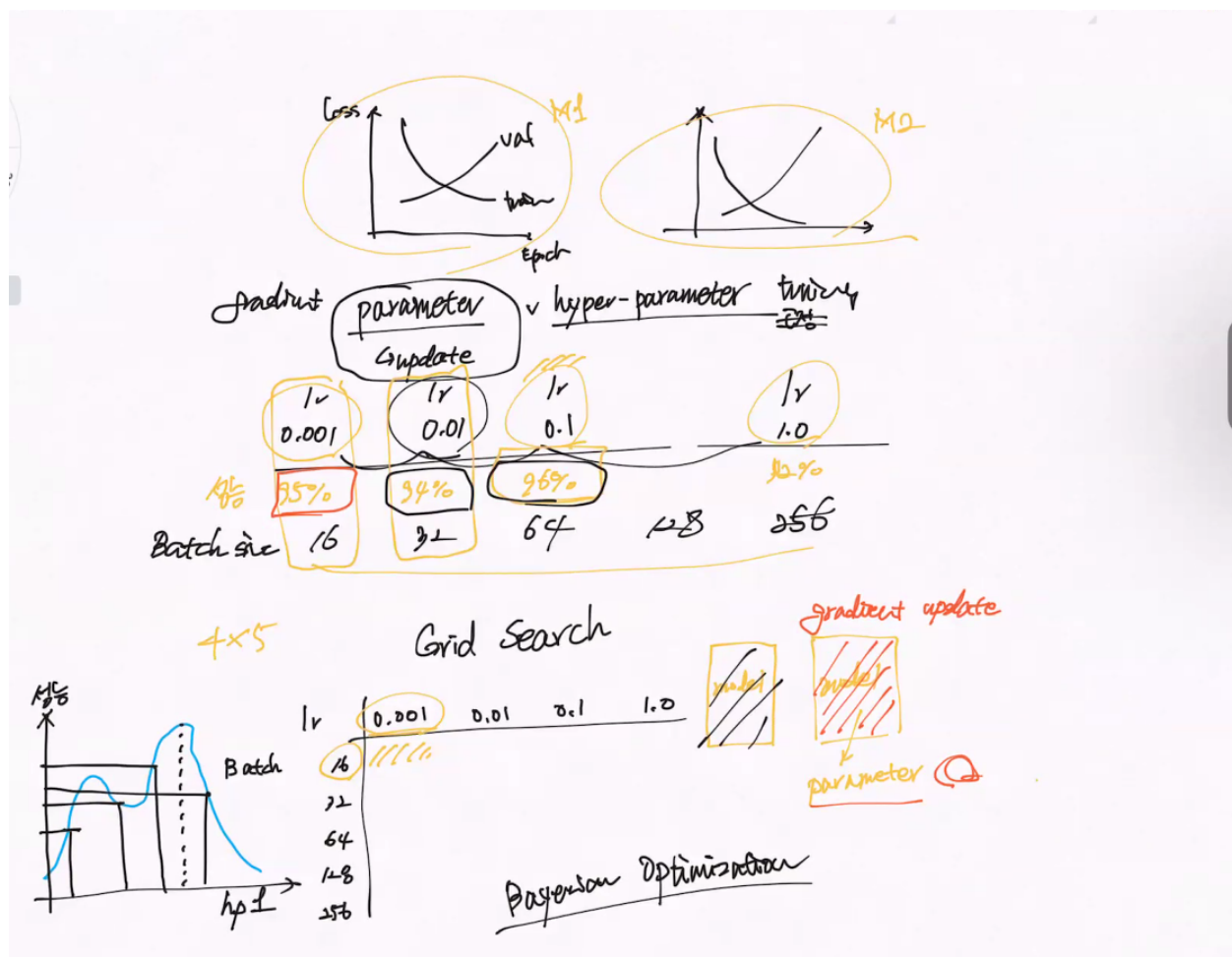


하이퍼파라미터 튜닝 (Hyperparameter tuning): 학습률, 배치 크기, Grid search

practice 인공지능,머신러닝 · 2024. 10. 5. 19:24

Q. 딥러닝에서 하이퍼파라미터 튜닝(Hyperparameter Tuning)이란 무엇이며, 대표적인 하이퍼파라미터에 대해 설명해 주세요.



일반적으로 모델을 학습한다는 것은, parameter를 업데이트 하는 것이다 (주어진 데이터로 역전파 등 사용)

모델의 학습 과정 전체에 영향을 미치는 값들이 있다. 예를 들어 학습이 얼마나 빨리 진행되는지 (학습률), 데이터를 몇번 학습할지 (epoch) 등이 있다

이러한 값들을 Hyper Parameter 라고 하며, 처음에 설정되어 일반적으로 학습 도중에는 변하지 않는다.

