

[Data Science] 데이터 사이언스 개념 - 9.신경망이 기초 — 나무늘보의 개발 블로그

노트북: 첫 번째 노트북

만든 날짜: 2021-01-22 오후 10:23

URL: <https://continuous-development.tistory.com/219?category=833358>

Data Science

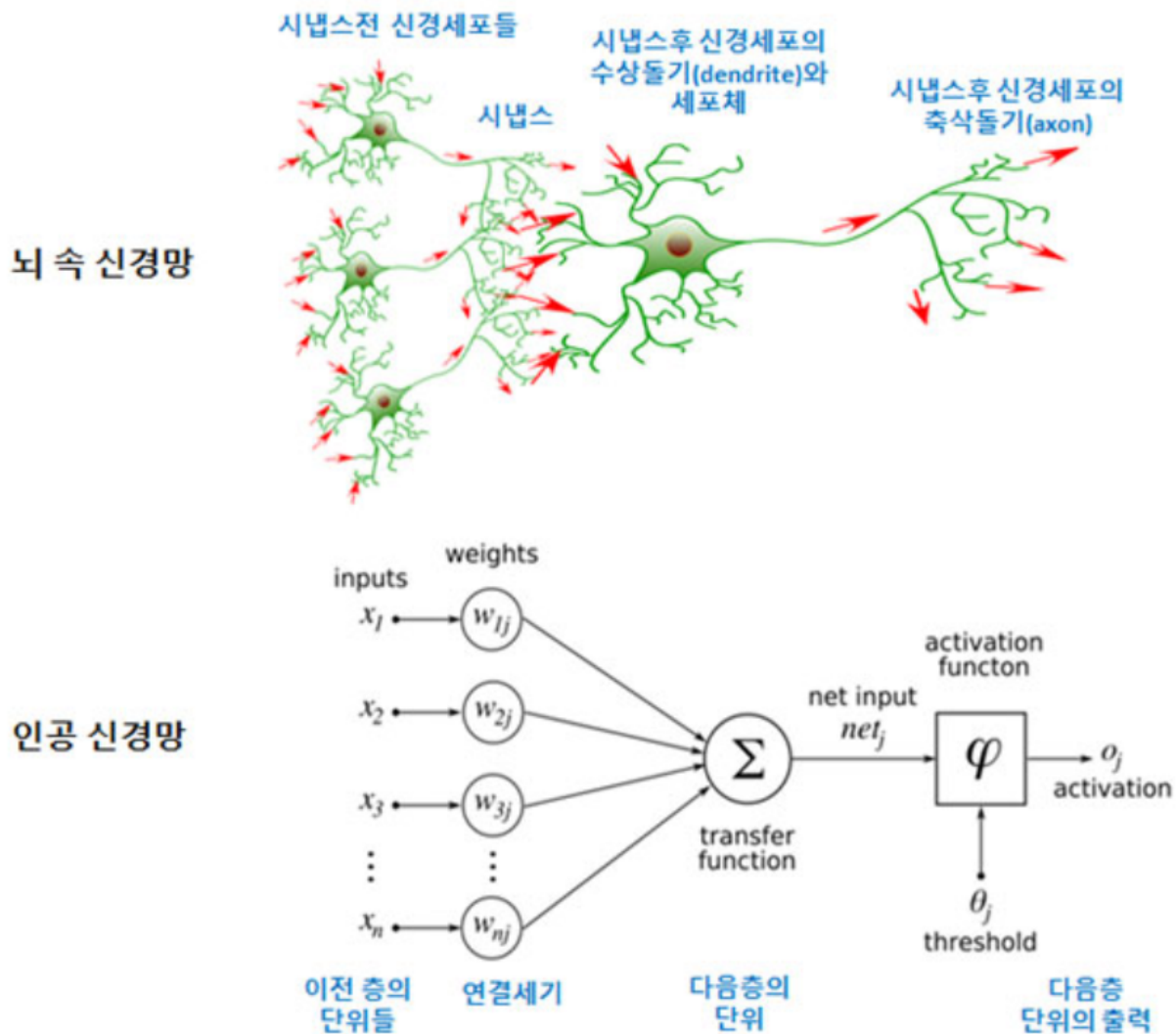
[Data Science] 데이터 사이언스 개념 - 9.신경망이 기초

2021. 1. 15. 10:51 수정 삭제 공개



신경망의 기초

1.신경망이란?



