Research on transformer fault diagnosis based on GWO-RF algorithm

Tianbing Wang¹, Lei Zhang¹ and Yufeng Wu²,*
¹China Railway Siyuan Survey and Design Group Co., Ltd., Wuhan 430063, China; ²School of Electric Power Engineering, South China University of Technology, Guangzhou 510640, China;

*Corresponding author: 202020113995@mail.scut.edu.cn

Abstract. Efficient fault diagnosis of power transformer can effectively ensure the safe and stable operation of power system. Considering that the effect of traditional Random Forest (RF) is seriously affected by the initial parameters, this paper proposes a RF algorithm based on Grey Wolf Optimization (GWO-RF) to improve the accuracy of transformer fault identification. The algorithm uses GWO to optimize the total number of decision trees and the depth of the maximum decision tree, effectively balancing the accuracy of the RF model with the ability to generalize. In this paper, the Dissolved Gas in oil is taken as the fault characteristic quantity, and 335 sets of data are used to form the fault set. Three different data sets are verified by GWO-RF, traditional RF and IEC three ratio method. The experimental results show the effectiveness of GWO-RF method.

Keywords: Transformer fault diagnosis, Random forest, Grey wolf optimization, Dissolved Gas in oil.

1. Introduction
The power transformer is one of the most important equipment in the power system. The accurate judgment and identification of transformer faults can effectively guarantee the safe and stable operation of the power system. It is of great significance to improve the accuracy of transformer fault diagnosis [1].

At present, the three-ratio method of Dissolved Gas Analysis (DGA) is still the most important fault diagnosis method [2, 3], but it has problems such as missed diagnosis in the application process. In response to these problems, intelligent fault diagnosis methods such as expert systems[4], support vector machines[5], fuzzy c-means clustering methods[, artificial neural networks, and Random Forest (RF)[6] have been proposed one after another. Among them, RF as a performance ensemble learning method, are more widely used in pattern recognition problems such as transformer fault diagnosis. RF combines the excellent performance of multiple decision trees, avoiding the limitations of a single classifier, and effectively improves the accuracy of transformer fault diagnosis. However, RF like other machine learning methods, requires reasonable settings for its hyperparameters during the model training process. Unreasonable parameter settings will easily cause missed diagnosis and misjudgment of transformer faults and affect the accuracy of detection results. And if the hyperparameters of RF are adjusted based
on experience, it will consume a lot of time and manpower. Indeed, these problems can be solved by directly optimizing RF parameters using artificial intelligence algorithms: The literature [7] propose a PSO-RF algorithm to predict photovoltaic processing, and achieve more ideal results than traditional RF. The algorithm optimizes the parameters of the random forest model based on the particle swarm optimization algorithm. The literature [8] proposes a Random Forest algorithm based improved grid search to make the classification performance of RF better. However, the above methods still have the defects of slow algorithm convergence and easy to fall into local optimum, which makes the selection of hyperparameters error, which affects the accuracy and speed of transformer fault diagnosis.

Aiming at the above problems, in this paper, a Random Forest algorithm based on Grey Wolf Optimizer (GWO-RF) is proposed for transformer fault diagnosis. The GWO algorithm is an intelligent group search algorithm with the characteristics of global optimization and rapid convergence. The proposed algorithm can use the global search capability of GWO to quickly search for the optimal the total number of decision trees and the maximum depth for RF algorithm, reduce the sensitivity of the RF algorithm to the hyperparameters, improve the accuracy of transformer fault diagnosis. Finally, this paper collects 335 sets of measured data from actual projects, and uses IEC three-ratio method, traditional RF algorithm, GWO-RF algorithm and other methods to test these data sets. The results show that the diagnostic model proposed in this paper performs better than other models in diagnostic accuracy.

2. Analysis of Transformer fault diagnosis based on GWO-RF algorithm

2.1. Random Forest Algorithm

The Random Forest Algorithm is an ensemble learning algorithm constructed on the basis of the traditional decision tree algorithm by applying statistical sampling principles. The random forest algorithm can be used for both classification problems can also be used for regression problems [9]. Specifically, RF belongs to the Bagging framework of ensemble learning, which uses an upper model (For classification problems, this model is usually a voter) to synthesize the performance of multiple decision trees $h(x, \Theta_k, k = 1, 2, \cdots, N)$ at the lower level to obtain credible prediction results. $\Theta_k$ represents the random variable of the $k$-th decision tree, which is independent and identically distributed and determines the relevant performance of the decision tree. The formation of RF is shown in Fig. 1.