이때, 모델을 잘 학습시키기 위해서는, 적절한 Hyper parameter 의 설정이 필수적이다.

-> 적절한 HP 를 어떻게 찾을 수 있을까?

동일한 인공지능 모델을 여러개 만드는데 (동일한 학습 데이터에서 학습), 이 과정에서 각 모델의 HP를 바꾼다

예시: 학습률 0.01 사용한 모델 1, 학습률 0.02 사용한 모델 2, 학습률 0.03 사용한 모델 3 등등

이 모델들은 다른 것은 모두 동일하지만 (학습 과정 등), 딱 학습률만 다르게 설정되어 있다.

-> 학습률이 다르니까, 동일한 데이터로 동일하게 학습했지만 결과는 다르게 나올 것

(이 과정에서 Batch size, Epoch, 등은 임의의 값으로 설정한 후 실험)

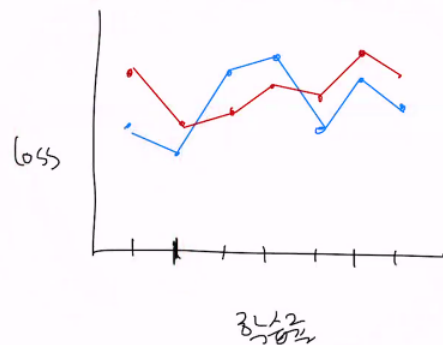
1. 모델의 학습에는 여러가지 HP가 있다
2. 적절한 HP를 선택하는 것은 중요하다
3. 처음에는 적절한 HP 값을 모르므로, 그냥 단순하게 모든 HP들을 기본값 (임의의 값)으로 설정한다 (파이토치 등에서 기본으로 설정된 값 등)
4. 그리고 HP 하나씩 여러가지로 바꾸어 보면서 (다른 것은 모두 동일한) 모델들을 학습한다 (동일한 학습 데이터에서)

예시: Batch size, Epoch 등은 기본값, 학습률은 0.01, 0.02, 0.03, 0.04 등으로 바꾸어 가며 여러개의 모델을 학습

5. 이후 여러 모델들 중, 가장 좋은 성과를 낸 모델을 찾는다. 그리고 이 모델의 학습률을 이후에 사용한다 (고정)
6. 동일한 과정을 반복해서, 다른 HP들 (Batch size, Epoch 등)의 최적의 값을 찾는다

아래그림) 모델을 학습한 후, 검증 데이터셋에 테스트를 한 후, 가장 낮은 loss(혹은 가장 좋은 결과)를 보여준 hyperparameter 값 선택

모델을 학습한 후, 검증 데이터셋에 테스트를 한 후, 가장 낮은 loss (혹은 가장 좋은 결과)를 보여준 hyperparameter 값 선택



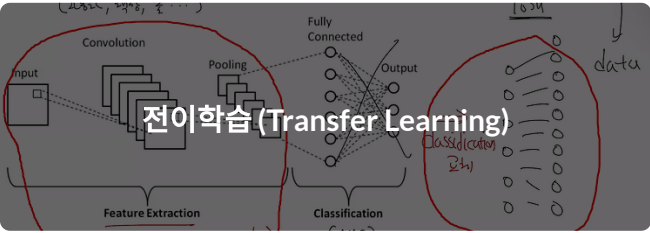
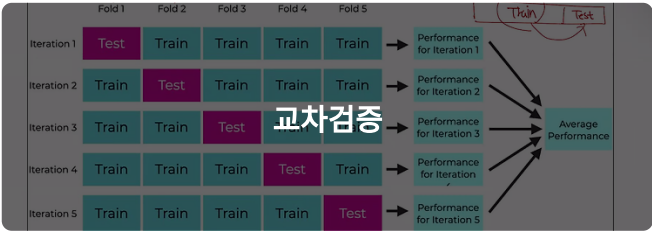
이러한 학습을 멈추고
→ hyper parameter

♡ 공감 📌

구독하기

'practice 인공지능,머신러닝' 카테고리의 다른 글

교차검증 (0)	2024.10.11
전이학습 (Transfer Learning). (0)	2024.10.11
Autoencoder (0)	2024.10.04
데이터 불균형 (Data Imbalance): Oversampling, Undersampling (1)	2024.09.28
transformer (1)	2024.09.28



Autoencoders

- = Auto-associators
- = Diabolo networks
- = Sandglass-shaped net

Autoencoder

데이터 불균형 (Data Imbalance):
Oversampling, Undersampling

multi class classification

자연어(NLP)

네이쳐2024 님의 블로그입니다.

구독하기 +

댓글 0



이름

비밀번호

내용을 입력하세요.



로그인