Feature Selection in MicroArray Datasets using Manta Ray Foraging Optimization Technique

Devendra K. Tayal[1] , Ramya Neelakantan[2] , Khushi Bhargava[3]

{dev_tayal2001@yahoo.com[1] , ramya014btcse19@igdtuw.ac.in[2] , khushi023btcse19@igdtuw.ac.in[3]}

Indira Gandhi Delhi Technical University for Women (IGDTUW)[1,2,3]

Abstract. In this document, we apply the Manta Ray Foraging Optimization technique for Feature Selection in the genetic dataset to identify the relevant features from a high dimensional feature space. MRFO is one of the recent meta-heuristic optimization techniques which has been used for Feature Selection. presence of cancerous tumors in the patient. [10]. This document analyzes the pre existing algorithms developed in the space of Feature Selection using the MRFO algorithm in the GEM dataset. The MRFO method uses three unique search techniques in three different ways. The local search ability of the algorithm is highly aided by the chain foraging behavior; the global search ability is determined by the manta ray cyclone foraging behavior; and the local search ability and convergence rate are greatly improved by the somersault foraging behavior. This document, two separate learning classifiers were used: K-NearestNeighbors (KNN) classifier and the Random Forest Classifier were used for calculating the fitness

Keywords: feature selection (FS), nature inspired algorithms (NIA), microarray dataset, genetic data, manta ray foraging optimization (MRFO).

1 Introduction

Feature Selection (FS) is a data pre-processing approach that is used to limit the number of input features that are utilized to train a machine learning model. The primary goal of Feature Selection is to lower the size of the feature set, while retaining the model’s performance and accuracy. This is accomplished by picking a subset of relevant features from the collection of features. GEM data is a dataset that simultaneously contains the molecular expression of thousands of genes. This massive amount of genetic data is fed into a Machine Learning system for analysis and thorough characterization of the categorization task. However, much of the genetic material is redundant and plays little to no part in the classification process. It degrades the prediction process by lowering accuracy and increasing the complexity and time necessary to analyze such massive amounts of data. As a result, feature selection becomes an unavoidable requirement when dealing with GEM datasets- which are high-dimensional and contain several redundant features. Data pre-processing refers to the manipulations performed on the data, either manually or automatically - using code, - to prepare the data for learning algorithms and to ensure a performance enhancement. It involves taking the raw data and converts it into a format that is understandable by the machines and learning algorithms. Data

ICASDMBW 2022, December 16-17, Delhi, India

DOI 10.4108/eai.16-12-2022.2326167

preprocessing is an important and necessary step before application before any learning algorithm is applied. This is because, if poor quality data is used to train the model, then the model also produces poor quality results, which are irrelevant - or harmful in some cases - for analysis. The primary steps involved in preprocessing of numeric data are:

1. Data quality assessment, including analysis of outliers in the data

2. Data cleaning, either manual or automatic

3. Data Transformation, including normalization of the data

4. Data Reduction, involving Feature Selection.

2 Literature review

Many bio and nature-inspired optimization techniques have been proposed over the past ten years. These algorithms are notably different from the natural or biological inspirations they drew from, making it simple and clear to identify each algorithm's category. The search behaviors of the algorithms, or updation of candidate solutions during the iterative phase, are also varied. Some recently proposed NIA are the whale optimization algorithm (WOA) [18], squirrel search algorithm (SSA) [14], virus colony search (VCS) [16], bat algorithm (BA) [28], fruit fly optimization algorithm (FOA)[22], butterfly optimization algorithm (BOA) [3], crow search algorithm (CSA) [4], grey wolf optimizer (GWO) [21]. Many scholars have used a variety of techniques to improve the performance of MRFO in the literature. In order to solve the thermal design problem, Turgut [25] combined the 10 chaotic maps with MRFO to create a novel chaotic MRFO that can successfully escape local optima and speed up algorithm convergence. By incorporating the "chaotic numbers generated by the Logistic map" in MRFO, Calasan et al. [6] improved the optimization capacity of the algorithm; this variant was used to estimate the parameters used for transformation. To address the problem of image segmentation, Houssein et al. [12] established an effective MRFO method using hybridizing opposition-based learning and the initialization steps of MRFO. The improved MRFO was proposed by Elaziz et al. [8] based on differential evolution as a method for selecting features to identify COVID-19 patients examining pictures of their chests. To predict and assess the impact of PM10 on public health, Yang et al. [27] introduced an MRFO based hybrid framework using grey prediction theory. The question is whether we need yet another novel optimization strategy because there are so many optimization algorithms in the literature that work fairly effectively. There can't be a single method that can handle all optimization issues, according to the No Free Lunch theorem for optimization [26]. Therefore, it follows that not all FS problems can be solved by the FS algorithms as they are now proposed. This inspired us to study and analyze a novel FS strategy based on the Manta Ray Foraging Optimization (MRFO) meta- heuristic algorithm, which was just recently put forth.The new meta-heuristic method known as "Manta Ray Foraging Optimization" (MRFO) [5] [2] models the natural foraging behavior of manta rays. Chain foraging and somersault foraging behaviors carry out local exploitation, while cyclone foraging activity carries out worldwide exploration. The application of this approach in several technical fields, such as geophysics [5], optimal allocation of energy [23], image processing [13], and use of electric power [7], demonstrates some optimization capabilities.

