

파라미터(원료, 연료): 석회석, 점토, 규석, 산화철 연료, 석고, 석탄

레이블: 이산화탄소 양

데이터를 생성할 때 위와 같은 비율을 바탕으로 랜덤으로 생성. 석회석에서 CO2가 가장 많이 발생한다고 알려져 있으므로 석회석의 양에 비례해서 이산화탄소 양이 결정되게끔 데이터셋을 생성.

(대외적)

실제 산업에서 센서로 데이터를 받아올 수 없으므로 가상으로 데이터를 생성함.

가상 데이터이지만 최대한 산업에서 직접 적용할 수 있을 만한 전문적인 지식을 바탕으로 데이터를 생성함.

전문가에게 원료나 연료 투입 비율, 그에 따른 이산화탄소 배출량 등과 같은 지식을 얻어서 데이터 셋을 생성함.

이러한 가상 데이터로 일단 모델을 만들고 모델이 실질적으로 사용가치가 있음을 증명한 다음 실제로 산업에 쓰이게 되면 그때 실제 데이터를 가지고 모델을 다시 만듬. (가상데이터 폐기)