MACHINE LEARNING
BARUCH COLLEGE
SPRING 2015
MIGUEL A. CASTRO
T.A.: LIQUN ZHU

Homework Assignment 1

1. Show that the single-Perceptron learning rule with hard limit activation converges (for linearly-separable problems). Is this an example of Supervised or Unsupervised Learning?
2. Show x XOR y = (x OR y) AND [NOT(x AND y)] (truth tables are good enough).
3. Why are the outputs of the OR and the NOT-AND Perceptrons guaranteed to be linearly separable? Does this hold in higher dimensions? Explain.
4. Can you think of a Perceptron *architecture* that would solve the XOR problem using an MLP with Hard-Limit activations? Discuss.
5. Can you think of a Perceptron *learning rule* that would converge in an *automated* fashion for the XOR problem using an MLP with Hard-Limit activations? Discuss.
6. Derive a MADALINE learning rule for a simple MADALINE with two inputs, two linear-activation nodes in the hidden layer and one linear-activation output node by back-propagating the error from the output node to the input weights using the chain rule of differentiation. The goal is to minimize the squared error: $e = \frac{1}{2}(t - y)^2$ using gradient descent.
7. Show that MADALINEs with one Hidden Layer cannot solve Linearly-Separable two-class problems by demonstrating that the separation hyperplane is given by
   $$\widetilde{\mathbf{W}}\mathbf{p} + \tilde{\mathbf{b}} = \mathbf{0},$$
   where $\widetilde{\mathbf{W}}$ and $\tilde{\mathbf{b}}$ are functions of the hidden-layer and output-layer weights and biases. Show that this result holds for an arbitrary number of hidden layers.
8. What are the two criteria required for ANNs to be universal mapping approximators? What condition must be met for automated learning on a multi-layer architecture?