

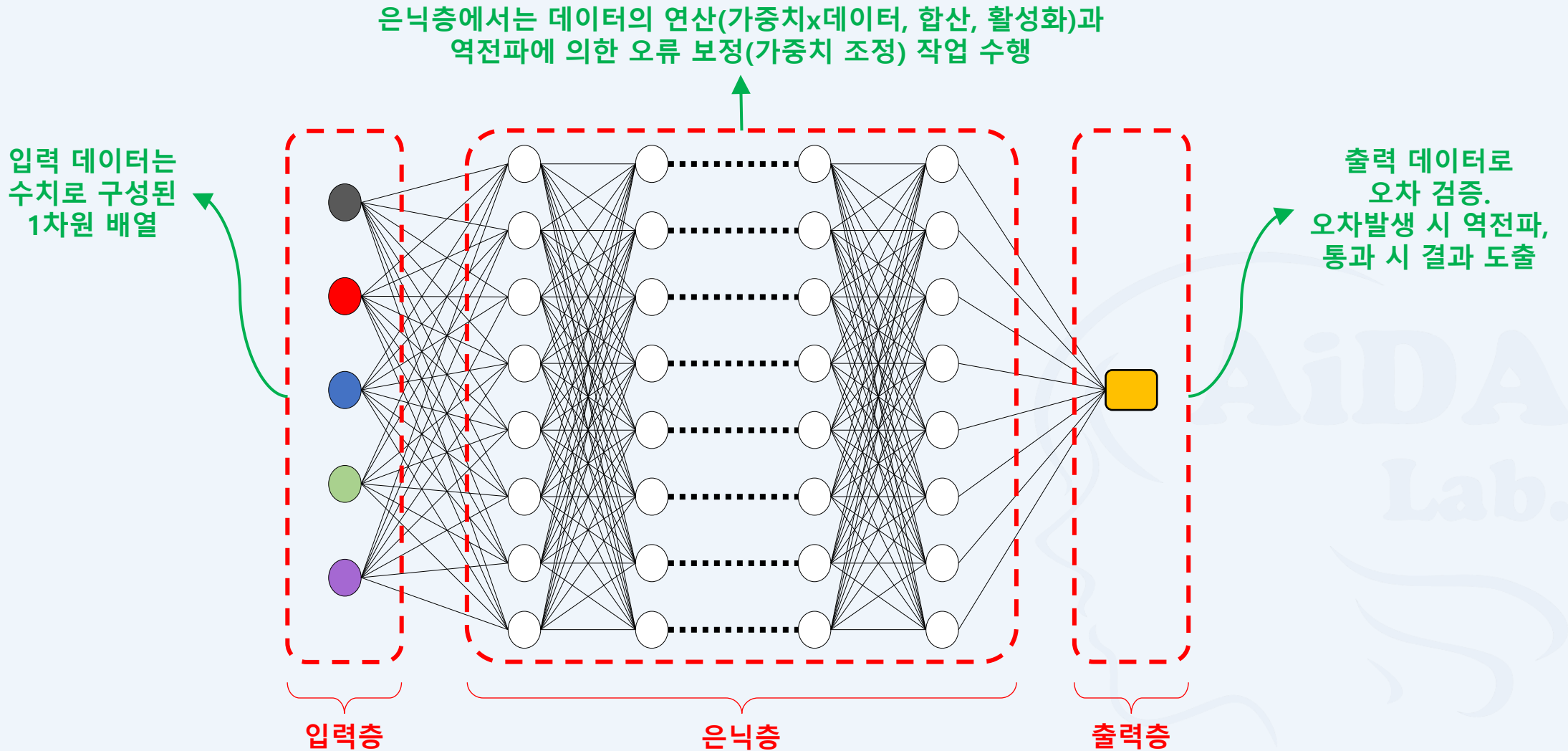
2021 인공지능 소수전공

33차시: 딥러닝 기본 모델

2021.07.29 18:30~19:15

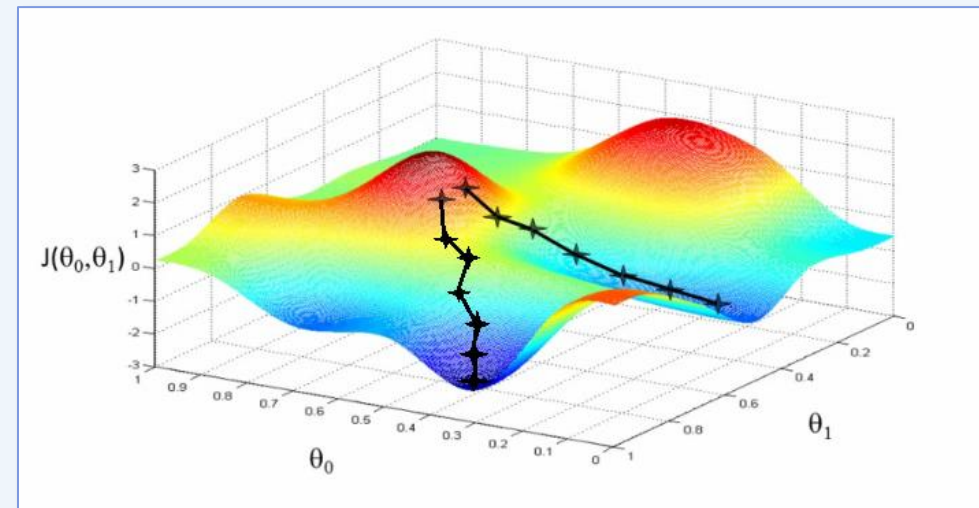
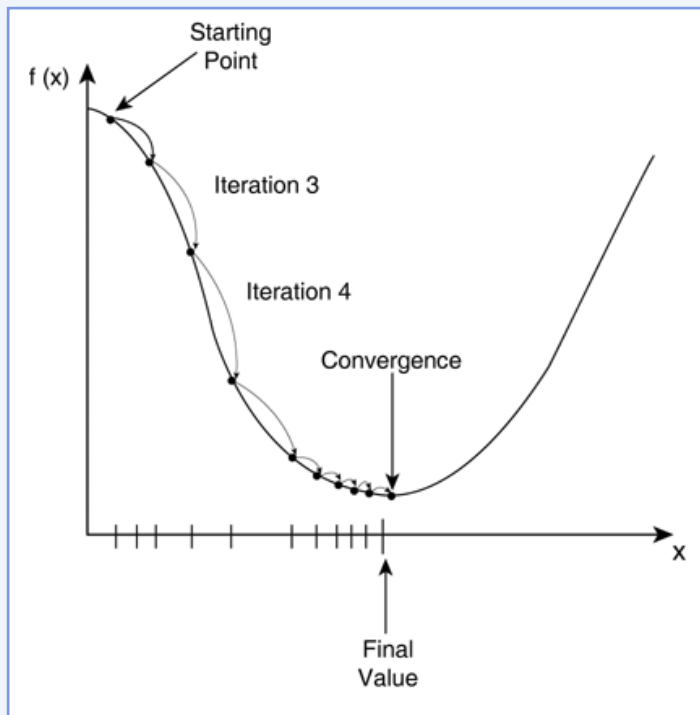
Seokhwan Yang

딥러닝 기본 모델



- 가장 많이, 기본적으로 사용되는 가중치 조정 방법
 - 경사 하강법 (Gradient Descent)

