

ĐỀ CƯƠNG MÔN HỌC PHÂN TÍCH DỮ LIỆU CHO TÀI CHÍNH

A. THÔNG TIN CHUNG VỀ MÔN HỌC

1. Tên môn học (tiếng Việt) : **PHÂN TÍCH DỮ LIỆU CHO TÀI CHÍNH**
2. Tên môn học (tiếng Anh) : **DATA ANALYSIS FOR FINANCE**
3. Mã số môn học : **DAT709**
4. Trình độ đào tạo : **Đại học**
5. Ngành đào tạo áp dụng : **Các ngành đào tạo bậc đại học**
6. Số tín chỉ : **03**
7. Phân bổ thời gian : **150 giờ**
 - Trực tiếp : **45 giờ**
 - o Lý thuyết : **45 giờ**
 - o Thực hành : **0 giờ**
 - Trực tuyến (nếu có) : **không quá 30% tổng số giờ giảng của môn học.**
 - Tự học, tự nghiên cứu : **105 giờ**
 - Khác : **0 giờ**
8. Khoa quản lý môn học : **Bộ môn Toán Kinh tế**
9. Môn học trước : **Học máy**
10. Mô tả môn học

Môn học thuộc khối kiến thức chuyên ngành và là môn học bắt buộc trong chương trình đào tạo. Môn học cung cấp cách thức phân tích dữ liệu. Ngày nay, dữ liệu được thu thập một cách liên tục, rộng khắp dần dẫn đến sự gia tăng nhanh chóng về khối lượng dữ liệu cũng như sự đa dạng và hỗn tạp về định dạng. Chiết xuất thông tin có giá trị từ một khối lượng dữ liệu lớn, hỗn tạp như vậy là thách thức lớn không nhỏ cho các giải pháp phân tích dữ liệu cổ điển. Những điều này đòi hỏi phải có cách tiếp cận dữ liệu một cách hiệu quả, thông minh, nhanh chóng và chính xác hơn. Môn học khai phá dữ liệu cung cấp kiến thức nền tảng về các thuật ngữ nhập môn lĩnh vực

khai thác dữ liệu cho tài chính; các vấn đề liên quan đến việc tiền xử lí dữ liệu cho tài chính; phương pháp phân loại dữ liệu, gom cụm dữ liệu và khai thác mẫu; ứng dụng của khai thác dữ liệu trong trường hợp cụ thể, với dạng dữ liệu có cấu trúc phức tạp như dòng dữ liệu, dữ liệu chuỗi thời gian, dữ liệu mạng xã hội.

11. Mục tiêu và chuẩn đầu ra của môn học

11.1. Mục tiêu của môn học

| Ký hiệu mục tiêu | Mô tả mục tiêu | Nội dung CĐR CTĐT phân bổ cho môn học | Ký hiệu CĐR CTĐT (PLOn) |
|------------------|--|---|-------------------------|
| (a) | (b) | (d) | (e) |
| CO1 | Môn học trang bị cho sinh viên kỹ năng làm việc độc lập và làm việc theo nhóm để giải quyết các bài toán liên quan đến khoa học dữ liệu cho tài chính. | Có năng lực làm việc theo nhóm, truyền đạt hiệu quả trong điều kiện làm việc thay đổi và môi trường hội nhập quốc tế. | PLO3 |
| CO2 | Người học hiểu, đánh giá được các công cụ dành cho phân tích dữ liệu cho tài chính. Người học vận dụng thành thạo các công cụ trực quan hoá dữ liệu như Matplotlib, Numpy, Numba, Pandas, SciPy, Scikit-Learn, StatsModels, . . . để từ đó xây dựng mô hình phù hợp cho các bài toán về khoa học dữ liệu trong kinh doanh. | Có khả năng vận dụng thành thạo kiến thức nền tảng và chuyên sâu để giải quyết các vấn đề chuyên môn | PLO6 |

| | | | |
|------------|--|---|-------|
| CO3 | Người học có khả năng vận dụng cao các kiến thức về khoa học dữ liệu và phân tích dữ liệu một cách hệ thống để giải quyết các vấn đề chuyên môn trong lĩnh vực kinh doanh, tài chính - ngân hàng | Có kiến thức và kỹ năng về khoa học dữ liệu ứng dụng trong kinh doanh | PLO11 |
|------------|--|---|-------|