3 Objective

The primary objective of the work presented in this document is to analyze the viability of the recently proposed MRFO algorithm for FS in the GEM dataset.

4 Research methodology

The primary objective of the work presented in this document is to analyze the viability of the recently proposed MRFO algorithm for FS in the GEM dataset.

4.1 Dataset used

The dataset used for analysis was collected from the genes expression microarray database that is made available to the public and kept in the GEO repository bank. (http://www.ncbi.nlm.nih.gov/gds/). The dataset used for this analysis is GDS4887 [11] [24] from the NCBI website. It contains data from "resected liver from hepatocellular carcinoma (HCC) patients with chronic hepatitis C (CHC), which represents interleukin-28B (IL-28B) SNP rs8099917 TT genotype and TG/GG genotype." 40 samples and 54,675 hepatocellular carcinoma characteristics identified in patients with chronic hepatitis C were included in GDS4887. The expression profiles of 20 carcinoma tumors and 20 non-tumor tissues were compared. The data is divided into rows and columns. The columns contain the genetic data for each individual, while the rows contain the numeric value for each genetic feature – identified by the ID reference and the gene identifier.

4.2 Dataset extraction

The dataset GDS4887 was extracted from the NCBI database as a .soft.gz zipped file. R software suite was used to extract the data from the file using the BiocLite function of the BiocManager package and to convert it into a .soft file. Data from this file was then extracted into a table in R, which was then exported into an Excel .csv file.

4.3 Data preprocessing

Preprocessing steps including data cleaning, scaling and normalization were performed on the dataset before selecting most relevant features using the Manta Ray Foraging Optimization technique. While conversion of the file from .soft to .csv for easy access to the data for python implementation, some of the genetic identifiers (for example MARCH2, MARC2 etc.) were automatically converted to date format. To correct this, the oct4th [1] Python package was used. It automatically converts the csv files into excel .xlsx files free of the gene name issue. Other minor corrections in the data were made manually. Data scaling was performed using the minmaxscaler function from the scikit learn library. It centered the data into a normal distribution with mean 0 and 1 standard deviation. In further steps, feature selection is applied by using the MRFO algorithm.

4.4 Feature selection

The process of creation of a subset of the most relevant features from the complete set of features, While developing a Machine Learning algorithm for high dimension datasets like the GEM dataset, care must be taken to substantially reduce the number of features considered for training, to prevent overfitting or incorrect training of the data. Thus, Feature Selection techniques must be applied to remove irrelevant or correlated features to reduce noise and redundancy in the dataset. In this document, we analyze the use of the Manta Ray Foraging Optimization algorithm, with some parameter tuning, to select the best features from the GEM dataset, in order to maximize the training accuracy using machine learning models.

Importance of Feature Selection Algorithms:

4.4.1 Reduce time required to train the model

Increasing the number of features for model training reduces the model accuracy. Thus, feature selection directly affects the training duration for a model.

4.4.2 Reduce or avoid the Curse of Dimensionality

In datasets like GEM, which have a large number of input features, there also exists a lot of redundancy and noise. It is highly likely that several features have a high correlation between them, that is, one feature directly affects the other, and therefore can be easily predicted by the presence of the other. In such cases, presence of these redundant features negatively affects the model accuracy. This is known as the Curse of Dimensionality. Removal of such redundant features can lead to a faster training time and higher model accuracy.

4.4.3 Reduce overfitting tendency by generalization

If the number of features used to train the model are restricted, the model is able to generalize better on the basis of the training data. This reduces the chances of overfitting of the model on the training data.

