|  |  |
| --- | --- |
| BỘ GIÁO DỤC VÀ ĐÀO TẠO  **TRƯỜNG ĐẠI HỌC KINH TẾ QUỐC DÂN** | **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**  **Độc lập – Tự do – Hạnh phúc** |

**ĐỀ CƯƠNG HỌC PHẦN TỔNG QUÁT**

*(Ban hành kèm theo quyết định số QĐ/ĐHKTQD, ngày tháng năm 2022)*

**1. THÔNG TIN TỔNG QUÁT (GENERAL INFORMATION)**

|  |  |
| --- | --- |
| ***- Tên học phần (tiếng Việt):*** | **Lập trình Phân tích dữ liệu** |
| ***- Tên học phần (tiếng Anh)*** | **Data science Programming** |
| ***- Mã số học phần*** | **CNTT1187** |
| ***- Thuộc khối kiến thức*** | **Kiến thức chuyên sâu** |
| ***- Số tín chỉ*** | **3 tín chỉ** (45 giờ tín chỉ; 50 phút/giờ) |
| ***+ Số tiết lý thuyết*** | **20** |
| ***+ Số tiết thảo luận/thực hành***  ***+ Số tiết tự học*** | **25**  **90** |
| ***- Các học phần tiên quyết:*** | **CNTT1116** |

**2. THÔNG TIN BỘ MÔN QUẢN LÝ VÀ GIẢNG VIÊN GIẢNG DẠY**

Bộ môn quản lý: **Bộ môn Công nghệ thông tin**

Địa chỉ: Phòng **1310 Nhà A1, Trường ĐH Kinh tế Quốc dân**

Giảng viên:

|  |  |  |
| --- | --- | --- |
| TS. Nguyễn Trung Tuấn | 0913537768 | tuannt@neu.edu.vn |
| ThS. Tống Thị Minh Ngọc | 0913002241 | ngoctm@neu.edu.vn |

**3. MÔ TẢ HỌC PHẦN (COURSE DESCRIPTIONS)**

Học phần Lập trình phân tích dữ liệu là học phần chuyên ngành, nhằm cung cấp các nội dung để sinh viên có thể hiểu rõ về dữ liệu và các đặc trưng cơ bản của dữ liệu. Sinh viên phân tích được bản chất dữ liệu, phân loại dữ liệu, đưa ra phương pháp phân tích dữ liệu phù hợp. Sinh viên có thể hiểu và sử dụng được phần mềm phân tích dữ liệu R (hoặc các phần mềm tương đương) để lập trình để thực hiện lập trình phân tích dữ liệu với những bài toán phân tích dữ liệu thông dụng, các bài toán phân tích dữ liệu nâng cao với các gói phần mềm tích hợp và kết nối dữ liệu từ các hệ quản trị cơ sở dữ liệu.

Học phần cung cấp cho người học các kiến thức về quy trình thu thập và xử lý dữ liệu, tiền xử lý dữ liệu, biểu diễn dữ liệu, một số phương pháp phân tích dữ liệu cơ bản và nâng cao. Học phần hướng dẫn cho người học các kỹ năng sử dụng phần mềm phân tích dữ liệu R để thực hiện các phép phân tích cơ bản, các phân tích nâng cao khác như phân cụm, cây quyết định, phân tích dữ liệu tài chính. Qua quá trình học tập, sinh viên sẽ rèn luyện thêm về kỹ năng làm việc độc lập, làm việc theo nhóm và kỹ năng trình bày vấn đề cũng như có ý thức đạo đức nghề nghiệp.

**4. TÀI LIỆU THAM KHẢO (LEARNING RESOURCES: COURSE BOOKS, REFERENCE BOOKS, AND SOFTWARES)**

**Giáo trình:**

[1] Nguyễn Văn Tuấn, Phân tích dữ liệu với R, Nhà xuất bản tổng hợp TP HCM, 2020.

**Tài liệu khác:**

[2] TS Ngô Văn Thứ, Giáo trình thống kê thực hành, Trường Đại học KTQD, 2012

[3] Data Mining - Concepts and Techniques - 3Rd - Han, Kamber and Pei - Morgan Kaufmann – 2012

[4] Phần mềm R và các tài liệu của phần mềm R tại địa chỉ https://cran.r-project.org.

