

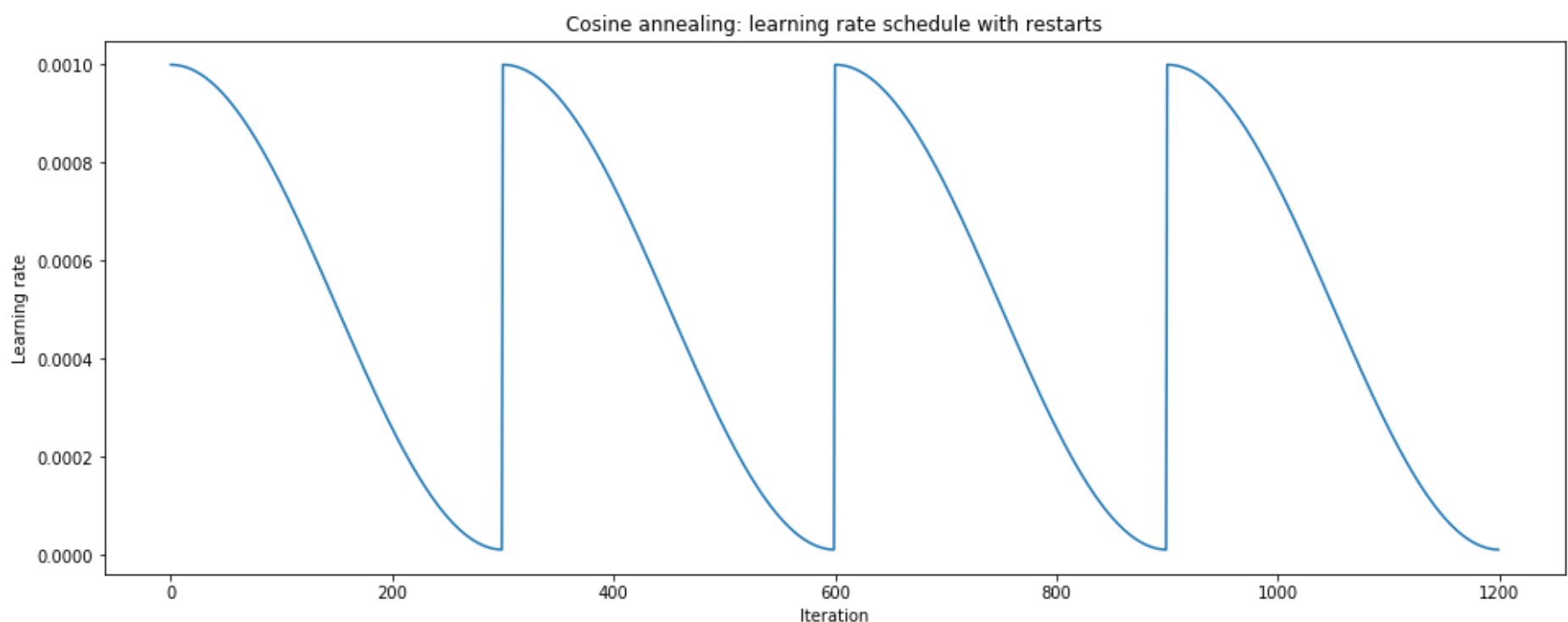
NNAdvanced

Learning Rate scheduler

Cosine annealing은 "SGDR: Stochastic Gradient Descent with Warm Restarts"에서 제안되었던 학습율 스케줄러로서, 학습율의 최대값과 최소값을 정해서 그 범위의 학습율을 코사인 함수를 이용하여 스케줄링하는 방법이다. Cosine annealing의 이점은 최대값과 최소값 사이에서 코사인 함수를 이용하여 급격히 증가시켰다가 급격히 감소시키 때문에 모델의 매니폴드 공간의 안장(saddle point)를 빠르게 벗어날 수 있으며, 학습 중간에 생기는 정체 구간들 또한 빠르게 벗어날 수 있도록 한다. 결과적으로 이러한 방법이 모델의 일반화 성능을 극대화시켜준다. 논문의 실험 결과에 따르면 Cosine annealing을 이용하여 학습한 wide residual network이 CIFAR-10, CIFAR-100 에서 Cosine annealing을 사용하지 않고 학습한 wide residual network에 비해 테스트 데이터에서 0.5%, 1.0%의 높은 성능을 보여주고 있다. Cosine annealing의 수식은 다음과 같다.

