Journal of Basrah Researches (Sciences) **49**(2), 171(2023)
DOI: https://doi.org/10.56714/bjrs.49.2.14

Development of A Sequential CNN Model With Three Hidden Layers For Diabetes Prediction

Haneen S. Al- Majidi, Abbas Mgharbel

Department of Computer Science, College of Science and Literature, American University of Culture and Education, Lebanon.

ARTICLE INFO

Received 20 August 2023
Accepted 07 December 2023
Published 30 December 2023

Keywords:

Deep Learning, Diabetes, Health Care, Convolutional Neural Network, Disease prediction.

Citation: H. S. Al- Majidi, A. Mgharbel, J. Basrah Res. (Sci.) **49**(2), 171 (2023). DOI:https://doi.org/10.56714/bjrs.49.2.14

ABSTRACT

The early diagnosis and treatment of diabetes can contribute to the mitigation of associated risks and effects. Therefore, it is imperative to anticipate and identify the ailment at an early stage through the utilization of dependable procedures that can offer forecasts characterized by a substantial level of dependability and precision. This study utilizes the Diabetes Readmission Dataset, comprising 101,766 records and encompassing 50 features. After selecting the most pertinent features, the dataset is partitioned into a training set and a test set. Subsequently, a sequential model employing deep learning, specifically a convolutional neural network (CNN) with three hidden layers, is constructed for prediction purposes. The correctness of the model was assessed through the utilization of performance testing metrics, resulting in a recorded accuracy rate of 99.53%. The findings of this study have the potential to inform the development of personalized treatment approaches that address the individualized requirements of patients, hence enhancing the quality of healthcare provide.

1. Introduction

Diabetes is a chronic metabolic disorder that is characterized by increased levels of glucose in the blood. This can occur when the pancreas produces little or no insulin (as in type 1 diabetes), when cells become resistant to the effects of insulin (as in type 2 diabetes), or when both of these factors are present. Over ninety percent of confirmed cases of diabetes worldwide are cases of type 2, with the remaining ten percent attributable to other types, such as type 1 and gestational diabetes [1].Diabetes is a non-communicable disease that is chronic and does not spread from person to person. Diabetes has consistently ranked as the leading cause of death on a global scale [2]. This could affect an additional 629 million individuals by the year 2045, representing a jump of 48 percent. Chronic hyperglycemia caused by diabetes puts a number of organs at risk for long-term damage, dysfunction, and even failure [3]. These organs include the heart, blood vessels, eyes, kidneys, nerves, and feet. Early diagnosis and treatment of diabetes can improve health outcomes and lessen the risk of complications for patients. Conventional methods, such as clinical threat record and biomarker analysis, are now in a less than ideal state for routine screening and prevention [4]. The low precision, expense, and intrusiveness of these methods all fall short of ideal.There is a lot of hope for this area

*Corresponding author email : hanwa.almajidi@gmail.com

ISSN: 1817-2695 (Print); 2411-524X (Online)
Online at: https://jou.iobrs.edu.iq

because of the success of deep learning models and other machine learning methods. A patient's medical history, laboratory results, and imaging studies can all be analyzed by these algorithms to predict the likelihood that a patient would get a particular disease [5]. These prediction methods aid doctors in making early diagnoses, which leads to quicker treatment and better health outcomes for patients. The examination can also be used to identify individuals who are at a high risk and who would profit from receiving preventative therapies or following-up appointments more frequently. Models of machine learning shouldn't be considered as a replacement for clinical expertise; rather, they should be seen as a tool to complement and improve upon clinical expertise. In order to guarantee an accurate diagnosis and the most appropriate therapy, it is essential to make use of any prediction tool in conjunction with the knowledge and expertise of the doctor. The tool may assist minimize costs associated with healthcare by minimizing the likelihood that more expensive treatments may be required in the event that an illness is not detected in a timely way, this method can save money.The research area provided a detailed evaluation of the most recent trend in deep learning technologies for diabetes research [6]. A comprehensive search was conducted, a collection of articles was chosen, and a summary of the most important material was written with a concentration on three topics: the diagnosis of diabetes, the management of glucose levels, and the identification of complications associated to diabetes. In these areas, a variety of different deep neural networks (DNN) architectures and learning algorithms have been deployed, and they have obtained better experimental performance than previous approaches that were more conventionally based on machine learning. On the other hand, various problems have been recognized as a result of reading the relevant literature. These challenges include the availability of data, the processing of features, and the interpretability of models. There is a significant possibility that these obstacles can be overcome by using the most recent developments in deep learning technology to enormous amounts of data that contain multiple modalities. Because of this, it is essential to carry out a large-scale study with the objective of developing and perfecting deep learning models for diabetes prediction. These models should take into account both clinical and demographic data. It is necessary to explain not only the clinical value of the models and their implications for the treatment and prevention of diabetes [7], but also the performance of the models based on a number of different metrics. Not only is it vital to discuss the clinical value of the models, but it is also important to address their implications for the treatment and prevention of diabetes. This research contributes to healthcare:

- To early detection and prompt treatment including preventing the illness from worsening.
- To help doctors come up with care plans that are tailored to each individual patient.
- Better diabetes management leads to lower healthcare costs.
- Enhancing patients' quality of life, reducing stress, and providing them with a sense of control over their own health are all goals of integrative medicine.

