

영어음성학 기말고사 필기

19/10/29

1. 데이터 - 소리/이미지/텍스트/숫자(주식, 온도, 날씨 등)의 4가지로 나뉘어짐.

- 여기서 숫자화 되지 않은 데이터들(소리, 이미지, 텍스트 데이터)이 어떻게 숫자화가 되는가?

- 행렬

- (영상이나 사진에서) 숫자들이 직사각형의 형태로 행(가로)과 열(세로) 속에 쳐 놓은 것이다.
- 예를들어 흑백 사진에서는 검정색이 0, 흰색이 10으로 포함된다면 회색은 1~9의 숫자값으로 직사각형의 형태로 배열되어 있는 것이다. - 2차원
- 컬러 사진일 경우에는 직사각형의 행렬이 여러 장이 있는 것이다. (예를 들어 RGB라면 한 직사각형 판은 얼마나 빨간 정도를 가지고 있는가, 다른 판은 초록, 다른 판은 파란 정도를 알려주는 것이다) - 3차원
- 여기에 시간까지 추가가 된다면 4차원이 되는 것이다.

- 벡터

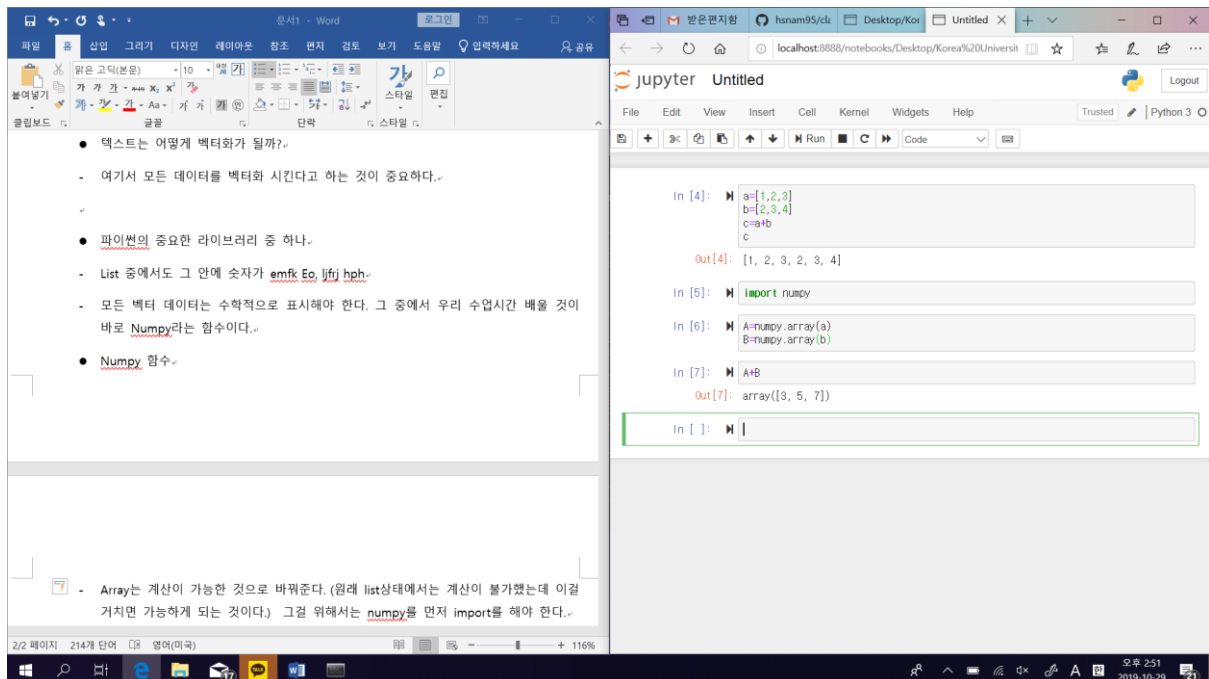
- 모든 데이터는 벡터의 형태로 되어야 한다. 행렬은 벡터는 아니다. 왜냐하면 직사각형 형태이기 때문이다.
- 이미지는 행렬이고, 이 것을 길게 쭉 늘어 놓은 것이 벡터라고 한다. (한 줄로)
- 모든 데이터는 벡터화 했을 때 다루기가 쉽다.
- 소리는 어떻게 영상과 비슷할까? (어떻게 벡터화가 되는 것일까?)
- 텍스트는 어떻게 벡터화가 될까?
- 여기서 모든 데이터를 벡터화 시킨다고 하는 것이 중요하다.

- 파이썬의 중요한 라이브러리 중 하나

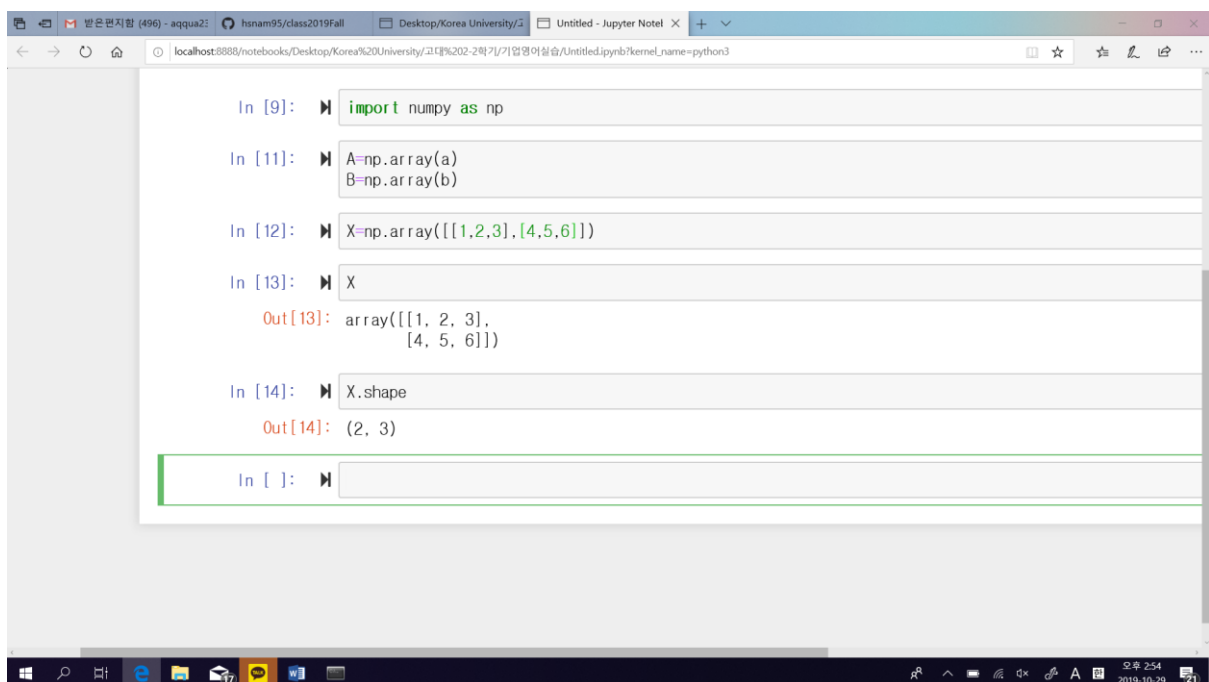
- List 중에서도 그 안에 숫자가 emfk Eo, ljfrj hph
- 모든 벡터 데이터는 수학적으로 표시해야 한다. 그 중에서 우리 수업시간 배울 것이 바로 Numpy라는 함수이다.

- Numpy 함수

- Array는 계산이 가능한 것으로 바뀌준다. (원래 list상태에서는 계산이 불가능했는데 이걸 거치면 가능하게 되는 것이다.) 그걸 위해서는 numpy를 먼저 import를 해야 한다.



- X.shape은 차원을 말해주는 것이다.



- 교수님 github 계정 들어가서 미리 예습 해보자.

19/10/31

1. Numpy / Import / From

- Numpy가 가장 큰 집합을 의미하는 것이다.
- 줄여서 쓰고 싶을 때는 `import numpy as np`라고 해서 밑에서는 계속 `np`라고 쓸 수 있다.
- `""` 온점은 포함관계로 되어 있는 것을 의미한다. (집합 안에 집합을 지칭)
- Numpy는 list와 비슷한데, 쓰는 이유는 수학적인 계산이 용이하기 때문이다. (따라서 이제는 list 대신에 numpy를 더 많이 써야 한다.
- 온점이 뒤에 적혀있다는 것은 데이터 타 | 입은



- Numby

`Np.linspace (0,10,6)` 0부터 10까지 6개로 나눠본다.

그래서 그 후의 값이 나오는 것이다.

Linear space는 공차(d-difference space)들이 같기 때문에 본다.

- Type 이야기를 했고.
- Type을 바꾸고싶으면 X.

2. Plotting

3. Hist = histogram. 바구니를 말하는 것이다. 히스토그램.

4. 여기까지가 중요했다. 그다음은 3번 Numpy I/O가 있는데 그게 또 중요하다. 알아서 니가 해 라고 할 때 -1이 된다. In [38]번에서! 괄호 맨 처음에 있는 숫자가 -1이든 4이든 값은 4로 같게 나온다.

5. 3번 Numpy I/O 바로 위에 있는 In [16]인가 그거는 할 필요 없다.

6. `savez`라는 함수를 쓰면 실제로 file로 저장이 되는 것이다.

7. In [42]번 ! 내 파이썬 사이트로 k서

- 이번줄 말에 복습해서 메모해서 정리하는 거 한번 해봐야 한다. 그리고 혼자서 정리 및설명을 할 수 있을 정도로 해라.
- 그리고 교수님 gitup 나중에 가서 수업자료 한번다운받아서 열어보기. 데이터를 서로 서로 전달을 할 때 데이터 숫자들을 파일로 전환해서 주는 거도.

8. 4번 Inspecting의 소제목이 등장!

9. type으로는 `numby.np.array`를 의미한다. 제일 위예가 ~이다.

10. Arrange 하나만 나오면 index가 나온다. 2개가 나오면 ~ 3개가 나오면 ~ (9부터 5까지 9,8,7,6으로 4개가 도출되는 것이다)

- In[9], [65], [13]은 다음에 알려주겠다고 했다.
- 5.2 Comparison은 수업 진행한다.

11. In [67], [68], [69]

- In [67] – np.arange랑 reshape.
- A=b 같으냐? 라고 묻는 것은 등호 2개 ==를 쓴다고 함. 질문형인 것임.
- 5.3번 Aggregate 는 매우 매우 중요한 것이다. 그래서 엄청 잘 봐야 한다.
- In [49]는 a / 70번에 a는 이미 처리 완료.

12. 함수면 반드시 괄호를 쳐야 한다. 자기 자신을 sum 하라고 하니까 일단 그렇게 한다. 뭘 sum을 해야 하는가 하면은 ... 이 것 쓰는 방법을 정확하게 이해해야 한다.

13. A가 3곱하기 3인데 이게 첫번째 차원에 해당되는게 가로행을 의미하는 것이고 두번째 차원은 세로열을 의미하는 것이다. 그래서 그 관점에서 실행을 하라고 명령을 하는 것이다.

- Broadcasting

- We-shape가 broadcasting의 일종이다.
- [52번에서] A의 차원이 5 이다
- 여기까지가 numpy고 다음시간에는 sound를 본격적으로 다룰 것이다. 엄청나게 import를 많이 한다. 많이 알아야 한다.
- 하나만 언급하고 넘어가겠다. sound라는 것이 continuous한 것인데 컴퓨터에 담으려면 하나의 값으로 담겨야 한다. 그래서 1초동안 얼마나 많은 값들이 뻑뻑하게 혹은 들성들성 있는 정도가 sampling rate라고 한다.
- 만약 10000이라고 rate를 해놓았다면 1초에 10000개의 값이 있는 것이다. Sampling rate 는 4만4천100 정도는 옛날의 CD같은 건데 사람이 구분할 수 있는 가장 높은... 것이다.
- 1초를 표현하는 데 소리를 표현할 수 있는 숫자를 의미한다. 그래서 이 개념을 사용하면서 다음 시간부터 적용할 것이다. Sampling rate를 잘 이해하고 오세요.