ISSN: 0974-5823 Vol. 7 No. 5 May, 2022

International Journal of Mechanical Engineering

Harris Hawk Optimizer based Enriched Deep Neural Network for early stage prediction for Diabetes Mellitus

M K Prakash[1], Dr. A. V. Ramani[2]

[1]Department of Computer Applications,

Sri Ramakrishna Mission Vidyalaya College of Arts and Science,

Coimbatore, Tamil Nadu, India

[2]Department of Computer Science,

Sri Ramakrishna Mission Vidyalaya College of Arts and Science,

Coimbatore, Tamil Nadu, India

Abstract

Predicting human diseases accurately remains a difficult task in the quest for improved and more appropriate treatment. Diabetes Mellitus is one of the most serious diseases, and it affects a large number of people. Diabetics are at a higher risk of developing ailments such as nerve damage, cardiovascular disease, diabetic retinopathy, stroke, vision problems, and so on. Machine learning has been applied to a variety of facets of medical health as a result of its rapid progress. Many literature works are in existence to predict the diabetes disease, the class imbalance is the major drawback to achieve highest accuracy of detection. While using deep learning models, the overfitting and weight decay are major issues which affects the prediction accuracy of diabetes. To overcome this problem and to improve the accuracy of the diabetes prediction, this proposed work focuses on developing behavioural based Deep Neural Network (DNN). The hyperparameters are the major factor which influences the learning rate of the deep neural network, hence in this proposed model Harris Hawk optimizer (HHO) is used for fine tuning parameter values of deep neural network with its prey searching strategy. The batch normalization is deployed to improve the quality of the raw dataset collected from Pima Indian Diabetes. The results proved that, the proposed Harris Hawk Optimizer based Deep Neural Network produced highest rate of accuracy 96.75% compared to Radial Basis Function, Support Vector Machine and KNN in early stage prediction of diabetes mellitus.

Keywords: Diabetes Mellitus, Deep Neural Network, Harris Hawk Optimizer, hyperparameter , prediction, behavioral model.

Introduction

Diabetes Mellitus is a kind of chronic condition or collection of metabolic diseases in which a person has an elevated blood glucose level in the body due to insufficient insulin synthesis or a failure of the body's cells to respond effectively to insulin [1]. In India, one in six people got affected by diabetes and its projected to increase constantly as stated by International Diabetes Federation [2]. Early detection and treatment of diabetes are the vital steps in preventing complications to other diseases like cardiovascular disorder, diabetic retinopathy, etc. Researchers and developers working on diabetes predictive models have two major obstacles due to the large number of studies that have been published in the field [3]. To begin with, past research has employed a wide range of machine learning algorithms, making it difficult to determine which is the best. Second, there is a lack of openness regarding the features involved to train the models, which decreases their interpretability and features are utterly important factor to doctors.

With the advent of big data and artificial intelligence it reshapes disease and patient care with a shift toward individualized diagnosis and therapy. As a result of this transformation, public health can become more proactive and preventative. In recent years, Deep Learning (DL) [4] is a novel research approach in machine learning, and it has made significant progress in disease prediction. Machine perception was originally used to construct the neural network. It differed from machine perception except that it brought together numerous hidden levels. The feature level grows as the network depth increases, which improves the model's representation capabilities [5]. Every layer in a deep learning model reflects a level of learnt knowledge. The layer closest to the input layer represents data at the lowest level, whereas the layer closest to the output layer reflects a higher level of differentiation with a more abstract principle. The model can be used for downscaling, regression, classification and clustering, the output layer can have several outputs.

This proposed work intends to present a novel deep neural network-based model for diabetes prediction by adapting the Harris Hawk optimizer to overcome the problem of hyperparameter fine tuning for combat against overfitting problem.

Related Work

Umair et al [6] developed IoT based diabetes prediction model to identify its presence at the early stage. The IoT based sensor is used to monitor the health condition of the person. They used three various classification algorithm and predictive analysis is also conducted using deep learning model.

Qawqzeh et al. [7] presented a photoplethysmogram based diabetes detection by applying logistic regression algorithm. This model classifies the nondiabetic and diabetic patient with cross validation model.

Gupta et al [8] in their work adapted machine learning algorithms for diabetes prediction. The k fold cross validation is applied to train and test the models. The feature selection is also done in order to improve the classification accuracy by considering the high information gain attributes.

Baha Ihnaini et al [9] designed a deep learning model along with data fusion approach for diabetes prediction. This method works as the recommendation system for smart health care. The irrelevant problem of computation complexity is handled using the data fusion model. The dataset is trained using ensemble machine learning algorithms to predict the diabetes.

Hasan et al [10] in their work utilized various algorithms such as Adaboost, SVM, random forest, kNN, decision tree for diabetes prediction. They used different parameters for finding the presence of diabetes at its early stage.

Arora and Pandey [11] adapted different deep learning algorithms for detecting the diabetic retinopathy. The dense layer of the DNN and CNN extracts the interest patterns and applied the activation function to predict the presence of diabetic retinopathy.

