MML - Review for Exam 4

We will have our fourth exam next Wednesday, April 23. This review sheet is again meant to help you succeed on that exam.

1. What is the sigmoid function $\sigma(x)$? More specifically,

   a. Write down the exact algebraic definition of $\sigma(x)$.
   b. Draw a graph of $\sigma(x)$.
   c. In the context of binary classification, how do we interpret the sigmoid?
   d. Show that $\sigma(x)$ satisfies the equation $\sigma' = \sigma(1 - \sigma)$.
   e. Suppose that the number x_0 is chosen so that $\sigma(x_0) = 0.8$. Compute $\sigma'(x_0)$.

2. What is the ReLU function $\text{ReLU}(x)$? More specifically,

   a. Write down the exact algebraic definition of $\text{ReLU}(x)$.
   b. Draw a graph of $\text{ReLU}(x)$.
   c. Suppose that the number x_0 is chosen so that $\text{ReLU}(x) = 0.8$. Compute $\text{ReLU}'(x)$.

3. Let's suppose that x and y represent two vectors of purely categorical data. Let's say

   $$x = [\text{red}, \text{yel}, \text{blu}, \text{blu}, \text{yel}, \text{yel}]$$
   $$y = [\text{blu}, \text{yel}, \text{red}, \text{blu}, \text{yel}, \text{yel}].$$

   How do you compute the Hamming distance between these two vectors and what is the value?

4. Consider the categorical data vector

   $$[\text{red}, \text{yel}],$$

   where the set of all possible values for the each entry is red, yel, and blu. Write down the one-hot encoding of that data vector.

5. Consider the vectors of *mixed* type

$$x = [\text{red}, \text{yel}, 8, 4]$$
$$y = [\text{blu}, \text{yel}, 2, 7].$$

Let's suppose that the numerical variable can take values from 0 to 10.

   a. What is the standard Gower distance between these vectors?
   b. Why might we like to use the Gower distance in K Nearest Neighbor algorithms?

6. Draw an expression graph for

$$f(x_1, x_2) = x_1^2 e^{-(x_1^2 + 2x_2^2)}.$$

Be sure to reuse node values as necessary.

7. Compute the convolution of the data D with the kernel K given by

$$D = \boxed{1}\boxed{1}\boxed{2}\boxed{2}\boxed{3}\boxed{3}$$

and

$$K = \boxed{2}\boxed{0}\boxed{2}.$$

8. Consider the two two-dimensional kernels

$$K_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad K_2 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -24 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

   a. Could these be appropriate for edge detection in image processing? Why or why not?
   b. What kind of difference might you expect in the behavior of these?

9. The neural network shown in figure 1 below consists of three layers:

   - the input layer,
   - one hidden layer, and
   - the output layer.

   Let's also suppose that the input layer has a ReLU activation and the output layer has a sigmoid activation.

   Note that the inputs are given. Use those inputs together with forward propagation to compute the value produced by this neural network.

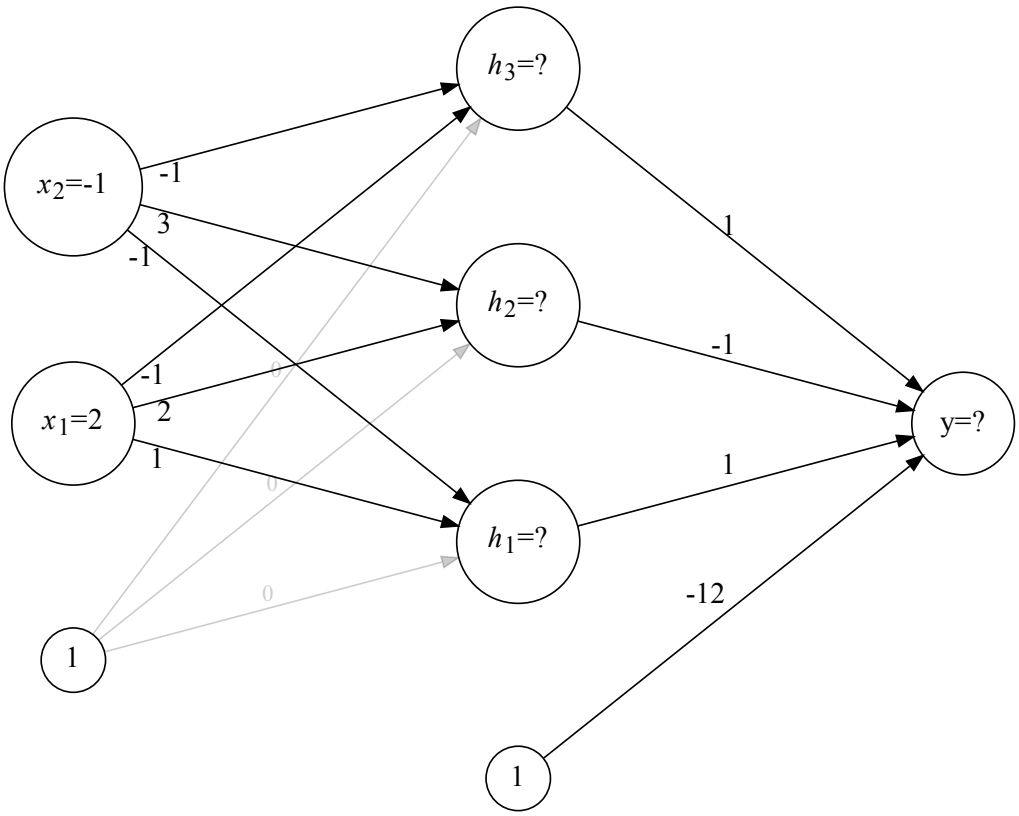

Figure 1: A neural network image