

1. Thông tin về giảng viên:

Họ và tên: **Phạm Thế Anh**
Chức danh, học hàm, học vị: PGS., TS. ngành Khoa học máy tính
Thời gian, địa điểm làm việc: VP Khoa CNTT & TT nhà A2, CSC
Địa chỉ liên hệ: Khoa CNTT & TT, Trường ĐHHĐ
Điện thoại: 0941070715 Email: phamtheanh@hdu.edu.vn

Họ và tên: **Lê Đình Nghiệp**
Chức danh, học hàm, học vị: Giảng viên, Thạc sỹ ngành CNTT
Địa điểm làm việc: VP Khoa CNTT & TT nhà A2, CSC
Địa chỉ liên hệ: Khoa CNTT & TT, Trường ĐHHĐ
Điện thoại: 0912698679 Email: ledinhnghiep@hdu.edu.vn

Họ và tên: **Lê Thị Hồng**
Chức danh, học hàm, học vị: Giảng viên, Thạc sỹ chuyên ngành KHMT
Địa điểm làm việc: VP Khoa CNTT & TT nhà A2, CSC
Địa chỉ liên hệ: Khoa CNTT & TT, Trường ĐHHĐ
Điện thoại: 0948921740 Email: lethihong@hdu.edu.vn

Họ và tên: **Nguyễn Đình Định**
Chức danh, học hàm, học vị: Giảng viên, Thạc sỹ ngành CNTT
Địa điểm làm việc: VP Khoa CNTT & TT nhà A2, CSC
Địa chỉ liên hệ: Khoa CNTT & TT, Trường ĐHHĐ
Điện thoại: 0948911357 Email: nguyendinhding@hdu.edu.vn

2. Thông tin chung về học phần:

- Tên ngành/khoá đào tạo: ĐH Công nghệ thông tin
- Tên học phần: **Học máy**
- Số tín chỉ: 3
- Học kỳ: 5
- Học phần: Bắt buộc: Tự chọn:
- Các môn học tiên quyết: Cấu trúc dữ liệu & giải thuật, Toán rời rạc, Lập trình nâng cao.

- Các môn học kế tiếp: Các môn chuyên ngành
- Giờ tín chỉ đối với các hoạt động:

Lý thuyết	Bài tập	Thực hành	Tự học
25	40	0	135

Địa chỉ bộ môn phụ trách học phần: BM Khoa học máy tính, khoa CNTT&TT phòng 203 nhà A2 cơ sở chính, Trường ĐH Hồng Đức.

3. Nội dung học phần

Nội dung học phần: Học phần cung cấp các phương pháp cơ bản để trích xuất tự động thông tin từ dữ liệu và sử dụng thông tin đã trích xuất để suy diễn và giải quyết các bài toán ra quyết định và phân lớp; các phương pháp biểu diễn tri thức và chiến lược học máy phổ biến (học có giám sát và học không có giám sát); các bộ phân lớp tiêu biểu (Decision Tree, SVM, Noron Networks); các thư viện mã nguồn mở (OpenCV, Python) chuyên dùng cho lập trình và phát triển các ứng dụng học máy.

Năng lực đạt được: người học có kỹ năng thu thập, xử lý dữ liệu, phân tích và vận dụng được các phương pháp học máy để giải quyết các bài toán về phân lớp và nhận dạng.

4. Mục tiêu của học phần

Mục tiêu	Mô tả	Chuẩn đầu ra CTĐT
1. Kiến thức	1.1. Nắm vững nguyên tắc và các thành phần của một hệ thống học máy	C11
	1.2. Hiểu được các phương pháp cơ bản thu thập dữ liệu và xử lý dữ liệu	C7, C10
	1.3. Hiểu được các phương pháp xây dựng và huấn luyện một bộ phân lớp học không giám sát (Decision Tree, SVM, Noron Networks), các kỹ thuật hồi quy tuyến tính	C5, C8, C9
	1.4. Hiểu rõ các phương pháp xây dựng và huấn luyện một bộ phân lớp học không giám sát (K-means, Meanshift)	C8, C9
2. Kỹ năng	2.1. Vận dụng thành thạo các phương pháp thu thập dữ liệu và xử lý dữ liệu phục vụ huấn luyện các bộ phân lớp	C16
	2.2. Vận dụng các bộ phân lớp để xây dựng một ứng dụng phân lớp, nhận dạng cụ thể	C14, C18, C20
	2.3. Sử dụng thành thạo các công cụ và môi trường phát triển ứng dụng liên quan lĩnh vực học máy (Winform, C++, OpenCV, Python)	C7, C18

	2.4.	Thành thạo các ứng dụng văn phòng để xây dựng các báo cáo theo yêu cầu đặc thù của môn học, xây dựng các bài thuyết trình lời cuốn, hấp dẫn và kỹ năng bảo vệ sản phẩm trước tập thể.	C16
3. Thái độ	3.1.	Cầu thị, ham học hỏi, chủ động tìm hiểu về các nội dung của học phần học máy; tích cực nghiên cứu, trao đổi với giảng viên về các thuật toán phân lớp, phân cụm, các vấn đề cài đặt, lập trình thuật toán liên quan.	C22
	3.2.	Đam mê thực hành, cần cù, sáng tạo, chăm chỉ cài đặt các thuật toán học máy và vận dụng xây dựng các sản phẩm thực tiễn.	C22
	3.3.	Tích cực trau dồi, chia sẻ kiến thức của môn học với sinh viên trong lớp, chủ động tổ chức các giờ tự học theo nhóm, làm bài tập lớn theo nhóm.	C16, C23
4. Năng lực	4.1.	Tổ chức, triển khai quy trình thu thập, xử lý dữ liệu; phân tích và vận dụng được các phương pháp học máy để giải quyết các bài toán về phân lớp và nhận dạng.	C14, C21, C22

5. Chuẩn đầu ra học phần

TT	Kết quả mong muốn đạt được	Mục tiêu	Chuẩn đầu ra CTĐT
A	Có kiến thức tổng quan các hệ thống học máy và phân lớp, các bộ phân lớp phổ biến (SVM, Cây quyết định, mạng nơ ron nhân tạo), các phương pháp học không có giám sát, các kỹ thuật hồi quy tuyến tính	1.1, 1.2, 1.3, 1.4, 4.1	C5, C7, C10, C8, C9, C11
B	Có kỹ năng ứng dụng các bộ phân lớp và khai thác các thư viện lập trình để xây dựng các ứng dụng thực tiễn liên quan đến bài toán phân loại, nhận dạng đối tượng. Thành thạo lập báo cáo tài liệu và xây dựng các bài thuyết trình dự án liên quan.	2.1, 2.2, 2.3, 2.4, 4.1	C7, C14, C16, C18, C20
C	Nhận thức được ý nghĩa môn học, đam mê học hỏi và tìm hiểu về các nội dung môn học, sẵn sàng chia sẻ kiến thức, trau dồi kỹ năng.	3.1, 3.2, 3.3	C16, C22, C23

6. Nội dung chi tiết học phần

CHƯƠNG 1. GIỚI THIỆU VỀ HỌC MÁY

1.1. Các khái niệm cơ bản

1.2. Các phương pháp học máy

- 1.2.1. Học máy có giám sát
- 1.2.2. Học máy không giám sát
- 1.3. Dữ liệu huấn luyện
 - 1.3.1. Vai trò của dữ liệu trong học máy
 - 1.3.2. Thu thập và gán nhãn dữ liệu
 - 1.3.3. Chuẩn hóa dữ liệu
- 1.4. Các ứng dụng học máy
 - 1.3.1. Phân lớp, nhận dạng đối tượng
 - 1.3.2. Hồi quy, dự đoán
 - 1.3.3. Khai phá thông tin từ dữ liệu

