**[오픈소스개발프로젝트 2021F]**

**출석 과제**

주의사항: 본문 글자 크기, 자간, 장평 등 서식 변경 금지

작성자: 홍진우 학번: 2018037044

|  |
| --- |
| 해당 주차의 동영상을 시청하고 내용을 요약하시오. (공백포함 600자 이상) |
| 이번주차에서 크게 json, numpy, matplotlib, pandas, scikiet-learn에 대해서 알아보았다. Json은 자바스크립트 기반의 웹 데이터이고 예전에 썼던 xml의 문제점 때문에 요즘은 json을 많이 사용한다. 파이썬의 리스트, 딕셔너리를 이어 붙인 것을 json이라고 한다. 저번에도 배웠지만 json파일을 가져와서 읽을 때 read, readline, readlines 방식이 있는데 read는 전체 readline은 한 줄을 뽑아서 읽는다. json파일에서 특정한 인덱스만 표시하고 싶을 때 특정 기호를 지정해서 추려낼 수 있다. numpy에서는 가장 기본 데이터 타입인 행렬이 있다. numpy에는 행렬의 차원을 나타내 주기도 하고 1행 4열 행렬을 2행 2열로 바꿔주고 전치행렬로 변환해 줄 수도 있다. [ : , :]를 이용하여 행렬의 데이터를 접근 할 수도 있다. 앞에 것이 행이고 뒤에 것이 열로 접근할 때 사용한다. 행렬을 계산할 수도 있다. Mathplotlib을 사용하여 우리가 수학시간에 배웠던 함수의 그래프를 그릴 수 있다. x, y 좌표쌍을 이용하여 print하면 좌표쌍이 그대로 나열되지만 plot을 그리면 함수의 그래프가 만들어지고 scatter을 사용하면 점선으로 표현 할 수 있다. Panda는 dataframe에서 데이터로 접근할 때 사용되는데 원하는 행이나 열을 필터해서 자기가 원하는 것만 볼 수 있다. 이번 수업시간엔 필요한 데이터들을 추출하는 방법에 대해서 살짝 맛을 본거 같다. |
| 동영상 내용에서 질문 사항이 있는 경우 작성하시오. |
| 1)  2) |