Graph Neural Network
and its application on Molecular Science

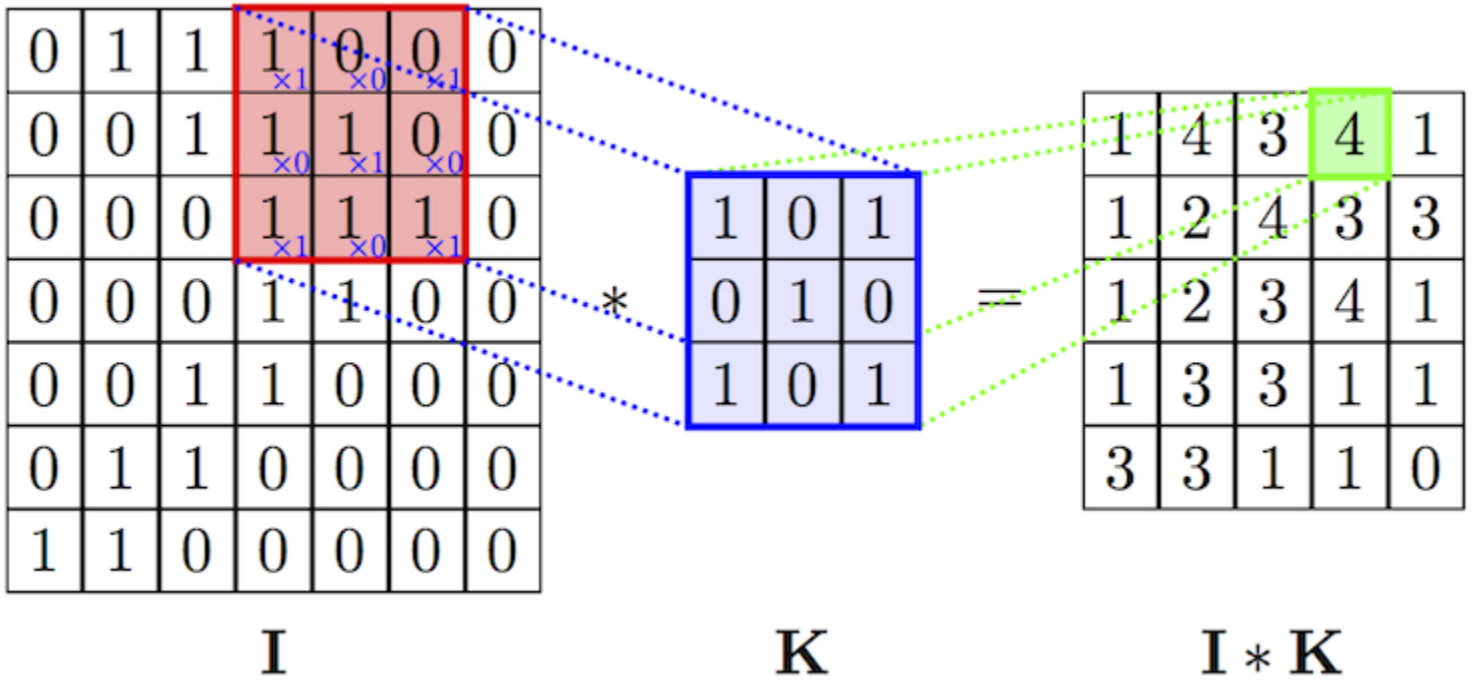
SEONGOK RYU

DEPARTMENT OF CHEMISTRY, KAIST

Introduction to GNNs

Convolutional Neural Network

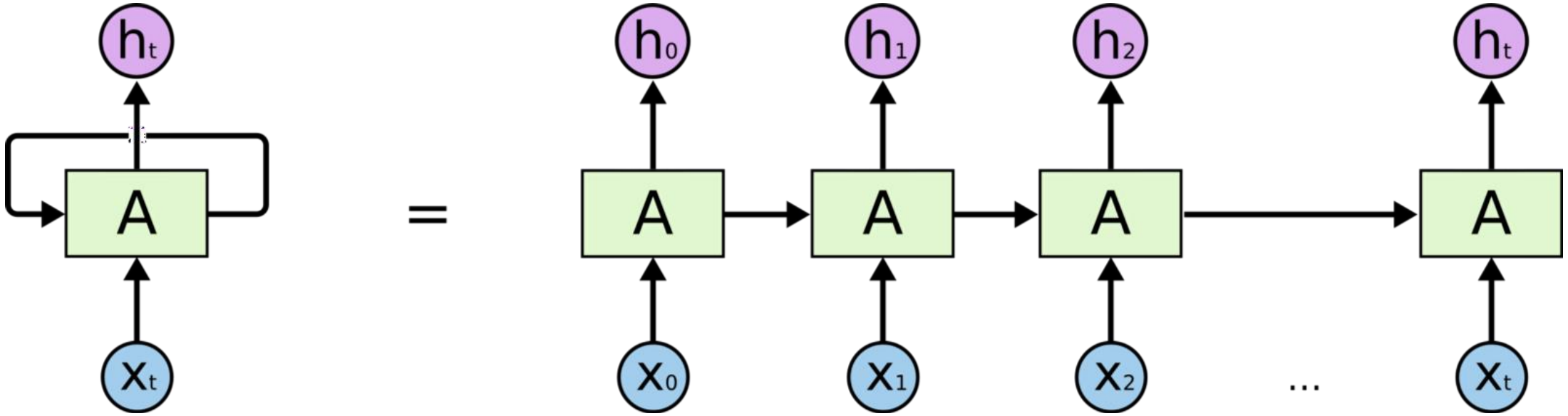
– Best solution for Vision Recognition



- Based on sampling theorem
- Works on regular grid

Recurrent Neural Network

– Best solution for Natural Language Processing

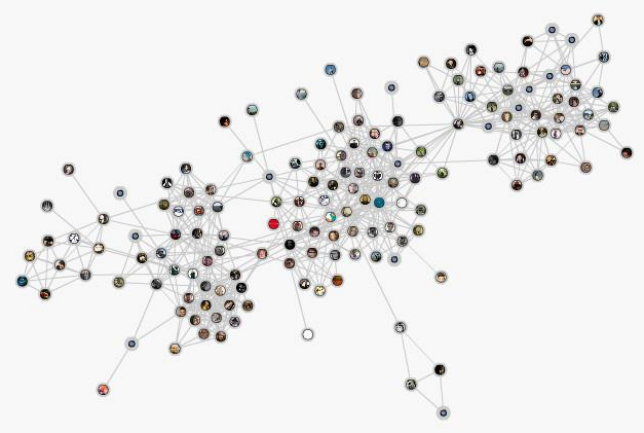


- Based on Markov process
- Suitable for finding relationship between elements in a sequence

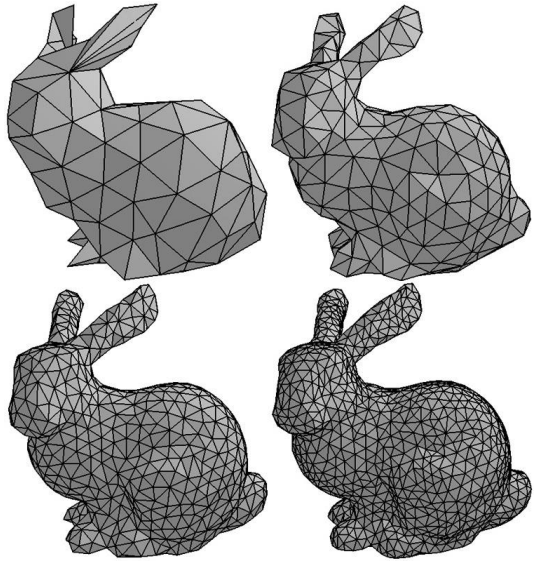
HOWEVER,

- There are other types of data as well as image, sequence, …

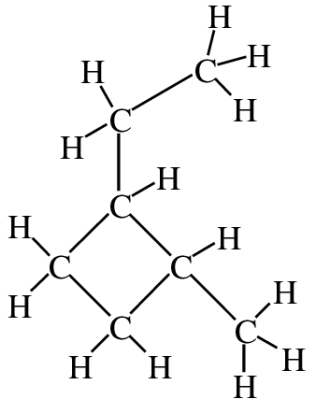
- Which have “irregular” data distribution

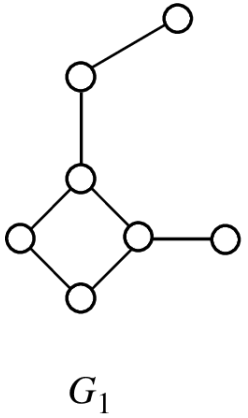


Social Graph
(Facebook, Wikipedia)



3D Mesh





Molecular Graph

All you need is GRAPH!

