

Multi-Objective Evolutionary Algorithm based on NSGA-II for Neural Network Optimization Application to the Prediction of Severe Diseases

Mansouria Sekkal^{*12}, Amina Benzina¹³, Lahouaria badir Benkrelifa¹

¹Electronics and Telecommunications Department, Faculty of Science and Technology, University of Ain Temouchent, Ain Temouchent, Algeria

² Biomedical Engineering Laboratory, Tlemcen University, Tlemcen, Algeria

³ Applied Materials Laboratory, Djillali Liabes University, Sidi Bel Abbes, Algeria

E-mail: mansouria.sekkal@univ-temouchent.edu.dz, amina.benzina@univ-temouchent.edu.dz, lahouaria.badir@univ-temouchent.edu.dz

*Corresponding author

Keywords: neural networks, multi-objective genetic algorithms, generic optimization, severe diseases.

Received: August 24, 2023

Neural networks have become essential classifiers in various domains, including medicine. The choice of topology, internal structure, and learning algorithm characterizes the type of neural network, resulting in incredible diversity among these networks. Until now, the most challenging problem to solve for classifiers in a neural network has been finding an optimal balance among three selected facets: architecture, synaptic weight, and input variables. To address this problem, we propose a multi-objective neuro-genetic system that simultaneously optimizes these three facets. To demonstrate the effectiveness of our approach, we have implemented and compared two types of classifiers - the classical neural classifier and the multi-objective neuro-genetic classifier - using several medical databases. The results obtained showcase the efficiency of our method, with correct classification rates of up to 100%, which is a very promising outcome. The comparison between the two approaches employed demonstrates the effectiveness of the multi-objective genetic approach

Povzetek: Predstavljen je nov večciljni evolucijski algoritem, zasnovan na NSGA-II, za optimizacijo nevronskih mrež pri napovedovanju hudih bolezni.

1 Introduction

Technological advancements have facilitated the acquisition and collection of extensive data, particularly in the medical field during patient examinations. Medical professionals can use this data to support their decision-making process. "Over the years, medical professionals have increasingly relied on medical diagnostic aids, considering them critical in several medical disciplines [1, 2].

In practice, several applications already exist to support physicians in their diagnostic work [3]. Moreover, artificial learning techniques are continually evolving and becoming more sophisticated [2].

The artificial learning approach most affected is that of artificial neural networks (ANNs). The initial application of neural systems was in the development of pattern recognition systems, such as character recognition, speech recognition, and image contour recognition. They are especially suitable for classification problems, particularly when sufficient data is available. The network's input layer connects directly to the data descriptors, and the output layer corresponds to the classification results.

Several research studies have utilized ANNs to detect and predict severe diseases.

Deepti Sharma and al 2022 aimed to improve breast cancer prediction using artificial neural networks and an extra tree classifier with feature ensemble learning. Results showed the study found that the proposed model achieved an accuracy of 98.42%, sensitivity of 98.90%, and specificity of 98.20% [4].

Hager Saleh and al. 2022 proposed an optimized deep learning approach for predicting breast cancer. The proposed approach combines a deep convolutional neural network (CNN) with a support vector machine (SVM) classifier. The authors optimized the proposed model using grid search and achieved an accuracy of 97.62% [5]. Summrina Kanwal et al. 2021 proposed an innovative method for optimizing shallow machine learning classifiers using artificial immune networks. The authors tested their approach on several medical datasets and demonstrated a significant improvement in classification performance compared to other traditional optimization methods [6].

Mohammed Alweshah et al. 2022, present a novel method for optimizing probabilistic neural networks using the African Buffalo algorithm. The authors tested their

method on several medical datasets and found it to be effective in solving classification problems [7]. Vijayalakshmi Saravanan et al. (2022) proposed using a multilayer perceptron (MLP) model to forecast the occurrence of diabetes mellitus in patients with high accuracy. The paper details the methodology used to train the MLP model and demonstrates that it outperforms other models in terms of accuracy. The authors suggest that using MLP models can significantly enhance the accuracy of diabetes mellitus forecasting, which could have critical implications for disease prevention and management [8]. S. Džaferović and al 2022 discusses the use of artificial neural networks for diagnosing Addison's disease. The study finds that ANNs can effectively diagnose the disease with high accuracy and sensitivity, outperforming traditional approaches. The author suggests that ANNs have potential for other medical applications [9].

2 Background

Neural networks present a great diversity. Indeed, their topology, input parameters, and learning algorithm define a type of neural network. Until now, several problems have remained difficult to solve. These problems are generally associated with learning, the choice of architecture, and the adjustment of input parameters. Generally, the difficulties are related to some rigid parameters to manage, such as the global minimum during learning, the optimal number of neurons for each layer in a multilayer network, the initial values of the network's connection weights during the learning phase, the number of hidden layers to be used, etc.). Poor decisions may result in the corresponding network performing poorly.

Professional's process structured patient data in the form of feature vectors to extract relevant information for medical diagnostics as the basis for problem-solving in this field. The quality of the diagnostic system depends directly on selecting the correct content for these vectors. In various cases, the high dimensionality of these vectors makes finding practical solutions to problems almost impossible. Therefore, reducing the size of the vectors to make them compatible with resolution methods is necessary and helpful. Sometimes, the relevant information for resolving complex phenomena with large descriptors can be represented by only a few features extracted from the initial data.

In the literature, many research works have addressed the area of selecting relevant variables from an input vector for a neural classifier: Choo Jun Tan et al. (2014) proposed Micro Modified Multi-Objective Genetic Algorithms (MMGA) to select input variables. They used this algorithm to perform an ensemble optimizer, which had two purposes: to specify a small number of input variables for classification and to evaluate the effectiveness of the proposed system. The proposed system was involved in human motion detection. [10]. On the other hand, Carlos Affonso et al. 2015 presented a classification approach of biological images through an artificial neural network hybridized with fuzzy logic

ensembles (FANN). The implemented approach improves the reinforcement learning process, focusing on selecting input variables by fuzzy sets. The FANNs employ for classification and fuzzy logic for image cratering [11].

Kabiru O. et al. 2015 proposed correlation-based algorithms to reduce the number of input variables of an ANN used in the predictive search of five separate oil wells. This approach improves search performance and decreases learning time. When selecting the relevant input variables of a neural classifier of different ceramic faults [12], Manasa Kesharaju et al. 2015 used an algorithm based on evolutionary systems and PCA (principal component analysis). They have compared the obtained results classification by the variables selected by this approach with three variable selection methods. The empirical results show that PCA combined with evolutionary systems gives better returns [13]. Antonio Marcio Ferreira Crespo et al. 2021 present an RReliefF-driven feature selection and iterative improvement of the neural network architecture [14].

