

# Bài tập lớn **Tính Toán Khoa Học**

## Phần 1. Thông tin chung

### I. Mục đích

Bài tập lớn giúp sinh viên củng cố kiến thức môn học, rèn luyện kỹ năng sử dụng MatLab và ứng dụng chúng trong các bài toán thực tế.

### II. Yêu cầu bài tập lớn

Sinh viên/Nhóm sinh viên cần hoàn thành các bước sau:

- Nghiên cứu nội dung đề tài đã chọn và tìm đọc các tài liệu liên quan.
- Cài đặt mô hình và thử nghiệm kết quả trên MatLab.
- Viết báo cáo theo mẫu (20-35 trang A4).
- Làm slide báo cáo và bảo vệ.

### III. Đăng ký

Các bạn có thể lựa chọn đề tài theo nhóm (tối đa 2 người một nhóm) hoặc từng cá nhân. Nhóm/cá nhân nào lựa chọn xong đề tài thì gửi thư về [sangdv@soict.hust.edu.vn](mailto:sangdv@soict.hust.edu.vn) để được xác nhận.

### IV. Đánh giá chấm điểm

Điểm cộng tối đa là +5 cho điểm giữa kì và +2 cho điểm cuối kì. Cụ thể có các thang sau:

Mức	Điểm cộng giữa kỳ	Điểm cộng cuối kỳ
1	5	2
2	4	1
3	3	0.5
4	2	0
5	1	0
6	0	0

## Phần 2. Danh sách các đề tài

### I. Thị giác máy tính (Computer Vision)

#### 1. Nghiên cứu bài toán bám vết đối tượng trên camera tĩnh

Nội dung: Xây dựng module phát hiện và bám vết các vật chuyển động trên camera tĩnh.

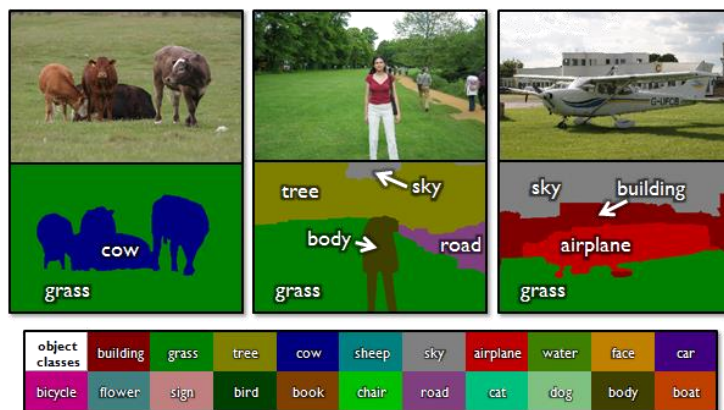
Một số ứng dụng:

- Xây dựng hàng rào ảo và cảnh báo khi có vật xâm phạm
- Đếm lưu lượng người ra vào tòa nhà, đếm lưu lượng xe cộ trên đường
- Cảnh báo đi ngược chiều
- Cảnh báo khi có đám đông tụ tập nơi công cộng
- Phát hiện vật bị bỏ quên ...
- Theo dõi bóng, đánh giá tốc độ sút bóng, quỹ đạo bóng trong bóng đá
- Đánh giá độ xoáy, tốc độ bóng, góc đánh trong môn thể thao golf

## 2. Nghiên cứu bài toán phân đoạn ảnh (Image segmentation)

Nội dung: Nghiên cứu bài toán phân đoạn ảnh thành các vùng khác nhau.

Ứng dụng: Ví dụ phân đoạn ảnh vệ tinh để tính toán tốc độ đô thị hóa, tốc độ xói mòn hay tốc độ chặt phá rừng...



## 3. Nghiên cứu bài toán khử nhiễu ảnh (Image denoising, Image restoration)

Nội dung: Nghiên cứu bài toán khử nhiễu cho các bức ảnh.

