

8 Hours

Facebook

HW_week

Bảng tin

Hàm List

numpy.d

Bài tập py

colab.research.google.com/drive/1AxioxrPvHBraZ5eJTtdDTY6tfnYlm-Jn#scrollTo=fLirtTkDTHnG

Laptop Python Học tập Online C++ Compiler Facebook YouTube

HW_week2.ipynb

Nhận xét

Chia sẻ

Chỉnh sửa

RAM

Ổ đĩa

Chỉnh sửa

• Load dữ liệu bank2.dat

[7] 1 from numpy import *
2 import numpy as np
3 dataset = np.array([vec.strip().split() for vec in open('/content/bank2.dat').readlines()])
4 data = dataset.astype(float)
5 print(data)

[[214.8 131. 131.1 9. 9.7 141.]
 [214.6 129.7 129.7 8.1 9.5 141.7]
 [214.8 129.7 129.7 8.7 9.6 142.2]
 ...
 [214.8 130.3 130.4 10.6 11.1 140.]
 [214.7 130.7 130.8 11.2 11.2 139.4]
 [214.3 129.9 129.9 10.2 11.5 139.6]]

• Xem mỗi cột dữ liệu là 1 vector:

1 vt_a = np.loadtxt('/content/bank2.dat')[0, :]
2 vt_b = np.loadtxt('/content/bank2.dat (ctrl + click)
3 vt_c = np.loadtxt('/content/bank2.dat')[1, :]
4 vt_d = np.loadtxt('/content/bank2.dat')[2, :]
5 vt_e = np.loadtxt('/content/bank2.dat')[3, :]
6 vt_f = np.loadtxt('/content/bank2.dat')[4, :]

EN 1:52 PM 10/13/2021

8 Hours

Facebook

HW_week

Bảng tin

Hàm List

numpy.d

Bài tập py

<

>

↺

colab.research.google.com/drive/1AxioxrPvHBraZ5eJTtdDTY6tfnYlm-Jn#scrollTo=fLirtTkDTHnG

🔗

📷

▶

🔍

🔖

📄

☰

Laptop Python Học tập Online C++ Compiler Facebook YouTube

CO

HW_week2.ipynb

☆

Nhận xét

Chia sẻ

H

Tệp chỉnh sửa Xem Chèn Thời gian chạy Công cụ Trợ giúp Lưu lần gần đây nhất...

+ Mã

+ Văn bản

✓ RAM

Ổ đĩa

✎ Chỉnh sửa

⬆

• Xem mỗi cột dữ liệu là 1 vecto:

[8] 1 vt_a = np.loadtxt('/content/bank2.dat')[:, 0]
2 vt_b = np.loadtxt('/content/bank2.dat')[:, 1]
3 vt_c = np.loadtxt('/content/bank2.dat')[:, 2]
4 vt_d = np.loadtxt('/content/bank2.dat')[:, 3]
5 vt_e = np.loadtxt('/content/bank2.dat')[:, 4]
6 vt_f = np.loadtxt('/content/bank2.dat')[:, 5]

• Tính giá trị trung bình mỗi cột

[10] 1 tb_a = np.mean(vt_a)
2 print("Giá trị trung bình của vecto A là:",tb_a)
3
4 tb_b = np.mean(vt_b)
5 print("Giá trị trung bình của vecto B là:",tb_b)
6
7 tb_c = np.mean(vt_c)
8 print("Giá trị trung bình của vecto C là:",tb_c)
9
10 tb_d = np.mean(vt_d)
11 print("Giá trị trung bình của vecto D là:",tb_d)
12

🏠

🔍

⏪

⏩

🔖

🕒

📦

⚙️

...

🏠

Ps

📄

🐛

🎯

🔗

🌐

📁

🎨

EN

1:53 PM

10/13/2021

8 Hours

Facebook

HW_week

Bảng tin

Hàm List

numpy.d

Bài tập py

colab.research.google.com/drive/1AxioxrPvHBraZ5eJTtdDTY6tfnYlm-Jn#scrollTo=fLirtTkDTHnG

Laptop Python Học tập Online C++ Compiler Facebook YouTube

HW_week2.ipynb

Nhận xét

Chia sẻ

H

Tệp chỉnh sửa Xem Chèn Thời gian chạy Công cụ Trợ giúp Lưu lần gần đây nhất...

