

- 収集したデータセット：
 - MNISTの手書き数字の画像データ内の学習用データ, 計52500データ
 - 1画像: 28×28の784次元のデータ
 - $X=0.7, 0.8, 0.9$ の3通りについて実験を行う
- 処理の流れ (pr2.py)
 1. まず、累積寄与率が $X\%$ 以上となるような最小の圧縮次元数を求める。
 - ・ `calc_ratio`関数により、これを行う。
 2. 1の結果から得た圧縮後次元数を使用し、784次元から次元の圧縮を行う
 - ・ `dim_reduc`関数により、これを行う。
 - ・ 出力は、圧縮後のベクトル空間である。
 3. 次に、テスト画像をランダムに選択する
 4. 圧縮前と圧縮後の空間において探索を行い、実行時間を計測する
 - ・ `search_nearest`関数により、これを行う
 - ・ 類似度はユークリッド距離により、比較を行う。
 - ・ 探索時間の比較は、`evaluate_time`関数により、これを行う。
 - ・ 出力は、圧縮後探索時間と圧縮前探索時間である。
 5. 最後に、圧縮前と圧縮後における探索時間の比を求める。
 - ・ `evaluate_time`関数により、これを行う。
 - ・ 出力は、(圧縮後探索時間/圧縮前探索時間)の比である。
 6. おまけ問題として、KD-treeによる探索も行う。
 - ・ `kd_tree`関数により、これを行う。
 - ・ 探索時間の比較は、`evaluate_time`関数により、これを行う。
 - ・ 出力は、圧縮後探索時間と圧縮前探索時間である。

次頁に、実行結果を示す。

```
$ python pr2.py
```

```
*****
```

```
X = 0.7
```

```
圧縮後次元数 : 26
```

```
次元圧縮後のベクトル空間 : [[ 988.90710734 -471.52864654 112.9557262 ...,  
-49.36082628
```

```
71.64388922 55.66302635]  
[ 526.79743802 612.07523648 -94.39698078 ..., 75.13220738  
-135.52447335 -101.83174049]  
[ 1016.44584305 -525.7773642 135.50750426 ..., 16.50342667  
-161.0949815 6.01662794]  
...,  
[-1339.61874604 -227.93015144 -140.01571365 ..., -186.35346666  
-131.0254716 -17.47776811]  
[ 600.92929486 157.64174971 425.23025267 ..., -135.81732385  
117.01729583 51.63410869]  
[ 682.76923822 229.98631667 -69.47222976 ..., 149.13114624  
218.34703722 151.18018392]]
```

```
圧縮前探索時間 : 1.40270495415
```

```
圧縮後探索時間 : 0.671653032303
```

```
探索時間の比 : 0.478827019408
```

```
#####おまけ問題(KD-treeによる探索)#####
```

```
圧縮前探索時間 : 6.16186881065
```

```
圧縮後探索時間 : 0.459137916565
```

```
探索時間の比 : 0.0745127705042
```

```
*****
```

```
X = 0.8
```

```
圧縮後次元数 : 43
```

```
次元圧縮後のベクトル空間 : [[ 988.90710734 -471.52864654 112.9557262 ...,  
-4.59657176
```

```
40.21317491 -59.53603319]  
[ 526.79743802 612.07523648 -94.39698078 ..., -81.7394206  
-15.45832574 11.53208735]  
[ 1016.44584305 -525.7773642 135.50750426 ..., 3.37977921  
-95.85207155 34.93035884]  
...,  
[-1339.61874604 -227.93015144 -140.01571365 ..., 93.65535323  
-136.7835402 55.08159799]  
[ 600.92929486 157.64174971 425.23025267 ..., 182.7302889  
3.24409362 -17.58561121]  
[ 682.76923822 229.98631667 -69.47222976 ..., 55.86081178  
42.0158807 -180.45467588]]
```

```
圧縮前探索時間 : 1.40227007866
```

圧縮後探索時間 : 0.670114994049
探索時間の比 : 0.47787869416
#####おまけ問題(KD-treeによる探索)#####
圧縮前探索時間 : 6.01073813438
圧縮後探索時間 : 0.468101978302
探索時間の比 : 0.0778776196594

X = 0.9

圧縮後次元数 : 87
次元圧縮後のベクトル空間 : [[988.90710734 -471.52864654 112.9557262 ...,
17.47058206
5.8543509 20.0119059]
[526.79743802 612.07523648 -94.39698078 ..., -32.35653415
2.49748644 28.03456466]
[1016.44584305 -525.7773642 135.50750426 ..., 88.44196794
-32.92282703 -3.13582808]
...,
[-1339.61874604 -227.93015144 -140.01571365 ..., 8.80284478
-120.47372549 58.01389002]
[600.92929486 157.64174971 425.23025267 ..., -69.28765419
49.1512246 -13.23393927]
[682.76923822 229.98631667 -69.47222976 ..., -48.22289436
87.6459023 -73.74629032]]

圧縮前探索時間 : 1.43486189842
圧縮後探索時間 : 0.741986036301
探索時間の比 : 0.517113205889
#####おまけ問題(KD-treeによる探索)#####
圧縮前探索時間 : 5.8956849575
圧縮後探索時間 : 0.511266946793
探索時間の比 : 0.0867188376716