



HCMUTE

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT THÀNH PHỐ HỒ CHÍ MINH

INFO UTE LIBRARY

ĐỀ CƯƠNG CHI TIẾT MÔN HỌC

PHÂN TÍCH DỮ LIỆU LỚN

BIG DATA ANALYSIS

(Ngành Kỹ thuật Dữ liệu - Chương trình đào tạo đại học 132TC)



BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
THÀNH PHỐ HỒ CHÍ MINH

-----**-----

INFO UTE LIBRARY



HCMUTE

**ĐỀ CƯƠNG CHI TIẾT MÔN HỌC
PHÂN TÍCH DỮ LIỆU LỚN
BIG DATA ANALYSIS**

(Ngành Kỹ thuật Dữ liệu - Chương trình đào tạo đại học 132TC)

HCMUTE.EDU.VN – THÁNG 9 NĂM 2019

MỤC LỤC

LỜI NÓI ĐẦU	4
GIỚI THIỆU CÁC DỊCH VỤ UTE Library	5
GIỚI THIỆU CÁC CƠ SỞ DỮ LIỆU TRỰC TUYẾN	9
ĐỀ CƯƠNG CHI TIẾT MÔN HỌC	10
GIỚI THIỆU CÁC CHƯƠNG TRÌNH ĐÀO TẠO	16

LỜI NÓI ĐẦU

Công nghệ thông tin (IT) là một trong những động lực quan trọng bậc nhất của sự phát triển, cùng với một số ngành công nghệ cao khác đang làm biến đổi sâu sắc đời sống kinh tế, văn hóa, xã hội của thế giới hiện đại. Việc ứng dụng và phát triển CNTT ở nước ta đã góp phần giải phóng sức mạnh vật chất, trí tuệ và tinh thần, thúc đẩy công cuộc đổi mới, phát triển nhanh và hiện đại hóa các ngành kinh tế, tăng cường năng lực cạnh tranh của các doanh nghiệp, hỗ trợ có hiệu quả cho quá trình chủ động hội nhập kinh tế quốc tế, nâng cao chất lượng cuộc sống của nhân dân, đảm bảo an ninh quốc phòng góp phần thực hiện thắng lợi sự nghiệp công nghiệp hóa, hiện đại hóa đất nước.

Hiện nay, ngành CNTT là một trong những ngành học được chú trọng trong hệ thống đào tạo của các trường Đại học trong đó có trường Đại học Sư phạm Kỹ thuật Tp. Hồ Chí Minh. Nó được xem là ngành đào tạo mũi nhọn hướng đến sự phát triển của công nghệ và khoa học kỹ thuật trong thời đại số hóa.

Chương trình đào tạo ngành CNTT trang bị cho người học những kiến thức nền tảng và chuyên sâu về lĩnh vực công nghệ thông tin để nâng cao tay nghề nhằm phát triển khả năng sửa chữa, xây dựng, cài đặt, bảo trì các phần cứng của máy tính cũng như nghiên cứu và phát triển các ứng dụng phần mềm. Ngoài ra cũng được trang bị kiến thức về an toàn và bảo mật thông tin mạng, một trong những lĩnh vực quan trọng được quan tâm hàng đầu trên thế giới hiện nay.

Để đáp ứng nhu cầu tìm kiếm và sử dụng tài liệu học tập các môn học chuyên ngành “Công nghệ thông tin” của sinh viên. Thư viện Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh đã biên soạn tài liệu thông tin về đề cương chi tiết môn học nhằm cung cấp cho sinh viên có thể nghiên cứu, tìm hiểu, học tập về các môn học chuyên ngành.

Tài liệu thông tin về đề cương chi tiết môn học ngành “Công nghệ thông tin” nhằm hướng dẫn tìm kiếm và sử dụng tài liệu học tập các môn học chuyên ngành “Công nghệ thông tin” sẽ mang đến cho sinh viên nắm được phương pháp học nhanh nhất và đạt hiệu quả.

