

선형 회귀 적용하기

환경과 집값의 변동을 보여주는 데이터셋 -> 머신러닝의 선형 회귀를 테스트하는 가장 유명한 데이터로 쓰이고 있음

이번에는 하나의 정답을 맞히는 것이 아니라 수치를 예측하는 문제!

-> 주어진 환경 요인과 집값의 변동을 학습해서 환경 요인만 놓고 집값을 예측하는 것!

머신러닝 혹은 딥러닝을 위해 주어진 데이터의 답을 구하는 문제는 크게 두 유형의 모델이 있음 -> 여러 개 중에 정답을 맞히거나 (분류), 성적 같은 수치를 맞히는 것 (회귀)

1. 데이터 확인하기

(실습코드는 github 참고)

주어진 데이터는 총 샘플의 수가 506개이고, 칼럼 수는 14개이므로 13개의 속성과 1개의 클래스로 이루어져 있음

마지막 칼럼은 클래스로 구분된 게 아니라 가격이 나와 있음

2. 선형 회귀 실행

선형 회귀 데이터는 마지막에 참과 거짓을 구분할 필요가 없음. 출력층에 활성화 함수를 지정할 필요도 없음

flatten() 함수 - 데이터 배열이 몇 차원이든 모두 1차원으로 바꿔 읽기 쉽게 해 주는 함수
range('숫자') - 0부터 '숫자-1'만큼 차례대로 증가하며 반복되는 값을 만들

참고 자료

<http://daplus.net/python-pandas%EC%97%90%EC%84%9C-%EA%B3%B5%EB%B0%B1%EC%9C%BC%EB%A1%9C-%EA%B5%AC%EB%B6%84-%EB%90%9C-%EA%B0%92%EC%9C%BC%EB%A1%9C-%ED%8C%8C%EC%9D%BC%EC%9D%84-%EC%9D%BD%EB%8A%94-%EB%B0%A9%EB%B2%95/>

https://velog.io/@u_jinju/Python-flatten-%ED%95%A8%EC%88%98