

CHƯƠNG TRÌNH GIÁO DỤC ĐẠI HỌC

CHƯƠNG TRÌNH ĐÀO TẠO NGÀNH CÔNG NGHỆ THÔNG TIN

ĐỀ CƯƠNG CHI TIẾT HỌC PHẦN

DIT0550 – NHẬP MÔN PHÂN TÍCH DỮ LIỆU VÀ HỌC SÂU INTRODUCTION TO DATA ANALYTICS AND DEEP LEARNING

1 Thông tin về học phần

1.1 Số tín chỉ: 03 (2 LT + 1 TH)

1.2 Số tiết và/hoặc số giờ đối với các hoạt động học tập:

- Số tiết lý thuyết trên lớp: 2 TC * 15 tiết (**30 tiết**)
- Số tiết thực hành, thực tập trên lớp: 1 TC * 30 tiết (**30 tiết**)
- Số giờ kiến tập, thực tập tại doanh nghiệp, làm tiểu luận, bài tập lớn, làm đồ án, đồ án/khoá luận tốt nghiệp (hoạt động thực tiễn ngoài lớp học): 3 TC * 45 giờ (**135 giờ**)
- Số giờ tự học của sinh viên: **60 giờ**

1.3 Học phần thuộc khối kiến thức:

<input type="checkbox"/> Kiến thức giáo dục đại cương	Kiến thức giáo dục chuyên nghiệp <input checked="" type="checkbox"/>		
<input type="checkbox"/> Bắt buộc	<input type="checkbox"/> Kiến thức cơ sở khối ngành	<input type="checkbox"/> Kiến thức cơ sở ngành	<input checked="" type="checkbox"/> Kiến thức ngành
<input type="checkbox"/> Tự chọn	<input type="checkbox"/> Bắt buộc <input type="checkbox"/> Tự chọn	<input type="checkbox"/> Bắt buộc <input type="checkbox"/> Tự chọn	<input type="checkbox"/> Bắt buộc <input checked="" type="checkbox"/> Tự chọn

1.4 Học phần tiên quyết: **Cơ sở dữ liệu (DIT0050), Cấu trúc Dữ liệu và Giải thuật (DIT0030)**

1.5 Học phần học trước: **Lập trình Python nâng cao (DIT0540)**

1.6 Học phần được giảng dạy ở học kỳ thứ: **V, VII** Khóa: **K25, K24**

1.7 Ngôn ngữ và tài liệu giảng dạy:

Tiếng Việt: Bài giảng, tài liệu học tập, hướng dẫn thực hành

Tiếng Anh: Tài liệu tham khảo (Ebook, Website)

1.8 Đơn vị phụ trách:

- Bộ môn/Ngành: Mạng máy tính và IoT / Công nghệ Thông tin
- Khoa: Công nghệ Thông tin

2 Mục tiêu và chuẩn đầu ra của học phần

2.1 Mục tiêu của học phần

Kiến thức: Sau khi kết thúc học phần, sinh viên hiểu được mô hình phân tích dữ liệu, các đoạn mã và thư viện lập trình Python dùng trong phân tích dữ liệu trên máy tính, trong cơ sở dữ liệu, trên môi trường mạng và điện toán đám mây. Với các kiến thức cơ bản về phân tích dữ liệu, sinh viên có thể tiếp cận mô hình học sâu và dự đoán dựa trên các mạng nơ ron (neural network) bằng các công cụ được tích hợp trong các thư viện học sâu như Keras, Scikit-Learning. Sinh viên sẽ hiểu được các kỹ thuật và công cụ cơ bản cần thiết để sử dụng

trong mô hình học sâu, tích lũy kinh nghiệm với các kiểu dữ liệu học sâu phổ biến, hiểu kiến trúc mô hình, các bộ dữ liệu nâng cao để cải thiện độ chính xác của mô hình, quá trình học chuyển giao giữa các mô hình để đạt được kết quả với ít dữ liệu và tính toán hơn, từ đó có được có hiểu biết cơ bản khi thực hiện một dự án ứng dụng với nền tảng học sâu.

Kỹ năng: Sinh viên có thể thiết kế, phát triển và cải tiến một số chương trình máy tính sử dụng các mô-đun tích hợp, ứng dụng trong quá trình nhập dữ liệu, làm sạch, chuẩn bị dữ liệu để phân tích, tóm tắt dữ liệu, vận hành các đường ống dữ liệu. Sinh viên làm quen và sử dụng thành thạo công cụ Pandas DataFrames, mảng đa chiều Numpy và thư viện SciPy để làm việc với các bộ dữ liệu khác nhau. Sinh viên cũng được cung cấp kỹ năng tải, thao tác, phân tích và trực quan hóa dữ liệu với Pandas (thư viện Python mã nguồn mở). Vận dụng mạng nơ ron nhân tạo (neural network), mạng nơ ron tích chập (convolution neural network), mạng nơ ron hồi quy (recurrent neural network) vào các bài toán trong thực tế như: bài toán nhận dạng hình ảnh, bài toán về dữ liệu chuỗi thời gian, bài toán xử lý ngôn ngữ tự nhiên.

