

An Optimized Artificial Neural Network Approach Based on Sperm Whale Optimization Algorithm for Predicting Fertility Quality

Engy EL-SHAFEIY^{1*}, Ali EL-DESOUKY¹, Sally EL-GHAMRAWY²

¹ Computer and system Department, Faculty of Engineering Mansoura University, Mansoura, Egypt
engy.elshafeiy@gmail.com (*Corresponding author), ali_i_eldesouky@yahoo.com

² Computers Engineering Department, MISR Higher Institute for Engineering and Technology, Mansoura, Egypt
saly_elghamrawy@yahoo.com

Abstract: This research introduces a new approach for predicting reproductive health by using the Sperm Whale Optimization algorithm (SWA) with Artificial Neural Networks (ANN-SWA). SWA is an emerging method with a powerful potential in tackling optimization difficulties based on its adaptability in searching mechanisms. ANN-SWA consists of four phases. The first phase is conditioned by the fertility disease which is a complex and multifactorial condition with increasing incidence. The fertility data is trained (90 cases) and the approach is then used to test findings in the test set (10 cases). In the second phase, the multilayer perceptron (MLP) is used to determine the maximum fitness function by getting the global minimum and hence, it revokes the ANN trapped in local. In the third phase, it optimizes and controls the parameters (weights and biases) to ensure rapid convergence with accuracy. In the fourth phase, ANN-SWA is used to predict the fertility quality and determine the accuracy. The results are verified by comparing them with optimization and classification algorithms. The quantitative and qualitative outcomes show that the proposed approach is able to outperform the current algorithms on the fertility dataset in the convergence rate of classification. The results demonstrate that an artificial neural network based on SWA achieved more than 99.96% of the accuracy of the approach.

Keywords: Sperm Whale Algorithm, Artificial neural network, Machine learning fertility quality, Biomarkers in male infertility.

1. Introduction

According to the World Health Organization at global level one in ten men aged 17–75 years had experienced infertility. Infertility is a reverse key to Venereal health and fertility which is renowned in global public health and World Organizations [10]. Infertility is defined as the inability to conceive after one year of frequent sexual intercourse without the use of contraceptive methods. The male genital system consists of the outside and the internal genital organs. The scrotum contains the principal internal genital organs that are the testes, i.e., the male genital glands. The gonads of the man are the testes and are responsible for the production of gametes, i.e., haploid cells (sperm) that are necessary for reproduction. The investigation of male fertility can be a critical diagnostic challenge for the practitioner. Health care providers can access an approach that covers a significant spectrum of relevant clinical issues related to male fertility disorders. There is no recognized or consistent definition of infertility, so the availability of fertility treatments are likely to have had an impact on influential processing factors. Studies have found that men in later age cohorts are more susceptible for infertility diseases and seeking help in healthcare services [3]. In the last five years, the artificial intelligence

in reproductive health has been able to detect many diseases. This is estimated in many studies, which turns out that the significant breakthrough in machine learning and the recognition of natural language has been able to discover and diagnose many reproductive diseases. Artificial neural network (ANN) [15] is based on nonparametric and intelligent mathematical models inspired by the biological nervous system. In the last three decades, ANN has been widely investigated and applied to classification, pattern recognition [5], regression and forecasting problems. The efficiency of the ANN is profoundly affected by its learning process. For multilayer perceptron (MLP) neural networks, which are the most common and applied ANN [11], there are two main categories of supervised training methods: gradient-based and stochastic processes. The back-propagation algorithm and its variants are considered standard examples of gradient-based methods and the most popular among researchers. Despite, there are three principal problems in the gradient-based techniques: the tendency to be trapped in local minima, slow convergence, and high dependency on the initial parameters. As robust alternatives to the gradient-based above-discussed approaches, heuristic research algorithms have been proposed

in literature for optimizing the MLP networks. In opposition to gradient algorithms, metaheuristics show greater efficiency in avoiding local minima. Swarm and evolutionary based algorithms [16], which are the pair starting families of metaheuristics algorithms are amongst the most investigated methods by researchers in training MLP networks. These natures of algorithms [12] are population-based, in which some feasible random solutions are generated, developed, and updated until a satisfying answer is achieved or some power iterations are given. These algorithms include randomness as the original mechanism to transfer from a local search for a global search, and hence, they are added suitable for global optimization. Evolutionary algorithms were used in the supervised learning of MLP networks in three various basic schemes: the automatic purpose of the network construction, optimizing the connection weights and biases of the net and developing the learning rules. It is necessary to consider here that simultaneous optimization of the construction and connection weights of the MLP network can drastically improve the number of parameters so that it can be analyzed as a large-scale optimization problem. Leap progress in the field of metaheuristics was accomplished by presenting the swarm search calculations as the very much trusted Particle Swarm Optimization (PSO). Eberhart and Kennedy introduced PSO in 1995. PSO calculation recreated the social conduct of winged creatures rushing in nature. In writing, PSO and its varieties were seriously examined in preparing MLP to be arranged. The outcomes demonstrate that PSO has top neighborhood optima shirking in streamlining MLP systems. Different cases that fall under the same classification are Ant Colony Optimization (ACO) and the Artificial Bee Colony (ABC). Examples of their application to the issue can be found in [21,4,19]. In the leading work of David Gil et al., three sorts of machine learning devices including decision trees (DT), multilayer perceptron (MLP) artificial neural system and support vector machines (SVM) were used to predict the seminal quality. A neural system approach giving MLP was additionally proposed to assess semen parameters. As it is basic in bioinformatics and medicinal finding [18], most accessible first quality forecast information

