Expert Systems with Applications 40 (2013) 4648–4652

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Particle Swarm Optimization of MLP for the identification of factors related to Common Mental Disorders

Teresa B. Ludermir [a,*], Wilson R. de Oliveira [b]

[a] Center of Informatics, Federal University of Pernambuco, Recife, Brazil
[b] Federal Rural University of Pernambuco, Recife, Brazil

ARTICLE INFO

Keywords:
Common Mental Disorder
Neural Network Optimization
Particle Swarm Optimization

ABSTRACT

Social class differences in the prevalence of Common Mental Disorder (CMD) are likely to vary according to time, culture and stage of economic development. The present study aimed to investigate the use of optimization of architecture and weights of Artificial Neural Network (ANN) for identification of the factors related to CMDs. The identification of the factors was possible by optimizing the architecture and weights of the network. The optimization of architecture and weights of ANNs is based on Particle Swarm Optimization with early stopping criteria. This approach achieved a good generalization control, as well as similar or better results than other techniques, but with a lower computational cost, with the ability to generate small networks and with the advantage of the automated architecture selection, which simplify the training process. This paper presents the results obtained in the experiments with ANNs in which it was observed an average percentage of correct classification of individuals with positive diagnostic for the CMDs of 90.59%.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The Common Mental Disorders (CMDs), and among them the anxiety and depression have been pointed out as the common causes of morbidity in developed countries as much as in the developing ones, as the example of Brazil. These mental disorders represent a high social and economic charge because they are disabled, they constitute important cause of lost of workdays and they take a substantial use of health care services (Ludermir & Lewis, 2003). The use of techniques that may lead to an identification of the factors that present the larger possibility of being related to these CMDs it is of great relevance to assist within the process of decision taking around the planning and intervention of public health care. Artificial Neural Networks (ANNs) have been largely used in the health care field and they are known because they generally obtain a good precision result (Marcano-Cedeño et al., 2013; Babu & Suresh, 2013). When they are applied to epidemiological data the ANNs have also had acceptance (Chernbumroong, Cang, Atkins, & Yu, 2013).

With this research we intend, mainly, to experimentally display that a MLP network trained with Particle Swarm Optimization (PSO) (Kennedy & Eberhart, 1995) with early stopping criteria is able to identify the factors related to the CMDs. Global search techniques, such as Tabu Search (TS) (Glover, 1989), Evolutionary Algorithms (EAs, like Genetic Algorithm - GA) (Eiben & Smith, 2003), Differential Evolution (DE) (Storn, 1999), Particle Swarm Optimization (PSO) (Kennedy & Eberhart, 1995) and Group Search Optimization (GSO) (He, Wu, & Saunders, 2009), are widely used in scientific and engineering problems, and these strategies have been combined with ANNs to perform various tasks, such as connection weight initialization, connection weight training and architecture design. PSO has some advantages with respect to evolutionary algorithms. PSO for example has no complicated operators as evolutionary algorithms and it has less parameters which need to be adjusted (Kennedy & Eberhart, 1995).

The obtained results in the experiments with ANNs were compared with the ones presented by Ludermir and Lewis (2003) who applied the logistics regression method, using the same data basis to analyze the independence of each variable association with the CMDs. On the statistic analysis for the identification of the factors related to the CMDs, it was estimated the simple and adjusted odds-ratios, whose statistic significance was evaluated by the Students t-test, considering the 95% confidence interval and values of $p <= 0.05$.

The remainder of the article is organized as follows: Section 2 presents the standard PSO algorithm and the proposed methodology, PSO–PSO:WD. Section 3 describes the data basis and Section 4 presents the experimental setup of this work. Section 5 presents and analyzes the results obtained from the experiments and in Section 6 we summarize our conclusions and future works.

* Corresponding author. Tel.: +55 81 21268430.
E-mail addresses: tbl@cin.ufpe.br (T.B. Ludermir), wilson.rosa@gmail.com (W.R. de Oliveira).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.02.003

2. Particle Swarm Optimization

This section presents the basic algorithm of Particle Swarm Optimization and the algorithm based on the interleaved execution of two PSO algorithms.

