

Efficient Feature Selection Using Hybrid Slime Mould- Grey Wolf Optimization

Dr. S. Shajun Nisha¹, A. Ameer Rashed Khan² and Dr. M. Mohamed Sathik³

¹ Assistant Professor & Head, PG & Research Department of Computer Science,
SadakathullahAppa College, RahmathNagar,Tirunelveli,India

² Reg No.18211192161021, Ph.D Research Scholar, PG & Research Department of Computer Science, SadakathullahAppa College, RahmathNagar, Tirunelveli, India. Affiliation of Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli 627012. ameerkhan.a1694@gmail.com

³ Principal, SadakathullahAppa College, RahmathNagar, Tirunelveli, India.

Article Info

Page Number: 10492 - 10499

Publication Issue:

Vol 71 No. 4 (2022)

Abstract

Feature selection becomes a prominent approach, especially when the records sets incorporate multiple variables and functions. It is the process of reducing the input data into an essential model, by disposing the unimportant variables and enhances the accuracy as well as the performance of type. In this paper Hybrid slime mould- grey wolf optimization algorithm is proposed for efficient feature selection by incorporating set of rules which could deal with the classical feature selection short comings. This algorithm is tested over prominent datasets with higher variety of distinct variables such as Diabetics, Alzheimer, Heart, Liver, Zoo, Breast Cancer. Four essential characteristics which makes feature selection is essential are; to simplify the model by way of lowering the range of parameters, subsequent to lower the training time, to lessen overfilling by using improving generalization, and to keep away from adding extra dimensionality. The proposed algorithm is compared with the state-of-the-art techniques Naïve Base (NB), support Vector Machines (SVM), K Nearest Neighbors (KNN), the best accuracy of the version is the exceptional classifier. Our experiments show case the comparative examine at distinct views. Furthermore, critical evaluation metrics Accuracy, Precision, Recall, F-Measure, Time, RMSE, MAE are used to evaluate the performance. Experimental consequences exhibits that SVM achieves a higher performance in all test corporations.

Keywords: Hybrid Slime Mould- Grey Wolf Optimization, KNN, SVM, NB.

Article History

Article Received: 15 September 2022

Revised: 25 October 2022

Accepted: 14 November 2022

Publication: 21 December 2022

1. INTRODUCTION

Feature selection issue becomes a concern in real world. In order to solve these issues, the researchers propose several algorithms which can solve the problems by giving better performances. And these algorithms mimic the animals or insects' behavior in nature in the process of finding food. The feature selection methods are broadly divided into three groups that are wrapper based; filter based and embedded based models. Filter based methods use ranking technique which ranks the features. A

filtering based approach provides good generalization ability and low computational complexity. Algorithm that absence of local minima, it depends on few parameters and an adequate generalization to new objects[1, 2], SVM has seen its prosperity and SVM techniques has been applied to many fields for the past 20 years. It has exerted an indispensable role in pattern recognition[3], disease diagnosis[4], forecasting[5], etc. while wrapper methods[6] In the other side, the wrapper method achieve superior classification accuracy, but requires more computational power. The drawbacks and complementarily of the two methods lead to the development of the hybrid method, such as the SAGA[7] and the normalized mutual information feature selection method which used a genetic algorithm to form a hybrid method called GAMIFS[8] and Wrapper feature selection using a Fuzzy Random Forest ensemble based on cross-validation[9] Researchers have studied various types of feature selection. The first key method is to measure the goodness of a feature subset determining an optimal one. Its depending on evaluation criteria, feature selection methods can be divided into the following categories, Filter methods, Wrapper Method, Embedded Method, and Hybrid method [10, 11, 12], Hybrid search technique combines the wrapper and later approaches. Another algorithm is proposed in [13] Bio-inspired [14] According to Yu and Liu [15] important and relevant features are segregated into necessary and unnecessary features. ReliefF and its variant feature-selection algorithms are used in the binary classification that Kira and Rendell proposed in 1992 [16], IBM [17] defined “big-data analytics” techniques that can handle datasets from large and diverse backgrounds and multiple types. It does not matter whether it is structured and unstructured or streaming and varies according to sizes. In this research paper the feature selection method will be discussed[18], In 2018, Singh and Hachimi [19] integrated the spiral equation of WOA into GWO and apply it to structural design problems. A similar method proposed in [20] and applied to data clustering algorithms. Despite the merits of this hybrid, the authors used it for continuous optimization; hence, it is not able to deal with problems with binary variables. In this paper, FS contributes to enhancing generalization by reducing over-fitting, computation time, and storage, enhancing the classification accuracy by creating an accurate predictive model among other things [21,22], Ashraf Darwish, whale optimizations algorithm [23]. The graphical flow of the proposed work is shown in figure 1.

