جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة ديالى / كلية العلوم
قسم علوم الحاسوب

نموذج تكيفي مبني على خوارزميات ارشادية متقدمة

الاطروحة مقدمة الى قسم علوم الحاسوب / كلية العلوم/ جامعة ديالى وهي جزء من متطلبات نيل درجة الماجستير في علوم الحاسوب

الباحث

تقى علي محمد

بإشراف

أ.م. د منتظر خميس مصطفى

1444 هـ

2023 م

Chapter One

General Introduction

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Overview

Artificial neural networks are a machine learning algorithm that imitates the structure and function of the human brain [1]. They are composed of interconnected nodes, or "neurons," which analyze information and make predictions based on patterns in data [2]. Neural networks have been used in a wide range of applications, [3,4] from image recognition and natural language processing to financial forecasting and medical diagnosis[5]. While they can be powerful tools for solving complex problems, they also require large amounts of data and computing power to train effectively [6]. MLP networks, or Multilayer Perceptron networks, are a type of artificial neural network that is commonly used in machine learning applications. Theyconsist of multiple layers of interconnected nodes, with each node performing a simple mathematical operation on its inputs and passing the result to the next layer[7]. MLP networks are often used for tasks such as classification and regression, and they can be trained using a variety of algorithms such as back propagation. While MLP networks have been aroundfor several decades, they continue to be an important tool in the field of machine learning due to their versatility and effectiveness[8]. Tuning an MLPnetwork! That can be a tricky task. When it comes to optimizing the performance of a multilayer perceptron (MLP) neural network, there are several factors to consider. One of the most important is the choice of

activation function for each neuron in the network. Different activation functions can have a significant impact on the network's ability to learn and generalize from training data. Another key factor is the number of hidden layers and neurons in each layer. Too few hidden layers or neurons can result in under fitting, while too many can lead to overfitting. It is important to strike a balance between complexity and generalization ability. Other factors that can affect MLP performance include learning rate, batch size, regularization techniques, and optimization algorithms such as stochastic gradient descent. Overall, tuning an MLP network requires careful experimentation and analysis to find the optimal combination of parameters for a given task [9]. The following Figure (1) show the General MLP Architecture.

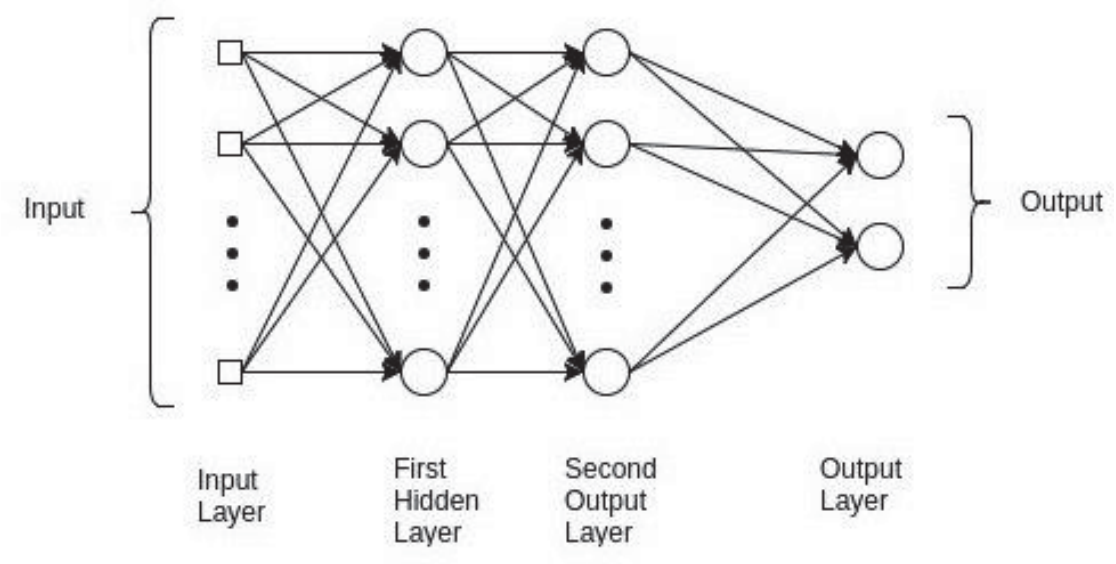

Figure (1.1). General MLP Architecture [10].

