Solution of Bio-Medical Problem by Genetic Algorithm

Narinder Singh*, Singh SB and Sharandeep Singh

Department of Mathematics, Punjabi University, Patiala-147002, India

*Corresponding author: Singh N, Department of Mathematics, Punjabi University, Patiala-147002, India, Tel: +91-7888785729; E-mail: narindersinghgoria@ymail.com

Citation: Singh N, Singh SB, Singh S (2018) Solution of Bio-Medical Problem by Genetic Algorithm. J Biomedical Sci. Vol.7 No.1:2.

Abstract

In operation research and computer science, a genetic algorithm (GA) is a most powerful meta-heuristic approach, its inspired by the process of natural selection. This approach is usually applied to generate superior quality results to standard and real life functions. Several number of researcher has been solved most number of real applications related to different fields with the help of this technique. After Inspired of these researchers, has been also solved the Breast Cancer and Iris data set problems in this article using some recent metaheuristics of nature inspired. For verification, the solutions are compared with some of the most well-known evolutionary trainers: Particle Genetic Algorithm (GA), Swarm Optimization (PSO), Ant Colony Optimization (ACO), Differential Evolution (DE), Personal Best Position Particle Swarm Optimization (PBPPSO), Evolutionary Strategy (ES), Biogeographical Based Optimization (BBO) and Population based Incremental Learning (PBIL). The numerical and statistical solutions show that GA algorithm is able to provide very competitive solutions in terms of improved local optima avoidance. The solutions also reveal a high level of accuracy in classification.

Keywords Breast cancer; Iris dataset; Bio medical problems; Nature inspired algorithms

Introduction

Neural Networks are one of the best inventions in the area of Artificial and Computational Intelligence. They mimic the neurons of human brain to mostly solve dataset and classification real life problems.

Several numbers of Neural Networks developed in the literature: Kohonen self-organizing network [1], recurrent neural network [2], Feedforward network [3], Spiking neural networks [4] and Radial basis function (RBF) network [5]. In Feedforward Neural Networks the information is cascaded in one direction throughout the networks.

In general, the approach that provides knowledge for a neural network is known as trainer. A trainer is answerable for training neural networks to find the best accuracy for new sets of given inputs. In the supervised education, a guiding method firstly provides Neural Networks with a set of samples called training samples. The guide then improves the structural constants of the Neural Network in every training sample in order to increase the accuracy. Once the training phase is finished, the guide is omitted and Neural Network is ready to apply. The guide can be measured as the most significant component of any Neural Networks.

There are two kinds of education techniques in the literature: (i) stochastic and (ii) deterministic. In such methods, the guiding phase solutions in the same accuracy if the guiding samples remain consistent. The guides in this group are mostly mathematical global optimization methods that aim to search the efficient solutions (least error). In contrast, stochastic guides use stochastic global optimization methods in order to maximize accuracy of a Neural Network.

Stochastic technique

The advantage of the stochastic techniques is high local optima avoidance, but they are mostly much slower than deterministic techniques. The literature proves that stochastic techniques have gained much attention recently, which is due to the high local optima avoidance.

Deterministic technique

The advantages of the deterministic guides are: speed and simplicity. This technique generally guides it toward an optimum and starts with a solution. The convergence is extremely rapid, but the superiority of the obtained result extremely depends on the initial result.

Some of the most popular multi-solution trainers in the literature are: Particle Swarm Optimization (PSO) [6,7], Genetic Algorithm (GA) [8,9], Ant Colony Optimization (ACO) [10,11], Differential Evolution (DE) [12,13], Personal Best Position Particle Swarm Optimization (PBPPSO) [14], Evolutionary Strategy (ES) [15], Biogeographical Based Optimization (BBO) [16] and Population based Incremental Learning (PBIL) [17]. The main reason as to why such nature inspired approaches have been employed as training techniques is their high performance in terms of approximating the global optimum. This also

motivates my attempts to investigate the efficiencies of recent metaheuristics in training Feedforward Neural Networks.

The rest of the article is organized as follows.

The literature review of meta-heuristics is presented in section 2. The genetic algorithm (GA) has also been discussed in section 3. Section 4 presents the experimental setup of parameter of biomedical problems. Results and discussion are provided in Section 5 respectively.Finally, the conclusion and future work of the work is summarized at end of the paper.

