



ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

---

## ĐỀ CƯƠNG MÔN HỌC

Môn: LẬP TRÌNH PYTHON CHO MÁY HỌC  
(Tiếng Anh: PYTHON FOR MACHINE LEARNING)

### **Chương trình đào tạo:**

Cử nhân Trí tuệ nhân tạo;

Cử nhân Khoa học máy tính.

*(Bản Cập nhật tháng 01 năm 2023)*

Người biên soạn: TS. Nguyễn Vinh Tiệp

**TP. HCM - 2023**

## 1. THÔNG TIN CHUNG

### 1.1 Thông tin về giảng viên:

- Họ và tên: Nguyễn Vinh Tiệp
- Chức danh, học vị: Tiến sĩ
- Email: tiepvn@uit.edu.vn

### 1.2 Thông tin về môn học

Tên môn học (tiếng Việt):	Python cho máy học
Tên môn học (tiếng Anh):	Python for machine learning
Đối tượng học tập	Những người muốn có kiến thức về trí tuệ nhân tạo
Mã môn học:	CS116
Thuộc khối kiến thức:	Chuyên ngành
Thuộc nhóm môn học	<input checked="" type="checkbox"/> Đại trà <input checked="" type="checkbox"/> Chất lượng cao <input checked="" type="checkbox"/> Tài năng
Khoa/Bộ môn phụ trách:	Bộ môn Trí tuệ Nhân tạo
Số tín chỉ:	4
Lý thuyết:	3
Thực hành:	1
Môn học tiên quyết:	
Môn học trước:	IT001 - Nhập môn lập trình; IT002 - Lập trình hướng đối tượng.

## 2. MÔ TẢ TÓM TẮT NỘI DUNG MÔN HỌC

Môn học trang bị cho người học kiến thức cơ bản về lập trình Python và một số thư viện nền tảng như numpy, sklearn, pandas, matplotlib. Đồng thời, người học được trang bị kiến thức cơ bản về các bước xây dựng mô hình máy học và kỹ năng sử dụng Python để cài đặt một số thuật toán cơ bản trong máy học như các thuật toán học có giám sát, học không giám sát, một số thuật toán tối ưu siêu tham số... phục vụ cho việc giải quyết một số bài toán thực tế.

## 3. MỤC TIÊU VÀ CHUẨN ĐẦU RA MÔN HỌC

### 3.1 Mục tiêu môn học:

**3.1.1. Mục tiêu chung:** Trang bị cho người học những kiến thức cơ bản về ngôn ngữ lập trình Python trong việc áp dụng vào các mô hình máy học phổ biến hiện nay nhằm giải quyết một số bài toán.

### 3.1.2. Mục tiêu cụ thể

Mục tiêu	Nội dung mục tiêu môn học
CG1	Làm việc ở mức độ cá nhân và cộng tác nhóm để trình bày và giải quyết một số thuật toán học không giám sát và có giám sát.
CG2	Hiểu và giải thích được các khái niệm, thuật ngữ liên quan tới các quy trình xây dựng mô hình máy học, một số phương pháp phân tích, tiền xử lý dữ liệu, một số mô hình máy học có giám sát, không giám sát, đánh giá mô hình.
CG3	Ứng dụng các lý thuyết, mô hình và thuật toán học có giám sát và không giám sát vào giải quyết các bài toán trong thực tế.
CG4	Có khả năng tự học một số vấn đề để vận dụng vào việc trả lời các câu hỏi trong quá trình học
CG5	Có thái độ, quan điểm và nhận thức đúng đắn về môn học.

### 3.2 Chuẩn đầu ra:

Sau khi hoàn thành môn học này người học có thể có được các khả năng sau:

CĐRMH	Mô tả CĐRMH	Ánh xạ CDR CTĐT	Cấp độ CĐRMH về NT, KN, TB
CLO1	Hiểu, giải thích được thuật ngữ tiếng Anh chuyên ngành của môn học liên quan đến ngôn ngữ lập trình Python, các quy trình xây dựng mô hình máy học, các thuật toán học không giám sát, có giám sát, các phương pháp tối ưu tham số mô hình máy học và ensemble learning.	LO2.2	NT3

CLO2	Biết được cách cài đặt bằng ngôn ngữ Python các bước phân tích, tiền xử lý dữ liệu, khởi tạo các mô hình máy học không giám sát, có giám sát, tối ưu siêu tham số và đánh giá mô hình.	LO3.3, LO4.2	KN3
CLO3	Rèn luyện kỹ năng làm việc nhóm và sử dụng ngoại ngữ (đọc hiểu, trình bày).	LO5.1, LO6.2	KN2
CLO4	Có thái độ, quan điểm và nhận thức đúng đắn về môn học.	LO8.2	TĐ2

