ECS 170 Homework 6

Hardy Jones
999397426
Professor Davidson
Winter 2014

1. What condition must hold on the training data so that a perceptron can accurately represent a function which classifies the training data perfectly?

   The training data must be linearly separable.

2. Name two properties why the objective function of linear units is desirable.

   (a) This function is differentiable.

   (b) This function will find the global minimum.

3. Construct a network of linear units that is capable of representing the XOR function of two inputs.

   We cannot construct the network with a single perceptron. We need at least one hidden layer, depending on what primitive functions are available.

   It helps to look at how to construct XOR.

A	B	A ⊕ B	A ∧ B	A ∨ B	¬(A ∧ B)	¬(A ∧ B) ∧ (A ∨ B)
0	0	0	0	0	1	0
0	1	1	0	1	1	1
1	0	1	0	1	1	1
1	1	0	1	1	0	0

   So, if we have functions for AND, OR, and NOT we can construct XOR.

   So given two input nodes x_1 and x_2 we construct hidden nodes x_3 to x_6, and output o_7,

   Intermediate nodes:

$$
\begin{aligned}
x_3 &= AND(x_1, x_2) \\
x_4 &= OR(x_1, x_2) \\
x_5 &= NOT(x_3) \\
x_6 &= NOT(NOT(x_4)) \qquad (*)
\end{aligned}
$$

   Output node:

$$o_7 = AND(x_5, x_6)$$

   *: this is the Identity function.

4. Suppose the inputs are given by x_1 and x_2, and the activation functions at each unit is given by the function g. Write out the values o_5 and o_6 at the output nodes of figure 1 in terms of the weights $w_{i,j}$ and the inputs x_k.

   It helps to work backwards from the output to the input. Let's start with o_6.

$$\begin{aligned} o_6 &= g(w_{3,6} \cdot x_3, w_{4,6} \cdot x_4) \\ &= g(w_{3,6} \cdot g(w_{1,3} \cdot x_1, w_{2,3} \cdot x_2), w_{4,6} \cdot g(w_{1,4} \cdot x_1, w_{2,4} \cdot x_2)) \end{aligned}$$

   And now o_5.

$$\begin{aligned} o_5 &= g(w_{3,5} \cdot x_3, w_{4,5} \cdot x_4) \\ &= g(w_{3,5} \cdot g(w_{1,3} \cdot x_1, w_{2,3} \cdot x_2), w_{4,5} \cdot g(w_{1,4} \cdot x_1, w_{2,4} \cdot x_2)) \end{aligned}$$