Introduction to
Graph Machine Learning

Rizal Fathony
rizal@fathony.com

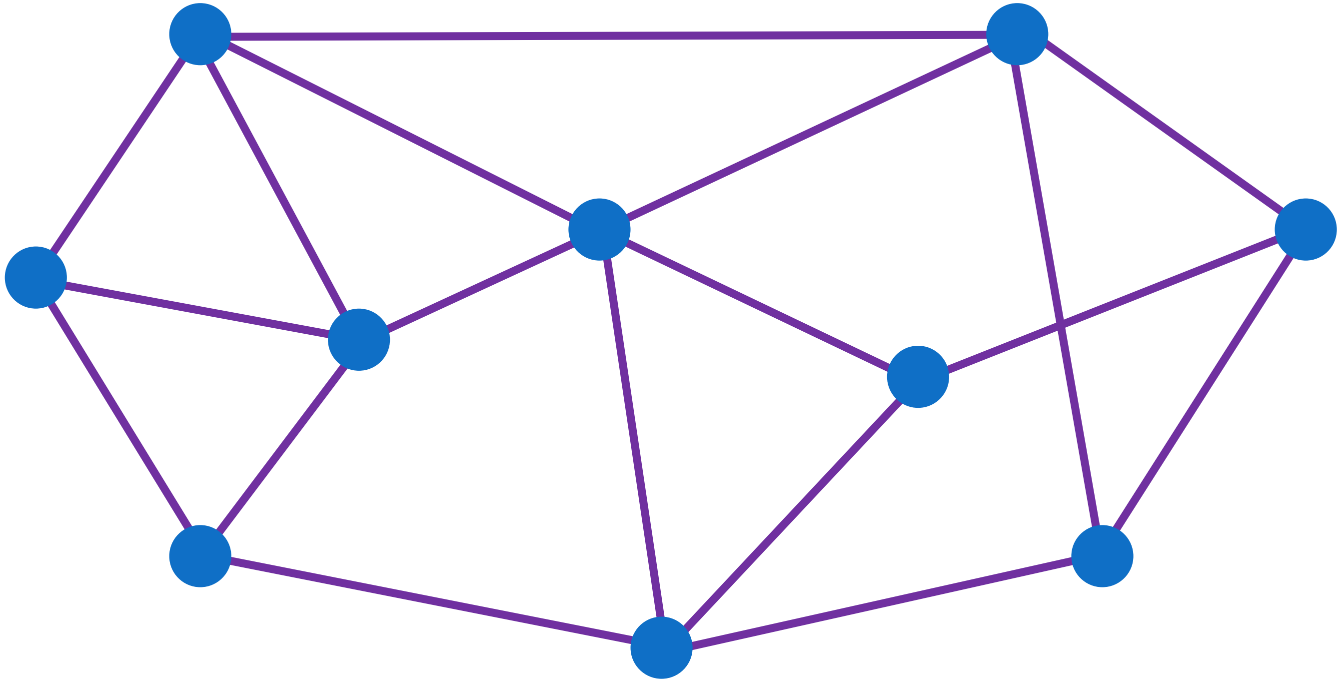
*Some of the materials are modified from Juree Leskovec's graph presentations

What is a Graph?

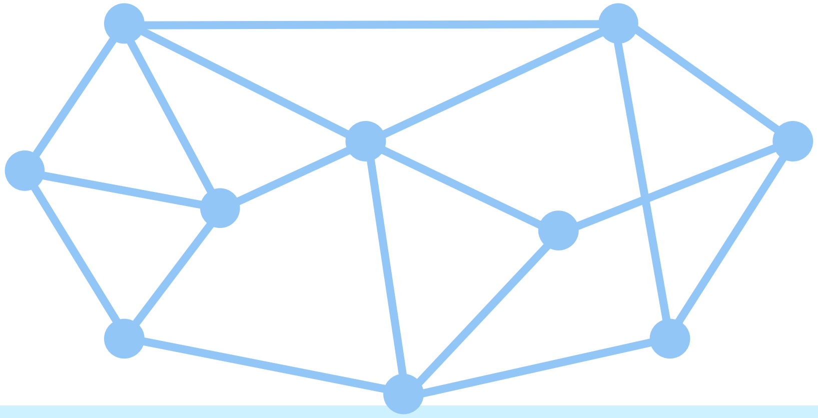
Mathematical structure for pairwise relations

Nodes

Edges

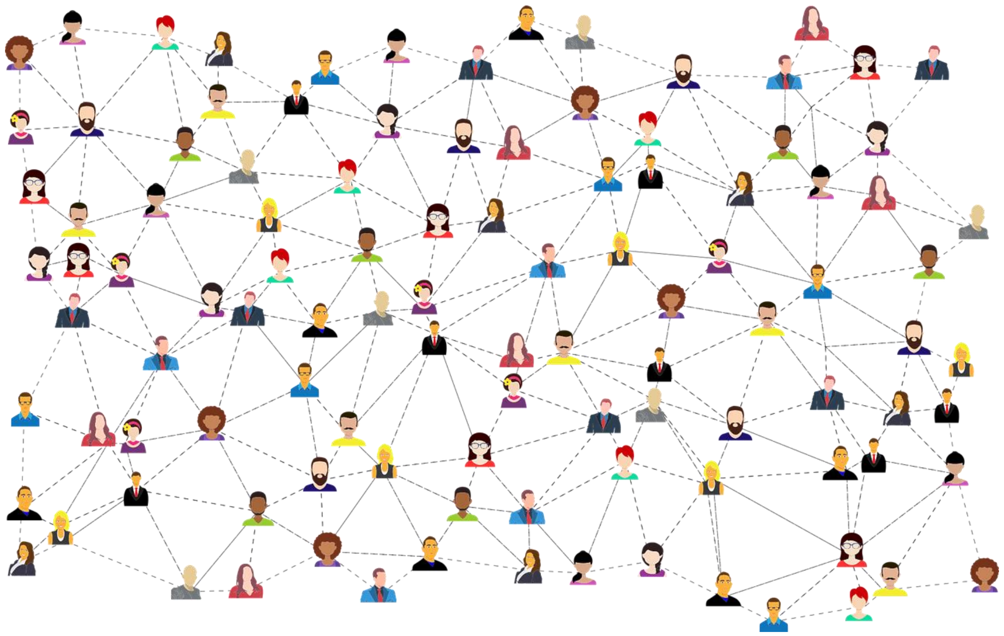


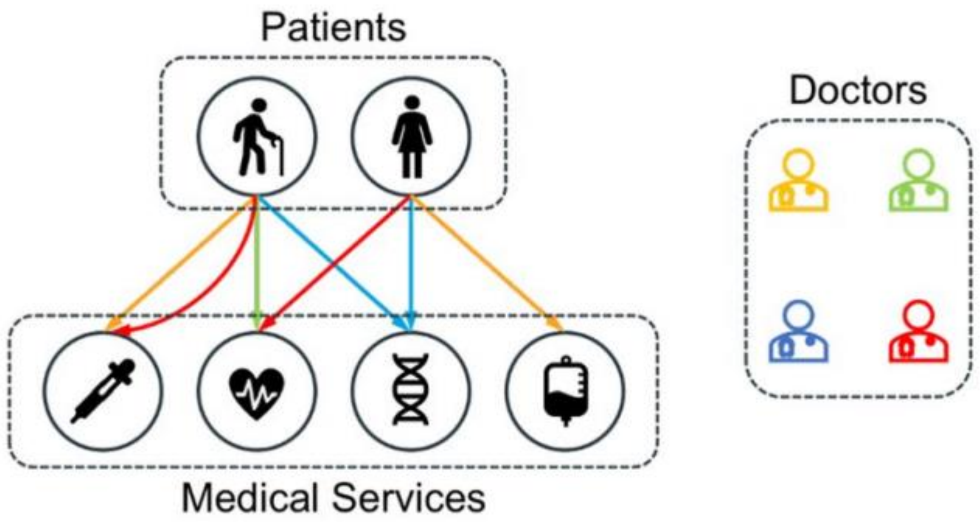
Why Graphs?



Graph is a general method for describing and modeling complex systems

Many Data are Graphs





Bipartite Graph

Social Networks

Nodes: Person/Account

Edges: Friendship/Follows

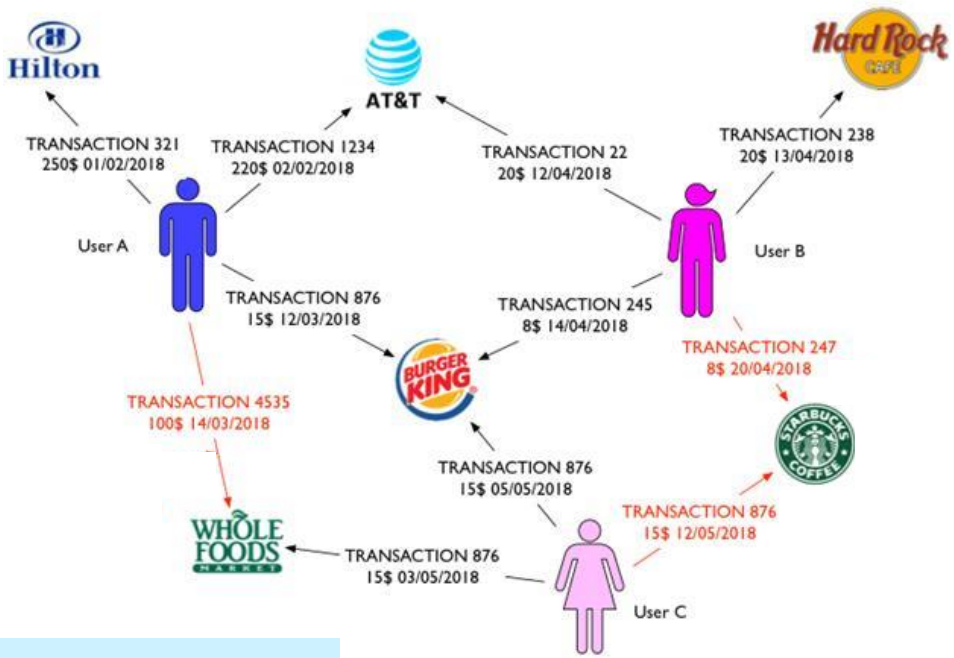
Health Records

Nodes: Patient, Medical Service

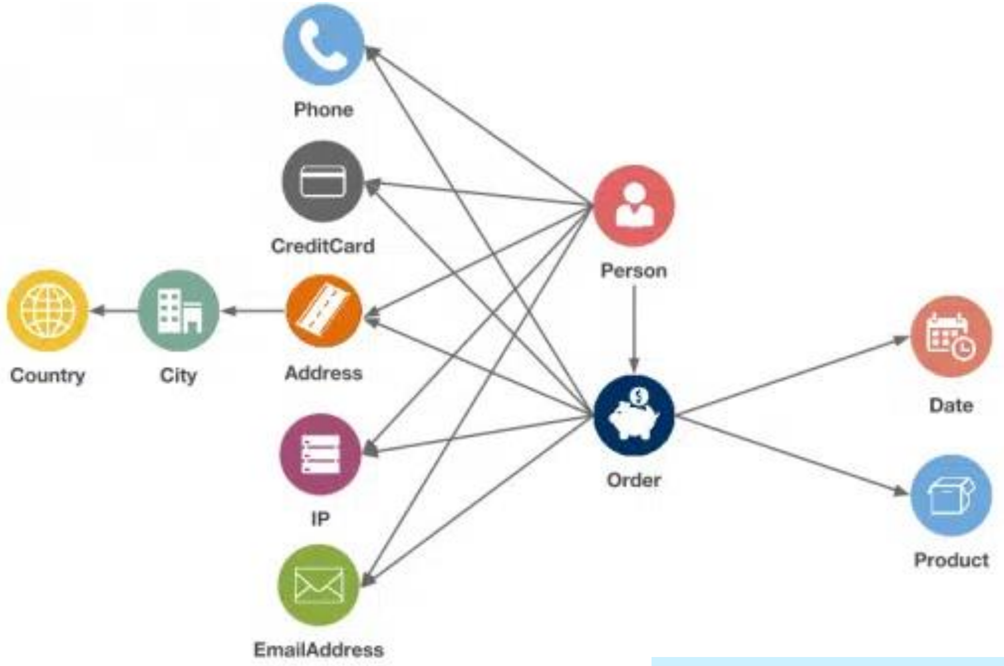
Edges: Treatment by Doctors

https://brandblast.com/strategy/avatars-and-favicons/
https://www.nature.com/articles/s41598-021-85255-w.pdf

