How to Cite:
Khan, A. A. R., Nisha, S. S., & Sathik, M. M. (2022). Hybrid slime mould - Grey wolf optimization algorithm for efficient feature selection. *International Journal of Health Sciences*, *6*(S1), 7657–7663. https://doi.org/10.53730/ijhs.v6nS1.6667

Hybrid slime mould - Grey wolf optimization algorithm for efficient feature selection

A. Ameer Rashed Khan
Research Scholar, Research Department of Computer Science, Sadakathullah Appa College, Tirunelveli, India
Corresponding author email: ameerkhan.a1694@gmail.com

S. Shajun Nisha
Assistant Professor, Research Department of Computer Science, Sadakathullah Appa College, Tirunelveli, India

M. Mohamed Sathik
Assistant Professor, Research Department of Computer Science, Sadakathullah Appa College, Tirunelveli, India

***Abstract*---**Selection of features is an effective method for minimizing the amount of data features in order to improve machine learning classification performance. Choosing a set of attributes is a high-level procedure for selecting a collection of relevant features. To boost the classifier's performance, use a dimensional dataset. We outline a typical feature selection problem in this work in order to minimize the amount of role and responsibilities while improving accuracy. Different classification dataset from the Machine learning repository have been used to test SMA with GWO as a feature selection strategy. The feature selections for UCI repository datasets include Bat optimization, Cuckoo search optimization, Slime mould optimization, Whale optimization, Particle swarm optimization, and Grey wolf optimization.

***Keywords*---**Slime Mould Algorithm, Grey Wolf Optimization, Feature selection, Particle swarm optimization.

Introduction

Problems with working optimization Meta - heuristic optimization techniques are capable of achieving optimal or near-optimal results in a reasonable amount of time by chance the fashionable value of the given variables that satisfies the highest or marginal intrinsic worth without violating the constraints. With the continuous advancement of intelligent machines, Metaheuristic optimization

Manuscript submitted: 18 Feb 2022, Manuscript revised: 09 March 2022, Accepted for publication: 27 April 2022

techniques are capable of achieving optimal or near-optimal results in a reasonable amount of time. As a result, they're commonly employed to solve difficulties including charge planning, picture separation, feature selecting, and parameter optimization. Modelling physical miracles or biological conditioning in nature helps metaheuristic algorithms identify optimal outcomes. Artificial intelligence (AI) has grown into a vital tool for addressing real-world engineering difficulties. Optimization techniques are a sort of machine learning that is gaining traction in a wide range of industries and societies. Metaheuristic algorithms are classified by physical occurrence, human learning patterns, natural evolution law, animal living behaviours, and swarming. The SMA method [24] is a recent metaheuristic algorithm that observes humanity's greatest sticky mould methodology for global optimization using a stochastic strategy. It offers a number of advantages, including a small number of parameters, robustness, great capacitating exploratory, and the ability to leverage inclination. In SMA, there are very few variables to consider. Metaheuristic algorithms were applied in many AI applications, such as health care.

Related Work

Data mining is the fastest-growing sector in information technology, owing to the huge amounts of data gathered every day and the need to turn that data into valuable knowledge [1]. Pre-processing steps in data mining include incorporation, transformation, reduction, knowledge display, and pattern evaluation [2]. Feature selection is one of the most important pre-processing procedures, and it tries to eliminate unnecessary and duplicated attributes from a dataset. In general, feature selection algorithms are divided into two categories: wrapper techniques [3], and direct approaches [4].

Methods that are not dependent on classifications and work exclusively on date belong to the first category. The discipline of optimization has recently gotten a lot of interest from academics, notably in the hybrid metaheuristics sector [5]. For example, the _rst hybrid met heuristic feature selection technique was introduced in 2004 [6,7,8] employing information retrieval methodologies and or the GA algorithm. PSO has been used with different meta - heuristic algorithms addressing continuous search space issues in the literature. For example, in [9],[10], a composite PSO with GA (PSOGA) was developed. Other related efforts include a PSO with DE (PSODE) [11], a hybrids PSO and Centrifugal Search Algorithm (GSA) (PSOGSA) [27], and a hybrids PSO and Geophysical Search Algorithm (GSA) (PSOGSA). Furthermore, in [12], PSO was combined with the Bacterial Foraging Optimization method to improve power system stability. These hybrid techniques are designed to complement one other's strengths in order to increase exploitation potential while lowering the risk of falling into a local optimum [13][14]. GWO has also gotten a lot of interest in the hybrids meta - heuristic algorithms field [16].

