

CHUYÊN ĐỀ NÂNG CAO LĨNH VỰC KHOA HỌC DỮ LIỆU

Tên khóa chuyên đề: MATHEMATICAL FOUNDATION FOR DATA SCIENCE
(CƠ SỞ TOÁN CHO KHOA HỌC DỮ LIỆU)

Thời lượng: 23 giờ

Giảng viên: PGS. TS. Nguyễn Thanh Bình

1. Mục tiêu khóa học

Củng cố kiến thức về các chủ đề chính tạo nền tảng cho việc tìm hiểu và nghiên cứu sâu hơn về phân tích dữ liệu, thống kê và các mô hình học máy. Nội dung gồm đại số tuyến tính, tính toán, xác suất thống kê giúp học viên áp dụng các khái niệm và kỹ thuật để phân tích và giải quyết vấn đề.

2. Kết quả đầu ra

Thông hiểu, nắm vững các kiến thức toán cơ bản dùng trong Khoa học dữ liệu.

3. Khung chương trình

- Củng cố các kiến thức cơ bản về toán (đại số tuyến tính, tính toán, giải tích cơ bản) cần cho quá trình phân tích dữ liệu nâng cao.
- Củng cố các kiến thức cơ bản, các kỹ thuật, phương pháp thống kê thường dùng trong khoa học dữ liệu.

4. Nội dung chi tiết

Buổi	Module	Topic
1	Linear Algebra (Đại số Tuyến tính)	Ma trận, các phép tính, tính chất quan trọng. Không gian Euclide, không gian vectơ.
2		Một số phép phân rã ma trận quan trọng, ánh xạ tuyến tính và các ứng dụng liên quan. Hệ phương trình tuyến tính, bài toán trị riêng trên ma trận.
3	Calculus (Giải tích)	Dãy số, giới hạn, hàm liên tục, tính chất hội tụ
4		Khái niệm về đạo hàm, tính khả vi, tích phân
5		Các phương pháp ước lượng, xấp xỉ hàm số.
6	Statistics (Thống kê)	Các khái niệm cơ bản của xác suất Các đặc trưng về độ phân tán dữ liệu Phân phối thường gặp.
7		Định lý thống kê: Law of Large Numbers, Central Limit Theorem. Kiểm định giả thuyết thống kê - Một số loại kiểm định thường dùng
8		Phân tích phương sai (ANOVA) Hồi quy tuyến tính - Hồi quy Logistic. Các tiêu chí để lựa chọn mô hình.