Many Data are Graphs



Directed Graph



Heterogenous Graph

Financial Transactions

Nodes: Customer, Merchant

Edges: Transaction/Payment

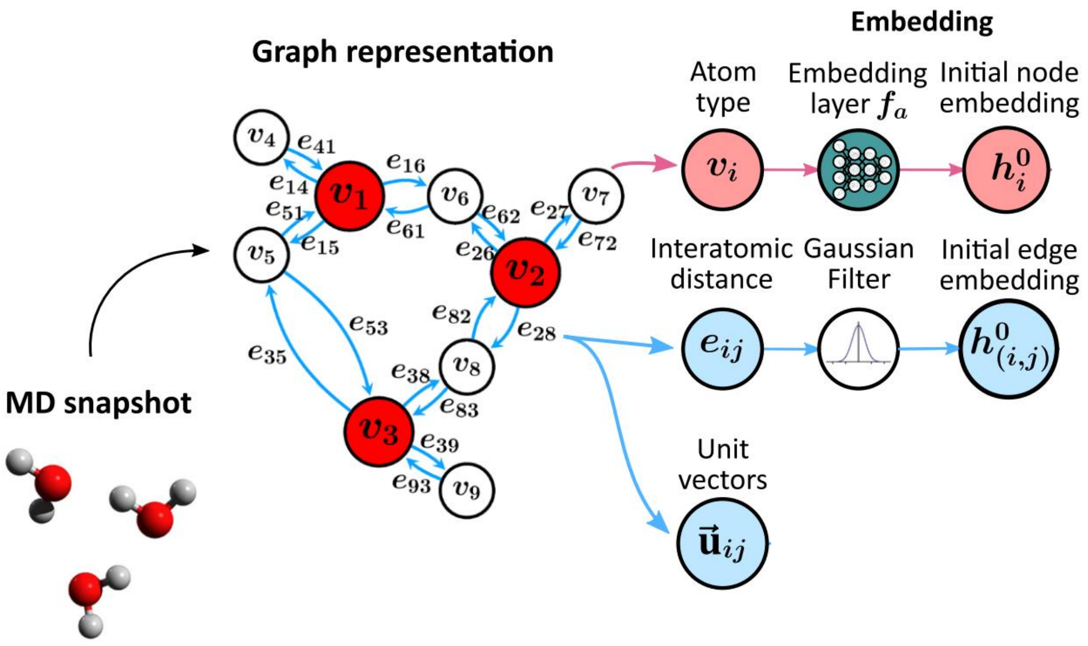
E-Commerce Data

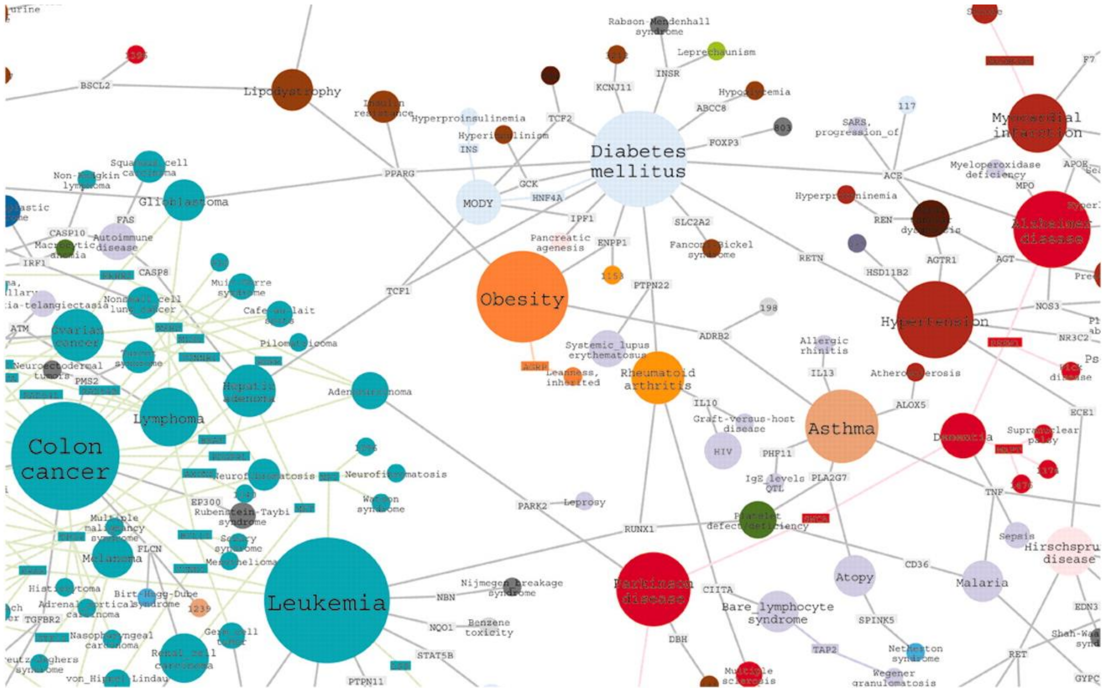
Nodes: Person, Product, Credit Cards, …

Edges: Has Phone, Has Address, Orders, …

https://livebook.manning.com/book/graph-powered-machine-learning/chapter-3/v-8/
https://7wdata.be/data-analysis/fraud-detection-in-retail-with-graph-analysis/

Many Data are Graphs





Molecular Modeling

Nodes: Atom

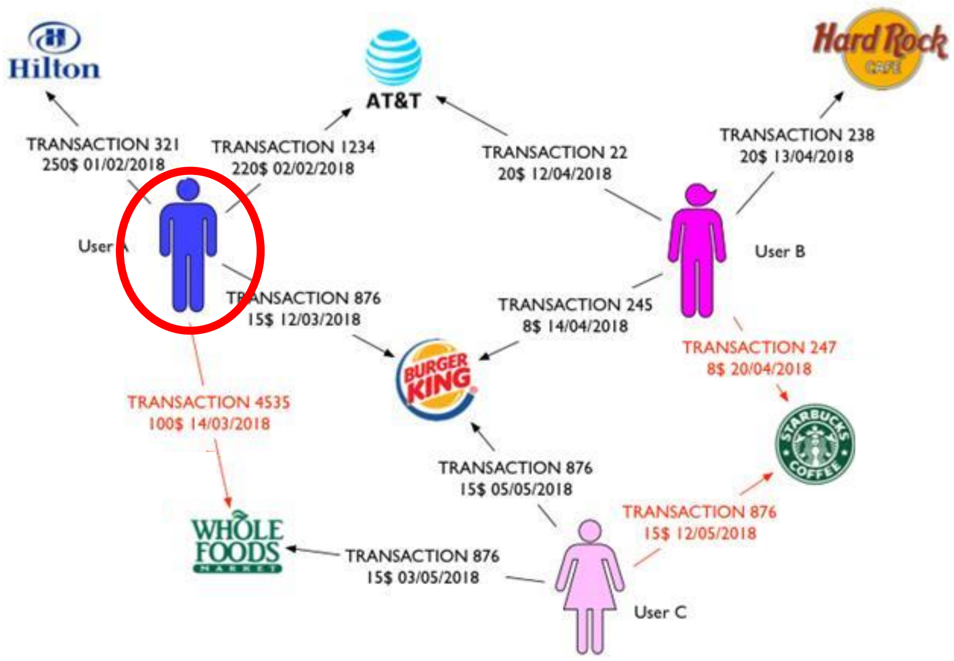
Edges: Chemical Bond

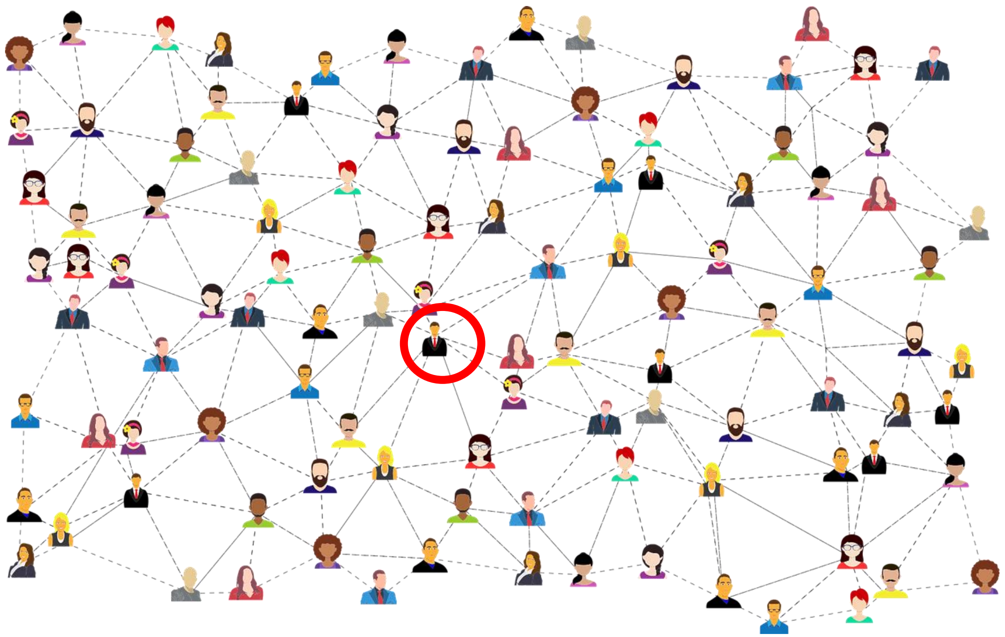
Human Disease Network

Nodes: Disease

Edges: Genetic Link

Task Example: Node Prediction





Input: Transactional Graph

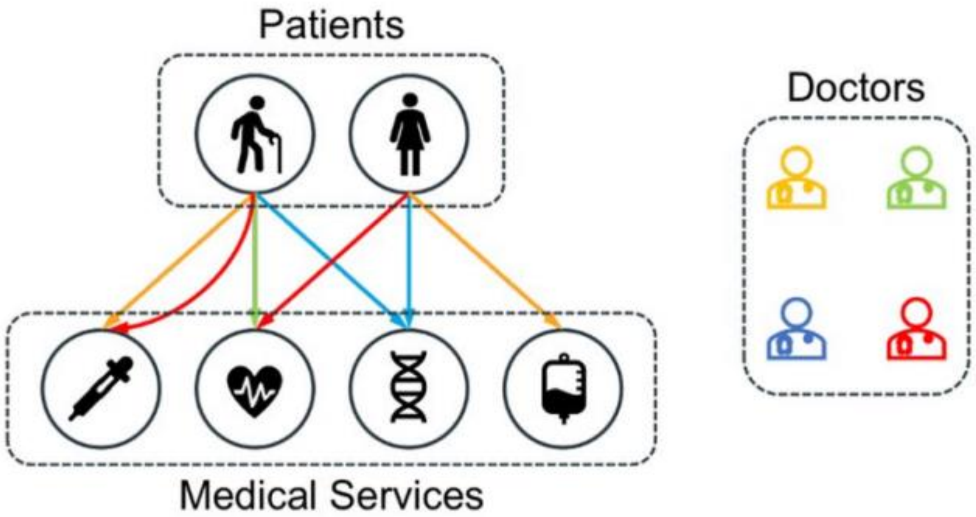
Task: Find user that use stolen credit card in the transactions

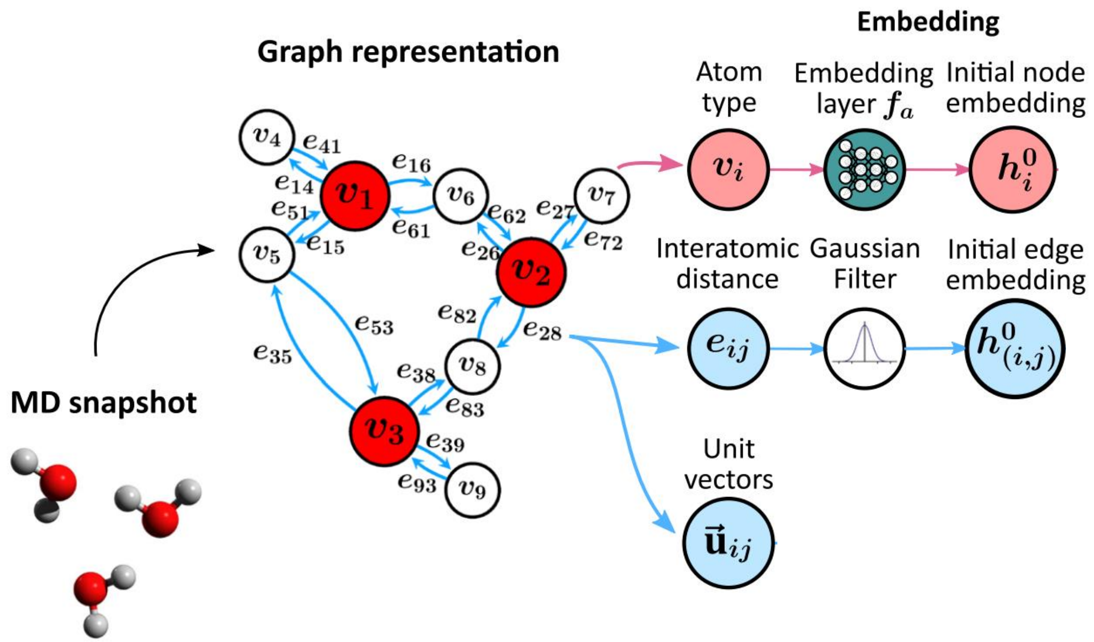
Input: Social Network

Task: Identify fake user with influence power

https://livebook.manning.com/book/graph-powered-machine-learning/chapter-3/v-8/
https://brandblast.com/strategy/avatars-and-favicons/

Task Example: Edge Prediction





Input: Health Records Graph

Task: Predict if a patient need to see a doctor for medical treatment

Input: Molecular Graph

Task: Predict how strong the chemical bonds' force for a given molecule

https://eng.uber.com/uber-eats-graph-learning/
https://www.nature.com/articles/s41524-021-00543-3.pdf

What's Next?

