

Chapter 6

Applications in Machine Learning

This chapter presents the fundamental concepts regarding the application of PSO on machine learning problems. The main objective in such problems is the training of computational models for performing classification and simulation tasks. It is not our intention to provide a literature review of the numerous relative applications. Instead, we aim at providing guidelines for the application and adaptation of PSO on this problem type. To achieve this, we focus on two representative cases, namely the training of artificial neural networks, and learning in fuzzy cognitive maps. In each case, the problem is first defined in a general framework, and then an illustrative example is provided to familiarize readers with the main procedures and possible obstacles that may arise during the optimization process.

INTRODUCTION

Machine learning is the field of artificial intelligence that deals with algorithms that render computational models capable of learning and adapting to their environment. From an abstract viewpoint, machine learning is the procedure of extracting information in the form of patterns or rules from data. This purpose requires the use of computational methods. Human interaction can also be beneficial within a collaborative framework to the algorithms. However, the elimination of this necessity still remains the main challenge in the development of intelligent systems.

There are different types of machine learning procedures, based on the desired outcome as well as on the degree of human intervention:

- a. **Supervised learning:** The algorithm builds a mapping between a set of presented input data and a set of desired output. This is possible by altering the parameters of the computational model so that the produced error between input and output is minimized.

DOI: 10.4018/978-1-61520-666-7.ch006

- b. **Unsupervised learning:** The algorithm tunes the computational model to regularities of the available data without a task-oriented measure of quality. This is possible through competition among the modules of the computational model.
- c. **Semi-supervised learning:** This is a combination of the two previous approaches that employs both learning with explicit input-output examples, as well as non-labeled examples.
- d. **Reinforcement learning:** The algorithm learns an input-output mapping by continuously interacting with the environment, which in turn admits an impact from every taken step, providing feedback to the model.

Several additional learning subtypes, which are outside the scope of the book at hand, can be distinguished.

Supervised learning constitutes a very prosperous application field for evolutionary algorithms and PSO, due to the existence of explicit performance measures. These measures usually come in the form of objective functions in the parameters of the computational model. Thus, training procedures aim at the detection of parameter values that minimize the model's error in learning a set of presented examples.

In this context, PSO has been applied for training artificial neural networks and fuzzy cognitive maps. The rest of this chapter is dedicated to an overview of these applications.

TRAINING ARTIFICIAL NEURAL NETWORKS WITH PSO

In the following sections, we briefly present the problem of neural network *training* for the most common case of feedforward neural networks. We also present an illustrative example for the logical XOR classification task. Further applications are also reported.

The Multi-Layer Perceptron Model

Artificial neural networks (NNs) are computational models based on the operation of biological neural networks, which constitute the information processing mechanism of the human brain. Their structure is based on the concept of the *artificial neuron*, which resembles biological neurons, as their main processing unit. The artificial neuron constitutes a nonlinear mapping between a set of input and a set of output data. Thus, if the input data are represented as a vector, $Q = (q_1, q_2, \dots, q_m)^T$, the artificial neuron implements a function:

$$y = F\left(b_i + \sum_{i=1}^m w_i q_i\right),$$

where F is the *transfer function*; w_i , $i = 1, 2, \dots, m$, are the *weights*; and b_i is a *bias*. In order to retain a compact notation, we will henceforth represent the bias, b_i , as a weight, w_0 , with an auxiliary constant input, $q_0 = 1$.

The training of the neuron to learn an input-output pair, $\{Q, y\}$, is the procedure of detecting proper weights so that the output y is obtained if Q is presented to the neuron. Obviously, this procedure can be modeled as an error minimization problem:

17 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/applications-machine-learning/40633

Related Content

The Use of Evolutionary Algorithm-Based Methods in EEG Based BCI Systems

Adham Atyabi, Martin Luerssen, Sean P. Fitzgibbon and David M W Powers (2013). *Swarm Intelligence for Electric and Electronic Engineering* (pp. 326-344).

www.irma-international.org/chapter/use-evolutionary-algorithm-based-methods/72835

Adaptive Neuro-Fuzzy Control Approach Based on Particle Swarm Optimization

Gomaa Zaki El-Far (2012). *Innovations and Developments of Swarm Intelligence Applications* (pp. 81-98).

www.irma-international.org/chapter/adaptive-neuro-fuzzy-control-approach/65807

Particle Swarm Optimization Algorithm in Electromagnetics- Case Studies: Reconfigurable Radiators and Cancer Detection

Arezoo Modiri and Kamran Kiasaleh (2013). *Swarm Intelligence for Electric and Electronic Engineering* (pp. 72-99).

www.irma-international.org/chapter/particle-swarm-optimization-algorithm-electromagnetics/72824

Opinion on Different Classification Algorithms Used in Internet of Things Environment for Large Data Set

Akhil Bansal, Piyush Kumar Shukla and Manish Kumar Ahirwar (2019). *International Journal of Organizational and Collective Intelligence* (pp. 51-60).

www.irma-international.org/article/opinion-on-different-classification-algorithms-used-in-internet-of-things-environment-for-large-data-set/218272

Data Mining Approaches for Geo-Spatial Big Data: Uncertainty Issues

Frederick E. Petry (2012). *International Journal of Organizational and Collective Intelligence* (pp. 52-71).

www.irma-international.org/article/data-mining-approaches-for-geo-spatial-big-data/85333