

Bài thực hành 1. LẬP TRÌNH PYTHON CƠ BẢN

1. CÁC KIỂU DỮ LIỆU CÓ CẤU TRÚC TRONG PYTHON

Array là một cấu trúc đặc biệt dùng để biểu diễn cho 1 vector. Python cung cấp thư viện Numpy để xử lý đối với dữ liệu dạng vector.

Để sử dụng Numpy, ta phải import vào bằng lệnh:

```
import numpy as np
```

Chiều của các vector có thể khác nhau:

- 1-D: `arr = np.array([1, 2, 3, 4, 5])`
- 2-D: `arr = np.array([[1, 2, 3], [4, 5, 6]])`
- 3-D: `arr = np.array([[[1, 2], [4, 5]], [[1, 3], [5, 6]]])`

Ngoài ra, kiểu dữ liệu **set** và **tuple** cũng khá thường được sử dụng để biểu diễn các dữ liệu có cấu trúc phức tạp.

- Ví dụ về **set**: `s = {1,1,-2,3,1,4,6}`
- Ví dụ về **tuple**: `t = (1,2)`

Điểm khác biệt chính giữa array và set:

- Các phần tử trong **array** có thứ tự, và có thể trùng nhau.
- Các phần tử trong **set** không có thứ tự, và không trùng.

Có thể tận dụng đặc điểm của list và set để **loại bỏ phần tử trùng nhau**.

Ví dụ:

```
a = np.array([1, 2, 3, 1, 1, 2, 4, 5, 5])
```

```
b = list(set(a))
```

Kết quả sẽ là: 1,2,3,4,5.

Mặt khác, tuple được dùng để hoán vị giá trị của 2 biến với nhau. Cụ thể:

```
a = 2
```

```
b = 3
```

```
(a,b) = (b,a)
```

Kết quả sẽ là: a = 3 và b = 2.

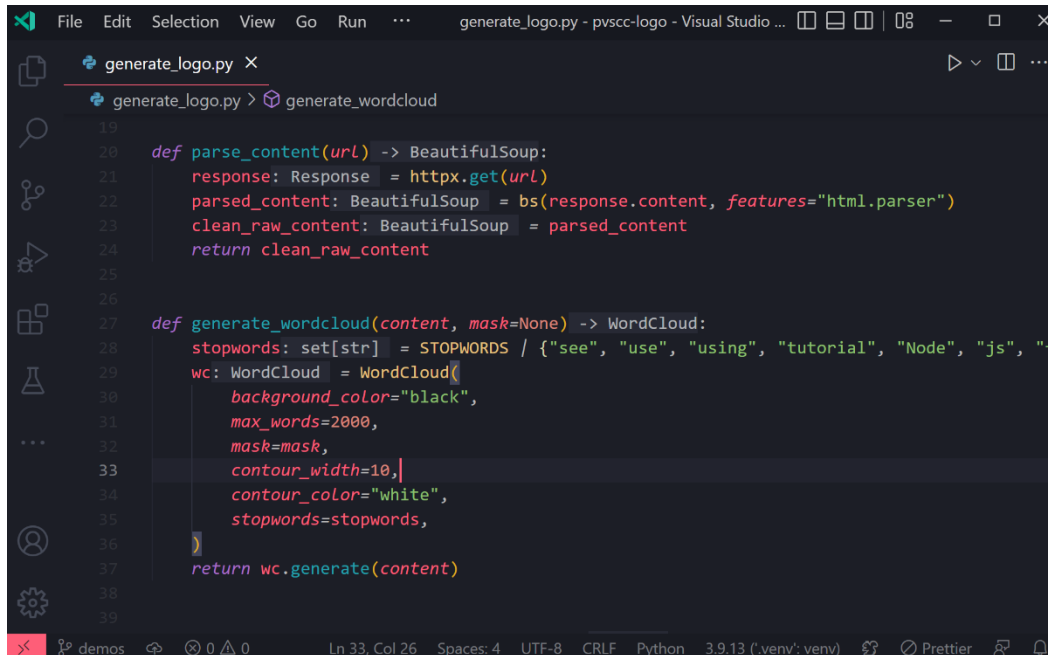
2. MÔI TRƯỜNG PYTHON VÀ CÁC THƯ VIỆN LẬP TRÌNH

Có 2 dạng môi trường chính để làm việc với Python:

- Dạng 1: File code thuần (có đuôi .py).
- Dạng 2: File code kết hợp văn bản trình bày (có tài liệu gọi là nhúng code vào văn bản).