Literature Review

Several researcher and scientists have proposed numerous population based meta-heuristics to find the best possible solution of the different types of real life application. The most popular meta-heuristics techniques are discussed in this section.

Evolutionary Strategy has been presented by Ingo [15]. This are a sub-class of nature inspired search approach belonging to the class of Evolutionary methods. In which approach use recombination, mutation and selection applied to a population of individuals containing agent's solutions in order to evolve iteratively superior and superior results.

Population-based incremental learning (PBIL) was introduced by Shumet, B. in [17,18]. It is a global optimization method and an estimation of distribution variant. Population based incremental learning approach is an extension to the Genetic Algorithm (GA) achieved through the re-examination of the accuracy of the Genetic Algorithm (GA) in terms of competitive learning. It is easier than a GA technique and in most number of cases leads to superior and better qualities of global optimal solutions than a standard GA algorithm.

The Particle Swarm Optimization variant was originally developed by Eberhart, and Kennedy, James [19,20]. Its fundamental judgment was primarily inspired by the simulation of the social behaviour of animals such as bird flocking and fish schooling. While searching for food, the birds are either scattered or go together before they settle the position where they can find the food. While the birds are searching for food from one position to another, there is always a bird that can smell the food very well, that is, the bird is observable of the position where the food can be found, having the correct food resource message. Because they are transmitting the message, particularly the useful message at any period while searching the food from one position to another, the birds will finally flock to the position where food can be found.

Storn et al. [21] has been presented a heuristic approach called Differential Evolution (DE). It is newly heuristic variant mainly have three importance advantages; fast convergence, searching the best global minimum regardless of the initial constant values and using few control constants. This approach is a population based technique like GA using similar operators, mutation, crossover and selection. Its performance of existing approach was tested on several standards functions and accuracy also verified in the terms of convergence rate, solution quality and rate of success.

The Ant colony Optimization (ACO) variant was originated by Marco Dorigo [22]. This algorithm is based on the behaviour of ants seeking a path between their colony and source of food. The main idea has since diversified to solve a wider class of numerical functions and improved the quality of the optimal solutions.

Biogeographical Based Optimization (BBO) is an evolutionary method and meta-heuristics, which is inspired by the biogeographic concepts: extinction of species, the migration of species between islands and speciation (the evolution of new species). The approach is originally proposed by Simon [16]. Its accuracy was evaluated based on fourteen standard test functions, and then was verified to solve a real life application like sensor selection function for aircraft engine health estimation. This approach did fine and showed that is an excellent approach as compared to the other meta-heuristics. Since then, a several number of scientists have been conducted, some of them to solve practical functions.

N. Singh et al. [14] were developed a new particle swarm optimization method. In which this variant a novel philosophy of modifying the velocity update equation of Standard Particle Swarm Optimization (SPSO) variant was applied. The modification has been done by vanishing the best term in the velocity update equation of SPSO. The accuracy of the existing variant was tested on numerious standard functions. It is concluded that the existing meta-heuristic performs superior rather than Standard PSO in terms of accuracy and quality of optimal solution.

Genetic Algorithm (GA)