Khanam et al [12] constructed different layers of neural network with varying epoch to predict diabetes. The dataset is collected form UCI repository. They stated that SVM and logistic regression achieves more accuracy compared to other state of art techniques.

Nazin Ahmed et al [13] focused on developing an effective classification approach to predict the diabetes by utilizing the clinical dataset. They improved the quality of dataset by applying pre-procesisng, label encoding and normalization. Several classification models re used to predict the performance in diabetes detection. They detected and ranked various risk factors related to diabetes by applying the feature selection models.

Sharma and Shah [14] reported in their survey about the usage of machine learning models and deep learning models for predicting the presence of diabetes. They discussed about the data inadequacy and its related challenges in detail. The parameters used for analysing the performance of the classifiers is mean absolute error, root mean square error and ROC.

Bothra and Rishab [15] designed big data approach for discovering the hidden pattern of information, unknown connections, allowing us to learn from the data and forecast the outcome which assist the healthcare industry to predict the presence of diabetes with huge volume of datasets. They adapted different machine learning methods to determine which algorithm produced best accuracy.

Jian et al [16] used various supervised classification models to forecast and categories eight diabetes problem. To deal with missing values and uneven data, some critical preprocessing processes were used. In addition, the top 5 and 10 attributes for each complication were chosen using feature selection. They used binary classification model to detect each complication related to diabetes.

Preliminaries

Deep Neural Network (DNN) is a sort of Artificial Neural Network that uses numerous layers as an intermediator between the input and output layers [17]. In general, neural networks are inspired by human brain functionality, and so they are made up of neurons as nodes, weights, and bias as parameters that supply signal to the link between one layer and the next. A multifaceted non-linear association is created using DNN. This design generates a compositional manner in which items are represented as layers. As indicated in the picture, DNN works as a feedforward network, with data flowing from the input layer to the intermediate levels and lastly toward the output layers without backtracking.

The DNN creates a virtual neurons map and assigns arbitrary weighting factors to the links seen among layers preceding and following [18]. The weights and input values are multiplied and added to use the activating function to create an output between 0 and 1. The observed output is compared to the expected output, and if there is a discrepancy, it is called an error. A backpropagation method is used to alter the weight values, such that the parameters become more influential until the accepted rate of accuracy is discovered. The DNN with input layer, numerous intermediate dense layers, and output layer as depicted in the diagram.

Methodology: Harris Hawk Optimization based Deep Neural Network for Diabetes Miletus Prediction

