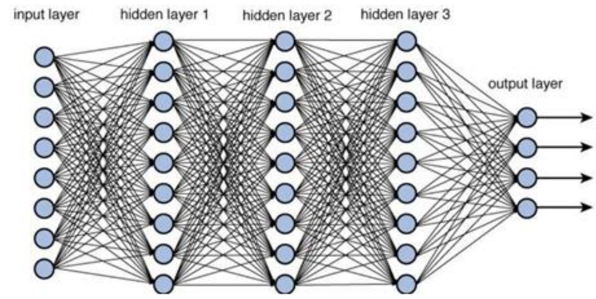

Fig.3: Deep Neural Network architecture.

The activation function is critical in DNN. Activation function must incorporate complex non-linear mapping features in order to gain the much-needed non-linearity quality that allows them to approximate any function. They are also important in squashing the infinite linearly weighted sum of neurons. It is also vital to avoid significant value accumulation high up the processing hierarchy. There are several types of DL algorithms available, including:

A. Convolutional Neural Network (CNN)

A CNN is an image processing architecture with multiple layers that perform various activities as needed. The primary layer is a convolution layer, which is highly useful for generating pattern recognition filters. The ReLU layer is used to normalize picture data to be in a positive vector. The flatten and fully connected layers follow, with the flatten converting the 2-D array to a 1-D array, which is then sent to the fully connected layers, which act similarly to ANNs and yield as many results as desired (figure 4).

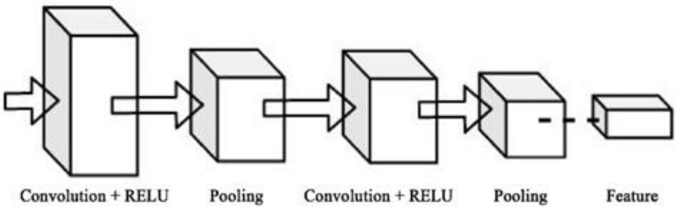

Fig.4: CNN Structure

B. Recurrent Neural Network (RNN)

RNN is similar to an ANN, except instead of a set of constant feed forward connections in nodes, it also has feed backward connections. Both sorts of connections might occur between different nodes of the layers (figure 5). It also provides accessibility to contextual layers and hidden units, with just the input and output layers remaining same.

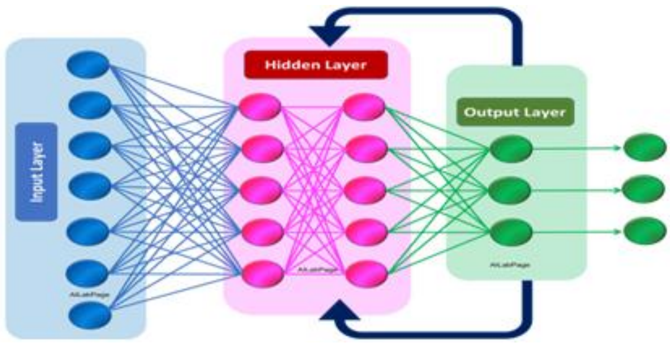

Fig.5: Architecture of RNN

C. Long Short-Term Memory (LSTM)

It is Recurrent Neural Network derived architecture. It essentially eliminates the features where neurons occur and replaces them with a new technology known as memory cell.

The memory cell can hold its value for a brief or lengthy period of time depending on the inputs (figure 6). The three gates that regulate the information flow to and from a memory cell are the forget gate, the input gate, and the output gate.

