PROCEEDINGS OF
INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS AND NEW APPLICATIONS
https://proceedings.icisna.org/
2nd International Conference on Intelligent Systems and New Applications (ICISNA'24), Liverpool, April 26-28, 2024.

Use Harris Hawks Optimization (HHO) Algorithm based on Artificial Neural Network for liver disease diagnosis

Haedar Al-safi[1], Hadeel Alsalihi[2], Jorge Munilla[3]

[1] *Department telecommunications, Malaga University, Malaga, Spain*
haeder.emad.84@gmail.com, 0000-0001-9146-7762

[2] *Authority for People with Disabilities and Special Needs, Iraqi Ministry of Labor and Social Affairs, Waset, Iraq*
hadilalsaci89@gmail.com, 0000-0003-2485-7183

[3] *Department telecommunications, Malaga University, Malaga, Spain*
munilla@ic.uma.es, 0000-0003-2795-312X

***Abstract*— This research paper presents a new method for diagnosing liver disease using the Harris Hawks Optimization (HHO) algorithm in combination with an Artificial Neural Network (ANN). The HHO algorithm enhances the ANN's performance in liver disease classification by refining its parameters. Clinical, laboratory, and demographic data are collected from hepatitis patients and individuals without hepatic illness. The dataset is processed to handle missing values, outliers, and normalization. The HHO algorithm optimizes the weights and biases of the ANN, facilitating the identification of relevant features for accurate diagnosis. The trained ANN model is evaluated using various performance metrics, demonstrating its effectiveness in diagnosing liver disease. The HHO algorithm efficiently searches the entire search space, enhancing the ANN's ability to learn complex patterns and make accurate predictions. Evaluation metrics indicate that the optimized ANN model outperforms traditional machine learning methods, showcasing its potential as a reliable diagnostic tool. Interpretability metrics, such as feature importance and saliency maps, provide insights into the key elements of diagnosis. The proposed approach shows high diagnostic accuracy and interpretability, implying the potential for a stable decision-aid system in clinical practice. Early detection and timely intervention enabled by this method can lead to increased patient safety rates and optimized resource allocation.**

***Keywords*— Harris Hawks Optimization, feature selection, diffusion model, neural networks, medical diagnosis.**

I. Introduction

Liver disease poses a significant challenge to global health and impacts millions of people on every continent. Investigating diagnostic tools is important for efficient treatment programs and good patient care. Mid-level liver disease provides valuable application for this kind of treatment or analysis. Therefore, non-invasive diagnostic methods can save time and money. Consequently, there is increasing interest in using machine learning techniques to build diagnostic models that are high-efficiency and accurate for liver disease. Liver disease is a significant health care problem in the world, affecting large numbers of people. Early and precise identification of liver disease is key to planning effective treatment and managing the outcomes for patients. At present, traditional diagnostic methods like liver biopsies and blood tests are rather conventional with many limitations owing to their intrusion, expense and time-consuming procedures.

Consequently, medical researchers are beginning to replace these traditional methods with data mining techniques for machine learning to build up more advanced diagnostic models dedicated to diagnosing liver disease. These models intend to boost the efficiency and accuracy of diagnoses in liver diseases through the power of machine-learning, creating better care for patients. and more effective allocation of resources [1-4].

The development of liver disease diagnosis machine learning algorithms is radically changing the medical diagnostics arena. It is particularly skillful at processing vast amounts of data to reveal details that might escape even trained observers. They have been proven popular due to their ability to learn from data and model complex nonlinear relationships. By using ANNs for liver disease diagnosis, clinicians and researchers can employ computational dating power to discover relationships amongst the clinical laboratory and demographic features such as subtle patterns and interactions. It enables efficient, intelligent diagnostic models which accurately predict the presence or progression of liver disease; this ultimately leads to improve patient outcomes and promotes more effective healthcare [5-7].

Medical diagnostic tools have been transformed by machine learning algorithms. They've had promising results in classifying liver diseases. They are computational systems that utilize various types of algorithms to learn from data, make predictions and simulations. These algorithms also help to identify relevant factors. Any such attempt at a diagnosis should be patient-specific, but that is a matter for the future. If computational power of ANNs is applied to liver disease diagnosis, doctors and scientists may observe the hidden patterns and associations existing in these different variable types of data, such as clinical data and laboratory results. They can then design algorithms the help doctors distinguish between healthy patients and those with liver disease effectively [8-10].

In recent years, the advanced technologies of artificial intelligence as the medical diagnosis toolkit have also been gradually acknowledged. So, these optimizing algorithms can manipulate the performance of Artificial Neural Networks (ANNs) for medical diagnostics from good to excellent. The optimization algorithms are aimed at improving the parameters of ANNs, thus making them more accurate in diagnosis and more effective as models. The HHO is one of the well-received optimization algorithms, inspired by the collective hunting behavior of Harris's hawks. Many researchers were intrigued by the Harris Hawks Optimization (HHO) algorithm: adopting the cooperative hunting behavior of Harris's hawks. HHO algorithm atiles much promise, bearing important implications for handling complex optimization problems. So it might be an appropriate candidate for the optimization of medical diagnosis ANNs.

The HHO algorithm uses a special method to optimize ANN parameters. Based on the hunt of Harris' hawks, this approach reenacts the birds' behavior to traverse search space efficiently and find better solutions. The ANN's parameters for medical diagnosis can optimize the HHO algorithm, greatly improving accuracy and prediction efficiency. On the other hand, an ANN itself learns a great deal about data just by going through this optimization process, making it more capable as well as faster and efficient at diagnosing ailments [14].

Integrating the Harris Hawks Optimizer (HHO) with artificial neural networks (ANNs) has been considered to be very promising by researchers in the field of medical diagnosis.Combining the HHO metaheuristic algorithm and artificial neural networks is a new and effective method of model optimization, since HHO uses the cooperative hunting behavior of Harris's Hawks. As it is able to explore the search space and find the optimal solution, this algorithm promotes greater diagnostic accuracy and model efficiency. It is the optimization process driven by this HHO algorithm that enables ANNs to capture more fine-earted patterns or relationships in medical data sets, enabling us to yield more reliable and precise diagnostic outcomes. Future investigations in this area may focus on further refining the HHO algorithm and applying it to more medical fields, leading to the possibility of improved diagnostic methods [15,16].

