

CHƯƠNG TRÌNH ĐÀO TẠO TRÌNH ĐỘ ĐẠI HỌC
NGÀNH ĐÀO TẠO: KHOA HỌC DỮ LIỆU VÀ TRI TUỆ NHÂN TẠO

ĐỀ CƯƠNG CHI TIẾT HỌC PHẦN ĐỒ ÁN
(DÀNH CHO HỌC PHẦN CƠ GIẢNG DẪY THEO ĐỒ ÁN)
TH94011: TRIỂN KHAI, TÍCH HỢP VÀ VẬN HÀNH HỆ THỐNG DỮ LIỆU LỚN
(BIG DATA SYSTEMS: DEPLOYMENT, INTEGRATION, AND OPERATION)

I. Thông tin về học phần

- Học kì: 7
- Tin chỉ: 03 (Lý thuyết 02 – Thực hành 01 - Tự học 09)

- Giờ tin chỉ đối với các hoạt động học tập: 45 tiết
- + Học lý thuyết trên lớp: 30 tiết
- + Làm bài tập trên lớp: 15 tiết

- Giờ tự học: 135 tiết
- Đơn vị phụ trách:

- Bộ môn: Khoa học máy tính
- Khoa: Công nghệ thông tin

- Học phần thuộc khối kiến thức:

Đại cương ☐	Cơ sở ngành ☐	Chuyên sâu ☒
Bắt buộc ☐	Tự chọn ☐	Bắt buộc ☒
Tự chọn ☐	Bắt buộc ☐	Tự chọn ☐

- Học phần học song hành:
- Học phần tiên quyết:
- Ngôn ngữ giảng dạy: tiếng Anh ☐ Tiếng Việt ☒

II. Chuẩn đầu ra của chương trình đào tạo và mục tiêu, kết quả học tập mong đợi của học phần
* Các chuẩn đầu ra và chỉ báo của chương trình đào tạo mà học phần đóng góp:

Chuẩn đầu ra	Chỉ báo chuẩn đầu ra của chương trình đào tạo
Sau khi hoàn tất chương trình, sinh viên có thể:	
Kiến thức chuyên môn	
CDR2. Phân tích được các vấn đề cơ bản về công nghệ liên quan đến xây dựng và vận hành hệ thống tin	2.2. Phân tích được các vấn đề cơ bản về công nghệ liên quan đến vận hành hệ thống thông tin.
CDR3. Xây dựng được hệ thống khai phá dữ liệu và hệ thống tri tuệ nhân tạo.	3.1 Xây dựng được hệ thống xử lý và phân tích dữ liệu.
Kỹ năng chung	

Chỉ bảo chuẩn đầu ra của chương trình đào tạo	Sau khi hoàn tất chương trình, sinh viên có thể:	CDR5. Sử dụng kỹ năng tư duy logic, kỹ năng phân tích và kỹ năng làm việc độc lập, làm việc nhóm vào việc giải quyết các vấn đề chuyên môn trong lĩnh vực Khoa học dữ liệu và Trí tuệ nhân tạo	Kỹ năng chuyên môn	CDR6. Kết hợp các kỹ năng cơ bản để xây dựng và vận hành hệ thống thông tin.	CDR7. Phát triển hệ thống khai phá dữ liệu và trí tuệ nhân tạo để giải quyết các vấn đề trong nhiều lĩnh vực	Năng lực tự chủ và trách nhiệm	
				5.1. Vận dụng kỹ năng tư duy logic, kỹ năng phân tích để giải quyết hoặc đánh giá các vấn đề chuyên môn.			
				6.2. Kết hợp các kỹ năng cơ bản để vận hành hệ thống thông tin.			
					7.1. Phát triển hệ thống khai phá dữ liệu để giải quyết các vấn đề trong nhiều lĩnh vực.		
							8.2. Đề xuất các ý tưởng khởi nghiệp

*** Mục tiêu:** Học phần nhằm cung cấp cho sinh viên

- Về kiến thức:

- o Sinh viên có thể xác định được vai trò, tầm quan trọng của các hệ thống dữ liệu lớn
- o Sinh viên hiểu được khi nào cần các hệ thống về dữ liệu lớn.
- o Sinh viên hiểu được thuật ngữ và các vấn đề chính trong các hệ thống dữ liệu lớn
- Về kỹ năng: Sinh viên có thể cài đặt và triển khai và vận hành các hệ thống dữ liệu lớn
- Về năng lực tự chủ và trách nhiệm: Hình thành tinh thần khởi nghiệp và học tập suốt đời.