Mặc dù chúng tôi đã cố gắng, song việc biên soạn tài liệu chắc chắn không tránh khỏi thiếu sót, rất mong nhận được những ý kiến đóng góp từ phía người sử dụng để lần biên soạn sau được hoàn chỉnh hơn.

Thư viện ĐHSPKT TP.HCM
028.389 69 920

thuvien@hcmute.edu.vn
thuvienspkt@hcmute.edu.vn
facebook.com/hcmute.lib

GIỚI THIỆU CÁC DỊCH VỤ UTE LIBRARY

Đồng hành cùng độc giả trên con đường chinh phục tri thức

Nơi cung cấp nguồn lực thông tin khoa học kỹ thuật và giáo dục phục vụ cho nhu cầu đào tạo của Nhà trường.

Cung cấp thông tin

- ✓ Nội dung phong phú
- ✓ Đa dạng loại hình
- ✓ Cập nhật thường xuyên

Hình thức phục vụ

- ✓ Đọc tại chỗ
- ✓ Mượn về nhà
- ✓ Khai thác tài nguyên số 24/24
- ✓ Các dịch vụ học tập trực tuyến

Các loại hình dịch vụ

1. Phục vụ trực tiếp tại thư viện: được hướng dẫn tận tình với hệ thống phòng đọc & Giáo trình mở, có thể tìm đọc tài liệu dạng giấy, CD-ROM, CSDL trực tuyến,...
2. Hỗ trợ công tác biên soạn xuất bản giáo trình và tài liệu học tập, kỹ yếu hội thảo (phục vụ hoạt động NCKH, chương trình đào tạo 150 tín chỉ, bổ sung hồ sơ xét các chức danh khoa học,...).
3. Thiết kế website phục vụ hoạt động học tập, nghiên cứu khoa học, hội thảo: Xây dựng template chung, Thiết lập các trang con (sub pages), Cấp quyền và chuyển giao quyền quản trị trang con.
4. Xuất bản kỷ yếu hội thảo: Tư vấn, thiết kế, dàn trang, Thiết kế các hình ảnh, nhãn hiệu liên quan đến hội nghị (logo hội nghị, banner, poster...), Đăng ký và xin giấy phép xuất bản kỷ yếu, Giám sát các tài liệu liên quan đến chương trình như thư, thông tin hội nghị, tài liệu tham khảo,...
5. Xây dựng các sản phẩm phục vụ hội thảo và các hoạt động học tập: CD-ROM chương trình và kỷ yếu, Ứng dụng tự chạy giới thiệu tài liệu, ghi đĩa CD- ROM, DVD, USB,...
6. Cung cấp thông tin theo yêu cầu (danh mục tài liệu, tài liệu chuyên ngành, kết quả nghiên cứu khoa học, tiêu chuẩn, phát minh sáng chế...).
7. Mô tả, tạo và chọn danh mục “Tài liệu tham khảo” cho luận văn, luận án, đề tài nghiên cứu khoa học theo chuẩn quốc tế.
8. Chuyển đổi tài liệu (từ tài liệu dạng giấy sang file PDF, từ file PDF sang file Word).
9. Cung cấp các thiết bị hỗ trợ cho việc học tập và nghiên cứu,...
10. Hỗ trợ, tư vấn và phối hợp với Thư viện các trường trong việc phát triển các sản phẩm đặc thù của từng đơn vị.

Với nhiều hình thức phục vụ phong phú, thuận tiện cho người học NGUỒN TÀI NGUYÊN THÔNG TIN

- CSDL Giáo trình và Tài liệu học tập
- CSDL Luận văn, Luận án
- CSDL các bài báo khoa học (các bài báo nghiên cứu của học viên theo chương trình đào tạo sau đại học được đăng trên các tạp chí chuyên ngành)
- CSDL Sách tham khảo Việt văn
- CSDL Sách tham khảo Ngoại văn

Địa chỉ liên hệ:

Thư viện Trường Đại học Sư phạm Kỹ thuật Tp. Hồ Chí Minh,
Số 1-3 Võ Văn Ngân, Phường Linh Chiểu, Quận Thủ Đức, Tp. Hồ Chí Minh.

Điện thoại: (+84 028) 3896 9920, 3721223 (nội bộ 8223)

Email: thuvienspkt@hcmute.edu.vn

<http://thuvien.hcmute.edu.vn>

<http://thuvienso.hcmute.edu.vn>

GIỚI THIỆU CÁC WEBSITE PHÁT HÀNH UTE EBOOK

“UTE EBOOK” là những tác phẩm chỉ có thể dùng các công cụ điện tử như máy vi tính, máy trợ giúp kỹ thuật số cá nhân (thiết bị đọc, máy tính bảng và điện thoại thông minh như iPhone, iPad, Samsung Galaxy, HTC Tablet,...) để xem, đọc, và truyền tải.

“UTE EBOOK” là hệ thống phân phối sách điện tử, cung cấp đến người dùng các nội dung sách, giáo trình, truyện, tạp chí, tài liệu, chuyên đề và các thể loại văn bản đọc, audio khác trên di động, thông qua các kênh website, wapsite và client giúp khách hàng có thể cập nhật thông tin kiến thức mọi lúc mọi nơi.

Số	Tên đơn vị phát hành	Website	Truy cập nhanh kho giáo trình
1	Nhà Xuất Bản Tổng Hợp Thành Phố Hồ Chí Minh	https://sachweb.com	https://bit.ly/2LSRzXU
2	Công Ty Cổ Phần Tin Học Lạc Việt	http://sachbaovn.vn	https://bit.ly/2Zx8YZn
3	Công Ty TNHH Sách Điện Tử Trẻ (YBOOK)	http://ybook.vn/	https://bit.ly/2GHF21Q
4	Công Ty Cổ Phần Dịch Vụ Trực Tuyến	https://read.alezaa.com	https://goo.gl/4MM7RM

	VINAPO		
5	Công Ty Cổ Phần Thương Mại Dịch Vụ Mê Kông COM	http://reader.vinaboo.com	https://goo.gl/i6Qpb1
6	Thư viện Trường Đại học Sư phạm Kỹ thuật TP. HCM	http://thuvien.hcmute.edu.vn/	http://thuvien.hcmute.edu.vn/

DỊCH VỤ HỖ TRỢ XUẤT BẢN ĐIỆN TỬ:
**Sách chuyên khảo, giáo trình, sách tham khảo, sách hướng dẫn, tài
liệu phục vụ đào tạo, Ký yếu hội thảo, tập san, tạp chí**

Nhằm hỗ trợ Quý Thư viện các Trường Đại học, Cao đẳng về việc xuất bản giáo trình, tài liệu học tập điện tử nội sinh,...đạt chất lượng cao phục vụ hoạt động đào tạo, giảng dạy; Thư viện Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh giới thiệu đến Quý Thư viện các Trường Đại học, Cao đẳng về “*Chương trình hỗ trợ xuất bản điện tử: Sách chuyên khảo, giáo trình, sách tham khảo, sách hướng dẫn, tài liệu phục vụ đào tạo, Ký yếu hội thảo, tập san, tạp chí*”.

1. Cơ sở xây dựng chương trình:

Căn cứ theo công văn Số: 4301/BGDDĐT-GDTX ngày 20 tháng 9 năm 2019 của Bộ trưởng Bộ Giáo dục và Đào tạo về việc xây dựng và phát triển tài nguyên giáo dục mở tại website: <https://itrithuc.vn>

2. Thời gian triển khai chương trình hỗ trợ

- Giai đoạn 05 năm (Từ ngày 01/10/2019 đến hết ngày 01/10/2024)

3. Tổ chức chương trình

- Thư viện Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh.