초기값에 따라
최저점이
달라질 수 있다

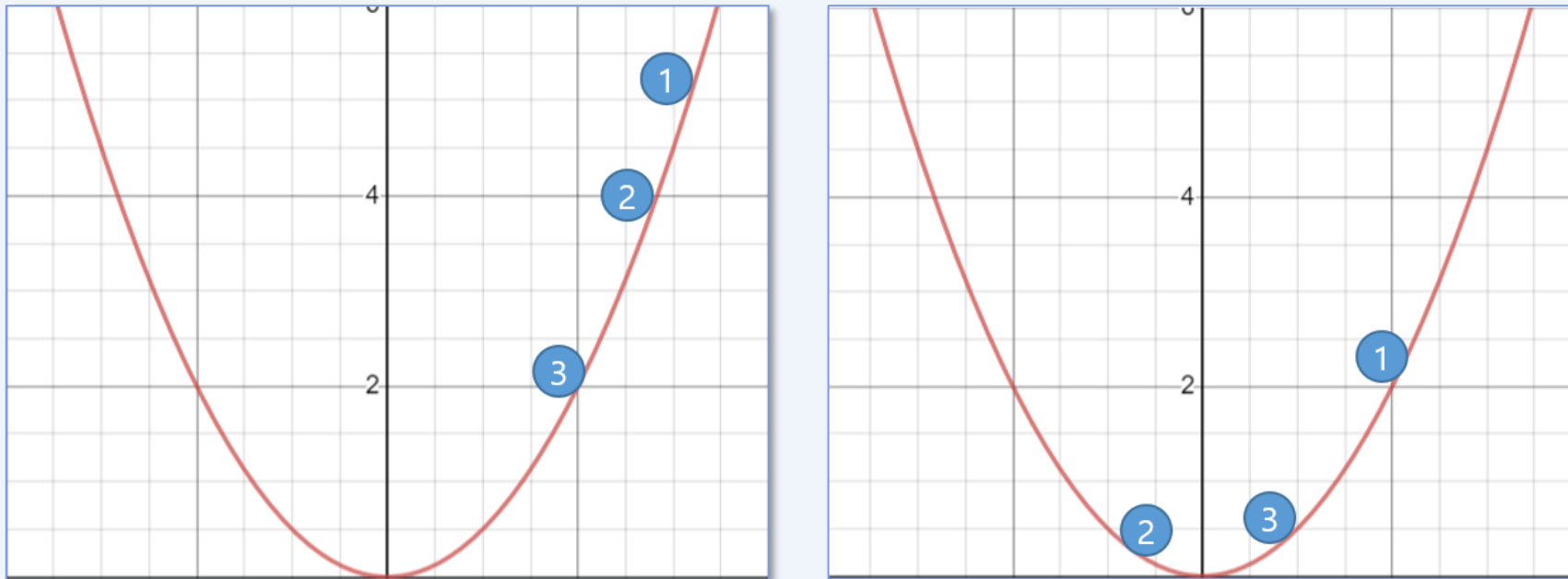


최저점은 하나가 아니다

이런 이유로...
같은 목적과 같은 데이터로 학습하더라도
모든 학습된 모델의 내부(가중치 집합)는 모두 다름

딥러닝 기본 모델-가중치의 조정

- 그 외의 가중치 조절 방법들..
 - SGD, Momentum, AdaGrad, RMSprop, Adam 등



공이 굴러가듯이 움직인다

어떻게 굴러가는지, 어디에서 튀어 오르는지, 관성의 적용을 어떻게 받는지... 등
이런 특징들을 활용 또는 참고하여 새로운 방식을 고려할 수도 있다

- 가중치 계산 방법으로 사용할 수 없는 것은?
 - 딥러닝 모델에서의 가중치 계산, 조정은 미분의 개념을 기반으로 움직임
 - 부드럽게 이어지지 않고 뾰족하거나 각진 형태로 인하여 미분이 불가능한 그라프의 형태를 가지는 계산 모델은 사용할 수 없음

- 출력층에서는 왜 오차를 측정하는가?
 - 신경망의 목적은 정확한 예측 결과를 얻는 것
 - 분류를 위한 모델은? 분류 역시 어떤 클래스가 가장 잘 일치할 것인가..를 예측, 계산하여 그 값이 가장 큰 것을 선택하는 것이므로 동일하다고 볼 수 있음
- 예측을 한 후에는 얼마나 잘 예측했는가 평가해야 함
 - 평가 방법으로 오차의 측정을 사용 → 가장 간단하고 쉬운 방법이므로
- 특히 가중치의 조정은 미분과 관련이 있다는 것은
→ **오차의 값은 양수만 사용**해야 한다는 의미

- 오차의 값은 왜 양수만 사용하는가?
 - 미분은 거리, 넓이를 이용한 개념 → 거리 또는 넓이는 음수가 없음
- 어떻게 양수만으로 오차를 처리할 것인가?
 - 실제로 오차를 측정하면 양수, 음수 모두 나올 수 있지만 각 값을 거리의 개념으로 바꿔서 사용
 - 절대값, 제곱 등을 이용하여 양수로 변환함

- 손실 함수 (Loss Function)

- 출력 값과 정답(기대 값)의 오차를 정의하는 함수
- 손실 함수는 데이터의 특성에 따라 변형, 새로 제안해서 사용 가능

- 종류

- 평균 제곱 오차 (MSE, Mean Squared Error)

- 가장 많이 사용됨. 출력 값과 기대 값의 차이를 제곱하여 평균한 값

$$E = \frac{1}{2} \sum_k (y_k - t_k)^2$$

- 큰 오차는 더욱 크게, 작은 오차는 더욱 작게 → **처리할 때에는 큰 오차에 더 집중**

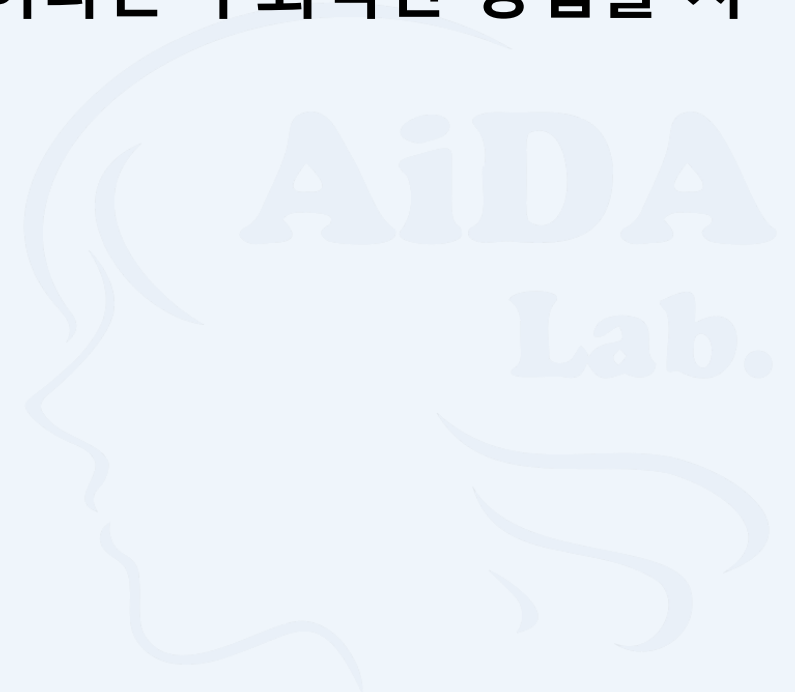
(전반적인 성능 향상에 더 좋음)

- 교차 엔트로피 오차 (Cross Entropy Error)

- 범주형 데이터의 분류에 주로 사용

$$E = - \sum_k t_k \log y_k$$

- 왜 손실함수를 사용하는가?
 - 학습의 궁극적인 목적은 높은 정확도를 끌어내는 매개변수를 찾는 것
 - 왜 "정확도"라는 지표를 놔두고 "손실함수의 값"이라는 우회적인 방법을 사용하는가?



