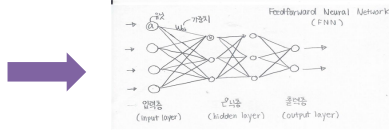


# 인공 신경망의 원리와 성능 개선 방법

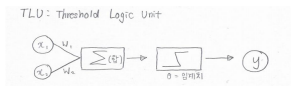
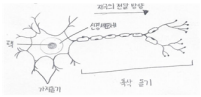
산업 및 에너지 부문

## 1 탐구의 동기 및 목적

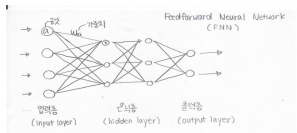


바야흐로 인공지능의 시대이다. 여러 번의 침체기를 지나 지금은 가장 빠르게 발달하고 있는 분야 중 하나인 인공지능은, SNS, 인터넷과 같은 곳에서 우리에게 맞춤 정보를 제공하고 가상 비서역할을 하는 등 우리 생활과 밀접한 관련이 있다. 게다가 알파고와 이세돌의 대국은 수많은 사람들을, 정치가들을, 기업가들을 인공지능이라는 가능성에 희망을 품게 만들었다. 단순히 인공지능의 사회적 파급효과만이 아닌 그 원리와 역사, 그리고 나아가 어떻게 하면 그 성능을 더 향상시킬 수 있을가에 의문을 가졌고, 탐구해보고자 하였다. 따라서 여러 인공 신경망 성능 개선 방안을 비교하고 실제로 구현하기 위해 본 탐구를 시작하게 되었다. 그러하기 위해, 우선 인공 신경망의 여러 메소드들을 탐구한 후 그 원리를 파악하고, Python을 통해 실제로 인공 신경망을 구현한 다음, 탐구한 메소드들을 실제 신경망에 적용하여 그 성능을 비교함으로써 인공지능의 성능을 개선하는 것을 목표로 한다.

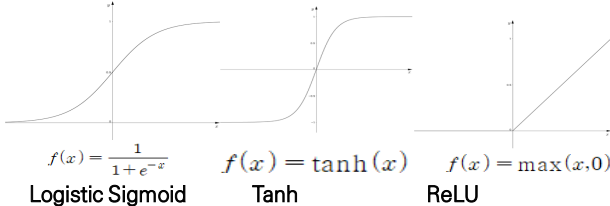
## 2 선행연구 고찰



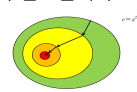
인공 신경망은 사람의 뇌를 모방하여, 전산상으로 구현된 뉴런을 연결하고, 문제에 대해 학습시켜 해결하게 한다. 이는 뇌가 시냅스의 연결을 조절하는 과정을 모방하였다.



신경망은 각각 입력된 신호에 가중치(신호의 세기를 조절, 시냅스의 역할)를 곱하고, 활성화 함수를 적용하여 다음 계층으로 신호를 발산한다: **순전파**



다음 계층으로 신호를 발산하기 전에, 가중치와의 곱에 **활성화 함수**를 적용하게 되는데, 이는 축삭돌기가 일정 세기 이상의 신호만을 발산하는 모습을 본 따왔다. 인공 신경망을 실제 문제 해결에 사용하려면, 가중치를 조절하여 문제 해결에 알맞게 '학습' 시켜야 한다. 학습은 주어진 데이터 세트를 이용해서 수행하는데, 편미분을 통해 주어진 입력에 대한 신경망의 출력이 주어진 기준 출력 값과 가까워지게끔 가중치를 조절한다. 이를 경사 감소법이라고 한다(E(u)를 최소화)



경사 감소법을 이용해 신경망의 오차를 최소화한다. 출력층의 오차함수는 아래와 같이 구한다(최소제곱 오차)

$$E(u) = (t_u - o_u)^2 \dots u \text{층의 오차}$$

출력층의 오차함수를 통해 구한 오차를 이용해 가중치를 보정하고, 이 오차를 이용해 이전 계층의 오차를 계산한다. 이전 계층으로 분배되는 오차는 그 유닛에 연결된 가중치에 비례한다: **역전파**



$$\sum_{i=0}^M w_{ij} \cdot e_j$$

오차함수를 위의 가중치 보정 식에 대입하면 아래와 같다(로지스틱 시그모이드)

$$\Delta w_{jk} = \alpha \times (e_k * \text{sigm}(o_k) * (1 - \text{sigm}(o_k)) \cdot o_j^f)$$

위와 같은 과정을 통해 신경망을 학습시킬 수 있었다. 학습된 신경망은 학습 데이터와 같은 유형의 문제를 해결할 수 있다(분류, 회귀 등)

## 3 인공 신경망의 성능 개선

### 0) MNIST Dataset

신경망의 성능을 개선하고, 개선된 성능을 평가하기 위해, 신경망을 이용한 이미지 인식 분야의 선구자인 Yann Lecun 교수님의 공용 신경망 학습용 데이터세트인 MNIST 데이터세트를 이용하여 손 글씨를 분류하게끔 학습시키고, 그 성능을 평가 및 비교한다.