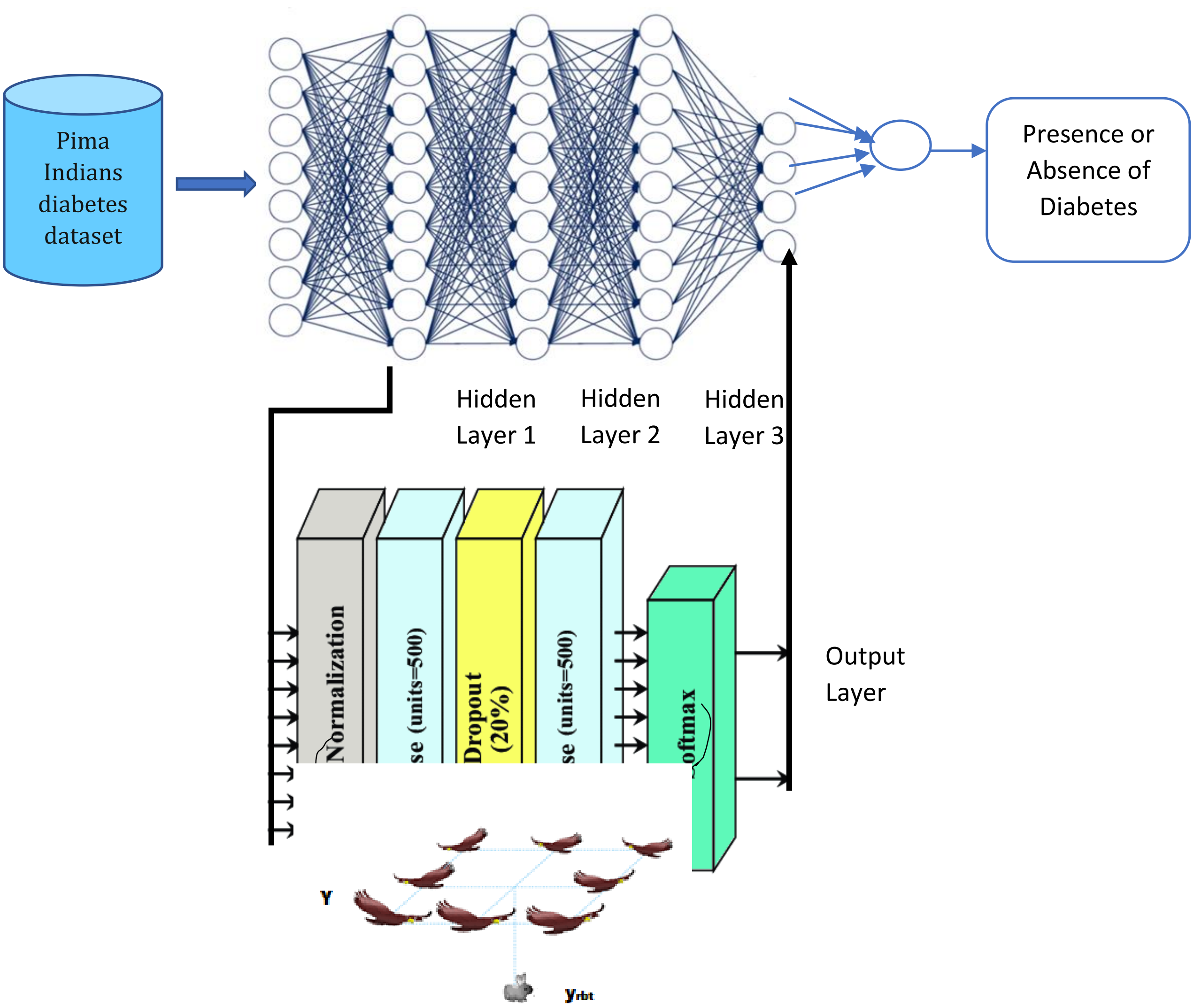

Figure 1 Overall Architecture of the HHO-DNN for Diabetes Prediction

The figure 1 illustrates the work flow of the proposed Harris Hawk Optimizer based Deep Neural Network for diabetes prediction. The dataset is collected from Pima Indians Diabetes dataset [21]. The raw dataset is fed as input to the deep neural network. Once the input layer receives the data, the batch normalization is applied to convert the attribute values to fall under same range of values. The hidden layers or dense layer performs the pattern discovery of the input data with the weight and bias parameter assigned to link which connects the two hidden layers. The learning rate and the weight decay problem is overcome by adapting the prey hunting behavior of Harris Hawk optimizer. The detailed working process of the proposed HHO-DNN is discussed in the following sections.

Dataset Description

To perform diabetes prediction, this work used Pima Indians Diabetes dataset collected from Kaggle repository [21]. This dataset comprised of 768 records with 8 independent variables and one depend variable. The detailed description is shown in the table 1.

Table 1: Pima Indians Diabetes Dataset Description

1 Number of times pregnant (NTP)	Numerical values
2 Plasma glucose concentration (PGC)	Numerical values
3 Diastolic blood pressure (DBP)	Numerical values in (mm Hg)
4 Triceps skin fold thickness (TSFT)	Numerical values in (mm Hg)
5 2-h serum insulin (2-HIS)	Numerical values in mm
6 Body mass index (BMI)	Numerical values in (mu U/ml)
7 Diabetes pedigree function (DPF)	Numerical values in (weight in kg/(height in m)^2)
8 Age	Numerical value
9 Diagnosis of diabetes	Yes = 1, No = 0

Limitations of Deep Neural Network

Overfitting is a common problem with the DNN model, which happens when the difference in accuracy is quite large. The DNN model remembers the outcomes from the training data and is unable to apply them to new data.

Deep learning models, unlike machine learning models, have a plethora of hyper factors that govern the network topology i.e number of hidden nodes and the variables that describes how the model is trained. The number of epochs was assigned as ten. After the parameter modification, the batch size is the number of sub samples sent to the network is set to 32. The loss function is computed using binary cross entropy, it computes the loss factor by averaging the class wise error and measuring how far away from the true value the forecast is for each of the classes. The standard DNN use the gradient descent for assigning and updating the values of weight, without considering any external knowledge. Thus, its random nature in assignment of hyperparameter may leads to overfitting and the model performance will be totally dropped.

The simplest and perhaps most common regularization method is to add a penalty to the loss function in proportion to the size of the weights in the model. Hence in this work weight regularization is focused to overcome the issue of overfitting in DNN. This would enable the network to transfer the inputs to the training dataset's outputs in a way that keeps the model's weights low. Weight regularization is a technique that has been demonstrated to be particularly active for both neural network and linear model. Hence, hyper parameter fine tuning is done in our suggested model to assist combat overfitting by incorporating the expertise of the Harris Hawk Optimizer algorithm.

Harris Hawks optimizer Method

Harris Hawks is a community strategy influenced by crowd intelligence, its major merits are being cooperative behaviour and surprise pounce hunting elegance nature [19]. During the hunting of the target each search agent or hawks jointly pounce prey by different positions as depicted in the figure 3. This methodology is implemented for finding the best target in the complex search space.

