



Early Prediction of Diabetes Using Deep Learning Convolution Neural Network and Harris Hawks Optimization

Murugadoss R^{1*}

¹Computer Science and Engineering Department,
St. Ann's College of Engineering & Technology, Chirala, INDIA

*Corresponding Author

DOI: <https://doi.org/10.30880/ijie.2021.13.01.009>

Received 17 February 2020; Accepted 10 October 2020; Available online 30 January 2021

Abstract: Owing to the gravity of the diabetic disease the minimal level symptoms for diabetic failure in the early stage must be forecasted. The prediction system instantaneous and prior must thus be developed to eliminate serious medical factors. Information gathered from the Pima Indian Diabetic dataset are synthesized through a profound learning approach that provides features for diabetic level information. Metadata is used to enhance the recognition process for the profound learned features. The distinct details retrieved by integrated machine and computer technology, including glucose level, health information, age, insulin level, etc. Due to the efficacious Hawks Optimization Algorithm (HOA), the data's insignificant participation in diabetic diagnostic processes is minimized in process analysis luminosity. Diabetic disease has been categorized with Deep Learning Convolution Networks (DLCNN) from among the chosen diabetic characteristics. The main objective of this paper is to evaluate the search and exploration capabilities of the Harris Hawks algorithm, the results are compared with those of commonly used alternative approaches such as PSOLDA, ACOMA, and SIRR algorithm. The process output developed is measured on the basis of test results in terms of error rate, sensitivity, specificity and accuracy.

Keywords: Metadata, Hawks Optimization Algorithm (HOA), Deep Learning Convolution Networks (DLCNN)

1. Introduction

Globally, diabetes has infected more than 246 million people, most of whom are women. This is projected to increase to over 380 million [1] by 2025, as per the WHO survey. The illness is named the fifth deadliest disease with no immediate cure in progress in the United States. Diabetes impact has been shown to have a more dangerous and damaging impact on people at the level of survival and life quality has decreased [2]. The WHO studies say that nearly one-third of diabetic people do not know the condition, so diabetes has to be controlled [3], but is unable to cure it. The individual with diabetes can experience serious medical issues such as nerve damage, heart attack, renal failure, and stroke. Statistics indicate that 8.8% of the world's population was projected to have diabetes in 2017. In 2045 this will probably rise to 9.9% [4].

Only approach to prevent problems is to recognize early [5]. Several scientists are experimenting in medicating diseases using various machine-learning algorithms such as J48, SVM, Naive Bayes, Decision Tree, Decision Table and others because research has shown that machine-learning algorithms work well for diagnosing specific diseases [6]. The ability to manage a vast quantity of data and merge data with several different sources and incorporate background data in the analysis gives data mining and machine learning algorithms strength [7]. In the scope of vital artificial intellectual problems and target recognition, it is also found that the reliability of standard learning algorithms is not at the affordable level, largely because of the large size of the information managed [8, 9].

The variety of machine learning strategies for automatic diabetes detection has indeed been suggested in a stereotactic way [10]. The use of training and data mining techniques for medical analysis is however a crucial technique for using vast quantities of data related to diabetes for awareness extraction [11, 12]. One of the key goals of the medical science analysis is the strong social influence of the specific disease, which ultimately provides enormous quantities of data [13]. In treatment, leadership and other associated health administration [14], computer and data mining methods are certainly of huge importance. Therefore, the existing literature on machine learning and data mining strategies to diabetes research is reviewed within the context of this study [15].

The work is focused on trying at proposing an efficient diabetes diagnostic method based on the approach of Hawks Optimization Algorithm [16] and the neural network of deep learning convolution [17]. Originally, the diabetes data was collected from the Pima Indian Diabetic dataset, and Hawks Optimization Algorithm uses the feature selection process to pick the optimal interface by using the position and endurance. Each function location is specified according to the parameters and must be constantly updated to achieve optimized functionalities in the memory basin. The selected features are eventually formed and ranked by neural networks depending on the proposed algorithm that encourages accelerating the process of diabetic prediction. The productive transition process increases the recognition rate as data from previous training are collected when new features are introduced into the diabetic recognition system.

The remaining article is structured as defies: Section 2 describes various works pertaining to the use of early detection of diabetes. Section 3 describes the proposed methodology on an automatic system for early detection of diabetes using Hawks Optimization Algorithm and deep learning convolution neural networks. Section 4 presents the experimental results. Section 5 defines the conclusion.

2. Related Work: A Brief Review

In this chapter, some of the various works submitted were discussed before here. In the first stage, a localized Diabetes dataset, formulated and obtained from Bombay Medical Hall, Upper Bazar Ranchi, India, has been developed and proposed by Choubey D.K et.al [18] for research consisting of four stages. In the second step, a neural network was used as the spatial diabetes data set classification methodology. In the third phase of the GA selection method, six features were extracted from among 12 apps. Finally, neural networks were used in the fourth stage to classify the appropriate GA attributes. The tests for the classification method used were tested with and without GA in this report. In this work, it was found that GA has been beneficial not just to eradicate important attributes, reduce costs and processing time, but also to improve ROC and reliability. The technique used can also be applied in many other medical conditions.

The research paper on diabetes diagnostics using soft computing techniques by Dilip Kumar Choubey et.al [19] has been reviewed in several studies. Diabetes is a chronic illness and a major threat to people's health globally. It's a disorder with an abnormally high blood glucose level. The body cannot produce or react properly to insulin essential to maintaining the glucose rate, causes diabetes. Soft computational approaches, such as SVMs, Ant Colony Optimization (ACO), rough environment, neural network, fuzzy logic, genetic algorithm, etc., could be used to detect diabetic diseases.