4. Đơn vị công bố phổ biến xuất bản phẩm điện tử

- Thư viện Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh & Các đơn vị liên kết.
- website: <http://thuvien.hcmute.edu.vn>
- website: <https://itrithuc.vn>
- website các đơn vị liên kết.

5. Cấu trúc của xuất bản điện tử giáo trình, tài liệu học tập nội sinh

Xuất bản điện tử giáo trình, tài liệu học tập nội sinh có cấu trúc như sau:

- Trang bìa.
- Trang bìa phụ.
- Lời nói đầu: Trình bày đối tượng sử dụng, mục đích yêu cầu khi sử dụng, cấu trúc nội dung, điểm mới của giáo trình, hướng dẫn cách sử dụng, phân công nhiệm vụ của tác giả biên soạn.

- Mục lục.
- Danh mục các từ viết tắt (nếu có).
- Danh mục hình (nếu có)..
- Danh mục bảng biểu (nếu có)..
- Nội dung chính: Trình bày các chương, mục, tiểu mục và nội dung chi tiết của từng chương, mục, tiểu mục, nội dung thảo luận xêmina, câu hỏi ôn tập, bài tập, các nhiệm vụ tự học và tài liệu học tập từng chương.
- Phụ lục (nếu có).
- Tài liệu tham khảo.

6. Để biết thêm thông tin chi tiết Quý Thầy/Cô liên hệ

- Thư viện Trường Đại Học Sư phạm Kỹ thuật TP. Hồ Chí Minh.
- Số 1-3 Võ Văn Ngân, Phường Linh Chiểu, Quận Thủ Đức, TP. Hồ Chí Minh.
- Email: luatvt@hcmute.edu.vn; thuvienspkt@hcmute.edu.vn
- Điện thoại: 08.38969920 hoặc 0909836920; 0906836920 (ThS. Vũ Trọng Luật)

Xin trân trọng giới thiệu đến Quý Thầy/Cô!

GIỚI THIỆU CÁC CƠ SỞ DỮ LIỆU TRỰC TUYẾN TÌM KIẾM VÀ SỬ DỤNG HIỆU QUẢ

Thư viện ĐH SPKT TP. HCM

<http://thuvien.hcmute.edu.vn>

thuvienspkt@hcmute.edu.vn

028.38969920

Nhằm tăng cường năng lực nghiên cứu của tập thể giảng viên và sinh viên của nhà trường, từng bước nâng cao chất lượng đào tạo và nâng vị thế của trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh lên tầm cao mới, xứng đáng là một trong những trường đại học hàng đầu của khu vực, nhà trường đã tiến hành mua quyền sử dụng tài khoản truy cập các cơ sở dữ liệu trực tuyến để giảng viên và sinh viên của trường khai thác. Việc khai thác các tư liệu điện tử, các CSDL trực tuyến sẽ góp phần tạo thêm nhiều cơ hội tiếp cận với các kho tư liệu học thuật và nghiên cứu có giá trị cao và được phổ biến trên toàn thế giới.

Để triển khai sử dụng và khai thác các cơ sở dữ liệu trực tuyến đạt hiệu quả, Thư viện xin thông báo đến Quý Thầy Cô & Các Bạn Sinh viên về nội dung triển khai và sử dụng CSDL như sau:

Thông tin triển khai sử dụng:

- Truy cập Web, không giới hạn số lượng người sử dụng, hỗ trợ việc truy cập cho 25.000 sinh viên và gần 1.000 giảng viên, cán bộ viên chức của nhà trường.

- **Dịa chỉ truy cập:** <http://cSDL.hcmute.edu.vn>
- Sử dụng email do nhà trường cấp để tạo tài khoản đăng ký:
theo địa chỉ email của cán bộ ...@hcmute.edu.vn
theo địa chỉ email của sinh viên ...@student.hcmute.edu.vn
- Hình thức truy cập: Theo IP đăng kí của nhà trường.
- **Lưu ý:** Việc khai thác các CSDL được nhà xuất bản xác minh người dùng với nhiều lớp bảo mật qua dài địa chỉ IP và username/password, quý thầy cô và bạn đọc chỉ có thể khai thác các CSDL này trên hệ thống máy tính được kết nối mạng trong trường.