2.1. Basic concepts

The Particle Swarm Optimization (PSO) is a population-based stochastic optimization technique developed by Kennedy and Eberhart in 1995. PSO attempts to model the flight pattern of a flock of birds (Kennedy & Eberhart, 1995). In PSO, each particle represents a candidate solution within a n-dimensional search space. The position of a particle i at iteration t is denoted by $\mathbf{x}_i(t) = [x_{i1}, x_{i2}, \ldots, x_{in}]$. In each iteration of the PSO, each particle moves through the search space with a velocity $\mathbf{v}_i(t) = [v_{i1}, v_{i2}, \ldots, v_{in}]$ calculated as follows:

$$v_{ij}(t+1) = wv_{ij}(t) + c_1 r_{j1}[y_{ij}(t) - x_{ij}(t)] + c_2 r_{j2}[\hat{y}_{ij}(t) - x_{ij}(t)], \quad (1)$$

where w is the inertia weight, $\mathbf{y}_i(t)$ is the personal best position of the particle i at iteration t and $\hat{\mathbf{y}}(t)$ is the global best position of the swarm at iteration t. The personal best position is named *pbest* and it represents the best position found by the particle during the search process until the iteration t. The global best position is named *gbest* and it represents the best position found by the entire swarm until the iteration t. The terms c_1 and c_2 are acceleration coefficients and are responsible for taking control of how far a particle can move in a single iteration. The terms r_{j1} and r_{j2} are random numbers sampled from an uniform distribution $U(0,1)$. The velocity is limited to the range $[\mathbf{v}_{min}, \mathbf{v}_{max}]$.

After updating velocity, the new position of the particle i at iteration $t+1$ is calculated using Eq. (2)

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1). \quad (2)$$

In Shi and Eberhart (1998), Shi and Eberhart proposed an adaptive inertia in which the parameter w reduces gradually as the iteration increases according to Eq. (3)

$$w(t) = w_{max} - t \times \frac{(w_{max} - w_{min})}{t_{max}}, \quad (3)$$

where w_{max} is the initial inertia weight, w_{min} is the final inertia weight and t_{max} is the maximum number of iterations. The inertia weight can control the degree of exploration of the search.

The standard PSO algorithm is presented in Algorithm 1. Rapid convergence in unimodal functions, with good success rate, and premature convergence in multimodal functions are properties frequently attributed to the standard PSO algorithm.

Algorithm 1. PSO Algorithm

Initialize the swarm S;
while stopping condition is false **do**
  for all particle i of the swarm **do**
    Calculate the fitness $f(\mathbf{x}_i(t))$;
    Set the personal best position $\mathbf{y}_i(t)$;
  end for
  Set the global best position $\hat{\mathbf{y}}(t)$ of the swarm;
  for all particle i of the swarm **do**
    Update the velocity $\mathbf{v}_i(t)$;
    Update the position $\mathbf{x}_i(t)$;
  end for
end while

2.2. The PSO–PSO Methodology

The methodology used to optimize weights and architectures of MLP neural networks is based on the interleaved execution of two PSO algorithms, one for weight optimization (inner PSO) and the other for architecture optimization (outer PSO). This approach was presented by Zhang and Shao in Zhang and Shao (2000), in which few details were given on the performance of the optimized neural networks.

In this methodology, the outer PSO simply searches for the number of hidden units for each of the considered hidden layers in the MLP network. In this work, we considered only network architectures of one hidden layer, but the extension for a more general case is straightforward. The inner PSO is responsible for the optimization of weights for each of the architectures (particles) present in the outer PSO. At the end of the inner PSO execution for an architecture of the outer PSO, the best set of weights found is recorded in the particle representing that architecture. The two processes are interleaved for a specific number of times.

The PSO–PSO methodology developed in this work used a PSO algorithm to search for architectures and a PSO with weight decay (PSO:WD) to search for weights. The PSO:WD algorithm was created in a previous work (Carvalho & Ludermir, 2006) and has more generalization control than the standard PSO. For the evaluation of performance of the particles of the two PSOs, we used different partitions of the example patterns set. The training set partition (50%) was used within the inner PSO to optimize weights while the validation set partition (25%) was used within the outer PSO to search for architectures. The remained data (25%) was used to test the final MLP network found by the process.

It should be noted that all the three partitions used in the methodology are disjoint. That restriction is related to the aim of improving the generalization control of the optimized networks. This can be done through the adjustment of the complexity (number of hidden units) of the networks guided by data examples different from the ones used to guide the search for synaptic weights.

The complete algorithm for the methodology created in this work is presented in Algorithm 2, in which the term A_i.*net* represents the vector used to record the best network found so far for the architecture A_i. Note that this term is updated with the best particle at the end of an execution of the inner PSO (line 7), and is used to assist a new execution of the inner PSO with previous good results (line 5). The PSO:WD algorithm is better described in Carvalho and Ludermir (2006) in which the standard PSO is combined with weight decay heuristic as an attempt to improve the generalization performance of the trained MLP networks.

3. Data basis description

Data collection was community-based through interviews and assessment of mental health status from a research made in the city of Olinda, Brazil by Ludermir and Lewis (2003). With that research Ludermir and Lewis (2003) determined the prevalence of the CMDs in that area, and analyzed the association with living and work conditions.

The study was developed with 621 adults of an aleatory domicile sample and the analysis of data was made using a statistic model of logistic regression.

متن کامل مقاله ← دریافت فوری

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات