2023 로봇스터디 7주차

기본 신경망 이론과 CNN

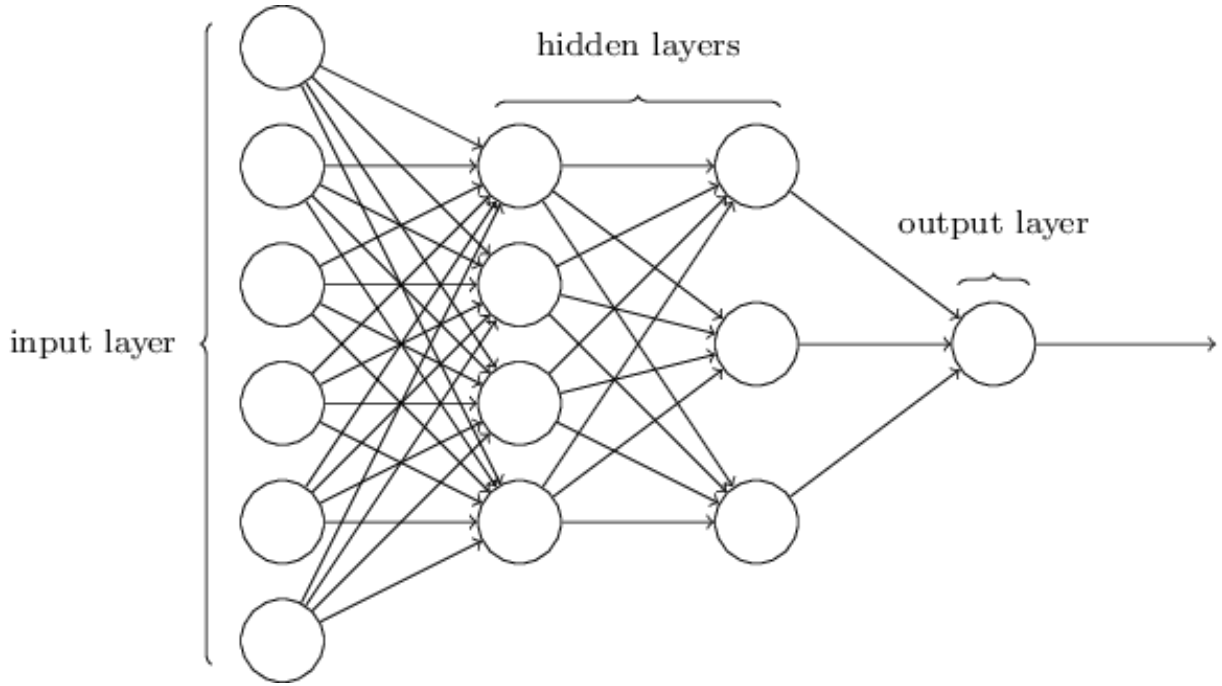
목차

1. 신경망 이론

1. 신경망 이론

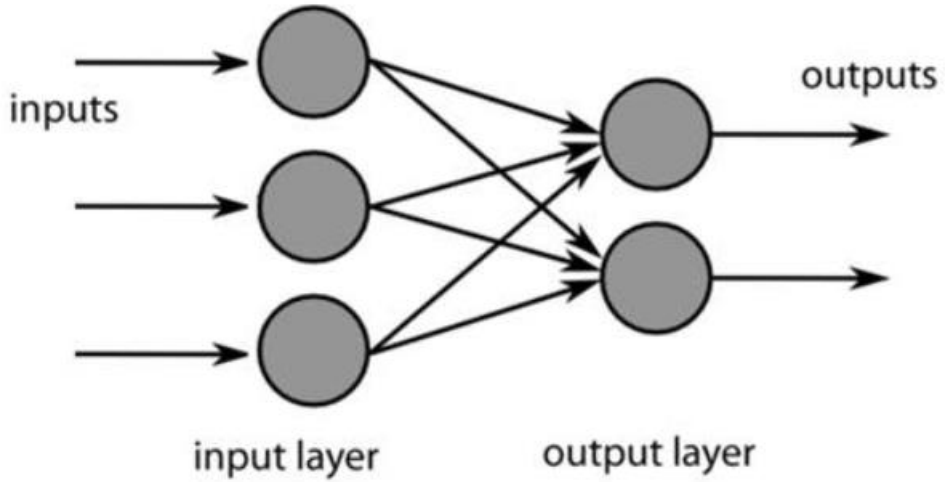
신경망(Neural Network)

인간의 뇌가 가지는 생물학적 특성 중 뉴런의 연결 구조

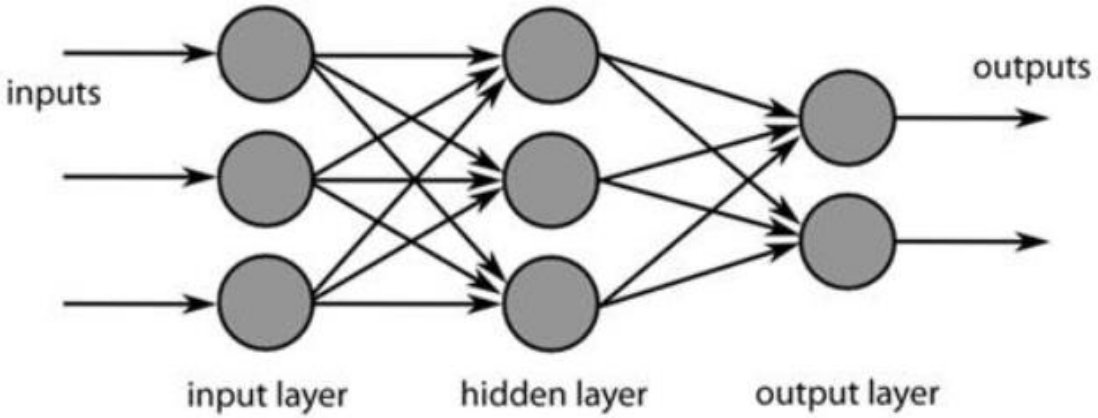


1. 신경망 이론

단층 신경망



다층 신경망



1. 신경망 이론

기울기(gradient)

텐서연산의 변화율(=경사도, 변화도)

1. 신경망 이론

손실 함수(Activation Functions)

지도학습 시 알고리즘이 예측한 값과 실제 정답의 차이를 비교하기 위한 함수

- '학습 중에 알고리즘이 얼마나 잘못 예측하는 정도'를 확인하기 위한 함수로써 최적화(Optimization)를 위해 최소화하는 것이 목적

- 목적 함수(Objective Function)라고도 부름

1. 신경망 이론

역전파(Back Propagation)

신경망의 각 노드가 가지고 있는 가중치(weight)와 편향(bias)을 학습시키기 위한 알고리즘

- 딥러닝에 있어 가장 핵심적인 부분
- 목표(target)와 모델의 예측결과(output)의 차이를 바탕으로 가중치와 편향을 뒤에서부터 앞으로 갱신해 나감
- 경사하강법(Gradient Descent)을 통해 오차를 최소화 함