Thái độ: Đi học chuyên cần, tự giác điểm danh, chuẩn bị bài học và thực hành theo yêu cầu và hướng dẫn của giáo viên, trợ giảng.

2.2 Chuẩn đầu ra của học phần và ma trận tương thích giữa CDR học phần (CELO) với CDR chương trình đào tạo (ELO):

Ký hiệu	CDR của học phần Hoàn thành học phần này, sinh viên có thể	CDR của CTĐT
	Kiến thức	
CELO1	Hiểu được nền tảng phân tích dữ liệu dựa trên các thuật toán, quy trình phát triển ứng dụng (thư viện phần mềm) dùng trong phân tích dữ liệu bằng ngôn ngữ lập trình Python	ELO1
CELO2	Làm quen với các khái niệm cơ bản của Học sâu, hiểu biết về các thư viện và công cụ sử dụng phổ biến trong Học sâu.	ELO1
CELO3	Có kiến thức cơ bản về quy trình phân tích dữ liệu, xây dựng và khai thác các thư viện và giải thuật trong Học sâu để giải quyết một số bài toán thực tế	ELO2
	Kỹ năng	
CELO4	Thực hiện được một số bài toán phân tích dữ liệu trên máy tính bằng các giải thuật và thư viện (ngôn ngữ lập trình Python)	ELO3
CELO5	Lựa chọn đúng các giải thuật học sâu cơ bản để giải quyết bài toán thực tế thông qua các thư viện Keras, Scikit-Learning	ELO4
	Năng lực tự chủ và trách nhiệm	
CELO6	Thái độ học tập nghiêm túc, tự giác hoàn thành các yêu cầu của môn học	ELO5

2.3 Ma trận chức năng (mức độ đóng góp của học phần cho các ELOs)

TT	Mã học phần	Tên học phần	Số tín chỉ	Kiến thức			Kỹ năng		Năng lực tự chủ và trách nhiệm	
				ELO 1	ELO 2	ELO3	ELO 3	ELO4	ELO 5	
1	DIT0550	Nhập môn phân tích dữ liệu và học sâu	3	H	H	S	S	N	S	N

- *N: Không đóng góp (none supported)*
- *S: Có đóng góp (suppoorted)*
- *H: Đóng góp quan trọng (highly supported)*

3 Mô tả vắn tắt nội dung học phần

- + Sử dụng môi trường điện toán đám mây **Dataiku DSS** <https://academy.dataiku.com> và Blended Learning LAB (<https://fit.vlu.edu.vn>) để phân tích dữ liệu.
- + Học và sử dụng được các tính năng cơ bản và nâng cao trong thư viện Numpy, Pandas.
- + Sử dụng các công cụ phù hợp để tải dữ liệu, làm sạch dữ liệu, chuyển đổi dữ liệu, trộn dữ liệu và thay đổi hình dạng dữ liệu (reshape).
- + Tạo các đồ thị trích xuất thông tin với thư viện Matplotlib.
- +Áp dụng chức năng groupby trong thư viện Pandas để thực hiện chức năng slide, dice và tóm tắt dữ liệu.
- +Phân tích và thao tác với kiểu dữ liệu chuỗi thời gian (time series).
- +Học cách giải quyết các vấn đề phân tích dữ liệu trong thế giới thực thông qua các ví dụ chi tiết.
- +Học cách áp dụng mạng nơ ron nhân tạo (neural network), mạng nơ ron tích chập (convolution neural network), mạng nơ ron hồi quy (recurrent neural network) vào các bài toán trong thực tế như: bài toán nhận dạng hình ảnh, bài toán về dữ liệu chuỗi thời gian, bài toán xử lý ngôn ngữ tự nhiên.
- +Học sách sử dụng các thư viện trong python như Karas, Tensorflow, OpenCV để giải quyết các bài toán về mạng nơ ron.

4 Phương pháp giảng dạy và học tập

4.1 Phương pháp giảng dạy

Giảng dạy lý thuyết trên giờ học, sinh viên có thể đọc bài giảng trước và sau giờ, tự tìm hiểu, phân tích tình huống

Thảo luận các vấn đề thực tế theo gợi ý của giảng viên

4.2 Phương pháp học tập

Nghe giảng, tự đọc tài liệu, làm bài tập, viết và trình bày báo cáo, đồ án môn học.