CHƯƠNG 2. PHƯƠNG PHÁP HỌC MÁY CÓ GIÁM SÁT

- 2.1 Kỹ thuật hồi quy tuyến tính hàm một biến (Linear Regression)
 - 2.1.1. Xây dựng mô hình và hàm mục tiêu
 - 2.1.2. Quá trình học tham số
 - 2.1.3. Sai số học
- 2.2. Kỹ thuật hồi quy tuyến tính hàm nhiều biến
 - 2.2.1. Xây dựng mô hình và hàm mục tiêu nhiều biến
 - 2.2.2. Quá trình học tham số
 - 2.2.3. Đánh giá mô hình học
- 2.3. Kỹ thuật hồi quy Logistic
 - 2.3.1. Chuẩn bị dữ liệu
 - 2.3.2. Xây dựng hàm logic hồi quy
 - 2.3.3. Xây dựng mô hình hồi quy
 - 2.3.4. Huấn luyện mô hình
 - 2.3.5. Phân lớp đa đối tượng (multi-class)
- 2.4. Cây quyết định
 - 2.4.1. Biểu diễn giả thuyết
 - 2.4.2. Chọn thuộc tính làm bộ phân loại
 - 2.4.3. Các vấn đề trong học bằng cây quyết định
- 2.5. Các bộ phân lớp Boosting
 - 2.5.1. Bộ phân lớp Discrete AdaBoost
 - 2.5.2. Bộ phân lớp Real AdaBoost
 - 2.5.3. Bộ phân lớp LogitBoost
 - 2.5.4. Bộ phân lớp Gentle AdaBoost
- 2.6. Bộ phân lớp Random Trees
 - 2.6.1. Quá trình tạo Random Trees

- 2.6.2. Xử lý dữ liệu thiếu
- 2.6.3. Các tiêu chuẩn rẽ nhánh
- 2.6.4. Hàm hồi quy
- 2.6.5. Quá trình phân lớp
- 2.7. Bộ phân lớp SVM (Support Vector Machine)
 - 2.7.1. Giới thiệu bài toán phân lớp 2 nhãn (2-class)
 - 2.7.2. Bộ học SVM tuyến tính (Linear SVM)
 - 2.7.3. Bộ học SVM mềm (Soft margin)
 - 2.7.4. Mở rộng SVM cho phân lớp bài toán 1 nhãn (1-class)
 - 2.7.5. Mở rộng SVM cho phân lớp bài toán nhiều nhãn (N-class)
- 2.8. Bộ phân lớp Bayesian
 - 2.8.1. Học khái niệm
 - 2.8.2. Giả thuyết MAP
 - 2.8.3. Bộ phân loại đơn giản
 - 2.8.4. Bộ phân loại tối ưu
- 2.9. Phân lớp dựa trên mạng nơ ron nhân tạo
 - 2.9.1. Khái niệm nơ ron nhân tạo
 - 2.9.2. Các hàm mục tiêu (Loss Functions)
 - 2.9.3. Mạng nơ ron Perceptron
 - 2.9.3. Mạng Perceptron đa lớp với giải thuật lan truyền ngược
 - 2.9.4. Các mạng nhân chập CNN (Convolutional Neuron Networks)
 - 2.9.5. Các mạng CNN học sâu (Deep CNNs)

CHƯƠNG 3. PHƯƠNG PHÁP HỌC MÁY KHÔNG CÓ GIÁM SÁT

- 3.1. Giới thiệu
- 3.2. Các thuật toán phân cụm
 - 3.2.1. Thuật toán K-means
 - 3.2.2. Thuật toán MeanShift
- 3.3. Nguyên lý phân tích thành phần PCA (Principal Components Analysis)
 - 3.3.1. Vai trò của PCA trong máy học
 - 3.3.2. Xây dựng PCA dùng phương pháp hiệp phương sai (covariance)
 - 3.3.3. Mối tương quan giữa PCA và K-means

CHƯƠNG 4. ỨNG DỤNG CASE STUDY

- 4.1. Giới thiệu bài toán nhận dạng mặt người
- 4.2. Chuẩn bị dữ liệu và trích chọn đặc trưng
- 4.3. Huấn luyện các bộ học: Decision Tree, SVM, Boosting trees
- 4.4. Phân lớp và đánh giá

7. Học liệu:

+ *Học liệu bắt buộc*

[1] Lê Hoài Bắc, Tô Hoài Việt, Cơ sở trí tuệ nhân tạo, NXB Khoa học và kỹ thuật, 2014.

[2] Hoàng Văn Dũng, Giáo trình Nhận dạng và Xử lý ảnh, Nhà xuất bản Khoa học và Kỹ thuật, 2018.

+ *Học liệu tham khảo*

[3] Phạm Thế Anh (chủ biên), Nguyễn Mạnh An, Đỗ Năng Toàn. Giáo trình Xử lý ảnh. NXB Giáo dục, 2017.

8. Hình thức tổ chức dạy học

8.1. *Lịch trình chung:*

Nội dung	Hình thức tổ chức dạy học môn học						Tổng
	LT	TH	BT/ TL	Tự học, tự N/C	Tư vấn của GV	KT ĐG	
CHƯƠNG 1. GIỚI THIỆU VỀ HỌC MÁY	4	0	2	20	3		30
1.1. Các khái niệm cơ bản 1.2. Các phương pháp học máy 1.2.1. Học máy có giám sát 1.2.2. Học máy không giám sát	2	0	0	4	1		
1.3. Dữ liệu huấn luyện 1.3.1. Vai trò của dữ liệu trong học máy 1.3.2. Thu thập và gán nhãn dữ liệu 1.3.3. Chuẩn hóa dữ liệu	1	0	1	8	1		
1.4. Các ứng dụng học máy 1.3.1. Phân lớp, nhận dạng đối tượng 1.3.2. Hồi quy, dự đoán 1.3.3. Khai phá thông tin từ dữ liệu	1	0	1	8	1		
CHƯƠNG 2. PHƯƠNG PHÁP HỌC MÁY CÓ GIÁM SÁT	15	0	28	80	9	04	136

<p>2.1 Kỹ thuật hồi quy tuyến tính hàm một biến (Linear Regression)</p> <p>2.1.1. Xây dựng mô hình và hàm mục tiêu</p> <p>2.1.2. Quá trình học tham số</p> <p>2.1.3. Sai số học</p>	1		2	8	1		
<p>2.2. Kỹ thuật hồi quy tuyến tính hàm nhiều biến</p> <p>2.2.1. Xây dựng mô hình và hàm mục tiêu nhiều biến</p> <p>2.2.2. Quá trình học tham số</p> <p>2.2.3. Đánh giá mô hình học</p>	1.5		2	8	1	Bài kiểm tra số 01	
<p>2.3. Kỹ thuật hồi quy Logistic</p> <p>2.3.1. Chuẩn bị dữ liệu</p> <p>2.3.2. Xây dựng hàm logic hồi quy</p> <p>2.3.3. Xây dựng mô hình hồi quy</p> <p>2.3.4. Huấn luyện mô hình</p> <p>2.3.5. Phân lớp đa đối tượng (multi-class)</p>	1.5		3	8	1		
<p>2.4. Cây quyết định</p> <p>2.4.1. Biểu diễn giả thuyết</p> <p>2.4.2. Chọn thuộc tính làm bộ phân loại</p> <p>2.4.3. Các vấn đề trong học bằng cây quyết định</p>	1.5		2.5	8	1		
<p>2.5. Các bộ phân lớp Boosting</p> <p>2.5.1. Bộ phân lớp Discrete AdaBoost</p> <p>2.5.2. Bộ phân lớp Real AdaBoost</p> <p>2.5.3. Bộ phân lớp LogitBoost</p> <p>2.5.4. Bộ phân lớp Gentle AdaBoost</p>	2		4	8	1	Bài kiểm tra số 02	
<p>2.6. Bộ phân lớp Random Trees</p> <p>2.6.1. Quá trình tạo Random Trees</p> <p>2.6.2. Xử lý dữ liệu thiếu</p> <p>2.6.3. Các tiêu chuẩn rẽ nhánh</p> <p>2.6.4. Hàm hồi quy</p> <p>2.6.5. Quá trình phân lớp</p>	1.5		2.5	8	1		

