

INF3490 exercises - week 7 2013

Problem 1

Last week we used the activation function

$$g(h) = \begin{cases} 1, & h > 0 \\ 0, & \text{else} \end{cases}$$

Why is this not used with backpropagation?

Problem 2

In the lecture slides the backpropagation deltas are first presented as

$$\delta_k = (y_k - t_k) y_k (1 - y_k)$$

What does this tell us about which activation function that is used?

Problem 3

What is the minimum number of layers needed in order to approximate an arbitrary continuous function, and why?

Problem 4

Why do we use a validation set? Describe how do the three different cross-validation methods presented in the lecture slides work, and what their advantages and disadvantages are.

Problem 5

Design an MLP that would learn to hyphenate words correctly. You would have a dictionary that shows correct hyphenation examples for lots of words, and you would need to choose methods of encoding the inputs and outputs that say whether a hyphen is allowed or not between each pair of letters.