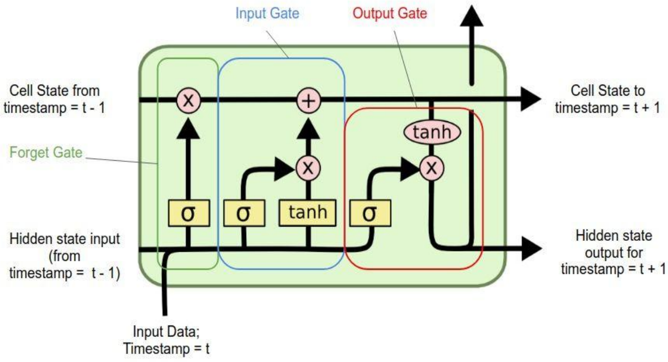

Fig.6: Architecture of LSTM

D. Deep Belief Network (DBN)

It is a basic network with an innovative training technique. It is multilevel, with each pair of layers linked to a limited Boltzmann machine (figure 7). Each Restricted Boltzmann machine is trained to reconstruct all sensory inputs received from the input layer and then transfer this information to the hidden layers via unsupervised learning and the output layers via supervised learning to provide an appropriately tailored outcome.

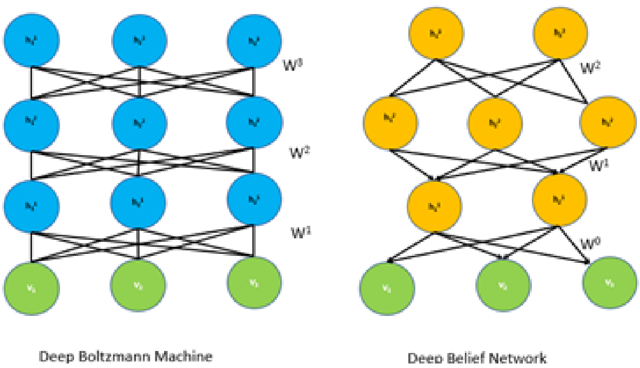

Fig.7: DBN Architecture

2. PIDD (PIMA INDIANS DIABETES DATABASE)

Table 1 displays the information required to envisage diabetes, in addition to the definitions and value for these variables. The presented technique is validated using diabetes data from the Pima Indians from the UCI repository. The fundamental reason for utilising the PIDD dataset is because the majority of the people in today's society lives anidentical ways of life, with a greater reliance on packaged food and a drop in physical activities. The information included diagnostic factors and measurements that allowed the patients to be diagnosed with any type of unceasingillness or diabetes in advance. PID participants are all females who are at least 21 years old.PID was made up of 768 samples, 268 of which were diabetic and 500 of which were not. The following are the eight most influential factors that led to diabetes prediction: the patient's previous pregnancies, BMI, level of insulin, age, Blood Pressure, thickness of Skin, Glucose, Diabetes Pedigree function with the label result.

Table.1: Attributes of the PIMA Indian dataset

ATTRIBUTE	DESCRIPTION	VALUE
Preg	Number of pregnancies	[0 – 17]
Plas	Plasma glucose concentration in an oral glucose tolerance test	[0-199]
Pres	Diastolic blood pressure	[0-122]
Skin	Triceps skin fold thickness	[0-99]
Insu	2-Hour serum insulin	[0-846]
Mass	Body mass index	[0-67]
Pedi	Diabetes pedigree function	[0-2.45]
Age	Age of an individual	[21-81]
class	Tested positive / negative	(0,1)

Figure 8 graphically depicts the various properties of the each attributes and their ranges in the PIMA dataset.

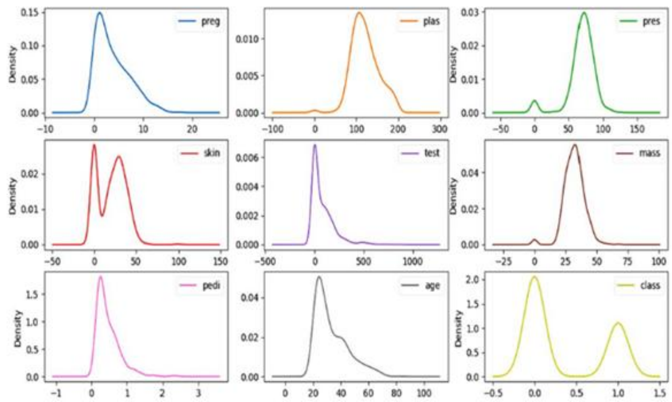

Fig.8: Charts displaying several parameters in the Pima Indian dataset

3 DIABETES PREDICTION BASED DL TECHNIQUES

Figure 9 depicts the frequency of use of DL approaches during the last six years.

