تنبؤ بإمكانية الإصابة بمرض السكري مع التعلم العميق: نموذج الهجين ESA و LSTM

د . محمد صالح الحصني _ جامعة درنه _كلية العلوم _ قسم حاسوب
د . اكرم خليل غيضان _ كلية التربية القبة _ جامعة درنة ـ قسم حاسوب
د . سلمى محمد بوبكر البرغثي _ كليه العلوم القية _ قسم الحاسوب
د . أسامه محمد عبدالونيس عوض _ كلية الهندسة القبة ـ جامعة درنة _ القسم العام

الملخـص :

يستخدم البحث الحالي نموذج التعلم العميق للتنبؤ بحالات السكري من قاعدة بيانات Pima Indians يجمع النموذج بين شبكات الذاكرة طويلة المدى (LSTM) والشبكات العصبية التلافيفية (CNN) من خلال تحليل دلالي واضح (ESA). يتم تطبيق النموذج المقترح على مجموعة البيانات باستخدام إعدادات مختلفة ، حيث وصلت دقة النموذج الهجين إلى 86.4٪. كانت أفضل دقة عند مقارنتها بالنماذج الأخرى باستخدام تصنيف واحد بواسطة LSTM أو.CNN

الكلمات الرئيسية : LSTM، CNN، ESA، التنبؤ بمرض السكري

A PREDICTION OF DIABETES POSSIBILITY WITH DEEP LEARNING: LSTM AND ESA HYBRID MODEL

Mohaned s.m.hasni Computer science Derna university
m.hasni@uod.edu.ly
Akram. Gihedan Derna university Computer science
Akram_kalil2010@yahoo.com
Salma .Albargathe Derna university Computer science
albrgate83@gmail.com
OUSAMA M ABDULWANES AWAD Derna university
College of Engineering osamaalawkle@gmail.com

Abstract

The current research uses a deep learning model for the prediction of diabetes cases from Pima Indians database. The model combines Long-short term memory networks (LSTM) and convolutional neural networks (CNN) through an explicit semantic analysis (ESA). The proposed model is applied to the dataset using different settings, where the accuracy of the hybrid model reached to 86.4%. It was the best accuracy when compared to other models using a single classification by either LSTM or CNN.

Keywords: LSTM, CNN, ESA, Diabetes prediction

INTRODUCTION

Since the diagnosis process of some diseases is difficult, the disease can be diagnosed using systems such as artificial neural network, image processing and deep learning by using clinical and physical data of various diseases such as diabetes [1]. Diabetes is one of the most common chronic diseases that can be seen in all age groups, both small and large. The pancreas is an organ responsible for producing the hormone insulin, a vital enzyme that regulates the sugar level in the human blood stream. Insulin is a vital and important component in balancing metabolism in the human body. It is known that diabetes occurs due to insufficient production of insulin or inappropriate use of the produced insulin [2]. However, the underlying cause of diabetes varies depending on whether it is Type 1 or Type 2 diabetes.

There are two different types of diabetes, type 1, and type 2 [3]. Due to the lack of insulin, Type 1 diabetes occurs. Your immune system, which normally fights harmful bacteria or viruses, attacks and destroys the insulin-producing cells in the pancreas. The only solution for this type of diabetes is to inject the required amount of insulin into the patient's body as a supplement. In type 2 diabetes, cells become resistant to the action of insulin and the pancreas cannot release enough insulin to overcome this resistance. It is believed that genetic and environmental factors also play a role in the development of type 2 diabetes. Being overweight is strongly linked to the development of type 2 diabetes. The prevalence of diabetes is increasing more rapidly in middle- and low-income countries. It is known that diabetes is among the main causes of blindness, kidney failure and heart attack.

A diabetic can lead a normal life with proper diagnosis and treatment. There are different measures for defining diabetes, such as A1c, random blood glucose, fasting blood glucose, and oral glucose tolerance test [4]. Diabetes identification based on a single parameter can lead to misdiagnosis and misleading decision making. Therefore, there is a need to combine different parameters to effectively diagnose diabetes. It will be more effective to identify and treat diabetes with data such as glucose, body mass index, diabetes pedigree, blood pressure, age, pregnancy, and skin thickness.

