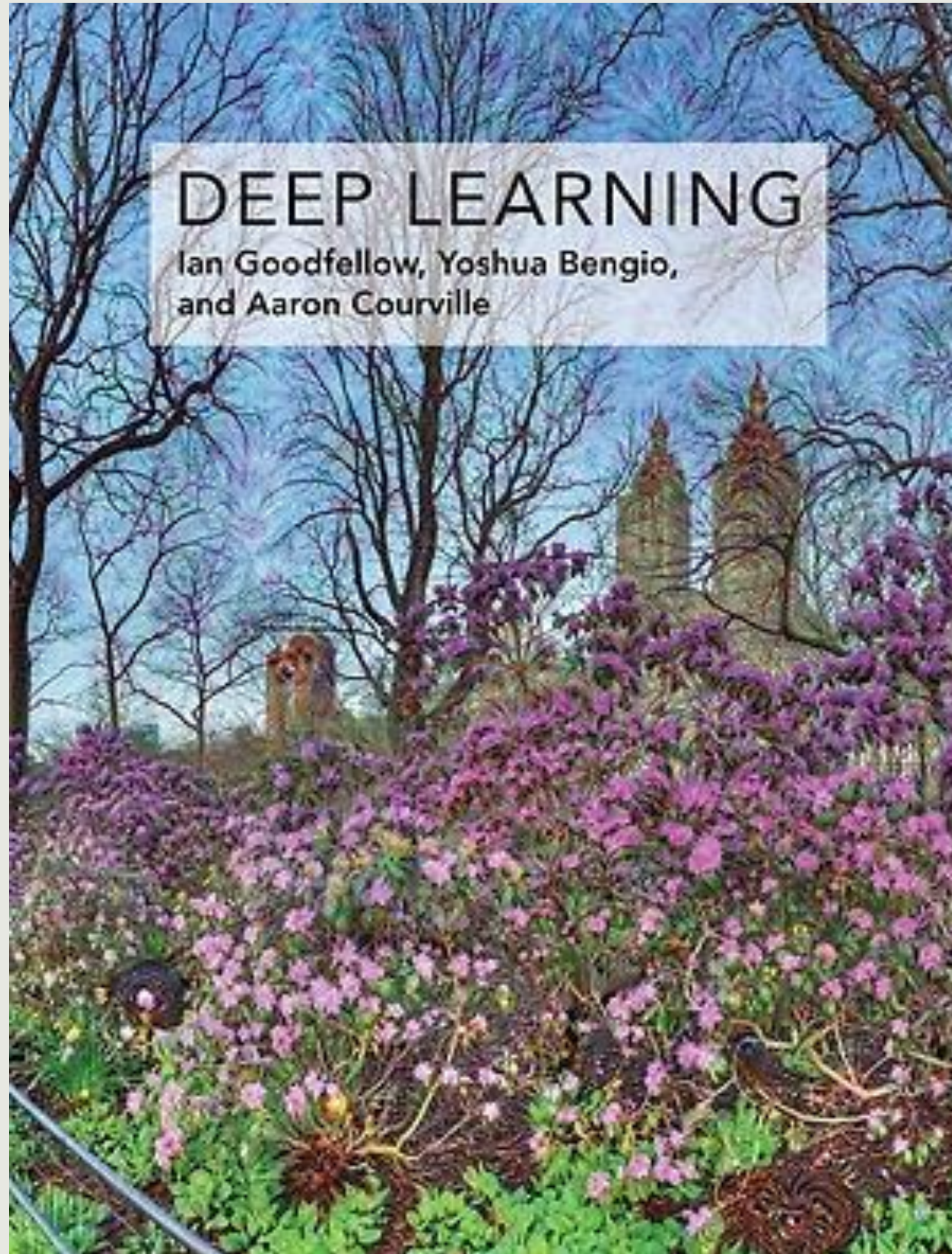


겨울방학 Session Introduction

2022-1 ESC

겨울방학 세션 교재



봄 학기 세션 Source

The Marginal Value of Adaptive Gradient Methods in Machine Learning

Ashia C. Wilson[#], Rebecca Roelofs[#], Mitchell Stern[#], Nathan Srebro[†], and Benjamin Recht[#]
{ashia,roelofs,mitchell}@berkeley.edu, nati@ttic.edu, brecht@berkeley.edu

[#]University of California, Berkeley

[†]Toyota Technological Institute at Chicago

How Does Batch Normalization Help Optimization?

Shibani Santurkar*
MIT
shibani@mit.edu

Dimitris Tsipras*
MIT
tsipras@mit.edu

Andrew Ilyas*
MIT
ailyas@mit.edu

Aleksander Madry
MIT
madry@mit.edu

Abstract

Batch Normalization (BatchNorm) is a widely adopted technique that enables faster and more stable training of deep neural networks (DNNs). Despite its pervasiveness, the exact reasons for BatchNorm's effectiveness are still poorly understood. The popular belief is that this effectiveness stems from controlling the change of the layers' input distributions during training to reduce the so-called

딥러닝 구현
라이브러리

 PyTorch

Objective

겨울방학

- 딥러닝 기본 내용
- Pytorch 연습

봄학기

- Architecture 학습
- Project 위주

과제 형식

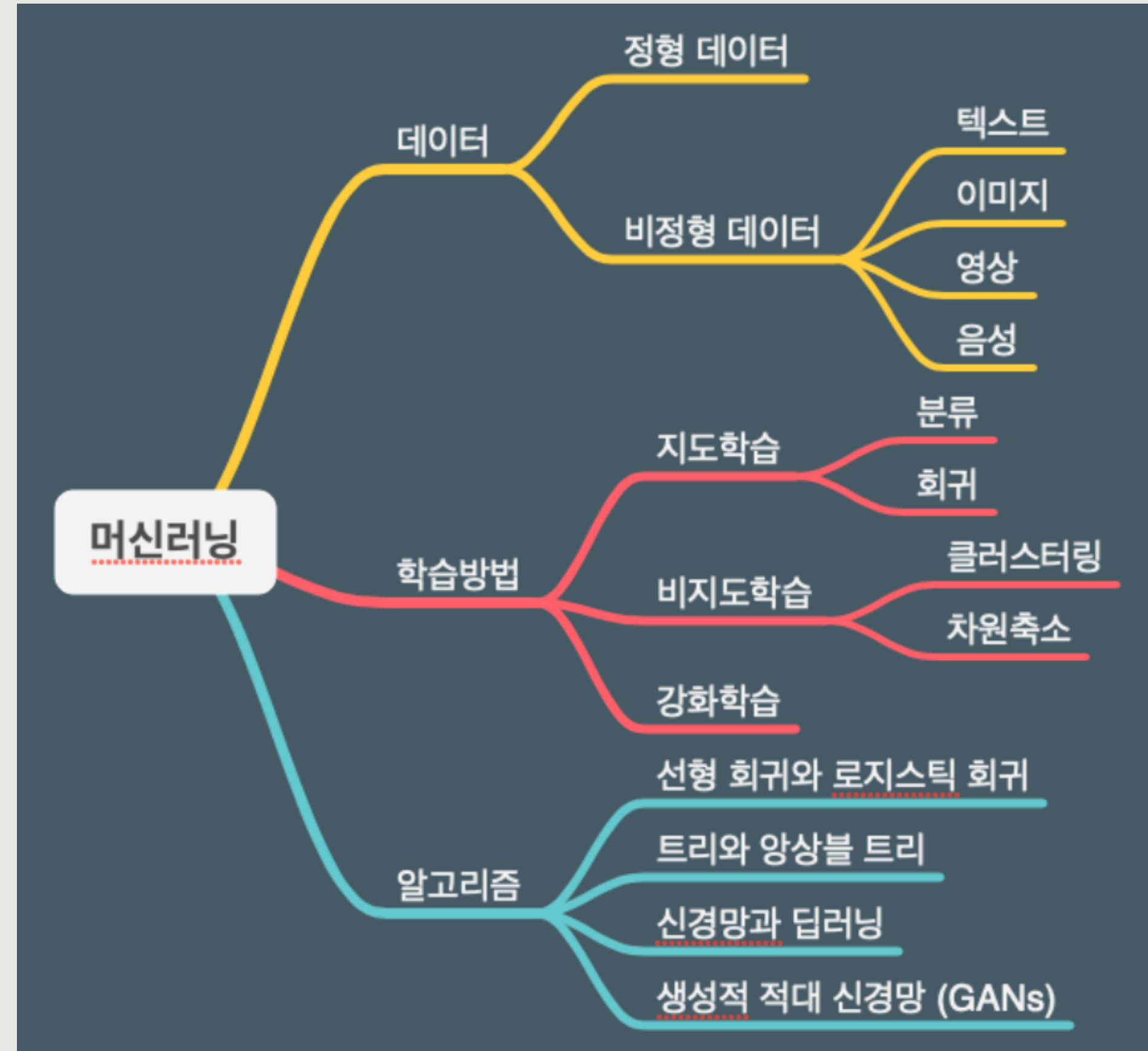
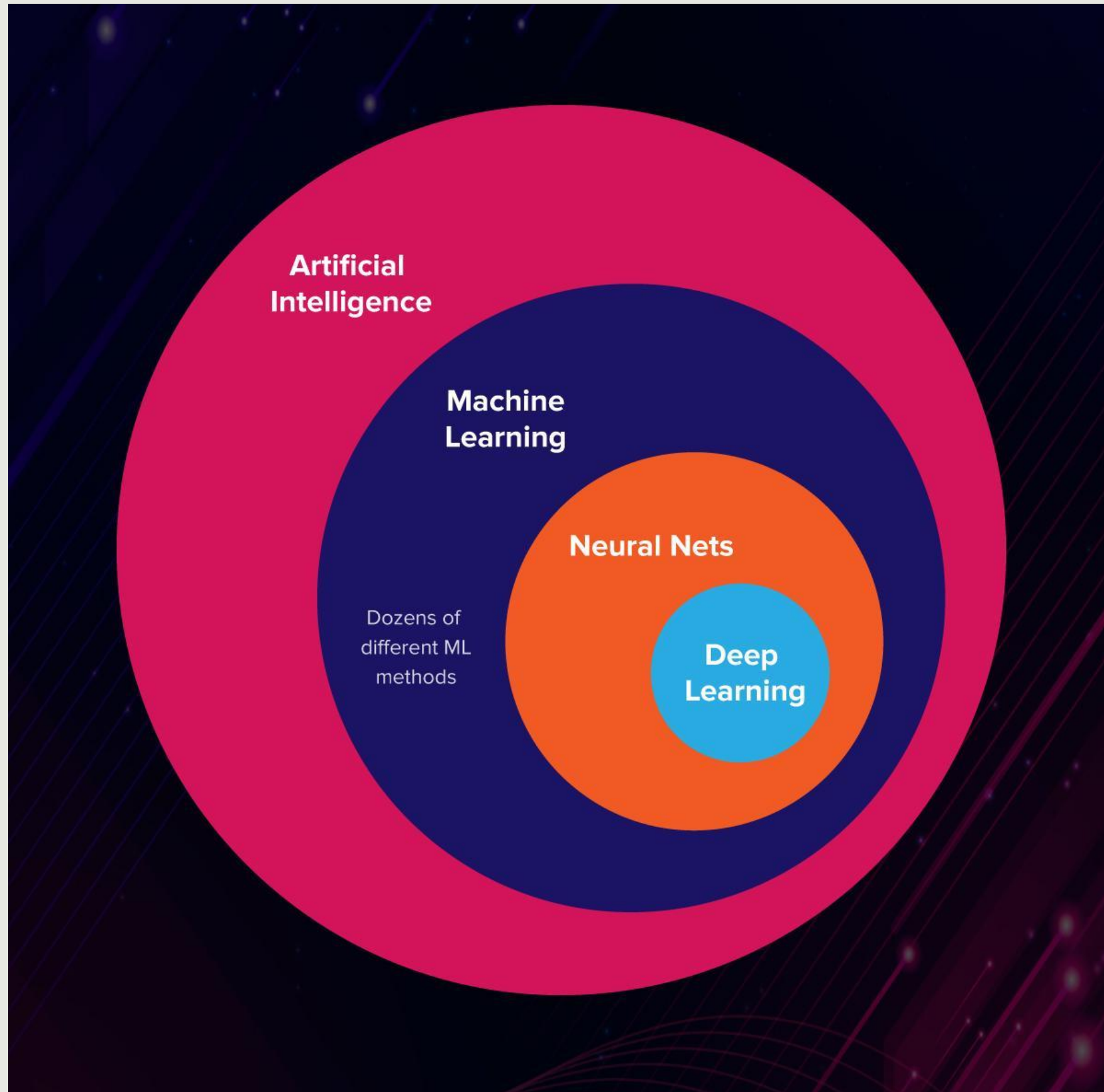
Week 2 & 3

- 자료 조사 & 정리
- 다음 세션 시작 전,
랜덤 1명 추천
: 세션 시작 전 발표

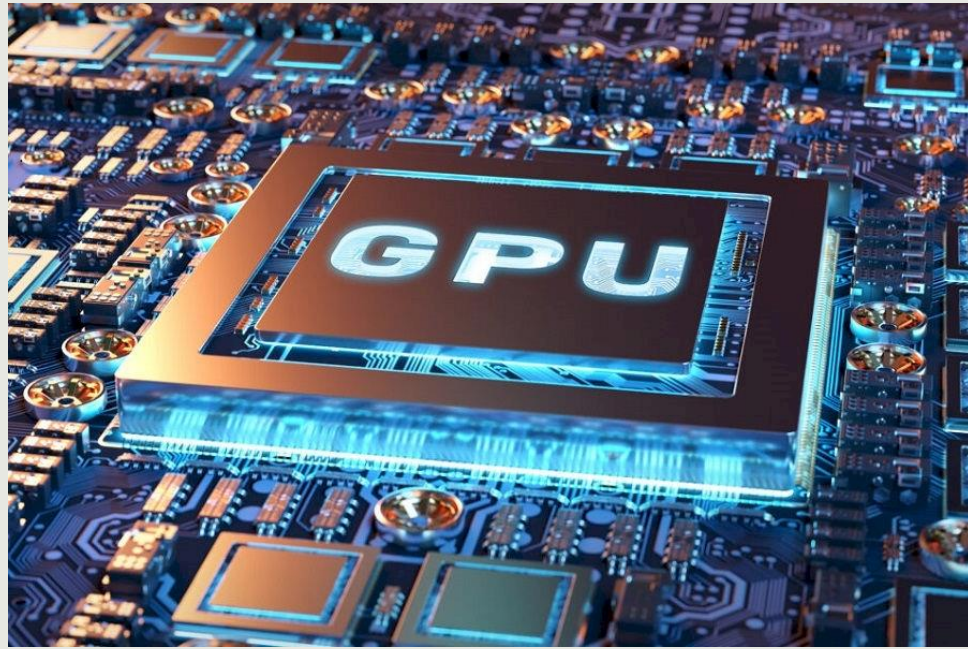
Week 4 & 6

- 실습 데이터 제공 : 학술부에서 제공
- 복습 스터디 조 단위의 실습
- 다음 세션 시작 전, 랜덤 1조 추천
: 세션 시작 전 결과물 발표
- 실습 형태는 자유
(ex) 구성원 간 협동 : 모델의 각 파트를 코딩 & 병합
각자 모델 전체를 직접 구현 : 좋은 성능의 코드를
뽑아서 발표