Methodology

Yang proposed the Bat algorithm (BA), which is a conceptual optimization technique. The sound waves behaviour of micro bats inspired this method. Each pulse in echolocation lasts only a few hundredths of a millisecond (up to roughly

8–10 ms). Li et al. suggested SMA in 2020, based on the natural foraging and moving behaviour of slime mould. Silicone mould seems to be a sub amoeba with no brain or nervous system. Its foraging behaviour more effective as a result of its veins. It may acquire the largest concentration of food using the oscillating modelling of veins, and this food location offers the best answer to optimization algorithm problems [17]. The effectiveness of SMA has indeed been proven, and it outperforms a number of well-known and cutting-edge techniques. The construction of SMA is basic and adaptable. It is simple to increase the performance of this flexible structure [18], [19].

The main idea of SMA is to take advantage of slime mold's foraging behaviour. Slime mould can approximate increased food, and slime mould can approximate high-quality food. Grey wolf optimization, evolutionary computation, bat improvement, particle swarm optimization improvement, lightning improvement, whale improvement, particle swarm optimization [20]. The following is the slime mould algorithm's position update:

Step 1: Begin
Step 2: Determine the size of the population
Step 3: Set the position of the slime moulds.
Step 4: Calculate the efficiency of all slime moulds.
Step 5: Sort the fitness features.
Step 6: Make sure the fitness feature is up to date.
Step 7: Determine the optimal location and fitness function.
Step 8: Finish

$$x_i(t+1) = r_1 \cdot (UB - LB) + LB \quad □□□ \quad r_2 < z \tag{1}$$

$$w_1 x_b + w_2 x_\beta + w_3 x_\gamma + v_b \cdot [w \cdot x_A(t) - x_B(t)] \quad □□□ \quad r_3 < p \tag{2}$$

$$v_c \cdot x_i\,(t) \quad □□□ \quad p \le r_3 \le 1 \tag{3}$$

(4)

$$x_i = \frac{x_\alpha + x_\beta + x_\gamma}{3} \quad □□□ \tag{5}$$

$$\varphi = \frac{1}{2} atan \quad □□□ \tag{6}$$

$$\emptyset = \frac{2}{\pi} acos \frac{1}{3} \cdot \mathrm{atan}(t) \quad □□□$$

$$w_1 = cos\theta, w_2 = \frac{1}{2} sin\theta \cdot cos\varphi, w_3 = 1 - w_3 - w_2 \quad □□□ \tag{7}$$

(i) Initializing Procedure

$$x_i = r_1(UB - LB) + LB \quad □□□$$

$r_1 \rightarrow random\ No\ in\ gauss\ distribution$

(ii) Iteration

$$x(t+1) = \begin{cases} r_2 \cdot (UB - LB) + LB & r_2 < z \\ x_b + v_b \cdot [w \cdot x_a(t) - x_b(t)] & r_3 < p \\ v_c \cdot x_i(t) & p \leq r_3 \leq 1 \end{cases}$$

An enhanced SMA technique with orthogonal learning capabilities is developed based on the original SMA technique. To increase the performance of the modified SMA, a chaotic initialization method and a boundary reset technique are implemented at the same time.

(iii) Based on the experimental results, the modified SMA and 10 sophisticated met heuristic algorithms, including the original SMA, are used to solve the issue, and overall performance in comparison and analysed.

(iv) To tackle the model estimation problem, the modified SMA method is utilised, and the simulated performance is evaluated to 10 other sophisticated techniques. The suggested algorithm's practicability and efficiency have also been demonstrated. By combining the GWO and SMA algorithms, we may substitute the best candidate with the averaging of alpha, beta, and delta prospects, and so the automatic update equation for the hybrid SMA-GWO strategy would've been the refreshing equations for individuals in a grey wolf swarm.

GWO Algorithm

When creating GWO, we use the completing as the alpha (α) to quantitatively describe the social structure of wolves. Consequently, the second or third best options have been identified named beta (β) and delta (δ) respectively.