The prediction and early diagnosis of the disease depends on a doctor's knowledge and experience, but in some cases, misdiagnoses can occur. The health sector collects a large amount of data on health services, but this data is not used effectively in undetected cases. Manual decisions can be quite dangerous for early disease diagnosis, as they are based on the not always accurate observation and judgment of the healthcare professional [5]. As a result, patients receive poor service; Therefore, different advanced mechanisms and various software-based programs are considered necessary for automatic diagnosis and early detection of diseases with better accuracy. For example, the development of machine learning algorithms helps to process large amounts of data and extract the underlying data model that facilitates decision making.

In this study, a combination of Explicit Semantic Analysis (ESA) and Long Short-Term Memory (LSTM) models was used to predict diabetes. Models were tested by trying different parameters to find the most suitable configuration for the models. In addition, the performance of the proposed model was compared with other methods in the literature. The second section of the article compares the existing literature to understand the differences between the used models. Furthermore, the methods used in the current research are explained, followed by a review of the used data set, experimental application, and the diabetes predictions. Finally, the findings of the prediction model are discussed with the relevant literature.

RELEVANT LITERATURE

There are many modern techniques in the literature created for the diagnosis of diabetes. The algorithms used in Dwivedi [6], in which the performance of machine learning algorithms for diabetes prediction were evaluated are support vector machine (SVM), artificial neural network (ANN), logistic Regression and K-nearest neighbor (K-NN). The performance of the system was evaluated in terms of accuracy, specificity, sensitivity, and precision. A ten-fold cross-validation method is preferred to separate test and training data. The highest accuracy was obtained using logistic regression at 78%. Heydari et al. [7] compared multiple classification algorithms for diabetes classification. The algorithms used are ANN, Bayesian classifier, Decision tree, SVM and K-NN. The best performance was obtained from ANN as 97.44%. With SVM, K-NN, Decision tree and Bayes classifiers, 81.19%, 90.85%, 95.03% and 91.60% accuracy were obtained, respectively. The dataset used for the system was collected from the University of Tabriz Medical Sciences and includes 2536 cases screened for type 2 diabetes.

Swapna et al. [8] developed a model to detect composite chronological features of ESA, LSTM, and input heart rate variability (HRV) data. SVM was applied to the detected features to classify the data. Compared to studies based on SVM, performance gains of 0.03% and 0.06% were obtained for the ESA and ESA-LSTM architecture, respectively. The study presented smart DVM model for diabetes diagnosis, where according to the authors, diabetes is a major health problem worldwide and if detected at an early stage, complications of type 2 diabetes can be reduced by up to 80%. In the proposed scenario, many data mining and machine learning algorithms are analyzed for diabetes prediction. Pima Indians Diabetes dataset from the National Institute of Diabetes and Digestive Diseases was used to evaluate the operation of the algorithm in Yuvaraj [9], where a new application of machine learning algorithms in Hadoop-based clusters is proposed for diabetes prediction. Random forest algorithm provided higher accuracy than decision tree and bayes algorithm in 4-node Hadoop cluster environment.

Wu et al. [10] proposed a new model for the detection of diabetes type-2 using K-means and logistic regression algorithms. The proposed methods consist of both set and class methods. The proposed methods increase the accuracy in predicting diabetes by 3%. The study developed a new diabetes classification model based on Convolutional Long Short-Term Memory (Conv-LSTM), three popular models such as ESA, traditional LSTM and Conv-LSTM were proposed, and the performances of the models were evaluated. In order to find the most suitable parameters for the applied models, hyperparameter optimization was carried out using the Grid search algorithm. In the first experiment, the dataset is split at different ratios for testing and training. In the first experiment, the Conv-LSTM-based model classified diabetes patients with the highest accuracy of 91.38%. In the second experiment, the cross-validation technique was applied. Using the cross-validation technique, the highest accuracy of 97.26% was obtained with the Conv-LSTM model. “Pima Indians Diabetes” data set and a methodology for diabetes prediction using different machine learning algorithms [2], the accuracies obtained by ANN, Bayes classifier, decision tree and deep learning are generally in the range of 90-98%. The best classification result was obtained from deep learning as 98.07%.

