Application of hybrid algorithm of bionic heuristic and machine learning in nonlinear sequence

Lipeng Qian, Changzheng Liu*, Jiaxin Yi, Shannan Liu

Departments of information science and technology, University of Shihezi, 659001, China
20182008105@shzu.edu.cn
*Corresponding author’s e-mail: liucz@sina.cn

Abstract: Aiming at the problems existing in the application of machine learning algorithms and the advantages of bionic heuristics with fast convergence and strong local optimization capabilities. In this paper, the typical bionic heuristic algorithm is used to construct the optimization system of machine learning models such as support vector machine (SVM), feedforward neural network (FFNN) and generalized regression neural network (GRNN). The hybrid algorithms are used to predict the discharges of the complex river flow series for 24 hours in the future, and the prediction results are quantitatively evaluated by using statistical indicators. The results show that, compared with the single algorithm, the hybrid algorithms have better global solution searching capability, higher prediction accuracy and fewer iterations. The response of practical problems to time scale should be fully considered in application as a result of longer training time and more parameters of hybrid algorithm. Hybrid algorithm can be used as an effective supplement to solve complex nonlinear problems.

1. Introduction

In practical production, many engineering problems need to solve the extreme value and maximum value in space dimension, such as robot path selection planning, logistics planning, circuit wiring, UAV flight route planning. Good planning methods can greatly reduce the cost of production. These problems are all NP-C problems, which have many feasible solutions. The methods to solve these problems include precise algorithm[1, 2], bionic heuristic algorithm[3-5], machine learning method[6, 7]. Both bionic heuristic algorithm and machine learning algorithm are inspired by biological behavior in nature, and the mathematical models are extracted. At present, the heuristic probability search mode is widely used to get rid of the local optimal situation[8, 9].

Various heuristic bionic algorithms have been applied in many fields of production practice, and the mixing of different algorithms can achieve better results[10, 11]. Many scholars have made contributions in the application of heuristic algorithms. On the optimization of neural network parameter structure, Zhang et al. [12] proposed a new memetic algorithm to optimize the parameters of extreme learning machine (ELM), and embedded the local search strategy into the global optimization framework to obtain the optimal network parameters. Support vector machine is a widely used machine learning algorithm in recent years. In order to construct an accurate SVM classifier, kernel function, kernel parameters and soft boundary constant (also known as regularization parameters) must be determined. For the selection of parameters, the application of intelligent heuristic algorithm provides some ideas. There are a lot of optimization schemes for SVM related parameters.
On the optimization of machine learning features, proposed an enhanced KNN rule based on evolutionary computation, unifying two classical weighting paradigms, that is, adjusting the contribution of each neighbor and the importance of data features at the same time to improve identify new instances well. In machine learning intensive simplification, prototype optimization, kernel function learning and other issues, the use of bionic intelligent optimization algorithms for optimization has achieved good results. The above shows that using different heuristic bionic algorithms to optimize machine learning methods may achieve better convergence results than independent algorithms. Although many scholars have done a lot of research on bionic intelligent optimization algorithms and machine learning algorithms, various algorithms have their own shortcomings, due to the lack of strict mathematical reasoning. Therefore, it is necessary to complement the advantages of multiple algorithms. Exploring the combination of typical machine learning algorithms and bionic intelligent optimization algorithms has certain theoretical significance for solving practical problems.

This paper explores the solution capabilities of a mixture of typical machine learning algorithms and heuristic bionic optimization algorithms, and applies them to the continuous prediction of river runoff in the next 24 hours. The statistics indicators are used to evaluate the prediction results of different algorithms. The advantages and disadvantages of the hybrid model are discussed, and finally the development prospect and direction of the hybrid model are given.

2. Material and Methods

2.1. Bionic intelligent optimization algorithm

Bionic heuristic algorithm is a computing technology and method developed by simulating the structural characteristics, evolutionary laws, behavior patterns and thinking structures of human, natural and other biological populations, and is used to solve various optimization problems. Although the bionic objects are different, they all show the characteristics of self-organization and self-adaptation. They have obvious advantages in solving combinatorial optimization problems which are difficult to obtain satisfactory solutions by traditional methods.