With the help of clinical and demographic data, the overarching purpose of this research is to evaluate whether or not deep learning can successfully be applied to the difficulty of diabetes prediction and to establish whether or not doing so is advantageous. In addition to this, it will make a contribution to the expanding field of research on the application of artificial intelligence in the medical care industry, which is a subject that will be advanced as a direct outcome of this inquiry.

2. State of the Art

A wide range of supervised learning algorithms are applicable in the medical field. Predicting patient outcomes, identifying disease risk factors, and improving diagnostic accuracy are all possible with the help of these algorithms. The use of predictive analytics in healthcare is difficult, but the positive effects on patient outcomes and the ability to make prompt treatment decisions are well worth the effort. The authors of [8] tested six different machine learning algorithms for their accuracy in diabetes diagnosis to investigate the potential of predictive analytics in healthcare. Two limitations of this investigation were the size of the dataset and the lack of some attribute values. Researchers in [9] developed a technique to distinguish between HRV signals from diabetics and those who are not afflicted with the disease. By combining convolutional neural networks with long short-term memory (LSTM). A support vector machine (SVM) was utilized for the purpose of attribute classification.The

authors of [10] created a model for early diabetes detection using three distinct categorization techniques that can be applied using machine learning: decision trees (DT), support vector machines, and naive bayes (NB). Data from the Pima Indians Diabetes Dataset (PIDD) were used in a number of different investigations. In order to improve the accuracy of the deep neural network's diabetes detection capabilities, it was trained using the same method outlined in [11], which included both 5- and 10-fold cross-validation during the training process. A cross-validation sample size of ten was utilized, and the results showed an accuracy of 97.11 percent.

The authors of [12] develop a framework that, by analyzing and contrasting a number of different machine learning algorithms, is capable of determining the likelihood that individuals suffer from diabetes with the highest degree of precision feasible. This is done in order to provide the most accurate diagnosis possible.Electronic health records, images, omics data, and text pose substantial obstacles for healthcare companies seeking to glean information and insight from this data. By using the most recent developments in machine learning technology to uncover previously unseen patterns, an earlier diabetes diagnosis may be possible. According to research [13], the accuracy of functional classifiers such Artificial Neural Network (ANN), NB, DT, and deep learning (DL) can be improved to around 90%-98%.A recently categorized diabetes dataset from South Asia (Bangladesh) was supplied by researchers in [14]. They also suggested using a weighted ensemble of machine learning classifiers. Hyperparameter optimization using a grid search was used to fine-tune the critical hyperparameters. Feature selection, K-fold cross-validation, and missing value imputation were also a part of the framework's design.The DNN was used in the model given in [15], and dropout regularization was used to prevent overfitting. To do this, the job was recast as a classification issue. Researchers optimized several parameters and used the binary cross-entropy loss function to create a deep neural network-based prediction model with good accuracy. The results of [16] prompted the idea that diabetes prediction could benefit from the application of machine learning expertise. An example of this is shown in J48 decision tree computation, also known as C4.5, is a decision tree algorithm that recursively partitions data into subsets to make binary decisions in a tree-like structure. The dataset that was analyzed had information from 768 different people, and each of those people had major 8 characteristics as well as a target column that indicated whether or not they were positive or negative. Weka was used to carry out the experiment, and the results that were presented reveal that the decision tree J48 calculation is significantly more effective despite requiring a significantly shorter amount of processing time. It is vital to evaluate the accuracy of each supervised classification method in order to compare the effectiveness of algorithms that work well on small data sets with those that function well on large data sets. The research [17] examines the outcomes of assessments performed by each program using correctly classified, wrongly classifieds, precision, and recall as measurement variables. The researchers in [18] focused on using a feedforward neural network, often known as FNN, to predict type 2 diabetes. There was evidence that estimation inaccuracy was caused by the effects of a variety of activation functions, hidden layer counts, as well as the hidden layer counts themselves. As a consequence of the unjustly increasing prevalence of diabetes mellitus, an increasing number of households are being unfairly impacted by the disease. Before being diagnosed with diabetes, the vast majority of diabetics have only a cursory understanding of the factors that put them at risk. The purpose of the research [19] is to investigate the application of data mining strategies for the early prediction of diabetes. The extraction of useful data in order to gain a deeper understanding of diabetes is a critical issue. The processes and techniques of data mining will be used to find the right procedures and methods for the accurate classification of the diabetes dataset and the extraction of valuable patterns. This will be accomplished by combining the two aspects of data mining. The risk of developing diabetes was calculated by the use of bioinformatics and other medical databases in the research [20]. In order to make the diagnosis of diabetes, mining techniques and the WEKA program were utilized.The researchers in [21] demonstrated that the XGBoost model may be applied to the screening of individuals at an early stage who are at an increased risk of developing type 2 diabetes, which is the ideal situation for the prevention and management of diabetes. The core qualities may be helpful in the development of diabetes prevention strategies that are more targeted. Data overfitting in the predictive model used to predict the occurrence of diabetes was reduced through the use of a dropout strategy, as recommended in [22]. In order to improve the accuracy of predictions, the suggested method used a novel type of deep neural network for diabetes prognosis. The proposed

