Những điều đã học được từ cuốn sách Machine Learning Mastery

# Chapter 4

Em học được cách đưa dữ liệu vào mô hình học máy.

1. Load file csv bằng thư viện tiêu chuẩn của Python
2. Load file csv bằng thư viện NumPy
3. Load file csv bằng thư viện Pandas

# Chapter 8

Em học được định nghĩa của chọn lọc đặc trưng cho mô hình máy học.

Em học được ba lợi ích của việc chọn lọc đặc trưng trước khi xây dựng mô hình là:

1. Giảm hiện tượng quá khớp (Overfitting): Dữ liệu dư thừa ít hơn đồng nghĩa với việc mô hình ít có khả năng đưa ra quyết định dựa trên nhiễu (noise).
2. Cải thiện độ chính xác: Dữ liệu gây hiểu lầm ít hơn sẽ giúp mô hình chính xác hơn.
3. Giảm thời gian huấn luyện: Dữ liệu ít hơn giúp thuật toán huấn luyện nhanh hơn.

Em học được các kỹ thuật chọn đặc trưng tự động.

1. Chọn đặc trưng đơn biến (Univariate Selection)
2. Loại bỏ đặc trưng đệ quy (Recursive Feature Elimination - RFE)
3. Phân tích thành phần chính (Principal Component Analysis - PCA)
4. Tầm quan trọng của đặc trưng (Feature Importance)

# Chapter 12

Em học được định nghĩa của spot-check.

Em học được 6 giải thuật máy học mà em có thể thử nghiệm nhanh các thuật toán phân loại.

1. Logistic Regression
2. Liner Discriminat Analysis
3. k-Nearest Neighbors
4. Naïve Bayes
5. Classification and Regression Trees
6. Support Vector Machine

Em học được cách thử nghiệm nhanh các thuật toán học máy trên một bài toán hồi quy

Em học được cách thử nghiệm nhanh bốn thuật toán hồi quy tuyến tính

Em học được cách thử nghiệm nhanh ba thuật toán hồi quy phi tuyến

# Chapter 16

Em học cách tinh chỉnh tham số (parameter tuning) cho các thuật toán học máy.

Em học được tầm quan trọng của việc tinh chỉnh tham số để cải thiện hiệu suất thuật toán.

Em học được cách sử dụng chiến lược tinh chỉnh tham số bằng Grid Search.

Em học được cách sử dụng chiến lược tinh chỉnh tham số bằng Random Search.