

ĐỀ CƯƠNG CHI TIẾT HỌC PHẦN

1. THÔNG TIN CHUNG VỀ HỌC PHẦN

1.1 Tên học phần (tiếng Việt):	Phân tích dữ liệu
1.2 Tên học phần (Tiếng Anh):	Data Analysis
1.3 Trình độ đào tạo:	Đại học
1.4 Định hướng đào tạo:	Ứng dụng
1.5 Mã học phần:	011710
1.6 Số tín chỉ:	3 (2LT+1TH)
1.7 Thuộc khối kiến thức:	Bổ trợ
1.8 Phân bổ thời gian (số giờ):	
- Lý thuyết:	30
- Bài tập, thảo luận:	0
- Thực hành:	30
- Tự học:	90
1.9 Bộ môn/Khoa phụ trách học phần:	Bộ môn Chuyên ngành - Khoa Công nghệ thông tin
1.10 Điều kiện tham gia học phần:	
- Điều kiện tiên quyết:	Không
- Học phần học trước:	Cấu trúc dữ liệu và giải thuật
- Học phần song hành:	Không

2. MÔ TẢ TÓM TẮT NỘI DUNG CỦA HỌC PHẦN

Học phần cung cấp cho sinh viên kiến thức và kỹ năng cơ bản để phân tích dữ liệu như các kiến thức tổng quan về phân tích dữ liệu, ngôn ngữ lập trình Python, tiền xử lý dữ liệu, phân tích và trực quan dữ liệu

3. MỤC TIÊU CỦA HỌC PHẦN:

3.1 Mục tiêu chung

Sau khi học học phần này, sinh viên hiểu và ứng dụng được những kiến thức và kỹ năng cần thiết để xử lý, thao tác, phân tích và trực quan hóa các kết quả dựa trên bộ dữ liệu. Sinh viên sử dụng được ngôn ngữ lập trình thông dụng trong lĩnh vực khoa học dữ liệu là ngôn ngữ Python.

3.2. Mục tiêu cụ thể

Kiến thức

- Hiểu cách lập trình Python để thực hiện các dự án phân tích dữ liệu.
- Nắm vững kiến thức về phân tích dữ liệu với các thư viện Numpy, Pandas, Matplotlib, Scikit-learn.
- Thực hiện được các bước cần tiến hành trong một dự án phân tích dữ liệu thực tế.

Kỹ năng

- Vận dụng thành thạo các thư viện trong Python để thu thập, tiền xử lý, làm sạch dữ liệu
- Có khả năng sử dụng Python để trực quan hóa dữ liệu

Năng lực tự chủ, tự chịu trách nhiệm

- Rèn luyện năng lực tự học, tự nghiên cứu, có khả năng làm việc nhóm.
- Ý thức về việc phát triển năng lực chuyên môn của cá nhân.

4. CHUẨN ĐẦU RA CỦA HỌC PHẦN:

4.1 Chuẩn đầu ra của học phần:

Ký hiệu	Nội dung
	Kiến thức
CLO1	Nhận diện về ngôn ngữ lập trình Python
CLO2	Hiểu các thư viện phân tích dữ liệu như Numpy, Pandas, Matplotlib, Scikit-learn
CLO3	Đánh giá được các bước cần thực hiện để phân tích dữ liệu
CLO4	Triển khai các thư viện đã được xây dựng vào Python để phân tích bài toán thực tế
	Kỹ năng
CLO5	Vận dụng thành thạo các thao tác cơ bản của ngôn ngữ python
CLO6	Vận dụng thành thạo các thư viện trong Python để thực hiện tiền xử lý, làm sạch dữ liệu
CLO7	Vận dụng thành thạo các thư viện trong Python để thực hiện trực quan hóa dữ liệu
CLO9	Kết hợp được Python và các thư viện hỗ trợ để thực hiện một dự án phân tích dữ liệu

	Năng lực tự chủ, tự chịu trách nhiệm
CLO10	Rèn luyện năng lực tự học, tự nghiên cứu, làm việc nhóm
CLO11	Hình thành ý thức nghiêm túc trong học tập

