

Workout report

M1 Harada Shunta

Structure

- Preparation
- Theory
 - Linear algebra
 - Bayes' theorem
- Practice
 - Preprocessing
 - Model
 - Validation
- Other

Structure

- Preparation
- Theory
 - Linear algebra
 - Bayes' theorem
- Practice
 - Preprocessing
 - Model
 - Validation
- Other

Part: Preparation

- Construction of work environment for Workout
 - Python, Jupyter notebook
 - Git , Github
- Skills

Progress management, How to report, How to make programming environment, How to use Git ...

Pick up: Progress management

Action list

- ☒ 草案づくり (このIssue)
- ☐ 仮スライド
- ☐ 先輩に相談

Structure

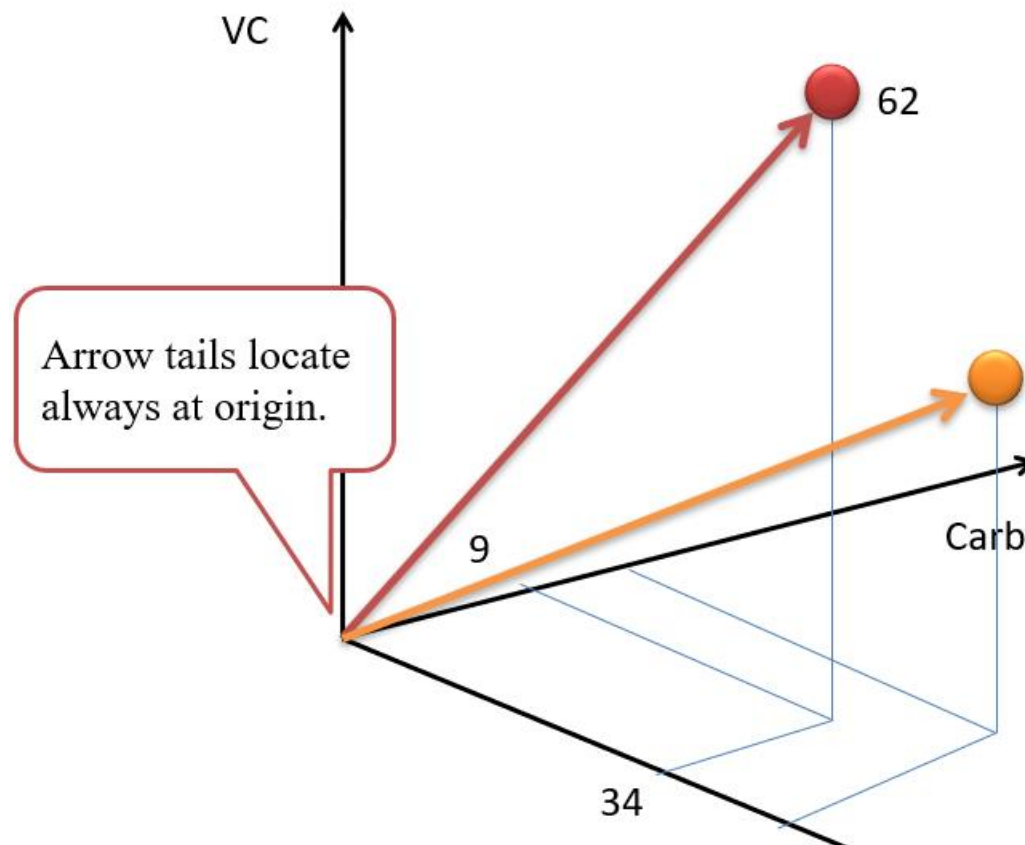
- Preparation
- Theory
 - Linear algebra
 - Bayes' theorem
- Practice
 - Preprocessing
 - Model
 - Validation
- Other


Part: Linear algebra


- ビギナー：ベクトル空間
ベクトル空間が何かを説明できる
ベクトル空間に標準内積を導入できる
- 常人：線形変換
写像が説明できる
射影が説明できる
線形変換の定義が説明できる
- 玄人：特異値分解と主成分分析
線形次元削減の定義を説明できる
データ行列の分散共分散行列を対角化して主成分を求めることができる
SVDを用いて主成分を求めることができる
求めた主成分を用いて次元削減ができる

