

Numpy배열 vs Python list 배열

202411325이충무

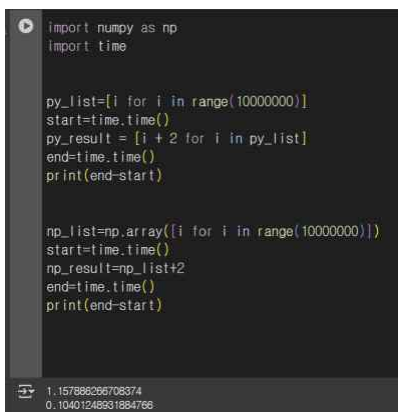
-Numpy

Numpy는 행렬 계산을 위한 파이썬 라이브러리의 모듈이다. Numpy는 C언어로 작성되어 있으며 Python list는 데이터가 흩어져 있는 반면에 Numpy는 연속되어 있는 공간에 데이터가 저장되어 있어 데이터를 찾기에 더 용이하다.

-구현 내용

해야할 것은 Numpy 배열과 Python list 배열을 생성하고 각 요소에 2를 더하는 코드를 구현하고 이에 따라 걸리는 실행시간 비교이다.

-Numpy 배열과 Python list 배열의 구현과 실행 결과



```
import numpy as np
import time

def list_operation():
    py_list=[i for i in range(10000000)]
    start=time.time()
    py_result = [i + 2 for i in py_list]
    end=time.time()
    print(end-start)

def numpy_operation():
    np_list=np.array([i for i in range(10000000)])
    start=time.time()
    np_result=np_list+2
    end=time.time()
    print(end-start)

list_operation()

numpy_operation()
```

1.157896266708374
0.10401248931884796

-실행 결과에 대한 분석

실행 결과 Python list를 사용한 것보다 Numpy 배열을 사용한 것의 시간 복잡도가 10배 이상 더 좋게 나온 것을 볼 수 있다. 이러한 이유로는 여러 가지 이유를 들 수 있다. 첫 번째, Python list를 사용한 것을 보면 for문을 사용해서 2를 각 요소에 반복해 넣어 주었다. 반면에, Numpy 배열을 사용하여 각 요소에 2를 더해준 것을 보면 단 한번의 연산으로 각 요소들에 2를 더해줄 수 있다. 이러한 연산 횟수의 차이로 인해 실행시간에 대한 차이가 난다. 두 번째, Numpy 배열이 C언어를 사용하는 것이다. C언어는 Python에 비해 저수준의 언어이다. 때문에 컴퓨터가 연산하는데 Python list보다 C언어를 사용하는 Numpy list가 더 용이하다. 세 번째, Numpy의 가장 큰 특징인 연속된 공간에 요소들을 저장하는 것이다. 이렇게 연속된 공간에 저장하게 되면 연산할 때 컴퓨터의 데이터에 대한 접근 속도가 빨라져 더 좋은 시간 복잡도를 가지게 된다.