method was tested on the PIDD dataset, where it performed with an accuracy of 88.41%.In the research [23], the authors presented a DNN framework for the categorization of diabetes-related data by employing stacked autoencoders. The DNN is constructed with stacked autoencoders and a softmax classifier that are cascaded together. A diabetic dataset consisting of Pima Indians was used for the training procedure of the model. The accuracy of the model was calculated to be 86.26%. However, the lack of patients and attributes is the main obstacle to using deep learning on the PIDD dataset. The trained models must be validated on a large population dataset to demonstrate the generalization of DNNs. In order to learn the representations from a large-scale dataset, the researchers [24] suggested a system called Deep Patient. Classification for diabetes diagnosis was 0.907% AUC. A large dataset consisting of 11,456 participants was used in [25]. Undiagnosed diabetes was detected with an AUC of 80.11% using a DNN model for screening. In addition, some research show promise for a non-invasive method of detecting diabetes. To assess the biomarkers in real-time breath signals for diabetes diagnosis and classification, research [26] presented a one-dimensional (1-D) CNN architecture. MOS sensors were used to gather the breath samples for further analysis of the VOCs. The sensor array took readings every 1000 seconds from a tiny gas chamber. After that, the signals were further processed by the CNN classifier, which, in comparison to principal component analysis (PCA), support vector machine, and singular value decomposition algorithm (SVD), can eliminate the requirement for feature selection and maximize the overall performance.The findings of the research [27] show the inadequacies of the approaches that are currently used, which make them less acceptable for using in diabetes diagnosis. The field of artificial intelligence as well as the field of medical research both make extensive use of deep learning. Because deep learning and other algorithms can increase the accuracy of diabetes diagnosis, they should be used to diagnose a variety of cardiac illnesses. One of these conditions is diabetes.The diabetes research community has adopted a substantial number of deep learning algorithms, covering a variety of architectures in supervised learning and unsupervised learning respectively. In especially for issues pertaining to clinical imaging, the most common use for these is in the form of CNN-based systems. CNNs are effective at extracting features from raw data, and they do so with minimal need for hand-engineered effort or specialized knowledge in the field of image processing [28]. As a result, the primary use of CNNs has been in the analysis of clinical scans and other types of medical pictures that are utilized in the process of diagnosing issues associated with diabetes. One other application for CNN-based architectures is to assist people living with diabetes in their day-to-day lives by determining the levels of macronutrients in food pictures.Due to the widespread use of the Pima Indians Diabetes Dataset, we want to expand upon the findings of past studies. We'll be working with a larger dataset for this project. Due to the simplicity of the prior models, increasing their complexity may result in better results. Some research findings were also inadequate, so we'll be working to improve upon them so that they can be applied more broadly.

3. Proposed Approach

In order to enhance the quality of the prediction, we will first perform an analysis on the dataset, then explain the proposed network architecture, and then implement the feature selection strategy. The procedures required to be taken in order to construct our proposed system are shown in Fig. 1.

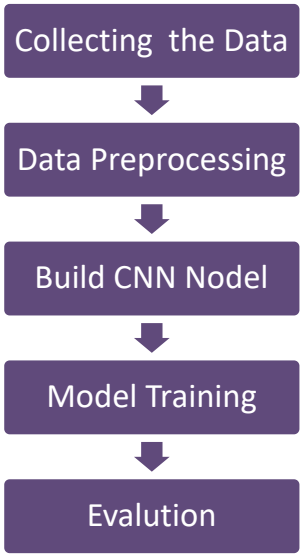

Fig. 1. The steps of the proposed system.

3.1. Collecting the Data

A data collection referred to as the Diabetes Readmission Dataset [29] provides information that is helpful regarding patients. The dataset contains 101,766 observations and 50 features, some of which include specifics of the patient's traits, medical history, and medications. Numerous studies involving diabetes and other aspects of healthcare that include machine learning and data analysis can make use of this dataset. The dataset, which was extracted from the electronic health record system of a big academic medical facility, is open to the use of anyone interested in conducting research on the topic. A bit field has been designated as the collection's target variable.Using this information in a model can be accomplished in a myriad of different ways. The dataset is composed of 29 different attributes, including information on patients' demographics, medical histories, diagnoses, prescriptions, and hospitalizations, as well as their medical histories and medical histories of their family members.

3.2. Data Preprocessing

In the beginning, we will examine the data to see if it contains any null values. Examining whether or not there are any null values is an essential part of the process of processing the data. It's possible that incomplete data or issues with data gathering could be indicated by null values. The correct treatment of null values is absolutely necessary in order to keep the analysis free of bias and mistakes. Utilizing the Pandas library that is available in Python is one strategy that may be used to search the dataset for null values. When a dataset has been loaded into a Pandas data frame, the isnull() method is able to locate any invalid values that may be contained within the dataset. This can be done after the dataset has been put into the Pandas dataset frame.The next step is to put into action the feature importance approach, a method for selecting features. Tree-based classifiers inherently group features by their relative importance. Extra Trees classifier [30] that extremely randomizes trees (Tree-based) classifiers use the results of multiple decision trees to determine the importance of a feature. The random forest (RF) classifier is distinct from the decision trees, which are constructed within the random forest itself. Additional trees in the forest are symbolic representations of individual decision trees which were produced by the primary training model. The trees are then presented with a random k feature from the set at each test node and tasked with selecting the most appropriate split criterion (often a genetic index) for the data. Using a random selection of attributes yields a large number of decision trees with no underlying pattern. We experimented with varying the number of features, however the database contains several that are useless for diabetes prediction and hence harmed the accuracy of our models. Due to the significance of these traits, the optimal number

of features was 9. That's why we settled on that total.We pick the best nine features as shown in Fig. 2 from the whole dataset. Then, using the StandardScaler method from the sklearn library in Python, the values of the features of interest can be standardized. We put 70% of the samples and targets through training, and just 30% of them go through testing.