### 1) MiniBatch

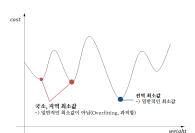
$$D_t = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$$

$$E_t(w) = \frac{1}{N_t} \sum_{n=1}^{N_t} E_n(w)$$

미니배치는 데이터세트의 레코드 하나만을 이용하여 가중치를 보정하는 확률적 경사감소법의 속도를 개선하는 방법이다. 데이터세트를 Nt개의 원소를 갖는 Minibatch 라 불리는 집합으로 나누고, 이 집합마다 오차를 구하여 가중치를 보정하는데, Nt의 값에 비례하여 속도가 빨라진다.

### 2) 가중치 감쇠 : Weight Decay

$$B_{\text{정제}}(w) = B(w) + \frac{\gamma}{2} \|w\|^2$$



학습 데이터 양	60,000
주기(epoch)	5회
학습률	0.1
미니배치 수	10
학습 시간	약 6분
정확도	68.69%

과적합은 오차함수를 최적화하는 과정에서 지역 최소 값에서 학습이 끝나버린 경우에 발생하는데, 학습 데이터에 대한 오차는 작지만, 테스트 데이터에 대한 오차는 크게 발생한다. 이는 가중치가 너무 크게 될 때 발생하는데, 이를 억제하기 위해 가중치 감쇠의 방법을 적용한다. 가중치 감쇠는 가중치가 지나치게 커지는 현상을 방지하는데, 가중치의 제곱 합과 감쇠상수의 곱을 오차함수에 더하여 가중치의 크기에 비례하게 작아지도록 한다. 감쇠상수는 주로  $10^{-4}$ ~ $10^{-8}$ 에서  $10^{-2}$  사이의 값을 사용한다. 벤치마크 결과를 해석하면, 연산량은 늘었지만 전역 최솟값에 더 가까이 도달하여 정확도가 올라갔음을 발견할 수 있다.

## 4 향후 전망과 활용

일련의 과정을 통해 신경망의 성능을 향상시켰으며, 성능 향상에 대한 지표는 아래와 같다.

학습 데이터 양	60,000	이 자료는 일반 Feed-Forward NN의 성능 자료이다.
주기(epoch)	5회	
학습률	0.1	
정확도	97.1%	
학습 데이터 양	60,000	이 자료는 가중치 감쇠를 적용한 신경망의 성능 지표로, 전역 최솟값에 가까워져 정확도가 올라갔음을 관측할 수 있다.
주기(epoch)	5회	
학습률	0.1	
정확도	97.25%	
학습 데이터 양	60,000	이 자료는 Minibatch를 적용한 신경망의 성능 지표로, 속도가 비약적으로 상승했으나 정확도가 떨어졌음을 관측할 수 있다. 큰 데이터셋에서 사용하는 등의 방법을 응용하면 성능 향상을 기대할 수 있을 것이다.
주기(epoch)	5회	
학습률	0.1	
정확도	68.69%	

인공지능 분야는 미래를 이끌어갈 주역이자, 현 시점에서 가장 양성해야 할 분야이다. 중국은 이미 초등학교에 인공지능 교과를 도입하는 것을 검토 중이며, 국가적으로 전폭적인 지원을 시행하고 있다. 이 결과로 중국은 인공지능 분야에서 굉장한 영향력을 행사하게 되었고, 이는 곧 국력의 강화를 가져왔다. 우리나라 또한, 인공지능 분야를 지원하고 이를 뒷받침할 연구를 시행해야 한다. 본 탐구에서 실험한 성능 개선 방법은 앞으로 발전할 인공지능 기술에 응용될 수 있다. 미니배치의 경우 속도 문제에 민감한 웹 관련 AI 기술이나 게임 AI에, 가중치 감쇠는 정확도에 민감한 주가예측 혹은 이미지/비디오 처리 등에 응용될 수 있다. 본 탐구의 실험 결과가 이와 같은 기술 발전의 지표로 사용하고자 한다