Ứng dụng: Ví dụ trong lĩnh vực y tế, các thiết bị chụp cắt lát bộ não, các thiết bị siêu âm thường có rất nhiều nhiễu. Bài toán khử nhiễu ảnh cho phép thu được hình ảnh sắc mịn, giúp cho bác sĩ dễ quan sát hơn.



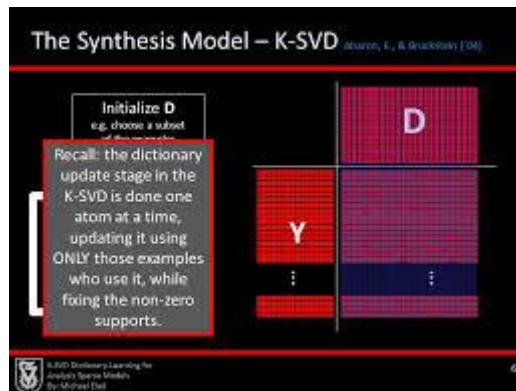
#### 4. Nghiên cứu bài toán phân loại ảnh (Image classification)

Nội dung: Nghiên cứu phương pháp phân chia một tập hợp ảnh thành các loại khác nhau dựa vào nội dung của các bức ảnh.

Ứng dụng: Trong các hệ thống truy vấn ảnh và lưu trữ dữ liệu ảnh như của google, facebook; Ứng dụng trong việc phát hiện mục tiêu và cảnh báo an ninh...

## II. Mã hóa thưa (sparse modeling) và ứng dụng

### 5. Mã hóa thưa và học từ điển dựa trên KSVD



Nội dung:

- Nghiên cứu KSVD và một số phương pháp học từ điển dựa trên KSVD như LC-KSVD và D-KSVD
- Ứng dụng chúng trong một số bài toán xử lý dữ liệu

### 6. Phương pháp Lasso và ứng dụng

Nội dung:

- Nghiên cứu phương pháp Lasso, Group Lasso

- Ứng dụng chúng trong một số bài toán xử lý dữ liệu lớn

## 7. Phương pháp Elastic Net và ứng dụng

Nội dung:

- Nghiên cứu phương pháp Elastic Net
- Ứng dụng chúng trong một số bài toán xử lý dữ liệu lớn

## 8. Phương pháp học từ điển Fisher Discriminative Dictionary Learning (FDDL) dựa trên tiêu chuẩn Fisher

Nội dung:

- Nghiên cứu phương pháp FDDL
- Ứng dụng chúng trong một số bài toán xử lý dữ liệu lớn
- 

$$\begin{matrix} \left[ \begin{array}{c} \text{blue} \\ \text{yellow} \\ \text{red} \\ \text{green} \\ \text{blue} \\ \text{red} \end{array} \right] = \left[ \begin{array}{cccccccccccc} \text{blue} & \text{yellow} & \text{red} & \text{green} & \text{blue} & \text{red} & \text{yellow} & \text{blue} & \text{yellow} & \text{red} & \text{blue} & \text{red} \\ \text{yellow} & \text{blue} & \text{red} & \text{green} & \text{blue} & \text{red} & \text{yellow} & \text{blue} & \text{yellow} & \text{red} & \text{blue} & \text{red} \\ \text{red} & \text{yellow} & \text{blue} & \text{green} & \text{blue} & \text{red} & \text{yellow} & \text{blue} & \text{yellow} & \text{red} & \text{blue} & \text{red} \\ \text{green} & \text{yellow} & \text{red} & \text{blue} & \text{blue} & \text{red} & \text{yellow} & \text{blue} & \text{yellow} & \text{red} & \text{blue} & \text{red} \\ \text{blue} & \text{yellow} & \text{red} & \text{green} & \text{blue} & \text{red} & \text{yellow} & \text{blue} & \text{yellow} & \text{red} & \text{blue} & \text{red} \\ \text{red} & \text{yellow} & \text{blue} & \text{green} & \text{blue} & \text{red} & \text{yellow} & \text{blue} & \text{yellow} & \text{red} & \text{blue} & \text{red} \end{array} \right] \left[ \begin{array}{c} \text{red} \\ \text{blue} \\ \text{yellow} \\ \text{blue} \end{array} \right] \\ \mathbf{y} \in \mathbb{R}^m & \mathbf{A} \in \mathbb{R}^{m \times n} & \mathbf{x} \in \mathbb{R}^r \end{matrix}$$