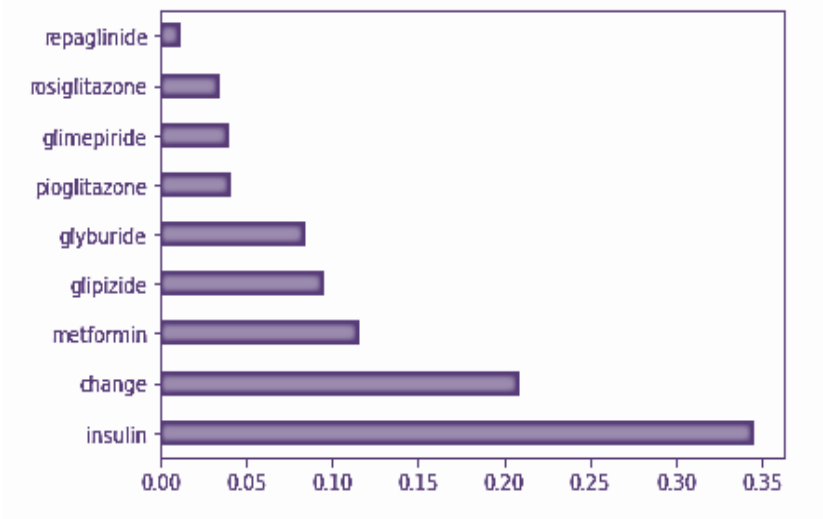

Fig. 2. Top nine features in the dataset.

3.3. Methodology

Unnecessary spikes in blood sugar might have severe consequences for one's health. Prediabetes can progress into full-blown diabetes if the necessary preventative measures aren't taken. As part of our study, we apply criteria derived from the dataset to make diabetes diagnoses. Accurate prediction of diabetes can help in therapy and management of the disease. High-risk individuals can be counseled on how to make behavioral changes, such as exercising more and eating healthier, that will lower their chances of developing diabetes. Furthermore, earlier management and medication for diabetes can improve outcomes and reduce the risk of complications from the condition, making early detection all the more important. we developed the model with 3 hidden layers. The building of the model is shown in Fig. 3. The input to the designed model is batches of samples or records, each containing 9 features.

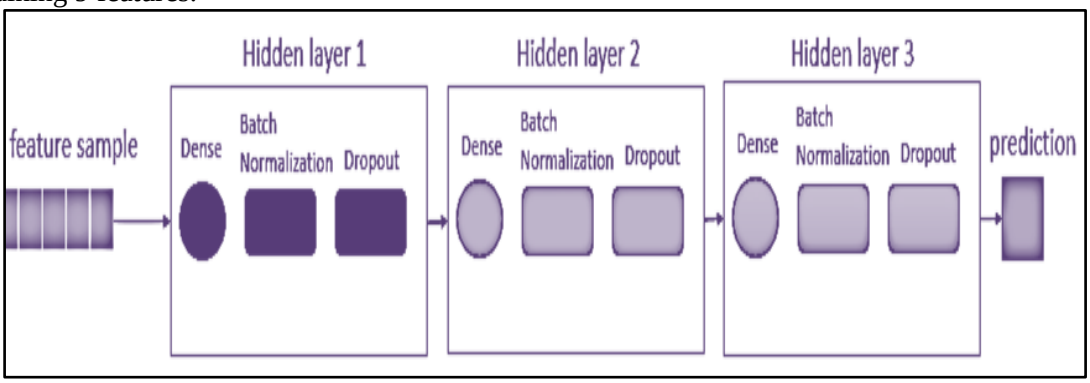

Fig. 3. Proposal CNN architecture and layers.

3.3.1. The Dense Layer

Because it is responsible for bridging the gap between the output of one layer and the input of the succeeding layer, the dense layer is a crucial component of deep learning models. As a direct consequence of this, the layer's output has been wired up to the layer's input [31]. This layer is considered to be fully linked since every neuron in the layer above it in the hierarchy is connected to every neuron in the layer below it in the hierarchy. Dense layers are composed of groups of neurons, with each neuron in the group getting input from all of the other neurons in its sublayer and sending a

single value as output to all of the neurons in its superlayer. Dense layers are also known as layers with high neuron density. Dense layers are also known as neural networks. The output of a neuron can be thought of as the weighted sum of its inputs, and these weights and biases are learnt throughout the course of training. Applications that typically use dense layers, among the many functions carried out by neural networks are image classification, natural language processing, and recommendation system operation. They are capable of being layered upon one another to create deep neural networks, which have the ability to learn complex representations from the input they are given.

3.3.2. The Batch Normalization Layer

Since it's one solution to overfitting, this layer plays a crucial role in the training process. When a model produces accurate predictions based on the data it was trained on, but not when applied to new data, it has overfit, which is an undesirable learning tendency. This occurs when the model is allowed to draw inferences from its own past performance. Batch normalization is a method for improving the stability and performance of neural networks. The incoming inputs are subjected to standardization and settlement processes. Since the inputs to a neural network are typically collected from the training dataset in batches, rather than all at once, the settlement procedure for those batches also typically occurs in batches [32]. In this sense, the concept of "Batch in Batch" normalization makes sense.