4.5 Manta Ray Foraging Optimization (MRFO) algorithm

4.5.1 A brief introduction

Manta Rays are a type of fish with a life span of 50 years. [29] They are typically 29 feet (8.8m) long, weigh about 1350 kgs and primarily eat plankton. There are 2 main species of Manta Rays: The Reef Manta Ray (which are usually 5.5m long and are found in the Indian Ocean), and the Giant Manta Ray (They are about 7m long and are found in high temperature oceans). The MRFO algorithm is a meta-heuristic, nature - inspired evolutionary algorithm first introduced by Zhao et. al. It is inspired by the intelligent behavior of Manta Rays for foraging. The algorithm is useful in solving high dimensional optimization problems.

4.5.2 Types of foraging techniques

4.5.2.1 Chain foraging

In the chain foraging technique, the Manta Ray observes the plankton (foodsource) and moves towards it by lining up in a chain-like structure. The Manta Rays form an organized line by lining up one behind the other when 50 or more of them begin foraging. Smaller male manta rays piggyback and swim on top of female ones. As a result, manta rays coming behind them will snag any plankton that was missed by the ones before them. The formula for chain foraging is seen in **Figure 1. [1]**

$$x_i^d(t+1) = \begin{cases} x_i^d(t) + r \cdot (x_{best}^d(t) - x_i^d(t)) + \alpha \cdot (x_{best}^d(t) - x_i^d(t)) & i = 1 \\ x_i^d(t) + r \cdot (x_{i-1}^d(t) - x_i^d(t)) + \alpha \cdot (x_{best}^d(t) - x_i^d(t)) & i = 2, \dots, N \end{cases} \tag{1}$$

$$\alpha = 2 \cdot r \cdot \sqrt{|log(r)|} \tag{2}$$

Fig. 1. Mathematical formula for chain foraging

4.5.2.1 Cyclone foraging

Several Manta Rays congregate in areas of high plankton concentration. They form a spiraling vortex in the cyclone's eye as their tail ends spiral together with their heads. This causes the filtered water to rise to the surface. As a result, the plankton are drawn into their wide jaws. The formula for cyclone foraging is seen in **Figure 2. [2]**

$$\begin{cases} X_i(t+1) = X_{best} + r \cdot (X_{i-1}(t) - X_i(t)) + e^{bw} \cdot \cos(2\pi w) \cdot (X_{best} - X_i(t)) \\ Y_i(t+1) = Y_{best} + r \cdot (Y_{i-1}(t) - Y_i(t)) + e^{bw} \cdot \sin(2\pi w) \cdot (Y_{best} - Y_i(t)) \end{cases} \tag{3}$$

where w is a random number in [0, 1].

This motion behavior may be extended to a n-D space. For simplicity, this mathematical model of cyclone foraging can be defined as

$$x_i^d(t+1) = \begin{cases} x_{best}^d + r \cdot (x_{best}^d(t) - x_i^d(t)) + \beta \cdot (x_{best}^d(t) - x_i^d(t)) & i = 1 \\ x_{best}^d + r \cdot (x_{i-1}^d(t) - x_i^d(t)) + \beta \cdot (x_{best}^d(t) - x_i^d(t)) & i = 2, \dots, N \end{cases} \tag{4}$$

$$\beta = 2e^{r_1 \frac{T-t+1}{T}} \cdot \sin(2\pi r_1) \tag{5}$$

Fig. 2. Mathematical formula for cyclone foraging

4.5.2.1 Somersault foraging

Manta rays perform a series of backward rolls when they locate a feeding source. They flip around and circle the plankton to attract it towards themselves. A somersault is thus a circular, random, movement which is frequent and localized. This enhances its ability to consume food. The formula for cyclone foraging is seen in **Figure 3. [1]**

$$x_i^d(t+1) = x_i^d(t) + S \cdot (r_2 \cdot x_{best}^d - r_3 \cdot x_i^d(t)), i = 1, \ldots, N$$

Fig. 3. Mathematical formula for somersault foraging

4.5.2 Pseudocode

Initialize the size of population N, the maximal number of iterations T and each manta ray $\boldsymbol{x}_i(t)$=$\boldsymbol{x}_l$+$rand$·($\boldsymbol{x}_u$-$\boldsymbol{x}_l$) for i=1,…,N and t=1. Compute the fitness of each individual f_i =f($\boldsymbol{x}_i$) and obtain the best solution found so far $\boldsymbol{x}_{best}$, Where $\boldsymbol{x}_u$ and $\boldsymbol{x}_l$ are the upper and lower boundaries of problem space, respectively .