Thông tin hỗ trợ kỹ thuật:

Trong quá trình sử dụng, quý thầy cô và bạn đọc cần hỗ trợ kỹ thuật xin vui lòng liên hệ:

Cô Trần Thị Ngọc Ý, Số ĐT 0919888975,
email: yttn@hcmute.edu.vn

Thư viện trân trọng thông báo đến Quý Thầy Cô và Các Bạn Sinh viên biết để sử dụng và khai thác các cơ sở dữ liệu trực tuyến đạt hiệu quả.

Trân trọng!

ĐỀ CƯƠNG CHI TIẾT MÔN HỌC

PHÂN TÍCH DỮ LIỆU LỚN

BIG DATA ANALYSIS

Môn học Phân tích dữ liệu lớn cung cấp cho sinh viên kiến thức về kiến trúc của các hệ thống và các công cụ phục vụ cho hoạt động phân tích dữ liệu lớn. Với mỗi công cụ, môn học giới thiệu các kiến thức cơ bản và nâng cao cũng như phương thức tối ưu hóa hiệu suất hệ thống sử dụng công cụ này. Cùng với các bài tập lập trình, môn học hướng đến mục tiêu giúp người học có thể hình thành ý tưởng, thiết kế và hiện thực hóa hoạt động phân tích dữ liệu trong các hệ thống dữ liệu lớn.

TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
TP. HỒ CHÍ MINH
KHOA CÔNG NGHỆ THÔNG TIN

Ngành đào tạo: Kỹ thuật Dữ liệu
Trình độ đào tạo: Đại học
Chương trình đào tạo: Kỹ thuật Dữ liệu

Đề cương chi tiết môn học

1. Tên môn học: PHÂN TÍCH DỮ LIỆU LỚN

Mã môn học: BDAN333977

2. Tên Tiếng Anh: BIG DATA ANALYSIS

3. Số tín chỉ: 3 (2+1) tín chỉ (2/1/8) (2 tín chỉ lý thuyết, 1 tín chỉ thực hành/thí nghiệm)

Phân bố thời gian: 10 tuần (2 tiết lý thuyết + 1*2 tiết thực hành + 8 tiết tự học/ tuần)

4. Các giảng viên phụ trách môn học:

1/

2/

5. Điều kiện tham gia học tập môn học

Môn học tiên quyết:

Môn học trước: Lập trình Python cho kỹ thuật dữ liệu

6. Mô tả môn học (Course Description)

Môn học cung cấp cho sinh viên kiến thức về kiến trúc của các hệ thống và các công cụ phục vụ cho hoạt động phân tích dữ liệu lớn. Với mỗi công cụ, môn học giới thiệu các kiến thức cơ bản và nâng cao cũng như phương thức tối ưu hóa hiệu suất hệ thống sử dụng công cụ này. Cùng với các bài tập lập trình, môn học hướng đến mục tiêu giúp người học có thể hình thành ý tưởng, thiết kế và hiện thực hóa hoạt động phân tích dữ liệu trong các hệ thống dữ liệu lớn.

