

## CHƯƠNG TRÌNH GIÁO DỤC ĐẠI HỌC

Trình độ đào tạo: Đại học      Ngành: Công nghệ thông tin (Chất lượng cao - Đào tạo theo cơ chế đặc thù, HTDN)      Mã số: 7480201DT

## ĐỀ CƯƠNG CHI TIẾT HỌC PHẦN

Tên học phần: Khoa học dữ liệu

Tên tiếng Anh: Data Science

1. Mã học phần:	
2. Ký hiệu học phần:	
3. Số tín chỉ:	03 TC (03 LT)
4. Phân bố thời gian:	
- Lý thuyết:	35 tiết
- Bài tập/Thảo luận:	10 tiết
- Thực hành/Thí nghiệm:	0 tiết
- Tự học:	90 tiết
5. Các giảng viên phụ trách học phần:	
- Giảng viên phụ trách chính:	TS. Ninh Khánh Duy
- Danh sách giảng viên cùng giảng dạy:	TS. Phạm Minh Tuấn, TS. Nguyễn Văn Hiệu, TS. Phạm Công Thắng, TS. Trương Ngọc Châu, TS. Huỳnh Hữu Hưng
- Bộ môn phụ trách giảng dạy:	Khoa CNTT
6. Điều kiện tham gia học phần:	
- Học phần tiên quyết:	
- Học phần học trước:	Đại số tuyến tính, Giải tích 1, Giải tích 2, Xác suất thống kê
- Học phần song hành:	Trí tuệ nhân tạo
7. Loại học phần:	<input checked="" type="checkbox"/> Bắt buộc <input type="checkbox"/> Tự chọn bắt buộc <input type="checkbox"/> Tự chọn tự do
8. Thuộc khối kiến thức	<input checked="" type="checkbox"/> Toán và KHTN <input type="checkbox"/> Cơ sở kỹ thuật/Cơ sở ngành <input type="checkbox"/> Chuyên ngành <input type="checkbox"/> Chung <input type="checkbox"/> Bổ trợ <input type="checkbox"/> Đồ án/Thực tập/Tốt nghiệp

## 9. Mô tả tóm tắt học phần:

Học phần cung cấp cho sinh viên các nguyên lý của khoa học dữ liệu (KHDL), các công cụ toán học thường dùng trong KHDL, và các thuật toán học máy cơ bản; các công cụ phần mềm nguồn mở nhằm thu thập, lưu trữ, xử lý, phân tích, trực quan hoá, mô hình hoá dữ liệu và ra quyết định dựa trên dữ liệu. Học phần này sử dụng các kiến thức toán học đã học trước đó và nhằm trang bị cho sinh viên các kiến thức và công cụ cơ bản trong KHDL. Sau khi kết thúc học phần, sinh viên có thể nắm vững các kỹ thuật học máy cơ bản, và khả năng áp dụng kiến thức vào triển khai các dự án KHDL với các loại hình dữ liệu và lĩnh vực khác nhau.

## 10. Mục tiêu của học phần:

### a. Kiến thức:

Trang bị cho sinh viên những kiến thức về các nguyên lý của KHDL, các công cụ toán học thường dùng trong KHDL, các kỹ thuật học máy cơ bản, và khả năng áp dụng kiến thức vào triển khai các dự án KHDL với các loại hình dữ liệu và lĩnh vực khác nhau.

### b. Kỹ năng:

Rèn luyện cho sinh viên kỹ năng tự học, kỹ năng giao tiếp, kỹ năng làm việc nhóm cùng với kỹ năng sử dụng các công cụ phần mềm nhằm thu thập, lưu trữ, xử lý, phân tích, trực quan hoá, mô hình hoá dữ liệu và ra quyết định dựa trên dữ liệu.

### c. Thái độ:

Giúp sinh viên hình thành thái độ học tập tích cực, đạo đức nghề nghiệp và sự tự chủ, tự chịu trách nhiệm trong giải quyết các vấn đề liên quan đến dữ liệu.

