

딥알못 탈출하기 #5

고려대학교 지능시스템 연구실

유민형 제작

Optimization Strategy

- 3. Stochastic Gradient Descent에서는 딥러닝에서 GD대신 SGD를 사용하는 이유에 대해 알아보았다.
- SGD를 바탕으로, 그 외에 많이 사용되는 최적화 기법을 알아보자.
- (문제1) 다른 기법을 알아보기에 앞서, SGD를 놔두고 왜 다른 최적화 기법을 사용하는 것일까? 다음 두가지 측면에서 설명하십시오.
 - efficiency:
 - optimality: (힌트 4. Function Approximator의 5번 문제)

Optimization Strategy

- (문제2) 다음 최적화 기법의 알고리즘을 적고, SGD와의 차이점과 각각의 관계를 설명하시오.
 - SGD
 - Momentum
 - Adagrad
 - RMSProp
 - Adam

Optimization Strategy

- (문제3) 위에 나온 최적화 알고리즘 외에도, 다른 것들이 많고, 최근에도 계속 연구되고 있다. 그렇다면 수많은 알고리즘 중에서 어떤 상황에 어떤 것을 사용해야 하는가? 또한 널리 통용되는 Baseline이 되는 것은 무엇인지 설명하십시오.