In the first stage the Genetic algorithms (GA) used to pick an attribute (feature), which decreases four attributes of eight attributes and in the second stage the Radial Basic Function Neural Network (RBF NN) was used to distinguish selected attributes between all attributes was proposed by Dilip Kumar Choubey et.al [20]. The observation results illustrated the quality of the existing Pima Indian Diabetes Dataset (PIDD) approach and offered a better clinical diagnosis of PIDD patients. GA added small elements, reduced costs and time for computation and increased identification ROC reliability. Other types of medical diseases could be used with the conceptual model.

Initial stage of the Genetic Algorithm (GA) was used for the collection of PMA Indian Diabetes Dataset (PIDD) and in the second phase, the J48Graft Decision Tree (J48graft DT) was used for the classification and projection of the chosen function by Dilip Kumar Choubey et.al [21] suggested an applied approach works in the 2 phases. Any illness with lower costs should be diagnosed early. Also one of these diseases is diabetes. The GA solution has increased the ROC of identification, which was not only intended to reduce storage space, expense and measurement times of the diagnostic process. The experimental results indicated that GA and J48graft DT could be used successfully to identify a diabetes disorder (74.7826 percent) and ROC (0.86 percent) identification accuracy.

A model incorporating genetic and fuzzy algorithms for selection and classification of functions has been suggested by E.P.Ephzibah et al [22]. Among developed and recently industrialized nations, diabetes had become the fourth major cause of death, and there was a significant indication that it has been approached crisis point. Behaviors of empirical symptoms and consequences of diagnostic tests were used for medical diagnosis. The affiliated costs and risks of these tests were different. The proposed system addresses the choice of attribute subsets in the automatic layout of the pattern classification. It was an endeavor to locate and choose a useful subset of patterns from a broader range of functionality. The proposed system increased the accuracy of the category with its flippant rules-based classification system.

J. Jayashree, S. Ananda Kumar [25] has introduced a swarm intelligent redundancy relevance (RR) along with convolution trained compositional pattern neural network for predicting the diabetic disease. In this work, diabetes data are collected from the Pima Indian Diabetes Database (PIDD), the dimension of data is reduced by mass intelligence RR techniques, and selected features are trained by layers of convolution networks that help accelerate the diabetes prognosis process. The missing value of the collected data is investigated using the random data coefficient residual normalization

procedure, which converts the noisy data by computing the mean, standard deviation value. The diabetes classification process is performed by the compositional pattern neural network.

A hybrid neuro-fuzzy model for the risk of developing hypertension has been presented by Melin, Patricia, et al. [30]. They conclude that the modular neural networks are an efficient technique, because, when dividing the problem into sub-problems, the complete problem becomes less complex. In each module, different architectures can be used for obtaining good results in simulating the behavior of the blood pressure of different patients. In this work, two Mamdani fuzzy inference systems are used as response integrators, which classify the blood pressure and heart rate level respectively. Also, the fuzzy inference system was used to classify the patient’s night profile and was tested, which correctly classifies all cases. This hybrid system combining neural networks and fuzzy logic is used to provide the final diagnosis to the patient.

A Gray Wolf Optimizer for Modular Neural Network (MNN) with Granular Approach has been proposed by Sanchez, Daniela et al [31]. The design of a modular granular neural network (MGNN) lies in finding the optimal parameters of its structure; These parameters include the number of subgroups, the percentage of data for the training phase, the number of hidden layers, the learning mechanism, the target error, and the number of their neurons. This method performs human recognition of the design of data recognition and modular neural networks, and a database of ear, iris, and face biometric measures is used to perform tests and comparisons against other works to demonstrate its effectiveness.

3. Proposed Methodology

A scheme of analysts depending on a Hawks optimization and deep learning convolution neural network was established to envisage patients ' diabetes to prevent risk factors pretty early. Mostly at the moment of the diabetic assertion, the mechanism functions process by process. First stage consists of the compilation of diabetic data, noise reduction or incompatible process and choice of core functionality. This first step is essential because the diabetic data generated usually consists of patient information, with few important or relevant details to diabetic prediction and certain data unrelated to the diabetic system to optimize computational complexity. Once the appropriate features have been extracted, it is now step 2 which uses the chosen features to predict unusual features by incorporating the neural network approach to deep learning convolutions. The prediction process uses large quantities of data while training the network, which categorizes diabetic features efficiently without delay and without major issues. The actual design of the diabetes forecast is included in the Figure 1 according to the description mentioned.

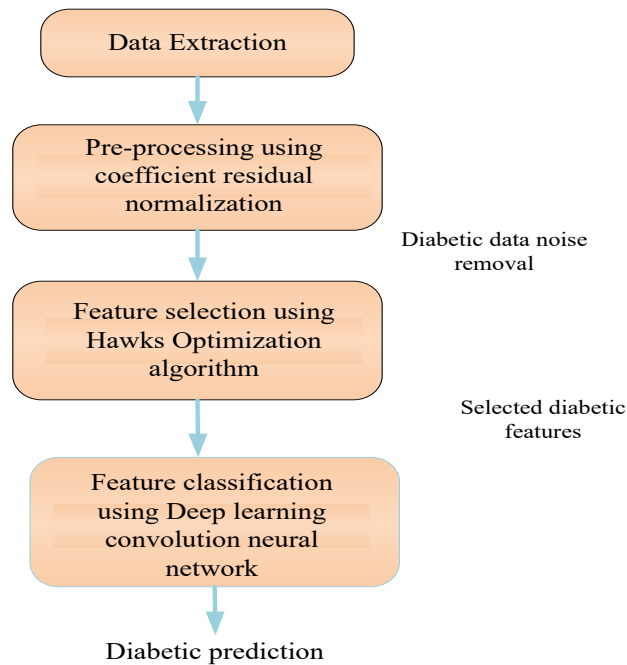


Fig. 1 - Overall structure of proposed methodology

3.1 Extraction of data and Diabetic data pre-processing coefficient for residual normalization

The first step of the mechanism of forecasting of diabetic disease is to collect data from the Pima Indian Diabetic list for patients with a glucose level, blood pressure, skin density, insulin, diabetes pedigree function, body mass index and personal details (<https://www.kaggle.com/uciml/pima-indiansdiabetes-database>). The data obtained makes it easier to

scrutinize the illness and the diabetes examined in order to reduce the health risks. The data obtained may include numerous incoherent and lacking details which limit the recognition system performance. Therefore, the data were evaluated and optimized by implementing a residual standardization method that tests any information contained in datasets effectively without adding uncertainty. People with diabetes are actively scrutinized for collection of data and special period information is collected that is maintained as a database. Throughout the data collection process, due to technical difficulties or manual completion processes, several data are ended up missing. First, it is found that the omitted value is substituted by the information coefficient variance estimation method.

