

# ICT リテラシー (情報技術論) B

-- 第 08 回 : 教師なし学習の代表的な手法 --  
(k平均法, 主成分分析)

栗野 俊一

講義内容の静止画・動画での撮影、及び SNS 等への転載を固く禁じます

2025/11/17 ICT リテラシー (情報技術論) B

# 伝言

---

## 私語は慎むように !!

□ 席は自由です

- できるだけ前に詰めよう
- コロナ対策のために、ソーシャルディスタンスをたもう

□ 色々なお知らせについて

- 栗野の Web Page に注意する事

<http://edu-gw2.math.cst.nihon-u.ac.jp/~kurino>

- google で「kurino」で検索

# 前回(第07回)の復習

---

ICTリテラシー(情報技術論)B

## 前回(第07回)の復習

講義内容の静止画・動画での撮影、及びSNS等への転載を固く禁じます

# 前回(第07回)の復習

---

## □ 前回(第07回)の復習

### ○ 講義：教師あり学習の代表的な手法

- ▷ k近傍法：予想したい点の近い k 点の結果から、予測する
- ▷ サポートベクターマシン：二つの集合を分離する(超)平面で分別する

### ○ 演習：python で、機械学習(1)

- ▷ Google Colab (jupyter notebook, Python)
- ▷ 機械学習の例（線形回帰、ロジスティック回帰）

# 今週(第08回)の概要

---

ICTリテラシー(情報技術論)B

## 今週(第08回)の概要

講義内容の静止画・動画での撮影、及びSNS等への転載を固く禁じます

# 今週(第08回)の予定

---

## □ 今週(第08回)の予定

- 講義: 教師なし学習の代表的な手法 (Text p.89, 7.2 節)
  - ▷ 教師なし学習の代表的な手法の一手法である k平均法, 主成分分析について学ぶ

# 今週(第08回)の目標

---

- 今週(第08回)の目標
  - k平均法, 主成分分析 の具体的なアルゴリズムを学ぶ

# 今週 (第 08 回)

---

## □ 前回 (第 07 回) の課題

- 振り返り課題-07
- 小テスト-07

## □ 今週 (第 08 回) の課題

- 振り返り課題-08
  - ▷ 提出期限は 1 週間
- 小テスト-08

# 教師なし学習の代表的な手法

---

ICT リテラシー (情報技術論) B

## 教師なし学習の代表的な手法

講義内容の静止画・動画での撮影、及び SNS 等への転載を固く禁じます

# 教師なし学習の代表的な手法

---

- 教師無し学習：学習データに、正解 Label がない
  - 予測しようとしている目的変数が明確でない(明確にできない...)場合に用いる
    - ▷ 学習データが持つであろう特徴を学習する
    - ▷ 特徴に基いて、予測をする
- ありなしの比較
  - 教師あり :  $y=f(x) = F(x,p)$  の  $f$  を求めるために  $p$  を定める
    - ▷  $y$  (正解 Label) の性質は解らないが、具体的な  $y$  は得られている
    - ▷ 予想したい値が具体的にイメージできる
  - 教師なし : 条件  $P$  を与え、 $P(y)$  を満す、 $Y=\{y|P(y)\}=\{y|y=f(x)\}$  を求める
    - ▷  $y$  (正解 Label) はないが、 $y$  が満して欲しい性質  $P$  が与えられている
    - ▷ 予想したい値は、グループ分け(他の要素との比較)のラベルに過ぎない
- クラスタ分類：集団を「似た者同士」のグループに分割する
  - どのグループに所属するかは意味がない
  - 新しい要素が、所属するグループが分れば、(同じグループに所属する)似た要素が見付かる

# k平均法, 主成分分析

---

ICT リテラシー (情報技術論) B

## k平均法, 主成分分析

講義内容の静止画・動画での撮影、及び SNS 等への転載を固く禁じます

# k平均法

---

## □ k平均法 : k-menas

○ 発想 : 対象(学習データの母集団)は、k 種類に分類できるとする

▷ 仮定 : 個々の分類の集団には中心(平均)があり、その集団要素は中心に近い

▷ 予測 : 集団の中心が得られれば、新しいデータの所属する要素は中心に近い

○ 手法 : k 個の中心をもとめる

▷ 仮の中心を k 個用意する

▷ 仮の中心で分類し、中心を、分類した集団の平均として再計算(修正)

▷ 仮の中心が変化しなくなった(収束した)ら、終了

○ 学習の対象 ( 内部状態 -> k 個の中心 )

▷ 中心による分類に「矛盾が生じない」ようにする (自らに相応しい対象を選択する)

○ 応用 : クラスタ分類

# 主成分分析

---

## □ 主成分分析

○ 発想：対象(学習データの母集団)は、幾つかの無関係な要素(主成分)の組み合せ

▷ 仮定：値を説明する要素に優劣(主となる成分)がある

▷ 予測：個々の要素の成分値が分れば、その要素を説明できる

○ 手法：

▷ 全体を最もよく説明する成分を一つみつける(回帰)

▷ それ以外の成分で、同じことを繰り返す

○ 学習の対象 ( 内部状態 -> 成分[の方向] )

▷ 少ない要素で、サンプルの状態をよく説明できる(誤差が少ない)

○ 応用：次元の縮小

# 推薦

---

## □ 応用 (推薦)

- 新しい顧客に対し、その好みとなる商品を提案する
  - ▷ 顧客が既存の客に似ていれば、その客の購入した製品を推薦する

## □ 協調フィルタリング

- 既存の対象をクラス分類 (教師なし学習)
  - ▷ 新しい対象を AI によってクラス分類
  - ▷ 同じクラスに所属する他の対象の性質を提示する

おしまい

---

ICT リテラシー (情報技術論) B

おしまい

講義内容の静止画・動画での撮影、及び SNS 等への転載を固く禁じます