

가) Activation function

A. 정의

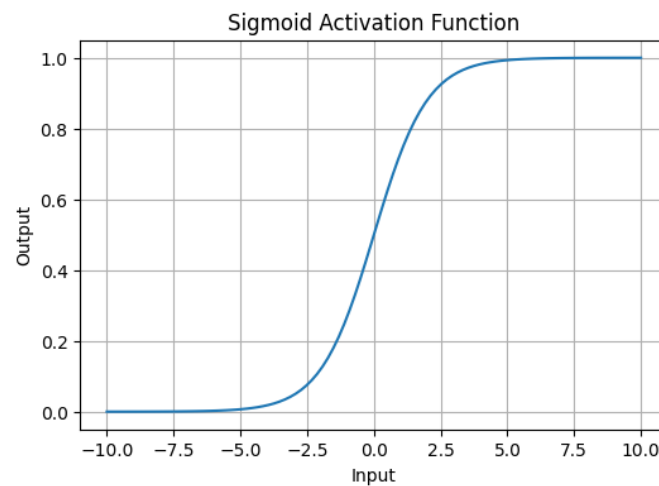
- 활성화 함수는 뉴런에 출력에 적용되는 함수
- 모델에 비선형성을 도입하여 데이터에서 복잡한 패턴을 학습하고 표현할 수 있도록 한다.
- 비선형성 특징이 없다면 신경망은 레이어 수에 관계없이 선형 회귀 모델처럼 동작한다.

B. 목적

- 매우 복잡한 데이터 분포를 모델링하고 딥러닝 작업을 할 수 있도록 한다.

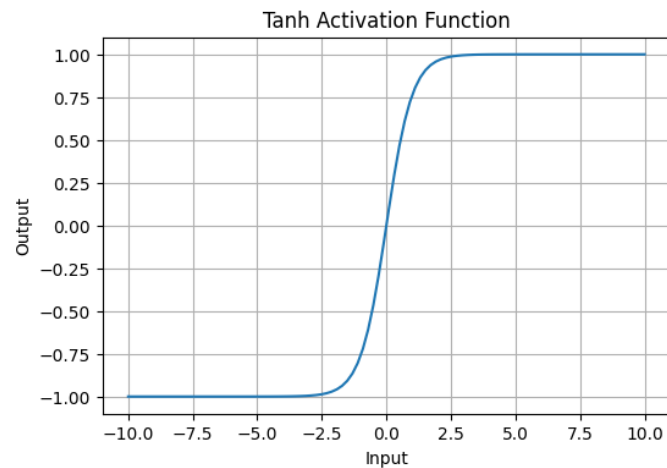
C. 종류

1) Sigmoid

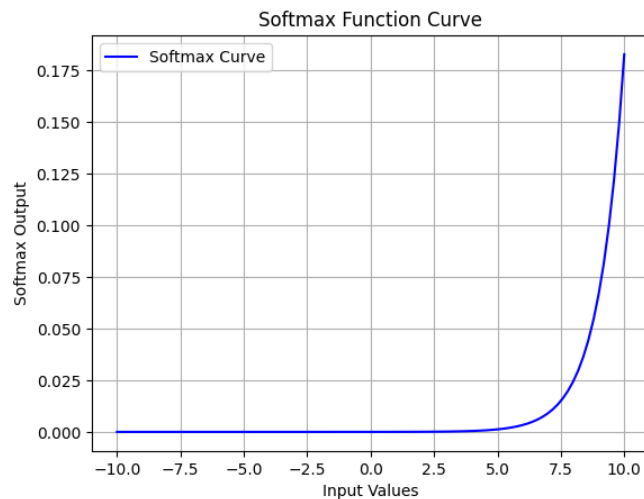


- S자 모양
- $Y = \frac{1}{1+e^{-x}}$
- 출력 범위 : 0~1 사이 >> 2진 분류에 용이
- Layer 많을수록 기울기 소실 >> 가중치의 변화가 없어짐 >> error값

1) Tanh



- Sigmoid 평행이동
- $\tanh(x) = 2 \times \text{sigmoid}(2x) - 1$
- 출력범위: -1~1 사이
- 기울기 손실
 - 1) Relu
- 가장 많이 사용하는 활성화 함수
- 출력 범위: $0 \sim \infty$
- 기본층에 사용
 - 1) Softmax



- 어느 클래스에 속하는지 식별하는 데 사용
- 출력 범위: $0 \sim \infty$
- $Y(x) = \log(1 + e^x)$
- Relu함수보다 기울기가 부드러워 다중 클래스 문제를 처리할 때 사용한다