3.3.3. The Dropout Layer

Getting the desired output from a neural network requires removing nodes from both the input layer and the hidden layer [33]. When a node is removed from a network, its connections both forward and backward are temporarily cut, creating a new network structure. Nodes in a graph are eliminated when the dropout probability reaches “p”.

“p” represents the dropout probability, which is a hyperparameter typically set to a specific value between 0 and 1. This probability determines the likelihood of any given node (neuron) being "dropped out" or temporarily removed from the network during training.The overfitting problem is significantly reduced thanks to the dropout layer, which is its principal contribution. As a result of dropout, neurons are unable to fix errors that have been created by other neurons, which in turn impairs co-adaptation. This is due to the fact that the existence of neurons is very ambiguous throughout each cycle. Therefore, by losing neurons at random, it promotes the layers to adopt a probabilistic method for handling the inputs. Overfitting can be avoided and the model's scope expanded in this way.

3.3.4. Proposed Algorithm

This algorithm will build a neural network model for binary classification with appropriate activation functions, regularization techniques (dropout and batch normalization), and configuration for training.

- model = Sequential (): Initialize a sequential neural network model.
- model.add(Input(shape=(9,))): Input layer with 9 nodes is added to the model.
- Hidden Layers: Three hidden layers are added to the model sequentially. Each hidden layer consists of the following components:
  model.add(Dense(64, activation='relu')): This line adds a dense (fully connected) layer with 64 units and ReLU (Rectified Linear Unit) activation function.
  model.add(BatchNormalization()): Batch normalization is used to normalize the activations of the previous layer, which can help stabilize training and improve convergence.
  model.add(Dropout(0.2)): Dropout is a regularization technique that randomly sets a fraction (in this case, 20%) of the neurons in the layer to zero during each training batch. This helps prevent overfitting.
  model.add(Dense(1, activation='sigmoid')): Adds the output layer with a single node and a sigmoid activation function.
- model.compile(optimizer=opt, loss='binary_crossentropy', metrics=['accuracy']): This step configures the training process for the model.

optimizer: The optimizer aims to minimize the loss function by adjusting the model's weights.
loss: It defines the loss function that measures the error between the predicted values and the actual target values during training. For binary classification problems, 'binary_crossentropy' is commonly used.
metrics: The metrics parameter is a list of evaluation metrics that you want to monitor during training include 'accuracy'.

- Use the model.fit method to train the model for 10 periods, with a batch size of 16.

3.3.5. Model Layers

The model starts with a sample batch fed into its input layer. There are 9 characteristics in each batch sample. They're placed inside a dense layer with 64 cores and a relu type activation mechanism. Due to the number of kernels in the preceding dense layer, its output is 64 characters long, and its values fall within the range [0, inf] since the relu function rounds down negative attribute values to zero and rounds up positive values. BatchNormalization is a layer that receives the dense layer's output. Overfitting is mitigated by a scaling of feature values in this layer. The output of the BatchNormalization layer is fed into the Dropout layer, which causes the nodes to have a probability of being deleted that is one-quarter as high as the original value. After that, the output is sent on to a dense layer that is equipped with 32 cores and an activation function of the relu type. The output of the dense layer that came before it is in the range [0, inf] because of the relu function, and it contains 32 different values as a result of the number of kernels that are included within it. The data is taken as input by BatchNormalization when it has been received. This layer helps prevent overfitting by standardizing the values of the features. The output of the BatchNormalization layer is inserted into the Dropout layer, which results in the nodes having a probability of 0.2 of being eliminated. A dense layer with 16 cores and a RELU-type activation function is the one that pulls in the output signal. This layer is known as the receiver. Because of the relu function, the output of the dense layer that came before it can take on one of sixteen possible values (one for each of the kernels). BatchNormalization receives the data and uses it as input. By normalizing feature values, this layer prevents overfitting. The nodes are removed with a probability of 0.2 by inserting the output of the BatchNormalization layer into the Dropout layer. The output is represented by a single nucleus in the final dense layer. A sigmoid activation function is also included, limiting the range of the prediction to the interval [0, 1].

4. Evaluation the results

4.1. Metrics for evaluation

In order to evaluate the performance of the model [34], we applied a variety of criteria. We will provide you with an overall summary of them.

4.1.1. Confusion matrix

After we have gained an understanding of the confusion matrix that they are predicated on, we will be able to compute the metrics that serve as their foundation. Using a confusion matrix is one way that researchers in the area of machine learning can determine how effective different classification algorithms are, which can provide a summary of the results. This evaluation can be done in order to determine how well the models perform. If the model is capable of multiclassification, the confusion matrix will take the shape of a n^2 matrix, where n is the total number of targets that are being considered. Compiling the model's test-set predictions and evaluating how well they perform in relation to the actual ground-truth labels of the dataset is how it accomplishes this goal. When talking about a confusion matrix, the following terminology is utilized:

- TP: The number of observations that were correctly classified as positive by the model.
- FP: The number of times a model incorrectly classifies negative input as positive.
- TN: The fraction of data points that were correctly labelled as "negative" by the model.
- FN: The number of observations that were incorrectly predicted by the model to be negative.