• 왜 손실함수를 사용하는가?

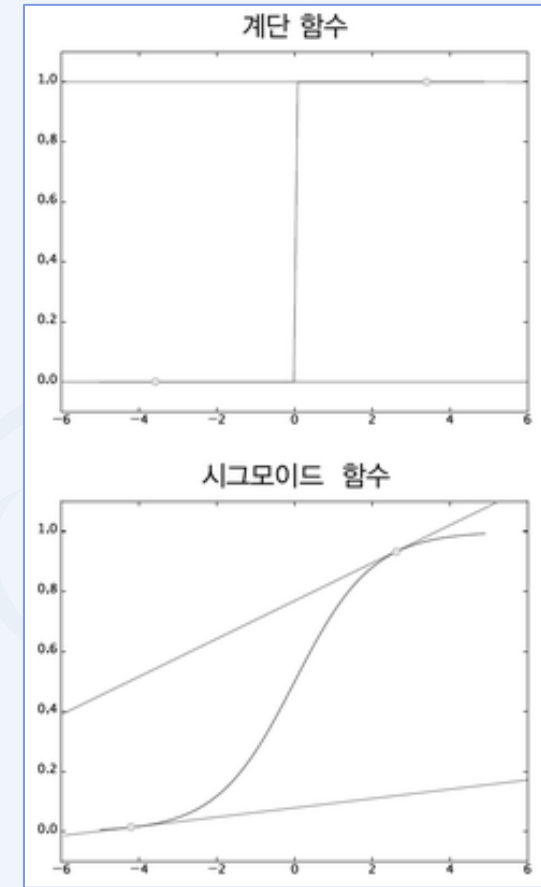
신경망 학습에서의 미분의 역할을 생각해 보면...

- 신경망 학습에서는 최적의 매개변수를 탐색할 때, 손실함수의 값을 가능한 작게 만드는 매개변수의 값을 찾음
- 이때 매개변수의 미분을 계산하고, 그 미분 값을 단서로 매개변수의 값을 서서히 갱신하는 과정을 반복함
 - 손실함수의 미분 값이 음수 \rightarrow 가중치 매개변수를 양의 방향으로 변화시켜 손실함수의 값을 줄일 수 있다.
 - 손실함수의 미분 값이 양수 \rightarrow 가중치 매개변수를 음의 방향으로 변화시켜 손실함수의 값을 줄일 수 있다.
 - 손실함수의 미분 값이 0 \rightarrow 어느 쪽으로도 움직이지 않으므로 갱신이 멈춘다.

- 정확도를 지표로 삼지 않는 이유
 - 미분 값이 대부분의 장소에서 0이 되어 매개변수를 갱신할 수 없다.
 - 정확도는 매개변수의 작은 변화에는 거의 반응을 보이지 않거나, 갑자기 변화한다.
 - 활성화 함수로 '계단 함수'가 아닌 '시그모이드 함수'를 사용하는 이유도 같다.

- 정확도를 지표로 삼지 않는 이유

- 계단함수는 대부분의 장소에서 기울기가 0 이지만, 시그모이드 함수의 기울기(접선)는 0이 아니다.
- 계단 함수는 한순간만 변화를 일으키지만, 시그모이드 함수의 미분은 연속적으로 변한다.
- 즉 시그모이드 함수의 미분은 어느 장소에서도 0이 되지 않는다.
- 이는 **신경망 학습에서 중요한 성질로, 기울기가 0이 되지 않는 덕분에 신경망이 올바르게 학습할 수 있다.**



- **과적합(Over Fitting)**

- 주어진 데이터로 학습을 너무 많이 하면 오히려 역효과!!

- 학습에 입력된 데이터는 완벽에 가깝게 처리함
 - 학습에 입력되지 않은 데이터는 제대로 처리되지 않음

- **원인: 잡음 데이터 (대부분)**

- 불필요한 정보가 많이 포함된 데이터로 학습이 반복됨에 따라 불필요한 정보가 분류의 기준에 포함되어 버리는 것이 원인

- **과적합(Over Fitting)의 해결 방안**

- **조기 종료**

- 적당한 선에서 학습을 종료시킴 → 데이터의 정규화와 관련

- **정규화 (데이터를 일반화 시키기)**

- 필요한 신호는 학습하고 잡음은 제거하는 효과
 - 모델의 학습 난이도를 높임으로써 학습 데이터의 세부 사항(잡음 포함)에 대한 일반화를 활용하도록 하는 기법의 일부로 사용됨

- 과적합(Over Fitting)의 해결 방안

- Drop Out

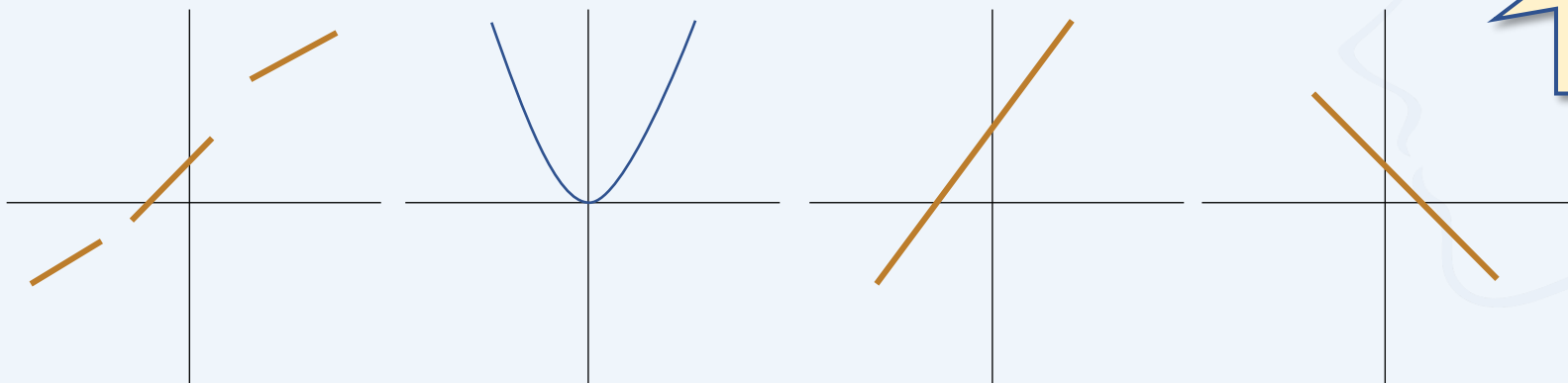
- 학습 중에 무작위로 선택한 뉴런을 0으로 설정
→ 군데군데 망의 연결고리를 잘라 내어 대형 신경망이 소형 신경망처럼 동작하게 만듦
- 소형 신경망에서는 과적합이 거의 발생하지 않음 (표현능력이 협소하기 때문)
- 대형 신경망(딥러닝 모델)을 Drop Out을 통해 소형 신경망처럼 동작하게 하여 과적합 발생률을 떨어뜨리는 방법

