

4장

1. 퍼셉트론

- 퍼셉트론 : 다수의 신호를 입력받아 하나의 신호로 출력 (1/0의 정보를 앞으로 전달)
- 게이트 종류
 - AND : 입력 중 어떤 하나라도 0을 가지면 작동을 멈춤
 - OR : 입력에서 둘 중 하나만 1이거나 둘 다 1일 때 작동
 - XOR : 입력 두 개 중 한 개만 1일 때 작동 → 학습 불가능

⇒ 이러한 문제 해결 : 다층 퍼셉트론

- 다층 퍼셉트론
 - 입력층과 출력층 사이에 하나 이상의 hidden layer을 둬
 - 비선형적으로 분리되는 데이터에 대해 학습할 수 있도록 고안됨

⇒ 입력층과 출력층 사이에 hidden layer이 여러 개 있는 신경망 : 심층 신경망 (딥러닝)

2. 딥러닝 구조

- 딥러닝 용어
 - 층 : 입력층 / 은닉층 / 출력층 으로 구분
 - 가중치
 - 입력값의 연산 결과를 조정하는 역할
 - 가중합 (전달 함수)
 - 각 노드에서 들어오는 값에 가중치를 곱한 후 다음 노드로 전달할 때 이들을 더한 값
 - 가중합을 활성화 함수로 전달 (전달 함수)
 - 활성화 함수

- 전달 함수에서 전달받은 값을 출력할 때 일정한 기준에 따라 출력값을 변화시키는 비선형 함수
- 신경망의 층을 더 깊이 있게 쌓기 위해 사용
- 시그모이드 함수
 - : 층이 많아질수록 기울기 소멸 문제 + zero-centered 문제 + exp() 계산 문제
 - : 보통 출력층에서만 사용 (이진분류 활성화함수)
- 하이퍼볼릭 탄젠트 함수
 - : 기울기 소멸 문제
- 렐루 함수
 - : 기울기 소멸 문제 해결 but 음수값일 때 0 반환
- 리키 렐루 함수
- 참고 사이트 : <https://syj9700.tistory.com/37#:~:text=활성화 함수는 이전 층,하는데 중요한 역할을 한다.>
- 소프트맥스 함수
 - : 입력값을 0~1 사이에 출력되도록 정규화 → 출력 값들의 총합이 항상 1이 되도록 함
 - : 보통 딥러닝에서 출력 노드의 활성화 함수로 많이 사용 (히든 레이어에서 사용할 경우 기울기를 제대로 찾지 못해 학습 효율성이 감소함)
 - : 다중 클래스 분류에서 사용되는 활성화 함수 (시그모이드의 일반형)
- 손실 함수
 - 학습을 통해 얻은 데이터의 추정치가 실제 데이터와 얼마나 차이나는지 평가하는 지표
 - 평균 제곱 오차
 - 크로스 엔트로피 오차
- 딥러닝 학습
 1. 순전파
 - 네트워크에 훈련 데이터가 들어올 때 발생
 2. 역전파

- Loss를 최소화하는 가중치로 업데이트
- chain rule을 이용해 가중치 업데이트

3. 딥러닝 문제점

- 과적합 문제
 - 훈련 데이터를 과하게 학습해 발생 → test data에 대해 accuracy 떨어지는 경우
 - 해결 방법 | Dropout
 - 학습 과정에서 임의로 일부 노드 제외해 편향되지 않도록 훈련
- 성능이 나빠지는 문제
 - 배치 경사 하강법 (BGD)
 - 전체 데이터셋에 대한 오류 → 기울기를 한번 계산해 가중치 업데이트
 - 확률적 경사 하강법 (SGD)
 - 임의로 선택한 데이터에 대해 기울기 계산
 - 미니배치 경사 하강법
 - 전체 데이터셋을 여러 개의 미니배치로 나눠 각 배치에서 기울기 계산한 후 가중치 업데이트
 - 가장 많이 사용
- 옵티마이저
 - 최적화하는 방법
 - SGD의 파라미터 변경 폭이 불안정한 문제를 해결하기 위해 학습 속도와 운동량을 조절
 - ADAGRAD → RMSProp → SGD+Momentum → ADAM