RITSUMEIKAN

機械学習 演習問題 解答例

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題の解答例

- これは講義中に実施した演習問題の解答例です
  - あくまでも「解答例」ですので、他にも正しい解答があることを理解した上で、この資料を参考にしてください

RITSUMEIKAN

機械学習 第１回 機械学習とは

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

- ビッグデータを用いた機械学習の応用事例を最低１つ考えなさい。

例） 翻訳システム
２ケ国語で記述された文章データを大量に集めて、そのビッグデータから翻訳のルール（規則）を機械学習によって導き出す。

演習問題 解答例

- ビッグデータを用いた機械学習の応用事例を最低１つ考えなさい。

感染症の拡大予想

もし
①とても優秀な機械学習の技術が存在する
②誰が、いつ、どこで、どのような症状にかかったのか？
という情報が記録されたビッグデータを迅速に集められる
が実現していれば、新型コロナウィルス（COVID-19）の流行をもっと早く予想できたかも？

演習問題

- 教師あり学習の識別問題と回帰問題のそれぞれについて応用事例を1つずつ考えなさい。
  - 講義スライドに記載されていない事例を考えてください。

演習問題 解答例

- 教師あり学習の識別問題と回帰問題のそれぞれについて応用事例を1つずつ考えなさい。
  - 講義スライドに記載されていない事例を考えてください。

識別：迷惑メールの判定（スパム検知など）
送信元のアドレスや文章などの情報をもとに「送られてきたメールが迷惑メールか？」を識別する。

回帰：株価の予測
過去の株価情報や経済状況をもとに、未来の株価を予測する。
高精度な予測システムを構築できれば、ビジネスチャンスに繋がるかも…

演習問題

- 教師なし学習のモデル推定とパターンマイニングのそれぞれについて応用事例を1つずつ考えなさい。

演習問題 解答例

- 教師なし学習のモデル推定とパターンマイニングのそれぞれについて応用事例を1つずつ考えなさい。

モデル推定：開発商品の改善点の抽出
商品に対する口コミ（人伝で伝達される評判や評価に関する情報）の文書をクラスタリングして、典型的な不満や要望を抽出する。

パターンマイニング：効果的な授業設計
過去の学生たちの学習状況やテスト成績をもとに、学生たちの成績を向上させる授業内容や課題を検討する。

演習問題

- 中間的学習の半教師あり学習と強化学習の
  それぞれについて応用事例を1つずつ考えなさい。

演習問題 解答例

- 中間的学習の半教師あり学習と強化学習の
  それぞれについて応用事例を1つずつ考えなさい。

半教師あり学習：口コミのPN判定（Positive-Negative判定）

ネットワーク上の口コミのPN判定では、正解付きデータを収集するのは大変（収集できても少量）だが、クローラプログラムを使えば、自動的に大量の教師なしデータを収集できる。ある製品に対する口コミのPN判定は、半教師あり学習が有効。

強化学習：囲碁の最善の一手（Alpha Goなど）

相手に勝利するまでの過程は様々です。強化学習は、盤面という状態に応じて、将来の得られる報酬が1番高い（勝利の可能性が高い）一手を導きます。

機械学習 第２回 機械学習の基本的な手順

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

※ヒント：標準化の方法

$$\text{標準化後の値} = \frac{\text{もとの値} - \text{その次元の平均値}}{\text{その次元の標準偏差}}$$

- 以下5名分の身長と体重のデータについて、身長と体重の平均値を0、標準偏差を1となるように標準化せよ。

標準化前

番号	身長 [cm]	体重 [kg]
1	160	54
2	166	58
3	168	60
4	172	62
5	184	66
平均		
標準偏差		

標準化後

番号	身長	体重
1		
2		
3		
4		
5		
平均	0	0
標準偏差	1	1

演習問題 解答例

- 以下5名分の身長と体重のデータについて、身長と体重の平均値を0、標準偏差を1となるように標準化せよ。

標準化前

番号	身長 [cm]	体重 [kg]
1	160	54
2	166	58
3	168	60
4	172	62
5	184	66
平均	170	60
標準偏差	8	4

標準化後（標準化後の特徴に単位なし）

番号	身長	体重
1	-1.25	-1.5
2	-0.5	-0.5
3	-0.25	0
4	0.25	0.5
5	1.75	1.5
平均	0	0
標準偏差	1	1

演習問題

- ある画像認識技術でバナナを検出することを考える。
  - 100個のリンゴ、30個のバナナ、70個のオレンジの画像データに対して、以下の識別結果が得られた。

1. 識別結果から混同行列を求めよ。
2. 混同行列から正解率、精度、再現率、F値を計算せよ。
   また、この認識技術の性能について考察せよ。

	識別結果	
	バナナ	バナナ以外
りんご	20	80
バナナ	15	15
オレンジ	10	60

	予測＋	予測－
正解＋		
正解－		

正解率

精度

再現率

F値

演習問題 解答例

- ある画像認識技術でバナナを検出することを考える。
  - 100個のリンゴ、30個のバナナ、70個のオレンジの画像データに対して、以下の識別結果が得られた。

1. 識別結果から混同行列を求めよ。

	識別結果	
	バナナ	バナナ以外
りんご	20	80
バナナ	15	15
オレンジ	10	60

混同行列

	予測＋	予測－
正解＋	(TP) 15	(FN) 15
正解－	(FP) 20+10=30	(TN) 80+60=140

演習問題 解答例

- ある画像認識技術でバナナを検出することを考える。
  - 100個のリンゴ、30個のバナナ、70個のオレンジの画像データに対して、以下の識別結果が得られた。

2. 混同行列から正解率、精度、再現率、F値を計算せよ。
   また、この認識技術の性能について考察せよ。

混同行列

	予測＋	予測－
正解＋	(TP) 15	(FN) 15
正解－	(FP) 30	(TN) 140

$$\text{正解率} = \frac{\mathrm{TP+TN}}{\mathrm{TP+FN+FP+TN}} = \frac{15+140}{15+15+30+140} = 0.775$$

$$\text{精度} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} = \frac{15}{15+30} = \frac{1}{3}$$

