

ĐỀ CƯƠNG CHI TIẾT HỌC PHẦN

1. Thông tin chung về học phần

- **Tên học phần:** Khai thác dữ liệu (Data mining)
- **Mã số học phần:** 1250174
- **Số tín chỉ học phần:** 4 tín chỉ
- Thuộc chương trình đào tạo của bậc, ngành: Đại học, ngành Công nghệ Thông tin
- **Số tiết học phần:**
 - Nghe giảng lý thuyết : 30 tiết
 - Làm bài tập trên lớp : 0 tiết
 - Thảo luận : 0 tiết
 - Thực hành, thực tập : 30 tiết
 - Hoạt động theo nhóm : 0 tiết
 - Thực tế: : 0 tiết
 - Tự học : 120 giờ
- **Đơn vị phụ trách học phần:** Bộ môn Hệ thống thông tin/Khoa Công Nghệ Thông Tin.

2. Học phần trước: Cơ sở dữ liệu, Xác suất thống kê, Các hệ quản trị Cơ sở dữ liệu.

3. Mục tiêu của học phần:

- Môn học này nhằm cung cấp cho sinh viên những kiến thức cơ bản về lĩnh vực Khai thác dữ liệu và những ứng dụng của nó vào các ngành khoa học khác.
- Môn học này bao gồm các nội dung tổng quát liên quan đến quy trình khám phá tri thức từ dữ liệu và các nội dung chuyên sâu liên quan đến các kỹ thuật thông dụng trong khai thác dữ liệu như khai thác tập phổ biến và luật kết hợp, phân loại, gom cụm, biểu diễn và đánh giá tri thức.

4. Chuẩn đầu ra:

	Nội dung	Đáp ứng CDR CTĐT
Kiến thức	4.1.1. Kiến thức nền tảng: Các khái niệm về lĩnh vực khai thác dữ liệu và các vấn đề liên quan. Biết và giải thích được các thuật ngữ tiếng Anh thuộc lĩnh vực dữ liệu lớn	K1
	4.1.2. Hiểu được ý nghĩa và ứng dụng của các bài toán trong khai thác dữ liệu đối với các vấn đề thực tế.	K2, K3
Kỹ năng	4.2.1. Sử dụng được một số công cụ trong việc triển khai và lập trình trên dữ liệu lớn.	S2
	4.2.2. Kỹ năng nghề nghiệp: Xây dựng được một số thành phần của một hệ thống khai thác dữ liệu ở quy mô nhỏ.	S1, S3
	4.2.3. Kỹ năng cá nhân: Vận dụng được một số vấn đề và các cách tiếp cận trong khai thác dữ liệu	
Thái độ	4.3.1. Tôn trọng quyền tác giả, sử dụng phần mềm hợp pháp	A1
	4.3.2. Đi học đều đặn, nghỉ học phải có lí do chính đáng và phải xin phép. Không đi trễ, về sớm.	A2, A3
	4.3.3. Nghiên túc nghe giảng viên giảng bài. Hoàn thành những bài thực tập do giảng viên yêu cầu	

5. Mô tả tóm tắt nội dung học phần:

Cung cấp các kiến thức cơ bản về khai phá dữ liệu và phát hiện tri thức: khái niệm, kiến trúc hệ thống và đặc trưng; các bài toán điển hình phân lớp, phân cụm, luật kết hợp; các thuật toán Bayes, cây quyết định, mạng neural...

6. Nội dung và lịch trình giảng dạy:

- Các học phần lý thuyết:

