



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

**ĐỀ CƯƠNG MÔN HỌC**  
**DS201 – Deep Learning trong Khoa học Dữ liệu**

**1. THÔNG TIN CHUNG (General information)**

Tên môn học (tiếng Việt):	<b>Deep Learning trong Khoa học Dữ liệu</b>
Tên môn học (tiếng Anh):	Deep Learning in Data Science
Mã môn học:	<b>DS201</b>
Thuộc khối kiến thức:	Đại cương <input type="checkbox"/> ; Cơ sở nhóm ngành <input type="checkbox"/> ; Cơ sở ngành <input type="checkbox"/> ; Chuyên ngành <input checked="" type="checkbox"/> ; Tốt nghiệp <input type="checkbox"/>
Khoa, Bộ môn phụ trách:	Khoa Khoa học và Kỹ thuật Thông tin, Bộ môn Khoa học Dữ Liệu
Giảng viên biên soạn:	TS. Đỗ Trọng Hợp, CN. Lưu Thanh Sơn Email: <a href="mailto:hopdt@uit.edu.vn">hopdt@uit.edu.vn</a> , <a href="mailto:sonlt@uit.edu.vn">sonlt@uit.edu.vn</a>
Số tín chỉ:	
Lý thuyết:	3 .....
Thực hành:	1 .....
Tự học:	8 .....
Môn học tiên quyết:	Không .....
Môn học trước:	Học máy thống kê (DS102) .....

**2. MÔ TẢ MÔN HỌC (Course description)**

Môn học cung cấp nền tảng về Deep Learning. Sinh viên sẽ được học các xây dựng một Neural Network cơ bản và cách nâng cao hiệu suất của Neural Network thông qua những phương pháp như tinh chỉnh siêu tham số, regularization, chuẩn hóa, tối ưu hóa. Sau đó, sinh viên sẽ được học những cấu trúc mạng như Convolutional Neural Network (CNN), Sequence Model và làm quen với các công cụ như Tensorflow và Keras. Cuối cùng sinh viên có thể sử dụng Deep Learning cho những bài toán cụ thể về xử lý ảnh, xử lý ngôn ngữ tự nhiên, và phân tích dữ liệu.

**3. MỤC TIÊU MÔN HỌC (Course goals)**

Sau khi hoàn thành môn học này, sinh viên có thể:

Ký hiệu	Mục tiêu môn học	Chuẩn đầu ra trong CTĐT
<i>G1</i>	Hiểu cách xây dựng một Neural Network, hiểu những vấn đề của Neural Network như underfitting, overfitting, vanishing/exploding gradient, học chậm,..., và biết những phương pháp để khắc phục những vấn đề này.	<i>LO 3, LO 5</i>
<i>G2</i>	Hiểu những cấu trúc và những kỹ thuật cao cấp như Convolutional Neural Network, Residual Network, Inception Network, Style Transfer Learning, Recurrent Neural Networks, Gated Recurrent Unit, Long Short Term Memory và các loại ứng dụng của chúng.	<i>LO 3, LO 5</i>
<i>G3</i>	Có thể vận dụng Deep Learning để giải quyết các bài toán về phân tích dữ liệu, xử lý ảnh, xử lý ngôn ngữ tự nhiên, xử lý giọng nói. Biết cách sử dụng các công cụ như Tensorflow và Keras để cài đặt Deep Learning.	<i>LO 5, LO 10</i>

#### 4. CHUẨN ĐẦU RA MÔN HỌC (Course learning outcomes)

CĐRMH	Mô tả CĐRMH (Mục tiêu cụ thể)	Mức độ giảng dạy
<i>G1.1</i>	Hiểu được cấu trúc mạng Neural Network, những khó khăn khi áp dụng Neural Network, và các phương pháp để khắc phục.	<i>TU</i>
<i>G1.2</i>	Áp dụng các kỹ thuật đã học để hiện thực một mạng neural trên một bộ dữ liệu cụ thể.	<i>IT</i>
<i>G2.1</i>	Hiểu những cấu trúc Neural Network cao cấp như Confolutional Neural Network (CNN), Residual Network, Inception Network, Recurrent Neural Network (RNN), Gated Recurrent Unit, Long Short Term Memory, và những biến thể khác của chúng.	<i>TU</i>
<i>G2.2</i>	Áp dụng các kỹ thuật và kiến thức đã học để xây dựng mạng neural CNN và RNN trên bộ dữ liệu cụ thể	<i>IT</i>
<i>G3.1</i>	Giới thiệu các bài toán áp dụng Deep Learning: bài toán phân loại, phát hiện vật thể, xử lý ngôn ngữ tự nhiên.	IT
<i>G3.2</i>	Hiện thực các bài toán trên bằng cách sử dụng Python và các công cụ như Tensorflow và Keras để cài đặt Deep Learning trên bộ dữ liệu cụ thể	<i>IT</i>