1.2 Problem statement

1-Tuning an MLP network is a tricky task. When it comes to optimizing the performance of a multilayer perceptron (MLP) neural network, there are several factors to consider. One of the most important is the choice of activation function for each neuron in the network. Different activation functions can have a significant impact on the network's ability to learn and generalize from training data.

2-An important aspect to consider is the quantity of concealed layers and neurons present in each layer. Too few hidden layers or neurons can result in under fitting, while too many can lead to over fitting.

3-It is important to strike a balance between complexity and generalization ability. Other factors that can affect MLP performance include learning rate, batch size, regularization techniques, and optimization algorithms such as stochastic gradient descent.

4- Overall, tuning an MLP network requires careful experimentation and analysis to find the optimal combination of parameters for a given task.

5- he pattern problem and the neurons number in the output layer equal to the number of class, for the choice of layers number and neurons in each layers and connections called architecture problem, our main objectives is to optimize it for suitable network with sufficient parameters and good generalization for classification or regression task.

1.3 The aims of study

1- Determine the most effective weights and biases
2- Developing Hybrid trainer of Multilayer perceptron with modern Meta-Heuristic Algorithms.
3- Update search space of ANN through great optimal balancing in the problem domain.

1.3 Literature Survey

Meta-heuristics are an essential part of figuring out how to solve optimization problems. Most of these algorithms are based on how people work together and how animals hunt for food [11]. And it is a fascinating field of study that is important for solving complex optimization problems. Since the first meta-heuristic was put forward, much progress has been made, and it has become more diverse [12]. Every day, people keep coming up with new algorithms. To address difficult optimization problems, meta- heuristic techniques are employed. Researchers have used natural and physical processes to develop meta-heuristics that have helped them find near-optimal or optimal solutions to several engineering problems. This paper will look at the most important algorithms and terms for optimizers and MLP terms with a comparison of previous studies [13].

- In 2012, Hamidzadeh and his colleagues developed DDC, a novel classification method that employs a "distance-based decision surface with closest neighbor projection. The DDC kernel type has been enhanced to account for the data's effective nonlinear nature". DDC, unlike certain classification algorithms such as SVM, does not need an optimization step or a traditional learning strategy such as the k-NN algorithm. DDC classifies the unclassified sample by computing the weighted average of distances to all training samples and assigning it to the class with the shortest obtained distance. This rule may be applied to create a decision surface formula. In most cases, DDC outperforms the k-NN and SVM algorithms in terms of accuracy. Also, we did not include the training step in DDC [14].

- In 2014, Mill. et al. dealt with the difficulties associated with standard training techniques, They suggested approach proposes employing "the Biogeography-Based Optimization (BBO) algorithm" to train "MLPs". The study comprises five classifications datasets and six function approximation databases to examine "the effectiveness of BBO in training MLPs. The BBO technique employs the balloon", iris, breast cancer, heart, sigmoid, cosine with one peak, sine with four peaks, sphere, Griewank, and Rosenbrock datasets. BBO should not, however, be utilized to train other "types of NNs" like recurrent, Kohonen, or Radial basis function (RBF) networks [15] .
- In 2014, Lee and Choeh created a method that builds models topredict the usefulness of reviews. The HPNN model, which they propose, uses a back-propagation multilayer perceptron neural network to calculate the amount of review usefulness based on a range of variables such as products, review attributes, and review textual characteristics. The goal of this research is to develop a technology that can determine the most useful reviews for a certain product. The Helpfulness Prediction Model Using A Neural Network approach (HPNN) Concentrate only on these outcomes by investigating what makes online customer reviews appear positive or negative. In response, good feedback can influence how consumers perceive online buying [16] .
- In 2015, for the first time, Seyedali Mirjalili developed a novel approach that employs "the Grey Wolf Optimizer (GWO)" to train "Multi-Layer Perceptron (MLP)". They used eight standard datasets to evaluate their technique, Five classifications samples and three function-approximation dataset are included. They compared their findings to those of other famous evolutionary trainers as "Particle Swarm Optimization (PSO)",

"Genetic Algorithm (GA)", "Ant Colony Optimization (ACO)", "Evolution Strategy (ES)", and "Population-based Incremental Learning (PBIL)" . "The GWO-based trainer was applied to five conventional classification datasets, including XOR, balloon, Iris, breast cancer, and heart, as well as three function-approximation datasets, sigmoid, cosine, and sine". Nevertheless, further study is required to establish how to apply "the GWO algorithm" to identify the appropriate number of [17].