7. Mục tiêu môn học (Course Goals)

Mục tiêu (Goals)	Mô tả (Goal description) (Môn học này trang bị cho sinh viên:)	Chuẩn đầu ra CDIO	Trình độ năng lực
G1	Kiến thức về kiến trúc hệ thống dữ liệu lớn và các công cụ phân tích dữ liệu lớn.	1.3	2
G2	Phân tích kiến trúc hệ thống dữ liệu lớn và tình huống áp dụng các công cụ phân tích dữ liệu lớn.	2.1	2
G3	Kỹ năng làm việc nhóm, giao tiếp và khả năng đọc hiểu các tài liệu kỹ thuật bằng tiếng Anh trong lĩnh vực dữ liệu lớn.	3.1 3.2 3.3	2 2 2
G4	Hình thành ý tưởng, thiết kế và hiện thực hóa một phần của hệ thống dữ liệu lớn.	4.3 4.4	3 3

8. Chuẩn đầu ra của môn học

Chuẩn đầu ra MH	Mô tả (Sau khi học xong môn học này, người học có thể:)	Chuẩn đầu ra CDIO	Trình độ năng lực
G1	G1.1 Trình bày được kiến trúc cho các hệ thống dữ liệu lớn và các thành phần của hệ sinh thái dữ liệu lớn hiện nay.	1.3	2
	G1.2 Trình bày được kiến thức chuyên ngành về các công cụ Pig, Hive, và Spark SQL.	1.3	2
G2	G2.1 Phân tích được các thành phần trong kiến trúc của một hệ thống dữ liệu lớn	2.1	2
	G2.2 Phân tích được tình huống áp dụng các công cụ Pig, Hive, và Spark SQL trong các hệ thống dữ liệu lớn	2.1	2
G3	G3.1 Làm việc trong các nhóm hình thành ý tưởng, thiết kế, triển khai và vận hành các hệ thống big data có sử dụng các công cụ Pig, Hive, và Spark SQL.	3.1 3.2	2 2
	G3.2 Giải thích được các thuật ngữ tiếng Anh trong lĩnh vực hệ thống dữ liệu lớn có liên quan đến các công cụ Pig, Hive, và Spark SQL.	3.3	2
G4	G4.1 Hình thành ý tưởng, thiết kế, và hiện thực hóa một phần hệ thống xử lý dữ liệu lớn ở cấp độ đơn giản sử dụng Pig, Hive, Spark SQL	4.3 4.4	3

9. Đạo đức khoa học:

Các bài tập ở nhà và đồ án phải được thực hiện từ chính bản thân sinh viên. Nếu bị phát hiện có sao chép thì xử lý các sinh viên có liên quan bằng hình thức đánh giá **0** (không) điểm quá trình và cuối kỳ.

10. Nội dung chi tiết môn học:

Tuần	Nội dung	Chuẩn đầu ra môn học	Trình độ năng lực	Phương pháp dạy học	Phương pháp đánh giá
1-2	<i>Chương 1: Kiến trúc hệ thống dữ liệu lớn</i>				
	A/ Các nội dung và PPGD chính trên lớp: (8) Nội Dung (ND) GD trên lớp 1.1. Giới thiệu một số ví dụ hệ thống dữ liệu lớn 1.2. Phân tích, so sánh các ví dụ hệ thống dữ liệu lớn 1.3 Giới thiệu kiến trúc các hệ thống dữ liệu lớn 1.4 Hoạt động và công cụ phân tích dữ liệu lớn.	G1.1 G2.1 G3.1 G3.2	2 2 2 2	**	**
3-6	B/ Các nội dung cần tự học ở nhà: (16) + Tìm hiểu ví dụ về hệ thống dữ liệu lớn.	G1.1	2		
	<i>Chương 2: Apache Pig</i> A/ Các nội dung và PPGD chính trên lớp: (16) Nội Dung (ND) GD trên lớp 2.1 Các kiến thức nền tảng về công cụ Pig. 2.2 Cơ bản về ngôn ngữ Pig Latin. 2.3 Ngôn ngữ Pig Latin nâng cao. 2.4 Tối ưu hóa hiệu suất hệ thống sử dụng Pig.	G1.2 G2.2 G3.1 G3.2 G4.1	2 2 2 2 2		
7-10	B/ Các nội dung cần tự học ở nhà: (32) + Bài tập lập trình Pig	G4.1	2		
	<i>Chương 3: Apache Hive</i>				
	A/ Các nội dung và PPGD chính trên lớp: (16) Nội dung GD lý thuyết: 3.1 Các kiến thức nền tảng về công cụ Hive 3.2 Cơ bản về ngôn ngữ	G1.2 G2.2 G3.1 G3.2 G4.1	2		