## 5. NỘI DUNG MÔN HỌC, KẾ HOẠCH GIẢNG DẠY (Course content, lesson plan)

### a. Lý thuyết

Buổi học (45 tiết)	Nội dung	CĐRMH	Hoạt động dạy và học	Thành phần đánh giá
Buổi 1, 2 (6 tiết)	<b>Chương 1: Giới thiệu về Neural Network</b> <ul style="list-style-type: none"> <li>- Giới thiệu về học có giám sát và Neural Network</li> <li>- Giới thiệu Logistic Regression</li> <li>- Giới thiệu Gradient descent</li> <li>- Mô hình Neural Network với 1 lớp ẩn</li> <li>- Hàm activation và gradient descent cho Neural Network 1 lớp ẩn</li> <li>- Cách khởi tạo tham số</li> <li>- Mô Hình Neural Network nhiều lớp</li> <li>- Forward propagation và backward propagation cho Neural Network nhiều lớp</li> </ul>	<i>G1.1</i>	<b>Dạy:</b> Dạy các nội dung như mô tả. Trả lời các câu hỏi về bài tập về nhà của sinh viên. <b>Học ở lớp:</b> Sinh viên nghe giảng, trả lời các câu hỏi của giảng viên, làm bài trắc nghiệm nhỏ đầu giờ. <b>Học ở nhà:</b> Làm các bài tập giao về nhà.	<i>A1, A3</i>
Buổi 3, 4, 5 (9 tiết)	<b>Chương 2: Nâng cao hiệu suất Neural Network</b> <ul style="list-style-type: none"> <li>- Hiện tượng Underfitting và Overfitting</li> <li>- Các phương pháp Regularization</li> <li>- Chuẩn hóa dữ liệu đầu vào</li> <li>- Hiện tượng vanishing/exploding gradient và cách khởi tạo tham số</li> <li>- Mini-batch gradient descent</li> <li>- Các thuật toán tối ưu như Momentum, RMSprop, Adam.</li> <li>- Quy trình điều chỉnh các siêu tham số</li> <li>- Batch normalization</li> <li>- Softmax regression</li> <li>- Giới thiệu Tensorflow</li> </ul>	<i>G1.1, G1.2</i>	<b>Dạy:</b> Dạy các nội dung như mô tả. Trả lời các câu hỏi về bài tập về nhà của sinh viên. <b>Học ở lớp:</b> Sinh viên nghe giảng, trả lời các câu hỏi của giảng viên, làm bài trắc nghiệm nhỏ đầu giờ. <b>Học ở nhà:</b> Làm các bài tập giao về nhà.	<i>A1, A3</i>
Buổi 6, 7, 8, 9 (12 tiết)	<b>Chương 3: Convolutional Neural Network</b> <ul style="list-style-type: none"> <li>- Cơ bản về Convolutional Neural Network</li> <li>- Trích xuất các đặc trưng trong ảnh sử dụng tích chập</li> <li>- Sử dụng padding</li> <li>- Strided convolution</li> <li>- Convolution over volumes</li> <li>- Cấu tạo của 1 lớp CNN</li> </ul>	<i>G2.1 G2.2 G3.1</i>	<b>Dạy:</b> Dạy các nội dung như mô tả. Trả lời các câu hỏi về bài tập về nhà của sinh viên. <b>Học ở lớp:</b> Sinh viên nghe giảng, trả lời các câu hỏi của giảng viên, làm bài trắc nghiệm nhỏ đầu	<i>A1, A3</i>

