

CSCI/ARTI 8950 Machine Learning

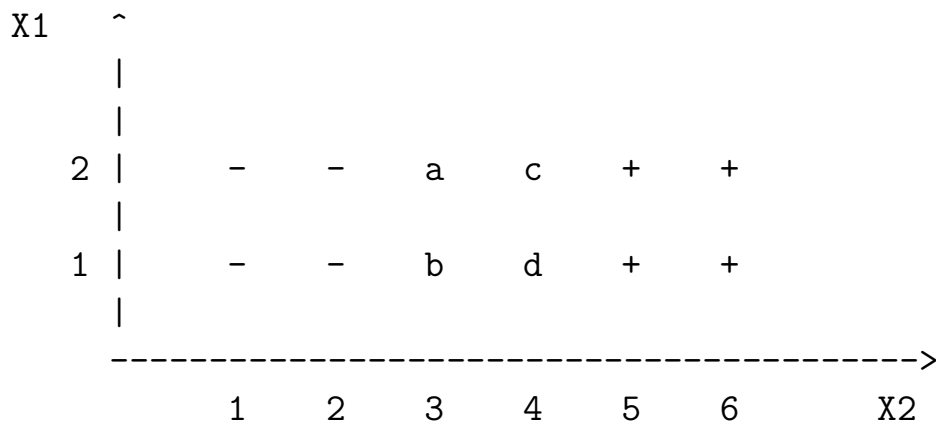
Assignment Number 3: Due 3/8/2022 (by elc)

1. [20 points][MID] 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	a
3	2	3	2	1	b
4	1	3	3	3	c
5	3	1	2	1	d
6	1	1	1	2	+

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

- (a) For the label assignment (a=-,b=+,c=+,d=-) how would the 1-Nearest Neighbor algorithm using Hamming Distance classify the following examples: (1,2,2,3) and (3,2,1,1)? The Hamming Distance is the number of positions that have different values in the two samples.
- (b) Propose a label assignment for a, b, c, and d that will make all samples correctly classifiable by the 3-Nearest Neighbor algorithm using Hamming Distance.
2. [20 points][MID] Consider the following diagram of a set of samples for machine learning:



- (a) For the label assignment (a=+,b=+,c=-,d=-) can all the given samples (including a, b, c, and d) be correctly classified by a properly trained (or computed in any possible way) Perceptron, whose inputs are X1,X2 and has 2 variable weights and a variable threshold? If your answer is YES, sketch one such Perceptron. If your answer is NO, briefly explain why.

- (b) Propose a label assignment for a, b, c, and d that makes all the given samples (including a, b, c, and d) correctly classifiable by the 3-nearest neighbor rule. This rule labels any sample with the majority label of its 3 nearest neighbors in Euclidean distance from the given samples (the sample itself is NOT considered one of the neighbors).

3. **[20 points][MID] Short answers please**

A hypothesis space is complete if it contains at least one hypothesis that can correctly classify all samples in any possible training set. An algorithm searches a hypothesis space completely if it finds all hypotheses in that space that can correctly classify all samples in a given training set.

- (a) Give the name of one algorithm that searches a complete space of hypotheses but searches it incompletely.
- (b) Give the name of one algorithm that searches an incomplete space of hypotheses and searches it incompletely.
- (c) Give one advantage to using support vector machines over perceptrons.
- (d) Give one advantage for using perceptrons over support vector machines.