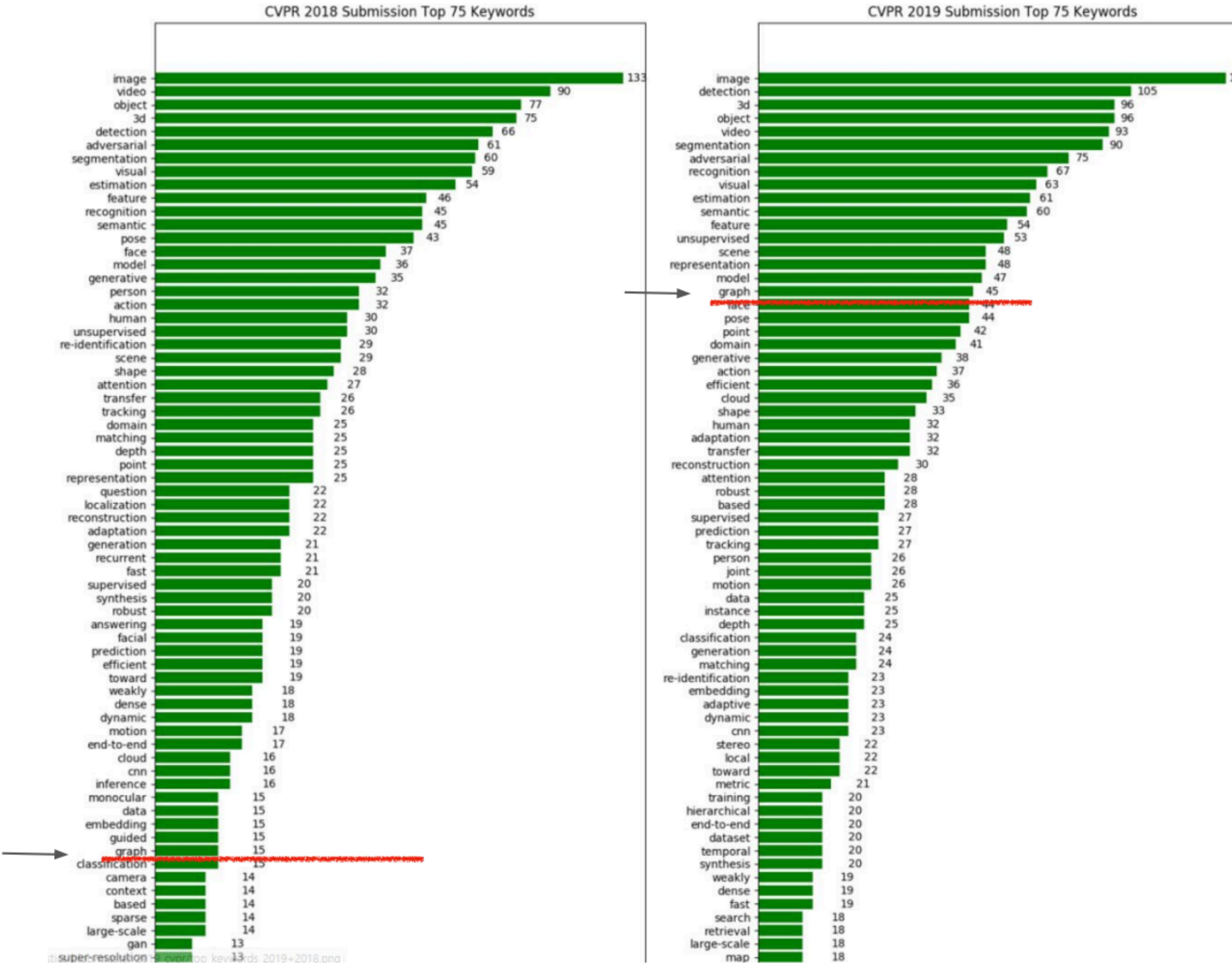


# **Graph Convolutional Network**

**Jaeyoung Cheong**

# 1. Trend

2019 CVPR



# 1. Trend

## SIGIR 2020 ACCEPTED PAPERS

The screenshot shows the 'Accepted Papers' section of the SIGIR 2020 website. At the top, there's a navigation bar with links for 'SUBMIT', 'PROGRAM', 'ATTEND', 'ORGANIZERS', 'SPONSORS', 'VOLUNTEERS', and 'PHOTO GALLERY'. Below the navigation is a large image of a traditional Chinese building at night with red lanterns. The main content area is titled 'Accepted Papers' and shows a grid of paper thumbnails. Each thumbnail includes the paper title and authors. A yellow arrow points from the text 'Ctrl+F: "recommend"' to the search bar on the right side of the grid.

Accepted Papers

Home / Program / Accepted Papers

Full Papers

Learning Efficient Representations of Mouse Movements to Predict User Attention in Sponsored Search  
Ioannis Arapakis: Telefonica Research; Luis A. Leiva: Aalto University

Measuring Recommendation Explanation Quality: The Conflicting Goals of Explanations  
Krisztian Balog: Google; Filip Radlinski: Google

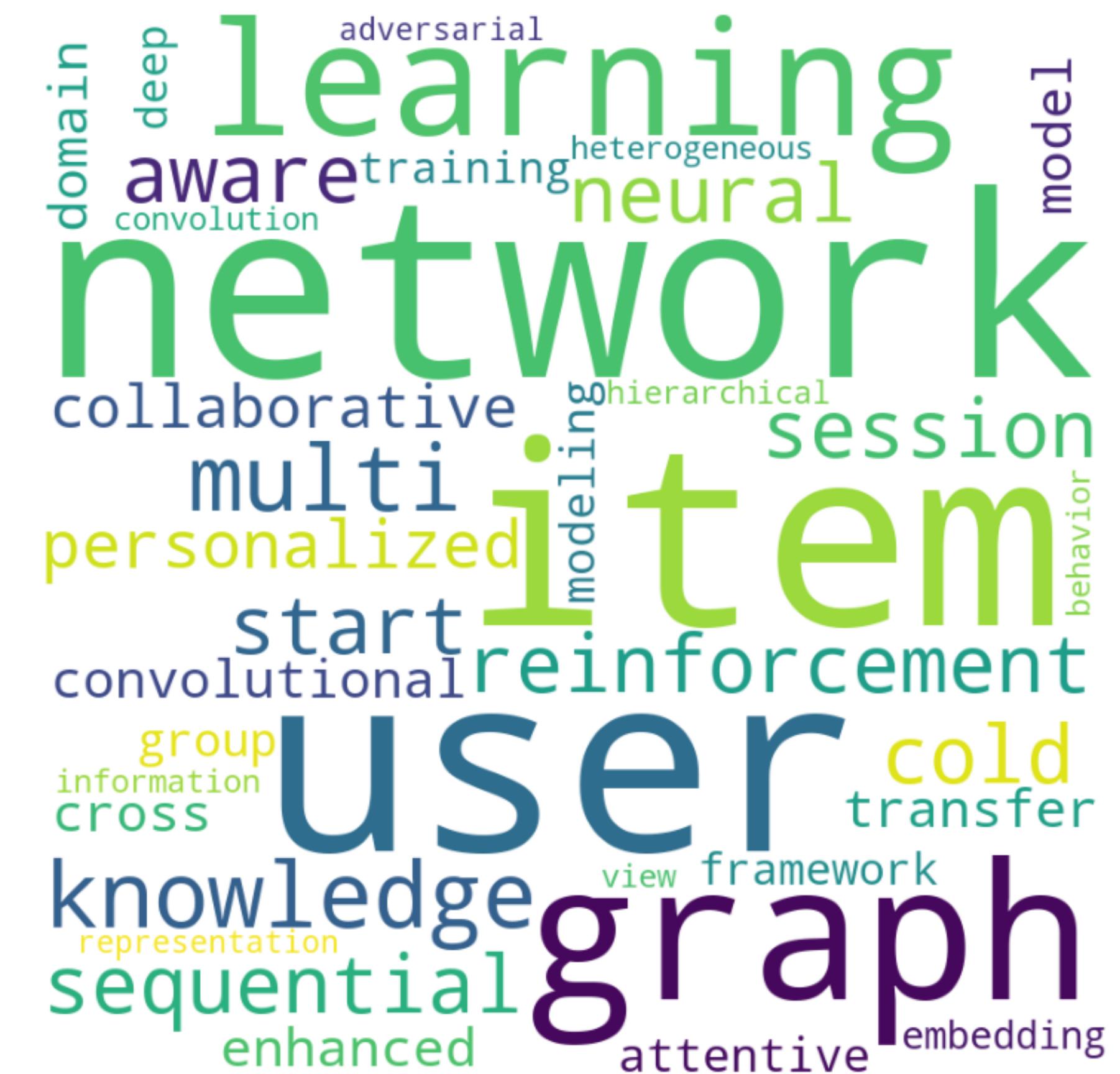
Bayesian Inferential Risk Evaluation On Multiple IR Systems  
Rodger Benham: RMIT University; Ben Carterette: Spotify; J. Shane Culpepper: RMIT University; Alistair Moffat: The University of Melbourne

Operationalizing the Legal Principle of Data Minimization for Personalization  
Asia J. Biega: Microsoft Research; Peter Potash: Microsoft Research; Hal Daumé III: Microsoft Research and University of Maryland; Fernando Diaz: Microsoft Research; Michèle Finck: Max Planck Institute for Innovation and Competition

Full Papers

Ctrl+F: 'recommend'

| page | count         |    |
|------|---------------|----|
| 0    | graph         | 21 |
| 1    | learning      | 18 |
| 2    | network       | 18 |
| 3    | knowledge     | 11 |
| 4    | multi         | 9  |
| 5    | item          | 9  |
| 6    | sequential    | 9  |
| 7    | cold          | 8  |
| 8    | start         | 8  |
| 9    | aware         | 7  |
| 10   | neural        | 7  |
| 11   | session       | 7  |
| 12   | reinforcement | 7  |
| 13   | personalized  | 6  |
| 14   | user          | 6  |
| 15   | users         | 6  |
| 16   | collaborative | 5  |
| 17   | convolutional | 5  |
| 18   | transfer      | 5  |
| 19   | cross         | 5  |
| 20   | domain        | 5  |
| 21   | enhanced      | 4  |
| 22   | model         | 4  |
| 23   | networks      | 4  |
| 24   | attentive     | 4  |



'recommendation' 같이 무정보적인 것은 불용어 처리함

# 1. Trend

## ACL 2020 ACCEPTED PAPERS

ACL 2020 Schedule Program Blog Registration Participants/FAQs Sponsors Calls

**Accepted Papers**  
Note that the titles/authors may change and papers may be withdrawn. For the final titles/authors, please refer to the proceedings on the anthology when they are out.

**Main Conference**  
There were 570 Long Papers and 208 Short Papers accepted.

**Long Papers**

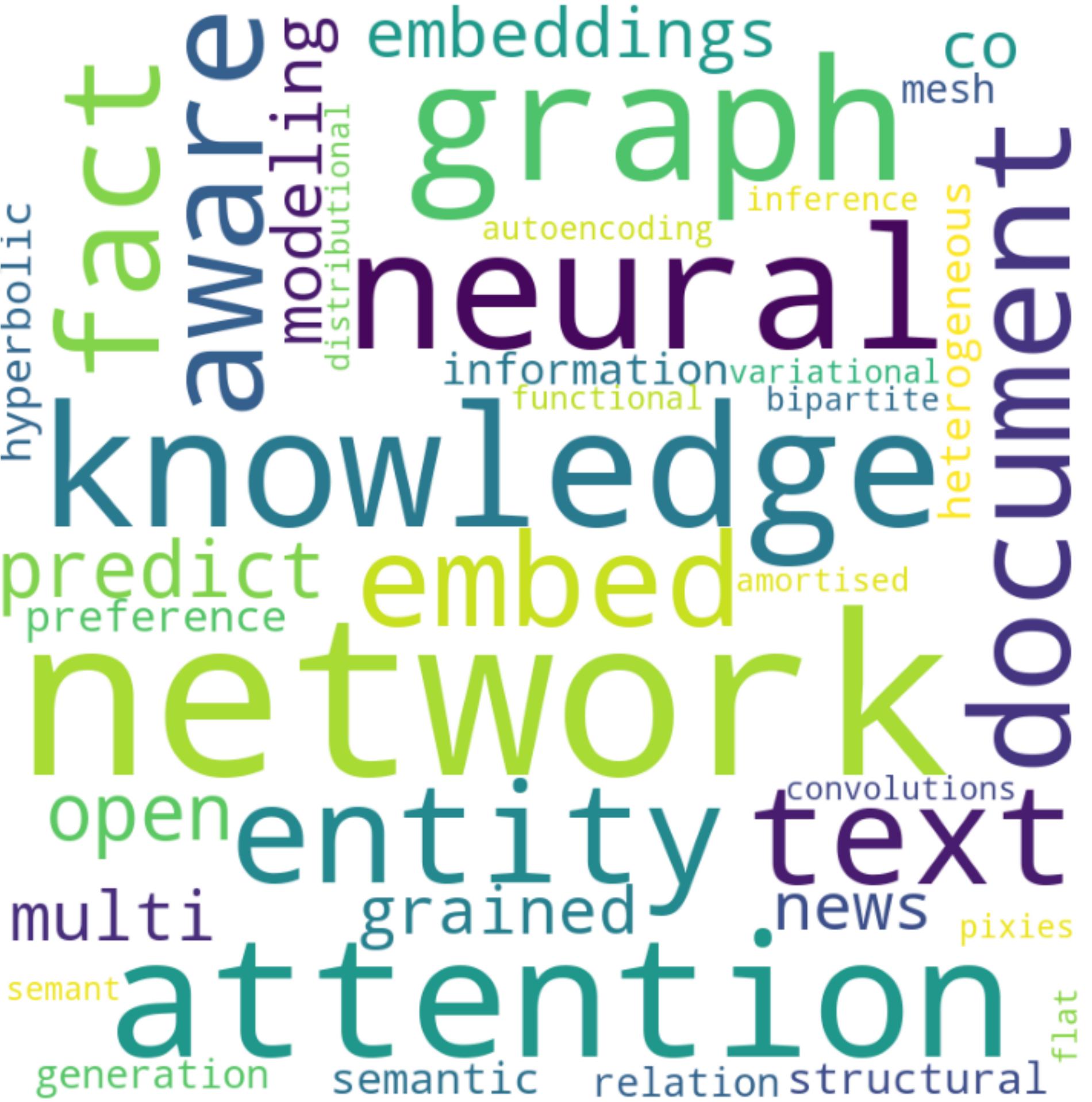
2kenize: Tying Subword Sequences for Chinese Script Conversion  
Pranav A and Isabelle Augenstein

A Batch Normalized Inference Network Keeps the KL Vanishing Away  
Qile Zhu, Wei Bi, Xiaojiang Liu, Xiyao Ma, Xiaolin Li and Dapeng Wu

**On this page**  
Main Conference  
Long Papers  
Short Papers  
System Demonstrations  
Student Research Workshop

Ctrl+F: 'graph'

| page | count      |    |
|------|------------|----|
| 0    | graph      | 30 |
| 1    | network    | 10 |
| 2    | knowledge  | 7  |
| 3    | attention  | 6  |
| 4    | networks   | 5  |
| 5    | neural     | 5  |
| 6    | entity     | 3  |
| 7    | document   | 3  |
| 8    | fact       | 3  |
| 9    | aware      | 3  |
| 10   | text       | 3  |
| 11   | embed      | 3  |
| 12   | predict    | 2  |
| 13   | open       | 2  |
| 14   | embeddings | 2  |
| 15   | modeling   | 2  |
| 16   | multi      | 2  |
| 17   | grained    | 2  |
| 18   | co         | 2  |
| 19   | news       | 2  |
| 20   | preference | 2  |



# 1. Trend

ACL & SIGIR 정리 2020 accepted papers

|    | page          | count |
|----|---------------|-------|
| 0  | graph         | 21    |
| 1  | learning      | 18    |
| 2  | network       | 18    |
| 3  | knowledge     | 11    |
| 4  | multi         | 9     |
| 5  | item          | 9     |
| 6  | sequential    | 9     |
| 7  | cold          | 8     |
| 8  | start         | 8     |
| 9  | aware         | 7     |
| 10 | neural        | 7     |
| 11 | session       | 7     |
| 12 | reinforcement | 7     |
| 13 | personalized  | 6     |
| 14 | user          | 6     |
| 15 | users         | 6     |
| 16 | collaborative | 5     |
| 17 | convolutional | 5     |
| 18 | transfer      | 5     |
| 19 | cross         | 5     |
| 20 | domain        | 5     |
| 21 | enhanced      | 4     |
| 22 | model         | 4     |
| 23 | networks      | 4     |
| 24 | attentive     | 4     |

## SIGIR

### Recommennder System

- Graph는 가장 HOT한 키워드
- 강화학습도 HOT
- Cold Start
- 개인화
- 협업 필터링
- 세션 기반 추천
- 신경망 기반 추천시스템

|    | page       | count |
|----|------------|-------|
| 0  | graph      | 30    |
| 1  | network    | 10    |
| 2  | knowledge  | 7     |
| 3  | attention  | 6     |
| 4  | networks   | 5     |
| 5  | neural     | 5     |
| 6  | entity     | 3     |
| 7  | document   | 3     |
| 8  | fact       | 3     |
| 9  | aware      | 3     |
| 10 | text       | 3     |
| 11 | embed      | 3     |
| 12 | predict    | 2     |
| 13 | open       | 2     |
| 14 | embeddings | 2     |
| 15 | modeling   | 2     |
| 16 | multi      | 2     |
| 17 | grained    | 2     |
| 18 | co         | 2     |
| 19 | news       | 2     |
| 20 | preference | 2     |

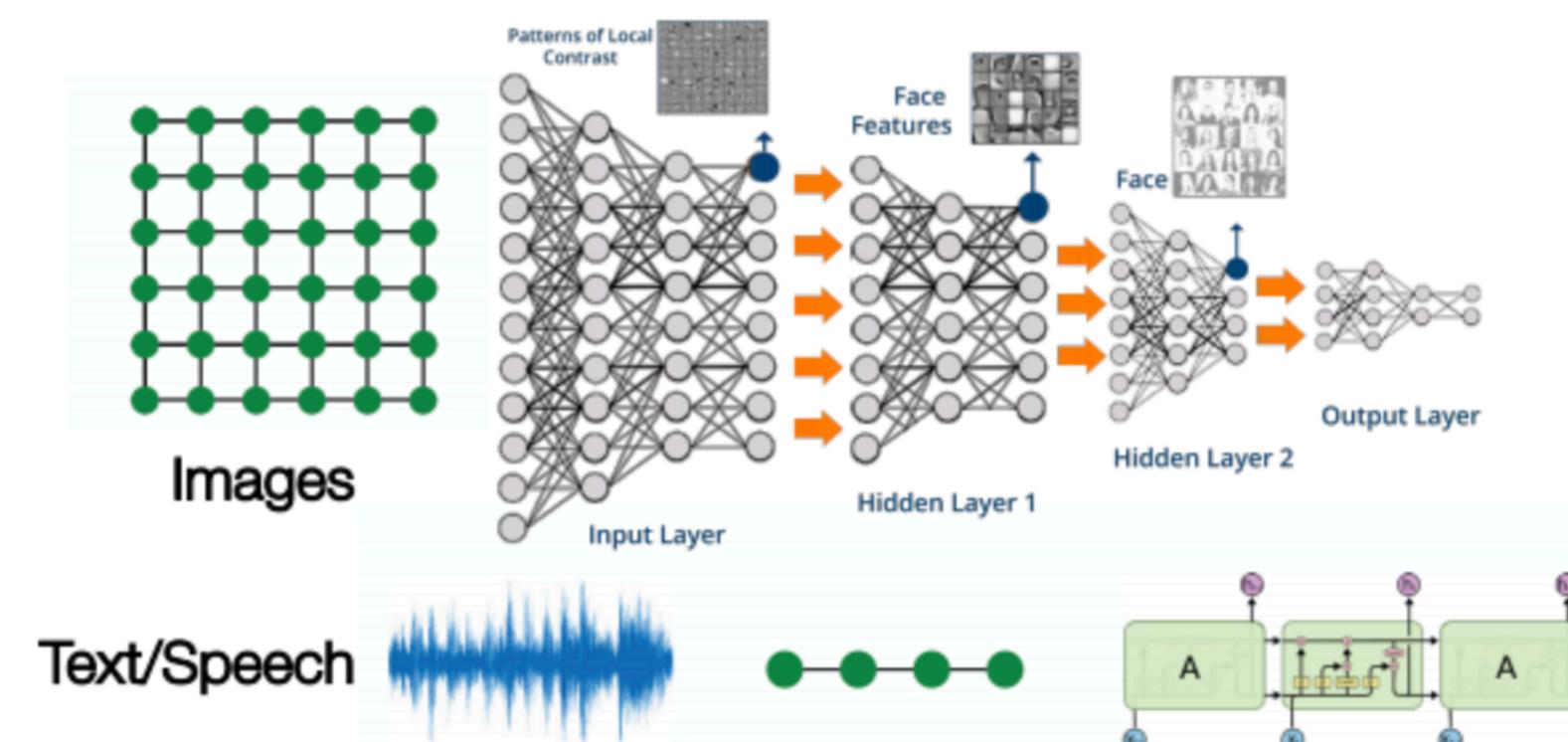
## ACL

### Graph

- GNN으로 NLP 다룸
- Graph Attention Net의 급부상

## 2. Motivation

- 대부분의 성공적인 딥 or 머신러닝 작업에서 사용되던 데이터는 **유클리디안 공간(grid 위에서 표현된 공간)**에서 표현되어진다.
- 하지만 데이터가 비유클리디안 공간에서 만들어지거나 **복잡한 관계를 가진 그래프로** 표현된 경우,  
기존의 머신러닝 알고리즘으로 이것을 처리하기 어렵다. (**그래프는 비유클리드 공간 위에서 표현됨**)
- e-커머스: 구매자와 상품 간의 복잡한 상호관계를 다룬다.
- 화학: 분자들은 그래프로 모델화된다.



Modern deep learning toolbox is  
designed for simple sequences & grids

### ConvGNNs

컴퓨터 비전에서 CNN의 성공이 동기가 되어, 그래프 데이터의 Convolution 개념에 대한 재정의가 이루어졌다.

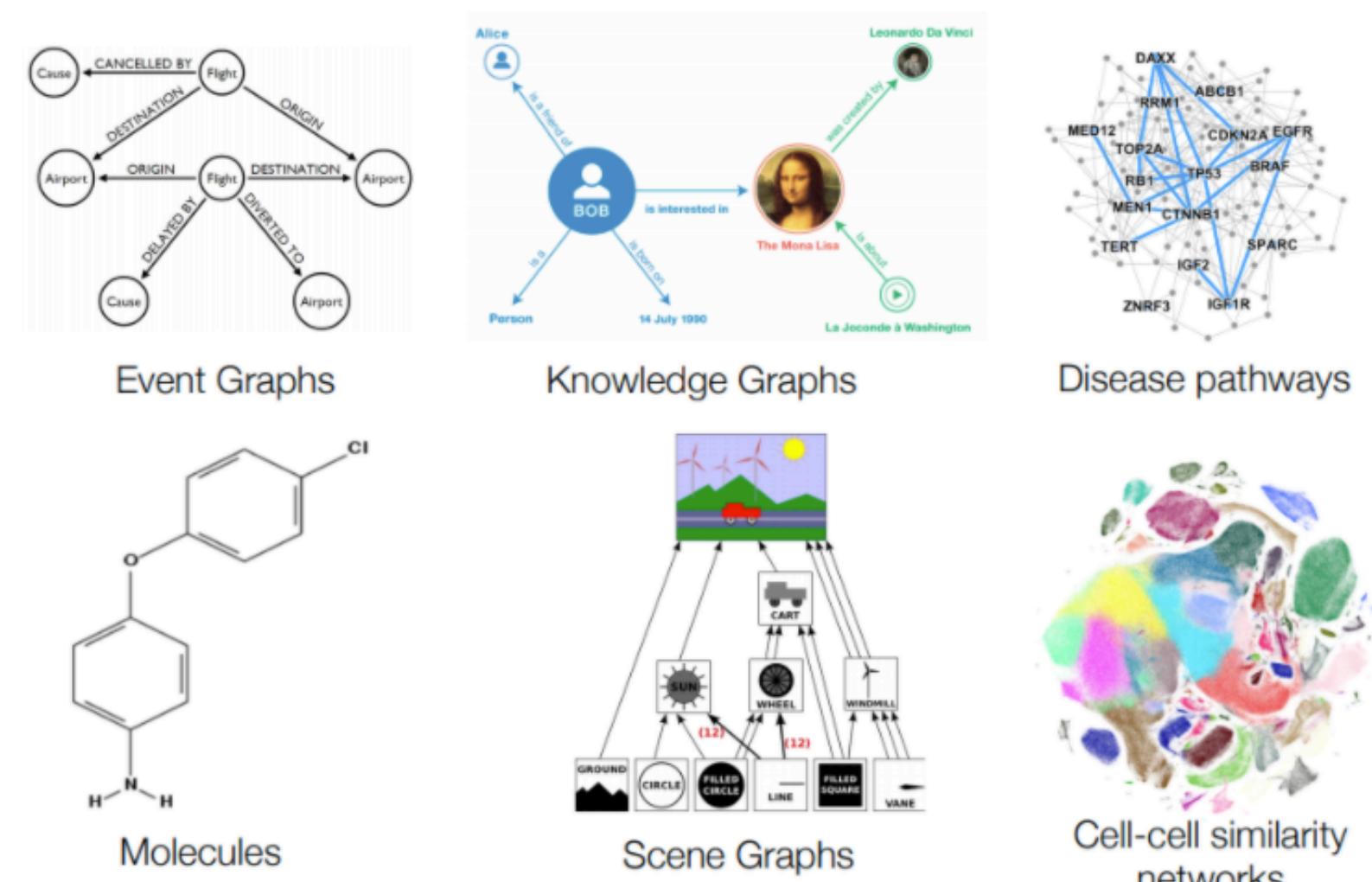
- the spectral-based approaches (Bruna et al 2013)
- The spatial-based approaches

### Reccurent GNNs

### AutoEncoder GNNs

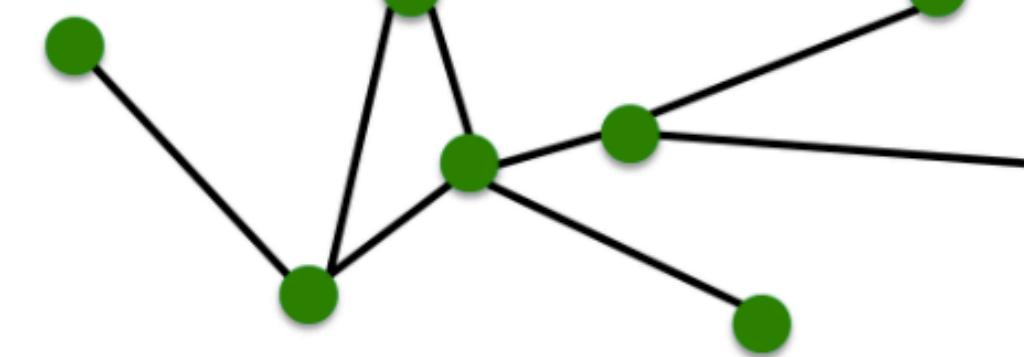
### Graph Attention Networks

...



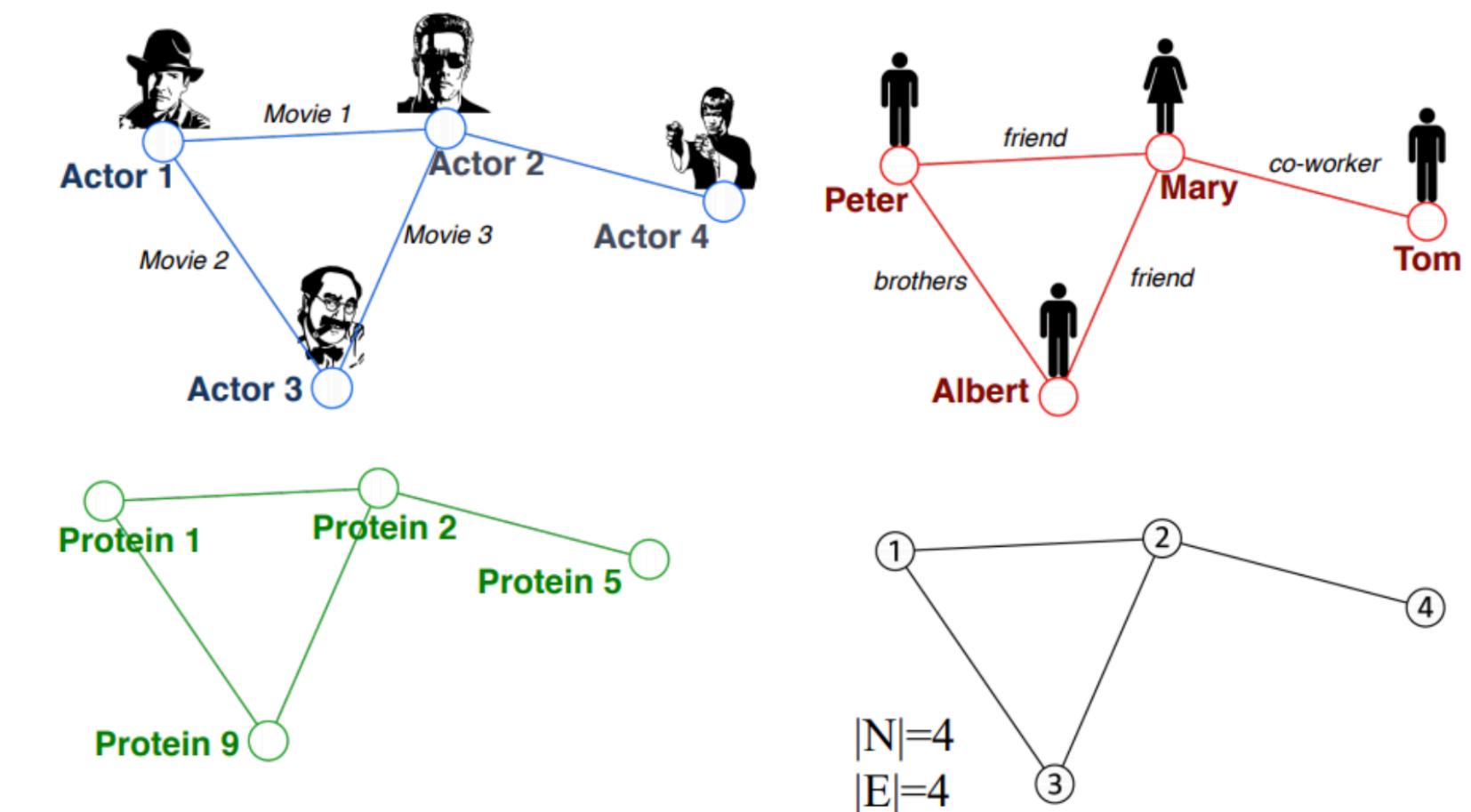
# 3. Graph

## 3.1 Def



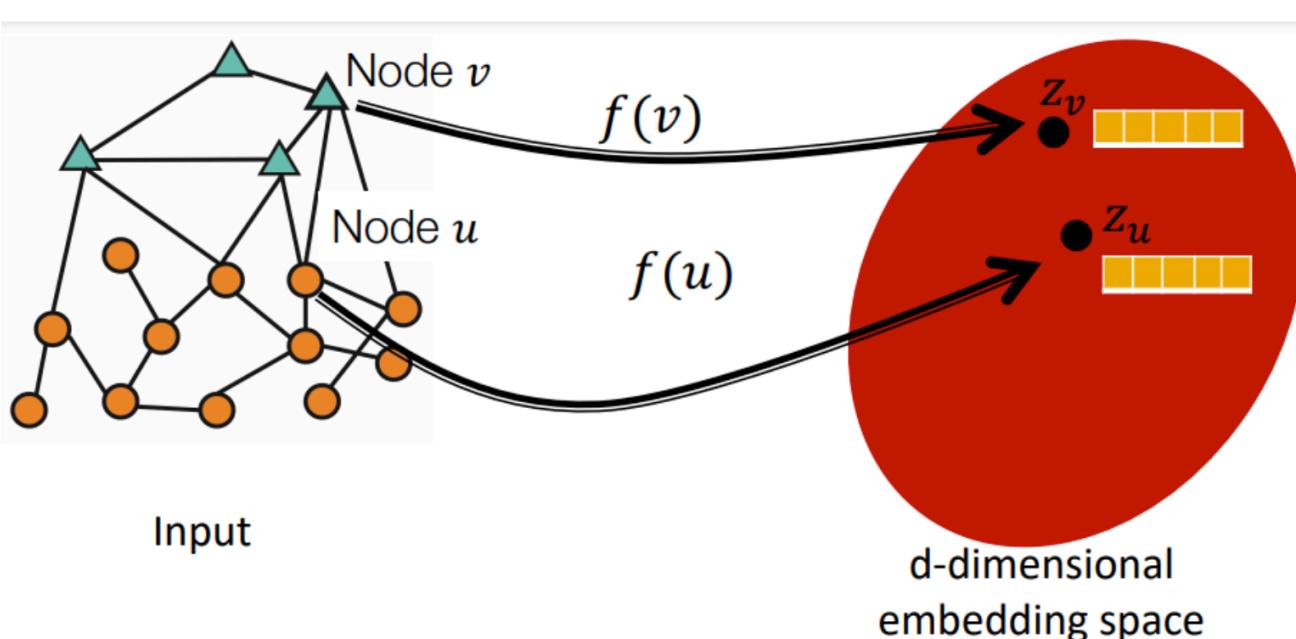
- **Objects:** nodes, vertices
- **Interactions:** links, edges
- **System:** network, graph

$$\begin{array}{c} N \\ E \\ G(N,E) \end{array}$$



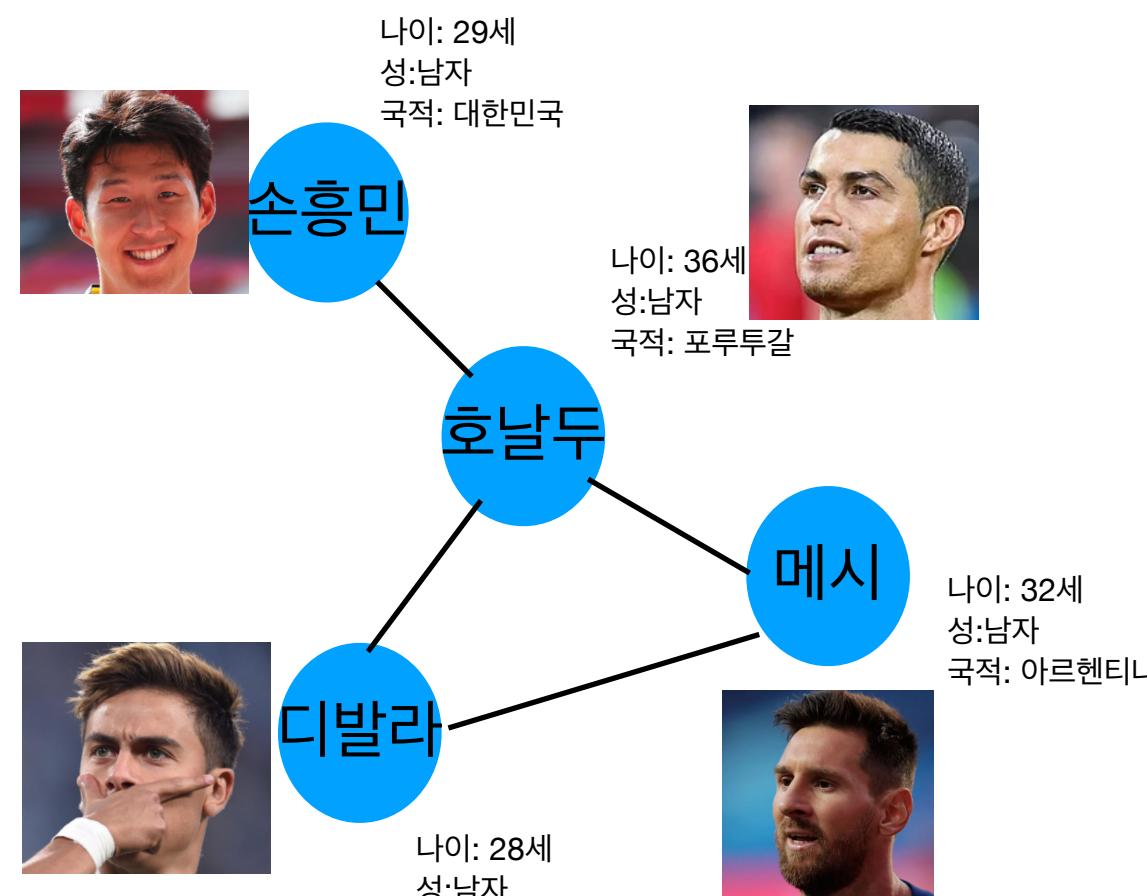
## 3.2 Embedding Node

그래프의 node를  $d$  차원으로 임베딩. 여기서 유사한 이웃을 가진 노드들은 서로 가까이 있게끔 임베딩된다.



# 3. Graph

## 3.3 How to represent graph into the matrix



|         | 국적 | Feature2 | Feature3 |
|---------|----|----------|----------|
| son     | 0  | 1        | 0        |
| Ronaldo | 1  | 1        | 1        |
| Messi   | 1  | 1        | 1        |
| Dybala  | 1  | 1        | 1        |

국적을 나타내는 Features라고 해보자

대한민국

포르투갈

둘다 아르헨티나

N(node) x F(features)의 **Node Feature Matrix**  $X_{ij}$

|         | son | Ronaldo | Messi | Dybala |
|---------|-----|---------|-------|--------|
| son     | 0   | 1       | 0     | 0      |
| Ronaldo | 1   | 0       | 1     | 1      |
| Messi   | 0   | 1       | 0     | 1      |
| Dybala  | 0   | 1       | 1     | 0      |

|         | son | Ronaldo | Messi | Dybala |
|---------|-----|---------|-------|--------|
| son     | 1   | 0       | 0     | 0      |
| Ronaldo | 0   | 3       | 0     | 0      |
| Messi   | 0   | 0       | 2     | 0      |
| Dybala  | 0   | 0       | 0     | 2      |

N x N(node)의 **Adjacent Matrix**  $A_{ij}$

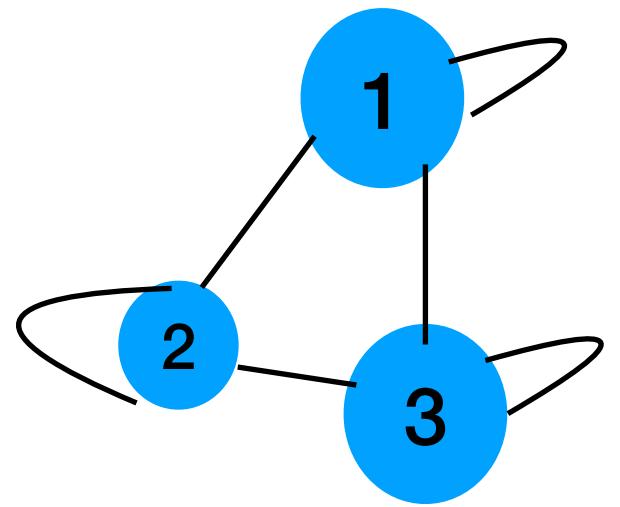
$D$  **Degree matrix**

# 4. Graph Convolutional Networks

## 4.1 The limits of adjacent matrix

1) 각 노드 자신에 대한 정보가 학습 시 업데이트되지 않는다.

→ 각 Mat에 I를 더해줌으로써, 각 노드 자기 자신에 대한 정보도 학습 시 업데이트할 수 있다.



|         | son | Ronaldo | Messi | Dybala |
|---------|-----|---------|-------|--------|
| son     | 1   | 0       | 0     | 0      |
| Ronaldo | 0   | 3       | 0     | 0      |
| Messi   | 0   | 0       | 2     | 0      |
| Dybala  | 0   | 0       | 0     | 2      |

D Degree matrix

|         | son | Ronaldo | Messi | Dybala |
|---------|-----|---------|-------|--------|
| son     | 1   | 0       | 0     | 0      |
| Ronaldo | 0   | 1       | 0     | 0      |
| Messi   | 0   | 0       | 1     | 0      |
| Dybala  | 0   | 0       | 0     | 1      |

+

|         | son | Ronaldo | Messi | Dybala |
|---------|-----|---------|-------|--------|
| son     | 1   | 0       | 0     | 0      |
| Ronaldo | 0   | 1       | 0     | 0      |
| Messi   | 0   | 0       | 1     | 0      |
| Dybala  | 0   | 0       | 0     | 1      |

+

|   |   |   |   |
|---|---|---|---|
| 2 | 0 | 0 | 0 |
| 0 | 4 | 0 | 0 |
| 0 | 0 | 3 | 0 |
| 0 | 0 | 0 | 3 |

$\sim$   
D

|         | son | Ronaldo | Messi | Dybala |
|---------|-----|---------|-------|--------|
| son     | 0   | 1       | 0     | 0      |
| Ronaldo | 1   | 0       | 1     | 1      |
| Messi   | 0   | 1       | 0     | 1      |
| Dybala  | 0   | 1       | 1     | 0      |

$A_{ij}$  Adjacent Matrix

|   |   |   |   |
|---|---|---|---|
| 1 | 1 | 0 | 0 |
| 1 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 |

$\sim$   
 $A_{ij}$

# 4. Graph Convolutional Networks

## 4.1 The limits of adjacent matrix

2) 기본적인 인접행렬  $A$ 는 정규화가 되어있지 않아 gradient explosion을 발생시키거나 feature matrix와 곱할 때( $AX$ )  $X$ 의 크기를 불안정하게 크게 만들 수 있다.

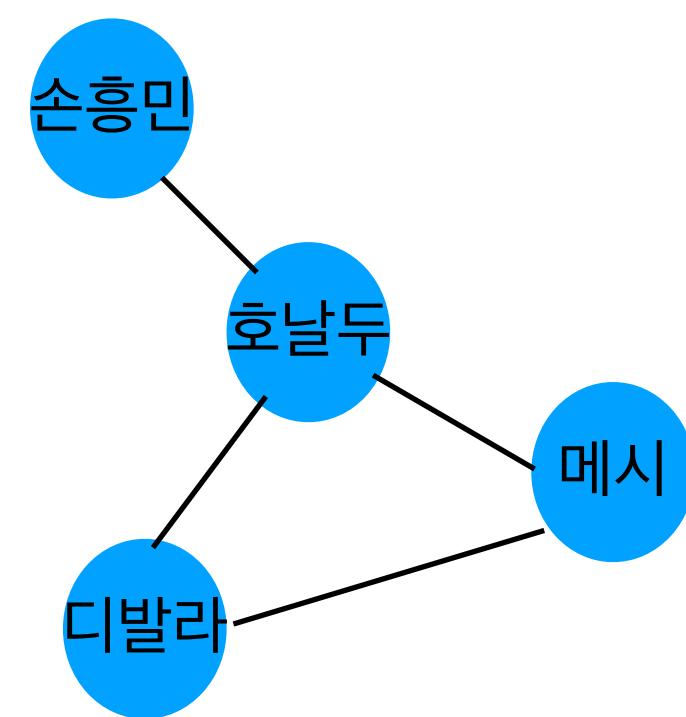
해소방안:

$$\tilde{A} \longrightarrow \hat{A} = \overbrace{D^{-0.5}}^{\sim} \overbrace{A}^{\sim} \overbrace{D^{-0.5}}^{\sim}$$

# 4. Graph Convolutional Networks

## 4.1 The limits of adjacent matrix

2) 기본적인 인접행렬  $A$ 는 정규화가 되어있지 않아 gradient explosion을 발생시키거나 feature matrix와 곱할 때( $AX$ )  $X$ 의 크기를 불안정하게 크게 만들 수 있다.



|   |   |   |   |
|---|---|---|---|
| 2 | 0 | 0 | 0 |
| 0 | 4 | 0 | 0 |
| 0 | 0 | 3 | 0 |
| 0 | 0 | 0 | 3 |

$$\sim_D$$

|     |     |     |     |
|-----|-----|-----|-----|
| 1/2 | 0   | 0   | 0   |
| 0   | 1/4 | 0   | 0   |
| 0   | 0   | 1/3 | 0   |
| 0   | 0   | 0   | 1/3 |

$$\widetilde{D}^{-1}$$

평균으로 정규화

고차 노드에 더 작은 수를 곱함으로써 영향력을 줄이기

$$\begin{array}{c}
 \begin{array}{|c|c|c|c|} \hline
 1/2 & 0 & 0 & 0 \\ \hline
 0 & 1/4 & 0 & 0 \\ \hline
 0 & 0 & 1/3 & 0 \\ \hline
 0 & 0 & 0 & 1/3 \\ \hline
 \end{array} * \begin{array}{|c|c|c|c|} \hline
 1 & 1 & 0 & 0 \\ \hline
 1 & 1 & 1 & 1 \\ \hline
 0 & 1 & 1 & 1 \\ \hline
 0 & 1 & 1 & 1 \\ \hline
 \end{array} \rightarrow \begin{array}{|c|c|c|c|} \hline
 1/2 & 1/2 & 0 & 0 \\ \hline
 1/4 & 1/4 & 1/4 & 1/4 \\ \hline
 0 & 1/3 & 1/3 & 1/3 \\ \hline
 0 & 1/3 & 1/3 & 1/3 \\ \hline
 \end{array} = \begin{array}{|c|c|c|c|} \hline
 1/2 & 0 & 0 & 0 \\ \hline
 0 & 1/4 & 0 & 0 \\ \hline
 0 & 0 & 1/3 & 0 \\ \hline
 0 & 0 & 0 & 1/3 \\ \hline
 \end{array} \\
 \sim D^{-1} \qquad \qquad \qquad \sim A_{ij} \qquad \qquad \qquad S \qquad \qquad \qquad \sim D^{-1} \qquad \qquad \qquad A
 \end{array}$$

# 4. Graph Convolutional Networks

## 4.1 The limits of adjacent matrix

|   |   |   |   |
|---|---|---|---|
| 2 | 0 | 0 | 0 |
| 0 | 4 | 0 | 0 |
| 0 | 0 | 3 | 0 |
| 0 | 0 | 0 | 3 |

$\tilde{D}$

|   |   |   |   |
|---|---|---|---|
| 1 | 1 | 0 | 0 |
| 1 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 |

$\tilde{A}_{ij}$

$$\hat{A} = \tilde{D}^{-1} \tilde{A} \tilde{D}^{-1}$$

이 같은 계산은 사실상 정규화를 두 번 하는 것이기 때문에 아래 같이 수정함으로써 rebalance 해주기

$$\hat{A} = D^{-0.5} \tilde{A} D^{-0.5}$$

Normalized Adjacency Matrix

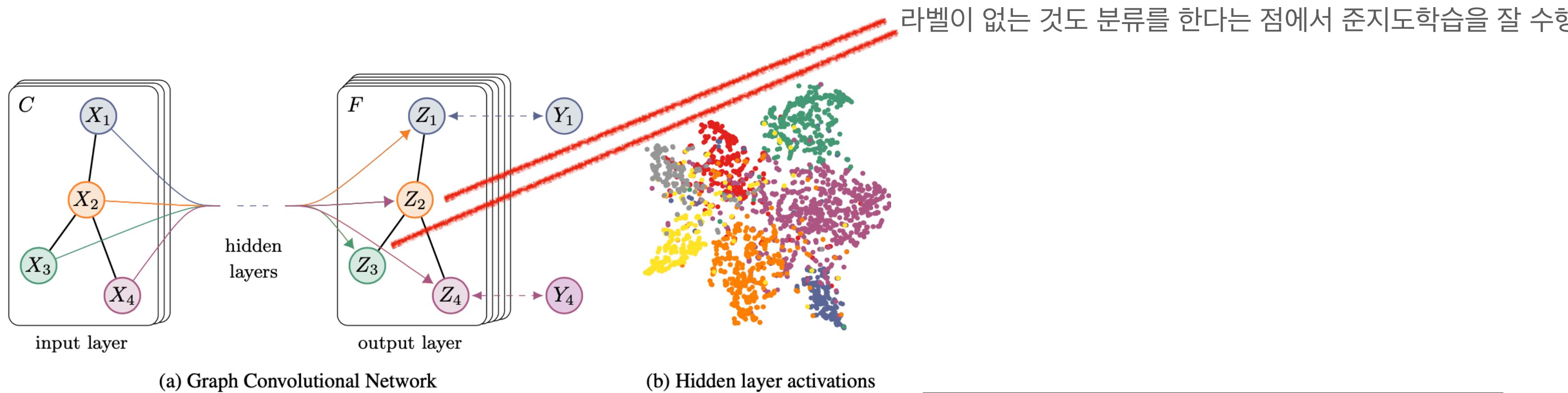
$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)})$

- 어떤 노드가 서로 연결되었는지
- 각 노드 간의 연결 강도가 얼마인지

- $= A$
- H: node feature matrix
  - W: weight

# 4. Graph Convolutional Networks

we consider a two-layer GCN for semi-supervised node classification on a graph with a symmetric adjacency matrix A



$$\hat{A} = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}}$$

$$Z = f(X, A) = \text{softmax}\left(\hat{A} \underbrace{\text{ReLU}\left(\hat{A} X W^{(0)}\right)}_{\text{Layer 1}} W^{(1)}\right)$$

Layer 1  
Layer 2

$$\mathcal{L} = - \sum_{l \in \mathcal{Y}_L} \sum_{f=1}^F Y_{lf} \ln Z_{lf},$$

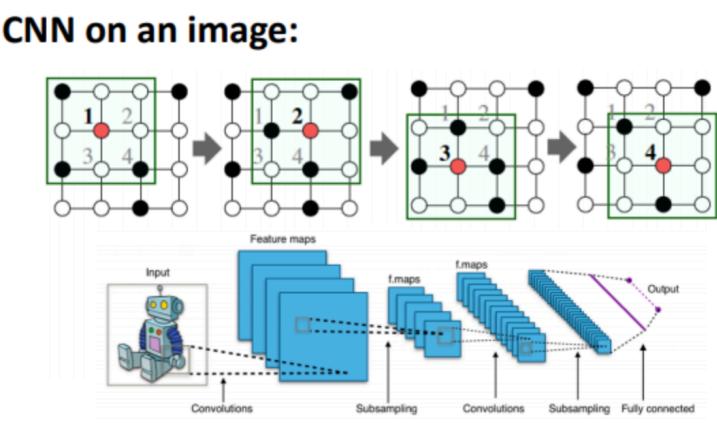
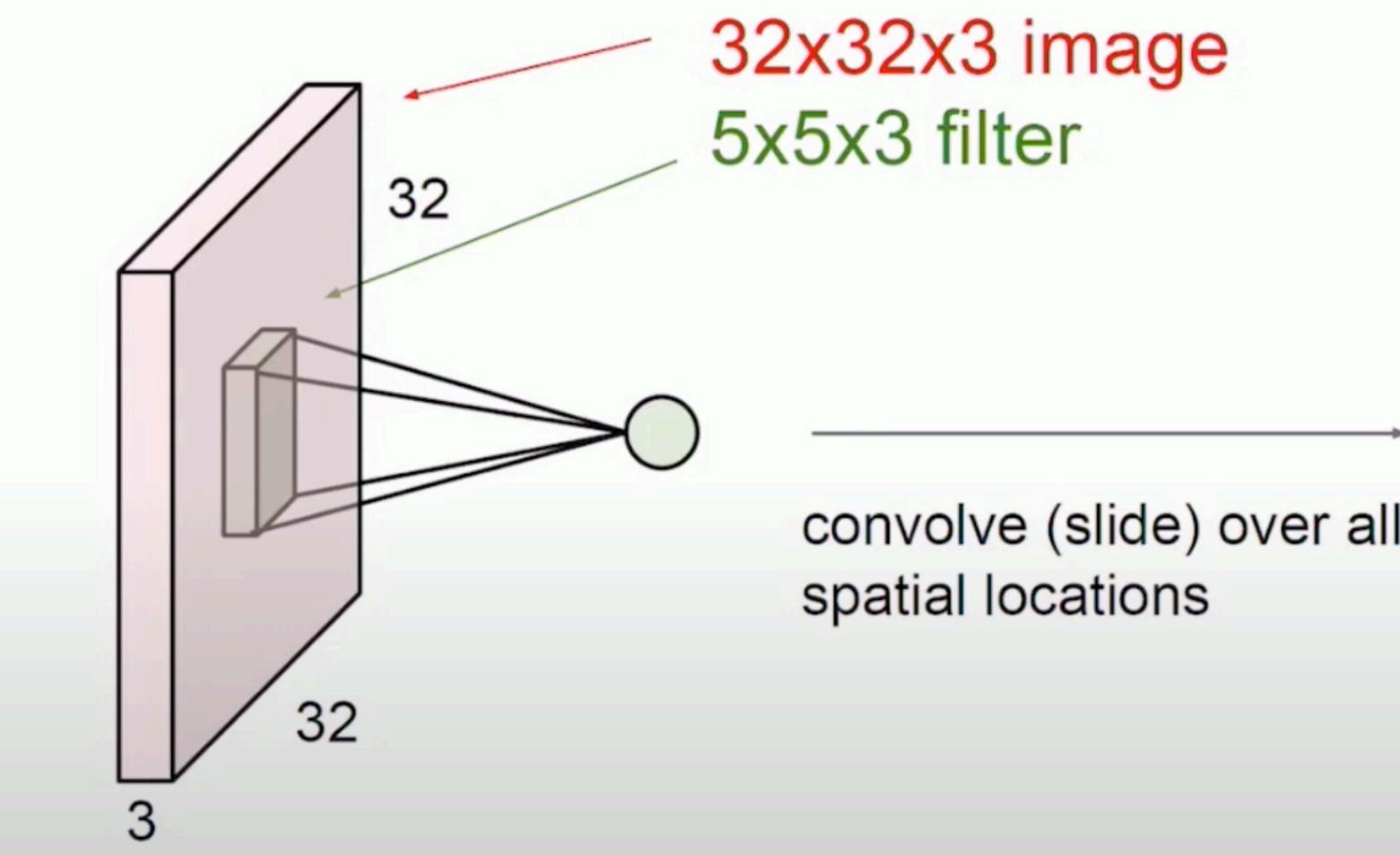
so that have labels

Y: 실제 값, Z: 예측 값: Y,Z의 크로스엔트로피

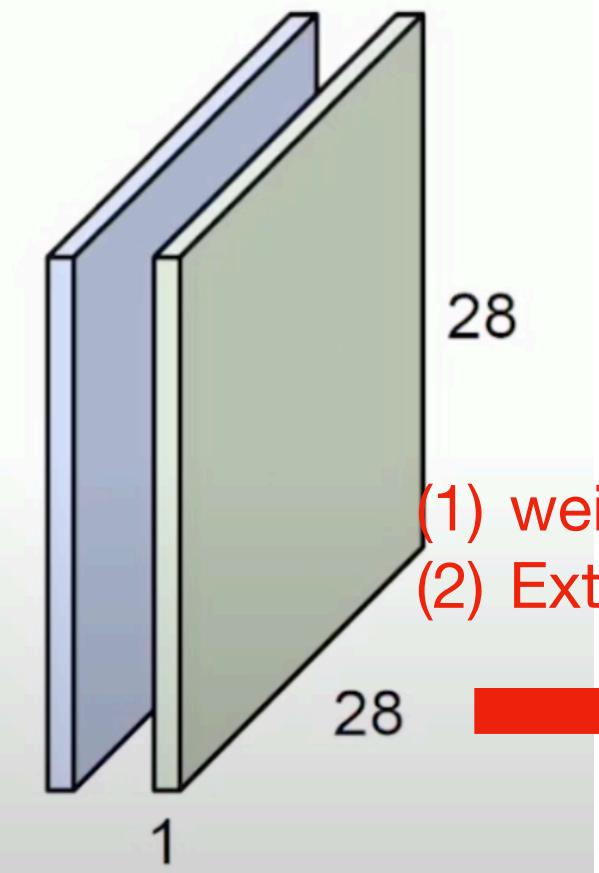
# 4. Graph Convolutional Networks

## 4.1 Convolution

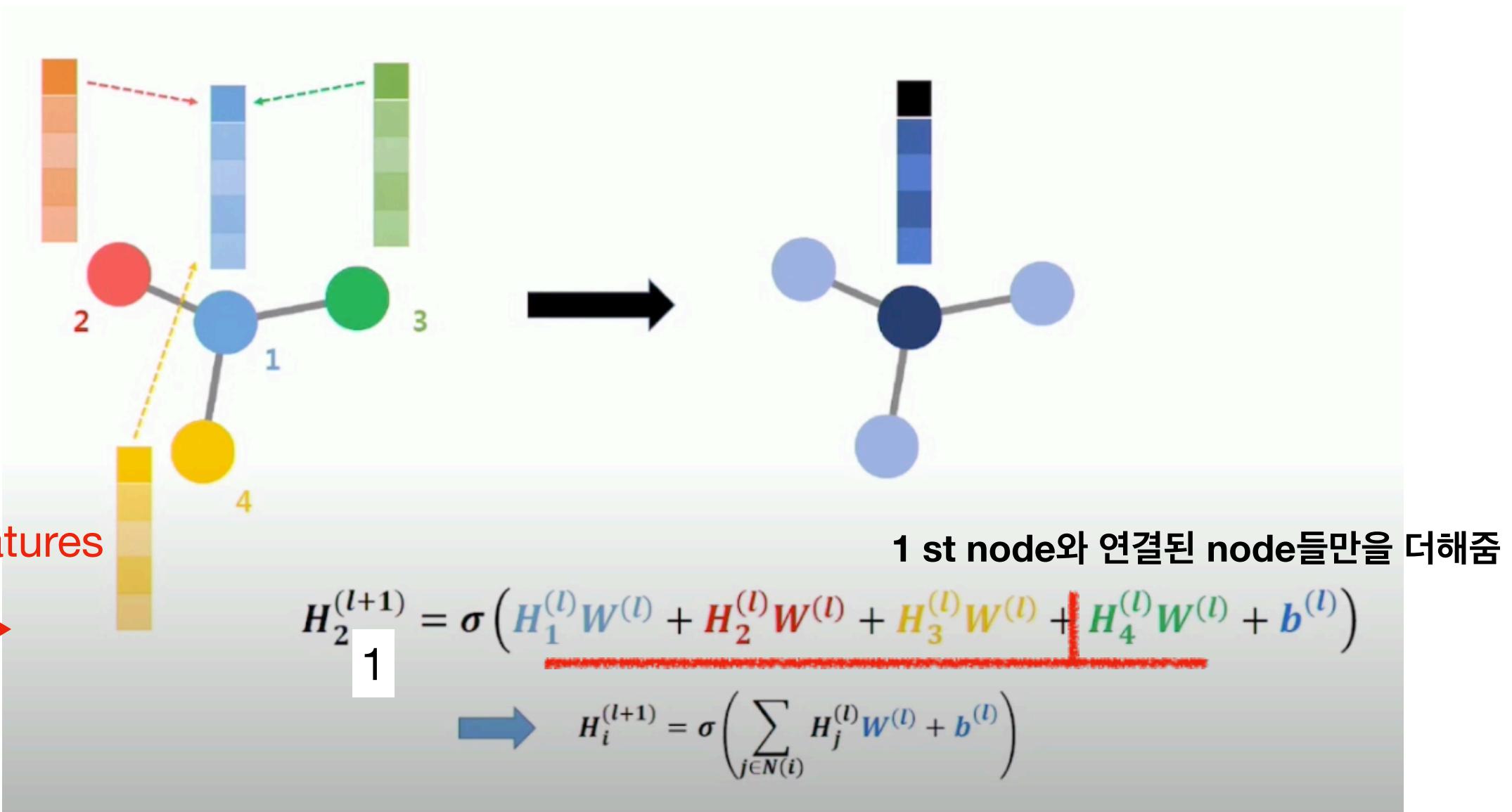
### Image case



### activation maps



### Graph case



- Filter: 한 층 당 25 개의 파라미터. 3 개의 층이니까 75 개 파라미터가 있다
  - 저 Filter는 정해진 stride 만큼 건너 뛰면서 전체 image를 조사
  - 몇 개의 필터를 사용했는가에 따라 activation map의 depth 결정된다
- **UPDATE:** Conv를 하나 거칠 때마다 (이미지의 특성을 담고 있는) activation map의 값이 업데이트

- **W:** Weight
- **H:** node feature matrix
- $H_n$ : n th node의 feature 값들
- $l$ : l th layer

- **UPDATE:** Conv를 하나 거칠 때마다 (노드의 특성을 담고 있는) activation map의 값이 업데이트

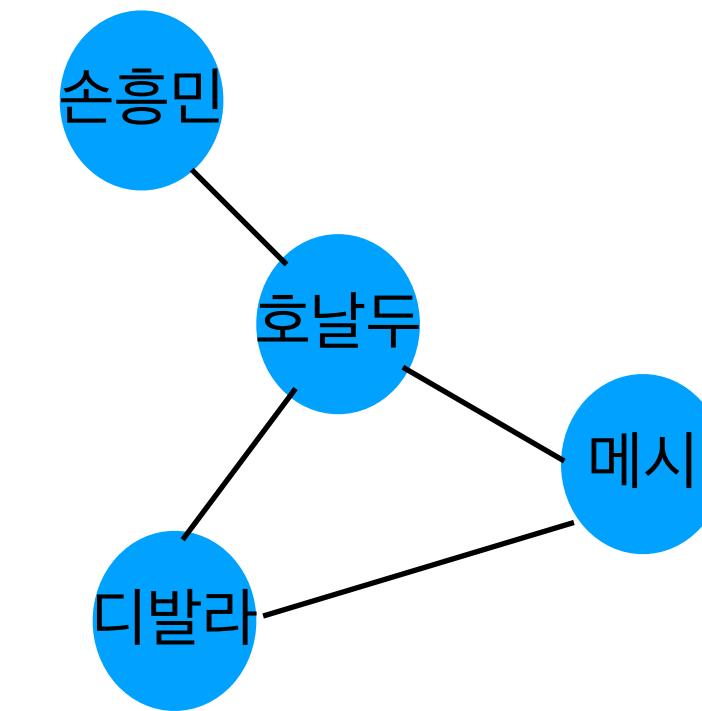
# 4. Graph Convolutional Networks

## 4.2 Compact Forms

$$H_i^{(l+1)} = \sigma \left( \sum_{j \in N(i)} H_j^{(l)} W^{(l)} + b^{(l)} \right) \longrightarrow \text{Adjacent Matrix}$$

$$H^{(l+1)} = \sigma \left( AH^{(l)} W^{(l)} + b^{(l)} \right)$$

learnable parameters are shared



|         | son | Ronaldo | Messi | Dybala |
|---------|-----|---------|-------|--------|
| son     | 1/2 | 1/8     | 0     | 0      |
| Ronaldo | 1/8 | 1/16    | 1/12  | 1/12   |
| Messi   | 0   | 1/12    | 1/9   | 1/9    |
| Dybala  | 0   | 1/12    | 1/9   | 1/9    |

Normalized  
Adjacent Matrix  
(Ahat 4,4)

|         | feature 1 | Feature 2 | Feature 3 | feature 4 | Feature 5 |
|---------|-----------|-----------|-----------|-----------|-----------|
| son     | 1         | 1         | 1         | 0         | 1         |
| Ronaldo | 1         | 0         | 1         | 1         | 0         |
| Messi   | 0         | 1         | 1         | 1         | 1         |
| Dybala  | 0         | 1         | 1         | 1         | 1         |

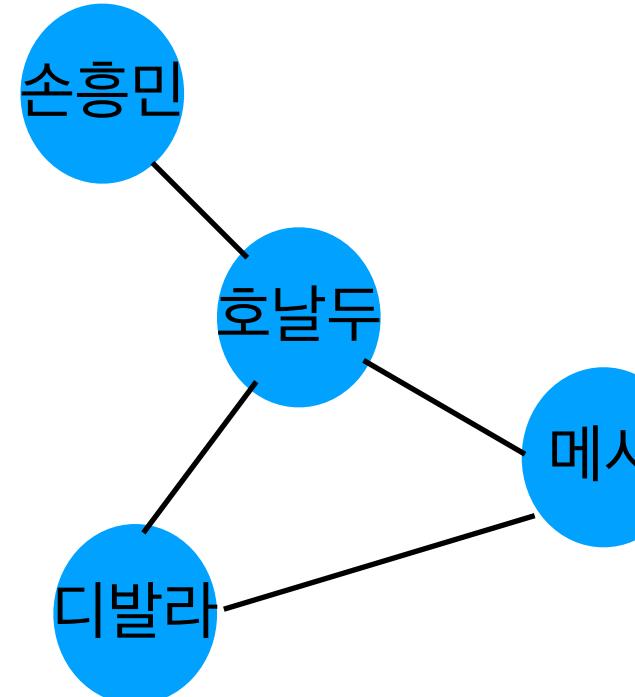
Node feature Matrix  
(X 4,5)

|          | Filter1 | Filter2 | ... | Filter64 |
|----------|---------|---------|-----|----------|
| feature1 | 1       | 1       | 1   | 0        |
| Feature2 | 1       | 0       | 1   | 1        |
| Feature3 | 0       | 1       | 1   | 1        |
| Feature4 | 0       | 1       | 1   | 1        |
| Feature5 | 1       | 1       | 0   | 1        |

Weight  
(W 5,64)

# 4. Graph Convolutional Networks

## 4.2 Compact Forms



$\sigma$

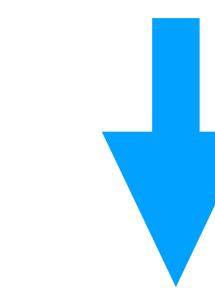
|         | son | Ronaldo | Messi | Dybala |
|---------|-----|---------|-------|--------|
| son     | 1/2 | 1/8     | 0     | 0      |
| Ronaldo | 1/8 | 1/16    | 1/12  | 1/12   |
| Messi   | 0   | 1/12    | 1/9   | 1/9    |
| Dybala  | 0   | 1/12    | 1/9   | 1/9    |

Ahat(4,4)

|         | feature 1 | Feature 2 | Feature 3 | feature 4 | Feature 5 |   | Filter1 | Filter2 | ... | Filter64 |
|---------|-----------|-----------|-----------|-----------|-----------|---|---------|---------|-----|----------|
| son     | 1         | 1         | 1         | 0         | 1         | * | 1       | 1       | 1   | 0        |
| Ronaldo | 1         | 0         | 1         | 1         | 0         |   | 1       | 0       | 1   | 1        |
| Messi   | 0         | 1         | 1         | 1         | 1         |   | 0       | 1       | 1   | 1        |
| Dybala  | 0         | 1         | 1         | 1         | 1         |   | 0       | 1       | 1   | 1        |

Node feature Matrix  
(X 4,5)

Weight  
(W 5,64)



|         | Filter1 | Filter2 | ... | Filter64 |
|---------|---------|---------|-----|----------|
| son     | 1       | 1       | 1   | 1        |
| Ronaldo | 1       | 1       | 0   | 1        |
| Messi   | 0       | 1       | 0   | 1        |
| Dybala  | 0       | 1       | 0   | 1        |

XW (4,64)