11.2. Chuẩn đầu ra của môn học (CĐR MH) và sự đóng góp vào chuẩn đầu ra của chương trình đào tạo (CĐR CTĐT)

| Ký hiệu CĐR MH (CLOi) | Nội dung CĐR MH | Mức độ theo thang đo | Đáp ứng mục tiêu môn học (COx) | Ký hiệu CĐR CTĐT (PLOn) |
|------------------------------|---|-----------------------------|---------------------------------------|--------------------------------|
| (a) | (b) | (c) | (d) | (e) |
| CLO1 | Ứng dụng thư viện Python vào xử lý dữ liệu, biến đổi dữ liệu thô thành dạng có thể sử dụng thống kê, dự đoán (dự báo), phân loại, thu giảm chiều hoặc gom cụm. | 3 | CO1, CO2 | PLO3 |
| CLO2 | Ứng dụng được các công cụ như Matplotlib Seaborn, ... vào trực quan hoá dữ liệu bằng các biểu đồ khác nhau. | 3 | CO2 | PLO6 |
| CLO3 | Ứng dụng được các mô hình thống kê và học máy sử dụng Statsmodels và sciKit-learn vào phân tích dữ liệu cho tài chính; PCA, t-SNE, UMAP, k-means vào thu giảm chiều dữ liệu và gom cụm dữ liệu. | 4 | CO1, CO3 | PLO6 |
| CLO4 | Ứng dụng các kiến thức về khoa học dữ liệu và học máy | 4 | CO2, | PLO11 |

| | | | | |
|-------------|---|---|----------|------|
| | vào mô hình hoá các bài toán thực tế. | | CO3 | |
| CLO5 | Xây dựng các giải pháp trực quan hoá, phân loại, dự báo, gom cụm dữ liệu ứng dụng trong lĩnh vực kinh doanh, tài chính ngân hàng. | 4 | CO2, CO3 | PLO6 |

11.3. Ma trận tích hợp giữa CDR MH (CLO_i), CDR CTĐT (PLO_n) và Chỉ số đánh giá kết quả thực hiện (PI_{n.k}):

| | PLO3 | | PLO6 | | PLO11 | |
|-------------|--------|--------|--------|--------|---------|---------|
| | PI 1.1 | PI 1.2 | PI 6.1 | PI 6.2 | PI 11.1 | PI 11.2 |
| CLO1 | 3 | 3 | | | | |
| CLO2 | | | 3 | 3 | | |
| CLO3 | | | 4 | 4 | | |
| CLO4 | | | | | 4 | 4 |
| CLO5 | | | 4 | 4 | | |

12. Phương pháp và hình thức dạy và học

12.1. Phương pháp dạy và học:

Nhằm giúp sinh viên đạt được các mục tiêu và chuẩn đầu ra của môn học, các chiến lược và phương pháp giảng dạy được áp dụng cụ thể như sau:

- Phương pháp Thuyết giảng (Lecture): Giảng viên trình bày tài liệu và giải đáp các thắc mắc của học viên.
- Phương pháp Bài giảng tương tác (Interactive lecture): Thuyết giảng trong đó có quãng thời gian nghỉ khoảng từ hai đến mười lăm phút để cho sinh viên thực hiện các hoạt động học tập (chẳng hạn như: trả lời các câu hỏi trắc nghiệm nhiều lựa chọn, giải quyết vấn đề, bài tập so khớp và điền vào chỗ trống trong tập ghi chú bài giảng, vấn đáp về một trường hợp điển hình nhỏ, bài tập chia sẻ suy nghĩ theo cặp hoặc thảo luận nhóm nhỏ) cứ sau 12 đến 20 phút.
- Phương pháp Ôn luyện (Recitation): Học sinh trả lời các câu hỏi về kiến thức và đọc hiểu.
- Phương pháp Thảo luận có định hướng (Directed discussion): Thảo luận trong lớp theo một bộ câu hỏi ít nhiều có tính trật tự mà giảng viên đặt ra để dẫn dắt học sinh đến những nhận thức hoặc kết luận nhất định hoặc để giúp họ đạt được một kết quả học tập cụ thể.

- Phương pháp giảng dạy thông qua Làm việc/học tập theo nhóm (Group work/learning): người học thực hiện hoạt động học tập hoặc tạo ra một sản phẩm theo nhóm nhỏ từ hai đến sáu người trong hoặc ngoài lớp học, dưới sự hướng dẫn cẩn thận của giảng viên.

12.2. Hình thức dạy và học:

Hình thức giảng dạy trực tiếp kết hợp trực tuyến, đảm bảo tổng thời gian giảng dạy trực tuyến không vượt quá 30% thời gian giảng dạy của cả môn học.

13. Quy định của môn học

- Tùy số sinh viên mà giảng viên quyết định số lượng thành viên nhóm học tập.
- Sinh viên tham dự lớp học phần phải tuân thủ quy tắc ứng xử của Nhà Trường; sinh viên phải đến lớp đúng giờ, đảm bảo thời gian học trên lớp, có thái độ nghiêm túc và chủ động, tích cực trong học tập, nghiên cứu.
- Sinh viên phải có ý thức chuẩn bị giáo trình, học liệu, máy tính cá nhân (khi cần thiết) để phục vụ quá trình học tập.

14. Tài liệu học tập

14.1. Tài liệu bắt buộc

[1] Scott, E. Page (2022). The model thinker: What you need to know to make data work for you. Basic Books.

