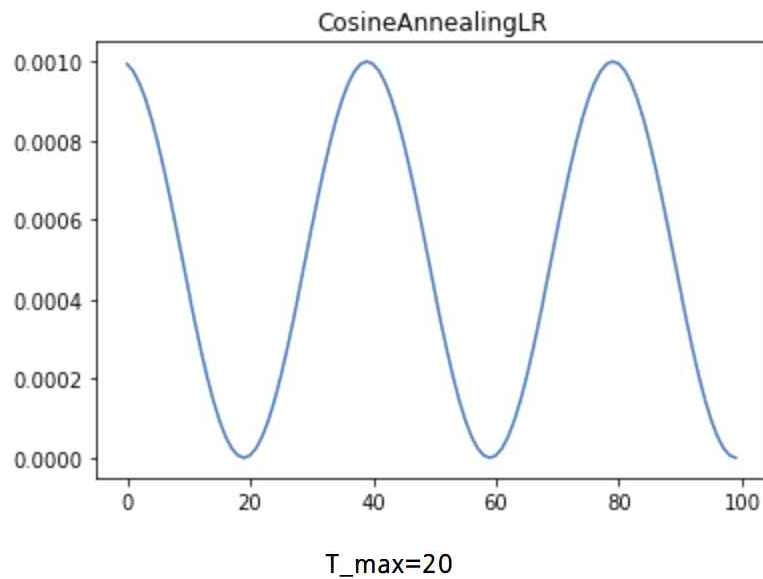


## Tobigs 6주차 과제 (Neural Network Advanced) - 21기 이소영

### Learning Rate Scheduler 소개

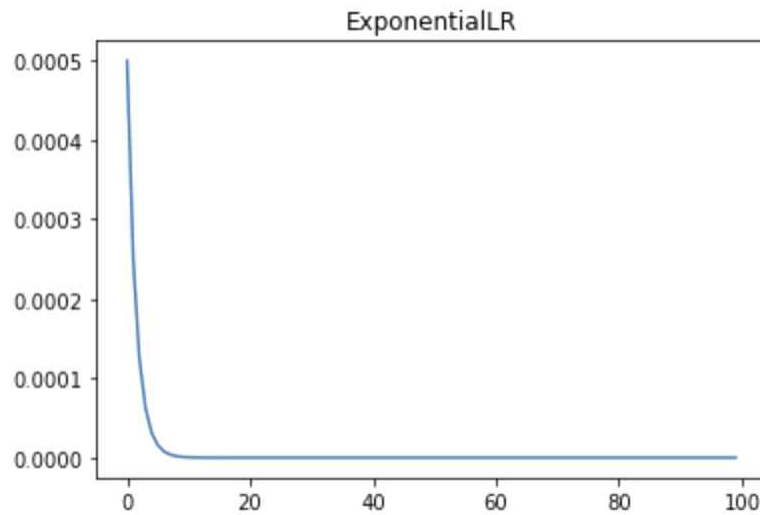
#### 1. Cosine Annealing LR

Cosine Annealing Learning Rate Scheduler는 학습률을 코사인 함수의 형태로 감소시키는 스케줄러이다. 이 스케줄러는 주로 학습률을 부드럽게 감소시켜주어 효과적으로 수렴하도록 돕는다. 즉, 훈련 초기에 빠르게 수렴하며 학습률이 안정화될 때까지 주기적으로 감소한다. 훈련을 통해 적절한 주기를 선택하여 모델의 성능을 향상시키며 활용 가능하다.



#### 2. Exponential LR

Exponential Learning Rate Scheduler는 각 훈련 단계에서 현재 학습률을 이전 학습률의 지수(exponential)로 취하는 스케줄러이다. 주어진 gamma 비율에 따라 현재 학습률을 계산하여 조절한다. 이 스케줄러는 초기 학습률을 효과적으로 관리하여 안정적으로 효율적인 모델 훈련을 도와준다. 적절한 gamma와 step\_size를 선택하여서 훈련을 통해 최적의 파라미터를 찾는 것이 중요하다.

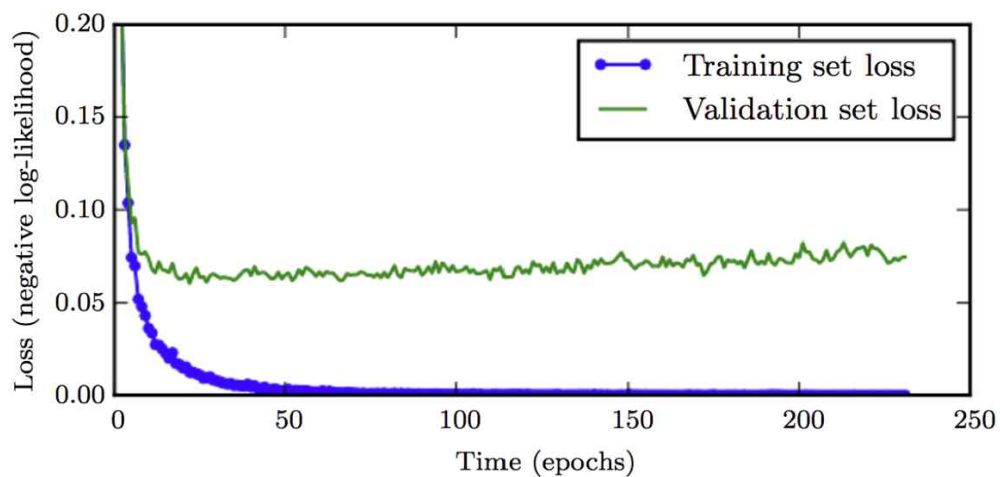


참고 : <https://sanghyu.tistory.com/113>

## Training Error와 Generalization Error 사이 간극을 줄이는 방안

### 1. Early Stopping

Epoch 수(학습 반복 횟수)가 늘어날수록 train set에 대한 오차는 줄어들지만, validation set의 오차가 증가하며 overfitting이 발생할 수 있다. 따라서, 이전 Epoch과 비교해서 오차가 증가하면 overfitting이 발생하기 적에 학습을 멈추는 것을 Early stopping이라고 한다. 다음 그림은 Early stopping이 필요한 경우이다.



## 2. Noise Robustness

Noise나 outlier와 같은 엉뚱한 데이터가 들어오더라도 흔들리지 않는 모델을 만들기 위한 방법으로 일부러 noise를 부여하는 방법이다. Layer 중간에 noise를 추가하는 게 parameter를 줄이는 것보다 강력할 수 있다.

예 :

- 1) weight에도 noise를 넣어준다. (히든 레이어를 dropout하는 것보다 덜 엄격한 느낌이 듦)
- 2) Classification 할 경우 라벨을 부드럽게(label-smoothing)한다. (1,0,0 -> 0.8, 0.1, 0)