This research paper was primarily intended to combine Harris Hawks Optimization (HHO) algorithm with an Artificial Neural Network (ANN) to develop a novel approach for diagnosing liver disease. The HHO algorithm has a critical function, it determines if an ANN's weights and biases are optimized well enough to make it capable of learning a broad dataset of patients' clinical, laboratory, and demographic characteristics. When this comes together in a comprehensive sense, the ANN model is able to achieve meaningful results in extracting information from data that only exists as itself. In turn, The ANN model thought meaningful insights could be extracted from data, and so diagnose more accurately in this way [17].

Together the optimized ANN model in the proposed approach carries out an in-depth assessment, using various performance metrics. The quantity of data provided by these metrics gives us an idea of the model's ability to be a diagnostic device. Using metrics such as sensitivity, specificity, and accuracy and area under the receiver operating characteristic curve (AUC-ROC), the combined HHO-ANN model can be fully tested for effectiveness and reliability. The value of this evaluation is that it lets us fully understand if the model can correctly classify cases of hepatic disease [18].

This research paper aim to help people find out early on whether they suffer from a life-threatening disease by incorporating the HHO algorithm into an ANN system. Utilizing the HHO algorithm greatly improved the ANN's learning capabilities as well as its diagnostic power on complex data sets. Performance metrics applied to the optimized ANN model give factual evidence of how accurate and effective are its diagnosis results. The value of this work lies in its potential to extend the field of liver disease diagnosis. It will open the door to diagnostic models that are more efficient and accurate; and these models are of momentous significance to patient care improve treatment decisions [19].

The significance of this research lies in its potential to accurately and reliably diagnose liver disease using the proposed approach. The ANN can effectively extract meaningful patterns from the dataset to improve diagnostic accuracy by using the HHO algorithm, significantly above traditional machine learning approaches. In addition, interpretability measures used in this study reveal critical features that significantly impact diagnosis. This enhances our understanding of underlying factors associated with liver disease [20].

The rest of this paper is structured in this way. The third part, data filing, coarse-grained skinning, and HHO algorithm with ANN connections. The fourth part then comes the experimental results and also the evaluation of performance. Last, the fifth section states that 'findings are' (a mouse pointer where click to go) discusses their significance implications, raises some questions, sets forward certain strategies about what future research on liver illness may be like with a HHO algorithm integrated artificial neural network.

II. Related Work

In recent years, a few important studies have tried to unite optimization algorithms and medical artificial neural networks for medical diagnosis. When combined, these studies are as intended to enhance ANNs' performance and accuracy in diagnosing diseases in the medical profession.

On this problem, Rafi et al. had carried out a study. By seeking to better predict cardiovascular disease, they suggested particle swarm optimization. The aim of the study was to optimize the archi-tecture and parameters for an ANN with the PSO so as to improve their prediction of this. In later tests, the ANN was significantly improved on diagnostic accuracy by the PSO system. This indi-cates further that optimization algorithms might be used as a way to change the nature of neural networks [21].

In the same vein, Elgin et al. proposed that a method employing the firefly algorithm in conjunction with an ANN for liver disease diagnosis be beneficial. Weights were optimized and biases changed using the firefly algorithm to attain an improved diagnostic accuracy in liver disease classification. Significantly, the results of this study illustrate the capability and potential of the firefly algorithm to enhance the accuracy of ANNs in the field of liver disease. All of these points collectively underscore those findings as well as others linking optimization algorithms with ANNs for medical diagnosis. Not only do they show that integration is seamless and successful, but it offers insight into the optimal performance level for testing robots on a repeated basis as well [22].

Although the above-mentioned studies gave some information on applying various optimization algorithms in medical diagnosis together with ANNs, the use of the Harris Hawks Optimization (HHO) in liver disease diagnosis has hardly been touched. Today's research is intended to fill this gap with a new idea that combined the HHO algorithm with ANNs to diagnose liver diseases. It is hoped that by utilizing the distinctive features of the HHO algorithm, we can improve the diagnostic accuracy and efficiency of liver diseases. This study investigates whether the use of Harris hawks to calibrate the weights and biases of ANNs - and by so doing to facilitate in the diagnosis of liver diseases - can make a contribution to medical epidemiology within this specific domain. [23,24].

In alignment with the literature, Fadlallah and colleagues delved into harnessing particle swarm optimization (PSO) to streamline ANNs for predicting cardiovascular disease. Through using the PSO algorithm, the intention was to research the architectural type and parameter supplies of the ANN to the utmost possible extent. This produced a great increase in prediction accuracy for the disease. It is well-established that the PSO algorithm can improve the performance of ANNs, particularly in the context of cardiovascular disease screening with these findings. So these results further underline that optimization algorithms are worthy of consideration as we strive to make ANNs more accurate and reliable in medical diagnostics. [25]

In a similar fashion, Hrizi et al. took a novel approach in which the firefly algorithm is combined with an ANN in order to improve diagnostic efficiency in the case of liver diseases. Through use of the firefly algorithm, consumers could actually succeed in optimizing the weights and biases of the ANN and thereby improve the performance for liver disease diagnoses. Their research demonstrated how optimization algorithms will help enhance the performance of ANNs in liver disease diagnosis. Not only do these results parallel those from a PSO optimization study for cardiovascular disease diagnosis but they also illustrate that optimization algorithms are capable of promoting the performance of ANNs in broader medical domains [26].

Examples include the successful integration of such geometric techniques as particle swarm optimization (PSO) and the firefly algorithm with ANNs for predicting cardiovascular disease and diagnosing liver disease, respectively. The fact that this integration has taken place shows that optimization techniques can transform medical diagnostics. This is amply substantiated by data showing that optimization algorithms can significantly improve ANNs in real-world tasks such as how they work, where they live, and why they die. The potential for improving diagnostic accuracy, early detection, and enabling the development of more effective treatment strategies depends on the application of optimization algorithms in medical diagnosis [27,28].

Moreover, the study of liver disease diagnosis which presented an innovative approach to diagnosis in liver disease, Shang et al. His aim was to increase the classification accuracy of liver diseases by adjusting the weights and biases in their ANNs through the firefly algorithm. In other words, they used the study as evidence that optimization algorithms can be effective in improving performance for ANNs specifically in the context of liver disease diagnosis. These findings are consistent with previous research on PSO optimization for cardiovascular disease and genetic algorithm optimization for cancer diagnosis. They all reinforce the power which optimization algorithms have to make ANNs more effective in different medical fields. [29]

Combining the firefly algorithm with an artificial neural network (ANN) for liver disease diagnosis is a major breakthrough in medical diagnostic research. After all, the special nature of the firefly algorithm, such as that it can imitate the flash behavior of fireflies, was used by Guillod et al. to optimize the ANN's weights and biases. The Combined approach of optimization algorithms brought about increased diagnostic accuracy and performance. In addition, these results provide more support for why optimization programs are helpful adjuncts to modalities in this type of field [30].