4.1.2. Accuracy

The "accuracy" of a data-labeling algorithm is the percentage of times it gets the label right. A prediction's accuracy is measured by how many instances it correctly predicted.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

4.1.3. Precision

The frequency with which favorable results were forecast and actually experienced. False positives are reduced when model precision increases in machine learning. The higher Precision's stock price, the greater this benefit. The majority of a model's true positive predictions must be in line with reasonable anticipations for the model to receive a high Precision score.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

4.1.4. Recall (Sensitivity)

It reflects how often our model correctly predicts something. The genuine positive recognition rate of the ML model is to be improved as part of the retrieval process's overall mission. It is important to keep in mind that a high percentage does not always indicate that the model is accurate. This is because, even if the recovery rate is high, the model will not rule out the possibility of a positive outcome. Therefore, it is impossible to evaluate the accuracy of the negative predictions.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

4.2. Results

After settling on an accuracy metric and a loss function (binary_crossentropy), we assemble the model. Optimizers for deep learning will make adjustments to the model's parameters in order to lower the loss function. Training a model involves searching for the optimal settings of its parameters to minimize loss. The loss function shows how accurately the model can predict values from a given dataset. Adam, an optimization algorithm, was used because it has the potential for faster convergence and better performance on noisy data compared to other optimization algorithms. The information in Fig. 4 was uncovered with the use of metrics, which we used.

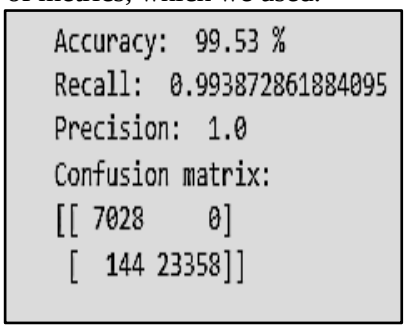

Fig. 4. Our obtained results.

4.3. Discussion

We found that boosting the number of model layers, the number of cores, and the application of optimization parameters improved the accuracy of the results, but also lengthened the training time. Table 1 displays comparisons between our proposed model and other research models employed in the same research setting as our current study. It shows the reliability of our recommended investigation and the outcomes of related searches.

Table 1. Comparing with related work.

Reference	Methods	Dataset	Accuracy
[8]	Several ML Methods	PIDD	77%
[9]	SVM, CNN	From HRV data analyse	95.7%
[10]	NB	PIDD	76.30%
[11]	DNN	PIDD	97.11%
[12]	NB	PIDD	74.28%
[13]	ANN, NB, DT, DL	PIDD	98.07%
[14]	XGBoost LightGBM, NB, RF, DT	Bangladesh Demographic and Health Survey (BDHS)	73.5%
[15]	DNN	Diabetes type dataset from the Data World repository	99.4112%
[16]	DT (J48)	PIDD	91.2%
[17]	SVM	PIDD	77.3%
[18]	FNN	PIDD	92.2%
[19]	DT (C5.0)	PIDD	74.63%
[20]	Several ML Methods	PIDD	75.03%
[21]	XGBoost	Health screening data of adults older than 65 years from 17 districts in Wuhan, China.	91.2%
[22]	DNN	PIDD	88.41%
[23]	DNN	PIDD	82.26%
Proposed model	Sequential model of 3 hidden layers	Diabetes Readmission Dataset	99.53%

It is evident, from the data that has been shown in the Table 1, that the model that has been proposed has the maximum accuracy that has been attained, which is 99.53%. This suggests that the model performs better than any other model that was considered in the previous research. Fig. 5 illustrates a comparison between the accuracy of our model and that of earlier studies. It demonstrates that the result we obtained by employing a serial CNN model with three hidden layers is superior to the results produced by those other studies.

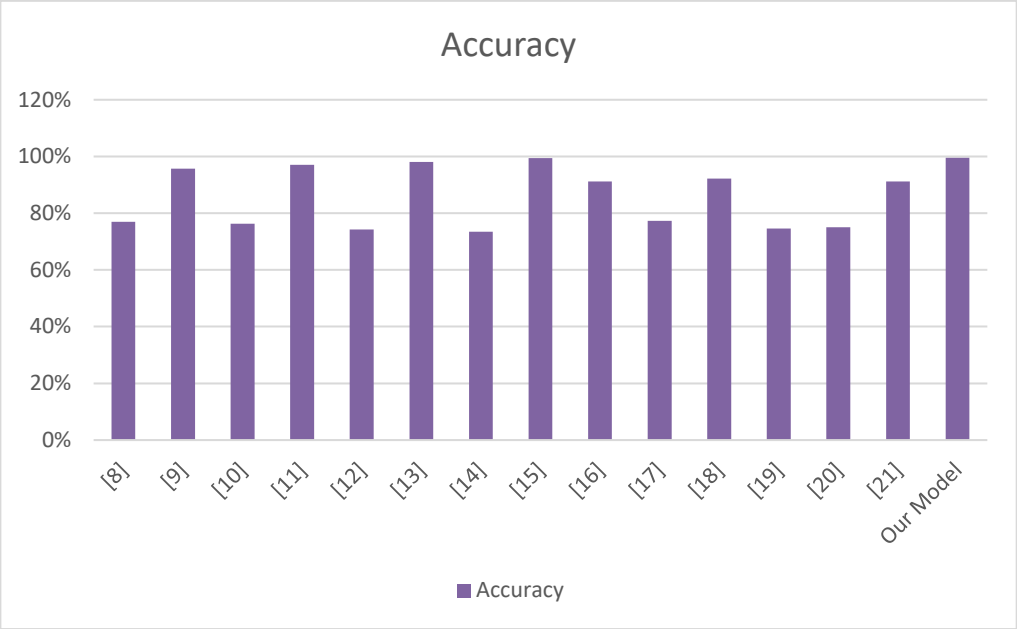

Fig. 5. The accuracy of our model compared to other studies.

