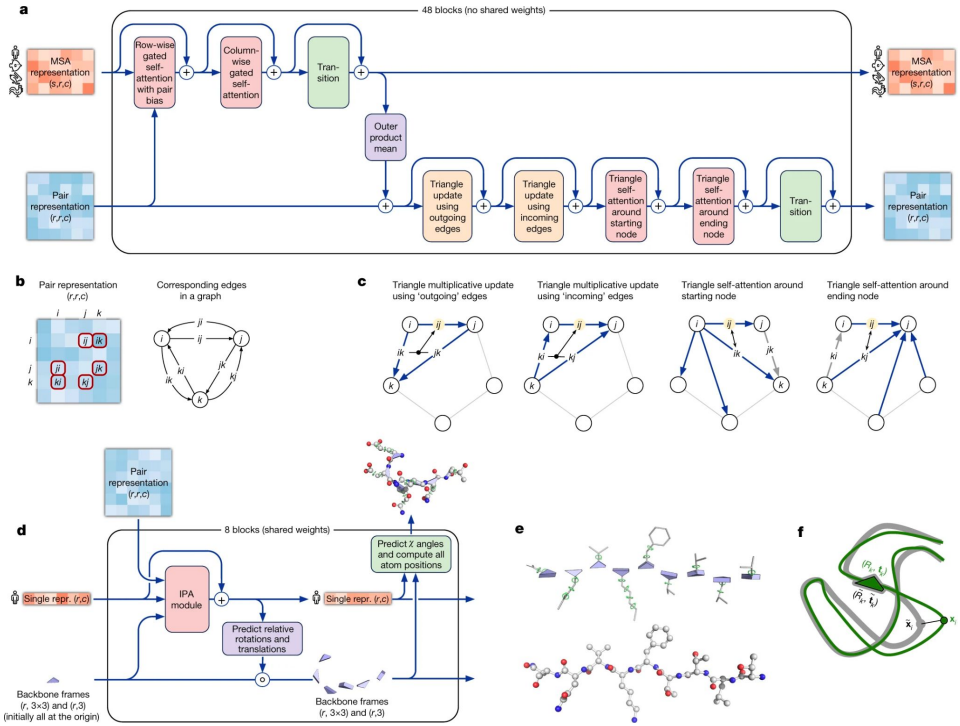
BioMedical Learning through Graph Neural Networks

Dr. Nisha Pillai
Postdoctoral Associate @ Mississippi State University

Source: @Nature, Deepmind

Graphs! Graphs!



Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589 (2021)

Today's Focus

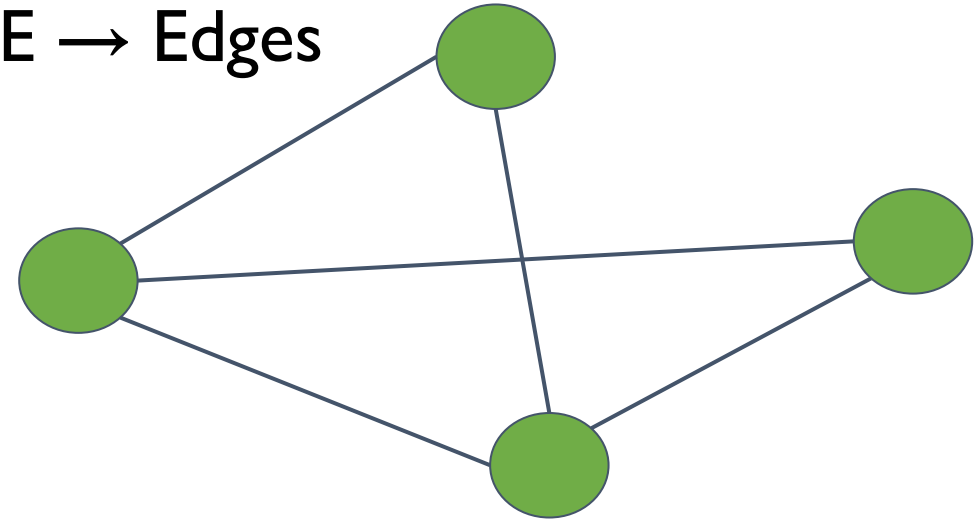
1. Terminology and representations
2. Popular architectures
3. An architecture demonstration
4. Related Work

Graphs

Real world objects, and their connections

G = {V, E}, V → Nodes, E → Edges

Directed / Undirected



Sanchez-Lengeling, et al., "A Gentle Introduction to Graph Neural Networks", Distill, 2021.

Representing Graphs

a) Edge lists → $((v_1, v_2)....(v_4, v_6))$

b) Adjacency Matrices

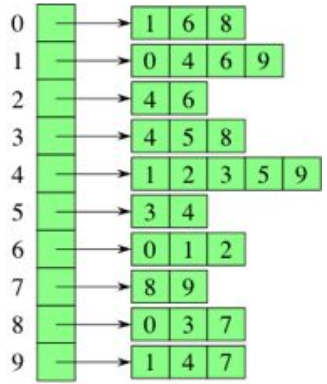
	0	1	2	3	4	5	6	7	8	9
0	0	1	0	0	0	0	1	0	1	0
1	1	0	0	0	1	0	1	0	0	1
2	0	0	0	0	1	0	1	0	0	0
3	0	0	0	0	1	1	0	0	1	0
4	0	1	1	1	0	1	0	0	0	1
5	0	0	0	1	1	0	0	0	0	0
6	1	1	1	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	1	1
8	1	0	0	1	0	0	0	1	0	0
9	0	1	0	0	1	0	0	1	0	0

[[0 1 0 0 0..]