<p>2.7. Bộ phân lớp SVM (Support Vector Machine)</p> <p>2.7.1. Giới thiệu bài toán phân lớp 2 nhãn (2-class)</p> <p>2.7.2. Bộ học SVM tuyến tính (Linear SVM)</p> <p>2.7.3. Bộ học SVM mềm (Soft margin)</p> <p>2.7.4. Mở rộng SVM cho phân lớp bài toán 1 nhãn (1-class)</p> <p>2.7.5. Mở rộng SVM cho phân lớp bài toán nhiều nhãn (N-class)</p>	2		4	12	1		
<p>2.8. Bộ phân lớp Bayesian</p> <p>2.8.1. Học khái niệm</p> <p>2.8.2. Giả thuyết MAP</p> <p>2.8.3. Bộ phân loại đơn giản</p> <p>2.8.4. Bộ phân loại tối ưu</p>	1.5		3	8	1	Bài kiểm tra giữa kỳ	
<p>2.9. Phân lớp dựa trên mạng nơ ron nhân tạo</p> <p>2.9.1. Khái niệm nơ ron nhân tạo</p> <p>2.9.2. Các hàm mục tiêu (Loss Functions)</p> <p>2.9.3. Mạng nơ ron Perceptron</p> <p>2.9.3. Mạng Perceptron đa lớp với giải thuật lan truyền ngược</p> <p>2.9.4. Các mạng nhân chập CNN (Convolutional Neuron Networks)</p> <p>2.9.5. Các mạng CNN học sâu (Deep CNNs)</p>	2.5		5	12	1		
CHƯƠNG 3. PHƯƠNG PHÁP HỌC MÁY KHÔNG CÓ GIÁM SÁT	5	0	8	20	2	01	36
<p>3.1. Giới thiệu</p> <p>3.2. Các thuật toán phân cụm</p> <p>3.2.1. Thuật toán K-means</p> <p>3.2.2. Thuật toán MeanShift</p>	2.5		4	10	1	Bài kiểm tra số 03	

3.3. Nguyên lý phân tích thành phần PCA 3.3.1. Vai trò của PCA trong máy học 3.3.2. Xây dựng PCA dùng phương pháp hiệp phương sai (covariance) 3.3.3. Mối tương quan giữa PCA và K-means	2.5		4	10	1		
CHƯƠNG 4. ỨNG DỤNG CASE STUDY	1	0	3	15	5	01	25
4.1. Giới thiệu bài toán nhận dạng mặt người 4.2. Chuẩn bị dữ liệu và trích chọn đặc trưng 4.3. Huấn luyện các bộ học: Decision Tree, SVM, Boosting trees 4.4. Phân lớp và đánh giá	1		3	15		Bài kiểm tra số 04	Báo cáo bài tập nhóm
Tổng	25	0	40	135	19	5	224

8.2. Lịch trình cụ thể cho từng nội dung:

Nội dung tuần 1 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	2 tiết Phòng học	CHƯƠNG 1. GIỚI THIỆU VỀ HỌC MÁY 1.1. Các khái niệm cơ bản 1.2. Các phương pháp học máy 1.2.1. Học máy có giám sát 1.2.2. Học máy không giám sát	- Hiểu được mục tiêu môn học và nguyên lý học động của một hệ thống học máy - Hiểu rõ sự khác biệt giữa phương pháp học có giám sát và học không giám sát - Hiểu được các ứng dụng thực tế của môn học	- Đọc tài liệu [2], [3] (Phần nội dung Các bộ phân lớp phổ biến) - Sử dụng Internet, Google để tra cứu các tài liệu khác.	A B

Lý thuyết	2 tiết Phòng học	<p>1.3. Dữ liệu huấn luyện</p> <p>1.3.1. Vai trò của dữ liệu trong học máy</p> <p>1.3.2. Thu thập và gán nhãn dữ liệu</p> <p>1.3.3. Chuẩn hóa dữ liệu</p> <p>1.4. Các ứng dụng học máy</p> <p>1.3.1. Phân lớp, nhận dạng đối tượng</p> <p>1.3.2. Hồi quy, dự đoán</p> <p>1.3.3. Khai phá thông tin từ dữ liệu</p>	<p>- Hiểu được vai trò của dữ liệu đối với các hệ thống học máy, phân biệt dữ liệu có nhãn và không có nhãn</p> <p>- Nắm được quy trình thu thập và gán nhãn dữ liệu, các kỹ thuật chuẩn hóa dữ liệu (đặc biệt dữ liệu ảnh)</p> <p>- Nắm được vai trò của học máy trong các ứng dụng thực tiễn (nhận dạng, chẩn đoán bệnh, tư vấn,...)</p>	- Đọc tài liệu [1], [3] (Phần nội dung Cây quyết định)	A B
Bài tập và Thảo luận	2 tiết Phòng học	<p>- Yêu cầu SV đưa ra các ứng dụng trong thực tế có ứng dụng học máy (học có nhãn và không nhãn)</p> <p>- SV phân tích vai trò của các học máy trong các ứng dụng thảo luận</p> <p>- Thảo luận về xu hướng phát triển của học máy, định hướng nghề nghiệp</p>	<p>- Hiểu rõ bản chất các hệ thống ứng dụng đưa ra</p> <p>- Nhận thực được sự ảnh hưởng của học máy và xu hướng công nghệ</p> <p>- Liên hệ với bản thân để có kế hoạch theo dõi đam mê lĩnh vực học máy</p>		A
Tự học	20 tiết học tại nhà, thư viện	<p>- Tìm hiểu về hệ thống học máy, sự khác nhau giữa học có nhãn và không nhãn</p> <p>- Tìm kiếm và tải một số bộ dữ liệu mẫu có sẵn, dùng các công cụ (Matlab, Excel) để phân tích và tiền xử lý dữ liệu</p>	<p>- Làm quen với một số bộ dữ liệu thường dùng để minh họa các hệ thống học máy</p> <p>- Làm quen với các kỹ thuật xử lý dữ liệu</p>		B C
KT-ĐG					

Tư vấn	Phòng bộ môn	Tư vấn về các vấn đề trong nội dung Chương 1 và các vấn đề liên quan	Thành thạo các kiểu bài tập trên cơ sở nội dung Chương 1	SV chuẩn bị các câu hỏi cần giải đáp	A B C
--------	--------------	--	--	--------------------------------------	-------------