Part: Linear algebra

- Pick up: Vector space




$$\begin{bmatrix} 34 \\ 9 \\ 62 \end{bmatrix}$$


$$\begin{bmatrix} 46 \\ 12 \\ 32 \end{bmatrix}$$

Structure

- Preparation
- Theory
 - Linear algebra
 - Bayes' theory
- Practice
 - Preprocessing
 - Model
 - Validation
- Other

Part: Bayes' theory

- ビギナー：ベイズルール
条件付き確率，周辺化，同時確率，独立の定義が書ける
条件付き確率の定義からベイズの定理を導出できる
ベイズの定理の式の各部分の名前が言える
- 常人：点推定
問題設定を説明できる
最尤推定でデータからモデルのパラメータを推定できる
MAP推定で（以下同）
- 玄人：分布推定
点推定との違いを説明できる
ベイズ推定でデータからモデルのパラメータの確率分布を推定できる

Part: Bayes' theory

- Pick up: Baye's theory

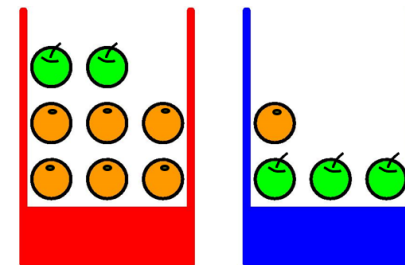
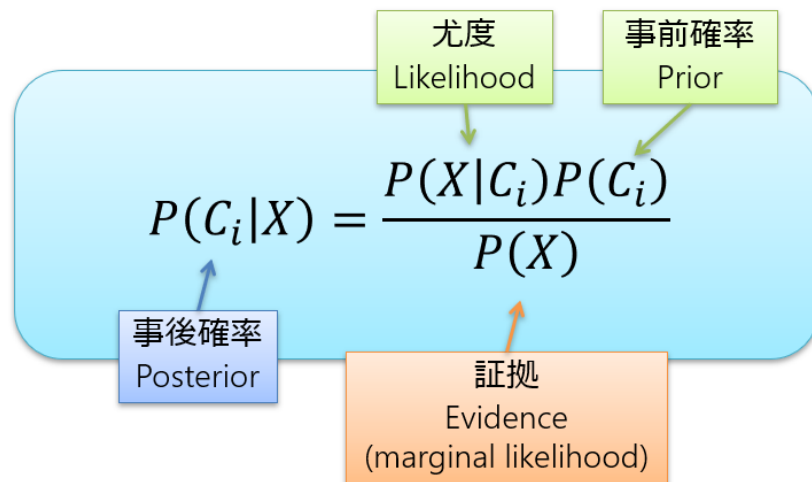
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A|B)P(B) = P(A \cap B)$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$P(B|A)P(A) = P(A \cap B)$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$



$$\text{Bern}(x|\mu) = \mu^x(1 - \mu)^{1-x}$$

Structure

- Preparation
- Theory
 - Linear algebra
 - Bayes' theorem
- Practice
 - Preprocessing
 - Model
 - Validation
- Other

Part: Preprocessing

- ビギナー & 常人

与えられたnumpy arrayを平均0, 分散1に標準化できる
sklearnやUCIなどの任意の公開データセットデータを「使える」
タスクに応じて適切なデータセットを調べられる
標準的な形式のファイルを読み込める

Part: Preprocessing

- Pick up: Category data encoding

One Hot Encode

```
[[ 'r', 'Male', 'A'], [ 'r', 'Female', 'B'], [ 'l', 'Female', 'AB'], [ 'l', 'Male', 'O'], [ 'l', 'Male', 'AB']]  
[ 'l', 'r']  
[ 'Female', 'Male']  
[ 'A', 'AB', 'B', 'O']  
  
array([[0., 1., 0., 1., 1., 0., 0., 0.],  
       [0., 1., 1., 0., 0., 0., 1., 0.],  
       [1., 0., 1., 0., 0., 1., 0., 0.],  
       [1., 0., 0., 1., 0., 0., 0., 1.],  
       [1., 0., 0., 1., 0., 1., 0., 0.]])
```

Label Encode

```
[ 'Havells', 'Philips', 'Syska', 'Eveready', 'Lloyd']
```

```
Eveready --> 0  
Havells --> 1  
Lloyd --> 2  
Philips --> 3  
Syska --> 4
```

Structure

- Preparation
- Theory
 - Linear algebra
 - Bayes' theorem
- Practice
 - Preprocessing
 - **Model**
 - Validation
- Other

Part: Model

- ビギナー：sklearnを用いた実装済みモデルの利用

irisデータセットに対してPCAをかける

Boston house-pricesに対して線形回帰を行い予測できる

sklearnの以下のサンプルコードを写経し自身の環境下で実行せよ
単回帰

PCA

- 常人：Kerasを用いたNNモデルの構築

任意のニューラルネットワークを構築できる

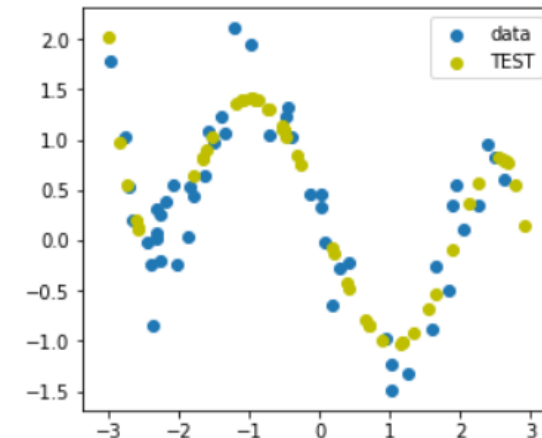
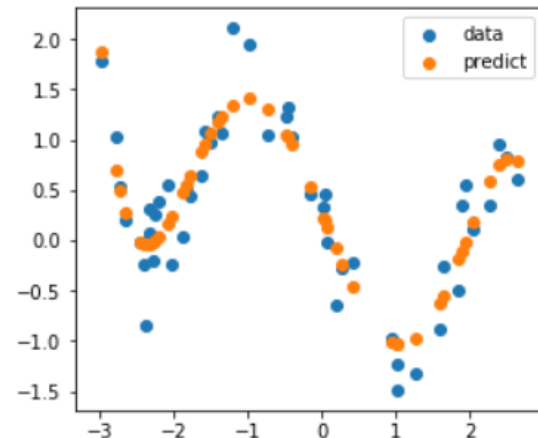
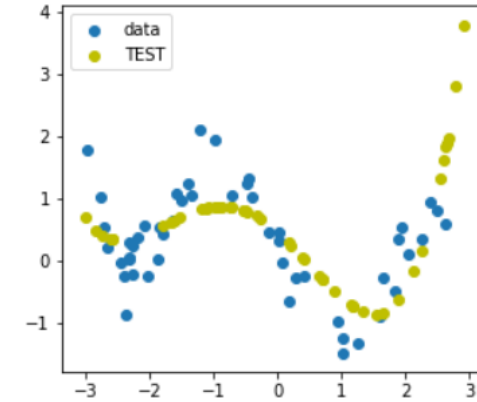
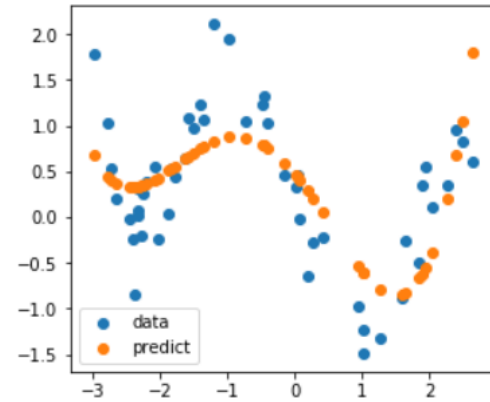
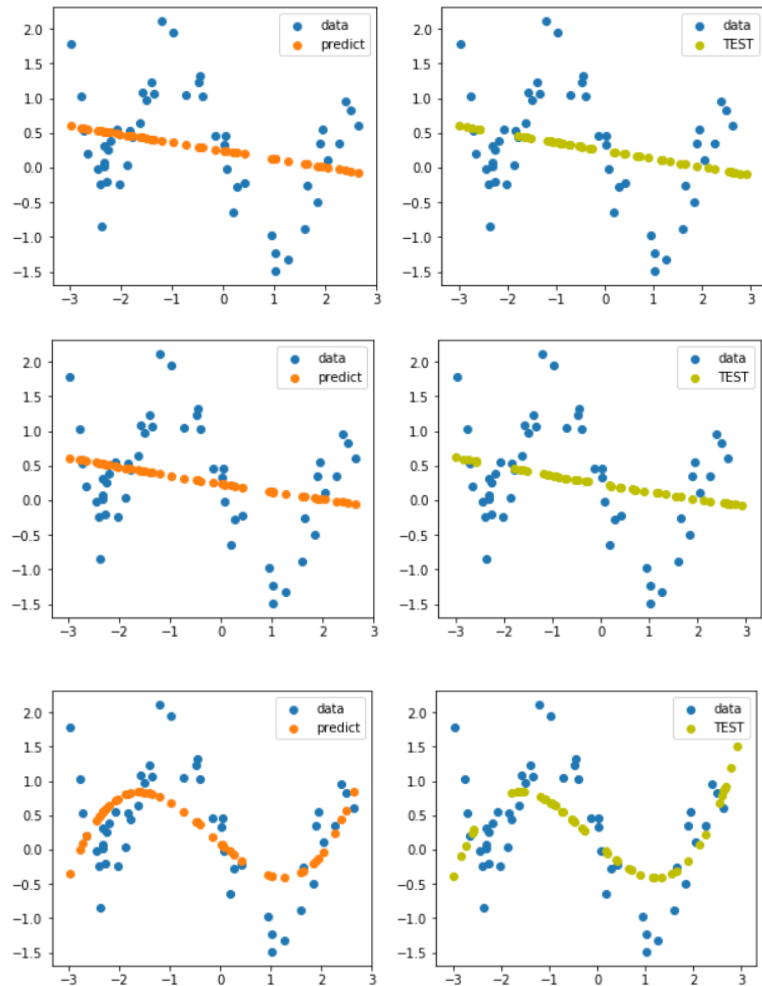
チュートリアルを実行しmnist datasetをClassificationせよ

CNNを構築しmnist datasetに対するClassificationの精度を向上させよ

Autoencoderを構築しmnistデータを用いて学習させよ．また入力データと再構成した画像を比較せよ． 任意の潜在変数からデコードした画像を表示せよ．

Part: Model

- Pick up: Using sklearn models



Structure

- Preparation
- Theory
 - Linear algebra
 - Bayes' theorem
- Practice
 - Preprocessing
 - Model
 - Validation
- Other

Part: Validation

- ビギナー: matplotlibによる描画

結果を出力できる

print
plot
scatter
histogram

- 常人: 教師あり学習のモデル評価・選択
教師あり学習モデルの評価・選択が出来る

Cross Varidation: CVができる

結果の平均分散を計算できる (バイアスバリエーション)

バイアスとバリエーションについて説明せよ.

cross Validationについて説明せよ.

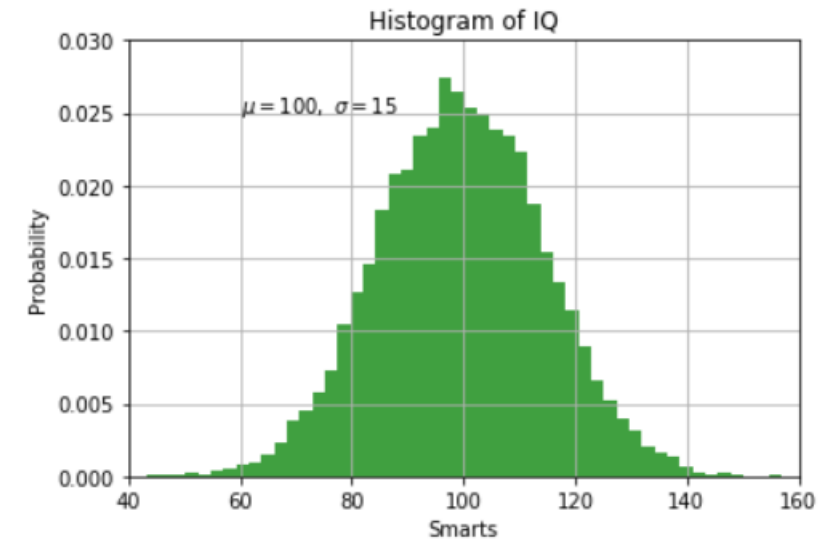
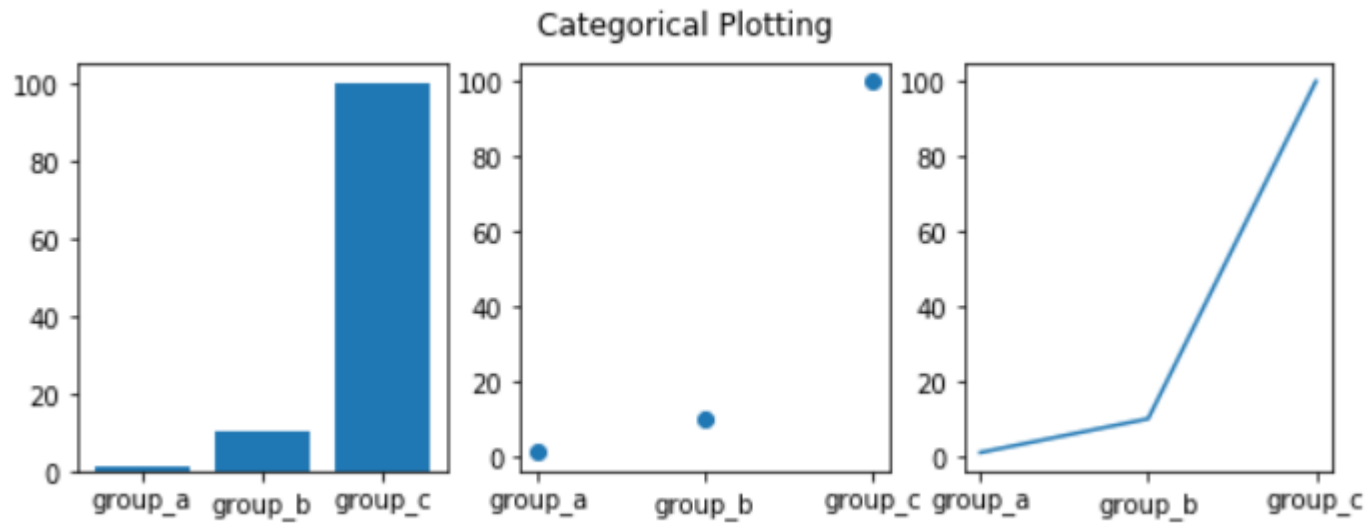
sklearnのBoston house-pricesデータセットをloadして線形回帰を用いて学習

学習させた結果をsklearnのcross validationで評価

学習結果の平均・分散を計算せよ

Part: Validation

- Visualization with matplotlib



Structure

- Preparation
- Theory
 - Linear algebra
 - Bayes' theorem
- Practice
 - Preprocessing
 - Model
 - Validation
- Other

Part: Other, Impressions

- How to manage task

How to organize emotions.

Keep a constant pace.

Resource management.

Can I make more time if I devise it?

Thank you for your listening