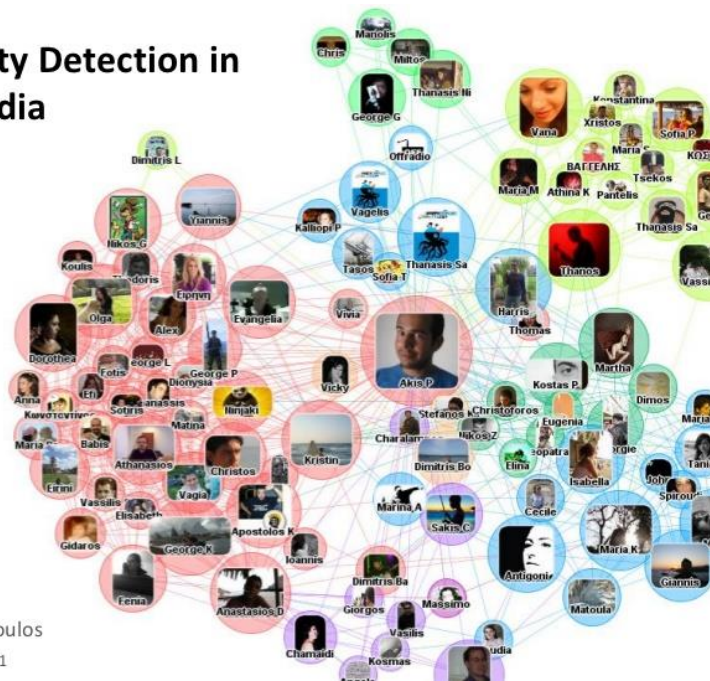
## III. Học máy (Machine Learning)

### 9. Khai phá mạng xã hội (Social Network Mining)

Nội dung:

- Nghiên cứu bài toán phân chia đồ thị và ứng dụng trong vấn đề khai phá các cộng đồng trong mạng xã hội
- Tìm hiểu lát cắt chuẩn
- Phân chia đồ thị dựa trên ma trận Laplac
- Thử nghiệm trên một số bộ dữ liệu lớn

