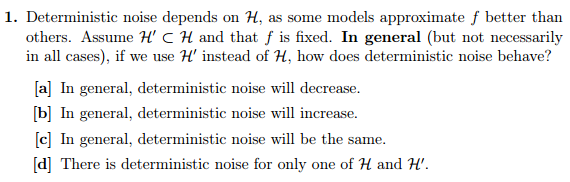
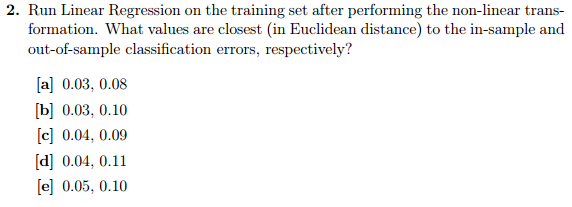
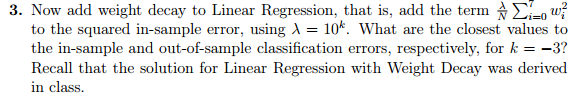
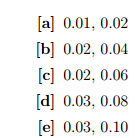
Learning from Data Homework 6

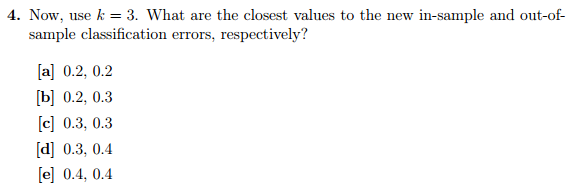


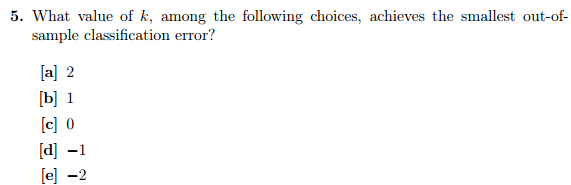


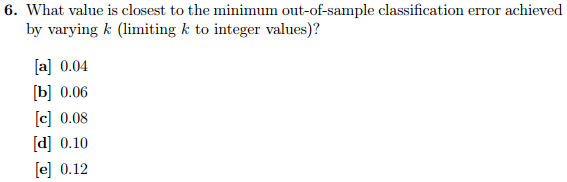


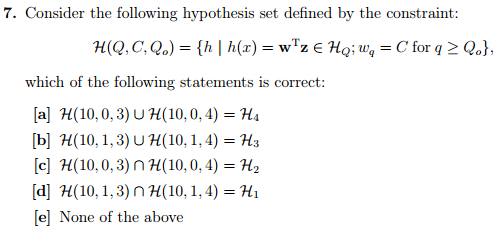


Performance out of sample improved from 0.84 to 0.8, but it is still closest to the exact same pair.

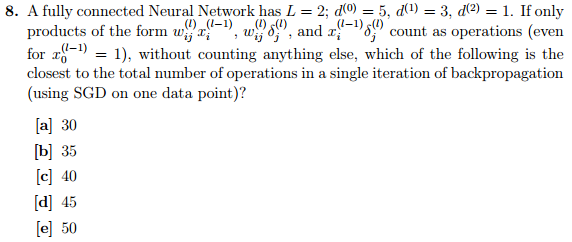








Correction: Thought was when dismissing C. Correct answer is C.



Notation: [source](http://web.cs.swarthmore.edu/~meeden/cs81/s10/BackPropDeriv.pdf)

is the activation amount put into nodes in layer

is the product of error and the derivation of the activation function in layer

is the weight of an output from layer to layer

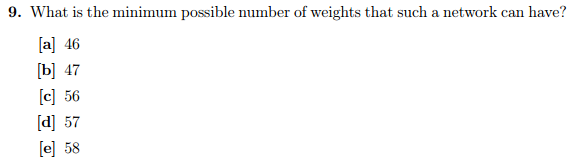
is the input layer

is the hidden layer

is the output layer

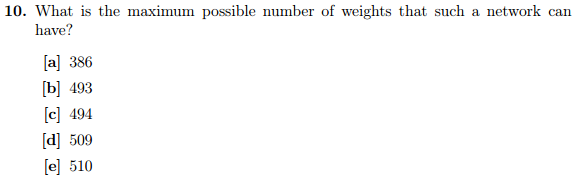
|  |  |
| --- | --- |
| # | Reasoning behind adding |
| 1 | ’s (1 node in this layer) |
| 3 | ’s nodes . can only map to 1 node, but can map to 3 (all nodes on layer 1), so 3\*1 |
| 3 | ’s nodes . can map to 3 nodes, to 1 per node |
| 15 | ’s nodes . maps to 3, to 5. |
| 15 | ’s nodes . can map to 3 nodes, to 5 per node. |
| 37 | Sum |

Correction: I forgot to include the 8 operations necessary to compute the output before SGD can be done (which is nonetheless part of the iteration). , D



with inputs at , hidden layer starting at , and outputs at . 2 hidden nodes per layer, including the intercept node . . 10 = # of nodes in input, 1 = # of nodes in every layer receiving activation, 2 = # of nodes sending activation in every layer, 18 = # of connections between two layers with 2 nodes sending and one receiving node in both.

is not possible because then one couldn’t have a neural network (one couldn’t have intercept nodes and traditional nodes in each hidden/input layer)



The above equation represents the number of weights in a neural network described by the specifications of the problem assuming that if you have fewer than 17 nodes per layer, you have the *same* number of nodes per layer. Latter assumption isn’t necessarily true, but seems to not impact the max configuration

reaches its maximum at with 510 weights.

Homework 6 Solutions

8/10 on first try

Incorrect Problems:

* #7, misunderstood notation (silly mistake, took as )
* #8, forgot to include operations necessary to compute output before backpropagation begins