Begin
Create the maned wolf population Xi ($i = 1, 2, \ldots, n$) from scratch
Create the variables a, A, and C
Calculate each search agent's fitness
 the most effective search agent
 the runner-up in the quest for a new job
 the third most effective search agent
$while$ (t) is the maximum number of iterations
for every search engine
Update the current search agent's location as described above equations
end for
$a, A, and\ C$ should all be updated
 Calculate the search agents' fitness
Update $X_\alpha, X_\beta, X_\delta$
  $t = t + 1$
end while
return
End

Table 1
Dataset description

Dataset	No. of Instances	No. of Features	No. of Classes
Diabetics	768	9	2

Dataset	No. of Instances	No. of Features	No. of Classes
Alzheimer	195	12	2
Heart	270	15	2
Bank Note	1372	4	2
Liver	1000	11	2
Zoo	101	19	7
Breast Cancer	699	11	2

Table 1. Dataset, Table 2 shows the different Evaluation Measures (best, worse, average, standard deviation, computational time) from the UCI Machine Learning repository. Figure 1 illustrates the different fitness performance indicators.

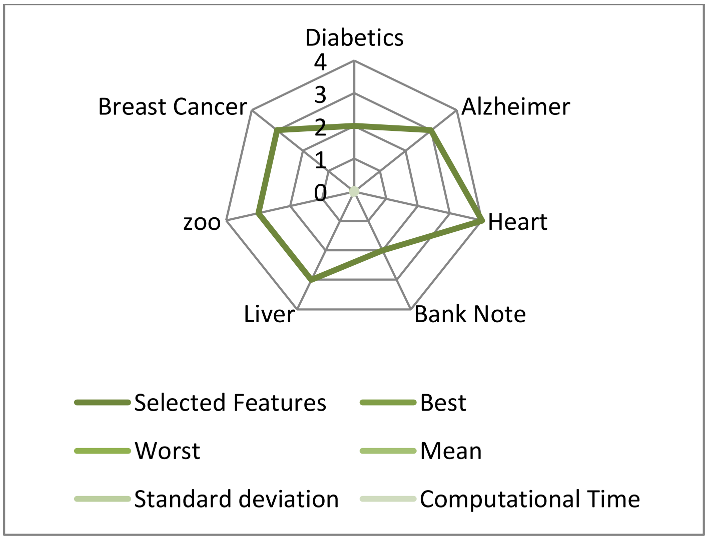

Fig. 1. Fitness Performance measures

Table 2
Evaluation Measures

Data set	Selected Features	Best	Worst	Mean	Standard deviation	Computational Time
Diabetics	2	0.0145	0.0395	0.0253	0.0027	0.0875
Alzheimer	3	0.0158	0.0375	0.0278	0.0038	0.0845
Heart	4	0.0152	0.0372	0.0295	0.0037	0.0796
Bank Note	2	0.0189	0.0394	0.0256	0.0039	0.0657
Liver	3	0.0197	0.0428	0.0256	0.0038	0.0845
Zoo	3	0.0125	0.0514	0.0371	0.0035	0.0687
Breast Cancer	3	0.0176	0.0645	0.0418	0.0037	0.0482

Conclusion

On benchmark functions, the proposed approach is compared to various recent met heuristic and high-performance algorithms. In terms of statistics, the results demonstrate that performs better than its competitors. It is capable of providing the best answer for the majority of difficulties. On the preponderance of datasets, the results showed that the suggested technique outperformed a wide variety of methods in terms of effectiveness and number of functions picked. Furthermore, a computational time comparison was made between the suggested binary hybrid technique and hybrid SMA, with GWO findings revealing that perhaps the propose method has a faster execution time. Furthermore, a comparison of statistical studies Mean, Best, and Worst Fitness was done, with the findings demonstrating that the suggested methodology outperformed existing state-of-the-art methodologies significantly. The suggested method's better findings show that it can regulate the trade-off between exploratory and exploitative behaviour throughout optimization rounds.

References

[1] Han, J., Pei, J., Kamber, M., “Data Mining: Concepts and Techniques. Amsterdam,” The Netherlands: Elsevier, 2011.

[2] Liu, H., Motoda, H.,“Feature Selection for Knowledge Discovery and Data Mining,” Springer, vol. 454, 2012.

[3] Dash, M., Liu, H.. “Feature selection for classification,” Intell. Data Anal., vol. 1, pp. 131-156, 1997.