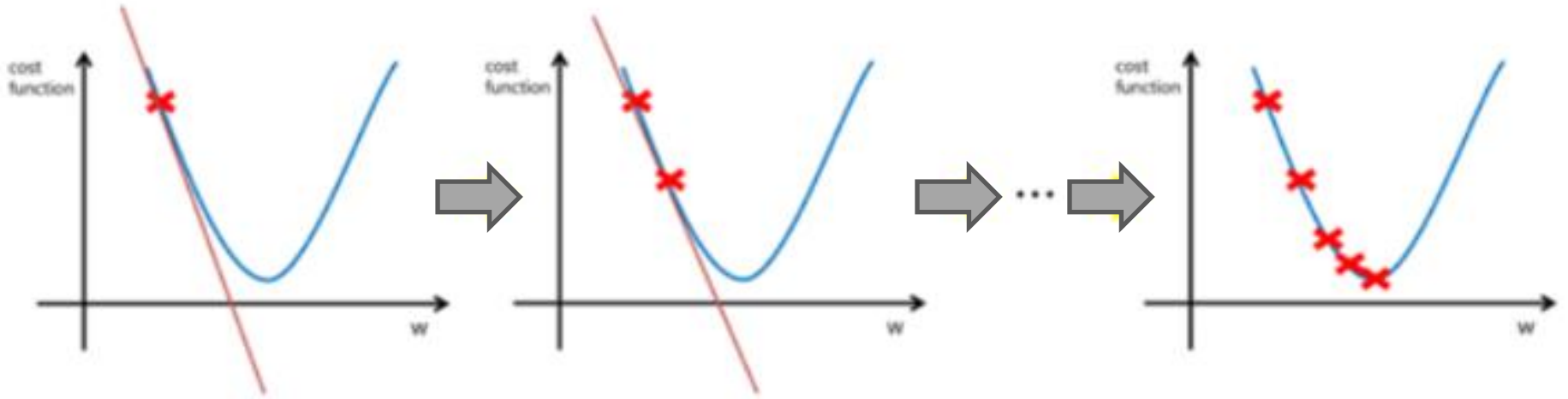
1. 신경망 이론

경사하강법(Gradient Descent)

함수의 기울기(gradient)를 이용해 경사의 반대 방향으로 계속 이동시켜

극 값에 이를 때까지 반복시키는 최적화 알고리즘

- 손실함수의 크기를 최소화하는 방향으로 파라미터를 업데이트하기 위해 사용
- 기울기가 최소값 일 때 최적의 파라미터를 찾는다

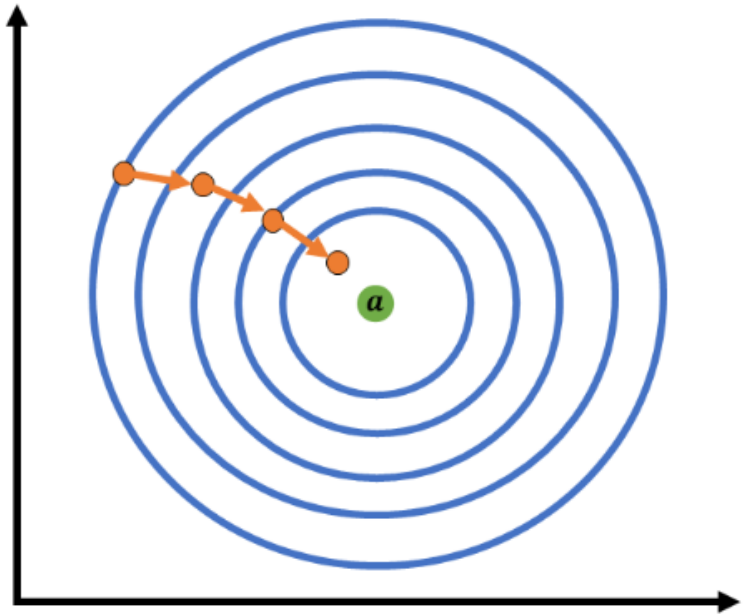


1. 신경망 이론

경사하강법(Gradient Descent)

- 전체 훈련 데이터셋을 대상으로 학습
- 한계 : 파라미터가 한 번 이동할 때마다 계산해야 할 값이 지나치게 많음

→ 학습 데이터셋이 커지면 커질수록 시간과 리소스 소모가 지나치게 큼

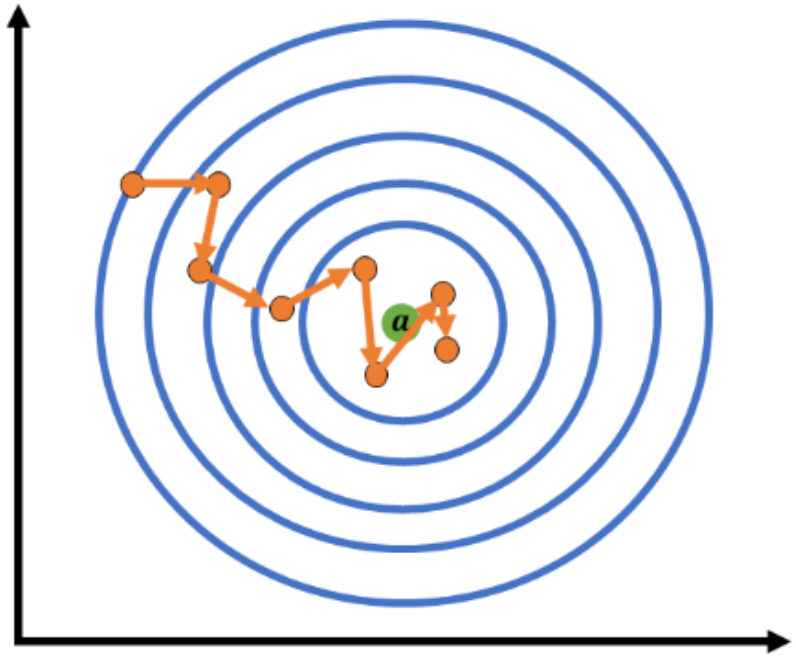


1. 신경망 이론

확률적 경사하강법(Stochastic gradient descent)

학습 데이터셋에서 무작위로 한 개의 샘플 데이터 셋을 추출하고 그 샘플에 대해서만 기울기를 계산하는 경사하강법

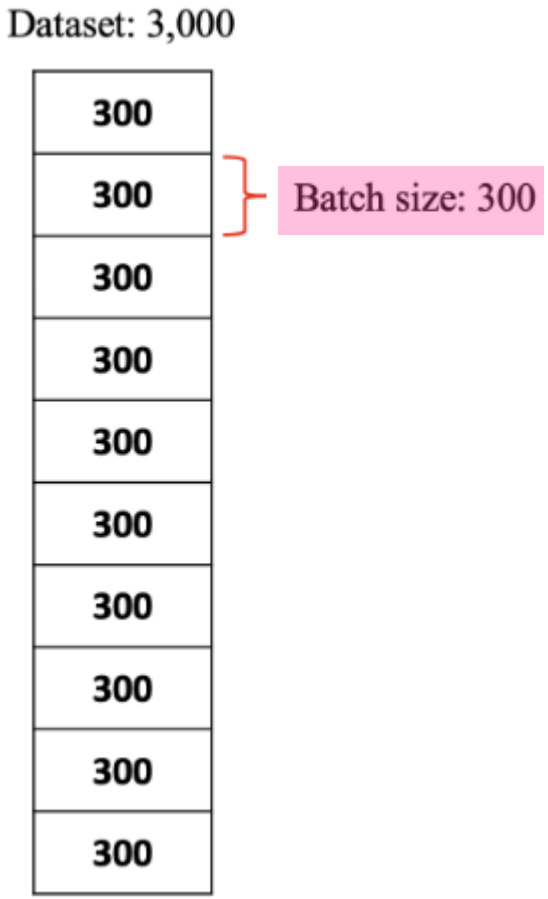
- 샘플 데이터셋에 대해서만 경사를 계산
- 매 반복에서 다뤄야 할 데이터 수가 매우 적어 학습 속도가 빠름