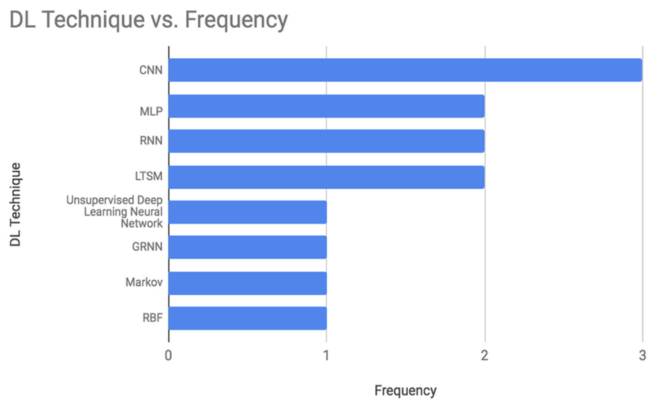

Fig.9: DL Techniques Frequency.

CNN is a subtype of MLP that was first employed in image processing. CNN performs effectively with grid topologicalinformation. The mainbenefit of CNN is that it automatically extracts features. It can also display the correlation between the elements for the given inputs [17]. CNN classification takes less time and produces higher accuracy, particularly in image processing. The computational cost, on the other hand, is significant, and it cannot perform well without a substantial amount of training data [10]. It has been used to predict diabetes with an accuracy of 77.9%-95.8%.

MLP is a supervised classifier that consists of numerous layers and basic linked neuron that provide a nonlinear mapping among output and input vectors. MLP can be applied to a variety of data formats, including text and images. It is appropriate for regression and prediction. MLP has numerous applications, including predicting ozone levels [13]. Moreover, because its neurons are entirely interconnected, the total number of parameters may be very enormous, increasing processing time. As a result, it may be insufficient for today's advanced computer vision sector. MLP was utilized to predict diabetes 2 times, achieving accuracy of 71.6% and 89.42%, accordingly. The second application integrated the General Regression Neural Network with the RBF.

RNN is a robust and powerful version of ANN. RNN analyzes input sequences each one at a time, keeping data on previous neurons in hidden neurons. As a result, the output of distinct neuron at varying time steps is fed into the neurons of the following time interval. They can handle multifaceted issues like guessing the next word in a document or correctly interpreting the words that comprise a sentence. This is owing to its skill to memorise. As a result, it can be more precise in prediction issues [4]. Its fundamental restriction is the problem of vanishing gradients. It occurs when information is distorted as a result of being multiplied by small values less than zero.As a result, calculating the weights and biases that reduce accuracy is complicated. RNN has been utilised in Diabetes prediction two times in the recent six years. The uppermost level of accuracy achieved was 82%.

RNN has been improved via LSTM. It consists of several memory blocks that can solve the vanishingincline problem. It is most commonly found in NLP applications. It performed well and shown advancements in speech recognition, language modelling, and computer vision. One of its key disadvantages is the long computational time necessary to determine the memory bandwidth for the processing units, making data training more difficult [13].In the previous six years, LSTM has been used to predict Diabetes twice. The first time it was merged with CNN, the accuracy was 94.9%. The last employed simply LSTM and obtained 58.9% precision.

The purpose of this study is to help the assortment of DLtechnique and characteristics for developing a new diabetes prediction models. The PIDD dataset is utilised in to predict diabetes in a patient. To attain high accuracy, experimental performance is compared across multiple measures.