4.2 Mối quan hệ giữa chuẩn đầu ra của học phần (CLOs) với chuẩn đầu ra của chương trình đào tạo (PLOs):

Chuẩn đầu ra của học phần	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8	PLO9	PLO10
CLO 1			R	R						
CLO 2			R	R						
CLO 3			R	R						
CLO 4			R	R						
CLO 5							R			
CLO 6							R			
CLO 7							R			
CLO 8							R			
CLO 9							R			
CLO 10									M	
CLO 11									M	
Tổng hợp học phần			R	R			R		M	

Chú thích: Mức độ đóng góp, hỗ trợ của CLOs với PLOs được xác định cụ thể như sau:

- *I (Introduced)* – Học phần có hỗ trợ đạt được PLO và ở mức giới thiệu/bắt đầu.
- *R (Reinforced)* – Học phần có hỗ trợ đạt được PLO và ở mức nâng cao hơn mức bắt đầu, có nhiều cơ hội được thực hành, thí nghiệm, thực tế,...
- *M (Mastery)* – Học phần có hỗ trợ cao đạt được PLO và ở mức thuần thục/thông hiểu
- *A (Assessed)* – Học phần quan trọng (hỗ trợ tối đa việc đạt được PLO) cần được thu thập minh chứng để đánh giá chuẩn đầu ra của chương trình đào tạo.

5. NỘI DUNG VÀ KẾ HOẠCH GIẢNG DẠY:

Thời gian	Nội dung	Hình thức tổ chức dạy-học					Phương pháp giảng dạy – học tập	Yêu cầu sinh viên chuẩn bị trước khi đến lớp
		GIỜ LÊN LỚP						
		Lý thuyết	Thực hành tích hợp	Nghiên cứu tình huống và thảo luận	Thực hành tại PM, PMP, DN	Tự học		
Tuần 1-2	Giới thiệu học phần: kế hoạch học tập, đề cương học phần, tài liệu tham khảo, cách đánh giá... Chương 1. Giới thiệu về lập trình Python cho phân tích dữ liệu 1.1. Tổng quan về Python 1.2. Các thư viện cơ bản cho phân tích dữ liệu 1.3. Cài đặt và thiết lập môi trường 1.4. Python cơ bản, IPython và Jupyter Notebook Câu hỏi và thảo luận Bài tập chương 1	4	0	0	3	11	Thuyết giảng, đặt vấn đề, phân tích bài toán và thực hành	Xem trước bài giảng chương 1 và các tài liệu tham khảo theo định hướng của GV
Tuần 2-4	Chương 2. Các kiểu dữ liệu tích hợp, Hàm và làm việc với tập tin 2.1. Các kiểu dữ liệu tích hợp sẵn 2.2. Cơ bản về hàm trong Python 2.3. Làm việc với tập tin và hệ điều hành Câu hỏi và thảo luận Bài tập chương 2	4	0	0	5	13	Thuyết giảng, đặt vấn đề, phân tích bài toán và thực hành	Xem trước bài giảng chương 2 và các tài liệu tham khảo theo định hướng của GV

Thời gian	Nội dung	Hình thức tổ chức dạy-học					Phương pháp giảng dạy – học tập	Yêu cầu sinh viên chuẩn bị trước khi đến lớp
		GIỜ LÊN LỚP						
		Lý thuyết	Thực hành tích hợp	Nghiên cứu tình huống và thảo luận	Thực hành tại PM, PMP, DN	Tự học		
Tuần 4-5	Chương 3. Thư viện Numpy 3.1. Giới thiệu tổng quan 3.2. NumPy ndarray: Mảng nhiều chiều 3.3. Tạo số ngẫu nhiên 3.4. Làm việc với mảng trên numpy 3.4. Các hàm trên Numpy 3.5. Các ma trận đặc biệt 3.6. Các phép toán trên ma trận Câu hỏi và thảo luận Bài tập chương 3	4	0	0	4	12	Thuyết giảng, đặt vấn đề, phân tích bài toán và thực hành	Xem trước bài giảng chương 3 và các tài liệu tham khảo theo định hướng của GV
Tuần 5-6	Chương 4. Thao tác dữ liệu với Pandas 4.1. Giới thiệu tổng quan 4.2. Các chức năng quan trọng 4.3. Thống kê trên pandas Câu hỏi và thảo luận Bài tập chương 4	4	0	0	4	12	Thuyết giảng, đặt vấn đề, phân tích bài toán và thực hành	Xem trước bài giảng chương 4 và các tài liệu tham khảo theo định hướng của GV