Buổi/ Tiết	Nội dung	Hoạt động của giảng viên	Hoạt động của sinh viên	Giáo trình chính	Tài liệu tham khảo	Ghi chú
1/3	Chương 1. Giới thiệu chung về khai phá dữ liệu (KPDL) 1.1. Nhu cầu KPDL 1.2. Các khái niệm KPDL 1.3. Một số ứng dụng 1.4. Các vấn đề chính trong KPDL	- Thuyết giảng	- Nghe giảng, ghi chú - Trả lời câu hỏi	Cuốn [1]: Phần 1	Cuốn [2]: Chương 1, 2	Giải quyết mục tiêu 4.1.1, 4.2.1
2/3	Chương 2. Tiền xử lý dữ liệu 2.1. Vấn đề về dữ liệu 2.2. Phân tích cấu trúc và thuộc tính dữ liệu 2.3. Tiền xử lý dữ liệu	- Thuyết giảng	- Nghe giảng, ghi chú - Trả lời câu hỏi	Cuốn [1]: Phần 1	Cuốn [2]: Chương 3	Giải quyết mục tiêu 4.1.1, 4.1.2, 4.2.1, 4.2.2
3/3	Bài tập	- Hướng dẫn và sửa bài tập	- Thảo luận, trả lời câu hỏi			Giải quyết mục tiêu 4.1.1, 4.1.2, 4.2.1, 4.2.2
4/3	Chương 3. Khai phá luật kết hợp 3.1. Các khái niệm cơ bản 3.2. Kỹ thuật khai thác mẫu phổ biến	- Thuyết giảng	- Nghe giảng, ghi chú - Trả lời câu hỏi	Cuốn [1]: Phần 2	Cuốn [2]: Chương 6	Giải quyết mục tiêu 4.1.3, 4.2.3
5/3	Bài tập	- Hướng dẫn và sửa bài tập	- Thảo luận, trả lời câu hỏi			
6/3	Chương 3. Khai phá luật kết hợp 3.3. Kỹ thuật khai thác luật kết	- Thuyết giảng	- Nghe giảng, ghi chú - Trả lời câu hỏi	Cuốn [1]: Phần 2	Cuốn [2]: Chương 6	Giải quyết mục tiêu

	hợp					4.1.3, 4.2.3
7/3	Bài tập	- Hướng dẫn và sửa bài tập	- Thảo luận, trả lời câu hỏi			
8/3	Chương 4. Phân lớp 4.1. Các khái niệm 4.2. Kỹ thuật phân lớp có giám sát 4.3. Kỹ thuật phân lớp bán giám sát	- Thuyết giảng	- Nghe giảng, ghi chú - Trả lời câu hỏi	Cuốn [1]: Phần 3	Cuốn [2]: Chương 7	Giải quyết mục tiêu 4.1.3, 4.2.3
9/3	Bài tập	- Hướng dẫn và sửa bài tập	- Thảo luận, trả lời câu hỏi			
10/3	Chương 5. Phân cụm 5.1. Khái niệm 5.2. Các độ đo cơ bản trong phân cụm 5.3. Các kỹ thuật phân cụm tiêu biểu 5.4. Gán nhãn cho cụm 5.5. Đánh giá phân cụm	- Thuyết giảng	- Nghe giảng, ghi chú - Trả lời câu hỏi	Cuốn [1]: Phần 3	Cuốn [2]: Chương 8	Giải quyết mục tiêu 4.1.3, 4.2.3

- Các học phần thực hành:

Buổi/ Tiết	Nội dung	Hoạt động của giảng viên	Hoạt động của sinh viên	Giáo trình chính	Tài liệu tham khảo	Ghi chú
1/3	Bài 1: Cài đặt và thao tác trên Weka	- Thuyết giảng - Hướng dẫn sinh viên thực hiện	- Nghe giảng, ghi chú - Trả lời câu hỏi - Làm bài tập	Tham khảo web Weka	Đọc thêm các tài liệu tham khảo	Giải quyết mục tiêu 4.1.1; 4.1.2; 4.1.3; 4.3.1; 4.3.2
2/3	Bài 2: Sử dụng Weka trong phân tích và thực thi các thuật toán	- Thuyết giảng - Hướng dẫn sinh viên	- Nghe giảng, ghi chú	Tham khảo web	Đọc thêm các tài liệu tham khảo	Giải quyết mục tiêu: 4.1.1; 4.1.2; 4.1.3;

	phân lớp	thực hiện - Sửa lỗi cho sinh viên và giải thích	- Trả lời câu hỏi - Làm bài tập	Weka		4.3.1; 4.3.2
3,4/6	Bài 3: Cài đặt thuật toán tiền xử lý dữ liệu cơ bản	- Thuyết giảng - Hướng dẫn sinh viên thực hiện - Sửa lỗi cho sinh viên và giải thích	- Nghe giảng, ghi chú - Trả lời câu hỏi - Làm bài tập	Cuốn [1], [2]	Đọc thêm các tài liệu tham khảo	Giải quyết mục tiêu: 4.1.1; 4.1.2; 4.1.3; 4.3.1; 4.3.2
5,6/6	Bài 4: Cài đặt thuật toán khai thác luật kết hợp	- Thuyết giảng - Hướng dẫn sinh viên thực hiện - Sửa lỗi cho sinh viên và giải thích	- Nghe giảng, ghi chú - Trả lời câu hỏi - Làm bài tập	Cuốn [1], [2]	Đọc thêm các tài liệu tham khảo	Giải quyết mục tiêu: 4.1.1; 4.1.2; 4.1.3; 4.3.1; 4.3.2
7, 8/6	Bài 5: Cài đặt thuật toán phân cụm	- Thuyết giảng - Hướng dẫn sinh viên thực hiện - Sửa lỗi cho sinh viên và giải thích	- Nghe giảng, ghi chú - Trả lời câu hỏi - Làm bài tập	Cuốn [1], [2]	Đọc thêm các tài liệu tham khảo	Giải quyết mục tiêu: 4.1.1; 4.1.2; 4.1.3; 4.3.1; 4.3.2
9/3	Bài 6: Nhúng source code vào Weka	- Thuyết giảng - Hướng dẫn sinh viên thực hiện - Sửa lỗi cho sinh viên và giải thích	- Nghe giảng, ghi chú - Trả lời câu hỏi - Làm bài tập	Cuốn [1], [2], tham khảo web Weka	Đọc thêm các tài liệu tham khảo	Giải quyết mục tiêu: 4.1.1; 4.1.2; 4.1.3; 4.3.1; 4.3.2
10	Thi					