Intelligent heuristic algorithms can be divided into two categories: single-solution-based methods and population-based methods. The single-solution-based method means that the search process starts from a candidate solution and continuously improves this single candidate solution in the subsequent iterations. The population-based method uses a set of solutions to perform optimization. The search process starts from a random initial population, and the population is improved in the iterative process. Compared with the single-solution-based method, the population-based method is more exploratory in that individuals help and share the search space information each other to avoid falling into local extremum. Typical algorithms include particle swarm optimization (PSO), genetic algorithm (GA), etc. NFL (no free lunch) theorem logically proves that a single intelligent heuristic algorithm cannot solve all optimization problems. In other words, a specific intelligent heuristic method may show better performance on one set of problems, but may show poor performance on another set of problems. Therefore, the current intelligent heuristic algorithm needs to be continuously improved, and new intelligent heuristic algorithm needed to be proposed.

To adopt intelligent heuristic algorithm, the problem of individual coding should be considered first, which will directly affect the accuracy and efficiency of the algorithm search. Intelligent heuristic algorithms mostly use vector coding, including numerical coding and symbol coding. Current heuristic algorithms can solve related parameter determination and type selection problems, but cannot solve multi-type combination and nesting problems. Heuristic algorithms based on nonlinear structure coding have emerged, such as genetic algorithm (GA), genetic programming (GP). The individual code length is determined by the actual problem dimension or accuracy.

The definition of fitness function is very important for intelligent heuristic algorithm modeling, and directly determines the convergence result. Fitness function is used to evaluate the merits and demerits of individuals in the iterative process, and guide the algorithm to search where to search. It is
usually defined according to the needs of practical problems. To define the fitness function, the constraint conditions and optimization function of the search target should be determined in advance, so as to improve the detection performance while feature selection. If multiple metrics are involved in the application, a weighted fitness function can be used to transform multi-objective optimization into a single-objective optimization problem [21]. The intelligent heuristic algorithm modeling process is as follows.

a. Set the parameters of the algorithm, such as the number of iterations, population size;
b. Initialize the population according to individual coding rules;
c. Find the best individual in the initial population according to the fitness function;
d. Generate a new generation of populations according to the rules of individual variation;
e. Calculate the best individual of the new generation population according to the fitness function;
f. Repeat steps d and e until the iterations is terminated or the optimal fitness value is met;
g. Output the optimal solution;

2.2. Machine learning algorithm

Inspired by human neurons, artificial neural network is designed to study how computers simulate or realize human learning behavior, acquire new knowledge or skills, and reorganize the existing knowledge structure to continuously improve its performance [28]. Machine learning algorithms have achieved good development, among which are widely used feedforward neural networks, support vector machines, generalized regression neural networks, etc.

Feedforward neural network (FFNN) is an intelligent dynamic model based on the human brain nervous system, composed of a large number of parallel processing systems, and has a strong capability to fit and estimate nonlinear functions[29]. The network structure is composed of the number of neurons, the layers (input layer, hidden layer, output layer) and the weight vector connecting each layer. In FFNN training, a suitable training algorithm is very important. Lavenberg-Marquardt algorithm is widely used because of its fast convergence speed[30]. The connection weights and thresholds between the layers are adjusted by the error back propagation algorithm, and the specific calculation of the back propagation algorithm is referred to lecture[31]. The mapping functions on the hidden layer and the output layer use the sigmoid function and the linear function respectively, and the function form is as follows:

\[
f(NET) = \frac{1}{1+\exp(-NET)}
\]

\[
O_k = \sum_{k=1}^{N} \sum_{j=1}^{M_k} (W_{kj}X_j + b_k)
\]

Where \( NET = W_{ij}X + b \) represents the output value of the hidden layer neurons, and \( b \) represents the threshold vector on the output layer.

Support vector machine (SVM) is an intelligent model based on statistical theory. It has been proved that it has high accuracy in regression, classification and time series prediction [32]. SVM establishes independent estimation functions for input vector X and output vector Y. In basic linear regression, linear functions are used to solve optimization problems. The approximation function \( f(x) \) is calculated when it is insensitive to the loss function value \( \varepsilon \) of all variable groups \( (x_i, y_i) \), and the loss function value represents the minimum error of prediction. The approximation of \( f(x) \) can be realized by Lagrange method, the weight vector W is the parameter connecting the training pairs, and the evaluation function to solve the optimization problem is as follows:

\[
f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) < x_i, x > + b
\]

Where \( \alpha_i \) and \( \alpha_i^* \) variable are a set of dual vectors, \( b \) is the offset vector, and \( < x_i, x > \) is the inner product of \( x_i \) and \( x \). In linear regression, this method does not need to calculate the \( W \) exactly. However, many problems are nonlinear models, and it is not appropriate to use linear regression. Therefore, the data is mapped in high dimension and expanded by kernel function, and processed in low dimension to avoid dimension disaster to solve nonlinear problems. Due to the strong simulation capability and flexibility of Gaussian kernel function, it is widely used, the form is as follows:
\[ K(x, x_i) = \exp(-|x - x_i|^2/2\sigma^2) \]

Where \( \sigma \) is the coefficient of the Gaussian kernel function, which has a great influence on the error between the model training result and the observation vector, so the parameters must be selected reasonably.