- In 2016, Ramchoun and his colleagues created a system that optimizes the design of Artificial Neural Networks using "the Genetic Algorithm". "This approach has been tried to discover the ideal number of hidden layers and connection weights in the Multilayer Perceptron", as well as the most favorable weights matrix after training. They also provided a novel modeling for the multilayer Perceptron design optimization issue with limitations. The results of testing using Iris data suggest that neural network topologies generalize well. It should be highlighted, however, "that the EBP technique has not been thoroughly evaluated on real-world issues such as diabetes, thyroid disease, and cancer from other databases" [18].
- In 2016, Aljarah and his team have introduced a novel training algorithm that utilizes the whale optimization algorithm (WOA), which has demonstrated its effectiveness in solving various optimization problems and surpassing existing algorithms. The team was inspired to evaluate its performance in training feedforward neural networks due to WOA's ability to avoid high local optima and achieve rapid convergence. "They approached the problem of training MLPs as a minimization problem". It is suggested that other kinds of ANNs be trained with WOA. It is worth thinking about how WOA-trained MLP can be used to solve classification

problems in engineering. Using the WOA-trained MLP to solve function approximation datasets can also be a useful contribution [19].

- In 2019, Hesami et al. used a complex computer software called MLP-NSGAII to identify the optimum approach to sterilize chrysanthemum in a lab. They examined seven distinct variables, including HgCl2 and Nano-silver, and used MLP to create models for how frequently the chrysanthemum was polluted and how well it survived (EV). Next they utilized NSGAII to determine the optimum method of sterilizing and determined which of the seven factors was most relevant. The results were excellent, with R2 values above 94%. Using technique GA, they discovered that employing 1.62% NaOCl for nearly 14 minutes was the optimum approach to achieve flawless EV and no contamination (MLP-NSGAII). Nevertheless, they did not compare various methods of performing this type of analysis [20] .
- In 2019, Using the grasshopper optimization technique, a team of researchers lead by Heidari Hesami created a novel training algorithm for multilayer perceptron neural networks (GOA). Because of its adaptive and flexible search methods, which allow it to avoid local optima and balance exploration and exploitation, the GOA algorithm is a potential tool for tackling optimization issues. The suggested GOAMLP model was validated using five large datasets from "individuals with breast cancer, Parkinson's disease, diabetes, coronary heart disease", and orthopedics. This strategy is also applicable to various types of neural networks and huge datasets [21].
- In 2020, Samadianfard and colleagues created a method that uses AI algorithms to discover the best weights in neural network layers. This is critical for properly extracting relevant characteristics from input data, which is required for developing exact models. The study's goal was to

forecast wind speed at 10 different sites using surrounding reference points as input data. They forecasted daily wind speed estimates for each target station using "MLP, MLP-WOA, and MLP-GA models", as well as Whale "Optimization and genetic algorithms". The approach made no use of any other climatic or atmospheric data. They used the IRIMO approach to examine the performance of MLP-WOA utilizing several statistical metrics. Nevertheless, no information on the data utilized in this study is supplied [22] .

