

Multi-Layer Perceptron (MLP)

- Perceptron의 한계: 선형 분리가 불가능한 경우 해결한 양의 오류 발생 (XOR 문제)

- MLP Idea

- ① 은닉층: 새로운 특징 공간으로 변환
- ② 시그모이드 활성화 함수
- ③ 오류 역전파

① - XOR 문제 해결: 퍼셉트론을 여러개 사용 (선형 결정 경계)
→ 두 perceptron을 순차로 결합

- 은닉층의 역할: 특징 추출 (특징 학습)

② - 활성화 함수

- 계단
- 코엑스 시그모이드
- 하이퍼볼릭 탄젠트
- 소프트 플러스
- 렉티파이어

② Model Architecture

③ - 오류 역전파

출력층의 결과와 원하는 타겟의 값과 차이를 구하여
그 오차 값을 각 레이어들을 지나며 역전파 해서
각 노드가 가진 변수들을 갱신하는 방법

Training & Testing MLP

- $\theta = \{U^1, U^2\}$ 를 최적화시키는 방법

- Parameter Update by SGD

• Gradient 계산

- 출력층: $\frac{\partial J}{\partial u_{kj}^2} = \delta_k z_j$
- 은닉층: $\frac{\partial J}{\partial u_{ji}^1} = \eta_j x_j$

- '미니배치' 스토캐스틱 경사 하강법

- Gradient의 잡음을 줄여 수렴이 빨라짐
- GPU를 사용한 병렬처리에도 유리
- 현대 기계학습은 미니배치를 널리 사용

Remarks on MLP

- 순수한 최적화 알고리즘으로는 높은 성능을 얻을 수 없음
 - 데이터 부족, 잡음, 미숙한 신경망 등
- 성능 향상을 위한 Heuristics가 필요
- MLP의 한계
 - 장음이 섞인 음성인식 성능 저하
 - 펄스 극소 인식 능력 저하
 - 바둑에서의 한계