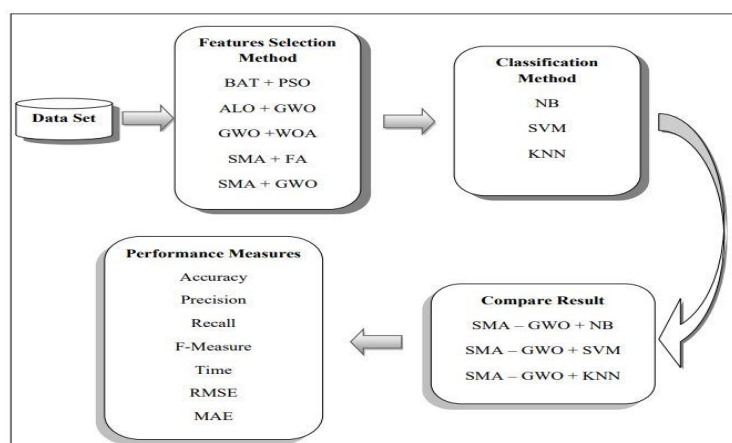


Figure 1. Flow of the proposed work

2. METHODOLOGY

2.1: HYBRID SLIME MOULD- GREY WOLF OPTIMIZATION

In system getting to know we're regularly interested in selecting the pleasant hypothesis (h) given facts (d). In a category hassle, our hypothesis (h) can be the class to assign for a brand new facts example (d). Bayes' Theorem gives a way that we will calculate the possibility of a hypothesis.

Bayes' Theorem is said as:

$$P(h | d) = (P(d | h) * P(h)) / P(d)$$

After calculating the posterior possibility for a number of extraordinary hypotheses, you may pick out the hypothesis with the very best chance. This is the maximum possibly speculation and can officially be referred to as the most a posteriori (MAP) hypothesis. This may be written as:

$$MAP(h) = \max(P(h | d))$$

or

$$MAP(h) = \max((P(d | h) * P(h)) / P(d))$$

or

$$MAP(h) = \max(P(d | h) * P(h))$$

The $P(d)$ is a normalizing term which allows us to calculate the possibility.

KNN (K nearest neighbor) algorithm is popularly acknowledged for numerous programs like genetics, forecasting, and so forth. The procedure is as follows:

Step 1: Pick the fee of ok associates ($k=n$)

Step 2: Discover the ok (n) nearest facts point for our new information point primarily based on Euclidean distance (which we talk later)

Step 3: Among these ok records factors remember the information points in every category. The distance metric we are using is Minkowski, the equation for it is given below

$$\left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p}$$

As per the equation, we have to select the p-value also.

$P = 1$ Manhattan Distance

$p = 2$ Euclidean Distance

$p = \text{infinity}$, Cheybchev Distance

3 RESULTS AND DISCUSSION

The datasets are randomly partitioned into three various equal quantities (e.g., validation, education, and trying out datasets). The following statistical measures are

examined from the validation records in every run the mean fitness function is a hallmark to the average fee of the fitness feature gained when set of rules run N instances, and it is calculated as follows:

$$mean = \frac{1}{N} \sum_{k=1}^N g_k^*$$

The imply fitness value won at run k. The best fitness feature is an indicator to the minimal value of fitness function won when set of rules run N times, and it is calculated as follows:

$$Best = \min_k g_k^*$$

In which g k the first-rate fitness fee won at run k. (five) the worst fitness feature Is an indicator to the maximum value of the fitness function when set of rules run N times, and it is calculated as follows:

Worst D max

$$Worst = \max_k g_k^*$$

Where g k the worst fitness (maximum) price received at run.