1. 신경망 이론

batch size(배치 사이즈)

: 파라미터를 한 번 업데이트 시킬 때 사용할 데이터 개수



1. 신경망 이론

Mini-Batch(미니배치)

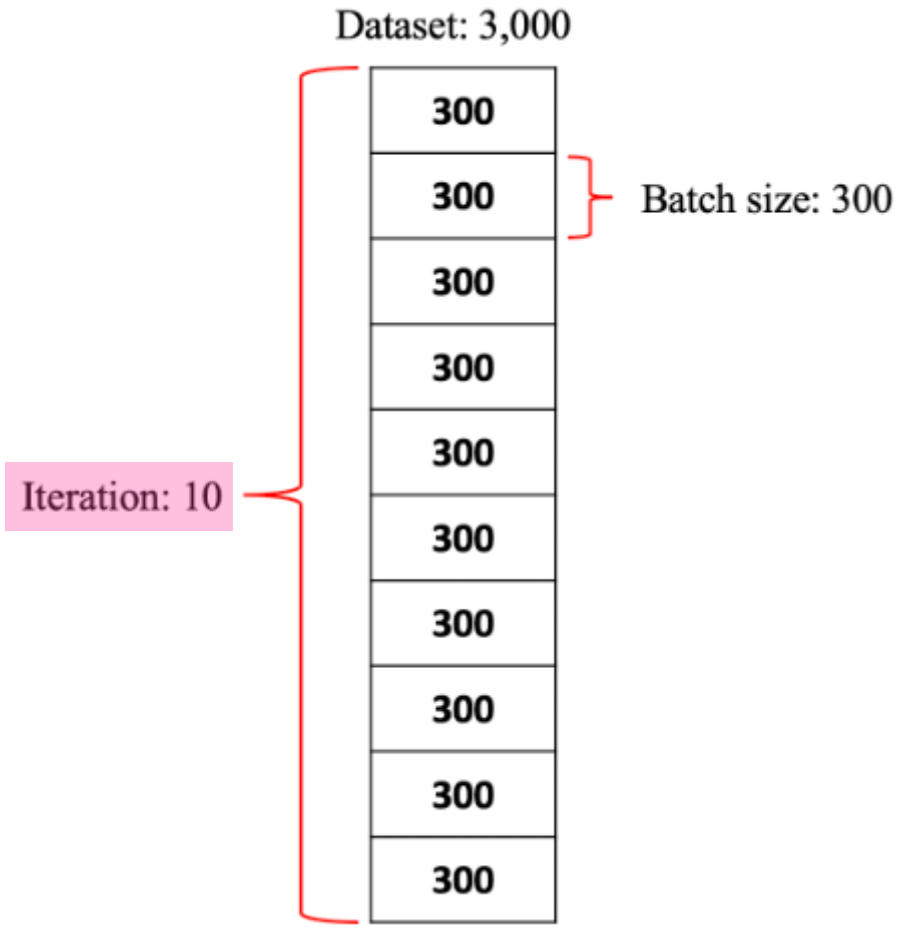
: 전체 데이터를 N등분하여 각각의 학습 데이터를 배치 방식으로 학습

- 확률적 경사 하강법과 배치를 섞은 것
- 신경망을 한 번 학습시키는데(Iteration) 걸리는 시간은 줄이면서 전체 데이터 반영하여 효율적으로 GPU 활용가능

1. 신경망 이론

Iteration(이터레이션)

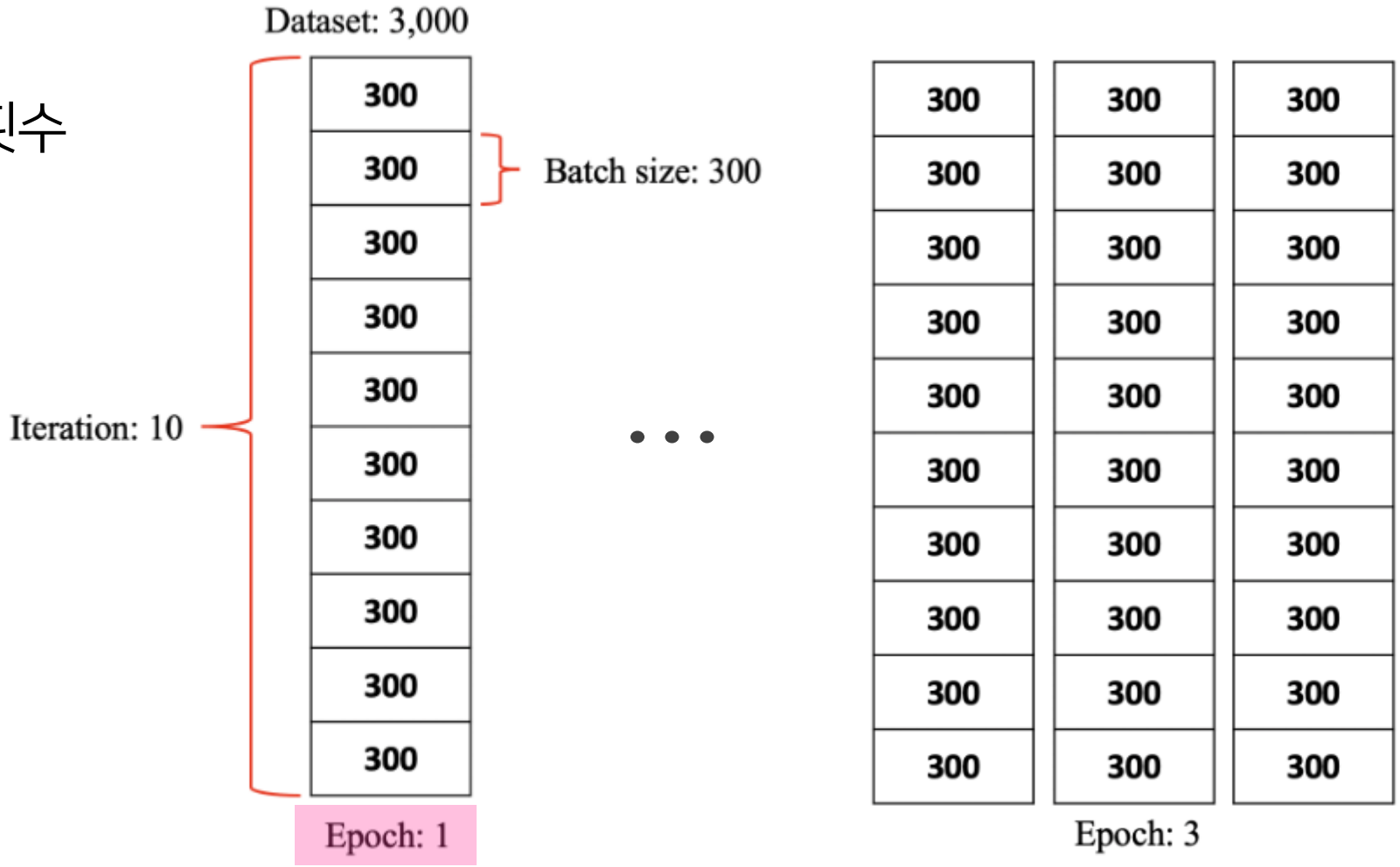
: 전체 데이터를 모델에 한 번 학습시키는데 필요한 배치의 수



1. 신경망 이론

epoch(에포크)