Generalized Regression Neural Network (GRNN) is a neural network with strong nonlinear mapping capability, high fault tolerance and robustness [33]. The structure has four layers. The first layer is the input layer, the number of neurons is equal to the dimension of the factor combination mode, which plays a role of signal source transmission, and has no weight connection with the mode layer; the second layer is the mode layer, the number of neurons is equal to the number of learning samples; the third layer is the summation layer, using two types of neurons for summation, one type is connected to the model layer neuron with a weight of 1, and the \( i^{th} \) neuron in the other type is connected to the summing layer. The weight of the connection between the \( j^{th} \) summing neuron in the middle is the \( j^{th} \) element in the \( i^{th} \) output sample \( Y_i \); the fourth layer is the output layer, which outputs the results of the coupled runoff forecast.

2.3. Hybrid algorithm

In the construction of machine learning model, for the construction of some models, the selection of individual parameters such as neural network weight and kernel parameters directly affects the accuracy and generalization of the model; feature optimization is an important step in the construction of machine learning model, and the appropriate feature set can not only save system resources but also accurately represent the original data; ensemble learning is an important research in the field of machine learning content, how to select the appropriate machine learning model and how to combine the models effectively are the important problems faced by ensemble learning. Intelligent Heuristic algorithms are used to solve the optimization problem in machine learning because they rely on relatively simple concepts and are easy to implement and do not require gradient information[34]. The intelligent heuristic algorithm assumes the optimization problem as a black box, which makes it highly flexible in solving various problems. The optimization process of intelligent heuristic algorithm is shown in Figure 1.

![Figure 1. The optimization process of intelligent heuristic algorithm.](image)

The machine learning modeling process based on the intelligent heuristic algorithm is shown in Figure 2. The training data based on certain features or samples is input into the core machine learning algorithm. With the help of the intelligent heuristic algorithm, the data model in the training set is analyzed and learned to obtain the trained model and sample features. The optimization process of the model involves multiple training of the model, which is determined by the number of iterations of the intelligent heuristic algorithm. Finally, input the test data into the trained algorithm model to get the final output result.
Models based on machine learning algorithms such as neural networks must be adjusted to achieve ideal results. Since no network model can be applied to all data sets, before applying machine learning methods to new data sets, an appropriate set of hyperparameters must be selected [35]. For the selection of parameters, the application of intelligent heuristic algorithms provides some ideas. In this paper, intelligent optimization algorithm is used to optimize the relevant parameters of machine learning model. The optimization parameters are shown in Table 1.

| Machine Learning Models | Bionic heuristic algorithm |
|-------------------------|---------------------------|
| Number of neurons, Weight and threshold between each layer | SVM, FFNN, GRNN | GA, PSO, ACCA, BA |
| Penalty parameter | SVM, FFNN | GA, PSO |
| Kernel parameter | SVM | GA, PSO, ACCA, BA |

Dimensionality reduction is a key step in building an effective machine learning model. Generally speaking, there are two kinds of dimensionality reduction methods: feature selection and feature extraction [36]. Feature selection can remove irrelevant or redundant features, and achieve the same or better results than the original sample set. Feature extraction attempts to reduce the dimension by combining the original features. This method tries to minimize the information loss, but it usually loses the original features [37]. A large number of experiments have shown that feature optimization methods based on intelligent heuristic algorithms can usually identify different feature subsets, so that the use of feature selection integration to construct higher-quality feature subsets is worthy of further study [38]. Each heuristic algorithm may have its own strengths and weaknesses when dealing with different data sets.

Integration reduction is the intermediate link between sub-model creation and sub-model integration. The main work of integration reduction is to reduce the number of learners to be integrated [39]. The idea of ensemble reduction is similar to the feature selection. Both use a certain method to select a subset of the original set. The biomimetic algorithm assigns a weight to each sub-classifier in the network, which indicates the possibility of including the neural network in the ensemble, and adds sub-models with a weight greater than the preset threshold to the integration.