1 Introduction

2 ML Algorithms

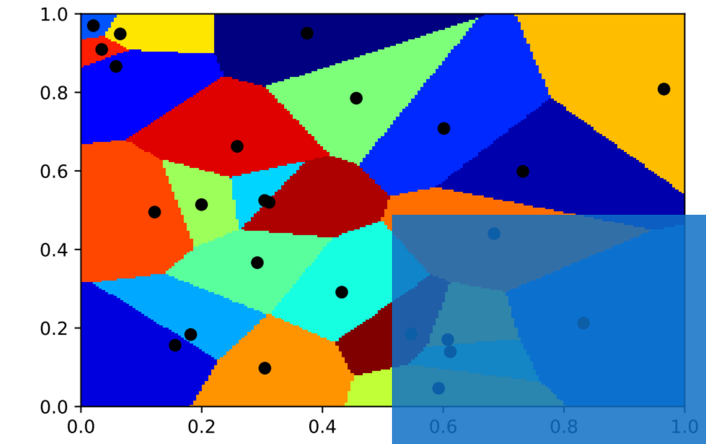
3 Node Embedding

4 Graph Neural Networks

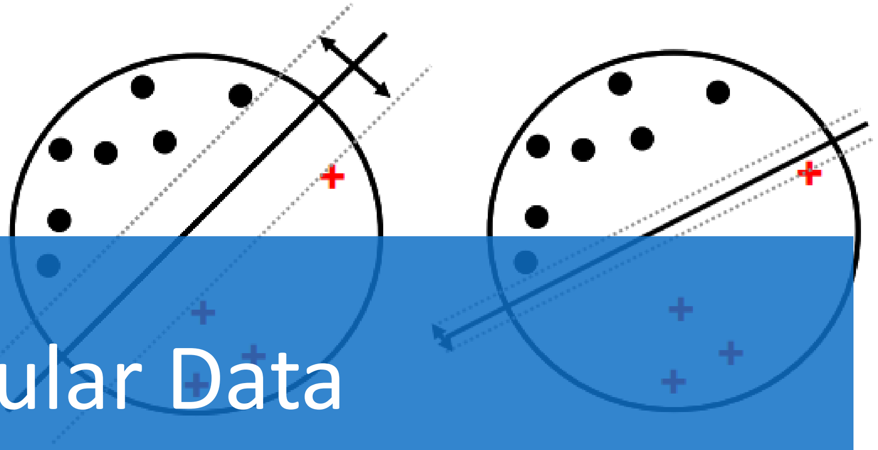
Machine Learning Algorithms

From Classical to Graph ML

Classical Machine Learning Algorithms

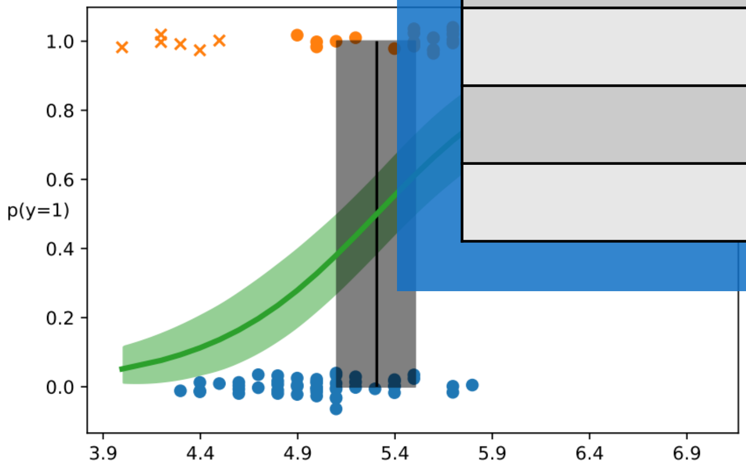


K-Nearest Neig

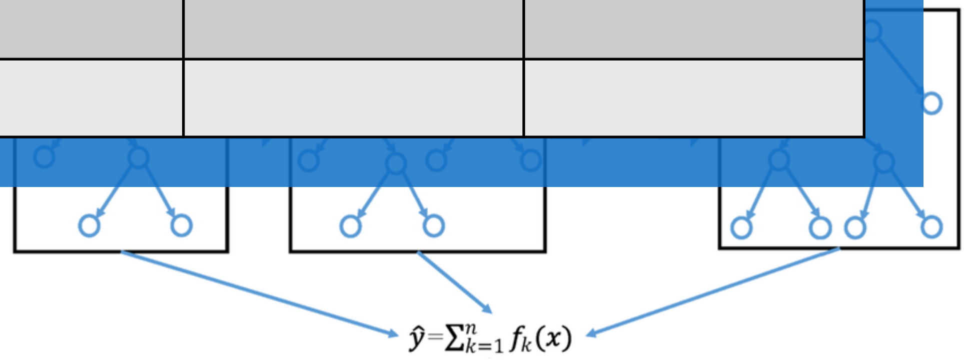


Tabular Data

Feature 1	Feature 2	Feature 3	Feature 4



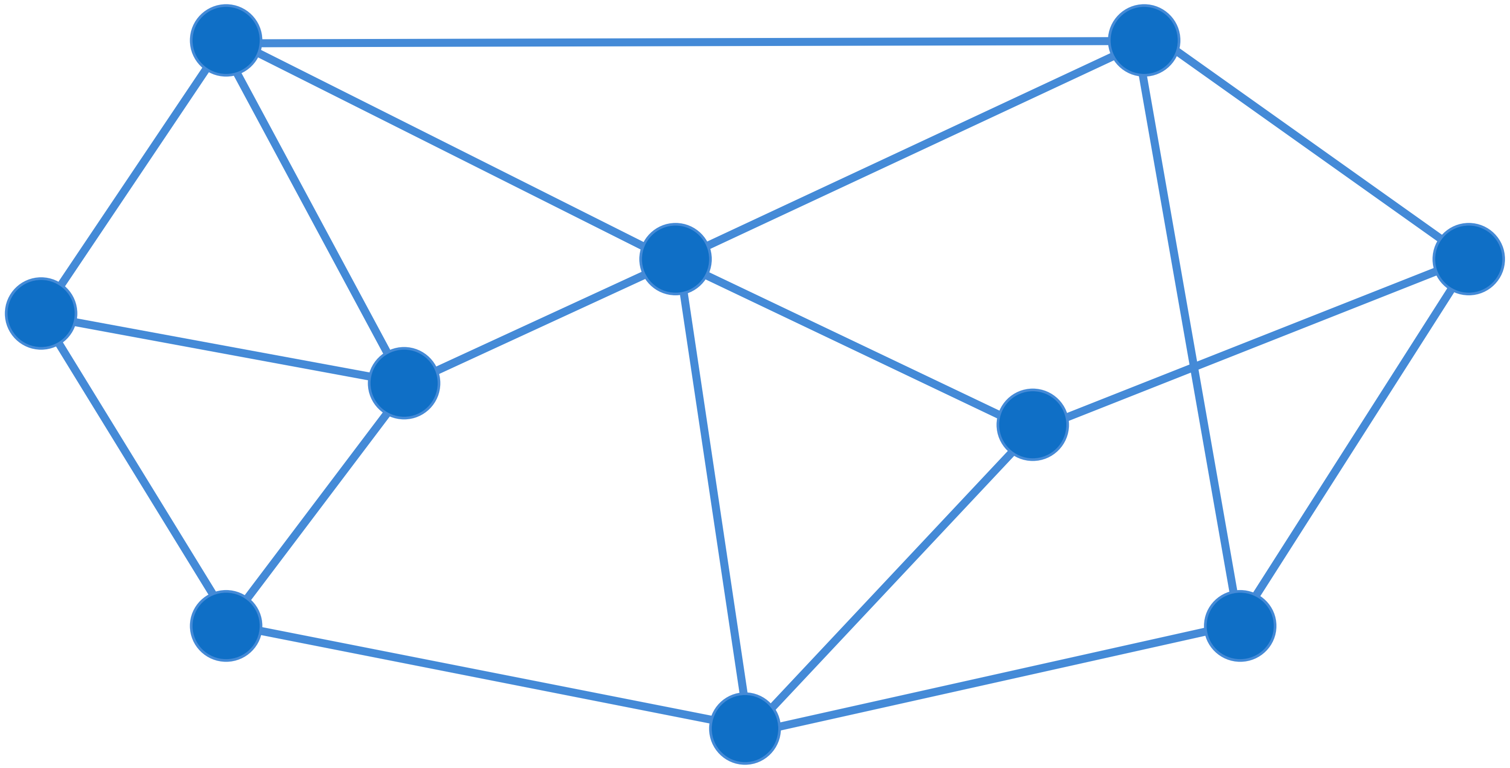
Logistic Regression



(Boosted) Decision Tree

https://probml.github.io/pml-book/book1.html

Classical ML for Graph Data?



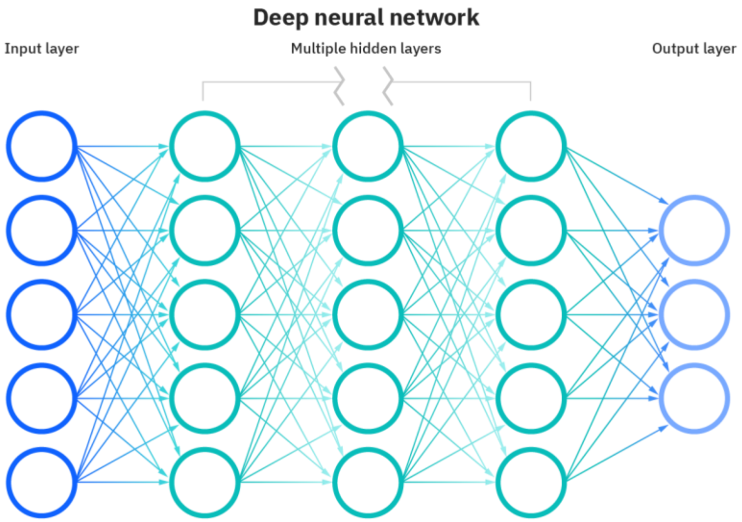
Graph

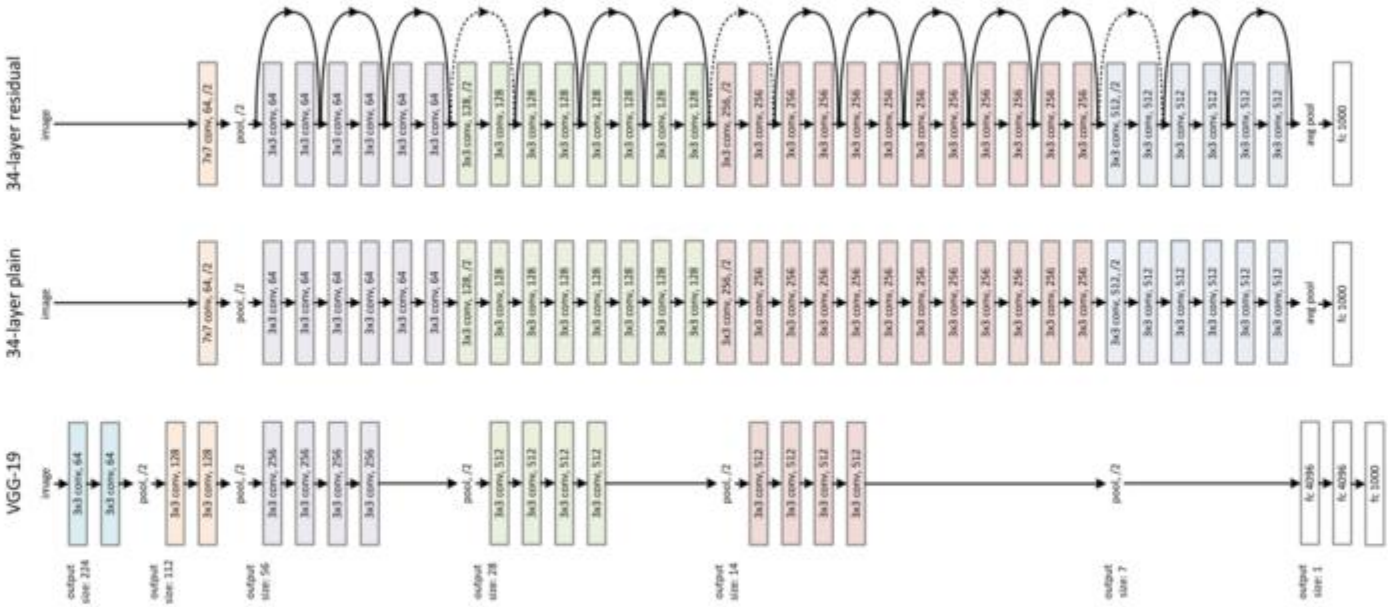
Classical ML for Graph Data?