The choice of ANN architecture is a significant problem that has become very important for researchers. During the learning phase, the user must choose the number of hidden neurons, the number of hidden layers, and their interconnection. People often perform this task in an ad-hoc way or use some simple heuristic rules. In fact, apart from an exhaustive approach, no one knows of a method to determine the optimal architecture for a given problem. However, implementing all the theoretical results on neural networks (computational power or generalization ability) only holds if the architecture is ideal. Shih-Hung Yang et al. 2012 designed an ANN used in prediction based on an evolutionary constructive pruning algorithm (ECPA). In the initial state, they start with a set of ANN with the simplest possible structure, a hidden neuron connected to an input node. The crossover and mutation operators increase the ANN population complexity [15]. In the same vein, Hong-Gui Han et al. 2014, proposed a pruning construction (PC) approach to optimize the neural network structure with a single hidden layer, the hidden neuron numbers described by their contributions; their calculation using a Fourier decomposition of the variance of the output of the hidden layer [16]. On the other hand, Haydee Melo, Junzo Watada 2016, proposed a structural learning algorithm based on a hybridization of a Gaussian with the Particle Swarm Optimization (PSO) method and a fuzzy approach to optimize the weights and structure of a multilayer perceptron. The proposed method has improved the learning and architecture of neural networks [17]. Rajesh Parekh 2000 presented two constructive learning algorithms, the MPyramid-real and MTiling-real, to the ANN construction; this technique eliminated several redundant neurons [18]. N. Dunkin, et al. 1997, and S.E. Fahlman and al. 1990, used constructive approaches to fix the ANN architecture [19, 20]. Tsopze Norbert and al. 2012 used a process based on Galois lattices to define the architecture of ANN [21]. Viet-Khoa Vo-Ho and al 2023, presents a study on the use of neural architecture search (NAS) for medical image applications. The study shows that NAS can significantly improve the performance of

medical image applications and provide more interpretable neural network architectures [22]. Tarun Kumar Gupta and al 2019, review various nature-inspired optimization techniques used for the optimization of artificial neural network (ANN) architecture. The study demonstrates the potential of these techniques for improving the performance of ANNs by finding the best architecture for a specific application. The article also discusses the challenges and limitations of nature-inspired optimization techniques and suggests further research to improve their effectiveness [23].

The problem of learning artificial neural networks and especially the issue of local minima has been a challenge for researchers since the nineties; for a multilayer network, we can observe the same phenomenon. When we randomly initialize the parameters to be optimized (the synaptic weights), starting the learning process at any point of the cost function, and can never be sure to reach the global minimum. Therefore, carrying out several learning phases from different initializations of the weights is a solution. Indeed, we increase our chances by starting to learn in a promising area. Therefore, it is necessary systematically carry out several learning phases from different initial configurations and to choose the one that converges to the lowest error. This approach ensures finding the global minimum on infinite structure numbers is practically not feasible. Juan-Manuel et al. 1998[24], and M. Karouia, and al.1995 [25] have used the simulated annealing minimization algorithm; this algorithm allows to obtain a better convergence toward the global minimum. To solve the same problem, C. Igel et al. 2003 used the Rprop algorithm for learning ANN; experimental results show that this algorithm is more suitable than the backpropagation algorithm [26]. Qun Dai et al. 2012 proposed a modified backpropagation algorithm that remarkably alleviated the local minima problem. This algorithm conducts a competition between the synaptic weights of ANN at iteration t and the synaptic weights at iteration $t-1$ when the output changes during the learning phase. They chose the best significances based on their performance calculated on a validation basis. [27]. On the other hand, Leong Kwan Li 2013 proposed a new optimization algorithm for a single hidden layer multilayer perceptron. The based principle is on a convex combination algorithm for synaptic weights in the hidden layer. This technique explores a continuum idea that combines the mutation and classical crossover strategies in genetic algorithms (GA) [28]. Alireza Askarzadeh et al. 2013 addressed the ANN learning problem with a recently invented Meta heuristic optimization algorithm named BMO "bird mating optimizer", which involved learning the weights of ANN to solve classification problems. They test the algorithm on three databases, Iris flower, breast cancer (Wisconsin breast), and diabetes (Indian Pima) [29]. Other works in the literature address the complexity problem and convergence time of backpropagation. B. Widrow et al. (2013) proposed a new algorithm called "No-Propagation" (No-Prop) for solving the complex problem of learning multilayer perceptrons. They initialized and fixed the weights of the neurons in the hidden layer with random values. They adjusted the

weights of the neurons in the output layer using the LMS algorithm of Widrow and Hoff, choosing the enormous slope to minimize the mean square error. The NO-Prop algorithm is much simpler and easier to implement than the backpropagation algorithm. In addition, it converges faster; but the backpropagation algorithm is more effective for complex architectures [30]. Ozan Kocadaji2015 applied a new hybrid Monte Carlo method with genetic algorithms and fuzzy logic on time series and analyzed the regression in the context of BNN (Bayesian neural network). This method minimized the learning time and gave good performance [31]. Y. S. Kong et al. 2019 applied PSO (Particle Swarm Optimization) to the datasets for ANN weights and biases adjustments, while they defined the mean square error (MSE) as an objective function [32]. Fehmi Burcin Ozsoydan and al 2022, propose a hyper-heuristic-based reinforcement-learning algorithm to train feedforward neural networks. The algorithm outperforms traditional optimization methods on benchmark datasets and has potential for improving the performance of neural networks in various applications. The article highlights the advantages of using reinforcement learning in training neural networks [33]. Adrian Bosire2019 investigates the application of the Artificial Bee Colony (ABC) algorithm to optimize Recurrent Neural Networks (RNNs) for traffic volume analysis. Demonstrating superior results compared to prior statistical or heuristic methods, this research highlights the efficacy of the ABC algorithm in training RNNs for forecasting purposes, with potential applicability across diverse domains [39].

The study by Banaz Anwer Qader and al. 2022 explores the application of the NEAT algorithm to train neural networks for efficient gameplay in the traditional board game Dama. Achieving or surpassing human-level performance, the research emphasizes the significance of abundant input information for optimal learning outcomes [40].

