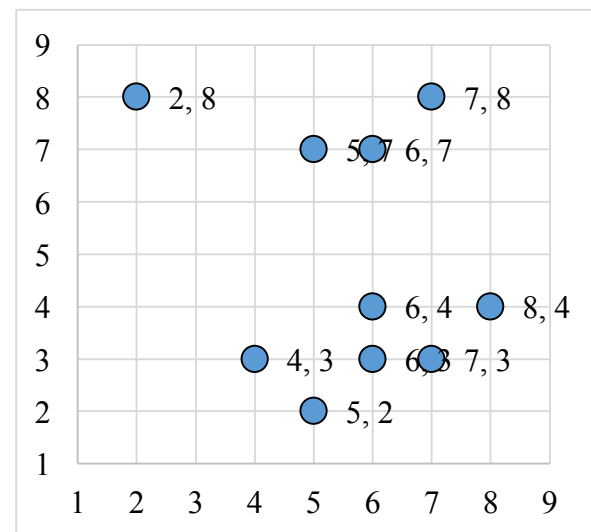# Homework 5

Handed Out: November 9, 2017
Due: November 28, 2017

General Instructions

- This assignment is due at 11:59 PM on the due date.
- We will be using Sakai (https://sakailogin.nd.edu/portal/site/FA17-CSE-40647-CX-01) for collecting this assignment. Contact TA if you face technical difficulties in submitting the assignment. We shall NOT accept any late submission!
- The homework MUST be submitted in pdf format. You can handwrite results and scan them into PDF. Name your pdf file as **YourNetid-HW5.pdf**.
- Please use Piazza if you have questions about the homework. Also feel free to send TA emails and come to office hours.

Clustering: K-Partitioning Methods and Kernel K-Means

Suppose we have 10 college *soccer* teams X1 to X10. We want to cluster them into 2 groups. For each *soccer* team, we have two features: One is # wins in Season 2016, and the other is # wins in Season 2017.

Team	# wins in Season 2016 (x-axis): x	# wins in Season 2017 (y-axis): y
X1	5	7
X2	6	7
X3	2	8
X4	7	8
X5	8	4
X6	6	4
X7	7	3
X8	6	3
X9	5	2
X10	4	3

[20'] Initialize with two centroids, (6, 4) and (6, 5). Use Manhattan distance as the distance metric. Please use K-Means to find two clusters.

[20'] Initialize with two centroids, (6, 4) and (6, 5). Use Euclidean distance as the distance metric. Please use K-Means to find two clusters.

[20'] Initialize with two centroids, (8, 7) and (2, 6). Use Manhattan distance as the distance metric. Please use K-Means to find two clusters.

[20'] Suppose we initialize with two medoids, (2, 8) and (8, 4). Use Euclidean distance as the distance metric. In K-Medoids clustering, given a non-medoid (5, 7), do we swap the medoid (2, 8) with (5, 7)?

[20'] Suppose the original two features are x and y. We use a kernel function to generate three new features: x^2, xy and y^2. Now we initialize with two centroids, (6, 4) and (6, 5), that are now (36, 24, 16) and (36, 30, 25). Use Manhattan distance as the distance metric in the new feature space. Please use Kernel K-Means to find two clusters.