Machine Learning (A Probabilistic Perspective) NOTE

by

Masato Naka

Powered by OpenBook

§1 Chapter 1 Introduction

1.1 Chapter 1 Introduction

学べる点：いくつかのConceptとこの本の特徴

1.3.1

model based clustering: we fit a probabilistic model to the data, rather than running some ad hoc algorithm

1.3.2

dimensionality reduction: a lower dimensional subspace which captures the "essence" of the data

1.3.4 Matrix completion

matrix completion: the goal of **inputation** is to infer plausible values for the missing entries

1. Collaborative Filtering
2. Market Basket Analysis

frequent itemset mining creates association rules

probabilistic approach: better predictive accuracy than association rules, less interpretable

Data Mining : there is more emphasis on interpretable models

Machine Learning: there is more emphasis on accurate models

1.4.1

Parametric model: does the model have a fixed number of parameters?:

Non-parametric model: does the number of paramters grow with the amount of training data?

1.4.2 non-parametric model: K-nearest neighbors

memory-based learningor **instance-based learning**

1.4.3 the curse of dimensionality

Hastie et al. 2009

1.4.4 Parametric models for classification and regression

1.4.5 Linear regression

non-linear function of the inputs,

basis function expansion

1.4.6 Logistic Regression

Logistic Regression : linear regression + **Bernoulli** distribution + pass through the **sigmoid** function

1.4.9 No Free Lunch Theorem

協賛企業募集