신경망 - 동물의 신경 시스템을 모방한 학습 모델의 총칭

시냅스를 매개로 다른 세포로부터 자극을 받아 들인 신경 세포는 자극이 어느 일정 수위를 넘으면 흥분 상태가 되고, 그 세포가 시냅스 결합을 가진 다른 신경 세포에 자극을 전달한다.

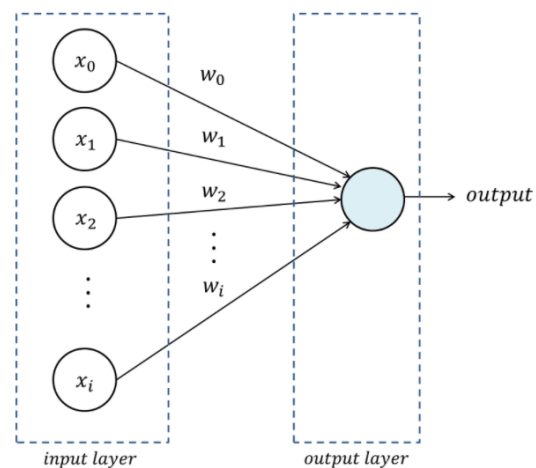
신경망은 이 신경 세포의 동작을 단순화해서 모방한 뉴런이라는 계산소자를 다수 결합해서 구성한 학습 모델이다.

가장 일반적인 것은 1940년대에 제안된 맥컬록-피츠 뉴런이다.

이 뉴런은 임계값을 넘었을 때 흥분 상태를 나타내는 1을 출력하고, 그 이외에는 0을 출력하는 함수다.

컴퓨터의 연산회로가 다수의 논리 회로를 조합해 다양한 연산을 할 수 있는 것 처럼 단순한 계산 소자인 뉴런을 복잡하게 구성함으로써 다양한 계산이 가능해진다.

2.단순 퍼셉트론



1950년대 로젠블라트가 제안한 퍼셉트론 알고리즘

단순 퍼셉트론 - 입력을 제외하고 2개 층으로 구성된 모델

입력에서부터 1층으로 결합은 랜덤값으로 고정된 것이고, 실제 학습 대상이 되는 것은 1층에서 2층으로의 결합 파라미터이다.

로젠블라트의 퍼셉트론 알고리즘은 2진 분류 문제에 대해 데이터를 하나씩 추출하고, 그 식별 오차를 이용해 파라미터를 갱신한다.

퍼셉트론 알고리즘의 중요한 성질은 풀어야 할 문제가 선형분리가 가능하면 최적해를 발견할 수 있지만, 반대로 선형분리가 불가능한 문제는 어떻게 할 수도 없다라는 점이다.

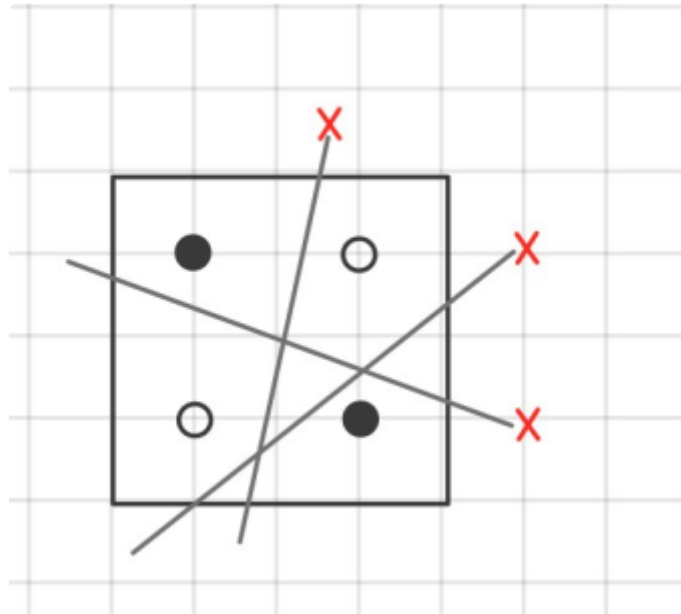
3.단순 퍼셉트론의 한계

XOR 함수를 통해 단순퍼셉트론의 한계를 볼 수 있다.