バナナと識別した結果の内、33.3%しか正しくない

$$\text{再現率} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} = \frac{15}{15+15} = 0.5$$

バナナのデータを入力しても正しく識別できたものは50%

$$\text{F値} = 2 \times \frac{\text{精度} \times \text{再現率}}{\text{精度} + \text{再現率}} = 0.4$$

正解率が0.775（77.5%）なので、一見、良い結果のように見えるが、精度や再現率などの結果から、正確にバナナを検出できていない

機械学習 第３回 識別 –概念学習-

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

- 以下の表は、喫茶店の飲み物の注文記録である。
  - FIND-Sアルゴリズムを用いて、「冷たい飲み物」を注文する客の仮説を立てよ。

喫茶店の飲み物の注文記録

	年齢	天気	気温	時間帯	飲み物
1	若年	雨	低	昼	冷
2	高齢	晴	高	朝	熱
3	若年	晴	低	昼	冷
4	高齢	雨	低	夜	熱
5	若年	曇	高	昼	熱
6	若年	曇	低	朝	冷

演習問題 解答例

- 冷たい飲み物を抽出する

	年齢	天気	気温	時間帯	飲み物
1	若年	雨	低	昼	冷
3	若年	晴	低	昼	冷
6	若年	曇	低	朝	冷

- 順番に共通項を見つける
  1. 「客1」と「客3」の共通点
     - 年齢＝若年、気温＝低、時間帯＝昼
  2. 「1.」と「客6」の共通点
     - 年齢＝若年、気温＝低

答え：年齢＝若年 ∧ 気温＝低い

演習問題

- 以下は、ある人がゴルフに参加するかを気象条件を特徴として決定するデータである。

	天候	気温	湿度	風	ゴルフ
1	晴	高温	多湿	なし	不参加
2	晴	高温	多湿	あり	不参加
3	曇	高温	多湿	なし	参加
4	雨	適温	多湿	なし	参加
5	雨	低温	標準	なし	参加
6	雨	低温	標準	あり	不参加
7	曇	低温	標準	あり	参加
8	晴	適温	多湿	なし	不参加
9	晴	低温	標準	なし	参加
10	雨	適温	標準	なし	参加
11	晴	適温	標準	あり	参加
12	曇	適温	多湿	あり	参加
13	曇	高温	標準	なし	参加
14	雨	適温	多湿	あり	不参加

- このデータのエントロピー$E(D)$を求めなさい。
- 各特徴量の情報獲得量を計算して、最初のデータ分割に用いる特徴を選択しなさい。

対数計算を用いるので、関数電卓などを活用してください。

演習問題 解答例

- エントロピー$E(D)$
  - 全14事例中、参加が9事例、不参加が5事例なので

$$E(D) = -\frac{9}{14}\log_2\frac{9}{14} - \frac{5}{14}\log_2\frac{5}{14}$$

$$\fallingdotseq -0.643 \times (-0.637) - 0.357 \times (-1.485)$$

$$\fallingdotseq \mathbf{0.94}$$

演習問題 解答例

- 情報獲得量

$$\mathrm{Gain}(D, 天候) = E(D) - \frac{5}{14}E(晴) - \frac{4}{14}E(曇) - \frac{5}{14}E(雨)$$

$$\fallingdotseq 0.94 - 0.357 \times \left(-\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5}\right)$$

$$-0.286 \times \left(-\frac{4}{4}\log_2\frac{4}{4} - \frac{0}{4}\log_2\frac{0}{4}\right) - 0.357 \times \left(-\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5}\right)$$

$$\fallingdotseq 0.94 - 0.357 \times 0.971 - 0.286 \times 0 - 0.357 \times 0.971$$

$$\fallingdotseq 0.247$$

同じ手順で、気温、湿度、風の情報獲得量を求めると、

$\mathrm{Gain}(D, 気温) \fallingdotseq 0.029$、$\mathrm{Gain}(D, 湿度) \fallingdotseq 0.151$、$\mathrm{Gain}(D, 風) \fallingdotseq 0.048$

最も情報獲得量の多い特徴は「天候」なので
この「天候」を特徴（質問）としてデータを分割すれば良い

機械学習 第４回 識別 –統計的手法-

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

- 外見では区別のつかない箱Aと箱Bがある
  - 箱A：白玉が1個、黒玉が3個入っている
  - 箱B：白玉が4個、黒玉が1個入っている

1. 箱A、箱Bのいずれかから白玉を取り出す確率を求めよ。それぞれの箱の選び方は等確率とする。
2. 箱A、箱Bのいずれかから玉を取り出すと、白玉であった。この白玉が箱A、箱Bから取り出された確率をそれぞれ求めよ。それぞれの箱の選び方は等確率とする。
3. 箱A、箱Bの選ばれる確率が 9 : 1 であったとき、1. と 2.で求めた確率は、どのように変化するか？

演習問題 解答例

1. 箱A、箱Bのいずれかから白玉を取り出す確率を求めよ。それぞれの箱の選び方は等確率とする。
   - 「箱Aを選ぶ」かつ「箱Aから白玉が出る」の確率：$\frac{1}{8}$
     - 箱Aを選ぶ確率を$P(A)$、箱Aから白玉が出る確率を $P(白|A)$とすると

$$P(A)P(白|A) = \frac{1}{2}\cdot\frac{1}{4} = \frac{1}{8}$$

   - 「箱Bを選ぶ」かつ「箱Bから白玉が出る」：$\frac{2}{5}$
     - 箱Bを選ぶ確率を$P(B)$、箱Bから白玉が出る確率を $P(白|B)$とすると

$$P(B)P(白|B) = \frac{1}{2}\cdot\frac{4}{5} = \frac{2}{5}$$

   - 白玉を取り出す可能性を全て加算すれば良いので

$$P(白) = P(A)P(白|A) + P(B)P(白|B) = \frac{1}{8} + \frac{2}{5} = \mathbf{\frac{21}{40}}$$