$$\text{Coefficient variance} = \frac{\sigma}{\mu} \tag{1}$$

Where σ is depicted in the diabetic list as a standard deviation of specific data. The standard value variance is measured as pursues variance.

$$\sigma = \sqrt{\frac{1}{k-1} \sum_{p=1}^k (a_p - \text{avgvalue})^2} \tag{2}$$

Where k is number of attributes in data list, a_p is the particular data in the list, μ is the average information value calculated accordingly,

$$\mu = \frac{1}{K} \sum_{p=1}^K a_p \tag{3}$$

The coefficient value in replacing the lost and unreliable data using equation 2 and 3 in the list shall be quantified. Everything has to be transformed as enforces after adding the missing value as it simplifies the prediction process. Then the standardization procedure is calculated,

$$\text{Residual normalization} = \frac{A_p - \mu_p}{\sigma_p} \tag{4}$$

From the Eqn.4 criterion, with standard deviation, the specific data value is standardized; the normalized data is a valuation of 0 to 1. Perhaps the entire process potentially substitutes noise information and extracts different features including minimum data, maximum data, standard deviation, mean, variance and other features. Every client has many data that is tough to handle with minimal difficulty and the features obtained are enormous. Given the importance of optimum functionalities, Hawks Optimization Algorithm (HOA) is used in the data set to identify the best functionality in the collection.

3.2 Feature Selection Using Hawks Optimization Algorithm (HOA)

The second step is the choice of the function; it is an essential step as it prefers the best functionality. Therefore, the functionality selection process mitigates information over appropriate and unnecessary data when assessing functions, as the overfitting data generates uncertainty during diabetic forecasting. This function also reduces the redundancy of the functions in the list by using HOA to lessen its data set dimensions. The exploration and exploitation phases of the proposed HOA are inspired by the exploration of a prey, focusing on surprise pounce and the Harris hawks' different attacking strategies. HOA is a population-based, gradient-free optimization technique. Therefore, it can be used for any optimization problem. Harris hawks' main trick is to capture a prey is ‘surprise pounce’; It is also known as the ‘seven kills’ strategy. In this clever strategy, many hawks attempt to co-operate from different directions and merge outside the cover of detected escape rabbit a simultaneously. This attack can be completed quickly by capturing the surprised prey in a few seconds. Harris's hawks can adapt to a variety of chase styles, depending on the environment and the prey's escape patterns. A switching trick occurs when the best hawk (leader) stops prey and loses, and the chase is pursued by one of the party members. These switches deal with different situations because they are beneficial for confusing the escape rabbit. The main advantage of these cooperative tactics is that Harris' hawks can continue to detect rabbit fatigue, which increases its vulnerability. Moreover, by confusing the escape prey, its defensive capabilities cannot be restored, and finally, the confronting team cannot escape the siege, because one of the hawks, often the most powerful and experienced, catches the tired rabbit effortlessly and shares it with the other party members. Originally, HOA is used to evaluate the attributes in order to solve the diabetes function selection issue in a global context. The method is used in

the exploratory and exploitative stages of the proposed HOA in the feature selection process, intrigued by an examination of Harris Hawks ' predators, surprise punch and various attack techniques. HOA is a structural, gradient-free method of optimization; thus, it can be implemented subject to valid formulations on any optimization problem. In the HOA exploratory process, we take the essence of Harris ' hawks into account, they locate and sense a beast with their strong eyes but sometimes the beast cannot be instantly identified. Thus, hawks wait, watch and calibrate the desert site, perhaps after several hours, to identify a prey. In HOA, the hawks of Harris are the solution and each step is regarded as the most appropriate prey or nearly the best solution to the candidate. HOA is a place where the hawks of the Harris perch at fairly random and wait patiently for a prey based on two approaches. When one assumes that each perching strategy has equal chances, they depend on the positions of the others in the family (as near to them when targeting) and the rabbit loosely based on Eqn. (5) For $x < 0.5$ or perch in high-lying trees arbitrary (random places within the home group) based on Eqn. (5) Condition of $x > 0.5$.

$$W(r+1) = \begin{cases} W_{rnd}(r) - b_1 |W_{rnd}(r) - 2b_2 W(r)| & x \geq 0.5 \\ (W_{rbt}(r) - W_{ap}(r)) - b_3(L_{lb} + b_4(U_{ub} - L_{lb})) & x < 0.5 \end{cases} \quad (5)$$

where $W(r+1)$ is the position vector of hawks in the next iteration r , $W_{rbt}(r)$ is the position of rabbit, $W(r)$ is the current position vector of hawks, b_1, b_2, b_3, b_4 and x are random numbers inside $(0,1)$, which are updated in each iteration, L_{lb} and U_{ub} show the lower and upper bounds of variables, $W_{rnd}(r)$ is a randomly selected hawk from the current population, and $W_{ap}(r)$ is the average position of the current population of hawks. These value changes randomly in each iteration to simulate the nature of rabbit motions.