14.2. Tài liệu tham khảo

[2] Lewinson, Eryk (2022). Python for Finance Cookbook: over 60 powerful recipes for effective financial data analysis, second edition. Packt Publishing.

[3] Makrides, A., Karagrigoriou, A., & Skiadas, C. H. (Eds.). (2020). Data Analysis and Applications 4: Financial Data Analysis and Methods. John Wiley & Sons.

14.3. Tài liệu khác

B. PHƯƠNG THỨC ĐÁNH GIÁ MÔN HỌC

1. Cơ cấu phương thức đánh giá môn học

| Loại hình đánh giá | Phương pháp đánh giá | Số lượng bài đánh giá | CĐR MH được đánh giá | Trọng số |
|-----------------------|--|-----------------------|-------------------------|----------|
| 1. Đánh giá quá trình | 1.1. Đánh giá tính chuyên cần | 3 | | 10% |
| | 1.2. Đánh giá quá trình 1: Kiểm tra trên phòng máy | 2 | CLO1, CLO2, CLO3 | 20% |
| | 1.3. Đánh giá quá trình 2: Bài tập lớn và bài tập về nhà | 2 | CLO1, CLO2, CLO3, CLO4, | 20% |

| | | | | |
|---------------------|---------------------------|---|------------------------|-----|
| | | | CLO5 | |
| 2. Đánh giá cuối kỳ | 2.1. Bài đánh giá cuối kỳ | 1 | CLO1, CLO2, CLO3, CLO4 | 50% |

2. Phương pháp và công cụ đánh giá

2.1. Đánh giá quá trình

2.1.1. Đánh giá tính chuyên cần

a. Hình thức đánh giá: sinh viên phát biểu, tự luận hoặc trắc nghiệm, trên lớp học hoặc trên các nền tảng học tập trực tuyến.

b. Thời điểm, cách thức thực hiện bài đánh giá:

Thời điểm: diễn ra trong suốt các buổi học của học phần.

Việc đánh giá kết quả học tập được thực hiện bằng đánh giá chuyên cần được thực hiện bằng phương thức điểm danh và ghi nhận quá trình tham gia học tập các nội dung của học phần. Điểm danh thực hiện trên danh sách lớp học phần chính thức do Trường cung cấp. Việc ghi nhận quá trình tham gia học tập các nội dung của học phần được thực hiện khi: (1) giảng viên mời đích danh sinh viên trả lời các câu hỏi hay phân công giải quyết các bài tập/chủ đề thảo luận (tham gia thụ động), (2) sinh viên tự giác (xung phong) trả lời các câu hỏi hay tham gia giải quyết các bài tập/chủ đề thảo luận (tham gia chủ động); sinh viên có tần suất tham gia trên 50% số buổi học với đa số lời đáp sát đáp án của vấn đề thì được xác định là chủ động tham gia rất tích cực vào quá trình học tập trên giảng đường.

c. Nội dung, kết cấu bài đánh giá

Nội dung của đánh giá chuyên cần gồm:

- tần suất hiện diện của sinh viên
- sự tham gia vào quá trình học tập tại giảng đường.

Kết cấu bài đánh giá: điểm danh, ghi nhận sự tham gia vào quá trình học.

d. Rubric đánh giá

| Tiêu chí đánh giá | CĐR | Trọng số | Thang điểm | | | | |
|----------------------------------|-----|----------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | | | Từ 0 đến 3,9 | Từ 4,0 đến 5,4 | Từ 5,5 đến 6,9 | Từ 7,0 đến 8,4 | Từ 8,5 đến 10 |
| Tần suất hiện diện của sinh viên | | 40% | Hiện diện giảng đường | Hiện diện giảng đường | Hiện diện giảng đường | Hiện diện giảng đường | Hiện diện giảng đường |

| | | | dưới 40% số buổi học | trên 50% số buổi học | trên 60% số buổi học | trên 70% số buổi học | trên 80% số buổi học |
|--|--|-----|---|---|---|--|--|
| Sự tham gia vào quá trình học tập tại giảng đường | | 60% | không tham gia vào quá trình thảo luận về nội dung bài học trong suốt học phần | tham gia thụ động vào quá trình thảo luận về nội dung bài học trong suốt học phần | tham gia chủ động vào quá trình thảo luận về nội dung bài học trong suốt học phần | chủ động tham gia tích cực vào quá trình thảo luận về nội dung bài học trong suốt học phần | chủ động tham gia rất tích cực vào quá trình thảo luận về nội dung bài học trong suốt thời học phần |

2.1.2. Đánh giá quá trình 1: Kiểm tra trên phòng máy

a. Hình thức đánh giá: sinh viên làm bài tập cá nhân lập trình trên phòng thực hành hoặc trên các nền tảng học tập trực tuyến.

b. Thời điểm, cách thức thực hiện bài đánh giá:

Thời điểm: một hoặc một phần buổi học do giảng viên và sinh viên thống nhất.

