0. 1

7기 권용호

* 활성화 함수 (Activation function)

간단하게 말하면, 입력을 받아 출력값이 활성화가 될지, 비활성이 될지를 결정하는데 사용되는 함수이다. 활성화 함수는 어떤 것을 사용하느냐에 따라 그 출력 값이 달라지기 때문에, 적절한 활성화 함수를 사용하는 것이 중요하다. **신경망에서는 활성화 함수로 선형함수가 아닌 비선형함수를 사용해야 한다.** Hidden Layer를 여러 개 다층으로 구성하고 활성화함수로 모두 선형함수를 이용하는 경우 단층으로 구성한 것과 동일하게 구성된다.

그러므로 무의미하다. 따라서 깊게 망을 구성하려면 1개 이상의 비선형함수를 이용하여야 한다. 또한 활성화 함수는 크게 4가지의 부류로 나눌 수 있다.

1. identity

2. binary step

3. soft step (= Logistic)

4. ReLU

손실 함수 (Loss function)

인공지능 **모델이 예측한 답과 원래 정답의 오차를 표현, 판단하는 함수**를 말한다. 인공지능의 성능을 증명하는 가장 중요한 지표 중 하나이다.

모델이 예측한 결과와 실제 테스트 케이스의 답은 다를 수 있는데, 이런 오차를 수식화하여 모델의 성능을 판단하는 것이 바로 손실함수이다. **인공지능은 이 손실함수의 값을 최소화하는 방향으로 학습을 진행한다.** 그렇기 때문에 어떠한 손실함수를 모델 학습의 기준으로 삼느냐가 매우 중요하다. 대표적인 손실함수로는 MSE, RMSE, MAE, Binary Crossentropy, Categorical Crossentropy 등이 있다.

경사 하강법 (Gradient descent)

gradient descent 방법은 steepest descent 방법이라고도 불리는데, 함수 값이 낮아지는 방향으로 독립 변수 값을 변형시켜가면서 최종적으로는 최소 함수 값을 갖도록 하는 독립 변수 값을 찾는 방법이다. gradient descent는 함수의 최소값을 찾는 문제에서 활용된다. 미분계수가 0인 지점을 찾는 방식이 아닌 gradient descent를 이용해 함수의 최소값을 찾는 주된 이유는

* 우리가 주로 실제 분석에서 맞딱드리게 되는 함수들은 닫힌 형태(closed form)가 아니거나 함수의 형태가 복잡해 (가령, 비선형함수) 미분계수와 그 근을 계산하기 어려운 경우가 많고,
* 실제 미분계수를 계산하는 과정을 컴퓨터로 구현하는 것에 비해 gradient descent는 컴퓨터로 비교적 쉽게 구현할 수 있기 때문이다.

추가적으로,

* 데이터 양이 매우 큰 경우 gradient descent와 같은 iterative한 방법을 통해 해를 구하면 계산량 측면에서 더 효율적으로 해를 구할 수 있다.

학습률 (Learning rate)

**학습률**인란, 현재점에서 다음점으로 얼만큼 이동할지, 다르게 말하면 모델이 얼마나 세세하게 학습을 할지를 말한다. **학습률**이 작다면 손실이 최적인 가중치를 찾는데 오랜 시간이 걸릴 것이며, **학습률**이 너무 크다면 최적점을 무질서하게 이탈할 수도 있다.

에포크 (Epoch)

Epoch(**에포크**) Epoch의 네이버 영어 사전 뜻은, "(중요한 사건·변화들이 일어난) 시대"라는 뜻이다. 훈련 데이터셋에 포함된 모든 데이터들이 한 번씩 모델을 통과한 횟수로, 모든 학습 데이터셋을 학습하는 횟수를 의미한다.