An appropriate architecture choice with a poor input vector choice and limited learning capabilities leads to poor performance. In the opposite case, an ideal choice of architecture or better learning with a non-optimal architecture will influence the results. Therefore, to solve this problem, it is necessary to find an optimal point between the different neural network parameters (architecture, synaptic weights, and input vector). In this study, we propose a new approach based on the hybridization of the multilayer perceptron with multi-objective genetic algorithms, to solve this problem and find an optimal point between the three spaces (architecture, input vector, synaptic weights). We propose this approach for predicting severe diseases.

3 Methodology

In this section, we provide a comprehensive overview of our methodology, focusing on the implementation of the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) in our neuro-genetic system and the comparative analysis

between the classical neural classifier and the multi-objective neuro-genetic classifier.

3.1 Non-dominated sorting genetic algorithm II (NSGA-II)

To begin, it is essential to delve into the technical intricacies of the NSGA-II, the cornerstone of our multi-objective optimization approach. NSGA-II is a genetic algorithm tailored for solving multiple-objective optimization problems. It operates in a step-wise fashion:

- **Initialization:** NSGA-II begins by generating an initial population of potential solutions. Each solution is represented as a chromosome, and these chromosomes make up the population.
- **Fitness evaluation:** The fitness of each solution is evaluated based on its performance concerning all the defined objectives. This is a crucial step in multi-objective optimization.
- **Non-dominated sorting:** Solutions are ranked based on Pareto dominance. A solution is non-dominated if no other solution performs better in all objectives. This results in the classification of solutions into different fronts.
- **Crowding distance assignment:** A crowding distance is calculated for each solution within the same front. This helps maintain diversity in the population by prioritizing solutions that contribute to a more evenly distributed Pareto front.
- **Selection:** A combination of non-dominated sorting and crowding distance is used for selecting the next generation of solutions.
- **Crossover and mutation:** The selected solutions are subjected to crossover and mutation operations, leading to a new population.
- **Termination criteria:** The algorithm iterates through these steps until a termination criterion is met, typically involving the number of generations or computational time.

Implementation in the Neuro-Genetic System

Our neuro-genetic system integrates the NSGA-II algorithm into the optimization process for neural network architectures. Specifically, NSGA-II is employed to optimize the structural and parametric parameters of artificial neural networks (ANNs). This includes determining synaptic weights, input vectors, hidden neurons, and connections between neurons. The NSGA-II algorithm ensures that the resulting network architectures are optimal with respect to multiple, often conflicting, objectives.

3.2 Rationale for choosing NSGA-II

The selection of NSGA-II as our optimization algorithm is grounded in its ability to efficiently explore the solution space and find a set of non-dominated solutions. This algorithm excels in handling multiple conflicting objectives, which is a common challenge in the context of neural network design and optimization.

- **Effectiveness in multi-objective optimization:** In contrast to traditional neural networks, where architecture design is often approached ad hoc, NSGA-II offers an effective means of addressing this issue. Traditional networks frequently rely on ad-hoc architectural choices, while training is conducted using backpropagation. However, this approach may encounter challenges, such as getting stuck in local minima during optimization, especially when dealing with complex neural network structures.
- **Scalability and efficiency:** NSGA-II's efficient non-dominated sorting and crowding distance calculation methods make it practical for larger and more complex optimization problems. This is particularly relevant when considering the impact of input vector size and the relevance of individual variables within those vectors. NSGA-II can efficiently explore architectural design choices and input parameter selection, accommodating the potential influence of vector size and irrelevant variables.
- **Diversity maintenance:** The algorithm's crowded tournament selection operator ensures diversity within the population. This diversity allows NSGA-II to explore a vast solution space and discover a set of solutions that best address the conflicting objectives. This feature is especially advantageous when dealing with architectural optimization and input vector selection, where multiple trade-offs need to be considered simultaneously.

Additionally, the choice of architecture in traditional neural networks is often subjective and may not fully capture the best configuration, leading to suboptimal performance. In contrast, the multi-objective neuro-genetic classifier, optimized by NSGA-II, simultaneously addresses multiple objectives, including classification accuracy, sensitivity, and specificity. This holistic approach helps to find a common ground between these objectives and produces neural network architectures that are more robust, efficient, and better suited to handle a wide range of input vector sizes and variable relevancy scenarios.

4 Implementation in the neuro-genetic system

In Figure 1, we present our multi-objective neuro-genetic system. Each chromosome of the population represents different constraints. The constraints present the different structural and parametric parameters of an ANN, which include:

- Synaptic weight
 - Input vector
 - Hidden neurons
 - Connections between neurons
- Encoding a network on a chromosome

