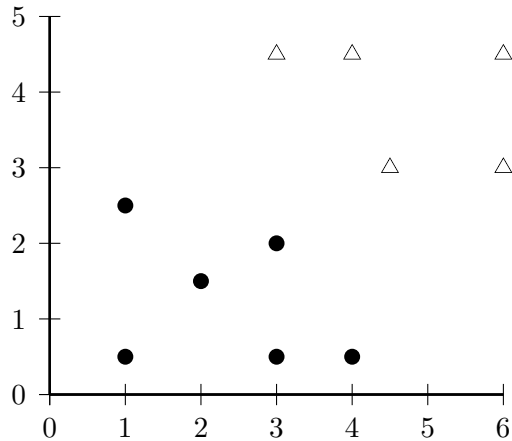# Information Retrieval: Assignment 5
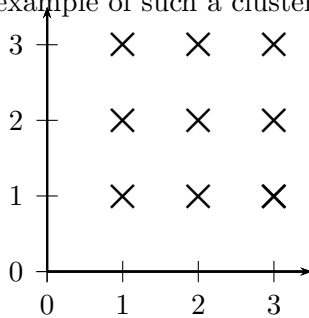
Problem 1. (20 points)

Indicate in the figure below what the linear maximum margin (SVM) classifier for the binary problem triangle vs dot is. Draw three lines: the two boundaries of the maximum margin and the maximum margin hyperplane. Which of the vectors are support vectors? You can solve this problem "visually" by drawing your solution into the figure.
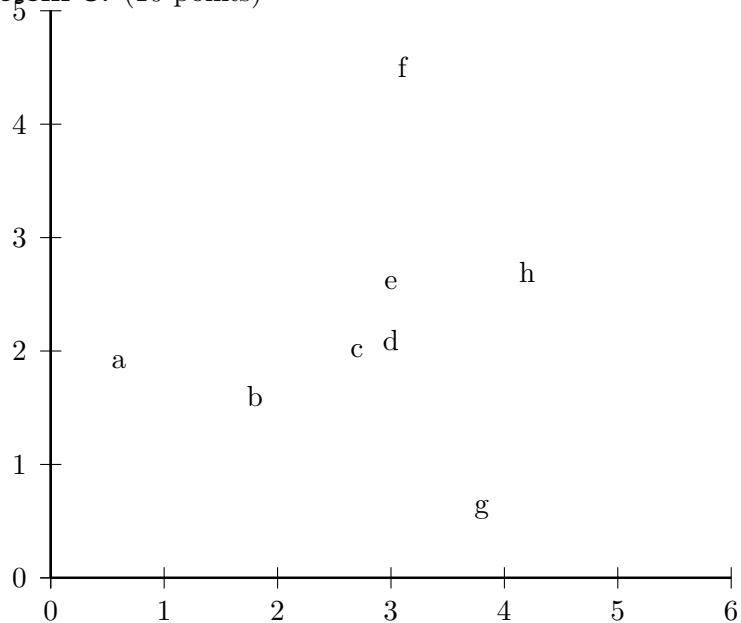

Problem 2. (10 points)

(i) Perform a 3-means clustering for the points below. Draw a different diagram for each iteration to show the assignments and the centroids. If a tie occurs during an assignment step, you can freely choose any of the possible assignments. (ii) There are several clusterings that 3-means can converge to in this case. Give an example of such a clustering that is different from the one in (i).
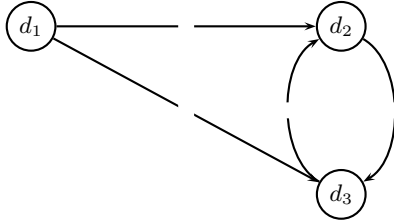
Problem 3. (10 points)

Compute (i) single-link and (ii) complete-link clusterings of the set of points and depict them as dendrograms. Make sure to indicate the merge value of each horizontal "merge" line (i.e., the similarity of the two clusters that are being merged in this step). Define the similarity of two points as $-(x_1-x_2)^2-(y_1-y_2)^2$. The points have the following coordinates: a: (0.6,1.9), b: (1.8,1.6), c: (2.7,2.0), d: (3.0,2.1), e: (3.0,2.6), f: (3.1,4.5), g: (3.8,0.6), h: (4.2,2.7).

Problem 4. (10 points)

advertiser	bid	CTR
A	$5.00	0.04
B	$1.00	0.1
C	$0.50	0.06
D	$2.00	0.03

Compute how much advertisers A, B, C, D have to pay for each click in a second price auction as described in class. The minimum amount per click is 0.005.

Problem 5. (10 points)

For this web graph, compute PageRank for each of the three pages. Assume that the PageRank teleport probability is 0.1. Hint: Using symmetries to simplify and solving with linear equations might be easier than using iterative methods.

Due date: Thursday, July 11, 2013, 12:15
Please turn in your assignment in class if possible. Email submissions are only accepted if you have a good reason why you cannot attend the review meeting. You will receive one extra point of credit if you use a staple or paper clip.