### 3.3 Mối quan hệ giữa Mục tiêu và Chuẩn đầu ra:

Mối quan hệ giữa mục tiêu môn học và chuẩn đầu ra môn học được thể hiện trong ma trận sau:

Các thành phần	Mục tiêu	Chuẩn đầu ra
Kiến thức	CG2, CG1	CLO1, CLO2
Kỹ năng	CG1, CG1	CLO3, CLO4
Thái độ	CG5	CLO5

## 4. YÊU CẦU ĐỐI VỚI NGƯỜI HỌC

- Sinh viên cần tuân thủ nghiêm túc các nội quy và quy định của Khoa và Trường.
- Đối với bất kỳ sự gian lận nào trong quá trình làm bài tập hay bài thi, sinh viên phải chịu mọi hình thức kỷ luật của Khoa/Trường và bị 0 điểm cho môn học này.

## 5. NỘI DUNG MÔN HỌC, KẾ HOẠCH GIẢNG DẠY

- Thời lượng: Mỗi buổi học lý thuyết là 3 tiết.

### 5.1 Lý thuyết

Buổi học	Nội dung	CĐRMH	Hoạt động dạy và học	Thời điểm kiểm tra
----------	----------	-------	----------------------	--------------------

1	<b>Giới thiệu môn học và ôn tập ngôn ngữ lập trình Python</b> <ol style="list-style-type: none"> <li>1. Giới thiệu môn học</li> <li>2. Các kiểu dữ liệu và toán tử cơ bản. Kiểu dữ liệu nâng cao</li> <li>3. Các câu lệnh rẽ nhánh</li> <li>4. Các câu lệnh lặp</li> <li>5. Hàm và cách truyền tham số</li> <li>6. Lập trình hướng đối tượng</li> </ol>	CLO1, CLO2	Giảng viên: - Thuyết giảng - GV đặt câu hỏi thảo luận, tổng hợp, tổng kết. - Giới thiệu đồ án cuối kỳ.  Sinh viên: - Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng. - Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ	A1, A4
2	<b>Lập trình với các thư viện Numpy và Matplotlib để xử lý tính toán và trực quan hóa dữ liệu.</b> <ol style="list-style-type: none"> <li>1. Kiểu dữ liệu và toán tử trên các kiểu dữ liệu cơ bản của Numpy</li> <li>2. Một số phương thức xử lý vector, ma trận và tensor với Numpy</li> <li>3. Trực quan hóa dữ liệu với thư viện Matplotlib: biểu đồ điểm, cột, tròn,...</li> </ol>	CLO1, CLO2	Giảng viên: - Thuyết giảng - GV đặt câu hỏi thảo luận, tổng hợp, tổng kết. - Hướng dẫn làm đồ án  Sinh viên: - Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng. - Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ	A1, A4
3	<b>Các bước xây dựng mô hình máy học</b> <ol style="list-style-type: none"> <li>1. Phân tích và tiền xử lý dữ liệu</li> <li>2. Xây dựng mô hình máy học</li> <li>3. Đánh giá mô hình</li> <li>4. Tối ưu hóa siêu tham số</li> <li>5. Triển khai mô hình</li> </ol> <b>Phân tích dữ liệu đầu vào</b> <ol style="list-style-type: none"> <li>1. Phân tích đơn biến.</li> </ol> Phân tích đa biến	CLO1, CLO2, CLO3	Giảng viên: Thuyết giảng GV đặt câu hỏi thảo luận, tổng hợp, tổng kết. Hướng dẫn làm đồ án  Sinh viên: Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng. Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ	A1, A4