Table1. Number of selected features obtained by different algorithms

Data Set	FS Algorithm	Selected Features	Best	Worst	Mean	Standard deviation	Computational Time
Data Set 1 (Diabetics) (9)	BAT-PSO	4	0.0195	0.1971	0.0677	0.0051	0.6741
	ALO-GWO	4	0.0183	0.0678	0.0331	0.0045	0.2471
	GWO-WOA	4	0.0176	0.0478	0.0335	0.0031	0.5213
	SMA-FA	3	0.0151	0.0476	0.0299	0.0028	0.1681
	SMA+GWO	2	0.0145	0.0395	0.0253	0.0027	0.0875
Data Set 2 (Alzheimer) (12)	BAT-PSO	4	0.0578	0.0761	0.0632	0.0067	2.1231
	ALO-GWO	3	0.0411	0.0782	0.0514	0.0059	1.4912
	GWO-WOA	4	0.0273	0.0398	0.0304	0.0045	1.0102
	SMA-FA	4	0.0269	0.0317	0.0293	0.0041	0.8079
	SMA+GWO	3	0.0158	0.0375	0.0278	0.0038	0.0845
	BAT-PSO	5	0.0246	0.037	0.027	0.0041	1.3124

Data Set 3 (Heart) (15)				5	9		
	ALO-GWO	3	0.0218	0.028 9	0.024 1	0.0040	0.4147
	GWO- WOA	4	0.0178	0.021 0	0.019 3	0.0038	1.0245
	SMA-FA	4	0.0168	0.048 1	0.034 2	0.0041	0.0865
	SMA+GW O	4	0.0152	0.037 2	0.029 5	0.0037	0.0796
Data Set 4 (Bank Note) (4)	BAT-PSO	4	0.0781	0.103 2	0.087 1	0.0072	3.3790
	ALO-GWO	3	0.3174	0.841 0	0.605 1	0.0058	1.5020
	GWO- WOA	2	0.2274	0.491 2	0.367 2	0.0049	0.3478
	SMA-FA	2	0.0462 7	0.051 0	0.048 8	0.0048	0.6871
	SMA+GW O	2	0.0189	0.039 4	0.025 6	0.0039	0.0657
Data Set 5 (Liver) (11)	BAT-PSO	4	0.7112	1.027 1	0.862 3	0.0091	1.0011
	ALO-GWO	3	0.3671	0.674 8	0.451 6	0.0075	0.8164
	GWO- WOA	2	0.1147	1.099 7	0.710 6	0.0071	1.0178
	SMA-FA	2	0.0354	0.091 4	0.067 4	0.0063	0.5173
	SMA+GW O	3	0.0197	0.042 8	0.025 6	0.0038	0.0845
Data Set 6 (Zoo) (19)	BAT-PSO	8	0.1078	0.378 4	0.251 7	0.0092	0.6807
	ALO-GWO	8	0.7088	2.564 1	1.845 1	0.1581	1.9710
	GWO- WOA	5	0.6710	1.317 9	0.971 8	0.0080	0.2713
	SMA-FA	4	0.0773	0.247 8	0.175 3	0.0053	0.2283
	SMA+GW O	3	0.0125	0.051 4	0.037 1	0.0035	0.0687
Data Set 7 (Breast Cancer) (11)	BAT-PSO	6	2.1371	3.518	2.732 5	0.1107	1.9700
	ALO-GWO	5	0.9747	1.102 8	0.990 7	0.8705	1.2157
	GWO- WOA	4	0.4723	0.787 5	0.596 1	0.0605	0.6517
	SMA-FA	4	0.0578	0.256 8	0.197 5	0.0108	0.1786

	SMA+GW O	3	0.0176	0.064 5	0.041 8	0.0037	0.0482
--	-------------	---	--------	------------	------------	--------	--------

Table 1 shows the number of selected features of the data sets obtained by the different algorithms respectively. It can be seen from the tables that proposed SMA+GWO obtains the best average number of selected features in the data sets which can be regarded as the best results in the test compared to other optimization algorithms. Table 2 depicts the Classification performance comparison of datasets in terms of metrics.