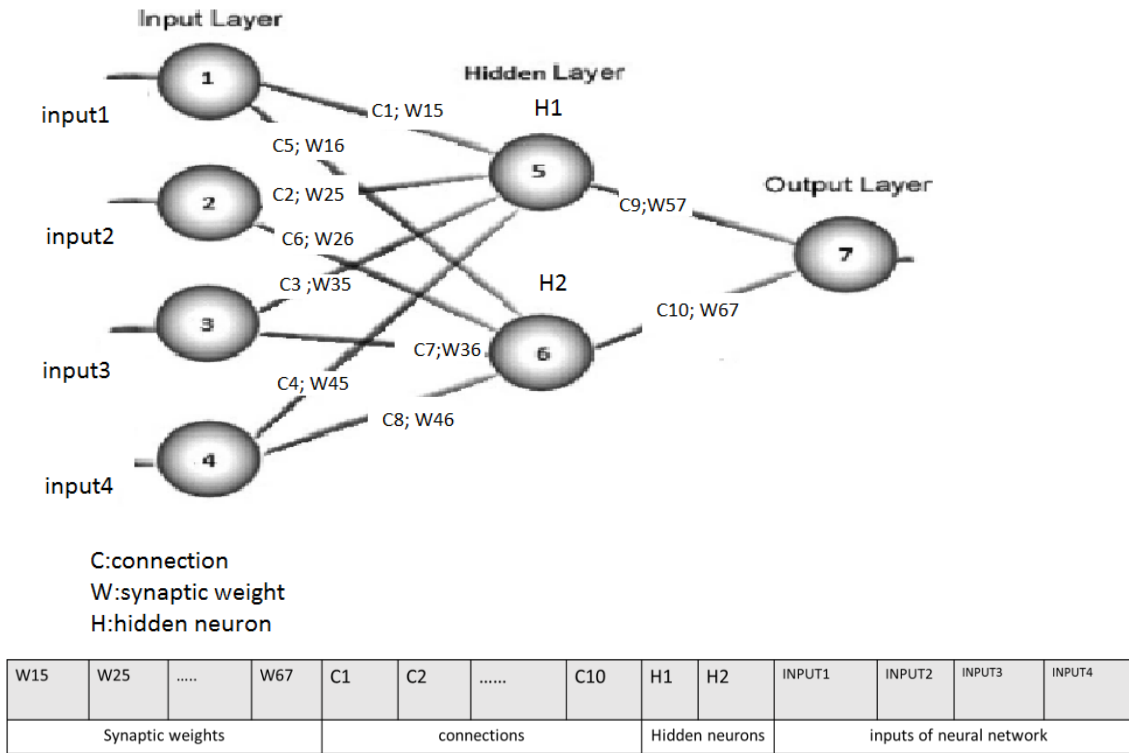


Figure 1: Multi-objective neuro-genetic system

5 Experimental results

To evaluate the effectiveness of our approach, we developed and compared two neural classifier models: a classical neural classifier (CNC) and a multi-objective neuro-genetic classifier (NGMOC). To demonstrate the We used these databases to train and test the classifiers. The number of input parameters varies between six and nineteen descriptors, depending on the problem. There are two class numbers: positive cases and negative cases. Table 2 indicates the number of cases we used to train and test each database.

Table 1 : Databases used

Data set	Bupa	Breast W	Pima
size	345	699	768
input	6	9	8
Positive case	200	241	268
Negative case	145	458	500
classes	2	2	2

Table 2 : Number of cases used for training and testing for each database

Data set		Bupa	Breast W	Pima
TRAIN	P	100	100	120
	N	70	100	200
TEST	P	100	141	148
	N	75	358	300

P: positive cases N: negative cases

5.1 Classical neural classifier (CNC)

In most neural network applications, the architecture, neuron count, and connections are determined through heuristics and testing. This section focuses on the training process of the classical neural classifier (CNC) using the backpropagation algorithm.

- **Backpropagation algorithm:** The backpropagation algorithm is a supervised learning method used to minimize the global squared error of the network. It employs the gradient descent algorithm to adjust the network's synaptic weights. Backpropagation has

gained popularity due to its simplicity and effectiveness in artificial neural networks (ANNs).

- **Adjusting parameters:** Before starting the learning phase, we adjust several parameters, such as setting the number of iterations to 1200, which ensures a satisfactory learning process. However, finding the right balance is crucial as an excessively large number of iterations may result in overlearning, while too few iterations may impede learning progress.
- **Random initialization:** We randomly choose the initial synaptic weights, introducing diversity into the network and preventing it from getting trapped in suboptimal solutions.
- **Error threshold:** The global error threshold is set to 0.001 after conducting various experimental trials. This threshold serves as a stopping criterion, indicating that the learning process will continue until the error falls below this value.
- **Learning rate:** We set the learning rate, also referred to as the learning step, to 0.3. This parameter determines the magnitude of adjustment applied to the synaptic weights in each iteration of the backpropagation algorithm. Choosing the right learning rate is essential for successful training. A higher value can result in faster convergence, but it may also lead to overshooting and instability.
- **Training process:** We continue the learning process until the error falls below the predefined threshold. This iterative approach allows the backpropagation algorithm to update the synaptic weights, gradually minimizing the error and improving the classifier's performance.

We train the classical neural classifier (CNC) using the backpropagation algorithm with the objective of minimizing the global squared error. Adjusting parameters such as the number of iterations, initial weights, error threshold, and learning rate plays crucial roles in achieving effective training.

Table 3: Structure of the classifiers for each database.

DATA SET	Bupa	Breast W	Pima
INPUT	6	9	8
FEATURES	6	9	8
HIDDEN NEURON	3	4	4
CONNEXIONS	21	40	36

5.2 Neuro-genetic multi-objective classifier (NGMOC)

In the field of artificial neural networks, optimizing the quadratic error is a common approach. However, this optimization method often fails to achieve high sensitivity and specificity rates, which, in turn, affect the overall classification accuracy. This limitation prevents reaching a 100% correct classification rate due to the unidirectional nature of the optimization process.

- **Objective of the NGMOC classifier:** The NGMOC classifier aims to optimize both the sensitivity and classification rate of the training dataset simultaneously, in order to achieve satisfactory results.
- **Diversity in neural networks:** Neural networks exhibit significant diversity, with their topology, internal structure, and learning algorithms defining different types of networks.
- **The challenge of finding the best point:** One of the most significant challenges in neural network optimization is finding the optimal point in a three-dimensional space that includes architecture, synaptic weights, and input variables. The goal of generic optimization is to solve these problems simultaneously and identify the best configuration.
- **Genetic multi-criteria and multi-objective optimization:** To address this challenge, we developed a classifier using genetic multi-criteria and multi-objective optimization techniques. Specifically, we utilized the MOGA (Multi-Objective Genetic Algorithms) along with the NSGA-II (Non-dominated Sorted Output Genetic Algorithm-II) algorithm.
- **Population size and constraints:** After conducting several tests, we determined that a population size of 980 chromosomes would be suitable. Each chromosome represents different constraints related to the structural and parametric parameters of the artificial neural network (ANN). These constraints include synaptic weights (authentic coding between -1 and 1), input vectors (authentic coding between 0 and 1), hidden neurons (binary coding), and connections between neurons (binary coding).
- **Fitness function:** The fitness function considered in this study includes two objectives:

$$Multi (1) = 1 - \text{sensitivity of the learning dataset}$$

$$Multi (2) = 1 - \text{the correct classification rate of the training set}$$

- **Genetic selection and evolution:** The NSGA-II algorithm assigns a rank based on dominance after loading the initial values of all chromosomes. The casino roulette selection method is employed to choose chromosomes for mutation and crossover operations. The mutation probability rate (Pm) is set to 0.01. The combination of mutated and crossed chromosomes forms a new population for the next generation. The probabilities undergo careful selection through multiple trials to ensure optimal

performance. The NSGA-II algorithm assigns efficiency based on the rank of chromosomes.

- **Continued evolution and objective:** The iterative process continues until the best solution is found, characterized by a minimal number of connections

and hidden neurons, as well as a relevant selection of input variables. The objective is for the MOGA to discover a common point that optimally balances the three surfaces of input variables, architecture, and learning.