Firstly Genetic Algorithm was developed by Holland [20] in the literature. This variant is inspired by Darwin's theory of evolution "survival of the fittest". Which variant every newly population is created by combination and mutation of the individuals in the earlier iteration. Hence the best individuals have a higher probability of participating in generating the new direction of the agent, the newly direction is likely to be superior than the earlier direction of the agent.
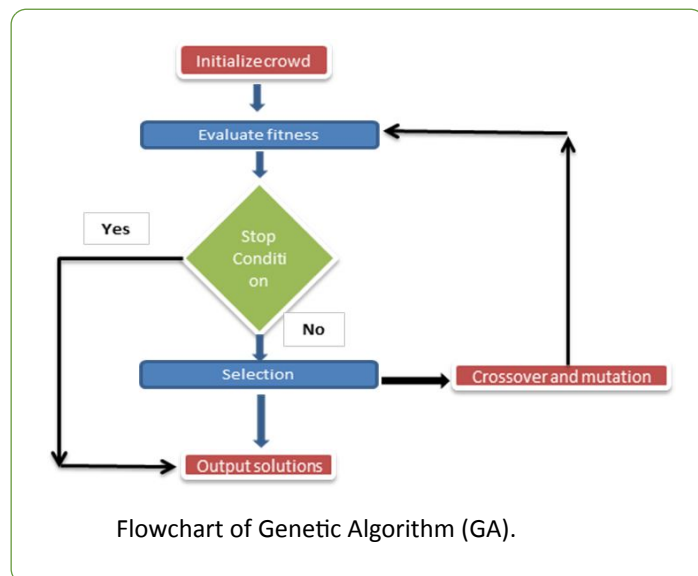
Darwin's idea of evolution is then modified to computational method to search best optimal value or solution to a function known objective function in natural fashion. A result created by GA is known a chromosome, while collection of chromosome is referred as a crowd. A chromosome is composed from genes and its value can be either binary, numerical, characters or symbols depending on the function want to be solved. These chromosomes will go through a procedure known fitness function to measure the suitability of result created by genetic algorithm with function. Some chromosomes in crowd will friend during process known crossover thus producing new chromosomes called children (offspring) which its genes composition are the grouping of their parent.

In a generation, some chromosomes will also mutation in their gene. The number of chromosomes which will undergo mutation and crossover is controlled by mutation rate value and crossover rate. Chromosome in the crowd that will maintain for the next iteration will be preferred based on Darwinian

evolution rule, the chromosome which has superior fitness result will have superior probability of being preferred once more in the next iteration. After numerous iterations, the chromosome value will converges to a certain value which is the best optimal solution for the function.

Algorithm: the pseudocode of a genetic algorithm

- Set constants.
- Choose encode technique.
- New crowd: generate a newly crowd by repeating following steps until the newly crowd is finished.
- while i < max$_{itr}$ and best $_{fitness}$ < maxfitness do
- **Fitness:** Evaluate the fitness $f(\chi)$ of every chromosome χ in the crowd.
- **Selection:** select two parent chromosomes from a crowd according to their fitness.
- **Crossover:** with a crossover probability cross over the parents to form new offspring. If no crossover was performed, offspring is the accurate duplicate of parents.
- **Mutation:** with a mutation probability mutate new children (offspring) at each locus as in the **Figure 1**.
- End while.
- Decode the individual with maximum fitness.
- return the best optimal value.

Flowchart of Genetic Algorithm (GA).

Experimental Setup

Table 1 The initial parameter of metaheuristics.

Classification datasets	Number of attributes	Number of training samples	Number of test samples	Number of classes
Breast Cancer	9	599	100	2
Iris	4	150	150	3

The numeral and statistical results are also compared with BBO, PSO, PBPPSO, GA, ACO, ES and Population-based Incremental Learning (PBIL) for verification. **Table 1** [18] shows the specifications of the datasets. It may be observed in **Table 3** that the easiest dataset is Iris dataset has 4 attributes, three classes and 150 training/test samples, and. In addition, the breast cancer dataset has 9 attributes, 2 classes, 100 test samples and 599 training samples.

Results and Discussion

For these experiments, the algorithms are coded in MATLAB R2013a, running on a Laptop with an 320 GB HDD, Intel HD Graphics, Pentium-Intel Core I, 15.6" 16.9 HD LCD, 3GB Memory and i5 Processor 430 M, and. In addition, to statistically asses the genetic algorithm compared with other methodologies, average and standard deviation are introduced.

The comprehensive information of parameters selection of test system is shown in section 2. All previous studies have been taken into account before applying the author's nature inspired metaheuristics techniques for finding the best possible solution of Breast Cancer and Iris dataset bio medical problems. Several number of metaheuristics has been tested on these dataset problems for three hundred times with various starting points and its performance have been compared with recent nature inspired algorithms. All the experimental results are demonstrated in **Tables 1 and 2**. The performance has been plotted using bar charts in **Figures 2-5**. On the basis of these Figures, have been represented the best classification accuracy of the meta-heuristics.

