ORIGINAL ARTICLE

International Journal of Applied Mathematics in Control Engineering

Journal homepage: http://www.ijamce.com

Feature Selection Based on Improved Black-winged Kite Optimization Algorithm

Jiabao Yang[a] | Jiawei Tan[a*]

[a]School of Mathematics and Statistics, Changchun University of Technology, Changchun, Jilin 130012, China

Correspondence
J. Tan, School of Mathematics and Statistics, Changchun University of Technology, Changchun, Jilin 130012, China
Email: tanjiawei@ccut.edu.cn

Article Info
Article history:
Received 18 April 2025
Accepted 20 June 2025
Available online 30 June 2025

Abstract

The application of intelligent optimization algorithms in feature selection has become an important research direction in the fields of machine learning and data mining. Feature selection aims to identify the subset of features that contribute most to model performance from high-dimensional data, in order to reduce computational complexity, eliminate redundant information, and improve model generalization. Intelligent optimization algorithms, by simulating the optimization mechanisms in nature, can efficiently explore the global optimal solution in complex search spaces. These algorithms typically combine the objective function to dynamically evaluate the quality of feature subsets and gradually approach the optimal feature combination through iterative optimization. The population initialization of the Black-winged Kite Optimization Algorithm is improved using the Tent chaotic map, and the migration and predation position updates of the Black-winged Kite are carried out through Levy flights. Additionally, the Golden Sine Algorithm strategy can also be used to obtain the optimal solution for the Black-winged Kite position update. Experiments show that the improved Black-winged Kite Optimization Algorithm demonstrates strong robustness and adaptability in feature selection tasks, making it particularly suitable for large-scale, high-dimensional, and nonlinear data scenarios.

KEYWORDS

Feature Selection, Black-winged Kite Optimization, Algorithm, Tent chaotic map, Levy flight strategy, Golden Sine Algorithm Strategy

1 | INTRODUCTION

The application of intelligent optimization algorithms in feature selection has become an important research direction in the fields of machine learning and data mining[1, 2, 3, 4, 5, 6]. Feature selection aims to identify the subset of features that contribute most to model performance from high-dimensional data, in order to reduce computational complexity, decrease redundant information, and improve model generalization ability. Intelligent optimization algorithms, by simulating natural optimization mechanisms, can efficiently explore the global optimum in complex search spaces.

Digital Object Identifiers: 10.62953/IJAMCE.405956

The Black-Winged Kite Optimization Algorithm (BKA) is a swarm intelligence optimization algorithm proposed by Wang et al. in 2024[7]. It is inspired by the attacking and migratory behaviors of black-winged kites in nature. The algorithm combines the Cauchy mutation strategy with the leader strategy to enhance convergence speed. However, it also faces issues such as weak global search ability, tendency to get trapped in local optima, and imbalance between global exploration and local exploitation. ZHAO H et al. improved the migration strategy based on the BKA algorithm[8]. MA H et al. further improved BKA by integrating multiple enhancement strategies, including a nonlinear convergence factor and adaptive t-distribution, which improve the algorithm's solution quality for optimization problems[9]. ZHAO M et al. used Tent chaotic mapping and nonlinear convergence factors to increase the optimization accuracy of the algorithm[10].

To further enhance the optimization capability of the Black-winged Kite Optimization Algorithm, this paper proposes a multi-improved Black-winged Kite Optimization Algorithm. The Levy flight strategy is introduced in the attack behavior phase to improve the global search ability of the algorithm, thereby reducing the likelihood of the algorithm getting trapped in local optima[11, 12]. These algorithms typically evaluate the quality of feature subsets dynamically based on the objective function and gradually approach the optimal feature combination through iterative optimization. Experiments demonstrate that intelligent optimization algorithms show strong robustness and adaptability in feature selection tasks, making them particularly suitable for large-scale, high-dimensional, and nonlinear data scenarios.

2 | RESEARCH CONTENT

2.1 | Black-winged Kite Optimization Algorithm

The Black-Winged Kite Optimization Algorithm primarily simulates the hunting and migration behaviors of the black-winged kite. The hunting behavior can be regarded as the process of searching for the optimal solution within a specific range, with the prey corresponding to the global optimal solution of the algorithm. The migration behavior corresponds to the algorithm exploring a larger area to avoid getting stuck in local optimal solutions and to find new potential solution regions. Population Initialization: First, generate a matrix and randomly assign the positions of each black-winged kite, using the following formula to evenly distribute their initial positions:

$$\mathbf{X}_i = BK_{lb} + \text{rand}\left(BK_{lb} - BK_{lb}\right) \tag{1}$$

Here, i is an integer between 1 and the number of individuals in the population, BK_{ub} and BK_{lb} are respectively the upper and lower limits of the black-winged kite,rand is an arbitrary value randomly generated from [0,1].