In the study, it was stated that the accuracy of the deep learning approach could be further increased by including ohmic data. An LSTM neural network application was developed for diabetes prediction in Massaro et al. [11]. A typical dataset used in the literature was used to validate the proposed model. The study focused on the procedure to be followed for an accurate LSTM data analysis using synthetic records. In addition, traditional ANN and LSTM approaches were compared. When the results of the LSTM approach were compared with the current ANN studies, an improvement of 6.5% was observed in the test set accuracy. Gill and Mittal [12] proposed a hybrid method for effective diabetes prediction, the proposed model was tested using the Pima diabetic dataset. In the first phase of the proposed hybrid model, the trait selection process was applied to select the most distinctive predictors reflecting the probability of developing diabetes. In the second stage, a two-layer classification was applied to the filtered data by combining the SVM and the neural network to increase the overall recognition rate of the model. An overall accuracy of 96.09% was obtained from the proposed hybrid model.

METHODS AND TECHNIQUES

Convolutional Neural Network (CNN)

Convolutional neural networks are in the group of multi-layered, feed-forward artificial neural networks that have been successfully applied in image analysis and computer vision, especially image object recognition. Convolutional networks are inspired by biological processes in the brain. The connectivity pattern between neurons is similar to the organization of the visual cortex [13]. ESAs are a neural network model that has proven to be highly effective in areas such as image recognition and classification [14]. ESA is a mathematical structure that typically consists of three types of layers. These layers are convolution, pooling and fully connected layers. The convolution and pooling layers perform the feature extraction, while the fully connected layers send the extracted features to the final output for classification.

Convolutions are necessary for a neural network to interpret pixels in an image as numerical values. The function of convolution layers is to convert the image into numerical values that the neural network can interpret and then extract the relevant patterns [15]. In the convolution layer, filters are passed over the input image. The convolution formula is given in equation 1. The "M" given in the equation represents the feature map, and the "w" refers to the (x,y) dimension convolution kernel.

$$M(i,j)=(R\times w)(i,j)=\sum_x \sum_y R(i-x,j-y)w(x,y) \quad (1)$$

After each convolution layer, the application of a nonlinear layer or activation layer is common in ESA. The purpose of this layer is to make a system that computes essentially linear operations nonlinear. The Rectified Linear Unit (ReLU) layer applies the function f(x) = max(0,x) to all incoming values. Basically, this layer changes all negative activations to 0. This layer increases the nonlinear properties of the model and the overall mesh without affecting the features from the convolution layer [16]. The necessary formula for ReLU is given in Equation 2.

$$\mathrm{Re}LU(x)=\max(x,0) \qquad (2)$$

The pooling layer is another building block of an ESA and is implemented after convolution and ReLU. Its function is to reduce the number of parameters and calculations in the network. The pooling layer operates independently of each feature map. The most common approach used in pooling is maximum pooling. Figure 1 shows an example of maximum pooling.

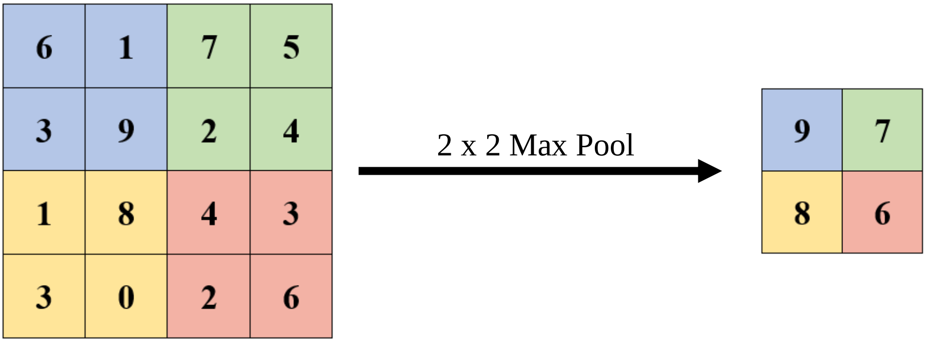


Figure 1. Maximum pooling

After the features that are subsampled by the convolution and pooling layers are created, these features are linked to the fully connected layer. Features are linked to one or more fully connected layers. Here, each input is connected to an output and each neuron has learnable weight [16]. The last fully connected layer typically has the same number of output nodes as the number of classes, and classification is done in this layer. Softmax is generally used, although different classifiers can be used in the last layer. The Softmax equation is given in 3. Softmax ensures that the output values of the neurons are in the range (0,1).