7. Nhiệm vụ của sinh viên:

Sinh viên phải thực hiện các nhiệm vụ như sau:

- Tham dự tối thiểu 80% số tiết học.
- Thực hiện đầy đủ các bài tập và được đánh giá kết quả thực hiện.
- Nộp đồ án phần kiểm tra giữa học kỳ.
- Nộp đồ án phần kết thúc học phần.
- Chủ động tổ chức thực hiện giờ tự học.

8. Đánh giá kết quả học tập của sinh viên:

8.1. Cách đánh giá

Sinh viên được đánh giá tích lũy học phần như sau:

TT	Thành phần	Điểm thành phần	Quy định	Trọng số điểm	Trọng số thành phần	Mục tiêu
1	Thực hành	Điểm chuyên cần	Tham dự ít nhất 80% số tiết học và số bài tập được giao	20%	40%	4.3
		Điểm thực hành	Thực hiện bài tập, thi	80%		4.2 4.1
2	Lý thuyết	Điểm thi kết thúc học phần	Thực hiện đồ án môn học		60%	4.1 4.2

8.2. Cách tính điểm

- Điểm đánh giá thành phần và điểm thi kết thúc học phần được chấm theo thang điểm 10 (từ 0 đến 10), làm tròn đến 0.5.
- Điểm học phần là tổng điểm của tất cả các điểm đánh giá thành phần của học phần nhân với trọng số tương ứng. Điểm học phần theo thang điểm 10 làm tròn đến một chữ số thập phân.

9. Tài liệu học tập:

9.1. Giáo trình chính:

[1] Nhập môn- Giáo trình phát hiện tri thức và khai phá dữ liệu, Nguyễn Đức Thuận, NXB Thông tin và Truyền thông, 2013

9.2. Tài liệu tham khảo:

[2] Data Mining: Concepts and Techniques (<http://web.engr.illinois.edu/~hanj/bk1/>), Jiawei Han and Micheline Kamber, Morgan Kaufmann Publishers, 2000.

[3] Nguyên lý các hệ cơ sở dữ liệu và cơ sở tri thức: Các hệ cơ sở tri thức, Ullman Jeffrey D., NXB Thống kê, 2000.

10. Hướng dẫn sinh viên tự học:

Tuần/ Buổi	Nội dung	Lý thuyết (tiết)	Thực hành (tiết)	Nhiệm vụ của sinh viên
1	Chương 1. Giới thiệu chung về khai phá dữ liệu (KPDL)	3	0	Đọc trước tài liệu Cuốn [1]: Phần 1, Cuốn [2]: Chương

				1, 2
2,3	Chương 2. Tiền xử lý dữ liệu	6	6	Cuốn [1]: Phần 1, Cuốn [2]: Chương 3
4,5,6	Chương 3. Khai phá luật kết hợp	9	9	Cuốn [1]: Phần 2, Cuốn [2]: Chương 6
8,9	Chương 4. Phân lớp	6	6	Cuốn [1]: Phần 3, Cuốn [2]: Chương 7
10	Chương 5. Phân cụm	3	3	Cuốn [1]: Phần 3, Cuốn [2]: Chương 8

Ngày... tháng.... Năm 2015
Trưởng khoa
(Ký và ghi rõ họ tên)

Ngày... tháng.... Năm 2015
Trưởng Bộ môn
(Ký và ghi rõ họ tên)

Ngày... tháng.... Năm 2015
Người biên soạn
(Ký và ghi rõ họ tên)

Trần Minh Thái
Ngày... tháng.... Năm 2015
Ban giám hiệu

Trần Minh Thái