: 전체 데이터셋을 학습한 횟수

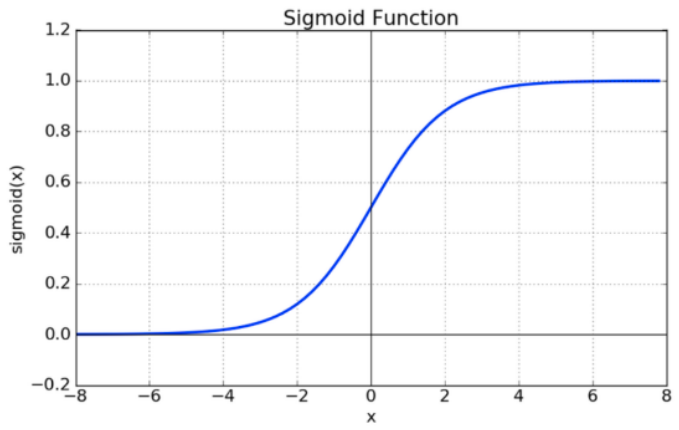


1. 신경망 이론

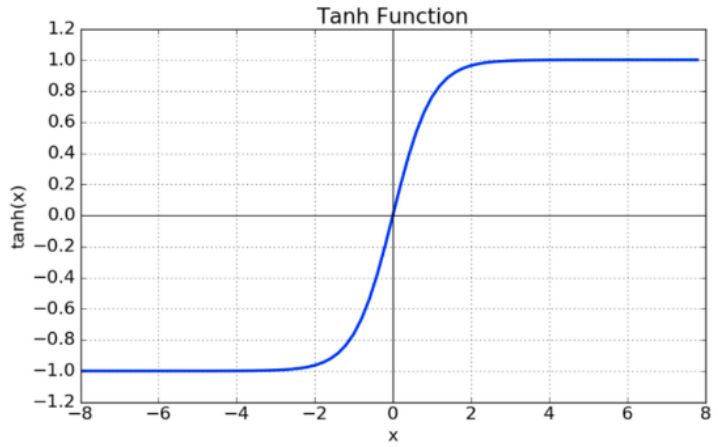
활성화 함수(Activation Functions)

입력 신호의 총합을 출력 신호로 변환하는 함수

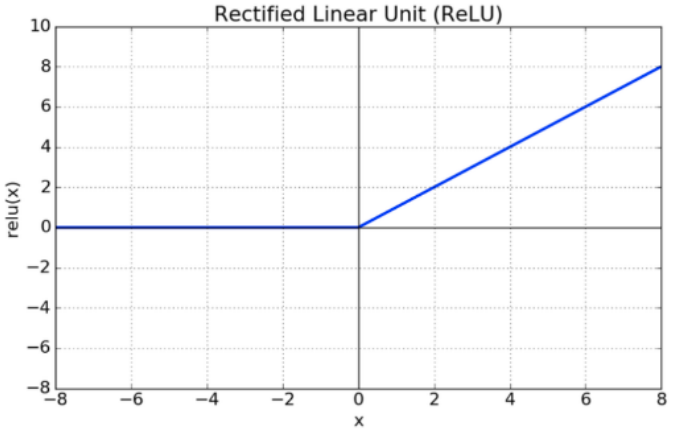
1. 시그모이드(sigmoid) 함수



2. tanh 함수



3. ReLU 함수



1. 신경망 이론

하이퍼파라미터

최적의 훈련 모델을 구현하기 위해 모델에 설정하는 변수로 학습률(Learning Rate), 에포크 수(훈련 반복 횟수), 가중치 초기화 등을 결정

- 개발자에 의해 임의로 조정 가능

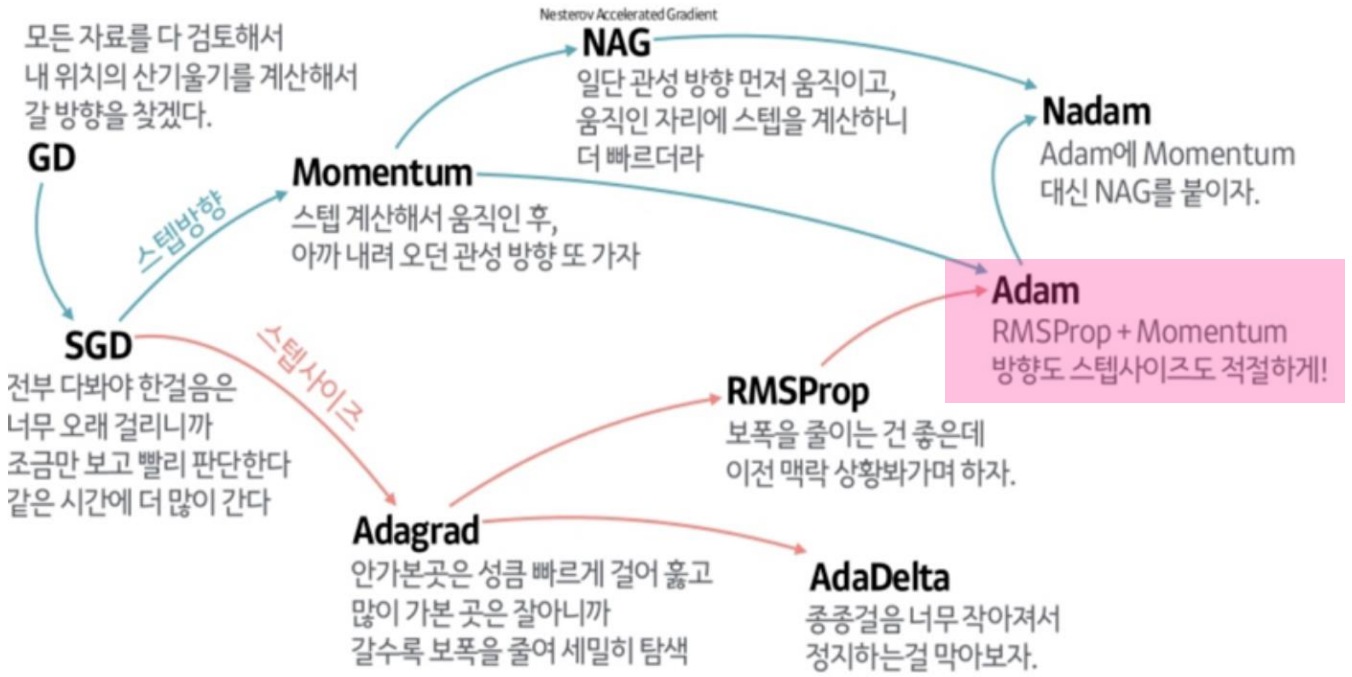
모델 파라미터	하이퍼파라미터
새로운 샘플이 주어지면 무엇을 예측할지 결정하기 위해 사용하는 것이며 학습 모델에 의해 결정	학습 알고리즘 자체의 파라미터 절대적인 최적값은 존재하지 않고 사용자가 직접 설정