4	<p><b>Tiền xử lý dữ liệu</b></p> <ol style="list-style-type: none"> <li>1. Một số loại dữ liệu đầu vào</li> <li>2. Chuẩn hóa dữ liệu: min-max, phân bố chuẩn, max absolute</li> <li>3. Xử lý dữ liệu bị khuyết</li> <li>4. Xử lý dữ liệu dạng phân lớp (không phải dạng số)</li> </ol> <p><b>Đánh giá mô hình máy học</b></p> <ol style="list-style-type: none"> <li>1. Quy trình đánh giá mô hình khách quan</li> <li>2. Đánh giá mô hình hồi quy</li> <li>3. Đánh giá mô hình phân lớp</li> </ol>	CLO1, CLO2, CLO3, CLO4	<p>Giảng viên:</p> <ul style="list-style-type: none"> <li>- Thuyết giảng</li> <li>- GV đặt câu hỏi thảo luận, tổng hợp, tổng kết.</li> <li>- Hướng dẫn làm đồ án</li> </ul> <p>Sinh viên:</p> <ul style="list-style-type: none"> <li>- Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng.</li> </ul> <p>Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ</p>	A1, A4
5	<p><b>Học không giám sát: các thuật toán gom cụm</b></p> <ol style="list-style-type: none"> <li>1. Thuật toán KMean</li> <li>2. Thuật toán DBSCAN</li> <li>3. SpectralClustering</li> <li>4. Cài đặt ứng dụng thuật toán gom nhóm</li> </ol>	CLO1, CLO2, CLO3, CLO4	<p>Giảng viên:</p> <ul style="list-style-type: none"> <li>- Thuyết giảng</li> <li>- GV đặt câu hỏi thảo luận, tổng hợp, tổng kết.</li> <li>- Hướng dẫn làm đồ án</li> </ul> <p>Sinh viên:</p> <ul style="list-style-type: none"> <li>- Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng.</li> <li>- Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ</li> </ul>	A1, A4
6	<p><b>Học không giám sát: các thuật toán giảm chiều dữ liệu</b></p> <ol style="list-style-type: none"> <li>1. Thuật toán PCA</li> <li>2. Ứng dụng thuật toán PCA để nén dữ liệu hình ảnh và âm thanh</li> <li>3. Ứng dụng thuật toán PCA với một số mô hình học có giám sát</li> </ol>	CLO1, CLO2, CLO3, CLO4	<p>Giảng viên:</p> <ul style="list-style-type: none"> <li>- Thuyết giảng</li> <li>- GV đặt câu hỏi thảo luận, tổng hợp, tổng kết.</li> <li>- Hướng dẫn làm đồ án</li> </ul> <p>Sinh viên:</p> <ul style="list-style-type: none"> <li>- Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng.</li> <li>- Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ</li> </ul>	A1, A4

7	<b>Học không giám sát: các thuật toán giảm chiều dữ liệu</b> <ol style="list-style-type: none"> <li>Hồi quy tuyến tính – Linear Regression</li> <li>KNN Regression</li> <li>Random Forest Regressor</li> <li>Cài đặt các ứng dụng của mô hình hồi quy</li> </ol>	CLO2, CLO3, CLO4, CLO5	Giảng viên: - Thuyết giảng - GV đặt câu hỏi thảo luận, tổng hợp, tổng kết. - Hướng dẫn làm đồ án  Sinh viên: - Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng. Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ	A1, A4
8	<b>Học có giám sát: các mô hình phân lớp</b> <ol style="list-style-type: none"> <li>Hồi quy luận lý – Logistic Regression</li> <li>KNN Classifier</li> <li>Random Forest Classifier</li> <li>Cài đặt các ứng dụng của mô hình phân lớp</li> </ol>	CLO1, CLO2, CLO3, CLO4	Giảng viên: - Thuyết giảng - GV đặt câu hỏi thảo luận, tổng hợp, tổng kết. - Hướng dẫn làm đồ án  Sinh viên: - Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng. Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ	A1, A4
9-10	<b>Học sâu: Mạng Convolutional Neural Network</b> <ol style="list-style-type: none"> <li>Kiến trúc mạng CNN</li> <li>Thay đổi thiết kế mạng CNN</li> <li>Trực quan hóa mạng CNN</li> </ol>	CLO1, CLO2, CLO3, CLO4	Giảng viên: - Thuyết giảng - GV đặt câu hỏi thảo luận, tổng hợp, tổng kết. - Hướng dẫn làm đồ án  Sinh viên: - Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng. Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ	A1, A4
11-12	<b>Siêu tham số và tối ưu siêu tham số</b> <ol style="list-style-type: none"> <li>Siêu tham số là gì</li> <li>Một số phương pháp tối ưu siêu tham số</li> <li>Cài đặt phương pháp Grid Search</li> </ol>	CLO1, CLO2, CLO3, CLO4	Giảng viên: - Thuyết giảng - GV đặt câu hỏi thảo luận, tổng hợp, tổng kết. - Hướng dẫn làm đồ án  Sinh viên:	A1, A4

			- Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng. Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ	
13	<b>Ensemble Model: Học tổng hợp</b> 1. Phương pháp Bagging và Boosting 2. Phương pháp XGBoost 3. Ứng dụng của phương pháp XGBoost trong các mô hình dự đoán	CLO1, CLO2, CLO3, CLO4	Giảng viên: - Thuyết giảng - GV đặt câu hỏi thảo luận, tổng hợp, tổng kết. - Hướng dẫn làm đồ án  Sinh viên: - Thảo luận nhóm, trình bày, thảo luận trước lớp, nghe giảng. Làm đồ án/bài tập lớn theo nhóm và báo cáo tiến độ	A1, A4
14-15	<b>Báo cáo đồ án cuối kỳ</b>  <b>Ôn tập</b>	CLO1, CLO2, CLO3, CLO4	Sinh viên: Trình bày seminar đồ án cuối kỳ.  Giảng viên: Nhận xét đánh giá các đồ án môn học.	A4

## 5.2 Thực hành

Buổi học	Nội dung	CĐRMH	Hoạt động dạy và học	Thời điểm kiểm tra
1	Lập trình cơ bản và nâng cao trên Python và các thư viện: Numpy, Matplotlib, Pandas.		Giảng viên: Hướng dẫn thực hành. Sinh viên: Thực hành cài đặt thuật toán.	A1
2	Cài đặt các mô hình không giám sát.		Giảng viên: Hướng dẫn thực hành. Sinh viên: Thực hành cài đặt thuật toán.	A1
3	Cài đặt các mô hình có giám sát cho bài toán hồi quy.		Giảng viên: Hướng dẫn thực hành.	A1



			Sinh viên: Thực hành cài đặt thuật toán.	
4	Cài đặt các mô hình có giám sát cho bài toán phân lớp.		Giảng viên: Hướng dẫn thực hành. Sinh viên: Thực hành cài đặt thuật toán.	A1
5	Cài đặt các phương pháp tối ưu mô hình và ensemble learning.		Giảng viên: Hướng dẫn thực hành. Sinh viên: Thực hành cài đặt thuật toán.	A1
6	Hướng dẫn SV thực hiện đồ án môn học.		Giảng viên: Hướng dẫn thực hiện đồ án môn học. Sinh viên: Đặt câu hỏi và trình bày kết quả.	A4

## 6. PHƯƠNG PHÁP VÀ PHƯƠNG TIỆN DẠY HỌC

- **Phương pháp giảng dạy chủ yếu:** Giảng viên đặt vấn đề, giải thích đề bài, trao đổi và kiểm tra giải pháp. Sinh viên nghe giảng, suy nghĩ và thảo luận nhằm thiết kế thuật toán cho các vấn đề
- **Học liệu:** các nguồn tài liệu được sử dụng trên lớp, các slides và tài liệu khác do giảng viên cung cấp.

## 7. KIỂM TRA, ĐÁNH GIÁ KẾT QUẢ HỌC TẬP

### 7.1 Khung thông tin về kiểm tra đánh giá kết quả học tập

	Thời điểm kiểm tra	Hình thức KTĐG	Công cụ TĐG	Trọng số	Thang điểm	Tiêu chí đánh giá
A1	Quá trình	Kiểm tra trên lớp	Bài tập lập trình trên lớp hoặc về nhà	30%	10	
A2	Giữa kỳ			0		
A3	Thực hành	Bài thực hành, bài tập đồ án	Báo cáo đồ án	30%	10	
A4	Cuối kỳ			40%	10	

	Bài thi viết	Làm bài viết trên giấy thi.	Câu hỏi tự luận	40%	10	Nội dung trong bài làm đảm bảo tính chính xác, đầy đủ và phù hợp với yêu cầu của nội dung câu hỏi.
--	--------------	-----------------------------	-----------------	-----	----	--

## 8. TÀI LIỆU HỌC TẬP, THAM KHẢO

- Python Data Science Handbook, Jake VanderPlas.
- <https://jakevdp.github.io/PythonDataScienceHandbook>

## 9. PHẦN MỀM HAY CÔNG CỤ HỖ TRỢ THỰC HÀNH

- Ngôn ngữ lập trình: Python
- Thư viện Anaconda
- Framework: Tensorflow / PyTorch
- Trình soạn thảo mã nguồn: Visual Code / Sublime / PyCharm

## 10. Ngày phê duyệt:

## 11. Cấp phê duyệt: Bộ môn Tin toán Đa phương tiện

*Tp.HCM, ngày 12 tháng 01 năm 2023*

**Trưởng Bộ môn**

(Ký và ghi rõ họ tên)

**Giảng viên biên soạn**

(Ký và ghi rõ họ tên)

Ngô Đức Thành

Nguyễn Vinh Tiệp