CIS 419/519: Quiz 3

September 30, 2019

1. You are training a binary classifier which predicts whether a test taken by a patient indicates if they have a rare disease (True) or not (False). Which one of the following performance measures would you like to optimize?

   (a) Precision, because It is important not to have many false negative examples

   (b) Recall, because It is important not to have many false negative examples

   (c) Accuracy, because It is important to know how many correct predictions you have

   (d) Recall, because It is important not to have many false positive examples

2. You are given a dataset D with P positive examples and N negative examples. In which of the following cases is the entropy of D the largest?

   (a) P = 1, N =69

   (b) P = 35, N = 35

   (c) P = 70, N = 0

   (d) P = 15, N = 65

3. Determine the recall, precision, and accuracy of a binary classifier given that its performance is provided in the following contingency table:

		Actual Label	
		True	False
Predicted Label	True	75	50
	False	25	50

   (a) Recall = 0.75, Precision = 0.75, Accuracy = 0.625

   (b) Recall = 0.75, Precision = 0.6, Accuracy = 0.75

   (c) Recall = 0.75, Precision = 0.6, Accuracy = 0.625

   (d) Recall = 0.6, Precision = 0.75, Accuracy = 0.625

4. You are tasked with learning a new function over 10 Boolean variables; you believe that this function evaluates to True if and only if a subset of at these variables (you don't know which, and how many) is 1. Your friend says that they have a good learning algorithm that can learn linear threshold units and suggest that you use it. Is this a good choice?

   (a) Yes, since the class of LTUs over 10 variables can express all the functions you care about

   (b) No, since the class of LTUs over 10 variables cannot express all the functions you care about. You should use Decision Trees

   (c) Yes, since all Boolean functions can be represented as LTUs.

   (d) No, since only neural networks can express the type of functions you care about

5. We run the ID3 algorithm for learning decision trees on 800 instances $< (A, B, C, D), y >$ where y is a binary label and A, B, C, D are binary attributes. It so happens that:

   (i) Half the data points have A=0, and they split evenly between positive (y=1) and negative (y=0) examples. But when A=1, all the examples are positive.

   (ii) Half the data points have B=0, but only 100 of them are negative (y=0) and the rest are positive (y=1) examples. Similarly, when B=1, 100 of them are negative, and the rest are positive.

   (iii) C and D take only the value 1, in all the examples.

   Determine which of the following statements is correct:

   (a) 75% of the examples are positive and A is chosen to be the root node.

   (b) 75% of the examples are positive and B is chosen to be the root node.

   (c) 75% of the examples are positive and there is a tie between C and D on who is the root node.

   (d) 50% of the examples are positive and there is a tie between C and D on who is the root node.