The kernel method is an effective learning method in the field of machine learning, mainly used to solve the problem of linear inseparability in low-dimensional space. The advantage of using the kernel method is that it can handle the feature space of any dimension without calculating its mapping to the feature space, avoiding complex inner product calculations [40]. The mathematical principle of kernel function has been introduced of SVM model in subsection 2.2. In order to determine the kernel function and related parameters, a large number of kernel methods are usually based on SVM. The objective function is transformed into different optimization problems and solved by different optimization methods. Using intelligent heuristic algorithm to solve the parameters of penalty term and Gaussian kernel function in SVM is a common scheme.
2.4. Analysis of non-stationary series

The time series is measured at distinct intervals \( t_3 < t_2 < \cdots < t_n \), and \( n \) groups of ordered sequences with time \( t \) as the parameter are obtained: \( x_{11}(1), \ldots, x_{1n}(t), x_{21}(1), \ldots, x_{2n}(t), \ldots, x_{n1}(1), \ldots, x_{nn}(t) \). The trend component and transient component of time series exhibited strong non-stationarity and randomness. The non-stationary series \( x_{in}(t) \) can be simplified as \( x(t) = D_1(t) + D_2(t)e(t) \), where \( D_1(t) \) and \( D_2(t) \) are functions determined at time \( t \), representing the mathematical expectation and variance of the change in \( x(t) \) over time, and \( e(t) \) is a stationary stochastic process. Therefore, the analysis of time series is a process based on extracting a deterministic function \( D(t) \) and processing a random function \( e(t) \). According to the internal variation law of time series, the physical causes of the evolution of the time series are deduced and the possible values of runoff in the future are predicted. Because the many time series are complex non-stationary random process, it is very important to fit the nonlinear process of the model in different periods when making prediction fitting. The general form of a time series model is \( Y_n = F(Y_{n-1}, Y_{n-2}, \ldots, \epsilon_n) \), where \( \epsilon_n = e(n) \) is the random disturbance term in the model. If the random disturbance term is a white noise, the model is denoted as

\[
Y_n = a_1Y_{n-1} + a_2Y_{n-2} + \cdots + a_pY_{n-p} + \epsilon_n.
\]

If it is not a white noise, a general autoregressive moving average process model is

\[
Y_n = a_1Y_{n-1} + a_2Y_{n-2} + \cdots + a_pY_{n-p} + \epsilon_n - c_1\epsilon_{n-1} - \cdots - c_q\epsilon_{n-q}.
\]

The main goal is to use a model with strong fitting capability and long short-term memory capability to explore the evolution law of a time sequence and then predict the future value.

The formation of river flow series is affected by many factors, which is a high-dimensional nonlinear system. High precision river flow prediction can provide valuable hydrological information for downstream reservoir discharge and flood prediction. The flow series of Manas River, which is the most representative in the northern Xinjiang, is selected as the research object. The data measured is obtained from the Construction Management Bureau of Kensvat hydropower station. The river flow series studied is from 2017 to 2019, including 13140 hours. After the outliers are eliminated by using domain identification method [41], a total of 13112 measured data are left.

2.5. Parameter calibration

After many experiments with different mixed models, the optimal parameters are obtained. The selection of various parameters in GA, PSO, AACA and BA is consistent with the actual situation of the predicted sequence. During training, all sequences are divided into \( n \) batches for processing to improve computational efficiency. All network connection weights are initialized as random numbers from the normal distribution \( N(0,0.01) \). 70% of the data in each batch of divided data is selected as the training set, and the remaining 30% is used as the verification set. When the training error of the verification set increases, the training will be stopped. In order to avoid over-fitting, a weight attenuation term \( C \) is added to the gradient term of each mixed model. In the training of all models, the training will be stopped when the error of verification begins to rise. In order to avoid the contingency of the results, the average was taken as the final prediction result after all the models were trained for 30 times.

Nash-Sutcliffe efficiency (NSE) was used to evaluate the performance of the model [42]. The range of NSE \((-\infty, 1]\) is close to 1, indicating that the better the performance of the model, the better it can reflect the statistical characteristics of the original sequence. The performance of the prediction results was evaluated by Pearson correlation coefficient (R), percentage bias (BIAS) and root mean square error (RMSE).

