# Python Machine Learning

## Chenghsi Hsieh

1. Coda: 最后的一个拼图

2. Giving Computers the Ability to Learn from Data: 赋予电脑资料学习的能力

## Algorithms for Classification

1. 感知器的学习规则

2. 以物件导向程式实作感知器

3. 以鸢尾花资料训练感知器

4. 以物件导向程式实作适应线性神经元

5. 资料标准化与随机梯度下降

## Classifiers Using Scikit-Learn

1. 以套件实作感知器

2. 以物件导向程式实作逻辑回归

3. 以套件实作逻辑回归

4. 支持向量机 1

- 5. 支持向量机 2
- 6. 实作支持向量机
- 7. 决策树
- 8. 随机森林与 K 最近邻分类器

#### Data Preprocessing

1. 处理遗失值及类别资料

2. 资料切割、特征缩放及选取有意义的特征 4. 以随机森林评估特征的重要性

- 3. 以物件导向程序实作循序特征选择演算法

## Dimensionality Reduction

1. 主成分分析

2. 线性判别分析 1

3. 线性判别分析 2

- 4. 核的概念
- 5. 核主成分分析 1
- 6. 核主成分分析 2

## Model Evaluation and Hyperparameter Tuning

- 1. 使用管线精简工作流程
- 2. K 折交叉验证、学习曲线与验证曲线
- 3. 格状搜寻与混肴矩阵

- 4. 接收操作特征图 (ROC Curve)
- 5. 多元分类计分指标与资料不平衡

#### Ensemble

- 1. 整体学习简介
- 2. 以物件导向程式实作多数决分类器
- 3. 以多数决原理作预测并评估整体学习分类器
- 4. 微调整体学习分类器

- 5. 袋装法
- 6. 适应强化法 (AdaBoost)
- 7. 由依序极小化误差函数看 AdaBoost 与 Gradient Boost 简介

### Regression

- 1. 回归简介与探索波士顿房价数据集
- 2. 以物件导向程式及套件实作回归模型
- 3. 随机样本共识回归 (RANSAC)、脊回归与 5. 决策树回归与随机森林回归
- LASSO 回归
- 4. 以线性回归配适曲线

#### Clustering Analysis

- 1. K-means 与 K-means++
- 2. 转折判断法与轮廓图

- 3. 阶层树分群
- 4. 以密度为基础的分群 (DBSCAN)

#### Supplementary Materials

1. Materials

September 22, 2025