- In 2020, Al-Badarneh and his associates dealt with uneven categorization, a novel training technique based on three evolutionary algorithms was created. "GWO-MLP, PSO-MLP, and SSA-MLP" models were utilized, each with its own fitness function such as accuracy, f1-score, and g-mean. The suggested methods were tested on ten unbalanced datasets, with "the average results of 30" separate runs, best results, and standard deviations determined for each statistic. The data showed that there is no evident advantage of one method over another. For dealing with unbalanced datasets, the g-mean and f-score fitness functions were found to give a competitive advantage over classification accuracy rate. The researchers propose using "neuro-evolutionary models with g-mean fitness functions" to improve memory of both primary concepts [23] .
- In 2020, a study conducted by Ece.et al. proposed a system that utilized single models and hybrid models to predict outcomes in thirteen categories. While the single models had the advantage of lower processing time, their lower accuracy was a significant limitation compared to the hybrid models, which were found to have higher accuracy and lower processing time, particularly in the case of "MLP-PSO over MLP-GA". The research outcomes were evaluated using RMSE and correlation coefficient

values, with Tanh (x) as the output function providing significantly improved accuracy for the developed models. In conclusion, while single MLPs have lower processing time, their lower accuracy is a disadvantage compared to hybrid models such as MLP-PSO and MLP-GA with population sizes of 125 and 50 respectively, which provided higher accuracy in testing steps. It is important to note that this conclusion does not address the fluctuations or beatings of the stock market [24].

Table (1.1): summary of Most Important Studies

RF.	year	Methods	Dataset	Optimization	limitation
[14]	2012	Distance-based decision classifier (DDC)	UCI	high dimensional space and hidden layer	not To apply other distances or new distances based on how the data is spread out in order to make DDC work better. For sample data to have a good projection line, it might be in the shape of a sphere. Using techniques like kernel and my be. One option for getting decision surfaces is to reduce the amount of data.
[15]	2014	Biogeography-Based Optimization (BBO)	benchmark datasets • Balloon • Iris • breast cancer • heart • sigmoid • sphere • Griewank • Rosenbrock • Sigmoid • cosine • Sine	w,b parameter	not training BBO in other types of NNs, like recurrence, Kohonen, or Radial basis function (RBF) networks.

[16]	2014	Helpfulness Prediction Model Using A Neural Network (HPNN)	Manual collected 28,699 cases	Null	Focus only on these results by looking into what makes online customer reviews seem good or bad. In turn, helpful feedback can change how people feel about online shopping.
[17]	2015	GWO	• UCI • XOR • Balloon • Iris • breast cancer • Heart • Sigmoid • cosine • sine .	define the optimal values for weights and biases	GWO algorithm has not been used to find the best number of hidden nodes, layers, and other MLP structural parameters. There should be more research into how to fine-tune this algorithm.
[18]	2016	EBP	Iris data	number of hidden layers and the number of neurons in each hidden layer and proces	Not much training on real problems like diabetes, thyroid, and cancer from other databases.
[19]	2016	WOA	• UCI • Blood • Breast cancer Diabetes • Hepatitis Vertebral • Diagnosis(I) • Diagnosis (II) • Parkinson • Liver • Australian • Credit Monk • Tic-tac-toe • Titanic • Ring • Twonorm • Ionosphere • Chess • Seed • Wine	best set of main controlling parameters (weights and biases)	It is suggested that other kinds of ANNs be trained with WOA. It is worth thinking about how WOA-trained MLP can be used to solve classification problems in engineering. Using the WOA-trained MLP to solve function approximation datasets can also be a useful contribution.

[20]	2019	GA (MLP-NSGAII). MLP (MLP-NSGAII). MLP	Collected data manually	values for significant model's parameters were tested based on a trial and error analysis	not focus on comparing and evaluating different multi-objective optimization algorithms in different areas of plant science, especially in plant tissue culture areas.
[21]	2019	GOA	• Breast cancer • Parkinsons • Diabetes • Saheart • Vertebral	w,b,parameter	GOA can be used with other kinds of NNs and big data sets.
[22]	2020	WOA	IRIMO	adding additional configuration in internal nodes. for controlling the mentioned complexity, eight neurons in the hidden layer were selected.	The data used is not described
[23]	2020	• (GWO) • (PSO) • (SSA)	• Iris0 • Glass1 • Vehicle2 • Segment0 • Yeast3 • Ecoli3 • Page-blocks0 • Yeast-02579-s68 • Vowel0 • Ecoli-0146s	the accuracy of ANN by updating the weights and bias values using GA and PSO estimate the weights of hidden and output layers of	Not looking into how the proposed meta-heuristic approach for optimizing Convolutional Neural Networks (CNN) could be used for complex tasks like classifying images and texts. Declaration.
[24]	2020	• MLP–GA • MLP-PSO	Matriks Information Delivery Services Inc.	Optimizing a set of weights and biases	Not address the stock market's beatings, that the stock market's fluctuations will be successfully addressed. So, the return difference is a problem with the current study.