.....

[0 1 0 0..]]

c) Adjacency Lists



[[1,6,8]

.....

[1,4,7]]

Source: Khan Academy

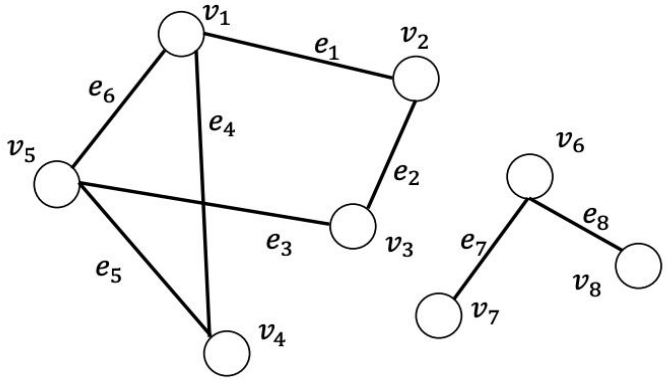
Properties

- Degree of a node → Number of neighbours
- Distance between Node A and B → shortest path
- Eccentricity → Maximum distance between A to anynode
- Connected components → Linked nodes
- Center of connected component

Connected Graphs

Trail → Walk with distinct edges

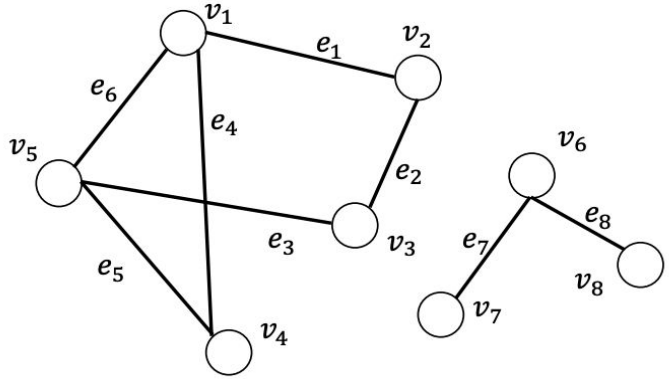
Path → Walk with distinct nodes



Source: Deep Learning on Graphs, Yao Ma and Jiliang Tang

Node Centrality

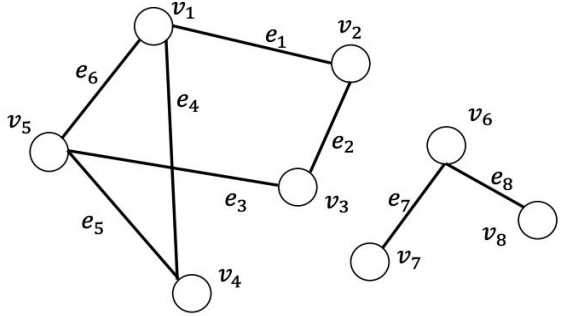
- Importance of a node
- Degree based → Number of neighbour nodes
- Eigenvector centrality
- Catz centrality
- Betweenness centrality



Source: Deep Learning on Graphs, Yao Ma and Jiliang Tang

Node Centrality

- Eigenvector centrality
  - largest eigenvalue of adjacency matrix
  - eigen vector as centrality score vector
- Betweenness centrality



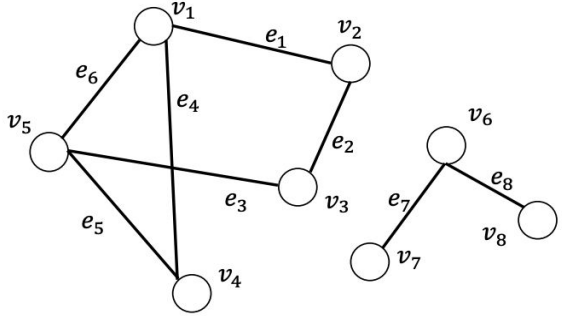
Number of shortest paths passing through a node

Source: Deep Learning on Graphs, Yao Ma and Jiliang Tang

Katz Centrality

Variant of eigenvector centrality

$$c_k(v_i) = \alpha \sum_{j=1}^{N} \mathbf{A}_{i,j} c_k(v_j) + \beta,$$



$$\mathbf{c}_k = (\mathbf{I} - \alpha \cdot \mathbf{A})^{-1} \boldsymbol{\beta}.$$

