International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Hybrid Binary Gray Wolf Optimization for finding Optimal Features in Classification Problems

R.S.Latha, G.R.Sreekanth, R.C.Suganthe, M.Geetha

Abstract: ***Finding the essential symptoms(features) is highly demanded in the area of medical applications. Binary Gray wolf optimizer (BGWO) is one of the latest bio-inspired optimization techniques, which simulate the hunting process of gray wolves in nature. In this work, the binary gray wolf optimization (BGWO) is applied to select important feature subset for classification purposes and to attain maximum accuracy with minimum number of features using various classification algorithms and data sets. The classification error rate and the number of features are considered in the objective function. The wolf with low error rate and minimal features is considered as the best wolf and this kind of problem is a minimization problem. In BGWO, at each iteration the positions of three wolves (alpha wolf , beta wolf, delta wolf) are identified. All the wolves move toward these three wolves to find out the target position. If the position of best solution (alpha wolf) is stuck in local minimum, the genetic algorithm (GA) is employed to get rid of it. The Hybrid binary gray wolf optimization with genetic algorithm (HBGWO) is used for classification problems in finding out the optimal feature subset with maximizing the classification accuracy while minimizing the number of selected features. The proposed HBGWO is used with the classification algorithms such as Naive Bayes, K-Nearest Neighbour and Decision Tree for different medical datasets. Results proved that the capability proposed HBGWO improves classification accuracy over BGWO.***

Keywords : Binary Gray wolf optimization, feature selection, Hybrid, Nature inspired optimization, Bio-inspired optimization, classification

I. INTRODUCTION

Machine learning is a technology design used to build intelligent systems. These systems have the ability to learn from past experience or analyze historical data. It provides results according to its experience. Data is the key concept of machine learning. Machine learning technology involves statistics and computer science fields. Statistics enables to draw inferences from the given data. Machine learning algorithm to develop the new algorithms and techniques.In machine learning feature selection plays an important role.

Manuscript published on November 30, 2019.
* Correspondence Author
R.S.Latha*, Assistant Professor(Sr.Grade), Department of CSE, Kongu Engineering College, Perundurai, Tamilnadu, India. Email : rslatha20@gmail.com
G.R.Sreekanth, Assistant Professor(Sr.Grade), Department of CSE, Kongu Engineering College, Perundurai, Tamilnadu, India. Email: grsreekanth25@gmail.com
R.C.Suganthe, Professor, Department of CSE, Kongu Engineering College, Perundurai, Tamilnadu, India. Email : rc.suganthe@gmail.com
M.Geetha, Assistant Professor(Sr.Grade), Department of CSE, Kongu Engineering College, Perundurai, Tamilnadu, India. Email : geetha@kongu.ac.in

When a dataset has too many features, it would not be ideal correct to include all of them in machine learning model. Some features may be irrelevant in identification. Instead of processing data with whole features to the learning algorithm directly, it is necessary to identify the important features mainly used for classification .

In many classification problems, it is difficult to learn good classifiers before removing these unwanted features due to the huge size of the data. Reducing the number of irrelevant or redundant features can drastically improve the running time of the learning algorithms and yields a more general classifier. This helps in clear understanding of a real-world classification problem. Thus, feature selection mainly helps in the training the classifier.

Feature selection provides a way for identifying the important features by selecting a subset of features and removing irrelevant (redundant) ones from the dataset. The feature selection objectives are data dimensionality reduction, improving prediction performance, and good data understanding.Feature selection has three methods such as Wrapper method, Filter method and Embedded method. In the proposed work , wrapper method is applied for feature selection in classification problems. The feature selection phase might be independent of the learning algorithm, like filter models, or it may iteratively utilize the performance of the learning algorithms to evaluate the quality of the selected features, like wrapper models. Several classification algorithms such as naive bayes, decision tree and k-nearest neighbour are studied in the proposed HBGWO algorithm with various UCI repository datasets [5] to attain maximum accuracy with minimal features and the results are compared with BGWO.

II. RELATED WORK

The classical optimization techniques have some restriction in solving the problems, so that evolutionary computation (EC) algorithms are the alternative for solving these limitations and searching for the optimum solution of the problems. Evolutionary computation (EC) algorithms are inspired from nature, social behavior, and biological behavior of (animals, birds, bat, wolves, etc.) in a group. Many researchers have proposed different computational methods, in order to mimic the behavior of these species to seek for their food (optimal solution) as a particle with special characteristics (position, fitness, and a speed vector). The nature inspired optimization algorithms found a sound scope in various domains in computer science field [6,7,8].