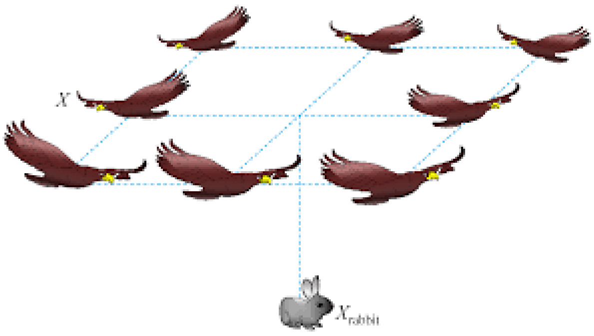

Figure 3 Harris Hawk Prey searching

Exploration stage. On the basis of the following equations, each search member randomly visits each location and waits to find a target

$$Y(s+1) = \begin{cases} Y_{rnd}(s) - g1|Y_{rnd}(s) - 2g_2Y(s)| & p \geq 0.5 \\ (Y_{rbt}(s) - Y_m(s)) - g_3(LB + g_4(UB - LB)) & p < 0.5 \end{cases}$$

The position vector of the hawk (search agent) is $Y(s+1)$ at the period s+1, the next generation of rabbit is Y_{rbt}. The recent position of search agent is Y(s), random numbers are p and g_1, g_2, g_3, g_4 whose values lies between (0,1). These values are updated during each iteration, the lower and the support bound for the search agent is defined by the parameters UB and LB. $Y_{rnd}(s)$ is the random search agent chosen and the current crowd's mean position is signified as $Y_{m.}$ The average position of the searching agent is formulated as

$$Y_m(t) = \frac{1}{N}\sum_{i=1}^{N} Y_i(s)$$

Where $Y_i(s)$ is the present location of each searching hawk, s is the iteration and N is the total number of search agent population.

Exploitation phase: During this phase the prey's energy is determined using the formula

$$Eg = 2Eg_0(1 - \frac{t}{Tr})$$

Whre Eg, T and Eg_0 are the target;s eschaping enrgy, total generations and energy's initial state as shown in the figure 4.

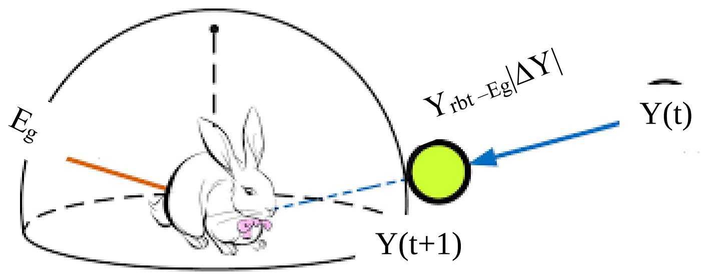

Fig :abbit) position and energy computation

Exploitation phase: There are two behavior's that must be simulated during the exploitation phase [20]. In soft besiege is the first behaviour, in which the rabbit's energy is still strong and it can flee quickly. In this case, Harris hawks try to follow it slowly and keep an eye on it until it becomes fatigued. While in hard besiege, the prey in this activity is exhausted and unable to flee. As a result, in this mode, the Harris hawks form closed circles to launch a surprise attack.

While $Eg \geq 0.5$ and $er \geq 0.5$, this condition signifies that the prey or rabbit has relatively high escaping energy (Eg) and the change of successful escape (er) is greater than 50%. Then Harris hawk exhibits soft besiege behaviour, the search agent's behaviors are accomplished using the following formula

$$Y(s+1) = \Delta Y(s) - Eg|JY_{rbt}(s) - Y(s)$$

$$\Delta Y(s) = Y_{rbt}(s) - Y(s)$$

The random jump strength of rabbit based on its escaping procedure is represented using J = 2(1-g5), $\Delta Y(s)$ refers to variation among location vector of the rabbit, g5 is a random number and s is the current position in the generation. To simulate the rabbit's nature, the value of J will be updated randomly during each iteration.

When $Eg \geq 0.5$ and $er < 0.5$ the it indicates that the rabbit has high energy. however, the likelihood of successfully fleeing is low, Harris hawk perform progressive soft besiege with rapid dives. The next move of the hawks is mathematically updated as

$$X = Y_{rbt}(s) - Eg|Y_{rbt}(s) - Y(s)|$$

In Hard besiege the current position of all agents are updated as

$$Y(s+1) = Y_{rbt}(s) - Eg|\Delta Y(s)|$$

The hawks will then determine which dive is better by comparing the present position to the prior dive. The hawks will take the previous dive if it is better. Otherwise, the hawks will use the levy flight to make a new dive.

$$H = X + K * LF(D)$$

X is the present location, K is the random vector, D is the dimension of the problem and LF is the levy flight. In both soft besiege and hard besiege behavior, it is necessary to replace the positions of all members.

$$Y(s+1) = \begin{cases} X, & F(X) < F(Y(s) \\ H, & F(H) < F(Y(s) \end{cases}$$

Harris Hawk Optimizer based Deep Learning Model for prediction of early diabetes mellitus

In this proposed work the prediction of diabetes mellitus is accomplished by developing a metaheuristic algorithm integrated deep learning model. The DNN comprised of eight input nodes which is fed directly into the network, it has five hidden layers are used with decreasing number of nodes in the succeeding layers as show in the figure 5.