likewise have the uneven class issue, i.e., the quantity of negative (dominant part) perceptions far minimizes the number of positive attitudes. Due to this skewed dispersion in information, numerous standard learning calculations regularly perform inadequately on the minority class. Because of the multifaceted nature of the issue itself, the information lopsidedness and its effect on expectation exactness in original quality need additionally judgement. Although modeling prediction for improved infertility disease in men has received much care, some factors regarding fertility have been largely ignored. Hence, the primary purpose of this paper is to investigate the probability of prediction of fertility in men using an approach which contains artificial intelligence neural network based on optimization Sperm Whale Algorithm SWA. This paper is structured as follows:

Section 2 introduces the theory and the methods used, the Sperm whale Algorithm (SWA optimizer) and ANN (such as MLP network). The proposed prediction ANN-SWA approach is explained in Section 3. Section 4 is devoted to results and discussions. Finally, the concluding remarks are included in Section 5.

2. Theory and Methods

2.1 Sperm Whale Algorithm (SWA)

The cachalot sperm whale (or *Physeter macrocephalus*) is a conspicuous user of echolocation and communication, with taut vocalization of 230 decibels (re 1 μ Pa at 1 m) underwater. Sperm whales eat squids but to get squids, they should go down to the profundities of the water (2000e3000 m profound) while they have to rise to the top to inhale [13]. That is the reason why they encounter two inverse shafts of their condition in each cycle of breathing and nourishing: the surface and, typically, the base of the ocean. Sperm whales, for the most part, go in gatherings of 6-9 and the males and females live in a similar gathering. The Sperm Whale Algorithm [6] is a way for the sperm whale to live in nature. A new effective multi-target is obtained by mathematical optimization algorithm. Every sperm whale encounters two inverse places in its breathing-feeding cycle (it needs to rise to the top to inhale and go down to

the seabed to feed). In this way, for each whale, one solution for its present place and another solution for its mirror, and the target capacities for the two focuses were assessed. In any case, the issue here was that the mirror impression of the best answer did not help in finding the ideal answer and just expanded the time required for discovering it. That was the reason why only the most exceedingly local answer was reflected. Exchange of information between the worst result positions can be transferred to the desired space through the search center reflection. The link between the two points is the connection of the worst sperm whale individual with the best sperm whale individual as shown in Figure 1(a). The retesting of the wrong answer of the SWA algorithm by the center reflection is shown in Figure 1(b). One of the colossal favorable circumstances of SWA over the others is that it utilizes the best and the worst answers at the same time. to get the optimal global values. The sperm whale calculation can maintain the global optimum and keep it from falling into local optimum by making distinctive subgroups as shown in Figure 2. In this algorithm, each answer exemplifies a sperm whale. When seeing a group of sperm whales taking the formation into social units, $m \times n$ number of answers were first made, assessed, and then requested as the underlying populace. This requested populace was separated into n Temporary Sub-Groups (TSG) each with m individuals, and one part was haphazardly chosen from every one of the brief subgroups for each Main Sub-Group (MSG). Figure 2 illustrates the gathering procedure. This procedure, the sperm whale gathering both male and female which exist in each gathering, guarantees, to some degree, the dispersion of solutions in various gatherings and counteracts early local stuck. The best individual and the worst individual in the population whale in each gathering were known as the X_{best} and the X_{worst} , individually, showing the following relationship:

$$X_{center} = X_{worst} + C \times X_{best} \tag{1}$$

$$X_{reflex} = 2X_{center} - X_{worst} \tag{2}$$

In the event that X_{reflex} is situated outside the inquiry space, c should diminish thusly:

$$c = r \times ci \tag{3}$$

where c_i is the beginning focus factor and r is the constriction coefficient that is under 1,

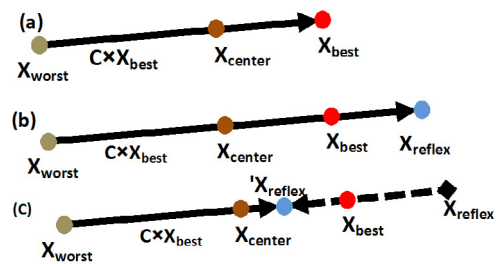


Figure 1. The Reflection of the worst answer in SWA

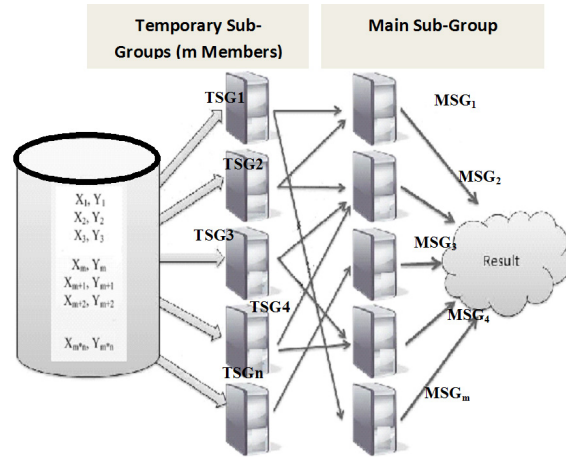


Figure 2. Grouping and sort process in SWA algorithm

where r and ci ought to be set as calculation parameters. Starting with multiple population common search by $m \times n$ number. The primary population is divided into n Temporary Sub-group (TSG), each containing m members. A new Main Sub-group (MSG) embodying n members is produced through a randomly selected TSG until m MSGs are generated for iterative optimization.