II. LITERATURE REVIEW

Past literature publications directly linked to the suggested methodology are described below in this part. DL algorithms are widely used in medical applications to anticipate various illnesses. Many researchers have attempted to create diabetes prediction systems using various DL models. This section discusses significant diabetes detection research. Previous reviews investigated DL approaches in diabetes, but with a very different focus.

Islam et al. [14] published a meta-analysis of DL models for detecting diabetic retinopathy (DR) in retinal fundus pictures. This review comprised 23 papers, with 20 of them also being included in the Meta - analysis.The researchersrecognized the model, datasets, and performance measures for each study and found that automated methods could perform DR screening. Chaki et al. [7] examined ML models in the identification of diabetes. The study covered 107 researches and classed them based on the model or classifier, the datasets, feature selection with four different types of features, and their performances. The authors discovered that text, shape, and texture factors resulted in improved results. Furthermore, they discovered that DNNs and SVMs outperformed RFs in classification.

In [21], Nai-Arun et al. used an algorithm to classify the risk of DM. To achieve the goal, the author used four well-known machine learning classification methods: DT, ANN, LR, and Naive Bayes.Bagging and boosting procedures are employed to enhance the robustness of the developed model. Experiment outcome reveal that the RF method outperforms all other algorithms tested. Pradeep and Naveen [22] developed several classification methods. The goal was to find a better way to forecast diabetic illness by looking at blood glucose levels before 2 hours. A large number of approaches and procedures are currently being used in the diagnosis of diabetes disease. The J48 DT method is used for feature identification in the reviewed technique for classification and prediction. The primary diagnosis yields the most cost-effective strategy. J48 algorithm is a common algorithm with good accuracy.

Guo et al. [11] created the Bayes Network to identify people with T2D. The database PIDD was used, which collects data from patients with and without T2D. The Weka software was used throughout the investigation. The technique produced correct results, demonstrating that the Bayes network developed to predict T2D is more dominant. T. MahboobAlam [31] mentions significant study in diabetes detection; diabetes is predicted utilising significant features, and the relationship between the various attributes is also characterised. The PCA method was employed to pick significant features.The scientists discovered a robust link between diabetes and BMI as well as glucose levels. Following this, the classification stage was conducted using RF,ANN and k-means clustering algorithms, with the highest accuracy attained by the ANN technique with a 76.3% success rate.

Swapna et al. [9] advocated using DL to identify diabetes. Diabetes was detected using CNN, LSTM, and the combination CNN-LSTM in this system. The heart rate was calculated from the acquired ECG signal, and this study demonstrates that diabetes may be identified using ECG data. CNN5-LSTM with SVM attained the highest accuracy of 96.1%. Kumari et al. [28] introduced a soft computing-based diabetes prediction system that employs an aggregation of three most widely employed supervised machine learning methods.For evaluation, they employed PIMA and breast cancer databases. They evaluated their performance with state-of-the-art individual and ensemble algorithms using RF, LR, and naive Bayes, and their method outperforms with 79.1% accuracy. Deep learning was proposed by Kamble and Patil [30] as a tool for detecting diabetes. The Boltzmann machine was employed to determine whether the patient was diabetic or not. Furthermore, the suggested method utilized a decision tree to determine if the patient was T1D or T2D. The dataset had 300 records. Min-max normalisation was used to normalise the dataset.The proposed method made use of theJ48 and LR algorithms. The findings demonstrate that the J48 was more accurate than the LR, with a 94.1% accuracy score for the J48 and 80.4% accuracy for the LR.

Gupta et al. [27] used naive Bayes and SVM techniques to classify diabetes. The dataset used was PIDD. They also used a feature selection technique and k-fold cross-validation to improve accuracy of the model. The experiments showed that the SVMsurpassed the naïve Bayes model.

