

CS102 – Máy học Thống kê

Chuẩn đầu ra môn học

- Nắm vững các thuật ngữ và một số thuật toán máy học thống kê.
- Biết cách lựa chọn thuật toán máy học phù hợp để giải quyết một bài toán thực tế.
- Biết cách đánh giá hệ thống máy học và các yếu tố ảnh hưởng đến nó.
- Sử dụng ngôn ngữ lập trình Python, các thư viện như keras, scikit-learn, numpy, pandas, scipy, matplotlib để hiện thực một hệ thống máy học giải quyết một bài toán cụ thể.

CS102 – Máy học Thống kê

Kiến thức học trước

- Tư duy lập trình và kiến thức lập trình căn bản
- Kiến thức cơ bản về đại số tuyến tính, giải tích

Nội dung chính

- Các thuật toán học có giám sát: Hồi quy (Regression) & Phân lớp (Classification)
- Tối ưu dùng Gradient Descent
- Tinh chỉnh các tham số (Tuning hyperparameters)
- Các thuật toán không giám sát: Phân nhóm (Clustering)
- Phát hiện bất thường (Anomaly Detection)
- Mô hình kết hợp (Ensemble Models)

Đánh giá môn học

- Quá trình: 20%
 - Các bài kiểm tra trắc nghiệm quá trình và vào tuần cuối (Nội dung: sách, slides bài giảng)
- Bài tập lập trình: 30% → hướng dẫn thực hành
- Đồ án cuối kỳ: 50%
 - Lựa chọn bộ dữ liệu khác những bộ dữ liệu trong chương trình học, xác định bài toán, lựa chọn thuật toán ML, thực hiện thí nghiệm, viết báo cáo, trình bày báo cáo (thuyết trình).
- KHÔNG thi giữa kì - cuối kì

Tài liệu tham khảo - Textbooks, MOOCs

- Books

- Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems [Aurélien Géron]
- Introduction to Machine Learning with Python: A Guide for Data Scientists