1. 신경망 이론

옵티마이저

손실함수의 최솟값을 찾아가는 것을 최적화하는 알고리즘

- 학습속도를 빠르고 안정적이게 하는 것을 목표



1. 신경망 이론

Adam (Adaptive Moment Esimation)

진행하던 속도에 관성을 주고, 최근 경로의 곡면의 변화량에 따른 적응적 학습률을 갖는 알고리즘

- 현재 가장 많이 사용되고 있는 최적화 알고리즘

```
# tensorflow2.x
tf.keras.optimizers.Adam(lr=0.001, beta_1=0.9, beta_2=0.99, epsilon=None)
```

2. 모델평가

2. 모델평가

<데이터 전처리>

왜곡된 분석결과를 방지하기 위해 분석에 적합하게 데이터를 가공하여

데이터의 품질을 올리는 과정

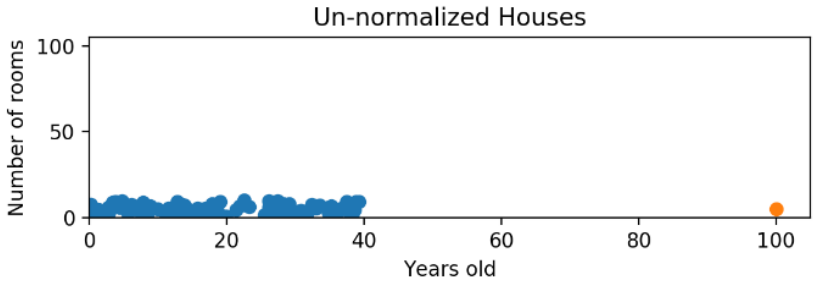
2. 모델평가

정규화(Data Normalization)

모델 학습데이터를 과도하게 학습하게 되어 학습데이터 이외의 새로운 데이터에 대해
예측을 하지 못하는 현상

1. Min-Max Normalization (최소-최대 정규화)

모든 feature에 0과 1 사이의 값으로 변환



-> 이상치(outlier) 문제 발생

2. Z-Score Normalization (Z-점수 정규화)

이상치(outlier)를 잘 처리하지만, 정확히 동일한 척도로 정규화 된 데이터를 생성하지는 않음

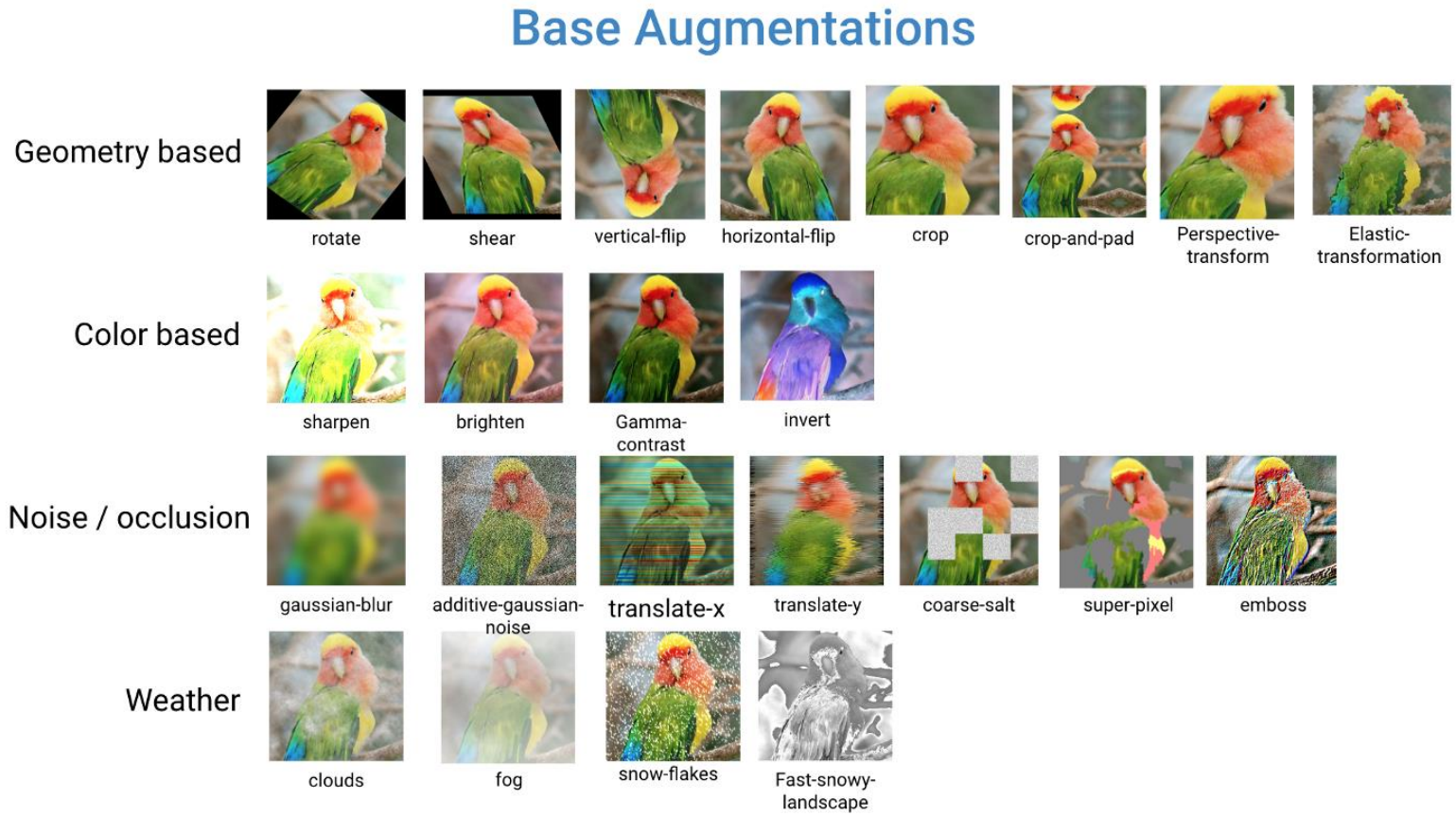
2. 모델평가

데이터 증강(Data Agumentation)

데이터의 양을 늘리기 위해 원본에 각종 변환을 적용하여 개수를 증강시키는 기법

- training dataset size 증가

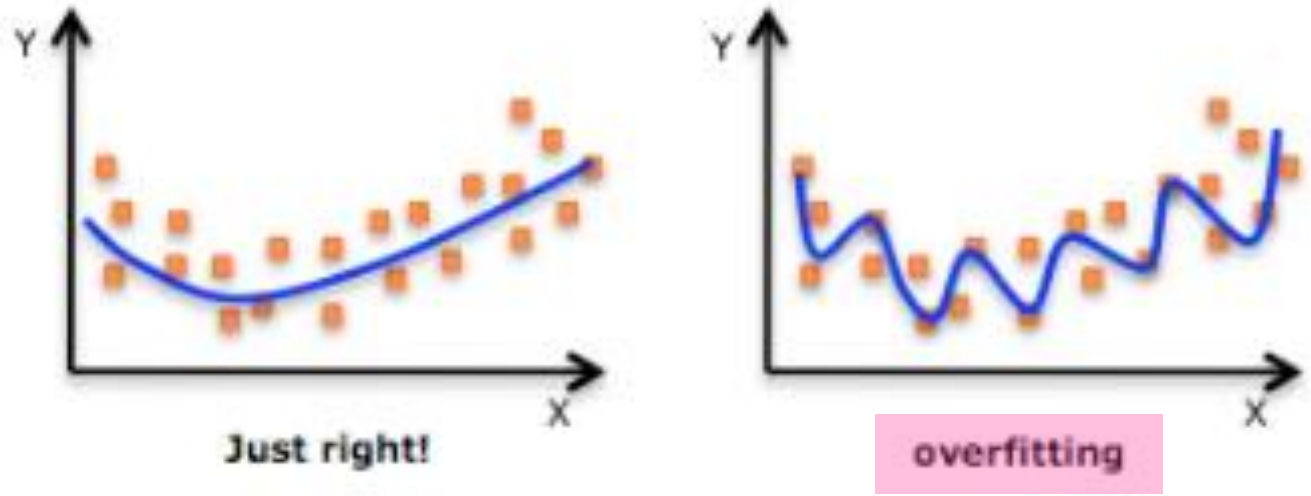
- rotation(회전)
- scaling(크기 줄이기)
- reflection(뒤집기)
- cropping(자르기)
- noise(노이즈 삽입)
- mixing image(모자이크)



