

241108 딥러닝

▼ 퍼셉트론

정의:

인공 신경망의 구성 요소

다수의 입력을 받아 하나의 출력값을 내보냄

출력값은 1 혹은 0

$$y = \begin{cases} 0 & w_1x_1 + w_2x_2 \dots \leq 0 \\ 1 & w_1x_1 + w_2x_2 \dots > 0 \end{cases}$$

학습 방법:

1. 좌표평면에 임의의 직선을 긋는다.
2. 좌표평면에 데이터를 하나씩 입력한다.
3. 입력값에 따른 모델의 예측값과 정답을 비교하여 틀린 경우 직선을 다시 긋는다.
4. 모든 학습용 데이터에 대해 2~3 과정을 반복한다.

한계:

데이터가 선형적으로 구분 가능해야만 함

예: XOR GATE의 경우 (0,0)과 (1,1)은 0, (1,0)과 (0,1)은 1

(0,0)과 (1,1) 사이에 존재하는 값만 1로 표시해야 함

선 하나로는 구분하지 못함 (선형적이지 않음)

▼ 다층퍼셉트론

입력층과 출력층 사이에 1개 이상의 은닉층이 있는 형태

2개 이상일 때 '심층신경망'이라고 부름

은닉층:

입력층과 출력층 사이에 위치

복잡한 특성, 패턴을 추출. 비선형적 데이터 특성을 학습 가능

보다 정교한 모델 구축 가능

▼ 활성화함수

퍼셉트론에서 입력값의 총합을 출력할지 말지 결정

출력한다면 어떤 값으로 변환하여 출력할지 결정

종류:

1. 계단 함수

입력값 총합이 0 이하라면 -1, 0보다 크다면 1을 출력

분리했는지 여부에만 관심을 가짐 (정확도를 생각하지 않음)

2. 시그모이드 함수

값이 작아질수록 0에 수렴

값이 커질수록 1에 수렴

출력이 0~1 사이로 확률 표현 가능

입력값이 커도 출력값 범위가 작아서 → 경사하강법 수행 시 기울기 소실 문제 발생

왜 비선형 함수?:

선형 함수 자체가 심층 신경망에 큰 도움이 되지 않음

(선형함수 사용시 결과값을 단 하나의 층으로 표현할 수 있게 됨. 신경망을 깊게 쌓는 의미 ↓)

▼ 순전파, 역전파

순전파:

신경망에서 입력데이터가 네트워크의 각 층을 **순차적**으로 통과하여 최종출력을 생성하는 과정

1. 첫번째 계층 → 가중합 사용

2. 가중합 완료됐다면 활성화함수(시그모이드) 적용

3. 활성화함수 결과값이 출력층 시그모이드 함수의 입력값이 됨

4. 출력층 시그모이드 함수 결과값이 해당 인공신경망의 최종 출력값

오차 계산을 위한 손실 함수:

평균 제곱 오차 MSE

$$\text{mean square error} = 1/2(y - y^t)^2$$

역전파:

예측값과 실제값의 차이를 줄이기 위한 과정

역행 이유: 오차를 기반으로 가중치값 업데이트 목적

1. 오차함수를 가중치로 미분 (오차가 가중치에 얼마나 민감한지 측정)

a. 연쇄법칙 사용

b. (수식)

c. (가중치 업데이트)

d. (앞부분 계산)

업데이트된 가중치로 순전파 계산 진행, 오차를 감소하는 것이 최종목적

과제:

계산 연습을 추천 → 업로드되는 계산 문제를 과제로 풀어오기

+ 하던대로 수업내용정리 PR하기

실제 가중치 초기부여: 랜덤부여 후 교정... 등등 방법 많음

데이터에 따라 달라지기도 하고 모델에 대한 분류여부/정확도 중 원하는 방향으로 지정...