Nội dung tuần 2 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	2 tiết Phòng học	<p>CHƯƠNG 2. PHƯƠNG PHÁP HỌC MÁY CÓ GIÁM SÁT</p> <p>2.1 Kỹ thuật hồi quy tuyến tính hàm một biến (Linear Regression)</p> <p>2.1.1. Xây dựng mô hình và hàm mục tiêu</p> <p>2.1.2. Quá trình học tham số</p> <p>2.1.3. Sai số học</p> <p>2.2. Kỹ thuật hồi quy tuyến tính hàm nhiều biến</p> <p>2.2.1. Xây dựng mô hình và hàm mục tiêu nhiều biến</p>	<ul style="list-style-type: none"> - Hiểu được ý nghĩa, chức năng của các mô hình hồi quy (tuyến tính) - Nắm được các khái niệm liên quan (hàm mục tiêu, sai số, biến độc lập, phụ thuộc) - Vận dụng để xây dựng mô hình hồi quy cho một số bộ dữ liệu 	<ul style="list-style-type: none"> - Đọc tài liệu [1], [2] (Phần Máy phân loại vector hỗ trợ) - Sử dụng Internet, Google để tra cứu các tài liệu khác. 	A B
Bài tập và Thảo luận	3 tiết Phòng học	<ul style="list-style-type: none"> - Giải các bài tập về hồi quy tuyến tính cho hàm một biến thông qua dữ liệu đơn giản (dự báo thời tiết) - Giảng viên minh họa trên máy tính một bộ dữ liệu thực tế (AirQuality) và kết quả của một số mô hình khác nhau - Thảo luận về kết quả của mỗi mô hình và sai số của phép hồi quy 	<ul style="list-style-type: none"> - Hiểu rõ bản chất của quá trình hồi quy và cập nhật trọng số mô hình - Thành thạo giải các bài tập về hồi quy đối với dữ liệu đơn giản - Biện luận và giải thích được sai số trên tập AirQuality 	<ul style="list-style-type: none"> - Tải bộ dữ liệu minh họa mô hình hồi quy AirQuality¹ 	A B

¹ http://rpubs.com/Nitika/linearRegression_Airquality

Tự học	20 tiết học tại nhà, thư viện	<ul style="list-style-type: none"> - Tìm hiểu từ Internet về các mô hình hồi quy một biến (mở rộng cho trường hợp dữ liệu phi tuyến) - Tìm kiếm và tải một số bộ dữ liệu mẫu có sẵn, dùng các công cụ (Matlab, Excel, OpenCV, Python) để chạy các mô hình hồi quy có sẵn - Vẽ biểu đồ hàm hồi quy và đánh giá kết quả 	<ul style="list-style-type: none"> - Làm quen với các công cụ hỗ trợ xây dựng mô hình hồi quy trực quan - Rèn luyện tư duy phản biện và đánh giá kết quả 	<ul style="list-style-type: none"> - Tìm hiểu về các công cụ hỗ trợ trong Excel, Matlab - Tự tìm thêm các bộ dữ liệu minh họa khác 	A B C
KT-ĐG	1 tiết Phòng học	Bài kiểm tra số 01: thi viết, thời gian 50 phút (1 tiết)	Kiểm tra kiến thức về xử lý và phân tích dữ liệu, bài tập về hồi quy đơn giản	SV chuẩn bị giấy kiểm tra và máy tính cầm tay	A B
Tư vấn	Phòng bộ môn	Tư vấn về các vấn đề của nội dung bài giảng, phạm vi và các trường hợp ứng dụng mô hình hồi quy, các dạng bài tập hồi quy	Thành thạo các dạng bài tập về hồi quy, hiểu được vai trò và cách dùng mô hình hồi quy	SV chuẩn bị các câu hỏi cần giải đáp	A B C

Nội dung tuần 3 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	2 tiết Phòng học	<p>CHƯƠNG 2. PHƯƠNG PHÁP HỌC MÁY CÓ GIÁM SÁT</p> <p>2.2.2. Quá trình học tham số</p> <p>2.2.3. Đánh giá mô hình học</p> <p>2.3. Kỹ thuật hồi quy Logistic</p> <p>2.3.1. Chuẩn bị dữ liệu</p> <p>2.3.2. Xây dựng hàm logic hồi quy</p> <p>2.3.3. Xây dựng mô hình hồi quy</p> <p>2.3.4. Huấn luyện mô hình</p> <p>2.3.5. Phân lớp đa đối tượng</p>	<ul style="list-style-type: none"> - Hiểu được nguyên tắc hồi quy tuyến tính nhiều biến - Phân tích, đánh giá độ chính xác, sai số của các mô hình - Nắm được nguyên tắc phân lớp nhiều đối tượng dựa - Vận dụng để xây dựng mô hình hồi quy cho một số bộ dữ liệu 	<ul style="list-style-type: none"> - Đọc tài liệu [1], [2] (Phần Máy phân loại vector hỗ trợ) - Tìm hiểu một số bộ dữ liệu nhiều biến - Sử dụng Internet, Google để tra cứu các 	A B

		(multi-class)		tài liệu khác.	
Bài tập và Thảo luận	4 tiết Phòng học	<ul style="list-style-type: none"> - Giải các bài tập về hồi quy tuyến tính cho hàm hai biến thông qua dữ liệu đơn giản (dự báo thời tiết với hai biến độc lập) - Giảng viên minh họa trên máy tính một bộ dữ liệu Boston. - Thảo luận về kết quả của mỗi mô hình và sai số của phép hồi quy 	<ul style="list-style-type: none"> - Hiểu rõ bản chất của quá trình hồi quy nhiều biến và vai trò hàm mục tiêu - Thành thạo giải các bài tập về hồi quy nhiều biến đối với dữ liệu đơn giản - Biện luận và giải thích được sai số trên tập dữ liệu Boston 	- Tải bộ dữ liệu minh họa mô hình hồi quy Boston ²	A B
Tự học	20 tiết học tại nhà, thư viện	<ul style="list-style-type: none"> - Tìm hiểu từ Internet về các mô hình hồi quy nhiều biến (mở rộng cho trường hợp dữ liệu phi tuyến) - Tìm kiếm và tải một số bộ dữ liệu mẫu có sẵn, dùng các công cụ (Matlab, Excel, OpenCV, Python) để chạy các mô hình hồi quy có sẵn - Vẽ biểu đồ hàm hồi quy và đánh giá kết quả 	<ul style="list-style-type: none"> - Làm quen với các công cụ hỗ trợ xây dựng mô hình hồi quy trực quan - Rèn luyện tư duy phản biện và đánh giá kết quả 	<ul style="list-style-type: none"> - Tìm hiểu về các công cụ hỗ trợ trong Excel, Matlab - Tự tìm thêm các bộ dữ liệu minh họa khác 	A B C
KT-ĐG					
Tư vấn	Phòng bộ môn	Tư vấn về các vấn đề của nội dung bài giảng, phạm vi và các trường hợp ứng dụng mô hình hồi quy, các dạng bài tập hồi quy	Thành thạo các dạng bài tập về hồi quy, hiểu được vai trò và cách dùng mô hình hồi quy	SV chuẩn bị các câu hỏi cần giải đáp	A B C

Nội dung tuần 4 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn	Chuẩn đầu ra
-----------	-------	----------------	-----------------	------------------	--------------