*** Kết quả học tập mong đợi của học phần:**

Học phần đóng góp cho Chuẩn đầu ra sau đây của CTĐT theo mức độ sau:

I – Giới thiệu (Introduction); P – Thực hiện (Practice); R – Củng cố (Reinforce); M – Đạt được (Master)

Mã HP	Tên HP	Mức độ đóng góp của học phần cho CDR của CTĐT					
		2.2	3.1	5.1	6.2	7.1	8.2
	Triển khai, tích hợp và vận hành hệ thống dữ liệu lớn	R	R	P	M	M	M

Ký hiệu	KQHTMB của học phần		Chỉ bảo CDR của CTĐT
K1	Trình bày được ứng dụng của các hệ thống dữ liệu lớn.	2.2	
K2	Giải thích được các vấn đề, công cụ và giải pháp trong các hệ thống dữ liệu lớn	3.1	
Kỹ năng			
K3	Triển khai hạ tầng lưu trữ dữ liệu lớn	5.1	
K4	Xử lý dữ liệu lớn	6.2	
K5	Quản trị vận hành hệ thống dữ liệu lớn	7.1	
Năng lực tự chủ và trách nhiệm			
K6	Có tinh thần khởi nghiệp trong lĩnh vực chuyên môn.	8.2	

Hoạt động đánh giá	KQHTMD được đánh giá	Trọng số (%)	Thời gian/Tuần học
Đánh giá quá trình			
Tham dự lớp	K1, K2	10%	1-9

Bảng 1: Kế hoạch đánh giá và trọng số

- Điểm kiểm tra cuối kì: 50%
- Điểm quá trình/Điểm kiểm tra giữa kì: 40%
- Điểm chuyên cần: 10 %

3. Phương pháp đánh giá

từng rubric

2. Điểm trung bình của học phần là tổng điểm của các rubric nhân với trọng số tương ứng của

1. Thang điểm: 10

VI. Đánh giá và cho điểm

- Thực hành: Sinh viên cần tham dự đầy đủ các buổi thực hành.
- Giảng viên, tài liệu tham khảo, đọc trước các nội dung được giao trước khi vào học;
- Chuẩn bị cho bài giảng: Tất cả sinh viên tham dự học phần này phải có giáo trình, bài giảng của
- Chuyên cần: Tất cả sinh viên tham dự học phần này bằng hình thức học trực tiếp hay học trực
- Tuyên đầu phải tham gia tối thiểu theo quy định đầy và học của Học viện;
- Khi sinh viên tham gia học trực tiếp trên giảng đường hay học trực tuyến qua MS Teams đều phải

V. Nhiệm vụ của sinh viên

- Học tập chủ động.
- Tham gia thảo luận.
- Làm bài tập
- Áp dụng cho sinh viên tham gia học trực tiếp trên giảng đường và học trực tuyến qua MS Teams:

2. Phương pháp học tập

- Phương pháp giảng dạy trực tuyến (E-learning).

thiết kế bài tập mẫu.

- Vấn đề, hướng dẫn và khuyến khích sinh viên giải quyết. Giảng viên cũng trình bày phân tích và
- Dạy học dựa trên tình huống: Giảng viên trình bày và giải thích các nội dung môn học, đặt ra các
- Phương pháp hỏi đáp: đặt câu hỏi và trả lời câu hỏi.
- Giảng dạy truyền thông: Giảng viên thực hiện giảng lý thuyết

Teams:

Áp dụng cho Giảng viên giảng dạy trực tiếp trên giảng đường và giảng dạy trực tuyến qua MS

1. Phương pháp giảng dạy

IV. Phương pháp giảng dạy và học tập

hành hệ thống dữ liệu lớn. Cung cấp những kỹ năng cơ bản trong thực hành xây dựng, triển khai và vận

Mô tả văn tài nội dung: Học phần gồm 4 chương với các nội dung: giới thiệu tổng quan về các hệ

chỉ II thuyết 02 – Tổng số tín chỉ thực hành 01 – Tổng số tín chỉ tự học: 09).