The construction of various sub-groups aims to prevent SWA from a premature end and from a fall into local optimum. MSG is created. Because each sperm whale undergoes two locations in breathing–feeding cycle (it breathes on the sea but takes food undersea) the objective function will be calculated by both two positions (current position and relative position). Assume that the worst and the best whales are named X_{worst} and X_{best} respectively, in Equations (1) and (2).

2.2 Artificial Neural Network

A good fit of the neurons in the human brain is tried through mathematically defined functions. The intracellular behaviour of brain neurons simulate and through the computational weights available on the lines synthesis of the synthetic neurons, synaptic function in natural neurons

become the network. To this end, the object, with a handful of real data measured, [17] trained, for new entries with attention to find the relationship found in the training stage, calculate the appropriate output. There are many limitations and problems that present the feedforward (FF) algorithm in training the MLP as exposed in [1]. Optimal Algorithms (OA) have demonstrated great effectiveness in solving this problem. An MLP consists of multiplied layers of neurons, ordinarily three layers (anywhere all layers wholly connected to the next one): an input data layer that draws outside inputs data, one hidden layer, and an output data layer which creates the classification decisions [8]. But for the input data layer, every neuron in the other layers is the computational component with a nonlinear activation function. The structuring neural network [2] for data that is the number of inputs, outputs, and hidden nodes. However, there are several different architectures that may be suitable for predicting problems, such as MLP. In MLP systems neurons in the layers are completely interconnected by numeric weights, every neuron containing the and initiation of capacities. The summation of work aggregates up to the result of sources info (weights, bias), and inclination as appeared in Eq. (4) where w_{ij} is the association weight interfacing X_i to neuron j , β_j is a predisposition term also, i is the aggregate number of neuron inputs. The yield of the summation work will be a contribution to the enactment of the work. For the most part, a nonlinear actuation work like the formed bent Sigmoid capacity is utilized. Sigmoid capacity appears in Eq. (5). In this manner, the yield of the neuron j can be portrayed as in Eq. (6).

$$y_j = \sum_{i=1}^n w_{ij} X_i + \beta_j \quad (4)$$

$$\alpha_j^i = f(x_j^i) = \begin{cases} -1, & \text{if } x_j^i < 0 \\ 1, & \text{if } x_j^i \geq 0 \end{cases} \quad (5)$$

$$y_j = f_j(\sum_{i=1}^n w_{ij} X_i + \beta_j) \quad (6)$$

At the point when the structure of ANN is made, the learning procedure of this approach is connected with a particular end goal to change the parameters of the approach (set of weights). In other words, these weights are tuned and refreshed to inexact the outcomes and limit some error criteria. Learning procedure or preparing procedure of the approach is an undertaken test that mirrors the quality of MLP and organizes the unraveled unique sorts of issues. One of the major MLP preparing approach is the administered

learning. The objective of the directed learning is to limit the error between the actual output and the target output and to find the subordinate of ANN target work with deference weights that traded between the layers. This strategy is wasteful at the point when the search space is huge and the technique is relevant just for differentiable target capacities. The performance of the approach through Mean Absolute Error (MAE) [20] and Root Mean Squared Error (RMSE) between y_i actual value and target value is showed in the following relationship:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

This research sets forth the neural network learning machine, to confirm how this network can be simply exploited to predict fertility quality. Also, this network presents very good generalized results with all samples of the dataset and is actually involved in the optimization of the predicted process. The overall methodology used for neural networks for prediction processes is depicted in the following section, in Figure 3, which illustrates the ANN-SWA approach (that uses both ANN and Sperm whale Algorithm (SWA)).

3. The proposed prediction approach

This section describes the proposed approach in detail. The approach for predicting the fertility quality for men, as illustrated in Figure 3, generally comprises the four phases of Training and Test, SWA Optimization, ANN and Fitness function and Predicting phase. More information on each stage is explained in the next sections Where X is the dataset $X_i \in \mathbf{R}^N$, $i=1,2,\dots,n$, is a sample; $C_i \in \mathbf{R}^N$, $i=1,\dots$, C_i is a center of X_i obtained from Eq.(1). N is the number of variables; i.e. the number of elements of X_i and C_i ; i is the number of samples where i -th sample is the input layer and the j -th is the number of the hidden layer. Let assume the training dataset as: $D = f(x_i, y_j)$ where $X=(X_1, X_2, X_3,\dots, X_n)$ and