• 앙상블

- 두 개의 신경망에 같은 데이터로 학습시키고 일부러 과적합을 유도하면
→ 동일한 잡음이 두 개의 신경망에서 발생하는 과적합의 공통 원인이 될 확률은 극히 낮음
- 각 신경망은 무작위로 설정된 망(가중치 등)으로 학습을 시작하고, 학습 데이터셋의 모든 이미지 사이의 오차를 구분할 잡음이 학습되면 학습이 종료되므로... 시작점이 다르면 각 가중치도 달라질 수 밖에 없음
- 각 신경망은 잡음을 학습하기 이전에 크고 넓은 범위의 특징부터 학습함(잡음도 포함)
- 학습이 진행되면서 서서히 잡음이 결정에 미치는 영향이 커지게 되므로
- 앙상블의 개념을 적용하면 각 신경망은 중요한 신호(데이터)만을 남기고 잡음은 상쇄됨

• 활성화 함수의 조건

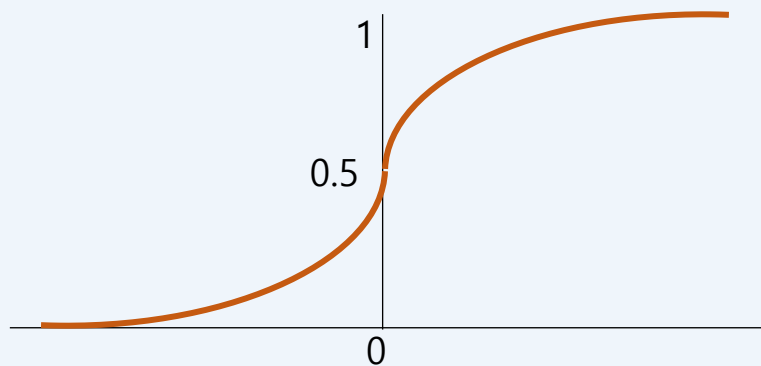
- 정의역(함수에 입력 가능한 값의 범위, 집합) 안에서 연속이며 무한해야 한다
- 단조 함수여야 한다 (방향을 바꾸지 않아야 한다)
- 비선형 함수여야 한다.
- 계산 효율이 좋아야 한다.



이런 함수들은
부적합

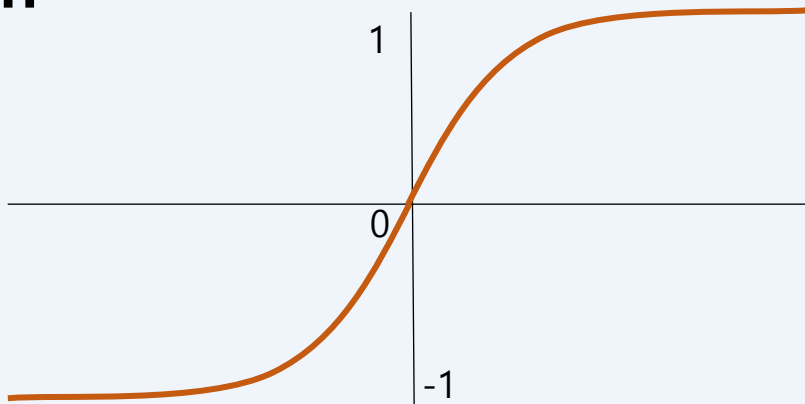
- 표준 은닉 활성화 함수

- Sigmoid



가장 많이 사용되어 왔고
가장 중요한 활성화 함수
(활성화 함수의 기본 형태)

- tanh

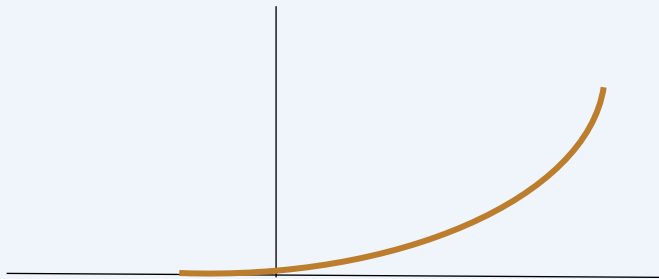


은닉층에서는 sigmoid 보다 tanh 함수가 더 좋음
음의 상관관계도 지원

- 표준 출력 계층 활성화 함수
 - 신경망의 목적에 따라 최선의 선택이 달라진다
 - 일반 데이터값 예측 → 활성화 함수 미적용
 - 서로 무관한 항목에 대한 예/아니오 확률 예측 → sigmoid
 - 여러 가능성 중 하나의 확률 예측 → softmax
- 최근 가장 많이 쓰이는 활성화 함수
 - Relu

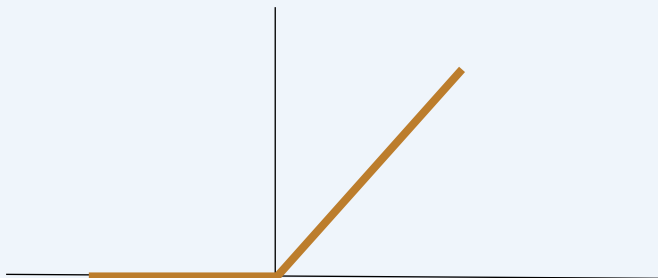
• 표준 출력 계층 활성화 함수

• Softmax



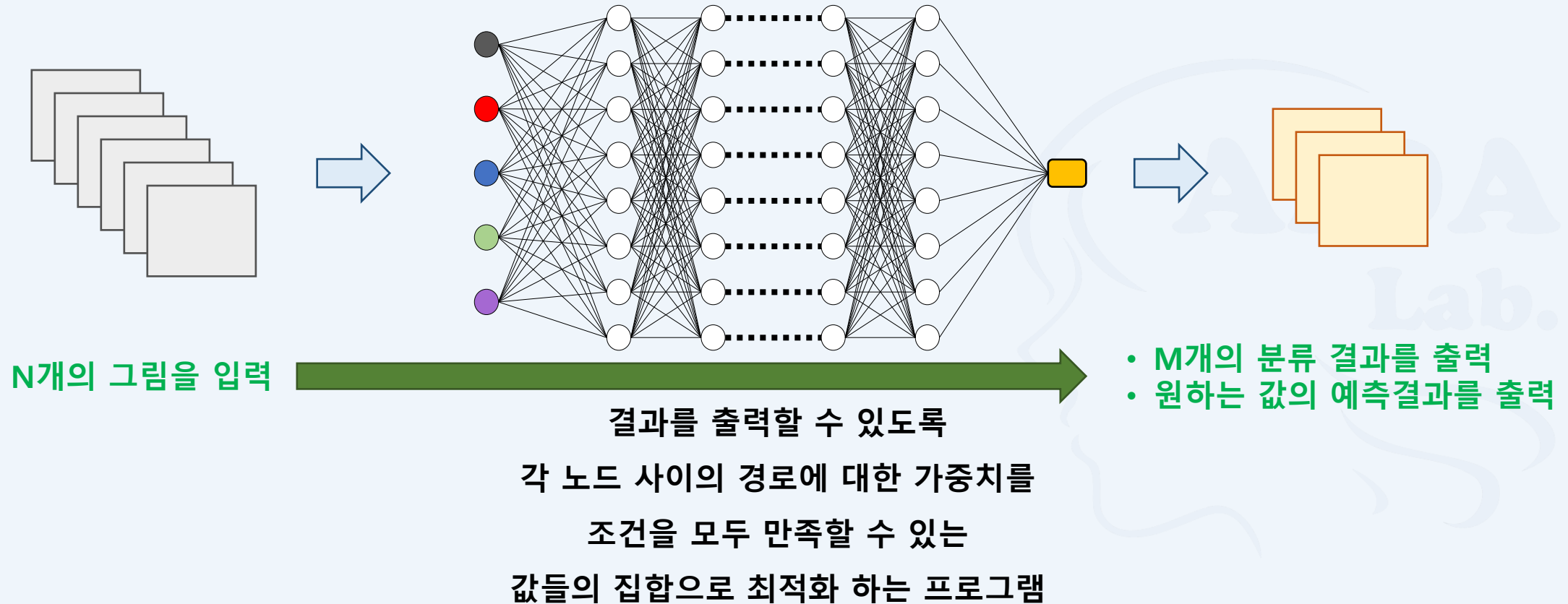
- k개의 값이 존재할 때 각각의 값의 편차를 확대시켜 큰 값은 상대적으로 더 크게, 작은 값은 상대적으로 더 작게 만든 후, 정규화 시키는 함수
- Softmax 함수를 거친 k개의 값의 총합은 1이 됨
- 지수증가를 기반으로 하는 함수

