

활성화 함수/ Loss Function에 관하여

출처:

- 핸드온 머신러닝
- <https://wikidocs.net/152768> , <https://wikidocs.net/152766> 한땀한땀 컴퓨터 비전 백과사전

표 10-1 회귀 MLP의 전형적인 구조

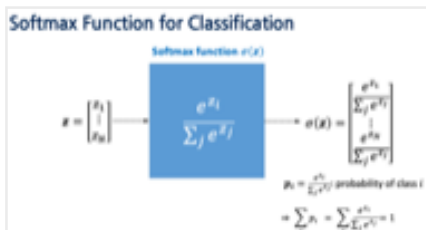
하이퍼파라미터	일반적인 값
입력 뉴런 수	특성마다 하나의 뉴런을 들어 MNIST의 경우 $28 \times 28 = 784$
은닉층 수	문제에 따라 다름, 일반적으로 1에서 5 사이
은닉층의 뉴런 수	문제에 따라 다름, 일반적으로 10에서 100 사이
출력 뉴런 수	예측 차원보다 하나
은닉층의 활성화 함수	ReLU 또는 SELU, 11장 참조
출력층의 활성화 함수	없음, 또는 (출력이 양수일 때) ReLU/softplus나 (출력을 특정 범위로 제한할 때) logistic/tanh를 사용
손실 함수	MSE나 (이상치가 있다면) MAE/Huber

표 10-2 분류 MLP의 전형적인 구조

하이퍼파라미터	이진 분류	다중 레이블 분류	다중 분류
입력층과 은닉층	회귀와 동일	회귀와 동일	회귀와 동일
출력 뉴런 수	1개	레이블마다 1개	클래스마다 1개
출력층의 활성화 함수	로지스틱 함수	로지스틱 함수	소프트맥스 함수
손실 함수	크로스 엔트로피	크로스 엔트로피	크로스 엔트로피

참고) 로지스틱 함수 == 시그모이드 함수, 크로스 엔트로피 == 소프트맥스 함수의 총

소프트맥스 함수란?



다중 레이블 분류와 다중 분류의 차이

- 다중 분류란 단일 레이블 분류를 의미하며 입력값 하나당 하나의 클래스에만 대응할 수 있다.
- 다중 레이블 분류란 하나의 입력값이 여러개의 클래스에 대응할 수 있다.

적어도 출력층의 활성화 함수는 라벨(클래스, Y)값을 포함하는 범위여야 한다.(그래야 정상적으로 학습된다)

활성화 함수는 왜 필요할까?

- 신경망에 비선형성을 더해줘 데이터를 설명하는 복잡한 형상의 경계면을 만들어 낼 수 있다.

☆

초기화 전략	활성화 함수	σ^2 (정규분포)
글로브	활성화 함수 없음, 하이퍼볼릭 탄젠트, 로지스틱 소프트맥스	$1 / fan_{in}$
He	ReLU 함수와 그 변종들	$2 / fan_{in}$
르문	SELU	$1 / fan_{in}$