Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Technology 4 (2012) 802 – 806

Procedia Technology

C3IT-2012

A New Meta-heuristic Bat Inspired Classification Approach for Microarray Data

Sashikala Mishra[a], Kailash Shaw[b], Debahuti Mishra[c]

[a,c]*Institute of Technical Education and Research, Siksha O Anusandhan Deemed to be University, Bhubaneswar, Odisha, India*

[b]*Gandhi Engineering College, Bhubaneswar, Odisha, India*

Abstract

The main objective of a classifier is to discover the hidden class level of the unknown data. It is observed that data size, number of classes and dimension of feature space and inter class separability affect the performance of any classifier. For a long time, efforts are made in improving efficiency, accuracy and reliability of classifiers for a wide range of applications. Different optimization algorithms such as Particle Swarm Optimization (PSO) and Simulated Annealing (SA) have been used to enhance the accuracy of classifiers. Bat is also a metaheuristic search algorithm which is use to solve multi objective engineering problem. In this paper, a model has been proposed for classification using bat algorithm to update the weights of a Functional Link Artificial Neural Network (FLANN) classifier. Bat algorithm is based on the echolocation behaviour of bats. The proposed model has been compared with FLANN, PSO-FLANN. Simulation shows that the proposed classification technique is superior and faster than FLANN and PSO-FLANN.

Keywords: Bat algorithm; FLANN; PSO-FLANN; BAT-FLANN

1. Introduction

High dimensionality of microarray data sets is a crucial issue to be considered while designing classifiers [5]. To handle the curse of high dimensionality, the data sets need to be pre-processed by reducing the redundant and irrelevant features. By removing such features or attribute we can also reduce the computational complexity. Principal Component Analysis (PCA) is used to deal with curse of dimensionality for micro array data set. The ultimate goal of any pattern recognition system is to achieve the best possible classification performance for a given problem domain. Meta heuristic algorithms like PSO [9][11] and SA are the powerful methods for solving many optimization problems. The fine adjustment of the parameters of the above techniques enhances the accuracy of the classifiers. In this paper, bat algorithm is used to update the weights of a FLANN classifier. Bat emits sound of various wavelength and frequency in the search of prey and direction [1]. Bat flies with velocity v at position x with different sound frequency f. Bat adjusts its velocity, direction and frequency on hearing echo signal. In this paper, a new meta heuristic bat algorithm has been formulated and also the whole working principle of the algorithm is explained. This paper is organized as follows; section 2 describes the related work, section 3 shows the schematic representation of the proposed model, section 4 contains the working

2212-0173 © 2012 Published by Elsevier Ltd.
doi:10.1016/j.protcy.2012.05.131

procedure of bat, section 5 gives the experimental evaluation and result; finally, section 6 deals with conclusion and future work.

2. Related Work

Xin-She Yang [1] proposed a bat algorithm and uses wavelength for the ease of implementation. Bats automatically adjust the wavelength of their emitted pulses and the rate of pulse emission. H. Hermann et al. [2] developed an alternative solution in the area of parametric and non parametric modeling of short time signals. The work presents the possibility of using the suggested parameterization methods in automatic species identification. J. D. Altringham [3] describes the behavior of bats.

3. Schematic Representation of Proposed Model

Proposed model consists of feature reduction techniques [6] for micro array data set which uses PCA. It is used to reduce the dimension and normalize the data set. 80 % of normalized data is sent to the BAT-FLANN classifier [7-8] for training and 20% is kept for testing. Weight is updated upon the pulse rate of the bat. Pulse rate is the controlling parameter of bat frequency which depends upon the error, bat position and velocity. Lastly the accuracy has been measured .The total procedure is illustrated in fig. 1.

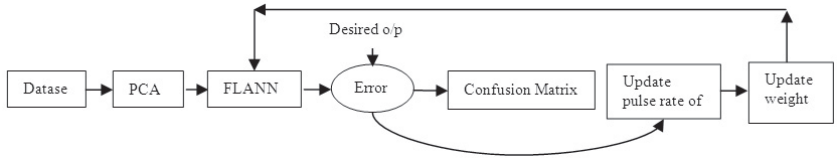

Fig.1. Proposed BAT-FLANN Model

4. Working Procedure of BAT

Bat sends signal with loudness of frequency 20 kHz to 200 kHz shown in fig 2. This signal when deflects back after striking the object to bat as echo signal given in fig 3, is used to calculate the distance *S*. The minimum distance from bat to any object is the destination of the bat [2]. Bat flies towards the minimum distance object. Bat reduces its pulse rate when it reaches nearer the object. Bat continues to do so till the position of the bat equals to any minimum distance object, i.e., distance becomes zero [3]. To calculate the distance, new position, frequency of a bat the following methods have been used.

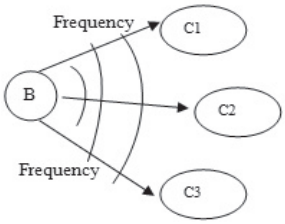

Fig.2. Bat sends sound signal with frequency *f*

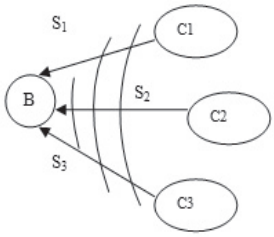

Fig.3. Echo signal use to calculate the distance *S*

4.1 Calculation of Frequency

Let us consider the dataset as *D* which is represented as *n* * *m* matrix, let the number of bats are $\{B_1, B_2,\ldots,B_k,\ldots,B_n\}$. Each row is considered as one bat. Every bat emits different sound signals with different wavelengths and pulse rate. Let us denote the frequency of sound as *f* for a bat *B*. The frequency