Therefore, the performance of RF depends on the performance of the underlying decision tree. The split criterion of a node is an important aspect that determines the performance of the decision tree. Considering that the Gini coefficient is excellent in multiple classification scenarios, this article uses it as the node classification criterion. The specific calculation method of the Gini coefficient is as follows


\[ gini(T) = 1 - \sum_{i=1}^{m} p_i^2 \]  

(1)

Where, \( T \) represents the \( T \)-th sample set; \( m \) is the total number of categories in the sample set; \( p_i \) indicates the probability of the \( i \)-th category; If sample set \( T \) is divided into \( l \) sample subsets \( T_1, T_2, \ldots, T_l \), and the number of samples in the subset are \( N_1, N_2, \ldots, N_l \), then the Gini index of this split is [10]

\[ gini_{split}(T) = \frac{N_1}{N} gini(T_1) + \frac{N_2}{N} gini(T_2) + \cdots + \frac{N_l}{N} gini(T_l) \]  

(2)

After determining the split formation mode of each decision tree, the hyperparameters in the random forest need to be further considered. Indeed, RF includes two key parameters: the number of decision trees to be constructed and the max depth of a single decision tree. The number of lower-level decision trees determines the possible search space of the entire random forest. The more the number, the larger the possible search space, but when the number reaches a certain level, the optimal result may not be obtained because the search space is too large. In addition, the maximum depth of a single decision tree is directly related to its generalization performance. As the depth of the decision tree deepens, the deviation of its performance on the sample data will be smaller, but its variance will be correspondingly larger, that is, the generalization performance will decrease, which indirectly affects the generalization of the entire random forest. Therefore, we have to choose the two hyperparameters of RF reasonably to make its performance reach the best.

2.2. Grey Wolf Optimizer Algorithm

All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper.

2.2.1. Social Dominant Hierarchy of Leadership in Grey Wolves. In the structure of the GWO algorithm, each wolf has its own social hierarchy. In grey wolves social, these grades are marked as: \( \alpha, \beta, \delta, \omega \), respectively. Among them, \( \alpha \) wolf is the best individual in the wolf pack, which determines the other pack actions. \( \beta \) and \( \delta \) are the entities second only to \( \alpha \). And when \( \alpha \) expire, these individuals will be updated to become the new \( \alpha \) wolf. As the lowest grade grey wolf, \( \omega \) always have to obeys the decisions made by other wolves. Generally speaking, in a high-ranking wolf pack, \( \alpha, \beta, \delta \) wolves make decisions, \( \omega \) wolves obey orders and perform hunting tasks.

2.2.2. Encircling behavior. After \( \alpha, \beta, \delta \) wolves give instructions, the wolves often take the process of encircling the prey first and then attacking. In order to describe this encircling behavior process, the mathematical model is established as follows

\[ D = |C \cdot X_p(l) - X(l)| \]  

(3)

\[ X(l+1) = X_p(l) - A \cdot D \]  

(4)

\[ A = 2a \cdot r_1 - a \]  

(5)

\[ C = 2 \cdot r_2 \]  

(6)
\[ a = 2 - l \times \left( \frac{2}{l_{\text{max}}} \right) \]  

(7)

Where, \( l \) is the current iteration number. \( l_{\text{max}} \) is the maximum number of iterations. \( A \) and \( C \) represent the coordination coefficient matrix. \( X_p(l) \) indicates the position vector of the prey. \( X(l) \) and \( X(l + 1) \) are respectively the position vector of each individual in the wolf pack after being affected by \( X_p(l) \). \( a \) represents the convergence factor, which linearly decreases from 2 to 0 during the iteration.

And \( r_1, r_2 \) represents a random vector between \([0, 1]\).