Artificial Bee Colony (ABC) is a numerical optimization algorithm based on foraging behavior of honeybees is applied in finding optimal features in [1]. In ABC, the employer bees try to and food source and advertise the other bees. The onlooker bees follow their interesting employer, and the scout bee y spontaneously to and the best food source.

In [2], Particle Swarm Optimization (PSO) technique is employed in finding useful features. The PSO algorithm works with a population (called a swarm) of candidate solutions (called particles). These particles are moved around in the search space towards the best particle. The movement of each particle is guided by its own best known position (pbest) in the search-space as well as the entire swarm's best known position(gbest). The authors applied this algorithm for binary classification problems.

The authors in [3] provide an overview of some of the feature selection methods. They provide a generic introduction to variable elimination that can be applied to a wide range of machine learning problems. Several methods like Filter, Wrapper and Embedded and feature selection techniques are applied on standard datasets to show the applicability of feature selection techniques.

In [4], the authors present a variety of methods to solve feature selection problems. The authors suggest that though evolutionary computation (EC) techniques got much attention and results in better performance, there are no standard guidelines on the strengths and weaknesses of the various approaches. They provide a comprehensive survey of the several EC techniques for feature selection. They also discuss current issues and challenges to identify promising areas for future research.

The author applied Binary Cuckoo Search based on the behavior of cuckoo birds in finding optimal features [9]. They conducted the experiments in the context of theft detection in power distribution systems among the datasets obtained from a Brazilian electrical power company. They proved the robustness of their work against several others nature-inspired optimization techniques.

In [10], the authors proposed feature selection technique based on the bats behaviour. They used wrapper approach that combines bats optimization with the Optimum-Path Forest classifier to maximize the accuracy and conducted experiments over five datasets to demonstrate their approach outperform some existing swarm-based techniques.

In [11], the feature selection approach is proposed based on the integration of a genetic algorithm and particle swarm optimization. The support vector machine(SVM) classifier is used in the fitness function. The experiment was done on the popular Indian Pines hyperspectral data set. The authors proved that their approach automatically selects the most informative features in terms of classification accuracy.

III. GRAY WOLF OPTIMIZATION

Gray wolf optimization (GWO) is a recently developed evolutionary algorithm, which says that the gray wolves have a successful reproduction more than hunting in the pack. Two gray wolves (male and female) that are in a higher position manages the other wolves in the pack. Gray wolf optimizer is one of the latest bio-inspired techniques, which simulates the hunting process of a pack of gray wolves in nature. The gray wolves are often live in a pack. The size of a pack is usually 5 to 12. The wolves follow very stringent rules in social dominant hierarchy. There wolves in the pack are categorized as alpha wolf , beta wolf and delta wolf where alpha wolf makes all decisions, beta wolf helps alpha wolf to make decision and delta wolf moves to the more dominant wolf. According to the optimizer, the role of each wolf is given below :

The alphas are leading the pack, the alpha wolves are responsible for making decisions. The alphas decisions are dictated to the pack

The betas are subordinate wolves that help the alpha in decision making or other activities. The beta can be either male or female, and he/she is probably the best candidate to be the alpha.

The omega wolves play the role of scapegoat. They have to give to all the other superior wolves. They are the last wolves allowed to eat. The delta wolves have to submit alphas and betas, but they dominate the omega. Scouts, elders, sentinels, hunters, and caretakers are fall in this category. Scouts are watching the boundaries of the territory and alert the pack in case of any danger. Sentinels are protecting and also assure the safety of the pack. Elders are the most experienced wolves which are alpha or beta.

Hunters help the alphas and betas by hunting prey and supplying food for the pack. Finally, the care takers are responsible for caring for the weak, ill, and wounded wolves in the pack.The wolf attracts towards the first three best solutions. In the BGWO, the pool of solutions is in binary form at any given time then the position of wolf is updated and crossover is done and best solution wolf act as a leader for the pack of wolves.

IV. PROPOSED HYBRID BINARY GRAY WOLF OPTIMIZATION (HBGWO)

To identify the important feature subset in the given dataset, a hybrid binary gray wolf optimization algorithm with genetic algorithm is proposed. The HBGWO operators play an important role in improvising the solutions as far the algorithm is concerned.

The design of proposed work is given below:

A. Wolf representation:

Each wolf is represented as an array of binary values. The length of the array is based on the number of features present in the dataset except the class label. The value one (1) indicates the feature is selected and zero (0) indicates the feature is not selected. For example, the wolf represented as 8-bit number 10110001 indicates that only the features 1,3,4 and 8 are selected and the remaining are unselected.

B. Initial Population:

Initial population is generated with random values of 1's and 0's based on the number of wolves and the dimension of the search space (i.e.) number of features in the dataset. In the proposed HBWGO, we have taken 25 wolves with different dimensions.