The accuracy of a model is determined by a number of different aspects, including as the dataset, the preprocessing techniques used, the hyperparameter tuning, and the particular problem that is being addressed. Consequently, obtaining a high accuracy score is a considerable accomplishment, and it shows that our model is well-suited for the work that is being done here.

5. Conclusion

The diagnosis of diabetes represents one of the most significant and time-sensitive technical challenges facing the medical diagnosis system. Diabetes can lead to a number of problems, all of which can be prevented with timely diagnosis and treatment: kidney failure, stroke, blindness, heart attack, and amputation of lower limbs. Deep learning is a subfield of machine learning that is used in artificial intelligence. It is able to learn new skills merely by seeing how other people employ them. Another possibility is that it is able to learn on its own without any assistance from human beings. The human brain could take years to learn to make sense of the same amount of unstructured and unlabeled data, whereas a machine could do so in a matter of days given the same circumstances. Deep neural networks were utilized in this study to make diabetes diagnoses and prognoses. This study made use of the technology of deep neural networks by employing a sequential model that contained three hidden layers and making use of medical data in order to identify diabetes based on a number of distinct characteristics. This was accomplished by utilizing a deep neural network model. When compared to other approaches that are being utilized to diagnose diabetes mellitus at the moment. The proposed method will be useful for both doctors and patients. Future study in this field could involve, for instance, the use of intelligent agents to make efficient decisions over the dataset and to increase the forecast's accuracy while decreasing the time required to do so. We could build a dashboard to show doctors and patients a graphical representation of the vitals summary. The model's generalizability might be determined by testing it on further data. The synergistic effect of using data from healthcare wearable devices in real time to improve diabetes prediction will be investigated. The internet app also can help sort people into those who do and do not have diabetes. For subsequent development, we propose to create an ensemble of multiple CNN models with different architectures or different initialization seeds and combine their predictions to improve performance.

References

[1] A. M. Egan, S. F. Dinneen, Medicine, 47 (2019). Doi:https://doi.org/10.1016/j.mpmed.2018.10.002

[2] S. Benbelkacem, B. Atmani, 2019 International Conference on Computer and Information Sciences (ICCIS) (2019).

[3] S. Saru, S. Subashree, International Journal of Emerging Technology and Innovative Engineering (2019).

[4] T. Zhu, K. Li, P. Herrero,P. Georgiou, IEEE Journal of Biomedical and Health Informatics, 25 (2020).

[5] A. Choudhury, D. Gupta, "A survey on medical diagnosis of diabetes using machine learning techniques." in Recent Developments in Machine Learning and Data Analytics: IC3 2018. Springer Singapore, (2019). Doi:https://doi.org/10.1007/978-981-13-1280-9_6

[6] M. D. Owens-Gary, X. Zhang, S. Jawanda, K. M. Bullard, P. Allweiss, B. D. Smith, Journal of General Internal Medicine **34**(2),320 (2019). Doi:https://doi.org/10.1007/s11606-018-4705-2

[7] T. Zhu, K. Li, P. Herrero, P. Georgiou, IEEE Journal of Biomedical and Health Informatics **25**(7), 2744 (2020). Doi:https://doi.org/10.1109/JBHI.2020.3040225

[8] M. A. Sarwar, N. Kamal, W. Hamid, M. A. Shah, 2018 24th International Conference on Automation and Computing (ICAC) (2018). Doi:https://doi.org/10.23919/IConAC.2018.8748992

[9] G. Swapna, R. Vinayakumar, K. P. Soman, ICT Express, 243 (2018). Doi:https://doi.org/10.1016/j.icte.2018.10.005

[10] D. Sisodia, D. S. Sisodia, Procedia Computer Science **132**, 1585 (2018). Doi:https://doi.org/10.1016/j.procs.2018.05.122

[11] S. I. Ayon, M. M. Islam, International Journal of Information Engineering and Electronic Business (2019).

[12] S. Shafi, G. A. Ansari, Proceedings of the International Conference on Smart Data Intelligence (ICSMDI 2021) (2021). Doi:http://dx.doi.org/10.2139/ssrn.3852590

[13] H. Naz, S. Ahuja, Journal of Diabetes & Metabolic Disorders,19 (2020). Doi:https://doi.org/10.1007/s40200-020-00520-5

[14] A. Dutta, M. K. Hasan, M. Ahmad, M. A. Awal, M. A. Islam, M. Masud, H. Meshref, International Journal of Environmental Research and Public Health **19**(19),12378 (2022).

[15] H. Zhou, R. Myrzashova, and R. Zheng, EURASIP Journal on Wireless Communications and Networking, 1 (2020). DOI: https://doi.org/10.1186/s13638-020-01747-2

[16] A. M. Posonia, S. Vigneshwari, D. J. Rani, 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) (2020). Doi:https://doi.org/10.1109/ICISS49785.2020.9316001

[17] M. Radja, A. W. R. Emanuel, 2019 5th International Conference on Science in Information Technology (ICSITech) (2019). Doi:https://doi.org/10.1109/ICSITech46713.2019.8987479

[18] A. A. Zulfikar, W. A. Kusuma, 2019 International Conference on Advanced Computer Science and Information Systems (ICACSIS) (2019). Doi:https://doi.org/10.1109/ICACSIS47736.2019.8979989.