Deep Learning



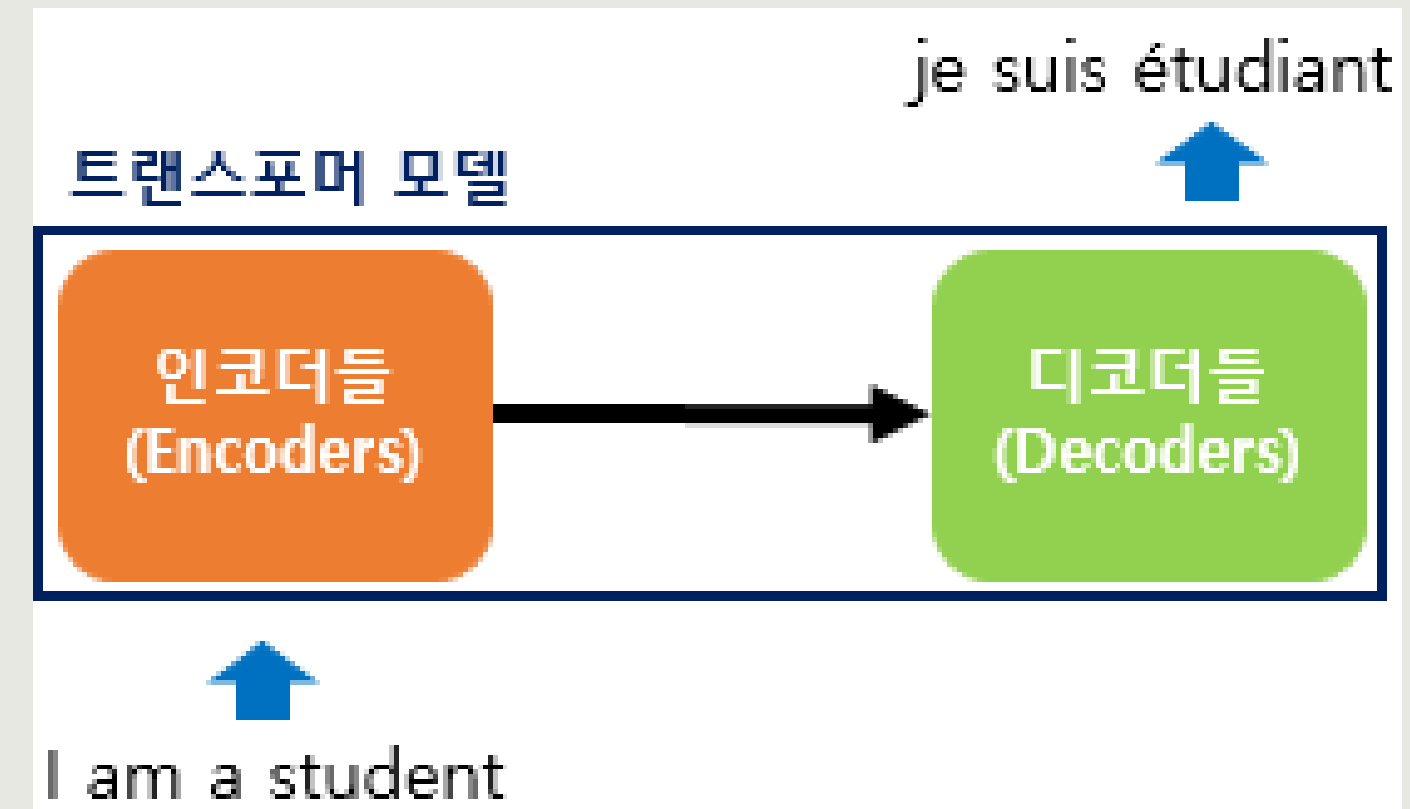
Deep Learning



Computing 기술 발전



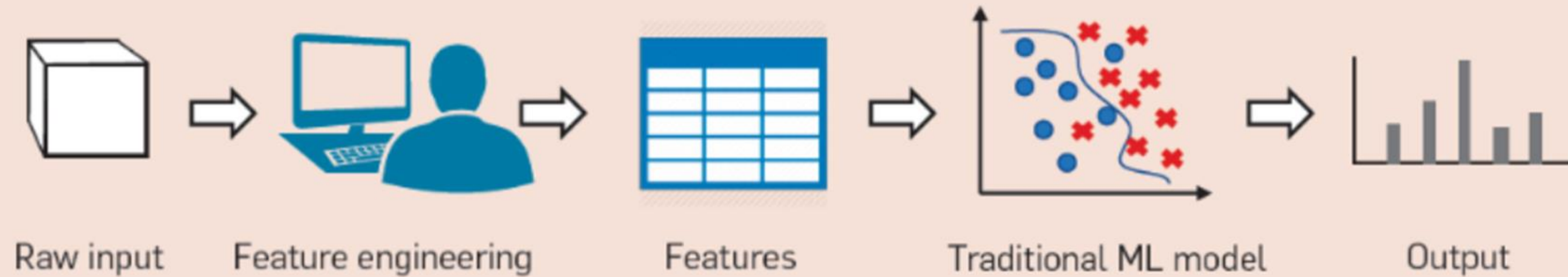
빅데이터



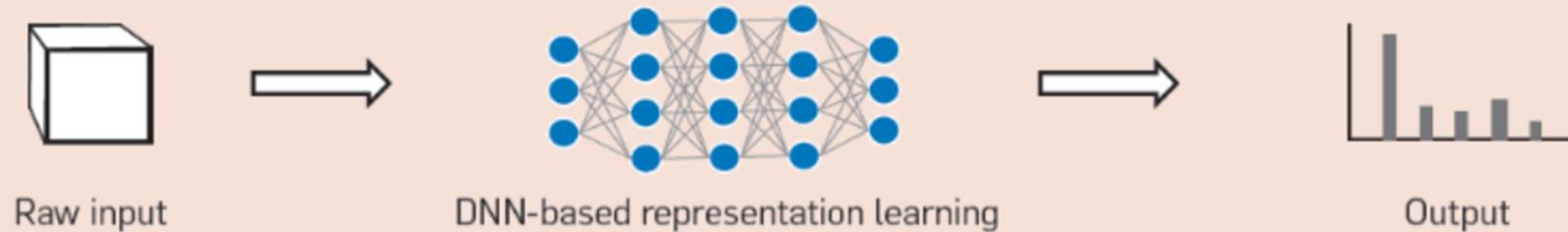
알고리즘 향상

General Machine Learning **vs** Deep Learning

Traditional machine learning

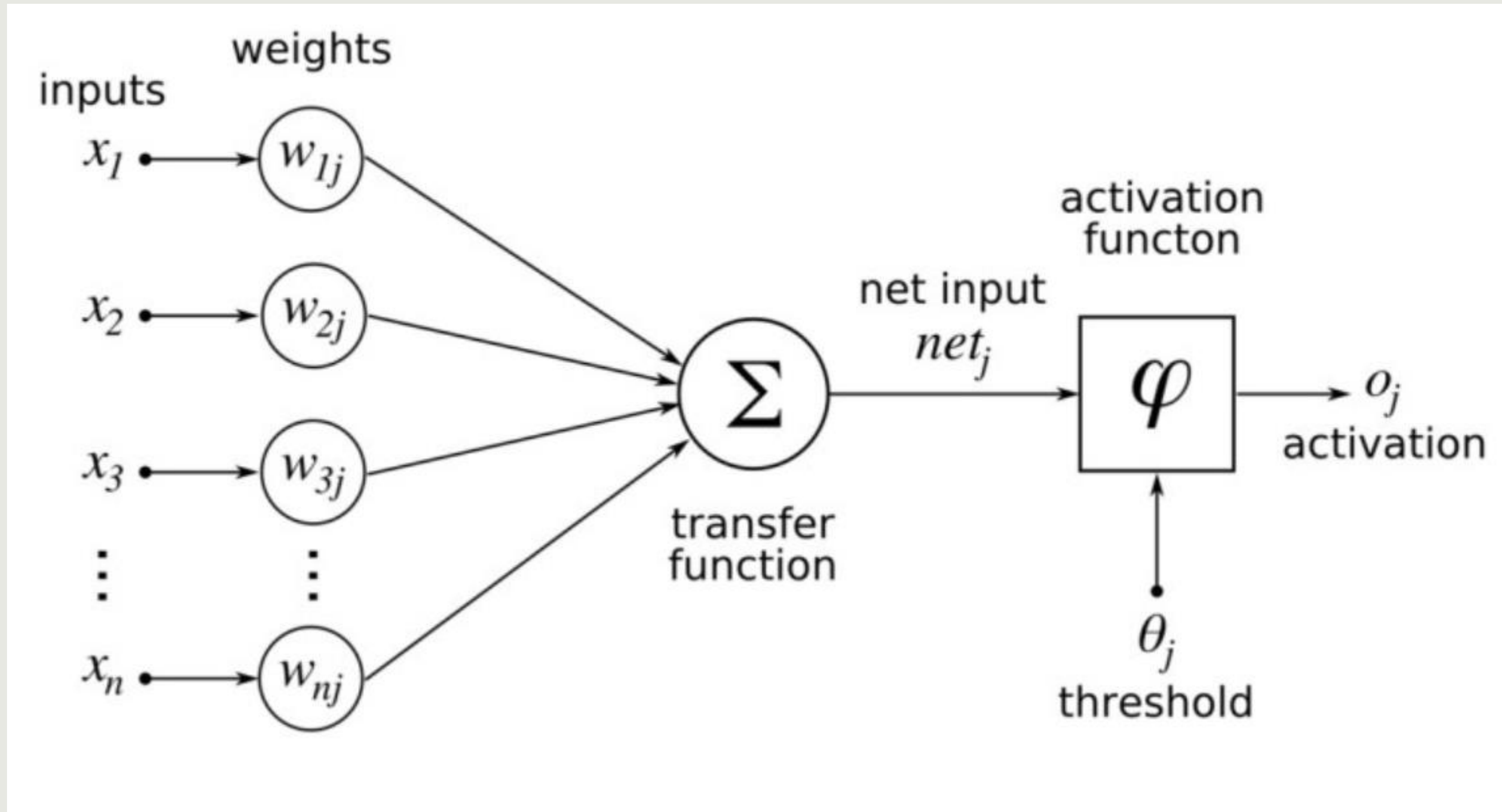


Deep learning



- Representation Learning
- End-to-end

Deep Learning



Deep Learning

