

초급2팀

머신러닝 개괄 & Keras

머신러닝 개괄

- 학습에는 지도학습 (supervised) 비지도학습(unsupervised) 이 있습니다.
- 지도학습 —> 고양이인지 개인지 나누는데 데이터를 개&고양이 이름과 같이 주고 학습 시키는 것.
- 비지도학습 —> 그냥 사진만 주고 스스로 개와 고양이를 분류 하게 하는 것.
- 저희 조는 지도학습으로 !!

머신러닝 개괄

- 우리가 가진 것 적당한 시간
- 우리가 없는 것 돈 그리고 자원
- 우리가 공부해야 할 방향 -?
- 이미 적당히 가공되어있는 데이터가 필요.

데이터 전처리

- 제한된 컴퓨팅 파워와 시간을 적절히 활용하여, 좋은 결과를 내기 위해서는 데이터 전처리(preprocessing)가 필요하다.
- 주어진 데이터의 값이 다른 데이터값과 차이가 너무 큰 경우 학습이 제대로 되지 않을 가능성이 높다.

데이터 전처리 예시

- 정규화
- 경향성
- 데이터 채우기/빼기
- 데이터 통합 등

데이터는 어디서 구하기 좋은가?

- 공공 데이터 포털
- data.world
- kaggle

논의한 예시 주제

- 전쟁 데이터 기반으로 전쟁 승패를 예측하기
 - 새로운 현대전 무기의 반영이 어려움
- 인스타에서 사진 크롤링하여 새로운 사진에 대하여 태그추가
 - api제공 안될것 같음.
- 날씨 풍속 풍향 계절 지역정보를 기반으로 미세먼지 예측
 - 아직까지는 태클 안들어왔음
- NLP를 활용한 자연어 처리
 - 아직 실력이 안될 것 같음

케라스?

- Keras, 이제 텐서플로우 위에도 쓸수있습니다. 이득.
- 텐서플로우를 바탕으로 빠르게 모델링을 해 볼 수 있는 프레임워크
- 텐서플로우 5줄 -> 케라스 짧은 1줄.
- 대개 최적화 되어있으나, 아주 세세한 최적화를 하려면 텐서플로우로 짜거나 ...뭔가 다른걸 해야함
- 하지만 우리 수준에서는 그 정도의 최적화를 할 문제는 없을 것이라 판단