WHILE stop criterion is not satisfied do

FOR i=1 TO N DO

IF $rand$ < 0.5 THEN //*Cyclone foraging*

IF t/ T_{max} < $rand$ THEN

$$\boldsymbol{x}_{rand} = \boldsymbol{x}_l + rand \cdot (\boldsymbol{x}_u - \boldsymbol{x}_l)$$

$$\boldsymbol{x}_i(t+1) = \begin{cases} \boldsymbol{x}_{rand} + r \cdot (\boldsymbol{x}_{rand} - \boldsymbol{x}_i(t)) + \beta \cdot (\boldsymbol{x}_{rand} - \boldsymbol{x}_i(t)) & i = 1 \\ \boldsymbol{x}_{rand} + r \cdot (\boldsymbol{x}_{i-1}(t) - \boldsymbol{x}_i(t)) + \beta \cdot (\boldsymbol{x}_{rand} - \boldsymbol{x}_i(t)) & i = 2, \cdots, N \end{cases}$$

ELSE

$$\boldsymbol{x}_i(t+1) = \begin{cases} \boldsymbol{x}_{best} + r \cdot (\boldsymbol{x}_{best} - \boldsymbol{x}_i(t)) + \beta \cdot (\boldsymbol{x}_{best} - \boldsymbol{x}_i(t)) & i = 1 \\ \boldsymbol{x}_{best} + r \cdot (\boldsymbol{x}_{i-1}(t) - \boldsymbol{x}_i(t)) + \beta \cdot (\boldsymbol{x}_{best} - \boldsymbol{x}_i(t)) & i = 2, \cdots, N \end{cases}$$

END IF.

ELSE //*Chain foraging*

$$\boldsymbol{x}_i(t+1) = \begin{cases} \boldsymbol{x}_i(t) + r \cdot (\boldsymbol{x}_{best} - \boldsymbol{x}_i(t)) + \alpha \cdot (\boldsymbol{x}_{best} - \boldsymbol{x}_i(t)) & i = 1 \\ \boldsymbol{x}_i(t) + r \cdot (\boldsymbol{x}_{i-1}(t) - \boldsymbol{x}_i(t)) + \alpha \cdot (\boldsymbol{x}_{best} - \boldsymbol{x}_i(t)) & i = 2, \cdots, N \end{cases}$$

END IF.

Compute the fitness of each individual $f(\boldsymbol{x}_i(t+1))$. IF $f(\boldsymbol{x}_i(t+1)) < f(\boldsymbol{x}_{best})$

THEN $\boldsymbol{x}_{best} = \boldsymbol{x}_i(t+1)$

//*Somersault foraging*

FOR i=1 TO N DO

$$\boldsymbol{x}_i(t+1) = \boldsymbol{x}_i(t) + S \cdot (r_2 \cdot \boldsymbol{x}_{best} - r_3 \cdot \boldsymbol{x}_i(t))$$

Compute the fitness of each individual $f(\boldsymbol{x}_i(t+1))$. IF $f(\boldsymbol{x}_i(t+1)) < f(\boldsymbol{x}_{best})$

THEN $\boldsymbol{x}_{best} = \boldsymbol{x}_i(t+1)$

END FOR.

END WHILE.

Return the best solution found so far $\boldsymbol{x}_{best}$.

Fig. 4. Pseudocode for MRFO algorithm

4.6 MRFO algorithm implementation for feature selection [9]

A discrete binary search space is an intrinsic property of a FS problem [20]. Therefore, to describe solutions in feature selection problems, we need a vector of 0s and 1s, where 0 denotes the feature that is not selected and 1 denotes the corresponding feature that is selected. The feature dimension of the dataset under consideration corresponds to the length of vector. A Manta Ray is therefore represented as a binary vector. By changing the variables in problems with continuous search space to binary variables, the problems can be transformed into binary problems. Position vectors are used to denote the position of each individual in the continuous search space. Update of positions in a discrete search space involves flipping between 0 and 1. As a result, a transfer function is used to the real manta ray position and transformed it to binary values [15]. The likelihood of changing the values of a binary solution is defined by the transfer function. Manta rays are forced to travel in a binary space via transfer functions. SShaped and V-Shaped transfer functions are two distinct types [19].