Table 2 Classification performance comparison (SMA+GWO)

Data Set	CLASSIFIER	Accuracy	Precision	Recall	F-Measure	Time	RMSE	MAE
Alzheimer's	NB	0.8682	0.9721	0.8345	0.7889	0.0165	5.5321	1.2214
	SVM	0.9797	0.9681	0.9565	0.9614	0.0156	5.5578	0.9637
	KNN	0.9884	0.9761	0.9764	0.9664	0.0123	5.4902	0.8779
Banknote	NB	0.8317	0.8434	0.8569	0.8553	0.0273	0.4048	0.1675
	SVM	0.9879	0.9965	0.9868	0.9842	0.0189	0.0431	0.0127
	KNN	0.9872	0.9986	0.9997	0.9965	0.0148	0.0454	0.0024
Breast Cancer	NB	0.8772	0.8697	0.8668	0.8543	0.1698	0.5828	0.3783
	SVM	0.9464	0.9547	0.9589	0.9136	0.1246	0.6043	0.2783
	KNN	0.9818	0.9768	0.9856	0.9487	0.0538	0.5357	0.2783
Diabetes	NB	0.7773	0.8054	0.8797	0.8451	0.0475	0.4745	0.2251
	SVM	0.6964	0.7768	0.7835	0.7754	0.0314	0.5505	0.3030
	KNN	0.9663	0.9824	0.9013	0.9452	0.0156	0.3879	0.1513
Heart	NB	0.9713	0.9784	0.9513	0.9742	0.0156	0.1483	0.1482
	SVM	0.9748	0.9635	0.9364	0.9651	0.0156	0.1367	0.1174
	KNN	0.9887	0.9812	0.9564	0.9754	0.0024	0.1139	0.0220
Liver	NB	0.8811	0.9181	0.9041	0.8954	0.0781	0.6459	0.4171
	SVM	0.9553	0.92354	0.9116	0.9641	0.0412	0.5904	0.3486
	KNN	0.9623	0.9854	0.9967	0.9823	0.0248	0.5503	0.3029
Zoo	NB	0.9312	0.9274	0.9345	0.9538	0.1472	1.6172	0.6923
	SVM	0.9463	0.9523	0.9453	0.9618	0.0289	1.5669	0.6239
	KNN	0.9848	0.9851	0.9745	0.9819	0.0154	1.3587	0.6154

The selected features from these hybrid feature selection approach is given as the input of the three classifiers like SVM, NB and KNN. Depending these analysis, the best classifier is chosen among them using varied performance metrics like accuracy, precision, recall, F-measure, RMSE, time and standard deviation. For each metrics, the KNN approach provides better results than SVM and NB methods. The SVM has overlapping issue hence it difficult to afford better classification results and also it takes large training time. Moreover, the SVM method cannot support large size of input data. Thus, these approaches attains reduced results as compared with KNN. KNN method is very easy to handle and it provides reduced error rate. Also KNN approach has the ability to process in large dimensionality of data. This makes the

system to achieve improved performance.

5. CONCLUSION

A hybrid version of slime mould optimization algorithm and grey wolf optimization was proposed and used to solve the problem of feature selection in this work. To confirm the effectiveness and the efficiency of the proposed method, 07 standard UCI benchmark datasets were employed. A set of evaluation measures were used to assess the proposed method. The proposed hybrid was compared with a number of existing hybrid optimization feature selection algorithms called BAT-PSO, ALO-GWO, GWO-WOA and SMA-FA. By the result it shows that SMA-GWO outperforms well while comparing to others and using classification method KNN it gives better accuracy than other two algorithms NB, SVM. The future work in this area can include designing a new nature-inspired algorithm which could have much more significant results than the established algorithms.