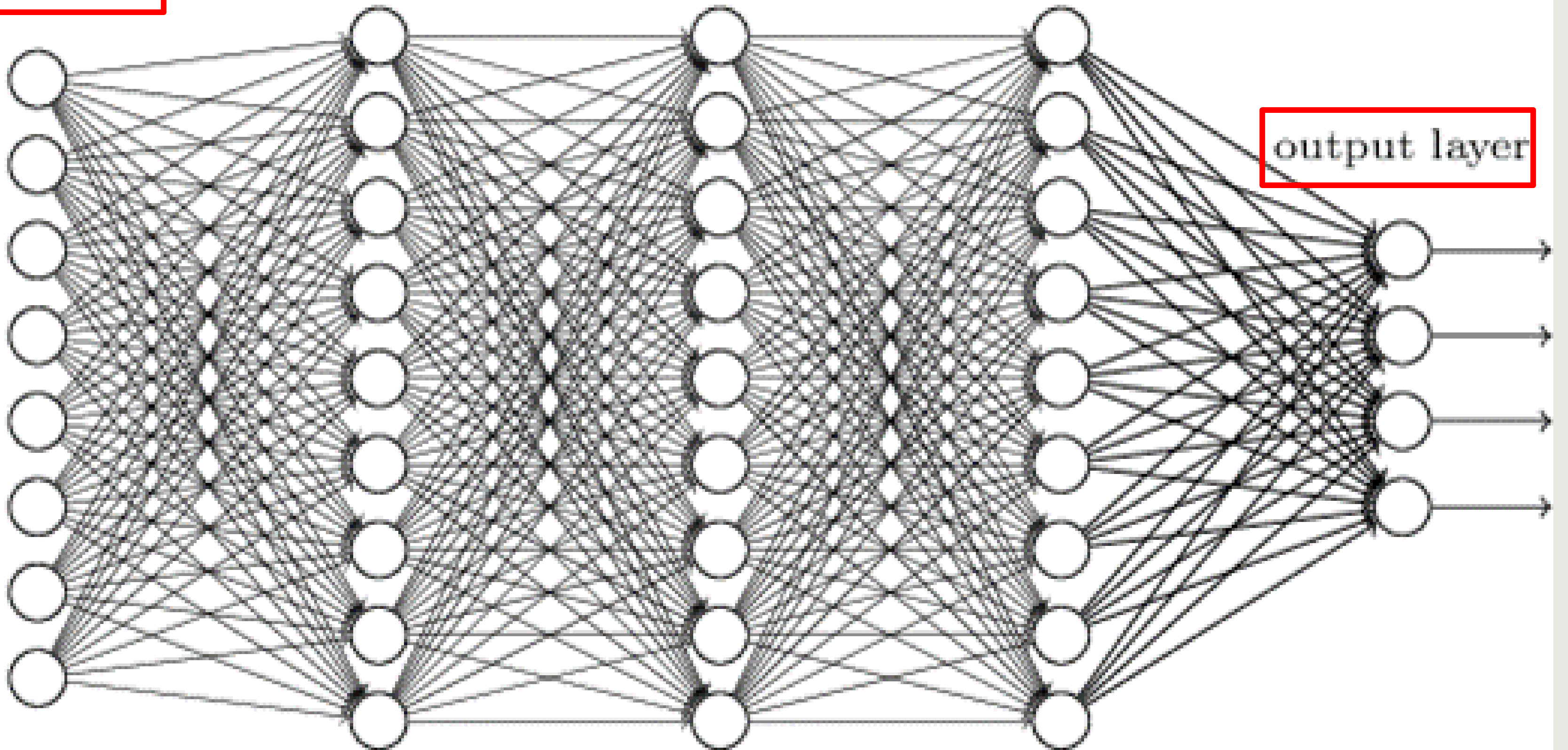
input layer

hidden layer 1

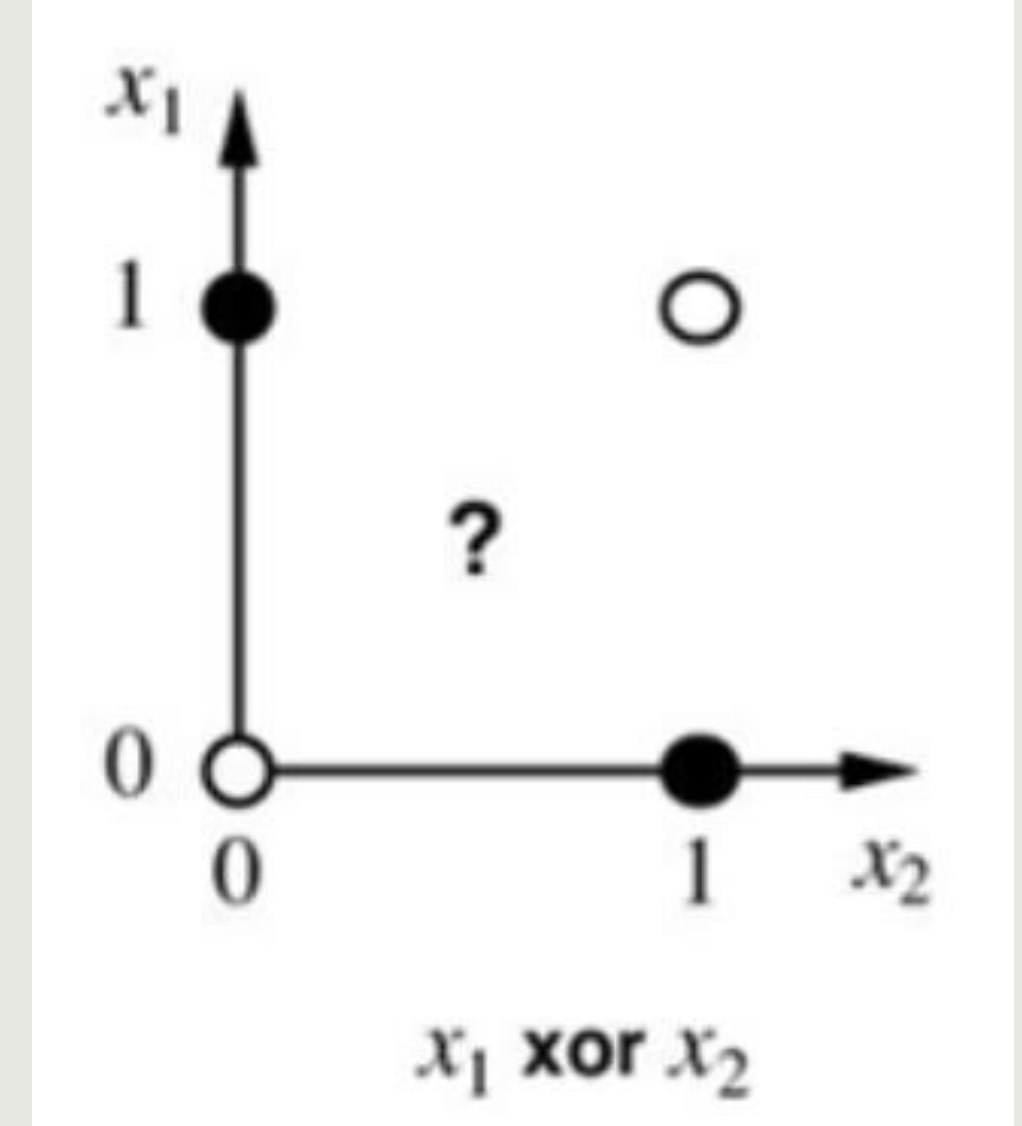
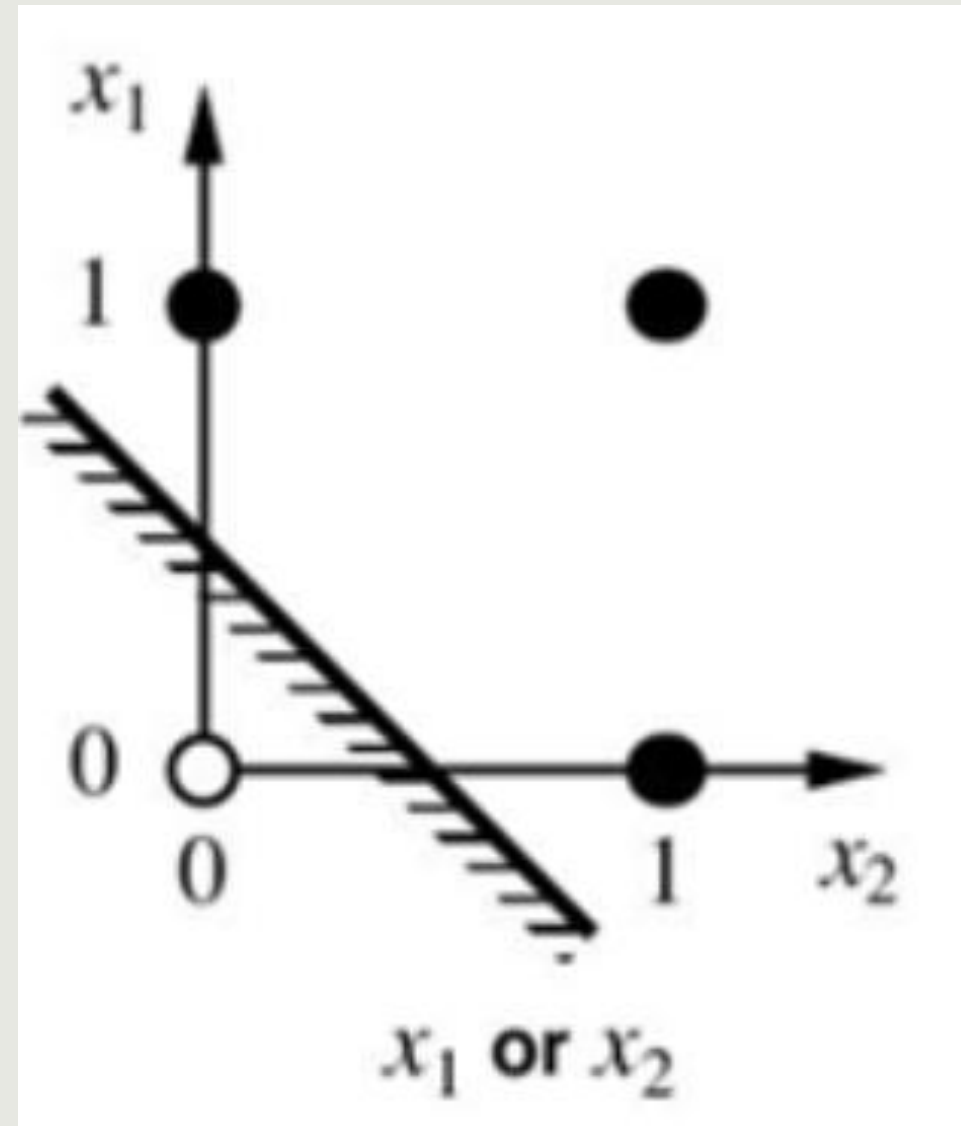
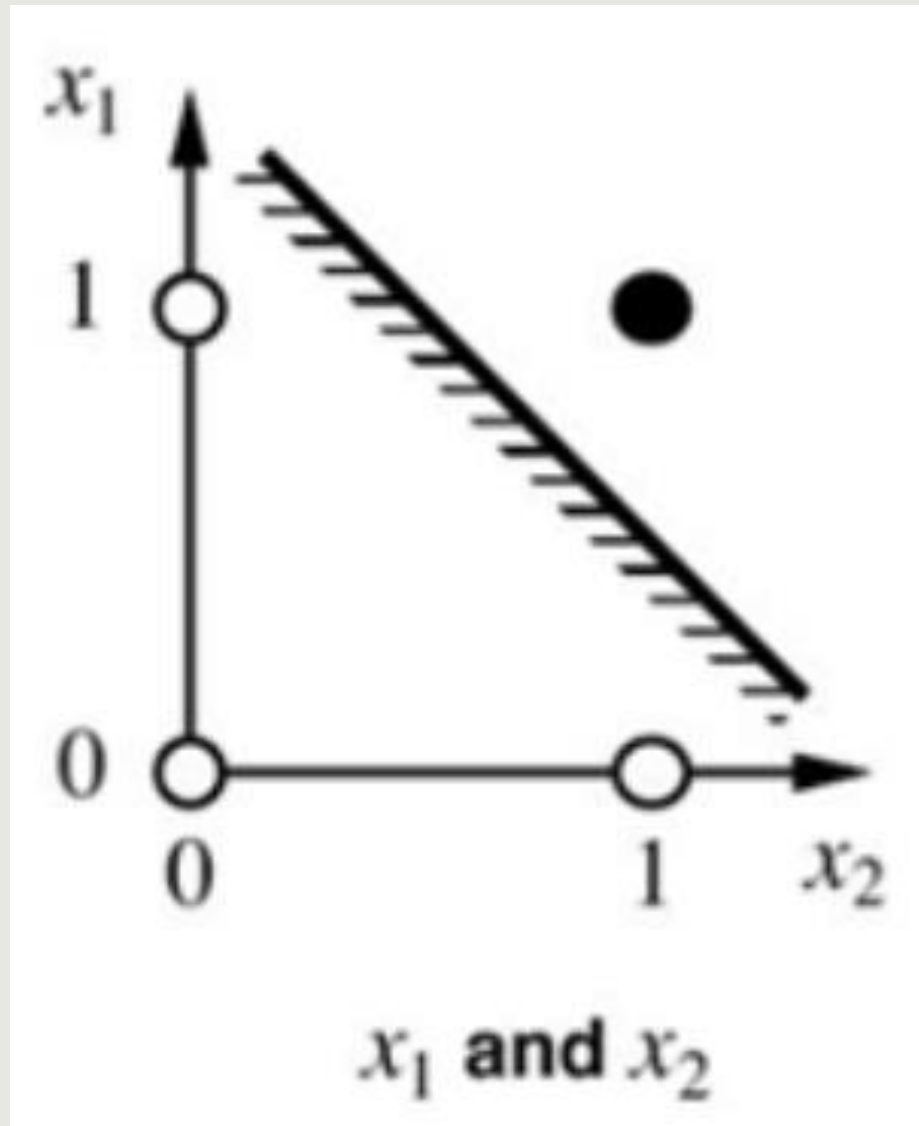
hidden layer 2

hidden layer 3

output layer

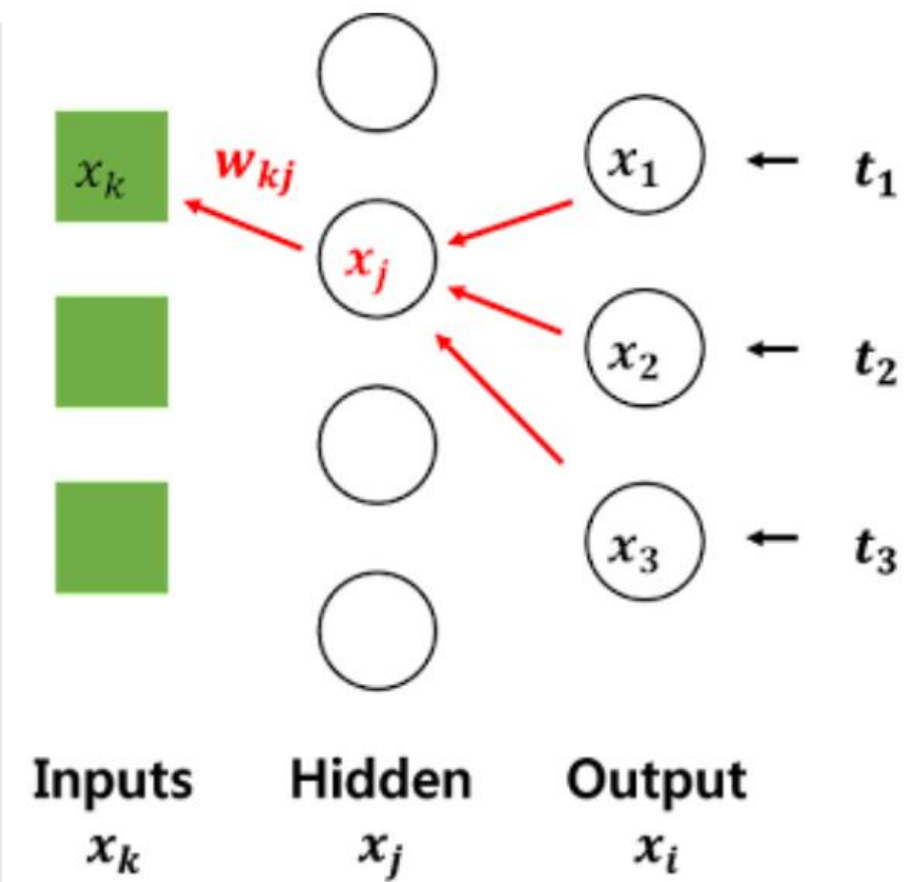
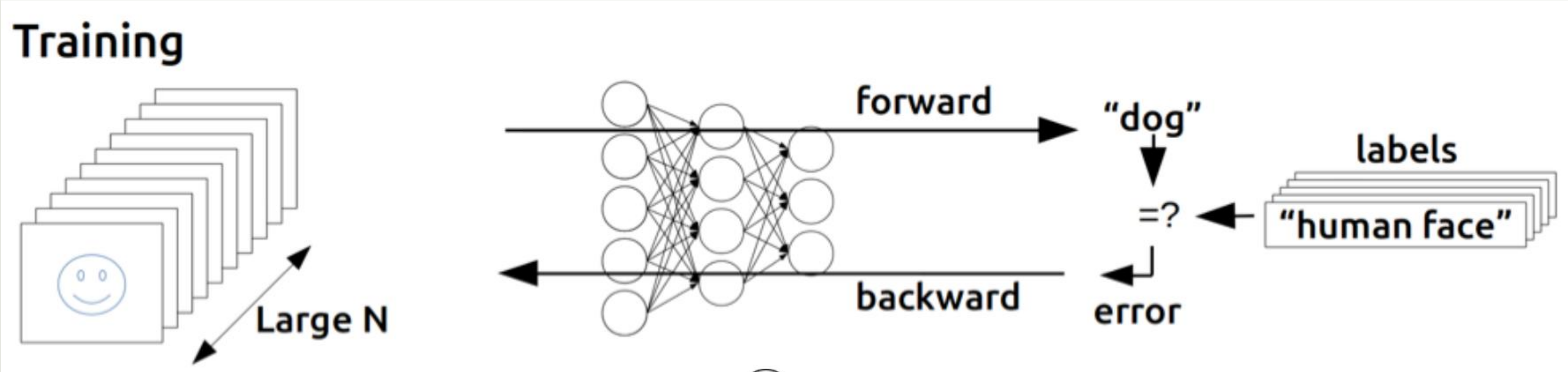


Example : XOR Problem

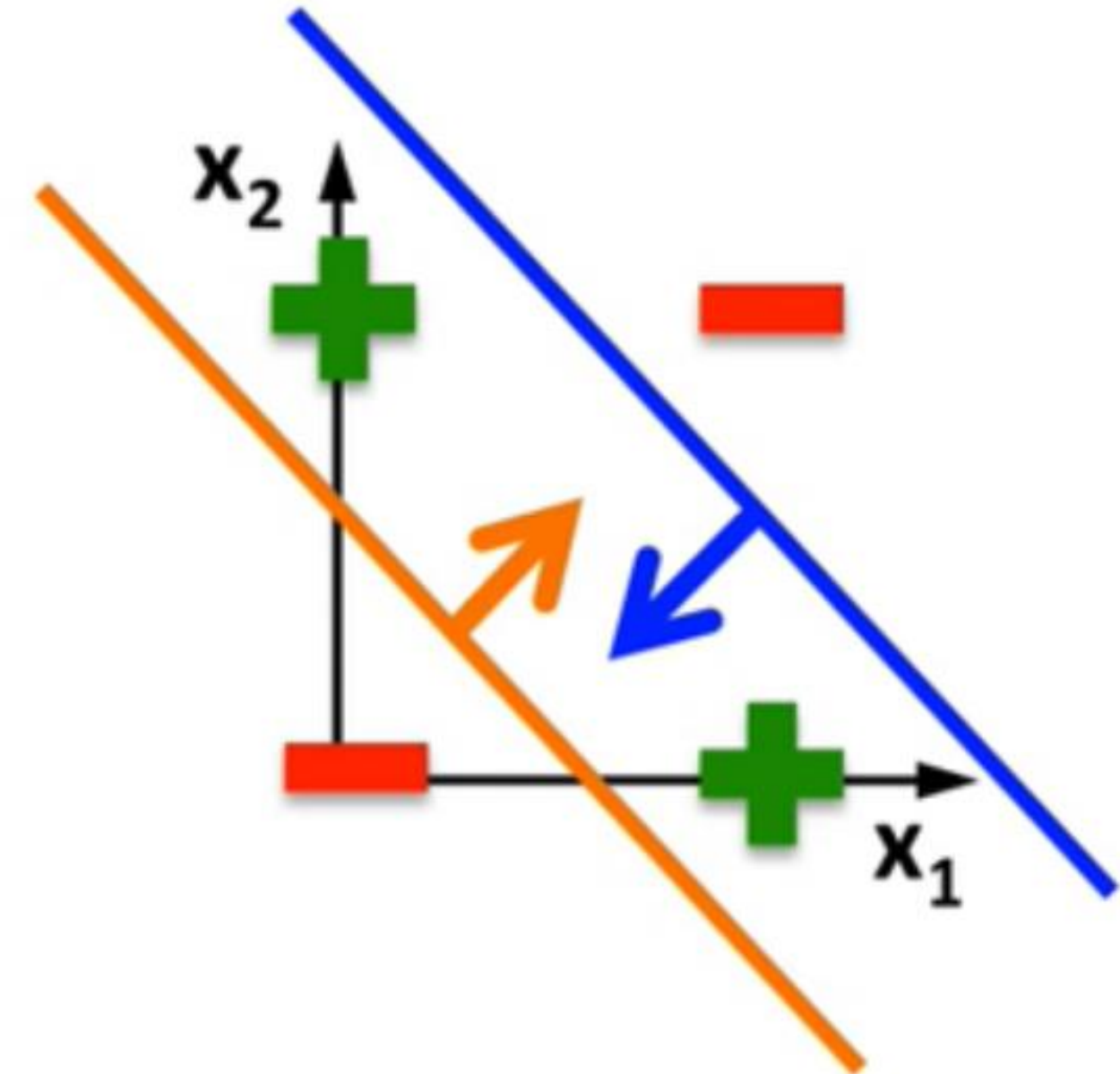
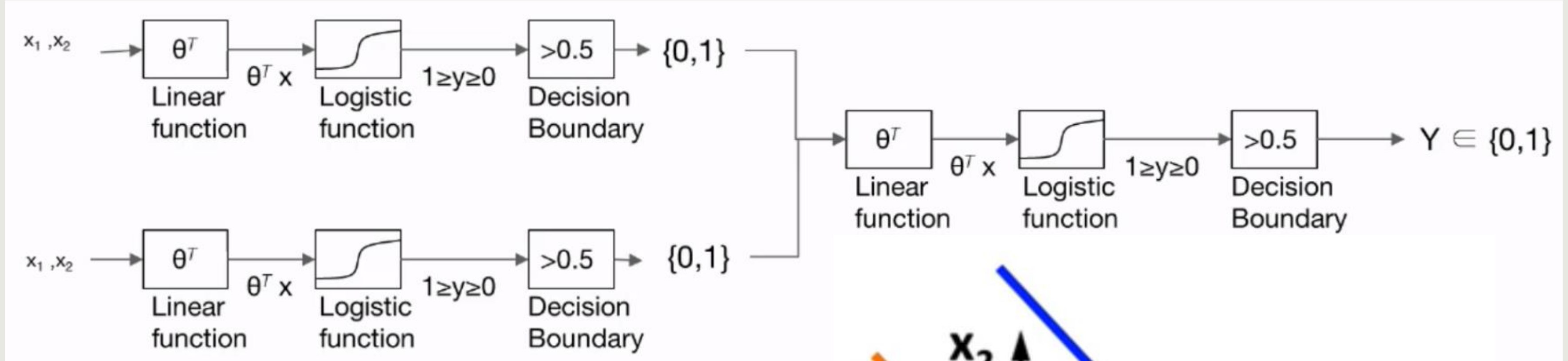