3. Results

Under the parameter calibration of subsection 2.5., the hourly flow, rainfall and monthly evapotranspiration from January 2017 to June 2019 are selected to train and calibrate the model, and the hourly flow in the continuous period of July 2019 is used as the test set of the model. The floe
trend in a continuous period of time includes a trend component and a transient component. The trend component reflects the overall change trend of the flow over a period of time, and the transient component can reflect the runoff that changes greatly at certain moments.

Figure 3, Figure 4 and Figure 5 respectively show the prediction results of flow within 24 hours after mixing SVM, FFNN and GRNN with typical heuristic bionic algorithms. From the error curve of validation set of different models (Figure 3-5(a)), the convergence effect of independent model is obviously worse than that of mixed model. Under the same error setting, the independent model has more iterations, and the capability to find the optimal solution is weaker than the hybrid model. From the prediction absolute error curve (Figure 3-5(c)), the overall trend of prediction error at different time nodes is stable and within the acceptable range. The spline diagram of the prediction results (Figure 3-5 (b)) shows that the prediction models are better for the trend component prediction of runoff variation. Although the prediction effect of each model is different in some flood seasons or alternation of dry seasons, the capability to retain the overall data characteristics of the measured series is acceptable. The prediction performance of the transient components is different, and the hybrid model performs significantly better with large changes. One reason may be that the hybrid model can better jump out of the local solution during the fitting, and the coupled model performs better.

Figure 3. Prediction based on support vector machine and bionic heuristic optimization algorithm: (a) error curve of verification set during training; (b) prediction results within 24 consecutive hours; (c) absolute error curve of prediction results.
Figure 4. Prediction based on feedforward neural network and bionic heuristic optimization algorithm: (a) error curve of verification set during training; (b) prediction results within 24 consecutive hours; (c) absolute error curve of prediction results.

Figure 5. Prediction based on generalized regression neural network and bionic heuristic optimization algorithm: (a) error curve of verification set during training; (b) prediction results within 24 consecutive hours; (c) absolute error curve of prediction results.
Table 2, Table 3 and Table 4 show the evaluation of the prediction results of different combination models. The results in Table 2 show that the NSE predicted by SVM and hybrid model for 24 hours is greater than 0.50. According to the literature [42], NSE is greater than or equal to 0.50, indicating that the prediction results of the model are acceptable, but the NSE values of several hybrid models are greater than the single SVM. Pearson correlation coefficient also shows that the combination of SVM and bionic optimization algorithm can achieve better prediction results. Table 3 and Table 4 also show the same results as Table 2. The coupling capability of the hybrid model to time series is stronger than that of the independent model, and the capability to jump out of the local solution is also stronger.

Table 2. Evaluation of prediction results under the hybrid of support vector machine and heuristic bionic optimization algorithm.

| Index | SVM | GASVM | PSOSVM | BASVM | AACASVM |
|-------|-----|-------|--------|-------|---------|
| NSE   | 0.67| 0.86  | 0.85   | 0.73  | 0.83    |
| R     | 0.89| 0.94  | 0.93   | 0.91  | 0.93    |
| BIAS  | 0.22| 0.12  | -0.11  | 0.22  | 0.093   |
| RMSE  | 28.43| 27.58 | 23.76  | 29.2  | 21.41   |

Table 3. Evaluation of prediction results under the hybrid of feedforward neural network and heuristic bionic optimization algorithm.

| Index | FFNN | GAFFNN | PSOFFNN | BAFFNN | AACAFFNN |
|-------|------|--------|---------|--------|-----------|
| NSE   | 0.71 | 0.82   | 0.86    | 0.82   | 0.89      |
| R     | 0.92 | 0.94   | 0.94    | 0.96   | 0.97      |
| BIAS  | 0.33 | 0.11   | -0.15   | -0.11  | -0.079    |
| RMSE  | 26.63| 23.32  | 20.37   | 21.38  | 21.79     |

Table 4. Evaluation of prediction results under the hybrid of generalized regression neural network and heuristic bionic optimization algorithm.

| Index | GRNN | GAGRNN | PSOGARNN | BAGRNN | AACAGRNN |
|-------|------|--------|----------|--------|-----------|
| NSE   | 0.73 | 0.82   | 0.85     | 0.83   | 0.89      |
| R     | 0.9  | 0.95   | 0.97     | 0.97   | 0.98      |
| BIAS  | 0.21 | 0.13   | 0.12     | -0.34  | -0.083    |
| RMSE  | 26.31| 29.34  | 31.07    | 22.32  | 18.25     |

In general, it can be seen from the researched prediction examples that the combination of machine learning models and bionic intelligent optimization algorithms can more easily jump out of the local optimum and converge to the global optimum in the coupling of nonlinear problems. Using heuristic algorithms to explore the initial weights of different neural networks can well improve the coupling performance of the algorithm.