CONCLUSION

Techniques for deep learning perform an essential part in healthcare by providing a scheme to assess medical information for disease identification. One of the most pressing real-world medical issues is the early identification of diabetes. To address this issue, systematic attempts are being made to build a system that will result in the prediction of diseases such as diabetes. In this study, a new DL model for diabetes mellitus prediction was proposed utilising CNN-Bi-LSTM and datasets from the PIDD. This study contributes to a clearer understanding of diabetes prediction and the progress of improved diabetes prediction approaches for overcoming diabetes through prompt forecast.

References

[1] American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2012; 35(1):S64–71. https://doi.org/10.2337/dc12-S064.

[2] Anna V, van der Ploeg HP, Cheung NW, Huxley RR, Bauman AE. Socio-demographic correlates of the increasing trend in prevalence of gestational diabetes mellitus in a large population of women between 1995 and 2005. Diabetes Care. 2008;31(12):2288–93.

[3] Avogaro P, Crepaldi G, Enzi G, Tiengo A. Associazione di iperlipidemia, diabetemellito e obesita` di mediogrado. ActaDiabetolLat. 1967; 4:36–41.

[4] Balaji, H.; Iyengar, N.; Caytiles, R.D. Optimal Predictive analytics of Pima Diabetics using Deep Learning. Int. J. Database Theory Appl. 2017, 10, 47–62.

[5] Barengo NC, Katoh S, Moltchanov V, Tajima N, Tuomilehto J. The diabetes-cardiovascular risk paradox: results from a Finnish population-based prospective study. Eur Heart J. 2008;29(15):1889–95

[6] Bellamy L, Casas JP, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. Lancet. 2009; 373:1773–9.

[7] Chaki J, Ganesh ST, Cidham SK, Theertan SA. Machine learning and artifcial intelligence based diabetes mellitus detection and self-management: a systematic review. J King Saud UnivComputInf Sci. 2020. https://doi.org/10.1016/j.jksuci.2020.06.013

[8] Darnton-Hill I, Nishida C, James WPT. A life-course approach to diet, nutrition and the prevention of chronic diseases. Public Health Nutr. 2004;7(1):101–21.

[9] G. Swapna, K. P. Soman, and R. Vijayakumar, “Automated detection of diabetes using CNN and CNN-LSTM network and heart rate signals,” Procedia Computer Science, vol. 132, no. Iccids, pp. 1253–1262, 2018.

[10] Goodfellow, I.; Bengio, Y.; Courville, A.; Bengio, Y. Deep Learning; MIT Press: Cambridge, MA, USA, 2016.

[11] Guo, Y., Bai, G., Hu, and Y.: Using bayes network for prediction of type-2 diabetes. In: International Conference for Internet Technology and Secured Transactions, pp. 471–472 (2012)

[12] H. Liu, H. Kou, C. Yan, L. Qi, Link prediction in paper citation network to construct paper correlated graph. EURASIP J. Wirel. Commun. Netw., 233 (2019)

[13] Hassan, A. Deep Neural Language Model for Text Classification Based on Convolutional and Recurrent Neural Networks. Ph.D. Thesis, University of Bridgeport, Bridgeport, CT, USA, 2018

[14] Islam MM, Yang H-C, Poly TN, Jian W-S, Li Y-CJ. Deep learning algorithms for detection of diabetic retinopathy in retinal fundus photographs: a systematic review and meta-analysis.ComputMethodsProgramsBiomed.2020;191:105320. https://doi.org/10.1016/j.cmpb.2020.105320

[15] Justin Ker, Lipo. P, Wang, Jai, Rao&Tchoyoson Lim 2018, ‘Deep learning applications in medical image analyses, IEEE Access, vol.6, pp. 9375-9389.

[16] Kumar, D.A., Govindasamy, R., 2015. Performance and Evaluation of Classification Data Mining Techniques in Diabetes. International Journal of Computer Science and Information Technologies, 6, 1312–1319

[17] Lekha, S.; Suchetha, M. Real-Time Non-Invasive Detection and Classification of Diabetes Using Modified Convolution Neural Network. IEEE J. Biomed. Health Inform. 2018, 22, 1630–1636.