5 Nhiệm vụ của sinh viên

Nhiệm vụ của sinh viên như sau:

- + Tham gia đầy đủ các buổi học trên lớp. Nếu vắng điểm danh 03 buổi sẽ bị cấm thi lần 1
- + Phải đi học đúng giờ theo quy định. Nếu đi trễ quá 15 phút sinh viên không được vào lớp;
- + Sinh viên đi học phải ăn mặc lịch sự, phù hợp với môi trường giáo dục. Không sử dụng điện thoại trong lớp, không được nói chuyện và làm việc riêng. Nếu giảng viên nhắc nhở mà sinh viên vẫn tiếp tục vi phạm thì có thể bị mời ra khỏi lớp và thông báo để Khoa xử lý;
- + Sinh viên làm việc nhóm, chuẩn bị bài thuyết trình, bài tập lớn, đồ án môn học
- + Lập các nhóm học tập (mỗi nhóm khoảng 6 SV), cử 1 nhóm trưởng;
- + Chuẩn bị bài trước khi đến lớp, như: làm bài tập, đọc trước tài liệu như đã hướng dẫn trong lịch trình giảng dạy.

6 Đánh giá và cho điểm

6.1 Thang điểm

Thang điểm **10** và quy đổi thành thang điểm chữ (A-F), thang điểm 4 theo quy chế đào tạo tín chỉ hiện hành của Trường Đại học Văn Lang.

6.2 Rubric đánh giá

Các tiêu chí và trọng số điểm đối với từng nội dung cần đánh giá được trình bày trong phần phụ lục đính kèm đề cương chi tiết này.

6.3 Kế hoạch đánh giá và trọng số thành phần đánh giá

Bảng 1 Ma trận phương pháp đánh giá để đạt CDR của học phần

CELOs	PHƯƠNG PHÁP ĐÁNH GIÁ					CÔNG CỤ ĐÁNH GIÁ	THỜI ĐIỂM ĐÁNH GIÁ
	Làm bài tập	Thảo luận, phát biểu	Thực hành tại PTN	Thi giữa kỳ	Thi cuối kỳ		
CELO 1		x		x	x	- Bài tập cuối chương - Thi trắc nghiệm cuối kỳ đề mở / đóng	- Cuối mỗi chương - Tuần 7 - Cuối kỳ
CELO 2		x		x	x	- Bài tập cuối chương - Thi trắc nghiệm cuối kỳ đề mở / đóng	- Cuối mỗi chương - Tuần 7 - Cuối kỳ
CELO 3		x			x	- Thi trắc nghiệm giữa kỳ đề mở / đóng	- Cuối mỗi chương - Tuần 7 - Cuối kỳ
CELO 4	x		x	x			- Hàng tuần
CELO 5	x		x	x		- Bài kiểm tra thường xuyên	- Hàng tuần
CELO 6	x		x	x			- Hàng tuần

Bảng 2 Trọng số thành phần đánh giá của học phần

TT	Thành phần	Trọng số (%)	Ghi chú
1	Dự lớp	5 %	
2	Thảo luận	5 %	
3	Thi giữa Học kỳ	20 %	
4	Báo cáo Đồ án	30 %	
5	Thi cuối Học kỳ	40 %	
	Tổng	100%	

7 Giáo trình và tài liệu học tập

7.1 Tài liệu học tập

[1] **Bài giảng** “Nhập môn Phân tích dữ liệu và học sâu” - BM Mạng máy tính và IoT. - Khoa CNTT. Trường Đại học Văn Lang – 2021 - 2022

[2] **Hướng dẫn Thực hành:** Phân tích dữ liệu trong môi trường điện toán đám mây **DataIku DSS** <https://academy.dataiku.com>

[3] **Hướng dẫn Đồ án môn học:** Chủ đề học sâu trong môi trường Phòng Thí nghiệm Học tập Kết hợp (Blended Learning LAB) của Khoa CNTT <https://fit.vlu.edu.vn>

7.2 Giáo trình chính

[1] **Python Data Analytics**. 2nd Edition, Fabio Nelli, Apress Media LLC, **2018**

[2] **Python for Data Analysis**. 2nd Edition, Wes McKinney, O'REILLY Media Inc., **2018**

7.3 Giáo trình và tài liệu tham khảo khác

[1] **Deep Learning with Keras**. Frank Millstein, 2018

[2] **Advanced Data Analytics Using Python**, Sayan Mukhopadhyay, Apress Media LLC, **2018**