演習問題 解答例

2. 箱A、箱Bのいずれかから玉を取り出すと、白玉であった。この白玉が箱A、箱Bから取り出された確率をそれぞれ求めよ。それぞれの箱の選び方は等確率とする。
   - 白玉が箱Aから取り出された確率：$P(A|白)$
   - 白玉が箱Bから取り出された確率：$P(B|白)$
   - ベイズの定理と1.の結果から

$$\boldsymbol{P(A|白)} = \frac{P(白|A)P(A)}{P(白)} = \frac{\frac{1}{4}\times\frac{1}{2}}{\frac{21}{40}} = \mathbf{\frac{5}{21}}$$

$$\boldsymbol{P(B|白)} = \frac{P(白|B)P(B)}{P(白)} = \frac{\frac{4}{5}\times\frac{1}{2}}{\frac{21}{40}} = \mathbf{\frac{16}{21}}$$

箱の選び方が等確率という条件で、白玉が出たときに、**箱Bを選択した**可能性が高いと考えられる

演習問題 解答例

3. 箱A、箱Bの選ばれる確率が 9 : 1 であったとき、1. と 2.で求めた確率は、どのように変化するか？

■ 1. と 2.の計算式に $P(A)=\frac{9}{10}$、$P(B)=\frac{1}{10}$ を代入して再計算

$$P(白)=P(A)P(白|A)+P(B)P(白|B)=\frac{1}{4}\times\frac{9}{10}+\frac{4}{5}\times\frac{1}{10}=\frac{61}{200}$$

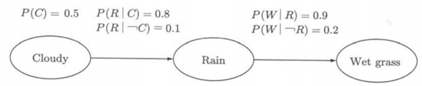
$$P(A|白)=\frac{P(白|A)P(A)}{P(白)}=\frac{\frac{1}{4}\times\frac{9}{10}}{\frac{61}{200}}=\frac{45}{61}$$

$$P(B|白)=\frac{P(白|B)P(B)}{P(白)}=\frac{\frac{4}{5}\times\frac{1}{10}}{\frac{61}{200}}=\frac{16}{61}$$

箱A、箱Bの選ばれる確率が9:1であったという条件だと
白玉が出たときに、**箱Aを選択した**可能性が高いと考えられる

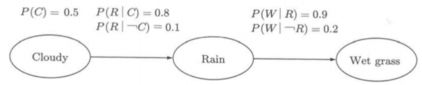
演習問題

□ 以下のhead-to-tail connectionのベイジアンネットワークを考える。



1. 事前確率$P(C)$から雨が降った確率$P(R)$を計算しなさい
2. 1. の結果から芝生が濡れている確率$P(W)$を計算しなさい
3. 「曇っている」ことが観測されたとき、その条件のもとで芝生が濡れている確率$P(W|C)$を計算しなさい

演習問題 解答例



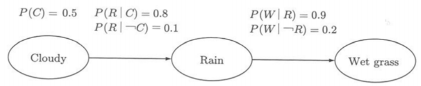
1. 事前確率$P(C)$から雨が降った確率$P(R)$を計算しなさい

■ 雨が降る全ての可能性を加算する

$$P(R)=P(R|C)P(C)+P(R|\neg C)P(\neg C)$$
$$=0.8\cdot 0.5+0.1\cdot 0.5=\mathbf{0.45}$$

「曇っているか？」がわからない状況で
雨が降っている可能性は45%である

演習問題 解答例



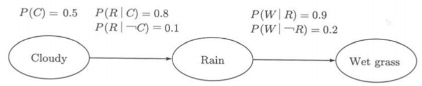
2. 1.の結果から芝生が濡れている確率$P(W)$を計算しなさい

■ 芝生が濡れる全ての可能性を加算する

$$P(W)=P(W|R)P(R)+P(W|\neg R)P(\neg R)$$
$$=0.9\cdot 0.45+0.2\cdot 0.55=\mathbf{0.515}$$

「雨が降っているか？」がわからない状況で
芝生が濡れている可能性は51.5%である

演習問題 解答例



3. 「曇っている」ことが観測されたとき、その条件のもとで芝生が濡れている確率$P(W|C)$を計算しなさい

■ 曇っている状況で、芝生が濡れる全ての可能性を加算すれば良いので

$$P(W)=P(W|R)P(R|C)+P(W|\neg R)P(\neg R|C)$$
$$=0.9\cdot 0.8+0.1\cdot(1-P(R|C))=\mathbf{0.74}$$

「曇っていること」がわかっていて、「雨が降っているか？」がわからない状況で、芝生が濡れている可能性は74%である

機械学習 第５回 識別 –生成モデルと識別モデル–

立命館大学 情報理工学部
福森 隆寛

Beyond Borders

演習問題

□ 1次元の特徴量xを用いて、入力データを正例、負例のクラスのどちらかに識別する問題を考える。

□ 右の表に示す4つの学習データが与えられている。パーセプトロンの学習アルゴリズムを適用して、識別関数の重みを学習する。

■ 学習係数：$\eta=2.0$

番号	特徴量 x	クラス
1	0.2	正例
2	1.2	正例
3	-0.5	負例
4	-1.0	負例

□ 識別関数：$g(x)=w_0+w_1x$

■ $g(x)>0$：正例、$g(x)<0$：負例

■ 重みの初期値：$w_0=1.0$、$w_1=-1.0$

演習問題

1. 番号１から順に学習アルゴリズムを適用するとき、最初に重み係数が更新されるのは何番のデータに対してか？
2. パーセプトロンの学習アルゴリズムによって最終的に決定される重みを求めよ。