$$\eta_t = \eta_{min}^i + \frac{1}{2} \left(\eta_{max}^i - \eta_{min}^i \right) \left(1 + \cos \left(\frac{T_{current}}{T_i} \pi \right) \right)$$

- η_t 는 Cosine annealing으로 계산된 현재의 학습율
- $\eta_{max}^i, \eta_{min}^i$ 는 학습 전에 설정하는 고정된 값으로 각 학습율의 최소값 최대값
- $T_{current}$ 은 현재 epoch, T_i 는 Cosine annealing을 실행하는 주기



reference : <https://ai4nlp.tistory.com/16>, <https://paperswithcode.com/method/cosine-annealing>,
<https://www.jeremyjordan.me/nn-learning-rate/>

Training & Generalization 사이 간극을 줄이는 방안

Test data에서도 좋은 성능을 보이기 위한 전략들

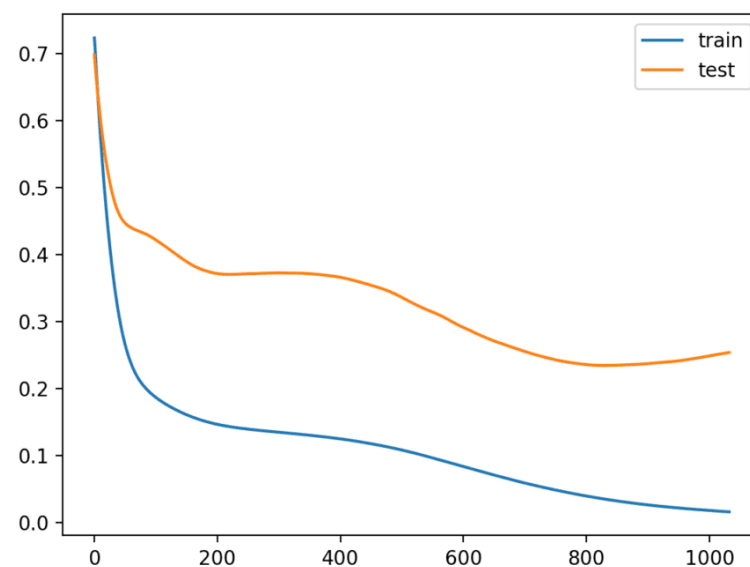
1. Cross fold validation

K-Fold : 가장 널리 사용되는 교차 검증 방법의 하나로, 데이터를 k개로 분할한 뒤, k-1개를 학습용 데이터 세트로, 1개를 평가용 데이터 세트로 사용하는데, 이 방법을 k번 반복하여 k개의 성능 지표를 얻어내는 방법이다. 하이퍼파라미터나 모델의 최적화 시, 최적의 조건을 찾는 데에 활용할 수 있다. 또 모든 데이터를 평가와 학습에 사용할 수 있어 알고리즘의 일반화 정도를 평가할 수 있다. 하지만 데이터 세트를 k번 나누어 학습하는 만큼, 일반적인 분할 방법에 비해 학습과 평가에 소요되는 연산 비용이 늘어난다는 단점이 있다.

2. Early Stopping

Early Stopping : Epoch 을 정하는데 많이 사용되는Early stopping은 무조건 Epoch 을 많이 돌린 후, 특정 시점에서 멈추는 것이다. 그 특정시점을 어떻게 정하느냐가 Early stopping 의 핵심이다. 일반적으로 hold-out validation set 에서의 성능이 더이상 증가하지 않을 때 학습을 중지시키게 된다.

오른쪽의 그래프는 총 4,000 epoch로 설정해 두었을때, 더이상 모델의 성능의 향상이 보이지 않을 때 epoch가 멈추는 것을 확인할 수 있다. 해당 모델은 epoch 1033까지 을 돌았다. 그리고 Test에서는 0.94의 성능을 보여주었다. 해당 점수는 epoch 880에서부터 변화가 없어서 1033에 끝난 것이었다.



reference : <https://seongyun-dev.tistory.com/52>, <https://machinelearningmastery.com/how-to-stop-training-deep-neural-networks-at-the-right-time-using-early-stopping/>