**5. MỤC TIÊU HỌC PHẦN (COURSE GOALS)**

**Bảng 5.1. Mục tiêu học phần**

|  |  |  |  |
| --- | --- | --- | --- |
| **TT** | **Mô tả**  **mục tiêu học phần** | **CĐR (PLO) của CTĐT** | **Mức độ** |
| **[1]** | **[2]** | **[3]** | **[4]** |
| G1 | Học phần cung cấp cho người học các kiến thức về quy trình thu thập và xử lý dữ liệu, tiền xử lý dữ liệu, biểu diễn dữ liệu, một số phương pháp phân tích dữ liệu cơ bản và nâng cao | PLO 1.5A | 3 |
| G2 | Học phần hướng dẫn cho người học các kỹ năng sử dụng phần mềm phân tích dữ liệu R để thực hiện các phép phân tích cơ bản, các phân tích nâng cao khác như phân cụm, cây quyết định, phân tích dữ liệu tài chính | PLO 2.2.2A | 4U, 4U, 4U |
| G3 | Học phần góp phần phát triển khả năng tự làm việc độc lập và làm việc theo nhóm, hiểu đúng đắn về trách nhiệm thực thi công việc và đạo đức nghề nghiệp. | PLO 3.1.1, PLO 3.1.2, PLO 3.2.3 | 3, 3, 3 |

**6. CHUẨN ĐẦU RA HỌC PHẦN (COURSE LEARNING OUTCOMES)**

**Bảng 6.1. Chuẩn đầu ra học phần (CLO)**

| **Mục tiêu** | **CLOs** | **Mô tả CLOs\*** | **Mức độ đạt được\*\*** |
| --- | --- | --- | --- |
| **[1]** | **[2]** | **[3]** | **[4]** |
| G1 | CLO1.1 | Hiểu về dữ liệu và các phương pháp tiền xử lý, thể hiện dữ liệu | 3 |
| CLO1.2 | Hiểu về các phương pháp phân tích dữ liệu cơ bản và nâng cao | 3 |
| G2 | CLO2.1 | Thành thạo trong việc sử dụng phần mềm phân tích dữ liệu R | 4U |
| CLO2.2 | Thành thạo trong phân tích dữ liệu cơ bản | 4U |
| CLO2.3 | Thành thạo trong phân tích dữ liệu nâng cao: phân cụm, phân lớp, cây quyết định, phân tích tài chính | 4U |
| G3 | CLO3.1 | Có khả năng tự học tập, tự tìm hiểu tài liệu để hoàn thành các bài học và bài tập được giao. Có khả năng làm việc theo nhóm. | 3 |
| CLO3.2 | Có ý thức tổ chức kỷ luật lớp học, tuân thủ các nội quy lớp học, có tinh thần hợp tác và trách nhiệm cá nhân đối với các hoạt động nhóm và hoạt động của lớp học, có đạo đức nghề nghiệp trong phân tích dữ liệu. | 3 |

**7. ĐÁNH GIÁ HỌC PHẦN (COURSE ASSESSMENT)**

**Bảng 7.1. Đánh giá học phần**

|  |  |  |
| --- | --- | --- |
| **Hình thức đánh giá** | **CLOs** | **Tỷ lệ (%)** |
| **[1]** | **[4]** | **[6]** |
| Chuyên cần | CLO1.1  CLO1.2  CLO2.1  CLO2.2  CLO2.3 | 10% |
| Đánh giá quá trình | CLO2.1  CLO2.2  CLO2.3 | 40% |
| CLO2.1  CLO2.2  CLO2.3 |
| Đánh giá cuối kỳ | CLO2.1  CLO2.2  CLO2.3 | 50% |