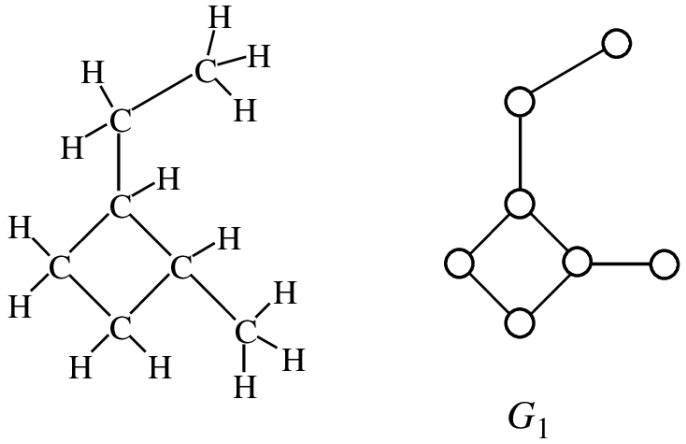
Graph Representation

$$\boldsymbol{Graph = G(X, A)}$$

X : Node, Vertex

- Individual person in a social network
- Atoms in a molecule

Represent elements of a system



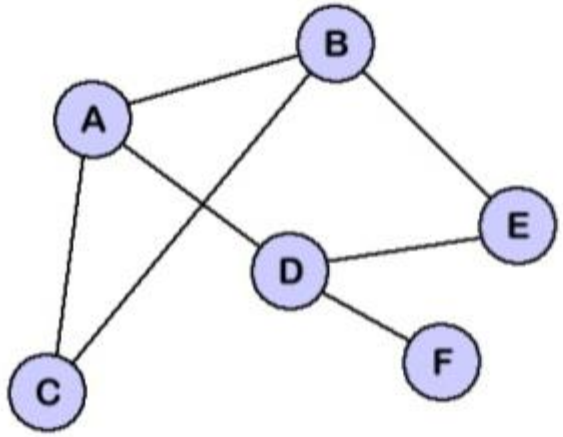
G_1

Graph Representation

$$\boldsymbol{Graph = G(X, A)}$$

A : Adjacency matrix

- Edges of a graph
- Connectivity, Relationship

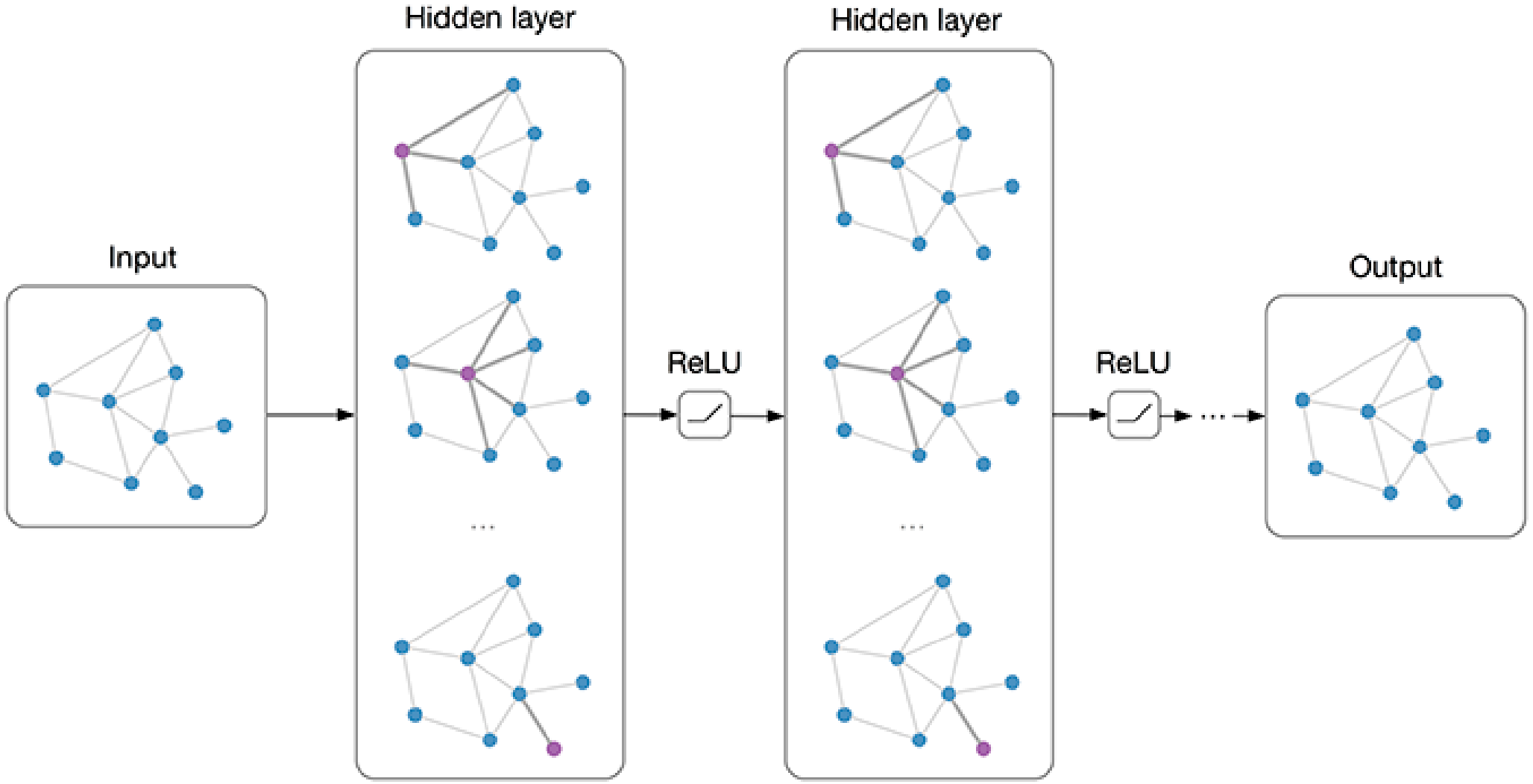


$$\begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Represent relationship, interaction between elements of the system

Graph Neural Networks

- Utilizing graphs for input of the neural networks



https://tkipf.github.io/graph-convolutional-networks/

Graph Neural Networks

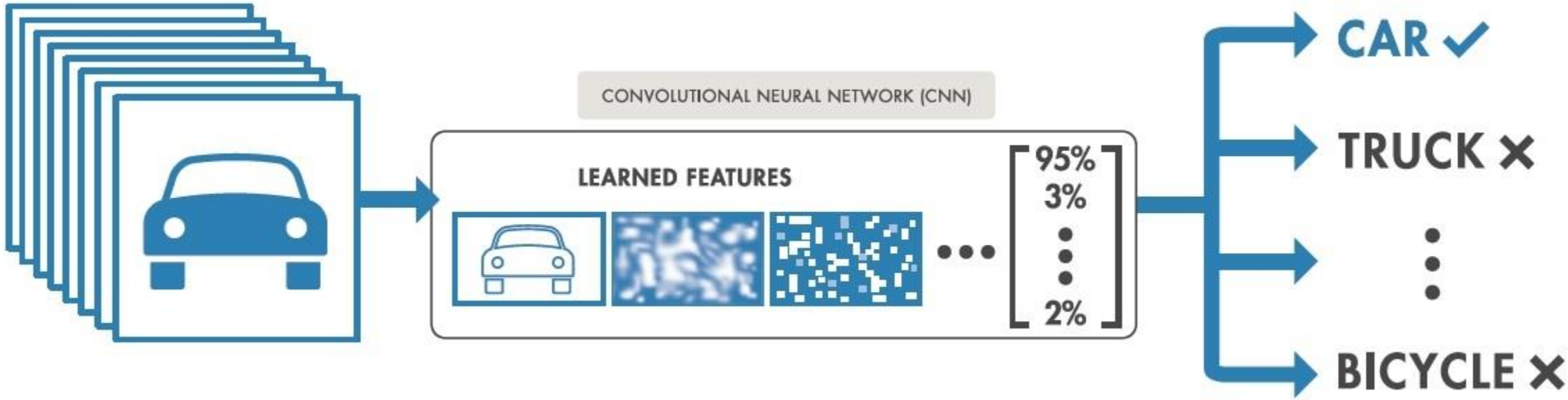
- What can we do with GNNs?

