

TRƯỜNG ĐẠI HỌC
TÀI CHÍNH NGÂN HÀNG HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN
BỘ MÔN TIN ỨNG DỤNG

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc Lập- Tự Do- Hạnh Phúc

ĐỀ CƯƠNG CHI TIẾT HỌC PHẦN KHAI PHÁ DỮ LIỆU LỚN TRONG KINH DOANH

Hệ đào tạo: Đại học chính quy

Ngành: Công nghệ thông tin

(Ban hành theo QĐ số 291/QĐ-ĐHTNH ngày 30/6/2025 của Hiệu trưởng)

1. Thông tin chung về học phần

- Tên học phần: **Khai phá dữ liệu lớn trong kinh doanh**
- + Tiếng Anh: **Big Data Mining in Business**
- Mã học phần: **DCT.02.40**
- Số tín chỉ: 3
- Vị trí của học phần trong CTĐT

STT	Khối kiến thức	Loại hình	
		Bắt buộc	Tự chọn
1	Kiến thức giáo dục đại cương		
2	Kiến thức cơ sở khối ngành		
3	Kiến thức cơ sở ngành		
4	Kiến thức chuyên ngành		x

- Học phần tiên quyết:
 - + Cơ sở lập trình (Mã số DCT.02.31)
 - + Hệ cơ sở dữ liệu và dữ liệu lớn (Mã số DCT.02.42)
- Học phần song hành:
- Phân bổ giờ tín chỉ đối với các hoạt động:

Hoạt động trên lớp			Hoạt động khác
Lí thuyết (LT)	Bài tập, Thực hành, Thảo luận (BT, TH, TL)	Kiểm tra (KT)	(tự học, làm bài tập, nghiên cứu, trải nghiệm, ...)
27	30	3	
60 tiết			90 tiết

(01 giờ chuẩn = 1 tiết LT, Kiểm tra hoặc 2 tiết Bài tập, Thực hành, Thảo luận trên lớp)

- Khoa, Bộ môn phụ trách học phần: Bộ môn Tin ứng dụng Viện CNTT
- Giảng viên phụ trách học phần (dự kiến):
 - 1) Họ và tên: ThS. Bùi Thị Thu Hiền
 - Chức danh: Giảng viên;
 - Thông tin liên hệ: 0985220287; Email: buithuhien1987@gmail.com

2) Họ và tên: ThS. Vũ Minh Tâm

Chức danh: Giảng viên

Thông tin liên hệ: ĐT: 0847242752; Email: yuminhtam.edu@gmail.com

2. Mô tả học phần

Học phần cung cấp cho sinh viên kiến thức về hoạt động kinh doanh và dữ liệu trong doanh nghiệp; các mô hình khai phá dữ liệu và dữ liệu lớn; các thuật toán khai phá dữ liệu và ứng dụng.

3. Mục tiêu của học phần:

Mục tiêu học phần	Mô tả mục tiêu học phần
PSO 1.1	Cung cấp cho sinh viên các kiến thức về khai phá dữ liệu, thuật toán khai phá dữ liệu kinh doanh.
PSO 2.1	Rèn luyện cho sinh viên kỹ năng xử lý các bài toán ứng dụng khai phá dữ liệu lớn.
PSO 3.1	Giúp sinh viên có năng lực lập kế hoạch, quản lý các hoạt động chuyên môn

4. Chuẩn đầu ra của học phần - CLO (Course Learning Outcomes)

Mục tiêu học phần ^[9]	CDR học phần ^[11]	Mô tả chuẩn đầu ra học phần ^[12] <i>Hoàn thành học phần này, người học thực hiện được:</i>	CDR của CTĐT ^[13]	Mức độ ^[14]
<i>CDR về kiến thức:</i>				
PSO 1.1	CLO 1.1	Vận dụng được kiến thức khai phá dữ liệu để xây dựng mô hình phân tích dữ liệu kinh doanh	PLO 1.2	3
<i>CDR về kỹ năng:</i>				
PSO 2.1	CLO 2.1	Xây dựng được mô hình phân tích dữ liệu kinh doanh nhờ kỹ thuật khai phá dữ liệu lớn	PLO 2.4	3
<i>CDR về năng lực tự chủ và trách nhiệm:</i>				
PSO 3.1	CLO 3.1	Có năng lực lập kế hoạch, quản lý các hoạt động trong học tập	PLO 3.4	3

Mức độ đóng góp theo thang Bloom.

5. Học liệu:

5.1. Tài liệu chính:

[1]. Đỗ Trung Tuấn, Phân tích thống kê và khai phá dữ liệu, NXB Đại học Quốc gia Hà Nội, 2018.