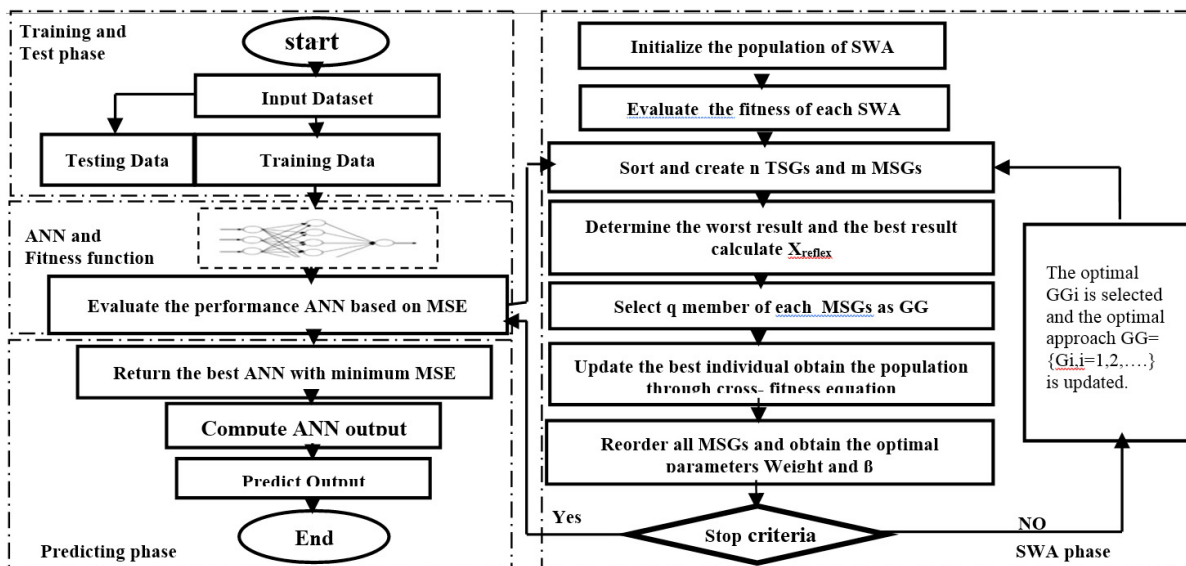


Figure 3. Flowchart for steps of building, training and evaluating the ANN-SWA prediction approach

$Y = \{0, 1\}$. (0 is “Normal” sperm quality and 1 is “abnormal” in this paper). The objective of the present paper is to predict the sperm quality result y on new test parameter X with the learning approach $Y = f(X)$ built for training dataset and to update the solution using the SWA to predict the best solution. The ANN-SWA begins with an initial population the so-called population of sperm solutions. As the ANN is used mostly for prediction goals, it receives input data and target data and predicts the relationship among these data (i.e., input data and target data mapping). In the ANN-SWA, the best-obtained solution at each iteration is assumed as target data and the purpose is to reduce the error among the target data and other predicated solutions. Based on the determined concept, the ANN-SWA is formed for minimizing problems (i.e., minimizing the error among the target and solutions). It is worth pointing out that this target solution has been updated at each iteration. In this paper, the structure of the ANN-SWA, together with some parts of its mathematical formulations and concepts, all come to develop a new optimization approach based on the neural network configuration. The detailed descriptions and the processes of the ANN-SWA are given in the following subsections as shown in Figure 3.

3.1 Training and Test Phase

In this phase, the optimization of an ANN approach is proposed for predicting the semen

quality of the lifestyle. The approach inputs comprise the following features which reflect the health status of the men: age, accident, hours spent sitting per day, season when the semen samples are produced, high fevers in the last year, smoking habit, childish diseases, surgeries and frequency of alcohol consumption. In the Training and Test phase, the real datasets for the fertility quality are divided into training set and testing set, containing initial parameters of ANN and initial data sample which are put into ANN-SWA approach to compute the best f fitness function value. Here, the training set is used to determine the approach parameters (weights and bias), and the test set is used to evaluate the accuracy and the network generalization.

3.2 ANN and Fitness Function

The proposed approach starts generating the initial sample of the sperm dataset, to construct the ANN. The optimum number of input nodes equals the input features in the sperm dataset. Regarding the number of the hidden nodes, there is no expected formula for choosing the optimum number of hidden neurons. Nevertheless, some general rules are used for determining the number of hidden neurons. An imperfect approach can be achieved by the geometric pyramid rule introduced by [8]. For a three layer network with i input and o output neurons, the hidden layer would have $\sqrt{i * o}$ neurons where W_{ij} is the weight from input