W1	W2	W40	C1	C2	C40	H1	H4	INPUT1	INPUT2	INPUT9
Synaptic weights				connections				Hidden neurons			inputs of neural network			
chromosome of BREAST W														
W1	W2	W21	C1	C2	C21	H1	H3	INPUT1	INPUT2	INPUT6
Synaptic weights				connections				Hidden neurons			inputs of neural network			
chromosome of BUPA														
W1	W2	W36	C1	C2	C36	H1	H4	INPUT1	INPUT2	INPUT8
Synaptic weights				connections				Hidden neurons			inputs of neural network			
chromosome of PIMA														

Figure 2: Chromosome coding for each database

6 Results and discussion

Several statistical criteria, including sensitivity, specificity, and correct classification rate, have been computed and are presented in tables 6 and 7. The classification rate performance is evaluated by presenting each example e_j from the test dataset to the classifier and comparing the resulting class $C(e_j) = s$ with the true class of e_j . The test dataset contains N objects, and we denote the number of correctly classified objects by the system as $N_{correct}$. We calculate the classification rate using the following formula:

$$Tclass = \frac{N_{correct} \cdot 100}{N}$$

$$Se(i) = \frac{TP(i)}{TP(i) + FN(i)}$$

$$sp(i) = \frac{TN(i)}{TN(i) + FP(i)}$$

The formulas mentioned above use the following terms:

- True Positives (TP): The number of positive instances correctly classified as positive.
- False Positives (FP): The number of negative instances incorrectly classified as positive.
- True Negatives (TN): The number of negative instances correctly classified as negative.
- False Negatives (FN): The number of positive instances incorrectly classified as negative.

Table 4: Performance of CNC

DATA SET	BUPA	BREAST W	PIMA
TP	77	139	122
TN	50	330	243
FP	25	28	57
FN	23	02	26
SE	77%	98.58%	82.43%
SP	66%	92.17%	81%
CC	72.55%	93.98%	81.47%
Hidden neurons	3	4	4
CONNEXIONS	21	40	36

Table 5: Performance of NGMOC

DATA SET	BUPA	BREAST W	PIMA
TP	87	141	142
TN	99	358	290
FP	6	00	10
FN	13	00	06
SE	87%	100%	95.94%
SP	88%	100%	96.66%
CC	87.42%	100%	96.42%
Hidden neurons	2	2	3
CONNEXIONS	12	18	24

In this study, we meticulously compared the performance of a classical neural classifier with a pioneering approach that amalgamates artificial neural networks (ANN) with MOGA multi-objective genetic algorithms. The outcomes unveiled substantial enhancements, particularly in mitigating the occurrences of false negatives and false positives. For example, in the context of the breast cancer database, our novel NGMOC classifier achieved an exemplary reduction in false negatives and false positives, both attaining a value of 0. Simultaneously, there was a noteworthy upsurge in true positives and true negatives. Tables 4 and 5 offer comprehensive comparisons encompassing key performance metrics, including sensitivity, specificity, correct classification rates, the number of hidden neurons, and connection numbers for

both classifiers. The outcomes unambiguously illustrate that the integration of neural networks with multi-objective genetic algorithms facilitates the convergence to a common point within a three-dimensional space. This convergence, in turn, yields heightened performance concerning sensitivity, specificity, and correct classification rates. Notably, the breast cancer classifier reached a flawless correct classification rate, sensitivity, and specificity of 100%, while maintaining a relatively streamlined structure with eighteen connections. Figure 3 illustrates the architecture of both classifiers for the breast cancer database, highlighting the remarkable simplicity of the architecture of the multi-objective neuro-genetic classifiers compared to classical neural classifiers. Furthermore, the simplified architecture of the MONGC classifier allowed for the characterization of the input variables for each database. Table 6 presents the detailed characterization of the input variables for the various databases used in the study.

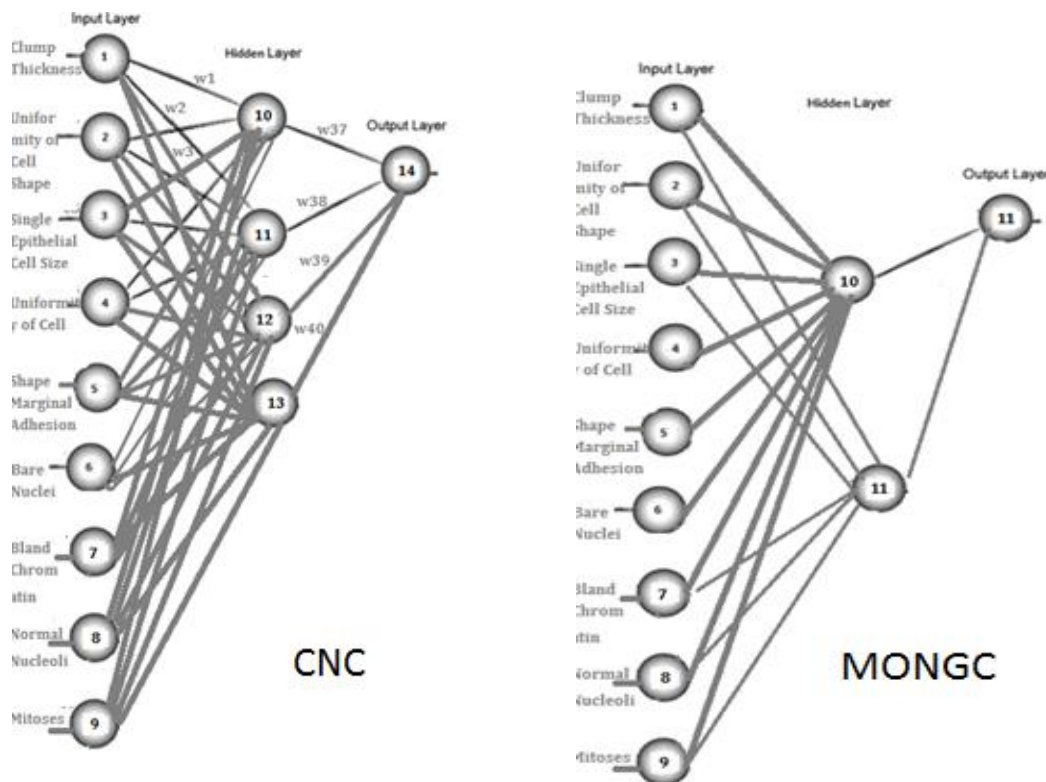


Figure 3: Comparison between the architecture of CNC and MONGC for breast cancer detection

Table 6: Characterization of the input variables of each database

PUBA	Characterization	Brest cancer	Characterization	PIMA	Characterization
MCV	0.4479	Clump Thickness	0.4	N PREG	0.123
ALKHAPHOS	0.1806	Uniformity of Cell	0.301	Glu	0.789
SGPT	0.9011	Single Epithelial Cell Size,	0.698	PAD	0.879
SGOT	0.5863	Uniformity of Cell Shape	0.936	Epai	0.894
GAMMAGT	0.7971	Shape Marginal Adhesion	0.923	INS	0.897
DRINKS	0.1254	Bare Nuclei	0.339	IMC	0.6754
		Bland Chromatin	0.430	Ped	0.234
		Normal Nucleoli	0.347	AGE	0.654
		Mitoses	0.945		

The NGMOC classifier assigned weights to different input variables, revealing their relative importance in the classification process. In the case of breast cancer, the classifier assigned lower weights to descriptors such as Clump Thickness, Uniformity of Cell Shape, Bare Nuclei, Normal Nucleoli, and Bland Chromatin, while placing crucial importance on descriptors such as Single Epithelial Cell Size, Uniformity of Cell Shape, Marginal Adhesion, and Mitoses. These results indicate the significance of these four descriptors for the automated detection of breast cancer.