+ Mã + Văn bản

✓ RAM 0 đĩa

Chỉnh sửa

Tính giá trị trung bình mỗi cột

1 tb_a = np.mean(vt_a)

2 print("Giá trị trung bình của vecto A là:",tb_a)

3

4 tb_b = np.mean(vt_b)

5 print("Giá trị trung bình của vecto B là:",tb_b)

6

7 tb_c = np.mean(vt_c)

8 print("Giá trị trung bình của vecto C là:",tb_c)

9

10 tb_d = np.mean(vt_d)

11 print("Giá trị trung bình của vecto D là:",tb_d)

12

13 tb_e = np.mean(vt_e)

14 print("Giá trị trung bình của vecto E là:",tb_e)

15

16 tb_f = np.mean(vt_f)

17 print("Giá trị trung bình của vecto F là:",tb_f)

Giá trị trung bình của vecto A là: 214.896

Giá trị trung bình của vecto B là: 130.12150000000003

Giá trị trung bình của vecto C là: 129.9565

Giá trị trung bình của vecto D là: 9.4175

Giá trị trung bình của vecto E là: 10.6505

Giá trị trung bình của vecto F là: 140.4835

Ps

U

Chrome

EN

1:53 PM

10/13/2021

8 Hours

Facebook

HW_week

Bảng tin

Hàm List

numpy.d

Bài tập py

colab.research.google.com/drive/1AxioxrPvHBraZ5eJTtdDTY6tfnYIm-Jn#scrollTo=fLirtTkDTHnG

LaptopPythonHọc tậpOnline C++ CompilerFacebookYouTube

HW_week2.ipynb

TậpChỉnh sửaXemChènThời gian chạyCông cụTrợ giúp

Nhận xétChia sẻ

+ Mã+ Văn bản

✓ RAMỔ đĩa

Chỉnh sửa

• Vecto trung bình mua:

✓
0 giây

[17] 1 mu = np.array([tb_a,tb_b,tb_c,tb_d,tb_e,tb_f])
2 print("Vecto trung bình mua là:",mu)

Vecto trung bình mua là: [214.896 130.1215 129.9565 9.4175 10.6505 140.4835]

• Tính ma trận $B = X - h\mu^T$ với h là ma trận số 1.

✓
0 giây

▶

1 mu_T = np.transpose(mu)
2 h = np.ones(6)
3 B = data - h*mu_T
4 print("Ma trận B là:\n",B)
5

Ma trận B là:
[[-0.096 0.8785 1.1435 -0.4175 -0.9505 0.5165]
[-0.296 -0.4215 -0.2565 -1.3175 -1.1505 1.2165]
[-0.096 -0.4215 -0.2565 -0.7175 -1.0505 1.7165]
...
[-0.096 0.1785 0.4435 1.1825 0.4495 -0.4835]
[-0.196 0.5785 0.8435 1.7825 0.5495 -1.0835]
[-0.596 -0.2215 -0.0565 0.7825 0.8495 -0.8835]]

Ps

U

Chrome

1:53 PM
10/13/2021

8 Hours

Facebook

HW_week

Bảng tin

Hàm List

numpy.d

Bài tập py

colab.research.google.com/drive/1AxioxrPvHBraZ5eJTtdDTY6tfnYIm-Jn#scrollTo=fLirtTkDTHnG

Laptop Python Học tập Online C++ Compiler Facebook YouTube

HW_week2.ipynb

Nhận xét

Chia sẻ

H

Tệp chỉnh sửa xem Chèn Thời gian chạy Công cụ Trợ giúp Lưu lần gần đây nhất

+ Mã + Văn bản

✓ RAM 0 đĩa

Chỉnh sửa

- Tính ma trận covariance $C = B^TB$, sau đó tìm giá trị riêng và vectơ riêng.

```
[37] 1 B_T = np.transpose(B)
      2 C = B_T.dot(B)
      3 print("Ma trận C là:\n",C)
```

Ma trận C là:

[28.2168	6.2572	4.5952	-20.546	-3.6896	16.7768]
[6.2572	25.93755	21.57705	42.94475	20.90285	-41.65905]
[4.5952	21.57705	32.49155	56.54225	25.86935	-47.85355]
[-20.546	42.94475	56.54225	415.28875	32.74325	-206.36225]
[-3.6896	20.90285	25.86935	32.74325	128.29995	-109.37335]
[16.7768	-41.65905	-47.85355	-206.36225	-109.37335	264.21555]