[18] Maniruzzaman M, Kumar N, Abedin MM, Islam MS, Suri HS, El-Baz AS, Suri JS. Comparative approaches for classifcation of diabetes mellitus data: machine learning paradigm. Comput Methods Programs Biomed. 2017; 152:23–34. https://doi.org/10.1016/j.cmpb.2017.09.004.

[19] Meigs JB, D’Agostino RB Sr, Wilson PW, Cupples LA, Nathan DM, Singer DE. Risk variable clustering in the insulin resistance syndrome: the Framingham Ofspring Study. Diabetes. 1997;46:1594–600

[20] Metzger BE, Lowe LP, Dyer AR, et al. Hyperglycemia and adverse pregnancy outcomes. N Engl J Med. 2008; 358:1991–2002.

[21] Nai-Arun, N., Moungmai, R., 2015. Comparison of Classifiers for the Risk of Diabetes Prediction. Procedia Computer Science 69, 132–142. doi:10.1016/j.procs.2015.10.014.

[22] Pradeep, K.R., Naveen, N.C.: Predictive analysis of diabetes using J48 algorithm of classification techniques. In: 2nd International Conference on Contemporary Computing and Informatics (IC3I), pp. 347–352 (2016)

[23] Q. Liu, G. Wang, F. Li, S. Yang, J. Wu, Preserving privacy with probabilistic indistinguishability in weighted social networks. IEEE Trans. Parall. Distrib. Syst. 28(5), 1417–1429 (2017)

[24] R. Miotto, F. Wang, S. Wang, X. Jiang, T. Dudley, Deep learning for healthcare: review, opportunities and challenges. Brief. Bioinform. 19(6), 1236–1246 (2017)

[25] R. P. Ambilwade and R. R. Manza, "Prognosis of diabetes using fuzzy inference system and multilayer perceptron," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), Noida, pp. 248-252, 2016

[26] Raha O, Chowdhury S, Dasgupta S, et al. Approaches in type 1 diabetes research: a status report. Int J DiabDevCtries. 2009; 29(2):85–101. https://doi.org/10.4103/0973-3930.53126.

[27] S. Gupta, H. K. Verma, and D. Bhardwaj, “Classification of diabetes using Na¨ıve Bayes and support vector machine as a technique,” Operations Management and Systems Engineering, Springer, Singapore, pp. 365–376, 2021

[28] S. Kumari, D. Kumar, and M. Mittal, “An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier,” International Journal of Cognitive Computing in Engineering, vol. 2, 2021.

[29] Sisodia, D., Sisodia, D.S., "Prediction of Diabetes using Classification Algorithms," in: International Conference on Computational Intelligence and Data Science (ICCIDS 2018), ELSEVIER. Procedia Computer Science, ISSN 1877-0509, vol 132.

[30] T. Kamble and S. Patil, “Diabetes detection using deep learning approach,” International Journal for Innovative Research in Science & Technology, vol. 2, no. 12, pp. 342–349, 2016.

[31] T. MahboobAlam, M.A. Iqbal, Y. Ali, A. Wahab, S. Ijaz, T. ImtiazBaig, A. Hussain, M.A. Malik, M.M. Raza, S. Ibrar, Z. Abbas, A model for early prediction of diabetes, Inform. Med. Unlocked 16 (2019) 100204, doi:10.1016/J.IMU.2019. 100204

[32] W. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5(4), 115–133 (1943)

[33] World Health Organization. Diagnostic criteria and classification of hyperglycemiafirst detected in pregnancy. World Health Organization. 2013. http://www.who.int/iris/handle/10665/85975.

[34] Y. Zhang, G. Cui, S. Deng, et al., Efficient query of quality correlation for service composition. IEEE Trans. Serv. Comput. (2018). https://doi.org/10.1109/TSC.2018.2830773