	<ul style="list-style-type: none"> <li>- Pooling layer</li> <li>- Mô hình chung của CNN</li> <li>- Các cấu trúc CNN điển hình (LeNet, AlexNet, VGG16)</li> <li>- Residual Network</li> <li>- 1x1 convolution</li> <li>- Inception network</li> <li>- Transfer learning</li> <li>- Giới thiệu Keras</li> <li>- Giới thiệu object classification, object localization, và object detection</li> <li>- Cài đặt sliding window với tích chập</li> <li>- Bouding box, Intersection Over Union, và anchor box</li> <li>- Thuật toán YOLO và R-CNN</li> <li>- Nhật dạng khuôn mặt</li> <li>- One-shot Learning</li> <li>- Siamese Network</li> <li>- Mô hình Deep Learning dùng để xác thực khuôn mặt</li> <li>- Hiểu về cách học và thông tin diễn đạt trong các các node và các lớp của CNN</li> <li>- Cách tạo ra bức ảnh nghệ thuật sử dụng Neural Style Transfer</li> </ul>		<p>giờ.</p> <p><b>Học ở nhà:</b> Làm các bài tập giao về nhà.</p>	
<p>Buổi 10, 11, 12, 13 (12 tiết)</p>	<p><b>Chương 4: Sequence model.</b></p> <ul style="list-style-type: none"> <li>- Sequence model và các ký hiệu</li> <li>- Mô hình Recurrent Neural Network</li> <li>- Backpropagation trong RNN</li> <li>- Các loại RNN</li> <li>- Language model và sequence generation</li> <li>- Hiện tượng vanishing gradient của RNN</li> <li>- Gated Recurrent Unit (GRU)</li> <li>- Long short term memory (LSTM)</li> <li>- Bidirectional RNN</li> <li>- Deep RNN</li> <li>- Word representation</li> <li>- Word embedding</li> <li>- Learning word embedding</li> <li>- Negative sampling</li> <li>- Bài toán Sentiment Classification</li> <li>- Vấn đề chọn câu trong thông dịch</li> <li>- Thuật toán Beam Search</li> <li>- Attention model và Neural</li> </ul>	<p><i>G2.1</i> <i>G2.2</i> <i>G3.1</i></p>	<p><b>Dạy:</b> Dạy các nội dung như mô tả. Trả lời các câu hỏi về bài tập về nhà của sinh viên.</p> <p><b>Học ở lớp:</b> Sinh viên nghe giảng, trả lời các câu hỏi của giảng viên, làm bài trắc nghiệm nhỏ đầu giờ.</p> <p><b>Học ở nhà:</b> Làm các bài tập giao về nhà.</p>	<p><i>A1, A3</i></p>

	machine translation - Speech recognition - Bài toán Trigger Word Detection			
Buổi 14 (3 tiết)	<b>Chương 5: Deep learning trong khoa học dữ liệu</b> - Recommendation system sử dụng Deep Learning - Multilayer Perceptron based Recommendation - Autoencoder based recommendation - Recommendation sử dụng CNN và RNN - Giới thiệu ứng dụng Deep Learning trong xử lý dữ liệu lớn - Semantic indexing - Discriminative tasks và semantic tagging - Những thử thách của Deep Learning trong khai phá dữ liệu và xử lý dữ liệu lớn	<i>G3.1</i> <i>G3.2</i>	<b>Dạy:</b> Dạy các nội dung như mô tả. Trả lời các câu hỏi về bài tập về nhà của sinh viên. <b>Học ở lớp:</b> Sinh viên nghe giảng, trả lời các câu hỏi của giảng viên, làm bài trắc nghiệm nhỏ đầu giờ. <b>Học ở nhà:</b> Làm các bài tập giao về nhà.	<i>A1</i>
Buổi 15 (3 tiết)	<b>Seminar đồ án môn học</b>	<i>G1.1,</i> <i>G1.2,</i> <i>G2.1,</i> <i>G2.2</i> <i>G3.1</i> <i>G3.2</i>	<b>Học ở lớp:</b> Sinh viên thuyết trình đồ án của mình trước lớp.	<i>A1, A3</i>