8 Nội dung chi tiết của học phần

8.1 Phần lý thuyết

Tuần	Nội dung	KQHTMD của HP
1	Chương 1: Nhập môn phân tích dữ liệu	CELO1
	A. Nội dung giảng dạy trên lớp: (3 giờ) 1.1 Phân tích dữ liệu: nhà phân tích, kỹ sư và nhà khoa học dữ liệu 1.2 Hiểu biết về dữ liệu và quy trình phân tích dữ liệu 1.3 Phân tích dữ liệu: định tính và định lượng 1.4 Dữ liệu mở 1.5 Ngôn ngữ lập trình và các thư viện Python Nội dung thảo luận: (3 giờ) - Thảo luận về nội dung Bài giảng - Giải đáp các thắc mắc thực hành trên trang học trực tuyến Dataiku (https://academy.dataiku.com) và Blended Learning LAB (https://fit.vlu.edu.vn)	
	B. Các nội dung sinh viên cần chuẩn bị tự học ở nhà: (3 giờ) - Đọc trước tài liệu trên Blended LAB - Đăng ký tài khoản cá nhân trên DataIku DSS	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: - Hoàn thành yêu cầu của giảng viên, trả lời các câu hỏi kiểm tra	
2	Chương 2: Các công cụ và thư viện phân tích dữ liệu trong Python	CELO1
	A. Nội dung giảng dạy trên lớp: (3 giờ) 2.1 Ngôn ngữ lập trình Python dùng trong PTDL 2.2 Thư viện SciPy 2.3 Thư viện NumPy 2.4 Thư viện Pandas	

Tuần	Nội dung	KQHTMD của HP
	<p>2.5 Thư viện Matplotlib</p> <p>Nội dung thảo luận: (3 giờ)</p> <ul style="list-style-type: none"> - Thảo luận về nội dung Bài giảng - Giải đáp các thắc mắc thực hành trên trang học trực tuyến Dataiku (https://academy.dataiku.com) và Blended Learning LAB (https://fit.vlu.edu.vn) 	
	<p>B. Các nội dung cần tự học ở nhà: (5 giờ)</p> <ul style="list-style-type: none"> - Đọc trước tài liệu học tập - Tham khảo tài liệu trên BL LAB, thực hành trên Dataiku DSS 	
	<p>C. Đánh giá kết quả học tập</p> <p>Phương pháp đánh giá:</p> <ul style="list-style-type: none"> - Hoàn thành yêu cầu của giảng viên, trả lời các câu hỏi kiểm tra - Chấm điểm bài tập theo Rubric do giảng viên đưa ra. 	
3	Chương 3 Thư viện NumPy	CELO2 CELO3
	<p>A. Nội dung giảng dạy trên lớp: (6 giờ)</p> <p>3.1 Quản lý mảng số (Ndarray)</p> <p>3.2 Các phép toán cơ bản</p> <p>3.3 Đánh chỉ số, phân đoạn, lặp</p> <p>3.4 Điều kiện và các mảng Boolean</p> <p>3.5 Thao tác hình và mảng</p> <p>3.6 Làm việc với tập tin</p> <p>Nội dung làm bài tập/thảo luận: (3 giờ)</p> <ul style="list-style-type: none"> - Thảo luận về nội dung Bài giảng - Tham khảo tài liệu trên BL LAB, thực hành trên Dataiku DSS 	
	<p>B. Các nội dung cần tự học ở nhà: (5 giờ)</p> <ul style="list-style-type: none"> - Đọc trước tài liệu học tập - Tham khảo tài liệu trên BL LAB, thực hành trên Dataiku DSS 	
	<p>C. Đánh giá kết quả học tập</p> <p>Phương pháp đánh giá:</p> <ul style="list-style-type: none"> - Hoàn thành yêu cầu của giảng viên, trả lời các câu hỏi kiểm tra - Chấm điểm bài tập theo Rubric do giảng viên đưa ra. 	
4	Chương 4 Thư viện Pandas	CELO2 CELO3
	<p>A. Nội dung giảng dạy trên lớp: (6 giờ)</p> <p>4.1 Giới thiệu thư viện Pandas</p> <p>4.2 Cấu trúc dữ liệu trong Pandas</p> <p>4.3 Đánh địa chỉ, sắp xếp, phép toán</p> <p>4.4 Làm việc với tập tin CSV và Text</p> <p>4.5 Làm việc với dữ liệu XML và JSON</p> <p>4.6 Làm việc với CSDL SQLite3, MongoDB</p> <p>4.7 Các giải thuật thao tác dữ liệu trong Pandas</p> <p>Nội dung làm bài tập/thảo luận: (3 giờ)</p> <ul style="list-style-type: none"> - Thảo luận về nội dung Bài giảng - Thực hành trên Dataiku 	
	<p>B. Các nội dung cần tự học ở nhà: (5 giờ)</p>	