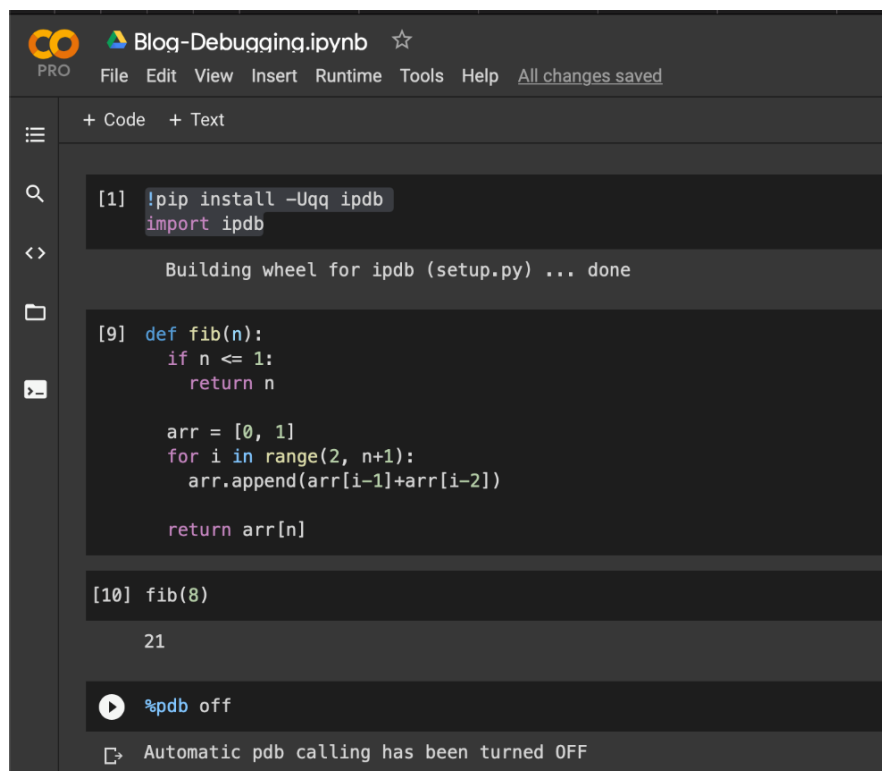
Các IDE thường dùng để code Python gồm:

- Dạng 1: PyCharm, Visual Studio, Visual Studio Code.
- Dạng 2: Google Colab, Jupyter Notebook.



```
19
20 def parse_content(url) -> BeautifulSoup:
21     response: Response = httpx.get(url)
22     parsed_content: BeautifulSoup = bs(response.content, features="html.parser")
23     clean_raw_content: BeautifulSoup = parsed_content
24     return clean_raw_content
25
26
27 def generate_wordcloud(content, mask=None) -> WordCloud:
28     stopwords: set[str] = STOPWORDS / {"see", "use", "using", "tutorial", "Node", "js", "f
29     wc: WordCloud = WordCloud(
30         background_color="black",
31         max_words=2000,
32         mask=mask,
33         contour_width=10,
34         contour_color="white",
35         stopwords=stopwords,
36     )
37     return wc.generate(content)
38
39
```

Viết mã nguồn Python trong Visual Studio Code



```
[1] !pip install -Uqq ipdb
import ipdb

Building wheel for ipdb (setup.py) ... done

[9] def fib(n):
    if n <= 1:
        return n

    arr = [0, 1]
    for i in range(2, n+1):
        arr.append(arr[i-1]+arr[i-2])

    return arr[n]

[10] fib(8)

21

%pdb off

Automatic pdb calling has been turned OFF
```

Viết mã nguồn Python trong Google Colab

Các thư viện lập trình thường dùng trong Máy học gồm:

- **Numpy**: Thư viện xử lý ma trận và vector.

- **Sklearn:** Thư viện thực hiện các mô hình máy học cơ bản, và các độ đo khác nhau.
- **Matplotlib:** Thư viện vẽ đồ thị.
- **Pandas:** Thư viện thao tác với dataframe. Hỗ trợ thao tác đọc và xử lý dữ liệu.
- **Tensorflow:** Thư viện hỗ trợ lập trình xử lý ma trận và vector phục vụ cho xây dựng mạng neural.

Để cài đặt thư viện trong Python, ta dùng lệnh:

`pip install <tên thư viện>`

3. LẬP TRÌNH HÀM

Cách lập trình sử dụng hàm trong Python:

```
def <tên hàm> (<tham số 1>, ..., <tham số n>):  
    // code xử lý  
    return <giá trị> (nếu có)
```

4. VÍ DỤ MINH HỌA

Sinh viên xem trong file Jupyter Notebook do giảng viên cung cấp.

5. BÀI TẬP

Sinh viên làm bài tập trên Wecode do giảng viên cung cấp.