

1. THÔNG TIN CHUNG

Tên học phần:	Nhập môn học máy và khai phá dữ liệu (Introduction to Machine Learning and Data Mining)
Mã số học phần:	IT3190
Khối lượng:	3(3-1-0-6) <ul style="list-style-type: none">- Lý thuyết: 45 tiết- Bài tập/BTL: 15 tiết bài tập lớn- Thí nghiệm: 0 tiết
Học phần tiên quyết:	
Học phần học trước:	<ul style="list-style-type: none">- IT3011 hoặc IT3010 (Cấu trúc dữ liệu và thuật toán)- IT3160 (Nhập môn Trí tuệ nhân tạo)
Học phần song hành:	Không

2. MÔ TẢ HỌC PHẦN

Học phần này nhằm cung cấp cho sinh viên các kiến thức cơ bản về học máy và khai phá dữ liệu. Học phần sẽ trình bày quá trình học máy và khai phá dữ liệu cùng với các kỹ thuật cơ bản đối với các bài toán của học máy và khai phá dữ liệu. Các kỹ thuật này được cài đặt trong các thành phần lõi hệ thống để đưa ra các dự đoán và phát hiện tri thức trong các ứng dụng khác nhau. Học phần cũng giới thiệu các ứng dụng điển hình của học máy và khai phá dữ liệu trong thực tế và các công cụ cũng như thư viện hữu ích.

Ngoài ra học phần cũng cung cấp cho sinh viên các kỹ năng làm việc nhóm, thuyết trình và thái độ cần thiết để làm việc và ứng dụng các kỹ thuật học máy và khai phá dữ liệu trong các công ty sau này.

3. MỤC TIÊU VÀ CHUẨN ĐẦU RA CỦA HỌC PHẦN

Sinh viên hoàn thành học phần này có khả năng:

Mục tiêu/CĐR	Mô tả mục tiêu/Chuẩn đầu ra của học phần	CĐR được phân bổ cho HP/ Mức độ (I/T/U)
[1]	[2]	[3]
M1	Hiểu và có khả năng thiết kế và quản lý các hệ thống ứng dụng kỹ thuật học máy và khai phá dữ liệu trong các tổ chức	1.1.1; 1.2.5; 1.1.4; 1.5.2; 2.3.3; 3.1.4
M1.1	Nhận diện và hiểu rõ các thành phần của hệ thống ứng dụng kỹ thuật học máy và khai phá dữ liệu	[1.1.1; 1.2.5; 1.1.4] (I)
M1.2	Nhận diện, so sánh và phân loại được các nhu cầu trong doanh nghiệp mà có thể dùng các công nghệ của Học máy và khai phá dữ liệu để giải quyết	[1.1.1; 1.2.5; 1.1.4] (T)
M1.3	Có khả năng thiết kế hệ thống ứng dụng kỹ thuật học máy và khai phá dữ liệu hỗ trợ xử lý dữ liệu, thông tin và tri thức trong tổ chức	[1.5.2; 2.3.3; 3.1.4] (TU)

Mục tiêu/CĐR	Mô tả mục tiêu/Chuẩn đầu ra của học phần	CĐR được phân bổ cho HP/ Mức độ (I/T/U)
M2	Nhận diện và làm chủ được các cơ hội trên thị trường do học máy và khai phá dữ liệu đem lại để phát triển tổ chức sẵn có và tạo ra các tổ chức mới	1.1.4; 1.3.3; 1.1.4; 1.3.3; 1.3.4; 1.6.3;
M2.1	Hiểu và vận dụng được các ứng dụng sử dụng học máy và khai phá dữ liệu đương đại nhằm hỗ trợ các hoạt động trong tổ chức	[1.1.4; 1.3.3; 1.6.3] (T)
M2.2	Nhận diện được các tác động của học máy và khai phá dữ liệu đối với tổ chức và môi trường hoạt động của tổ chức	[1.3.4] (U)
M3	Nhận diện các xu hướng phát triển của học máy và khai phá dữ liệu có khả năng hỗ trợ việc thay đổi các tổ chức	4.1.1; 4.1.2; 4.1.3
M3.1	Chủ động tìm hiểu và nhận diện các ứng dụng học máy và khai phá dữ liệu mới nhất	[4.1.1; 4.1.2] (T)
M3.2	Xác định được các cơ hội tương lai mà học máy và khai phá dữ liệu đem lại để phát triển tổ chức sẵn có	[4.1.3] (U)

4. TÀI LIỆU HỌC TẬP

Sách tham khảo

- [1] Tom Mitchell. *Machine Learning*. McGraw-Hill, 1997.
- [2] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT press, 2016.
- [3] Jiawei Han, Micheline Kamber, Jian Pei. *Data Mining: Concepts and Techniques* (3rd Edition). Morgan Kaufmann, 2011.
- [4] Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar. *Introduction to Data Mining* (2nd Edition). Pearson, 2017.
- [5] Trevor Hastie, Robert Tibshirani, Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (3rd Edition). Springer, 2017.