## 11. Chuẩn đầu ra của học phần:

Sau khi kết thúc học phần sinh viên có khả năng:

STT	Chuẩn đầu ra học phần (CLO)	Kiến thức	Kỹ năng	Thái độ	CDIO
1	Hiểu các nguyên lý của KHDL	Hiểu		Tiếp nhận	T
2	Hiểu và áp dụng các công cụ toán học thường dùng trong KHDL	Vận dụng		Tiếp nhận	U
3	Sử dụng được các công cụ phần mềm nhằm thu thập, lưu trữ, xử lý, phân tích, trực quan hoá, mô hình hoá dữ liệu và ra quyết định dựa trên dữ liệu	Vận dụng	Quan sát, Thực hành	Tiếp nhận	U
4	Ứng dụng các công cụ học máy vào các bài toán thực tế	Vận dụng	Xây dựng	Trao đổi	U
5	Trình bày các đặc điểm của dữ liệu và đánh giá hiệu quả các giải pháp khi triển khai một dự án KHDL	Phân tích, Đánh giá	Phát triển	Đề xuất	U

## 12. Mối liên hệ của CĐR học phần (CLOs) đến CĐR Chương trình đào tạo (PLOs):

PLO	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
Cấp độ	H	H	M	M		H		
CLO 1	X							
CLO 2	X	X						
CLO 3				X				

CLO 4	X	X						
CLO 5	X	X	X			X		

### 13. Nhiệm vụ của sinh viên:

Sinh viên phải thực hiện các nhiệm vụ sau đây:

- Tham gia ít nhất 80% số tiết học của lớp học phần;
- Tham gia các hoạt động làm việc nhóm theo qui định của lớp học phần;
- Tự tìm hiểu các vấn đề do giảng viên giao để thực hiện ngoài giờ học trên lớp;
- Hoàn thành tất cả bài đánh giá của học phần.

### 14. Đánh giá sinh viên:

Kết quả học tập của sinh viên được đánh giá bằng các thành phần: đánh giá quá trình và đánh giá cuối kỳ.

Thành phần ĐG	Bài đánh giá	Phương pháp đánh giá	Tiêu chí	Trọng số cấu phần	Trọng số thành phần	CĐR học phần
A1. Đánh giá quá trình	A1.1 Bài tập nhóm	Báo cáo theo nhóm	Trình bày khoa học, đúng đắn, logic	W1.1 50%	W1 50%	CLO 1, 2, 3
	A1.2 Kiểm tra giữa kỳ	Thi trên máy	Viết chương trình đúng đắn và logic	W1.2 50%		CLO 1, 2, 3, 4, 5
A2. Đánh giá cuối kỳ	A2.1 Kiểm tra cuối kỳ	Tiểu luận và bảo vệ	Trình bày khoa học, đúng đắn, logic	W2.1 100%	W2 50%	CLO 1, 2, 3, 4, 5

### 15. Kế hoạch giảng dạy và học

#### 15.1. Kế hoạch giảng dạy và học cho phần lý thuyết

Buổi (3 tiết)	Nội dung chi tiết	Hoạt động dạy và học	Bài đánh giá	CĐR học phần
1	Chương 1. Các nguyên lý của KHDL 1.1 Khái niệm về KHDL 1.2 Quy trình của KHDL 1.3 Các kiểu dữ liệu 1.4 Tổng quan về học máy	<b>Dạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời. <b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên. <b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.	A1.1, A1.2, A2.1	CLO 1, 2
2	Chương 2. Các công cụ toán học trong KHDL 2.1 Phân tích dữ liệu khám phá 2.2 Phân bố dữ liệu và phân bố lấy mẫu 2.3 Kiểm định giả thiết	<b>Dạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.	A1.1, A1.2, A2.1	CLO 1, 2

	2.4 Suy diễn Bayes	<p><b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.</p> <p><b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.</p>		
3	Chương 3. Lập trình Python 3.1 Các công cụ lập trình: Jupyter notebook và Google Colab 3.2 Cơ bản về Python	<p><b>Dạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.</p> <p><b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.</p> <p><b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.</p>	A1.1, A1.2, A2.1	CLO 1, 2
4	Chương 3. Lập trình Python cho KHDL 3.2 Cơ bản về Python (tt)	<p><b>Dạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.</p> <p><b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.</p> <p><b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.</p>	A1.1, A1.2, A2.1	CLO 3
5	Chương 3. Lập trình Python cho KHDL (tt) 3.2 Xử lý dữ liệu vector và ma trận với Numpy	<p><b>Dạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.</p> <p><b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.</p> <p><b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.</p>	A1.1, A1.2, A2.1	CLO 3
6	Chương 3. Lập trình Python cho KHDL (tt) 3.3 Thao tác dữ liệu với Pandas	<p><b>Dạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.</p> <p><b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.</p>	A1.1, A1.2, A2.1	CLO 3