Tables 2 and 3, illustrate the performance of BBO, PSO, PBPPSO, GA, ACO, ES and PBIL variants in the terms of minimum objective function value, maximum objective function value, average and standard deviation and classification rate. The solutions reveal that the best average and standard deviation belongs to Genetic algorithm. This proves that this variant has the maximum ability to avoid local optima, significantly superior than other meta-heuristics. Genetic algorithm also classifies this dataset with hundred percentage accuracy. The GA approach is an evolutionary technique which has a extremely high level of exploration. These optimal solutions reveal that the GA approach is capable to prove very competitive solutions compare to the other meta-heuristics. Similarly, the accuracy of the meta-huristics has been tested in the terms of minimum and maximum objective function values, the results reveals that, the genetic algorithm gives best optimal values in their terms of maximum and minimum objective values of the datasets functions as comparison to other meta-heuristics with least number of iterations.

Hence, **Tables 2 and 3** and **Figures 2-5**, it is clear that GA algorithm gives a better quality of solutions and signifies GA's approach higher efficiency to find the solution of these bio medical dataset problems as compared to other meta heuristics.
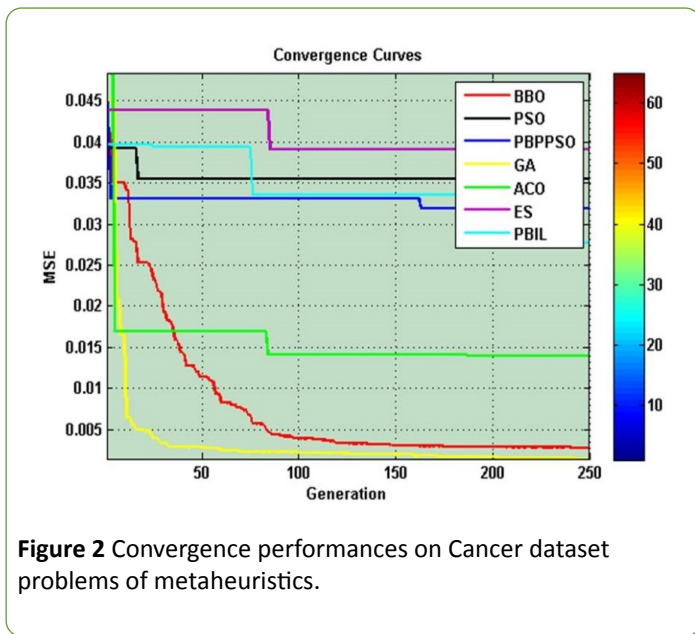
Figure 2 Convergence performances on Cancer dataset problems of metaheuristics.

Table 2 Solutions of Cancer dataset problem.

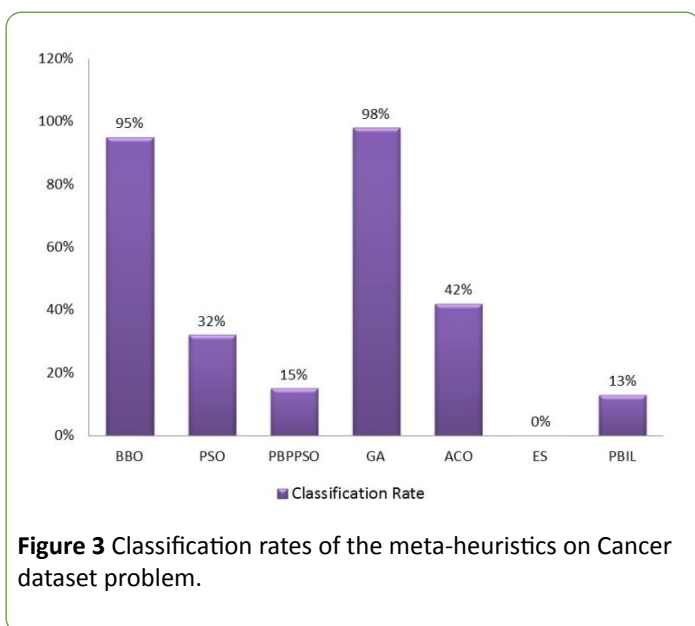| Algorithms	Min	Max	Ave.	S.D.	Classification Rate
BBO	0.0028	0.0409	0.0079	0.0089	95%
PSO	0.0354	0.047	0.0357	0.0012	32%
PBPPSO	0.0319	0.0466	0.0327	0.0011	15%
GA	0.0015	0.048	0.0034	0.0064	98%
ACO	0.014	0.048	0.0155	0.0044	42%
ES	0.0391	0.0439	0.0407	0.0023	0%
PBIL	0.0277	0.0397	0.0342	0.0042	13%

Figure 3 Classification rates of the meta-heuristics on Cancer dataset problem.

