IPTEK The Journal of Technology and Science, 33(2), 2022 (e/pISSN:2088-2033, 0853-4098)

DOI: 10.12962/j20882033.v33i2.13769 | Received 27 July, 2022; Revised 3 August, 2022; Accepted 3 Sept, 2022

ORIGINAL RESEARCH

FEATURE SELECTION USING HYBRID BINARY GREY WOLF OPTIMIZER FOR ARABIC TEXT CLASSIFICATION

Muhammad Bahrul Subkhi | Chastine Fatichah* | Agus Zainal Arifin

[1]Dept. of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Correspondence

*Chastine Fatichah, Dept of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. Email: chastine@if.its.ac.id

Present Address

Gedung Teknik Informatika, Jl. Teknik Kimia, Kampus ITS Sukolilo, Surabaya 60111, Indonesia

Abstract

Feature selection in Arabic text is a challenging task due to the complex and rich nature of Arabic. The feature selection requires solution quality, stability, convergence speed, and the ability to find the global optimal. This study proposes a feature selection method using the Hybrid Binary Gray Wolf Optimizer (HBGWO) for Arabic text classification. The HBGWO method combines the local search capabilities or exploratory of the BGWO and the search capabilities around the best solutions or exploits of the PSO. HBGWO method also combines SCA's capabilities in finding global solutions. The data set used Arabic text from islambook.com, which consists of five Hadith books. The books selected five classes: Tauhid, Prayer, Zakat, Fasting, and Hajj. The results showed that the BGWO-PSO-SCA feature selection method with the fitness function search and classification method using SVM could perform better on Arabic text classification problems. BGWO-PSO with fitness function and the classification method using SVM (C=1.0) gives a high accuracy value of 76.37% compared to without feature selection. The BGWO-PSO-SCA feature selection method provides an accuracy value of 88.08%. This accuracy value is higher than the BGWO-PSO feature selection and other feature selection methods.

KEYWORDS:

BGWO, BGWO-PSO, BGWO-PSO-SCA,Feature Selection, Text Classification

1 | INTRODUCTION

The Arabic text classification system is a challenging task due to the complex and rich nature of Arabic. Unlike English, Arabic has twenty-eight letters, two genders (feminine and masculine), many inflections, and is written from right to left. In Arabic, nouns are singular, double, or plural. Besides, Arabic has three grammatical cases: nominative, accusative, and genitive[1].

Text documents are usually converted into term-frequency vectors to perform document processing tasks in the classification[2]. Such a transformation raises the problem of high feature space dimensions because each unique token in each document will be represented as a dimension in the feature space. There are two approaches used to reduce dimensions: feature extraction (FE) and feature selection (FS). Both approaches aim to reduce the number of features in the feature space, but the FE method reduces the feature space by generating new feature combinations. In contrast, the FS method includes and removes features presented in the feature space without changing them. Research conducted by Chantar et al.[1] proposes the Binary Grey Wolf Optimizer (BGWO) as a Wrapper-based feature selection technique for classifying Arabic documents with high dimensions and problems in identifying the most relevant features. BGWO is used at the optimal feature subset selection stage. In contrast, Decision Tree, K-Nearest Neighbor (KNN), Naïve Bayes (NB), and Support Vector Machine (SVM) is used in the evaluation stage to assess the quality of the selected feature subset. This research shows that the proposed feature selection technique with an elite-based crossover scheme can perform better on Arabic document classification problems. Feature Selection generated by BGWO in Arabic document classification gives good performance in solution quality. However, feature selection requires solution stability, quality, convergence speed, and the ability to find the global optimal.

This study proposes a Hybrid feature selection method that combines BGWO-PSO with SCA for Arabic text classification. BGWO-PSO is used for the power of exploration and exploitation, and SCA is used for improving the best optimal global solutions. This research is expected to improve the performance of convergence, exploration, exploitation, and global solutions.

2 | PREVIOUS RESEARCHES

Research conducted by Singh and Singh[3] used exploration of PSO and exploitation of GWO[4–6] to solve optimization problems. Research of Al-Tashi et al.[7] using suitable operators to solve binary problems by combining BGWO[1, 8], and GWOPSO[3] produces Binary PSOGWO (BGWO-PSO). The experimental results of the proposed BGWO-PSO are better than other methods.

The research conducted by Singh and Singh[9] improves global convergence, exploration, and exploitation performance by accelerating search and letting the algorithm run multiple generations without any improvement through Hybrid Grey Wolf Optimizer (GWO)-Sine Cosine Algorithm (SCA)[9] or GWO-SCA[10]. The experimental results prove that the proposed Hybrid algorithm effectively solves real applications with or without a restricted and unknown search area.

Several studies have proposed a hybrid algorithm for feature selection by combining the exploitation and exploration capabilities between algorithms, including a study conducted by Salton and Buckley[2] that improves the performance of the Grey Wolf Optimizer (GWO) by Chantar et al.[1] combining Particle Swarm Optimization (PSO). Improved performance by using the exploitation capabilities of the PSO and the exploration capabilities of the GWO aims to significantly outperform the PSO and GWO algorithms in terms of solution quality, convergence speed, and the ability to find global optimal.

In the research conducted by Fauzi et al.[11] to solve the problem of weighting a word using TF, this study used a dataset of 13 Arabic Fiqh e-books. This work implemented classification using a vector space model (VSM). In the research, the method's best performance was obtained using the best 1000 features with 76% precision value, 74% recall value, and 75% F-Measure value.

3 | MATERIAL AND METHOD

3.1 | Grey Wolf Optimizer (GWO)

In the feature selection method of the grey wolf optimizer, the most suitable solution is called alpha. The second and third best solutions are named beta and delta. The rest of the candidate solutions are assumed to be omega. Groups of grey wolves in searching for prey surround the prey first. The behavior model for circling prey is mathematically written in Eq. 1 to 4[12].

$$\vec{X}(t+1) = \vec{X}_p(t) + \vec{A} \,.\, \vec{D} \tag{1}$$

Where $\vec{D}$ is defined in Eq. refeq:eq2, t is the number of current iterations, $\vec{A}$ and $\vec{C}$ are coefficient vectors, $\vec{X}_p$ is the position vector of prey, and $\vec{X}$ is the position vector of a grey wolf.

$$\bar{X} = |\vec{C} \,.\, \vec{X}_p(t) + \vec{X}(t)| \tag{2}$$

The vectors $\vec{A}$ and $\vec{C}$ are calculated in Eq. 3 and Eq. 4.

$$\vec{A} = 2a \, . \, \vec{r}_1 - a \tag{3}$$

$$\vec{C} = 2\vec{r}_2 \tag{4}$$

Where a is linearly derived from 2 to 0 during the iteration, r_1 and r_2 are random vectors in [1,0]. The alpha usually guides the hunt. Beta and delta may also occasionally participate in hunting. To mathematically simulate the hunting behavior of the grey wolf, alpha (best candidate solution), beta (second best candidate solution), and delta (third best candidate solution) were assumed to have better knowledge of potential prey locations. The first three best candidate solutions were obtained and required other search agents (including omega) to update their positions according to the best search agency positions. The Update for the wolf position can be seen in Eq. 5.

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{5}$$

Where $\vec{X}_1$, $\vec{X}_2$, and $\vec{X}_3$ are defined as in equations (6) to (8), respectively.

$$\vec{X}_1 = |\vec{X}_\alpha - \vec{A}_1 . \vec{D}_\alpha| \tag{6}$$

$$\vec{X}_2 = |\vec{X}_\beta - \vec{A}_2 . \vec{D}_\beta| \tag{7}$$

$$\vec{X}_3 = |\vec{X}_\delta - \vec{A}_3 . \vec{D}_\delta| \tag{8}$$

Where $\vec{X}_\alpha$, $\vec{X}_\beta$, and $\vec{X}_\delta$ are the three best solutions in the herd of wolves at a given iteration t, $\vec{A}_1$, $\vec{A}_2$, and $\vec{A}_3$ are defined in equation (3), and $\vec{D}_\alpha$, $\vec{D}_\beta$, $\vec{D}_\delta$ are expressed as equations (9) to (11) respectively.

$$\vec{D}_\alpha = |\vec{C}_1 \, . \, \vec{X}_\alpha - \vec{C}| \tag{9}$$

$$\vec{D}_\beta = |\vec{C}_2 \, . \, \vec{X}_\beta - \vec{C}| \tag{10}$$

$$\vec{D}_\delta = |\vec{C}_3 \, . \, \vec{X}_\delta - \vec{C}| \tag{11}$$

Where $\vec{C}_1$, $\vec{C}_2$, $\vec{C}_3$ is defined in Eq. 4.

3.2 | Hybrid Binary Grey Wolf Optimizer (BGWO) with PSO

Research by Singh and Singh[3] combines the exploration of PSO[13] and the exploitation of GWO to solve optimization problems. A study by Chantar et al.[1] uses suitable operators to solve binary problems by combining BGWO[12] and GWOPSO[3], which produces Binary PSOGWO (BGWO-PSO). The experimental results of the proposed BGWO-PSO are better than other methods.