Tabular

Neural Networks and Deep Learning

Multiple layers of learning





Multi Layer Perceptron (MLP)

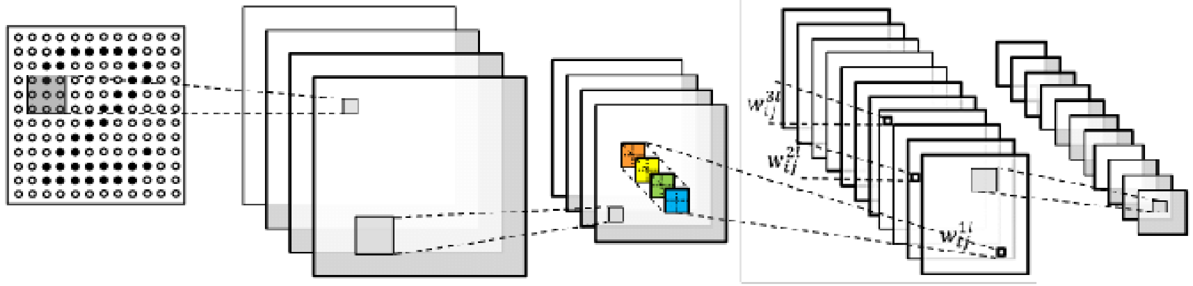
Residual Networks (ResNet)

Capable to learn from *"raw"* data

https://www.ibm.com/cloud/learn/neural-networks
https://medium.com/analytics-vidhya/deep-residual-learning-for-image-recognition-resnet-94a9c71334c9

Grids and Sequences

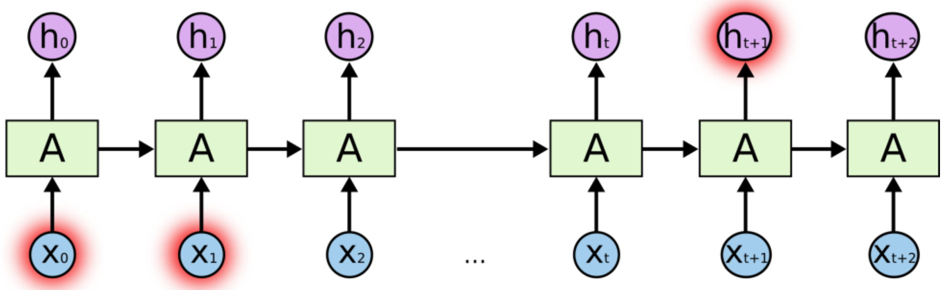
Grids



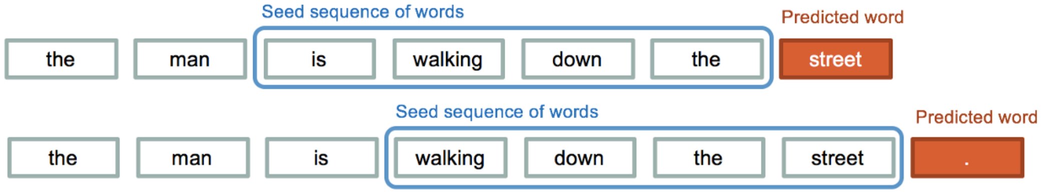
Convolutional Neural Networks (CNN)

Images

Sequences



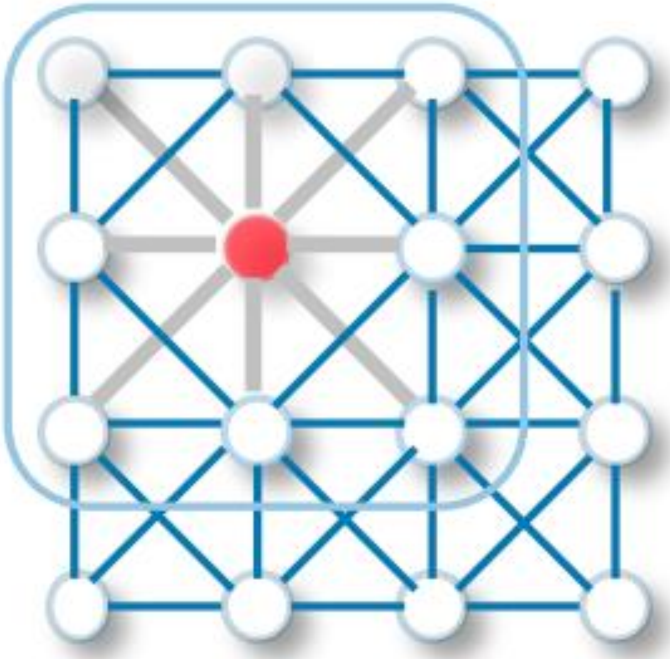
Recurrent Networks (RNN)



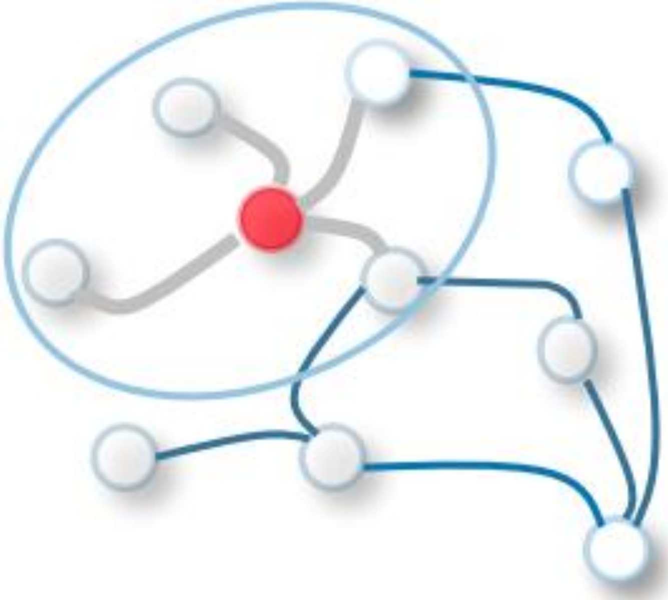
Text

https://www.researchgate.net/figure/Feed-forward-of-a-CNN-model-CONV-CONV-layers-are-the-main-components-of-a-CNN-model_fig1_305727136
https://kotakode.com/blogs/4437/Mengenal-LSTM-Networks---Part-3

Grid and Sequence as Graph



Grid Computation Flow



Graph Computation Flow

https://arxiv.org/pdf/1901.00596.pdf

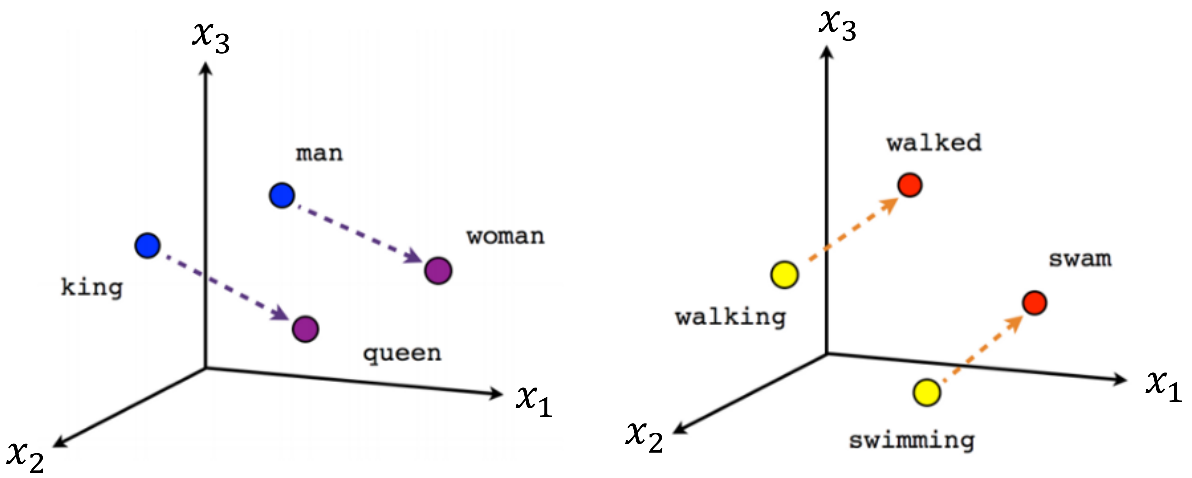
Node Embedding

Nodes + neighbors → numbers

Inspiration from word embedding: **word2vec**

Map **words** to **numerical** features

similar word → similar values
preserve word associations



Male-Female

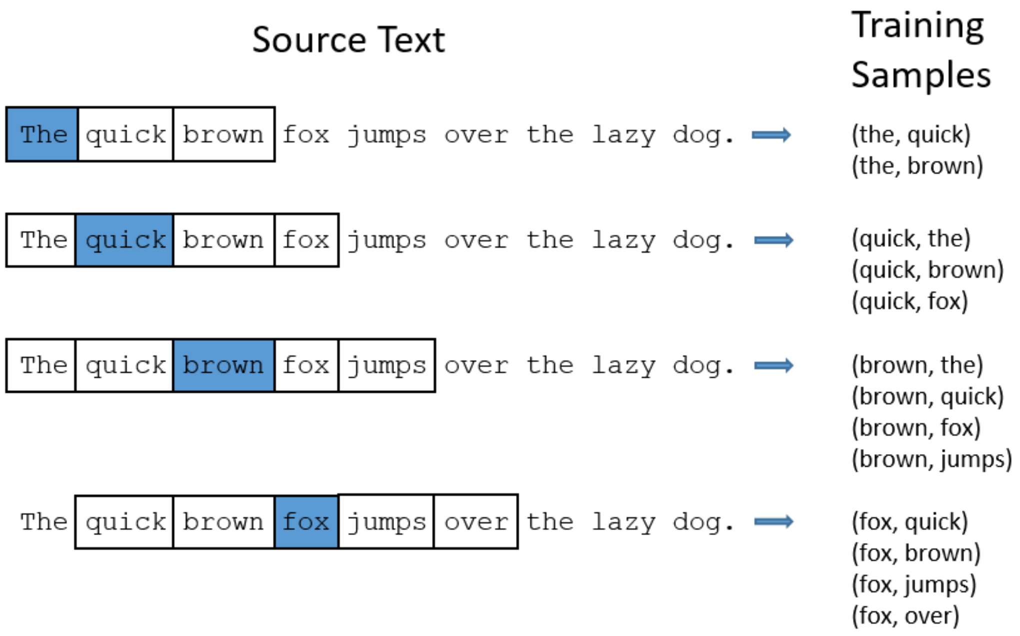
Verb tense

`king – man + woman ≈ queen`

`walked – walking ≈ swam - swimming`

word2vec training process:

predict the neighboring words

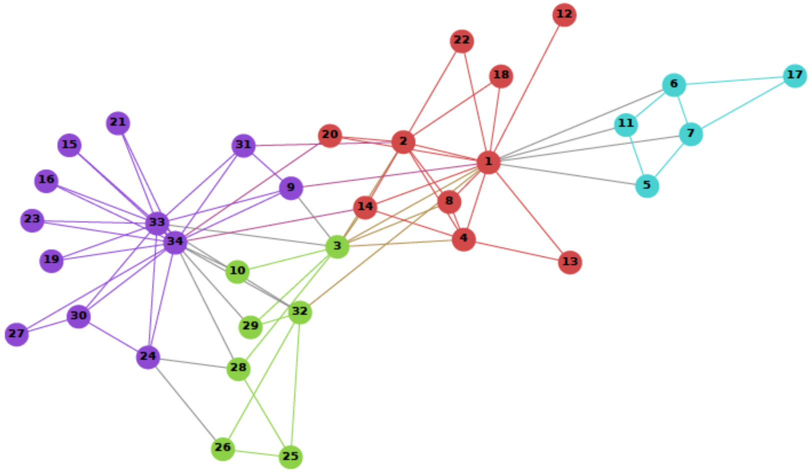


https://towardsdatascience.com/creating-word-embeddings-coding-the-word2vec-algorithm-in-python-using-deep-learning-b337d0ba17a8
http://mccormickml.com/2016/04/19/word2vec-tutorial-the-skip-gram-model/