[4] Guyon, I., Elisseeff, A.,”An introduction to variable and feature selection,” J. Mach. Learn. Res., vol. 3, pp. 1157-1182, 2003.

[5] Liu,H., Zhao, Z.,“Manipulating data and dimension reduction methods: Feature selection,” Encyclopedia Complexity Systems Science. New York, NY, USA: Springer, pp. 5348-5359, 2009.

[6] Liu, H., Motoda, H., Setiono, R., Zhao, Z.,“Feature selection: An ever-evolving frontier in data mining,” In Proc. 4th Workshop Feature Selection Data Mining, pp. 4-13, 2010.

[7] Zarshenas, A., Suzuki, K.,“Binary coordinate ascent: An efficient optimization technique for feature subset selection for machine learning,” Knowl-Based Syst., vol. 110, pp. 191-201, 2016.

[8] Talbi, E.G.,“Metaheuristics: From Design to Implementation,” Hoboken, NJ, USA: Wiley, vol. 79, 2009.

[9] Lai, C., Reinders, M. J. T., Wessels, L.,“Random subspace method for multivariate feature selection,” Pattern Recognit. Lett., vol. 27, pp. 1067-1076, 2006.

[10] Liu, H., Motoda, H., “Feature Extraction, Construction and Selection: A Data Mining Perspective,” Springer, vol. 453, 1998.

[11] Emary, E., Zawbaa,H. M., Hassanien, A. E.,“Binary grey wolf optimization approaches for feature selection,” Neuro computing, vol. 172, pp,. 371-381, 2016.

[12] Al-Tashi, Q., Rais, H., Jadid, S.,“Feature selection method based on grey wolf optimization for coronary artery disease classification,” In Proc. Int. Conf. Reliable Inf. Commun. Technol., pp. 257-266, 2018.

[13] Kabir, M. M., Shahjahan, M., Murase, K.,“A new local search-based hybrid genetic algorithm for feature selection,” Neuro computing, vol. 74, pp. 2914-2928, 2011.

[14] Kashef, S., Nezamabadipour,H.,“An advanced ACO algorithm for feature subset selection,'' Neurocomputing, vol. 147, pp. 271-279, 2015.

[15] Bello,R.,Gomez,Y., Nowe, A.,Garcia, M. M.,“Two-step particle swarm optimization to solve the feature selection problem,” In Proc. 7th Int. Conf. Intell. Syst. Des. Appl. (ISDA). Pp. 691-696, 2007.

[16] ZorarpacI, E., Özel, S. A., “A hybrid approach of differential evolution and artificial bee colony for feature selection,” Expert Syst. Appl., vol. 62, pp. 91-103, 2016.

[17] Mafarja,M. M., Eleyan, D.,Jaber, I.,Hammouri, A.,Mirjalili, S.,“Binary dragonfiy algorithm for feature selection,” In Proc. Int. Conf. New Trends Comput. Sci. (ICTCS), pp. 12-17, 2017.

[18] Mafarja, M. M., Mirjalili, S.,“Hybrid whale optimization algorithm with simulated annealing for feature selection,” Neurocomputing, vol. 260, pp. 302-312, 2017.

[19] Singh, N.,Singh, S. B., “Hybrid algorithm of particle swarm optimization and Grey Wolf optimizer for improving convergence performance,” J. Appl. Math., 2017.

[20] Kennedy J., Eberhart, R.,“Particle swarm optimization,” In Proc. ICNN Int. Conf. Neural Netw., pp. 1942-1948, 1995.

[21] Kennedy, J.,“The particle swarm: Social adaptation of knowledge,” In Proc. IEEE Int. Conf. Evol. Comput. (ICEC), pp. 303-308, 1997.

[22] Mirjalili, S., Lewis, A.,“Grey Wolf Optimizer,” Adv. Eng. Softw., vol. 69, pp. 46-61, 2014.

[23] Faris, H.,Aljarah, I.,Al-Betar,M. A., Mirjalili,S.,“Grey wolf optimizer: A review of recent variants and applications,” Neural Comput. Appl., vol. 30, pp. 413-435, 2018.

[24] Oh, I.S.,Lee, J.S., Moon, B.R.,“Hybrid genetic algorithms for feature selection,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, pp. 1424-1437, 2004.