# Introduction to Machine Learning - Homework 4

Prof. Alan Yuille

Spring 2014

**Due on Thursday 5/June. 2014. Hand in papers in class.**

**Question 1. K-means Algorithm**

Describe the K-means algorithm and its soft version. What is the difference between them?

**Question 2. EM Algorithm**

Briefly describe the EM algorithm. How does EM relate to K-means?

**Question 3. Inference and Learning Algorithms for probabilistic graphs**

Describe how to use dynamic programming to perform efficient inference on the probabilistic graph $p(\vec{y})$. (see handout lecture 15 section 3.1) And also describe how to do learning for model with latent variables $p(\vec{x}, \vec{h}|\lambda)$, where $\vec{x}$ is observed, $\vec{h}$ is latent and $\lambda$ is parameter. The distribution of latent variables $p(\vec{h})$ is a MRF. How is dynamic programming used for learning. (see handout lecture 15 section 3.2)

**Question 4. Multi-Class max-margin and Latent SVM.**

Describe how structure max-margin (structure SVM) can be applied to learning the parameters for a multi-class problem where the solution is of form $\vec{y}^* = \arg\max_{\vec{y}} \lambda \cdot \phi(\vec{x}, \vec{y})$, where $\vec{x}$ is the input and $\vec{y}$ is the output.

Discuss how the complexity of the graph structure (e.g., the number of closed loops) affects the difficulty of the learning.

Describe how to learn latent SVMs where some output variables $\vec{y}$ are observed, some variables $\vec{h}$ are hidden/latent, and the input data is $\vec{x}$. The solution is of form: $\vec{y}^*, \vec{h}^* = \arg\max_{\vec{y},\vec{h}} \lambda \cdot \phi(\vec{x}, \vec{y}, \vec{h})$.

How will the learning depend on the complexity of the graph structure and on the initial conditions?