2. 모델평가

오버피팅(overfitting)

모델 학습데이터를 과도하게 학습하게 되어 학습데이터 이외의 새로운 데이터에 대해 예측을 하지 못하는 현상



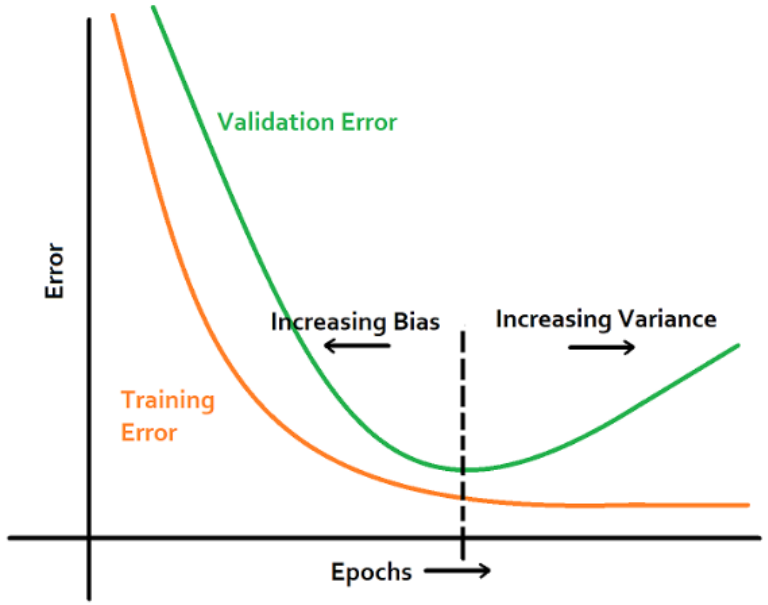
2. 모델평가

오버피팅(overfitting) 해결방안

- 더 많은 dataset 확보

- Dropout 기법 (추후 설명)

- EarlyStopping
  → Validation Acc 가 감소하는 지점에서 학습을 중단

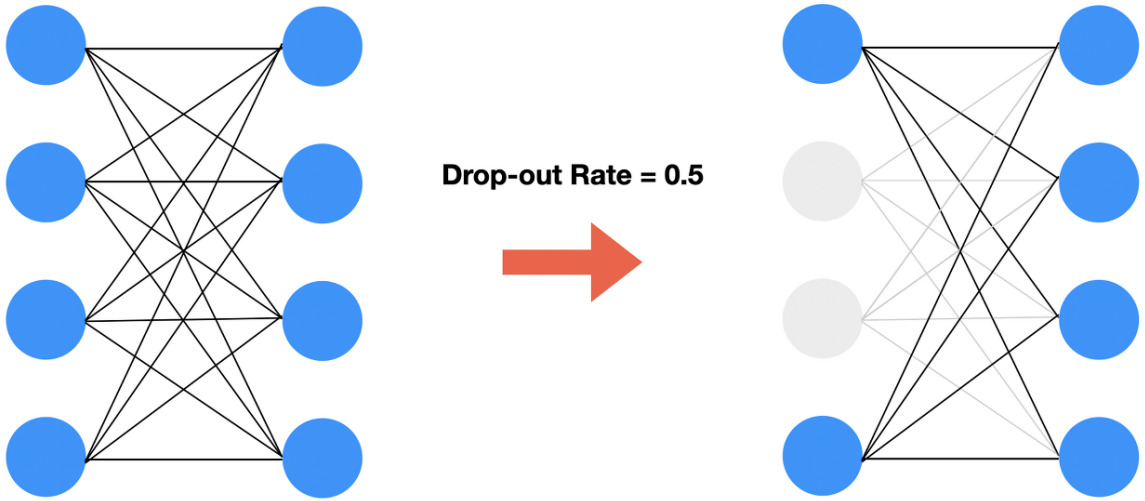


2. 모델평가

드롭아웃(Drop-out)

서로 연결된 연결망(layer)에서 0부터 1 사이의 확률로 뉴런을 제거(drop)하는 기법

- 과적합 방지



2. 모델평가

소프트맥스 함수(softmax)