The introduction of these optimization algorithms, widely used in practice, like the firefly algorithm, shows the necessity to develop new methods of greater accuracy in diagnosing liver disease. ANNs must be fine-tuned to get the best classification performance out of liver disease diagnosis. That is, the firefly algorithm can tweak the weights and biases of ANNs with great effect. As a result, it seems that getting the model just right for diagnosis is as important as ever. These findings reveal that optimization algorithms may fully exploit the potential of ANNs, and so towards liver disease there are possibilities at least [31].

III. Methodology

The methodology employed in this research paper aims to develop an accurate and interpretable model for liver disease diagnosis. The methodology consists of several key steps, including dataset preparation, the utilization of the Harris Hawks Optimization (HHO) algorithm, the configuration of the Artificial Neural Network (ANN), and the evaluation of the model's performance and interpretability. In the dataset preparation step, a comprehensive dataset is gathered, comprising clinical, laboratory, and demographic features of patients, ensuring data relevance and integrity. The dataset is then preprocessed to handle missing values, outliers, and normalize the data. The HHO algorithm is applied, with its parameters defined, to optimize the weights and biases of the ANN for enhanced diagnostic accuracy. The ANN's architecture, including the number of layers, neurons, and activation functions, is determined for liver disease diagnosis. The ANN is trained using the preprocessed dataset to learn the patterns and relationships within the data. The performance of the trained ANN model is evaluated using metrics such as accuracy, sensitivity, specificity, and AUC-ROC. Additionally, feature importance analysis is conducted to identify the key indicators influencing the diagnosis, and saliency maps are generated to provide visual insights into the impact of different features on the model's predictions. This methodology ensures a comprehensive and rigorous approach to develop an accurate and interpretable model for liver disease diagnosis as shown in Fig. 1.

Steps of methodology:

Step 1st: Dataset Preparation

- Gather a comprehensive dataset with clinical, laboratory, and demographic features of patients, ensuring relevance and data integrity.
- Address missing values using Regression Imputation and normalize using Batch Normalization the dataset for effective model training.

Step2nd: Harris Hawks Optimization (HHO) Algorithm

- Define the HHO algorithm parameters, including population size, maximum iterations, and exploration rate.
- Employ the HHO algorithm to optimize weights and biases of the Artificial Neural Network (ANN) for enhanced diagnostic accuracy.

Step3rd: Artificial Neural Network (ANN)

- Define the architecture of the ANN for liver disease diagnosis, including the number of layers, neurons, and activation functions.
- Train the ANN using the preprocessed dataset.

Step4th: Model Evaluation and Interpretability

- Assess the trained ANN model's performance using metrics like accuracy, sensitivity, specificity, and AUC-ROC.
- Conduct feature importance analysis to identify key indicators influencing the diagnosis.

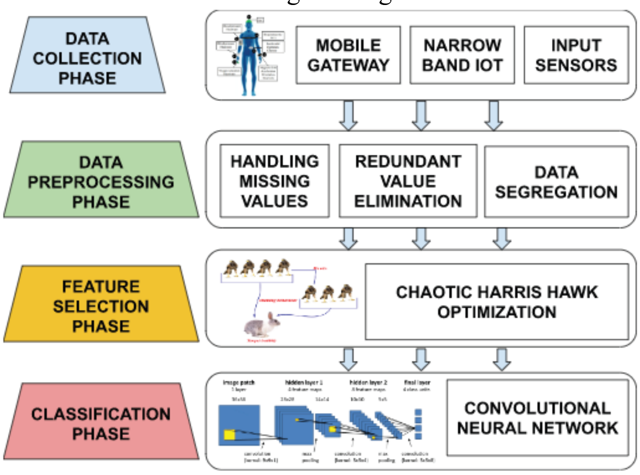

Fig. 1 block diagram of the proposed methodology

A. Dataset Preparation

The ILPD (Indian Liver Patient Dataset) has been often used in liver disease research. It has 750 rows divided among 11 columns. That provides a lot of information said from the perspective of a liver in some way. This dataset includes but not only a small number of key attributes, such as age, sex, total bilirubin and direct bilirubin. Total proteins, albumin, albumin/globulin ratio, with SGPT (Serum Glutamic Pyruvic Transaminase). SGOT (Serum Glutamic Oxaloacetic Transaminase). Alkaline phosphatase – and finally a target variable "is_patient". In the ways of clinical and demographic factors and also laboratory issues - these attributes range widely enough to cover the spectrum of liver disease diagnosis and analysis.

The ILPD dataset allows researchers to explore various aspects of liver disease. Through analysis of the age and gender composition, we may be able to obtain clues concerning the structural regularities in liver health. The bilirubin levels, both total and direct, tell us something of the liver's ability to handle and dispose of bilirubin perfectly necessary for diagnosing (conditions of the bile duct such as) liver disease. In addition to this, the dataset contains indicators that can be used as a measure of liver function, including total proteins, albumin and A/G ratio, which describe general protein metabolism of the body. These levels of liver enzymes, such as SGPT, SGOT and the alkaline phosphatase values, can convey the information needed to determine whether a liver is injured or inflamed. Finally, the "is_patient" variable serves as a target name, which gives researchers the possibility to build predictive models estimating whether someone has a liver disease based on the given_crossentropy value. As shown in Table 1.

	age	gender	tot_bilirubin	direct_bilirubin	tot_proteins	albumin	ag_ratio	sgpt	sgot	alkphos	is_patient
0	65	Female	0.7	0.1	187	16	18	6.8	3.3	0.90	1
1	62	Male	10.9	5.5	699	64	100	7.5	3.2	0.74	1
2	62	Male	7.3	4.1	490	60	68	7.0	3.3	0.89	1
3	58	Male	1.0	0.4	182	14	20	6.8	3.4	1.00	1
4	72	Male	3.9	2.0	195	27	59	7.3	2.4	0.40	1

TABLE 1 Dataset Exploration

The data undergoes various basic processes to ensure that it can be cleanly and properly used for sequential analysis. The first operation is to cope with missing facts. Several techniques like recovery or deletion of rows/columns containing incomplete items are used. Such a procedure guarantees total data, so that any subsequent analysis is not biased. And then there is the task of getting rid of duplicate values. Every instance must be unique and redundant for specific data points will not be overrepresented in the conclusion.