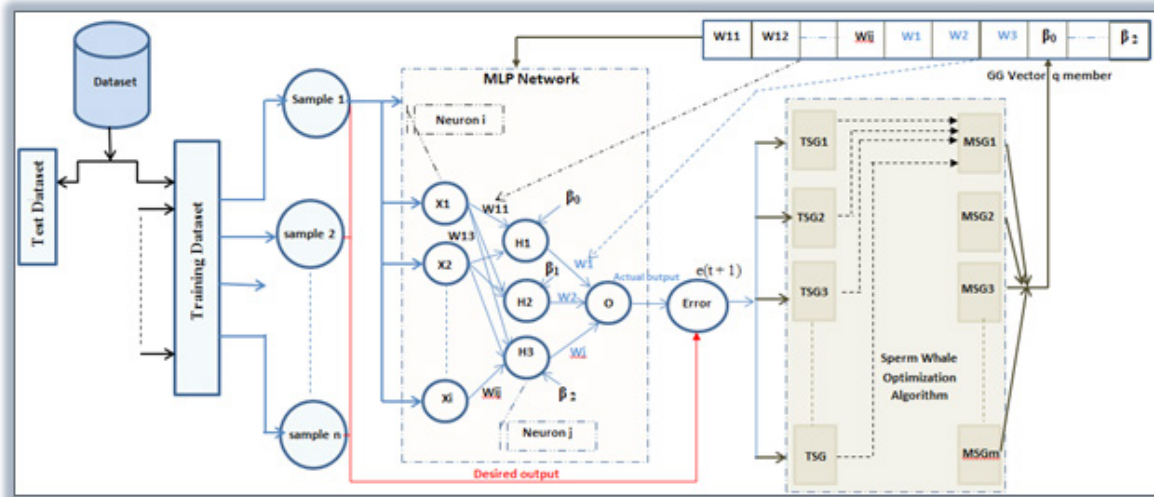


Figure 4. Schematic view of the performance of the ANN-SWA approach

X_i to hidden H_j and W_j is the weight from hidden H_j to output as shown in Figure 4. To measure the fitness value of the generated sperm agents, the mean square error (MSE) fitness function was used which is based on the calculation of the difference between the actual and predicted values by the generated agents (MLP) for all the training samples. MSE is shown in Eq. (9) where y is the actual value, \hat{y} is the predicted value, and n is the number of instances (number of patients) in the training dataset. In this approach, the SWA algorithm adjusts the weights of each connection and parameter in the network [14] reduce the value of the error function and to compute the error in Eq.(9). MLP tries to reduce the MSE (Mean squared error) represented as fitness function f . Where f is the objective function. In the whole paper, notations having a vector type have corresponded as vector values (array). Otherwise, the rest of the notations and parameters are considered scalar values. After determining the f for all solutions, then the best solution (in this article, a candidate solution with the minimum objective function value) considered as the target solution was found.

3.3 SWA Phase

The parameters (weights and bias) in ANN are updated based on the SWA algorithm that used a fitness function defined in Equation (9).

The samples are contained A set of input variables as an array of $1 \times n$, representing input data in the ANN is shown in Figure 4. This is defined as follows: $D = [X_1, X_2, X_3, \dots, X_n]$. In this approach, the search process undergoes multiple

sperm whale units to get the desired results. The approach of the ANN-SWA for training can be described in the following steps:

- In SWA phase set the initial population number, TSG number m , MSG number n , C refers to center factor, c_i refers to initial mirror center reflection factor and r refers to contraction coefficient number of sperm whales (q members) with better fitness functions are selected from each main subgroup as the Good Gang (GG) and, considering the sperm whales' limited vision, a local search is made around the members of this group. GG is the individual number q , where $c = r \times c_i$.

$$c = \text{random}() \times \text{Min}(c) \quad (10)$$

In the ANN-SWA approach, the initial optimal set of sample needs to be initialized as an empty Good Gang $GG = \{ \}$, load the initial optimal set of (weights, biases) for MLP networks.

- Initialize the predefined number of search agents by having $m \times n$ individuals After this the initial population needs to be divided into n temporary subgroups (TSG), and one is randomly selected from each of the temporary subgroups for every Main subgroup (MSG).
- Because of sperm whales' limited vision, a local search is carried out around the members of the (GG) which consists of q whales selected from each MSG. The search method is the following: iterations can continuously change members of GG within the field of definable vision Q (radius), and once a better fitness function f is gained, the original one in GG will be replaced.

- Determine the worst result of f and the best result of f , calculate X_{reflex} by equations 2 and when using equation 10 there is no necessity to verify whether X_{reflex} is beyond the search space. Consequently, not only does this performance speed up the operation to a certain extent, but it also decreases the number of parameters and dimensions.
- The optimization needs to be taken into the ANN-SWA approach one by one to calculate the variables corresponding to datasets GG_1, GG_2, \dots, GG_n in the sample composed by each of the remaining sample GG . Then, select the sample GG_i variables to optimal sample, and update it.

Pseudocodes of Approach ANN-SWA

```

Initialize the sperm whale population Xi (i = 1, 2, ..., n)
Initialize the parameters q, c, population size (Npop) and maximum number of iterations
Calculate the fitness f of each agent
Sort sperm population and create n TSGs
Create m MSG
Find the X_worst and X_best, calculate X_reflex
Update the X_reflex and reorder
Procedure sperm (population, the maximum number of iterations, q, c, N_pop)
t=1
while (t < maximum number of iterations)
  For i = 1 to N_pop
    If X_reflex ∈ random() × Min(c)
      Update the X_reflex and reorder
      Select q members from each MSG to establish GG
      For i=1:q
        For j=1:c
          if the best nearby solution is superior to X_j then
            replace X_j with X_neighborhood
          End if
        End for
      End for
      For k=q+1:n
        /* crossover occurs between the best answer of the Good Gang and
        the rest of the answers in the main subgroup, and between two
        generated children, one (randomly) takes the place of the mother
        solution. */
        Execute cross operation between X_i and X_k,
        generate the offspring population,
        replace the X_k with the best offspring individual
      End for
      Place all MSGs together and reorder them
    Else (c=r×c)
      End If
    End For
  End while
  Return X
  Calculate the fitness of each agent by Apply Eq. (9).
End procedure

