

Part. 05

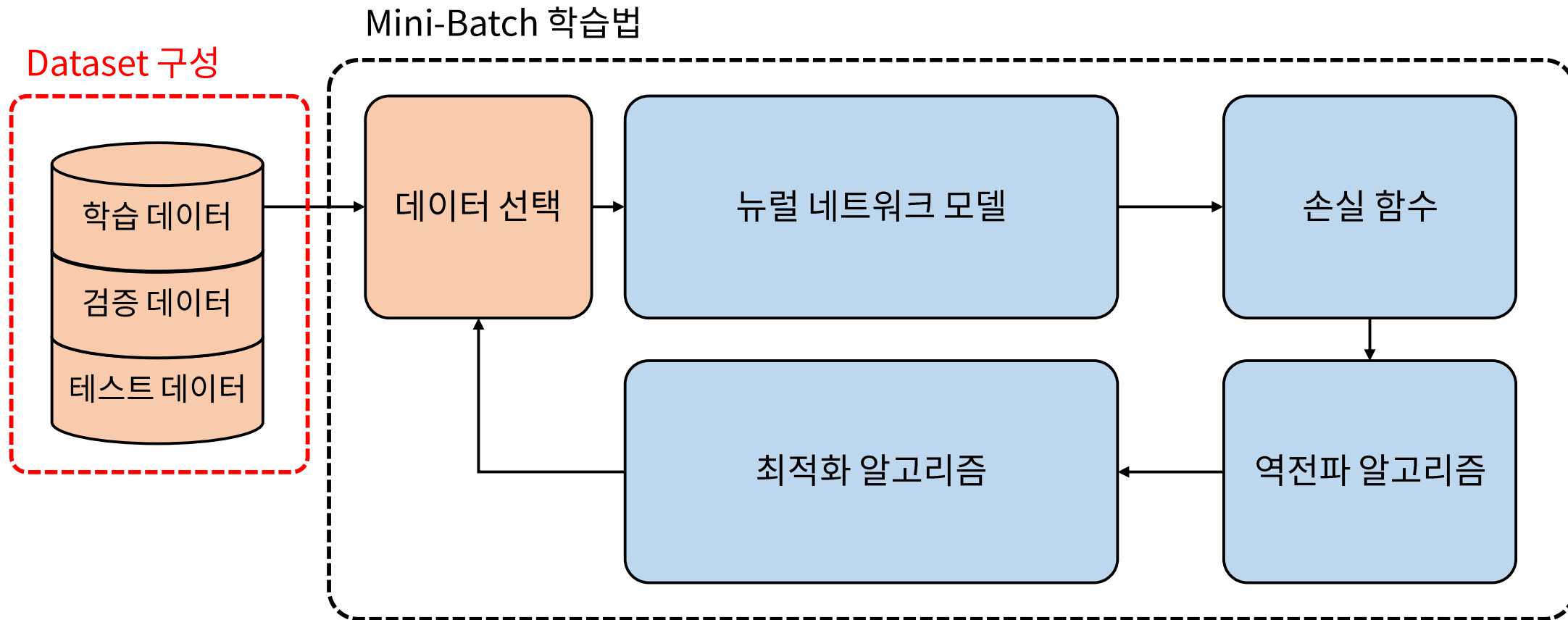
Training Neural Network

| 데이터셋의 이해

FASTCAMPUS
ONLINE

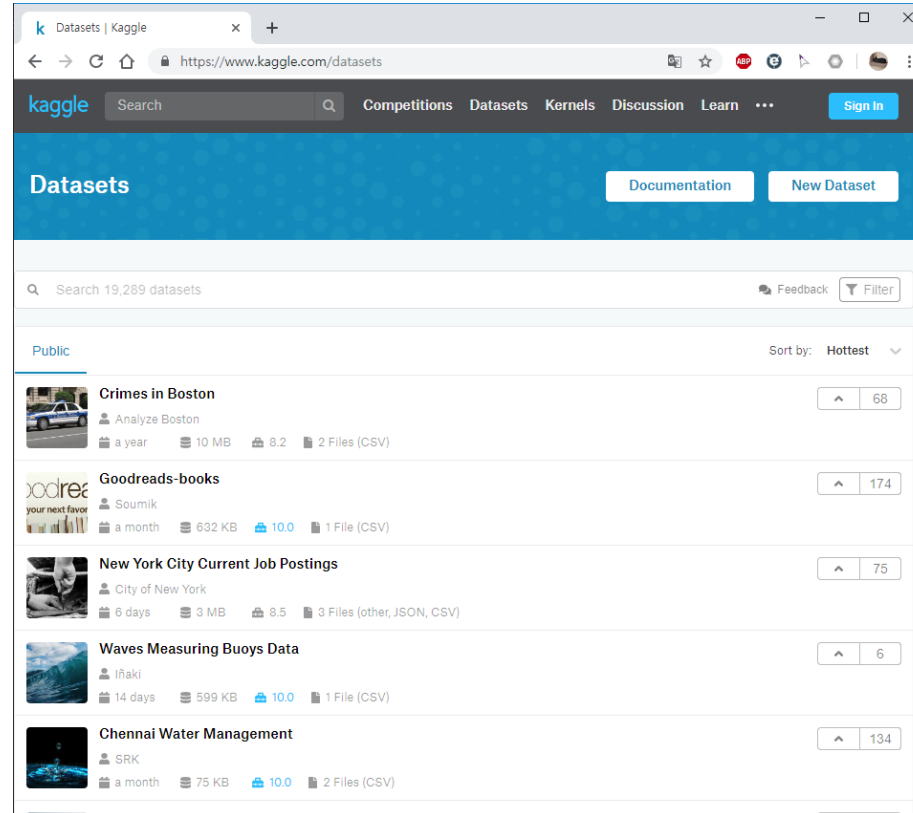
강사. 신제용

I 뉴럴 네트워크의 학습



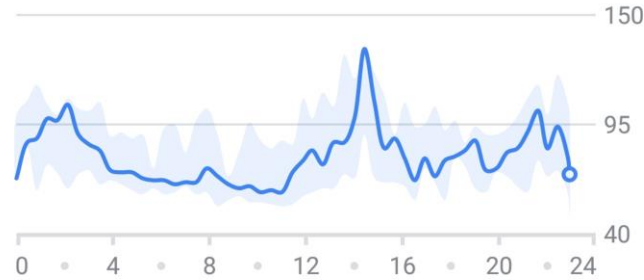
뉴럴 네트워크의 학습 과정

I 공개 데이터셋



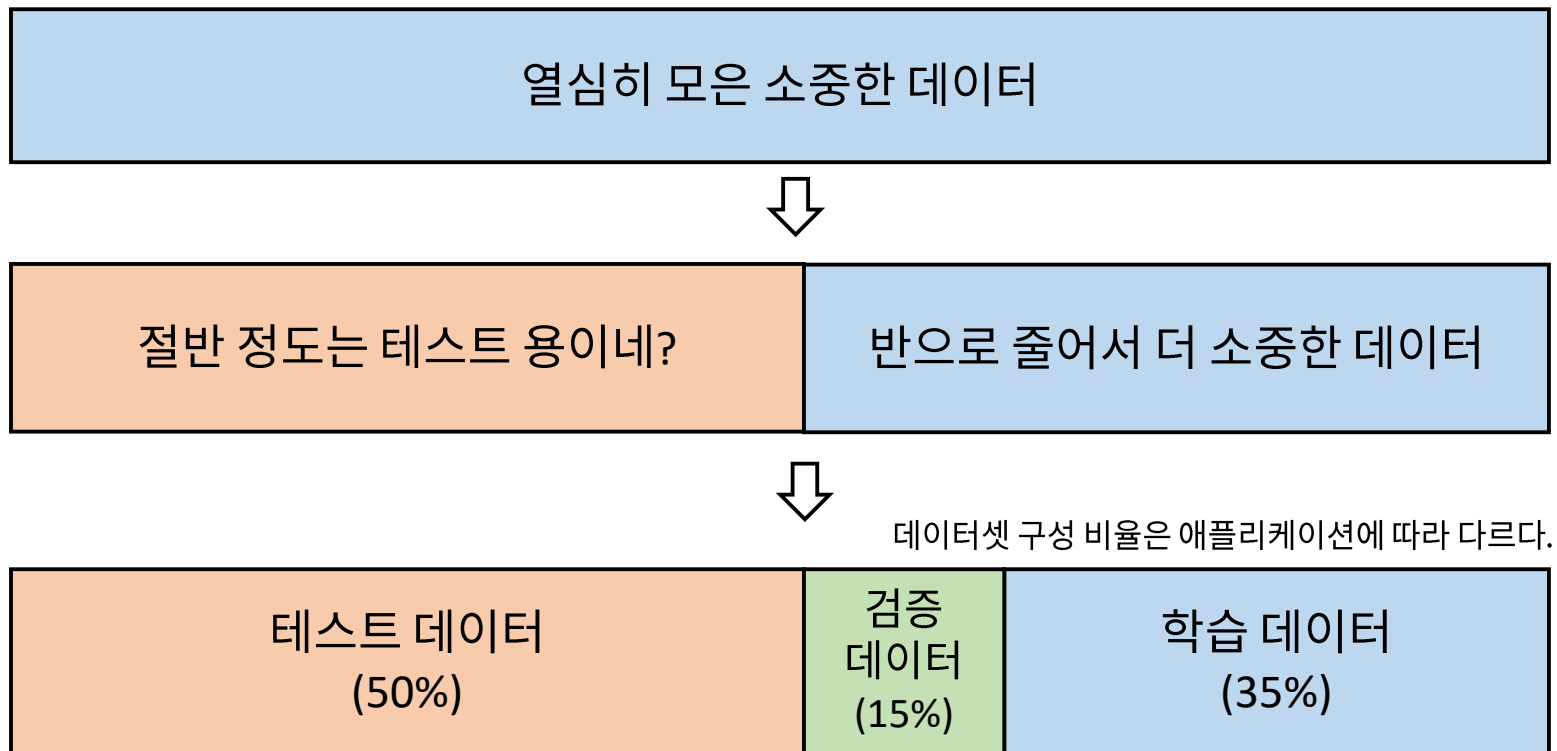
Data Scientist의 성지인 Kaggle에서는, 수많은 연구용 데이터셋을 구하고
현상금이 걸린 Challenge에 참여할 수 있는 기회가 주어진다.

I 소중한 산업용 데이터



산업 현장에서는 **무수히 많은 신호가 발생**한다. 하지만 그 중 극히 일부가 취득되어 데이터가 되며, 취득된 데이터 중 **딥러닝 학습에 유용한 데이터를 선별하는 일은 절대 간단하지 않다.**

I 데이터셋의 구성



사용 가능한 데이터셋은 매우 소중하다. 소중한 데이터를 모두 학습에 쓰지 않고, 나누어 사용하는 데에는 분명 다 이유가 있다.

I 각 데이터 셋의 의미

학습 데이터(Training Data) : 학습 과정에 보여지고, **실제 모델을 학습하는 데에 사용**되는 데이터

검증 데이터(Validation Data) : 학습 과정에 보여지는 데이터이지만, 모델 학습에 사용하지 않고 **학습이 잘 되고 있는지 검증**하는 데에만 사용

테스트 데이터(Test Data) : 학습 과정에서는 보여지지 않고, **학습을 마친 모델을 평가**하기 위해 단 한번만 보여지는 데이터

I 학습, 검증, 평가