² <https://fderyckel.github.io/machinelearningwithr/mlr.html>

				bị	HP
Lý thuyết	2 tiết Phòng học	<p>CHƯƠNG 2. PHƯƠNG PHÁP HỌC MÁY CÓ GIÁM SÁT</p> <p>2.4. Cây quyết định</p> <p>2.4.1. Biểu diễn giả thuyết</p> <p>2.4.2. Chọn thuộc tính làm bộ phân loại</p> <p>2.4.3. Các vấn đề trong học bằng cây quyết định</p> <p>2.5. Các bộ phân lớp Boosting</p> <p>2.5.1. Bộ phân lớp Discrete AdaBoost</p>	<ul style="list-style-type: none"> - Hiểu được nguyên tắc học có giám sát - Hiểu được thuật toán cây quyết định (ID3, C4.5), các độ đo Gain, Entropy. - Hiểu được các vấn đề khi huấn luyện cây quyết định (over fitting, under fitting). - Hiểu được thuật toán mở rộng Discrete AdaBoost - Vận dụng để huấn luyện cây QĐ, Discrete AdaBoost cho bộ dữ liệu cụ thể 	<ul style="list-style-type: none"> - Đọc tài liệu [2], [3] (Phần Cây quyết định) - Sử dụng Internet, Google để tra cứu các tài liệu khác. 	A B
Bài tập và Thảo luận	3 tiết Phòng học	<ul style="list-style-type: none"> - Giải các bài tập xây dựng cây quyết định thông qua dữ liệu đơn giản (dự đoán giới tính thông qua các đặc trưng cơ thể) - Bài tập xây dựng bộ phân lớp Discrete AdaBoost từ 3 cây quyết định - Thảo luận về kết quả Discrete AdaBoost với cây quyết định 	<ul style="list-style-type: none"> - Thành thạo giải các bài tập liên quan thuật toán ID3 hoặc C4.5, Discrete AdaBoost - Biện luận và giải thích được sự ưu việt, điểm yếu của cây QĐ và Discrete AdaBoost 	<ul style="list-style-type: none"> - Đọc tài liệu [2], [3] (Phần Cây quyết định) - Tìm hiểu các bài tập mẫu trong tài liệu tham khảo 	A B
Tự học	20 tiết học tại nhà, thư viện	<ul style="list-style-type: none"> - Tìm tài liệu tiếng Anh về ID3 và Discrete AdaBoost, đọc và dịch, viết báo cáo bằng tiếng Việt - Giải các bài tập khác về ID3 và Discrete AdaBoost 	<ul style="list-style-type: none"> - Hình thành kỹ năng tự tìm và làm việc với tài liệu Tiếng Anh, kỹ năng viết báo cáo và tổng hợp - Củng cố kiến thức về thuật toán cây QĐ 		A B C

KT-ĐG	1 tiết Phòng học	Bài kiểm tra số 02: thi viết, thời gian 50 phút (1 tiết)	Kiểm tra kiến thức về cây quyết định, các thuật toán ID3, C4.5	SV chuẩn bị giấy kiểm tra và máy tính cầm tay	A B
Tư vấn	Phòng bộ môn	Tư vấn về các vấn đề của nội dung bài giảng, tiềm năng ứng dụng của cây QĐ và kỹ thuật khắc phục các vấn đề khi huấn luyện, tư vấn các kiểu bài tập thường gặp	Hiểu rõ hơn về bản chất cây QĐ, thành thạo các dạng bài tập về cây QĐ	SV chuẩn bị các câu hỏi cần giải đáp	A B C

Nội dung tuần 5 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	2.5 tiết Phòng học	<p>CHƯƠNG 2. PHƯƠNG PHÁP HỌC MÁY CÓ GIÁM SÁT</p> <p>2.5.2. Bộ phân lớp Real AdaBoost</p> <p>2.5.3. Bộ phân lớp LogitBoost</p> <p>2.5.4. Bộ phân lớp Gentle AdaBoost</p> <p>2.6. Bộ phân lớp Random Trees</p> <p>2.6.1. Quá trình tạo Random Trees</p>	<ul style="list-style-type: none"> - Hiểu được sự khác nhau giữa các phiên bản cải tiến thuật toán Real AdaBoost, Gentle AdaBoost, và LogitBoost. - Hiểu được ý tưởng xây dựng rừng ngẫu nhiên - Vận dụng để huấn luyện các bộ phân lớp cho bộ dữ liệu cụ thể 	<ul style="list-style-type: none"> - Đọc tài liệu [2] (Phần Kỹ thuật Boosting) - Sử dụng Internet, Google để tra cứu các tài liệu khác. 	A B
Bài tập và Thảo luận	3.5 tiết Phòng học	<ul style="list-style-type: none"> - Giải các bài tập xây dựng các thuật toán Boosting đề cập trong bài giảng - Thảo luận, phân tích, đánh giá ưu nhược điểm của mỗi thuật toán 	<ul style="list-style-type: none"> - Thành thạo các dạng bài tập về Boosting - Biện luận và giải thích được sự ưu việt, điểm yếu của các bộ phân lớp đã học 	<ul style="list-style-type: none"> - Đọc tài liệu [2] (Phần Kỹ thuật Boosting) - Tìm hiểu các bài tập mẫu trong tài liệu tham khảo 	A B

Tự học	20 tiết học tại nhà, thư viện	<ul style="list-style-type: none"> - Tìm tài liệu tiếng Anh về Real AdaBoost, Gentle AdaBoost, và LogitBoost, đọc và dịch, viết báo cáo bằng tiếng Việt - Giải các bài tập khác về Boosting 	<ul style="list-style-type: none"> - Hình thành kỹ năng tự tìm và làm việc với tài liệu Tiếng Anh, kỹ năng viết báo cáo và tổng hợp - Củng cố kiến thức về thuật toán Boosting 		A B C
KT-ĐG					
Tư vấn	Phòng bộ môn	Tư vấn về các vấn đề của nội dung bài giảng, tiềm năng ứng dụng của Boosting, tư vấn các kiểu bài tập thường gặp	Hiểu rõ hơn về bản chất Boosting, thành thạo các dạng bài tập về Boosting	SV chuẩn bị các câu hỏi cần giải đáp	A B C

Nội dung tuần 6 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	2 tiết Phòng học	<p>CHƯƠNG 2. PHƯƠNG PHÁP HỌC MÁY CÓ GIÁM SÁT</p> <p>2.6.1. Quá trình tạo Random Trees</p> <p>2.6.2. Xử lý dữ liệu thiếu</p> <p>2.6.3. Các tiêu chuẩn rẽ nhánh</p> <p>2.6.4. Hàm hồi quy</p> <p>2.6.5. Quá trình phân lớp</p> <p>2.7. Bộ phân lớp SVM (Support Vector Machine)</p> <p>2.7.1. Giới thiệu bài toán phân lớp 2 nhãn (2-class)</p> <p>2.7.2. Bộ học SVM tuyến tính (Linear SVM)</p>	<ul style="list-style-type: none"> - Hiểu được ý tưởng xây dựng rừng ngẫu nhiên và kỹ thuật voting (từ các weak classifiers) - Nắm được thuật toán SVM phân lớp nhị phân (hai đối tượng) - Nắm được bản chất các hàm kernel dùng trong SVM (tuyến tính, phi tuyến) 	<ul style="list-style-type: none"> - Đọc tài liệu [2] (Phần Máy phân loại vector hỗ trợ) và [3] (Phần Bộ phân lớp Random Forest) - Sử dụng Internet, Google để tra cứu các tài liệu khác. 	A B
Bài tập và Thảo	4 tiết Phòng học	<ul style="list-style-type: none"> - Giải các bài tập xây dựng các thuật toán rừng ngẫu nhiên - Biến đổi toán học để chứng 	<ul style="list-style-type: none"> - Thành thạo các dạng bài tập về rừng ngẫu nhiên, 	<ul style="list-style-type: none"> - Tìm hiểu các bài tập 	A B