Example : XOR Problem

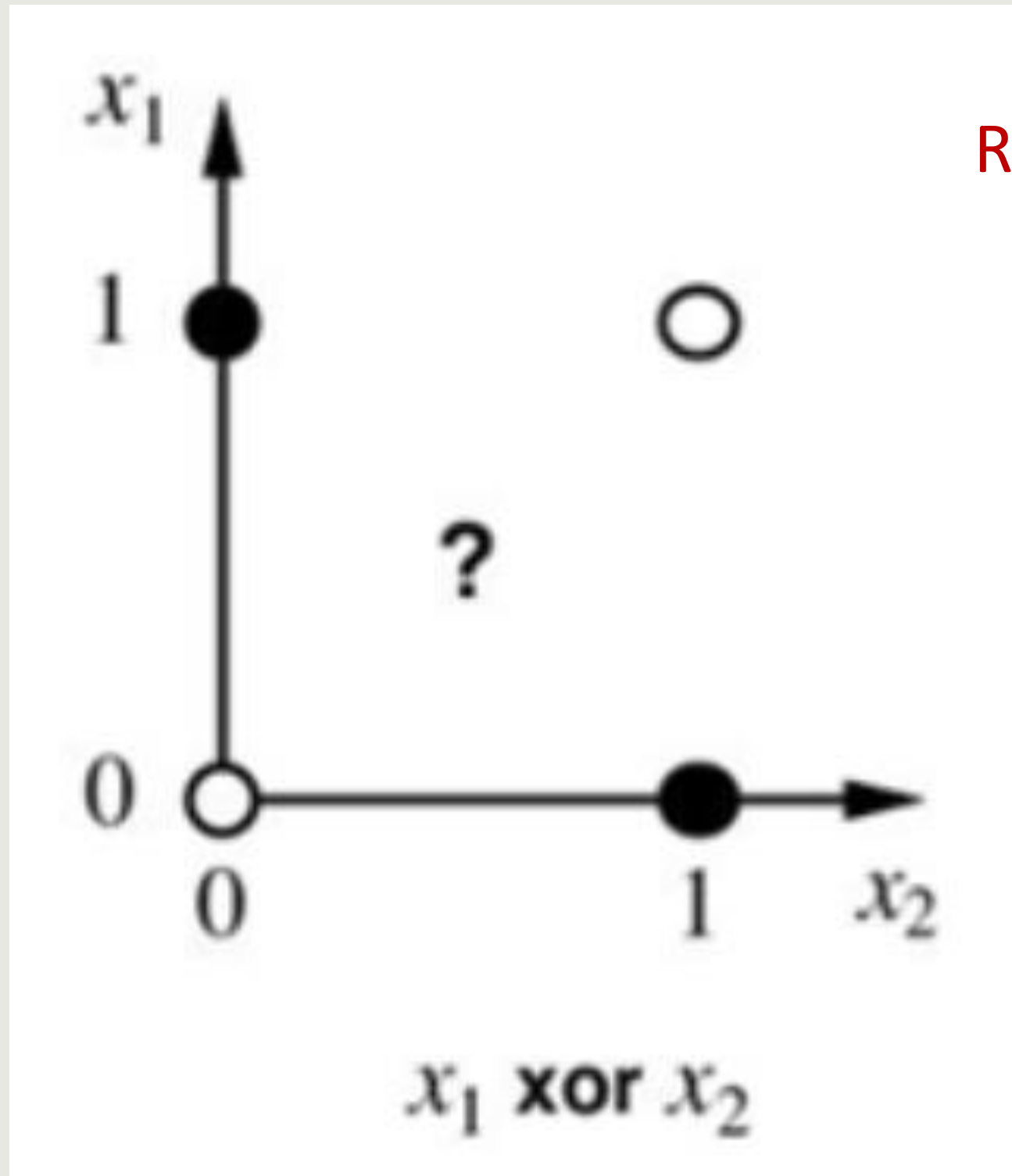
BackPropagation (역전파)



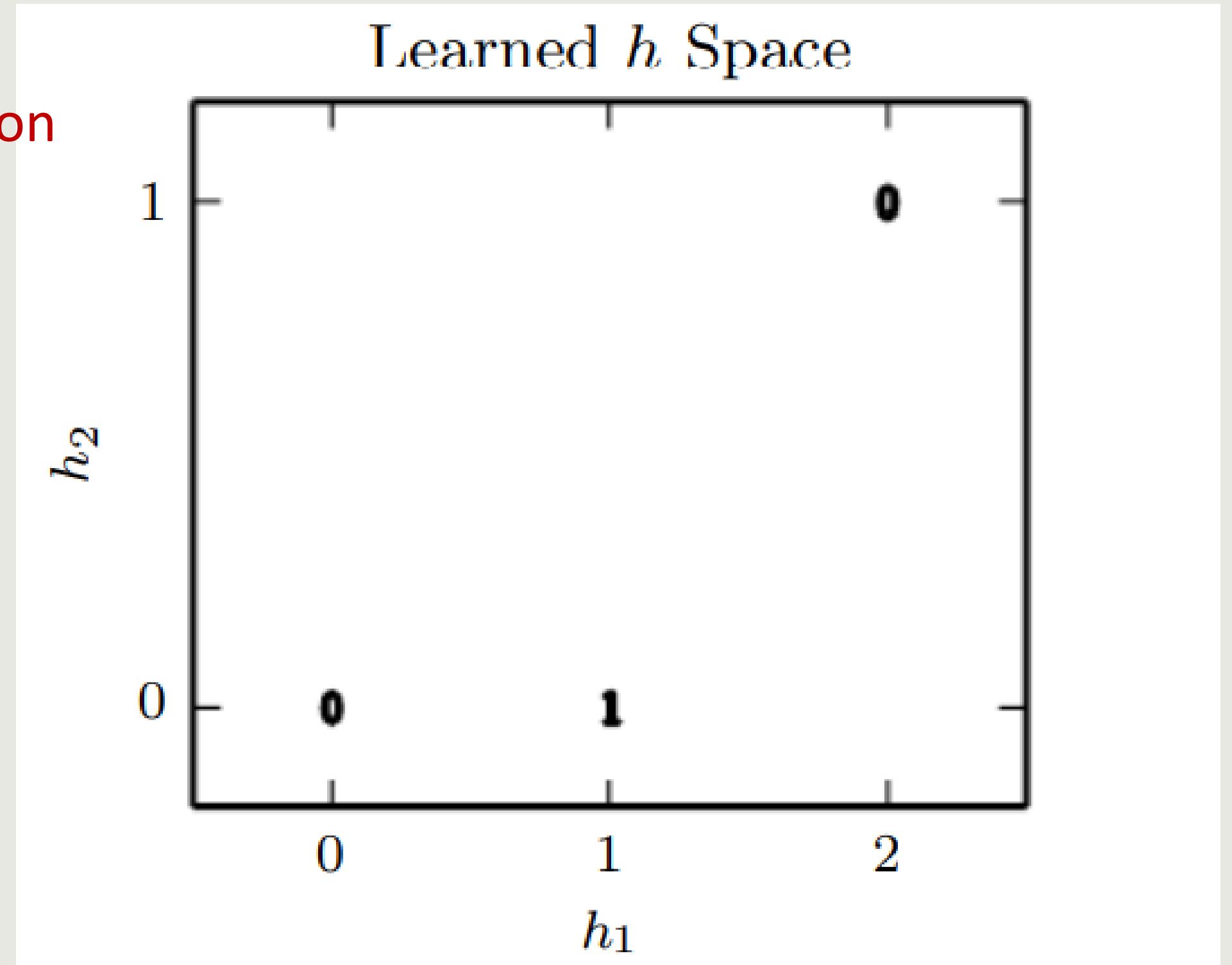
Example : XOR Problem



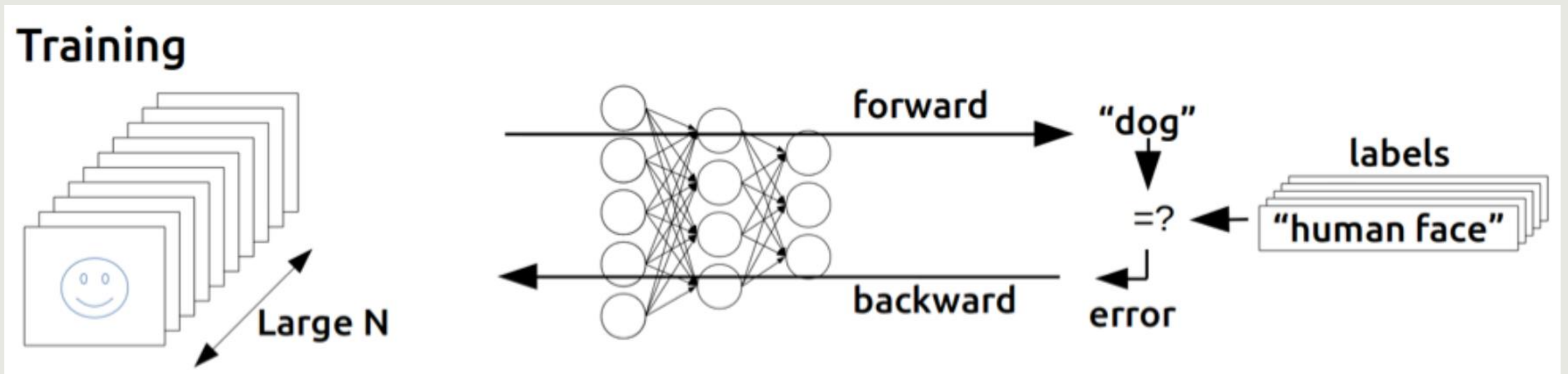
Example : XOR Problem



Representation
Learning



Optimization

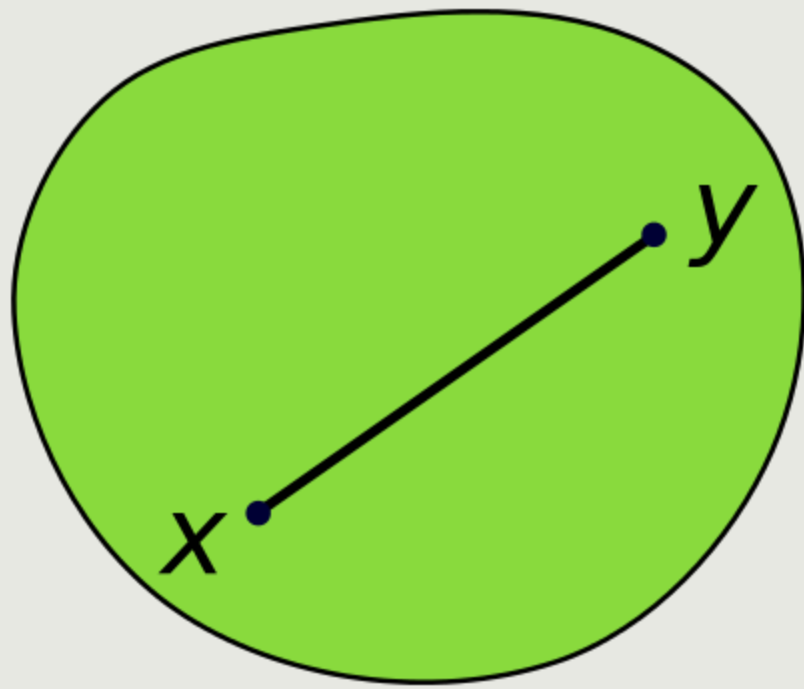


Loss Function => Optimization

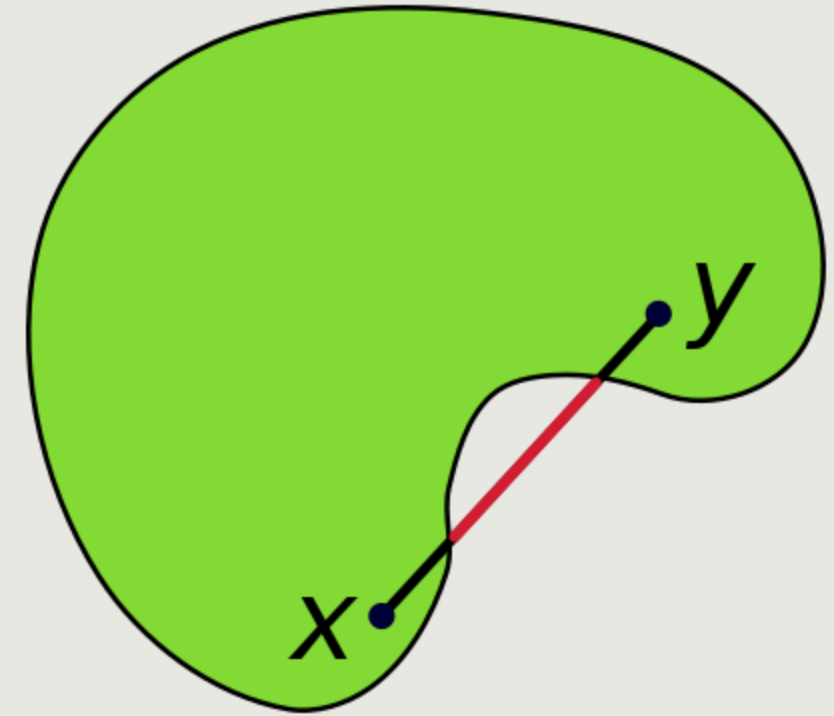
Convex

Non-Convex

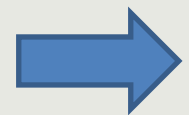
Convex Function & Convex Set



Convex Set



Non-Convex Set



A function f is a convex function if and only if an epigraph of a function f is a convex set.