Tuần	Nội dung	KQHTMD của HP
	<ul style="list-style-type: none"> - Đọc trước tài liệu học tập - Tham khảo tài liệu trên BL LAB, thực hành trên DataIku DSS 	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: <ul style="list-style-type: none"> - Hoàn thành yêu cầu của giảng viên, trả lời các câu hỏi kiểm tra - Chấm điểm bài tập theo Rubric do giảng viên đưa ra. 	
	Chương 5 Trực quan hóa dữ liệu với matplotlib	CELO3
5	A. Nội dung giảng dạy trên lớp: (6 giờ) 5.1 Giới thiệu thư viện matplotlib 5.2 Thiết kế Chart đơn giản 5.3 Cửa sổ vẽ (plotting windows) 5.4 Thêm các đặc trưng 5.5 Làm việc với thư viện NumPy 5.6 Các chart cơ bản: Pie, Bar, Collums, Contour, Polar Nội dung làm bài tập/thảo luận: (3 giờ) <ul style="list-style-type: none"> - Thảo luận về nội dung Bài giảng - Thực hành trên DataIku 	
	B. Các nội dung cần tự học ở nhà: (5 giờ) <ul style="list-style-type: none"> - Đọc trước tài liệu học tập - Tham khảo tài liệu trên BL LAB, thực hành trên DataIku DSS 	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: <ul style="list-style-type: none"> - Hoàn thành yêu cầu của giảng viên, trả lời các câu hỏi kiểm tra - Chấm điểm bài tập theo Rubric do giảng viên đưa ra. 	
6	Chương 6 Phân tích dữ liệu với học máy và học sâu A. Nội dung giảng dạy trên lớp: (6 giờ) 6.1 Học máy với thư viện Scikit-Learn 6.2 Học có giám sát và không giám sát 6.4 Dữ liệu học và dữ liệu test 6.5 Phân loại K-nearest Neighbors 6.6 Trí tuệ nhân tạo và học máy 6.7 Học máy và học sâu 6.8 Những khái niệm cơ bản trong học sâu Nội dung làm bài tập/thảo luận: (3 giờ) <ul style="list-style-type: none"> - Thảo luận về nội dung Bài giảng - Thực hành trên DataIku 	
	B. Các nội dung cần tự học ở nhà: (5 giờ) <ul style="list-style-type: none"> - Đọc trước tài liệu học tập - Tham khảo tài liệu trên BL LAB, thực hành trên DataIku DSS 	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: <ul style="list-style-type: none"> - Hoàn thành yêu cầu của giảng viên, trả lời các câu hỏi kiểm tra - Chấm điểm bài tập theo Rubric do giảng viên đưa ra. 	CELO3

Tuần	Nội dung	KQHTMD của HP
7	Chương 7 Học sâu với Tensor Flow	
	A. Nội dung giảng dạy trên lớp: (6 giờ) 7.1 Học sâu với mạng Neural Network và GPU 7.2 Các nguồn dữ liệu cho học sâu: Dữ liệu mở, IoT, Big Data 7.3 Các nền tảng Python trong học sâu 7.4 Mạng Neuron nhân tạo 7.5 TensorFlow và Lập trình khởi tạo với TensorFlow 7.6 Single Layer và Multi Layer Perception với TensorFlow Nội dung làm bài tập/thảo luận: (3 giờ) - Thảo luận về nội dung Bài giảng - Thực hành trên Dataiku	
	B. Các nội dung cần tự học ở nhà: (5 giờ) - Đọc trước tài liệu học tập - Tham khảo tài liệu trên BL LAB, thực hành trên Dataiku DSS	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: - Hoàn thành yêu cầu của giảng viên, trả lời các câu hỏi kiểm tra - Chấm điểm bài tập theo Rubric do giảng viên đưa ra.	
8	Chương 8 Hướng dẫn báo cáo đồ án môn học	
	A. Nội dung giảng dạy trên lớp: (6 giờ) 8.1 Bài toán học sâu với dữ liệu khí tượng 8.2 Bài toán học sâu với nhận dạng chữ viết tay 8.3 Bài toán học sâu với phân tích văn bản bằng NLTK (Natural Language Toolkit) 8.4 Bài toán học sâu với phân tích ảnh trong OpenCV 8.5 Bài toán học sâu với dữ liệu y khoa Nội dung làm bài tập/thảo luận: (3 giờ) - Thảo luận về nội dung Bài giảng - Thực hành trên Dataiku	
	B. Các nội dung cần tự học ở nhà: (5 giờ) - Đọc trước tài liệu học tập - Tham khảo tài liệu trên BL LAB, thực hành trên Dataiku DSS	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: - Hoàn thành yêu cầu của giảng viên, trả lời các câu hỏi kiểm tra - Chấm điểm bài tập theo Rubric do giảng viên đưa ra.	
9	Chương 9 Hướng dẫn ôn thi tự luận và vấn đáp	
	A. Nội dung giảng dạy trên lớp: (6 giờ) 9.1 Hướng dẫn ôn thi theo danh mục câu hỏi thi 9.2 Hướng dẫn trình bày báo cáo đồ án môn học cá nhân và theo nhóm 9.3 Hướng dẫn ôn thi các chứng chỉ Quốc tế liên quan Nội dung làm bài tập/thảo luận: (3 giờ) - Thảo luận về nội dung Bài giảng - Thực hành trên Dataiku	