### b. Thực hành

Buổi học (30 tiết)	Nội dung	CĐRMH	Hoạt động dạy và học	Thành phần đánh giá
Buổi 1 (5 tiết)	<b>Thực hành 1:</b> Mạng neural căn bản - Giới thiệu Google colab và Jupyter notebook. - Cài đặt mạng neural căn bản. - Ứng dụng cho bài toán phân lớp ảnh chó mèo.	<i>G1.2</i> <i>G2.2</i>	<b>Dạy:</b> Dạy cách cài đặt các thành phần để xây dựng một lớp Deep Neural Network với số hidden layer tùy chọn sử dụng python. <b>Học ở lớp:</b> Sinh viên thực hành theo hướng dẫn của giảng viên.	<i>A3</i>
Buổi 2 (5 tiết)	<b>Thực hành 2:</b> Tối ưu và cải thiện mạng neural - Initialization - Regularization	<i>G1.2</i> <i>G2.2</i>	<b>Dạy:</b> Dạy cách cài đặt các kỹ thuật để cải thiện mạng Neural Network	A3

	- Optimization		như Initialization, Regularization, Optimization sử dụng python <b>Học ở lớp:</b> Sinh viên thực hành theo hướng dẫn của giảng viên.	
Buổi 3 (5 tiết)	<b>Thực hành 3:</b> Mạng neural tích chập - Mạng tích chập cơ bản. - Ứng dụng mạng tích chập: hand sign detection - Neural Style transfers.	<i>G1.2</i> <i>G2.2</i>	<b>Dạy:</b> Dạy cách cài đặt mạng CNN cơ bản. <b>Học ở lớp:</b> Sinh viên thực hành theo hướng dẫn của giảng viên.	<i>A3</i>
Buổi 4 (5 tiết)	<b>Thực hành 4:</b> Ứng dụng mạng neural tích chập - Các dạng mạng tích chập: AlexNET, VGG. - Bài toán phân lớp ảnh (MNIST, Jewelry classification). - Vehicle detection by Yolo (optional)	<i>G1.2</i> <i>G2.2</i> <i>G3.2</i>	<b>Dạy:</b> Dạy cách sử dụng Keras để cài đặt một Neural Network <b>Học ở lớp:</b> Sinh viên thực hành theo hướng dẫn của giảng viên.	<i>A3</i>
Buổi 5 (5 tiết)	<b>Thực hành 5:</b> Mạng neural hồi quy - Mạng hồi quy cơ bản. - Ứng dụng: Emoji detection	<i>G1.2</i> <i>G2.2</i>	<b>Dạy:</b> Dạy cách cài đặt một Recurrent Neural Network cho bài toán xử lý ngôn ngữ tự nhiên <b>Học ở lớp:</b> Sinh viên thực hành theo hướng dẫn của giảng viên.	<i>A3</i>
Buổi 6 (5 tiết)	<b>Thực hành 6:</b> Ứng dụng của mạng neural hồi quy. - Bài toán phân lớp văn bản: Fake new detection. - Bài toán Named entities recognition: CoNLL 2003 NER dataset	<i>G1.2</i> <i>G2.2</i> <i>G3.2</i>	<b>Dạy:</b> Dạy cách cài đặt một Recurrent Neural Network cho bài toán xử lý ngôn ngữ tự nhiên <b>Học ở lớp:</b> Sinh viên thực hành theo hướng dẫn của giảng viên.	<i>A3</i>

## 6. ĐÁNH GIÁ MÔN HỌC (Course assessment)

Thành phần đánh giá	CĐRMH	Tỷ lệ (%)
---------------------	-------	-----------

A1. Quá trình (Kiểm tra trên lớp, bài tập, ...)	<i>G1.1, G1.2</i>	<i>20%</i>
A2. Giữa kỳ		<i>0%</i>
A3. Thực hành	<i>G1.2, G2.2, G3.2</i>	<i>30%</i>
A4. Đồ án	<i>G1.1, G1.2, G2.1, G2.2, G3.1, G3.2</i>	<i>50%</i>