The wolf updating mechanism is a function of three vector positions, namely x_1, x_2 and x_3 so each wolf goes to the first three best solutions. Furthermore, the position update can be modified into equation (12) to work in a binary space.

$$x_d^{t+1} = \begin{cases} 1, & if\ Sigmod\{\frac{x_1+x_2+x_3}{3}\} \geq rand \\ 0, & otherwhise \end{cases} \tag{12}$$

Where x_d^{t+1} is the binary update position at iteration t in dimension d and the rand is a random number taken from a uniform distribution ∈ [1,0], and $Sigmoid(a)$ is denoted as follows:

$$Sigmoid(a) = \frac{1}{1 + e^{-10(x-0.5)}} \tag{13}$$

The positions x_1, x_2 and x_3 were updated and calculated using Eq. 14 until Eq. 16 as follows:

$$(x_1^d) = \begin{cases} 1, & if \left(x_\alpha^d + bstep_\alpha^d\right) > 1 \\ 0, & otherwhise \end{cases} \tag{14}$$

$$(x_2^d) = \begin{cases} 1, & if \left(x_\beta^d + bstep_\beta^d\right) > 1 \\ 0, & otherwhise \end{cases} \tag{15}$$

$$(x_3^d) = \begin{cases} 1, & if \left(x_\delta^d + bstep_\delta^d\right) > 1 \\ 0, & otherwhise \end{cases} \tag{16}$$

Where $x_{\alpha,\beta,\delta}^d$ is the position vector of alpha, beta, delta wolf in dimension d, and $bstep\alpha, \beta, \delta^d$ is a binary step in dimension d which can be formulated as Eq. 17.

$$bstep_{\alpha,\beta,\delta}^d = \begin{cases} 1, & if\, cstep_{\alpha,\beta,\delta}^d \geq rand) \\ 0, & otherwhise \end{cases} \tag{17}$$

With a random value of rand derived from a uniform distribution of ∈ [1,0], d denotes the dimension, and *cstepd* continuous value of $bsteps_{\alpha,\beta,\delta}^d$. This component was calculated using the following Eq. 18.

$$bstep_{\alpha,\beta,\delta}^d = \frac{1}{1 + e^{-10(A_1^d D_{\alpha,\beta,\delta}^d - 0.5)}} \tag{18}$$

In BGWO-PSO, and based on the three best solution positions updated, exploration and exploitation are controlled by a constant weight of inertia which is mathematically modeled as Eq. 19[3].

$$\begin{aligned} \vec{D}_\alpha &= |\vec{C}_1.\vec{X}_\alpha - w * \vec{X}|, \\ \vec{D}_\beta &= |\vec{C}_2.\vec{X}_\beta - w * \vec{X}|, \\ \vec{D}_\delta &= |\vec{C}_3.\vec{X}_\delta - w * \vec{X}|, \end{aligned} \tag{19}$$

Thus, the velocity and position have been updated via Eq. 20.

$$v_i^{k+1} = w * \left(v_i^k + c_1 r_1 (x_1 - x_i^k) + c_2 r_2 (x_2 - x_i^k) + c_3 r_3 (x_3 - x_i^k)\right) \tag{20}$$

In equation (20), the three best solutions, $x1$, $x2$, and $x3$ are updated according to equation (14).

$$x_i^{k+1} = x_d^{t+1} + v_i^{k+1} \tag{21}$$

x_d^{t+1} and v_i^{k+1} were calculated based on Equation (12) and Equation (20), respectively.

3.3 | Hybrid Binary Grey Wolf Optimizer with SCA

Research conducted[9] combined GWO with SCA. In this variant, the alpha agent movement of the grey wolf algorithm was improved based on the sine cosine algorithm. The research was intended to improve global convergence, exploration, and exploitation performance by speeding up searches instead of letting variants run multiple generations without any improvements.