Tuần	Nội dung	KQHTMD của HP
	B. Các nội dung cần tự học ở nhà: (5 giờ) - Đọc trước tài liệu học tập - Tham khảo tài liệu trên BL LAB, thực hành trên Dataiku DSS	
	B. Các nội dung cần tự học ở nhà: (5 giờ) - Đọc trước tài liệu học tập - Tham khảo tài liệu trên BL LAB, thực hành trên Dataiku DSS	
10	Bảo vệ Đồ án và Thi cuối kỳ + Hình thức thi + Làm bài thi viết tự luận + Trình bày báo cáo đồ án + Trả lời câu hỏi vấn đáp	

8.2 Phần thực hành

Tuần	Nội dung	KQHTMD của HP
1	Bài Thực hành 1 + Đăng ký tài khoản cá nhân trên Dataiku Data Science Studio (https://academy.dataiku.com/) để học và thực hành Online + Chuẩn bị lộ trình học tập trên Blended Learning LAB + Lập Nhóm thực hành, khởi tạo Project trên Dataiku (phân tích dữ liệu) + Làm bài tập LAB 1 : <ul style="list-style-type: none"> - Thao tác với môi trường Dataiku DSS - Cấu trúc điều khiển - Cấu trúc lặp - Viết hàm trong Python - Nhập xuất trong Python 	CELO4
	A. Nội dung làm việc trong PTN/Phòng thực hành: (03 giờ) - Chuẩn bị môi trường thực hành Blended Learning LAB - Đọc bài giảng, nộp bài tập	
	B. Các nội dung cần tự học ở nhà: (1 giờ) + Xem trước các yêu cầu thực hành trên BL LAB và Dataiku DSS	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: - Kỹ năng trình bày, kết quả thực hành	

Tuần	Nội dung	KQHTMD của HP
2	<p>Bài Thực hành 2 + Làm bài tập LAB 2:</p> <ul style="list-style-type: none"> - Tìm hiểu nghiệp vụ dữ liệu - Nhập liệu bằng công cụ từ file excel - Xác định dữ liệu định tính và định lượng - Hiệu chỉnh các thang đo phù hợp và kiểu giá trị dữ liệu cho từng biến số - Hiệu chỉnh dữ liệu và xử lý dữ liệu thiếu - Chuyển đổi (transformation) dữ liệu theo khoảng cho trước - Tạo biến số phụ thuộc theo biến độc lập - Tạo biến định tính phân loại - Làm việc với tập tin CSV và Text - Làm việc với dữ liệu XML và JSON - Làm việc với CSDL SQLite3, MongoDB <p>A. Nội dung làm việc trong PTN/Phòng thực hành: (03 giờ) - Thực hành theo hướng dẫn của giảng viên</p> <p>B. Các nội dung cần tự học ở nhà: (1 giờ) + Xem trước các yêu cầu thực hành trên BL LAB và Dataiku DSS</p> <p>C. Đánh giá kết quả học tập Phương pháp đánh giá: - Kỹ năng trình bày, kết quả thực hành</p>	CELO4

Tuần	Nội dung	KQHTMD của HP
3	<p>Bài Thực hành 3 Làm bài tập LAB 3: Xử lý các vấn đề sau về dữ liệu</p> <ul style="list-style-type: none"> - Thiếu dòng tiêu đề ở file csv - Nhiều biến lưu ở một cột - Dữ liệu cột chứa các giá trị đơn vị không nhất quán - Dữ liệu có một dòng trống - Dữ liệu có các dòng trùng lặp - Các ký tự không phải ASCII - Giá trị bị mất - Tiêu đề cột là giá trị chứ không phải tên biến <p>A. Nội dung làm việc trong PTN/Phòng thực hành: (6 giờ) - Thực hành theo hướng dẫn của giảng viên</p> <p>B. Các nội dung cần tự học ở nhà: (3 giờ) + Xem trước các yêu cầu thực hành trên BL LAB và Dataiku DSS</p>	CELO4