TH94011 - Triển khai, tích hợp và vận hành hệ thống dữ liệu lớn (Tổng số tín chỉ 03: Tổng số tín

III. Nội dung tóm tắt của học phần

VII. Giao trình/ tài liệu tham khảo

hình thức đánh giá 0 (không) điểm qua trình và cuối kỳ.

chính bản thân sinh viên. Nếu bị phát hiện có sao chép thì xử lý các sinh viên có liên quan bằng

- *Yêu cầu về đạo đức:* Các bài tập, bài thực hành, bài kiểm tra, bài thi phải được thực hiện bởi qua học phần.

tham gia bài thi cuối kỳ. Sinh viên không tham gia bài thi cuối kỳ sẽ không đủ điều kiện để xét

- *Tham dự các bài thi:* Sinh viên không tham gia đánh giá giữa kỳ sẽ không được xét đủ điều kiện

4. Các yêu cầu, quy định đối với học phần

Tiêu chí	Trọng số (%)	Tốt 8.5 – 10 điểm	Khả 6.5 – 8.4 điểm	Trung bình 4.0 – 6.4 điểm	Kém 0 – 3.9 điểm	
Thái độ tham dự	50	Luôn chú ý và tham gia các hoạt động	Khả chú ý, có tham gia	Có chú ý, ít tham gia	Không chú ý/không tham gia	
Thời gian tham dự	50	- Không được vắng quá thời gian quy định trong Quy định dạy và học				

Rubric 1: Đánh giá quá trình (Tham dự lớp, Làm bài tập)

(Áp dụng cho cả hai tình huống học trực tiếp và trực tuyến)

KQHTMB	Chỉ báo thực hiện KQHTMB	Chỉ báo 1: Trình bày được các trường hợp sử dụng thực tế của các hệ thống dữ liệu lớn.	Chỉ báo 2: Trình bày được lĩnh vực ứng dụng	Chỉ báo 3: Trình bày được các vấn đề trong các hệ thống dữ liệu lớn.	Chỉ báo 4: Trình bày được các giải pháp cho các vấn đề trong các hệ thống dữ liệu lớn	Chỉ báo 5: Trình bày được các công cụ hỗ trợ trong các hệ thống dữ liệu lớn	Chỉ báo 6: Tích hợp được các công cụ quản lý	Chỉ báo 7: Trình khai thác được hệ thống dữ liệu lớn	Chỉ báo 8: Xây dựng và thi hành được các kịch bản xử lý dữ liệu lớn.	Chỉ báo 9: Thực hành được các kịch bản xử lý dữ liệu lớn.	Chỉ báo 10: Trình khai thác được các công cụ hỗ trợ quản trị các hệ thống dữ liệu lớn	Chỉ báo 11: Giám sát được các hệ thống dữ liệu lớn	Chỉ báo: Giải thích được ý tưởng đề xuất
K6													
K5													
K4													
K3													
K2													
K1													

Bảng 2. Chỉ báo thực hiện các kết quả học tập mong đợi của học phần

Hoạt động đánh giá	KQHTMB được đánh giá	Trọng số (%)	Thời gian/Tuần học
Đánh giá quá trình			
Kiểm tra giữa kỳ (đồ án môn học)	K1, K2, K6	40%	8
Đánh giá cuối kỳ			
Thi cuối môn học	K3, K4, K5	50%	Thi theo lịch học viên

*** Sách giáo trình/Bài giảng:**

1. Thomas Erl, Wajid Khattak, Paul Buhler (2015). Big Data Fundamentals: Concepts, Drivers & Techniques, First Edition. Pearson
2. Pradeep Menon (2022). Data Lakehouse in Action: Architecting a modern and scalable data analytics platform, First Edition. Packt Publishing.