**8. KẾ HOẠCH GIẢNG DẠY (LESSON PLAN)**

**Bảng 8.1. Kế hoạch giảng dạy**

|  |  |  |  |
| --- | --- | --- | --- |
| **Tuần** | **Nội dung giảng dạy\*** | **CLOs** | **Công cụ  đánh giá\*\*\*\*** |
| **[1]** | **[2]** | **[4]** | **[6]** |
| 1 | **Làm quen, giới thiệu học phần, logic của học phần và vai trò, vị trí của học phần trong CTĐT**  **Chương 1. Giới thiệu chung về dữ liệu**  1.1 Giới thiệu chung về dữ liệu  1.2 Giới thiệu về qui trình thu thập và xử lý thông tin | CLO1.1  CLO1.2 | Mức độ tham gia  Mức độ tương tác |
| 2 | **Chương 1. Giới thiệu chung về dữ liệu**  1.3 Phân loại dữ liệu  1.4 Tiền xử lý dữ liệu | CLO1.1  CLO1.2 | Mức độ tham gia  Mức độ tương tác  Chất lượng câu trả lời |
| 3 | **Chương 1. Giới thiệu chung về dữ liệu**  1.4 Tiền xử lý dữ liệu (tiếp theo)  1.5 Thể hiện dữ liệu | CLO1.1  CLO1.2  CLO3.1 | Mức độ tham gia  Mức độ tương tác  Chất lượng câu trả lời |
| 4 | **Chương 2. Ngôn ngữ R**  2.1 Giới thiệu R  2.2 Cài đặt R và các gói (package)  2.3 Văn phạm R | CLO1.2  CLO2.1  CLO3.1 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
| 5 | **Chương 2. Ngôn ngữ R**  2.4 Nhập và biên tập dữ liệu trong R  2.5 Sử dụng tính toán cơ bản trong R | CLO1.2  CLO2.1  CLO3.1 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
| 6 | **Chương 2. Ngôn ngữ R**  2.6 Trực quan hóa dữ liệu | CLO1.2  CLO2.1  CLO3.1 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
| 7 | **Chương 3. Kiểm định giả thuyết**  3.1 Mục đích, ý nghĩa của kiểm định  3.2 Kiểm định mối liên hệ hai biến định tính  3.3 Kiểm định trung bình 1 mẫu | CLO1.2  CLO2.1  CLO2.2  CLO3.1  CLO3.2 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
| 8 | **Chương 3. Kiểm định giả thuyết**  3.4 Kiểm định trung bình của hai mẫu độc lập  3.5 Kiểm định trung bình 2 mẫu cặp  3.6 Kiểm định trung bình nhiều mẫu | CLO1.2  CLO2.1  CLO2.2  CLO3.1  CLO3.2 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
| 9 | **Chương 4. Phân tích thống kê mô tả**  4.1 Lập bảng phân phối tần suất  4.2 Lập bảng tổng hợp nhiều biến  4.3 Lập bảng phân tích mối liên hệ giữa hai biến định tính | CLO1.2  CLO2.1  CLO2.2  CLO3.1  CLO3.2 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
|  | **Bài kiểm tra giữa kỳ (40%)** | CLO1.2  CLO2.1  CLO2.2 | Đề kiểm tra tự luận dạng lập trình và khai thác dữ liệu thực hiện trên máy tính |
| 11 | **Chương 4. Phân tích thống kê mô tả**  4.4 Phân tích hồi quy | CLO1.2  CLO2.1  CLO2.2  CLO3.1  CLO3.2 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
| 12 | **Chương 4. Phân tích thống kê mô tả**  4.5 Hồi quy logistics | CLO1.2  CLO2.1  CLO2.2  CLO3.1  CLO3.2 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
| 13 | **Chương 5. Các phương pháp phân tích nâng cao**  5.1 Phân lớp dữ liệu  5.2 Cây quyết định | CLO1.2  CLO2.1  CLO2.3  CLO3.1  CLO3.2 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
| 14 | **Chương 5. Các phương pháp phân tích nâng cao**  5.3 Phân cụm dữ liệu | CLO1.2  CLO2.1  CLO2.3  CLO3.1  CLO3.2 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
| 15 | **Chương 5. Các phương pháp phân tích nâng cao**  5.4 Phân tích dữ liệu tài chính | CLO1.2  CLO2.1  CLO2.3CLO3.1  CLO3.2 | Mức độ tham gia  Mức độ tương tác  Bài tập áp dụng |
| - | **Bài thi cuối kỳ** | CLO1.2  CLO2.1  CLO2.2  CLO2.3 | Theo yêu cầu bài thi |

**9. QUY ĐỊNH CỦA HỌC PHẦN (COURSE REQUIREMENTS AND EXPECTATION)**

**9.1. Quy định về điều kiện thi kết thúc học phần**

* Sinh viên được tham dự thi cuối kỳ/thi kết thúc học phần (50%) nếu có điểm chuyên cần (10%) đạt mức 5 điểm trở lên (thang 10).

**9.2. Quy định về tham dự lớp học**

* Sinh viên/học viên có trách nhiệm tham dự đầy đủ các buổi học. Trong trường hợp nghỉ học do lý do bất khả kháng thì phải có giấy tờ chứng minh đầy đủ và hợp lý. Mỗi buổi vắng mặt sẽ bị trừ 1 điểm đánh giá quá trình. Sinh viên vắng quá 3 buổi học dù có lý do hay không có lý do đều bị coi như không hoàn thành khóa học và phải đăng ký học lại.
* Sinh viên sẽ được cộng điểm cho mỗi lần phát biểu xây dựng bài, có thể bù đắp cho điểm chuyên cần, điểm kiểm tra.

**9.3. Quy định về hành vi lớp học**

* Học phần được thực hiện trên nguyên tắc tôn trọng người học và người dạy. Mọi hành vi làm ảnh hưởng đến quá trình dạy và học đều bị nghiêm cấm.
* Sinh viên phải đi học đúng giờ quy định. Sinh viên đi trễ quá 10 phút sau khi giờ học bắt đầu sẽ không được tham dự buổi học.
* Tuyệt đối không làm ồn, gây ảnh hưởng đến người khác trong quá trình học.
* Tuyệt đối không được ăn uống, nhai kẹo cao su, sử dụng các thiết bị như điện thoại, máy nghe nhạc trong giờ học.
* Máy tính xách tay, máy tính bảng chỉ được thực hiện vào mục đích ghi chép bài giảng, tính toán phục vụ bài giảng, bài tập, tuyệt đối không dùng vào việc khác.

|  |  |  |
| --- | --- | --- |
| **TRƯỞNG BỘ MÔN**  **TS. Phạm Xuân Lâm** | **TRƯỞNG KHOA/VIỆN**  **TS. Nguyễn Trung Tuấn** | **HIỆU TRƯỞNG**  **PGS.TS. Phạm Hồng Chương** |