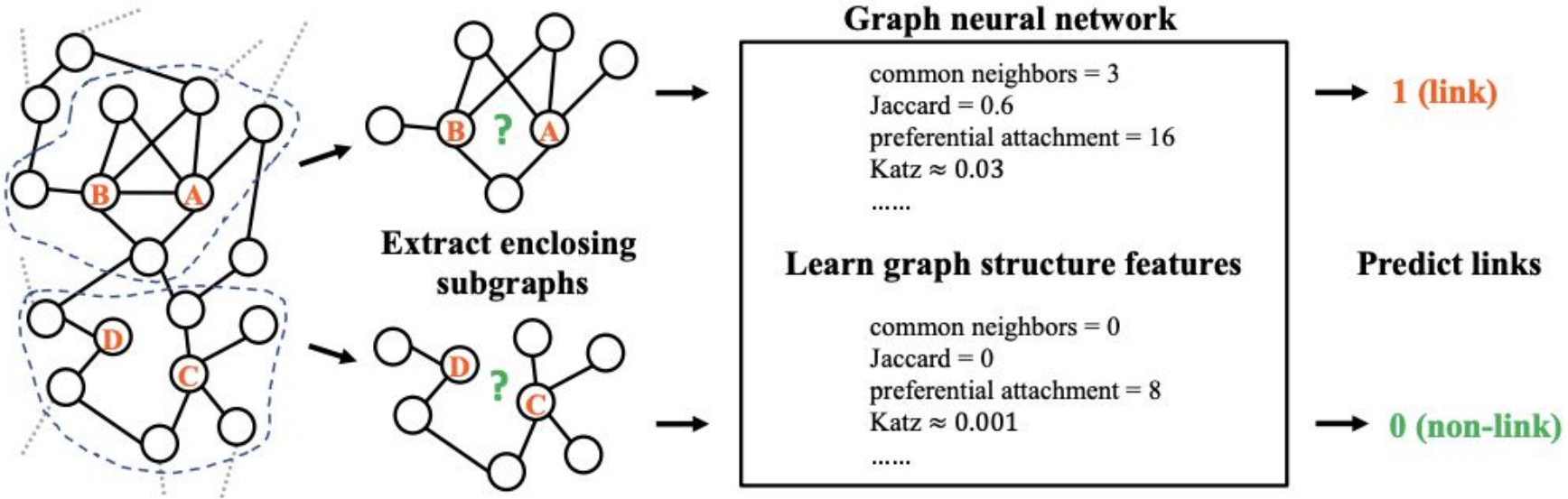
Source: Deep Learning on Graphs, Yao Ma and Jiliang Tang

Applications in Biology

- Protein-Protein Interaction (PPI)
  - Interactions among proteins, Proteins as nodes, interactions as edges
- Gene regulatory networks
  - Regulating gene expression, generating proteins from DNA sequence
- Metabolic networks
  - Metabolic reactions in a living organisms, enzymes as edges
- Drug drug interaction networks

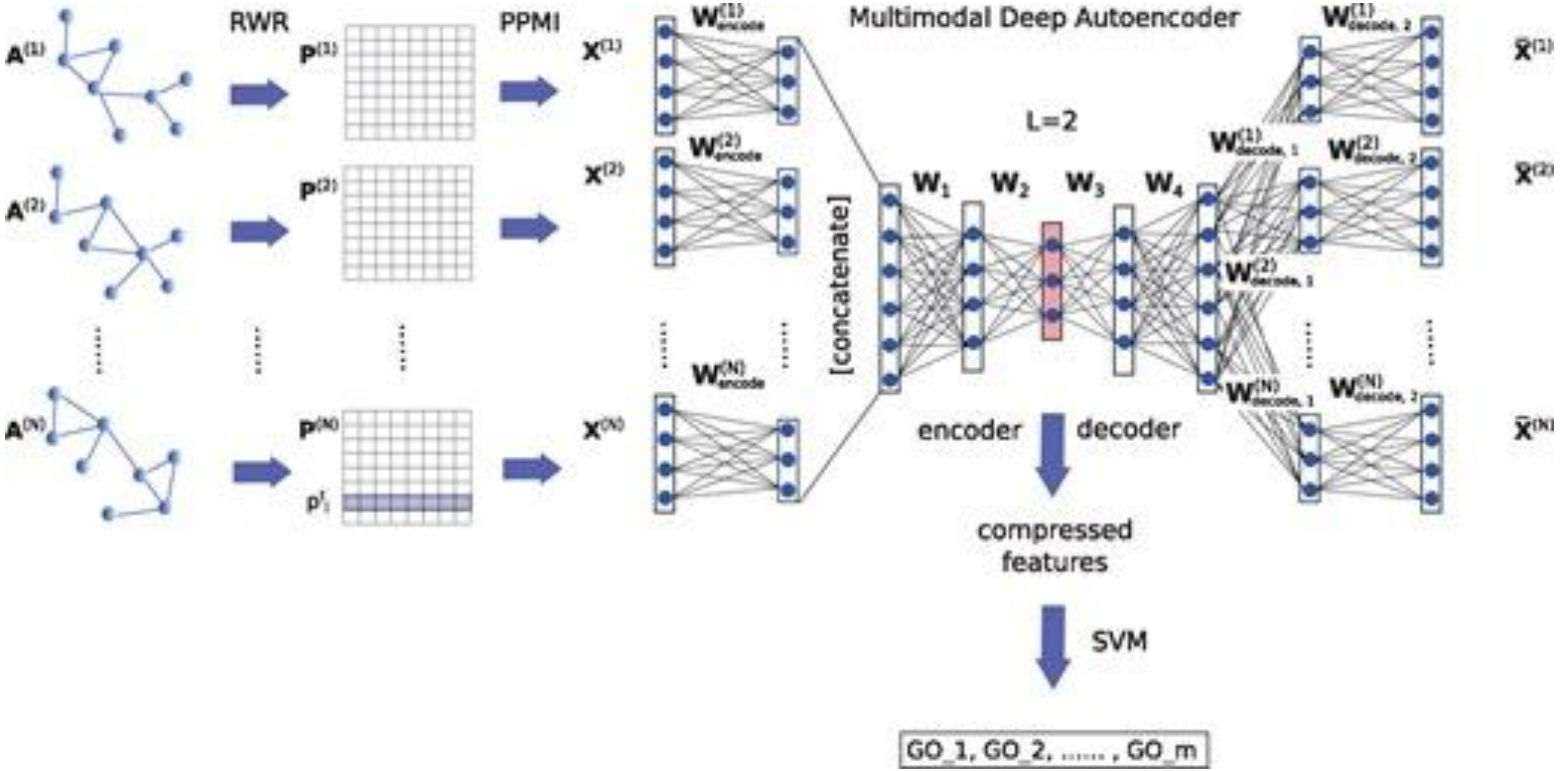
Giulia Muzio et al. Biological network analysis with deep learning, Bioinformatics, 2021

Link Prediction



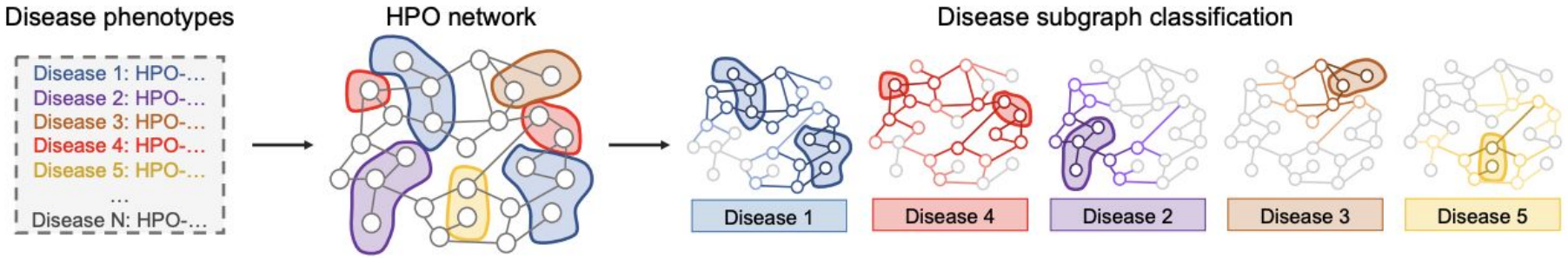
Zhang and Chen, Link Prediction Based on Graph Neural Networks, NeurIPS 2018

Node Classification



Gligorijević V, Barot M, Bonneau R. deepNF: deep network fusion for protein function prediction. *Bioinformatics*, 2018.

Graph Classification



Michelle M. Li, Kexin Huang, Marinka Zitnik, Graph Representation Learning in Biomedicine, 2021