During initialization, the individual with the highest fitness is considered as the leader X_L of the initial population, that is, the optimal position of the black-winged kite,The calculation formula for X_L:

$$f_{\text{best}} = \min\left(f\left(\mathbf{X_i}\right)\right) \tag{2}$$

$$X_L = X\left(\text{find}\left(f_{\text{best}} = f\left(\mathbf{X_i}\right)\right)\right) \tag{3}$$

f_{best} is the best fitness value in the population, here taking the minimum value as an example.

Attack behavior: The black-winged kite is a predator of small mammals and insects on the grasslands. During flight, it adjusts the angles of its wings and tail according to wind speed, quietly hovers to observe prey, and then swiftly dives to attack, exhibiting two states: hovering in the air and swooping toward the prey. The position update is as follows:

$$y_{t+1}^{j,j} = \begin{cases} y_t^{j,j} + n(1+\sin(r)) \times y_t^{i,j}, \text{p} < r \\ y_t^{j,j} + n \times (2r-1) \times y_t^{j,j}, \text{else} \end{cases} \tag{4}$$

$y_{t+1}^{j,j}$ and $y_t^{j,j}$ represent the positions of the i black-winged kite in the j dimension at the t and $(t+1)$ iteration steps, respectively.

r is a random number in the range of 0 to 1, p is a constant value usually taken as 0.9, and n is a nonlinear factor.

$$n = 0.05 \times e^{-2\times\left(\frac{t}{T}\right)^2} \tag{5}$$

t is the number of iterations completed so far, and T is the total number of iterations.

Migration behavior: Migration behavior is led by a leader. If the fitness value of the current population is less than that of a random population, the leader will give up leadership and join the migrating population. In this case, the leader abandons the leadership of the population. Conversely, if the fitness value of the current population is greater than that of the random population, the leader continues to lead the population to the destination. The specific formula is:

$$y_{t+1}^{i,j} = \begin{cases} y_t^{i,j} + C(0,1) \times \left(y_t^{j,j} - L_t^j\right), F_i < F_{ri} \\ \\ y_t^{j,j} + C(0,1) \times \left(L_t^j - m \times y_t^{j,j}\right), \text{else} \end{cases} \tag{6}$$

$$m = 2 \times \sin(r + \pi/2) \tag{7}$$

Here, L_t^j represents the optimal position of the black-winged kite in the j dimension up to the t iteration. F_i is the current position (fitness value) obtained by the black-winged kite at the t iteration. F_{ri} is the exact fitness value at a random position during the t iteration.

The one-dimensional Cauchy distribution is a continuous probability distribution with two parameters. The following equation is the probability density function of the one-dimensional Cauchy distribution:

$$f(x,\delta,\mu) = \frac{1}{\pi}\frac{\delta}{\delta^2 + (x-\mu)^2}, \quad -\infty < x < \infty \tag{8}$$

When $\delta = 1$, $\mu = 0$, it becomes the probability density formula in standard form. The formula is as follows:

$$f(x,\delta,\mu) = \frac{1}{\pi}\frac{1}{x^2+1}, \quad -\infty < x < \infty \tag{9}$$

2.2 | Improving the Black Wing Kite Optimization Algorithm

The Black-winged Kite Optimization Algorithm primarily updates formulas based on two components: position updating and weight adjustment.

$$X_i^{t+1} = X_i^t + W \cdot V_i^t \tag{10}$$

Where X_i^t is the position of individual i in generation t. W is the inertia weight, used to control the current position and speed. V_i^t velocity vector is defined as a combination of target attraction and disturbance updates.

The basic update rule for the velocity vector is:

$$V_i^t = c_1 \cdot r_1 \cdot \left(X_{\text{best}} - X_i^t\right) + c_2 \cdot r_2 \cdot \left(X_{\text{random}} - X_i^t\right) \tag{11}$$

Where X_{best} is the position of the global optimal individual. X_{random} represents the random individual positions in the population. c_1 and c_2 are weighting factors used to balance optimal attraction and random disturbance. r_1 and r_2 are random numbers with a uniform distribution, used to introduce randomness.

