

# ICT リテラシー (情報技術論) B

-- 第 11 回：手法の評価 --

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

2022/11/28 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とう

### □ 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 10 回) の復習

---

ICT リテラシー (情報技術論) B

# 前回 (第 10 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 10 回) の復習

---

## □ 前回 (第 10 回) の復習

- ニューロンの学習：出力の誤差を小さくするように、重みを調整する
  - ▶ 勾配降下法：微分可能な関数  $y=f(x)$  の極小値を求める方法 (完全ではない)
- バックプロパゲーション ( 誤差逆伝播法 )
  - ▶ ニューロンの出力の誤差から、入力側の誤差を計算
  - ▶ 深い層の誤差を浅い層の誤差に伝播させ、それでニューロンの学習をする

# 今週 (第 11 回) の概要

---

ICT リテラシー (情報技術論) B

## 今週 (第 11 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 11 回) の予定

---

## □ 今週 (第 11 回) の予定

### ○ 講義：手法の評価 (Text p.93, 7.5 節)

- ▶ Deep Learning で作られた AI の評価
- ▶ 学習結果の評価
- ▶ 混同行列
- ▶ 評価指標

# 今週 (第 11 回) の目標

---

## □ 今週 (第 11 回) の目標

- 学習結果の良否を判定するための基準について学ぶ。

# 今週 (第 11 回) の課題

---

- 前回 (第 10 回) の課題
  - Web Class「小テスト-10」
- 今週 (第 11 回) の課題
  - Web Class「小テスト-11」



# 手法の評価

---

## ICT リテラシー (情報技術論) B

# 手法の評価

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

# Deep Learning で作られた AI の評価

---

## □ Deep Learning の得失

- 学習データがあれば、AI を自動的に作ってくれる

  - ▶ 100% を保障するものではない

- Model の作り方は試行錯誤

  - ▶ 結果も最適とは限らない ( 勾配降下法の制限 )

  - ▶ 最初の(乱数で決める)重みの初期値によって、できが左右される

- 基本、大きな Model にしておけば、無難だが、時間がかかる

## □ 学習結果を評価する必要がある

# 学習結果の評価

---

## □ 過学習：学習データに適合し過ぎる学習をしてしまうこと

○ 例: 学習データはサンプリング (母集団の性質を学習したい...)

▶ 母集団に共通な性質ではなく、学習データ固有の性質を学習する

○ 例: 学習データは誤差を含む (滑らかな関数が、ぎざぎざになる)

▶ 目的の関数ではなく、それに誤差を再現する形で学習する

## □ 過学習の計測

○ 学習に利用するデータと独立な評価手段が必要

▶ 学習結果を評価するデータは、学習そのものには使わない

## □ 学習データの分割と評価用データ

○ 学習データを、学習用と評価用に分けておく (例: 8:2)

# 混同行列

---

## □ 混同行列：2 分類 (有る/無し) を区別する AI の判定結果の分類

○ 判定は正しかったのか？

▶ 正解か不正解の区別 ( True/False )

○ 期待しているものを予測したか？

▶ 期待していた予測か否かの区別 ( Positive/Negative )

▶ AI の予想結果が期待していた内容かどうかによって、対応が異なる

○ TP ( True/Positive ) : 期待したものを正しく判定

○ TN ( True/Negative ) : 期待していないものを正しく判定

○ FP ( False/Positive ) : 期待したものを間違って判定

○ FN ( False/Negative ) : 期待していないものを間違って判定

## □ 同じ間違った判定でも、FP と FN では意味合いが異なる

○ 例：癌検診 (癌を見付けたい)

▶ FP : 癌でなかったのに、癌だと予測 [人騒がせだが、真実を知ってホッとする]

▶ FN : 癌だったのに、癌でないと予測 [手遅れになり、人死が...]

○ 例：マーケットリサーチ (顧客の候補を探したい)

▶ FP : 買う気のない人を顧客と予測 [宣伝しても無駄になる]

▶ FN : 買う気のある人を顧客でない判断 [機会損失だが、現実的な損はない]

# 評価指標

---

## □ 評価指標：混同行列の結果を利用した評価指標

○ 正解率 (Accuracy) : 予測内容と無関係に正答の割合 :  $(TP+TN)/(TP+TN+FP+FN)$

▶ 評価の基本となる指標

○ 適合率 (Precision) : 期待した予想の内の正答の割合 :  $TP/(TP+FP)$

▶ 期待していない結果は、確実に予測して欲しい (結果を信じると無駄になる)

○ 再現率 (Recall) : 期待した事実の内の正答の割合 :  $TP/(TP+FN)$

▶ 期待していない事実は、確実に予測して欲しい (発見できないは困る)

○ F 値 : 適合率と再現率の調和平均 :  $(2 \times \text{適合率} \times \text{再現率}) / (\text{適合率} + \text{再現率})$

▶ 適合率と再現率を総合的に評価

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