**COMP 473/6731, FALL 2019** **CONCORDIA UNIVERSITY**

# PATTERN RECOGNITION
## ASSIGNMENT 4.
## Trees and Neural Networks.
Due: Dec. 1, 2019.

Do the following problems. [1]

1. Problem 8 (a), sect. 8.3, p. 439.
   Consider a two-category problem and the following training patterns, each having binary attributes:

| $\omega_1$ | $\omega_2$ |
|---|---|
| 0110 | 1011 |
| 1010 | 0000 |
| 0011 | 0100 |
| 1111 | 1110 |

   (a) (20 marks) Use the entropy impurity eq. (1), p. 398 to create by hand an unpruned classifier for these data.

2. Problem 10, sect. 6.3, p. 337.
   Express the derivative of a sigmoid in terms of the sigmoid itself in the following two cases (for positive constants $a$ and $b$):

   (a) (10 marks) A sigmoid that is purely positive: $f(net) = \frac{1}{1+\exp(a \cdot net)}$.

   (b) (10 marks) An anti-symmetric sigmoid: $f(net) = a \cdot \tanh(b \cdot net)$.

3. Problem 4, sect. 6.3, p. 344 (Computer Exercise).
   Write a backpropagation program for a 2-2-1 sigmoidal network with bias to solve the XOR problem (see Fig. 6.1 below).

   (a) (30 marks) Show the input-to-hidden weights and analyze the function of each hidden unit.

   (b) (30 marks) Plot the representation of each pattern as well as the final decision boundary in the $y_1y2$-space.

   (c) (20 marks) Although it was not used as a training pattern, show the representation $\mathbf{x} = \mathbf{0}$ in your $y_1y2$-space.

[1] see the textbook Duda, Hart, and Stork *Pattern Classification* 2nd edition, 2001

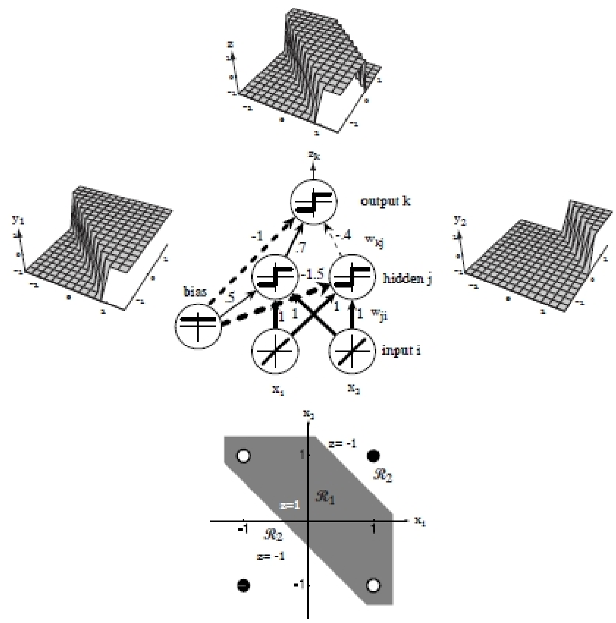


Figure 6.1: The two-bit parity or exclusive-OR problem can be solved by a three-layer network. At the bottom is the two-dimensional feature space $x_1 - x_2$, and the four patterns to be classified. The three-layer network is shown in the middle. The input units are linear and merely distribute their (feature) values through multiplicative weights to the hidden units. The hidden and output units here are linear threshold units, each of which forms the linear sum of its inputs times their associated weight, and emits a +1 if this sum is greater than or equal to 0, and −1 otherwise, as shown by the graphs. Positive ("excitatory") weights are denoted by solid lines, negative ("inhibitory") weights by dashed lines; the weight magnitude is indicated by the relative thickness, and is labeled. The single output unit sums the weighted signals from the hidden units (and bias) and emits a +1 if that sum is greater than or equal to 0 and a -1 otherwise. Within each unit we show a graph of its input-output or transfer function — $f(net)$ vs. $net$. This function is linear for the input units, a constant for the bias, and a step or sign function elsewhere. We say that this network has a 2-2-1 fully connected topology, describing the number of units (other than the bias) in successive layers.