Similarly, for the Bupa database, the classifier emphasized the variable SGPT over ALKHAPHOS and drinks. In the case of the PIMA database, the best-selected input variables were INS, Epai, and PAD. Overall, our approach has improved the performance of neural classifiers in medical diagnostics by utilizing an optimal architecture and effectively characterizing the input variables.

The histogram provides a clear visual representation of the comparative performance of CNC and NGMOCO classifiers across three distinct datasets: bupa, breast, and pima. Notably, the NGMOCO classifier consistently outperforms the CNC counterpart, achieving higher correct classification rates in all cases. The superior accuracy demonstrated by NGMOCO, especially in the 'breast' dataset, highlights its potential to significantly

enhance diagnostic precision in medical applications. These results underscore the relevance of integrating artificial neural networks with multi-objective genetic algorithms for improved classification, with promising implications for medical decision-making.

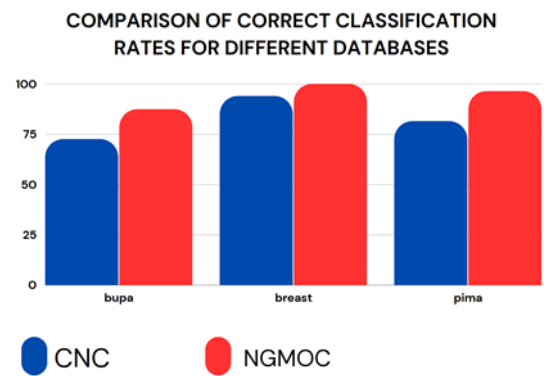


Figure 4: Comparison of correct classification rate for different databases

7 Comparing optimization approaches for neural networks

Table 7: Comparative analysis of neural network optimization approaches

Author and Year	Approach	Optimization Method	Optimized Parameters
Antonio Marcio Ferreira Crespo et al. (2021)[14]	RReliefF-Driven Feature Selection and Iterative Neural Network Architecture Optimization	RReliefF for feature selection	Features and Architecture
Viet-Khoa Vo-Ho et al. (2023) [22].	Neural Architecture Search (NAS)	NAS for neural architecture optimization	Architecture
Tarun Kumar Gupta et al. (2019) [23].	Nature-Inspired Optimization Techniques for ANN Architecture	Nature-inspired optimization techniques	Architecture
Y. S. Kong et al. (2019)[32]	ANN Optimization with PSO (Particle Swarm Optimization)	PSO for adjusting ANN weights and biases	Weights and Biases
Fehmi Burcin Ozsoydan et al. (2022)[33]	Hyper-Heuristic Reinforcement Learning Algorithm for Feedforward Neural Networks	Hyper-heuristic-based reinforcement learning method	Architecture

Our approach - Neuro-Genetic Multi-Objective Classifier (NGMOC)	Classification with Optimization of Architecture, Synaptic Weights, Input Vectors, Hidden Neurons, and Connections Between Neurons	NSGA-II for simultaneous optimization of parameters	Architecture, Weights, Input Hidden Connections	Synaptic Vectors, Neurons,
---	--	---	---	----------------------------

Table 7 provides a comparative analysis of various neural network optimization approaches, each addressing specific aspects of neural network design and performance improvement. While these methods have made valuable contributions to the field, the "Neuro-Genetic Multi-Objective Classifier (NGMOC)" stands out by offering a more comprehensive and holistic solution. It simultaneously optimizes multiple parameters, including architecture, synaptic weights, input vectors, hidden neurons, and connections between neurons, using the NSGA-II algorithm. By addressing a wider range of optimization objectives and parameters, NGMOC significantly enhances the capabilities of neural network classifiers. It results in improved classification performance, as evidenced by the notable reduction in false negatives and false positives. Furthermore, the method excels in terms of sensitivity, specificity, and correct classification rate.

This comprehensive approach demonstrates the potential to overcome the limitations of more specialized techniques and presents a versatile solution for neural network optimization. NGMOC, as outlined in the table, offers a novel and advanced way to optimize neural network performance, making it a valuable contribution to the field. To understand these differences, we can consider the nature of our multi-objective neuro-genetic approach. Unlike many of the studies listed, our approach simultaneously optimizes the architecture, synaptic weights, and input variables, making it a holistic and versatile solution applicable to a wide range of tasks. This comprehensive optimization may contribute to the achievement of high correct classification rates. Additionally, our approach leverages genetic algorithms, which are capable of exploring a broader search space for optimal solutions. This can be especially advantageous when dealing with complex and diverse input data. The genetic algorithms in our system offer a level of adaptability and exploration that classical methods often lack.

Novelty of the multi-objective neuro-genetic system:

Our multi-objective neuro-genetic system stands out in the field due to its ability to handle complex problems, adapt to various domains, and provide consistently high accuracy. The combination of neuro-genetic techniques allows for not only effective feature selection but also optimization of the neural network architecture and synaptic weights.

This approach addresses the limitations of classical neural classifiers, which are often designed for specific applications and may not generalize well to other domains or handle complex input data. By achieving up to 100% correct classification rates, our system demonstrates its potential as a novel and versatile solution for a wide range of classification tasks, making it a valuable addition to the field of neural network optimization.

However, it is essential to acknowledge the limitations of our approach. One major limitation is that the optimization process may require more computational resources and time compared to single-objective approaches. Additionally, fine-tuning multiple parameters can be challenging, and the effectiveness of our method may vary depending on the specific dataset and problem.

Another limitation to consider is related to the hybridization of the NSGA algorithm with the multi-layer perceptron. This hybridization may pose challenges when dealing with large databases. The multi-layer perceptron, while powerful, may have its limitations when handling large datasets. In the future, an important perspective will be to extend this hybridization to deep learning convolutional networks. These networks, particularly in the field of computer vision, have demonstrated outstanding performance, but their integration with our approach will require thorough adaptation and exploration.