Đánh giá kết quả học tập bằng Bài kiểm tra cá nhân được thực hiện bằng cách tổ chức kiểm tra tập trung tại phòng thực hành có máy tính hỗ trợ làm bài: lập trình trên máy và được sử dụng tài liệu (trong trường hợp cần thiết, có thể thay thế bằng hình thức kiểm tra online; trong trường hợp này, giảng viên sẽ thông báo chi tiết đến sinh viên ít nhất 1 tuần trước ngày kiểm tra về thời gian giao đề, thời gian thực hiện bài kiểm tra, phương thức làm bài và nộp bài qua mạng internet).

c. Nội dung, kết cấu bài đánh giá

Nội dung đánh giá của Bài kiểm tra cá nhân là khối lượng kiến thức kiểm tra tương ứng với khối lượng kiến thức của tiến độ dạy học đã được quy định. Đề kiểm tra do giảng viên soạn, chịu trách nhiệm về chuyên môn;

Kết cấu bài đánh giá: tối thiểu có 02 câu hỏi; thời gian kiểm tra tối đa bằng thời gian thi hết học phần.

d. Rubric đánh giá

| Tiêu chí đánh giá | CĐR | Trọng số | Thang điểm | | | | |
|---|------|----------|----------------------------|--------------------------------|------------------------------------|---|---|
| | | | Từ 0 đến 3,9 | Từ 4,0 đến 5,4 | Từ 5,5 đến 6,9 | Từ 7,0 đến 8,4 | Từ 8,5 đến 10 |
| Gọi đúng thư viện để thao tác trên dữ liệu dạng bảng | CLO1 | 20% | Có nộp kết quả làm bài tập | Chương trình chạy không có lỗi | Chương trình chạy đúng 1/3 số test | Chương trình chạy đúng 2/3 số test case | Chương trình chạy đúng tất cả test case |
| Sử dụng đúng các thư viện hỗ trợ trực quan hoá dữ liệu | CLO2 | 40% | Có nộp kết quả làm bài tập | Chương trình chạy không có lỗi | Chương trình chạy đúng 1/3 số test | Chương trình chạy đúng 2/3 số test case | Chương trình chạy đúng tất cả test case |
| Sử dụng đúng các thư viện hỗ trợ các mô hình học máy cơ bản | CLO3 | 40% | Có nộp kết quả làm bài tập | Chương trình chạy không có lỗi | Chương trình chạy đúng 1/3 số test | Chương trình chạy đúng 2/3 số test case | Chương trình chạy đúng tất cả test case |

2.1.3. Đánh giá quá trình 2: Bài tập lớn và bài tập về nhà

a. Hình thức đánh giá: bài báo cáo và sản phẩm (chương trình máy tính) do một cá nhân hoặc nhóm không quá 5 thành viên thực hiện (khi cần thiết, giảng viên xem xét việc tăng thành viên của nhóm). Quy cách của báo cáo được giảng viên quy định trong 03 buổi học đầu tiên.

b. Thời điểm, cách thức thực hiện bài đánh giá:

Thời điểm: giảng viên và sinh viên thống nhất chọn thời điểm làm bài và nộp bài.

Giảng viên tổ chức cho sinh viên hình thành các nhóm bài tập lớn. Đề tài của bài tập lớn, thời gian và phương thức nộp được giảng viên thông báo nhóm sinh viên trực tiếp tại lớp hoặc qua email trong tuần học đầu tiên của học phần. Nhóm sinh viên tổ chức thực hiện bài tập trong thời gian tự học tại nhà; sau đó nộp bài tập đến giảng viên theo thời gian quy định. Giảng viên chấm, trả điểm

và cung cấp nhận xét cho các nhóm rút kinh nghiệm, đúc kết tri thức cần lĩnh hội.

c. Nội dung, kết cấu bài đánh giá

Nội dung đánh giá của hình thức đánh giá kết quả học tập bằng bài tập lớn theo nhóm và bài tập về nhà thực hiện bởi từng cá nhân, là khối lượng kiến thức của học phần đã quy định trong Đề cương này, thể hiện cụ thể qua các chủ đề của các bài tập

Kết cấu bài đánh giá: tối thiểu có 03 yêu cầu đối với một đề tài bài tập nhóm, tối thiểu 02 câu hỏi đối với bài tập về nhà.