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 10)	90
dense_2 (Dense)	(None, 50)	550
dense_3 (Dense)	(None, 10)	510
dense_4 (Dense)	(None, 5)	55
dense_5 (Dense)	(None, 1)	6

Total params: 1,211 ; Trainable params: 1,211

Figure 5 Model of the Deep Neural Network in prediction of Diabetes Mellitus

The attribute vector of pima dataset is passed directly to the input nodes of the deep network. The hidden layer comprised of hidden nodes where each of the node combined with activation function and linear combination to produce output, which in turn connected to the succeeding hidden layer. Various layers have different activation functions. The features are then obtained and merged to create a new feature vector. To determine the confidence of each relationship, the softmax classifier obtains the latest feature representation. The output vector can be obtained via the classifier. The number of classes is the output vector's dimension, and the confidence of each classification is equal to the quantity of every aspect. The input feature passes through the input nodes at the bottom of the deep learning network during the training phase, where the weights are initialized with random values. The weight vectors are then fine-tuned one by one by applying the Harris Hawk optimizer. The fundamental purpose of the training is to reduce the process loss function while increasing its accuracy function of diabetes prediction.

Algorithm Harris Hawk Optimizer based Deep Neural Network (HHO-DNN) for prediction of diabetes

Input: D = ($I_1, I_2, .. I_n$) represents the matrix of diabetes mellitus dataset with n instances

CL = ($cl_1, cl_2, .., cl_n$) class variable of their corresponding instances

M is the number of attributes in the dataset

Tr: Training dataset , CS = Candidate dataset, L = Tr $\cup$ CS

IW : Input weight , TW = weight of training set, CW = weight of the candidate set GIW = gradient weight corresponding to IW

Procedure

Begin

1. Initialize Bias learning rates = 0.01, Learning Rate= 0.1, Momentum= 0.9
2. While |Tr| $\leq$ M+1 do
3. Assign CW = 0
4. Update hidden layer weight and Input weight IW
a. Call HHO ()
5. Obtain average after multiple dropout
6. Compute O = $arg \max_{cs \in CS} \|GR_{L_{cs}}\|$
7. Update learning rate using HHO ()
8. Conduct TR = TR $\cup$ L_O and CS = CS / L_O

End

Output: presence or absence of diabetes

Algorithm 2: Harris Hawks Optimizer for hyperparameter regularization

Input: Initialize Population size P, Maximum Iteration T

Ouput: Best Weight vectors and its fitness value

Begin

1. Initialize hawks location Y_i (i.1,2,..,n)
2. While (stopping criteria is not met) do

3. Calculate the fitness values of each hawks
4. Set the best position for Y_{rbt} // prey
5. For all search agent (Y_i)

{

Assign the initial energy Eg_0 =(2 * rnd()-1)

Set Jump strength J = 2(1-g5)

Update the energy using equation

$$Eg = 2Eg_0(1 - \frac{t}{Tr})$$

IF $Eg \geq 1$ then
Modify the position using the equation as follows

$$(s+1) = \begin{cases} Y_{rd}(s) - g1|Y_{rnd}(s) - 2g_2Y(s)| & p \geq 0.5 \\ Y_{rbt}(s) - Y_m(s)) - g_3(LB + g_4(UB - LB)) & p < 0.5 \end{cases}$$

If $Eg < 1$ then
If rnd ≥ 0.5 and $Eg \geq 0.5$ then

Modify the position as

$$Y(s+1) = \Delta Y(s) - Eg|JY_{rbt}(s) - Y(s)$$

$$\Delta Y(s) = Y_{rbt}(s) - Y(s)$$

If rnd ≥ 0.5 and $Eg < 0.5$ then
Modify the position vector as

$$Y(s+1) = Y_{rbt}(s) - Eg|\Delta Y(s)|$$

If rnd < 0.5 and $Eg \geq 0.5$ then
Compute the position vector as

$$Y(s+1) = \begin{cases} X, & F(X) < F(Y(s) \\ H, & F(H) < F(Y(s) \end{cases}$$

}

End {while}

Return Y_{rbt} {weight}
End

The algorithm describes that the hyperparameter of DNN is fine tuned to overcome the overfitting problem in diabetes prediction by adapting Harris Hawk optimizer. The weight vectors are initialized as the rabbits in the HHO algorithm. Initially sample of hawks are selected as search agents and searchers the location of the prey, and depending on the energy of the rabbit the hawks changes its strategy of movement and attack the prey. The possible best weight values are discovered by the hawks using the fitness value of each weights in the vector and the best suited are passed to the DNN. Thus, the proposed (HHO-DNN) produced better accuracy in early prediction of diabetes.

Results and Discussions

In this section, the performance of the proposed Harris Hawk optimizer based deep neural network (HHO-DNN) for diabetes prediction is discussed in detail. The proposed model HHO-DNN is implemented using python code. The dataset used for diabetes prediction is collected from Pima Indians diabetes data from the Kaggle repository [20]. The dataset comprised of 768 records with 8 attributes and one class variable. The HHO-DNN is compared with the standard classification models namely kNN, SVM and RBF. The evaluation metrics used for assessing the performance of diabetes prediction is accuracy, precision, recall and f-measure.