	HiveQL 3.3 Ngôn ngữ HiveQL nâng cao 3.4 Tối ưu hóa hiệu suất hệ thống sử dụng Hive				
	B/ Các nội dung cần tự học ở nhà: (32) + Bài tập lập trình Hive	G4.1	2		
11-14	Chương 4: Spark SQL A/ Các nội dung và PPGD chính trên lớp: (16) Nội dung GD lý thuyết: 4.1 Cơ bản về Spark SQL 4.2 Spark SQL nâng cao 4.3 Xây dựng ứng dụng sử dụng Spark SQL 4.4 Tối ưu hóa hiệu suất hệ thống sử dụng Spark SQL	G1.2 G2.2 G3.1 G3.2 G4.1			
	B/ Các nội dung cần tự học ở nhà: (32) + Bài tập lập trình Spark SQL	G4.1			
15	Ôn tập & Báo cáo đồ án môn học A/ Các nội dung và PPGD chính trên lớp: (4) Nội dung GD lý thuyết: Ôn tập Báo cáo đồ án môn học	G4.1			
	B/ Các nội dung cần tự học ở nhà: (8) + Chuẩn bị báo cáo đồ án môn học	G4.1			

** Tham khảo văn bản hướng dẫn để xác định phương pháp giảng dạy và phương pháp đánh giá giúp đạt được trình độ năng lực mong muốn.

11. Đánh giá kết quả học tập:

- Thang điểm: 10

- Kế hoạch kiểm tra như sau:

Hình thức KT	Nội dung	Thời điểm	Chuẩn đầu ra đánh giá	Trình độ năng lực	Phương pháp đánh giá	Công cụ đánh giá	Tỉ lệ (%)
Bài tập							40
BT#1	Thuyết trình về hệ thống dữ liệu lớn	Tuần 1-2	G1.1 G2.1 G3.1	2 2 2	**	Thuyết trình	10
BT#2	Kiểm tra lý thuyết chương 2, 3, 4	Tuần 3 - 14	G1.2 G2.2 G3.2	2 2 2		Kiểm tra trên	30

							lớp	
	Bài tập lớn (Project)							
BL#1	Bài tập lập trình chương 2, 3, 4	3 - 14	G4.1	3			Bài tập về nhà	30
	Tiểu luận - Báo cáo							
TL#1	Các nhóm sinh viên được yêu cầu đưa ra ý tưởng, thiết kế, hiện thực hóa một phần hệ thống phân tích dữ liệu lớn sử dụng các công cụ Pig, Hive và Spark SQL (Nhóm SV lựa chọn đề tài rồi thông báo cho GV, GV gợi ý, hướng dẫn SV cách làm). Bài tiểu luận được các nhóm SV báo cáo vào tuần cuối cùng.	Tuần 1-15	G3.1 G4.1	2 3			Tiểu luận - Báo cáo	30
Thi cuối kỳ								

CDR môn học	Hình thức kiểm tra						
	BT #1	BT #2	...	BL #1	...	TL #1	Thi cuối kỳ
G1.1	X						
G1.2		X					
G2.1	X					X	
G2.2		X				X	
G3.1	X					X	
G3.2		X					
G4.1				X		X	

12. Tài liệu học tập

- Sách, giáo trình chính:

1. **Gates A.**, Programming Pig, O'Reilly Media, 2011.
2. **Capriolo E. Wampler D. & Rutherford J.**, Programming Hive, O'Reilly Media, 2012.
3. **Sarkar A.**, Learning Spark SQL, Packt Publishing, 2017.

