

Natural Computing Assessment

Exam Nos. B156771, [insert exam number here]

November 12, 2020

1 Neural Network Training with PSO

1.1 Choosing the Fitness Function

Let $X \subset [-6, 6]^2$ and $F : X \rightarrow \{0, 1\}$ be the function we are trying to approximate via a neural network. Furthermore, let $f_v : X \rightarrow [0, 1]$ be the prediction function, where v is the weight vector of the network. We choose the fitness function of the PSO to be the Mean Squared Error of the network on the training set X , as given by,

$$C_{f_v}^{X,F}(w) = \frac{1}{|X|} \sum_{i=1}^{|X|} |F(x_i) - f_w(x_i)|^2. \quad (1)$$

This fitness function was chosen for a variety of reasons, not least of which is simplicity of implementation. The same function can be used to train a neural network using Stochastic Gradient Descent, providing us a baseline on which to test the PSO trained model.

1.2 Defining the Search Space

The search space of a PSO is given by $[a, b]^D$, where $a, b \in \mathbb{R}$, and $D \in \mathbb{N}$ is the dimension of the space. We will start by determining the value for D . We define the shape of a neural network by a sequence $(a_n)_{n=1}^l$ where $l \in \mathbb{N}$ defines the number of layers, and each $a_n \in \mathbb{N}$ gives the number of neurons in layer n . The dimension of the weight vector for a would be the number of edges connecting each sequential pair of layers, plus the biases, given by the number of neurons in the latter of the pair of layers. Succinctly, the dimension D of the search space is given by,

$$D = \left(\sum_{i=1}^{l-1} a_{i+1} (a_i + 1) \right) - 1, \quad (2)$$

where the final decrement addresses the unnecessary final bias term.

For this task, we set the shape to be $(6, 8, 1)$, meaning that the dimension of the search space $D = 43$. To ensure that we are notified of a divergent swarm, we set $-a = b = 10^{10}$, meaning that the search space for the PSO will be $[-10^{10}, 10^{10}]^{43}$.

1.3 Training the Model

1.4 Restriction to Linear Inputs

1.5 Particle Swarm or Gradient Descent