The core of introducing dynamic adjustment lies in the adaptive variation of inertia weight and weighting factors. The following provides the improved formula derivation:

(1) Adaptive adjustment of inertia weight: The inertia weight controls the population's global exploration and local exploitation capabilities, and can be dynamically adjusted through linear or nonlinear decay functions:

$$W(t) = W_{\max} - \frac{t}{T} \cdot (W_{\max} - W_{\min}) \tag{12}$$

Among them: W_{max} and W_{min} are the maximum and minimum values of the inertia weight. t is the current number of iterations. T is the maximum number of iterations.Usually $W_{max} = 0.9$, $W_{min} = 0.4$.

Nonlinear adaptive adjustments, to more flexibly control W, can introduce a nonlinear function:

$$W(t) = W_{\min} + (W_{\max} - W_{\min}) \cdot \exp\left(-\frac{\beta \cdot t}{T}\right) \tag{13}$$

Where β controls the decay rate, usually taken as $\beta \in [2,5]$.

(2) Adaptive adjustment of weighting factors: In the early stages of optimization, it is necessary to enhance global search capabilities, while in the later stages of optimization, more focus is placed on local exploitation. The dynamic adjustment rules are as follows:

$$c_1(t) = c_{1,\min} + \frac{t}{T} \cdot \left(c_{1,\max} - c_{1,\min}\right) \tag{14}$$

$$c_2(t) = c_{2,\max} - \frac{t}{T} \cdot \left(c_{2,\max} - c_{2,\min}\right) \tag{15}$$

Here, $c_{1,\min}$, $c_{2,\min}$, $c_{1,\max}$, $c_{2,\max}$ is the threshold value of the weight factor.

2.3 | Population initialization based on Tent chaotic map

A chaotic map is a mapping method that can exhibit complex, nonlinear behavior, characterized by sensitivity to initial conditions and unpredictability. Among them, the Tent chaotic map stands out for its simple form and excellent chaotic performance. Its advantage lies in its ability to generate uniformly distributed chaotic sequences, enhancing the algorithm's global search capability.

Using the Tent chaotic map to optimize the initial population can improve the population diversity of the algorithm and overcome the drawback of getting stuck in local optima. As shown in the formula, it is generated through the Bernoulli transformation:

$$X_{n+1} = \begin{cases} 2x_n & 0 \le x_n \le 0.5 \\ 2(1 - x_n) & 0.5 \le x_n \le 1 \end{cases} \tag{16}$$

$$X_{n+1} = (2x_n)\,\mathrm{mod}1 \tag{17}$$

2.4 | Local search position update mechanism based on Levy flight

The Levy flight strategy can enhance the algorithm's flexibility and diversity in exploring the search space by adjusting the randomness of step size and direction. Therefore, the Levy flight strategy is introduced, defined as:

$$\text{Levy} : u = t^{-\lambda}, 1 < \lambda \le 3 \tag{18}$$

The specific formula for Levy flight path is $s = \mu / \mid \nu \mid^{1/\beta}$.

$$\sigma_\mu = \left[\frac{\Gamma(1+\beta) \cdot \sin(\pi \times \beta/2)}{\Gamma(1+\beta)/2] \cdot \beta/2^{(1+\beta)/2}}\right]^{1/\beta} \tag{19}$$

After obtaining the Levy flight path Levy(λ), by replacing the random number r used for individual position updates in the local search of the original algorithm with Levy(λ), the improved position update formula can be obtained:

$$x_i^{t+1} = \omega \cdot x_i^t + \left(\text{Levy}(\lambda)^2 \cdot x_j^t - x_k^t\right) \cdot f_i \tag{20}$$

2.5 | Golden Sine Algorithm Strategy

The Golden Sine Algorithm (Gold-SA) is a new meta-heuristic optimization algorithm proposed in 2017. The design of this algorithm is inspired by the sine function in mathematics, simulating the spatial search process of solutions to the optimization problem through the scanning of the sine function within the unit circle. At the same time, the golden sine algorithm also introduces the principle of the golden ratio, using the golden ratio coefficient to reduce the search space, making the search process more efficient and precise. The main process of the golden sine algorithm includes initializing the search range interval, calculating the function values at the two golden section points, and determining the step size and search direction for the next iteration based on the principle of the sine function. During the iterative process, the algorithm divides the search interval into two parts based on the golden ratio principle and selects the more favorable interval for the next round of search. In this way, the golden section algorithm can gradually approach the optimal solution of the problem. The main process of the Gold-SA algorithm is the update process of its solutions. First, the positions of n individuals are randomly generated, assuming that each solution to the optimization problem corresponds to the position of an individual in the search space, and x_i^t represents the solution vector of the i-th individual in the individual space.