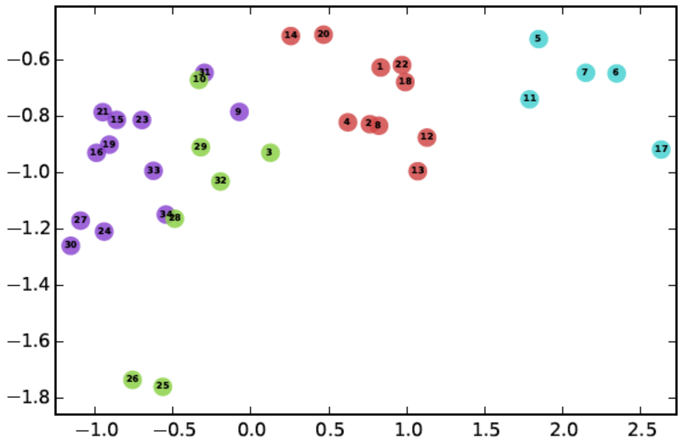
node2vec : a node embedding algorithm

Map nodes in a graph to
numerical features (embedding)

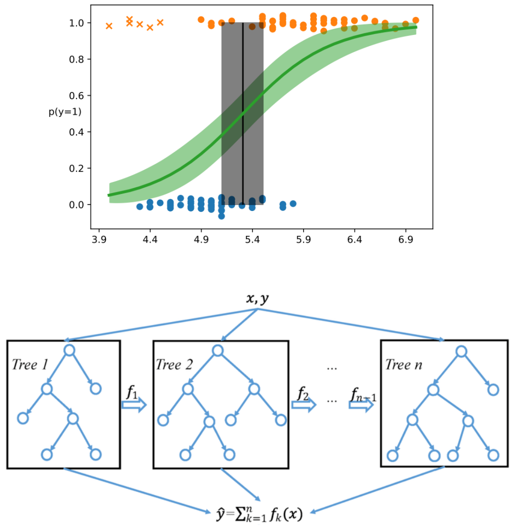
similar nodes → **similar** embeddings



Graph



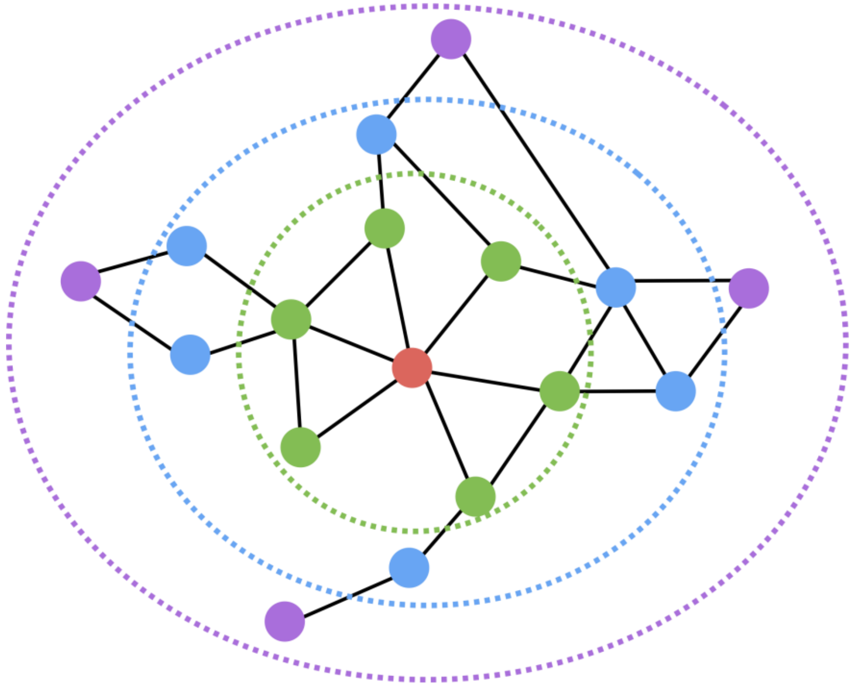
Embedding



Classifier
(ML models)

K-Hop Similarity

Neighboring nodes achievable in k-hop should have similar embedding



- **Red:** Target node
- **Green**: 1-hop neighbors
    - **A** (i.e., adjacency matrix)
- **Blue**: 2-hop neighbors
    - $\mathbf{A}^2$
- **Purple**: 3-hop neighbors
    - $\mathbf{A}^3$

Objective:

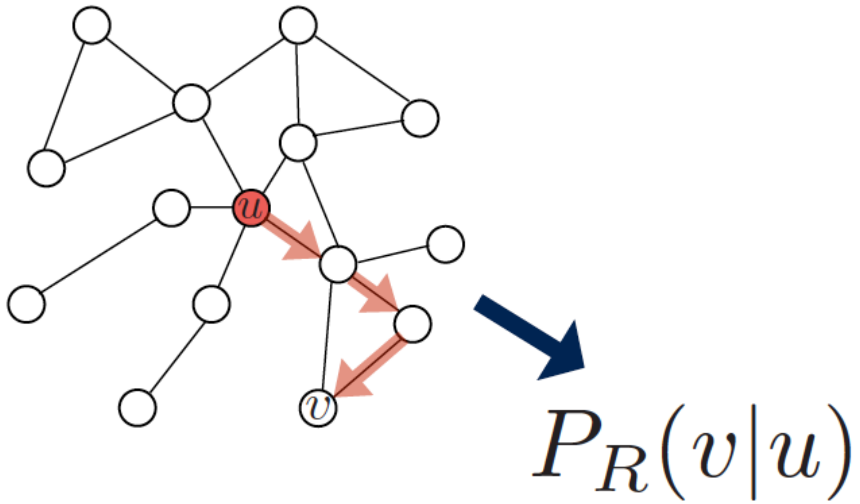
$$\mathcal{L} = \sum_{(u,v) \in V \times V} \| \mathbf{z}_u^\top \mathbf{z}_v - \mathbf{A}_{u,v}^k \|^2$$

Random Walk Similarity

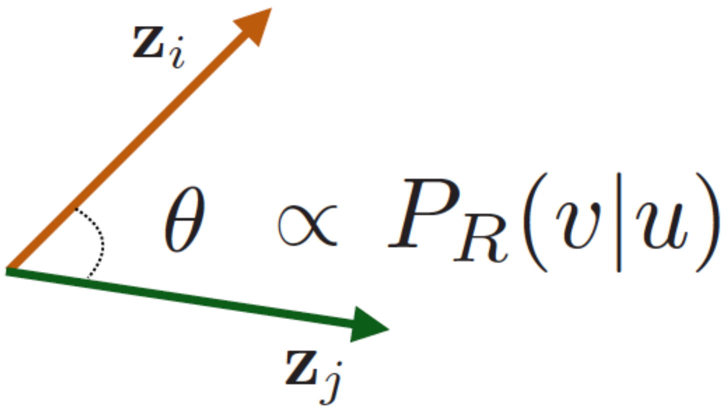
Random walk: start from node u, repeatedly jump (walk) to a neighboring node

$P_R(v|u)$: probability of visiting node v from random walks starting from node u

Embedding similarity should approximate $P_R(v|u)$



A random walk from u to v

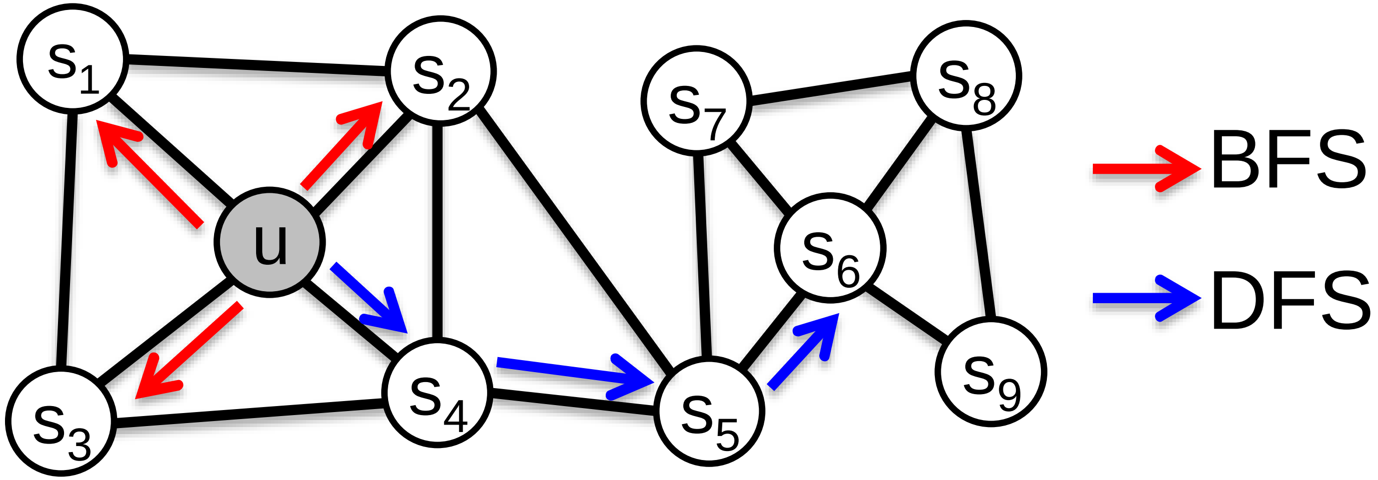


Cosine similarity

node2vec： biased random walk similarity

Biased random walk to encourage:
local and **global** views

local microscopic view → breadth first search (BFS) walk
global macroscopic view → depth first search (DFS) walk

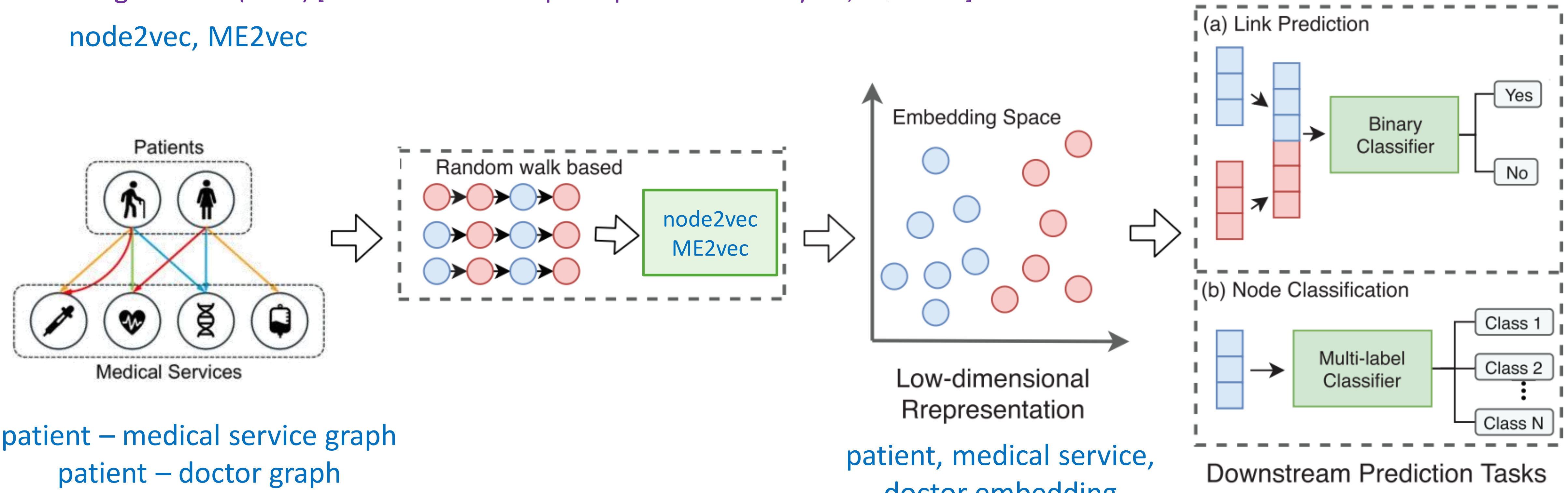


BFS and DFS biased random walks

Application: Health Records

Leveraging graph-based hierarchical medical entity embedding for healthcare applications
Tong Wu et.al (2021) [Nature Scientific Reports | Advanced Analytics, IQVIA Inc]

node2vec, ME2vec



patient – medical service graph
patient – doctor graph

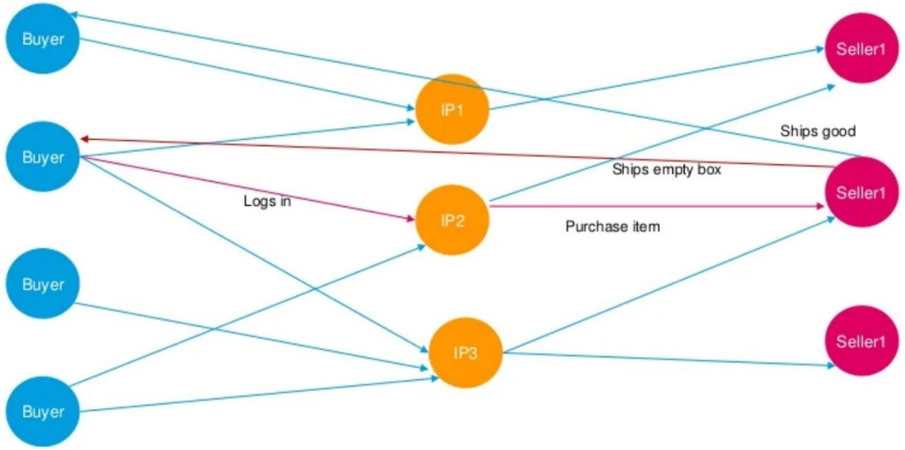
patient, medical service, doctor embedding