Popular Datasets

Drugbank

Protein Interaction Network Analysis

MUTAG

Molecular INTeraction

PubChem's BioAssay database

DREAM4

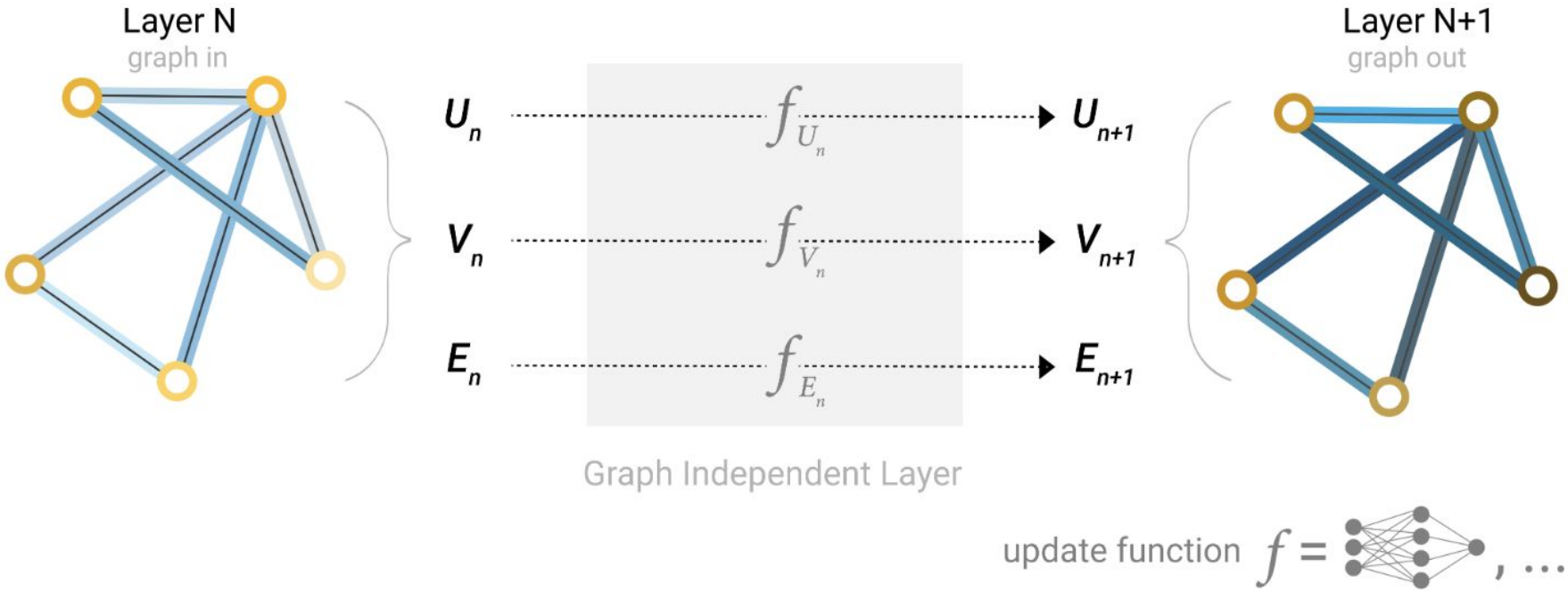
Protein Data Bank

Tox21

Human Protein Reference Database

Giulia Muzio et al. Biological network analysis with deep learning, Bioinformatics, 2021

Graph Neural Network Model



Sanchez-Lengeling, et al., "A Gentle Introduction to Graph Neural Networks", Distill, 2021.

Graph Predictions

- Pooling → Gather all information in matrix, Average/Sum
- Node prediction with no node features → Pooling with edges
- Edge prediction with no properties → Pooling with nodes
- Global Average Pooling → Graph embedding

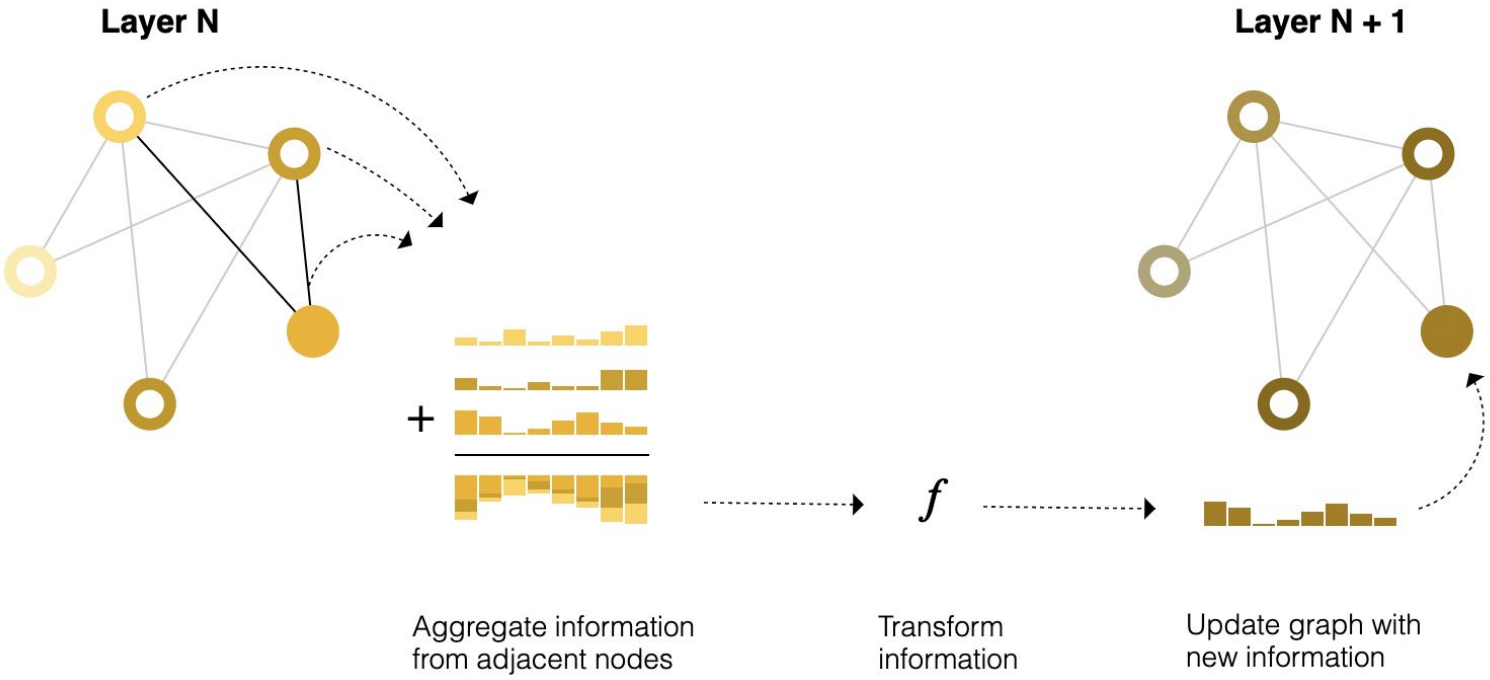
Sanchez-Lengeling, et al., "A Gentle Introduction to Graph Neural Networks", Distill, 2021.

Message Passing

- Gather neighbour node information
- Aggregate them
- Update function → an NN method

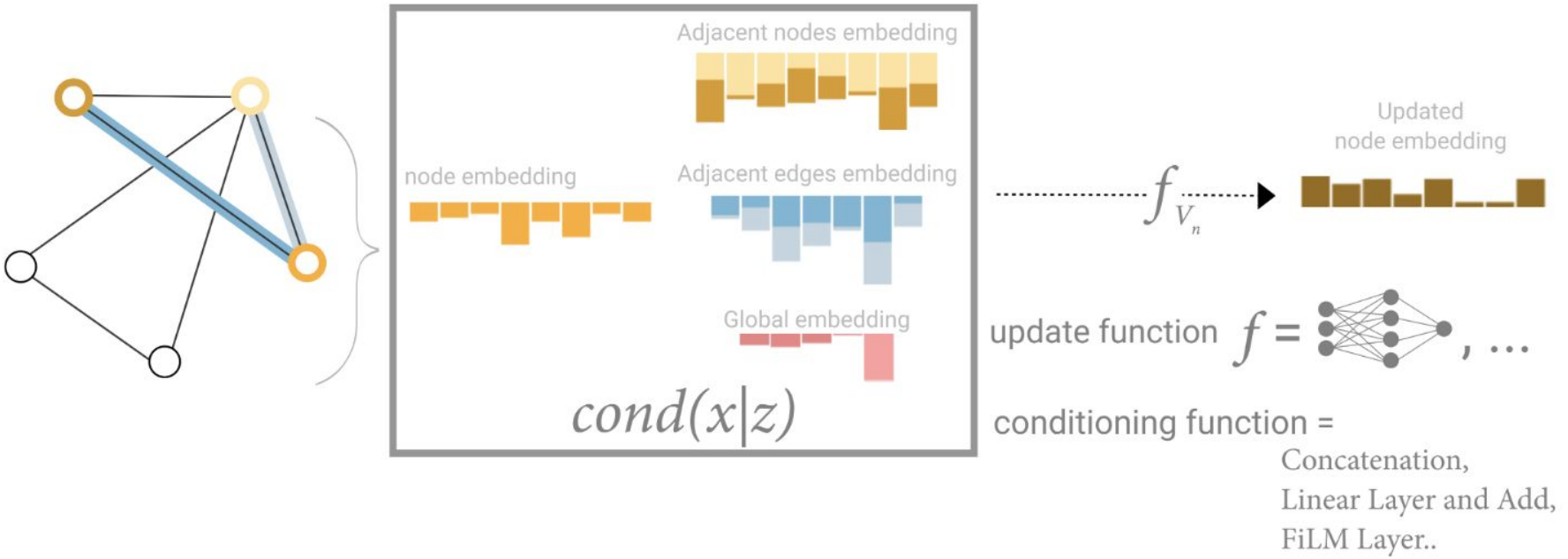
Sanchez-Lengeling, et al., "A Gentle Introduction to Graph Neural Networks", Distill, 2021.

Message Passing



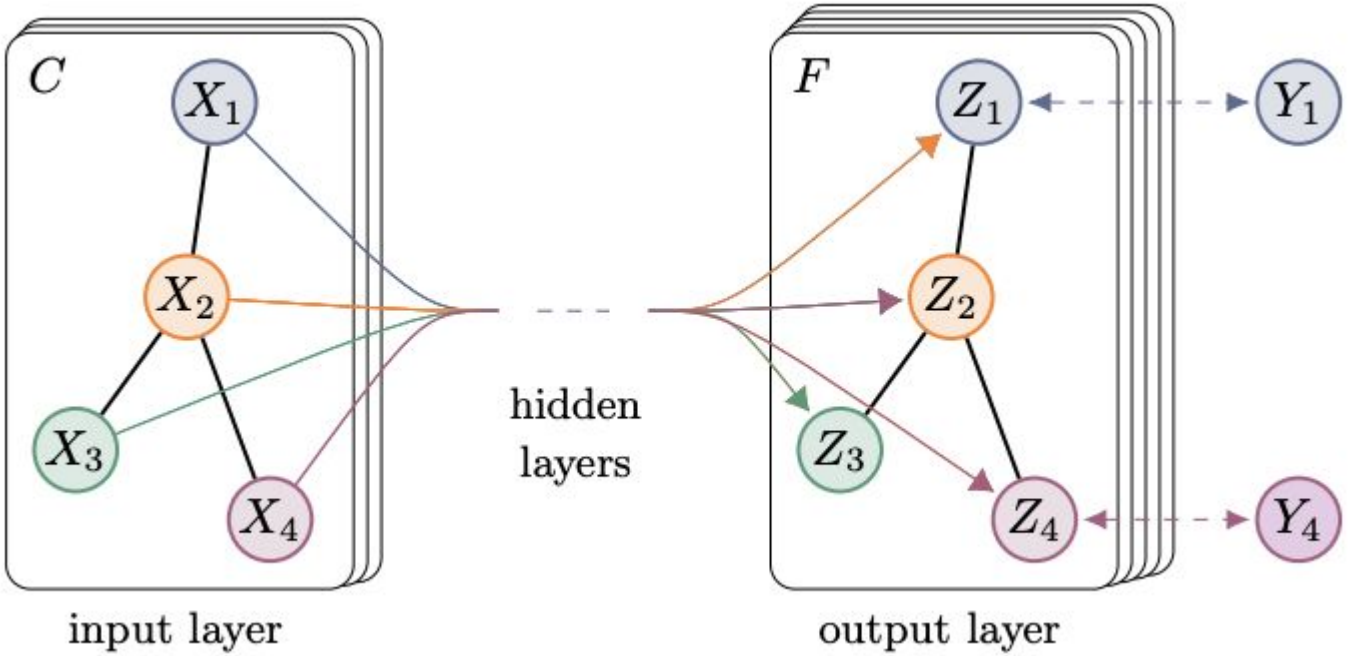
Sanchez-Lengeling, et al., "A Gentle Introduction to Graph Neural Networks", Distill, 2021.