Convex Optimization & Non-Convexity

- Convexity : locally minimal point \Rightarrow global minimal point
- Strict convexity \Rightarrow **unique** global minimal point

But, Deep Learning의 대부분의 Loss function은 Non-convex

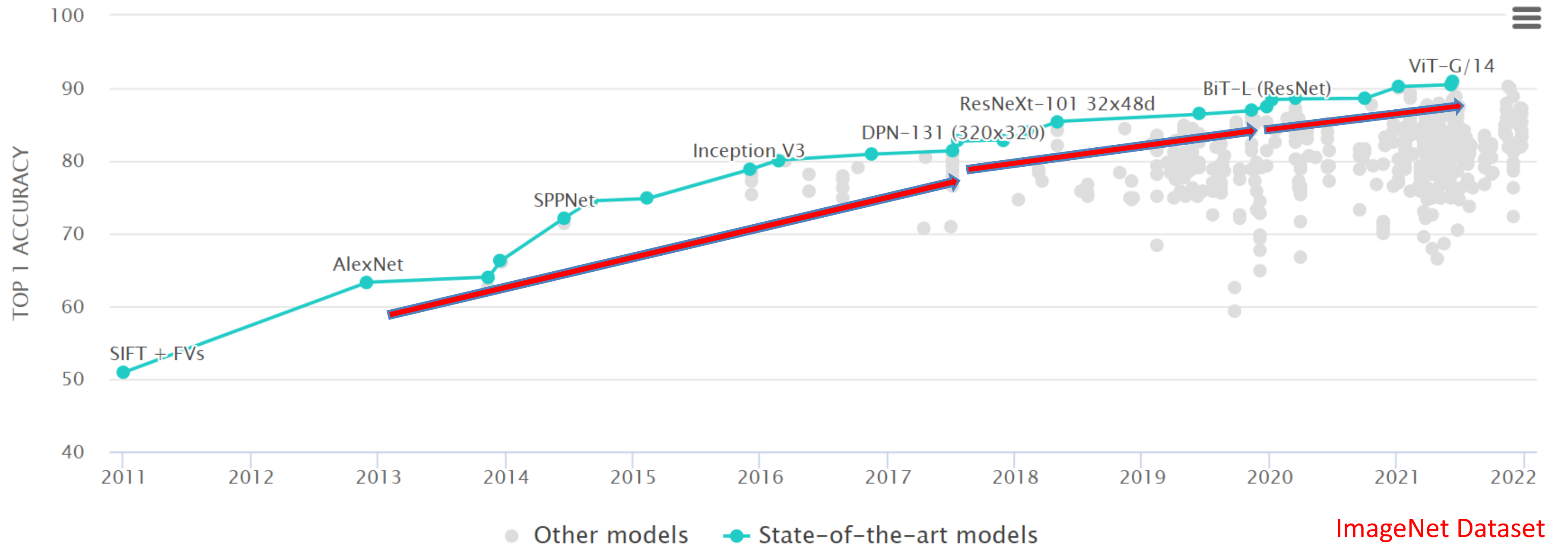
\Rightarrow Convex Optimization : Convex Relaxation

\Rightarrow Deep Learning : 다양한 Optimization 기법 발달

(고차원에서 local minima가 거의 없을 수도 있다는 논문도 존재)

Deep Learning 발전 양상

<https://paperswithcode.com/sota>



Deep Learning Architecture - CNN

Input Image
 5×5

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Filter (Kernel)
 3×3

1	0	1
0	1	0
1	0	1

Convolution:

4	3	4
2	4	3
2	3	4

Result
 3×3

The value **3** is the inner product of the patch

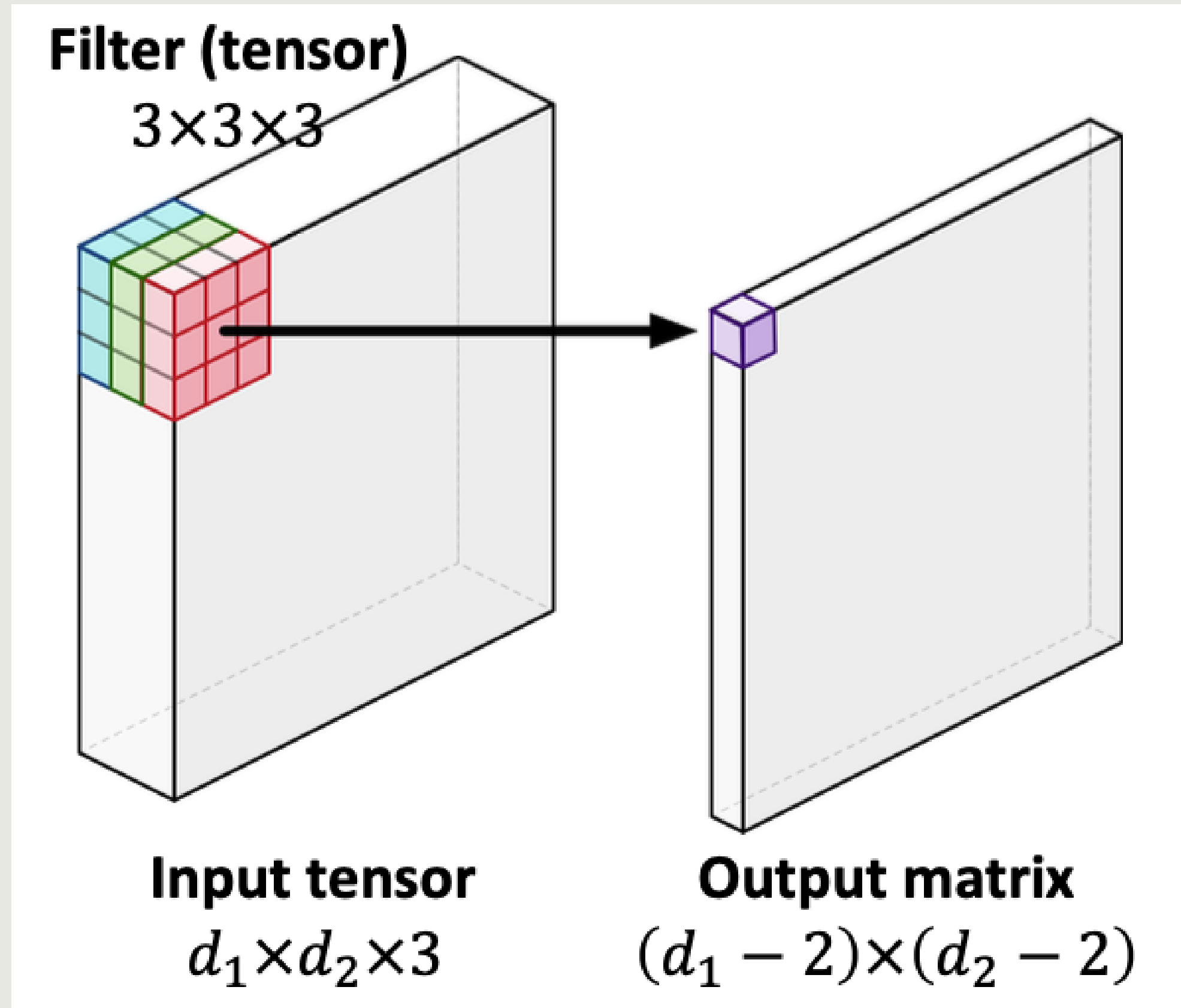
1	1	0
1	1	1
0	1	1

and the filter

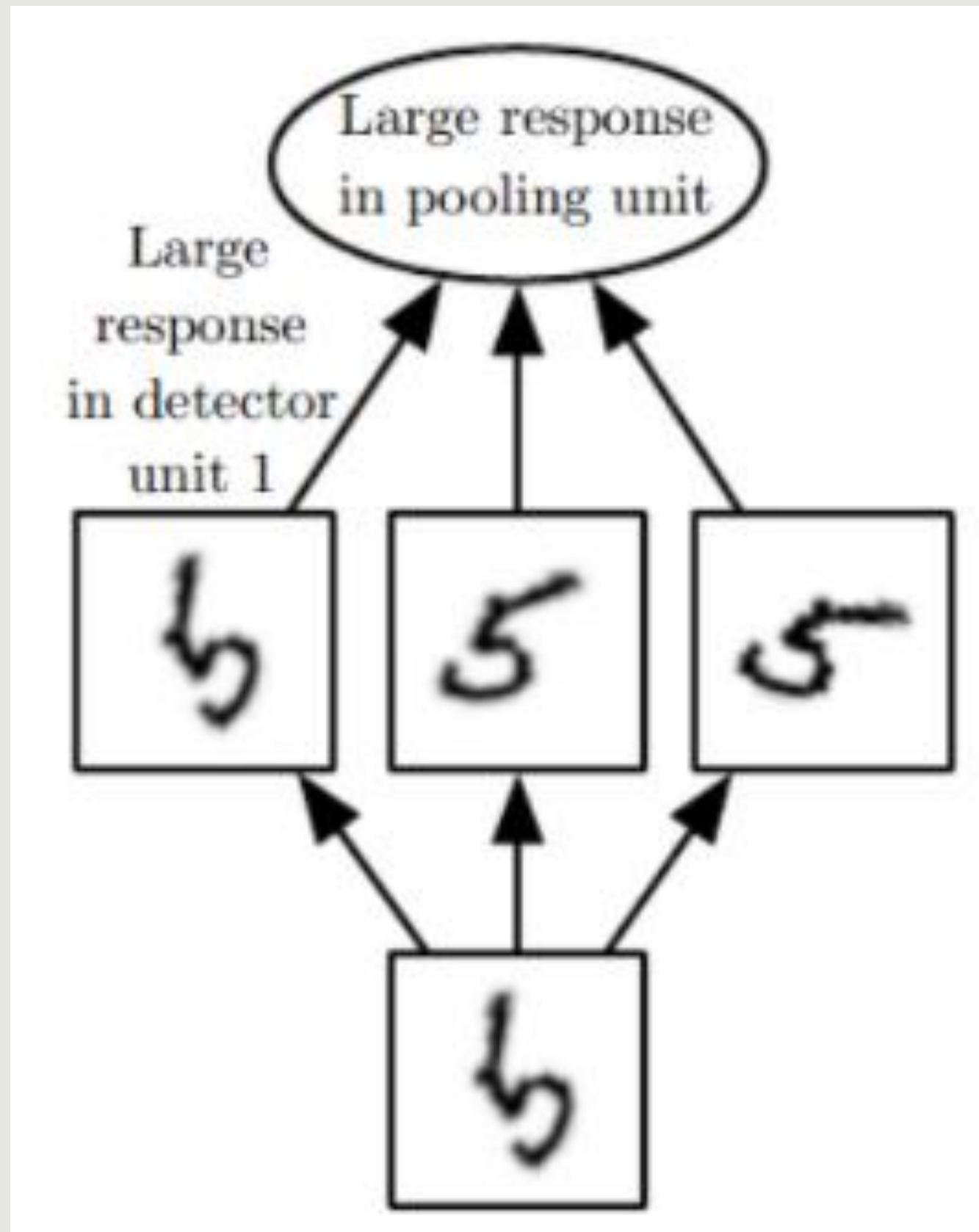
1	0	1
0	1	0
1	0	1

.