The technique proposed by Mirjalili[14] called Sine Cosine Algorithm (SCA), only based on Sine and Cosine functions, was applied for the exploitation and exploration phases in the global optimization function. The Sine Cosine Algorithm (SCA) generates different initial random agent solutions. It requires the keys to fluctuate outward or toward the best solution using a mathematical model based on the sine and cosine functions.

$$\vec{x}_i^{t+1} = \vec{x}_i^t + rand_1 \times sin(rand_2) \times |rand_3 \times l_i^t - \vec{x}_i^t| \tag{22}$$

$$\vec{x}_i^{t+1} = \vec{x}_i^t + rand_1 \times cos(rand_2) \times |rand_3 \times l_i^t - \vec{x}_i^t| \tag{23}$$

Where $\vec{x}_i^t$ is t iteration in i dimensions, $rand_1, rand_2, rand_3 \in [0, 1]$ is a random number, and l_i is the targeted optimal global solution. The $0.5 < rand_4 \geq 0.5$ condition is in equation (22), and equation (23) is used for exploitation and exploration.

$$\vec{x}_i^{t+1} = \begin{cases} \vec{x}_i^t + rand_1 \times sin(rand_2 \times rand_3 \times l_i^t - \vec{x}_i^t|, & rand_4 < 0.5 \\ \vec{x}_i^t + rand_1 \times cos(rand_2 \times rand_3 \times l_i^t - \vec{x}_i^t, & rand_4 \geq 0.5) \end{cases} \tag{24}$$

In this algorithm, the position, speed, and accuracy of the grey wolf agent convergence (alpha) apply equation (24) for position updates. The SCA algorithm is used to balance the exploration and exploitation process and expand the convergence performance of the grey wolf optimization algorithm. Equation (25) applies the SCA algorithm in updating the position of the grey wolf (alpha)[9].

$$\vec{x}_\alpha \begin{cases} rand() \times sin(rand \times \vec{c}_1 \times \vec{x}_a - \vec{x}, & rand_4 < 0.5 \\ rand() \times cos(rand \times \vec{c}_1 \times \vec{x}_a - \vec{x}, & rand_4 \geq 0.5 \end{cases} \tag{25}$$

$$\vec{x}_1 = \vec{x}_\alpha - \vec{a}_1.\vec{d}_\alpha \tag{26}$$

$$\vec{x}_2 = \vec{x}_\beta - \vec{a}_1.\vec{d}_\beta \tag{27}$$

$$\vec{x}_3 = \vec{x}_\delta - \vec{a}_1.\vec{d}_\delta \tag{28}$$

$$\vec{x} = \frac{\vec{x}_1 + \vec{x}_2 + \vec{x}_3}{3} \tag{29}$$

Increasing the global solution to the grey wolf (alpha) can be done using equation (26). Where $\vec{x}_\alpha$ is a value in the alpha position, $\vec{a}_1$ is a random value, and $\vec{d}_\alpha$ is the result of the conditions for applying the SCA algorithm with equation (25). Equations 26 to 29 are used to evaluate the act of hunting prey, where the potential position of the prey is well known by α, β, δ[9].

3.4 | Fitness Function

The best combination of features is the combination with the maximum classification performance and the minimum number of selected features[12]. The K-Nearest Neighbor (KNN) method[13] is a simple and easy method to implement and a very general classifier. KNN is used as a classification to ensure the goodness of the selected features[12]. Thus, the goodness of a subset is determined based on two main criteria, namely the number of features selected in the subset and the error rate of KNN obtained

TABLE 1 The parameter setting used in this study.

Parameter	Value
α	0.99
β	1-α
Iteration	100
Search Agents	10
Inertia weight for BPSO	0.9
C1 and C2 parameters for the velocity of BPSO	2

TABLE 2 The data distribution used in this study.

Data	Training	Test	Feature Size
Small	590	254	2893
Big	3208	1376	5081

TABLE 3 Time, feature selection results, and fitness function with KNN (k=5) on the feature selection.

Feature Selection	Time	Result	Fitness Function
BGWO	45 seconds	751	0.322
BPSO	39 seconds	1423	0.309
BGWO-PSO	54 seconds	772	0.377
BGWO-PSO-SCA	49 seconds	1209	0.429

by using a subset of features selected from training data with the BGWO algorithm. The lower the error rate and the minimum number of features selected, the better the feature subset.

Minimization problems can be done by using error rate rather than classification quality and using selected feature ratios instead of unselected feature sizes. The minimization problem can be formulated as in equation (30).

$$Fitness = \alpha E_R(D) + \beta \frac{|R|}{|C|} \tag{30}$$

Where $E_R(D)$ is the error rate of the classification process, R is the length of the selected feature subset, C is the total number of features, and $\alpha \in [0, 1]$ and $\beta = 1 - \alpha$ are constants to control the importance of classification accuracy and feature reduction.

