

1. Scheduler 소개

Cosine Annealing Scheduler

정의: 총 학습 횟수 내에서 lr을 조절하는 방법 중 하나로, Cosine 함수를 기반으로 한 이 scheduler는 학습 진행도에 따라 lr을 주기적으로 감소시킨다.

작동원리: 학습 초기에는 높은 lr로 시작하여, 각 epoch마다 lr을 cosine 함수의 형태에 따라 감소시킨다. 그러나 특정 주기마다 lr을 초기값으로 재설정하며, 이러한 조정은 local minimum을 피하고, global minimum을 찾는 데 도움을 준다.

위의 scheduler 사용 시 모델이 다양한 학습 속도를 겪으며, local minima에 갇히지 않도록 돕는다.

2. Training Error와 Generalization Error 사이의 간극을 줄이는 방안

Early Stopping: validation dataset 내에서 성능이 개선 되지 않을 때 학습을 조기에 중지시킨다. 이는 모델이 train dataset에 overfitting되는 것을 방지한다.