演習問題 解答例

1. 番号１から順に学習アルゴリズムを適用するとき、最初に重み係数が更新されるのは何番のデータに対してか？

■ 初期状態：$g(x)=w_0+w_1x=-1.0x+1.0$

■ 正例のデータの場合：$g(x)>0$、負例のデータの場合：$g(x)<0$ である必要がある

■ 番号１（正例のデータ：$x=0.2$）
- $g(0.2)=-1.0\cdot 0.2+1.0=0.8>0$ → 修正なし

■ 番号２（正例のデータ：$x=1.2$）
- $g(1.2)=-1.0\cdot 1.2+1.0=-0.2<0$ → **修正あり**

答え：番号２

演習問題 解答例

2. パーセプトロンの学習アルゴリズムによって最終的に決定される重みを求めよ。

・番号１（正例）：$g(0.2) = -1.0 \cdot 0.2 + 1.0 = 0.8 > 0$ → 修正なし
・番号２（正例）：$g(1.2) = -1.0 \cdot 1.2 + 1.0 = -0.2 < \mathbf{0}$ → **修正あり**
$w_0' = w_0 + \eta \cdot 1.0 = 1.0 + 2.0 \cdot 1.0 = 3.0$
$w_1' = w_1 + \eta \cdot x = -1.0 + 2.0 \cdot 1.2 = 1.4$

・パターン番号３（負例）：$g(-0.5) = 1.4 \cdot (-0.5) + 3.0 = 2.3 > \mathbf{0}$ → **修正あり**
$w_0' = w_0 + \eta \cdot 1.0 = 3.0 - 2.0 \cdot 1.0 = 1.0$
$w_1' = w_1 + \eta \cdot x = 1.4 - 2.0(-0.5) = 2.4$

・パターン番号４（負例）：$g(-1.0) = 2.4 \cdot (-1.0) + 1.0 = -1.4 < 0$ → 修正なし
・パターン番号１（正例）：$g(0.2) = 2.4 \cdot 0.2 + 1.0 = 1.48 > 0$ → 修正なし
・パターン番号２（正例）：$g(1.2) = 2.4 \cdot 1.2 + 1.0 = 3.88 > 0$ → 修正なし
・パターン番号３（負例）：$g(-0.5) = 2.4 \cdot (-0.5) + 1.0 = -0.2 < 0$ → 修正なし

答え： $w_0 = 1.0, w_1 = 2.4$

演習問題

- 以下の重みが与えれたロジスティック識別器を考える

$$p(\oplus\,|\boldsymbol{x}) = \frac{1}{1+\exp\{-\hat{c}(\boldsymbol{x})\}} = \frac{1}{1+\exp\{-(w_0 + \boldsymbol{w}\cdot\boldsymbol{x})\}}$$

$$w_0 = -4.18$$
$$\boldsymbol{w} = \{w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8\}$$
$$= \{0.06, 0.02, -0.01, 0, 0, 0.04, 0.47, 0.01\}$$

このロジスティック識別器に、以下のデータが入力されたとき、何%の確率で正例に分類されるか？

$$\boldsymbol{x} = \{6.0, 148.0, 72.0, 35.0, 0.0, 33.6, 0.627, 50.0\}$$

※ 関数電卓を活用してください

演習問題 解答例

$$\hat{c}(\boldsymbol{x}) = w_0 + \boldsymbol{w}\cdot\boldsymbol{x}$$
$$= -4.18 + 0.06 \times 6.0 + 0.02 \times 148.0$$
$$-0.01 \times 72.0 + 0 \times 35.0 + 0 \times 0.0$$
$$+0.04 \times 33.6 + 0.47 \times 0.627 + 0.01 \times 50.0$$
$$\cong 0.559$$

$$p(\oplus\,|\boldsymbol{x}) = \frac{1}{1+\exp\{-\hat{c}(\boldsymbol{x})\}} = \frac{1}{1+\exp\{-0.559\}}$$
$$\cong 0.636$$

入力データ$\boldsymbol{x}$は、約63.6%の確率で正例に分類される

機械学習 第６回 回帰

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

- 右表のような身長と体重のデータが与えられた

- 身長xから体重$\hat{c}(x)$を予測する線形回帰式が
$\hat{c}(x) = 0.625x - 48.604$
であるときの決定係数と相関係数を計算せよ

身長と体重データ

番号	身長 [cm]	体重 [kg]
1	147.9	41.7
2	163.5	60.2
3	159.8	47.0
4	155.1	53.2
5	163.3	48.3
6	158.7	55.2
7	172.0	58.5
8	161.2	49.0
9	153.9	46.7
10	161.6	52.5

演習問題 解答例

番号	身長 x_i	体重 y_i	$\hat{c}(x)$	$(y_i - \hat{c}(x_i))^2$	$(y_i - \tilde{y})^2$
1	147.9	41.7	43.8	4.6	90.8
2	163.5	60.2	53.6	43.8	80.5
3	159.8	47.0	51.3	18.2	17.9
4	155.1	53.2	48.3	23.7	3.9
5	163.3	48.3	53.5	26.6	8.6
6	158.7	55.2	50.6	21.3	15.8
7	172.0	58.5	58.9	0.2	52.9
8	161.2	49.0	52.1	9.9	5.0
9	153.9	46.7	47.6	0.8	20.5
10	161.6	52.5	52.4	0.0	1.6

