Machine Learning *CS 689, Spring 2015*

HW11: Clustering, dimensionality reduction

Hand in via moodle at: `https://moodle.umass.edu/course/view.php?id=20836`. Remember that only PDF submissions are accepted. We encourage using LaTeX to produce your writeups. See `hw00.tex` for an example of how to do so. You can make a `.pdf` out of the `.tex` by running "`pdflatex hw00.tex`". You'll need mydefs.sty and notes.sty which can be downloaded from the course page.

1. Suppose you have N points. Consider a brute force algorithm for agglomerative clustering using the single-link (minimum distance) criteria. For a pair of sets A and B, to compute the distance between them the algorithm compares all pairs of distances between the elements of set A and set B to find the minimum. The time taken to do this is $O(|A||B|)$ where $|A|$ is the size of set A. What is the complexity of this algorithm (in "big O" notion) for clustering N points?

2. A dataset D has 64 points. Each point in D is a vector of length 6830, in other words there are 6830 features. If we do Principal Component Analyses (PCA) of this dataset D, how many principal components with non-zero variance would we get? Explain why?