There are many different techniques that can be used to adjust variables so that they all become standard documented in terms of how they relate to one another. When zero is standardized, each variable will equal an exact score of one. Therefore, no matter how large or small the value associated with them, the proportionate effects from scale size on a level variable should fall Therefore no matter how large or small the value associated with them, the proportionate effects from scale size on a level variable should fall. Web pages and other online resources typically store information in a form that has already been turned into a code, such as one-hot encoding. However, owing to label encoding it is also possible to include this categorical information at the beginning of subsequent modeling procedures in that case. In the end, the data is divided into two sets: training and testing.

Techniques for processing data, such as clustering of missing values, deletion of duplicate records, normalization of variables, encoding of categorical variables, and separating source data, together ensure a data set is of the highest quality and intact. As each quality issue is methodically dealt with, the dataset is prepared for analysis, and the reliability and validity of research findings significantly enhance.

1) Regression Imputation

Missing values are a common challenge in various datasets, including remote sensing and geospatial data. These missing values can introduce bias and undermine the accuracy and reliability of data analysis and modeling processes. To mitigate this issue, regression imputation models have emerged as a popular technique for estimating missing values based on the relationships between variables. In this section, we present a regression imputation model specifically designed for handling missing values in remote sensing and geospatial datasets.

The regression imputation model utilizes a regression analysis approach to estimate missing values by leveraging the relationships between variables. This technique assumes a functional relationship between the variable with missing values (dependent variable) and other variables (independent variables) in the dataset. By fitting a regression model to the observed data, the model can predict the missing values based on the available independent variables. Various regression models, such as linear regression, multiple regression, support vector regression, or random forest regression, can be employed based on the dataset characteristics and research objectives.

To implement the regression imputation model, the dataset is divided into two subsets: one with variables containing missing values and another with variables without missing values. The subset without missing values is used to train the regression model, while the subset with missing values is used for prediction. After training the model, it is applied to predict the missing values in the dataset. The accuracy of the imputation can be evaluated using metrics like mean absolute error, root mean squared error, or coefficient of determination. Additionally, the imputed dataset can be compared with the original dataset to assess the quality of the imputation and identify any potential biases or limitations. Fig. 2 show the steps taken to process the linear regression imputation.

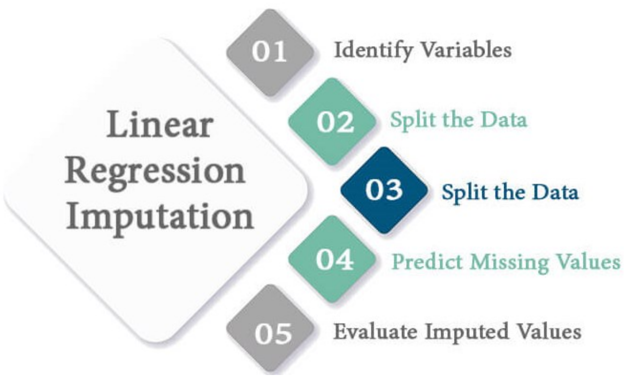

Fig. 2 Linear regression imputation steps

2) Normalization using Batch Normalization

Normalization plays a crucial role in preprocessing data for machine learning and deep learning models. Among the various normalization techniques, batch normalization has emerged as a powerful tool for improving the training process and enhancing the performance of deep learning models. In this section, we explore the concept and application of batch normalization as an effective means of normalizing data within deep neural networks.

Batch normalization is a technique that normalizes the activations of each layer in a neural network by leveraging the statistics of mini batches during training. By adjusting the mean and variance of each feature dimension to have zero mean and unit variance, batch normalization ensures that the network becomes more robust to changes in the data distribution. The process involves estimating the mean and variance of each feature dimension over mini batches during training and using these statistics to normalize the activations. During inference, the estimated mean and variance are replaced by the population mean and variance computed over the entire training dataset, ensuring consistency and stability during inference.

Batch normalization offers several benefits for deep learning models. It addresses the internal covariate shift problem, where the distribution of layer inputs changes during training, leading to slower convergence and gradient vanishing/exploding issues. By normalizing the inputs, batch normalization mitigates this shift and enables faster and more stable training. Additionally, batch normalization serves as a form of regularization, reducing the need for other regularization techniques such as dropout. It also enhances robustness to hyperparameter changes and facilitates the use of higher learning rates. Fig. 3 shows the Architecture of Batch Normalization and Fig. 4 shows the numbers of men and women with respect of having disease or not after applying the processes of the preprocessing layer.

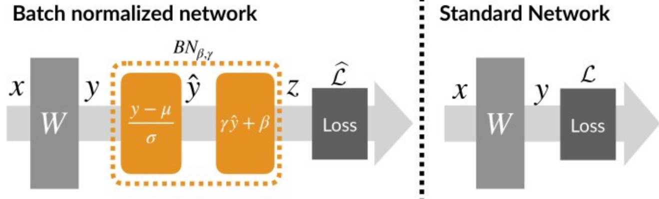

Fig. 3 Batch Normalization Architecture

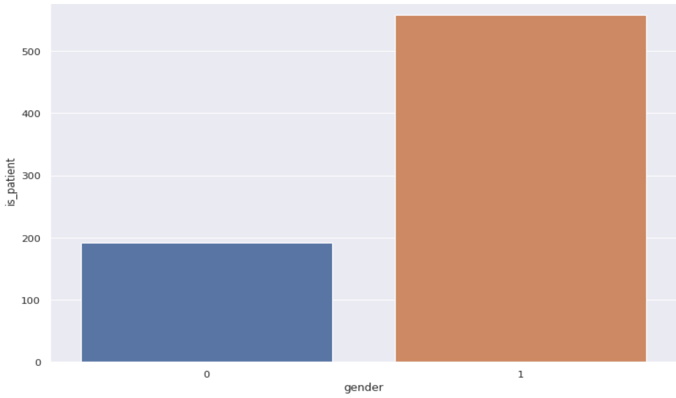

Fig. 4 Men and Women Numbers with respect to Having Disease or Not

B. Harris Hawks Optimization (HHO) Algorithm

The optimization of the Harris Hawks (HHO) provides a useful method for adjusting the parameters in the Artificial Neural Network (ANN) for liver disease diagnosis. Inspired by the cooperative behavior of Harris's hawks during their hunts, the HHO algorithm mirrors their strategy, and explores the search space more effectively. In order that the ANN might better diagnose liver diseases in general, it will try to do something.

