

BE530: Medical Deep Learning

2021학년도 1학기, 중간고사

학과: AI컴퓨터공학부

구분: 학사과정

학번: 201511837

이름: 이상민

답안:

- (1) 검증데이터셋의 손실함수 값이 훈련 데이터셋의 손실함수 값보다 크다는 것은 모델의 성능에 학습 데이터셋이 큰 영향을 주지 못하고, 검증데이터셋을 통한 Hyperparameter 조정이 이 모델의 성능에 영향을 많이 주고 있다는 의미이기 때문에 Hyperparameter가 많이 사용된다는 것을 의미하므로 학습데이터 양에 비해 모델의 용량이 크다는 것을 알 수 있습니다.

(2) 문제를 해결하기 위해서는 데이터 증강 등 여러 방법을 사용하여 학습데이터의 양을 늘리거나, 학습데이터셋에 맞게 모델의 용량을 줄임으로서 문제를 해결할 수 있습니다.
- (2) 모든 활성화 함수에 대해 Batch Normalization을 적용한다.

Data를 비선형으로 바꾸기 위해 tanh or ReLU와 같은 Activation function을 사용함으로써 가중치가 전달되지 않는 Vanishing gradient 문제가 발생하였습니다.

이를 해결하기 위해서는 신경망 안에 포함되어 학습 시 평균과 분산 계산한 후 정규화를 통해 scale, shift 연산을 실행하는 Batch Normalization을 통해서 Covariate Shift 현상을 막아 학습 시 현재 layer의 입력은 모든 이전 layer의 파라미터의 변화에 영향을 받게 되며, 망이 깊어짐에 따라 이전 layer에서의 작은 파라미터 변화가 증폭되어 뒷 단계에 큰 영향을 끼치게 될 수도 있어 Vanishing gradient 문제를 해결할 수 있습니다.

추가적으로, Batch Normalization을 사용할 경우, propagation시 parameter의 scale 영향을 받지 않기 때문에 learning rate를 높게 설정할 수 있어 빠른 학습이 가능해집니다.
- 2번: $f(x) = 0.9x + 1$

활성화 함수로 2와 같은 선형 활성화 함수를 사용한다면, 아무리 깊게 Layer를 쌓아도 $y=x$ 와 같은 직선함수와 다를 바가 없기 때문에 Deep한 Network의 이점이 전혀 존재하지 않기 때문입니다.

4.

Layer	Activation Volume Dimensions	Number of parameters
Input	32x32x3	0
CONV3-8	32x32x8	216
Leaky ReLU	32x32x8	216
POOL-2	16x16x8	0
BATCHNORM	16x16x8	0
CONV3-16	16x16x16	432
Leaky ReLU	16x16x16	432
POOL-2	8x8x16	0
FLATTEN	1024	0
FC-10	1x1x10	10240