Table 3 Solutions of Iris dataset problem.

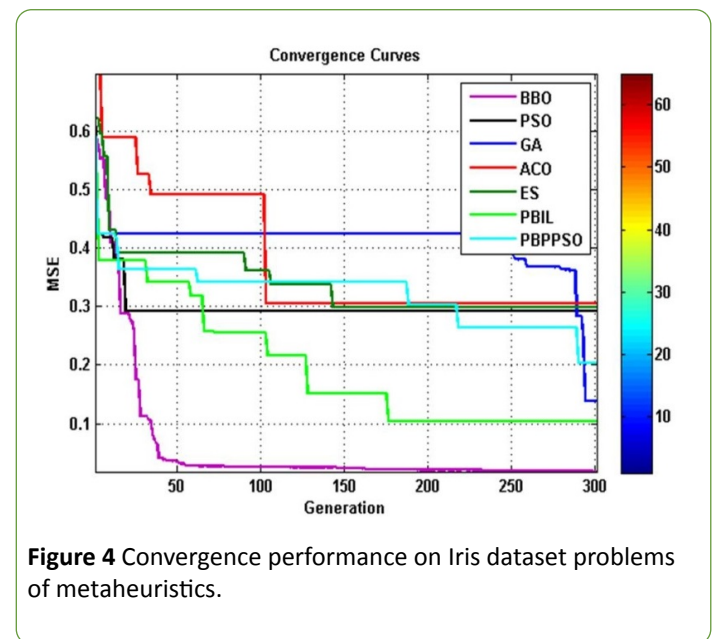| Algorithms	Min	Max	Ave.	S.D.	Classification Rate
BBO	0.1377	0.4239	0.4034	0.0514	34%
PSO	0.2927	0.6167	0.3002	0.0337	22.67%
PBPPSO	0.2021	0.6201	0.3221	0.0516	56.67%
GA	0.0178	0.6208	0.0579	0.1103	89.33%
ACO	0.3057	0.6084	0.3789	0.107	15.33%
ES	0.2978	0.6205	0.34	0.0591	42.6667
PBIL	0.1048	0.551	0.1953	0.103	64%

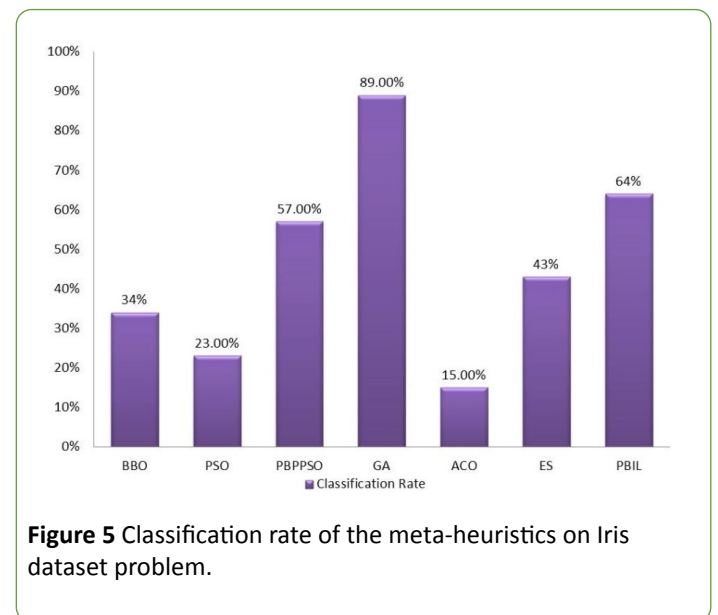**Figure 4** Convergence performance on Iris dataset problems of metaheuristics.

Figure 5 Classification rate of the meta-heuristics on Iris dataset problem.