4 | RESULT

This study aims to combine the exploitation capabilities of PSO and BGWO exploration for feature selection to improve the performance of Arabic text classification more optimally and combine SCA's ability to find global solutions in BGWO-PSO. The proposed algorithm can be seen in Algorithm 1, and the overall process of the system built to achieve the research objectives can be seen in Fig. 1 , using the parameter settings as shown in Table 1 . The process was started by preprocessing the Hadith data to produce a set of very informative words to represent the text as a feature vector. Furthermore, the feature selection process aims to improve accuracy and computational efficiency in text classification by eliminating irrelevant features. The feature selection method proposed in this study is BGWO-PSO-SCA.

The Arabic text of Hadith data was taken from research [15]. The distribution of the number of data or classes is shown in Table 2 . The data was divided into five classes: Tauhid, prayer, zakat, fasting, and hajj. Two types of data in Table 2 are imbalanced data.

4.1 | Small Imbalance Dataset

The feature selection trial was carried out with 844 imbalanced data. The feature selection process was time-consuming, indicating the time required for each feature selection method process. Feature size is the number of features selected in the feature

Algorithm 1 Psedocode BGWO-PSO-SCA

Initialization the population
Initialize A, a and C
Find the fitness of each search member $\vec{x}_\alpha$ are the best search agents $\vec{x}_\beta$ are the $2^n d$ best search agents $\vec{x}_\delta$ are the 3^{rd} best search agents
while (t<Maximum number of iterations) **do**
  for every search member **do**
    Update the velocity using Eq. 20
    Update the position of agents into a binary position based on Eq. 21
  end for
  Update $\vec{a}$ and $\vec{c}$
  Calculate the fitness of all search member
  Update the position of $\vec{x}_\beta$, $\vec{x}_\delta$, and $\vec{x}_\alpha$ as below:
  if $rand() < 0.5$ **then** cos based on Eq. 25
  elsesin based on Eq. 25
  end if
  Evaluation position use Eq. 26 - 29
end while
return $\vec{x}_a$

TABLE 4 Time, feature selection results, and fitness function with SVM (C=1.0) in the feature selection.

Feature Selection	Time	Result	Fitness Function
BGWO	02 minutes 10 seconds	586	0.185
BPSO	03 minutes 19 seconds	1424	0.230
BGWO-PSO	02 minutes 35 seconds	615	0.189
BGWO-PSO-SCA	01 minute 23 seconds	943	0.272

TABLE 5 Time, feature selection results, and fitness function with SVM (C=10) in the feature selection.

Feature Selection	Time	Result	Fitness Function
BGWO	02 minutes 32 seconds	790	0.158
BPSO	03 minutes 16 seconds	1416	0.203
BGWO-PSO	02 minutes 41 seconds	723	0.177
BGWO-PSO-SCA	01 minute 37 seconds	1474	0.231

TABLE 6 Accuracy results with the feature selection method using KNN (k=5) on the fitness function.

Classification	BGWO	BPSO	BGWO-PSO	BGWO-PSO-SCA
KNN (k=5)	69.685	61.417	61.417	56.299
SVM (C=1.0)	69.685	66.535	71.259	65.354
SVM (C=10)	73.622	72.047	76.377	66.929

selection and getting the best fitness function value. The test results for each classification change in feature selection are shown in Table 3 , Table 4 , and Table 5 .

The feature selection process was complete; then, the classification process was used to measure accuracy. The results of the accuracy of the PSO feature selection method and GWO, which has become a binary version, as well as the Hybrid Binary Gray Wolf Optimizer with Particle Swarm Optimization, are shown in Table 6 . The classification process on the fitness function used KNN (k=5).

The next test was to change the fitness function using the SVM classification method, using parameters C=1.0 and C=10. Tables 7 and Table 8 are the results of trials using feature selection. The data used is 844.