Prediction Tasks:

1. Predict patient diagnostic [node classification]
2. Predict if patient need to see a doctor [link prediction]
3. Readmission prediction [node classification]

https://www.nature.com/articles/s41598-021-85255-w.pdf
https://arxiv.org/abs/1906.05017

Application: PayPal's Collusion Fraud Prevention



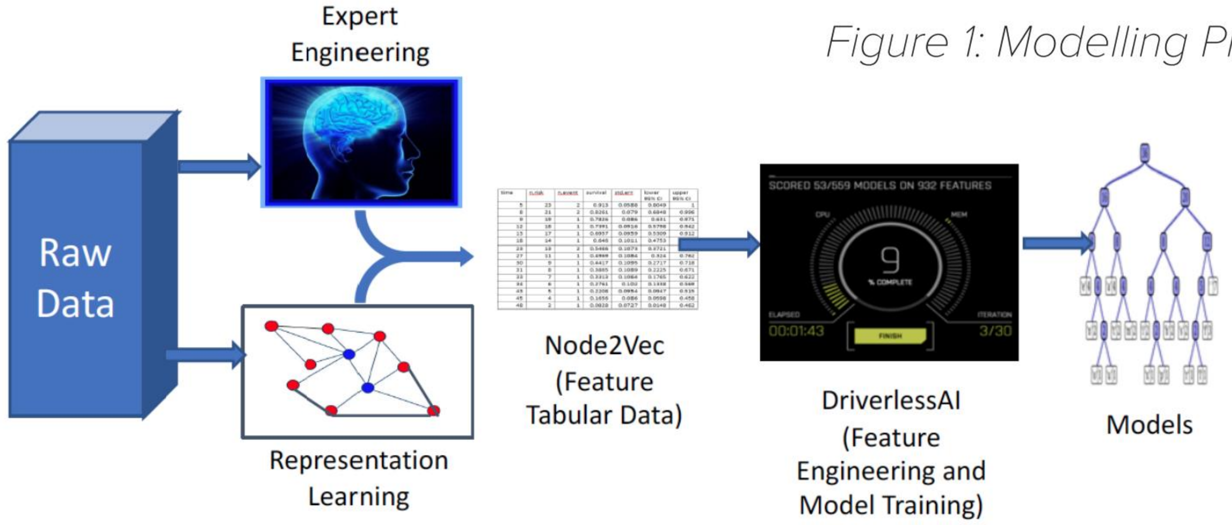


Figure 1: Modelling Pipeline

DATA

Training Data:

- Subset of one year's transactions.
- 1.5 billion edges, .5 million nodes.

Test Data:

- 3 months

Number of Features:

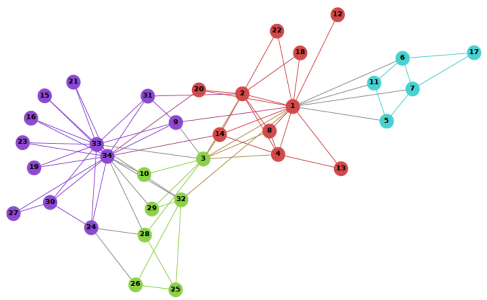
- 400-600

https://www.slideshare.net/Hadoop_Summit/graph-representation-learning-to-prevent-payment-collusion-fraud
https://www.h2o.ai/content/dam/h2o/en/marketing/documents/2020/01/PayPal-Customer-Case-Study-rnd2-1.pdf

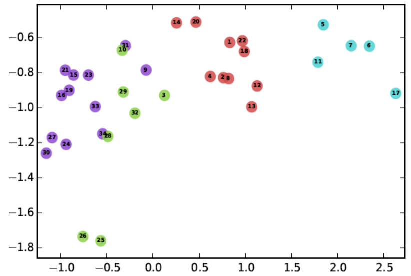
Graph Neural Networks

End-to-end learning for graph data

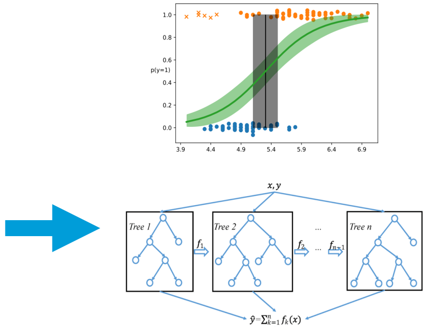
Node Embedding Limitations



Graph



Embedding



Classifier
(ML models)

Solving problems in two steps

Learn embedding first, then learn predictive model

Do not consider node features

Embeddings are generated solely based on the graph structure

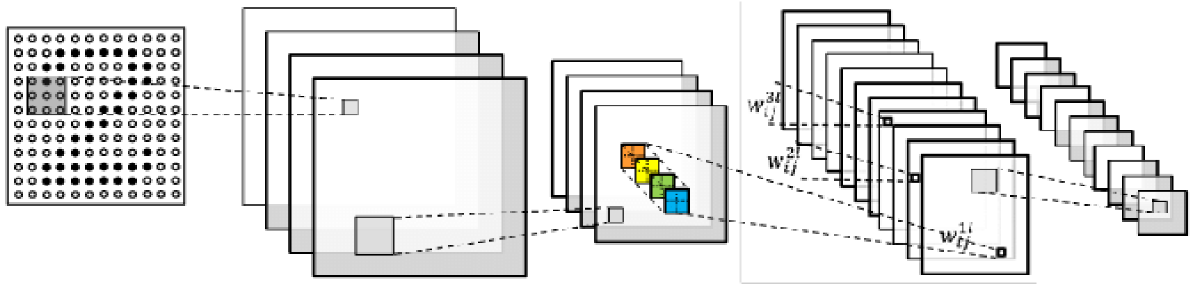
Transductive learning (instead of inductive)

Impossible to generate new embedding for new nodes not seen in the training

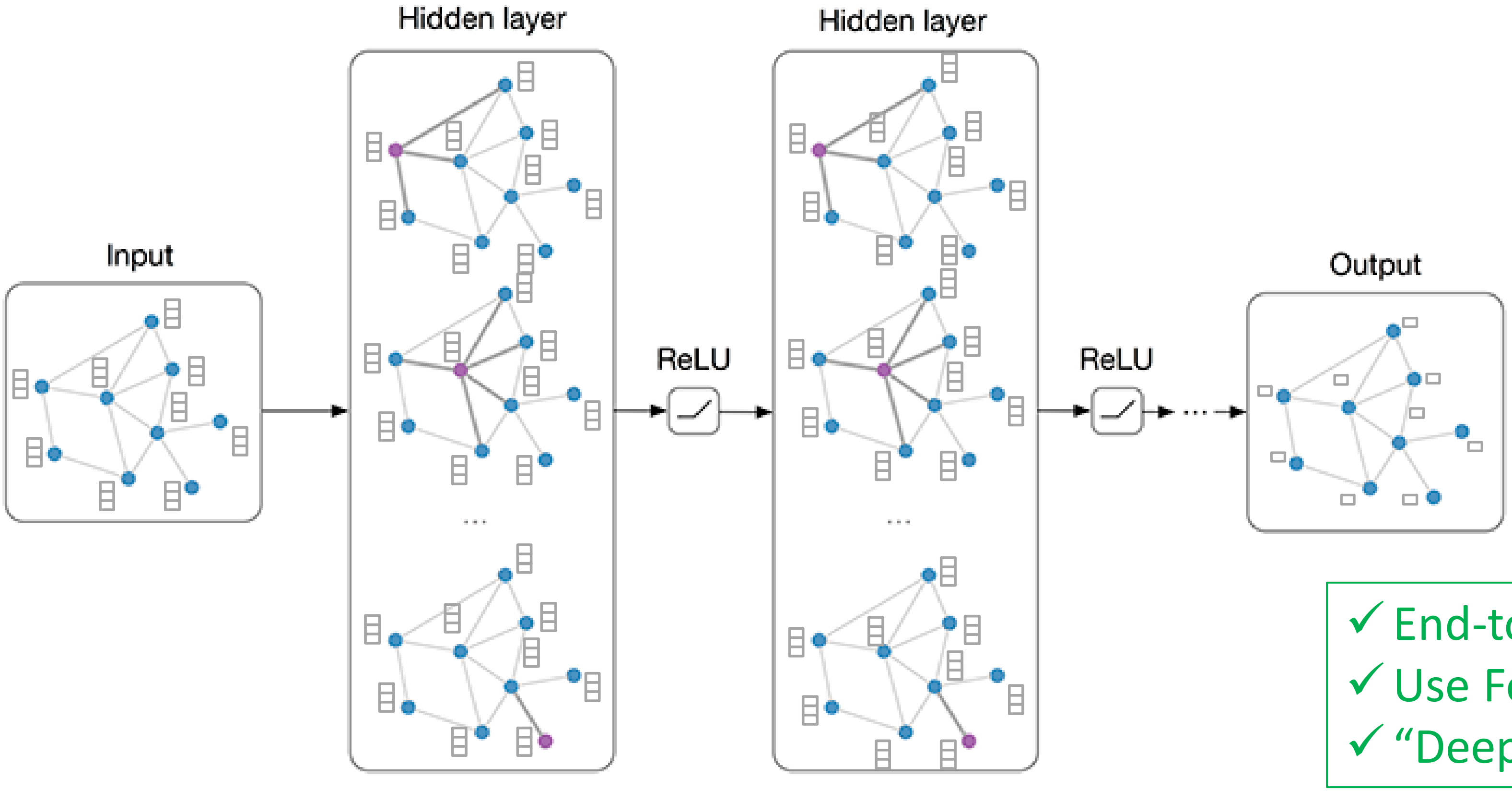
"Shallow" learning

Unable to take advantage of the representation power of deep neural networks

Graph Convolutional Networks



Convolutional Neural Networks (CNN)

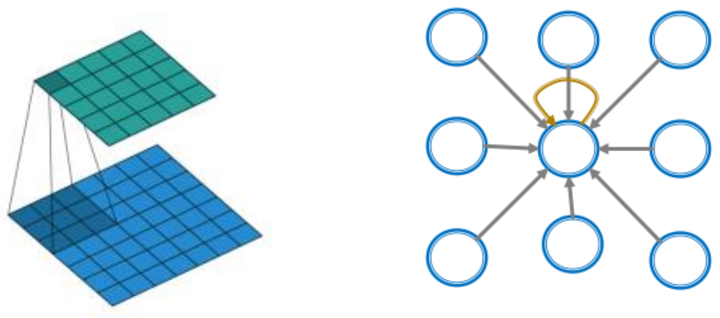


Graph Convolutional Networks (GCN)

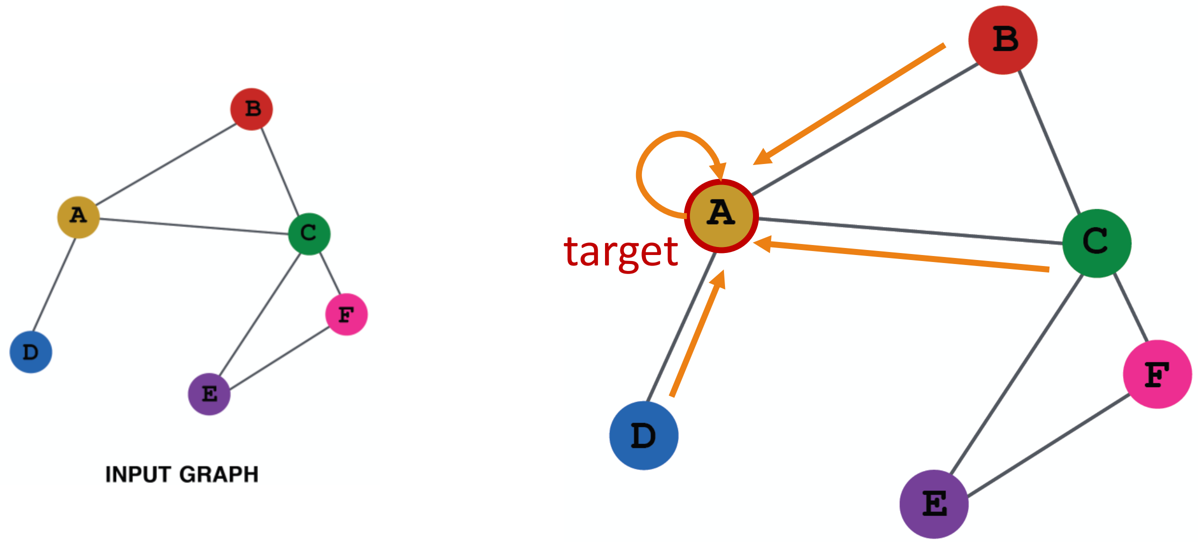
✓ End-to-end Learning
✓ Use Features
✓ “Deep” Learning