4. Discussion

Different from the independent model, although different models have good performance to couple the characteristics of the original sequence data in the trend component of the sequence, there are significant differences between the hybrid model and the independent model in the transient component. The prediction and evaluation results in Table 2, Table 3 and Table 4 show that the hybrid model has stronger capability to couple nonlinear problems. The main reason is that the bionic intelligent optimization algorithm provides more reasonable initial weight and threshold for FFNN by
using the powerful global search capability, which makes the neural network better avoid falling into local optimum in training. The overall results of the prediction example show that the hybrid model is superior to the independent model in nonlinear problem processing.

However, the training time of the model increases obviously after the hybrid of machine learning and bionic heuristic algorithm, especially when dealing with some complex problems, the hybrid model has great disadvantages. In addition, the parameters that need to be considered and set during training are much more than those of the independent model, which is also the main reason restricting the application of the hybrid model. For example, in the flight route planning of UAV, the response of UAV is greatly reduced due to the long training time. During the attitude adjustment and flight, the response of too long time will have a negative impact on the flight of UAV. Therefore, how to effectively reduce the number of parameter settings of hybrid model? How to further adjust the algorithm to speed up the training speed of the model? How to further improve the search element of global solution? The in-depth study and discussion of these problems will have very important theoretical and practical significance.

Using bionic heuristic algorithm to optimize machine learning method is a relatively new field. In addition, there are many new research fields on bionic heuristic algorithm. The no free lunch theorem [43] points out that there is no optimization method suitable for all problems. Researching new intelligent heuristic algorithms adapted to specific application scenarios is the direction that researchers have been working hard on. Considering the advantages and disadvantages of different algorithms, combining the search strategies of different algorithms to generate a new algorithm is also a hot research topic [44]. The related parameters of Intelligent Heuristic algorithm will affect the performance of the algorithm in varying degrees. How to adjust the parameters to improve the performance of the algorithm is an important task [9]. Although the intelligent heuristic algorithm is very effective in practice, it has not been proved by mathematical analysis in theory. Trying to use some mathematical tools such as Markov chain theory, dynamic system and other methods to prove the effectiveness of the algorithm is the direction of scholars' efforts [38, 45]. There are many optimization problems to be solved in the fields of science, engineering, industry, and commerce. Researching suitable heuristic algorithms in certain application scenarios is a problem that researchers should consider [8, 9].

5. Conclusions

In this paper, the bionic heuristic algorithm is used to optimize the parameter structure, feature optimization, ensemble reduction and kernel function learning of neural network. Effectively mix machine learning algorithms widely used in engineering and typical bionic intelligent optimization algorithms. Taking the complex non-stationary river flow sequence as an example, the coupling and prediction performance of the hybrid model are discussed. In the 24-hour continuous flow prediction, the hybrid model has higher accuracy and better convergence than the independent model. The research example proves the performance of the hybrid model in dealing with nonlinear problems. However, the longer training time and more parameter setting are the main factors restricting its application. Therefore, in the application, the response requirements of the actual problem to the time scale should be fully evaluated to select the appropriate training algorithm.

Acknowledgements
This work was financially supported by Fund Project of XJPCC (2019AB001, 2017CD010 and 2017CA018).

References
[1] Ahmadi, M. H., Hosseinzade, H., Sayyaadi, H., et al. (2013) Application of the multi-objective optimization method for designing a powered Stirling heat engine: Design with maximized power, thermal efficiency and minimized pressure loss. Renewable Energy, 60: 313-322.
[2] Nedic, A., Ozdaglar, A. and Parrilo, P. A. (2010) Constrained Consensus and Optimization in
Multi-Agent Networks. IEEE Transactions on Automatic Control, 55(4): 922-938.

[3] Wang, D., Tan, D. and Liu, L. (2018) Particle swarm optimization algorithm: an overview. Soft Computing, 22(2): 387-408.

[4] Wang, J., Du, P., Niu, T., et al. (2017) A novel hybrid system based on a new proposed algorithm-Multi-Objective Whale Optimization Algorithm for wind speed forecasting. Applied Energy, 208: 344-360.