2.2.3. Hunting behavior. After completing the encirclement of the prey, the wolves will start hunting operations under the guidance of \( \alpha, \beta \) and \( \delta \) wolf. Specifically, in each iteration, the best three solutions obtained will be divided into \( \alpha, \beta \), and \( \delta \) wolf in turn, and other gray wolves will change their positions according to \( \alpha, \beta \), and \( \delta \) wolf. The specific expression is

\[
D_{\alpha} = |C_1 \cdot X_\alpha(l) - X(l)|
\]

(8)

\[
D_{\beta} = |C_1 \cdot X_\beta(l) - X(l)|
\]

(9)

\[
D_{\delta} = |C_1 \cdot X_\delta(l) - X(l)|
\]

(10)

\[
X_1 = X_\alpha(l) - A_1 \cdot D_{\alpha}
\]

(11)

\[
X_2 = X_\beta(l) - A_2 \cdot D_{\beta}
\]

(12)

\[
X_3 = X_\delta(l) - A_3 \cdot D_{\delta}
\]

(13)

\[
X(l + 1) = (X_1 + X_2 + X_3)/3
\]

(14)

Where, \( X_1, X_2, X_3 \) indicate the movement directions of other wolves in the pack after being affected by the three wolves \( \alpha, \beta, \) and \( \delta \). According to formula (14), the final moving direction of the wolf pack is obtained, and the location of the wolf pack is updated iteratively. When convergence is reached, the wolf pack no longer moves and the current position of a wolf is the optimal solution. The GWO algorithm location update method is shown in the Fig. 2.
2.3. Transformer fault diagnosis based on GWO-RF algorithm

As mentioned above, the number of decision trees at the lower level and the maximum depth of a single tree affect the accuracy of RF model. Besides, GWO has a good global optimization ability and convergence properties. Therefore, this paper proposes a GWO-RF algorithm, which uses GWO to optimize two important hyperparameters of RF, and applies it to transformer fault diagnosis. The basic steps of GWO-RF are as follows:

(1) Step 1: Data processing. Input transformer monitoring data and preprocess it.

(2) Step 2: Initialization. Initialize the wolf pack \( P_i \), that is, randomly generate \( n \) pieces of coded data, expressed as follows

\[
P = \begin{bmatrix}
  n_{\text{estimators}}_1 & \text{max}_\text{depth}_1 \\
  n_{\text{estimators}}_2 & \text{max}_\text{depth}_2 \\
  \vdots & \vdots \\
  n_{\text{estimators}}_k & \text{max}_\text{depth}_k \\
  \vdots & \vdots \\
  n_{\text{estimators}}_N & \text{max}_\text{depth}_N 
\end{bmatrix}
\]

\[
(15)
\]

Where, \( N \) is the total number of individuals in the wolf pack. \( n_{\text{estimators}}_k, \text{max}_\text{depth}_k \) indicate the number of decision trees and the maximum depth of a single tree in the \( k \)-th individual, respectively. \( \text{max}_\text{estimators} \) and \( \text{max}_\text{depth} \) represents the upper bound of \( n_{\text{estimators}} \) and \( \text{max}_\text{depth} \). Correspondingly, \( \text{min}_\text{estimators} \) and \( \text{min}_\text{depth} \) is the lower bound of \( n_{\text{estimators}} \) and \( \text{max}_\text{depth} \).

(3) Step 3: Calculate fitness function. In order to determine the hierarchical relationship of each individual in the wolf pack, it is necessary to set a reasonable fitness function to quantify the degree of individual fitness. The fitness function \( f(P) \) defined in this paper is as follows

\[
f(P) = J(U, P)
\]

(16)
Where, $J(U, P)$ is objective function value, where it represents the accuracy of fault diagnosis.

(4) Step 4: Update the location of the wolves. Sort the individual wolves according to the fitness function value, and select $\alpha, \beta,$ and $\delta$ wolves. And update the location of the wolves according to formulas (3)-(14)

(5) Step 5: Convergence decision. After each iteration, we need to determine whether the algorithm has converged. If the algorithm converges, the algorithm proceeds to Step 6. Otherwise, it returns to Step 3 to continue iterating. When the discrete degree of the fitness value of the wolf pack is low, the global search ability of GWO begins to decline, and the algorithm tends to converge. Therefore, this paper uses the variance of the wolf pack fitness as the criterion. When the variance is smaller than the threshold, the algorithm converges.

(6) Step 6: Training RF. Set the RF model parameters according to the final a wolf position, and execute the RF algorithm to output the predict result.

3. Case Study

3.1. Feature Selection
In the field of transformer fault diagnosis, DGA is one of the most effective and widely used methods. By analysing the type of gas generated in transformer oil and its volume fraction, it is possible to determine whether abnormalities and specific fault types have occurred. The gases contained in transformer oil are $\text{H}_2$, $\text{CH}_4$, $\text{C}_2\text{H}_6$, $\text{C}_2\text{H}_4$, $\text{C}_2\text{H}_2$, $\text{CO}$ and $\text{CO}_2$. However, except the transformer failure, the solid insulation of the transformer and external infiltration will change the volume fraction of $\text{CO}$ and $\text{CO}_2$ too. On other hand, because the volume fraction of $\text{CO}$ and $\text{CO}_2$ is usually much larger than that of hydrogen hydrocarbon gas, using them as input features will affect the final accuracy of fault diagnosis. Therefore, $\text{CO}$ and $\text{CO}_2$ are not suitable as typical fault gases. Besides, the three ratios of dissolved gas $\text{C}_2\text{H}_2/\text{C}_2\text{H}_4$, $\text{CH}_4/\text{H}_2$ and $\text{C}_2\text{H}_4/\text{C}_2\text{H}_6$ have an important contribution to fault diagnosis. So, this paper selects $\text{H}_2$, $\text{CH}_4$, $\text{C}_2\text{H}_6$, $\text{C}_2\text{H}_4$, $\text{C}_2\text{H}_2$, $\text{C}_2\text{H}_4/\text{C}_2\text{H}_6$, $\text{CH}_4/\text{H}_2$ and $\text{C}_2\text{H}_4/\text{C}_2\text{H}_6$ as the feature quantity.

3.2. The selection of Data Samples
In order to verify the performance of GWO-RF algorithm proposed in this paper, we select 334 fault DGA data samples from the actual operating records of the transformer for simulation testing, which include 5 states of transformer: High-energy discharge fault (HD), Low-energy discharge fault (LD), High-temperature overheating fault (HT), Low-temperature overheating fault (LT) and Normal state (NS). A part of measure data samples and the distribution of samples are shown in Table 1. According to the ratio of 4:1, randomly select the data to form the training set and the test set.
### Table 1 DGA data samples of power transformer and its Distribution.

| Transformer State | Diagnosis gas content (μL/L) | Number of samples |
|-------------------|------------------------------|-------------------|
|                   | \( \text{H}_2 \) | \( \text{CH}_4 \) | \( \text{C}_2\text{H}_6 \) | \( \text{C}_2\text{H}_4 \) | \( \text{C}_2\text{H}_2 \) | \( \text{C}_2\text{H}_2/\text{C}_2\text{H}_4 \) | \( \text{CH}_4/\text{H}_2 \) | \( \text{C}_2\text{H}_4/\text{C}_2\text{H}_6 \) |
| HD                | 217.5 | 40 | 4.9 | 51.8 | 67.5 | 1.30 | 0.18 | 10.57 | 95 |
|                   | 1678 | 652.9 | 80.7 | 1005.9491.9 | 0.42 | 0.39 | 12.46 |  |  |
|                   | 60 | 40 | 6.9 | 110 | 70 | 0.64 | 0.67 | 15.94 |  |
| LD                | 345 | 112.25 | 27.5 | 51.5 | 58.75 | 1.14 | 0.33 | 1.87 | 58 |
|                   | 565 | 34 | 47 | 0 | 0.00 | 0.16 | 1.38 |  |  |
|                   | 115.9 | 75 | 14.7 | 25.3 | 6.8 | 0.27 | 0.65 | 1.72 |  |
| HT                | 172.9 | 334.1 | 172.9 | 812.5 | 37.7 | 0.05 | 1.93 | 4.70 | 109 |
|                   | 25.1 | 411.9 | 1320.9 | 1832.8 | 18.4 | 0.01 | 16.41 | 5.71 |  |
|                   | 274 | 376 | 55 | 1002 | 17 | 0.02 | 1.37 | 18.22 |  |
| LT                | 181 | 262 | 210 | 528 | 0 | 0.00 | 1.45 | 2.51 | 35 |
|                   | 160 | 130 | 33 | 96 | 0 | 0.00 | 0.81 | 2.91 |  |
|                   | 170 | 320 | 53 | 520 | 3.2 | 0.01 | 1.88 | 9.81 |  |
| NS                | 7.5 | 5.7 | 3.4 | 2.6 | 3.2 | 1.23 | 0.76 | 0.76 | 37 |
|                   | 220 | 340 | 42 | 480 | 14 | 0.03 | 1.55 | 11.43 |  |
|                   | 80 | 10 | 4 | 1.5 | 0 | 0.00 | 0.13 | 0.38 |  |