5. CÁCH ĐÁNH GIÁ HỌC PHẦN

Điểm thành phần	Phương pháp đánh giá cụ thể	Mô tả	CĐR được đánh giá	Tỷ trọng
[1]	[2]	[3]	[4]	[5]
A1. Điểm quá trình (*) (Một trong hai lựa chọn: A1.1 - A1.2, hoặc A1.3)	Đánh giá quá trình			40%
	A1.1. Kiểm tra giữa kỳ A1.2. Bài tập về nhà theo nhóm	Tự luận Báo cáo	M1.1; M1.2; M2.1; M2.2	20% 20%
	A1.3. Bài tập lớn theo nhóm	Thuyết trình/Báo cáo	M1.3; M3.1; M3.2	40%
A2. Điểm cuối kỳ	A2.1. Thi cuối kỳ	Thi viết hoặc trắc nghiệm	M1÷M2	60%

* Điểm quá trình sẽ được điều chỉnh bằng cách cộng thêm điểm chuyên cần. Điểm chuyên cần có giá trị từ -2 đến +1, theo Quy chế Đào tạo đại học hệ chính quy của Trường ĐH Bách khoa Hà Nội.

6. KẾ HOẠCH GIẢNG DẠY

Tuần	Nội dung	CDR học phần	Hoạt động dạy và học	Bài đánh giá
[1]	[2]	[3]	[4]	[5]
1	Chương 1: Giới thiệu chung về học máy và khai phá dữ liệu 1.1. Học máy và các bài toán học máy 1.2. Giới thiệu về khai phá dữ liệu 1.3. Các công cụ và thư viện của học máy và khai phá dữ liệu	M1.1 M2.2 M3.2	Giảng bài	A1.2 A1.3
2	Chương 2: Thu thập và tiền xử lý dữ liệu 2.1. Thu thập dữ liệu 2.2. Trục quan hóa dữ liệu 2.3. Làm sạch dữ liệu 2.4. Tích hợp dữ liệu 2.5. Biến đổi dữ liệu 2.6. Lựa chọn đặc trưng	M1.1 M1.2	Đọc trước tài liệu; Giảng bài;	A1.2 A2.1
3	Chương 3: Hồi qui 3.1. Giới thiệu bài toán hồi qui 3.2. Mô hình tuyến tính 3.3. Một số phương pháp	M1.2; M1.3; M2.1; M2.2	Đọc trước tài liệu; Giảng bài;	A2.1
4	Chương 4: Phân cụm 4.1. Giới thiệu bài toán phân cụm 4.2. Tiêu chuẩn đánh giá chất lượng 4.3. Phân cụm phân hoạch (K-mean và các biến thể)	M1.2; M1.3;	Đọc trước tài liệu; Giảng bài;	A1.2 A2.1
5	4.4. Phân cụm phân cấp	M1.2; M1.3; M2.1; M2.2	Đọc trước tài liệu; Giảng bài;	A1.1 A1.2 A2.1
6	Chương 5: Phân loại 5.1. Giới thiệu bài toán phân loại 5.2. Cây quyết định và rừng ngẫu nhiên	M1.2; M1.3;	Đọc trước tài liệu; Giảng bài;	A1.1 A1.2
7	5.3. Mạng nơ ron nhân tạo	M1.2; M1.3;	Đọc trước tài liệu; Giảng bài;	A1.1 A2.1
8	5.4. Máy véc tơ hỗ trợ	M1.2; M1.3;	Giảng bài;	A1.2

Tuần	Nội dung	CDR học phần	Hoạt động dạy và học	Bài đánh giá
[1]	[2]	[3]	[4]	[5]
9	5.5. Lựa chọn tham số và đánh giá hiệu năng	M1.2; M1.3; M2.1; M2.2	Giảng bài;	A1.3 A1.2
10	5.6. Học dựa trên xác suất	M1.2; M1.3;	Giảng bài;	A2.1
11	Chương 6: Khai phá dữ liệu 6.1. Cơ bản về khai phá dữ liệu 6.2. Một số thách thức	M3.1; M3.2	Giảng bài;	A2.1 A1.2
12	6.3. Khai phá tập mục thường xuyên và các luật kết hợp 6.4. Thuật toán Apriori	M3.1; M3.2	Giảng bài;	A2.1 A1.2
13	Chương 7: Mở rộng 7.1. Hiệu chỉnh và vai trò 7.2. Thảo luận ứng dụng học máy và khai phá dữ liệu trong thực tế	M3.1; M3.2	Giảng bài; Trao đổi, chia sẻ kinh nghiệm	A1.2
14	Báo cáo bài tập nhóm	M3.1; M3.2	Báo cáo bài tập nhóm	A1.2
15	Tổng kết và ôn tập			

7. QUY ĐỊNH CỦA HỌC PHẦN

(Các quy định của học phần nếu có)

8. NGÀY PHÊ DUYỆT:

Chủ tịch Hội đồng

Nhóm xây dựng đề cương

Thân Quang Khoát

Nguyễn Thị Kim Anh

Ngô Văn Linh

Đinh Viết Sang

Trịnh Anh Phúc

Phạm Văn Hải

Đỗ Tiến Dũng

Nguyễn Đức Anh

Nguyễn Duy Hiệp

Trần Việt Trung

Lê Thanh Hương

9. QUÁ TRÌNH CẬP NHẬT

Lần cập nhật	Nội dung điều chỉnh	Ngày tháng được phê duyet	Áp dụng từ kỳ/khóa	Ghi chú
1			
2			