• Relu (Rectified Linear Unit, 정류된 선형 유닛)



- 데이터가 0보다 작으면 무조건 0
- 데이터가 0보다 크면 입력값 그대로!! (주로 선형증가의 형태가 된다)

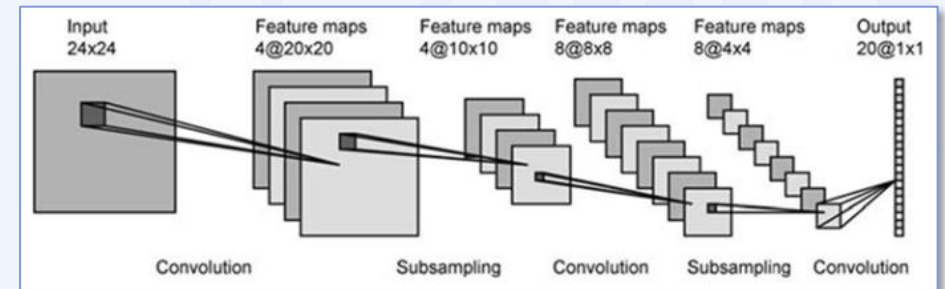
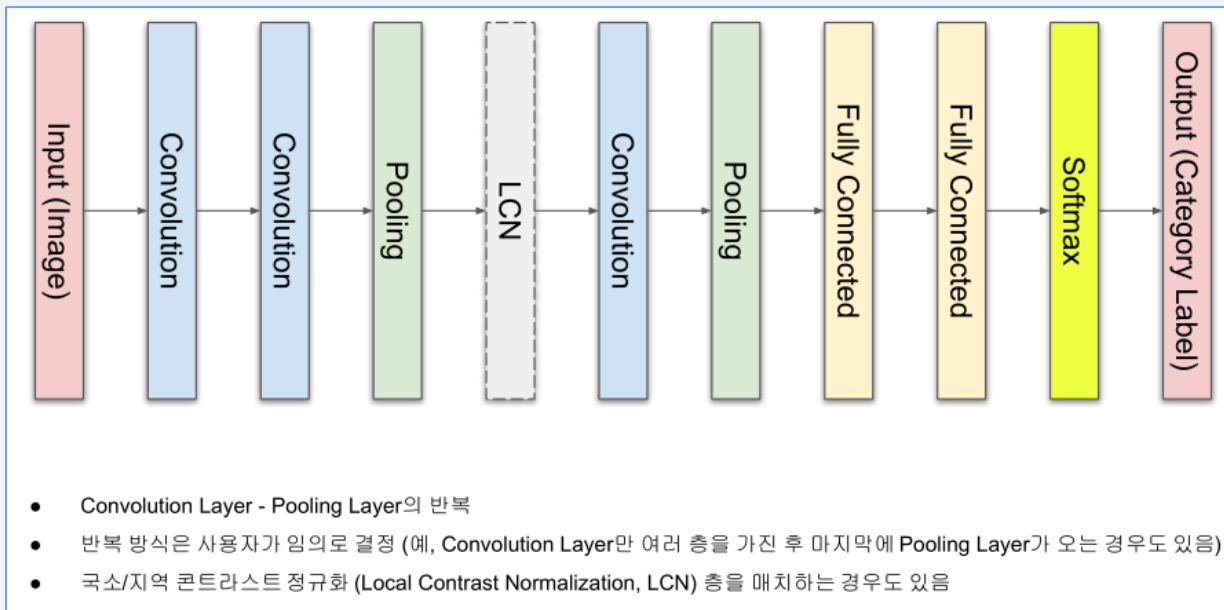
- 엄밀하게 따지면 딥러닝이란...
- 다수의 노드에 적용되는 최적화 프로그램이다.



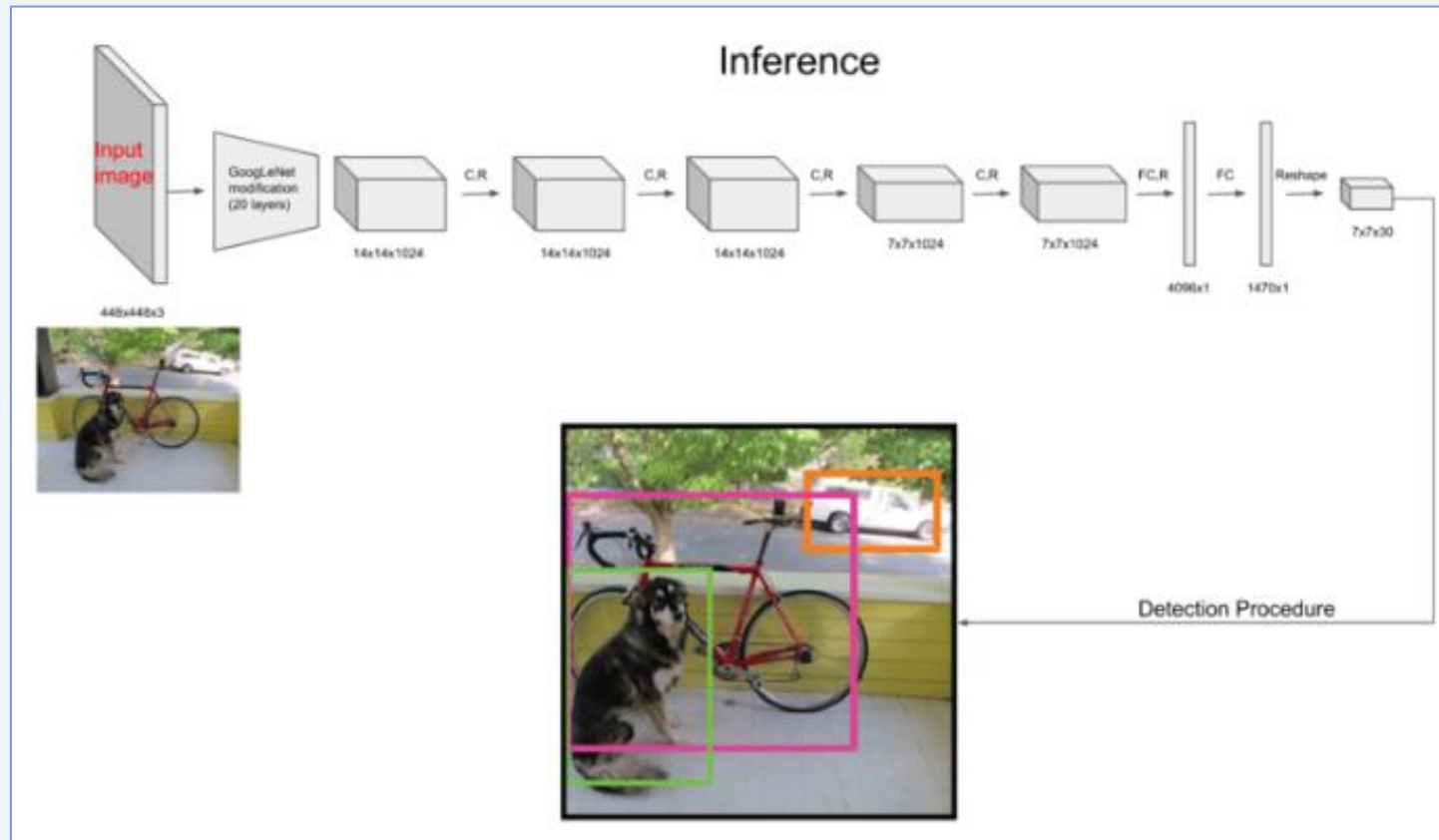
- 그렇다면 이것은 인공지능이 아니지 않나?
 - 애초에 인간의 두뇌 역시 다양한 입력의 결과를 올바르게 출력하기 위한 최적화 과정을 처리함
 - 많이 사용될 수록 굽어지고 민감해지는 각 신경세포 간의 시냅스 결합 강도를
 - 뉴런 사이의 가중치로 모델링하여 구현한 것일 뿐!!