	C. Đánh giá kết quả học tập Phương pháp đánh giá: - Kỹ năng trình bày, kết quả thực hành	
--	---	--

Tuần	Nội dung	KQHTMD của HP
4	Bài Thực hành 4 Làm bài tập LAB 4: <ul style="list-style-type: none"> - Làm sạch dữ liệu - Phân tích dữ liệu - Data Wrangling - Khai thác thông tin hữu ích 	CELO4
	A. Nội dung làm việc trong PTN/Phòng thực hành: (6 giờ) - Thực hành theo hướng dẫn của giảng viên	
	B. Các nội dung cần tự học ở nhà: (3 giờ) + Xem trước các yêu cầu thực hành trên BL LAB và Dataiku DSS	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: - Kỹ năng trình bày, kết quả thực hành	

Tuần	Nội dung	KQHTMD của HP
5	Bài Thực hành 5 Làm bài tập LAB 5: <ul style="list-style-type: none"> - Thống kê dữ liệu - Trình bày dữ liệu - Trực quan hóa dữ liệu theo nhóm phân loại - Trực quan hóa dữ liệu nâng cao - Mô tả dữ liệu và khảo sát dạng phân phối - Bài tập tổng hợp 	CELO4
	A. Nội dung làm việc trong PTN/Phòng thực hành: (6 giờ) - Thực hành theo hướng dẫn của giảng viên	
	B. Các nội dung cần tự học ở nhà: (3 giờ) + Xem trước các yêu cầu thực hành trên BL LAB và Dataiku DSS	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: - Kỹ năng trình bày, kết quả thực hành	

Tuần	Nội dung	KQHTMD của HP
6	Bài Thực hành 6 Làm bài tập LAB 6: <ul style="list-style-type: none"> - Nhận dạng chữ viết tay bằng mạng nơ ron nhân tạo. - Nhận dạng chữ viết tay bằng mạng nơ ron tích chập. 	CELO4
	A. Nội dung làm việc trong PTN/Phòng thực hành: (6 giờ) - Thực hành theo hướng dẫn của giảng viên	
	B. Các nội dung cần tự học ở nhà: (3 giờ) + Xem trước các yêu cầu thực hành trên BL LAB và Dataiku DSS	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: - Kỹ năng trình bày, kết quả thực hành	

Tuần	Nội dung	KQHTMD của HP
7	Bài Thực hành 7 Làm bài tập LAB 7: <ul style="list-style-type: none"> - Bài toán học sâu với dữ liệu khí tượng. - Bài toán học sâu với phân tích văn bản bằng NLTK (Natural Language Toolkit). 	CELO4
	A. Nội dung làm việc trong PTN/Phòng thực hành: (6 giờ) - Thực hành theo hướng dẫn của giảng viên	
	B. Các nội dung cần tự học ở nhà: (3 giờ) + Xem trước các yêu cầu thực hành trên BL LAB và Dataiku DSS	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: - Kỹ năng trình bày, kết quả thực hành	

Tuần	Nội dung	KQHTMD của HP
8	Bài Thực hành 8 Làm bài tập LAB 8: <ul style="list-style-type: none"> - Bài toán học sâu với phân tích ảnh trong OpenCV. - Bài toán học sâu với dữ liệu y khoa. 	CELO4
	A. Nội dung làm việc trong PTN/Phòng thực hành: (6 giờ) - Thực hành theo hướng dẫn của giảng viên	
	B. Các nội dung cần tự học ở nhà: (3 giờ) + Xem trước các yêu cầu thực hành trên BL LAB và Dataiku DSS	

	C. Đánh giá kết quả học tập Phương pháp đánh giá: - Kỹ năng trình bày, kết quả thực hành	
--	---	--

Tuần	Nội dung	KQHTMD của HP
9 - 10	Hướng dẫn Đồ án môn học	CELO5
	A. Nội dung làm việc trong PTN/Phòng thực hành: (6 giờ) + Chọn chủ đề viết và trình bày báo cáo Đồ án môn học + Làm theo hướng dẫn của giảng viên	
	B. Các nội dung cần tự học ở nhà: (3 giờ) Xem trước các yêu cầu trên BL LAB và Dataiku DSS	
	C. Đánh giá kết quả học tập Phương pháp đánh giá: kỹ năng viết báo cáo và trình bày - Báo cáo Đồ án Môn học - Kỹ năng trình bày, giải quyết vấn đề	

9 Yêu cầu của giảng viên đối với học phần

- Phòng học: Phòng học lý thuyết có máy chiếu, bảng viết, micro
- Phương tiện phục vụ giảng dạy: + Phòng thực hành BL LAB kết nối Internet
+ Môi trường học tập Online Elearning, Dataiku