Conclusion and Future Work

In this article, genetic algorithm (GA) was applied to two standard datasets like breast cancer and Iris. For verification, the best optimal solution of the genetic approach was compared to

six recent nature inspired algorithms i.e. BBO, PSO, PBPPSO, ACO, ES and PBIL. The simulation optimal solutions proved that the genetic approach is capable to be extremely efficient in terms of improved local optima avoidance. The solutions also reveal a high level of accuracy in classification. This article also discussed and identified the best reasons for poor and strong performances of the metaheuristics.

The future work will be concentrated on two parts: (i) balloon dataset, XOR dataset, heart dataset, feature selection, Structural Damage Detection, composite functions, the gear train design problem, aircraft's wings, bionic car problem, mechanical engineering functions and Cantilever beam (ii) Developing new approach based population based nature inspired techniques for these tasks. To end with, we expectation that this work will encourage young researchers and scientists, who are working on these concepts.

References

1. Kohonen T (1990) The self-organizing map. Proc IEEE 78: 1464-1480.

2. Dorffner G (1996) Neural networks for time series processing. in Neural Network World 6: 447-468.

3. Bebis G, Georgiopoulos M (1994) Feed-forward neural networks. Potentials, IEEE 13: 27-31.

4. Ghosh-Dastidar S, Adeli H (2009) Spiking neural networks. Int J Neural Syst 19: 295-308.

5. Park J, Sandberg IW (1993) Approximation and radial-basis-function networks. Neural Comput 5: 305-316.

6. Mendes R, Cortez P, Rocha M, Neves J (2002) Particle swarms for feedforward neural network training. learning 6: 1-5.

7. Gudise VG, Venayagamoorthy GK (2003) Comparison of particle swarm optimization and backpropagation as training algorithms for neural networks. In: Proceedings swarm intelligence symposium 3: 110-117.

8. Mirjalili S, Mohd Hashim SZ, Moradian Sardroudi H (2012) Training feedforward neural networks using hybrid particle swarm optimization and gravitational search algorithm. Appl Math Comput 218: 11125-11137.

9. Whitley D, Starkweather T, Bogart C (1990) Genetic algorithms and neural networks: Optimizing connections and connectivity. Parallel comput 14: 347-361.

10. Blum C, Socha K (2005) Training feed-forward neural networks with ant colony optimization: an application to pattern classification. In: 5th international conference on, Hybrid Intelligent Systems 5: 6.

11. Socha K, Blum C (2007) An ant colony optimization algorithm for continuous optimization: application to feed-forward neural network training. Neural Comput Appl 16: 235-247.

12. Ilonen J, Kamarainen JK, Lampinen J (2003) Differential evolution training algorithm for feed-forward neural networks. Neural Process Lett 17: 93-105.

13. Slowik A, Bialko M (2008) Training of artificial neural networks using differential evolution algorithm. In: 2008 Conference on, Human System Interactions 60-65.

14. Singh N, Singh SB (2012) Personal Best Position Particle Swarm Optimization. Journal of Applied Computer Science & Mathematics 6: 12.

15. Ingo R (1971) Evolutionary Strategy. Optimierung technischer Systeme nach Prinzipien der biologischen Evolution (PhD thesis). Reprinted by Fromman-Holzboog 1973.

16. Simon D (2008) Biogeography-based optimization. IEEE Transactions on Evolutionary Computation. 12: 702-713.

17. Baluja S (1994) Population-Based Incremental Learning: A Method for Integrating Genetic Search Based Function Optimization and Competitive Learning. Technical Report, Pittsburgh, PA: Carnegie Mellon University (CMU-CS-94-163).

18. Mirjalili S (2015) How effective is the Grey Wolf optimizer in training multi-layer perceptrons. Applied Intelligence 43: 150-161.

19. Kennedy J, Eberhart R (1995) Particle Swarm Optimization. Proceedings of IEEE International Conference on Neural Networks 1942-1948.

20. Holland JH (1992) Genetic Algorithms. Scientific American 267: 66-72.

21. Storn R, Price K (1997) Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization 11: 341-359.

22. Dorigo M, Birattari M, Stutzle T (2006) Ant Colony Optimization. Computational Intelligence Magazine, IEEE 1: 28-39.