- **Node classification**
- **Link prediction**
- **Node2Vec, Subgraph2Vec, Graph2Vec**
- **Learning physics law from data**

https://tkipf.github.io/graph-convolutional-networks/

Graph Neural Networks

- Node state : Feature extracted from the Graph Neural Network

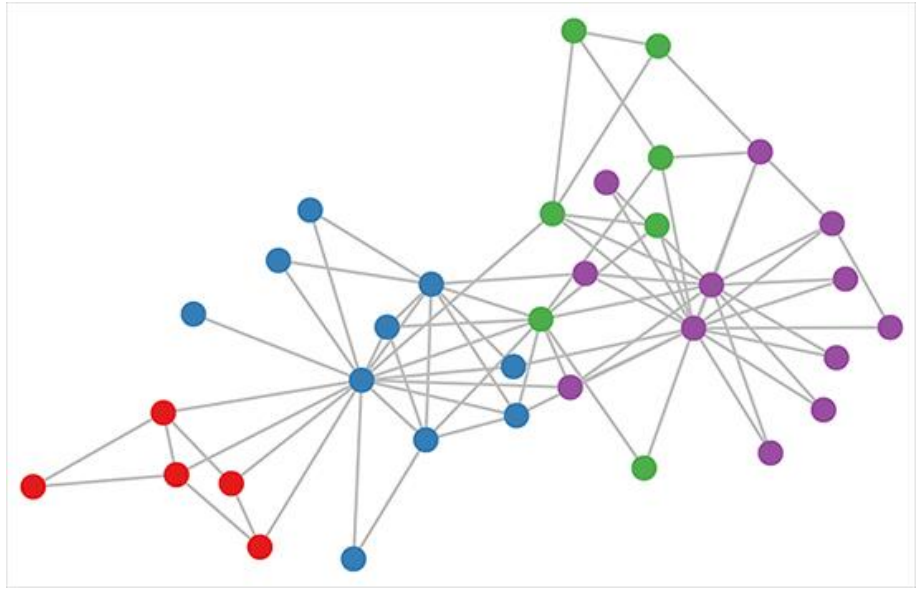


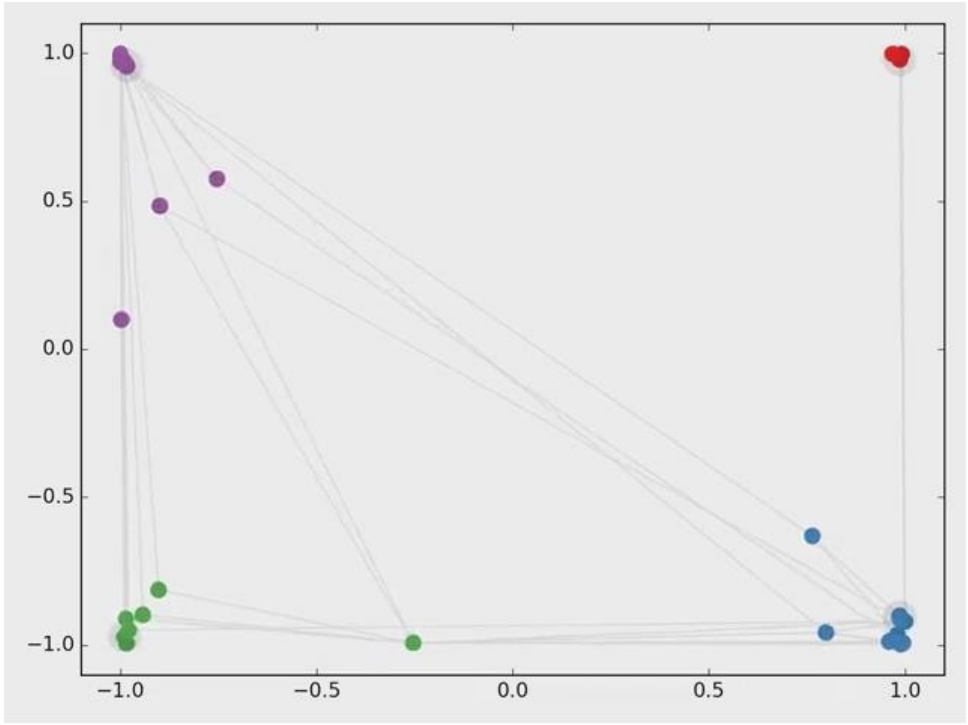
CNNs learn features from convolution operations and classify the objects