Therefore, future research should focus on developing more efficient optimization algorithms to mitigate these limitations and further enhance the robustness of our approach, while exploring the application of our method to the domain of deep learning convolutional networks.

8 Ethical considerations

The deployment of our optimized neural network approach for the prediction of severe diseases raises significant ethical considerations. It is crucial to acknowledge that the use of advanced technologies in the medical field comes with ethical responsibilities to ensure accuracy, confidentiality, and impact on patients. First and foremost, it is essential to ensure that the data used to train and test the model are collected, stored, and used ethically, following relevant regulations on privacy protection and data security. Informed consent from patients must be obtained when necessary, and data should be anonymized to prevent the unauthorized disclosure of personal information. Furthermore, the use of these models for the prediction of severe diseases should be carried out by qualified healthcare professionals. The results of these models should be interpreted with caution, and final medical decisions should not solely rely on the model's predictions. We also advocate for full transparency regarding how the models were trained, the data that fueled them, and the limitations of these models. This builds trust among healthcare professionals, patients, and the public. Ultimately, while our approach offers substantial potential to enhance severe disease prediction, maintaining strict ethical standards is imperative to ensure that these technologies are used responsibly and ethically for the benefit of patients.

9 Conclusion

This paper introduces an innovative approach to tackling medical multi-classification problems using a combination of artificial neural networks (ANN) and multi-objective genetic algorithms. Through the synergy of these two techniques, our objective is to pinpoint a convergence point within the ANN space, effectively addressing three critical aspects: architecture, learning, and input variables.

Our proposed method, the Neuro-Genetic Multi-Objective Classifier (NGMOC), showcases remarkable advancements in recognition rates when applied to diverse databases, outperforming conventional classifiers with optimized architectures and input variable characterizations. The empirical results underscore a clear and substantial enhancement in performance, underscoring the effectiveness of our approach in meeting the challenges of medical multi-classification.

References

- [1] R. A. Miller (1994). *Medical diagnostic decision support systems - past, present, and future: a threaded bibliography and brief commentary*. Journal of the American Medical Informatics Association.1(1):8-27 <https://doi.org/10.1136/jamia.1994.95236141>
- [2] E. Coiera (2003). *Guide to health informatics Hodder& Stoughton Educational*, UK, 2nd edition.
- [3] M. Huguier and A. Flahault (2003). *Biostatistiques au quotidien*. Elsevier, 2003.
- [4] Sharma, D., Kumar, R., & Jain, A. (2022). *Breast cancer prediction based on neural networks and extra tree classifier using feature ensemble learning. Measurement: Sensors*, 24, 100560. <https://doi.org/10.1016/j.measen.2022.100560>
- [5] Hager Saleh and al (2022). 'Predicting Breast Cancer Based on Optimized Deep Learning Approach'. *Computational Intelligence and Neuroscience* <https://doi.org/10.1155/2022/1820777>
- [6] Kanwal S, Hussain A, Huang K. (2021). *Novel artificial immune networks-based optimization of shallow machine learning (ML) classifiers. Expert Systems with Applications*. 165: 113834. <https://doi.org/10.1016/j.eswa.2020.113834>
- [7] M. Alweshah and al. (2022). *African buffalo algorithm: training the probabilistic neural network to solve classification problems. J King Saud Univ, Comput Inf Sci*. Volume 34, Issue 5 , Pages 1808-1818, May 2022. <https://doi.org/10.1016/j.jksuci.2020.07.004>
- [8] Vijayalakshmi S and al, *Reliable diabetes mellitus forecasting using artificial neural network multilayer perceptron*, *Artificial Intelligence and Machine Learning for EDGE Computing*, Pages 121-131, 2022 <https://doi.org/10.1016/B978-0-12-824054-0.00013-7>
- [9] S. Džaferović (2022), *Diagnosis of Addison's disease Using Artificial Neural Network IFAC-PapersOnLine* Volume 55, Issue 4, Pages 68-73. <https://doi.org/10.1016/j.ifacol.2022.06.011>
- [10] Choo Jun Tan, Chee Peng Lim, Yu-N Cheah (2014). *A multi-objective evolutionary algorithm-based ensemble optimizer for feature selection and classification with neural network models. Neurocomputing*, Volume 125, 11, Pages 217–228. <https://doi.org/10.1016/j.neucom.2012.12.057>
- [11] Carlos Affonso, Renato Jose Sassi, Ricardo Marques Barreiros (2015). *Biological image classification using rough-fuzzy artificial neural network; Expert Systems with Applications*. Volume 42, Issue 24, Pages 9482-9488. 30 December 2015. <https://doi.org/10.1016/j.eswa.2015.07.075>
- [12] Kabiru O. Akande, Taoreed O. Owolabi, Sunday O. Olatunji (2015). *Investigating the effect of correlation-based feature selection on the performance of neural network in reservoir characterization. Journal of Natural Gas Science and Engineering*, Volume 27, Part 1, Pages 98-108. <https://doi.org/10.1016/j.jngse.2015.01.007>
- [13] Manasa Kesharaju, Romesh Nagarajah (2015) *Feature selection for neural network-based defect classification of ceramic components using high frequency ultrasound; Ultrasonics*, Volume 62, Pages 271-277. <https://doi.org/10.1016/j.ultras.2015.05.027>
- [14] Antonio Marcio Ferreira Crespo, Chun Wang, Thiago Marques Ferreira Crespo, Li Weigang, Alexandre Barreto (2021). *Learning framework for carbon emissions predictions incorporating a RReliefF driven features selection and an iterative neural network architecture improvement. SN Applied Sciences* 3:460, 2021. <https://doi.org/10.1007/s42452-021-04411-z>
- [15] Shih-Hung Yang, Yon-Ping Chen. *An evolutionary constructive and pruning algorithm for artificial neural networks and its prediction applications; Neurocomputing*, Volume 86, Pages 140-149, 1 June 2012 <https://doi.org/10.1016/j.neucom.2012.01.024>
- [16] Hong-Gui Han, Li-Dan Wang, Jun-Fei Qiao (2014). *Hierarchical extreme learning machine for feedforward neural network. Neurocomputing*, Volume 128, Pages 128-135, 27 <https://doi.org/10.1016/j.neucom.2013.01.057>
- [17] Haydee Melo, Junzo Watada. *Gaussian - PSO with fuzzy reasoning based on structural learning for training a Neural Network. Neurocomputing*, Volume 172, Pages 405-412, 8 January 2016. <https://doi.org/10.1016/j.neucom.2015.03.104>
- [18] R. Parekh, J. Yang, et V. Honavar (2000). *Constructive Neural-Network Learning Algorithms for Pattern Classification. IEEE Transactions on Neural Networks*, volume 11(2) 2000, pages 436–451. <https://doi.org/10.1016/j.neucom.2015.03.104>
- [19] Dunkin, N., J. Shawe-Taylor, and P. Koiran (1977) *A New Incremental Learning Technique*. In *Neural Nets Wirm Vietri-96. Proceedings of the eighth Italian workshop on neural nets*, 112–118. London: Springer. http://dx.doi.org/10.1007/978-1-4471-0951-8_8
- [20] S. E. Fahlman and C. Lebiere (1990). *The cascade-correlation learning architecture*. In D. S. Touretzky, *Advances in Neural Information Processing Systems*, , Morgan Kaufmann, San Mateo, Denver 1989, 1990.