Deep Learning Architecture - CNN



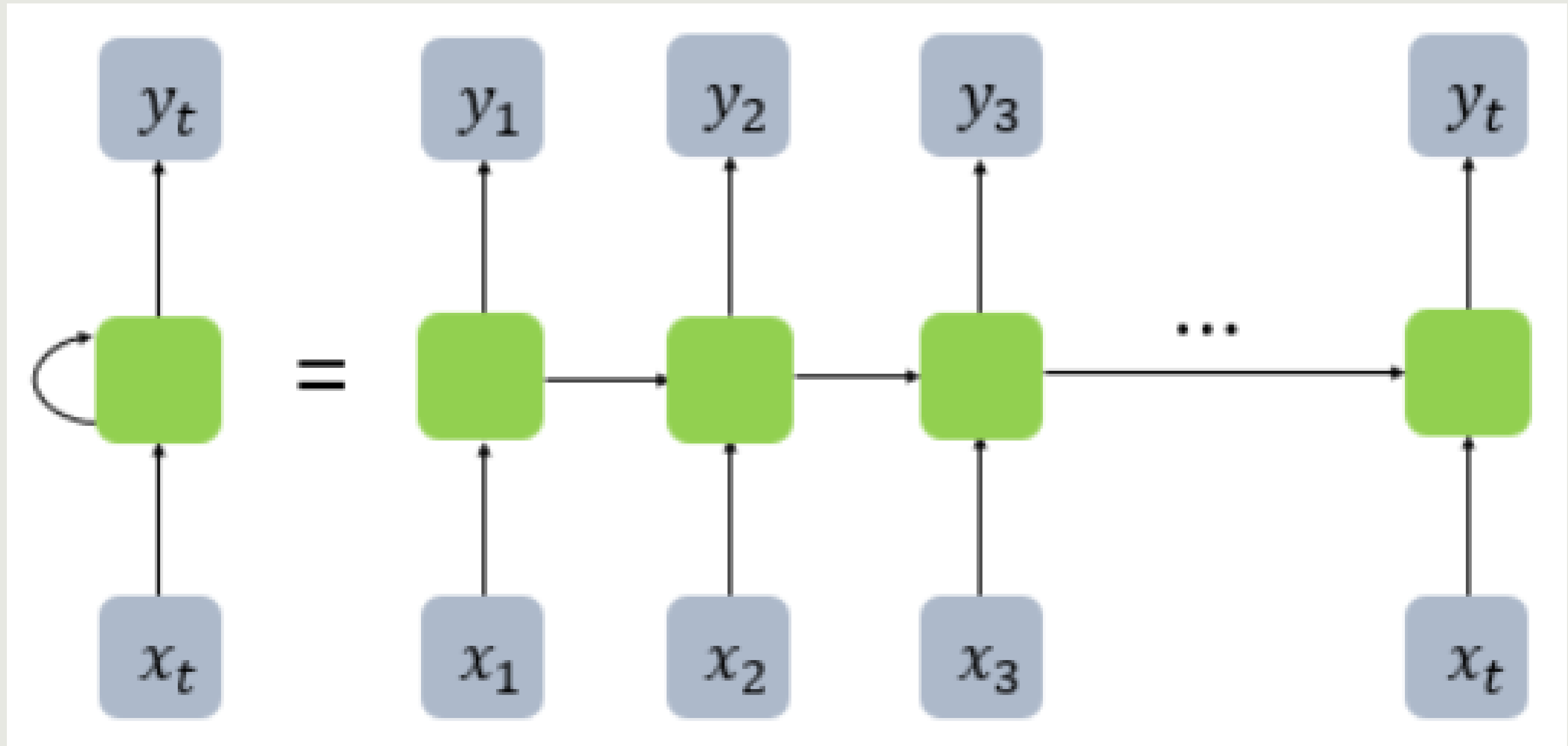
Deep Learning Architecture - CNN



- Scale Invariance

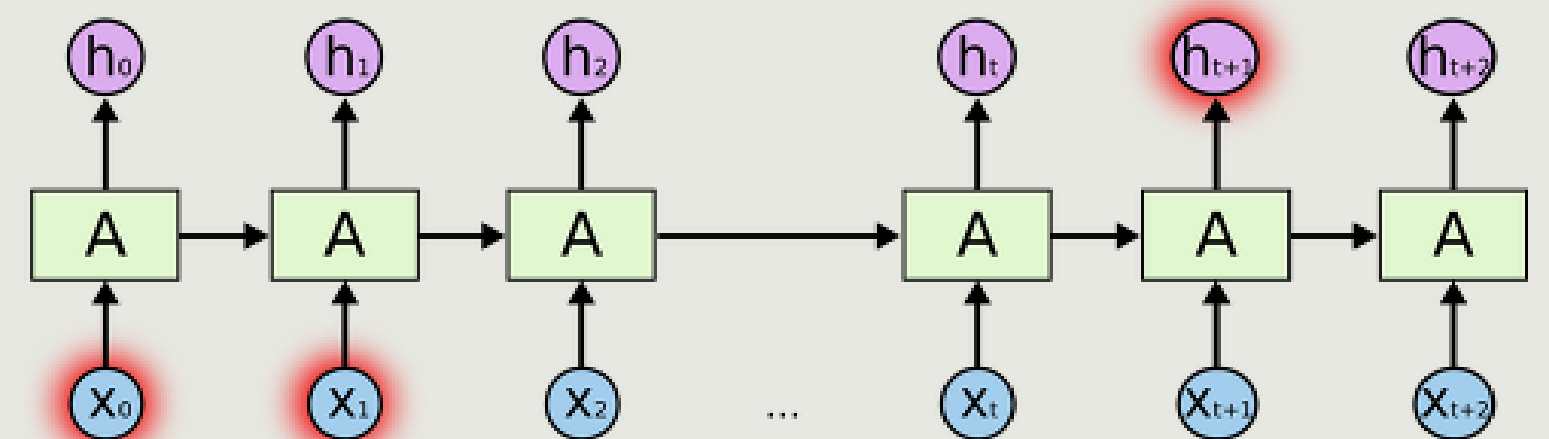
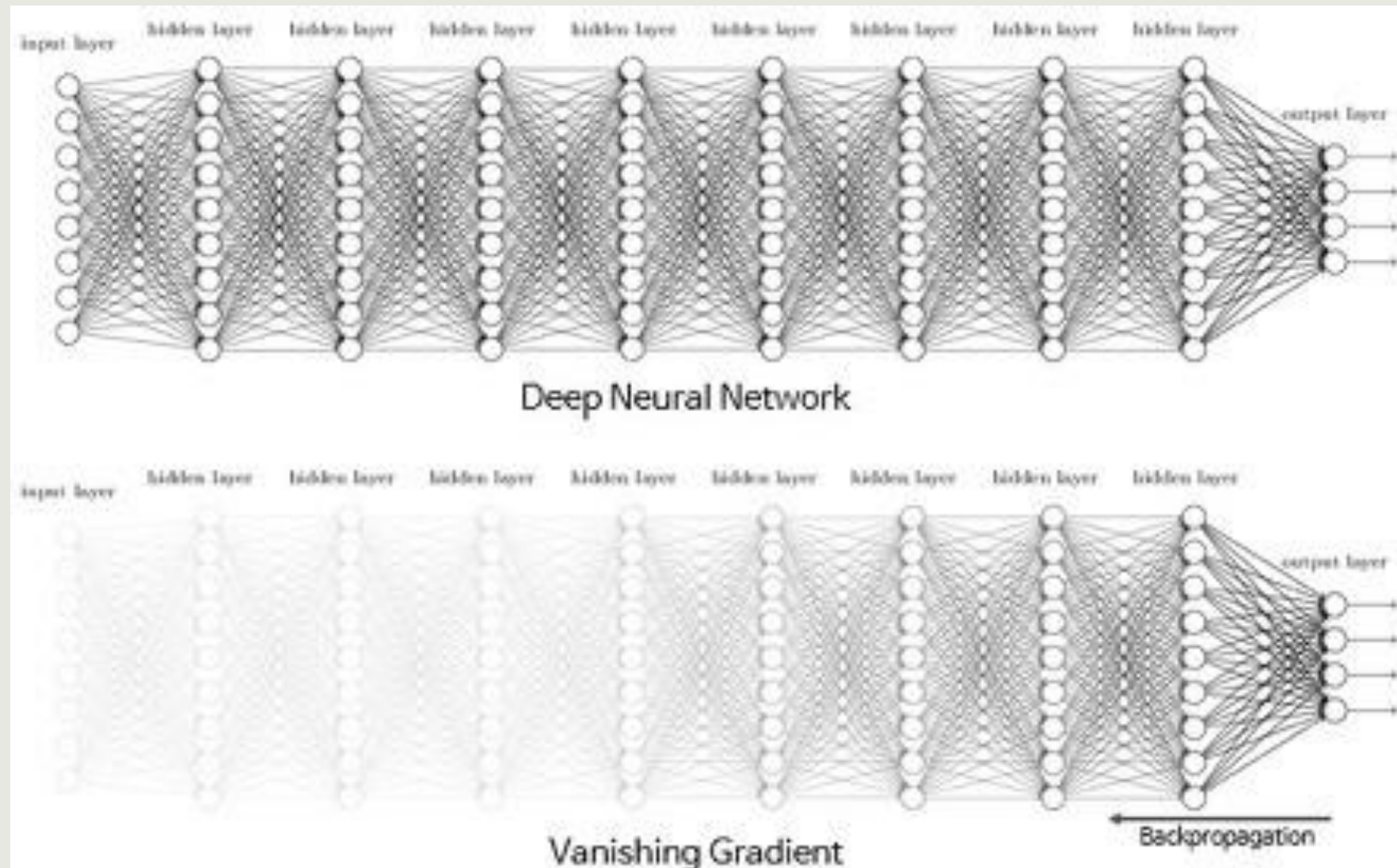
Max Pooling

Deep Learning Architecture - RNN

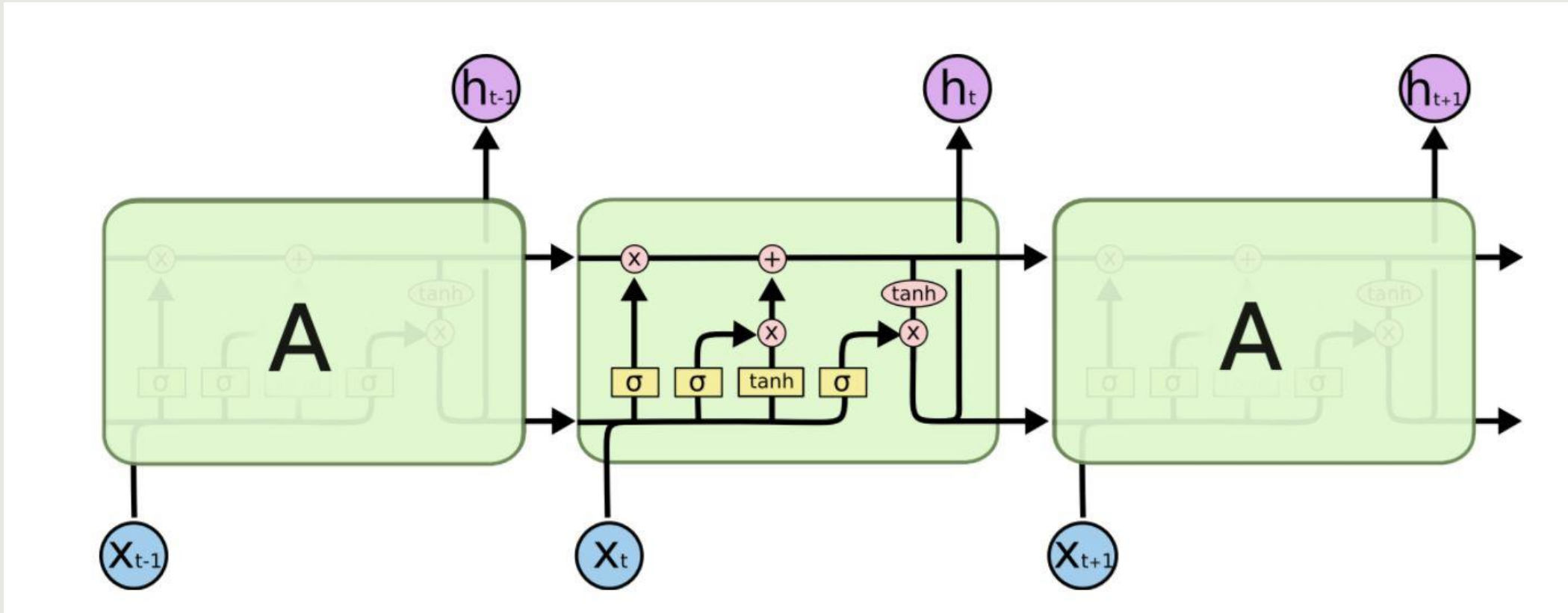


Deep Learning Architecture - RNN

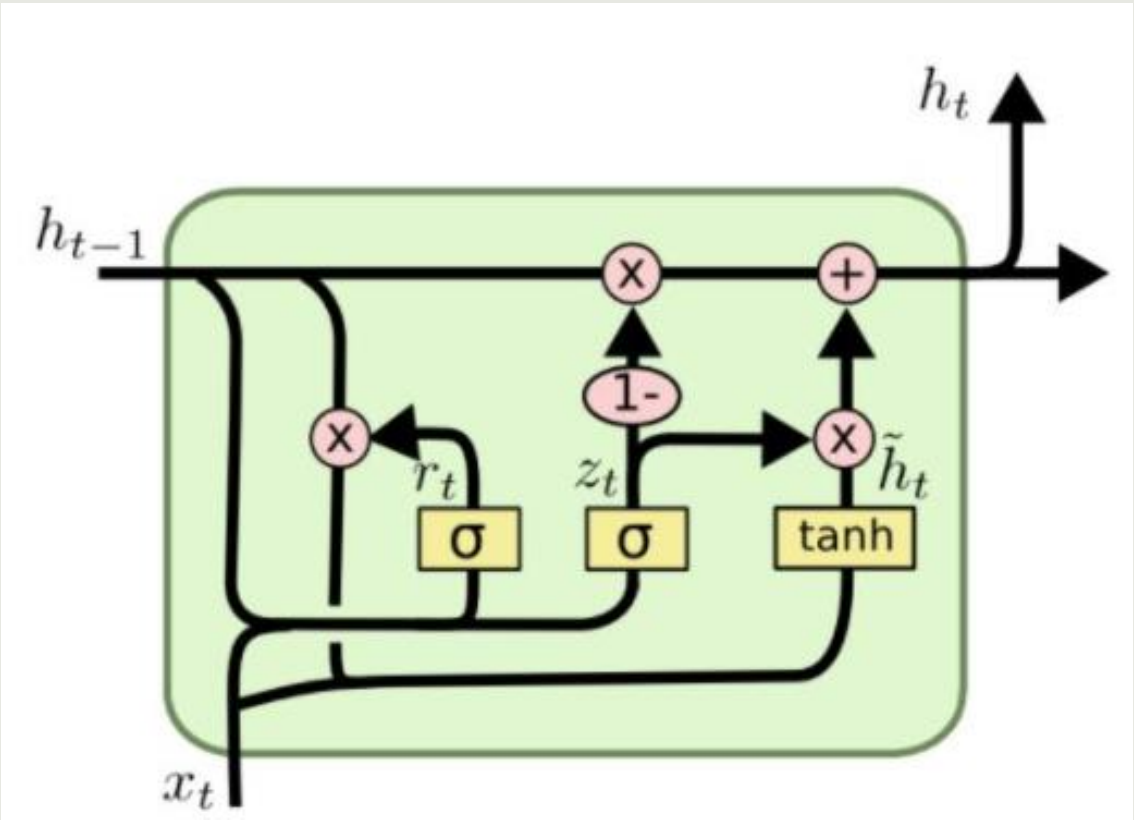
- Long-Term Dependency
- Vanishing / Exploding Gradient Problem



Deep Learning Architecture - RNN

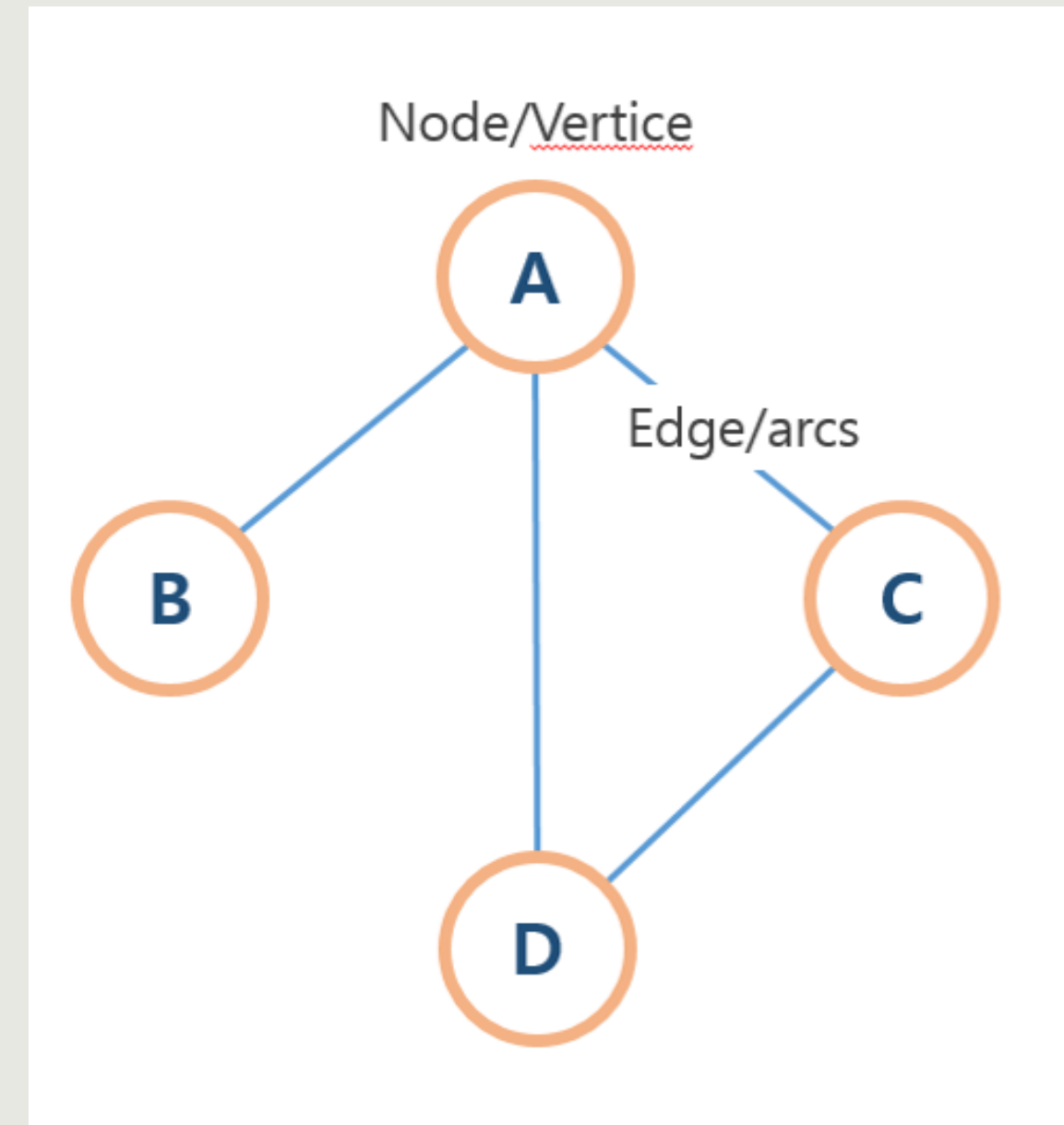
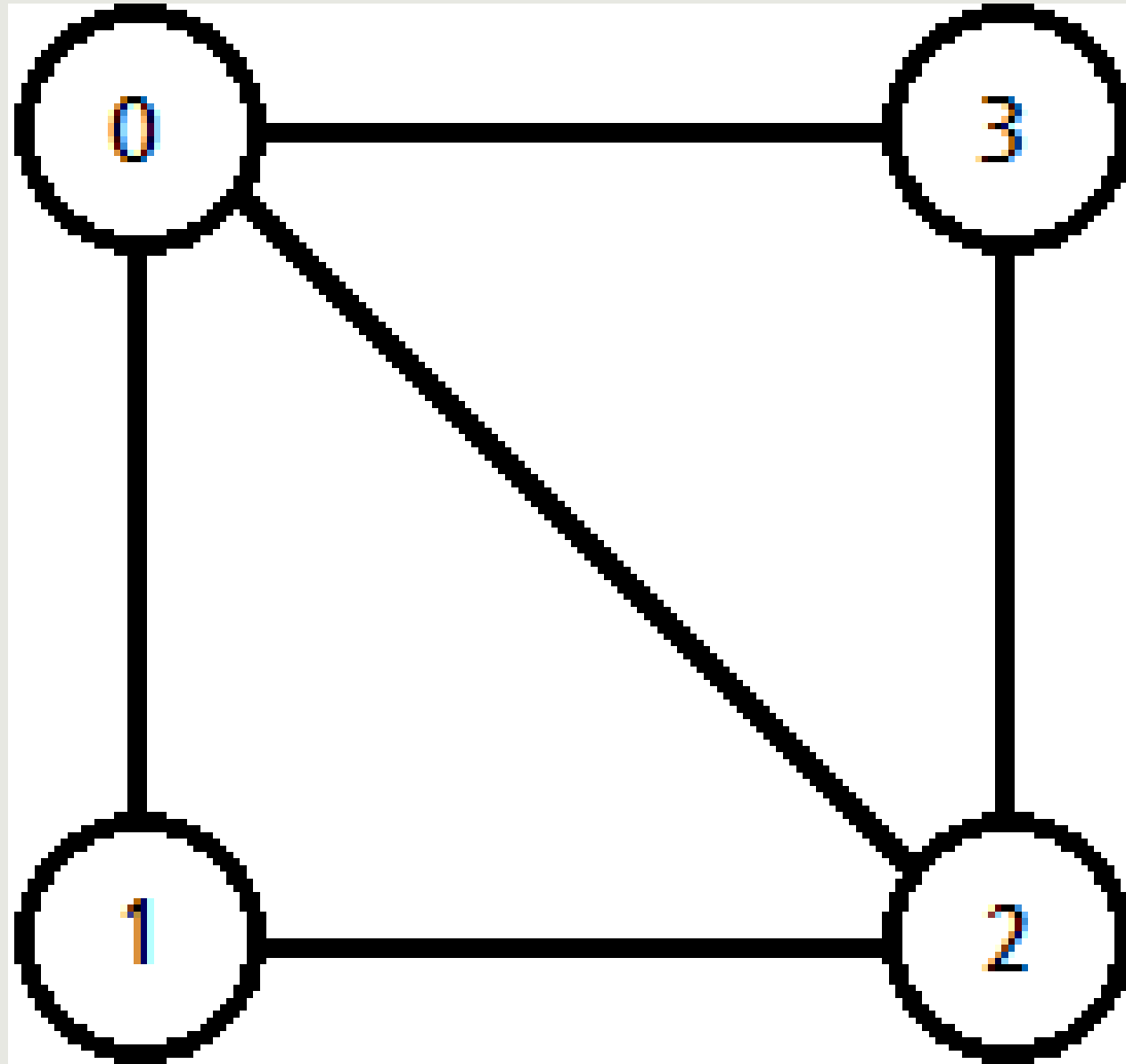