5.2. Tài liệu tham khảo:

[1]. Galit Shmueli, Peter C. Bruce, Peter Gedeck, Nitin R. Patel. DATA MINING FOR BUSINESS ANALYTICS, John Wiley & Sons, 2020.

6. Phương pháp dạy và học áp dụng cho học phần

STT	CDR	Phương pháp dạy học
1	CLO 1.1: Vận dụng được kiến thức khai phá dữ liệu để xây dựng mô hình phân tích dữ liệu kinh doanh	1. Thuyết trình tích cực: Áp dụng cho các nội dung kiến thức lý thuyết của các học phần.
		2. Hướng dẫn học tập cộng tác: Thiết kế các hoạt động học tập yêu cầu sinh viên làm việc nhóm với phân công vai trò rõ ràng và phản hồi thường xuyên.
		3. Tích hợp công nghệ trong giảng dạy: Sử dụng các phần mềm trình diễn, các phương tiện trình chiếu, các công cụ minh họa và hệ thống quản lý học tập để tăng hiệu quả truyền đạt của giảng viên và tiếp thu kiến thức của sinh viên.
		4. Hướng dẫn tự học
2	CLO 2.1: Xây dựng được mô hình phân tích dữ liệu kinh doanh nhờ kỹ thuật khai phá dữ liệu lớn	1. Dạy thực hành: Hướng dẫn sinh viên thực hiện bài tập thực hành theo nhóm để trau dồi kỹ năng, củng cố kiến thức, qua đó hình thành, phát triển các năng lực của người học. 2. Học theo dự án: Thiết kế các dự án học tập liên quan đến vấn đề thực tiễn, yêu cầu sinh viên vận dụng kiến thức đa lĩnh vực để thực hiện.
3	CLO 3.1: Có năng lực lập kế hoạch, quản lý các hoạt động chuyên môn	1. Hướng dẫn học tập cộng tác: Thiết kế các hoạt động học tập yêu cầu sinh viên làm việc nhóm với phân công vai trò rõ ràng và phản hồi thường xuyên. 2. Hướng dẫn tự học
		❖ Phương pháp khác

7. Nội dung chi tiết học phần (Kế hoạch giảng dạy)

Bài	Nội dung	Số tiết			CLO	Nhiệm vụ của SV
		LT	BT, TL	KT		
1	+ Giới thiệu đề cương chi tiết + Chương 1 Giới thiệu 1.1 Phân tích kinh doanh 1.2 Khai phá dữ liệu 1.3 Khai phá dữ liệu và các điều khoản liên quan 1.4 Dữ liệu lớn 1.5 Khoa học dữ liệu 1.6 Nhiều phương pháp khác nhau 1.7 Thuật ngữ và ký hiệu	2	1		CLO 1.1	Tự nghiên cứu chương 1 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)

2	Chương 2 Tổng quan về quy trình khai phá dữ liệu 2.1 Giới thiệu 2.2 Ý tưởng cốt lõi trong khai phá dữ liệu 2.3 Các bước khai phá dữ liệu 2.4 Các bước sơ bộ 2.5 Sức mạnh dự đoán và trang bị quá mức	2	1		CLO 1.1	Tự nghiên cứu chương 2 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
3	2.6 Xây dựng mô hình dự đoán 2.7 Sử dụng Python để khai phá dữ liệu trên máy cục bộ 2.8 Giải pháp khai phá dữ liệu tự động	1	2		CLO 1.1 CLO 2.1	Tự nghiên cứu chương 2 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
4	Chương 3 Trực quan hóa dữ liệu 3.1 Giới thiệu 3.2 Dữ liệu thử nghiệm 3.3 Biểu đồ cơ bản: Biểu đồ thanh, Biểu đồ đường và Biểu đồ phân tán 3.4 Trực quan hóa đa chiều 3.5 Hình dung chuyên biệt 3.6 Trực quan hóa và vận hành chính, theo mục tiêu khai phá dữ liệu	1	2		CLO 1.1 CLO 2.1	Tự nghiên cứu chương 3 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
5	Chương 4 Giảm chiều dữ liệu 4.1 Giới thiệu 4.2 Phân tích tương quan 4.3 Giảm số lượng danh mục trong các biến phân loại 4.4 Chuyển đổi một biến phân loại thành biến số 4.5 Giảm kích thước bằng mô hình hồi quy 4.6 Giảm kích thước bằng cách sử dụng cây phân loại và cây hồi quy + Kiểm tra bài 1 (1 tiết)	1	1	1	CLO 1.1	Tự nghiên cứu chương 4 giáo trình và ôn tập kiểm tra (4,5 tiết)
6	Chương 6 Hồi quy tuyến tính bội 6.1 Giới thiệu 6.2 Mô hình giải thích và dự đoán	1	2		CLO 1.1	Tự nghiên cứu chương 6 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)