Thời gian	Nội dung	Hình thức tổ chức dạy-học					Phương pháp giảng dạy – học tập	Yêu cầu sinh viên chuẩn bị trước khi đến lớp
		GIỜ LÊN LỚP						
		Lý thuyết	Thực hành tích hợp	Nghiên cứu tình huống và thảo luận	Thực hành tại PM, PMP, DN	Tự học		
Tuần 6-8	Chương 5. Làm việc với dữ liệu 5.1. Giới thiệu tổng quan 5.2. Tải và lưu trữ dữ liệu 5.3. Tiền xử lý dữ liệu 5.4. Làm sạch dữ liệu Câu hỏi và thảo luận Bài tập chương 5	4	0	0	4	12	Thuyết giảng, đặt vấn đề, phân tích bài toán và thực hành	Xem trước bài giảng chương 5 và các tài liệu tham khảo theo định hướng của GV
Tuần 9-10	Chương 6. Trực quan hóa dữ liệu 6.1. Giới thiệu về matplotlib 6.2. Vẽ biểu đồ trên matplotlib 6.3. Các thư viện khác Câu hỏi và thảo luận Bài tập chương 6	5	0	0	5	15	Thuyết giảng, đặt vấn đề, phân tích bài toán và thực hành	Xem trước bài giảng chương 6 và các tài liệu tham khảo theo định hướng của GV
Tuần 11-12	Chương 7. Phát triển mô hình Machine Learning cơ bản 7.1. Giới thiệu tổng quan 7.2. Thư viện Scikit-learn 7.3. Quy trình xây dựng mô hình Machine learning 7.4. Một số mô hình Machine learning Câu hỏi và thảo luận Bài tập chương 7	5	0	0	5	15	Thuyết giảng, đặt vấn đề, phân tích bài toán và thực hành	Xem trước bài giảng chương 7 và các tài liệu tham khảo theo định hướng của GV

Thời gian	Nội dung	Hình thức tổ chức dạy-học					Phương pháp giảng dạy – học tập	Yêu cầu sinh viên chuẩn bị trước khi đến lớp
		GIỜ LÊN LỚP						
		Lý thuyết	Thực hành tích hợp	Nghiên cứu tình huống và thảo luận	Thực hành tại PM, PMP, DN	Tự học		
	Tổng cộng	30	0	0	30	90		

6. HỌC LIỆU:

6.1 Tài liệu chính:

[1]. Wes McKinney (2022), *Python for Data Analysis*, O'Reilly

6.2 Tài liệu tham khảo:

[1]. Python documentation, <https://docs.python.org/>

[2]. Numpy documentation, <https://numpy.org/doc/>

[3]. Matplotlib: Visualization with Python, <https://matplotlib.org/>

[4]. Scikit-learn User Guide, https://scikit-learn.org/stable/user_guide.html

7. PHƯƠNG PHÁP ĐÁNH GIÁ KẾT QUẢ HỌC TẬP HỌC PHẦN

7.1 Đánh giá quá trình: trọng số 50% trong điểm học phần

TT	PHƯƠNG PHÁP ĐÁNH GIÁ	MÔ TẢ	CÁC CDR ĐƯỢC ĐÁNH GIÁ	CẤU TRÚC ĐIỂM THÀNH PHẦN
1	Chuyên cần, đóng góp vào bài học	Đánh giá tính chuyên cần và tích cực tham gia quá trình học tập	CLO10, CLO11	20%
2	Bài tập cá nhân	Thống kê số lần thảo luận/thực hiện bài tập mẫu. Đánh giá mức độ tiếp thu kiến thức và kỹ năng vận dụng kiến thức đó vào bài tập ứng với từng buổi học	CLO5, CLO6, CLO7, CLO8, CLO9, CLO10	15%
2	Bài tập nhóm	Mỗi nhóm được chọn một đề tài/một phần nội dung để tìm hiểu trước sau đó báo cáo với giảng viên	CLO1 – CLO4 CLO10, CLO11	15%

TT	PHƯƠNG PHÁP ĐÁNH GIÁ	MÔ TẢ	CÁC CDR ĐƯỢC ĐÁNH GIÁ	CẤU TRÚC ĐIỂM THÀNH PHẦN
3	Bài kiểm giữa kỳ	Đánh giá mức độ tiếp thu và tổng hợp một cách logic khối kiến thức làm cơ sở vận dụng vào bài toán	CLO1 – CLO9	50%
	TỔNG			100%