- Sách (TLTK) tham khảo:

13. Ngày phê duyệt lần đầu:

14. Cấp phê duyệt:

Trưởng khoa

Trưởng BM

Nhóm biên soạn

15. Tiến trình cập nhật ĐCCT

Lần 1: Nội Dung Cập nhật ĐCCT
lần 1: ngày tháng năm

(người cập nhật ký và
ghi rõ họ tên)

Tổ trưởng Bộ môn:

GIỚI THIỆU CÁC CHƯƠNG TRÌNH ĐÀO TẠO

1. Ngành Công nghệ Thông tin (7480201D, 7480201C, 7480201A)

Sinh viên có kiến thức chuyên môn vững chắc về lĩnh vực Công nghệ thông tin (Công nghệ phần mềm, Mạng máy tính, Hệ thống thông tin); Sinh viên được chú trọng học tập cả về lý thuyết lẫn kỹ năng thực hành nhằm đảm bảo Sinh viên ra trường có thể tiếp cận và làm việc ngay tại các công ty.

Ra trường, Sinh viên có thể làm việc tại: Các công ty phần mềm, các công ty cung cấp giải pháp mạng, các công ty cung cấp giải pháp hệ thống thông tin trong nước cũng như quốc tế,... Có thể tham gia nghiên cứu và giảng dạy tại các Viện nghiên cứu, Trường Đại học, Cao đẳng, Trung cấp chuyên nghiệp, cũng như các Trường phổ thông; có thể làm việc tại các công ty ngoài lĩnh vực CNTT nhưng có áp dụng triển khai CNTT trong hoạt động, quản lý, sản xuất.

Hiện nay nhu cầu về kỹ sư phần mềm tại Việt Nam là rất lớn bởi sự chuyển hướng đầu tư của các công ty nước ngoài vào Việt Nam. Ngoài ra, SV sau khi hoàn thành chương trình đại học, có khả năng học tiếp lên các bậc học cao hơn (thạc sĩ, tiến sĩ).

Được đào tạo tại khoa Công nghệ Thông tin và khoa Đào tạo Chất lượng cao.

2. Ngành Kỹ thuật dữ liệu (7480203D)

Ngành Kỹ thuật dữ liệu trang bị cho sinh viên cho sinh viên kiến thức cơ sở ngành về công nghệ thông tin, và kiến thức chuyên sâu về phân tích, thiết kế, xây dựng các mô hình vật lý và mô hình lô-gic cho dữ liệu, xây dựng hạ tầng dữ liệu (data infrastructure, data sets) và các công cụ phục vụ việc thiết kế, dự đoán các sản phẩm tương lai.

Các kiến thức được trang bị cho sinh viên như: Cấu trúc dữ liệu và giải thuật, lập trình hướng đối tượng, lập trình web, trí tuệ nhân tạo, Khai phá dữ liệu, học máy, mô hình rủi ro tài chính, v.v...

Sinh viên tốt nghiệp có thể làm việc trong các công ty Công nghệ thông tin, cũng như bộ phận công nghệ thông tin của các công ty ở các lĩnh vực khác. Các vị trí có thể đảm nhận trong các tổ chức gồm:

- Data Engineer (Kỹ sư dữ liệu).
- Big data Engineer (Kỹ sư quản trị dữ liệu lớn).
- Lập trình viên, chuyên viên phát triển phần mềm.
- Chuyên viên quản trị cơ sở dữ liệu, quản trị web.
- Trưởng dự án phát triển hệ thống thông tin.

- Cán bộ nghiên cứu và ứng dụng Công nghệ thông tin ở các viện, trung tâm nghiên cứu và các trường đại học, cao đẳng.
- Giảng viên Công nghệ thông tin ở các các trường đại học, cao đẳng, trung học phổ thông, trung học chuyên nghiệp và dạy nghề.



TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT THÀNH PHỐ HỒ CHÍ MINH

www.hcmute.edu.vn

ISBN: 978-604-73-2175-9



9 786047 321759