$$x_i^t = \left(x_{i,1}, x_{i,2}, \ldots, x_{i,D}\right)(i = 1,2,\ldots,n; t = 1,2,\ldots,t_{\max}) \tag{21}$$

$$x_i^{t+1} = x_i^t \cdot |\sin(r_1)| + r_2 \cdot \sin(r_1) \cdot \left|\theta_1 \cdot g^* - \theta_2 \cdot x_i^t\right| \cdot f_i \tag{22}$$

In the formula: r_1 and r_2 are random numbers, and r_1 determines the distance an individual moves in the next iteration of the algorithm, $r_1 \in [0,2\pi]$. r_2 determines the direction of position updates in the next iteration, $r_2 \in [0,\pi]$. θ_1 and θ_2 are coefficients obtained from the golden ratio, which can reduce the search space and guide individuals towards the optimal value.

3 | EXPERIMENTAL ANALYSIS

This article selects the following 11 test functions to evaluate the performance of the improved Black-winged Kite Optimization Algorithm, in order to verify the algorithm's optimization effectiveness. The names of the test functions and the relevant parameters are shown in Table 1.

To enhance the persuasiveness of the experiment, it is necessary to compare the optimization performance of Particle Swarm Optimization (PSO), Simulated Annealing (SA), Genetic Algorithm (GA), Differential Evolution Algorithm (DEA), Gradient Descent (GD), and the improved Black-winged Kite Algorithm (BFO). The population size of all algorithms is 30, and the maximum number of iterations is 500. To reduce the statistical error of the algorithm's experimental results, each test function is run independently 30 times.

From the analysis of Table 2, it can be seen that for F1, F2, F3, F5, F6, F8, and F11, the improved Black-winged Kite optimization algorithm can find the theoretical optimal values in both unimodal and multimodal test functions, and it also performs better in the computation of results, outperforming other algorithms. For the discrete function F4 and the unimodal function F9, although it does not perform better than gradient descent, the improved Black-winged Kite algorithm performs well enough compared to other algorithms. In 11 test functions, the improved Black-winged Kite algorithm was able to find the optimal solution, which fully demonstrates the high stability and reliability of the Black-winged Kite algorithm; its various metrics show advantages over other algorithms, enhancing both global search capability and robustness. This makes the algorithm more efficient, stable, and flexible when solving complex optimization problems.

4 | CONCLUSION

Whether on unimodal benchmark functions or multimodal benchmark functions, the improved Black-winged Kite exhibits higher convergence accuracy, converges the fastest, and achieves excellent optimization results. The optimization effect is also very satisfactory; after improvement, it can converge quickly, indicating that the introduction of the Tent chaotic map

TABLE 1 Benchmark Test Function Information

Type	Function Name	Test function	Dimension
	F1:Sphere Function	$f(x)=\sum_{i=1}^{n} x_i^2$	30
	F5:Rosenbrock	$f(x)=\sum_{i=1}^{n-1}\left[100\cdot\left(x_{i+1}-x_i^2\right)^2+\left(x_i-1\right)^2\right]$	30
Unimodal function	F2:Schwefel's Problem 2.22	$f(x)=\sum_{i=1}^{n}\lvert x_i\rvert+\prod_{i=1}^{n}\lvert x_i\rvert$	30
	F6:Shifted Sphere Function	$f(\mathbf{x})=\sum_{i=1}^{n}\left(x_i-\mathbf{c}_i\right)^2$	30
	F7:Weighted Sphere Function with Random Noise	$f(x)=\sum_{i=1}^{n} w_i\cdot x_i^2+\varepsilon$	30
	F9:Rastrigin	$f(x)=10n+\sum_{i=1}^{n}\left[x_i^2-10\cos\left(2\pi x_i\right)\right]$	30
Multimodal function	F8:Schwefel Function	$f(x)=418.9829-x\cdot\sin(\sqrt{\mid x\mid})$	30
	F10:Ackley Function	$f(x)=-a\cdot\exp\left(-b\cdot\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right)-\exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos\left(c\cdot x_i\right)\right)+a+\exp(1)$	30
	F11:Griewank	$f(x)=1+\frac{1}{4000}\sum_{i=1}^{n} x_i^2-\prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right)$	30
	Himmelblau	$f(x,y)=\left(x^2+y-11\right)^2+\left(x+y^2-7\right)^2$	30
	F3: Sum of Different Powers	$f(\mathbf{x})=\sum_{i=1}^{n}\lvert x_i\rvert^{i+1}$	30
Discrete function	F4: Max Function	$\mathrm{Max}\left(a_1,a_2,\ldots,a_n\right)=\max\left(a_1,a_2,\ldots,a_n\right)$	30

makes the population more uniform, gives the population in the algorithm the advantage of initial positioning, increases the algorithm's likelihood of searching for high-quality solutions on a global scale, and enhances the algorithm's solving efficiency and accuracy. The improved algorithm exhibits significantly higher optimization accuracy than other algorithms and does not experience stagnation, indicating that the introduction of the Golden Sine Algorithm can compensate for the slow convergence of the algorithm in the later stages of iteration.