XOR 함수란 2개의 0과 1을 취하는 변수를 입력으로 하고 0이나 1을 출력으로 반환하는 함수이다.

Input A	Input B	XOR Output
0	0	0
0	1	1
1	0	1
1	1	0

선형분리 가능한 문제란 직선으로 분리를 할 수 있는지이다.



단순 퍼셉트론은 위 그림과 같이 간단한 문제도 풀수 없다.(직선하나로 검은색 원과 흰색 원을 나눌수 없다.)

이 단순 퍼셉트론의 한계는 학습 대상이 단지 한 층으로만 제한되었던 것이 문제의 원인이다.

4.다층 퍼셉트론

1980년대에 럼멜하트와 힌튼, 르쿤 등에 의해 생겨난 모델이다.

입력층, 중간층, 출력층으로 된 2층의 신경망이다.(입력층은 보통 세지 않는다.)

다층 퍼셉트론에서는 중간층의 2개 뉴런을 입력이 0,0 이 아닌지, 또는 1, 1이 아닌지 각각 식별하도록 훈련한다.

이것으로 XOR 함수를 전체로서 구현 할 수 있다.

배론 등이 제시한 시그모이드 활성화 함수를 가진 다층 퍼셉트론은 충분히 많은 중간층 뉴런을 준비하면 임의의 함수를 임의의 정밀도로 근사할 수 있다 라고하는 정리이다.

이런 성과들을 통해 XOR 함수와 관련된 신경망에 대한 비판은 완전히 뒤집혔다.

5.다층 퍼셉트론 학습

다층 퍼셉트론 학습에서 중심적인 역할을 담당하는 기법이 오차역전파와 후술할 확률적 강사 하강법이다.

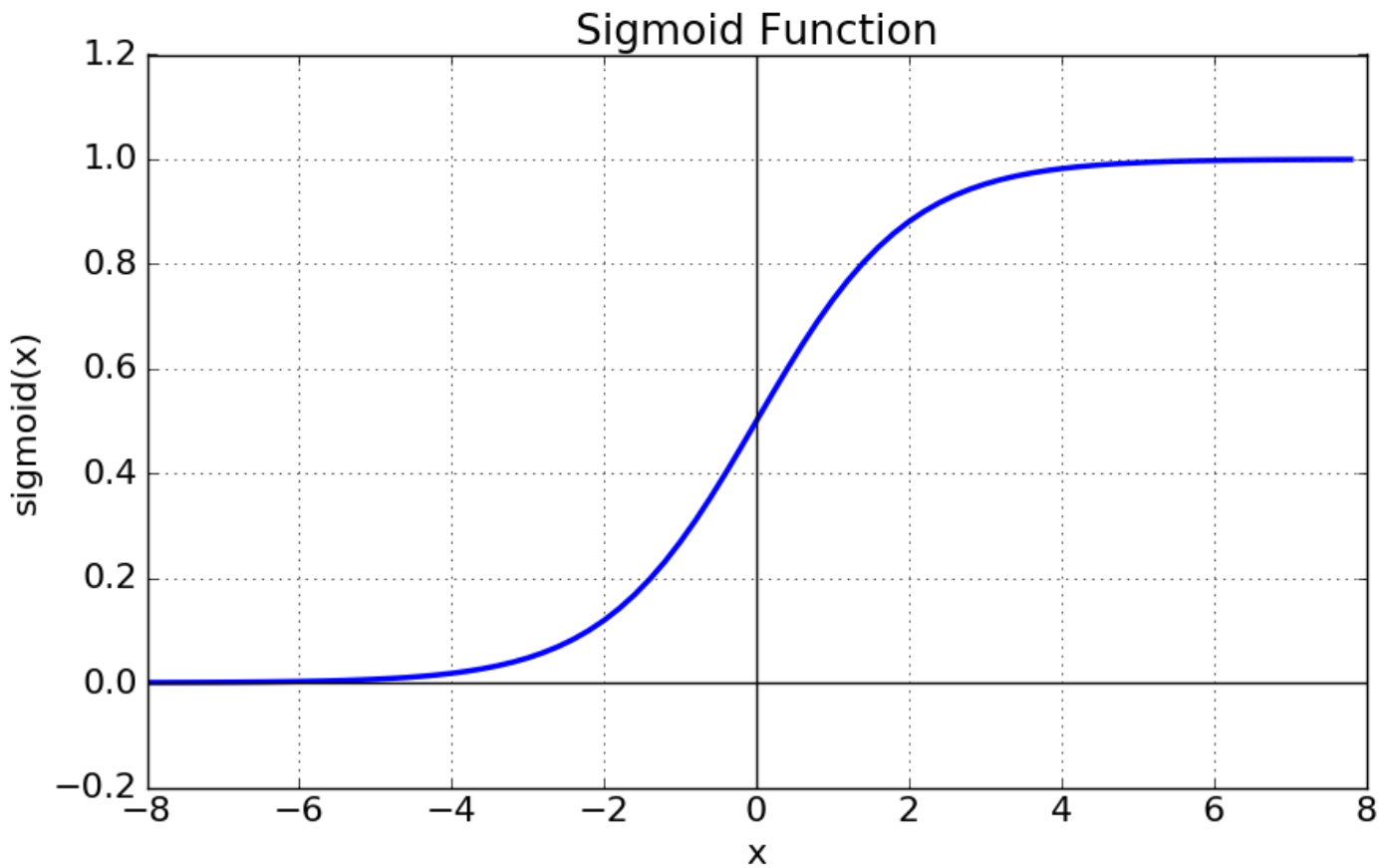
단수 퍼셉트론과 다층 퍼셉트론에서의 중요한 차이중 하나는 시그모이드 활성화 함수이다.