• CNN (Convolutional Neural Network) 모델

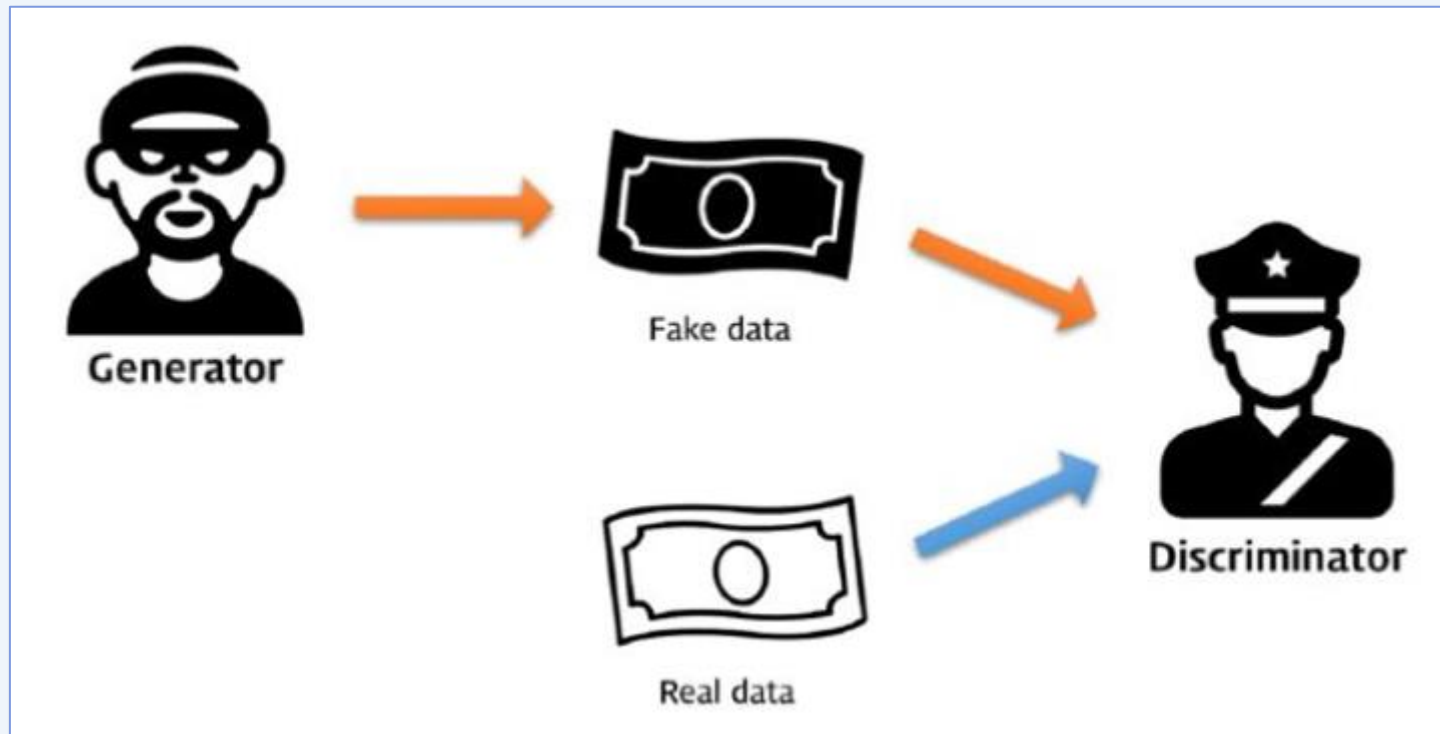
- 최근의 영상 인식, 처리 분야에 있어서 가장 기본이 되는 모델
- 눈과 뇌에서 처리되는 신경과학적 시각처리 방식에서 고안한 모델