A basic structure for generating random locations within the home range of the group (L_{lb} and U_{ub}) is proposed. The very first rule creates arbitrary location and certain other hawks strategies. In Eqn. (5) second rule, there is the distinction of the best condition up until now, with the group's average distance plus an arbitrarily-scaled mechanism based on a variety of different factors. The scaling coefficient of b_3 increases the randomness of rule once b_4 is close to 1. We contribute an arbitrarily proportioned length of the momentum to the L_{lb} in this rule. Then we consider the element as a random scaling coefficient to generate additional patterns in diversification and exploring specific regions. Distinct optimizing rules can be established, but we have used the least intrusive rule to imitate the conduct of hawks. Hawks reach their average position by Eqn. (6):

$$W_{ap}(r) = \frac{1}{M} \sum_{g=1}^M W_g(r) \quad (6)$$

Where $W_g(r)$ indicates the location of each hawk in iteration r and M denotes the total number of hawks. The median position can be determined by various means, but the least intrusive rules are being used.

The HOA algorithm can migrate from exploration to exploitation and can then modify between various exploitative conduct based on a loss of energy from the prey. During the escape, the energy of a beast drops significantly. The energy of a prey is styled on the following template:

$$EC = 2EC_0(1 - \frac{h}{H}) \quad (7)$$

Where, EC signifies the prey's escape energy, H shall be the maximum number of iterations, and EC_0 shall be its initial energy level. In HOA, at each repetition, EC_0 shifts randomly within the interval $(-1, 1)$. The rabbit is self-destructing when the value of EC_0 decreases from 0 to -1 , while the value of EC_0 rises from 0 to 1, mean that rabbits become stronger. A declining trend during the iterations is in the vibrant escaping energy EC . When the energy manages to escape $|EC| \geq 1$, the hawks look for a rabbit's position in a different region, therefore the HOA investigates the exploration process, and when $|EC| < 1$, the algorithm tries to use the alternatives neighborhood during the activity processes. Simply put, exploration takes place when $|EC| \geq 1$, while later exploitation happens when $|EC| < 1$.

In exploitation activity, the hawks of the Harris carry out the surprise shot by targeting the envisaged beast spotted in the preceding phase. But preys often try to flee risky circumstances. Therefore, in unusual situations various ways of pursuit exist. The HOA proposes four strategic options for modeling the assaulting stage, as per the escape behavior of the Harris hawks and the chasing strategy. The preys are always trying to escape dangerous situations. Presume e is an

opportunity to escape a target safely ($b < 0.5$) or not ($b \geq 0.5$) before surprise pounce. Regardless of what the prey does, the hawks are hard or weak in order to catch the prey. This implies that they annihilate the prey tenderly or severely based on the accumulated energy of the prey from various directions. The counterfeit attacks the desired target in real life scenarios to boost their chances to kill the rabbit cohesively by doing the trick shot. The escaping prey loses more and more energy in a couple of minutes, so the hawk intensifies the assembly process to capture the depleted prey without difficulty. The EC parameter is used for shaping this approach and allowing the HOA to switch from soft and hard-biting processes.

Soft besiege arrives as $b \geq 0.5$ and $|EC| \geq 1$, the rabbit does have sufficient energy and tries to avoid those false hops unexpectedly, although it probably won't. In such actions, the hawks of Harris cover the rabbit in a soft way and then expel the surprise. The following rules applied to this behavior:

$$W(r+1) = \Delta W(r) - EC|GW_{rbi}(r) - W(r)| \tag{8}$$

$$\Delta W(r) = W_{rbi}(r) - W(r) \tag{9}$$

Where $\Delta W(r)$ is the vector position of a rabbit different from the current position of iteration r , the b_5 is a random number inside (0,1), and $G = 2(1 - b_5)$. The discrete jump resilience of a rabbit symbolizes a random jumping force throughout the escape process. The J value adjustments pick a random each time the rabbit motions are simulated.

The prey is so flustered and has low escape energy when $b < 0.5$ and $|EC| < 1$ hard besiege occurs. Therefore, the hawks of Harris hardly cover the anticipated target, to carry out the surprise blasting. The current position is updated with Eqn. 10 in this scenario:

$$W(r+1) = W_{rbi}(r) - EC|\Delta W(r)| \tag{10}$$

From the modified app location, optimal or maximized fitness value specific apps are chosen from list of feature set. Although the hawks' optimization system assigns effective features, it must be independent of duplication as well as pertinent to the diabetic product selection process. The HOA approach chooses customized functionality, the dimensionality of the feature set is minimized using an efficient fitness algorithm. The average fitness of hawks tracks how the average fitness of the entire population varies during the optimization process. The integration metric reveals how the rabbit's fitness value (the best solution) varies during optimization. The function is chosen according to the position, direction of functionalities and fitness value. In order to assess the redundancy and relevance of the feature, the designated attributes are also reviewed. The parameters eradicate data fit and thus minimize the dimensionality of the data. The mutual importance of the chosen characteristic is calculated to estimate how appropriate to the diabetic classification. The feature's mutual importance is then examined appropriately as follows:

$$MTL(F, cl) = \frac{1}{|F|} \sum_{sp_k \in F} M(sp_k, cl) \tag{11}$$

In eqn. (11), $MTL(F, cl)$ is defined as the feature's mutual significance, F is chosen as a function set with a fitness value; cl is a common feature class set, and sp_k is a relevant feature. It is used for determining how the chosen features are relevant to the diabetic characteristic set by the calculated mutual importance of features. In accordance with mutual significance, the despondent value of all corresponding data measured as described is determined by calculating the mean value:

$$Q(y) = \frac{1}{|Y|^2} \sum_{sp_k, sp_l \in Y} L(sp_k, sp_l) \tag{12}$$

The following general feature selection criteria must be defined by using mutual importance (Eqn. 11) and redundancy of features (Eqn. 12).

$$= \max_f \left[\frac{1}{|F|} \sum_{sp_k \in F} M(sp_k, cl) - \frac{1}{|Y|^2} \sum_{sp_k, sp_l \in Y} L(sp_k, sp_l) \right] \tag{13}$$

The selected features are consistently investigated and relevant in accordance with the criteria. While the features satisfy this condition (Eqn. 13), an automated function was chosen to forecast the patient's diabetic diagnosis.