[5] Zakariazadeh, A., Jadid, S. and Siano, P. (2014) Multi-objective scheduling of electric vehicles in smart distribution system. Energy Conversion and Management, 79: 43-53.

[6] Brunton, S. L., Noack, B. R. and Koumoutsakos, P. (2020) Machine Learning for Fluid Mechanics. Annual Review of Fluid Mechanics, Vol 52, 52: 477-508.

[7] Qi, C., Ly, H.-B., Chen, Q., et al. (2020) Flocculation-dewatering prediction of fine mineral tailings using a hybrid machine learning approach. Chemosphere, 244.

[8] Akhmedova, S., Semenkin, E. and Iop (2017) Investigation into the efficiency of different bionic algorithm combinations for a COBRA meta-heuristic. V International Workshop on Mathematical Models and Their Applications 2016, 173.

[9] Li, G., Liu, P., Le, C., et al. (2019) A Novel Hybrid Meta-Heuristic Algorithm Based on the Cross-Entropy Method and Firefly Algorithm for Global Optimization. Entropy, 21(5).

[10] Qi, C., Chen, Q. and Kim, S. S. (2020) Integrated and intelligent design framework for cemented paste backfill: A combination of robust machine learning modelling and multi-objective optimization. Minerals Engineering, 155.

[11] Rubini, L. J. and Perumal, E. (2020) Hybrid Kernel Support Vector Machine Classifier and Grey Wolf Optimization Algorithm Based Intelligent Classification Algorithm for Chronic Kidney Disease. Journal of Medical Imaging and Health Informatics, 10(10): 2297-2307.

[12] Zhang, H., Wang, H., Li, N., et al. (2020) Time-optimal memetic whale optimization algorithm for hypersonic vehicle reentry trajectory optimization with no-fly zones. Neural Computing & Applications, 32(7): 2735-2749.

[13] Zhu, H. and Liu, T. (2020) Rotor Displacement Self-Sensing Modeling of Six-Pole Radial Hybrid Magnetic Bearing Using Improved Particle Swarm Optimization Support Vector Machine. Ieee Transactions on Power Electronics, 35(11): 12296-12306.

[14] Beaudoin, A., Bernier, P. Y., Guindon, L., et al. (2014) Mapping attributes of Canada's forests at moderate resolution through kNN and MODIS imagery. Canadian Journal of Forest Research, 44(5): 521-532.

[15] Song, J., Dong, F., Zhao, J., et al. (2019) An Efficient Multiobjective Design Optimization Method for a PMSLM Based on an Extreme Learning Machine. Ieee Transactions on Industrial Electronics, 66(2): 1001-1011.

[16] Du, Y. and Yang, N. (2019) Analysis of image processing algorithm based on bionic intelligent optimization. Cluster Computing-the Journal of Networks Software Tools and Applications, 22(2): S3505-S3512.

[17] Jaddi, N. S. and Abdullah, S. (2020) Global search in single-solution-based metaheuristics. Data Technologies and Applications, 54(3): 275-296.

[18] Akselsson, A., Lindgren, H., Georgsson, S., et al. (2020) Pregnancy outcomes among women born in Somalia and Sweden giving birth in the Stockholm area - a population-based study. Global Health Action, 13(1).

[19] Kastaun, S., Brown, J. and Kotz, D. (2020) Association between income and education with quit attempts, use of cessation aids, and short-term success in tobacco smokers: A social gradient analysis from a population-based cross-sectional household survey in Germany (DEBRA study). Addictive behaviors, 111: 106553-106553.

[20] Li, X. (2010) Niching Without Niching Parameters: Particle Swarm Optimization Using a Ring Topology. Ieee Transactions on Evolutionary Computation, 14(1): 150-169.

[21] Magnier, L. and Haghighat, F. (2010) Multiobjective optimization of building design using TRNNSYS simulations, genetic algorithm, and Artificial Neural Network. Building and
[22] Mallipeddi, R. and Suganthan, P. N. (2010). Ensemble of Constraint Handling Techniques. IEEE Transactions on Evolutionary Computation. 14: 561-579.

[23] Ghadimi, N., Akbarimajd, A., Shayeghi, H., et al. (2018) Two stage forecast engine with feature selection technique and improved meta-heuristic algorithm for electricity load forecasting. Energy, 161: 130-142.