d. Rubric đánh giá

Bài tập về nhà:

| Tiêu chí đánh giá | CDR | Trọng số | Thang điểm | | | | |
|---|------|----------|----------------------------|--------------------------------|------------------------------------|---|---|
| | | | Từ 0 đến 3,9 | Từ 4,0 đến 5,4 | Từ 5,5 đến 6,9 | Từ 7,0 đến 8,4 | Từ 8,5 đến 10 |
| Gọi đúng thư viện để thao tác trên dữ liệu dạng bảng | CLO1 | 20% | Có nộp kết quả làm bài tập | Chương trình chạy không có lỗi | Chương trình chạy đúng 1/3 số test | Chương trình chạy đúng 2/3 số test case | Chương trình chạy đúng tất cả test case |
| Sử dụng đúng các thư viện hỗ trợ trực quan hoá dữ liệu | CLO2 | 40% | Có nộp kết quả làm bài tập | Chương trình chạy không có lỗi | Chương trình chạy đúng 1/3 số test | Chương trình chạy đúng 2/3 số test case | Chương trình chạy đúng tất cả test case |
| Sử dụng đúng các thư viện hỗ trợ các mô hình học máy cơ bản | CLO3 | 40% | Có nộp kết quả làm bài tập | Chương trình chạy không có lỗi | Chương trình chạy đúng 1/3 số test | Chương trình chạy đúng 2/3 số test case | Chương trình chạy đúng tất cả test case |

Bài tập lớn:

| Tiêu chí đánh giá | CDR | Trọng số | Thang điểm |
|-------------------|-----|----------|------------|
|-------------------|-----|----------|------------|

| | | | Từ 0 đến 3,9 | Từ 4,0 đến 5,4 | Từ 5,5 đến 6,9 | Từ 7,0 đến 8,4 | Từ 8,5 đến 10 | |
|------------------------------|------|-----|---|--|---|--|---|--|
| Phân lập trình | CLO2 | 50% | Có nộp kết quả làm bài tập | Chương trình chạy không có lỗi | Chương trình chạy đúng 1/3 số test | Chương trình chạy đúng 2/3 số test case | Chươn g trình chạy đúng tất cả test case | |
| Cơ sở lý luận của báo cáo | CLO3 | 30% | Không sử dụng lý luận khoa học liên quan | Có sử dụng nhưng chưa đúng lý luận khoa học liên quan | Trình bày lý luận khoa học liên quan, nhưng chưa đủ sức thuyết phục | Sử dụng đúng lý luận khoa học liên quan, có sức thuyết phục | Sử dụng chính xác, rất thuyết phục về lý luận khoa học liên quan | |
| Văn phong khoa học | CLO4 | 10% | Hành văn tối nghĩa ở các phần nội dung chính | Hành văn lủng củng khiến cho người đọc khó có thể hiểu nội dung | Hành văn lủng củng nhưng người đọc vẫn có thể hiểu nội dung | Hành văn tốt, đôi khi có lỗi diễn đạt | Hành văn rõ ràng, mạch lạc | |
| Lỗi đạo văn | | | Trừ 100% điểm | | | | | |
| Hình thức của báo cáo | CLO5 | 10% | Không định dạng theo bất cứ tiêu chí nào | Có lỗi: không căn lề, không thống nhất định dạng | Có lỗi: Không định dạng toàn văn bản, | Có lỗi: thiếu đánh số trang; thiếu bìa hoặc trình bày bìa | Định dạng đúng tất cả các tiêu chí yêu cầu | |

| | | | | | | | |
|--|--|--|--|------------------------------|---------------------------------------|-----------------|--|
| | | | | đoạn văn, sai khô giấy | không thống nhất font chữ | sai quy định | |
|--|--|--|--|------------------------------|---------------------------------------|-----------------|--|

2.2. Bài đánh giá cuối kỳ

- Đề thi do Khoa/Bộ môn quản lý học phần ra đề hoặc trích xuất từ Ngân hàng câu hỏi thi/Ngân hàng đề thi của trường (nếu có)
- Hình thức đánh giá:

| Hình thức đánh giá | | Lựa chọn |
|--|----------------------------|----------|
| Vấn đáp | | |
| Tự luận | | |
| Trắc nghiệm khách quan | | |
| Trắc nghiệm khách quan kết hợp tự luận | | |
| Thực hành trên máy tính | | |
| Tiểu luận/đồ án/bài tập lớn (cá nhân) | Nộp bài không thuyết trình | |
| | Nộp bài có thuyết trình | |
| Tiểu luận/đồ án/bài tập lớn (nhóm) | Nộp bài không thuyết trình | X |
| | Nộp bài có thuyết trình | |

- Thời gian làm bài (tối thiểu – tối đa): từ 1 đến 7 ngày làm việc.
- Mô tả về kết cấu đề thi: đề tài dự án khoa học dữ liệu trong kinh doanh: Đề tài có ít nhất 03 yêu về việc vận dụng các kiến thức đã học để giải quyết một bài toán thực tế. Sinh viên làm dự án theo nhóm. Nội dung yêu cầu của dự án phải phản ánh khía cạnh thực tế của việc ứng dụng khoa học dữ liệu trong kinh doanh. Tùy theo sĩ số lớp mà giảng viên quyết định mỗi nhóm bao nhiêu sinh viên.
- Mô tả về phạm vi nội dung của đề thi: chương 2 đến chương 6.
- Rubric đánh giá/hoặc Ma trận đề: Đánh giá tương tự như phần 2.1.3 (Bài tập lớn).