- YOLO (You Only Look Once)
 - 지나치게 느린 CNN을 대신하기 위하여 개발된 모델



- GAN (Generative Adversarial Networks)
 - 동시에 두 개의 모델을 훈련하는 머신러닝의 한 종류



- GAN (Generative Adversarial Networks)
 - GAN을 기반으로 다양한 모델이 개발됨



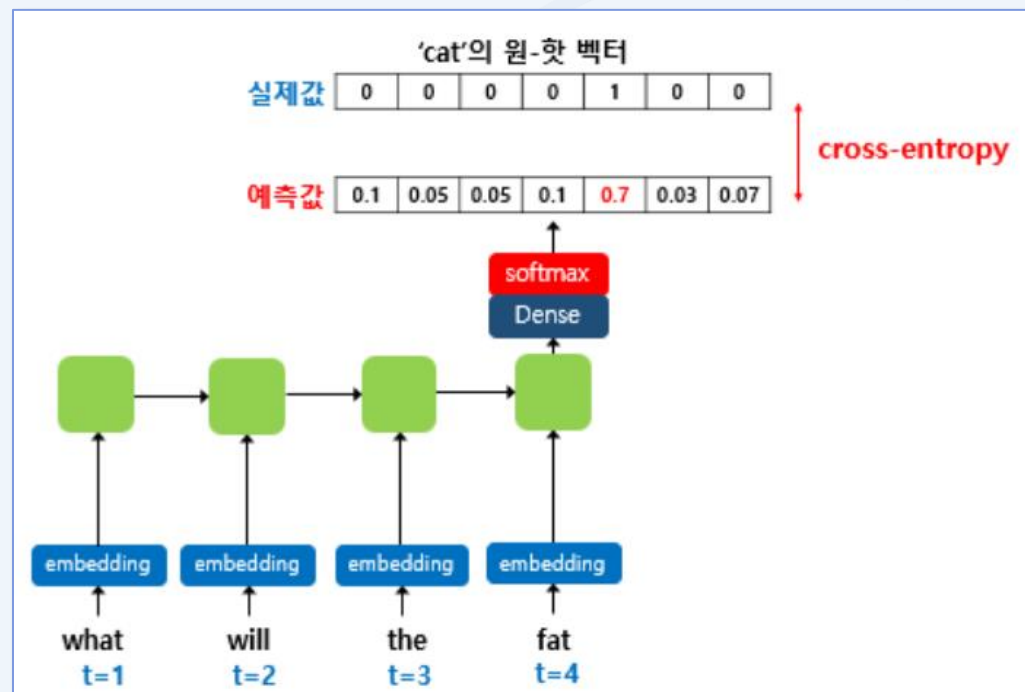
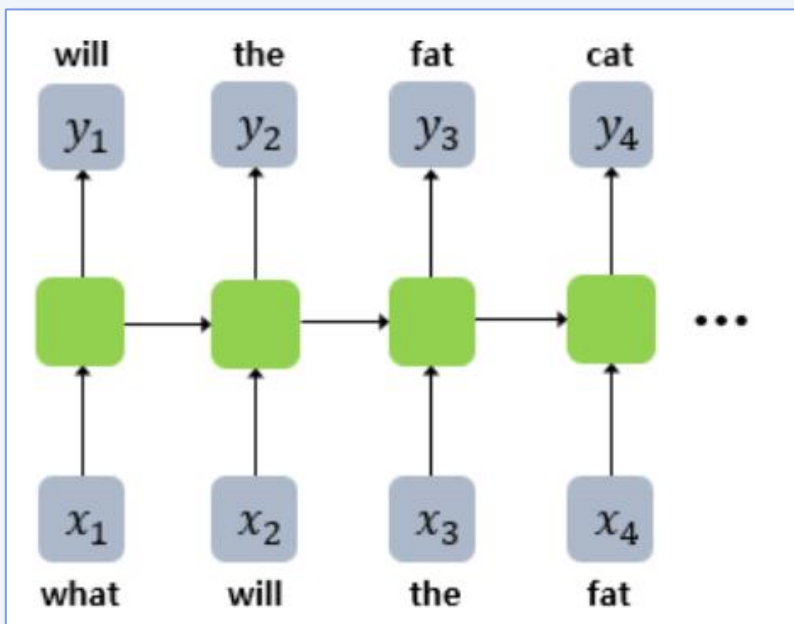
| GAN을 통해 유명인 사진을 바탕으로 만들어진 허구의 인물 <출처: 엔비디아>



| GAN을 활용해 간단한 스케치만으로 시제품을 디자인할 수 있다. <출처: Berkeley AI Research (BAIR)>

- RNN (Recurrent Neural Network)

- 전체 네트워크 안에서 순환적으로 데이터를 처리하는 신경망 모델
- 과거의 데이터를 끊임없이 참조하여 현재의 데이터를 학습하는 모델



예술, 문화 등과 관련된 딥러닝 적용 사례

Art Palette

특정 영상의 색깔과 유사한 색깔의 영상을 검색



<https://artsexperiments.withgoogle.com/artpalette/colors/889e93-3d5d5e-cdc99e-42311d-7a3625>

예술, 문화 등과 관련된 딥러닝 적용 사례

LIFE tags

특정 영상의 사물을 인식하고, 인식된 태그를 LIFE 지의 사진을 검색하여 결과를 제시

COWBOY

WIKIPEDIA DEFINITION

A cowboy is an animal herder who tends cattle on ranches in North America, traditionally on horseback, and often performs a multitude of other ranch-related tasks. The historic American cowboy of the late 19th century arose from the vaquero traditions of northern Mexico and became a figure of special significance and legend. A subtype, called a wrangler, specifically tends the horses used to work cattle. In addition to ranch work, some cowboys work for or participate in rodeos. Cowgirls, first defined as such in the late 19th century, had a less-well-documented historical role, but in the modern world work at identical tasks and have obtained considerable respect for their achievements. Cattle handlers in many other parts of the world, particularly South America and Australia, perform work similar to the cowboy. The cowboy has deep historic roots tracing back to Spain and the earliest European settlers of the Americas. Over the centuries, differences in terrain and climate, and the influence of cattle-handling traditions from multiple cultures, created several distinct styles of equipment, clothing and animal handling. As the ever-practical cowboy adapted to the modern world, his equipment and techniques also adapted, though many classic traditions are preserved.

MORE RELATED LABELS

 Pack Animal	 Bird	 Fish	 Horse Tack	 Horse
 Mine	 Mine	 Horse Supplies	 Animal Sports	 Stallion
 Working Animal	 Horse Trainer	 Saddle	 Horse Harness	 Livestock
 Equestrianism	 Mustang Horse	 Equestrian Sport		