$$\operatorname{soft\,max}(x)j = \frac{e^{xi}}{\sum_{n=1}^{N} e^{x} n} \qquad \text{for } j = 1, \ldots\ldots, N \tag{3}$$

Long Short-Term Memory (LSTM)

Long Short-Term Memory networks are often referred to as “LSTM” and are a special type of Recurrent Neural Networks (RNN). The LSTM neural network has a complex structure called the LSTM cell in its hidden layer. The LSTM cell shown in Figure 2 has three gates, the input gate, the forget gate, and the output gate, which control the flow of information through the cell and neural network.

The LSTM model is arranged as a chain structure [17]. However, the repeating module has a different structure. Instead of a single neural network like a standard RNN, it has four interactive layers with a unique method of communication. A typical LSTM network consists of memory blocks called cells. Cell states are transferred to the next cell. The cell state is the main data flow chain that allows data to proceed essentially unchanged.

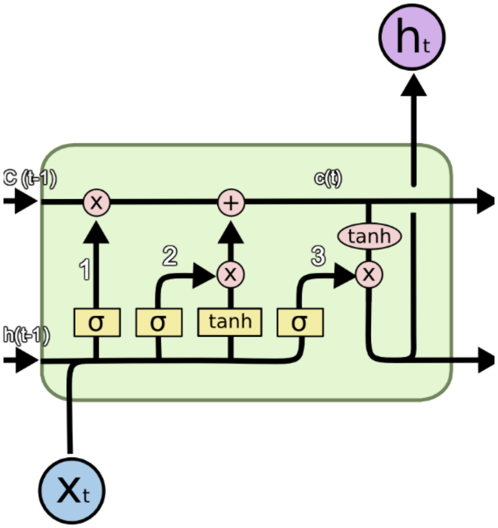


Figure 2. LSTM architecture (cell 1: forget; cell 2: store; cell 3: output)

However, some linear transformations may occur. Data can be added to or subtracted from the cell state via sigmoid gates. A gate is similar to a layer or a set of matrix operations containing different individual weights. The first step in creating an LSTM mesh is to identify the information that is not needed and that will be removed from the cell. This data definition and exclusion is decided by the sigmoid function, which takes the output of the last LSTM unit (h_{t-1}) at time t-1 and the current input (X_t) at time t. In addition, the sigmoid function determines which part of the old output should be eliminated. This gate is called the forget gate (or f_t); where h_{t-1} is a vector ranging from 0 to 1 and corresponding to each number in the cell state, C_{t-1}.

$$f_t = \sigma(W_f\,[h_{t-1}, X_t] + b_f) \tag{4}$$

Where, σ is the sigmoid function and Wf and bf are the weight matrices and bias of the forget gate, respectively.

The following step decides and stores the information from the new entry (X_t) and also updates the cell state. This step consists of two parts, the sigmoid layer and the second tanh layer. First, the sigmoid layer decides whether to update new information (0 or 1), and second, the tanh function decides on the importance levels (-1 to 1), giving weight to the passed values. The two values are multiplied to update the new cell state. This new memory is then added C_{t-1} to the old memory, resulting in C_t.

$$i_t = \sigma(W_i\left[h_{t-1}, X_t + b_i\right]) \tag{5}$$

$$N_t = \tanh(W_n\left[h_{t-1}, X_t\right] + b_n) \tag{6}$$

$$C_t = C_{t-1} f_t + N_t i_t \tag{7}$$

Where, C_{t-1} , C_t , *t-1* and *t* are the cell states at times. The sigmoid layer decides which parts of the cell state will reach the output. Next, the output of the sigmoid gate (O_t) is multiplied by the new value generated from the cell state (C_t) by the tanh layer.

$$O_t = \sigma(W_0\left[h_{t-1}, X_t\right] + b_0) \tag{8}$$

$$h_t = O_t \tanh(C_t) \tag{9}$$

Where, W_0 and b_0 are the weight matrices and bias of the output gate, respectively.

Proposed Method

The proposed method for the detection of diabetes consists of three steps. First, the features obtained from the samples in the data set were passed through the designed ESA. Then, the features obtained from ESA were given as input to LSTM. Finally, the samples in the data set were classified using the Softmax classifier. The proposed method is given in figure 3. In addition, the same experiments were performed using ESA and LSTM separately.