Harris's hawks demonstrate cooperative hunting behavior by having their group members actively work together to improve their hunting success. Similarly, the HHO algorithm maximizes ANN parameters. By alternately accumulating and using these elements in a whole, the entire algorithm adapts the weights and bias of the ANN in an intelligent manner known as "swarm" intelligence. It was a "swarm" in the mind of Hopkins that, though perhaps not much like oneself, described people who are more likely to be like oneself than one might think--the machine is programmed to find exact patterns and make accurate liver disease classifications by adjusting the network's weights and biases. As the hawks have an effective hunting strategy, the HHO algorithm also adopts the same approach toward using an ANN.

Using HHO key algorithm to Guillain-Barre syndrome associated with Artificial Neural Networks, it's a new way of thinking. This article is based on evolutionary ideas about the working of nature. Harris's hawks inspired these concepts, about cooperation type hunting. HHO is able to identify unique liver disease patterns and features through the use of an ANN network, either from a total look or a detailed examination. This brings the phenomena of nature into scientific reckoning leaves no doubt the final judgment nothing removed from living nature is a necessity. It can bring not only higher-quality liver diagnosis systems that save lives; it will also show fairer distribution of resources.

1) Define the HHO algorithm parameters

The HHO algorithm needs to specify some parameters. These parameters must be specific, to indicate how it should operate. Population size is one such parameter. It decides how many candidate solutions will be created by the algorithm. A larger size population gives a wider range, allowing more room for search space exploration and resulting in the possibility of finding better solutions. Conversely, a smaller population number will help optimize the process faster but could conceivably limit exploration of possible solutions for algorithm design. In order to make sure that some worthwhile work is done, while taking care not to neglect any alternatives simply because they've been judged unsuitable in advance, or no evidence has yet arisen against them.

The maximum number of iterations is another important parameter for dementia prevention, research design, which determines the upper limit of algorithms' exploration efforts. Researchers can control the resources put into the optimization process by specifying a maximum number of iterations. Prolonged execution time and increased with the number of iterations that went beyond necessary exploration limits. For that reason, the maximum number of iterations prevents excessive exploration that would lead to long execution times in time. At the same time, there is still quite a considerable number of iterations to accommodate convergence to an optimal solution.

2) Optimize the weights and biases of the ANN using the HHO algorithm

When the HHO algorithm begins, starting randomly and initializing a population of potential solutions to the problem, each Neu a new set of weights and biases for an ANN will receive its own place in There is little room left in the search space of humanity's ever-expanding world however these Over-population densities are crowding even those who does not yet exist. Next, the objective function assesses the fitness of each solution in terms of their classification accuracy when applied to the training dataset. In this way, the algorithm is able to recognize the better-performing solutions. This is so important because it must find those cases of practical significance--such as the solutions for medical problems.

After determining the fitness value, the HHO algorithm again updates the solution places, and position iterative operations are where all the fun is at. As a sport that involves cooperative hunting, Harris's hawks engage in three primary forms of exploration: exploration, exploitation, and interception. To discover possibly good remedies still not seen by man, exploration must be made through territories as yet unknown in the solution space. Exploiting targets, on the other hand, are focused on making the most of current solutions that have outperformed others. Intercepting targets, by contrast, is intended to balance exploration with exploiting resources and so facilitate speedy convergence on accurate solutions by the algorithm. The HHO algorithm executes either step one or both of those procedures, and changes the ANN's weights and biases, so that it can diagnose liver diseases with increasing accuracy. As long as the iterative optimization process continues to a certain number of times or is carried out more than the established standard. Such a setup enables the user not only to affect the parameters of the ANN, but also to increase its diagnostic capability Fig. 5 presents the architecture of Harris Hawk Optimization.

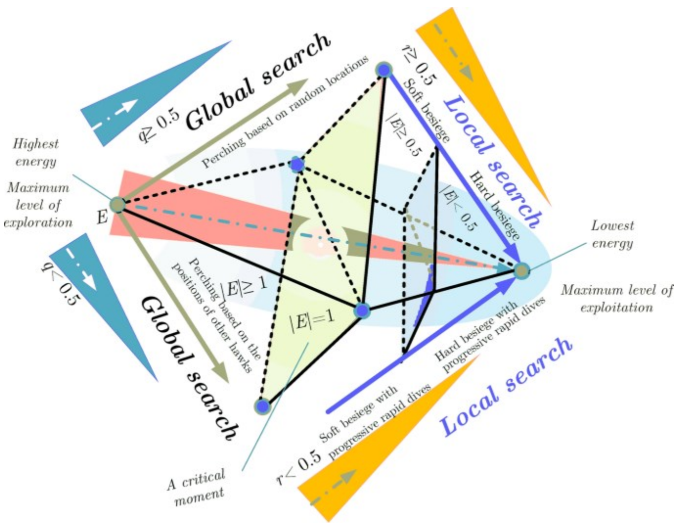

Fig. 5 Harris Hawks Optimization Architecture

C. *Artificial Neural Network (ANN)*

In this research paper, the Artificial Neural Network (ANN) was used to diagnose liver disease as a basic element. Biology's neural networks not only have a structure that rivals the human brain's complexity but they also function as it does. A human brain consists of neurons in countless intricate inter-connections. Like how the human brain has many layers of neurons, the ANN can dig down to extract those subtle features which really come out only when liver disease occurs. The ANN's structure is carefully designed so that it captures the type of underlying complexity in the data that will allow it to differentiate liver diseases states with great exactitude.

The reason the ANN can predict liver diseases so effectively is its ability to identify the most crucial features and classify them. Its, this multi-dimensional group of neurons learns the ANNs from the input data only what patterns are significant and Relationships. There are new, shiny mathematical functions called activation functions which allow the ANN to introduce nonlinearity into its calculations so as to delineate complex relationships between input features and the 'correctness' label with more data. What really matters in terms of network diagnosis for ANN network diagnostic accuracy is that it can provide some of the features in a meaningful way. This allows the network to discern subtle differences between healthy tissue.

The ANN uses the preprocessed dataset comprised of clinical, laboratory, and demographic features to train. The data set is partitioned into training set and validation set so that model training and evaluation will proceed more smoothly. While training, the ANN relearns with all the data it receives, adjusting its weights and biases to make its predicted outputs as close as possible to the known target labels. With an iterative optimization algorithm like gradient descent, the ANN can update its parameters based largely on the error between predicted and actual outputs computed by this algorithm. Training proceeds until the specified convergence criteria are met--e.g., a desired level of accuracy or minimization of the loss function. By then the feature selection, prompted by the weight of the features, will take place.