3.3 Diabetic Classification Using Deep Learning Convolution Neural Networks (DLCNN)

The final stage of the diabetic prediction process is to classify the methods used by Deep learning convolution neural networks (DLCNN). An effective training process enhances the detection rate because data from previous learning are collected when new functionality is added to the diabetic detection system. The training process relies on the fact of the physiological principle which relies on the relations in the training space among neurons or characteristics. It seems that every feature plays a particular role, so that each characteristic pays attention to its input characteristics to the scenario. The network uses different layers during the features training process like input, hidden and output between the three layers, input layer has several sub layers, namely convolutions, pooling, fully connected and normalized layers, which were used to effectively train attributes. The designated features are originally perceived as an input to the next hidden layer of the input layer. The data generated is analyzed in respect of transmission area in a convolution network. In comparison to these sources, the metadata is gathered attributable to the profound learning process, which helps to identify the functionality.

The convolution layer system outputs the responsive field, which results in the cluster output, indicating that the same inputs as the same cluster are generated. Pertain to the max pooling function to the concentration layer which is used to examine each cluster. This implies selecting from each feature cluster the maximum value of attributes. For a long time, the network analyzed the over-fitting of data by increasing the file capacity of the information it transmitted. The processed data is projected into the function space in the pooling layer and chooses the highest feature value used to pass to the fully connected layer which calculates the output value. Matrix multiplication is used in the completely connected layer to determine the output significance as authentication function. The process continues until the chosen functions are educated that are retained in the database for the prediction. Eventually, new diabetic characteristics are studied in order to predict normal and abnormal attributes in a Deep learning convolution neural network. The extracted characteristics are synthesized to the input level, which is sent to the hidden layer that devours the weighted input and is transmitted through the authentication function to estimate the output of the hidden layer.

Accomplish a vector input convolution operation using the analogous and comprehensive eqn 14.

$$Z_d^c = \left(\sum_{f=1}^m \sum_{s=-u}^u \sum_{t=-u}^u D_f^{c-1}(k-s, l-t) * V_{f,d}^c(s, t) + Y_d^c \right) \quad (14)$$

where m represents number of feature maps in last layer, f and d denotes feature map indices of current layer and previous layer respectively, l denotes the layer, $*$ denotes convolution operation, Y and u represents the bias and size of filter respectively. Initially D_f^0 represents the input image on which first convolution is to be performed and D_f^1 represents the input on which second convolution is to be performed, which can be obtained after applying pooling on D_f^0 .

$$X_d^c(k, l) = \left(\frac{1}{4} \sum_{s=0}^v \sum_{t=0}^v Z_f^c(2k-s, 2l-t) \right) \quad (15)$$

Where v is the window width of the enclosure. Perform a considerable number of iterations for convergence and pooling as required. Transfer the result acquired in the last pooling layer through a fully connected classification layer and measure the performance generated by equation 16.

$$Actual\ Output = \sigma(wgt \times opt + Y) \quad (16)$$

Where opt is the final output vector obtained after last pooling operation, wgt is the weight vector of fully connected layer.

For sigmoid activation function,

$$\Delta \hat{B}(k) = (\hat{B}(k) - B(k)) \bullet \hat{B}(k)(1 - \hat{B}(k)) \quad (17)$$

Computation of Δwgt ,

$$\Delta wgt = \Delta \hat{B}(k) \times opt(l) \tag{18}$$

Computation of ΔY ,

$$\Delta Y = \frac{\partial P}{\partial Y(k)} \tag{19}$$

On the basis of Eqn. (16), the output is determined in order to predict precisely the situation of new diabetes outputs in the expert model, which is compatible with the training data in the database. The weights and bias values are regulated consistently during the classification process so as to mitigate the system error rate. Following weight and reproductive value verification, the component is used to make adjustments the new test characteristics to train. Productive training and testing classifications categorizes normal and abnormal diabetic traits.

4. Experimental Results

4.1 Performance Metrics

Certain metrics are computed with True Negative (T_N), True positive (T_P), False Negative (F_N) and False positive (F_P).

4.1.1 Accuracy

The measure of overall usefulness/ effectiveness of the classification technique are called accuracy. The equation for accuracy is

$$Accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{20}$$

4.1.2 Specificity

To recognize patterns of a negative class, it is used to measure the classifier ability. It is computed as follows.

$$specificity = \frac{T_N}{T_N + F_P} \tag{21}$$

4.1.3 Sensitivity

To recognize patterns of a positive class, it is calculated to measure the classifier ability. It is computed as follows.

$$sensitivity = \frac{T_P}{T_P + F_N} \tag{22}$$

4.1.4 Error rate

The number of all incorrect predictions divided by the total number of the dataset is calculated as error rate.

$$Error\ rate = \frac{F_P + F_N}{T_P + T_N + F_P + F_N} \tag{23}$$

4.2 Performance Analysis

To test the data using the proposed classification techniques, we consider 768 data items, were 10% of data are used for data validation and 90% of the models are used for training and testing. Training sets are used to train the models and to evaluate the best model parameters and test sets are used to test the model performance, and finally the validation set ensures the effectiveness and validity of the selected model. Due to the use of the methods, the specificity and sensitivity of their algorithm in the experimental set of 192 cases was 76%. The proposed DCNN with a 0.1 split verification had 88.41% accuracy, which is the new recorded accuracy of the PIDD dataset.

