

Bài giảng, tài liệu

Bài 1: Tổng quan về Học máy (Machine Learning)

- slides: [Giới thiệu về Học máy](#) (31/8, 2.14 MB) ([Outline](#))
 - Sách: ISL Chương 2
 - [Sử dụng R](#) (Nguyễn Chí Dũng)
 - Học máy (HXX): Chương 1 & 2.
 - Các ứng dụng của Học máy
 - Xe không người lái (video)
 - Nhân tài Việt Nam tại Google làm về Học máy ([Lê Việt Quốc- BBC video](#))
 - "How Credit Card Companies Spot Fraud Before You Do" (article, U.S. News)
 - "Mining Electronic Records for Revealing Health Data" (article, New York Times)
 - "The Algorithm That's Hunting Ebola" (article, IEEE Spectrum)
- **tùy chọn**
 - Slide giới thiệu ngôn ngữ R của nhóm sinh viên ĐH Cần Thơ ([tại đây](#))
 - Bài giảng học máy của Nguyễn Nhật Quang ([slides](#))
 - Giới thiệu ngôn ngữ R của PGS. Đỗ Thanh Nghị ([tại đây](#)).
 - <http://www.listendata.com/2016/10/datatable-tutorial.html>
 - <http://www.r2d3.us/visual-intro-to-machine-learning-part-1/>
 - <https://github.com/ZuzooVn/machine-learning-for-software-engineers>
 - https://jalamar.github.io/visual-interactive-guide-basics-neural-networks/?utm_content=buffer79526&utm_medium=social&utm_source=facebook.com&utm_campaign=buffer
 - ML cơ bản của Nguyễn Xuân Khánh
 - ML cơ bản của Vũ Hữu Tiệp
 - Bài giảng R của UCLA Department of Statistics: [Basic_R_course](#), [Intermediate_R](#)
 - Tài liệu ngôn ngữ R của GS. Nguyễn Văn Tuấn ([tài liệu](#), Video)
 - tutorial code
 - Dữ liệu: Auto.data, Auto.csv
 - 10 thuật toán Học máy cần biết; R: Best Programming Language for Machine Learning

Bài 2: Học máy có giám sát

- slides: [Học máy có giám sát và Hồi quy tuyến tính](#) (2MB)
 - Chú ý: Chi tiết về hồi quy logic sẽ trình bày ở phía cuối khóa học
 - ISL: Chương 3 (linear regression)
- **tùy chọn** Nhắc lại về véc-tơ riêng/giá trị riêng ([xem tại đây](#)) (video, 10 phút)

Bài 3: Kỹ thuật kiểm tra chéo, hiệu chỉnh mô hình, mô hình thưa

- slides: [Kiểm tra chéo, Hiệu chỉnh mô hình và mô hình thưa](#) (1MB)
 - ISL: Mục 5.1 (cross-validation)
 - ISL: Mục 6.2 (lasso and ridge regression)

Bài 4: Hồi quy Logistic và Máy Véc-tơ hỗ trợ

- slides: [Hồi quy Logistic và Máy véc-tơ hỗ trợ](#) (1,473 MB, 17/9)
 - ISL: Mục 4.3 (logistic regression), [video giới thiệu mô hình hồi quy Logistic của GS. Tuấn](#)
 - ISL: Chương 9 (support vector machines)

- Textbook về SVM của TS. Nguyễn Đức Dũng ([tại đây](#))
- [Logistic của Vũ Hữu Tiệp](#)
- Andrew Ng's CS 229 course notes on SVMs available here for those interested in the mathematics behind SVMs.

Bài 5: Cây phân loại và hồi quy (CART)

- slides: [Cây phân loại và hồi quy](#) (0.77 MB)
 - ISL: Mục 8.1, 8.3.1, and 8.3.2 (CART sử dụng "tree")
 - [Why do Decision Trees Work?](#)
 - **Không bắt buộc** 10-phút R video minh họa cây CART: lệnh "rpart"
 - Mã nguồn R về CART and Dữ liệu thử nghiệm

Bài 6: Bootstrap và phương pháp học máy kết hợp (Ensemble Methods) (bao gồm Random Forests)

- slides: [Phương pháp Bootstrap và các phương pháp Học máy kết hợp](#) (0.77 MB, Boosting, Bagging, and Random Forests)
- Đọc trước
 - ISL: Mục 5.2 and 5.3.4 (The Bootstrap)
 - ISL: Mục 8.2.1, 8.2.3, and 8.3.4 (Boosting and Bagging)
 - ISL: Mục 8.2.2 and 8.3.3 (Random Forests)
 - <http://mlwave.com/kaggle-ensembling-guide/>

Bài 7: Học máy không giám sát

- slides: [Kỹ thuật phân cụm và Giảm chiều dữ liệu](#) (3 MB)
 - ISL: Chương 10 (Phương pháp phân cụm K-means 10.3.1)
 - ISL: Chương 10 (Phương pháp phân cụm phân cấp 10.3.2)
 - Học máy (HXH): Chương 6.
 - ISL: Chương 10 (PCA 10.2, clustering 10.3)
 - ISL: Mục 10.4 "Lab 1: Principal Component Analysis"
 - [Minh họa K-means](#) của tác giả Ông Xuân Hồng
 - [K-means của Vũ Hữu Tiệp](#)
 - các mã nguồn ví dụ (thảo luận trong bài giảng)
 - [Mã nguồn PCA](#) trong R và dataset tương ứng (chemical levels in olive oils from different regions)

ISL viết tắt tên sách *An Introduction to Statistical Learning with Applications in R*

ESL viết tắt tên sách tham khảo *The Elements of Statistical Learning*

Nội dung bài giảng tham khảo từ CME 250 của ĐH Stanford và IOM 530 của ĐH Southern California, Hoa Kỳ.