입력받은 값을 0~1 사이의 출력이 되도록 정규화하여 출력 벡터들의 총합이 항상 1이 되는 특성을 가진 함수

- 딥러닝에서는 출력 노드의 활성화 함수로 많이 사용됨

$$\mathrm{softmax}_i(x) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}$$

$$\text{where } x \in \mathbb{R}^n.$$

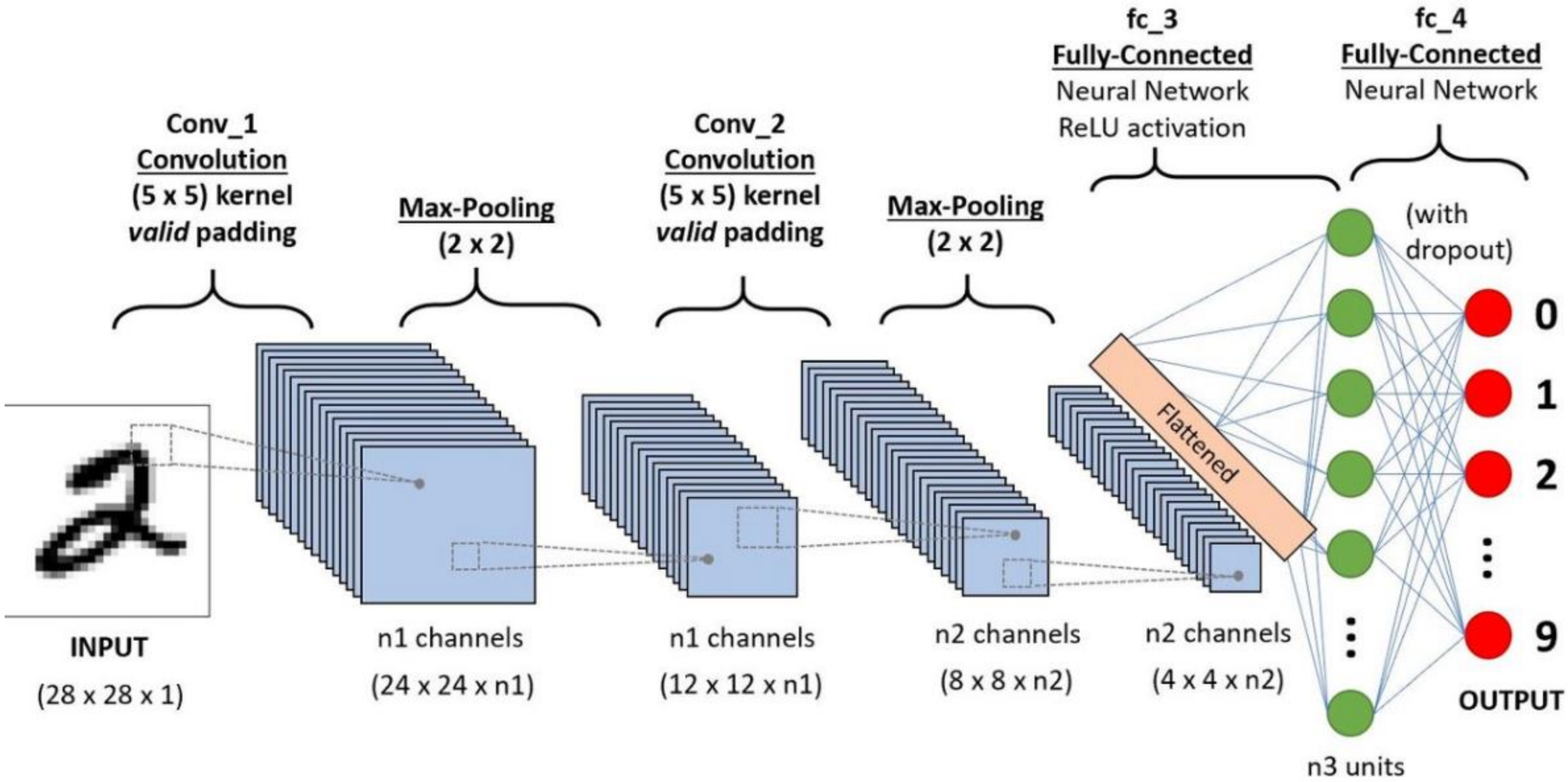
2. 모델평가

크로스엔트로피 함수(Cross-Entropy)

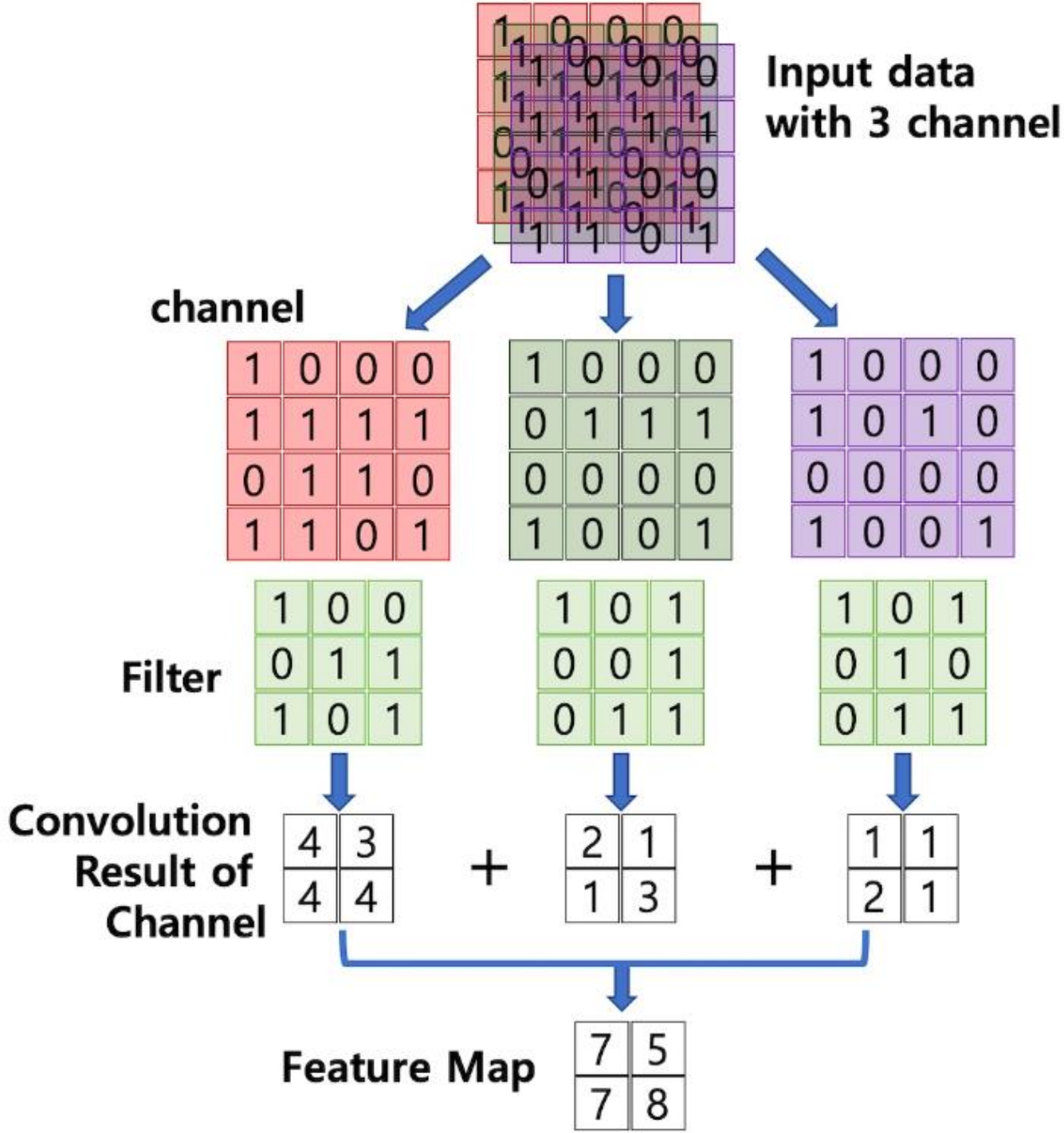
실제 데이터의 확률 분포와, 학습된 모델이 계산한 확률 분포의 차이를 구하는데 사용되는 손실 함수