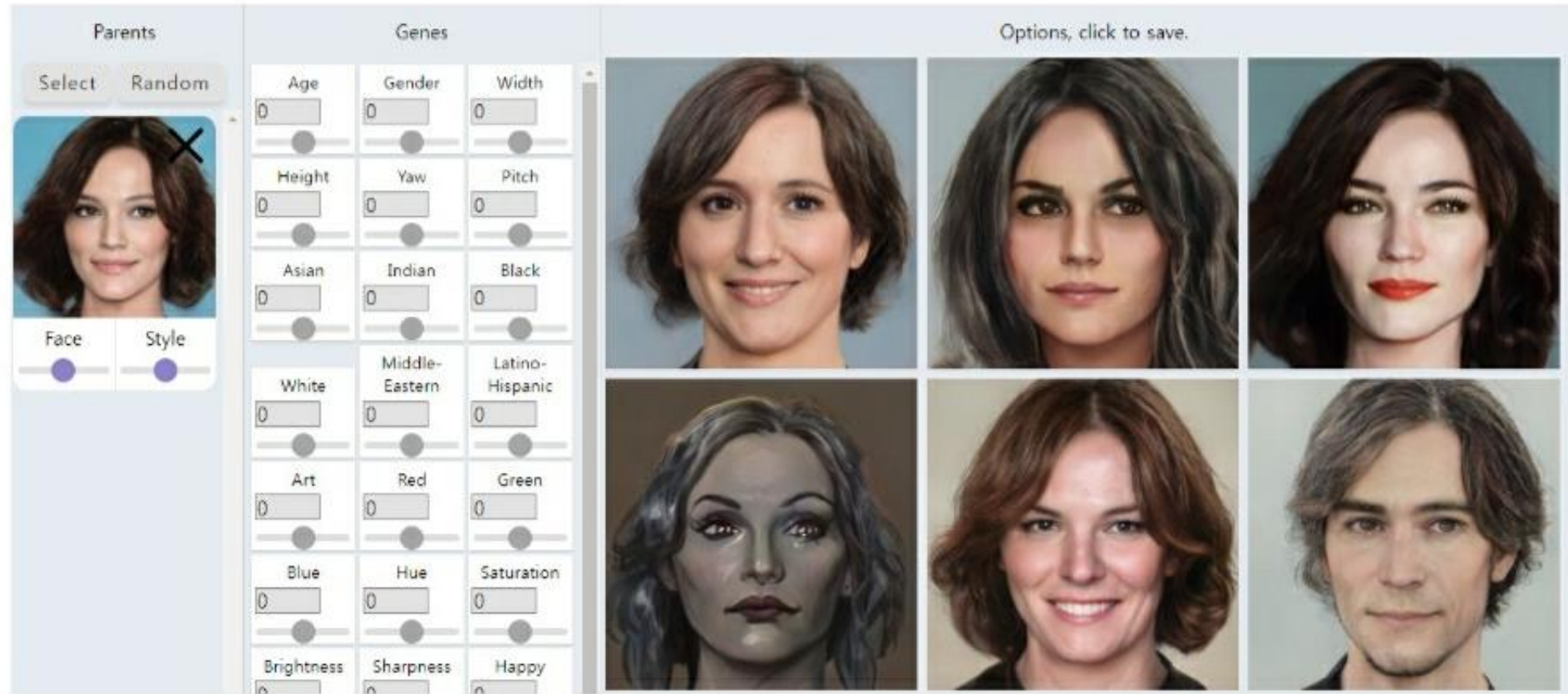


<https://artsexperiments.withgoogle.com/lifetags/>

예술, 문화 등과 관련된 딥러닝 적용 사례

Artbreeder

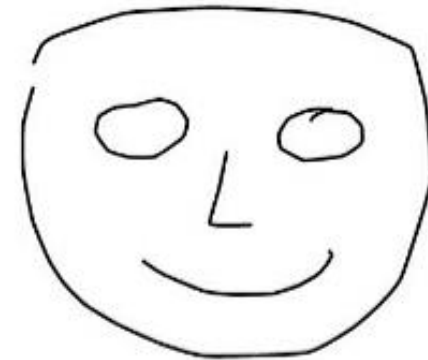
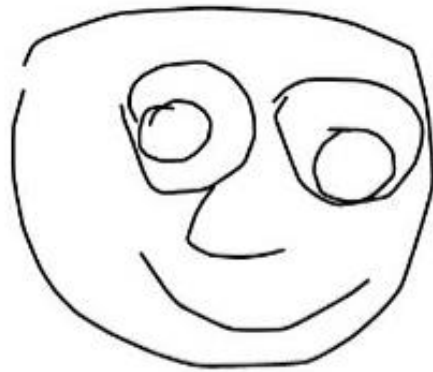
기반 이미지로 새로운 이미지를 생성한다.



<https://artbreeder.com/browse>

Sketch rnn

사람들이 그린 낙서를 기반으로 사용자가 그린 그림을 마저 그린다.



https://magenta.tensorflow.org/assets/sketch_rnn_demo/index.html

예술, 문화 등과 관련된 딥러닝 적용 사례

NSYNTH

복수의 음원 소스로 새로운 소리를 생성

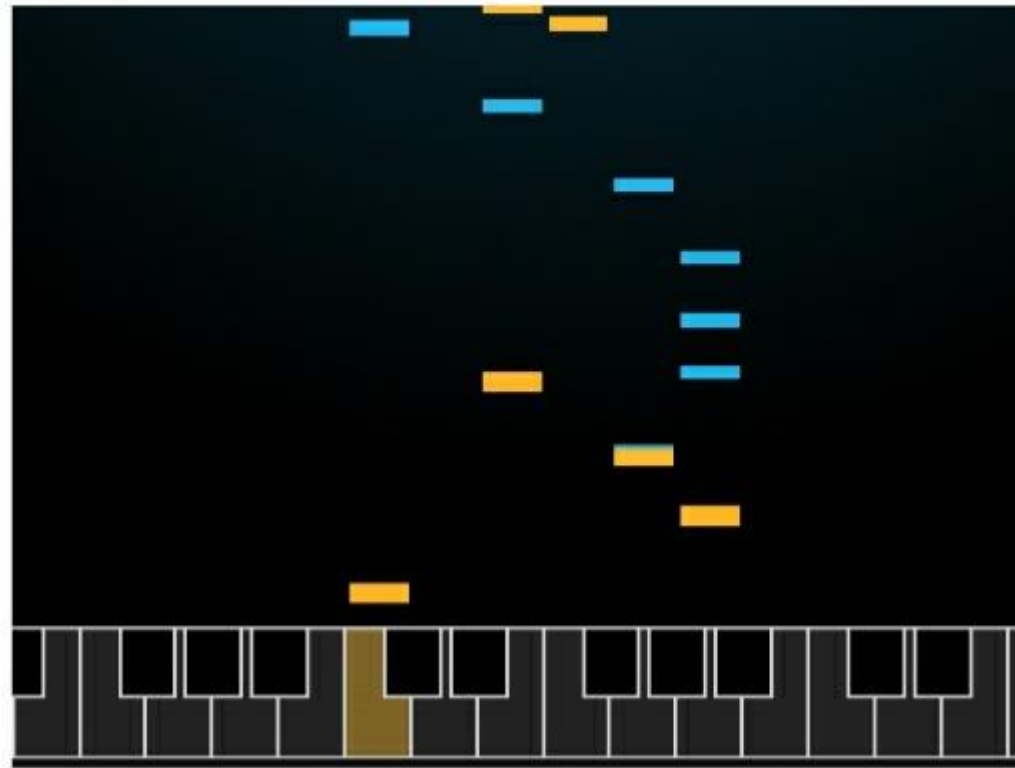


<https://nsynthsuper.withgoogle.com/>
<https://www.youtube.com/watch?v=0fjopD87pyw>

예술, 문화 등과 관련된 딥러닝 적용 사례

AI Duet

사용자의 음악 연주에 AI가 호응 연주



<https://experiments.withgoogle.com/ai/ai-duet/view/>

예술, 문화 등과 관련된 딥러닝 적용 사례

MuseNet 음악 스타일 변경

Compose in the style of Chopin -
starting with
Adele's Someone Like You -

SHOW ADVANCED SETTINGS



↓ DOWNLOAD • TWEET • RESET

<https://openai.com/blog/musenet/>

예술, 문화 등과 관련된 딥러닝 적용 사례

FontJoy 서로 어울리는 폰트를 제안



<https://fontjoy.com/>

예술, 문화 등과 관련된 딥러닝 적용 사례

Assisted Melody 입력한 몇 개의 멜로디에 대한 이후의 멜로디를 제안해서 작곡을 도와 준다.

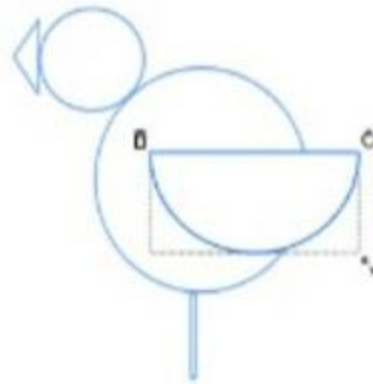


<https://experiments.withgoogle.com/assisted-melody>

예술, 문화 등과 관련된 딥러닝 적용 사례

Draw To Art Shape

모양을 주면 이에 대한 그림을 그려준다.



Gadwall Decoy
Samuel W. Ford
National Gallery of Art, Washington DC



<https://experiments.withgoogle.com/draw-to-art-shape>