7.2 Đánh giá kết thúc học phần: trọng số 50% trong điểm học phần

PHƯƠNG PHÁP ĐÁNH GIÁ	MÔ TẢ	CÁC CDR ĐƯỢC ĐÁNH GIÁ	CẤU TRÚC ĐIỂM THÀNH PHẦN
Đồ án học phần: Thực hiện một dự án phân tích dữ liệu thực tế. Đề tài được đăng ký hoặc được Giảng viên [GV] giao vào đầu học kỳ, nộp theo quy định. Thực hiện theo cá nhân.	Đánh giá mức độ tiếp thu kiến thức và vận dụng trong các chương	CLO1 – CLO4	20%
	Đánh giá Kỹ năng tư duy logic trong hoạch định, thiết lập và triển khai một dự án phân tích dữ liệu	CLO5 – CLO9	60%
	Đánh giá về mức tự chủ tự chịu trách nhiệm: tương tác của sinh viên; Khả năng tự học, tự nghiên cứu	CLO10, CLO11	20%
Tổng			100%

8. QUY ĐỊNH CỦA HỌC PHẦN

8.1 Quy định về tham dự lớp học

Tham gia lớp học nghiêm túc, tuân thủ các quy định chung của nhà trường.

8.2 Quy định về hành vi lớp học

- Nghiêm túc thực hiện các bài tập, bài kiểm tra trong mỗi buổi học.
- Chuẩn bị bài trước khi đến lớp theo yêu cầu của giảng viên
- Sinh viên không làm việc riêng trong thời gian lên lớp, chỉ sử dụng các thiết bị phục vụ cho nội dung học tập và hoạt động tương ứng trong từng buổi học

8.3 Quy định về học vụ

- Sinh viên vắng vào các buổi có làm bài đánh giá cá nhân, nhóm sẽ bị 0 điểm đối với bài đánh giá đó.

- Thành viên nhóm thuyết trình không được vắng, nếu vắng sẽ không được làm bài kiểm tra thay thế.
- Các bài tập nhóm, cá nhân phải tuân thủ theo mẫu và thời hạn nộp theo quy định của giảng viên.

9. HOLISTIC RUBRIC

Rubric 1: Đánh giá chuyên cần, thái độ

Tiêu chí	Tỷ lệ	Mức chất lượng			
		Rất tốt	Tốt	Đạt	Không đạt
		10 - 8.5	8.4 - 7.0	6.9 - 5.0	4.9 - 0.0
Mức độ tham dự theo TKB	50%	Tham dự >90% buổi học	Tham dự 80 - 90% buổi học	Tham dự 70 - 80% buổi học	Tham dự <70% buổi học
Mức độ tham gia các hoạt động học tập	50%	Nhiệt tình trao đổi, phát biểu, trả lời nhiều câu hỏi	Có đặt/trả lời > 2 câu hỏi	Có đặt/trả lời ít nhất 1 câu hỏi	Không tham gia thảo luận, trả lời, đóng góp

Rubric 2: Đánh giá bài tập cá nhân

Tiêu chí	Tỷ lệ	Mức chất lượng			
		Rất tốt	Tốt	Đạt	Không đạt
		10 - 8.5	8.4 - 7.0	6.9 - 5.0	4.9 - 0.0
Bài tập	100%	- Giải quyết đúng 100% yêu cầu của bài tập - Nộp bài đúng thời gian quy định	- Giải quyết đúng 70%-80% yêu cầu bài tập - Nộp bài đúng thời gian quy định	- Giải quyết đúng 50% - 70% yêu cầu bài tập - Nộp bài quá thời gian quy định	- Giải quyết đúng <50% yêu cầu bài tập - Nộp bài quá thời gian quy định