*** Tài liệu tham khảo khác:**

1. Bill Inmon, Ranjeet Srivastava (2023). Rise of the Data Lakehouse: Building the Data Lakehouse, First Edition. Technics Publications
2. Ghislain Fourny (2023). The Big Data Textbook: From clay tablets to data lakehouses, First Edition. Independently published.

*** Tài liệu tham khảo trực tuyến:**

- Giảng viên gửi bài giảng bản mềm, các tài liệu tham khảo trực tuyến vào nhóm của môn học trên MS Teams.

VIII. Nội dung chi tiết của học phần

Tuần	Nội dung	KQHTMĐ
1	Chương 1. Tổng quan về các hệ thống dữ liệu lớn A/Các nội dung chính trên lớp: (5 tiết) 1.1. Giới thiệu về môn học, cách học, tài liệu tham khảo 1.2. Giới thiệu về dữ liệu lớn 1.3. Vai trò của các hệ thống dữ liệu lớn 1.4. Phân tích một vài trường hợp sử dụng B/Các nội dung cần tự học ở nhà: (15 tiết) Tìm hiểu các trường hợp sử dụng khác trên internet và liên hệ thực tế.	K1 K1, K5
2,3	Chương 2. Lưu trữ dữ liệu lớn A/Tóm tắt các nội dung chính trên lớp: (12 tiết) Nội dung GD lý thuyết: (5 tiết) 2.1. Vấn đề với lưu trữ dữ liệu lớn 2.2. Giải pháp lưu trữ dữ liệu lớn 2.3. Công cụ lưu trữ dữ liệu lớn Nội dung GD thực hành (7 tiết) - Triển khai giải pháp lưu trữ dữ liệu lớn B/Các nội dung cần tự học ở nhà: (36 tiết) Thực hành thêm tại nhà	K2, K3 K2, K3, K6
4,5	Chương 3. Xử lý dữ liệu lớn A/Tóm tắt các nội dung chính trên lớp: 17 tiết) Nội dung GD lý thuyết: (7 tiết) 3.1. Mô hình xử lý phân tán với dữ liệu lớn 3.2. Mô hình truy vấn dữ liệu lớn Nội dung GD thực hành (10 tiết)	K2, K4

(Áp dụng cho cả hai tình huống học trực tiếp và trực tuyến)

Rubric 1: Làm việc nhóm

4. Rubric đánh giá thực hiện đồ án

- Địa điểm thực hiện đồ án: linh hoạt theo đề tài nhóm thực hiện.
- + Giai đoạn 5: Viết báo cáo phân tích kết quả.
- + Giai đoạn 4: Giả lập vận hành
- + Giai đoạn 3: Triển khai thử nghiệm
- + Giai đoạn 2: Phân tích và đề xuất giải pháp công nghệ
- + Giai đoạn 1: Thành lập nhóm, tìm hiểu về bài toán
- Thời gian thực hiện: 6 tuần và được chia làm các giai đoạn:
- Số lượng sinh viên 2-3 sv/nhóm:

3. Tổ chức thực hiện đồ án:

- túc, chủ động tìm hiểu các kiến thức bổ sung trong quá trình học tập.
- Về các mục tiêu khác (thái độ học tập): Thái độ hợp tác trong nhóm; Làm việc theo nhóm, nghiêm túc, chủ động tìm hiểu các kiến thức bổ sung trong quá trình học tập.
- Về kỹ năng: Sinh viên có thể triển khai, giám sát vận hành hệ thống tích hợp dữ liệu lớn.
- Kiến thức: Sinh viên biết phân tích và lựa chọn được công nghệ phù hợp cho bài toán thực tế.

3. Kết quả học tập mong đợi từ đồ án

- Tên các chủ đề: triển khai và phân tích vận hành hệ thống tích hợp dữ liệu cho cơ quan/tổ chức.
- Sản phẩm và yêu cầu đối với sản phẩm: Hoàn thành hệ thống, giả lập quá trình vận hành, phân tích lý do lựa chọn công nghệ và kết quả vận hành.