7	6.3 Ước lượng phương trình hồi quy và dự đoán 6.4 Lựa chọn biến trong hồi quy tuyến tính	1	2		CLO 1.1 CLO 2.1	Tự nghiên cứu chương 6 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
8	Chương 7 Phương pháp k- láng giềng gần nhất (k-NN) 7.1 Bộ phân loại k-NN (Kết quả phân loại) 7.2 k-NN cho kết quả bằng số 7.3 Ưu điểm và nhược điểm của thuật toán k-NN	2	1		CLO 1.1 CLO 2.1	Tự nghiên cứu chương 7 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
9	Chương 8 Phân loại Bayes đơn giản 8.1 Giới thiệu 8.2 Áp dụng Bộ phân loại Bayes đầy đủ 8.3 Ưu điểm và nhược điểm của phương pháp Bayes đơn giản	1	2		CLO 1.1	Tự nghiên cứu chương 8 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
10	Chương 9 Cây phân loại và cây hồi quy 9.1 Giới thiệu 9.2 Cây phân loại 9.3 Đánh giá hiệu suất của cây phân loại	1	2		CLO 1.1	Tự nghiên cứu chương 9 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
11	9.5 Quy tắc phân loại từ cây 9.6 Cây phân loại cho nhiều hơn hai lớp 9.7 Cây hồi quy 9.8 Cải thiện dự đoán: Rừng ngẫu nhiên và cây được tăng cường 9.9 Ưu điểm và điểm yếu của cây	1	2		CLO 1.1	Tự nghiên cứu chương 9 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
12	Chương 10 Hồi quy cung ứng 10.1 Giới thiệu 10.2 Mô hình hồi quy cung ứng 10.3 Ví dụ về chấp nhận khoản vay cá nhân	2	1		CLO 1.1	Tự nghiên cứu chương 10 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
13	10.4 Đánh giá hiệu suất phân loại 10.5 Hồi quy cung ứng cho phân loại nhiều lớp + Kiểm tra bài 2 (1 tiết)	2		1	CLO 1.1	Tự nghiên cứu chương 10 giáo trình và ôn tập kiểm tra (4,5 tiết)

14	Chương 11 Mạng nơ ron 11.1 Giới thiệu 11.2 Khái niệm và cấu trúc của mạng nơ ron 11.3 Lắp mạng vào dữ liệu 11.4 Yêu cầu đầu vào của người dùng	2	1		CLO 1.1	Tự nghiên cứu chương 11 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
15	11.5 Khám phá mối quan hệ giữa các yếu tố dự đoán và kết quả 11.6 Học sâu 11.7 Ưu điểm và điểm yếu của mạng nơ ron	2	1		CLO 1.1 CLO 2.1	Tự nghiên cứu chương 11 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
16	Chương 13 Các phương pháp kết hợp: Mô hình tập hợp và nâng cao 13.1 Tập hợp 13.2 Mô hình nâng cao 13.3 Tóm tắt	2	1		CLO 1.1	Tự nghiên cứu chương 13 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
17	Chương 14 Quy tắc kết hợp và lọc cộng tác 14.1 Quy tắc kết hợp 14.2 Lọc cộng tác 14.3 Tóm tắt	2	1		CLO 1.1	Tự nghiên cứu chương 14 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
18	Chương 15 Phân tích cụm 15.1 Giới thiệu 15.2 Đo khoảng cách giữa hai bản ghi 15.3 Đo khoảng cách giữa hai cụm	1	2		CLO 1.1 CLO 2.1	Tự nghiên cứu chương 15 giáo trình và làm bài tập theo hướng dẫn của GV (4,5 tiết)
19	15.4 Phân cụm theo phân cấp 15.5 Phân cụm không phân cấp: Thuật toán k-Means + Kiểm tra bài 3 (1 tiết)		2	1	CLO 2.1	Làm bài tập và ôn tập kiểm tra (4,5 tiết)
20	+ Tổng kết học phần + Hướng dẫn ôn tập		3			Ôn tập theo hướng dẫn của GV (4,5 tiết)
	Tổng cộng	27	30	3		90

8. Nhiệm vụ của người học

- Tham dự giờ lên lớp: Tối thiểu 80% số tiết học trên lớp, trong phòng thực hành có sự giảng dạy, hướng dẫn trực tiếp của giáo viên;
- Bài tập, thảo luận:

- + Đọc tài liệu, chuẩn bị và tham gia thảo luận theo hướng dẫn của giáo viên;
- + Thực hiện đầy đủ các bài tập được giao;
- Làm bài kiểm tra định kỳ;
- Tham gia thi kết thúc học phần.