Rubric 3: Đánh giá bài tập nhóm

Tiêu chí	Tỷ lệ	Mức chất lượng			
		Rất tốt	Tốt	Đạt	Không đạt
		10 - 8.5	8.4 - 7.0	6.9 - 5.0	4.9 - 0.0
Tổ chức và phối hợp tham gia thực hiện của các thành	10%	Từ 80% đến 100% thành viên tham gia thực hiện	Từ 60% đến 80% số lượng thành viên tham gia thực	Từ 40% đến 60% số lượng thành viên trở lên tham gia thực hiện	< 40% số lượng thành viên tham gia

viên trong nhóm		nhiệm vụ nhóm	hiện nhiệm vụ nhóm	nhiệm vụ nhóm	thực hiện nhiệm vụ nhóm
Hình thức bài báo cáo kết quả làm việc nhóm	15%	Đẹp, rõ ràng các bộ cục (mở đầu, nội dung chính, kết luận), không lỗi chính tả	Đầy đủ các bộ cục (mở đầu, nội dung chính, kết luận) rõ ràng, font chữ đúng quy định, còn lỗi chính tả.	Các bộ cục (mở đầu, nội dung chính, kết luận) chưa rõ, font chữ đúng quy định, còn lỗi chính tả.	Sơ sài, không đầy đủ các bộ cục (mở đầu, nội dung chính, kết luận), font chữ không đồng bộ, nhiều lỗi chính tả.
Nội dung	50%	Đáp ứng 100% yêu cầu của bài tập nhóm	Đáp ứng 70% - 80% yêu cầu của bài tập nhóm	Đáp ứng 50% - 70% yêu cầu của bài tập nhóm	Đáp ứng <50% yêu cầu của bài tập nhóm
Thuyết trình và trả lời câu hỏi	25%	Trả lời đúng từ 85% đến 100% số câu hỏi được đặt ra.	Trả lời đúng từ 70% đến dưới 85% số câu hỏi được đặt ra.	Trả lời đúng từ 50% đến dưới 70% số câu hỏi được đặt ra.	Trả lời đúng dưới 50% số câu hỏi được đặt ra

Rubric 4: Đánh giá kiểm tra giữa kỳ

Tiêu chí	Tỷ lệ	Mức chất lượng			
		Rất tốt	Tốt	Đạt	Không đạt
		10 - 8.5	8.4 - 7.0	6.9 - 5.0	4.9 - 0.0
Bài tập	100%	- Giải quyết đúng 100% yêu cầu của bài tập - Nộp bài đúng thời gian quy định	- Giải quyết đúng 70%-80% yêu cầu bài tập - Nộp bài đúng thời gian quy định	- Giải quyết đúng 50% - 70% yêu cầu bài tập - Nộp bài quá thời gian quy định	- Giải quyết đúng <50% yêu cầu bài tập - Nộp bài quá thời gian quy định

Rubric 5: Đánh giá cuối kỳ

Tiêu chí	Tỷ lệ	Mức chất lượng			
		Rất tốt	Tốt	Đạt	Không đạt
		10 - 8.5	8.4 - 7.0	6.9 - 5.0	4.9 - 0.0
Đánh giá kỹ năng thiết kế báo cáo	10%	Giải quyết đúng 85% - 100% yêu cầu	Giải quyết đúng 70% - 84% yêu cầu	Giải quyết đúng 50% - 70% yêu cầu	- Giải quyết đúng <50% yêu cầu
Đánh giá khả năng trình bày báo cáo đúng chuẩn công việc	10%	Giải quyết đúng 85% - 100% yêu cầu	Giải quyết đúng 70% - 84% yêu cầu	Giải quyết đúng 50% - 70% yêu cầu	- Giải quyết đúng <50% yêu cầu
Đánh giá kỹ năng trích dẫn và trình bày tài liệu tham khảo	10%	Giải quyết đúng 85% - 100% yêu cầu	Giải quyết đúng 70% - 84% yêu cầu	Giải quyết đúng 50% - 70% yêu cầu	- Giải quyết đúng <50% yêu cầu
Đánh giá mức độ tiếp thu kiến thức; Kỹ năng tư duy logic trong hoạch định, thiết lập và triển khai dự án phân tích dữ liệu	70%	Giải quyết đúng 85% - 100% yêu cầu	Giải quyết đúng 70% - 84% yêu cầu	Giải quyết đúng 50% - 70% yêu cầu	- Giải quyết đúng <50% yêu cầu

Thành phố Hồ Chí Minh, ngày 30 tháng 8 năm 2023

TUQ. Hiệu trưởng
Trưởng phòng QLKH

PGS.TS Phan Thị Hằng Nga

Trưởng khoa

TS. Trương Thành Công

Tr. Bộ môn

TS. Trương Thành Công