5 Analysis

5.1 Using MRFO for feature selection

The MRFO algorithm was applied to the dataset after preprocessing the data in order to select the best feature subset from the set of 54675 features. The primary component of FS techniques is the evaluation of the optimal feature subset. This makes the selection of the fitness function a crucial task. Since MRFO is a wrapper-based FS technique, a machine learning algorithm (classifier) is used to evaluate the feature subset. In this document, 2 separate learning classifiers were used: K-NearestNeighbors (KNN) classifier and the Random Forest Classifier were used as the fitness function. Both classifier methods were first used on the complete set of features to calculate the initial accuracy. The MRFO algorithm was then applied for Feature Selection. The algorithm generated an array of features, consisting of zeros and ones, with a total width equal to the number of features in the dataset - a binary array. 1 means that the feature is a part of the feature subset and 0 otherwise. This selection of optimal features was then passed to the models once again to calculate the new accuracy based on the reduced subset of features.

5.2 Comparison of accuracies of the models

The Random Forest Classifier and the KNearest Neighbors classifier were used as fitness functions in the MRFO algorithm to obtain the optimal feature subset. The following results were observed

Table 1. Comparison of accuracy before and after MRFO using random forest and KNN as fitness functions.

Algorithm used	Iteration Number	Initial Accuracy	Number of features selected	Accuracy after feature selection
K - Neighbors Classifier	1	50	27	100
	2	62.5	10	100
	3	87.5	14	100
Random Forest	1	62.5	5	87.5
	2	87.5	15	100
	3	87.5	9	100

6 Results / findings

Experimental work that has already been done on challenging benchmark problems has shown that MRFO is very reliable and effective for FS problems. Using KNearestNeighbors and Random Forest algorithm in the fitness function gave an accuracy of 100 percent for almost all iterations. On average. KNN selects 0.03 percent of features, while Random Forest selects 0.01 percent of the total available features. However, one of the major disadvantages of the algorithm used is that MRFO suffers from weak exploration capabilities, which leads to premature convergence to the local optimum. The imbalance between exploration and exploitation for hybrid problems also poses an issue.

7 Implications of Research

Manta Ray Foraging Optimization(MRFO) provides strong global optimization capabilities for both confined and unconstrained situations [17]. The algorithm is ideal for dealing with practical issues which call for precise final solutions with minimal computing effort. The binary version of MRFO is used for feature selection in the dataset presented. [9] However, 2 open problems were found. One is to devise an efficient tactic to adapt the steps of the movement of each individual based on their neighbor's position. Another is to employ a way to share the information amongst the individuals to increase efficiency in the search process. Once these issues are resolved, the features selected by the algorithm can be used to improve training time and accuracy of ML models for prediction of the presence of cancerous tumors in the patient. [10].

8 Conclusion

This document analyzes the pre existing algorithms developed in the space of Feature Selection using the MRFO algorithm in the GEM dataset. This strategy is highly recommended, owing to its simplicity and limited parameters to be configured. The challenges faced by the MRFO algorithm mean more opportunities for further research, which may be leveraged to improve its performance in the FS process.

9 Funding

This work is funded under the Data Science Research of Interdisciplinary Cyber-Physical Systems (ICPS) Programme of the Department of Science and Technology (DST) [Sanction Number T-54], New Delhi, Government of India, India

References

[1] Robert Aboukhalil. URL: https://oct4th.sandbox.bio/.

[2] Hesham Alhumade et al. “Optimal parameter estimation methodology of solid oxide fuel cell using modern optimization”. In: Mathematics 9.9 (2021), p. 1066.

[3] Sankalap Arora and Satvir Singh. “Butterfly optimization algorithm: a novel approach for global optimization”. In: Soft Computing 23.3 (2019), pp. 715–734.

[4] Alireza Askarzadeh. “A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm”. In: Computers & structures 169 (2016), pp. 1– 12.

[5] Ubong C Ben et al. “Novel methodology for interpretation of magnetic anomalies due to twodimensional dipping dikes using the Manta Ray Foraging Optimization”. In: Journal of Applied Geophysics 192 (2021), p. 104405.

[6] Martin P Ćalasan et al. “Notes on parameter estimation for single-phase transformers”. In: IEEE Transactions on Industry Applications 56.4 (2020), pp. 3710–3718.

[7] Serhat Duman, Adem Dalcalı, and Harun Özbay. “Manta ray foraging optimization algorithm–based feedforward neural network for electric energy consumption forecasting”. In: International Transactions on Electrical Energy Systems 31.9 (2021), e12999.

[8] Mohamed Abd Elaziz et al. “New machine learning method for image-based diagnosis of COVID-19”. In: Plos one 15.6 (2020), e0235187.