Figure 2 shows that the Hawks Optimization Algorithm (HOA) achieves high precision and also minimizes the redundancies of information when evaluating features in the app space, by choosing best diabetic features appropriately. The choice accuracy achieved is greater than the alternative approaches used are PSOLDA [23], ACOMA [24], and SIRR [25]. Even if the HOA approach chooses customized functionality, the dimensionality of the feature set is minimized using an efficient fitness algorithm. The choice of characteristics enhances the method of diabetic detection minimizes the error rate.

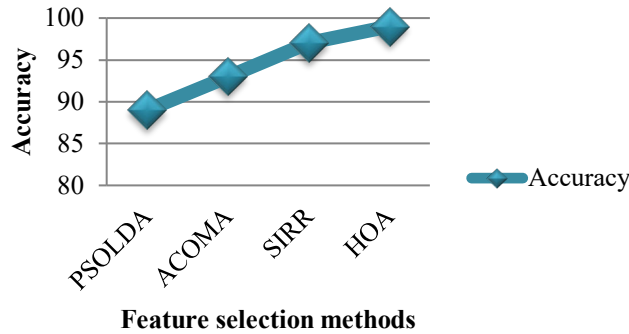


Fig. 2 - Accuracy based on different feature selection methods

In Figure 3, Deep learning convolution neural networks (DLCNN) classifies incoming diabetic features with high specificity (99.14%) that is maximum compared to other methods such as Convolution Trained Compositional Pattern Neural Network (CTCPNN) (98.58%), Learning Vector Quantization Optimized with Ant Colony (LVQAC) (95.36%) and Memetic Optimized Deep Learning Neural Network (MODLNN) (97.62%). From the outcomes it is evidently shown that the technique effectively implemented to eliminate the dimensionality of the feature set and categorizes the incoming features in an economical way into normal and abnormal.

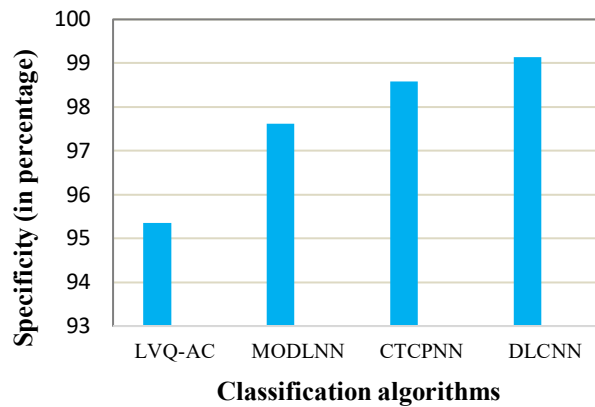


Fig. 3 - Specificity based on classifiers

Figure 6 demonstrates that the sensitivity value of different diabetic prediction methods from the different approach, Deep learning convolution neural networks ensures high value 99.02% that is higher collated to conventional methods such as Convolution Trained Compositional Pattern Neural Network (CTCPNN) (98.42%), LVQ-AC (96.32%) and MODLNN(97.43%).

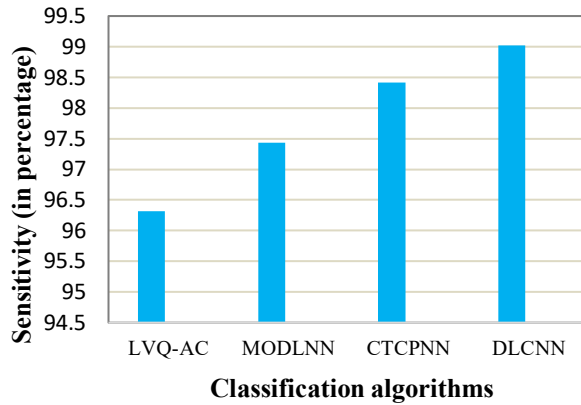


Fig. 4 - Sensitivity based on classifiers

In Figure 5 deep learning convolution Neural networks achieve a minimal error rate of 0.0108 relative to the CTCPNN [25], Learning Vector Quantization Optimized with Ant Colony (LVQAC) [26] and Memetic Optimized Deep Learning Neural Network (MODLNN) [27] in the creation of a neural structure sequence. The effective training method for the estimation of the functions, weight and bias minimizes the error rate. The decreased error rate and the efficient activation and training functions enhance the identification process evaluated with responsiveness and accuracy metrics, since the DLCNN approach evaluates how efficiently the input diabetic features organize relevant data from the qualified data set.

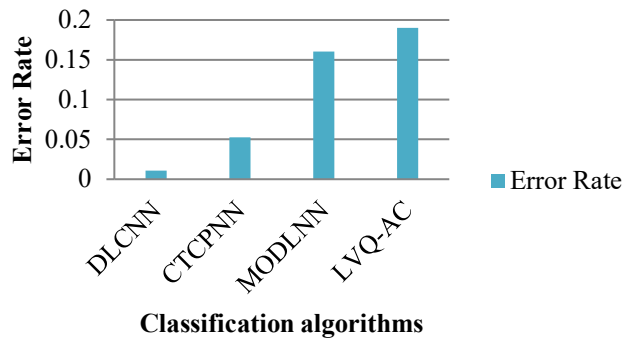


Fig. 5 - Error rate based on classifiers

Figure 6 demonstrates that sensitivity and specificity value of different diabetic prediction methods from the different approach, Deep learning convolution neural networks ensures high value (specificity-99.14%, sensitivity-99.02%) that is higher collated to conventional methods such as Convolution Trained Compositional Pattern Neural Network (CTCPNN) (specificity-98.58%, sensitivity-98.42%), LVQ-AC (specificity-95.36%, sensitivity-96.32%) and MODLNN(specificity-97.62%, sensitivity-97.43%). The increased tolerance and accuracy level increases the overall prediction system, which is analyzed using precision metrics.