2. Mô tả chung về đồ án:

- Giả lập vận hành, phân tích logs và đánh giá.
- Triển khai hệ thống tích hợp dữ liệu từ 3-5 nguồn dữ liệu khác nhau.
- trung tâm tích hợp dữ liệu.
- Nhằm khai thác sức mạnh của dữ liệu (lớn) các cơ quan/tổ chức lớn đã và đang xây dựng các

1. Lý do lựa chọn chủ đề/nội dung thực hiện đồ án

Đồ án là một phần bắt buộc của học phần. Thông tin về đồ án như sau:

IX. Đồ án

	6	A/Tóm tắt các nội dung chính trên lớp: (11 tiết) Nội dung GD lý thuyết: (5 tiết) 4.1. Sự phức tạp trong quản trị, vận hành hệ thống dữ liệu lớn 4.2. Thiết kế hệ thống dữ liệu lớn 4.3. Các công cụ quản trị, vận hành hệ thống dữ liệu lớn Nội dung GD thực hành (6 tiết) - Thực hành triển khai công cụ quản trị giám sát hệ thống dữ liệu lớn	B/Các nội dung cần tự học ở nhà: (33 tiết) Đọc tài liệu, và thực hành tại nhà	K2, K5, K6
		Chương 4. Quản trị, vận hành hệ thống dữ liệu lớn	B/Các nội dung cần tự học ở nhà: (51 tiết) Đọc tài liệu, tìm hiểu thực hành thêm	K2, K4, K6

Tiêu chí	Trọng số %	Tốt 8.5 – 10 điểm	Khá 6.5 – 8.4 điểm	Trung bình 4.0 – 6.4 điểm	Kém 0 – 3.9 điểm
Câu trực	20	Câu đối, hợp lý	Khá cân đối, hợp lý	Tương đối cân đối, hợp lý	Không cân đối, thiếu hợp lý
Nội dung	15 15 15 15	Đặc tả rõ bài toán và đề xuất công nghệ hợp lý	Hệ thống hoạt động được	Giả lập được các tình huống xảy ra	Bảo cáo phân tích vận hành

Rubric 4. Đánh giá Bảo cáo thu hoạch

Tiêu chí	Trọng số (%)	Tốt 8.5 – 10 điểm	Khá 6.5 – 8.4 điểm	Trung bình 4.0 – 6.4 điểm	Kém 0 – 3.9 điểm
Nội dung	40	Phong phú hơn yêu cầu	Đầy đủ theo yêu cầu	Khá đầy đủ, còn thiếu 1 nội dung quan trọng	Thiếu nhiều nội dung quan trọng
Kỹ năng trình bày	30	Dẫn dắt vấn đề và lập luận logic, thuyết phục	Trình bày rõ ràng nhưng chưa logic, lập luận khá thuyết phục	Khó theo dõi nhưng vẫn có thể hiểu được các nội dung quan trọng	Trình bày không rõ ràng, người nghe không thể hiểu được các nội dung quan trọng
Trả lời câu hỏi	30	Các câu hỏi đặt đúng đều được trả lời đầy đủ, rõ ràng, và thỏa đáng	Trả lời đúng đa số câu hỏi đặt đúng và nêu được định hướng phù hợp đối với những câu hỏi	Trả lời đúng đa số câu hỏi đặt đúng nhưng chưa nêu được định hướng phù hợp đối với những câu hỏi chưa trả lời được	Không trả lời được đa số câu hỏi đặt đúng

(Áp dụng cho cả hai tình huống học trực tiếp và trực tuyến)

Rubric 2. Đánh giá THUYẾT TRÌNH (theo nhóm)