luận		<p>minh hàm nhân của thuật toán SVM</p> <ul style="list-style-type: none"> - Yêu cầu SV tìm mối liên hệ qua phương trình toán học giữa hồi quy và SVM - Thảo luận, phân tích, đánh giá ưu nhược điểm của mỗi thuật toán 	<p>hàm nhân SVM</p> <ul style="list-style-type: none"> - Hiểu rõ cơ sở toán học của SVM - Biện luận và giải thích được sự ưu việt, điểm yếu của các bộ phân lớp đã học 	<p>mẫu trong tài liệu tham khảo</p>	
Tự học	20 tiết học tại nhà, thư viện	<ul style="list-style-type: none"> - Cài đặt các thư viện (python, opencv, C++) và thực hành huấn luyện các bộ SVM, Random Trees 	<ul style="list-style-type: none"> - Hình thành kỹ lập trình, vận dụng các thư viện liên quan - củng cố kiến thức về SVM, Random Trees và vận dụng để huấn luyện một bộ phân lớp cụ thể 	<ul style="list-style-type: none"> -Cài đặt các thư viện (python, opencv, C++) - Tảibộ dữ liệu huấn luyện trong tài liệu tham khảo 	A B C
KT-ĐG					
Tư vấn	Phòng bộ môn	<p>Tư vấn về các vấn đề của nội dung bài giảng, tiềm năng ứng dụng của SVM, Random Trees, tư vấn các kiểu bài tập thường gặp, các ứng dụng trong thực tế</p>	<p>Hiểu rõ hơn về bản chất SVM, Random Trees, thành thạo các dạng bài tập về Random Trees, SVM, các ứng dụng của SVM</p>	<p>SV chuẩn bị các câu hỏi cần giải đáp</p>	A B C

Nội dung tuần 7 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	1.5 tiết Phòng học	<p>CHƯƠNG 2. PHƯƠNG PHÁP HỌC MÁY CÓ GIÁM SÁT</p> <p>2.7.3. Bộ học SVM mềm (Soft margin)</p> <p>2.7.4. Mở rộng SVM cho phân lớp bài toán 1 nhãn (1-class)</p>	<ul style="list-style-type: none"> - Hiểu được nguyên tắc phân lớp 1-lớp tương và n-lớp tương từ bộ SVM nhị phân (one-versus-one và one-versus-rest) 	<ul style="list-style-type: none"> - Đọc tài liệu [2] (Phần Máy phân loại vector hỗ trợ) và tài liệu [3] 	A B

		<p>2.7.5. Mở rộng SVM cho phân lớp bài toán nhiều nhãn (N-class)</p> <p>2.8. Bộ phân lớp Bayesian</p> <p>2.8.1. Học khái niệm</p> <p>2.8.2. Giả thuyết MAP</p>	<ul style="list-style-type: none"> - Vận dụng nguyên lý bỏ phiếu (voting) sử dụng trong SVM - Nắm được cơ sở toán học và mô hình phân lớp Bayesian 	<p>(Phần bộ phân lớp SVM)</p> <ul style="list-style-type: none"> - Sử dụng Internet, Google để tra cứu các tài liệu khác. 	
Bài tập và Thảo luận	3.5 tiết Phòng học	<ul style="list-style-type: none"> - Giải các bài tập đơn giản về SVM 1-lớp và 3-lớp để hiểu được ý tưởng thuật toán - Yêu cầu SV phân tích độ phức tạp (tính toán, bộ nhớ) và độ chính xác của các kỹ thuật one-vs-one, one-vs-rest - Thảo luận các vấn đề khi huấn luyện SVM 1 lớp, giải pháp xử lý - Thảo luận các vấn đề khi dữ liệu bị thiếu, sai, không chính xác 	<ul style="list-style-type: none"> - Thành thạo các dạng bài tập về SVM 1-lớp và nhiều lớp - Đánh giá độ phức tạp và độ chính xác của SVM 1-lớp và nhiều lớp. - Nhận thức được các vấn đề thường gặp về dữ liệu huấn luyện và cách khắc phục 	<ul style="list-style-type: none"> - Tìm hiểu các bài tập mẫu trong tài liệu tham khảo 	A B
Tự học	20 tiết học tại nhà, thư viện	<ul style="list-style-type: none"> - Tìm tài liệu phân tích các vấn đề về dữ liệu (thiếu, không chính xác, ...) và giải pháp khắc phục - Tìm hiểu cơ sở toán cho mô hình mạng Bayesian, lĩnh vực ứng dụng 	<ul style="list-style-type: none"> - Hình thành kỹ năng tự tìm kiếm tài liệu và học trên Internet - củng cố kiến thức về xây dựng và huấn luyện bộ phân lớp khi thiếu dữ liệu - củng cố kiến thức về thuật toán SVM, Bayesian 	<ul style="list-style-type: none"> - Tìm hiểu tài liệu từ các nguồn miễn phí trên Internet - Đọc thêm trong tài liệu môn học 	A B C
KT-ĐG	1 tiết Phòng học	Bài kiểm tra giữa kỳ: thi viết, thời gian 50 phút (1 tiết)	Kiểm tra kiến thức tổng hợp về cây quyết định, hồi quy, SVM	SV chuẩn bị giấy kiểm tra và máy tính cầm tay	A B C
Tư vấn	Phòng bộ môn	Tư vấn về các vấn đề về xử lý dữ liệu khi huấn luyện, các	- củng cố kiến thức và kỹ năng về các	SV chuẩn bị các câu	A

		model SVM 1-lớp, nhiều lớp, Bayesian	bộ phân lớp SVM, Bayesian	hỏi cần giải đáp	B C
--	--	--------------------------------------	---------------------------	------------------	--------

Nội dung tuần 8 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	3 tiết Phòng học	<p>CHƯƠNG 2. PHƯƠNG PHÁP HỌC MÁY CÓ GIÁM SÁT</p> <p>2.8.2. Giả thuyết MAP</p> <p>2.8.3. Bộ phân loại đơn giản</p> <p>2.8.4. Bộ phân loại tối ưu</p> <p>2.9. Phân lớp dựa trên mạng nơ ron nhân tạo</p> <p>2.9.1. Khái niệm nơ ron nhân tạo</p> <p>2.9.2. Các hàm mục tiêu (Loss Functions)</p>	<ul style="list-style-type: none"> - Hiểu được nguyên lý xác suất MAP để xây dựng mô hình Bayesian - Hiểu được nguyên lý hoạt động của nơ ron nhân tạo và mạng nơ ron nhân tạo - Nắm được một số hàm mục tiêu cơ bản (Sigmoid, Softmax) 	<ul style="list-style-type: none"> - Đọc tài liệu [1] [2] (Mạng Nơ ron) và [3] (Phản mạng nơ ron nhân tạo) - Đọc thêm tại địa chỉ MAP³ 	A B
Bài tập và Thảo luận	3 tiết Phòng học	<ul style="list-style-type: none"> - Giải các bài tập xây dựng mạng nơ ron với thông số cụ thể - Tính toán các bài tập xây dựng mô hình xác suất MAP - Thảo luận, phân tích, đánh giá mô hình nơ ron và các bộ phân lớp đã giới thiệu 	<ul style="list-style-type: none"> - Thành thạo các dạng bài tập về mạng nơ ron và mô hình MAP - Có năng lực đánh giá, phân tích các mạng đã giới thiệu - củng cố kiến thức về MAP, nơ ron 	<ul style="list-style-type: none"> - Tìm hiểu các bài tập mẫu trong tài liệu tham khảo 	A B
Tự học	20 tiết học tại nhà, thư viện	<ul style="list-style-type: none"> - Tìm tài liệu tiếng Anh về MAP và mạng NN, đọc và dịch, viết báo cáo bằng tiếng Việt - Giải các bài tập khác về MAP, NN. 	<ul style="list-style-type: none"> - Hình thành kỹ năng tự tìm và làm việc với tài liệu Tiếng Anh, kỹ năng viết báo cáo và tổng hợp - củng cố kiến thức về thuật toán MAP, NN 		A B C

