BRSU

Neural Networks
Assignment 5

Bastian Lang

November 14, 2015

This report contains the summary, ex3.1 and ex3.2 without any programming/plotting.

1 Outline

- Introduction
  - Three distinct characteristics of MLP:
    * Model of each neuron includes nonlinear activation function that is differentiable everywhere
    * One or more layers of hidden units
    * High degree of connectivity
  - Organization of the chapter
- Some Preliminaries
  - Signals in an MLP:
    * Function Signals: Propagating forward through the network
    * Error Signals: Propagating backwards through the network
  - Computations for each neuron:
    * Function signal
    * Gradient vector of error surface
  - Notation
- Back-Propagation Algorithm
  - Error signal at output neuron: $e_j(n) = d_j(n) - y_j(n)$
  - Instantaneous value of total error energy: $\xi(n) = 0.5\sum_{j\epsilon C} e_j^2(n)$
  - Induced Local Field $v_j(n) = \sum_{i=0}^{m} w_{ji}(n) y_i(n)$
  - Function signal $y_i(n) = \phi_j(v_j(n))$
  - Derivation of delta rule
  - Local gradient for output neuron: $\delta_j(n) = e_j(n)\phi'_j(v_j(n))$
  - Local gradient for hidden neuron: $\delta_j(n) = \phi'_j(v_j(n))\sum_k \delta_k(n) w_{kj}(n)$
  - Weight correction: $\Delta w_{ji}(n) = \eta * \delta_j(n) * y_i(n)$
  - Two passes of Computation
    * Forward pass: Compute net output
    * Backward pass: Adjust weights according to error
  - Activation Function
    * Logistic Function
    * Hyperbolic tangent function

    - Rate of Learning
        * Including momentum term for stability
    - Sequential and Batch Modes of Training
        * Epoch: Complete presentation of training set
        * Sequential Mode of back-propagation learning: Update weights directly
        * Batch Mode of back-propagation learning: Update weights only after an epoch
    - Summary of the Back-Propagation Algorithm
        * Cycle for sequential updating:
            · Initialization
            · Presentations of Training Examples
            · Forward Computation
            · Backward Computation
            · Iteration

- Summary of the back-propagation algorithm

- XOR problem
    - Special case of classifying points in the unit hypercube
    - Elemental perceptron cannot solve XOR
    - Example model using McCulloch-Pits neurons

- Heuristics for making the back-propagation algorithm perform better
    - Methods to improve back-prop
        * Sequential versus batch update
        * Maximizing information content
            · Use of example that results in largest training error
            · Use of example that is radically different than previous ones
        * Activation Function
            · Better antisymmetric than non-symmetric
        * Target values
            · Offset of target values to lie within function range
        * Normalizing the inputs
        * Initialization
        * Learning from hints
        * Learning rates