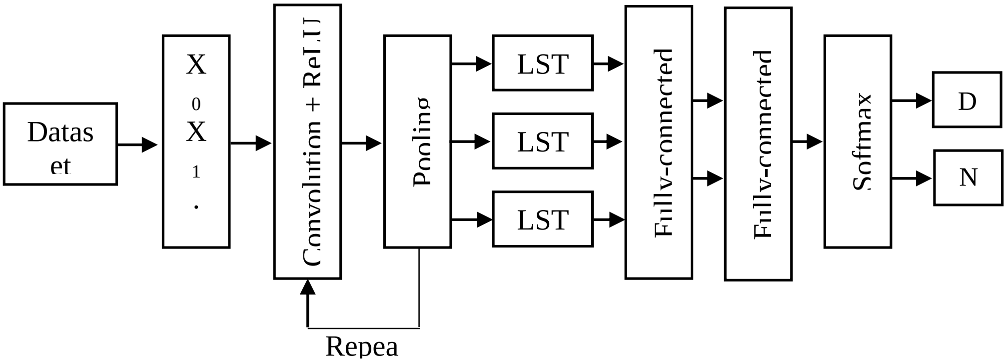


Figure 3. Proposed method process (D: Diabetic; ND: Not Diabetic)

The topology of the designed ESA consists of 1 input layer, 2 convolution layers, 2 pooling layers, 2 fully connected layers and 1 output layer. There are 6 layers in total in the designed architecture. There are 128 convolution filters in the first convolution layer and 64 convolution filters in the second convolution layer. ReLU was used as the activation function. In addition, after each convolution and ReLU operation, a maximum pooling of 2x1 was made. After the convolution layers, the data was passed through the LSTM layer. Finally, the model is completed by completely connecting the connection layers and Softmax. The designed ESA architecture is given in Table 1.

Table 1. ESA architecture of the model

Name of Layer	Type	Specifications
Input	Input vector	
conv1	Convolution	128 convolution filters with the same size
relu1	ReLU	
pool1	Maximum pooling	2 x 1
conv2	Convolution	64 convolution filters with the same size
relu2	ReLU	
pool2	Maximum pooling	2 x 1
fc1	Fully connected layer	1024 neurons
relu7	ReLU	
drop7	Dropout	50%
fc2	Fully connected layer	512 neurons
Output	Softmax	2. Class

EXPERIMENT AND FINDINGS

Utilized Dataset

The performance of the proposed model was evaluated using Pima Indians Diabetes, a popular dataset from the UCI machine learning repository [18]. The dataset includes records of 768 female patients aged at least 21 years, 268 of whom were diabetes positive, and the rest were diabetes negative. The dataset has eight variables, such as Number of Pregnancies, Glucose, Blood Pressure, Body Mass Index, Skin Thickness, Insulin, Diabetes Pedigree, and Age, and a target variable called outcome, to predict diagnostically whether a patient has diabetes. The description of the data set is given in Table 2.

Table 2. Features of Pima Indian utilized dataset

Feature	Explanation	Limits
Diabetes category	Binary system with zero indicating no diabetes (n=259) and one indicating diabetes (n = 483).	0 or 1
Age	Age of data providers	22 - 79
Pedigree tree of diabetes	A feature used in prognosis of diabetes	0.065 – 2.49
Insulin level	Blood insulin level averaged after measuring it every 2 hours	40 - 832
Thickness of skin	Measured through a skinfold method	51 - 106
Body Mass Index (BMI)	A body quality index calculated by dividing the body weight in kilograms by its height in meters.	17.2 – 56.8
Blood pressure	Diastolic reading	69-142
Glucose level	Oral glucose concentration test	0-185
Pregnancy	Number of pregnant data providers	13

Findings

The performance criteria of the proposed method are based on accuracy, precision, and F-score ratios.