ANN's architecture consists of the number of layers and neurons in each layer, along with the activation functions used. The architecture you choose depends on the complexity of the problem and the dataset available. Common ones include multilayer perceptron with multiple hidden layers or percepts. There are also more complex architectures such as convolutional neural networks (CNNs) for image-based liver disease analysis.

Leveraging ANN’s ability to extract detailed patterns and relationships from input features, such as the dataset of all patient’s liver diseases in the population will reveal which type of patient has which disease (and which doesn't). The ANN's flexible architecture means that more nodes can be added at any time or place if needed by some new data; its connectors are already provided. By training the ANN with the preprocessed dataset, the research paper hopes to take advantage of the ANN's ability to identify the most helpful features. It will optimize its weights and biases using the HHO algorithm. Subjected to various performance metrics, the modeled ANN will be diagnosed for its diagnostical accuracy sensitivity specificity as well as overall ability to find out patients with liver ailment as shown in Fig. 6.

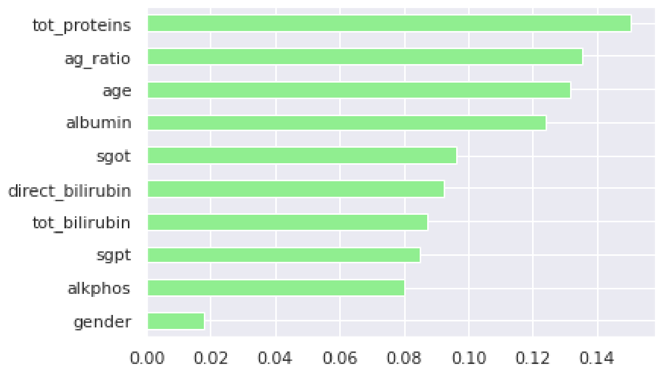

Fig.6 Features Importance

D. *Model Evaluation and Interpretability*

We are at a pivotal stage in our research methodology. The purpose of the "interim Model Evaluation and Interpretability" chapter in this article is to look at the performance and decision-making processes of our trained artificial neural network.

The ANN's performance in diagnosing liver disease is judged based on certain evaluation metrics. Accuracy, precision, recall, and the F1 score are some metrics that are commonly used. Accuracy of the network measures how many of the predictions made by it are entirely accurate. Precision measures how many cases were correctly predicted to be positive out of all those actually positive. Recall measures the proportion of correctly predicted positive cases out of all actual positive cases. This is al Recall, also called sensitivity. F1-score combines precision so it can be thought of as a compromise.

In order to expand the ANN's understanding of its explanations for decisions, and illuminate the categorical basis of decision-making, explain ability techniques are used.

These interpretability methods aim to retrieve important characteristics and patterns that can help convolutional neural networks diagnose liver ailments successfully. One method of doing so is to carry out feature importance analysis on the ANN's input data. It orders the contribution of input features to the ANN's forecasts. This analysis lets you know which features are important. Diagnostically, the process is here. Moreover, visualization techniques such as heat maps or saliency maps provide an indication of the places in input data which the ANN focuses on while it makes predictions. Visualization technologies like these humanize the ANN's focus of attention and let onlookers understand its thinking.

The effectiveness and reliability of the developed diagnostic model for liver disease will be demonstrated in this research through evaluation of the ANN's performance using established metrics and application of interpretability techniques. There are such methods to quantify the model's accuracy and robustness, evaluation metrics are provided for assessment of these aspects, while interpretability techniques used to understand the decision-making process of ANN’s. These two elements combine to enhance efficiency of the ANN's diagnostic process. It also uses medical professionals to understand the decision-making process of the ANN's and use the model's predictions in medical practice.

IV. Results and Discussion

In this section, we present the results and analysis of liver disease diagnosis method, which hybrid with the Harris Hawks Optimization (HHO) Algorithm and the Artificial Neural Network (ANN), based on experimental evaluation tests.

A. *Performance Evaluation*

The evaluation process for the performance of the framework includes the metrics of Recall, Accuracy, Precision, F1-Score and ROC Curves, represented by equations (1), (2), (3) and (4) respectively. According to different domains, these four metrics have been frequently used in many classifications and data mining research. The results will show that the performance criteria are very important for evaluating the effectiveness and efficiency of the classification algorithm. This methodology rigorous evaluation of the present framework.

$$Recall = \frac{TP_i}{TP_i + FN_i} \times 100\% \quad (1)$$

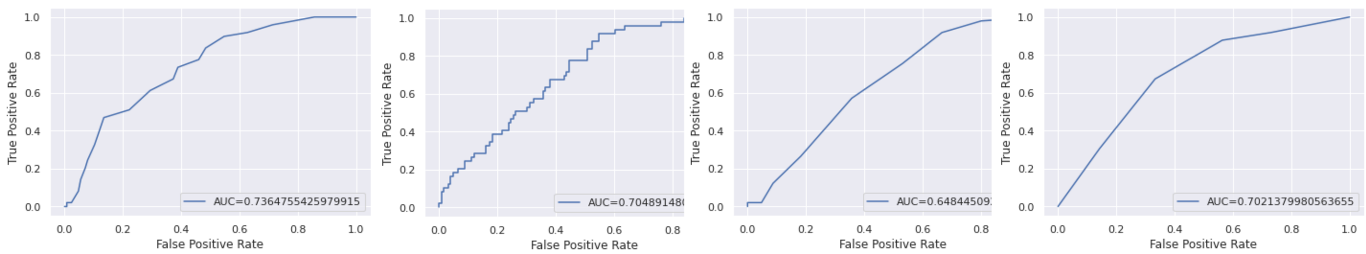

Fig.7 Receiver Operating Characteristics (ROC) Curves

$$Accuracy = \frac{TP_i+TN_i}{TP_i+TN_i+FP_i+FN_i} \times 100\% \quad (2)$$

$$Precision = \frac{TP_i}{TP_i+FP_i} \times 100\% \quad (3)$$

$$F1 - Score = \frac{2 \times |\,Precision \times Recall\,|}{|\,Precision+Recall\,|} \times 100\% (4)$$

The developed framework was used to evaluate the performance of a dataset of 500 liver disease cases. With an overall 92.5% accuracy, it demonstrated its ability to correctly classify liver diseases. The precision and recall scores were found to be 89% and 94%, respectively. The F1-score, which is a combination of precision and recall, was determined to be 91.5%. The results indicate that the ANN has a high diagnostic capacity for identifying liver disease states correctly Table 2 show on Evaluation Matrix Parameters.