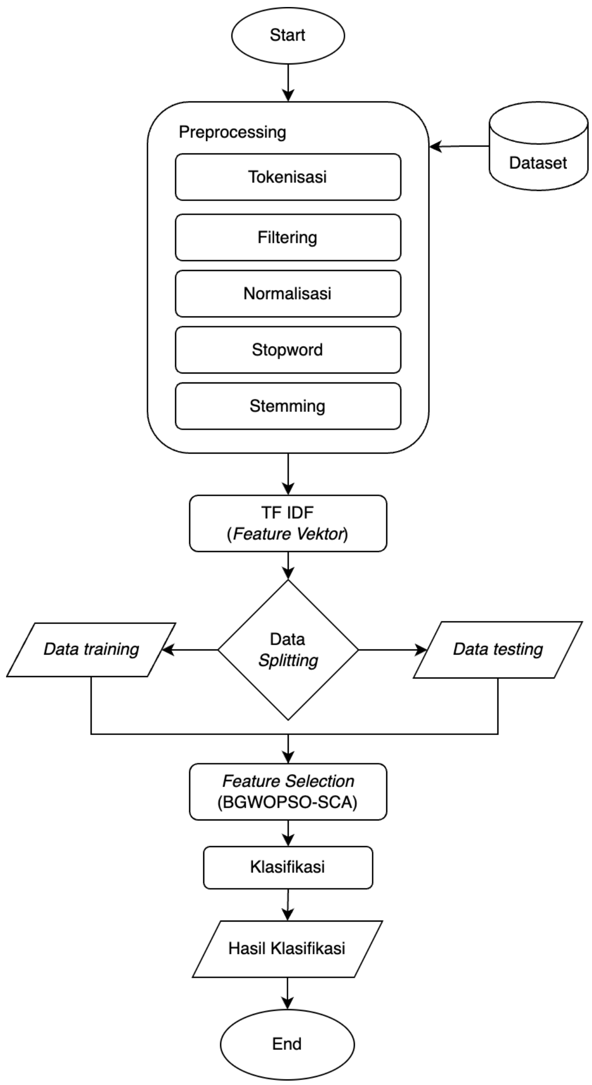

FIGURE 1 The Arabic text classification process.

TABLE 7 Accuracy results with the feature selection method using SVM (C=1.0) on the fitness function.

Classification	BGWO	BPSO	BGWO-PSO	BGWO-PSO-SCA
KNN (k=5)	55.511	60.236	51.181	53.937
SVM (C=1.0)	83.858	81.102	81.496	84.251
SVM (C=10)	77.952	80.314	78.346	81.496

Table 9 The Highest Accuracy Value for each Change in the Classification Method on the Fitness Function.

From the results of trials carried out using 844 Hadith data in Tables 6 to 8 , it can be concluded that the highest accuracy value of each classification changes in the fitness function. The highest accuracy value is shown in Table 9 . From that table, the highest accuracy value uses the BGWO-PSO-SCA feature selection method, where a fitness function and the classification method use SVM (C=10).

4.2 | Big Imbalance Dataset

This trial uses 4584 data; the feature selection process takes time and produces the number of selected features. The time required for the feature selection process is shown in Table 10 . The fitness function uses KNN with k=5. Tables 11 and Table 12

TABLE 8 Accuracy results with the feature selection method using SVM (C=10) on the fitness function.

Classification	BGWO	BPSO	BGWO-PSO	BGWO-PSO-SCA
KNN (k=5)	47.244	41.338	41.338	50.787
SVM (C=1.0)	72.834	74.803	77.165	77.165
SVM (C=10)	85.039	82.677	81.496	85.433

TABLE 9 The data distribution used in this study.

Feature Selection	Fitness Function	Classification	Accuracy (%)
BGWO-PSO	KNN (k=5)	SVM (C=10)	76.377
BGWO-PSO-SCA	SVM (C=1.0)	SVM (C=1.0)	84.251
BGWO-PSO-SCA	SVM (C=10)	SVM (C=10)	85.433

TABLE 10 Time, Feature Selection Results, and Fitness Function with KNN (k=5) in the Feature Selection.

Feature Selection	Time	Result	Fitness Function
BGWO	05 minutes 05 seconds	1879	0.296
BPSO	04 minutes 58 seconds	2496	0.284
BGWO-PSO	05 minutes 24 seconds	1618	0.298
BGWO-PSO-SCA	03 minutes 27 seconds	1913	0.382

TABLE 11 Time, Feature Selection Results, and Fitness Function with SVM (C=1.0) in the Feature Selection.

Feature Selection	Time	Result	Fitness Function
BGWO	42 minutes 37 seconds	1721	0.141
BPSO	63 minutes 45 seconds	2545	0.151
BGWO-PSO	54 minutes 49 seconds	1930	0.163
BGWO-PSO-SCA	18 minutes 48seconds	2574	0.197

TABLE 12 Time, feature selection results, and fitness function with SVM (C=10) in the Feature Selection.

Feature Selection	Time	Result	Fitness Function
BGWO	64 minutes 48 seconds	1817	0.152
BPSO	76 minutes 11seconds	2545	0.144
BGWO-PSO	59 minutes 05 seconds	1766	0.146
BGWO-PSO-SCA	39 minutes 30 seconds	2475	0.183

are the time required and the number of results from the feature selection method in which the fitness function uses the SVM classification method with parameters C = 1.0 and C = 10.