- False positives (FP): samples from the negative class, predicted as positive.
- False negatives (FN): samples predicted as negative with true class positive.
- True positives (TP): correctly predicted samples belonging to the positive class.
- True negatives (TN): correctly predicted samples belonging to the negative class.

$$Accuracy = \frac{|TN| + |TP|}{|FN| + |FP| + |TN| + |TP|} \tag{10}$$

Precision (P) is a measure that estimates the probability that a positive prediction will be correct. The precision measurement is given in Equation 11.

$$Percision(P) = \frac{|TP|}{|TP| + |FP|} \tag{11}$$

The F-score is a congruent mean of positive predictive rate and sensitivity measures and is calculated as shown in Equation 12.

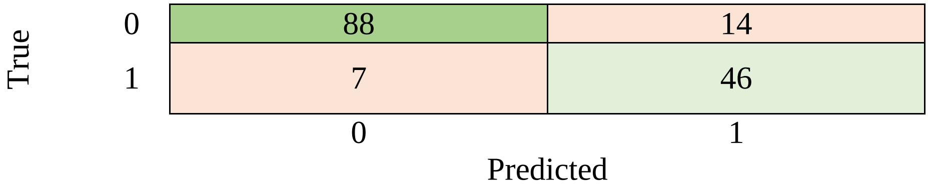
$$F-score=\frac{2\times|TP|}{2\times|TP|+|FP|+|FN|} \tag{12}$$

Learning rate, epoch number and minibatch size are the most important hyper parameters that affect the performance of ESA. In the designed architecture, the minibatch size is 64, the maximum epoch number is 20, and the learning rate is set to 1e^{-4}. In order to determine the effect of the size of the data allocated for training and testing on the performance of the network, this data was divided into two different ways. In the first experiment, 70% of the data is used for training and 30% for testing; In the second experiment, 80% of the data was used for training and 20% for testing.

Table 3 shows the classification results obtained from ESA, LSTM and ESA+LSTM (combined). By dividing the data by 70%-30% for training and testing, 82.5% accuracy was obtained from ESA, 83.8% from LSTM and 85.2% from ESA+LSTM. By dividing the data by 80%-20% for training and testing, 83.3% accuracy from ESA, 85.2% from LSTM and 86.5% from ESA+LSTM were obtained. Subsequently, the confusion matrix of the best classification results is shown in Figure 4.

Table 3. Results of the two types of classification applied to the data

Parameter	70% training / 30% testing			80% training / 30% testing		
	ESA	LSTM	Combined	ESA	LSTM	Combined
Accuracy	82.5	83.8	85.2	83.3	85.2	86.5
Precision	83.1	84.5	84.9	83.6	85.3	87.0
F-score	83.6	84.2	58.1	83.7	85.4	88.2



	0	1
0	88	14
1	7	46

Figure 4. The best obtained confusion matrix from the proposed model

The results obtained from other methods used in the literature were compared with the findings we obtained for performance evaluation. Some important studies are given in Table 4. The comparison with the applied model in the literature suggests that the proposed method outperformed the other three methods: deep learning architecture, LSTM, and firefly algorithms (FA).

Table 4. Comparison of the proposed model with other models suggested in the literature

Study	Accuracy
Current Research	86.4%
Massaro et al. [11]	86.0 %
Haritha et al. [19]	81.0%
Ashiquzzaman et al. [20]	81.0%

CONCLUSIONS

The current research aimed to develop a prediction model for the measurement of diabetes. A large proportion of the human population suffers from diabetes. The application of the developed model was made on the Pima Indian Diabetes dataset. ESA and LSTM architectures are used together in the developed model. During the experiment, the features in the data set were given as input to the designed ESA+LSTM model. Softmax is embedded in the last layer of the architecture designed for classification. In addition, the same experiments were repeated using the ESA and LSTM architectures separately. The data set was divided into 70% training, 30% testing in the first experiment, and 80% training and 20% testing in the second experiment. An overall accuracy of 86.5% was obtained from the proposed model combining LSTM and ESA. 83.3% and 85.2% accuracy were obtained from ESA and LSTM, respectively. Comparing the developed model with the state-of-the-art models shows the effectiveness of the proposed model in diabetes prediction. In future studies, it is planned to help healthcare professionals in the early diagnosis of diabetes and to develop applications using deep learning algorithms.

REFERENCES

[1] H. M. Deberneh and I. Kim, "Prediction of Type 2 Diabetes Based on Machine Learning Algorithm," *International Journal of Environmental Research and Public Health,* vol. 18, 2021.

[2] Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju and H. Tang, "Predicting Diabetes Mellitus With Machine Learning Techniques," *Frontiers in Genetics,* 2018.

[3] A. T. Kharroubi and H. M. Darwish, "Diabetes mellitus: The epidemic of the century," *World Journal of Diabetes,* vol. 6, no. 6, p. 850–867, 2015.

[4] K. E. Kreider, "The Diagnosis and Management of Atypical Types of Diabetes," *Journal of Nurse Practitioners,* vol. 15, no. 2, pp. 171-176, 2018.

[5] S. Wild, C. Fischabacher and J. McKnight, "Using Large Diabetes Databases for Research," *Journal of Diabetes Science and Technology,* vol. 10, no. 5, pp. 1073-1078, 2016.

[6] A. K. Dwivedi, "Analysis of computational intelligence techniques for diabetes mellitus prediction," *Neural Computing and Applications,* vol. 30, pp. 3837-3845, 2018.

[7] M. Heydari, M. Teimouri, Z. Heshmati and S. M. Alavinia, "Comparison of various classification algorithms in the diagnosis of type 2 diabetes in Iran," *International Journal of Diabetes in Developing Countries*, vol. 36, pp. 167-173, 2016.

[8] G. Swapna, R. Vinayakumar and K. P. Soman, "Diabetes detection using deep learning algorithms," *ICT Express*, vol. 4, no. 4, pp. 243-246, 2018.

[9] N. Yuvaraj and K. R. SriPreethaa, "Diabetes prediction in healthcare systems using machine learning algorithms on Hadoop cluster," *Cluster Computing,* vol. 22, pp. 1-9, 2019.

[10] H. Wu, S. Yang, Z. Huang, J. He and X. Wang, "Type 2 diabetes mellitus prediction model based on data mining," *Informatics in Medicine Unlocked,* vol. 10, pp. 100-107, 2018.

[11] A. Massaro, V. Maritati, D. Giannone, D. Convertini and A. Galiano, "LSTM DSS Automatism and Dataset Optimization for Diabetes Prediction," *Applied Sciences*, vol. 9, no. 17, 2019.

[12] N. S. Gill and P. Mittal, "A computational hybrid model with two level classification using SVM and neural network for predicting the diabetes disease," *Journal of Theoretical and Applied Information Technology*, vol. 87, no. 1, pp. 153-157, 2016.

[13] Y. Fu and C. Aldrich, "Flotation froth image recognition with convolutional neural networks," *Minerals Engineering*, vol. 132, pp. 183-190, 2019.

[14] M. Ibbotson and Y. J. Jung, "Origins of Functional Organization in the Visual Cortex," *Frontiers in Systems Neuroscience*, 2020.

[15] D. C. CiresAn, U. Meier, J. Masci, L. M. Gambardella and J. A. Schmidhuber, "Flexible, high performance convolutional neural networks for image classification," in *IJCAI'11: Proceedings of the Twenty-Second international joint conference on Artificial Intelligence*, Barcelona, 2011.

[16] N. Wang, Y. Zhang, Z. Li, Y. Fu, W. Liu and Y. G. Jiang, "Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images," *Lecture Notes in Computer Science*, pp. 55-71, 2018.

[17] G. V. Houdt, C. Mosquera and G. Napoles, "A review on the long short-term memory model," *Artificial Intelligence Review*, vol. 53, pp. 5929-5955, 2020.

[18] UCI Machine Learning, "Pima Indians Diabetes Database: Predict the Onset of Diabetes Based on Diagnostic Measures," Kaggle, 06 October 2016. [Online]. Available: https://www.kaggle.com/uciml/pima-indians-diabetes-database. [Accessed 05 January 2021].

[19] R. Haritha, D. S. Babu and P. Sammulal, "A Hybrid Approach for Prediction of Type-1 and Type-2 Diabetes using Firefly and Cuckoo Search Algorithms," *International Journal of Applied Engineering Research*, vol. 13, no. 2, pp. 896-907, 2018.

[20] A. Ashiquzzaman, A. K. Tushar, R. Islam and J. M. Kim, "Reduction of Overfitting in Diabetes Prediction Using Deep Learning Neural Network," in *7th iCatse International Conference on IT Convergence and Security*, 2017.