9. Đánh giá kết quả học tập và cho điểm

9.1. Thang điểm đánh giá

Đánh giá theo thang điểm 10, sau đó được quy đổi sang thang điểm chữ và thang điểm 4 theo Quy chế đào tạo hiện hành.

9.2. Phương pháp, hình thức đánh giá

Thành phần đánh giá	Trọng số (%)	Phương pháp đánh giá	Công cụ đánh giá	CLO	Trọng số CLO (%)
(1)	(2)	(3)	(4)	(5)	(6)
Đánh giá chuyên cần	10	Quan sát có cấu trúc: Quan sát và ghi nhận quá trình học tập của sinh viên về tham dự các buổi học, tinh thần, thái độ học tập.	Rubric	CLO 3.1	100%
Đánh giá quá trình					
03 bài kiểm tra tự luận 1 tiết	30	Đánh giá kết hợp: Kết hợp các hình thức đánh giá như trắc nghiệm khách quan, tự luận.	Rubric	CLO1.1, CLO2.1	40% 60%
Đánh giá kết thúc					
Bài tự luận 90 phút	60	Tự luận	Rubric	CLO1.1, CLO2.1	40% 60%

9.3. Các Rubric đánh giá kết quả học tập

9.3.1 Rubric đánh giá chuyên cần

Tiêu chí đánh giá	Mức chất lượng	Thang điểm
– Tham gia học tập hoặc các hoạt động từ 80% thời gian – Thực hiện công việc đúng tiến độ với chất lượng tốt – Tham gia tích cực các hoạt động của doanh nghiệp – Ý thức, thái độ học tập và chấp hành quy định của trường tốt và được đánh giá cao	Mức A (Vượt quá mong đợi)	8,5 - 10
– Tham gia học tập hoặc các hoạt động từ 80% thời gian – Thực hiện công việc được giao đúng tiến độ nhưng chất lượng còn một số hạn chế – Tham gia tích cực các hoạt động của lớp – Ý thức, thái độ thực tập và chấp hành quy định của trường tốt	Mức B (Đáp ứng được mong đợi)	7,0 - 8,4
– Tham gia học tập hoặc các hoạt động từ 80% thời gian – Thực hiện công việc đúng tiến độ nhưng chất lượng công việc còn nhiều hạn chế nhưng sửa chữa kịp thời theo hướng dẫn của giáo viên	Mức C (Đạt, song cần cải thiện)	5,5 - 6,9

– Chưa tích cực tham gia các hoạt động của doanh nghiệp – Ý thức, thái độ thực tập và chấp hành quy định của trường chưa cao		
– Tham gia học tập hoặc các hoạt động dưới 80% thời gian – Thực hiện công việc không đúng tiến độ, chất lượng công việc không đáp ứng yêu cầu – Chưa tích cực tham gia các hoạt động của doanh nghiệp – Ý thức, thái độ thực tập và chấp hành quy định của trường không tốt.	Mức D (Chưa đạt)	4,0 - 5,4

9.3.2 Rubric đánh giá bài kiểm tra tự luận trong kỳ và thi tự luận hết học phần

Tiêu chí đánh giá của từng câu hỏi trong đề thi	Mức chất lượng	Thang điểm
– Nội dung đủ, trả lời đúng 100% câu hỏi. – Trình bày rõ ràng, diễn đạt ngắn gọn, súc tích, logic. – Mắc 1-2 lỗi về thuật ngữ chuyên môn. – Còn vài lỗi chính tả.	Mức A (Vượt quá mong đợi)	85 - 100
– Trả lời đúng 70-80% câu hỏi. – Trình bày rõ ràng, diễn đạt logic. – Mắc ít lỗi (3-5) lỗi về thuật ngữ chuyên môn. – Có khá nhiều lỗi chính tả.	Mức B (Đáp ứng được mong đợi)	70 - 84
– Trả lời đúng 50-60% câu hỏi. – Trình bày không rõ ý, chưa logic. – Mắc lỗi về thuật ngữ chuyên môn (trên 5 lỗi). – Nhiều lỗi chính tả.	Mức C (Đạt)	55 - 69
– Trả lời đúng 40-50% câu hỏi. – Không làm hết câu hỏi, bỏ nội dung hơn 60%. – Trình bày tối nghĩa, diễn đạt không rõ ý. – Mắc nhiều lỗi về thuật ngữ chuyên môn. – Nhiều lỗi chính tả.	Mức D (Đạt, song cần cải thiện)	40 - 54
– Các trường hợp còn lại	Mức F (Không đạt)	Dưới 40

Viện trưởng

Trưởng Bộ môn

Người soạn đề cương

TS. Phùng Văn Ổn

ThS. Vũ Minh Tâm

TS. Phùng Văn Ổn