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

GCN Convolution Operator



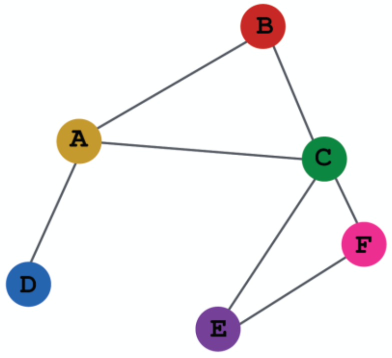
CNN layer with 3x3 filter computation flow



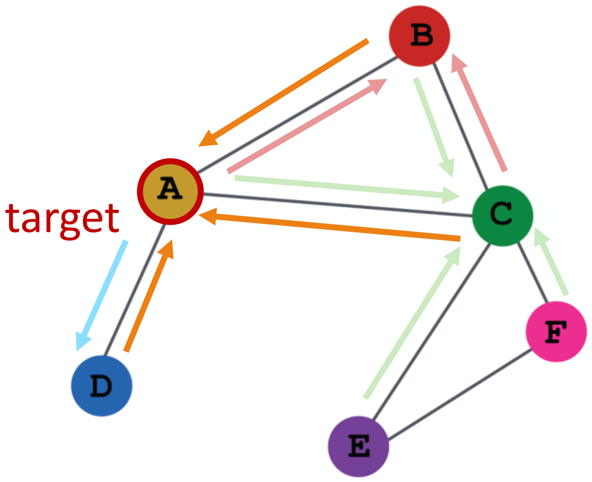
GCN computation flow

http://web.stanford.edu/class/cs224w/

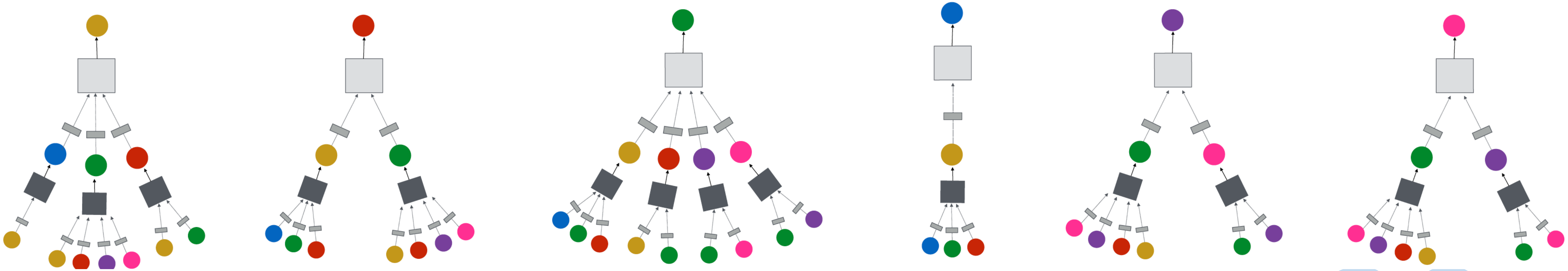
GCN Computation Flow



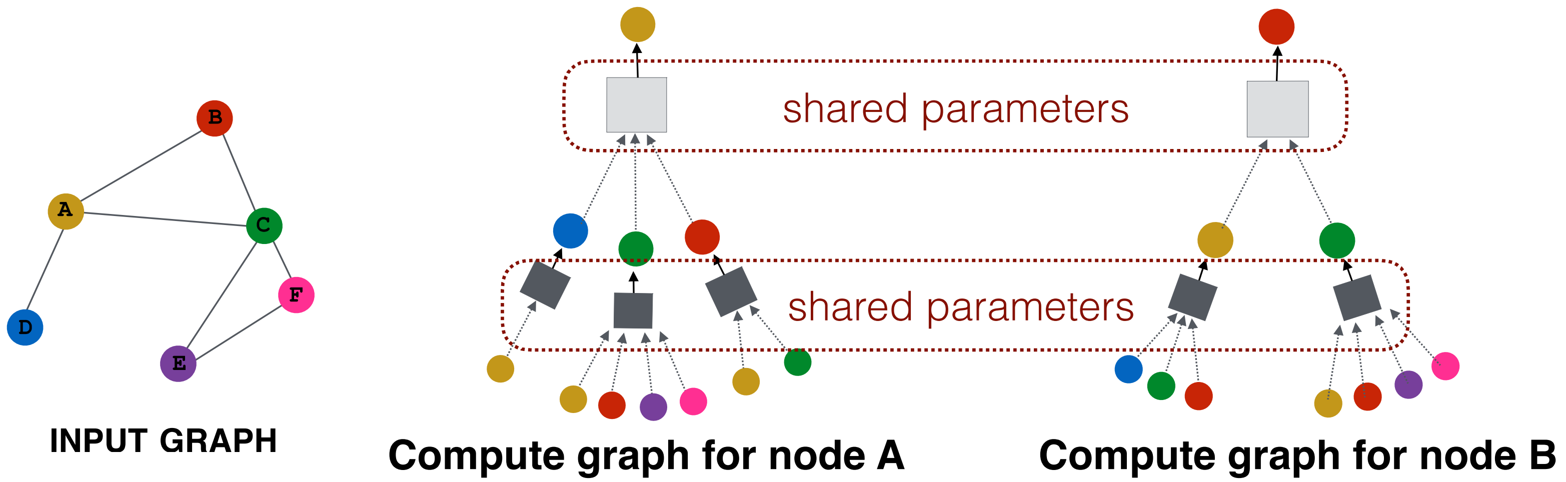
INPUT GRAPH



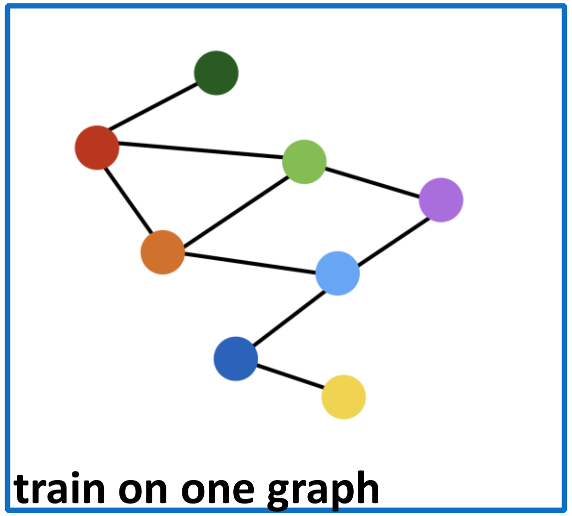
2 layers GCN computation flow

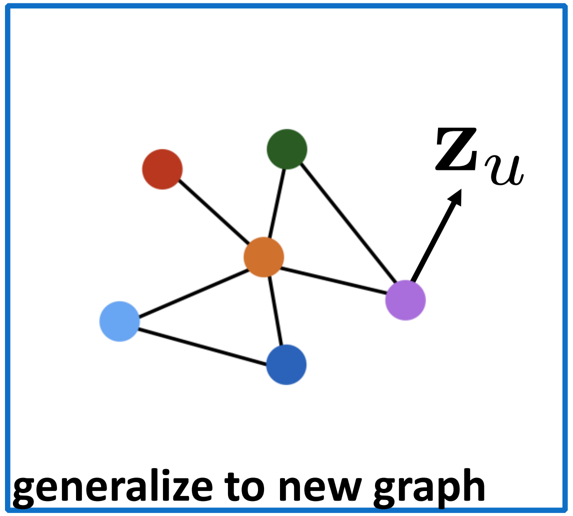


Inductive Capability



Weight Sharing

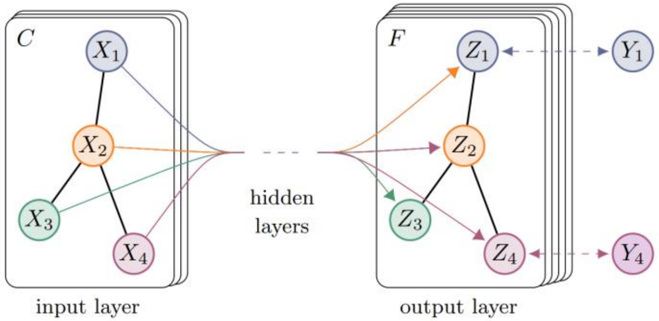




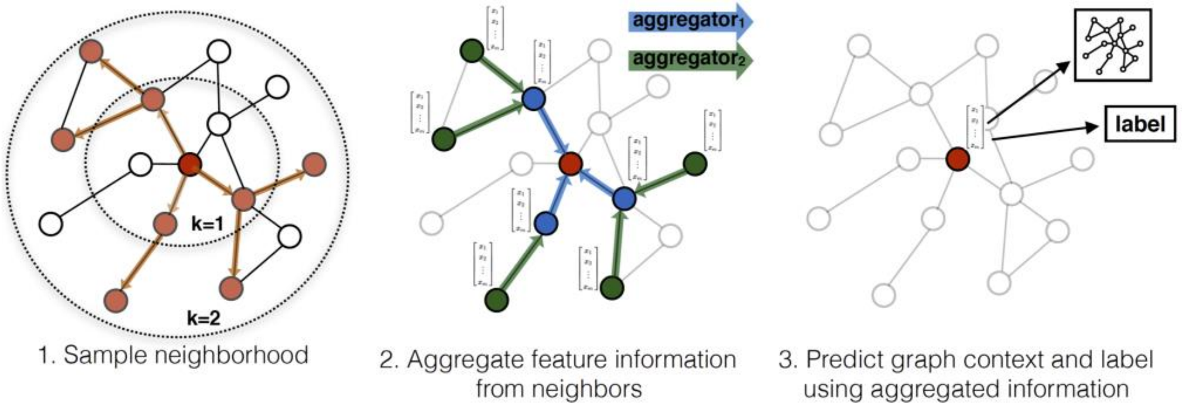
✓ Inductive

Applicable to unseen nodes

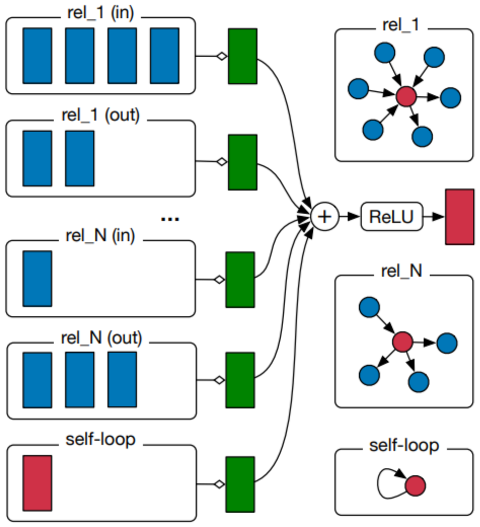
Some Flavors of Graph Neural Networks



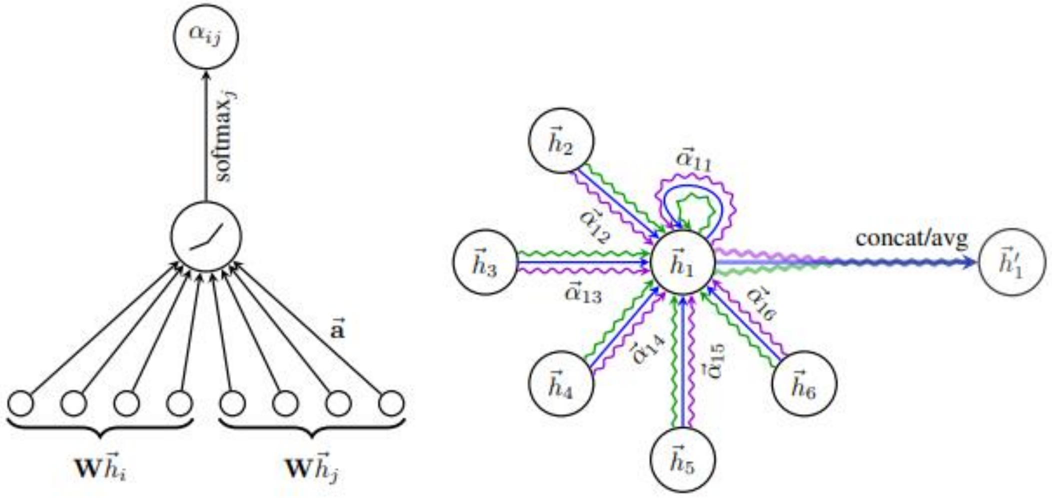
Graph Convolutional Networks (GCN)