### 3.3. Calculation example results and analysis

In addition to the GWO-RF algorithm our proposed, this paper also selects several methods commonly used in transformer fault diagnosis for simulation, such as BP neural network, three-ratio method, SVM, traditional RF, to further prove the Superiority of our method. The parameter settings of the above methods are shown in Table 2. Fig 3 shows the accuracy of diagnosis results by different methods. Obviously, the diagnostic accuracy of GWO-RF on the training samples set and test samples set is higher than other methods. It can be seen from the results that the performance of the GWO-RF method on the training set and the test set is the best. Although the accuracy of BP on the training set is only 0.59% behind that of GWO-RF, its performance on the test set is far inferior to that of GWO-RF. Obviously, the phenomenon of over-fitting occurred in the learning process of BP network, and the idea of ensemble learning in GWO-RF and RF effectively avoided this situation. In addition, compared to traditional RF, GWO-RF performs better, which shows that using GWO method to adjust RF hyperparameters can effectively improve the accuracy of transformer fault diagnosis.

### Table 2 DGA data samples of power transformer and its Distribution.

| Method       | Parameter Setting                                                                 |
|--------------|-----------------------------------------------------------------------------------|
| BP           |                                                                                   |
| Three-Ratio  | Refer to Coding rules of the IEC three-ratio method, as shown in Table 3.          |
| SVM          | Regularization parameter C: 1.0; Kernel: RBF; Gamma: 1/8.                           |
| RF           | \( \text{n\textunderscore estimators: 60; max\textunderscore depth: 5;} \)         |
| GWO-RF       | \( \text{Number of wolves: 30; Max Iteration: 200; Epsilon: 10^{-7;}} \)          |
|              | \( \text{Upper Bound of n\textunderscore estimators: 150; Lower Bound of n\textunderscore estimators: 50;} \) |
|              | \( \text{Upper Bound of max\textunderscore depth: 20; Lower Bound of max\textunderscore depth: 4;} \) |
Table 3 Coding rules of the IEC three-ratio method.

| Fault Types | \( \text{C}_2\text{H}_2/ \text{C}_2\text{H}_4 \) | \( \text{CH}_4/ \text{H}_2 \) | \( \text{C}_2\text{H}_4/ \text{C}_2\text{H}_6 \) |
|-------------|-----------------|-----------------|-----------------|
| HD          | <0.1            | >1              | 1-4             |
| LD          | <0.2            | >1              | >4              |
| HT          | 0.6-2.5         | 0.1-1           | >2              |
| LT          | >1              | 0.1-0.5         | >1              |

Fig. 3 The Accuracy of different methods on train data set and test data set.

Table 4 The Accuracy of different methods on train data set and test data set.

| Methods     | Accuracy (%) | On train data set | On test data set | Total       |
|-------------|--------------|-------------------|------------------|-------------|
| GWO-RF      | 98.8764      | 80.597            | 95.2096          |             |
| RF          | 86.1423      | 77.6119           | 84.4311          |             |
| SVM         | 59.5506      | 64.1791           | 60.479           |             |
| Three-Ratio | 67.4157      | 68.6567           | 67.6647          |             |
| BP          | 98.2833      | 73.2673           | 93.2651          |             |

In order to further evaluate the performance of different methods, this article counts the accuracy of the above methods for identifying different types of faults. The results are shown in the Table 5 and Figure 4. It can be seen from Fig. 4 (a) that the recognition accuracy of GWO-RF in the training set for samples of any state is all above 96.26%, which is the best among all methods. It can be seen from Figure 4 (b) that the recognition accuracy of GWO-RF for HT, LT, and LD is not optimal (the difference from the optimal method is about 8.26%, 12.50%, 4.76%, respectively). But according to the results of Figure 3, the overall performance of GWO-RF in the test set is still the best. Note that all methods have large errors in the identification of LT and LD, which shows that the difference in data between LT and LD is not obvious. To achieve accurate classification of these two states, other features besides the gas content of the transformer oil must be added, such as vibration signals and temperature data.
Table 5 The Accuracy of different states on train data set and test data set.