- Tìm giá trị riêng và vectơ riêng.

```
[50] 1 from numpy import linalg as la
      2
      3 w, v = la.eig(C)
      4
      5 # Danh sách trị riêng
      6 print("Eigenvalues: ")
      7 print(w)
      8 # Danh sách vector riêng
```

Windows

Ps

U

Chrome

EN

1:54 PM

10/13/2021

8 Hours X Facebook X HW_week X Bảng tin X Hàm List X numpy.d X Bài tập py X

colab.research.google.com/drive/1AxioxrPvHBraZ5eJTtdDTY6tfnYlm-Jn#scrollTo=fLirtTkDTHnG

Laptop Python Học tập Online C++ Compiler Facebook YouTube

HW_week2.ipynb

Tệp Chỉnh sửa Xem Chèn Thời gian chạy Công cụ Trợ giúp Lưu lần gần đây nhất...

Nhận xét Chia sẻ H

+ Mã + Văn bản

✓ RAM 0 đĩa

Chỉnh sửa

• Tìm giá trị riêng và vectơ riêng.

```
[50] 1 from numpy import linalg as la
      2
      3 w, v = la.eig(C)
      4
      5 # Danh sách trị riêng
      6 print("Eigenvalues: ")
      7 print(w)
      8 # Danh sách vector riêng
      9 print("Eigenvectors: ")
     10 print(v)
```

Eigenvalues:
[597.06066928 186.18848402 48.43932757 38.73709015 16.95715839
7.06742059]

Eigenvectors:
[[-0.04377427 -0.01070966 -0.3263165 -0.56169177 0.75257278 0.09809807]
[0.11216159 -0.07144697 -0.25896137 -0.45545879 -0.34680082 -0.76651197]
[0.13919062 -0.06628208 -0.34473274 -0.41532963 -0.53465173 0.63169678]
[0.76830499 0.56307225 -0.2180222 0.18610822 0.09996771 -0.02221711]
[0.2017661 -0.65928988 -0.55668568 0.45069851 0.10190229 -0.03485874]
[-0.57890193 0.48854255 -0.59176285 0.25844832 -0.08445895 -0.04567946]]

Ps

1:54 PM 10/13/2021

8 Hours

Facebook

HW_week

Bảng tin

Hàm List

numpy.d

Bài tập py

colab.research.google.com/drive/1AxioxrPvHBraZ5eJTtdDTY6tfnYlm-Jn#scrollTo=fLirtTkDTHnG

Laptop Python Học tập Online C++ Compiler Facebook YouTube

HW_week2.ipynb

Tệp

Chỉnh sửa

Xem

Chèn

Thời gian chạy

Công cụ

Trợ giúp

Lưu lần gần đây nhất

Nhận xét

Chia sẻ

H

+ Mã

+ Văn bản

✓ RAM 0 đĩa

Chỉnh sửa

- Xếp giá trị riêng và vectơ riêng theo thứ tự giảm dần.

```
[51] 1 #Xếp giá trị riêng theo thứ tự giảm dần
      2 print(sorted(w, reverse=True))
      3 #Xếp vectơ riêng theo thứ tự giảm dần
      4 v.sort()
      5 print(v)
```

```
[597.0606692840747, 186.18848402065893, 48.439327574877325, 38.7370901459286, 16.95715838515131, 7.067420589]
[[-0.56169177 -0.3263165 -0.04377427 -0.01070966 0.09809807 0.75257278]
 [-0.76651197 -0.45545879 -0.34680082 -0.25896137 -0.07144697 0.11216159]
 [-0.53465173 -0.41532963 -0.34473274 -0.06628208 0.13919062 0.63169678]
 [-0.2180222 -0.02221711 0.09996771 0.18610822 0.56307225 0.76830499]
 [-0.65928988 -0.55668568 -0.03485874 0.10190229 0.2017661 0.45069851]
 [-0.59176285 -0.57890193 -0.08445895 -0.04567946 0.25844832 0.48854255]]
```

- Tính tỉ lệ phần trăm tích lũy giá trị riêng so với tổng tất cả giá trị riêng rồi so sánh với ngưỡng PCA (thường $\geq 90\%$).

1

7 - Paint

EN 1:55 PM 10/13/2021