**a. Rubric của thành phần đánh giá A1**

<b><i>Đánh giá quá trình học tập tại lớp</i></b>	<b><i>Giỏi (9-10đ)</i></b>	<b><i>Khá (7-8đ)</i></b>	<b><i>TB (6-7đ)</i></b>	<b><i>Yếu (4-5đ)</i></b>	<b><i>Kém (&lt;3đ)</i></b>
<i>Mức độ chuyên cần</i>	<i>Tích cực tham gia đóng góp và xây dựng các bài giảng. Chủ động trả lời câu hỏi và làm 100 %bài tập về nhà.</i>	<i>Tham gia trả lời câu hỏi tại lớp và hoàn thành 100% bài tập về nhà.</i>	<i>Hoàn thành 70% các bài tập về nhà.</i>	<i>Hoàn thành 40% các bài tập về nhà.</i>	<i>Không tham gia vào bất cứ hoạt động nào tại lớp.</i>
<b><i>Các bài trắc nghiệm kiến thức trên lớp</i></b>	<b><i>Giỏi (9-10đ)</i></b>	<b><i>Khá (7-8đ)</i></b>	<b><i>TB (6-7đ)</i></b>	<b><i>Yếu (4-5đ)</i></b>	<b><i>Kém (&lt;3đ)</i></b>
<i>Điểm trung bình các bài trắc nghiệm</i>	<i>Trung bình đúng 90-100%</i>	<i>Trung bình đúng 70-80%</i>	<i>Trung bình đúng 60-70%</i>	<i>Trung bình đúng 40-50%</i>	<i>Trung bình đúng dưới 30%</i>

**b. Rubric của thành phần đánh giá A2**

Không có

**c. Rubric của thành phần đánh giá A3**

	<b><i>Giỏi (9-10đ)</i></b>	<b><i>Khá (7-8đ)</i></b>	<b><i>TB (6-7đ)</i></b>	<b><i>Yếu (4-5đ)</i></b>	<b><i>Kém (&lt;3đ)</i></b>
<i>Làm bài tập thực hành hằng tuần</i>	<i>Làm 6 bài tập</i>	<i>Làm 5 bài tập</i>	<i>Làm 4 bài tập</i>	<i>Làm 3 bài tập</i>	<i>Làm dưới 2 bài tập</i>
<i>Báo cáo thực hành</i>	<i>Nội dung đầy đủ, trình bày rõ ràng</i>	<i>Đáp ứng được 80% nội dung yêu cầu</i>	<i>Đáp ứng được 50% nội dung yêu cầu</i>	<i>Đáp ứng được 30% nội dung yêu cầu</i>	<i>Không đáp ứng hoặc sai nội dung yêu cầu</i>

**d. Rubric của thành phần đánh giá A4**

<b>Đồ án cuối kì</b>	<b>Giỏi (9-10đ)</b>	<b>Khá (7-8đ)</b>	<b>TB (6-7đ)</b>	<b>Yếu (4-5đ)</b>	<b>Kém (&lt;3đ)</b>
<i>Nội dung</i>	<i>Thỏa các tiêu chí sau:</i> - Có tính mới - Có sự cải tiến so với hệ thống có sẵn - Bài toán có độ khó - Kết quả đáp ứng yêu cầu của nội dung - Có sự so sánh giữa các phương pháp/thuật toán - Trình bày rõ ràng	- Bài toán có độ khó tương đối - Kết quả đáp ứng yêu cầu của nội dung. - Có sự cải tiến tương đối so với hệ thống đã có	Đáp ứng được 50 đến 70% yêu cầu của nội dung	Đáp ứng 1/4 yêu cầu nội dung	Sai nội dung hoặc không đáp ứng nội dung theo yêu cầu

## 7. QUY ĐỊNH CỦA MÔN HỌC (Course requirements and expectations)

- Sinh viên đến lớp học theo quy định chung của nhà trường.
- Sinh viên làm bài tập và thực hiện đồ án môn học đầy đủ.

## 8. TÀI LIỆU HỌC TẬP, THAM KHẢO

### Tài liệu tham khảo

1. *Deep Learning Specialization* by Andrew Ng, Coursera.
2. Nielsen, Michael A. Neural networks and deep learning. Vol. 2018. San Francisco, CA: Determination press, 2015.

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

1. Python (<https://www.python.org/>)
2. Keras (<https://keras.io/>)
3. Google colab (<https://colab.research.google.com/>)

Tp.HCM, ngày 24 tháng 08 năm 2020

**Trưởng khoa/bộ môn**

(Ký và ghi rõ họ tên)

**Giảng viên biên soạn**

(Ký và ghi rõ họ tên)

Đỗ Trọng Hợp