REFERENCE

- [1] V. N. Vapnik, "Statistical learning theory," 1998.
- [2] N. Christianini and S. J. Taylor, "An introduction to support vector machines (and othre kernel-based learning methods)," 2000.
- [3] A. Mueller, G. Candrian, V. A. Grane, J. D. Kropotov, V. A. Ponomarev, and G.-M. Baschera, "Discriminating between ADHD adults and controls using independent ERP components and a support vector machine: a validation study," *Nonlinear biomedical physics*, vol. 5, p. 5, 2011.
- [4] B. Magnin, L. Mesrob, S. Kinkingnéhun, M. Pélégriani-Issac, O. Colliot, M. Sarazin, et al., "Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI," *Neuroradiology*, vol. 51, pp. 73-83, 2009.
- [5] O. Kisi and M. Cimen, "A wavelet-support vector machine conjunction model for monthly streamflow forecasting," *Journal of Hydrology*, vol. 399, pp. 132-140, 2011.
- [6] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial intelligence*, vol. 97, pp. 273-324, 1997.
- [7] I. A. Gheyas and L. S. Smith, "Feature subset selection in large dimensionality domains," *Pattern Recognition*, vol. 43, pp. 5-13, 2010.
- [8] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized mutual information feature selection," *Neural Networks, IEEE Transactions on*, vol. 20, pp. 189-201, 2009.
- [9] L. Breiman, Random forests, *Machine Learning* 45 (1) (2001) 5{32}
- [10] S. M. Vieira, J. M. C. Sousa, U. Kaymak, Fuzzy criteria for feature selection, *Fuzzy set and System* 189 (1) (2012) 1{18.
- [11] Y. Saeys, I. Inza, P. Larra~naga, A review of feature selection techniques in bioinformatics, *Bioinform-matics* 23 (19) (2007) 2507{2517.
- [12] G. George, V. Raj, Review on feature selection tehcniques and the impact of svm for cancer classfication using gene expresion pro¯le, *International Journal of Computer Science & Engineering Survey* 2 (3) (2011) 16{27
- [13] S. M. Vieira, J. M. C. Sousa, U. Kaymak, Fuzzy criteria for feature selection, *Fuzzy set and System* 189 (1) (2012) 1{18.}
- [14] Xumei Fan, William Sayers, Shujun Zhang, Zhiwu Han, Luquan Ren, Hassan

Chizari, Review and Classification of Bio-inspired Algorithms and Their Applications, J Bionic Eng 17 (2020) 611–631

- [15] Yu. Lei, Huan Liu, Efficient feature selection via analysis of relevance and redundancy, J. Mach. Learn. Res. 5 (Oct. 2004) 1205–1224.
- [16] K. Kira, L.A. Rendell, The feature selection problem: traditional methods and a new algorithm, in: Aaai, 1992, pp. 129–134.
- [17] “What is Big Data Analytics?” [Online]. Available: <http://www-01.ibm.com/software/data/infosphere/hadoop/what-is-big-data-analytics.html>.
- [18] Jianyu Miao, Lingfeng Niub, A Survey on Feature Selection, Information Technology and Quantitative Management (ITQM 2016), Procedia Computer Science 91 (2016) 919 – 926
- [19]. Singh N, Hachimi H. A new hybrid whale optimizer algorithm with mean strategy of grey wolf optimizer for global optimization, Mathematical and Computational Applications. 2018;23(1):14.
- [20]. Jadhav AN, Gomathi N. Wgc: hybridization of exponential greywolf optimizer with whale optimization for data clustering. Alex, Eng J. 2017.
- [21]. Aljarah I, Ala’M AZ, Faris H, Hassonah MA, Mirjalili S, Saadeh H. Simultaneous feature selection and support vector machine optimization using the grasshopper optimization algorithm. Cogn Comput. 2018;10:1–18.
- [22]. Pang J, Zhao Y, Xu J, Gu Y, Yu G. Super-graph classification based on composite subgraph features and extreme learning machine. Cogn Comput. 2018;10:1–15.
- [23]. Ashraf Darwish, Bio-inspired computing: Algorithms review, deep analysis, and the scope of applications, Future Computing and Informatics Journal 3 (2018) 231-246