Retrieval Number: D8251118419/2019©BEIESP
DOI:10.35940/ijrte.D8251.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

C. Fitness Function

Since the proposed HBGWO is a minimization optimization problem, the fitness function is coined to reduce the error rate and the number of optimal feature selected. The wolf resulting in minimum error rate with minimal feature subset is considered as the best wolf. For a wolf (W), the fitness value is calculated as

$$Fitness(W) = \alpha\big(1 - errorrate(W)\big) + \beta \frac{No.features\ selected\ in\ W}{Total\ No.of\ features\ in\ W}$$

where α, β are constant are set as 0.7 and 0.3.

D. Selection:

With the initial population considered, the fitness value for each wolf is computed using the fitness function. Using rand function random variables are generated and important features are selected to maximize the accuracy and minimize the number of selected features. The fitness value of every wolf differs because of the different set of features are selected.

E. Crossover:

This is a crucial and unique operation of the GWO technique. The purpose of this operation was to get optimal fitness valued wolf. As in GA, the crossover step is applied on the three best wolves (alpha, beta and delta) which resembles exploration phase in the optimization technique.

F. Mutation:

Another unique operation related to GA is the mutation operation. Mutation is performed when the fitness remains same for certain number of iterations. The mutation operation changes only a very few features and this exploitation phase is controlled by the mutation rate. In the proposed work, the mutation rate set as 0.1.

The figure 1 gives the flowchart of the proposed HBGWO approach for feature selection problem.

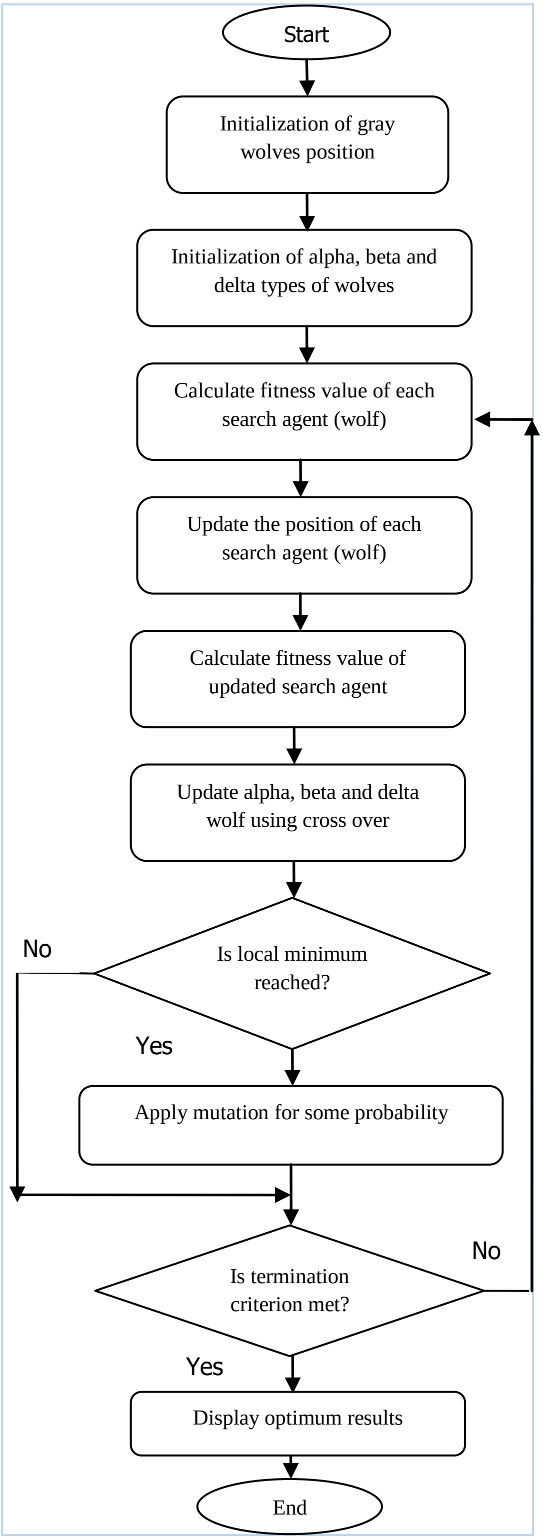

Figure 1 . Flowchart of proposed HBGWO

V. RESULTS AND DISCUSSION

In our implementation we have initially kept 25 wolves and 500 iterations and the results are measured as an average of 10 runs. We are loading different datasets for analyzing the performance of HBGWO. Segregating the dataset for training and testing is in the 70:30 ratio. Out of whole dataset, 70% records are considered for training and 30 % are for testing.