각 노드의 고차원 정보를 담은 행렬

$$H^{(l+1)} = \sigma(AH^{(l)}W^{(l)} + b^{(l)})$$

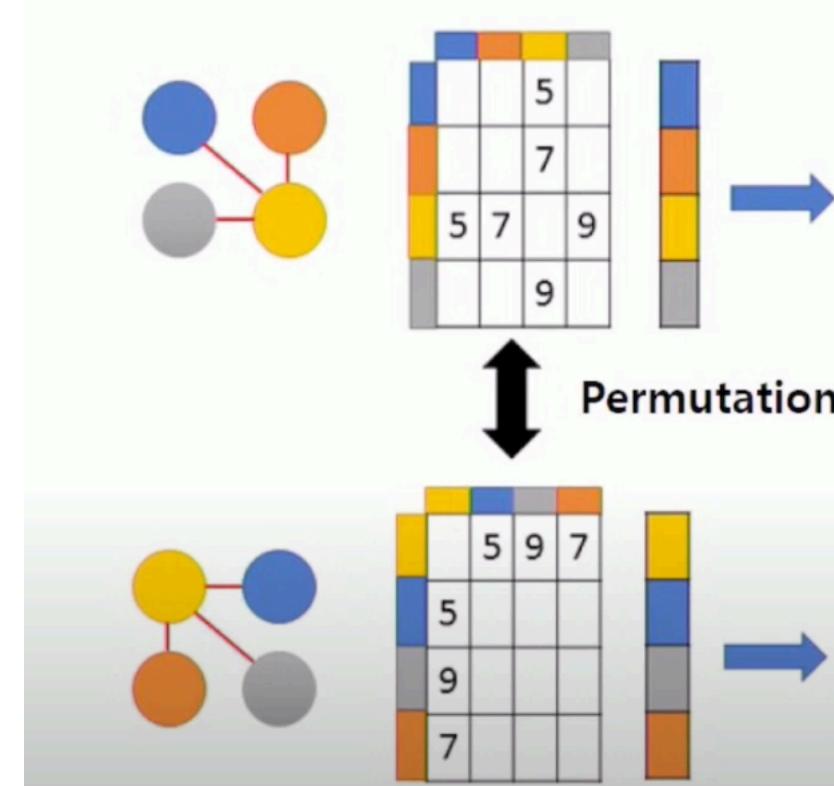
learnable parameters are shared

# 4. Graph Convolutional Networks

4.3 Readout Layer

Permutation Invariance

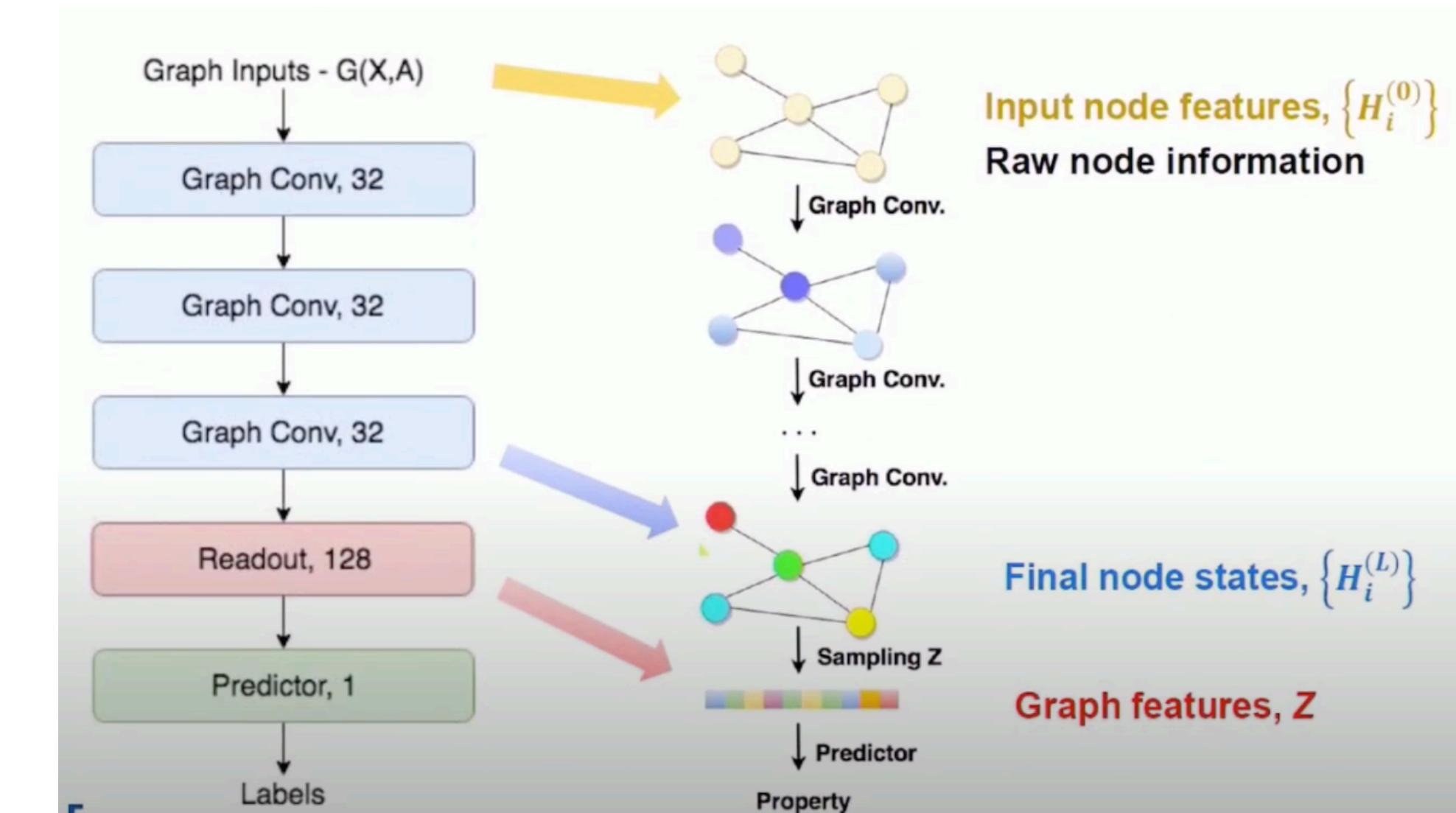
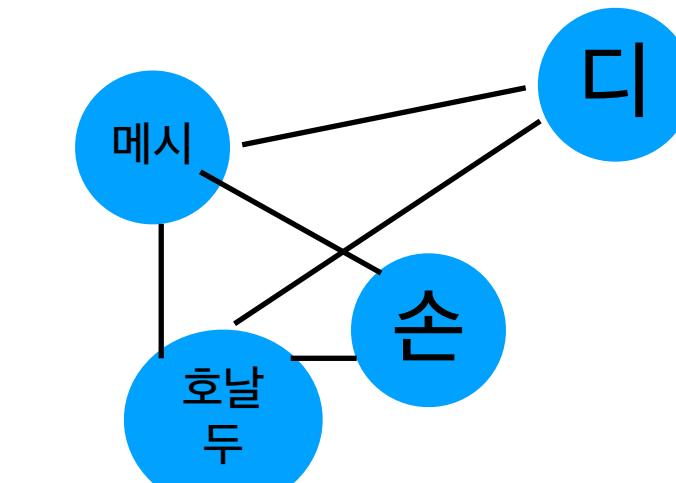
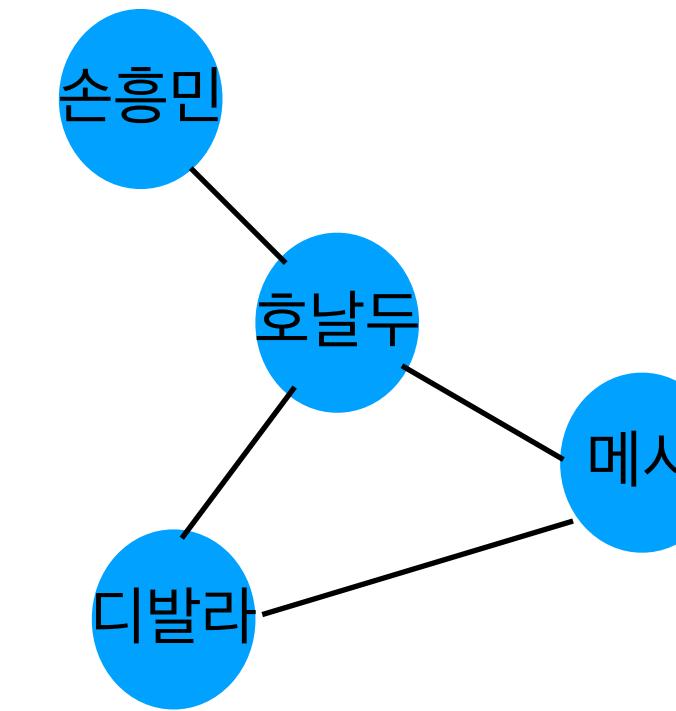
## Pytorch-geometric stellagraph



### Node-wise summation

$$z_G = \tau \left( \sum_{i \in G} \text{MLP} (H_i^{(L)}) \right)$$

- 노드의 순서가 바뀐다고 해서 그래프가 바뀌지는 않는다.
- 하지만 우리가 표현하는 **node feature matrix**는 그래프 모양이 바뀜에 따라 달라진다.
- 이것에 invariance를 부여하기 위해 readout layer를 거친다.



|         | Feature1 | Feature2 | Feature3 |
|---------|----------|----------|----------|
| son     | 0        | 1        | 0        |
| Ronaldo | 0        | 1        | 1        |
| Messi   | 1        | 1        | 1        |
| Dybala  | 1        | 1        | 1        |



|         | Feature1 | Feature2 | Feature3 |
|---------|----------|----------|----------|
| Ronaldo | 0        | 1        | 1        |
| son     | 0        | 1        | 1        |
| Messi   | 1        | 1        | 1        |
| Dybala  | 1        | 1        | 1        |

# References

- [1] cs224w lecture1, 8
- [2] [#30.Lec] Basic of Graph Convolution Network - 딥러닝 훌로서기
- [3] A Comprehensive Survey on Graph Neural Networks 2019
- [4] [https://tykimos.github.io/warehouse/2019-7-4-ISS\\_2nd\\_Deep\\_Learning\\_Conference\\_All\\_Together\\_jtkim\\_file.pdf](https://tykimos.github.io/warehouse/2019-7-4-ISS_2nd_Deep_Learning_Conference_All_Together_jtkim_file.pdf)