Global Representation



Sanchez-Lengeling, et al., "A Gentle Introduction to Graph Neural Networks", Distill, 2021.

Graph Convolutional Networks



Thomas N. Kipf, Max Welling, "Semi-Supervised Classification with Graph Convolutional Networks, 2016.

Graph Convolutional Networks

Inputs: Initial feature vector for nodes, adjacency list

$$H^{l+1} = f(H^l, A),\ \text{L - number of layers}$$

Let the "f" be ReLU,

$$H^{l+1} = ReLU(AH^lW^l),\quad W \text{ is the weight matrix,}$$

ReLU is a non-linear activation function, To normalize, $1/D_i^{1/2}$ A $1/D_j^{1/}$

Implementations

Pytorch:

https://pytorch-geometric.readthedocs.io/en/1.3.2/_modules/torch_geometric/nn/conv/gcn_conv.html

Tensorflow:

https://github.com/stellargraph/stellargraph/blob/develop/stellargraph/layer/gcn.py

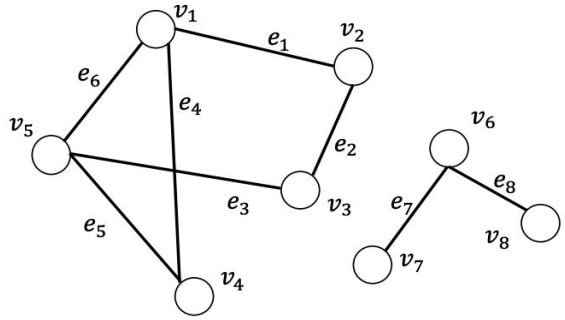
Deep Walk

Adjacent nodes ⟷ Similar representations

Transductive algorithm

1. Random walks of "L" length, "k" times
2. SkipGram method

   Gradient descent to update neighbour node



GraphSage

Dynamic graph learning - No need to retrain

Context based similarity representations

1. Random Walks
2. Learning weights using aggregate function
3. Neural network layer to learn weights

GraphSage

Algorithm 1: GraphSAGE embedding generation (i.e., forward propagation) algorithm

Input : Graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$; input features $\{\mathbf{x}_v, \forall v \in \mathcal{V}\}$; depth K; weight matrices $\mathbf{W}^k, \forall k \in \{1, ..., K\}$; non-linearity σ; differentiable aggregator functions $\text{AGGREGATE}_k, \forall k \in \{1, ..., K\}$; neighborhood function $\mathcal{N} : v \rightarrow 2^{\mathcal{V}}$

Output : Vector representations $\mathbf{z}_v$ for all $v \in \mathcal{V}$

$\mathbf{h}_v^0 \leftarrow \mathbf{x}_v, \forall v \in \mathcal{V}$;
for $k = 1...K$ **do**
  for $v \in \mathcal{V}$ **do**
    $\mathbf{h}_{\mathcal{N}(v)}^k \leftarrow \text{AGGREGATE}_k(\{\mathbf{h}_u^{k-1}, \forall u \in \mathcal{N}(v)\})$;
    $\mathbf{h}_v^k \leftarrow \sigma\left(\mathbf{W}^k \cdot \text{CONCAT}(\mathbf{h}_v^{k-1}, \mathbf{h}_{\mathcal{N}(v)}^k)\right)$
  end
  $\mathbf{h}_v^k \leftarrow \mathbf{h}_v^k / \|\mathbf{h}_v^k\|_2, \forall v \in \mathcal{V}$
end
$\mathbf{z}_v \leftarrow \mathbf{h}_v^K, \forall v \in \mathcal{V}$

Hamilton, Will, Zhitao Ying, and Jure Leskovec. “Inductive representation learning on large graphs.” , NeuIPS, 2017.

Graph Attention Networks

Attention mechanism → higher weight for important nodes

1. Linear transformation, $W^{l}H^{l}$
2. Additive attention
   a. Concatenate 2 neighbour features
   b. Multiply with weight vector, a.
   c. Apply LeakyReLU
3. Normalize with softmax
4. GCN

Graph Encoder

To find a representative embedding

Encoder layer → Graph convolutional layers, non-linear activations

Decoder layer - Graph convolutional layers

Reconstruction loss to find the difference.