| Tiêu chí đánh giá | CDR | Trọng số | Thang điểm | | | | |
|-------------------|------|----------|------------------------------|-------------------------|-------------------------|----------------------------------|--------------------------|
| | | | Từ 0 đến 3,9 | Từ 4,0 đến 5,4 | Từ 5,5 đến 6,9 | Từ 7,0 đến 8,4 | Từ 8,5 đến 10 |
| Phân lập trình | CLO2 | 50% | Có nộp kết quả làm bài | Chương trình chạy | Chương trình chạy | Chương trình chạy đúng 2/3 | Chươn g trình chạy |

| | | | | | | | |
|---------------------------|------|----------------------|--|---|---|--|--|
| | | | tập | không có lỗi | đúng 1/3 số test | số test case | đúng tất cả test case |
| Cơ sở lý luận của báo cáo | CLO3 | 30% | Không sử dụng lý luận khoa học liên quan | Có sử dụng nhưng chưa đúng lý luận khoa học liên quan | Trình bày lý luận khoa học liên quan, nhưng chưa đủ sức thuyết phục | Sử dụng đúng lý luận khoa học liên quan, có sức thuyết phục | Sử dụng chính xác, rất thuyết phục về lý luận khoa học liên quan |
| Văn phong khoa học | CLO4 | 10% | Hành văn tối nghĩa ở các phần nội dung chính | Hành văn lung củng khiến cho người đọc khó có thể hiểu nội dung | Hành văn lung củng nhưng người đọc vẫn có thể hiểu nội dung | Hành văn tốt, đôi khi có lỗi diễn đạt | Hành văn rõ ràng, mạch lạc |
| Lỗi đạo văn | | Theo quy chế đào tạo | | | | | |
| Hình thức của báo cáo | CLO5 | 10% | Không định dạng theo bất cứ tiêu chí nào | Có lỗi: không căn lề, không thống nhất định dạng đoạn văn, sai khổ giấy | Có lỗi: Không định dạng toàn văn bản, không thống nhất font chữ | Có lỗi: thiếu đánh số trang; thiếu bìa hoặc trình bày bìa sai quy định | Định dạng đúng tất cả các tiêu chí yêu cầu |

- Quy định về việc sử dụng tài liệu: sinh viên được sử dụng tài liệu nhưng cần ghi rõ nguồn trích dẫn/tham khảo.
- Các quy định khác về bài đánh giá: Không có.

C. NỘI DUNG CHI TIẾT GIẢNG DẠY

| Thời lượng (giờ) | Nội dung giảng dạy | CDR MH | Hoạt động Phương pháp dạy và học | Bài đánh giá | Học liệu |
|------------------|---|---|---|---|----------------------|
| (a) | (b) | (c) | (d) | (e) | (f) |
| 10 | <p>CHƯƠNG 1: GIỚI THIỆU VỀ PHÂN TÍCH DỮ LIỆU CHO TÀI CHÍNH</p> <p>Tổng quan về các loại model (hồi quy, phân lớp, gom cụm)</p> <p>Vòng đời của một dự án phân tích dữ liệu: CRISP-DM</p> <p>Thông kê và kiểm định giả thuyết trong phân tích dữ liệu</p> <p>Bài tập vận dụng</p> <p>Tổng kết</p> | <p>CLO1</p> <p>CLO2</p> <p>CLO3</p> <p>CLO4</p> <p>CLO5</p> | <p>Trực tiếp (offline) hay trực tuyến (online)</p> <p>GIẢNG VIÊN:</p> <p>Giới thiệu môn học (chú ý: giáo trình, công bố lịch kiểm tra giữa kỳ, giao đề tài bài tiểu luận nhóm).</p> <p>Giảng bài, vấn đáp về các khái niệm cơ bản.</p> <p>Giao sinh viên chuẩn bị bài Chương 2.</p> <p>Giao bài tập về nhà</p> <p>SINH VIÊN:</p> <p>Nghe giảng, tham gia thảo luận, phát biểu ý kiến xây dựng bài.</p> <p>Trả lời câu hỏi của giảng viên.</p> | <p>B.1.1</p> <p>B.1.2</p> <p>B.1.3</p> <p>B.2.1</p> | <p>[1], [2], [3]</p> |
| 10 | <p>CHƯƠNG 2: CÁC MÔ HÌNH TOÁN CHO PHÂN TÍCH DỮ LIỆU TÀI CHÍNH</p> <p>Phân phối xác suất và phân phối đa một (multimodal distributions)</p> <p>Mô hình sự không chắc chắn:</p> | <p>CLO1</p> <p>CLO2</p> <p>CLO3</p> <p>CLO4</p> <p>CLO5</p> | <p>Trực tiếp (offline) hay trực tuyến (online)</p> <p>GIẢNG VIÊN:</p> <p>Giới thiệu môn học (chú ý: giáo trình, công bố lịch kiểm tra giữa kỳ, giao đề tài bài tiểu luận nhóm).</p> <p>Giảng bài, vấn đáp về các khái niệm cơ</p> | <p>B.1.1</p> <p>B.1.2</p> <p>B.1.3</p> <p>B.2.1</p> | <p>[1], [2], [3]</p> |

