

CSCI/ARTI 8950 Machine Learning

Assignment Number 5: Due 11/12/2013 (in class)

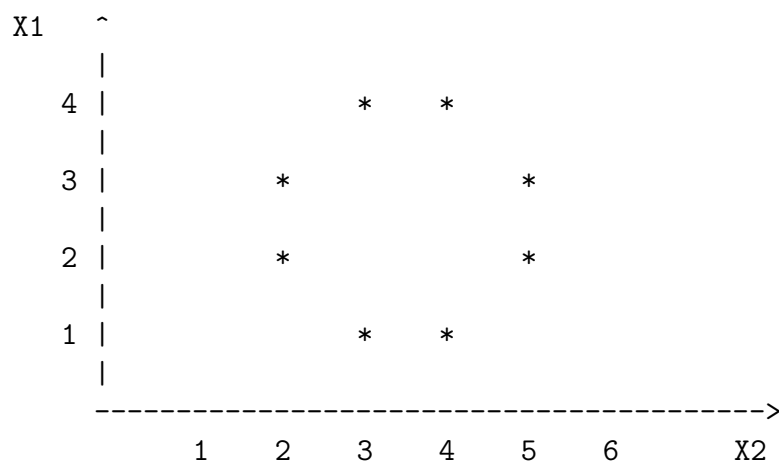
1. [20 points][FIN] Consider the following training set of samples for machine learning:

Example	A1	A2	A3	A4	label
1	1	2	2	2	+
2	1	1	1	1	−
3	2	3	2	1	+
4	1	3	3	3	−
5	3	1	2	1	+
6	1	1	1	2	+

The attributes **A1** through **A4** are integers with values in the range [1,2,3] each.

- (a) Give a classifier in the GIL format that can correctly classify all the training examples.
 - (b) How would the GIL classifier in Part (a) above classify the following examples: (1,2,2,3) and (3,2,1,1)?
 - (c) Of the following two methods, which do you think will be a more **efficient** learning method for this problem and why?
 - i. Using the GIL classifier system learning.
 - ii. Using ID3 to learn a decision tree.
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. [20 points][FIN] Consider the following diagram of a set of 8 instances for machine learning:



- Consider the hypothesis space H_1 consisting of all possible **circles** in the plane (i.e. each hypothesis h in H_1 is a circle which classifies all points in it as positive and all points outside it as negative). Does H_1 shatter the given set of instances? Briefly justify your answer.
- Consider the hypothesis space H_2 consisting of all possible **rectangles** in the plane. Does H_2 shatter the given set of instances? Briefly justify your answer.
- Based **only** on your answers to parts (a) and (b) above, what can you conclude about the VC dimensions of H_1 and H_2 ?