

Chapter 2

Swarm–Based Nature–Inspired Metaheuristics for Neural Network Optimization

Swathi Jamjala Narayanan
VIT University, India

Boominathan Perumal
VIT University, India

Jayant G. Rohra
VIT University, India

ABSTRACT

Nature-inspired algorithms have been productively applied to train neural network architectures. There exist other mechanisms like gradient descent, second order methods, Levenberg-Marquardt methods etc. to optimize the parameters of neural networks. Compared to gradient-based methods, nature-inspired algorithms are found to be less sensitive towards the initial weights set and also it is less likely to become trapped in local optima. Despite these benefits, some nature-inspired algorithms also suffer from stagnation when applied to neural networks. The other challenge when applying nature inspired techniques for neural networks would be in handling large dimensional and correlated weight space. Hence, there arises a need for scalable nature inspired algorithms for high dimensional neural network optimization. In this chapter, the characteristics of nature inspired techniques towards optimizing neural network architectures along with its applicability, advantages and limitations/challenges are studied.

INTRODUCTION

Swarm based Optimization is a strategy that considers several agents collectively working to intelligently achieve a goal in the most optimal manner. Nature inspired techniques might include the consideration of these agents to be a flock of birds, a school of fish, a swarm of bees etc. Metaheuristic approaches are framed to form an analogy between the nature and computational systems and hence implement a

DOI: 10.4018/978-1-5225-2857-9.ch002

relevant behavior as a paradigm to perform the required task. Since 1990, several nature inspired meta-heuristic techniques have been proposed. There exist several applications areas where the metaheuristics like swarm based or evolutionary optimization algorithms play vital role. Several NP-hard optimization problems like Traveling Salesman Problem, Quadratic Assignment Problem, Graph problems are also solved using nature inspired techniques.

Classification is the task of assigning an object to a pre-defined class or group (Duda, 1973). Classifier can be considered as a mapping function of the form $Y_i = f(X_i, \Phi_1, \Phi_2, \dots, \Phi_N)$, where $f(\cdot)$ is the classifier that maps the object X_i to class Y_i based on parameters $\Phi_1, \Phi_2, \dots, \Phi_N$ that are related attributes of object X_i . Classification is widely used in business, science, industry, and medicine and addresses many real world problems such as bankrupt prediction, credit scoring, medical diagnosis, handwritten character recognition, and speech recognition.

In traditional statistical classifiers, classification decision depends on posterior probability which is derived based on the assumptions on underlying probability model. Prior knowledge required on data properties and model capabilities limits the scope of statistical classifiers in many real world problems. Emergence of Neural Network, a non-linear model that models real world complex problems provides solution for the conventional statistical classifiers. The advantage of neural networks lies in the following theoretical aspects. First, neural networks are data driven self-adaptive methods in that they can adjust themselves to the data without any explicit specification of functional or distributional form for the underlying model. Second, they are universal functional approximators in that neural networks can approximate any function with arbitrary accuracy (Cybenko 1989; Hornik, 1991; Hornik et al, 1989).

Since any classification procedure seeks a functional relationship between the group membership and the attributes of the object, accurate identification of this underlying function is doubtlessly important. Third, neural networks are nonlinear models, which makes them flexible in modelling real world complex relationships. Finally, neural networks are able to estimate the posterior probabilities, which provide the basis for establishing classification rule and performing statistical analysis. Neural networks are considered as data driven self-adaptive methods and universal functional approximators that estimates posterior probability with arbitrary accuracy.

To improve the performance of the neural networks by optimizing its parameters, the authors (Werbos, 1990, 1994; Williams et al., 1986; Gupta & Sexton, 1999; Wilamowski, 2002) have suggested that back propagation using gradient descent methods is the most widely used neural network training method to optimize the neural network parameters in supervised learning strategy. In recent years, many improved learning algorithms have been developed that aim to remove the shortcomings of the gradient descent based systems.

Zell (2002) developed The Stuttgart Neural Network Simulator (SNNS), which uses many different algorithms including Error Back Propagation developed by Fahlman (1988), Resilient Error Back Propagation developed by Riedmiller, & Braun (1993), Back percolation, Delta-bar-Delta, Cascade Correlation developed by Fahlman & Lebiere (1989) etc. All these algorithms are derivatives of steepest gradient search; hence the ANN training is relatively slow. To have a fast and efficient training method, second order learning algorithms are developed. The most effective method is Levenberg Marquardt (LM) algorithm proposed by Hagan & Menhaj (1994), which is a derivative of the Newton method. This is quite multifaceted algorithm since both the gradient and the Jacobian matrix is calculated. The LM algorithm was developed only for layer-by-layer ANN topology, which is far from optimal. LM algorithm is ranked as one of the most efficient training algorithms for patterns that are both small and medium

29 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/swarm-based-nature-inspired-metaheuristics-for-neural-network-optimization/187679

Related Content

Searching for Self-Replicating Systems

Eleonora Bilotta and Pietro Pantano (2010). *Cellular Automata and Complex Systems: Methods for Modeling Biological Phenomena* (pp. 185-209).

www.irma-international.org/chapter/searching-self-replicating-systems/43221

An Observer Approach for Deterministic Learning Using Patchy Neural Networks with Applications to Fuzzy Cognitive Networks

H. E. Psillakis, M. A. Christodoulou, T. Giotis and Y. Boutalis (2011). *International Journal of Artificial Life Research* (pp. 1-16).

www.irma-international.org/article/observer-approach-deterministic-learning-using/52974

Hengam a MapReduce-Based Distributed Data Warehouse for Big Data: A MapReduce-Based Distributed Data Warehouse for Big Data

Mohammadhossein Barkhordari and Mahdi Niamanesh (2018). *International Journal of Artificial Life Research* (pp. 16-35).

www.irma-international.org/article/hengam-a-mapreduce-based-distributed-data-warehouse-for-big-data/211933

Existence of Positive Solutions for Generalized p-Laplacian BVPs

Wei-Cheng Lian, Fu-Hsiang Wong, Jen-Chieh Lo and Chih-Chih Yeh (2011). *International Journal of Artificial Life Research* (pp. 43-53).

www.irma-international.org/article/existence-positive-solutions-generalized-laplacian/52978

Energy Resources and Their Consumption

Harpreet Kaur Channi (2022). *Applications of Nature-Inspired Computing in Renewable Energy Systems* (pp. 267-276).

www.irma-international.org/chapter/energy-resources-and-their-consumption/294396