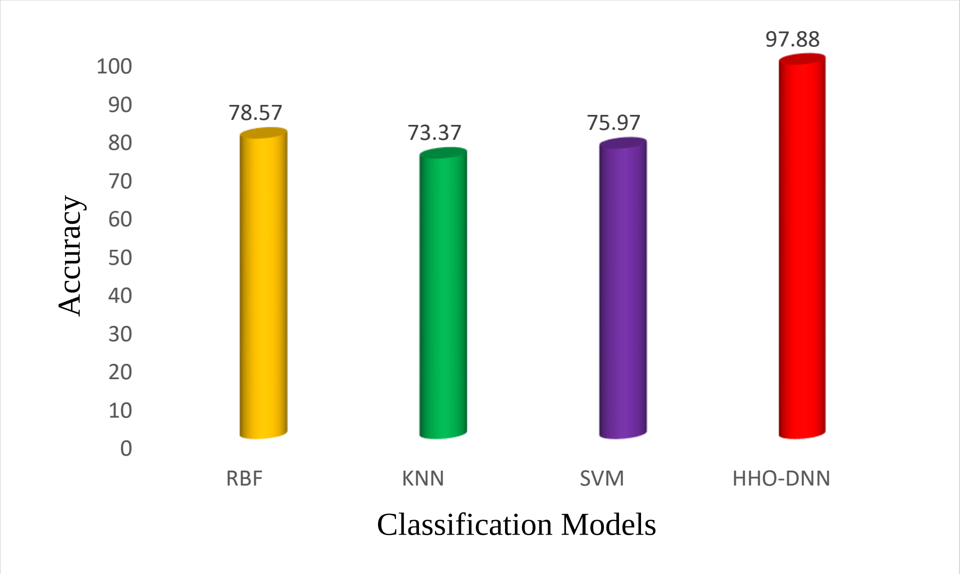

Figure 6 comparative analysis based on accuracy

The figure 6 explores the accuracy-based performance comparison of four different classification models to predict diabetes diseases. It is examined from the result that proposed HHO-DNN Network achieves highest rate of accuracy by predicting highest number of patients with or without diabetes. The existing algorithms produce less accuracy as they suffer from class imbalance problem. Whereas proposed DNN improves its learning rate by intellectually handling small data size of instances with presence of diabetes using Harris Hawk optimizer algorithm.

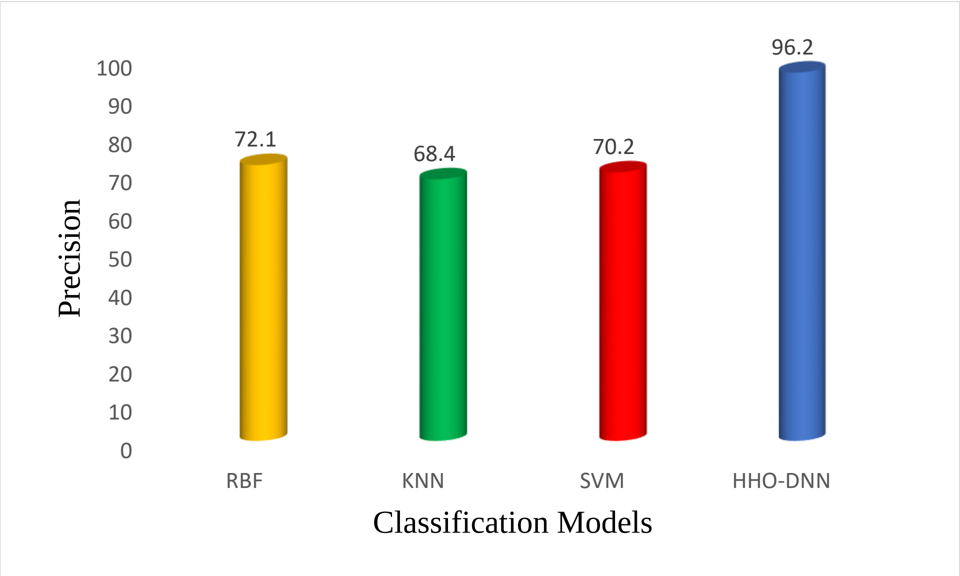

Figure 7 comparative analysis based on precision

The precision rate generated by four various classification algorithms to classify the presence or absence of diabetes is displayed in the figure 7. To handle the problem of overfitting and class imbalance among instances of normal and diabetes patients. This work instead of using the gradient descent algorithm for training the datasets, it uses intelligence for understanding the smallest set of instances with equal importance and achieves best outcome in prediction of diabetes patients with highest rate of correctness. Hence, proposed HHO-DNN attains best precision value compared to the other existing models. The HHO algorithm fine tunes the hyperparameters of DNN based on the fitness values and only the best suited values are assigned to the parameters.

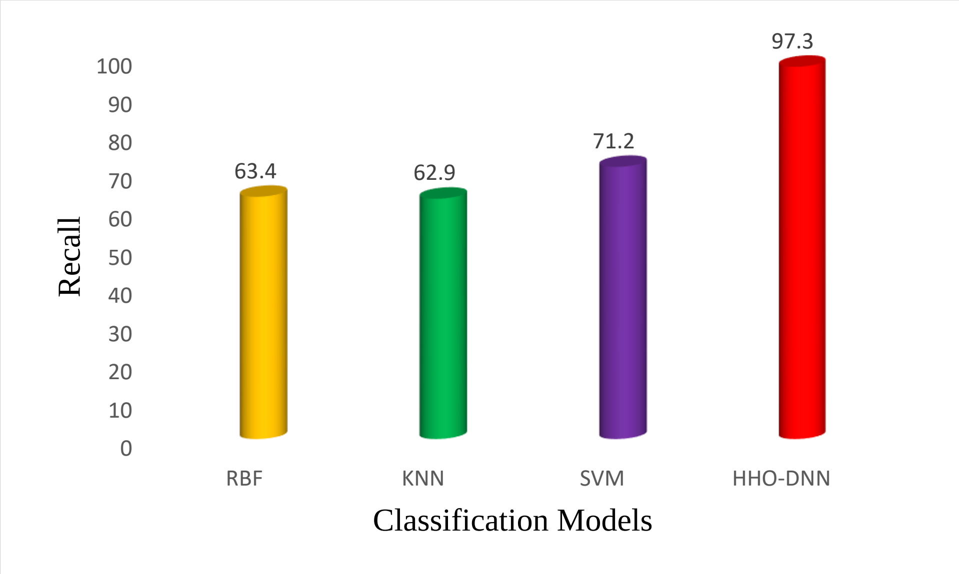

Figure 8 comparative analysis based on Recall

The figure 8 shows recall rate generated by proposed HHO-DNN and three standard classification models for predicting diabetes. The highest ratio of diabetic presence is accomplished by the newly devised DNN model. In this work the normalization. Overfitting is a problem that predictive models frequently confront. When the disparity in accuracies is quite large, overfitting occurs. The HHO-DNN model remembers the outcomes from the training data and is unable to apply them to new data. This proposed work added normalization layers in DNN to assist combat overfitting and the Harris Hawk searching strategy of its prey is used in this work to select the best suitable values for the parameters involved in DNN for prediction of diabetes. Thus HHO-DNN achieve better recall value compared with other three classification models.

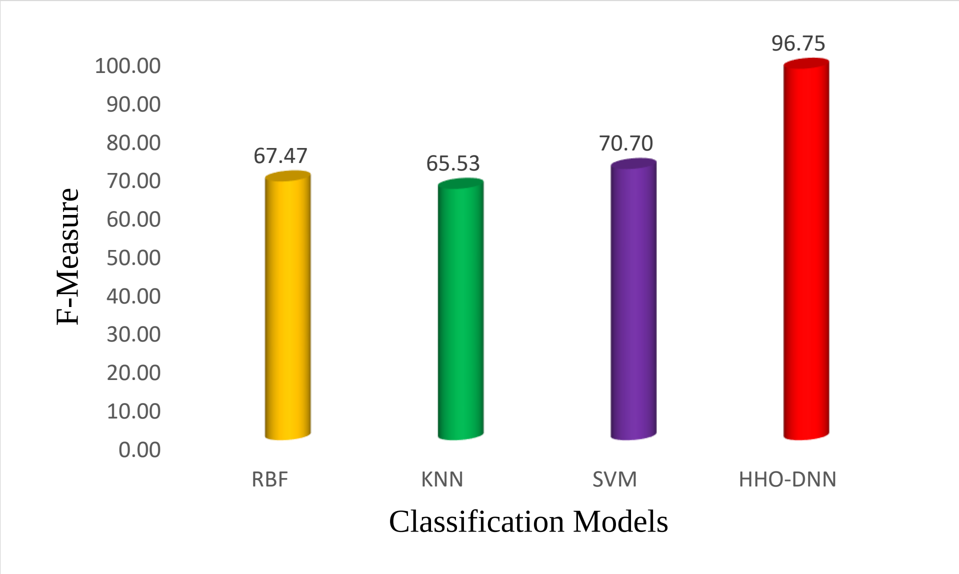

Figure 9 comparative analysis based on F-Measure

The F-measure of four various classification models involved in diabetes patient prediction is exposed in the figure 9. F-measure reflects the influence of precision and recall value, proposed HHO-DNN generates highest rate of F-measure compared to other three diabetic prediction model. The nature of soft and hard besiege searching strategy the hawks which works as a search agent and choose the best fittest value for the parameters in DNN to predict the presence or absence of diabetes more precisely.