Tiêu chí	Trọng số (%)	Tốt 8.5 – 10 điểm	Khá 6.5 – 8.4 điểm	Trung bình 4.0 – 6.4 điểm	Kém 0 – 3.9 điểm
Thái độ tham gia	40	Tích cực kết nối các thành viên trong nhóm	Kết nối tốt với thành viên khác	Có kết nối còn lơ là, phải nhắc nhở	Không kết nối
Ý kiến đóng góp	60	Sáng tạo/rất hữu ích	Hữu ích	Tương đối hữu ích	Không hữu ích

Hình thức trình bày (font chữ, cỡ, lề, format...)	20	Nhất quán về format trong toàn bài. Không có lỗi chính tả	Vài sai sót nhỏ về format. Một vài lỗi nhỏ	Vài chỗ không nhất quán. Lỗi chính tả khá nhiều	Rất nhiều chỗ không nhất quán. Lỗi rất nhiều và do sai chính tả và typing cầu thả
---	----	---	--	---	---

IX. Yêu cầu của giảng viên đối với học phần

- Phương tiện phục vụ giảng dạy: Máy chiếu, bảng viết
- Các phương tiện khác: Loa trợ giảng

- Phương tiện phục vụ giảng dạy E- learning: phần mềm dạy trực tuyến MIS Teams, máy tính, hệ thông máy chủ và hạ tầng kết nối mạng Internet với bảng thông đáp ứng nhu cầu người dùng, không để xảy ra nghẽn mạng hay quá tải. Phòng học trực tuyến đầy đủ ánh sáng, cách âm tốt, thông thoáng, ngăn nắp, gọn gàng, sạch sẽ.

TRƯỜNG BỘ MÔN

(Ký và ghi rõ họ tên)

[Signature]

Trần Đức Hùng

TRƯỜNG KHOA

(Ký và ghi rõ họ tên)

[Signature]

Shim Quang Dũng



(Ký và ghi rõ họ tên)

KT. GIÀM ĐỐC

PHÓ GIÀM ĐỐC

Nguyễn Văn Hùng

(Ký và ghi rõ họ tên)

GIẢNG VIÊN BIÊN SOẠN

Hà Nội, ngày 21 tháng 6 năm 2024

PHÓ GIÀM ĐỐC

Shim Văn Dũng

PHỤ LỤC

THÔNG TIN VỀ ĐỘI NGŨ GIẢNG VIÊN GIẢNG DẠY HỌC PHẦN

Giảng viên phụ trách học phần

Họ và tên: Nguyễn Văn Hoàng	Học hàm, học vị: Thạc sĩ
Địa chỉ cơ quan: Khoa CNTT, Học viện Nông nghiệp Việt Nam	Điện thoại liên hệ:
Email: nvhoang@vnua.edu.vn	Trang web: www.vnua.edu.vn/khoa/fita
Cách liên lạc với giảng viên: email, điện thoại; hoặc đặt lịch gặp trực tiếp tại tầng 3 Nhà Hành Chính.	

Giảng viên giảng dạy học phần

Họ và tên:	Học hàm, học vị:
Địa chỉ cơ quan:	Điện thoại liên hệ:
Email:	Trang web: (Đưa tên website của Khoa; website cá nhân – nếu có)
Cách liên lạc với giảng viên: (nêu rõ hình thức liên lạc giữa sinh viên với giảng viên/trợ giảng)	

X. CÁC LẦN CẢI TIẾN (ĐỀ CƯƠNG ĐƯỢC CẢI TIẾN HÀNG NĂM THEO QUY ĐỊNH CỦA HỌC VIỆN)

- Lần 1: 07/2020
 - Xây dựng đề cương chi tiết mới để mở ngành
- Lần 2: 07/2021
 - Rà soát và cập nhật nội dung kiến thức học phần, rà soát tài liệu tham khảo
- Lần 3: 07/2022
 - Rà soát và cập nhật tài liệu tham khảo
- Lần 4: 07/2023
 - Rà soát và cập nhật tài liệu tham khảo
- Lần 5: 07/2024
 - Cập nhật CDR của CTĐT
 - Cập nhật đóng góp bảng IPRM
 - Thay đổi mã môn học
 - Rà soát và cập nhật nội dung học phần và tài liệu tham khảo