³ <https://machinelearningmastery.com/maximum-a-posteriori-estimation/>

KT-ĐG					
Tư vấn	Phòng bộ môn	Tư vấn về các vấn đề của nội dung bài giảng, tiềm năng ứng dụng của MAP, NN trong thực tế	Hiểu rõ hơn về bản chất MAP, NN	SV chuẩn bị các câu hỏi cần giải đáp	A B C

Nội dung tuần 9 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	2 tiết Phòng học	<p>CHƯƠNG 2. PHƯƠNG PHÁP HỌC MÁY CÓ GIÁM SÁT</p> <p>2.9.3. Mạng nơ ron Perceptron</p> <p>2.9.3. Mạng Perceptron đa lớp với giải thuật lan truyền ngược</p> <p>2.9.4. Các mạng nhân chập CNN (Convolutional Neuron Networks)</p> <p>2.9.5. Các mạng CNN học sâu (Deep CNNs)</p>	<ul style="list-style-type: none"> - Hiểu được khái niệm về mạng Perceptron và nguyên tắc hoạt động - Nắm được thuật toán tính đạo hàm lan truyền ngược và cập nhật trọng số mô hình - Sự phát triển của các mạng DNNs và tiềm năng ứng dụng 	<ul style="list-style-type: none"> - Đọc tài liệu [1] [2] (Một số kiến trúc mạng tích chập học sâu) và [3] (Phần mạng nơ ron nhân tạo) - Sử dụng Internet, Google để tra cứu các tài liệu khác. 	A B
Bài tập và Thảo luận	4 tiết Phòng học	<ul style="list-style-type: none"> - Bài tập xây dựng các mạng DCNN đơn giản - Thảo luận, phân tích, đánh giá xu hướng phát triển của mạng DCNNs và thế mạnh, hạn chế của mạng, tiềm năng ứng dụng 	<ul style="list-style-type: none"> - Thành thạo các dạng bài tập về DCNN - Biện luận và giải thích được sự ưu việt, điểm yếu của các mạng DCNNs 	<ul style="list-style-type: none"> - Tìm hiểu các bài tập mẫu trong tài liệu tham khảo 	A B

Tự học	20 tiết học tại nhà, thư viện	<ul style="list-style-type: none"> - Tìm tài liệu về DCNN, Perceptron, mạng lan truyền ngược BP - Viết báo cáo tìm hiểu, phân tích, đánh giá về các mạng nơ ron đã trình bày 	<ul style="list-style-type: none"> - Hình thành kỹ năng tự tìm và làm việc với tài liệu trên Internet, kỹ năng viết báo cáo và tổng hợp - củng cố kiến thức về các thuật toán đã học 		A B C
KT-ĐG					
Tư vấn	Phòng bộ môn	Tư vấn về các vấn đề của nội dung bài giảng, DCNN, Perceptron, mạng lan truyền ngược BP	Hiểu rõ hơn về bản chất của DCNN, Perceptron, mạng lan truyền ngược BP	SV chuẩn bị các câu hỏi cần giải đáp	A B C

Nội dung tuần 10 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	2 tiết Phòng học	<p>CHƯƠNG 3. PHƯƠNG PHÁP HỌC MÁY KHÔNG CÓ GIÁM SÁT</p> <p>3.1. Giới thiệu</p> <p>3.2. Các thuật toán phân cụm</p> <p>3.2.1. Thuật toán K-means</p> <p>3.2.2. Thuật toán MeanShift</p>	<ul style="list-style-type: none"> - Hiểu được khái niệm phân cụm - Nắm được nguyên lý hoạt động của các kỹ thuật phân cụm - Hiểu được các thuật toán phân cụm K-Means, MeanShift - Phân tích và đánh giá được ưu và nhược điểm của K-means (tắc cục bộ) và cách khắc phục 	<ul style="list-style-type: none"> - Đọc tài liệu [2](Phân Phương pháp phân đoạn ảnh) [3] (Các phương pháp phân vùng ảnh) - Sử dụng Internet, Google để tra cứu các tài liệu khác. 	A B
Bài tập và Thảo	3 tiết Phòng học	<ul style="list-style-type: none"> - Giải các bài tập về phân cụm K-means với các cụm cho trước, số vòng lặp cho trước 	<ul style="list-style-type: none"> - Thành thạo các dạng bài tập về phân cụm K-means, 	<ul style="list-style-type: none"> - Tìm hiểu các bài tập mẫu trong 	A B

luận		<ul style="list-style-type: none"> - Thảo luận, phân tích, đánh giá ưu nhược điểm của thuật toán K-means - Đề xuất các cách xử lý tắc nghẽn cục bộ của K-means - Đánh giá độ phức tạp tính toán 	MeanShift <ul style="list-style-type: none"> - Hình thành từ duy biện luận và giải thích được sự ưu việt, điểm yếu của K-means, MeanShift 	tài liệu tham khảo	
Tự học	20 tiết học tại nhà, thư viện	<ul style="list-style-type: none"> - Tìm tài liệu về phân cụm không giám sát và K-means, viết báo cáo - Ứng dụng các công cụ có sẵn (Matlab, OpenCV, Python) để trực quan hóa kết quả phân cụm K-means, MeanShift 	<ul style="list-style-type: none"> - Hình thành kỹ năng tự tìm tài liệu và viết báo cáo - củng cố kiến thức về thuật toán K-means, MeanShift - Rèn luyện kỹ năng thực hành và lập trình 		A B C
KT-ĐG	1 tiết Phòng học	Bài kiểm tra số 03: thi viết, thời gian 50 phút (1 tiết)	Kiểm tra kiến thức về K-means và đánh giá độ phức tạp thuật toán	SV chuẩn bị giấy kiểm tra và máy tính cầm tay	A B C
Tư vấn	Phòng bộ môn	Tư vấn về các vấn đề của nội dung bài giảng, tiềm năng ứng dụng của K-means, MeanShift tư vấn các kiểu bài tập thường gặp	Hiểu rõ hơn về bản chất K-means, MeanShift thành thạo các dạng bài tập liên quan	SV chuẩn bị các câu hỏi cần giải đáp	A B C

Nội dung tuần 11 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	2 tiết Phòng học	<p>CHƯƠNG 3. PHƯƠNG PHÁP HỌC MÁY KHÔNG CÓ GIÁM SÁT</p> <p>3.3. Nguyên lý phân tích thành phần PCA</p> <p>3.3.1. Vai trò của PCA trong máy học</p> <p>3.3.2. Xây dựng PCA dùng phương pháp hiệp phương sai</p>	<ul style="list-style-type: none"> - Hiểu được khái niệm phân tích thành phần chính (PCA), eigen vector, hiệp phương sai (covariance) - Nắm được nguyên lý hoạt động của các PCA - Vai trò PCA trong 	<ul style="list-style-type: none"> - Đọc tài liệu [2] [3] (Phần nguyên lý phân tích thành phần) và tìm các tài liệu PCA 	A B