```

3.4 Predicting Phase

Two important aspects are taken into consideration when the approach is designed: the representation of the search agents for the sperm and the selection of the fitness function. In ANN-SWA, each search agent is encoded as a one-dimensional vector to represent a candidate neural network. calculating the fitness value of each iteration to be gained. Determine whether the stop conditions are or not. Repeat until the maximum number of iterations is reached. Finally, the MLP network with the f fitness function (minimum MSE value) is tested on the unseen part of the dataset (test/validation samples). Parts: a set of weights connecting the input layer with the hidden layer, a set of weights

connecting the hidden layer with the output layer, and a set of biases. The length of each vector equals the total number of weights and biases in the network, and it can be calculated using Eq. 11 where i is the number of input and H is the number of neurons in the hidden layer. To measure the fitness value of the generated sperm agents, Equation (10) is used.

$$\text{Individual length} = (i \times H) + (2 \times H) + 1 \quad (11)$$

The SWA algorithm has been used to optimize ANN classifier which selects the best parameters that give the highest accuracy and determine the accuracy of a prediction the sperm quality. The training set for the fertility quality contains the initial parameters of ANN and the initial data sample used to compute the best f fitness function value. The ANN-SWA is a Multilayer perceptron optimization approach having self-feedback. In the proposed approach, there are no sensitive initial parameters (excluding the common user parameters: the maximum number of iteration and population size). Hence, the user does not test to fine-tune or check a various set of initial parameters while he applies the publicly available fertility quality expression data. Experiments are executed over fertility dataset which is available at UCI repository [20]. The data consist of hundred instances recorded in ten columns, nine of them containing input features as season number, age, diseases in childhood, accidents or seriously previous life events, surgeries, fever rates in last years, the rate of alcohol, in case of a drinker person sitting hours, smoking or not and the last one is the predicted column, these first nine input parameters cover environmental, health, lifestyle factors which predict the fertility. All experiments are performed on Core i5-2400 with 8 GB RAM, 500 GB hard disk running with a Window 7. Fertility dataset contains 100 samples; where 88 samples are normal and 12 samples are abnormal.

4. Performance analysis and implementation

This section explains experimental results obtained from the proposed approach.

4.1 Description of the Dataset

This is a dataset which contains fertility instances. Fertility is the found data which includes 100 instances and 10 attributes. In this research, the fertility quality parameter was used that

Table 1. Statistical error parameters of the proposed approach for the estimation of fertility in training and testing datasets.

Performance		Best	Average	Worst	MSE
ANN	Training	4.04E+00	5.43E+00	7.78E+00	4.97E+00
	Testing	3.44E+00	4.53E+00	6.99E+00	4.07E+00
ANN-GA	Training	0.9216E+00	1.8921E+00	3.81E+00	1.8853E+00
	Testing	0.8124E+00	1.223E+00	3.01E+03	1.2213E+00
Decision-tree	Training	6.94E+00	7.43E+00	8.78E+00	5.37E+00
	Testing	5.83E+00	6.95E+00	9.88E+00	5.20E+00
SVM	Training	0.0814E+02	0.2721E+02	0.9410E+02	0.0924E+02
	Testing	0.8813E+00	0.1913E+02	0.8812E+02	0.0852E+02
ANN-SWA	Training	1.52E-01	0.43E+00	1.65E+00	1.10E-01
	Testing	1.31E-01	2.12E-02	5.27E-02	1.38E-01

contains 9 normalized variables about life habit. In this dataset, health status and a semen quality result as normal or abnormal [9]. Moreover, the number of positive and negative samples is 7:1. In the involving new ideas or methods by David Gil et al. [7], three type of machine learning including multilayer perceptron (MLP), decision trees (DT) and support vector machines (SVM) were applied in this current dataset.

4.2 The Measure in Samples

Accuracy is equivalent to Eq. (12).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (12)$$

The Recall is computed as given in Eq. (13).

$$Recall(TPR) = \frac{TP}{TP+FN} \quad (13)$$

Exactness is called positive prescient value (PPV) and is determined as given in Eq. (14).

$$Precision = \frac{TP}{TP+FP}$$

4.3 Performance Test of the Fertility Quality

To validate the proposed approach for the fertility quality, 10-fold crossvalidation is performed using parameters such as accuracy and sensitivity. It is not enough to measure the performance of the machine learning approach only on the basis of accuracy. 10 different values of k are taken for cross-validation. The purpose of cross-validation is to test the performance of the approach in different types of scenarios. The accuracy is calculated as the percentage deviation of the predicted target with the actual target with a passable error. Figure 5 illustrates two graphs 5(a) and 5(b) which are charted for fertility quality, against a common parameter of

sensitivity. Each figure has 2 datasets: testing and training. Values of training data show an almost consistent straight line. Similarly, values of testing data, show an almost consistent straight line but with lower values than that of training data. This graph suggests that the proposed approach displays consistency and satisfactory results against the sensitivity parameter. All the data was recorded in ANN-SWA in each one of the ten folds and the last mean variable scores are given in Figure 5. The parameters used in the experiment are shown below: The topology of the MLP: 9 input neurons; 1 hidden layers; and one output neuron. Transfer function for hidden layer is, transfer function for output layer in equation 6 and Population Size: 10. There are various algorithms used to predict fertility in men such as decision tree, ANN, ANN based on genetic and Support Vector Machines. The performance of the ANN-SWA was tested using the aforementioned classifiers. These classifiers were compared regarding accuracy, specificity and sensitivity as shown in Figure 6. The initial weight vector of the ANN has been optimized gradually to achieve minimum error. Some studies have been performed for fertility quality, but the Sperm Whale Algorithm optimization wasn't applied as an optimization technique to optimize the parameters of the classifier. Moreover a comparative analysis of the proposed approach ANN-SWA with ANN, Support Vector Machine (SVM) and decision tree was performed. Results state that proposed ANN-SWA approach outperforms existing methods for fertility quality and comparison graphs plotted out have been shown in Figure 6. Results portrays that the proposed approach is robust. Moreover,

it classifies the fertility samples with a higher accuracy. The experimental results proved that the proposed ANN-SWA achieved massive accuracy, sensitivity, and specificity in predicting the male fertility.