- [21] TSOPZE Norbert. *Galois lattice and neural networks* (2012): *a constructive approach to neural network architecture*; PhD thesis, University of Artois, France .
- [22] Viet-Khoa Vo-Ho, Kashu Yamazaki, Hieu Hoang, Minh-Triet Tran, and Ngan Le (2023). *Neural architecture search for medical image applications*. In *Meta-Learning with Medical Imaging and Health Informatics Applications*, pages 369–384. Elsevier, 2 <https://doi.org/10.1016/B978-0-32-399851-2.00029-6>
- [23] Gupta T.K., Raza K.(2019). Chapter 7 *Optimization of ANN Architecture: A review on Nature-Inspired Techniques*, Editor(s): Nilanjan Dey, Surekha Borra, Amira S. Ashour, Fuqian Shi, Machine Learning in Bio-Signal Analysis and Diagnostic Imaging, Academic Press, Pages 159–182. <https://doi.org/10.1016/B978-0-12-816086-2.00007-2>
- [24] J.-M. Torres-Moreno and M.B. Gordon (1998). *Efficient adaptive learning for classification tasks with binary units*, *Neural Computation*, vol. 10, no. 4, pp. 1007 - 1030. <https://doi.org/10.1162/089976698300017601>
- [25] M. Karouia, R. Lengellé, and Denoeux T.(1995). *Performance analysis of a MLP weight initialization algorithm*. In Michel Verleysen, editor, *European Symposium on Artificial Neural Networks*, Brussels.
- [26] C. Igel and M. Husken. Sylvain Tertois (2003). *Reduction des effets des non-linearites dans une modulation multi porteuse à l'aide de reseaux de neurones*. PhD thesis, Rennes 1.
- [27] Qun Dai, Ningzhong Liu (2012). *Alleviating the problem of local minima in Backpropagation through competitive learning*. *Neurocomputing*, Volume 94, Pages 152–158. 1 October <https://doi.org/10.1016/j.neucom.2012.03.011>
- [28] Leong Kwan Lia, Sally Shaob, Ka-Fai Cedric Yiu (2013), *A new optimization algorithm for single hidden layer feedforward neural network*; *Applied Soft Computing*, Volume 13, Issue 5, Pages 2857–2862. <https://doi.org/10.1016/j.asoc.2012.04.034>
- [29] Alireza Askarzadeh, Alireza Rezazadeh (2013). *Artificial neural network training using a new efficient optimization algorithm*, *Applied Soft Computing*, Volume 13, Issue 2, Pages 1206–1213. <https://doi.org/10.1016/j.asoc.2012.10.023>
- [30] Bernard Widrow, Aaron Greenblatt, Youngsik Kim, Dookun Park (2013). *The No-Prop algorithm. A new learning algorithm for multilayer neural networks*. *Neural Networks* Volume 37, Pages 182–188. <https://doi.org/10.1016/j.neunet.2012.09.020>
- [31] Ozan Kocadağlı (2015). *A novel hybrid learning algorithm for full Bayesian approach of artificial neural networks*. *Applied Soft Computing* Volume 35, Pages 52–65. <https://doi.org/10.1016/j.asoc.2015.06.003>
- [32] Y. S. Kong, S. Abdullah, D. Schramm, M. Z. Omar and S. M. Haris(201ç) ; *Design of artificial neural network using particle swarm optimisation for automotive spring durability* ; *Journal of Mechanical Science and Technology* 33 (11) 5137~5145. <http://dx.doi.org/10.1007/s12206-019-1003-9>
- [33] Fehmi Burcin Ozsoydan and İlker Gölcük (2022), 'A hyper-heuristic based reinforcement-learning algorithm to train feedforward neural networks. *Engineering Science and Technology*, an International Journal Volume 35, 101261;
- [34] Kumar L., Kumar K., Chhabra D. (2022). *Experimental investigations of electrical discharge micro-drilling for Mg-alloy and multi-response optimization using MOGA-ANN*, *CIRP Journal of Manufacturing Science and Technology*, 38 , pp. 774-786. 2022 . <https://doi.org/10.1016/j.cirpj.2022.06.014>
- [35] Jiabing Wang, Linlang Zeng, Kun Yan (2023). *Multi-objective optimization of printed circuit heat exchanger with airfoil fins based on the improved PSO-BP neural network and the NSGA-II algorithm*, *Nuclear Engineering and Technology* Volume 55, Issue 6, Pages 2125-2138. <https://doi.org/10.1016/j.net.2023.02.029>
- [36] David Panzoli (2003). *Simulation comportementale par réseau de neurones et apprentissage par algorithme génétique*. DEA Informatique de l'image et du langage .
- [37] V. Maniezzo (1993). *Searching among space search: hastening the genetic evolution of feedforward neural networks in Proceedings of ANNGA*.
- [38] A; Asuncion , D. J. Newman, (2007) . *UCI machine learning repository*. Irvine, CA: University of California, School of Information and Computer Science. <http://www.ics.uci.edu/mllearn/MLRepository.html>
- [39] Adrian Bosire (2019), *Recurrent Neural Network Training using ABC Algorithm for Traffic Volume Prediction*, Vol 43, No 4 (2019), *Informatica* 43 (2019), 551–559. <https://doi.org/10.31449/inf.v43i4.2709>
- [40] Banaz Anwer Qade and al (2022), *Evolving and training of neural network to play dama board game using neat algorithm*, Vol 46, No 5 (2022), *Informatica* 46 (2022) 29–37. <https://doi.org/10.31449/inf.v46i5.3897>