| Thời lượng (giờ) | Nội dung giảng dạy | CĐR MH | Hoạt động Phương pháp dạy và học | Bài đánh giá | Học liệu |
|------------------|---|--------------------------------------|--|----------------------------------|------------------|
| | Entropy Dempster-Shafer Theory Markov model Học quần hợp (ensemble model) cho phân tích dữ liệu tài chính Ví dụ và bài tập minh hoạ Tổng kết | | bản. Giao sinh viên chuẩn bị bài Chương 2. Giao bài tập về nhà SINH VIÊN: Nghe giảng, tham gia thảo luận, phát biểu ý kiến xây dựng bài. Trả lời câu hỏi của giảng viên. | | |
| 5 | CHƯƠNG 3: MÔ HÌNH HỒI QUY Hồi quy tuyến tính Case study 1: xây dựng thuật toán hồi quy tuyến tính cho tập dữ liệu thực tế (ví dụ giá nhà ở thành phố Boston) Hồi quy logistic Case study 2: xây dựng thuật toán hồi quy logistic cho tập dữ liệu thực tế (ví dụ Banking dataset) Bài tập Tổng kết Các nội dung sinh viên tự nghiên cứu: Hồi quy tuyến tính đa biến Hồi quy logistic đa biến Ứng dụng logistic cho bài toán phân lớp đa | CLO1 CLO2 CLO3 CLO4 CLO5 | Trực tiếp (offline) hay trực tuyến (online) GIẢNG VIÊN: Kiểm tra bài cũ. Giảng bài mới. Giao sinh viên chuẩn bị bài Chương 3. Giải đáp các câu hỏi của sinh viên. Hướng dẫn sinh viên tự học, tự nghiên cứu. Giao bài tập về nhà và bài tập lớn Hướng dẫn làm bài tập lớn SINH VIÊN: Nghe giảng, tham gia thảo luận, phát biểu ý kiến xây dựng bài. - Trả lời câu hỏi của giảng viên. | B.1.1 B.1.2 B.1.3 B.2.1 | [1], [2], [3] |

| Thời lượng (giờ) | Nội dung giảng dạy | CDR MH | Hoạt động Phương pháp dạy và học | Bài đánh giá | Học liệu |
|------------------|--|---|---|---|---------------------------|
| | lớp | | | | |
| 5 | <p>CHƯƠNG 4: QUÁ KHỚP (OVERFITTING) & TƯ DUY PHÂN TÍCH DỮ LIỆU</p> <p>Quá khớp (overfitting) Các kỹ thuật khắc phục overfitting Giới thiệu về tư duy phân tích Những thang đo cho các mô hình hồi quy Những thang đo cho các mô hình phân loại Khung giá trị kỳ vọng Bài tập Tổng kết</p> <p>Các nội dung sinh viên tự nghiên cứu: Các kỹ thuật tiền xử lý dữ liệu: chuẩn hoá min-max, chuẩn hoá softmax làm đầy dữ liệu xử lý các điểm dữ liệu bất bình thường (outliers) Các kỹ thuật tăng cường dữ liệu</p> | <p>CLO1 CLO2 CLO3 CLO4 CLO5</p> | <p>Trực tiếp (offline) hay trực tuyến (online)</p> <p>GIẢNG VIÊN: Kiểm tra bài cũ. Giảng bài mới. Giao sinh viên chuẩn bị bài Chương 4. Giải đáp các câu hỏi của sinh viên. Hướng dẫn sinh viên tự học, tự nghiên cứu. Giao bài tập về nhà và bài tập lớn Hướng dẫn làm bài tập lớn</p> <p>SINH VIÊN: Nghe giảng, tham gia thảo luận, phát biểu ý kiến xây dựng bài. - Trả lời câu hỏi của giảng viên.</p> | <p>B.1.1 B.1.2 B.1.3 B.2.1</p> | <p>[1], [2], [3]</p> |
| 5 | <p>CHƯƠNG 5: MÔ HÌNH PHÂN LỚP (CLASSIFICATION)</p> <p>Ôn tập về xác suất Thang đo đồng dạng (similarity)</p> | <p>CLO1 CLO2 CLO3 CLO4</p> | <p>Trực tiếp (offline) hay trực tuyến (online)</p> <p>GIẢNG VIÊN: Kiểm tra bài cũ. Giảng bài mới.</p> | <p>B.1.1 B.1.2 B.1.3 B.2.1</p> | <p>[1], [2], [3]</p> |