We have used MATLAB 2016 for conducting our experiments. Initial positions of wolves are assigned using rand function. Initially the alpha, beta and delta wolf fitness scores are set as infinity. Fitness is calculated and alpha, beta and delta score and position are updated as shown in the flowchart. If the alpha score value remains same for 20 iterations, then mutation operation is applied and the mutation rate is set as 0.1. Overall best score is taken as alpha score and alpha position tells which important features are selected in the given dataset.

The dataset used in proposed HBGWO is Diabetics dataset, Breast Cancer dataset and Heart Disease dataset. In diabetics dataset has 769 instances and 9 attributes such as Pregnancies, Glucose, Blood Pressure, Skin Thickness, Insulin, BMI, Diabetes Pedigree Function, Age and outcome where outcome is the class label having 0s and 1s where 0 indicate the patient is free from diabetes and 1 indicate the patient has diabetes.

Breast cancer data consists of 116 instances with 10 attributes such as Age, BMI, Glucose, Insulin, HOMA, Leptin, Adiponectiion, MCP. Heart disease dataset has 30 instances and 14 attributes. The classification algorithms used here are Naive Bayes, K Nearest Neighbour and Decision tree. The results of Hybrid BGWO algorithm are compared with that of BGWO and it is given the table 1

Table 1. Comparison of BGWO with HBGWO

Dataset	**Classification Algorithm**	**Accuracy %**	
		Binary Gray Wolf Optimization (BGWO)	**Hybrid Binary Gray Wolf Optimization with GA (HBGWO)**
Diabetes	Naive bayes	**80.45**	**82.52**
	KNN	78.38	79.57
	Decision tree	77.40	79.79
Breast cancer	Naive bayes	73.75	75.45
	KNN	**91.25**	**95.83**
	Decision tree	83.33	83.33
Heart disease	Naive bayes	85	87.74
	KNN	**87.21**	**89.32**
	Decision tree	83.28	85.08

VI. CONCLUSION

The main objective of the proposed work is to maximise the accuracy with the selection of minimal number of features. Compared to existing algorithm, the proposed HBGWO produce better results in terms of classification accuracy and the number of features selected. For diabetes data set, Naïve bayes classifier works better compared to KNN and Decision Tree Classifiers. For the Breast cancer and heart disease datasets, KNN produces better classification accuracy than the others both in BGWO and HBGWO. This shows that the proposed HBGWO algorithm can be better suited in medical field for identification of diseases.

REFERENCES

1. Banharnsakun, Anan, Tiranee Achalakul, and Booncharoen Sirinaovakul. "The best-so-far selection in artificial bee colony algorithm." Applied Soft Computing 11.2 (2011): 2888-2901
2. Unler, Alper, and Alper Murat. "A discrete particle swarm optimization method for feature selection in binary classification problems." European Journal of Operational Research 206.3 (2010): 528-539.
3. Chandrashekar, Girish, and Ferat Sahin. "A survey on feature selection methods." Computers & Electrical Engineering 40.1 (2014): 16-28.
4. Xue, Bing, et al. "A survey on evolutionary computation approaches to feature selection." IEEE Transactions on Evolutionary Computation 20.4 (2016): 606-626.
5. S. Mirjalili, S. M. Mirjalili, A. Lewis, "Grey Wolf Optimizer", Advances in Engineering Software, Vol. 69, pp. 46-61, 2014.
6. Sreekanth, G. R., R. C. Suganthe, and R. S. Latha. "An Energy Efficient and Stable Routing for MANET using Particle Swarm Optimization." (2015).
7. Latha, R. S., and G. Murugesan. "Encounter Based Clustering with Cuckoo Search for Ad Hoc Communication in WirelessHealth Informatics." Journal of Medical Imaging and Health Informatics 6.8 (2016): 1983-1989.
8. R.S. Latha, R & Murugesan, G. (2016). Multi-Metric Clustering in Mobile Ad-HocNetworks using Firefly Optimization Algorithm. Asian Journal of Research inSocial Sciences and Humanities. 6. 630. 10.5958/2249.
9. Rodrigues, Douglas, et al. "BCS: a binary cuckoo search algorithm for feature selection." 2013 IEEE International Symposium on Circuits and Systems (ISCAS2013). IEEE, 2013.
10. Nakamura, Rodrigo YM, et al. "BBA: a binary bat algorithm for feature selection." 2012 25th SIBGRAPI conference on graphics, patterns and images. IEEE, 2012.
11. Ghamisi, Pedram, and Jon Atli Benediktsson. "Feature selection based on hybridization of genetic algorithm and particle swarm optimization." IEEE Geoscience and remote sensing letters 12.2 (2015): 309-313.

Retrieval Number: D8251118419/2019©BEIESP
DOI:10.35940/ijrte.D8251.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication