CSCI/ARTI 8950 Machine Learning

Assignment Number 6: Due 4/22/2010 **(in class)**

1. **[15 points][FIN]** Consider the following training set of samples for machine learning:

Example	Sentence	Label
1	This test is easy	+
2	This problem is easy	+
3	Two is more than three	−
4	Swimming is easy	−

Apply the Naive Bayes method to the above examples. You should:

(a) Identify the vocabulary.

(b) For every word Wi in the vocabulary, compute P(Wi ‖ +) and P(Wi ‖ −). (Hint, many probabilities are equal)

(c) Use the Naive Bayes classifier to classify the following sentences:

i. This problem is more than easy

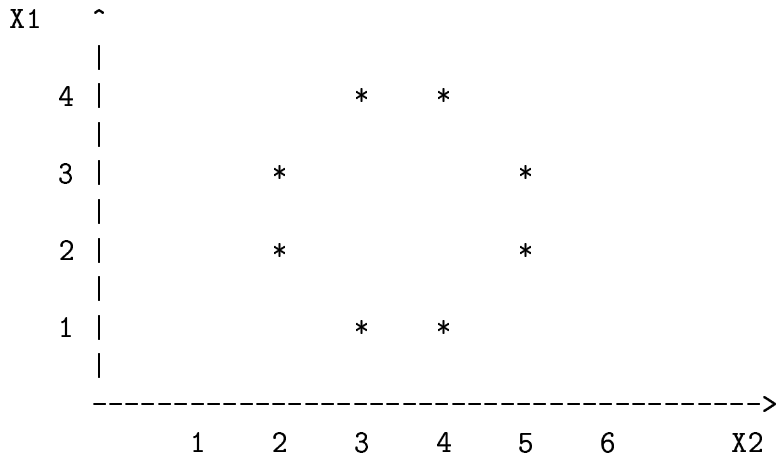
ii. No test is easy

2. **[20 points][FIN]**

(a) Propose a lazy version of the back-propagation algorithm for training neural networks. What are the advantages and disadvantages of your algorithm, compared to the original back-propagation algorithm?

(b) Propose an eager version of the nearest neighbour algorithm for classification. What are the advantages and disadvantages of your algorithm, compared to the original nearest neighbour algorithm?

3. **[15 points][FIN]** Consider the following diagram of a set of 8 instances for machine learning:



(a) Consider the hypothesis space H1 consisting of all possible **circles** in the plane (i.e. each hypothesis h in H1 is a circle which classifies all points in it as positive and all points outside it as negative). Does H1 shatter the given set of instances? Briefly justify your answer.

(b) Consider the hypothesis space H2 consisting of all possible **rectangles** in the plane. Does H2 shatter the given set of instances? Briefly justify your answer.

(c) Based **only** on your answers to parts (a) and (b) above, what can you conclude about the VC dimentions of H1 and H2?