LSTM

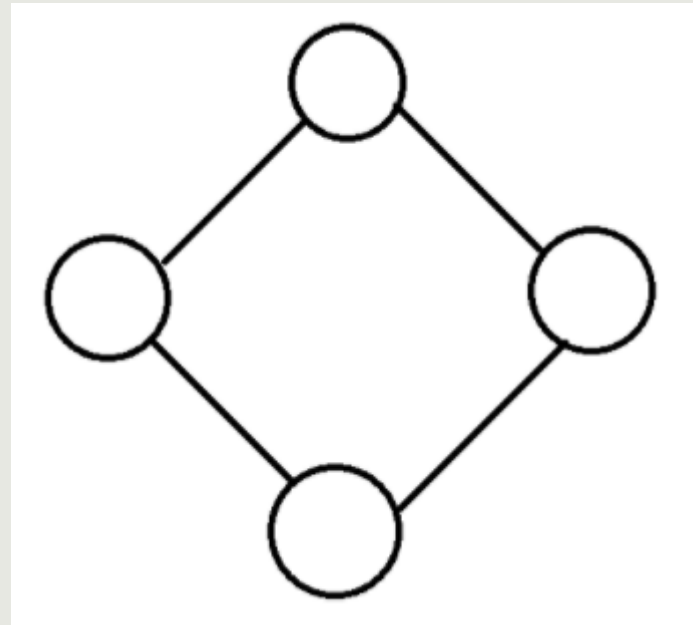


GRU

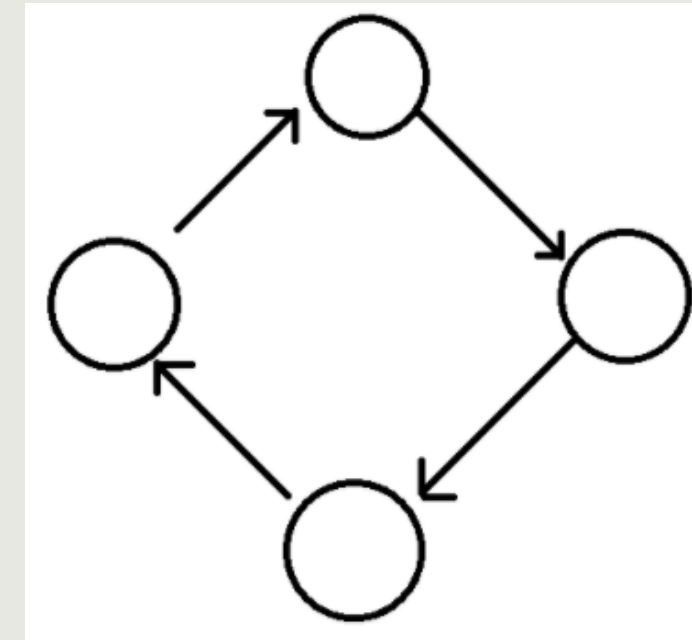
Deep Learning Architecture - GNN



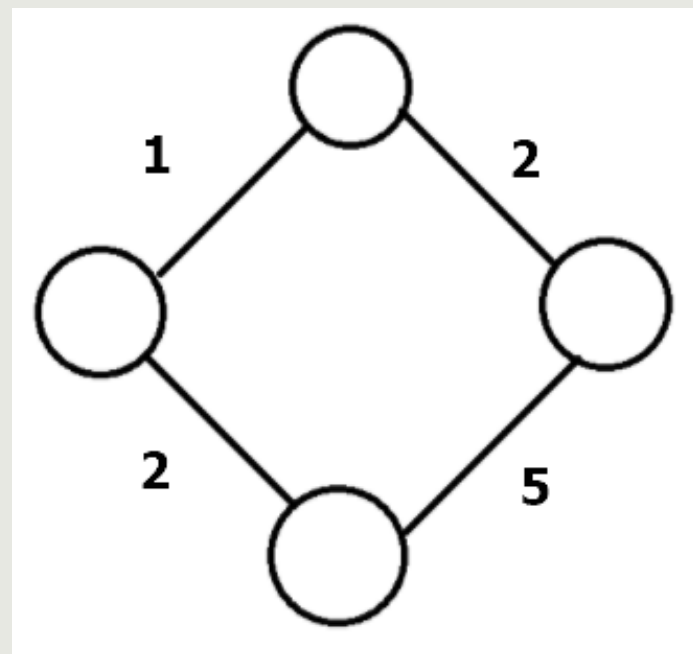
Deep Learning Architecture - GNN



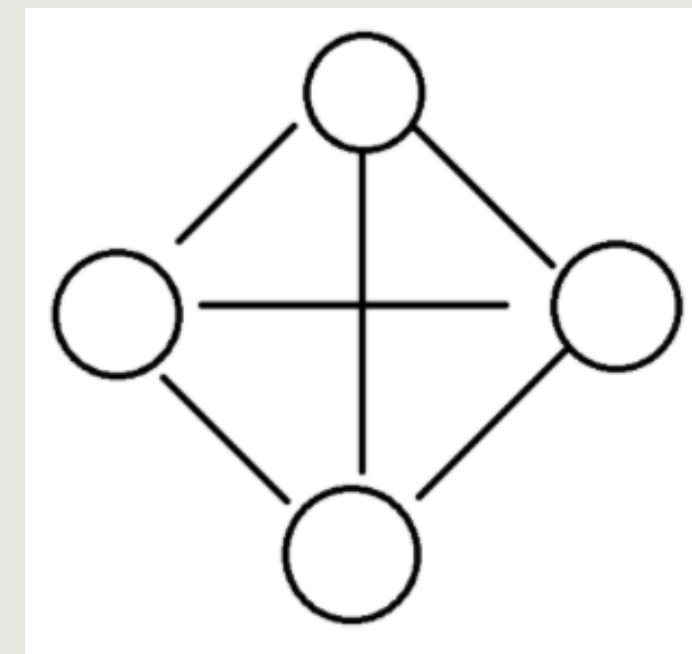
Undirected Graph



Directed Graph

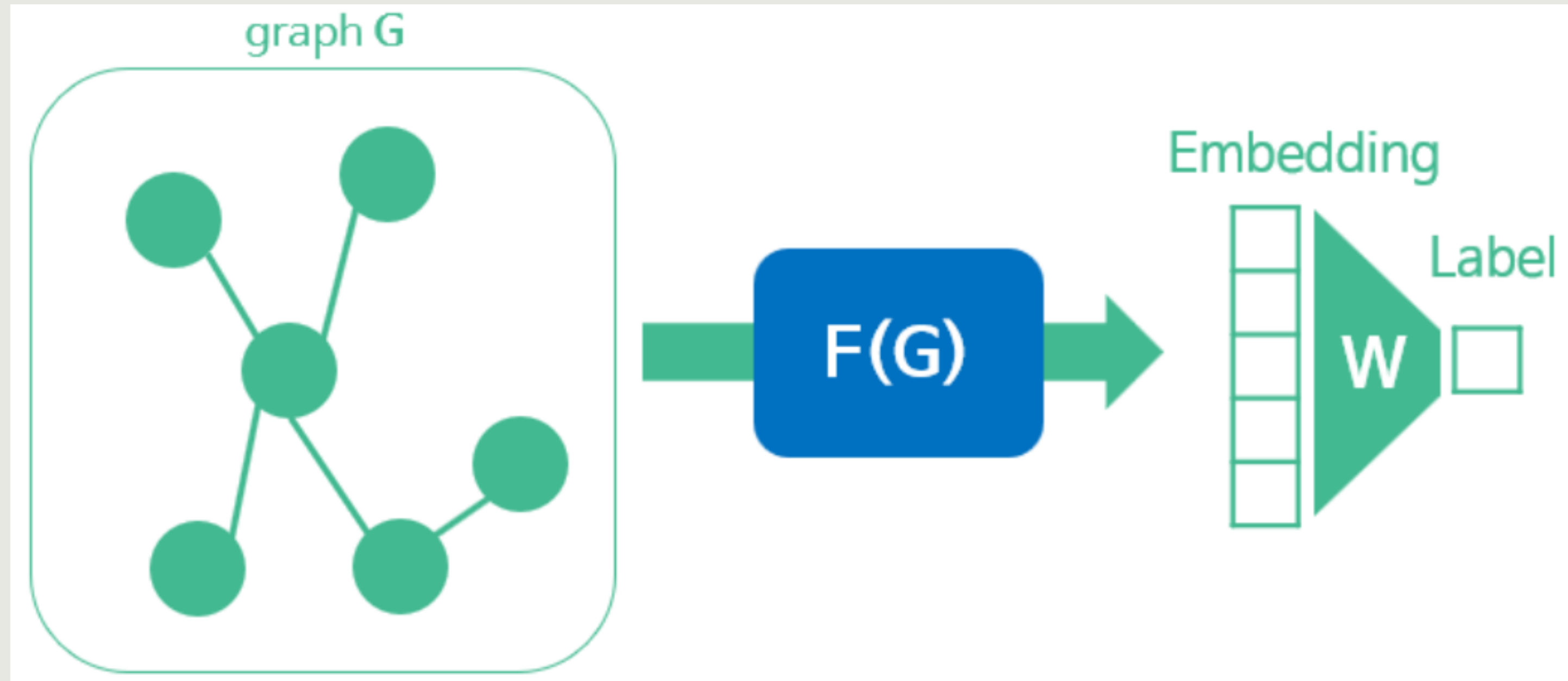


Weighted Graph

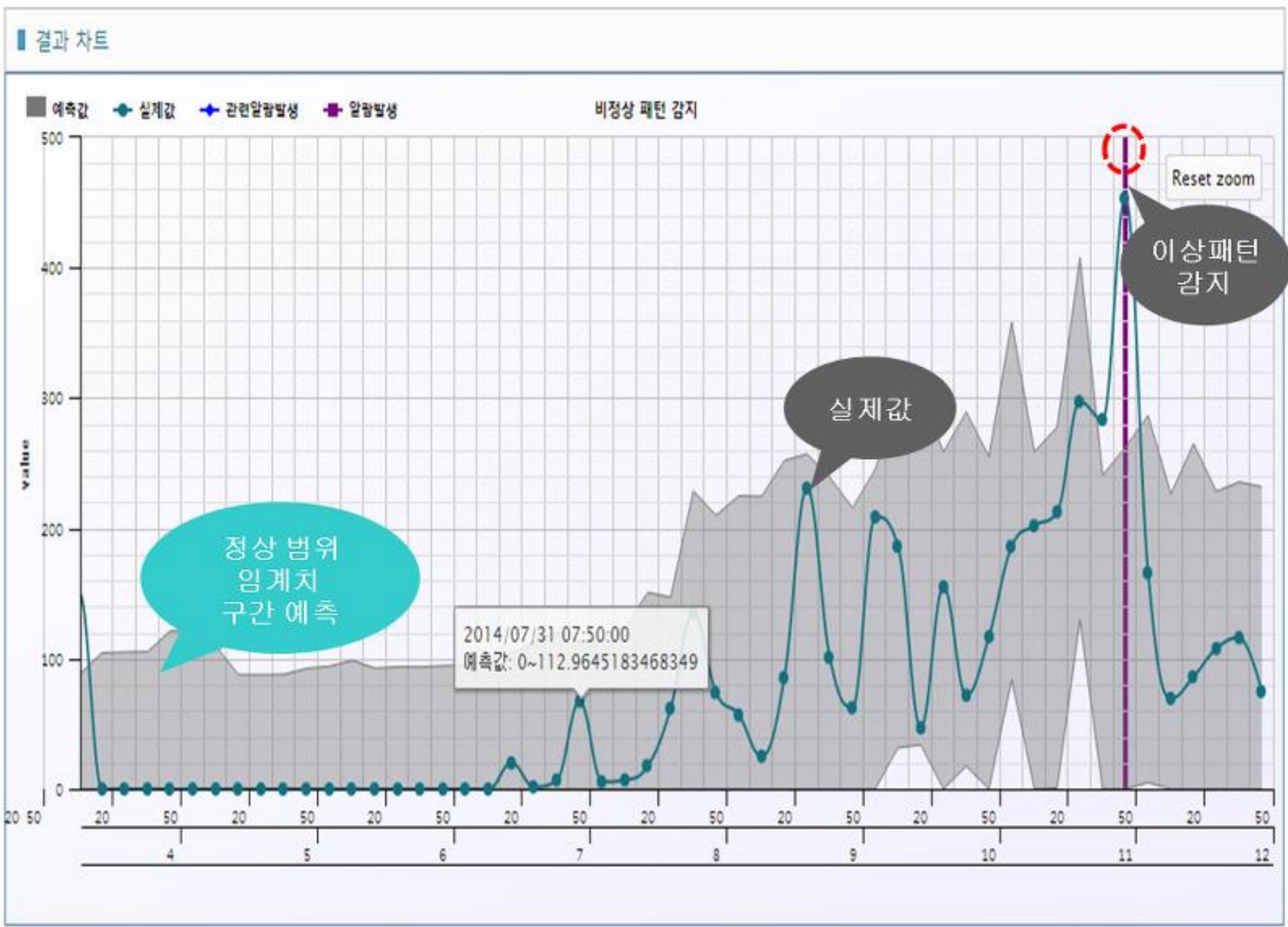
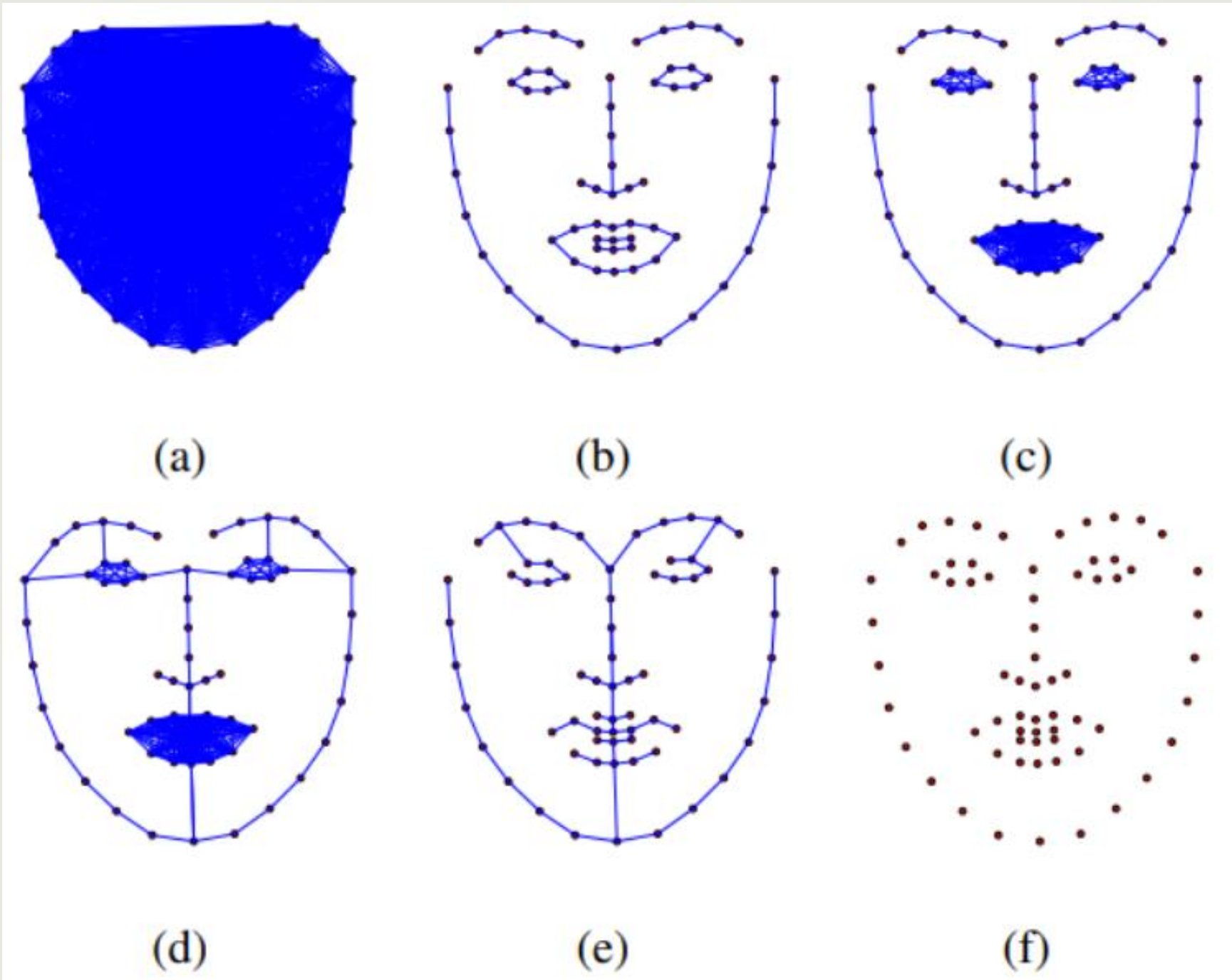