Conclusion

This paper focuses on developing an enriched deep neural network for predicting the diabetes presence at its early stage to improve the life quality of the persons. In standard deep learning and machine learning algorithms, selection of parameter values creates a chaos effect in accurate detection. The proposed work adapts the intelligence of Harris Hawk prey searching behavior which balances both local and global optima by applying soft and hard besiege. The deep neural network learning rate is improved during the training phase of the model by handling the weight decay and overfitting problem. The Harris Hawk optimizer assigns augmented values to the parameters involved in process of predicting diabetes. The simulation results proved that the performance of the proposed HHO-DNN accomplishes better accuracy rate while comparing with other state of arts in diabetes prediction.

References

1. Abdulhadi, N.; Al-Mousa, A. Diabetes Detection Using Machine Learning Classification Methods. In Proceedings of the 2021, International Conference on Information Technology (ICIT), Amman, Jordan, 14–15 July 2021; pp. 350–354
2. S. Bae, T. Park, Risk prediction of type 2 diabetes using common and rare variants. Int. J. Data Min. Bioinform. 20(1), 77–90 (2018)
3. Kumar P.S, Umatejaswi V, Diagnosing Diabetes using Data Mining Techniques. International Journal of Scientific and Research Publications 7,705–709, 2017.
4. Tejas N. Joshi, Prof. Pramila M. Chawan, "Diabetes Prediction Using Machine Learning Techniques. Int. Journal of Engineering Research and Application, Vol. 8, Issue 1, 2018, pp.-09-13
5. Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, and H. Tang, "Predicting diabetes mellitus with machine learning techniques," Frontiers in Genetics, vol. 9, 2018.
6. Umair Muneer Butt, Sukumar Letchmunan, Mubashir Ali, Fadratul Hafinaz Hassan, Anees Baqir, Hafiz Husnain Raza Sherazi, "Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications, Journal of Healthcare Engineering, vol. 2021, Article ID 9930985, 17 pages, 2021.
7. Y. K. Qawqzeh, A. S. Bajahzar, M. Jemmali, M. M. Otoom, A. Thaljaoui, "Classification of diabetes using photoplethysmogram (PPG) waveform analysis: logistic regression modeling," BioMed Research International, vol. 2020, 6 pages, 2020.
8. S. Gupta, H. K. Verma, and D. Bhardwaj, "Classification of diabetes using Naïve Bayes and support vector machine as a technique," Operations Management and Systems Engineering, Springer, Singapore, 2021.
9. Hasan M.K, Alam M.A, Das, Hossain, Hasan M, Diabetes Prediction Using Ensembling of Different Machine Learning Classifiers. IEEE Access 2020, 8, 76516–76531
10. Baha Ihnaini, M. A. Khan, Tahir Abbas Khan, Sagheer Abbas, Mohammad Sh. Daoud, Munir Ahmad, Muhammad Adnan Khan, "A Smart Healthcare Recommendation System for Multidisciplinary Diabetes Patients with Data Fusion Based on Deep Ensemble Learning", Computational Intelligence and Neuroscience, vol. 2021, 11 pages, 2021
11. M. Arora, M. Pandey, "Deep neural network for diabetic retinopathy detection," in Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), pp. 189–193, Faridabad, India, February 2019.
12. Khanam, Jobeda Jamal Foo, Simon. (2021). A comparison of machine learning algorithms for diabetes prediction. ICT Express. 7. 10.1016/j.icte.2021.02.004.
13. Nazin Ahmed, Rayhan Ahammed, Manowaru lIslam, Ashraf Uddin, Arnisha Akhter, Al-Amin Talukder, Bikash Kumar Paul, Machine learning based diabetes prediction and development of smart web application International Journal of Cognitive Computing in Engineering Vol 2, 2021, 229-241.
14. Sharma, T., Shah, M. A comprehensive review of machine learning techniques on diabetes detection. Vis. Comput. Ind. Biomed. Art 4, 30 (2021).
15. Bothra, Rishab. (2021), Diabetes Prediction Using Machine Learning Algorithms, International Journal of Engineering Applied Sciences and Technology, 2021, Vol. 6, Issue 5, ISSN No. 2455-2143, Pages 151-154
16. Jian Y, Pasquier M, Sagahyroon A, Aloul, F. A Machine Learning Approach to Predicting Diabetes Complications. Healthcare 2021, 9, 1712.
17. Barbedo, J. G. A. (2018b). Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Comput. Electron. Agric. 153, 46–53
18. Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., et al. (2018). Opportunities and obstacles for deep learning in biology and medicine. J. R. Soc. Interface. 15:20170387
19. Heidari, A.A.; Mirjalili, S.; Faris, H.; Aljarah, I.; Mafarja, M.; Chen, H. Harris Hawks optimization: Algorithm and applications, Future Gener. Comput. Syst. 2019, 97, 849–872
20. Alabool, H.M., Alarabiat, D., Abualigah, L, Harris Hawks optimization: a comprehensive review of recent variants and applications. Neural Computer Applications 33, 8939–8980 (2021).
21. https://www.kaggle.com/uciml/pima-indians-diabetes-database