1. (10 pts total, 5 pts each) Consider the learning data shown in Figure 18.3 of your book (both 2nd & 3rd ed.). Your book (Section 18.3, "Choosing Attribute Tests") shows that *Gain(Patrons)* ≈ 0.541 while *Gain(Type)* = 0. Calculate *Gain (Alternate)* and *Gain(Hungry).*

a. (5 pts) *Gain (Alternate)* =

b. (5 pts) *Gain(Hungry)* =

2. (15 pts total, 5 pts each) Consider an ensemble learning algorithm that uses simple majority voting among M learned hypotheses. Suppose that each hypothesis has error $\varepsilon > 0$ and that the errors made by each hypothesis are independent of the others'. Show your work.

a. (5 pts) Calculate a formula for the error of the ensemble algorithm in terms of M and ε.

b. (5 pts) Evaluate it for the cases where M = 5, 10, and 20 and ε = 0.1, 0.2, and 0.4.

c. (5 pts) If the independence assumption is removed, is it possible for the ensemble error to be worse than ε? Produce either an example or a proof that it is not possible.

3. (35 pts total, 5 pts off for each wrong answer, but not negative) Label as TRUE/YES or FALSE/NO.

a. (5 pts) Suppose that you are given two weight vectors for a perceptron. Both vectors, w1 and w2, correctly recognize a particular class of examples. Does the vector w3 = w1 − w2 ALWAYS correctly recognize that same class?

b. (5 pts) Does the vector w4 = w1 + w2 ALWAYS correctly recognize that same class?

c. (5 pts) Does the vector w5 = cw1 where c = 42 ALWAYS correctly recognize the same class?

d. (5 pts) Does the vector w6 = dw2 where d = −117 ALWAYS correctly recognize the same class?

e. (5 pts) Now suppose that you are given two examples of the same class A, x1 and x2, where $x1 \neq x2$. Suppose the example x3 = 0.5x1 + 0.5x2 is of a different class B. Is there ANY perceptron that can correctly classify x1 and x2 into class A and x3 into class B?

f. (5 pts) Suppose that you are given a set of examples, some from one class A and some from another class B. You are told that there exists a perceptron that can correctly classify the examples into the correct classes. Is the perceptron learning algorithm ALWAYS guaranteed to find a perceptron that will correctly classify these examples?

g. (5 pts) An artificial neural network can learn and represent only linearly separable classes.

h. (5 pts) Learning in an artificial neural network is done by adjusting the weights to minimize the error, and is a form of gradient descent.

i. (5 pts) An artificial neural network is not suitable for learning continuous functions (function approximation or regression) because its transfer function outputs only 1 or 0 depending on the threshold.