The trial was done by increasing the amount of Hadith data aimed at whether there was a significant increase or change in results using 4584 data. The accuracy test results are shown in Tables 13 to 15 .

From the results of trials carried out using 4584 Hadith data in Table 13 to Table 15 , it can be concluded that the highest accuracy value of each classification changed in the fitness function. The highest accuracy value is shown in Table 16 . The highest accuracy value uses the BGWO-PSO-SCA feature selection method with a fitness function, and the classification method uses SVM (C=10).

5 | DISCUSSION

This study succeeded in conducting feature selection using the Hybrid Binary Gray Wolf Optimizer feature selection method by combining the exploitation capabilities of PSO and exploration capabilities of BGWO and SCA to find the best global optimal solution. The combined feature selection method can select features locally or exploit by proving that the number of selected features is more than GWO and less than PSO and can find optimal solutions from exploration capabilities as evidenced by more

TABLE 13 Accuracy Results with the Feature Selection Method Using KNN (k=5) on the Fitness Function.

Classification	BGWO	BPSO	BGWO-PSO	BGWO-PSO-SCA
KNN (k=5)	70.421	71.802	70.130	83.284
SVM (C=1.0)	78.052	83.357	78.997	82.848
SVM (C=10)	78.27	83.648	80.450	83.848

TABLE 14 Accuracy Results with The Feature Selection Method Using SVM (C=1.0) on the Fitness Function.

Classification	BGWO	BPSO	BGWO-PSO	BGWO-PSO-SCA
KNN (k=5)	60.828	69.985	63.517	65.116
SVM (C=1.0)	86.046	85.247	83.866	84.883
SVM (C=10)	83.720	86.046	82.340	83.648

TABLE 15 Accuracy Results with the Feature Selection Method Using SVM (C=10) on the fitness function.

Classification	BGWO	BPSO	BGWO-PSO	BGWO-PSO-SCA
KNN (k=5)	60.537	50.581	64.680	68.313
SVM (C=1.0)	82.848	82.558	82.412	83.793
SVM (C=10)	84.956	85.901	85.537	88.081

TABLE 16 The Highest Accuracy Value for Each Change in the Classification Method on the Fitness Function.

Feature Selection	Fitness Function	Classification	Accuracy (%)
BGWO-PSO-SCA	KNN (k=5)	SVM (C=10)	83.848
BGWO	SVM (C=1.0)	SVM (C=1.0)	86.046
BGWO-PSO-SCA	SVM (C=10)	SVM (C=10)	88.081

selected features than just the combination of GWO and PSO. In this study, the BGWO-PSO algorithm proposed by Chantar et al.[1] is used for feature selection of the Arabic text of Hadith and inspired by Singh and Singh[9] by combining GWO with SCA to find the best global solution that aims to improve and speed up the global search.

A trial using the feature selection method proposed in this study is BGWO-PSO-SCA. In the trial, the researchers used two different amounts of data, including small imbalance data and extensive imbalance data. Each trial was carried out by changing the evaluation method on the fitness function and classification method using SVM with parameters C=1.0 and C=10 and using KNN with parameter k=5. The trial using small imbalance data showed that the highest accuracy value was 85.433% in BGWO-PSO-SCA with fitness function and SVM classification method (C=10), established in Table 9 . Tables 3 to 5 are the results of the time required and the number of feature selections in the feature selection method process. The experiment was carried out to overcome the global search, which has a broad scope, so that the direction and speed are improved by using the feature selection algorithm of SCAs. The time required by combining the SCA feature selection method is relatively low compared to other feature selection methods. More features are selected than BGWO-PSO.

Trials using a large imbalance dataset are shown in Table 13 to Table 15 . Table 16 shows that with a significant imbalance dataset, the BGWO-PSO-SCA feature selection method, the accuracy value obtained is 88.08%. The test results of the BGWO-PSO-SCA feature selection method are the highest accuracy values compared to other feature selections and the accuracy values without feature selection. The BGWO-PSO-SCA feature selection method gets the highest accuracy value, using the SVM classification method (C=10) and the fitness function with SVM (C=10). It can be seen in Table 10 to Table 12 that the combined objective of SCA is achieved by proving that the time required is faster and the selected features are more than BGWO-PSO but not more than PSO.

6 | CONCLUSION

This study successfully implemented the Hybrid Binary Gray Wolf optimizer for Arabic text classification, using Arabic text of Hadith data. The proposed feature selection method is BGWO-PSO, combining SCA capabilities according to the objectives to increase the direction of motion and speed in global search or exploration. In this case, it is proven by the number of selected features more than BGWO-PSO, but not more than with PSO, and the speed is increased compared to other feature selection algorithms. The feature selection method in the GWO has a fitness function value used to evaluate finding the best position for the grey wolf. This study uses the SVM classification method to assess the selected features to produce higher accuracy than KNN.