平均体重 $\tilde{y} \cong 51.2$　　合計：149.0　合計：297.4

決定係数

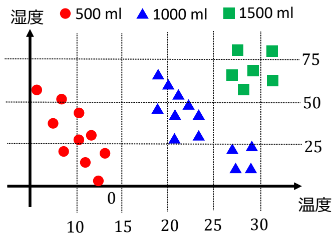
$$R^2 = 1 - \frac{\sum_{i=1}^{N}\{y_i - \hat{c}(x_i)\}^2}{\sum_{i=1}^{N}(y_i - \tilde{y})^2} = 1 - \frac{149.0}{297.4} \cong \mathbf{0.5}$$

相関係数

$$R = \sqrt{R^2} = \sqrt{0.5} \cong \mathbf{0.706}$$

演習問題

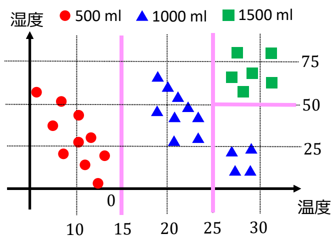
- 以下の学習データが与えられたとき、温度と湿度から1日あたりのビールの消費量を予測する回帰木を作成せよ



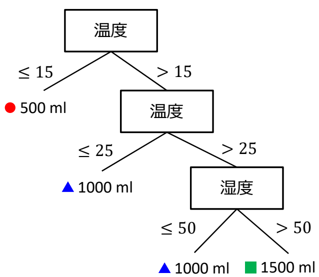
温度・湿度・1日あたりのビールの消費量の関係

演習問題 解答例

出力値の近いデータが集まるように、
特徴の値によって学習データを分割するのがポイント



温度・湿度・1日あたりのビールの消費量の関係



機械学習 第７回 サポートベクトルマシン

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

- SVMは2クラスのデータに対する分類器である。
ここで、3クラス以上のデータに対する多クラス識別問題にSVMを適用するには、どうすれば良いか考えよ

演習問題 解答例

- SVMは2クラスのデータに対する分類器である。
ここで、3クラス以上のデータに対する多クラス識別問題にSVMを適用するには、どうすれば良いか考えよ

- 解答例
  - cクラスの場合（$c > 2$）
あるクラスとそれ以外のデータを識別するc個のSVMを作成し、結果として最もスコアの高いクラスを選択する

演習問題

- カーネル関数の1つである多項式カーネルは、以下の式で与えられる
$$K(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{\phi}(\boldsymbol{x})^T \boldsymbol{\phi}(\boldsymbol{x}') = (\boldsymbol{x}^T \boldsymbol{x}' + 1)^p$$

- 2次元特徴ベクトルを $\boldsymbol{x} = (x_1, x_2)$、$p = 2$として $\boldsymbol{\phi}(\boldsymbol{x})$を求めよ

演習問題 解答例

$$K(\boldsymbol{x}, \boldsymbol{x}') = (\boldsymbol{x}^T \boldsymbol{x}' + 1)^2$$
$$= (x_1 x_1')^2 + (x_2 x_2')^2 + 2x_1 x_1' x_2 x_2' + 2x_1 x_1' + 2x_2 x_2' + 1$$
$$= \left((x_1)^2, (x_2)^2, \sqrt{2}x_1 x_2, \sqrt{2}x_1, \sqrt{2}x_2, 1\right) \cdot \left((x_1')^2, (x_2')^2, \sqrt{2}x_1' x_2', \sqrt{2}x_1', \sqrt{2}x_2', 1\right)$$
$$\therefore \boldsymbol{\phi}(\boldsymbol{x}) = \left(x_1^2, x_2^2, \sqrt{2}x_1 x_2, \sqrt{2}x_1, \sqrt{2}x_2, 1\right)$$

- 第3項（$\sqrt{2}x_1 x_2$）：特徴の積の項が加わっている
- 2つの特徴が同時に現れるときに大きな値となる

→ 共起の情報が加わったということ

機械学習 第８回 ニューラルネットワーク

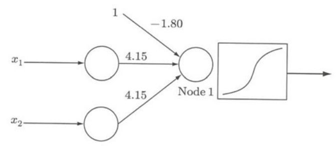
立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

- 以下のニューラルネットワークのノードが与えられたとする
  1. 入力x_1とx_2が取りうる値が0と1のどちらかであるとき、x_1とx_2の全ての組み合わせに対する出力を計算せよ
  2. 1.の結果より、どのような入力が与えられると高い出力が得られやすいか考察しなさい



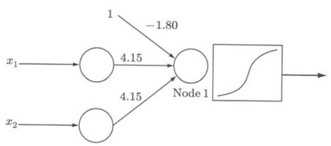
ニューラルネットワークのノード

入力		出力
x_1	x_2	
0	0	
0	1	
1	0	
1	1	

演習問題 解答例

1. 入力x_1とx_2が取りうる値が0と1のどちらかであるとき、x_1とx_2の全ての組み合わせに対する出力を計算せよ

x_1, x_2 に対する出力の値は、以下の式で算出できる
$$f(x_1, x_2) = \frac{1}{1 + \exp\{-(1 \cdot (-1.80) + 4.15 \cdot x_1 + 4.15 \cdot x_2)\}}$$
$f(x_1 = 0, x_2 = 0) \cong 0.142$ $\quad f(x_1 = 0, x_2 = 1) \cong 0.913$
$f(x_1 = 1, x_2 = 0) \cong 0.913$ $\quad f(x_1 = 1, x_2 = 1) \cong 0.998$