Complete Graph

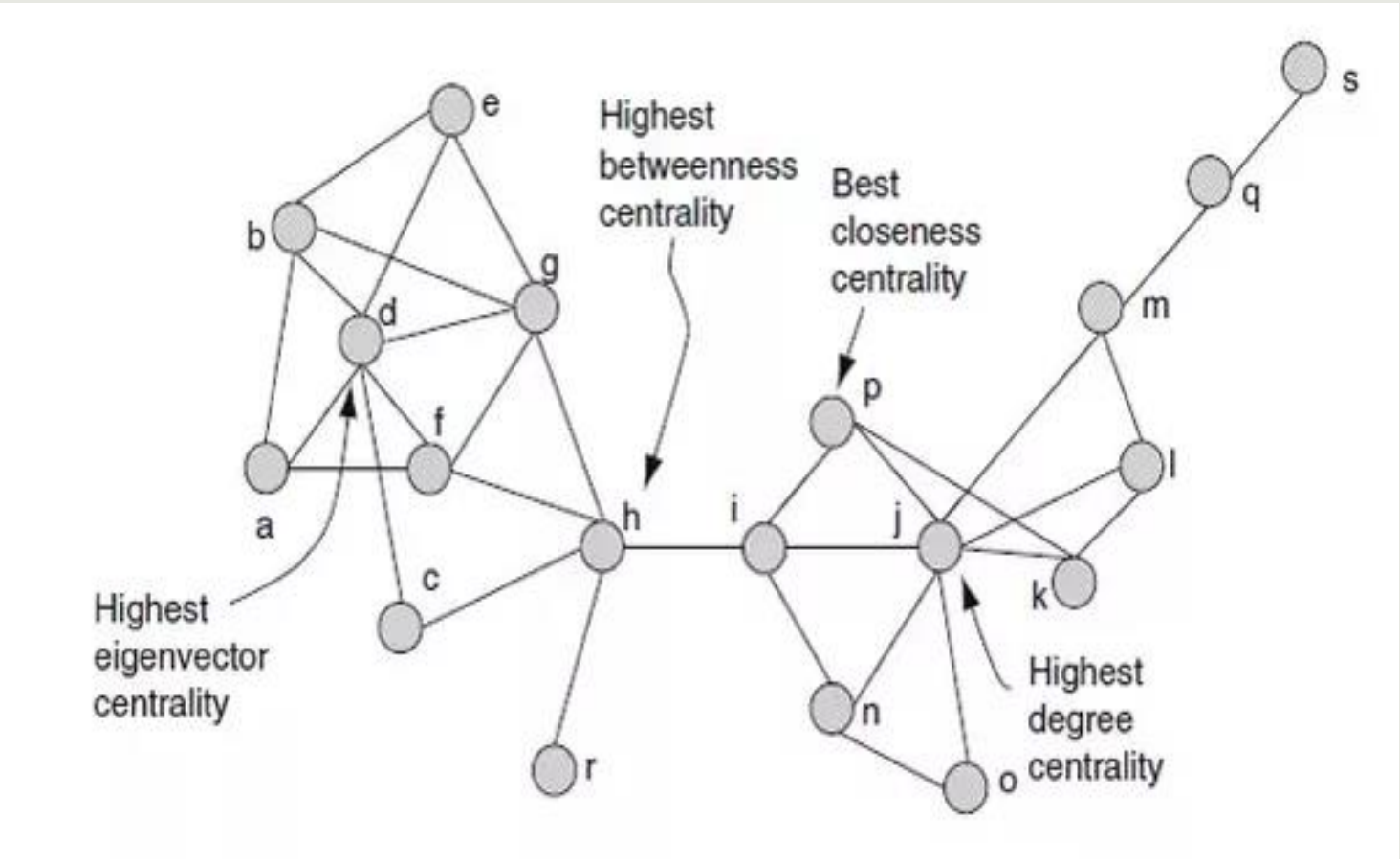
Deep Learning Architecture - GNN



Deep Learning Architecture - GNN

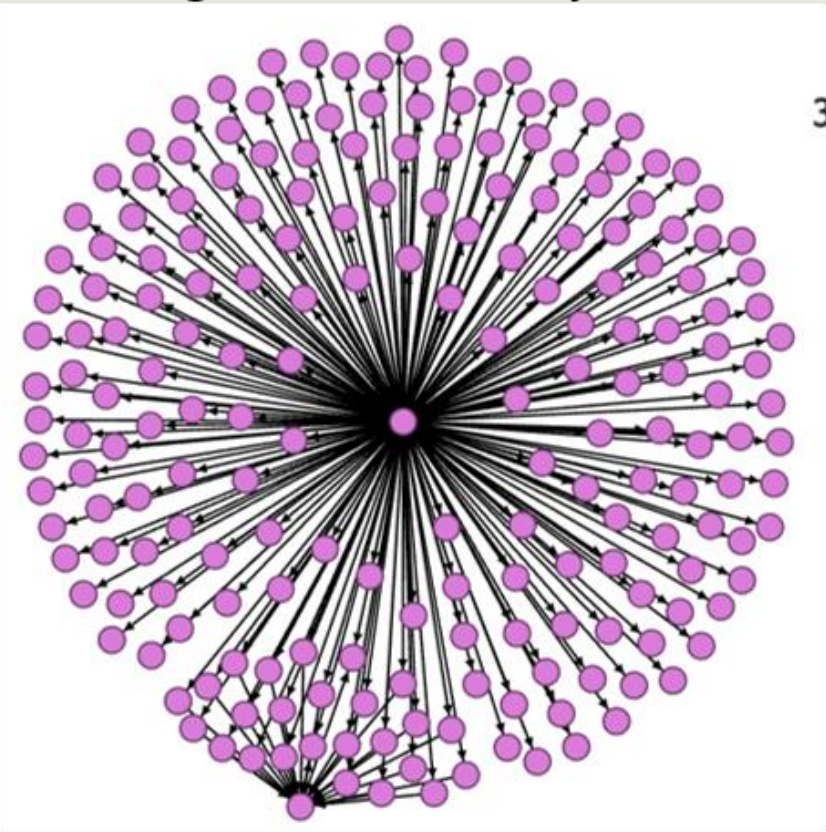


Deep Learning Architecture - GNN

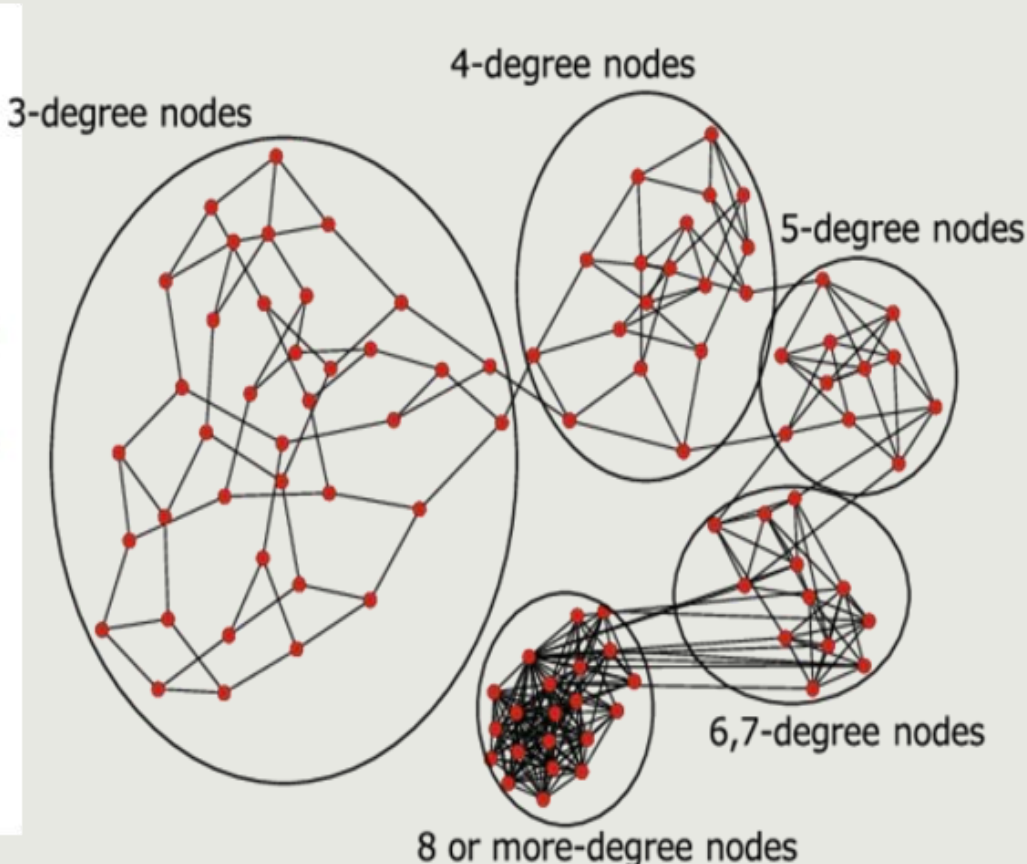


Node Importance

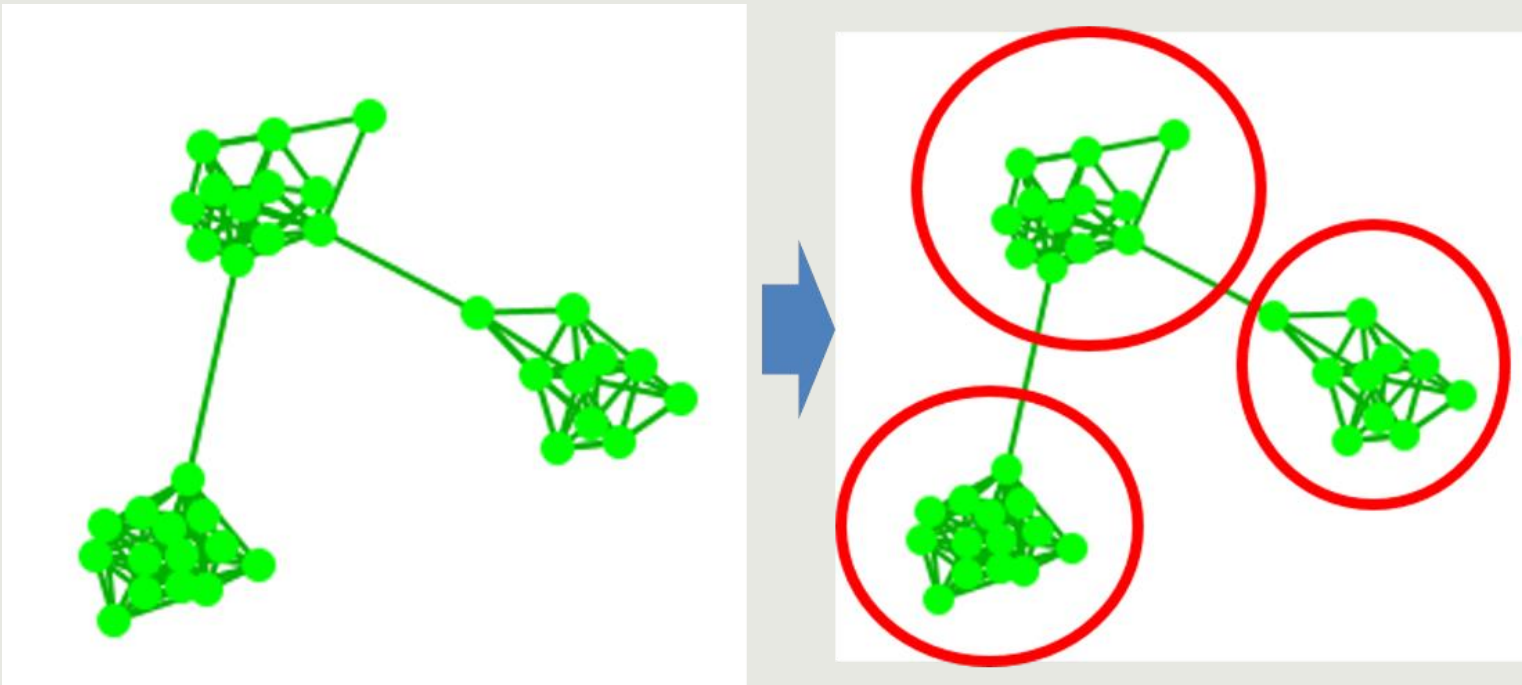
Low degree assortativity network



High degree assortativity network



Network Structure Analysis



Network Clustering

감사합니다