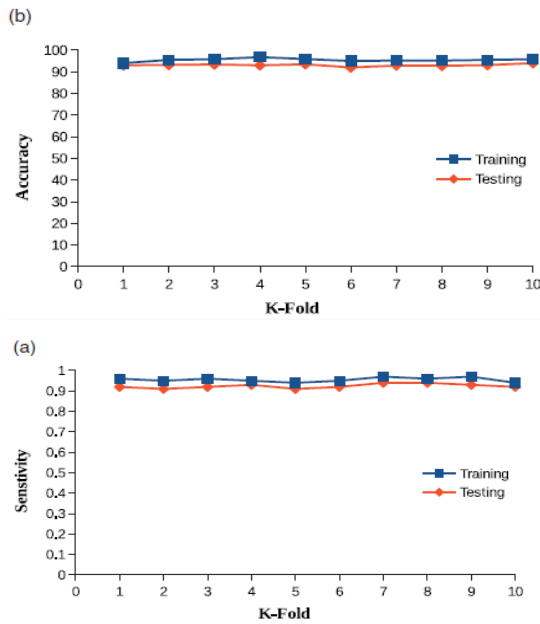


Figure 5. 10-fold effectiveness of classification sensitivity using the proposed approach

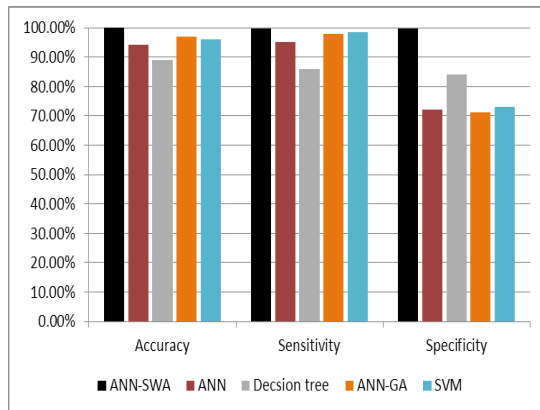


Figure 6. Performance comparison of fertility approach

Furthermore, the importance of ANN-SWA in accelerating classifications operation has been proved by experimental results as shown in Figures 7 a, 7 b, 7c. SWA expression values of machine learning techniques were used to predict the fertility in an effective manner.

Additionally, for the first time, a system of forecasting based on neural network machine was used with the best possible accuracy.

The curve shows the relationship between fertility quality sensitivity and specificity for every possible crossing. The average relative error, the MSE and the worst error.

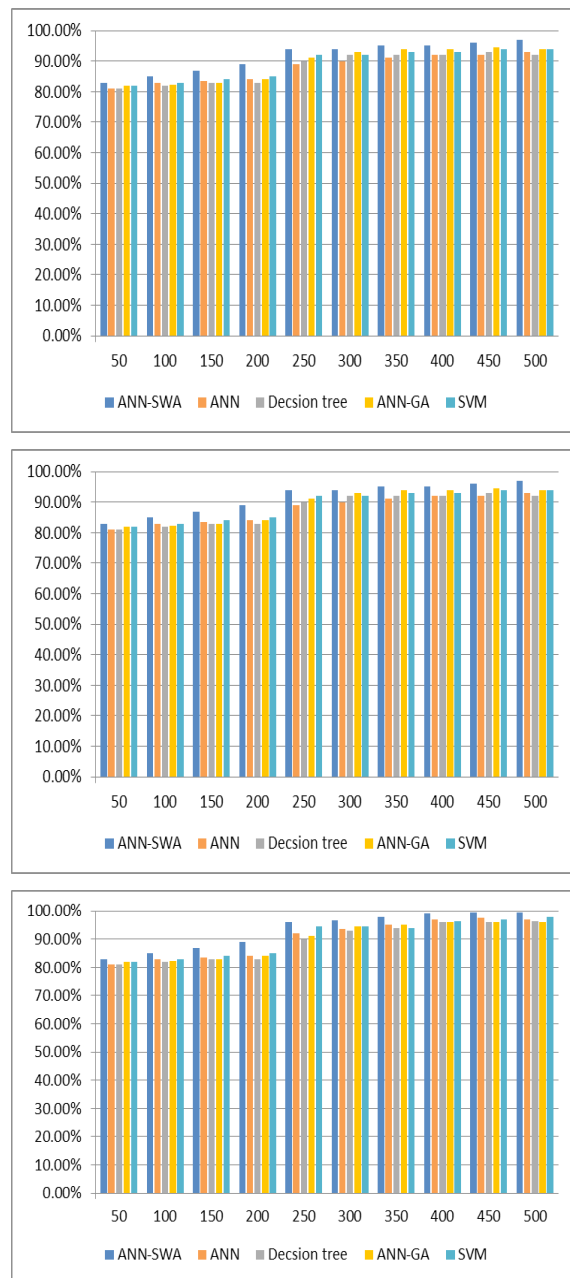


Figure 7 a, b, c The accuracy, sensitivity and specificity of the different fertility approach