[19] K. M. Varma, D. B. Panda, Journal of Emerging Technology and Innovation Research **6**(6),522 (2019).

[20] A. K. Sahoo, C. Pradhan, H. Das, Nature Inspired Computing for Data Science (2020). Doi:https://doi.org/10.1007/978-3-030-33820-6_8

[21] Q. Liu, M. Zhang, Y. He, L. Zhang, J. Zou, Y. Yan, Y. Guo, Predicting the risk of incident type 2 diabetes mellitus in Chinese elderly using machine learning techniques, In Journal of Personalized Medicine **12**(6), 905 (2022).

[22] A. Ashiquzzaman, A. K. Tushar, M. R. Islam, D. Shon, K. Im, J. H. Park, IT Convergence and Security 2017 (2018). Doi:https://doi.org/10.1007/978-981-10-6451-7_5

[23] K. Kannadasan, D. R. Edla, V. Kuppili, Clinical Epidemiology and Global Health **7**(4), 530 (2019).Doi:https://doi.org/10.1016/j.cegh.2018.12.004

[24] R. Miotto, L. Li, B. A. Kidd, J. T. Dudley, Scientific reports **6**(1), 1 (2016).

[25] K. S. Ryu, S. W. Lee, E. Batbaatar, J. W. Lee, K. S. Choi, H. S. Cha, Applied Sciences **10**(1), 421 (2020).

[26] S. Lekha, M. Suchetha, IEEE journal of biomedical and health informatics **22**(5), 1630 (2017). Doi:https://doi.org/10.1109/JBHI.2017.2757510
[27] F. Anwar, M. Y. Ejaz, A. Mosavi, Informatics in Medicine Unlocked **21**(100482), (2020). Doi:https://doi.org/10.1016/j.imu.2020.100482
[28] K. Huang, A. Hussain, Q. F. Wang, R. Zhang, Deep learning: fundamentals, theory and applications (2019).
[29] F. Tamin, N. M. S. Iswari, 2017 4th International Conference on New Media Studies (CONMEDIA) (2017). Doi:https://doi.org/10.1109/CONMEDIA.2017.8266024
[30] M. Shardlow, "An analysis of feature selection techniques", The University of Manchester (2016).
[31] A. M. Javid, S. Das, M. Skoglund, S. Chatterjee, ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2021). Doi:https://doi.org/10.1109/ICASSP39728.2021.9414269
[32] V. Michalski, V. Voleti, S. E. Kahou, A. Ortiz, P. Vincent, C. Pal, D. Precup, arXiv preprint arXiv:1908.00061 (2019).
[33] S. Park, N. Kwak, Computer Vision–ACCV 2016: 13th Asian Conference on Computer Vision, Revised Selected Papers, 189 (2017). Doi:https://doi.org/10.1007/978331954184-6_12
[34] K. Ong, S. C. Haw, K. W. Ng, Proceedings of the 2019 2nd International Conference on Computational Intelligence and Intelligent Systems (2019). Doi:https://doi.org/10.1145/3372422.3372444

Journal of Basrah Researches (Sciences) **49**(2), 171(2023)
DOI: https://doi.org/10.56714/bjrs.49.2.14

تطوير نموذج CNN متسلسل مع ثلاث طبقات مخفية للتنبؤ بمرض السكري

حنين صباح الماجدي ، عباس مغربل

قسم علوم الحاسوب ، كلية العلوم والآداب ، الجامعة الأمريكية للثقافة والتعليم ، لبنان.

معلومات البحث

الاستلام 20 اب 2023
القبول 07 كانون الأول 2023
النشر 30 كانون الأول 2023

الكلمات المفتاحية

التعلم العميق ، السكري ، الرعاية الصحية ،الشبكة العصبية التلافيفية ، التنبؤ بالأمراض.

Citation: H. S. Al- Majidi, A. Mgharbel, J. Basrah Res. (Sci.) **49**(2), 171 (2023).
DOI:https://doi.org/10.56714/bjrs.49.2.14

الملخص

يمكن أن يساهم التشخيص المبكر لمرض السكري وعلاجه في التخفيف من المخاطر والمضاعفات المرتبطة به. لذلك، من الضروري توقع المرض وتحديده في مرحلة مبكرة من خلال استخدام إجراءات يمكن الاعتماد عليها و يمكنها تقديم تنبؤات تتميز بمستوى كبير من الموثوقية والدقة. تستخدم هذه الدراسة مجموعة بيانات إعادة القبول لمرض السكري، والتي تضم 101,766 سجلًا وتشمل 50 ميزة. بعد تحديد الميزات الأكثر صلة، تم تقسيم مجموعة البيانات إلى مجموعة تدريب ومجموعة اختبار. بعد ذلك، تم إنشاء نموذج متسلسل يستخدم التعلم العميق، وتحديدًا الشبكة العصبية التلافيفية (CNN) ذات ثلاث طبقات مخفية، لأغراض التنبؤ. تم تقييم صحة النموذج من خلال استخدام مقاييس اختبار الأداء، مما اعطى معدل دقة مسجل قدره 99.53٪. نتائج هذه الدراسة لديها القدرة على إثراء تطوير أساليب العلاج الشخصية التي تلبي المتطلبات الفردية للمرضى، وبالتالي تعزيز جودة الرعاية الصحية المقدمة.

*Corresponding author email : hanwa.almajidi@gmail.com

ISSN: 1817-2695 (Print); 2411-524X (Online)
Online at: https://jou.iobrs.edu.iq