

# レポート課題(4回目)

「Rで学ぶデータサイエンス」12.3節の問題(p135)

Rに組み込まれているFisherとAndersonによるアヤメの分類データセットirisを用いて深層学習により、以下のことを行いなさい  
なお、

- Rのライブラリー: h2oを使用する(他のパッケージの使用も可)  
(pythonの場合には適当な深層学習ライブラリを使用すること)
- Irisのデータの80%を学習データに、20%を実験データとする  
(交差検証でも可)
- 評価尺度: 実験データで正しくSpeciesを判定できた割合(判別精度)

課題1: 学習データについて、Speciesを目的変数、他の4変数を説明変数として判別関数を決定しなさい(モデルの出力)

課題2: 判別関数を実験データの判別に適用しなさい((平均)判別精度)

## 12.3 ◆ 問題

R に組み込まれている Fisher と Anderson によるアヤメの分類データセット `iris` を用いる。同じ問題を、すでに他の手法で判別している。ここでは、深層学習を用いて以下のことを行いなさい。

```
> data(iris)
> head(iris)
```

	sepal.Length	sepal.width	Petal.Length	Petal.width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa

- ① 学習データについて、Species を目的変数、他の 4 変数を説明変数として判別関数を決定しなさい。
- ② 判別関数を実験データの判別に適用しなさい。

# レポート提出方法

- Manaba+Rのレポート機能を用いてレポートを提出する
- レポートの先頭に氏名と学生証番号を記入すること  
形式は自由
- レポートのファイル形式はpdf(pdf以外は減点)
- レポートの締切日時  
15回目の講義日前日(1/17)の17:00
- レポートに記述すべき項目
  - 課題1と課題2  
実行結果のキャプチャー画面とその説明文他
  - R(Python等も可)のソースプログラム
  - 更なる分析結果(詳細な評価,他の手法との比較,他のデータへの適用等)
  - 考察
  - 感想