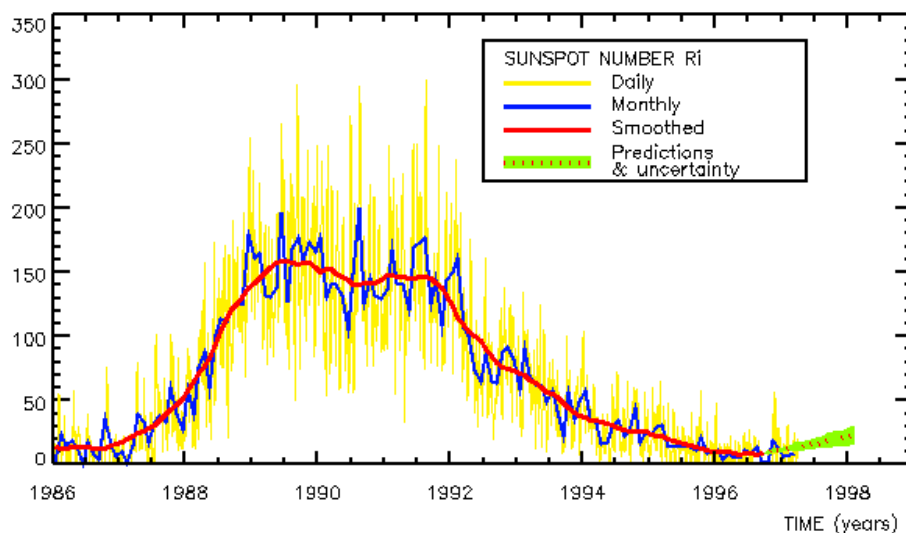
## Community Detection in Social Media



Symeon Papadopoulos  
CERTH-ITI, 22 June 2011

## 10. Nghiên cứu các mô hình dự báo dữ liệu chuỗi thời gian (Time series analysis)

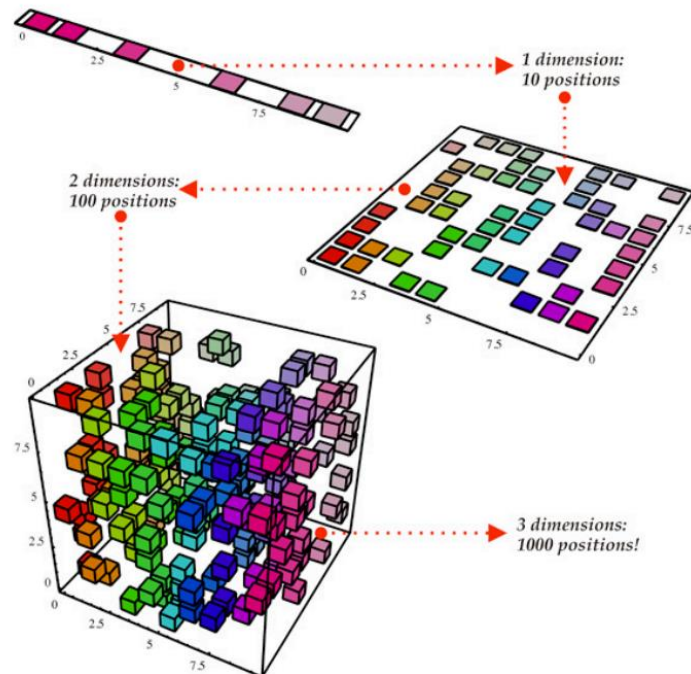
Nội dung: Nghiên cứu các mô hình phân tích dữ liệu chuỗi thời gian để dự báo sự thay đổi của dữ liệu trong tương lai. Ứng dụng trong việc dự báo giá vàng, chứng khoán, GDP, lượng mưa ... và rất nhiều lĩnh vực khác.



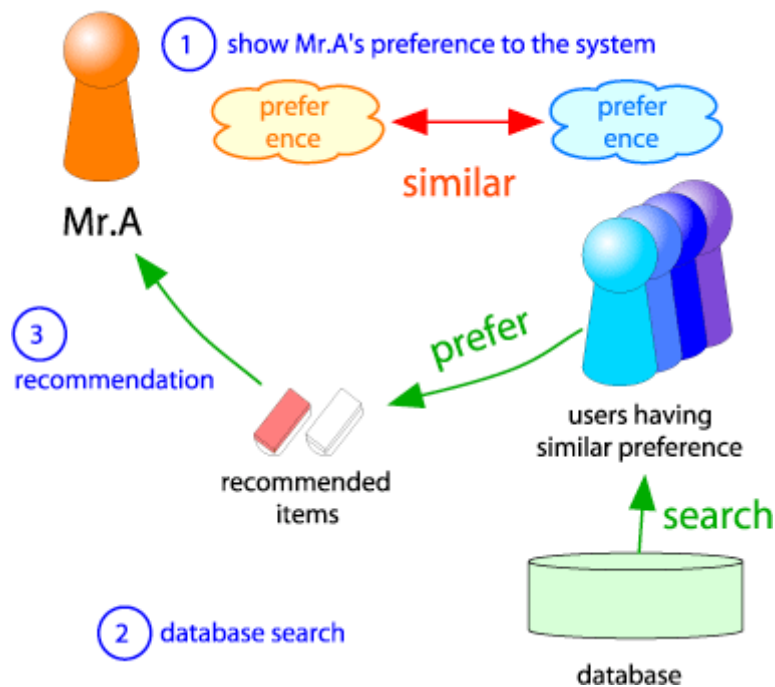
## 11. Nghiên cứu các phương pháp giảm chiều dữ liệu (dimension reduction) và ứng dụng.

Nội dung: Nghiên cứu cách xác định các đặc trưng quan trọng nhất trong dữ liệu đa chiều để phục vụ cho quá trình phân tích.