Using the feasible approach, the ANN-SWA has been equipped for solving fertility quality. The results presented in this paper show that the ANN-SWA provides superior performance than the before utilized approaches which were tested for the particular the dataset. The proposed ANN-SWA is a simple search model inspired by sperm optimization algorithm. The individual structure of ANN as a new model for developing the optimization method and further improvements in this area can be investigated. Furthermore, randomizing some of the control parameters may have positive effects on the performance of the ANN.

5. Conclusion

This paper introduced a new intelligent approach using neural network based on Sperm Whale Algorithm (SWA). This approach is used to predict the fertility quality in reproductive health for men of different ages. The Sperm Whale Algorithm is used to determine the optimal parameters from the ANN classifier for the dataset used in this paper. The results of the proposed ANN-SWA approach are compared with others methods. It showed the highest performance of the proposed approach and provided the most promising results. For the present ANN-SWA approach, it has been empirically shown that the addition of SWA improves the prediction of fertility.

REFERENCES

1. Aziz, A. S. A., Hassanien, A. E., Hanaf, S. E. O. & Tolba, M. F. (2013, December). Multi-layer hybrid machine learning techniques for anomalies detection and classification approach. In *2013 13th International Conference on Hybrid Intelligent Systems (HIS)* (pp. 215-220). IEEE.
2. Basheer, I. A. & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application, *Journal of microbiological methods*, 43(1), 3-31.
3. Doll, R., Peto, R., Wheatley, K., Gray, R. & Sutherland, I. (1994). Mortality in relation to smoking: 40 years' observations on male British doctors, *Bmj*, 309(6959), 901-911.
4. Dorigo, M., Bonabeau, E. & Theraulaz, G. (2000). Ant algorithms and stigmergy, *Future Generation Computer Systems*, 16(8), 851-871.
5. El-Shafeiy, E. A. & El-Desouky, A. I. (2017). A Big Data Framework for Mining Sensor Data Using Hadoop, *Studies in Informatics and Control*, 26(3), 365-376.
6. Ebrahimi, A. & Khamehchi, E. (2016). Sperm whale algorithm: An effective metaheuristic algorithm for production optimization problems, *Journal of Natural Gas Science and Engineering*, 29, 211-222.
7. Gil, D., Girela, J. L., De Juan, J., Gomez-Torres, M. J. & Johnsson, M. (2012). Predicting seminal quality with artificial intelligence methods, *Expert Systems with Applications*, 39(16), 12564-12573.
8. Heiat, A. (2002). Comparison of artificial neural network and regression models for estimating software development effort, *Information and software Technology*, 44(15), 911-922.
9. <<https://archive.ics.uci.edu/ml/datasets/Fertility>>.
10. <<http://www.who.int/reproductivehealth/topics/infertility/perspective/en/>>.
11. Huang, G. B., Zhu, Q. Y. & Siew, C. K. (2004, July). Extreme learning machine: a new learning scheme of feedforward neural networks. In *Proceedings of the 2004 IEEE International Joint Conference on Neural Networks, Vol. 2* (pp. 985-990). IEEE.
12. Ionescu, V. S., Demian, H. & Czibula, I. G. (2017). Natural Language Processing and Machine Learning Methods for Software Development Effort Estimation, *Studies in Informatics and Control*, 26(2), 219-228.
13. Möhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A. & Lund, A. (2003). The monopulsed nature of sperm whale clicks, *The Journal of the Acoustical Society of America*, 114(2), 1143-1154.
14. Mirjalili, S., Mirjalili, S. M. & Lewis, A. (2014). Let a biogeography-based optimizer train your multi-layer perceptron, *Information Sciences*, 269, 188-209.
15. Nilsson, N. J. & Nilsson, N. J. (1998). *Artificial intelligence: a new synthesis*. Morgan Kaufmann.
16. Piotrowski, A. P. (2014). Differential evolution algorithms applied to neural network training suffer from stagnation, *Applied Soft Computing*, 21, 382-406.
17. Rossett, A. (1987). *Training needs assessment*. Educational Technology.
18. Sotiropoulos, D. N. & Tsihrintzis, G. A. (2016). *Machine Learning Paradigms: Artificial Immune Systems and Their Applications in Software Personalization, Vol. 118*. Springer.
19. TSai, P. W., Pan, J. S., Liao, B. Y. & Chu, S. C. (2009). Enhanced artificial bee colony optimization, *International Journal of Innovative Computing, Information and Control*, 5(12), 5081-5092.
20. Wang, W. & Lu, Y. (2018, March). Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. In *IOP Conference Series: Materials Science and Engineering (Vol. 324, No. 1, p. 012049)*. IOP Publishing.
21. Zhu, G. & Kwong, S. (2010). Gbest-guided artificial bee colony algorithm for numerical function optimization, *Applied mathematics and computation*, 217(7), 3166-3173.