| States | Accuracy of different methods (%) |  |  |  |  |  |
|--------|----------------------------------|--|--|--|--|--|
|        | GWO-RF | RF | SVM | Three-Ratio | BP |
|        | Train | Test | Train | Test | Train | Test | Train | Test |
| HT     | 100.000 | 89.286 | 96.296 | 89.614 | 96.429 | 88.889 | 85.142 | 100.000 | 85.366 |
| LT     | 96.296 | 50.000 | 70.370 | 37.500 | 3.704 | 0.000 | 62.500 | 95.652 | 58.333 |
| HD     | 100.000 | 83.333 | 96.104 | 83.333 | 75.000 | 66.667 | 74.026 | 72.222 | 98.508 | 67.857 |
| LD     | 97.959 | 66.667 | 61.225 | 66.667 | 4.082 | 0.000 | 26.531 | 11.111 | 97.727 | 71.429 |
| NS     | 96.970 | 100.000 | 87.879 | 100.000 | 60.606 | 100.000 | 60.600 | 75.000 | 96.774 | 83.333 |

Fig. 4 The Accuracy of different states ((a) on train data set; (b) on test data set.)

4. Conclusions
Aiming at the problem that the performance of the Traditional Random Forest Algorithm is too dependent on hyperparameter settings, this paper proposes a GWO-RF algorithm, which uses the Grey Wolf Optimization algorithm to determine the total number of decision trees and the maximum depth of the RF algorithm for transformer fault diagnosis. The simulation results on 334 samples prove that the method proposed in this paper can select the hyperparameters of random forest effectively, reduce the error of transformer fault diagnosis caused by unreasonable hyperparameter settings. In addition, compared with other methods, the GWO-RF has better accuracy.

Acknowledgments
This work was financially supported by Key Equipment Prediction and Health Management (PHM) System for Power Supply and Transformation (2019C16) fund.

References
[1] Guo, C., Wang, Y., Wang, M., Yin, Y., & Zhang, J. (2013). Time-varying outage model for transformers representing internal latent fault. Proceedings of the Csee, 33(1), 63-69.
[2] F Jakob, & Dukarm, J. J. (2015). Thermodynamic estimation of transformer fault severity. IEEE Transactions on Power Delivery, 30(4), 1941-1948.
[3] Bhowmick, S., & Nandi, S. (2015). Online detection of an interturn winding fault in single-phase distribution transformers using a terminal measurement-based modeling technique. IEEE Transactions on Power Delivery, 30(2), 1007-1015.
[4] Shi, R. F., Shi, Y. F., Jun, M. U., Hui, L. I., & Liu, X. H. (2014). Power transformer fault diagnosis
expert system with dissolved gas analysis in oil. Proceedings of the CSU-EPSA.

[5] Sun, H., Fang, J., Wang, Z., Sun, L., Liu, Q., & Wang, S. et al. (2008). Fault diagnosis model of dga for power transformer based on fcm and svm. Proceedings of SPIE - The International Society for Optical Engineering, 7127, 71271N-71271N-7.

[6] Shah, A. M., & Bhalja, B. R. (2016). Fault discrimination scheme for power transformer using random forest technique. Iet Generation Transmission & Distribution, 10(6), 1431-1439.

[7] Wang, X., D Luo, & Li, C. (2019). ABC-SVM and PSO-RF Model for Photovoltaic Forecasting Based on Big Data. 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD).

[8] Wen, B., Dong, W., Xie, W., & Jun, M. A. (2018). Parameter optimization method for random forest based on improved grid search algorithm. Computer Engineering and Applications, 54(10): 154-157.

[9] Wang, Y., Xia, S. T., Tang, Q., Jia, W., & Zhu, X. (2018). A novel consistent random forest framework: bernoulli random forests. IEEE Transactions on Neural Networks and Learning Systems, PP (99), 1-14.

[10] Breiman, L. (2001) Random Forests, Machine Learning, 45(1), 5-32.

[11] Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. Advances in Engineering Software, 69(3), 46–61.