5 | FUNDING

This work was supported by Scientific Research Project of Jilin Provincial Department of Education (No.JJKH20220663KJ).

REFERENCES

[1] Dantzig, G B. Linear Programming and Extensions. Princeton: Princeton University Press; 1963.

[2] Gao Y, Yang Q, Wang X, et al. Overview of New Swarm Intelligent Optimization Algorithms. Journal of Zhengzhou University(Engineering Science) 2022;43(03):21–30.

[3] Zhang J, Zhang X, Du W, et al. Research on Urban Bus Path Model Based on Particle Swarm Optimization. Application Research of Computers 2007;(01):131–132+147.

[4] Wang Q, Hu R. Improved grasshopper optimizing multi-layer perceptrons neural netword and its application on data classification. Computer Engineering and Design 2022;43(12):3443–3452.

[5] Tian J, Liu X. Research on Financial Security Risk Perception Monitoring under Intelligent Optimization Algorithm: Based on Regional Data of China's Social Network During COVID-19. China Journal of Econometrics 2023;3(02):570–588.

TABLE 2 Benchmark Function Optimization Results Comparison

Function	BFO	BFO++	GA	PSO	SA	GD	DE
F1	2.67e-11	0.00	0.00	1.28e-16	1.25e-16	2.14e-16	1.34
F2	1.69	0.00	0.00	2.21e-03	1.44e-07	1.43e+1	2.61
F3	0.34	0.00	0.00	1.45e-11	1.79e-12	4.95e-13	2.23
F4	0.13	0.00	0.00	2.15e-05	7.17e-06	2.04	0.60
F5	3.38	6.25e-27	4.53e-28	11.35e-10	1.75e-10	4.38e-11	5.63
F6	1.64e-17	6.16e-32	6.16e-32	1.25e-16	6.78e-15	6.95e-14	3.70
F7	0.12	0.50	0.97	7.27e-02	5.66e-02	2.31+02	2.91
F8	-19.75	-19.64	-19.65	-19.67	-19.67	-19.73	-12.38
F9	2.99	0.00	0.00	2.49e-14	2.49e-14	36.81	16.50
F10	5.97e-03	4.44e-16	4.44	1.99e-08	1.58e-08	5.29	2.92
F11	1.48e-02	0.00	0.00	4.10e-02	7.40e-03	5.91e-02	0.24

[6] Martens D, Baesens B, Fawcett T. Editorial survey: swarm intelligence for data mining. Machine Learning 2011;82:1–42.
[7] Wang J, Wang W, Hu X, et al. Black-winged kite algorithm: a nature-inspired meta-heuristic for solving benchmark functions and engineering problems. Artificial Intelligence Review 2024;57(4):98.
[8] Zhao H, Li P, Duan S, et al. Inversion of image-only intrinsic parameters for steel fibre concrete under combined rate-temperature conditions: an adaptively enhanced machine learning approach. Journal of Building Engineering 2024;94:109836.
[9] Ma H, As'arry A, Feng Y, et al. Improved black-winged kite algorithm and finite element analysis for robot parallel gripper design. Advances in Mechanical Engineering 2024;16(10):16878132241288402.
[10] Zhao M, Su Z, Zhao C, et al. Improved black-winged kite algorithm based on chaotic mapping and adversarial learning. Journal of Physics: Conference Series 2024;2898(1):012040.
[11] IACCA G, DOS SANTOS JUNIOR V C, DE MELO V V. An improved Jaya optimization algorithm with Lévy flight. Expert Systems with Applications 2021;165:113902.
[12] He P, Wu W. Levy flight-improved grey wolf optimizer algorithm-based support vector regression model for dam deformation prediction. Frontiers in Earth Science 2023;11:1122937.

Jiabao Yang received his bachelor's degree from Changchun University of Technology in 2021, and is currently enrolled in Changchun University of Technology pursuing his master's degree, with his primary research interests in big data analytics.

Jiawei Tan received his doctor's degree from Jilin University in 2020, and Currently, he works at Changchun University of Technology, dedicated to researching artificial intelligence, intelligent optimization algorithms, and their applications in engineering.