The results showed that feature selection in the Arabic text classification of Hadith resulted in better accuracy scores. BGWO-PSO with fitness function and SVM classification (C=1.0) gives a higher accuracy value of 76.37% compared to without feature selection. The BGWO-PSO-SCA feature selection method provides an accuracy value of 85.43% for small imbalanced data. This accuracy value is higher than the BGWO-PSO and other feature selection methods. The BGWO-PSO-SCA feature selection method on the large imbalance dataset gets the highest accuracy of the other feature selection methods, which is 88.081%. Regarding the feature selection method for classifying Arabic texts of Hadith, this study concludes that the BGWO-PSO-SCA feature selection method can improve both local and global search ability.

CREDIT

Muhammad Bahrul Subkhi: Data Curation, Writing – Original Draft, Validation, and Investigation. **Chastine Fatichah:** Conceptualization, Methodology, Formal analysis, Writing – Review & Editing, and Supervision. **Agus Zainal Arifin:** Conceptualization, Methodology, Formal analysis, Writing – Review & Editing, and Supervision.

References

1. Chantar H, Mafarja M, Alsawalqah H, Heidari AA, Aljarah I, Faris H. Feature selection using binary grey wolf optimizer with elite-based crossover for Arabic text classification. Neural Computing and Applications 2020 8;32(16):12201–12220.
2. Salton G, Buckley C. Term-weighting Approaches in Automatic Text Retrieval. Information Processing & Management 1988;24(5):513–523.
3. Singh N, Singh SB. Hybrid Algorithm of Particle Swarm Optimization and Grey Wolf Optimizer for Improving Convergence Performance. Journal of Applied Mathematics 2017;2017.
4. Mirjalili S, Mirjalili SM, Lewis A. Grey Wolf Optimizer. Advances in Engineering Software 2014;69:46–61.
5. Kohli M, Arora S. Chaotic grey wolf optimization algorithm for constrained optimization problems. Journal of Computational Design and Engineering 2018 10;5(4):458–472.
6. Ghareb AS, Bakar AA, Hamdan AR. Hybrid feature selection based on enhanced genetic algorithm for text categorization. Expert Systems with Applications 2016 5;49:31–47.
7. Al-Tashi Q, Kadir SJA, Rais HM, Mirjalili S, Alhussian H. Binary Optimization Using Hybrid Grey Wolf Optimization for Feature Selection. IEEE Access 2019;7:39496–39508.
8. Panwar LK, K SR, Verma A, Panigrahi BK, Kumar R. Binary Grey Wolf Optimizer for large scale unit commitment problem. Swarm and Evolutionary Computation 2018 2;38:251–266.
9. Singh N, Singh SB. A novel hybrid GWO-SCA approach for optimization problems. Engineering Science and Technology, an International Journal 2017 12;20(6):1586–1601.
10. Sahu PC, Prusty RC, Panda S. Approaching hybridized GWO-SCA based type-II fuzzy controller in AGC of diverse energy source multi area power system. Journal of King Saud University - Engineering Sciences 2020 3;32(3):186–197.
11. Fauzi MA, Arifin AZ, Yuniarti A. Arabic book retrieval using class and book index based term weighting. International Journal of Electrical and Computer Engineering 2017 12;7(6):3705–3710.

12. Emary E, Zawbaa HM, Hassanien AE. Binary grey wolf optimization approaches for feature selection. Neurocomputing 2016 1;172:371–381.

13. Chuang LY, Chang HW, Tu CJ, Yang CH. Improved binary PSO for feature selection using gene expression data. Computational Biology and Chemistry 2008;32(1):29–38.

14. Mirjalili S. SCA: A Sine Cosine Algorithm for solving optimization problems. Knowledge-Based Systems 2016 3;96:120–133.

15. Ariyanto ADP. Deteksi Interelasi Antar Kitab Hadis Menggunakan Word Embedding dan Ensemble Learning. PhD thesis, Master Programme in Computer Science, Institut Teknologi Sepuluh Nopember; 2022.

How to cite this article: Subkhi M.B., Faticha C., Arifin A.Z. (2022), Feature Selection Using Hybrid Binary Grey Wolf Optimizer for Arabic Text Classification, *IPTEK The Journal of Technology and Science*, *33(2):105-116*.