Demonstration

Brain Decoding

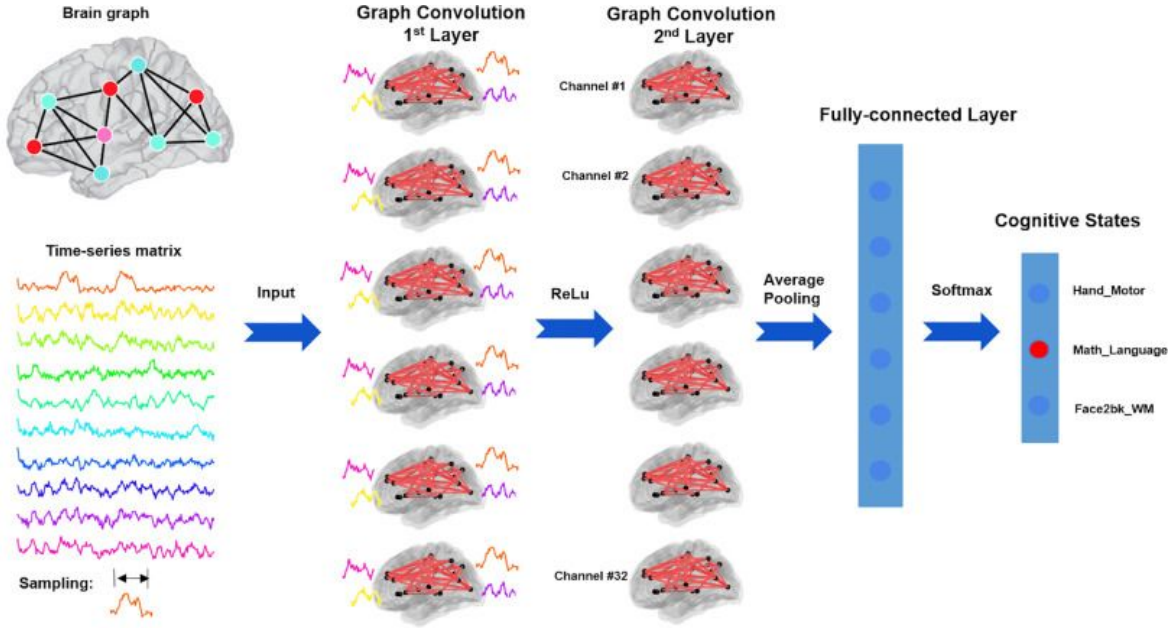
- Multidomain brain decoder for brain response
- Graph classification
- Human Connectome Project task-fMRI database
  - 1200 participants, 21 different experimental conditions
  - 6 cognitive domain: emotion, language, motor, relational, social, and working memory.

Yu Zhang, Loïc Tetrel, Bertrand Thirion, Pierre Bellec

"Functional annotation of human cognitive states using deep graph convolution" , 2021

.

Brain Decoding



6 GCN layers with 32 graph filters

Global average pooling layer,

2 fully connected layers

Softmax layer

Yu Zhang, Loïc Tetrel, Bertrand Thirion, Pierre Bellec

"Functional annotation of human cognitive states using deep graph convolution" , 2021

Autism Spectrum Disorder

Graph node classification

Autism Brain Imaging Data Exchange (ABIDE) dataset

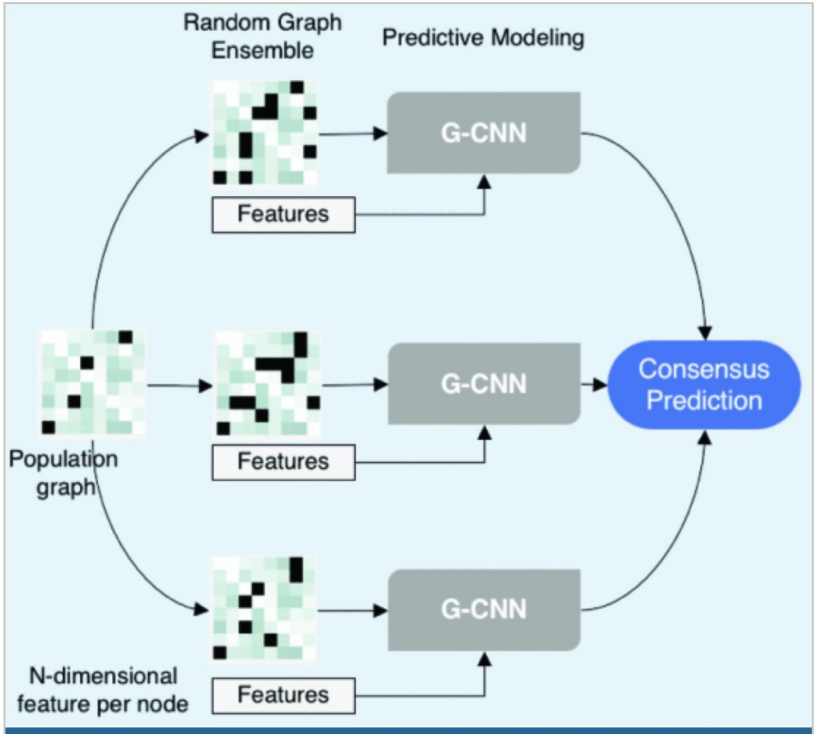
Gender and site information → Population Graph

Imaging features → Node feature vector

Anirudh, R., and Thiagarajan, J

Bootstrapping graph convolutional neural networks for autism spectrum disorder classification, 2019

Autism Spectrum Disorder



Anirudh, R., and Thiagarajan, J

Bootstrapping graph convolutional neural networks for autism spectrum disorder classification, 2019

Conclusion

Introduction to graph notations

Basic architectures

Some application