1.5 Thesis overview

The rest section of this dissertation is organized as following:

Chapter Two: Defines essential terms, and provides general review of Metaheuristic search techniques Classical Mechanics Based Meta- Heuristics, Sine Cosine Algorithm (SCA), WAO Whale Optimization Algorithm, 2 GWO Grey Wolf Optimizer, GTO Artificial Gorilla Troops Optimizer, AVOA "African Vultures Optimization Algorithm ,and optimization , Multilayer perceptron ,Advantage Of MLP, Steps to train the MLP, Classification, Evaluation Measurements.

Chapter Three: This chapter provides the research methodology utilized in this thesis, and thesis architecture, Initial Population, Fitness Function, details of the Proposed GTO_MLP Approach, The Proposed AVOA_MLP Approach.

Chapter Four: This chapter presents the experiment result and then the decision of hybridization strategies the Proposed GTO_MLP Approach, The Proposed AVOA_MLP Approach. Dataset and parameter, the results of the proposed method, Evaluation criteria, statistical analysis, and development of the Proposed GTO_MLP Approach with MSE, The Proposed AVOA_MLP Approach with MSE, and The experimental setup Function-approximation datasets.

Chapter Five: In conclusion, our research has yielded several key findings and recommendations for future work. To summarize, we have identified important conclusions and provided suggestions for future endeavors.

الخلاصة

يمكن أن يكون تحسين شبكة الإدراك متعدد الطبقات (MLP) باستخدام خوارزمية الكشف عن مجريات الأمور أمرًا صعبًا للغاية ، ولكن لديه القدرة على تحسين الأداء العام للشبكة بشكل كبير. تشبه خوارزميات الاستدلال الفوقي مساعدين مفيدين يمكن أن يساعدوا في اكتشاف أفضل الإعدادات للجوانب المختلفة للشبكة العصبية ، مثل مدى الأهمية التي يجب إعطاؤها لعوامل معينة (الأوزان والتحيزات) ، ومدى سرعة التعلم ، وعدد الخلايا العصبية التي يجب أن يكون في كل طبقة ، وعدد الطبقات التي يجب إخفاؤها.

يتطلب ضبط شبكة MLP باستخدام خوارزمية metauristic دراسة دقيقة وتجريبًا للعثور على مجموعة مثالية من المعلمات التي يمكنها تحسين أدائها. في هذه الرساله ، نقدم نماذج مدرب تكيفية تعتمد على محسن قوات الغوريلا الاصطناعية (GTO) وخوارزمية تحسين النسور الأفريقية (AVOA) لضبط شبكة MLP. يمكن أن تساعد خوارزميات GTO و AVOA في تحسين شبكة MLP من خلال تعديل أوزانها وتحيزاتها ، مما قد يؤدي إلى نتائج أكثر دقة وأداء أفضل. خمس مجموعات بيانات قياسية للتصنيف (سرطان الثدي ، XOR ، القلب ، البالون ، القزحية) تم قياس كفاءة التقنيات المقترحة من خلال مقارنتها بالمعايير المعمول بها.

يتم استخدام Optimizers و GTO و AVOA لأول مرة كمدربين متعدد الطبقات ، علاوة على ذلك ، تم تقييم فعالية النهج المقترح من خلال مقارنته بتقنيتين تحسين شائعة الاستخدام: خوارزمية محسن الحوت (WOA) ، وخوارزمية محسن جيب التمام (SCA). أظهرت النتائج التجريبية أن كلاً من خوارزميات GTO_MLP و AVOA_MLP تسمح لشبكة "MLP" لاستكشاف مناطق مختلفة من مساحة البحث والتقارب نحو حل أمثل يؤدي إلى حل المشكلة المحلية المثلى ومعدل تصنيف بين 90٪ إلى 100٪ تحقيق معدل تصنيف مرتفع.