		(covariance) 4.4. Phân lớp và đánh giá	các thuật toán rút gọn đặc trưng, rút gọn chiều, huấn luyện máy học	trên Internet	
Bài tập và Thảo luận	4 tiết Phòng học	<ul style="list-style-type: none"> - Giải các bài tập về tính toán covariance trên tập dữ liệu nhỏ - Thảo luận ý nghĩa của PCA - Đánh giá độ phức tạp tính toán PCA - Các giải pháp thay thế PCA - Ứng dụng PCA trong nén thông tin 	<ul style="list-style-type: none"> - Thành thạo các dạng bài tập về phân cụm covariance - củng cố kiến thức về PCA và ứng dụng 	<ul style="list-style-type: none"> - Tìm hiểu các bài tập và phân tích về PCA từ nguồn khác trên Internet 	A B
Tự học	20 tiết học tại nhà, thư viện	<ul style="list-style-type: none"> - Tìm tài liệu về PCA và viết báo cáo - Ứng dụng các công cụ có sẵn (Matlab, opencv) để trực quan hóa kết quả phân tích dữ liệu dùng PCA 	<ul style="list-style-type: none"> - Hình thành kỹ năng tự tìm tài liệu và viết báo cáo - củng cố kiến thức về thuật toán PCA - Rèn luyện kỹ năng thực hành và lập trình 	Tải dữ liệu minh họa PCA tại ⁴	A B C
KT-ĐG					
Tư vấn	Phòng bộ môn	Tư vấn về các vấn đề PCA và tiềm năng ứng dụng	Hiểu rõ hơn về bản chất PCA và phân tích thành phần quan trọng của dữ liệu	SV chuẩn bị các câu hỏi cần giải đáp	A B C

Nội dung tuần 12 (LT+BT: 6 tiết)

Hình thức	TG/ĐD	Nội dung chính	Mục tiêu cụ thể	Yêu cầu SV chuẩn bị	Chuẩn đầu ra HP
Lý thuyết	2 tiết Phòng học	3.3.3. Mối tương quan giữa PCA và Kmeans CHƯƠNG 4. ỨNG DỤNG CASE STUDY 4.1. Giới thiệu bài toán	<ul style="list-style-type: none"> - Hiểu rõ về PCA và K-means - Liên hệ được cơ sở lý thuyết học máy và các bài toán, ứng dụng 	<ul style="list-style-type: none"> - Đọc tài liệu [3] (Chương 5, Một số ứng dụng minh 	A B

⁴ <http://www2.imm.dtu.dk/courses/02450/DemoPCA.html>

		nhận dạng mặt người 4.2. Chuẩn bị dữ liệu và trích chọn đặc trưng 4.3. Huấn luyện các bộ học: Decision Tree, SVM, Boosting trees 4.4. Phân lớp và đánh giá	trong thực tiễn - Hiểu rõ nguyên lý hoạt động của một ứng dụng điển hình (nhận dạng mặt người) - Đánh giá độ chính xác của các bộ phân lớp khác nhau	họa)	
Bài tập và Thảo luận	1 tiết Phòng học	- Thảo luận về ứng dụng các bộ phân lớp cho ứng dụng cụ thể về nhận dạng mặt người - Đánh giá độ phức tạp và sai số của các bộ phân lớp - Thảo luận các cải tiến nâng cao độ chính xác	- Rèn luyện kỹ năng vận dụng cơ sở học máy vào bài toán cụ thể trong thực tế	- Tìm hiểu các bài tập mẫu trong tài liệu tham khảo	A B
Tự học	20 tiết học tại nhà, thư viện	- Tìm tài liệu về các ứng dụng nhận dạng trong thực tế (y tế, giáo dục, an ninh, giao thông,...) - Chọn một chủ đề và viết báo cáo	- Hình thành kỹ năng tự tìm tài liệu và viết báo cáo - Rèn luyện kỹ năng thực hành và lập trình		A B C
KT-ĐG	3 tiết Phòng học	Bài kiểm tra số 04 (báo cáo bài tập nhóm)	Sinh viên báo cáo kết quả bài tập nhóm đã thực hiện		A B C
Tư vấn	Phòng bộ môn	Tư vấn các nội dung, vấn đề sinh viên thắc mắc về môn học, thi kết thúc học phần	Củng cố kiến thức học phần và hướng dẫn ôn thi cuối kỳ	SV chuẩn bị các câu hỏi cần giải đáp	A B C

9. Chính sách đối với phân học

Yêu cầu đối với người học:

- Người học phải đầy đủ tư liệu để tự nghiên cứu và chuẩn bị bài trước khi đến lớp tối thiểu là tài liệu [1], [2].
- Hiện diện trên lớp theo quy định (không nghỉ quá 20% tổng số tiết TC).
- Người học phải tham gia đầy đủ các bài tập nhóm, bài kiểm tra - đánh giá định kỳ trong quá trình học và bài kiểm tra kết thúc học phần
- Kết thúc học phần, người học phải hoàn thành một dự án, thực hiện báo cáo dự án và thi vấn đáp để lấy điểm thi học phần.

10. Phương pháp, hình thức kiểm tra - đánh giá kết quả học tập học phần

10.1. Kiểm tra- đánh giá thường xuyên:

- Đánh giá quá trình học và tự học của người học trong các giờ học. .
- Báo cáo 01 bài tập nhóm: báo cáo trong thời gian 5-10 phút/báo cáo.
- Điểm: từ 0 đến 10.
- Điểm trung bình của các bài kiểm tra có trọng số 0,3.
- Tiêu chí kiểm tra đánh giá:
 - + Với các bài tập nhóm: các nhóm phải thực hiện phân công thành viên thực hiện bài tập nhóm một cách khoa học, hiệu quả, thực hiện đúng và đầy đủ các yêu cầu của các bài tập nhóm.
 - + Với bài kiểm tra: sinh viên phải theo dõi bài trên lớp, hiểu và vận dụng kiến thức, kỹ năng được trang bị từ bài giảng để làm các bài tập thực hành.

10.2. Kiểm tra – đánh giá giữa kỳ:

- Kiểm tra đánh giá giữa kì: 1 bài kiểm tra thời gian 50 phút.
- Điểm: từ 0 đến 10.
- Điểm của bài kiểm tra giữa kỳ có trọng số 0,2.

10.3. Kiểm tra – đánh giá cuối kì:

- Hình thức: Thi viết
- Điểm: từ 0 đến 10, làm tròn đến một chữ số thập phân.
- Thời gian: 120 phút
- Địa điểm: Theo lịch thi.
- Trọng số: 0,5

11. Các yêu cầu khác :

- Bố trí lịch học, thời gian học theo đúng lịch trình cụ thể (mục 8.2)
- Các yêu cầu đối với học phần: Giờ lý thuyết phải được học tại phòng chức năng có đầy đủ các thiết bị: máy tính, máy chiếu Projector, nếu phòng học lớn cần có thêm Micro, loa.

Ngày 15 tháng 08 năm 2019

Trưởng khoa

Phạm Thế Anh

Ngày 10 tháng 08 năm 2019

Trưởng bộ môn

Trịnh Viết Cường

Giảng viên

Phạm Thế Anh