나) 최적화 함수

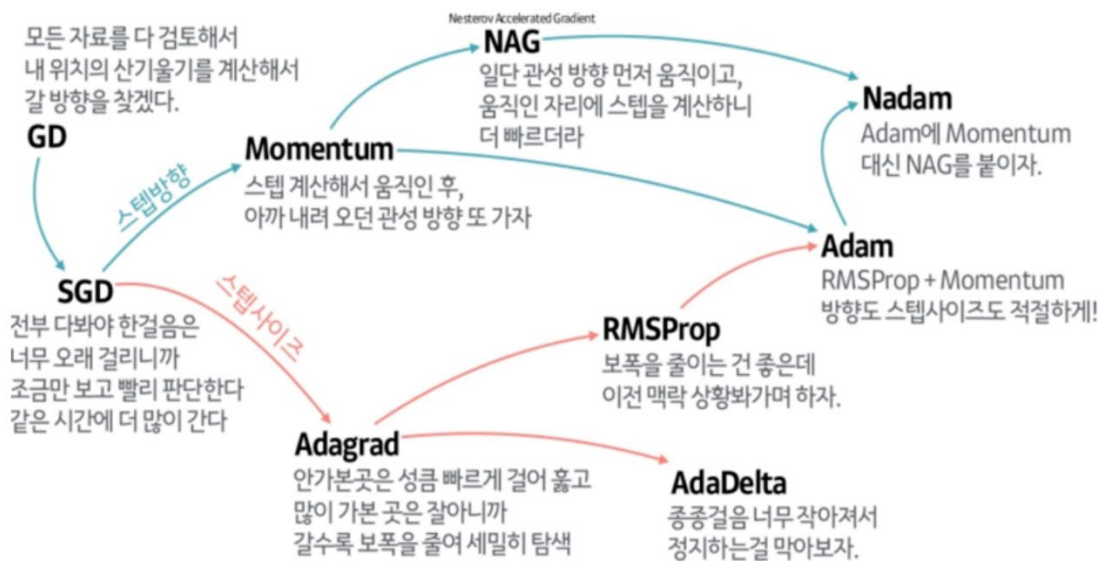
A. 정의

- 손실함수의 결과값을 최소화하는 함수 >> 정확도, 속도
- 에포크의 가중치를 수정하고 손실 함수를 최소화
- 가중치 및 학습률과 같은 신경망의 속성을 조절하는 알고리즘

+ 딥러닝 용어

- 에포크 : 전체 학습 데이터 세트 횟수
- 샘플 : 데이터 세트의 단일 행
- 배치 : 모델 매개변수를 업데이트하기 위해 취해야 할 샘플 수
- 학습률 : 모델 가중치를 얼마나 업데이트 해야 하는지의 척도의 매개변수
- 비용 함수/손실 함수 : 비용 함수는 예측 값과 실제 값의 차이인 비용을 계산
- 가중치/편향 : 두 뉴런 간 신호를 제어하는 모델의 학습 가능한 매개변수

B. 종류



1) Adam

- 일반적으로 사용 (계산이 빠름, 매개변수가 적음)

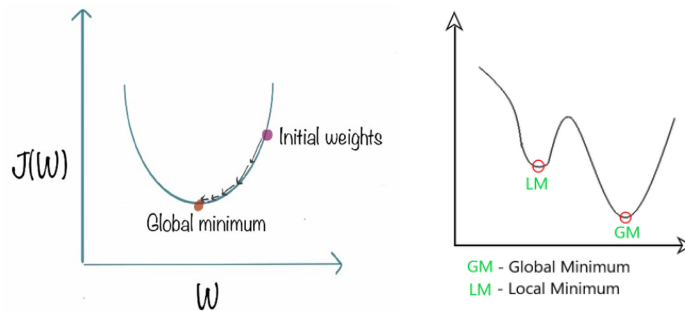
1) GD

- 미적분을 사용하여 값을 지속적으로 수정하고 최소값을 달성하는 모델
- $x_{new} = x - \alpha * f'(x)$ (α : 반복에서 각 기울기에 대해 얼마나 이동할 지 나타내는 크기)
- 계수 초기화 >> 비용 평가 및 수정 >> 계수 업데이트 >> 과정 반복 >> 최소값 도달

3) SGD

- 기본 모델에서 무작위 데이터를 배치하여 한 번에 데이터 세트에서 몇 개의 샘플만 고려하여 학습
- $w := w - \eta \nabla Q_i(w)$

4) RMSProp



3) K- Fold

A. 정의

- 전체 데이터를 K개 부분(or Fold)로 나누는 방법
- 폴드는 각 한 번씩 테스트 데이터로 사용, 나머지는 k-1개의 폴드는 훈련 데이터로 사용된다. (3, 5, 10)

B. 원리

1) 데이터를 k개의 폴드로 나눈다.

2) 각 폴드에 대해 수행한다.

A. 선택한 폴드를 테스트 데이터로, 나머지를 훈련 데이터로 사용

B. 모델을 훈련하고 테스트 데이터로 성능을 평가한다.

3) k개의 폴드에 대한 성능을 평균내어 최종 성능을 얻는다.