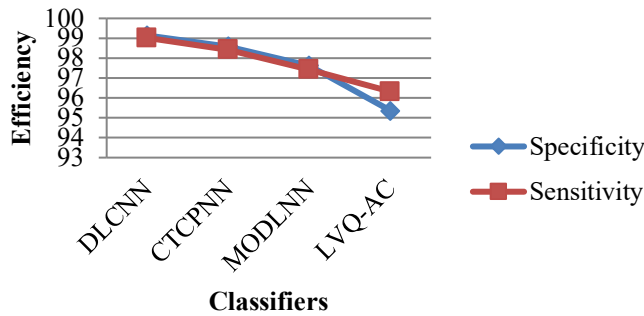


Fig. 6 - Efficiency based on classifiers

In Figure 7, Deep learning convolution neural networks (DLCNN) classifies incoming diabetic features with high accuracy (99.35%) that is maximum compared to other methods such as Convolution Trained Compositional Pattern Neural Network (CTCPNN) (98.41%), Learning Vector Quantization Optimized with Ant Colony (LVQAC) (95.84%) and Memetic Optimized Deep Learning Neural Network (MODLNN) (97.27%). From the outcomes it is evidently shown that the technique effectively implemented eliminates the dimensionality of the feature set and categorizes the incoming features in an economical way into normal and abnormal.

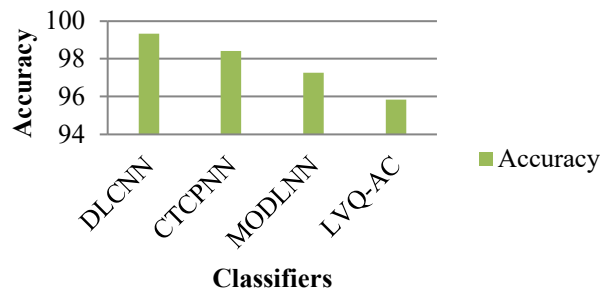


Fig. 7 - Accuracy based on classifiers

5. Conclusion

The paper hence analyses the diabetic condition of the patient with the hawks’ optimization algorithm and the deep learning convolution neural networks. Initially, diabetes data from the Pima Indians dataset are obtained as it consists of a large amount of data used to examine the process of diabetes prediction. The data are explored using a residual standardization process coefficient to replace the loud data with the mean, standard deviation value. The optimized characteristics are then selected in exploration and chasing mode, in which function is chosen according to the position, direction of functionalities and fitness value. In order to assess the redundancy and relevance of the feature, the designated attributes are also reviewed. The parameters eradicate data fit and thus minimize the dimensionality of the data. The chosen features were trained by specific hidden layers of deep learning approach which used to evaluate the new incoming diabetic functionality that is categorized using Deep learning convolution neural networks which have multiple activation mechanism when categorizing features. Then the performance of the proposed approach is evaluated based on the metrics and the experimental result show that the specificity, sensitivity, accuracy and efficiency of the proposed system achieves the value of 99.14%, 99.02%, 99.35% and 99.08% respectively, which is high when compared to the different diabetic prediction methods from the different approach such as Deep learning convolution neural networks ensures high value that is higher collated to conventional methods such as CTCPNN, LVQ-AC and MODLNN. The increased tolerance and accuracy level increases the overall prediction system, which is analyzed using precision metrics. This proposed hybrid system that combines the Deep Learning Convolution Neural Network and the Harris Hawks Optimization can be used to provide an early prediction of the patient's diagnosis. The risk diagnosis results presented by a patient in developing hypertension can help patients know that they need to make changes to their lifestyle to improve their health. In future work, we use the proposed hybrid models to better address the uncertainty of the decision-making process involved in the diagnosis of diabetes. In addition, meta-heuristic algorithms can be used to optimize both the deep learning convolutional neural network and the Harris Hawks optimization involved in the proposed hybrid model.