10. Biên soạn và cập nhật đề cương chi tiết

10.1. Đề cương được biên soạn vào năm học: Năm học 2021-2022

10.2. Đề cương được chỉnh sửa lần thứ: Biên soạn lần đầu

10.3. Nội dung được chỉnh sửa, cập nhật, thay đổi ở lần gần nhất: Tháng 08 / 2021

Tp. HCM, ngày 20 tháng 08 năm 2021

TRƯỞNG KHOA

TRƯỞNG BỘ MÔN

NGƯỜI BIÊN SOẠN

TS. Hoàng Lê Minh

TS. Trương Khắc Tùng

ThS. Nguyễn Thị Mỹ Linh

HIỆU TRƯỞNG

PGS. TS. Trần Thị Mỹ Diệu

PHỤ LỤC 3a: GIẢNG VIÊN PHỤ TRÁCH HỌC PHẦN

Giảng viên cơ hữu

Họ và tên: Hoàng Lê Minh	Học hàm, học vị: Tiến sĩ
Địa chỉ cơ quan: 45 Nguyễn Khắc Nhu, Phường Cô Giang, Quận 1, Tp. HCM	Điện thoại liên hệ: 0902703969
Email: minh.hl@vlu.edu.vn	Trang web: https://fit.vlu.edu.vn/1162001/
Cách liên lạc với giảng viên: Email hoặc gặp trực tiếp vào Thứ Hai và Thứ Sáu hàng tuần, lúc 14g00 - 15g30	

Họ và tên: Th.S Nguyễn Thị Mỹ Linh	Học hàm, học vị: Th. S
Địa chỉ cơ quan: 45 Nguyễn Khắc Nhu, Phường Cô Giang, Quận 1, Tp. HCM	Điện thoại liên hệ: 0909135500
Email: linh.ntm@vlu.edu.vn	Trang web: https://fit.vlu.edu.vn/
Cách liên lạc với giảng viên: Email hoặc gặp trực tiếp	

Giảng viên thỉnh giảng của môn học (nếu có)

Họ và tên: Nguyễn Thế Huỳnh	Học hàm, học vị: Tiến sĩ
Địa chỉ cơ quan: Deep Learning Institute Công ty NVIDIA	Điện thoại liên hệ:
Email:	Trang web:
Cách liên lạc với giảng viên:	

Họ và tên: Th.S Phan Hồ Việt Trường	Học hàm, học vị: Th. S
Địa chỉ cơ quan: 45 Nguyễn Khắc Nhu, Phường Cô Giang, Quận 1, Tp. HCM	Điện thoại liên hệ:
Email:	Trang web: https://fit.vlu.edu.vn/
Cách liên lạc với giảng viên: Email hoặc gặp trực tiếp	

PHỤ LỤC 3b: RUBRIC ĐÁNH GIÁ

Rubric 1 Đánh giá thực hành trong Phòng Thí nghiệm / Môi trường điện toán đám mây

Tiêu chí	Trọng số (%)	Tốt Từ 8 – 10 đ	Khá Từ 6 – dưới 8 đ	Trung bình Từ 4 – dưới 6 đ	Yếu dưới 4 đ
Nắm rõ mục tiêu và nội dung	5%	Trả lời đầy đủ rõ ràng và chính xác	Trả lời đúng nhưng còn sai sót nhỏ	Còn sai sót quan trọng	Không trả lời được
Hiểu mô hình	5%	Trả lời đầy đủ rõ ràng và chính xác	Trả lời đúng nhưng còn sai sót nhỏ	Còn sai sót quan trọng	Không trả lời được
Hiểu quy trình, phương pháp thực hiện	20%	Trả lời đầy đủ các vấn đề đặt ra, rõ ràng và chính xác	Trả lời đúng nội dung yêu cầu nhưng còn sai sót nhỏ	Trả lời đúng một số câu hỏi, còn sai sót quan trọng	Không trả lời được đa số câu hỏi
Chuẩn bị	10%	Tính toán số liệu cần, có bảng ghi kết quả thí nghiệm	Tính toán số liệu cần, không có bảng ghi kết quả thí nghiệm	Tính toán số liệu cần nhưng chưa đúng, thiếu bảng ghi kết quả TN	Chưa tính toán số liệu cần, không có bảng ghi kết quả thí nghiệm
Kỹ năng thực hành trong PTN	50%	Chủ động, thực hiện đúng yêu cầu, kỹ năng phân tích chính xác	Chủ động, thực hiện đúng yêu cầu, vẫn còn một số nội dung chưa thuần thực	Thực hiện đúng yêu cầu, nhưng chưa chủ động, chưa thuần thực	Thực hiện không đúng yêu cầu, vi phạm quy định PTN, làm sai
Sự phối hợp trong nhóm	10%	Phân công công việc trong nhóm và phối hợp tốt	Có phân công nhưng phối hợp chưa tốt	Phân công nhưng chưa hợp lý	Không có sự phân công trước khi thực hành
	100%				