[9] Kushal Kanti Ghosh et al. “S-shaped versus V-shaped transfer functions for binary Manta ray foraging optimization in feature selection problem”. In: Neural Computing and Applications 33.17 (2021), pp. 11027–11041.

[10] Ibrahim Hayatu Hassan et al. “An improved binary manta ray foraging optimization algorithm based feature selection and random forest classifier for network intrusion detection”. In: Intelligent Systems with Applications 16 (2022), p. 200114.

[11] Yuji Hodo et al. “Association of Interleukin-28B Genotype and Hepatocellular Carcinoma Recurrence in Patients with Chronic Hepatitis CIL-28B Genotype Is Associated with HCC Recurrence with CHC”. In: Clinical cancer research 19.7 (2013), pp. 1827–1837.

[12] Essam H Houssein, Marwa M Emam, and Abdelmgeid A Ali. “Improved manta ray foraging optimization for multi-level thresholding using COVID-19 CT images”. In: Neural Computing and Applications 33.24 (2021), pp. 16899–16919.

[13] Essam H Houssein et al. “An efficient ECG arrhythmia classification method based on Manta ray foraging optimization”. In: Expert Systems with Applications 181 (2021), p. 115131.

[14] Mohit Jain, Vijander Singh, and Asha Rani. “A novel nature-inspired algorithm for optimization: Squirrel search algorithm”. In: Swarm and evolutionary computation 44 (2019), pp. 148–175.

[15] Sangwook Lee et al. “Modified binary particle swarm optimization”. In: Progress in Natural Science 18.9 (2008), pp. 1161–1166.

[16] Mu Dong Li et al. “A novel nature-inspired algorithm for optimization: Virus colony search”. In: Advances in engineering software 92 (2016), pp. 65–88.

[17] Yingying Liao, Weiguo Zhao, and Liying Wang. “Improved manta ray foraging optimization for parameters identification of magnetorheological dampers”. In: Mathematics 9.18 (2021), p. 2230.

[18] Majdi Mafarja and Seyedali Mirjalili. “Whale optimization approaches for wrapper feature selection”. In: Applied Soft Computing 62 (2018), pp. 441–453.

[19] Majdi Mafarja et al. “S-shaped vs. V-shaped transfer functions for ant lion optimization algorithm in feature selection problem”. In: Proceedings of the international conference on future networks and distributed systems. 2017, pp. 1–7.

[20] Seyedali Mirjalili and Andrew Lewis. “S-shaped versus V-shaped transfer functions for binary particle swarm optimization”. In: Swarm and Evolutionary Computation 9 (2013), pp. 1– 14.

[21] Seyedali Mirjalili, Seyed Mohammad Mirjalili, and Andrew Lewis. “Grey wolf optimizer”. In: Advances in engineering software 69 (2014), pp. 46–61.

[22] Wen-Tsao Pan. “A new fruit fly optimization algorithm: taking the financial distress model as an example”. In: Knowledge-Based Systems 26 (2012), pp. 69–74.

[23] Abdullah M Shaheen et al. “Economic power and heat dispatch in cogeneration energy systems using manta ray foraging optimizer”. In: IEEE Access 8 (2020), pp. 208281–208295.

[24] Anggrainy Togi Marito Siregar et al. “Comparison of supervised models in hepatocellular carcinoma tumor classification based on expression data using principal component analysis (PCA)”. In: AIP Conference Proceedings. Vol. 2264. 1. AIP Publishing LLC. 2020, p. 030002.

[25] Oguz Emrah Turgut. “A novel chaotic manta-ray foraging optimization algorithm for thermo-economic design optimization of an air-fin cooler”. In: SN Applied Sciences 3.1 (2021), pp. 1–36.

[26] David H Wolpert, William G Macready, et al. No free lunch theorems for search. Tech. rep. Technical Report SFI-TR-95-02-010, Santa Fe Institute, 1995.

[27] Wendong Yang et al. “A novel framework for forecasting, evaluation and early-warning for the influence of PM10 on public health”. In: Atmosphere 12.8 (2021), p. 1020.

[28] Xin-She Yang and Amir Hossein Gandomi. “Bat algorithm: a novel approach for global engineering optimization”. In: Engineering computations (2012).

[29] Weiguo Zhao, Zhenxing Zhang, and Liying Wang. “Manta ray foraging optimization: An effective bio-inspired optimizer for engineering applications”. In: Engineering Applications of Artificial Intelligence 87 (2020), p. 103300.