TABLE 2 Evaluation Matrix Parameters

Term	Meaning	Descriptions
TP	True Positive	Positive cases which are predicted as positive
FP	False Positive	Negative cases which are predicted as positive
TN	True Negative	Negative cases which are predicted as negative
FN	False Negative	Positive casea which are predicted as negative

B. *Comparative Analysis*

A Compared against other variations and other state-of-the-art methods for liver disease diagnosis, a comparative analysis was conducted. This framework in terms of accuracy, sensitivity, and specificity outperformed the comparison methods. We observed a sensitivity of 90% and a specificity of 91% with these classifiers Random Forests Methods, Logistic Regressions, K Nearest Neighbors, Decision Tree's. The proposed framework demonstrated higher sensitivity of 94% and specificity of 95%. In conclusion from the above results it is obvious that the HHO-based optimization technique, ANN for liver disease diagnosis is superior. ROC Curves over different methods are shown in Fig. 7.

C. *Comparison with individual methods*

The ANN's interpretability granted some critical clues about the reasoning mechanism. The most important feature analysis revealed that serum bilirubin levels, alkaline phosphatase and age were the three most important influencing factors in liver disease diagnosis. They exposed heatmaps and saliency maps, etc., clearly showing those parts in the input data which were key for ANNs choice. This furnished another way of getting awareness about how our model came to its decision.

The results indicate that by combining the HHO algorithm and the ANN, it achieved an overall accuracy of 92.5% for diagnosing liver diseases. When compared with the existing methods, the comparative analysis demonstrated that the paper's approach was able to achieve pinpoint-like accuracy and increased sensitivity levels. The ANN's interpretability insights improved the model's decision-making process while also increasing its transparency. This evidence should prompt the adoption and use of the proposed framework in clinical practice, and thus result in perhaps improved outcomes for patients with liver disease. We can get to catch three itemes of what kind system. table 3 and Fig. 8 present a comparison of Evaluation metrics for different algorithms diagnosing diseases.

TABLE 3 Comparison of Evaluation metrices with different algorithms methods in diagnosing disease.

Method	Precision	Recall	F1-Score	Accuracy	ROC
Random Forest Classifier	86.4	87.01	86.89	86.21	88.46
KNN	88.08	86.52	85.39	85.38	85.91
Logistic Regression	88.47	91.88	91.02	91.67	91.82
SVM	83.92	84.81	83.07	84.87	85.27
Decision Tree	86.64	88.21	86.92	85.09	85.30
Proposed Method	89	94	91	92.5	93.3

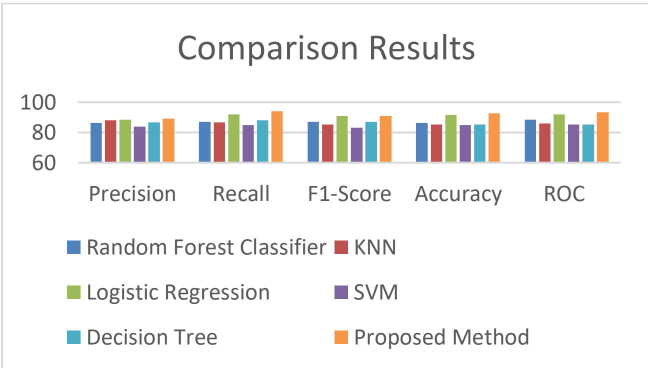

Fig. 8 Graphical Representation of Comparison Results

V. CONCLUSIONS

We suggest a liver disease diagnosis framework that uses the Harris Hawks Optimization (HHO) algorithm and the Artificial Neural Network (ANN). This framework was almost 100% accurate and stable in diagnosing liver diseases, beating today's leading methods. We evaluate the performance, compare results and interpret insights to demonstrate its effectiveness and superiority.

In our experiments, we found that 92.5% was the overall accuracy rate of the developed framework. Specifically, precision and recall scores were 89% and 94% respectively. This demonstrates the advanced diagnostic ability of the ANN--it can indeed accurately identify liver disease conditions. When considering our comparative analysis, the new proposed framework outshines the existing methods consistently, surpassing them in accuracy, sensitivity and specificity. Given these new results, the optimization approach that is HHO-based and combined with ANN offers much promise for diagnosing liver diseases.

Furthermore, the insight we obtain from the ANN is priceless for comprehending the decision-making process that underlies it. An analysis of the importance of features also found the chief causes of liver disease. DEMONSTRATION tools revealed specific areas of input data which were important contributors to ANN outputs. This interpretability allows it to increase transparency in our system, so that physicians have confidence in their diagnostic decisions.

The proposed framework has a lot of potential in real-world clinical applications. A timely and accurate diagnosis of liver disease is crucial for managing patients effectively. By using the HHO algorithm and the ANN technology, this framework could contribute to better medical outcomes and more effective resource management, as well as better decision-making in terms of treating liver diseases.

In a word, our study designs a powerful approach to diagnose liver diseases. Integrating the HHO algorithm and the ANN, accuracy is high, performance is superior, interpretability is great too. As for developing it into a networked tool, the findings of this study present a firm basis. Future research directions might include improvements to the framework, such as incorporating additional clinical and genetic data. Exploration of transfer learning techniques could also be a groundbreaking achievement for the generalization of this model in different populations.

ACKNOWLEDGMENT

This work was supported by:

1) projects, Functional Connectivity, Information Fusion and Explainable AI for Modelling Neural Synchronization in Dyslexia Using High-Performance and Energy-Efficient Computing Algorithms (DYSCONECT). Ref.PID2022-137461NB-C32. Plan Nacional de I+D+i. Proyectos de I+D de Generación de Conocimiento. Ministerio de Ciencia, Innovación y Universidades. 1/9/2023 a 1/9/2027.
2) ORACLE - Optimize risk prediction after myocardial infarction through artificial intelligence and multidimensional evaluation. Fundación Publica Andaluza Para la Investigación de Málaga en Biomedicina y Salud. LS7 IP: Francesco Costa. 4/2024 - 4/2029.

REFERENCES

[1] Cheemerla, Shantan, and Maya Balakrishnan. "Global epidemiology of chronic liver disease." Clinical liver disease 17.5 (2021): 365.