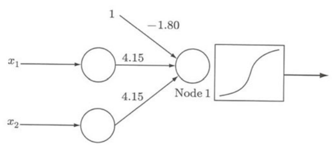
ニューラルネットワークのノード

入力		出力
x_1	x_2	
0	0	0.142
0	1	0.913
1	0	0.913
1	1	0.998

演習問題 解答例

2. 1.の結果より、どのような入力が与えられると高い出力が得られやすいか考察しなさい

x_1, x_2のどちらかに1が含まれていると高い値が出力される
つまり、このノードは論理演算（OR関数）が表現されている



ニューラルネットワークのノード

入力		出力
x_1	x_2	
0	0	0.142
0	1	0.913
1	0	0.913
1	1	0.998

演習問題

- シグモイド関数を活性化関数にすると多層ニューラルネットワークの重みを学習するときに、勾配消失問題が発生しやすい理由を考えよ

演習問題 解答例

- シグモイド関数を活性化関数にすると
  多層ニューラルネットワークの重みを学習するときに、
  勾配消失問題が発生しやすい理由を考えよ

シグモイド関数の微分値（0に近い値）を使って重みを更新するから

最急勾配法による重みの更新式

$$w_j \leftarrow w_j + \eta \sum_{x_i \in D} (y_i - o_i) o_i (1 - o_i) x_{ij}$$

$$o_i = \frac{1}{1 + \exp(-h_i)}$$

シグモイド関数
（値が取りうる範囲：0～1）

1以下　1以下

$$\frac{\partial o_i}{\partial h_i} = o_i(1 - o_i)$$

シグモイド関数の微分
1以下の値の積（0に近づく）

機械学習 第９回 深層学習

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

- 以下の0から7までの数を自己写像する
  ニューラルネットワーク（オートエンコーダ）を考える
  - 入力層・出力層：8次元ベクトル
    - 表現する数に対応する特徴のみが1、その他は0
      - 例：「3」という数に対する入力ベクトルは「0,0,1,0,0,0,0,0」
  - 中間層のノード数：3
- このネットワークを学習した結果、
  中間層では、入力がどのように表現されているか
  （入力からどのような情報が獲得されたか）考えよ

演習問題 解答例

- このネットワークを学習した結果、中間層では、
  入力がどのように表現されているか考えよ
- 解答例
  - 「2進エンコーディングの概念」が獲得されている（かも）
  - 0から7の数は、3桁以内の2進数で表現できる
    - 8ノードの入力情報が、より少ない3ノードの中間情報で表現
    - 「3」を例に考えると
      - 入力層：「0,0,1,0,0,0,0,0」、中間層（2進数）：「0,1,1」

オートエンコーダは、このようにして得られた中間層の値を入力として、1階層上にずらして、同様の表現学習を行う。この手順を積み重ねると、入力に近い側では単純な特徴が、階層が上がっていくにつれて複雑な特徴が学習される

演習問題

- 畳み込みニューラルネットワークを用いたモデルを構築することを考える。このネットワークを構築するために、事前に決めておくべきパラメータは何か考えよ
  - 1つ目は「活性化関数を何にするか？」
  - それ以外に事前に決めておかないといけないものは何？

演習問題 解答例

- 畳み込みニューラルネットワークを用いたモデルを構築することを考える。このネットワークを構築するために、事前に決めておくべきパラメータは何か考えよ
  - 活性化関数
  - 畳み込み層のフィルタサイズとフィルタ数
  - プーリングの種類（平均なのか？最大値なのか？）
  - 層の構成（畳み込み層、全結合層の数など）
  - ドロップアウトの割合
  - 重みを更新する式の学習係数
  - 更新回数
  - 事前学習の有無　　など

機械学習 第１０回 アンサンブル学習

立命館大学 情報理工学部

福森 隆寛

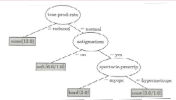
Beyond Borders

演習問題

- バギングで作成する識別器は不安定な方が良い。
  このアルゴリズムで用いる不安定な識別器は、
  どのようなものがあるか考えよ。

演習問題 解答例

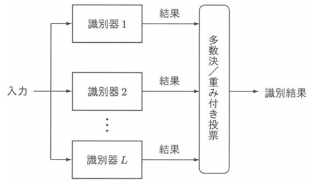
- 解答例
  - 決定木
    - 得られる情報が多い特徴を木の上方のノードに配置するように構成



  - 決定木は過学習を起こしやすい
    - 過学習は、モデルが学習データに特化しすぎて、未知データに対して性能が下がる現象
  - 全ての事例のエラーが無くなるように決定木を作成すると、木が成長しすぎて、学習データに適応しすぎた過学習になりがち
    - 言い換えると、学習データを変えると、決定木の構成も異なる

演習問題

- この講義ではアンサンブル学習によって識別問題を解く方法を紹介した。これを回帰問題に適用するにはどうすれば良いか考えなさい。



アンサンブル学習による識別問題

演習問題 解答例

- この講義ではアンサンブル学習によって識別問題を解く方法を紹介した。これを回帰問題に適用するにはどうすれば良いか考えなさい。

- 解答例
  - アンサンブル学習で、複数の回帰モデルを生成
  - それぞれのモデルの出力値の平均値や中央値を用いて最終的な回帰の結果を得る

機械学習 第１１回 モデル推定

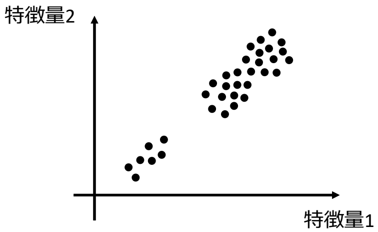
立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

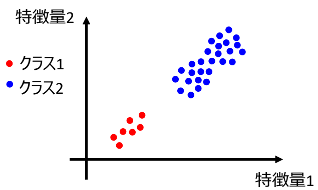
- 以下の教師なしデータに対して、分割数を2の場合と3の場合のそれぞれで分割最適化クラスタリングを適用したとき、どのような結果が期待されるか？

