# Homework V

I. Consider the training data given below, x is the attribute and y is the class variable.

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| y | A | A | A | A | B | A | A | A | A | B | B | B | B | A | B | B | B | B |

a. What would be the classification of a test sample with x = 4.2 according to 1-NN?

b. What would be the classification of a test sample with x = 4.2 according to 3-NN?

c. What is the "leave-one-out" cross validation error of 1-NN. If you need to choose between two or more examples of identical distance, make your choice so that the number of errors is maximized. **[10 Points]**

II. We have data from a questionnaires survey (to ask people opinion) and objective testing with two attributes(acid durability and strength) to classify whether a special paper tissue is good or not. Here are the four training examples

| X1 = Acid durability ( in seconda) | X2 = Strength ( Kg/sq meter) | Y = Classification |
|---|---|---|
| 7 | 7 | Bad |
| 7 | 4 | Bad |
| 3 | 4 | Good |
| 1 | 4 | Good |

Now the factory produces a new tissue that pass laboratory test X1 = 3 and X2 = 7. Without another expensive survey, can we guess the classification of the new tissue using K-nearest neighbor algorithm using k= 3? **[10 Points]**

III. Draw a neural network that represents the function f(x1; x2; x3) defined below. You can only use two types of units: linear units and sign units. Recall that the linear unit takes as input weights and attribute values and outputs $w_0 + \sum_i w_i x_i$, while the sign unit outputs +1 if $w_0 + \sum_i w_i x_i$, > 0 and -1 otherwise.

| x1 | x2 | x3 | f(x1,x2,x3) |
|---|---|---|---|
| 0 | 0 | 0 | 10 |
| 0 | 0 | 1 | -5 |
| 0 | 1 | 0 | -5 |
| 0 | 1 | 1 | 10 |
| 1 | 0 | 0 | -5 |
| 1 | 0 | 1 | 10 |
| 1 | 1 | 0 | 10 |
| 1 | 1 | 1 | 10 |

You have to write down the precise numeric weights (e.g., -1, -0:5, +1, etc.) as well as the precise units used at each hidden and output node.

IV. Consider a two-layer feedforward ANN with two inputs a and b, one hidden unit c, and one output unit d. This network has five weights ( $w_{ca}$, $w_{cb}$, $w_{c0}$, $w_{dc}$, $w_{d0}$), where $w_{x0}$ represents the threshold weight for unit x. Initialize these weights to the values (.1,.1,.1,.1,.1), then give their values after each of the first two training iterations of the Backpropagation algorithm. Assume learning rate Ƞ =.3, momentum α = 0.9, incremental weight updates, and the following training examples: **[10 Points]**

| a | b | d |
|---|---|---|
| 1 | 0 | 1 |
| 0 | 1 | 0 |

V. Revise the BACKPROPAGATION algorithm in Table 4.2 in Tom Mitchell book so that it operates on units using the squashing function $tanh$ in place of the sigmoid function. That is, assume the output of a single unit is $o = tanh(\vec{w} * \vec{x})$ ). Give the weight update rule for output layer weights and hidden layer weights. Hint: $tanh'(x) = 1 - tanh^2(x)$.

**[10 Points]**