[24] Aktulga, H. M., Fogarty, J. C., Pandit, S. A., et al. (2012) Parallel reactive molecular dynamics: Numerical methods and algorithmic techniques. Parallel Computing, 38(4-5): 245-259.

[25] Williams, D. (2020) Predictive coding and thought. Synthese, 197(4): 1749-1775.

[26] Mahesh, A. and Sandhu, K. S. (2020) A genetic algorithm based improved optimal sizing strategy for solar-wind-battery hybrid system using energy filter algorithm. Frontiers in Energy, 14(1): 139-151.

[27] Abedi, M., Chiong, R., Noman, N., et al. (2020) A multi-population, multi-objective memetic algorithm for energy-efficient job-shop scheduling with deteriorating machines. Expert Systems with Applications, 157.

[28] Khashei, M. and Bijari, M. (2010) An artificial neural network (p, d, q) model for timeseries forecasting. Expert Systems with Applications, 37(1): 479-489.

[29] Sun, W. and Wang, Y. (2018) Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network. Energy Conversion and Management, 157: 1-12.

[30] Lv, C., Xing, Y., Zhang, J., et al. (2018) Levenberg-Marquadt Backpropagation Training of Multilayer Neural Networks for State Estimation of a Safety-Critical Cyber-Physical System. IEEE Transactions on Industrial Informatics, 14(8): 3436-3446.

[31] Whittington, J. C. R. and Bogacz, R. (2019) Theories of Error Back-Propagation in the Brain. Trends in Cognitive Sciences, 23(3): 235-250.

[32] Peng, X. (2010) TSVR: An efficient Twin Support Vector Machine for regression. Neural Networks, 23(3): 365-372.

[33] Li, H.-z., Guo, S., Li, C.-j., et al. (2013) A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. Knowledge-Based Systems, 37: 378-387.

[34] Zhang, P., Wu, H.-N., Chen, R.-P., et al. (2020) Hybrid meta-heuristic and machine learning algorithms for tunneling-induced settlement prediction: A comparative study. Tunneling and Underground Space Technology, 99.

[35] Ahmadlou, M. and Adeli, H. (2010) Enhanced probabilistic neural network with local decision circles: A robust classifier. Integrated Computer-Aided Engineering, 17(3): 197-210.

[36] Wang, G., Zhang, B., Liao, W., et al. (2020) Estimation of functional regression model via functional dimension reduction. Journal of Computational and Applied Mathematics, 379.

[37] Nie, F., Xu, D., Tsang, I. W.-H., et al. (2010) Flexible Manifold Embedding: A Framework for Semi-Supervised and Unsupervised Dimension Reduction. IEEE Transactions on Image Processing, 19(7): 1921-1932.

[38] Pan, Q.-K., Suganthan, P. N., Tasgetiren, M. F., et al. (2010) A self-adaptive global best harmony search algorithm for continuous optimization problems. Applied Mathematics and Computation, 216(3): 830-848.

[39] Hudson, I. L. (2021) Data Integration Using Advances in Machine Learning in Drug Discovery and Molecular Biology. Methods in molecular biology (Clifton, N.J.), 2190: 167-184.

[40] Li, X. and Wu, B. (2017) A new reproducing kernel method for variable order fractional boundary value problems for functional differential equations. Journal of Computational and Applied Mathematics, 311: 387-393.

[41] Zheng, H., Wang, R., Yang, Y., et al. (2020) Intelligent Fault Identification Based on Multisource Domain Generalization Towards Actual Diagnosis Scenario. IEEE Transactions on Industrial Electronics, 67(2): 1293-1304.
[42] Moriasi, D. N., Gitau, M. W., Pai, N., et al. (2015) HYDROLOGIC AND WATER QUALITY MODELS: PERFORMANCE MEASURES AND EVALUATION CRITERIA. Transactions of the Asabe, 58(6): 1763-1785.

[43] Mallipeddi, R. and Suganthan, P. N. (2010) Ensemble of Constraint Handling Techniques. Ieee Transactions on Evolutionary Computation, 14(4): 561-579.

[44] Zhang, Q., Xia, D., Liu, K., et al. (2020) A general model of decision-theoretic three-way approximations of fuzzy sets based on a heuristic algorithm. Information Sciences, 507: 522-539.

[45] Berbeglia, G., Cordeau, J.-F. and Laporte, G. (2010) Dynamic pickup and delivery problems. European Journal of Operational Research, 202(1): 8-15.