

ĐỀ CƯƠNG MÔN HỌC

DỮ LIỆU LỚN

(Big Data)

Mã số MH : XXXXXXXXXX

- Số tín chỉ	:	Tc (LT.BT&TH.TựHọc): 3	TCHP:	
- Số tiết	:	Tổng: 60	LT: 30	BT: TH: ĐA: BTL/TL: 30
- Đánh giá	:	Kiểm tra giữa kỳ	20%	90 phút
Thang điểm		Bài tập lớn	30%	Bài tập lớn theo nhóm
10/10		Thi cuối kỳ	50%	Thi viết 120 phút
- Môn tiên quyết	:	-		MS:
- Môn học trước	:	-		MS:
- Môn song hành	:	-		MS:
- CTĐT ngành		Khoa học máy tính		
Mã ngành	:	60.48.01		
- Ghi chú khác	:	Học viên không được thiếu bất kỳ cột điểm nào mới được tính điểm sau cùng của môn học.		

❖ Mục tiêu của môn học:

Cung cấp cho học viên kiến thức về xử lý và ứng dụng của Big Data (dữ liệu lớn), là những tập hợp dữ liệu phát triển nhanh chóng và rộng khắp dưới nhiều hình thức khác nhau, làm cho chúng vượt quá khả năng xử lý của những hệ thống cơ sở dữ liệu truyền thống.

❖ Nội dung tóm tắt môn học:

Big data và các ứng dụng liên quan từ góc nhìn kinh doanh và kỹ thuật. Các đặc trưng cơ bản về big data: dung lượng lớn, vận tốc lớn và tính đa dạng. Giới thiệu các giải pháp công nghệ trong việc quản lý và phát triển ứng dụng với big data. Khai phá dữ liệu với big data. Bảo mật và bảo vệ tính riêng tư với big data. Những hướng nghiên cứu và ứng dụng mới có liên quan.

❖ Các hiểu biết, các kỹ năng cần đạt được sau khi học môn học

Học viên hiểu về vai trò của big data và các ứng dụng có liên quan đối với các tổ chức. Có khả năng tìm hiểu để vận dụng công nghệ cho giải pháp quản lý và phát triển ứng dụng với big data. Có khả năng tự học và nghiên cứu các chủ đề mới có liên quan đến big data.

❖ Tài liệu tham khảo chính:

[1] P.C. Zikopoulos, C. Eaton, D. DeRoos, T. Deutsch, G. Lapis: Understanding Big Data. Mc Graw Hill, 2012

- [2] K. Davis, D. Patterson: Ethics of Big Data. O'Reilly, 2012.
 [3] T. Craig, M. E. Ludloff: Privacy and Big Data, O'Reilly, 2011.

❖ Hướng dẫn cách học - chi tiết cách đánh giá môn học:

- Tham dự giờ giảng trên lớp + làm bài tập đầy đủ
- Về kiểm tra giữa kỳ (tuần kiểm tra, thời lượng....): không có kiểm tra giữa kỳ
- Về thực hiện báo cáo tiểu luận/BT Lớn, phần thực hành: nhóm 2-4 người, đánh giá: 30%
- Thuyết trình trên lớp: nhóm 2-4 người, đánh giá: 20%
- Cách tổ chức thi cuối kỳ: 50%, viết và trắc nghiệm
- Ghi chú về điều kiện cấm thi, cách tổng kết điểm: sinh viên không được thiếu bất cứ cột điểm nào của môn học mới được tính điểm sau cùng của môn

❖ Nội dung chi tiết:

PHẦN GIẢNG DẠY TRÊN LỚP: (30 tiết)

Tuần	Nội dung	Tài liệu	Ghi chú
1	Giới thiệu dữ liệu lớn và các ứng dụng 1.1 Dữ liệu lớn là gì? 1.2 Tại sao dữ liệu lớn lại quan trọng? 1.3 Những ứng dụng của dữ liệu lớn	[1]	
2,3	Những đặc trưng cơ bản của dữ liệu lớn 2.1 Giới thiệu ba đặc trưng cơ bản (V3) 2.2 Dữ liệu có dung lượng lớn 2.3 Dữ liệu có vận tốc lớn 2.4 Dữ liệu có tính đa dạng cao 2.5 Những giải pháp kỹ thuật ứng với các đặc trưng của dữ liệu lớn	[1,2,3]	
4,5	Những giải pháp kỹ thuật để quản lý và phát triển ứng dụng với dữ liệu lớn 3.1 Kho dữ liệu lớn 3.2 Dòng dữ liệu lớn 3.3 Case studies	[1, internet]	
6,7	Phát triển ứng dụng dữ liệu lớn với Hadoop 4.1 Các thành phần của Hadoop 4.2 Đưa dữ liệu lớn vào Hadoop 4.3 Phát triển ứng dụng với Hadoop 4.4 Case studies	[1]	
8,9	Bảo mật và bảo vệ tính riêng tư với dữ liệu lớn 5.1 Bảo mật và tính riêng tư dữ liệu 5.2 Tính riêng tư người dùng 5.3 Bảo vệ tính riêng tư trong các ứng dụng với dữ liệu lớn 5.4 Bảo vệ tính riêng tư trong khai phá dữ liệu lớn 5.5 Những nguyên tắc sử dụng dữ liệu lớn	[2,3]	
10	Những hướng nghiên cứu và ứng dụng mới với dữ liệu lớn Ôn tập	[1,2,3, internet]	

**PHẦN BÀI TẬP, TIỂU LUẬN NGOẠI KHÓA, HV ĐI NGHIÊN CỨU THỰC TIỄN
NGOÀI TRƯỜNG: (30 tiết)**

TT	Nội dung	Số tiết	Địa điểm	TLT K
1	Hadoop và những công nghệ liên quan cho việc quản lý dữ liệu lớn	10		[1,2,3, internet]
2	Bảo mật và bảo vệ tính riêng tư cho dữ liệu lớn trong Hadoop	5		[1,2,3, internet]
3	Một ứng dụng minh họa cho việc quản lý dữ liệu lớn	15		[1,2,3, internet]

Tp.Hồ Chí Minh, ngày tháng năm 2013

**KT.TRƯỞNG KHOA
PHÓ TRƯỞNG KHOA**

CB PHỤ TRÁCH LẬP ĐỀ CƯƠNG

TS. PHẠM TRẦN VŨ

PGS.TS ĐẶNG TRẦN KHÁNH