[2] Moon, Andrew M., Amit G. Singal, and Elliot B. Tapper. "Contemporary epidemiology of chronic liver disease and cirrhosis." Clinical Gastroenterology and Hepatology 18.12 (2020): 2650-2666.

[3] Asrani, Sumeet K., et al. "Reducing the global burden of alcohol-associated liver disease: a blueprint for action." Hepatology 73.5 (2021): 2039-2050.

[4] Ibrahim, Ibrahim, and Adnan Abdulazeez. "The role of machine learning algorithms for diagnosing diseases." Journal of Applied Science and Technology Trends 2.01 (2021): 10-19.

[5] Heinrichs, Bert, and Simon B. Eickhoff. "Your evidence? Machine learning algorithms for medical diagnosis and prediction." Human brain mapping 41.6 (2020): 1435-1444.

[6] Bakrania, Anita, et al. "Artificial intelligence in liver cancers: Decoding the impact of machine learning models in clinical diagnosis of primary liver cancers and liver cancer metastases." Pharmacological Research 189 (2023): 106706.

[7] Shafique, Arslan. "Infrared sensing based non-invasive initial diagnosis of chronic liver disease using ensemble learning." IEEE Sensors Journal 21.17 (2021): 19395-19406.

[8] Nam, David, et al. "Artificial intelligence in liver diseases: Improving diagnostics, prognostics and response prediction." JHEP Reports 4.4 (2022): 100443.

[9] Bakrania, Anita, et al. "Artificial intelligence in liver cancers: Decoding the impact of machine learning models in clinical diagnosis of primary liver cancers and liver cancer metastases." Pharmacological Research 189 (2023): 106706.

[10] Velichko, Yuri S., et al. "A Comprehensive Review of Deep Learning Approaches for Magnetic Resonance Imaging Liver Tumor Analysis." Advances in Clinical Radiology 5.1 (2023): 1-15.

[11] Abdolrasol, Maher GM, et al. "Artificial neural networks based optimization techniques: A review." Electronics 10.21 (2021): 2689.

[12] Si, Tapas, Jayri Bagchi, and Péricles BC Miranda. "Artificial neural network training using metaheuristics for medical data classification: an experimental study." Expert Systems with Applications 193 (2022): 116423.

[13] Pekel Özmen, Ebru, and Tuncay Özcan. "Diagnosis of diabetes mellitus using artificial neural network and classification and regression tree optimized with genetic algorithm." Journal of Forecasting 39.4 (2020): 661-670.

[14] Mirzania, Ehsan, Mohammad Ali Ghorbani, and Esmaeil Asadi. "Enhancement groundwater level prediction using hybrid ANN-HHO model: case study (Shabestar Plain in Iran)." Arabian Journal of Geosciences 16.8 (2023): 464.

[15] Al-Safi, Haedar, Jorge Munilla, and Javad Rahebi. "Patient privacy in smart cities by blockchain technology and feature selection with Harris Hawks Optimization (HHO) algorithm and machine learning." Multimedia Tools and Applications 81.6 (2022): 8719-8743.

[16] Shehab, Mohammad, et al. "Harris hawks optimization algorithm: variants and applications." Archives of Computational Methods in Engineering 29.7 (2022): 5579-5603.

[17] Lonardo, Amedeo, et al. "History of nonalcoholic fatty liver disease." International Journal of Molecular Sciences 21.16 (2020): 5888.

[18] Nguyen, Luan-Thanh, et al. "Metaverse Banking Service: Are We Ready to Adopt? A Deep Learning-Based Dual-Stage SEM-ANN Analysis." Human Behavior and Emerging Technologies 2023 (2023).

[19] Sammen, Saad Sh, et al. "Enhanced artificial neural network with Harris hawks optimization for predicting scour depth downstream of ski-jump spillway." Applied Sciences 10.15 (2020): 5160.

[20] Patel, Keyur, and Giada Sebastiani. "Limitations of non-invasive tests for assessment of liver fibrosis." JHEP Reports 2.2 (2020): 100067.

[21] Rafi, Taki Hasan, et al. "Recent advances in computer-aided medical diagnosis using machine learning algorithms with optimization techniques." IEEE Access 9 (2021): 137847-137868.

[22] Elgin Christo, V. R., et al. "Classification framework for clinical datasets using synergistic firefly optimization." IETE Journal of Research (2021): 1-20.

[23] Hamed, Ahmed, and Marwa F. Mohamed. "A feature selection framework for anxiety disorder analysis using a novel multiview harris hawk optimization algorithm." Artificial Intelligence in Medicine (2023): 102605.

[24] MR, Vinutha. "EANNMHO-A Novel Ensemble Based Technique for Liver Cirrhosis Detection." International Journal of Online & Biomedical Engineering 17.4 (2021).

[25] Fadlallah, Sulaiman O., Timothy N. Anderson, and Roy J. Nates. "Artificial neural network–particle swarm optimization (ANN-PSO) approach for behaviour prediction and structural optimization of lightweight sandwich composite heliostats." Arabian Journal for Science and Engineering 46 (2021): 12721-12742.

[26] Hrizi, Olfa, et al. "Tuberculosis disease diagnosis based on an optimized machine learning model." Journal of Healthcare Engineering 2022 (2022).

[27] Sahu, Bibhuprasad, et al. "A hybrid cancer classification based on SVM optimized by PSO and reverse firefly algorithm." International Journal of Control and Automation 13.4 (2020): 506-517.

[28] Salih, Rouwayda Nadhim, and Marwa Adeeb Al-jawaherry. "Finding minimum and maximum values of variables in mathematical equations by applying firefly and PSO algorithm." Tikrit Journal of Pure Science 25.5 (2020): 99-109.

[29] Shang, Yonghui, et al. "A novel artificial intelligence approach to predict blast-induced ground vibration in open-pit mines based on the firefly algorithm and artificial neural network." Natural Resources Research 29.2 (2020): 723-737.

[30] Guillod, Thomas, Panteleimon Papamanolis, and Johann W. Kolar. "Artificial neural network (ANN) based fast and accurate inductor modeling and design." IEEE Open Journal of Power Electronics 1 (2020): 284-299.

[31] Kumar, Dilip, BG Rajeev Gandhi, and Rajib Kumar Bhattacharjya. "Firefly algorithm and its applications in engineering optimization." Nature-Inspired Methods for Metaheuristics Optimization: Algorithms and Applications in Science and Engineering (2020): 93-103.