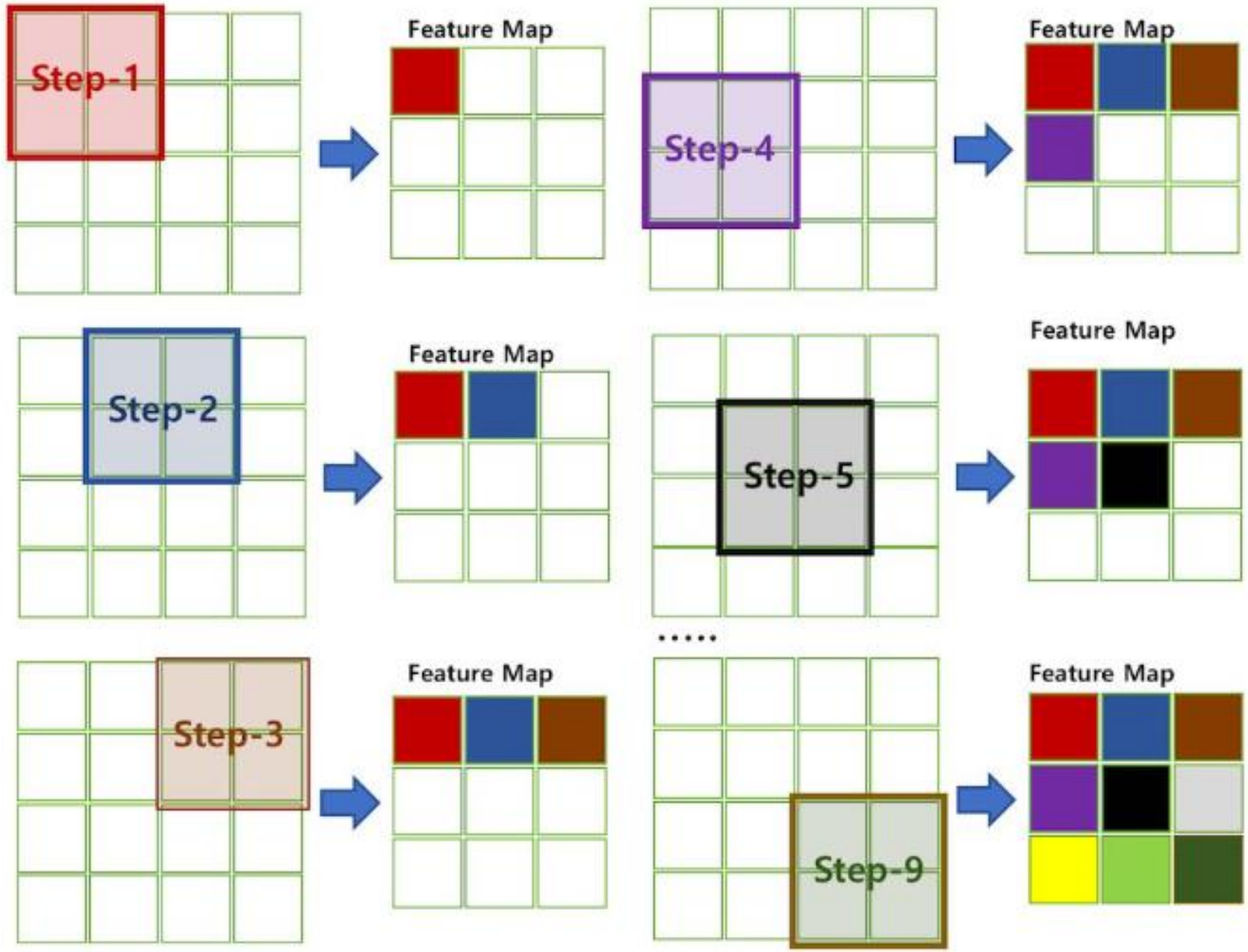
$$H_p(q) = -\sum_{i=1}^{n} q(x_i) log p(x_i)$$

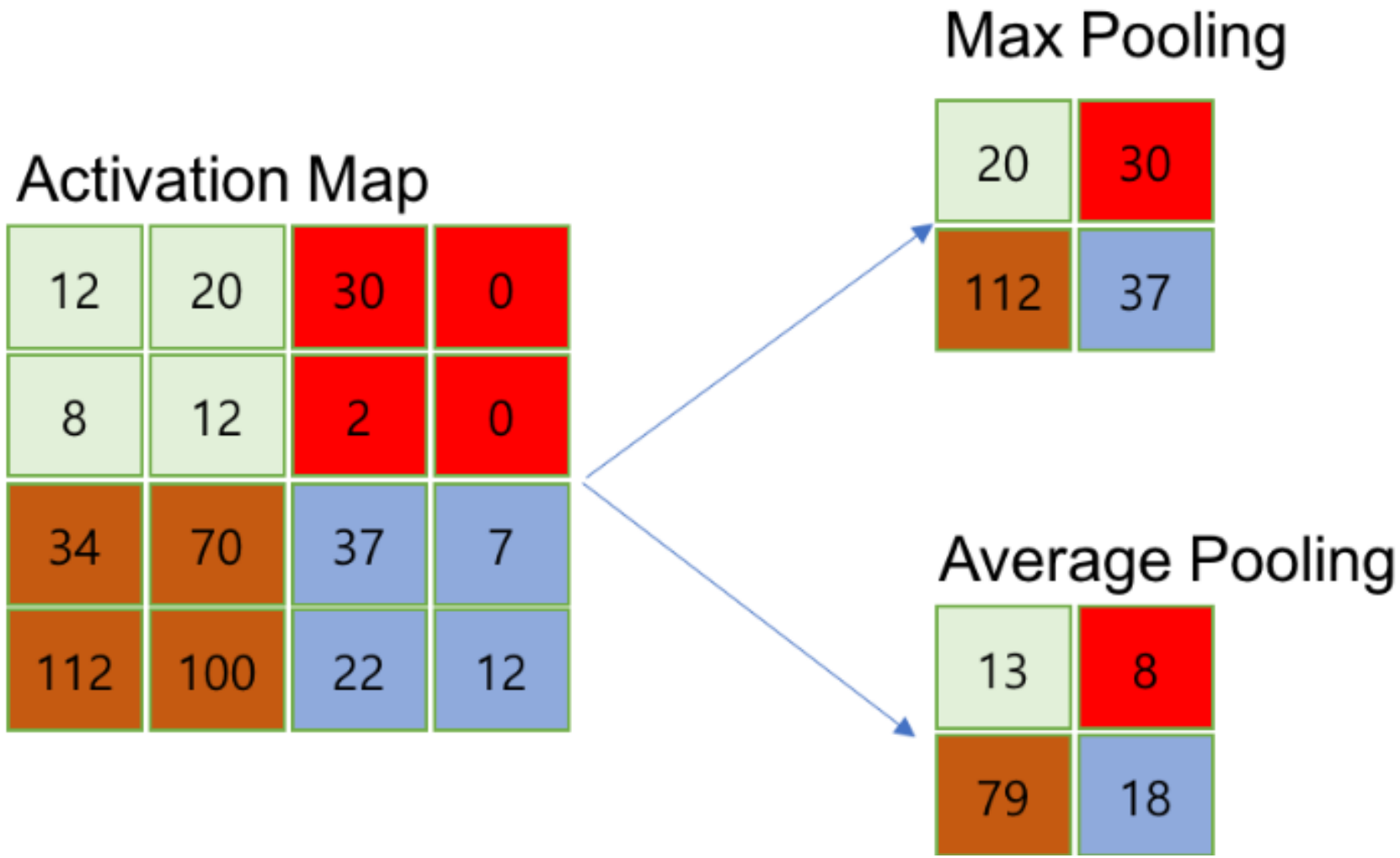
3. CNN



전체 CNN의 과정입니다.


Input data
with 3 channel
channel
1 0 0 0
1 1 1 1
0 1 1 0
1 1 0 1
1 0 0 0
0 1 1 1
0 0 0 0
1 0 0 1
1 0 0 0
1 0 1 0
0 0 0 0
1 0 0 1
Filter
1 0 0
0 1 1
1 0 1
1 0 1
0 0 1
0 1 1
1 0 1
0 1 0
0 1 1
Convolution
Result of
Channel
4 3
4 4
+
2 1
1 3
+
1 1
2 1
Feature Map
7 5
7 8


Step-1
Feature Map
Step-4
Feature Map
Step-2
Feature Map
Step-5
Feature Map
.....
Step-3
Feature Map
Step-9
Feature Map


Activation Map
12
20
30
0
8
12
2
0
34
70
37
7
112
100
22
12
Max Pooling
20
30
112
37
Average Pooling
13
8
79
18

감사합니다