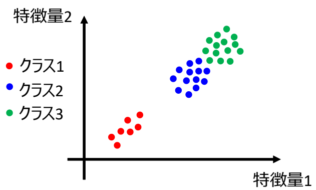


演習問題 解答例

- 以下の教師なしデータに対して、分割数を2の場合と3の場合のそれぞれで分割最適化クラスタリングを適用したとき、どのような結果が期待されるか？

- 解答例





演習問題

- k-meansアルゴリズムのクラスタリング結果を用いて新たなデータが観測されたときに、そのデータが属するクラスタを決める方法を考えなさい

- k-meansアルゴリズムのクラスタリング結果に基づく教師なし学習では、良い識別器を作ることができない理由を考えなさい

演習問題 解答例

- k-meansアルゴリズムのクラスタリング結果を用いて新たなデータが観測されたときに、そのデータが属するクラスタを決める方法を考えなさい

- 解答例
  - k-meansアルゴリズムの結果は、各クラスの平均ベクトルなので、新しく観測したデータと、各クラスタの平均ベクトルとの距離を求めて、最も近いクラスタを分類結果とする

演習問題 解答例

- k-meansアルゴリズムのクラスタリング結果に基づく教師なし学習では、良い識別器を作ることができない理由を考えなさい

- 解答例
  - この距離計算は、全クラスタの分散が等しいことを前提にしており、クラスタごとにデータの広がり方が異なる場合には適切な結果とならない
  - クラスタの事前確率も考慮されていない
  - 事後確率を計算する確率モデルを作り、その確率が最大となるモデルパラメータをEMアルゴリズムで学習する

機械学習 第１２回 パターンマイニング

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

- 以下のスーパーマーケットの売り上げ記録がある
  - 商品点数：{ミルク、パン、バター、雑誌} の4点
  - トランザクション：6件（6回分の購入記録）

1. このデータは「疎らなデータ」、「密なデータ」のどちらであるか考えよ
2. 項目集合{ミルク、パン}の支持度を示す support({ミルク、パン})を求めなさい
3. 商品点数が1000種類であったとき、可能な項目集合の数はいくつか？
   ヒント：商品は「買った」「買っていない」の2種類に分類できると考える

No.	ミルク	パン	バター	雑誌
1	t	t		
2		t		
3				t
4		t	t	
5	t	t	t	
6	t	t		

スーパーマーケットの売り上げ記録
（tはその商品が買われたことを意味する）

演習問題 解答例

1. このデータは「疎らなデータ」、「密なデータ」のどちらであるか考えよ

- 解答例
  - 疎らなデータ
  - トランザクションの特徴ベクトルのほとんどの次元の値は空白で、少数の次元のデータが埋まっている
    - スーパーマーケットで揃えている商品の種類数：数千～数万
    - 1回の買い物で1人の客が買う商品の種類数：多くて数十

スーパーマーケットの売り上げ記録

No.	ミルク	パン	バター	雑誌
1	t	t		
2		t		
3				t
4		t	t	
5	t	t	t	
6	t	t		

演習問題 解答例

2. 項目集合{ミルク、パン}の支持度を示す support({ミルク、パン})を求めなさい

- 解答例
  - support({ミルク、パン}) $= \frac{T_{\{\text{items}\}}}{T} = \frac{3}{6} = \mathbf{0.5}$
    - T：全トランザクション件数＝6
    - T_{items}：項目集合itemsが出現するトランザクション件数
      - $T_{\{\text{items}\}} = T_{\{\text{ミルク、パン}\}} = 3$

No.	ミルク	パン	バター	雑誌
1	t	t		
2		t		
3				t
4		t	t	
5	t	t	t	
6	t	t		

スーパーマーケットの売り上げ記録

演習問題 解答例

3. 商品点数が1000種類であったとき、可能な項目集合の数はいくつか？

- 解答例
  - $\mathbf{2^{1000} - 1}$**通り**
    - 1000点の商品に対して{買った、買わなかった}の値を割り振るので、2^{1000}通りとなり、この中から何も買われなかった場合の1通りを除く
  - この結果からも、まともに全ての可能な組み合わせについて計算することは難しいことがわかる
    - 支持度を計算する項目集合をいかにして絞り込むかが重要！

演習問題

- 「ハム→卵」という規則について、以下の情報が得られたとする
  - 項目集合 {ハム、卵} の支持度が0.1
  - 「ハム→卵」という規則の確信度が0.7、リフト値：5
- 上記の条件において、以下の項目を求めなさい
  1. ハムと卵を同時に購入している客の割合
  2. ハムを購入した客の中で卵も購入する客の割合
  3. 「ハムを既に買った客が卵を買う確率」は「任意の客が卵を買う確率」の何倍か？

演習問題 解答例

- 「ハム→卵」という規則について、以下の情報が得られたとする
  - 項目集合 {ハム、卵} の支持度が0.1
  - 「ハム→卵」という規則の確信度が0.7、リフト値：5
- 解答例
  1. ハムと卵を同時に購入している客の割合：10 %（支持度）
  2. ハムを購入した客の中で卵も購入する客の割合：70 %（確信度）
  3. 「ハムを既に買った客が卵を買う確率」は「任意の客が卵を買う確率」の何倍か？：5倍（リフト値）

演習問題

- 実際に協調フィルタリングが用いられている事例を調べなさい

演習問題 解答例

- 実際に協調フィルタリングが用いられている事例を調べなさい
- 解答例
  - Amazonオンラインショップ
  - ある商品を検索すると
    - 同時に購入される商品
    - 他にチェックしている商品

  が一緒に表示される

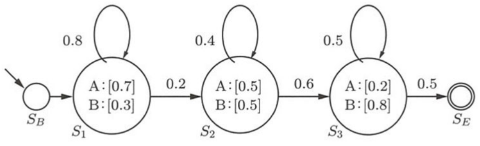
機械学習 第１３回 系列データの識別

立命館大学 情報理工学部
福森 隆寛

Beyond Borders

演習問題

- 下図に示すHMMが与えられているとき、特徴ベクトル系列 “AAABB” が出力されるような状態遷移系列を全て求めよ

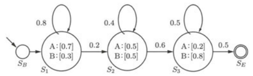


HMM

演習問題 解答例

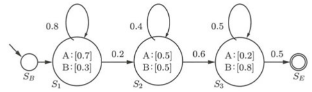
- 下図に示すHMMが与えられているとき、特徴ベクトル系列 “AAABB” が出力されるような状態遷移系列を全て求めよ

