Probabilistic and Unsupervised Learning (COMP0086)

**Description**

**Aims:**
This module provides students with an in-depth introduction to statistical modelling and unsupervised learning techniques. It presents probabilistic approaches to modelling and their relation to coding theory and Bayesian statistics. A variety of latent variable models will be covered including mixture models (used for clustering), dimensionality reduction methods, time series models such as hidden Markov models which are used in speech recognition and bioinformatics, independent components analysis, hierarchical models, and nonlinear models. The course will present the foundations of probabilistic graphical models (e.g. Bayesian networks and Markov networks) as an overarching framework for unsupervised modelling. We will cover Markov chain Monte Carlo sampling methods and variational approximations for inference. Time permitting, students will also learn about other topics in probabilistic (or Bayesian) machine learning.

**Learning outcomes:**
To be able to understand the theory of unsupervised learning systems; to have in-depth knowledge of the main models used in Unsupervised Learning; to understand the methods of exact and approximate inference in probabilistic models; to be able to recognise which models are appropriate for different real-world applications of machine learning methods.

**Content:**
- Basics of Bayesian learning and regression;
- Latent variable models, including mixture models and factor models;
- The Expectation-Maximisation (EM) algorithm;
- Time series, including hidden Markov models and state-space models;
- Spectral learning;
- Graphical representations of probabilistic models;
- Belief propagation, junction trees and message passing;
- Model selection, hyperparameter optimisation and

**Key information**

<table>
<thead>
<tr>
<th>Year</th>
<th>2019/20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credit value</td>
<td>15 (150 study hours)</td>
</tr>
<tr>
<td>Delivery</td>
<td>PGT L7, Campus-based</td>
</tr>
<tr>
<td>Reading List</td>
<td>View on UCL website</td>
</tr>
<tr>
<td>Tutor</td>
<td>Prof Maneesh Sahani</td>
</tr>
<tr>
<td>Term</td>
<td>Term 1</td>
</tr>
<tr>
<td>Timetable</td>
<td>View on UCL website</td>
</tr>
</tbody>
</table>

**Assessment**

- Written examination (main exam period): 50%
- Coursework: 17%
- Coursework: 17%
- Coursework: 16%

**Find out more**

For more information about the department, programmes, relevant open days and to browse other modules, visit [ucl.ac.uk](http://ucl.ac.uk)

**Disclaimer:** All information correct as of June 2019. Please note that aspects of the module may be subject to change. UCL will make best efforts to inform applicants of major changes.
Gaussian-process regression;

**Requisites:**
In order to be eligible to select this module, a student must be registered on a programme for which it is a formally-approved option or elective choice AND must have (i) a good background in statistics, calculus, linear algebra, and computer science; and (ii) good competency in Matlab or Octave (or be taking a class on Matlab/Octave, or be willing to learn it on their own).

**Audit:**
Any student or researcher at UCL meeting the requisite conditions is welcome to attend the lectures. Students who wish to formally register on the module should consult with the Module Leader.

**Preparatory Material:**
Students wishing to take the module should thoroughly review the maths in the cribsheet provided at www.gatsby.ucl.ac.uk/teaching/courses/ml1-2008/cribsheet.pdf before the start of the module.