		<b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.		
7	Chương 3. Lập trình Python cho KHDL (tt) 3.4 Trực quan hoá dữ liệu với Matplotlib và Seaborn	<b>Đạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.  <b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.  <b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.	A1.1, A1.2, A2.1	CLO 3
8	Kiểm tra giữa kỳ (thi trên máy)	<b>Đạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.  <b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.  <b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.	A1.1, A1.2, A2.1	CLO 3
9	Chương 4. Thu thập và lưu trữ dữ liệu 4.1 Đọc và ghi dữ liệu từ file 4.2 Thu thập dữ liệu từ web 4.3 Sử dụng các API 4.4 Tương tác với cơ sở dữ liệu	<b>Đạy:</b> Nghe sinh viên trình bày báo cáo theo nhóm và đánh giá. <b>Học ở lớp:</b> Trình bày báo cáo theo nhóm trên lớp, trao đổi và trả lời các câu hỏi của giảng viên. <b>Tự học:</b> Ôn tập.	A1.2	
10	Chương 5. Xử lý dữ liệu 5.1 Thống kê mô tả về dữ liệu 5.2 Làm sạch dữ liệu 5.3 Chuẩn hoá dữ liệu	<b>Đạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.  <b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.  <b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.	A1.1, A2.1	CLO 3
11	Bài tập Chương 4 và Chương 5			
12	Chương 6. Phân tích dữ liệu chuỗi thời gian	<b>Đạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.	A1.1, A2.1	CLO 1, 2, 3, 4, 5

		<p><b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.</p> <p><b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.</p>		
13	Chương 7. Phân tích dữ liệu nhiều chiều	<p><b>Dạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.</p> <p><b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.</p> <p><b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.</p>	A1.1, A2.1	CLO 1, 2, 3, 4, 5
14	Chương 8. Học máy và ứng dụng 8.1. Giới thiệu scikit-learn và Tensor Flows	<p><b>Dạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.</p> <p><b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.</p> <p><b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.</p>	A1.1, A2.1	CLO 1, 2, 3, 4, 5
15	Chương 8. Học máy và ứng dụng (tt) 8.2. Một số ứng dụng thực tế	<p><b>Dạy:</b> Giảng bài kết hợp trình chiếu slide. Đặt câu hỏi cho sinh viên suy nghĩ và trả lời.</p> <p><b>Học ở lớp:</b> Nghe, thảo luận, trao đổi và trả lời các câu hỏi của giảng viên.</p> <p><b>Tự học:</b> Đọc trước bài giảng, tham khảo thêm tài liệu.</p>	A1.1, A2.1	CLO 1, 2, 3, 4, 5
16	Báo cáo Bài tập nhóm	<p><b>Dạy:</b> Nghe sinh viên trình bày báo cáo theo nhóm và đánh giá.</p> <p><b>Học ở lớp:</b> Trình bày báo cáo theo nhóm trên lớp, trao đổi và trả lời các câu hỏi của giảng viên.</p> <p><b>Tự học:</b> Ôn tập.</p>	A1.1	
	Kiểm tra cuối kỳ	Thi cuối kỳ	A2.1	

**16. Tài liệu học tập:****16.1 Sách, bài giảng, giáo trình chính:**

- [1] Nguyễn Văn Tuấn, *Phân tích dữ liệu với R*, NXB Tổng Hợp TPHCM, 2014.
- [2] Jake VanderPlas, *Python Data Science Handbook: Essential Tools for Working with Data*, O'Reilly Media, 2016.

**16.2 Sách, tài liệu tham khảo:**

- [1] Joel Grus, *Data Science from Scratch: First Principles with Python*, O'Reilly Media, 2019.
- [3] Alan J. Izenman, *Python for Data Analysis*, Springer, 2013.
- [4] Gareth James, Daniela Witten, et al., *An Introduction to Statistical Learning: with Applications in R*, Springer, 2013.

**17. Đạo đức khoa học:**

- Sinh viên phải tôn trọng giảng viên và sinh viên khác.
- Sinh viên phải thực hiện quy định liên chính học thuật của Nhà trường.
- Sinh viên phải chấp hành các quy định, nội quy của Nhà trường.

**18. Ngày phê duyệt: 30/12/2020****19. Cấp phê duyệt:**

Trưởng khoa	Phụ trách CTĐT	Giảng viên biên soạn
PGS. TS. Nguyễn Tấn Khôi	TS. Đặng Hoài Phương	TS. Ninh Khánh Duy