- 解答例
  - 系列 “AAABB” が出力されるということは、以下の制約がつく
    - 最初の記号Aは状態S_1から出力
    - 最後の記号Bは状態S_3から出力
  - このような制約の中で、考えられる状態遷移を全て洗い出せばよい



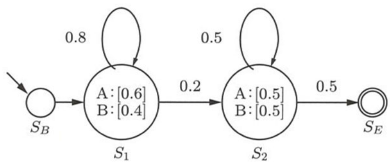
演習問題 解答例

- 解答例（続き）
  - 系列 “AAABB” を出力できる状態遷移系列は以下の6通り
    - $S_B \to S_1 \to S_1 \to S_1 \to S_2 \to S_3 \to S_E$
    - $S_B \to S_1 \to S_1 \to S_2 \to S_2 \to S_3 \to S_E$
    - $S_B \to S_1 \to S_1 \to S_2 \to S_3 \to S_3 \to S_E$
    - $S_B \to S_1 \to S_2 \to S_2 \to S_2 \to S_3 \to S_E$
    - $S_B \to S_1 \to S_2 \to S_2 \to S_3 \to S_3 \to S_E$
    - $S_B \to S_1 \to S_2 \to S_3 \to S_3 \to S_3 \to S_E$



演習問題

- 下図に示すHMMが与えられているとき、特徴ベクトル系列 “AAB” の出力確率を求めよ

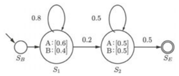


HMM

演習問題 解答例

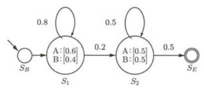
- 下図に示すHMMが与えられているとき、特徴ベクトル系列 “AAB” の出力確率を求めよ

- 解答例
  - 系列 “AAB” が出力されるということは、以下の制約がつく
    - 最初の記号Aは状態S_1から出力
    - 最後の記号Bは状態S_2から出力
  - 系列長: 3、状態数: 2より、S_1とS_2のどちらかで1回自己遷移
  - 系列 “AAB” を出力できる状態遷移系列は以下の2通り
    - $S_B \to S_1 \to S_1 \to S_2 \to S_E$
    - $S_B \to S_1 \to S_2 \to S_2 \to S_E$



演習問題 解答例

- 解答例（つづき）



  - $S_B \to \overset{(A)}{S_1} \to \overset{(A)}{S_1} \to \overset{(B)}{S_2} \to S_E$ の確率

$$P(S_B \to S_1)P(A|S_1)P(S_1 \to S_1)P(A|S_1)P(S_1 \to S_2)P(B|S_2)P(S_2 \to S_E)$$
$$= 1.0 \times 0.6 \times 0.8 \times 0.6 \times 0.2 \times 0.5 \times 0.5 = 0.0144$$

  - $S_B \to \overset{(A)}{S_1} \to \overset{(A)}{S_2} \to \overset{(B)}{S_2} \to S_E$ の確率

$$P(S_B \to S_1)P(A|S_1)P(S_1 \to S_2)P(A|S_2)P(S_2 \to S_2)P(B|S_2)P(S_2 \to S_E)$$
$$= 1.0 \times 0.6 \times 0.2 \times 0.5 \times 0.5 \times 0.5 \times 0.5 = 0.0075$$

  - これらは同時には起こりえない事象なので、最終的に求める確率は、これらの和となる
  - 系列 “AAB” の出力確率 $= 0.0144 + 0.0075 = \mathbf{0.0219}$

RITSUMEIKAN

機械学習 第１４回 半教師あり学習

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

- 人間の成長過程は、半教師あり学習に似ていると言われている。その理由を、実際の事例を挙げながら述べなさい。

演習問題 解答例

- 人間の成長過程は、半教師あり学習に似ていると言われている。その理由を、実際の事例を挙げながら述べなさい。

- 解答例
  - 幼い頃、大人たちに「これは猫だよ」、「これは犬だよ」と何度か教えてもらっただけで、あとは自分で大量の動物を見ながら、猫や犬を正確に認識していく
    - 大人たちから教えてもらった少量の情報 → 正解付きデータ
    - 自分自身が見た大量の動物 → 正解なしデータ

演習問題

- 半教師あり学習を活用できる場面を考えなさい

演習問題 解答例

- 半教師あり学習を活用できる場面を考えなさい

- 解答例
  - 以下のような条件を満たす状況を考えるのがポイント
    - 用意できる正解データが少ない
    - 収集したデータに正解情報を付与するのに大きなコストがかかる
  - 医療用画像の分類・認識
    - ラベル付けされた医療用画像は少量
      - その病気と診断された患者の数だけしか集められない
    - 画像に正解情報を付与できる人が限られる（コストが大きい）
      - 基本的に医療用画像から病気を診断できるのは医師（専門家）だけ

機械学習 第１５回 強化学習

立命館大学 情報理工学部

福森 隆寛

Beyond Borders

演習問題

- 「強化学習」と「教師あり/教師なし学習」の違いを考えなさい

演習問題 解答例

- 「強化学習」と「教師あり/教師なし学習」の違いを考えなさい
- 解答例
  - 教師信号が間接的
    - 「何が正解か？」ではなく、報酬（連続値）が時々与えられる
  - 報酬が遅れて与えられる
    - 将棋の勝利、迷路のゴール
  - 探求が可能
    - エージェントが自分で学習の対象を選べる
  - 状態が確定的でない場合がある
    - 確率分布でそれぞれの状態にいる確率を表せる