| Thời lượng (giờ) | Nội dung giảng dạy | CĐR MH | Hoạt động Phương pháp dạy và học | Bài đánh giá | Học liệu |
|------------------|--|---|--|---|----------------------|
| | <p>k-láng giềng gần nhất</p> <p>Thuật toán Naïve Bayes</p> <p>Ước lượng hợp lý cực đại</p> <p>Case study 1: Áp dụng thuật toán phân lớp trên tập dữ liệu thực tế (ví dụ Irish Flowers)</p> <p>Học cây quyết định</p> <p>Case study 2: Xây dựng thuật toán phân lớp cho tập dữ liệu thực tế (ví dụ Banking dataset)</p> <p>Bài tập</p> <p>Tổng kết</p> <p>Các nội dung sinh viên tự nghiên cứu:</p> <p>Các thuật toán khác nếu bỏ giả thuyết “Naïve” trong tên thuật toán Naive Bayes</p> <p>Các độ đo tương tự phổ biến được dùng trong học máy</p> | CLO5 | <p>Giao sinh viên chuẩn bị bài Chương 5.</p> <p>Giải đáp các câu hỏi của sinh viên.</p> <p>Hướng dẫn sinh viên tự học, tự nghiên cứu.</p> <p>Giao bài tập về nhà và bài tập lớn</p> <p>Hướng dẫn làm bài tập lớn</p> <p>SINH VIÊN:</p> <p>Nghe giảng, tham gia thảo luận, phát biểu ý kiến xây dựng bài.</p> <p>- Trả lời câu hỏi của giảng viên.</p> | | |
| 5 | <p>CHƯƠNG 6: MÔ HÌNH GOM CỤM (CLUSTERING)</p> <p>Gom cụm dữ liệu</p> <p>Phân cụm k-means</p> <p>Case study: xây dựng thuật toán gom cụm cho tập dữ liệu thực tế (ví dụ phân cụm khách hàng của một siêu thị)</p> | <p>CLO1</p> <p>CLO2</p> <p>CLO3</p> <p>CLO4</p> <p>CLO5</p> | <p>Trực tiếp (offline) hay trực tuyến (online)</p> <p>GIẢNG VIÊN:</p> <p>Kiểm tra bài cũ.</p> <p>Giảng bài mới.</p> <p>Giao sinh viên chuẩn bị bài Chương 6.</p> <p>Giải đáp các câu hỏi của sinh viên.</p> <p>Hướng dẫn sinh viên tự học, tự nghiên</p> | <p>B.1.1</p> <p>B.1.2</p> <p>B.1.3</p> <p>B.2.1</p> | <p>[1], [2], [3]</p> |

| Thời lượng (giờ) | Nội dung giảng dạy | CĐR MH | Hoạt động Phương pháp dạy và học | Bài đánh giá | Học liệu |
|------------------|---|--------------------------------------|---|----------------------------------|------------------|
| | Bài tập Tổng kết Các nội dung sinh viên tự nghiên cứu: Các kỹ thuật thu giảm số chiều dữ liệu: LDA PCA t-SNE UMAP Ứng dụng thu giảm số chiều trong trực quan hoá dữ liệu và minh hoạ cho bài toán phân cụm. | | cứu. Giao bài tập về nhà và bài tập lớn Hướng dẫn làm bài tập lớn SINH VIÊN: Nghe giảng, tham gia thảo luận, phát biểu ý kiến xây dựng bài. - Trả lời câu hỏi của giảng viên. | | |
| 5 | CHƯƠNG 7: REINFORCEMENT LEARNING IN FINANCE Tổng quan về reinforcement learning Bài toán multi-armed bandit và ứng dụng trong tài chính Bài tập Tổng kết | CLO1 CLO2 CLO3 CLO4 CLO5 | Trực tiếp (offline) hay trực tuyến (online) GIẢNG VIÊN: Kiểm tra bài cũ. Giảng bài mới. Hướng dẫn sinh viên ôn tập. Thông báo một phần điểm quá trình Hướng dẫn làm bài tập lớn SINH VIÊN: Nghe giảng, tham gia thảo luận, phát biểu ý kiến xây dựng bài. Trả lời câu hỏi của giảng viên. Xem điểm, đề nghị điều chỉnh sai sót (nếu | B.1.1 B.1.2 B.1.3 B.2.1 | [1], [2], [3] |

| Thời lượng (giờ) | Nội dung giảng dạy | CDR MH | Hoạt động Phương pháp dạy và học | Bài đánh giá | Học liệu |
|------------------|--------------------|--------|--|--------------|----------|
| | | | có) và xác nhận điểm. Trao đổi với giảng viên, cả lớp về các nội dung trong học phần còn cần được giải đáp. | | |

TRƯỞNG BỘ MÔN


Nguyễn Minh Hải

NGƯỜI BIÊN SOẠN


Võ Đức Vĩnh

TRƯỞNG KHOA/BỘ MÔN


Nguyễn Minh Hải

HIỆU TRƯỞNG


Nguyễn Đức Trung