단순 퍼셉트론에서는 계단형 활성화 함수가 이용 됐다.

하지만 이것들은 미분 불가능한 점을 포함하는 함수와는 어울리지 않는다.

시그모이드 활성화 함수는 계단 함수를 매끄럽게 하는 함수이고, 어디에서나 미분을 실행할 수 있는 함수이다.

어느 방향으로 파라미터를 움직이면 목적함수가 개선되어 학습이 촉진되는지 쉽게 알 수 있기 때문에 시그모이드 활성화 함수를 이용한 신경망은 매우 다루기 쉽다.



다층 퍼셉트론에서의 파라미터에 관한 미분을 효율적으로 하는 방법이 오차역전파이다.

기어에 비유하면 파라미터를 각 기어의 회전량이라고하고, 파라미터에 관한 미분은 각 기어의 출력에 대한 회전비에 해당한다.

회전비는 출력에서 입력을 향해 차례로 각 기어의 회전비를 곱하는 것으로, 효율적으로 계산할 수 있다.

이 방법이 오차 역전파이다. 입력에서 출력으로 전달되는 회전과는 역방향으로 계산이 진행되므로 역전파라고 불리어진다.

6.다층 퍼셉트론의 한계

근본적인 문제로서 **다층 퍼셉트론의 목적 함수의 비볼록성**이 있다.

경사법으로 최적해를 구하기가 사실상 불가능했다. 또한 **기울기 소실, 폭 발등** 학습이 불안정해지거나 전혀 학습이 진행되지 않는 등

학습 자체의 어려움등이 있었다.

또한 다층 퍼셉트론의 **과적합**을 실용적으로도 이론적으로도 억제하기 어려웠다.

본 내용은 그림으로 배우는 DataScience 데이터 과학을 참고한 내용입니다

'Data Science' 카테고리의 다른 글

[Data Science] 데이터 사이언스 개념 - 10.딥러닝

[Data Science] 데이터 사이언스 개념 - 9.신경망이 기초

[Data Science] 데이터 사이언스 개념 - 8.토픽 모델 / 네트워크 분석

[Data Science] 데이터 사이언스 개념 - 7.비지도 학습

[Data Science] 데이터 사이언스 개념 - 6.분류문제

[Data Science] 데이터 사이언스 개념 - 5.앙상블 학습

다중 퍼셉트론

다층 퍼셉트론의 한계

단순 퍼셉트론

단순 퍼셉트론의 한계

신경망



나아무늘보

혼자 끄적끄적하는 블로그 입니다.