- Nghiên cứu một số phương pháp giảm chiều như PCA, SVD, CUR...
- Ứng dụng trong một số bài toán cụ thể như hệ thống gợi ý hay phương pháp mã hóa thưa để phân loại dữ liệu.



## 12. Hệ thống gợi ý (Recommendation System)



Nội dung: Hệ thống gợi ý nghiên cứu sở thích người dùng dựa trên dữ liệu thu thập được. Sau đó hệ thống sẽ phân tích các mặt hàng tiềm năng mà người dùng

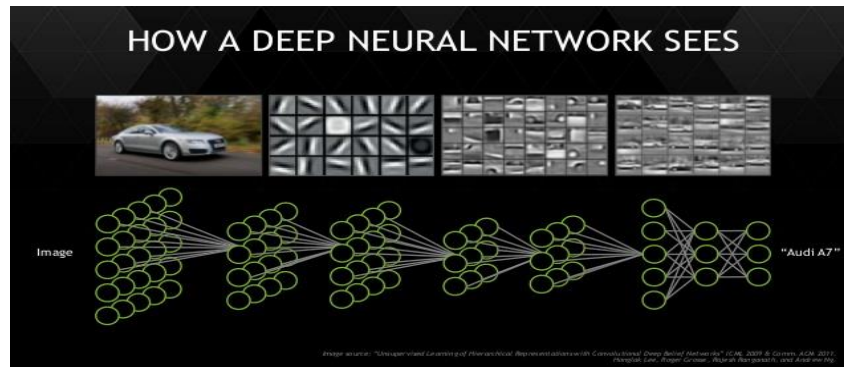


có thể thích để gợi ý cho họ. Ví dụ: Hệ thống gợi ý video của Youtube, gợi ý sách của amazon, gợi ý phim của Netflix ...

Nghiên cứu các phương pháp sử dụng trong hệ thống gợi ý:

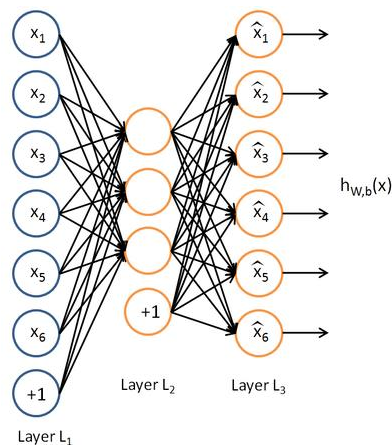
- Gợi ý dựa trên nội dung
- Gợi ý dựa trên học cộng tác (collaborative filtering)
- Mô hình nhân tố ẩn (latent factor model)

### 13. Mô hình CNN trong bài toán phân loại



Nội dung: Nghiên cứu mô hình học sâu hiện đại nhất hiện nay (Convolutional Neural Network) và ứng dụng nó trong một số bài toán phân loại dữ liệu.

### 14. Mô hình Autoencoder trong bài toán phân loại



Nội dung: Nghiên cứu mô hình học sâu tự mã hóa (autoencoder) và ứng dụng nó trong một số bài toán phân loại dữ liệu.

- Nghiên cứu mô hình autoencoder cơ bản
- Nghiên cứu mô hình denoising autoencoder
- Nghiên cứu mô hình sparse autoencoder

- Nghiên cứu mô hình stacked autoencoder

## **IV. Một số bài toán khác**

### **15. Xây dựng tool quản lý phòng và thời khóa biểu cho ĐHBKHN**

Nội dung:

- Cho phép nhập thời khóa biểu, lịch thi, thông tin phòng từ file excel
- Cho phép truy vấn phòng trống theo thời gian mong muốn (tuần, thứ, buổi, tiết hoặc ngày, buổi, tiết). Có thể truy vấn theo từng tòa nhà hoặc theo số lượng chỗ ngồi ...
- Cho phép kiểm tra lịch thi và thời khóa biểu có bị trùng phòng trùng tiết hay không.
- Cho phép kiểm tra các môn học đã sắp xếp đủ số tiết so với khối lượng yêu cầu hay chưa.