References

- [1] Stewart, W., Ricci, J., Chee, E., Hirsch, A., & Brandenburg, N. (2007). Lost Productive Time and Costs Due to Diabetes and Diabetic Neuropathic Pain in the US Workforce. *Journal Of Occupational And Environmental Medicine*, 49, 672-679. doi: 10.1097/jom.0b013e318065b83a
- [2] Kumar, K., Gupta, A., & Kumar, A. (2014). Attitude of health care professionals about the diabetes from India. *Journal Of Social Health And Diabetes*, 02, 092-095. doi: 10.4103/2321-0656.130795
- [3] Nelson, M. (2010). Drug treatment of elevated blood pressure. *Australian Prescriber*, 33, 108-112. doi: 10.18773/austprescr.2010.055
- [4] Maria Rotella, C., Pala, L., & Mannucci, E. (2013). Role of Insulin in the Type 2 Diabetes Therapy: Past, Present and Future. *International Journal Of Endocrinology And Metabolism*, 11. doi: 10.5812/ijem.7551
- [5] Adult Onset Type 1 Diabetes Mellitus Versus Type 2 Diabetes Mellitus: A Case Study. (2009). *The Internet Journal Of Advanced Nursing Practice*, 10(1). doi: 10.5580/cdb
- [6] Mohamed Shakeel, P., Baskar, S., Sarma Dhulipala, V., Mishra, S., & Jaber, M. (2018). Maintaining Security and Privacy in Health Care System Using Learning Based Deep-Q-Networks. *Journal Of Medical Systems*, 42. doi: 10.1007/s10916-018-1045-z
- [7] Butalia, S., Kaplan, G., Khokhar, B., & Rabi, D. (2016). Environmental Risk Factors and Type 1 Diabetes: Past, Present, and Future. *Canadian Journal Of Diabetes*, 40, 586-593. doi: 10.1016/j.jcjd.2016.05.002
- [8] Dolinski, S. (2000). *Irwin and Rippe's Intensive Care Medicine, Fourth Edition*. By Irwin, Rippe, and Cerra. Philadelphia, Lippincott Williams & Wilkins, 1998. Pages: 2912. List Price: \$225.00. *Anesthesiology*, 92, 1511. doi: 10.1097/0000542-200005000-00069
- [9] 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2019. (2018). *Diabetes Care*, 42(Supplement 1), S13-S28. doi: 10.2337/dc19-s002
- [10] Shakeel, P. M., Baskar, S., Dhulipala, V. S., & Jaber, M. M. (2018). Cloud based framework for diagnosis of diabetes mellitus using K-means clustering. *Health information science and systems*, 6, 16. doi: 10.1007/s13755-018-0054-0
- [11] Patil, B. M., Joshi, R. C., & Toshniwal, D. (2010). Hybrid prediction model for type-2 diabetic patients. *Expert systems with applications*, 37, 8102-8108.
- [12] Li, Y., Li, H., & Yao, H. (2018). Analysis and Study of Diabetes Follow-Up Data Using a Data-Mining-Based Approach in New Urban Area of Urumqi, Xinjiang, China, 2016-2017. *Computational and mathematical methods in medicine*, 2018. doi: 10.1155/2018/7207151
- [13] Mishra, T., & Kapoor, V. (2016). Financial trends prediction using the back propagation neural network and YQL. *Int. J. Comput. Appl*, 144(10), 20-26. doi: 10.5120/ijca2016910451
- [14] Jain, D., & Singh, V. (2018). Feature selection and classification systems for chronic disease prediction: A review. *Egyptian Informatics Journal*, 19, 179-189. doi: 10.1016/j.eij.2018.03.002
- [15] Daliri, M. R. (2012). Feature selection using binary particle swarm optimization and support vector machines for medical diagnosis. *Biomedizinische Technik/Biomedical Engineering*, 57, 395-402. doi: 10.1515/bmt-2012-0009
- [16] Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., & Chen, H. (2019). Harris hawks optimization: Algorithm and applications. *Future generation computer systems*, 97, 849-872. doi: 10.1016/j.future.2019.02.028
- [17] Indolia, S., Goswami, A. K., Mishra, S. P., & Asopa, P. (2018). Conceptual understanding of convolutional neural network—a deep learning approach. *Procedia computer science*, 132, 679-688. doi: 10.1016/j.procs.2018.05.069
- [18] Roopaei, M., Rad, P., & Jamshidi, M. (2017). Deep learning control for complex and large scale cloud systems. *Intelligent Automation & Soft Computing*, 23, 389-391. doi: 10.1080/10798587.2017.1329245
- [19] Choubey, D. K., & Paul, S. (2016). Classification techniques for diagnosis of diabetes: a review. *International Journal of Biomedical Engineering and Technology*, 21, 15-39. doi: 10.1504/ijbet.2016.076730
- [20] Choubey, D. K., & Paul, S. (2017). GA_RBF NN: a classification system for diabetes. *International Journal of Biomedical Engineering and Technology*, 23, 71-93. doi: 10.1504/ijbet.2017.082229
- [21] Choubey, D. K., & Paul, S. (2015). GA_J48graft DT: a hybrid intelligent system for diabetes disease diagnosis. *International Journal of Bio-Science and Bio-Technology*, 7, 135-150. doi: 10.14257/ijbsbt.2015.7.5.13
- [22] Ephzibah, E. P. (2011). Cost effective approach on feature selection using genetic algorithms and fuzzy logic for diabetes diagnosis. *arXiv preprint arXiv:1103.0087*. doi: 10.5121/ijsc.2011.2101
- [23] Han, F., & Zhu, J. S. (2013). Improved particle swarm optimization combined with backpropagation for feedforward neural networks. *International Journal of Intelligent Systems*, 28, 271-288. doi: 10.1002/int.21569
- [24] XU, J. (2008). Hybrid genetic ant colony algorithm for traveling salesman problem. *Journal of Computer Applications* 28, 2084-2087. doi: 10.3724/sp.j.1087.2008.02084
- [25] Jayashree, J., & Kumar, S. A. (2019). Hybrid swarm intelligent redundancy relevance (RR) with convolution trained compositional pattern neural network expert system for diagnosis of diabetes. *Health and Technology*, 1-10. doi: 10.1007/s12553-018-00291-3
- [26] Cao, F., Che, S., & Zhao, J. (2017). A weighted hybrid training algorithm of neural networks for robust data regression. *International Journal of Machine Intelligence and Sensory Signal Processing*, 2, 51-66. doi: 10.1504/ijmissp.2017.10009124

- [27] Suganuma, M., Kobayashi, M., Shirakawa, S., & Nagao, T. (2019). Evolution of Deep Convolutional Neural Networks Using Cartesian Genetic Programming. *Evolutionary computation*, 1-23. doi: 10.1162/evco_a_00253
- [28] Guzmán, Juan Carlos, et al. "Optimal genetic design of type-1 and interval type-2 fuzzy systems for blood pressure level classification." *Axioms* 8.1 (2019): 8.
- [29] Miramontes, Ivette, et al. "Optimal design of interval type-2 fuzzy heart rate level classification systems using the bird swarm algorithm." *Algorithms* 11.12 (2018): 206.
- [30] Melin, Patricia, Ivette Miramontes, and German Prado-Arechiga. "A hybrid model based on modular neural networks and fuzzy systems for classification of blood pressure and hypertension risk diagnosis." *Expert Systems with Applications* 107 (2018): 146-164.
- [31] Sánchez, Daniela, Patricia Melin, and Oscar Castillo. "A grey wolf optimizer for modular granular neural networks for human recognition." *Computational intelligence and neuroscience* 2017 (2017).
- [32] Sánchez, Daniela, Patricia Melin, and Oscar Castillo. "Optimization of modular granular neural networks using a firefly algorithm for human recognition." *Engineering Applications of Artificial Intelligence* 64 (2017): 172-186.