https://tkipf.github.io/graph-convolutional-networks/

Graph Neural Networks

- Node states : Feature extracted from the Graph Neural Network



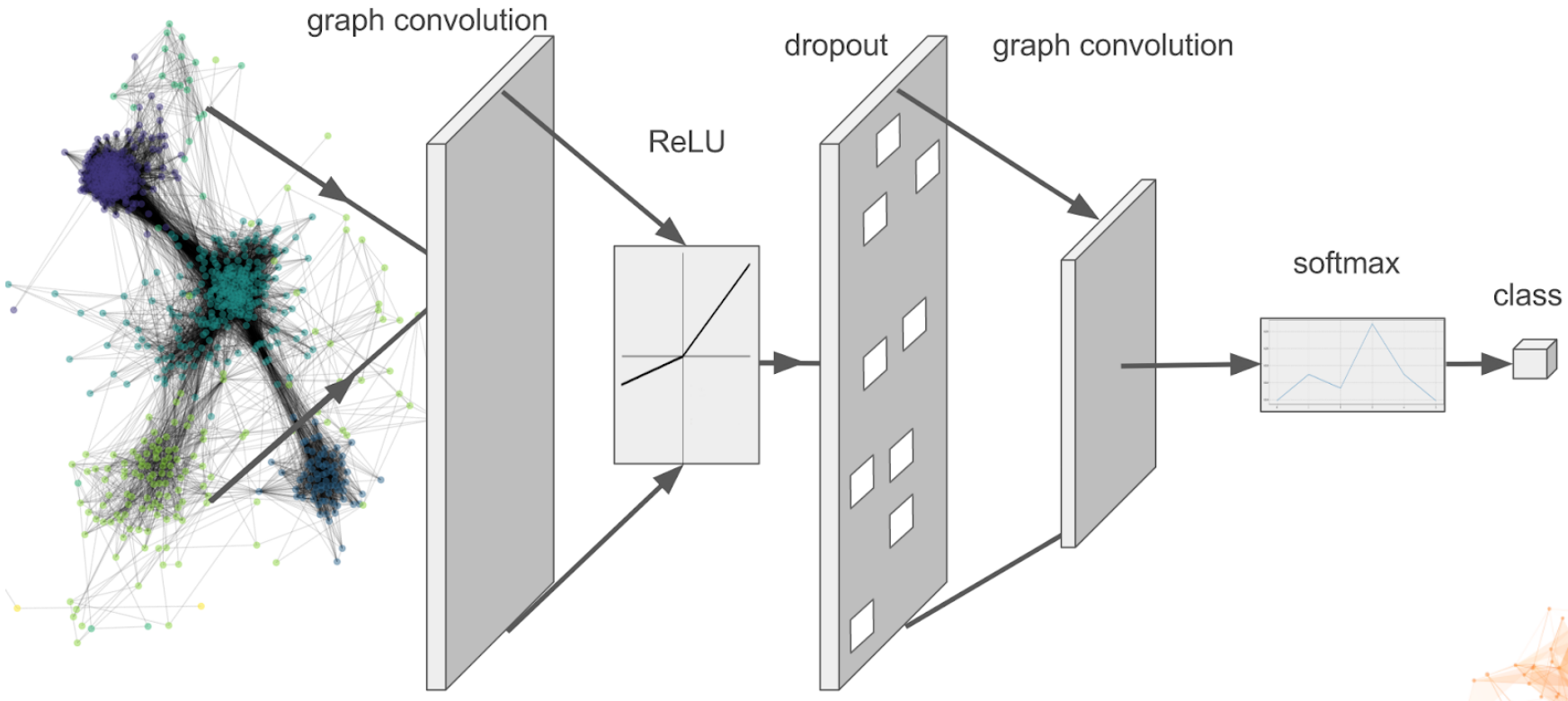


Update ‘hidden node states’, in other words, learn node features

https://tkipf.github.io/graph-convolutional-networks/

Graph Neural Networks

- Node states : Feature extracted from the Graph Neural Network



Then, we can do classification, regression,

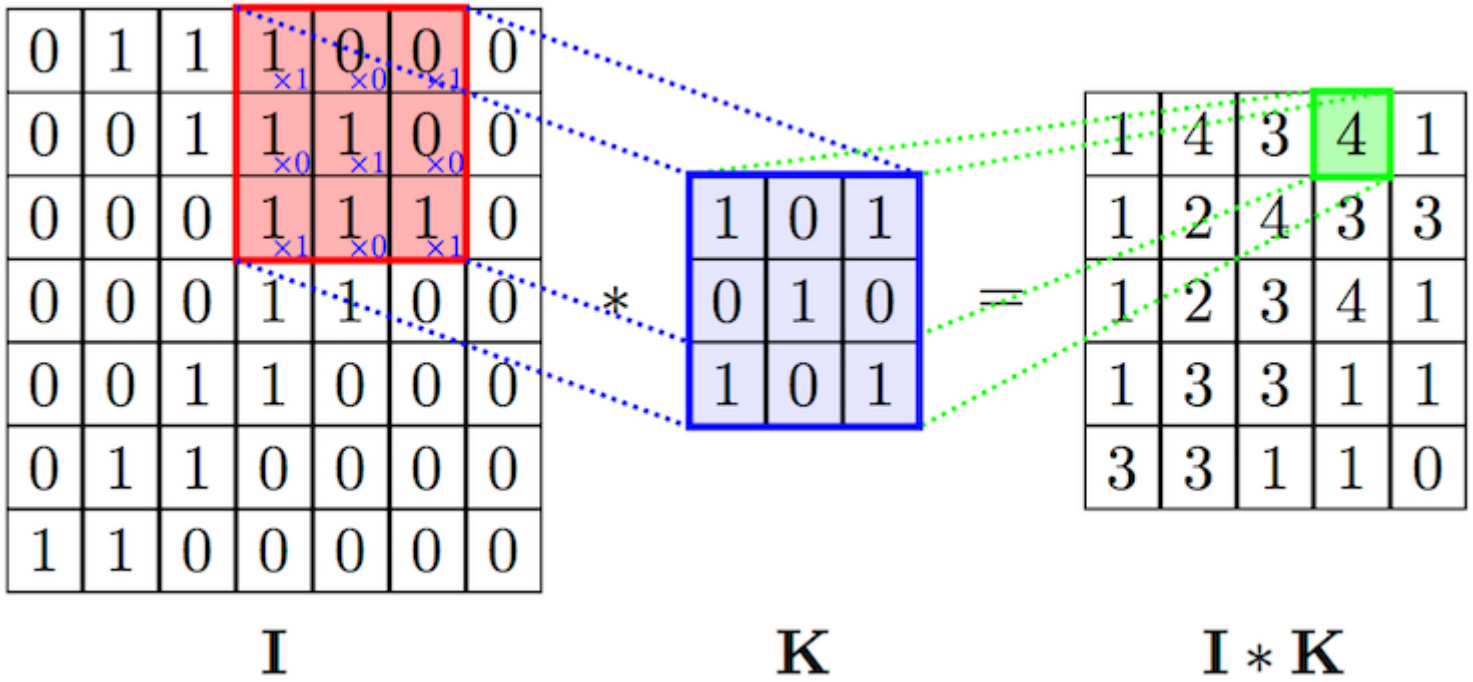
and other works based on updated hidden node states

https://www.experoinc.com/post/node-classification-by-graph-convolutional-network

Graph Convolution Network and beyond

Convolution Neural Network

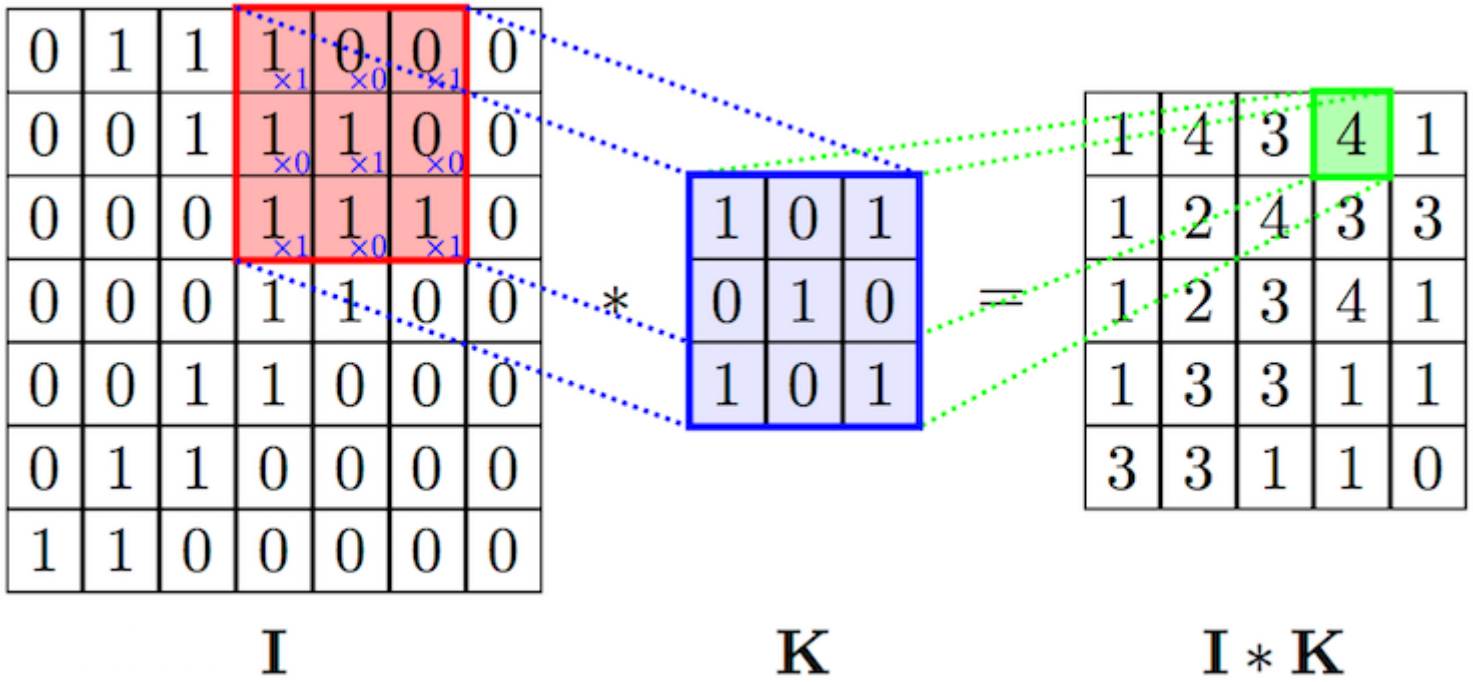
- Learning weight parameter for sampling points on a regular grid



What NN learns

Convolution Neural Network

- Learning weight parameter for sampling points on a regular grid

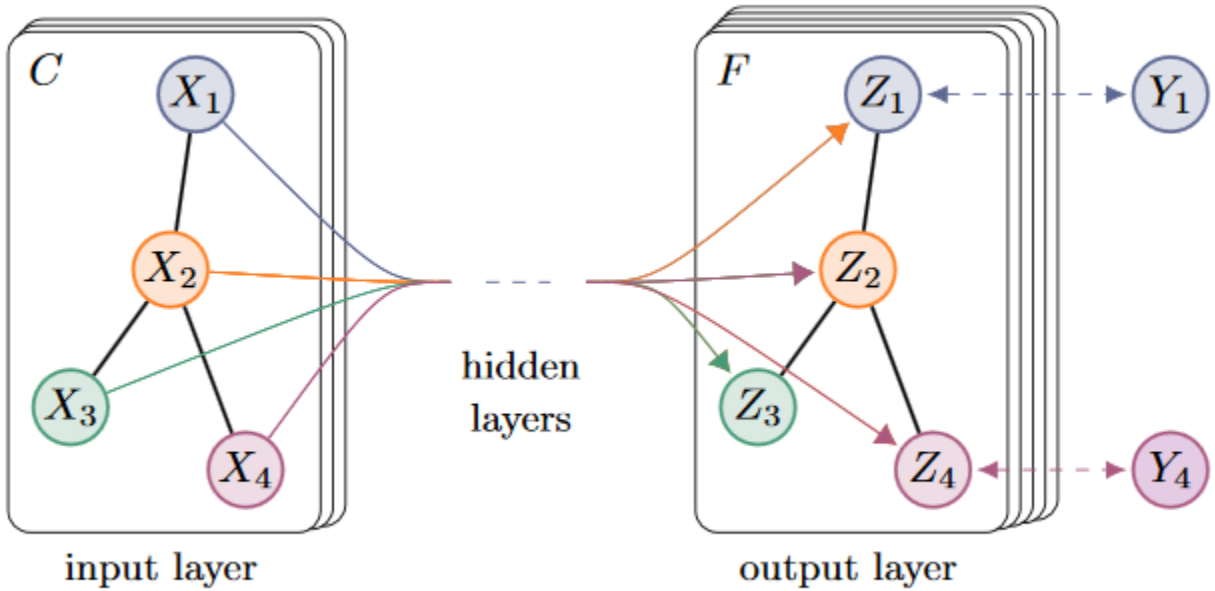


What NN learns

$$X_i^{(l+1)} = \sigma\left(\sum_{j\in[i-k,i+k]} W_j^{(l)} X_j^{(l)} + b_j^{(l)}\right)$$

Graph Convolution Network

- For Irregular representations? Like Graph structure!

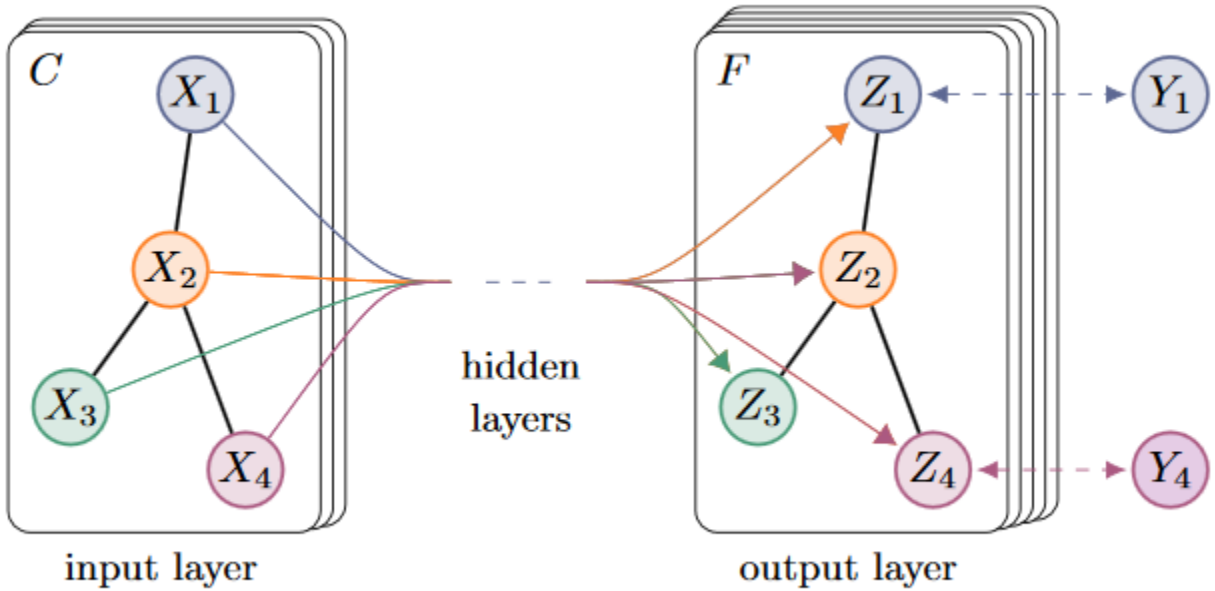


(a) Graph Convolutional Network

$$X_i^{(l+1)} = \sigma\Big(\sum_{j \in [N(i)]} W_j^{(l)} X_j^{(l)} + b_j^{(l)}\Big)$$

Graph Convolution Network

- For Irregular representations? Like Graph structure!

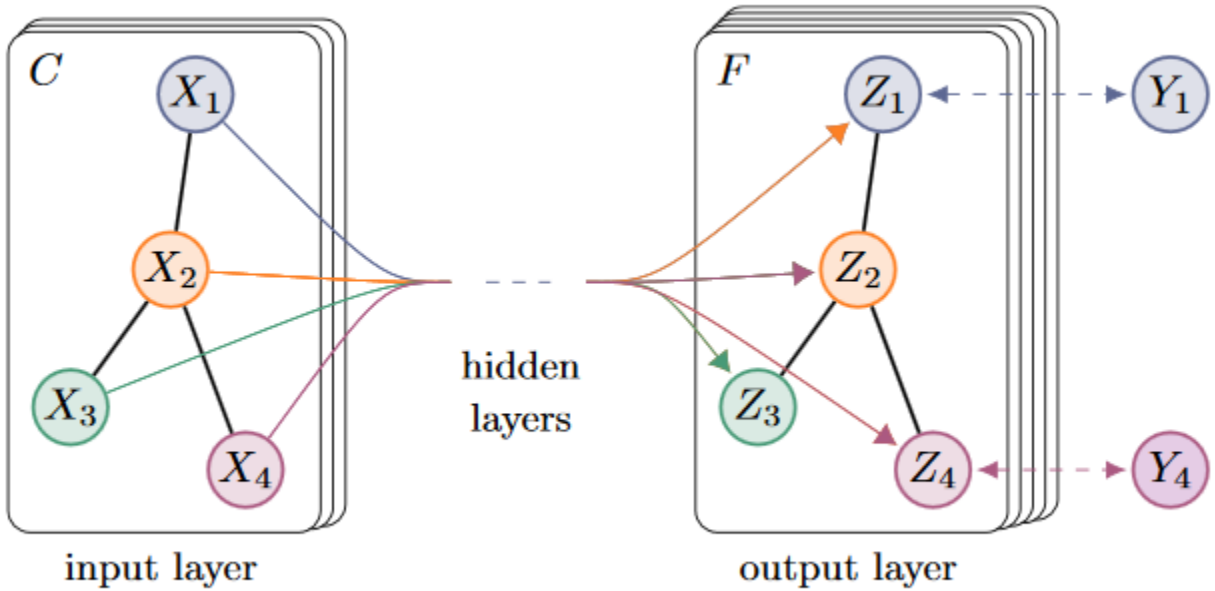


(a) Graph Convolutional Network

$$\boldsymbol{X_2^{(l+1)} = \sigma(X_1^{(l)} W^{(l)} + X_2^{(l)} W^{(l)} + X_3^{(l)} W^{(l)} + X_4^{(l)} W^{(l)})}$$

Graph Convolution Network

- For Irregular representations? Like Graph structure!



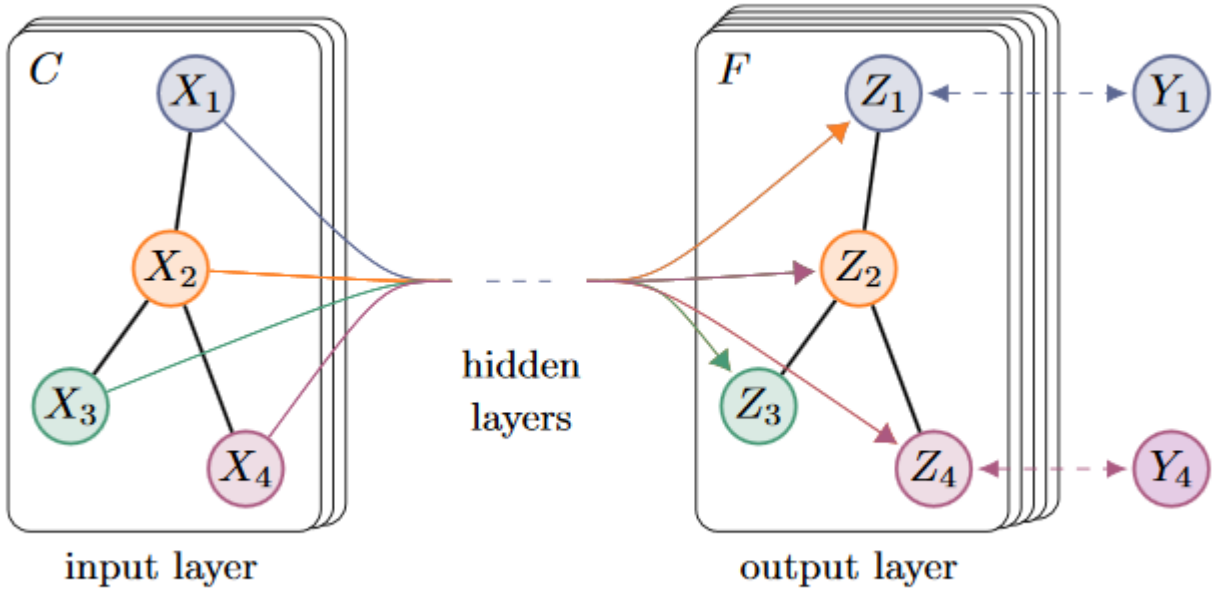
(a) Graph Convolutional Network

$$X_i^{(l+1)} = \sigma(AX^{(l)}W^{(l)})$$

A : adjacency matrix – connectivity matrix between nodes

Graph Attention Network – beyond GCN

- Attention mechanism again



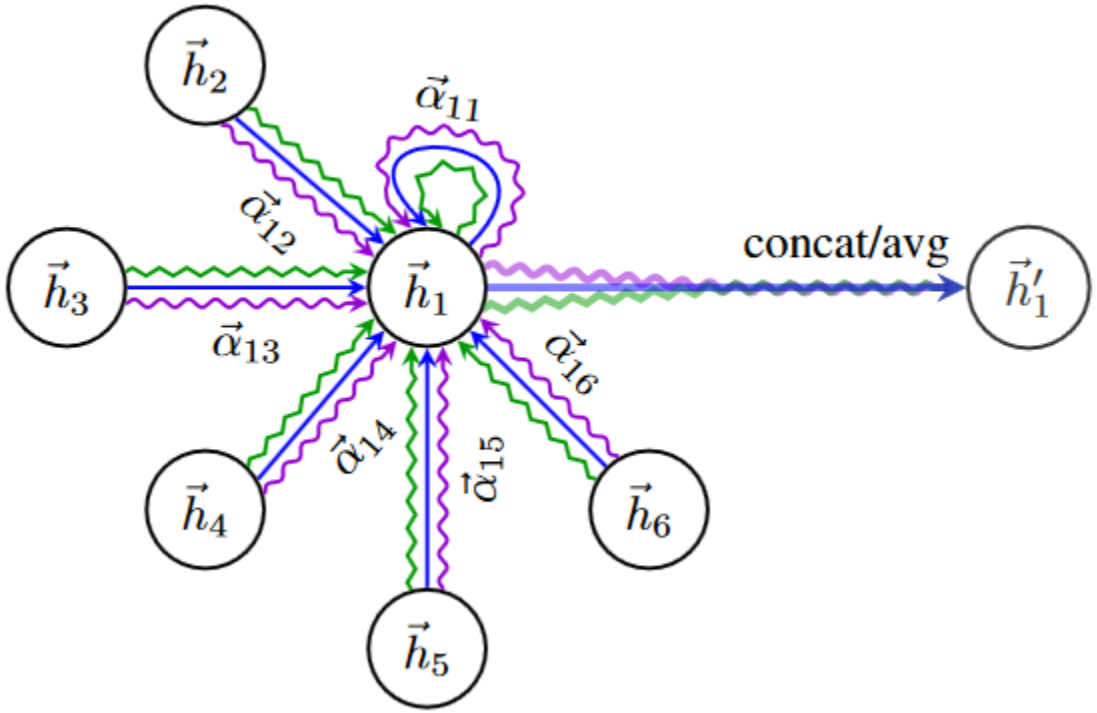
(a) Graph Convolutional Network

$$X_2^{(l+1)} = \sigma(X_1^{(l)}W^{(l)} + X_2^{(l)}W^{(l)} + X_3^{(l)}W^{(l)} + X_4^{(l)}W^{(l)})$$

- It simply sum node states with **same weights**
- However, the neighborhoods must be considered with **different importances**.
- All we need is **an attention**, again!

Graph Attention Network – beyond GCN

- Attention mechanism again



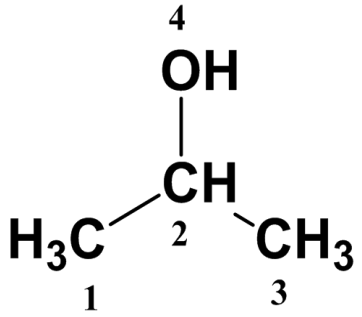
- It simply sum node states with **same weights**
- However, the neighborhoods must be considered with **different importances**.
- All we need is **an attention**, again!

$$X_i^{(l+1)} = \sigma\Big(\sum_{j \in [NN(i)]} \alpha_{ij} W_j^{(l)} X_j^{(l)} + b_j^{(l)} \Big) \qquad \alpha_{ij} = f(X_i^{(l)} W^{(l)}, X_j^{(l)} W^{(l)})$$

Applications on Molecular Science

Molecules

- A set of atoms consists the molecule



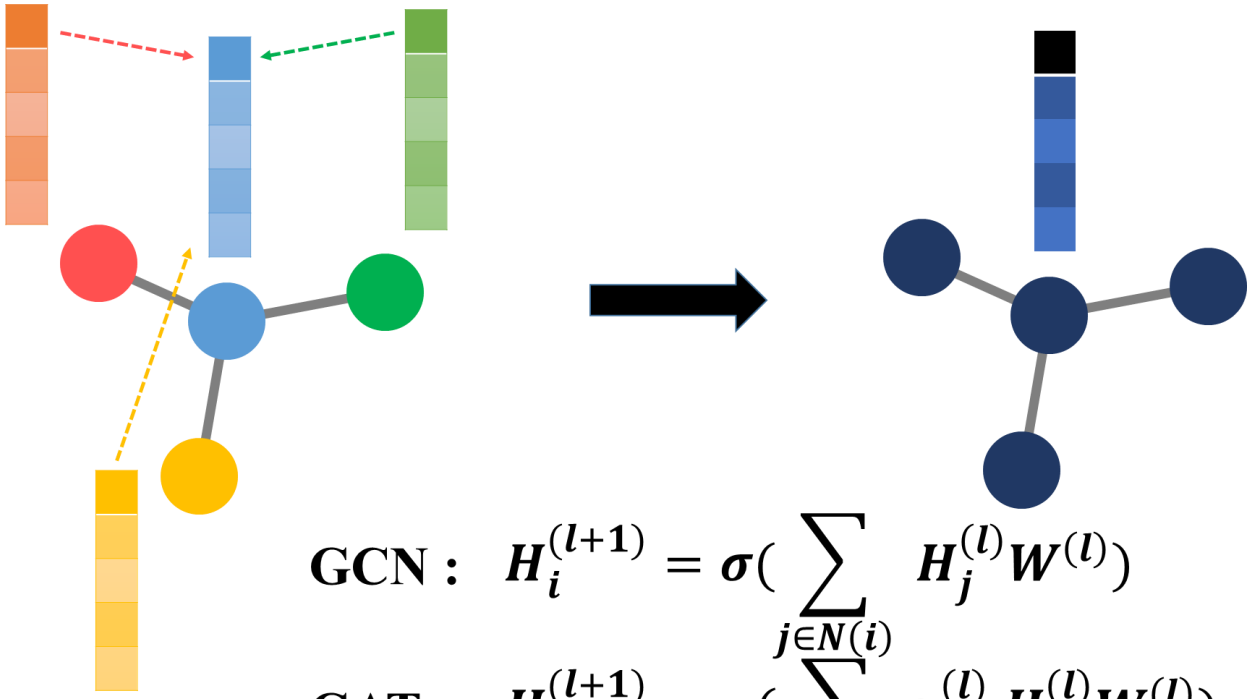
$$X_1 = \begin{bmatrix} 6 \\ 3 \\ 4 \\ 0 \end{bmatrix} \quad \cdots \quad X_4 = \begin{bmatrix} 8 \\ 1 \\ 4 \\ 0 \end{bmatrix}$$

Atom type
# of Hs.
 # Valence
Aromaticity

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

GCN and attention mechanism

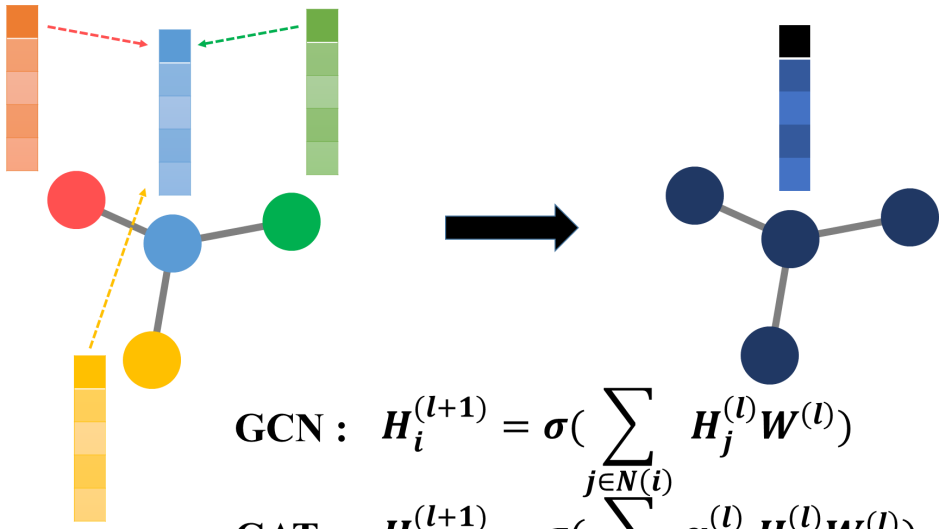
- We treat “physical interaction between atoms”, which is central principle of molecular science, using GCN and attention mechanism



$$\text{GCN}: \quad H_i^{(l+1)} = \sigma\Big(\sum_{j \in N(i)} H_j^{(l)} W^{(l)}\Big)$$

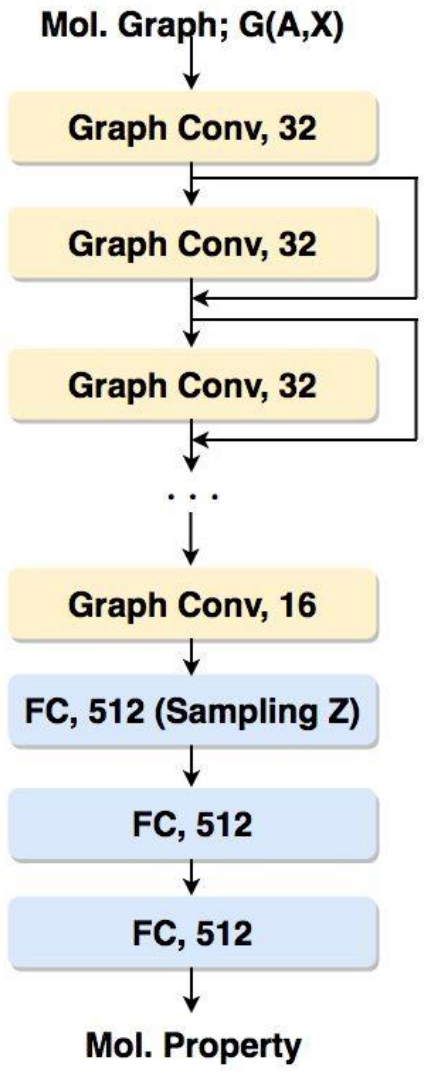
$$\text{GAT}: \quad H_i^{(l+1)} = \sigma\Big(\sum_{j \in N(i)} \alpha_{ij}^{(l)}\, H_j^{(l)} W^{(l)}\Big)$$

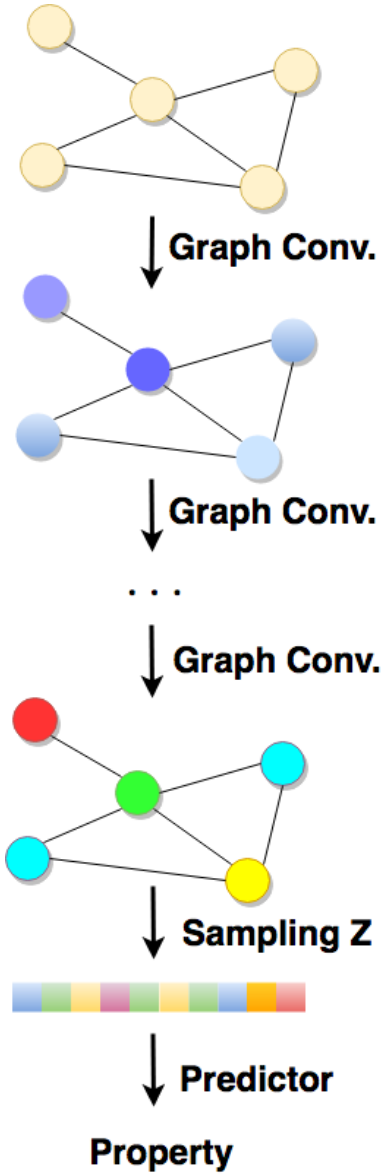
Our Neural Network for the molecular system



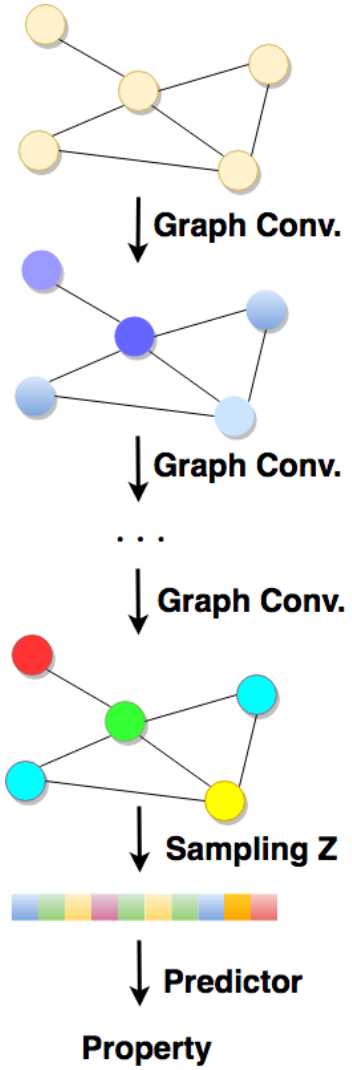
$$\text{GCN}: \quad H_i^{(l+1)} = \sigma\Big(\sum_{j \in N(i)} H_j^{(l)} W^{(l)}\Big)$$

$$\text{GAT}: \quad H_i^{(l+1)} = \sigma\Big(\sum_{j \in N(i)} \alpha_{ij}^{(l)}\, H_j^{(l)} W^{(l)}\Big)$$



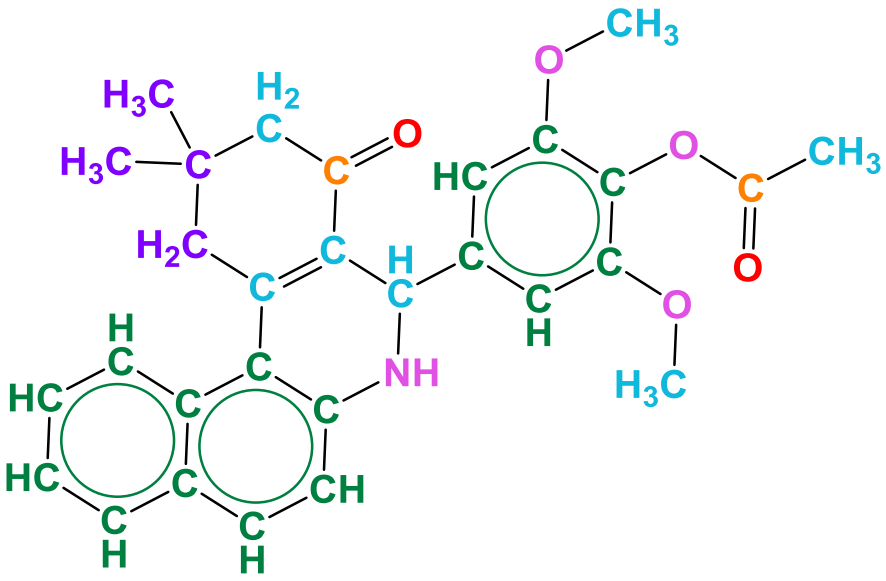


Our Neural Network for the molecular system



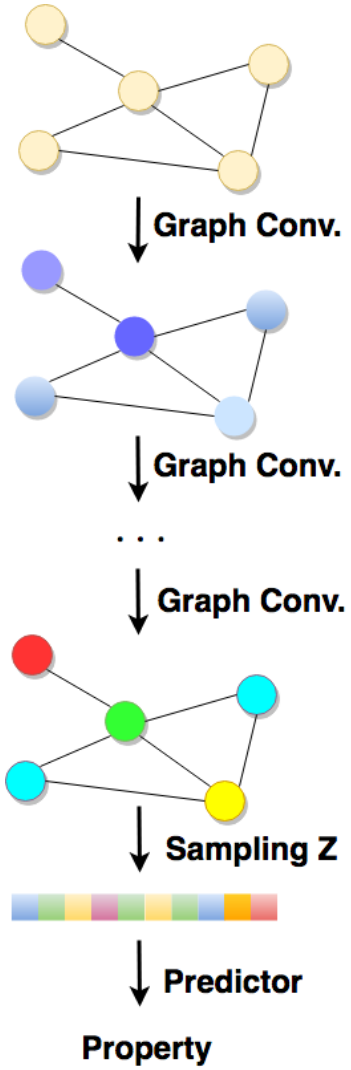
Final Node states

Learning solubility of molecules



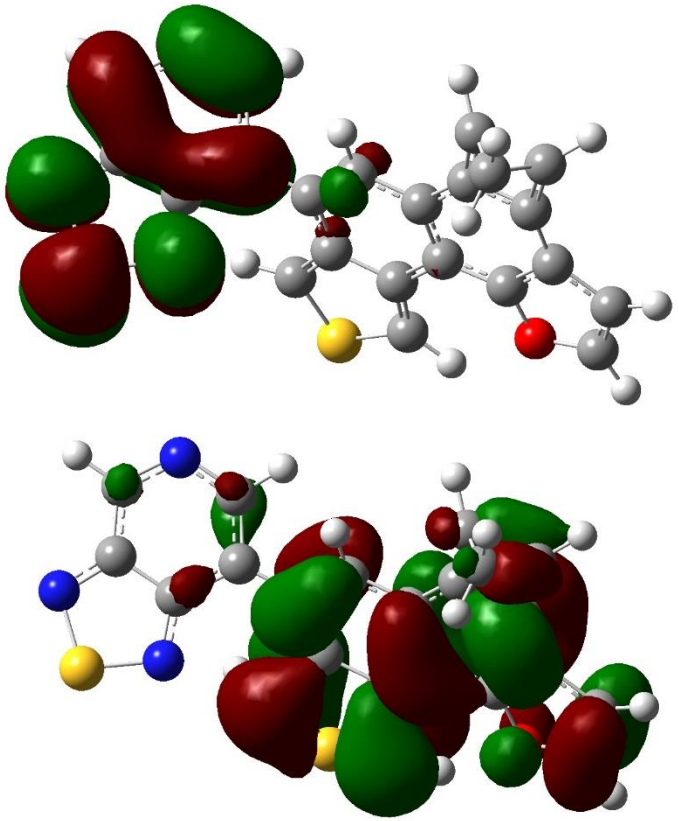
The neural network recognize
several functional groups differently

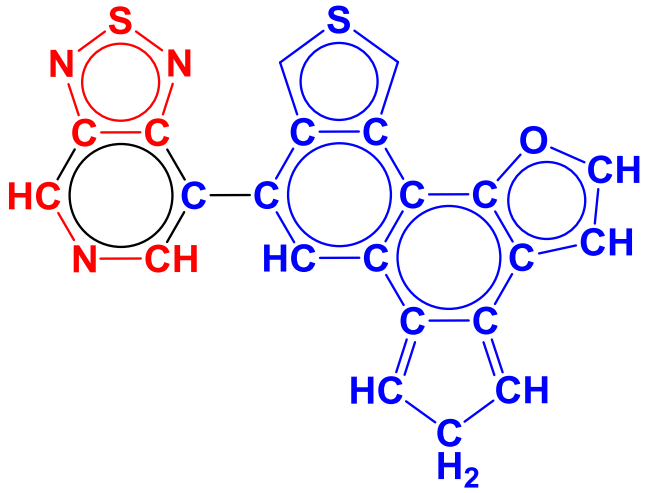
Our Neural Network for the molecular system



Final Node states

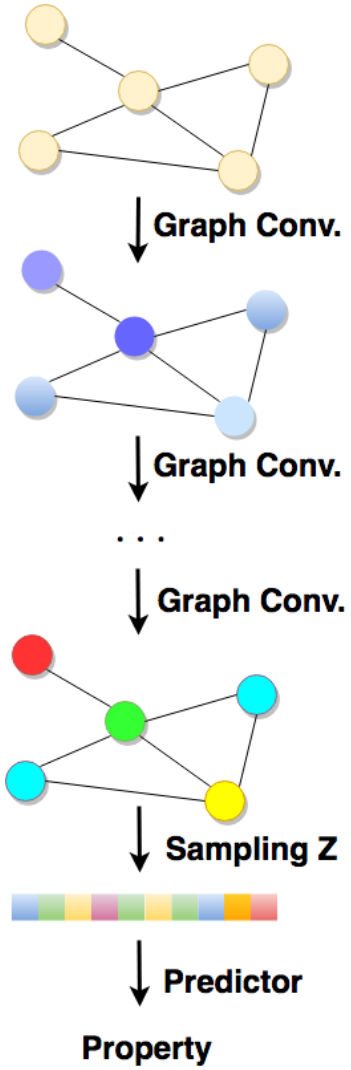
Learning photovoltaic efficiency (QM phenomena)





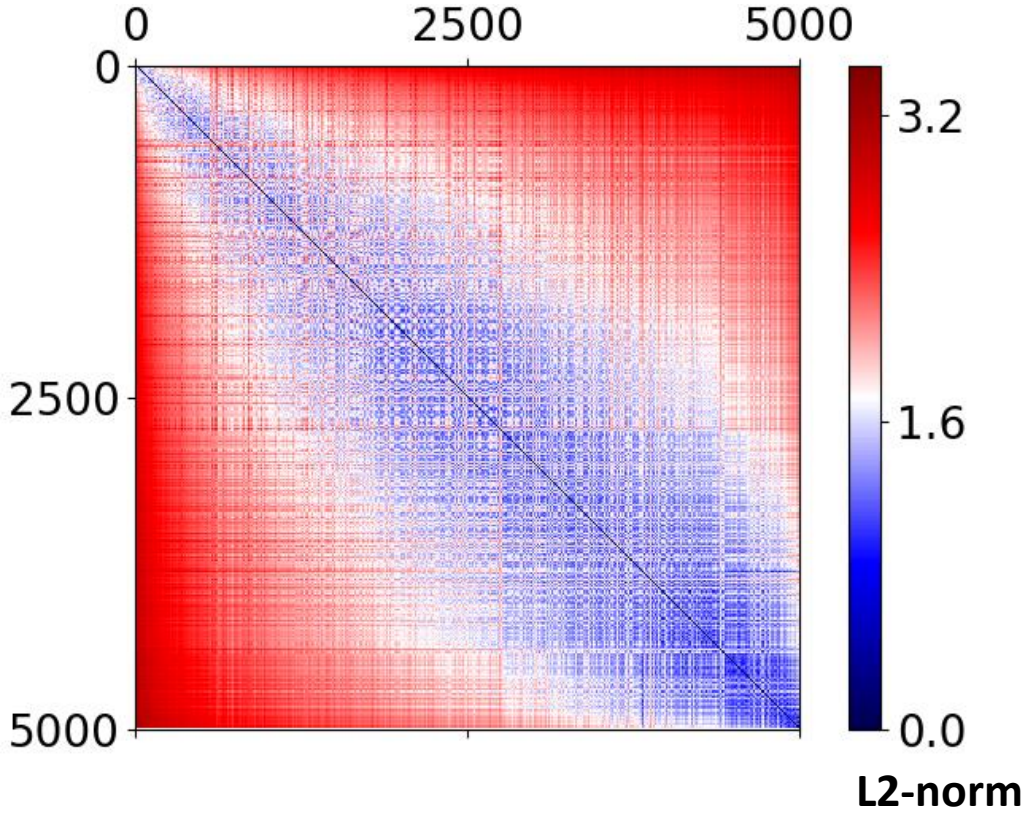
Interestingly, The NN also can differentiate nodes according to the **quantum mechanical characteristics**.

Our Neural Network for the molecular system



Graph Features

Learning solubility of molecules



Ascending order

Similar molecules are located closely in the graph latent space

Our Neural Network for the molecular system

- Molecules can be represented by graph structures.
- We can precisely predict molecular properties using **graph convolution with attention mechanism**.
- The neural network can classify atoms (nodes) ***according to the chemistry knowledge***.
- Also similar molecules are located closely in graph latent space.
- Not only ***prediction***, but also ***interpretable results*** for molecular science
- We have devised ***generative models*** for *de novo* molecular design.