GraphSAGE [Sample and Aggregate]



Relational GCN (R-GCN)



Graph Attention Networks (GAT)

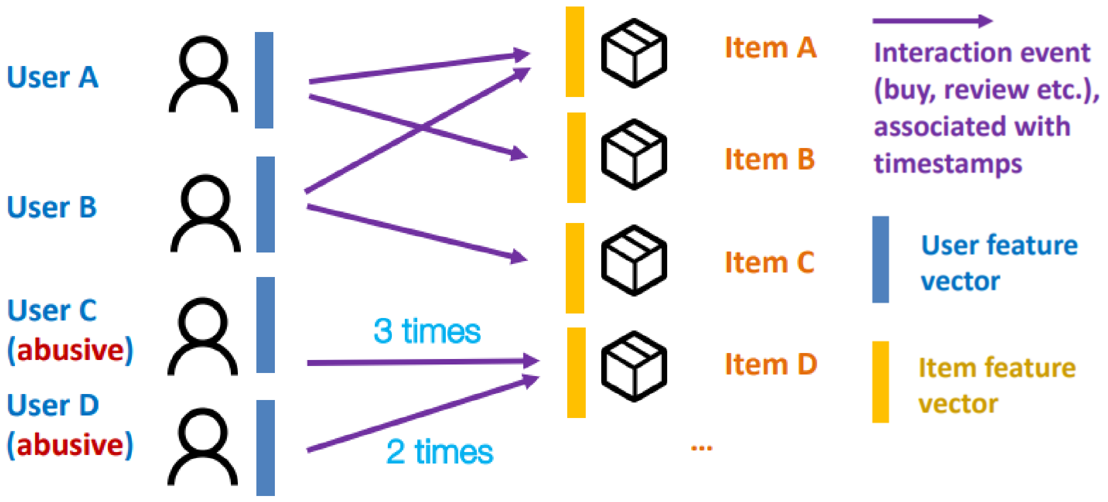
Application: Abuse Detection in Web

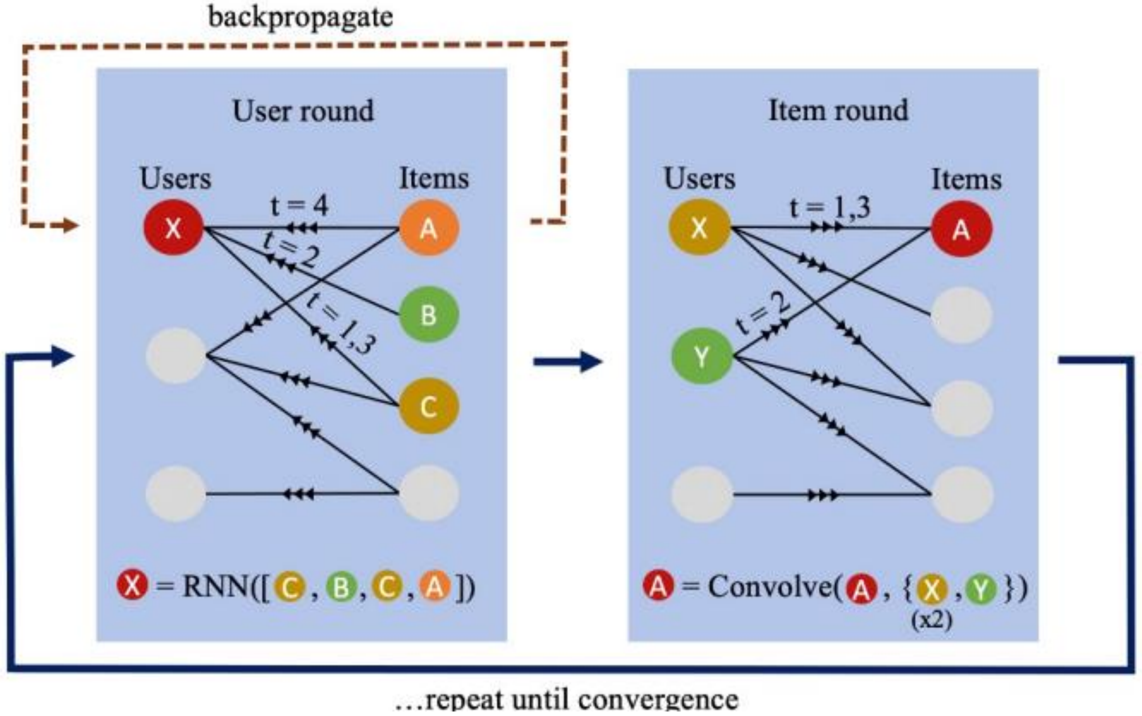
Bipartite Dynamic Representations for Abuse Detection (Andrew Wang, et.al, 2021) [KDD | Stanford U, Purdue U, Amazon]

WIKIPEDIA
The Free Encyclopedia

Trolling, propagating misinformation, offensive language

Fake reviews or purchases to inflate product rankings





Architecture GCN + RNN

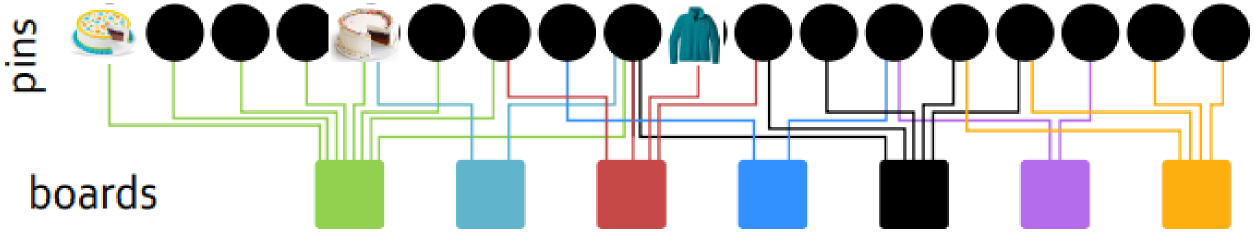
http://snap.stanford.edu/graphlearning-workshop/slides/stanford_graph_learning_Abuse_detection.pdf
https://cs.stanford.edu/people/jure/pubs/bydin-kdd21.pdf

Application: PinSAGE, Pinterest's Recommendation System

Large-scale GCN/GraphSAGE implementation

Contextual Image Recommendation

Graph Convolutional Neural Networks for Web-Scale Recommender Systems (Rex Ying et.al, 2018) [KDD | Pinterest, Stanford]



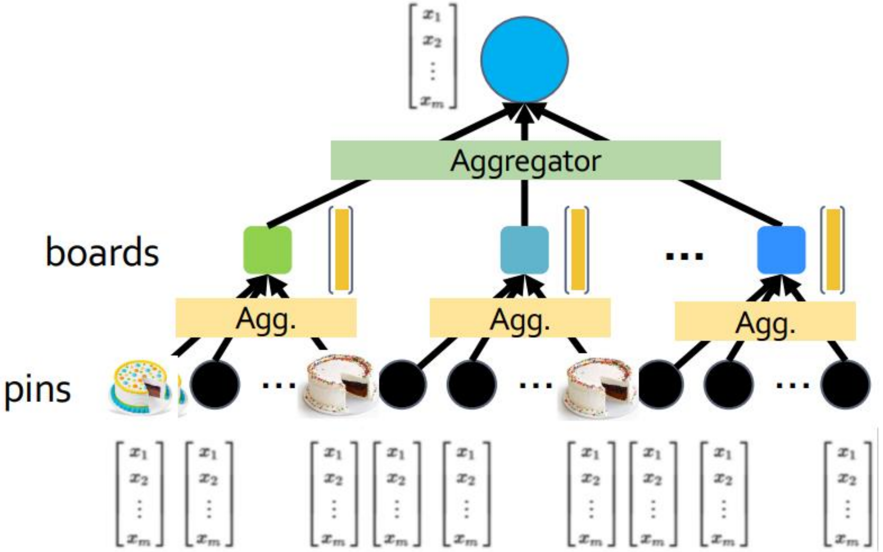
Pin
(image + desc.)

Recommend related pins

7.5 billion training data

1.2 billion positive pairs

6.5 billion negative pairs



Features: image embedding + text embedding

https://medium.com/pinterest-engineering/pinsage-a-new-graph-convolutional-neural-network-for-web-scale-recommender-systems-88795a107f48
https://arxiv.org/pdf/1806.01973.pdf

Implementation

Tools and Frameworks

Tools for Graph Neural Networks

Deep Learning Frameworks

PyTorch

TensorFlow

PyG

DGL DEEP GRAPH LIBRARY

Spektral

Graph Neural Network Frameworks

Code Example

PyTorch PyG

PyG is PyTorch-on-the-rocks:

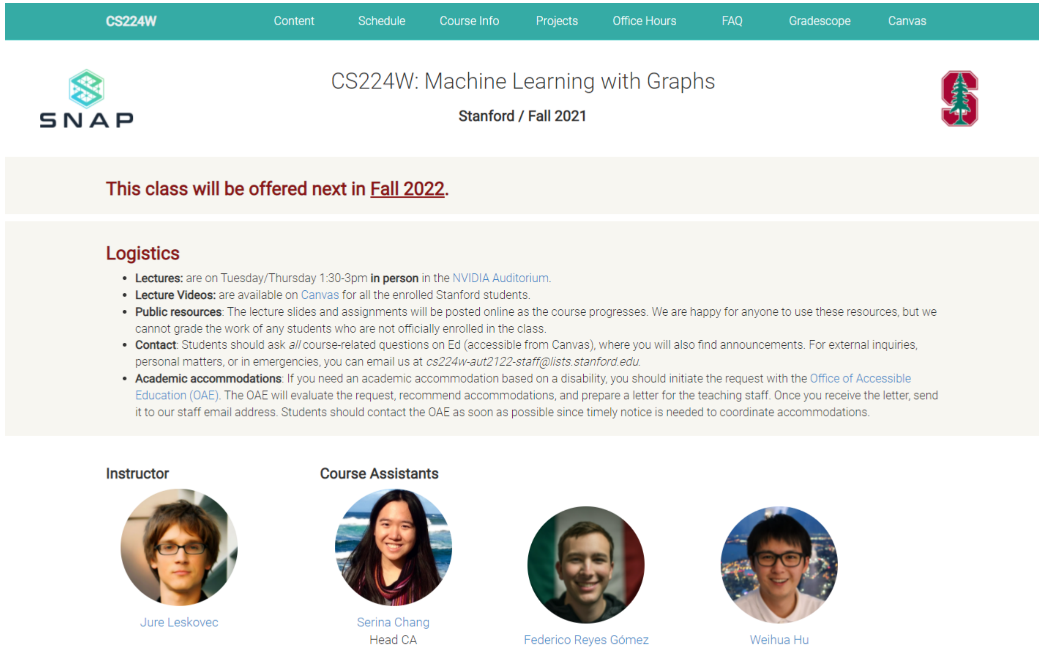
```
from torch.nn import Conv2d

class CNN(torch.nn.Module):
    def __init__(self):
        self.conv1 = Conv2d(3, 64)
        self.conv2 = Conv2d(64, 64)

    def forward(self, input):
        h = self.conv1(input)
        h = h.relu()
        h = self.conv2(h)
        return h
```

```
from torch_geometric.nn import GCNConv

class GNN(torch.nn.Module):
    def __init__(self):
        self.conv1 = GCNConv(3, 64)
        self.conv2 = GCNConv(64, 64)

    def forward(self, input, edge_index):
        h = self.conv1(input, edge_index)
        h = h.relu()
        h = self.conv2(h, edge_index)
        return h
```


http://snap.stanford.edu/graphlearning-workshop/slides/stanford_graph_learning_PyG_2_0.pdf

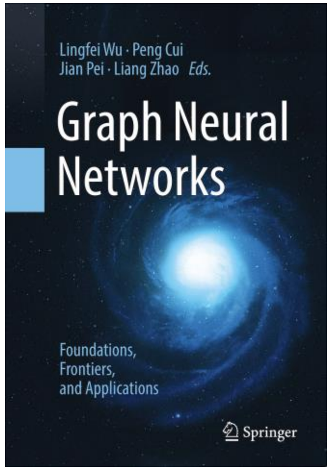
Learning & Resources

Stanford's Machine Learning with Graphs class



Course Slides, Video Lectures

Comprehensive resources for Graph ML from
Jure Leskovec, one of the authorities on Graph ML

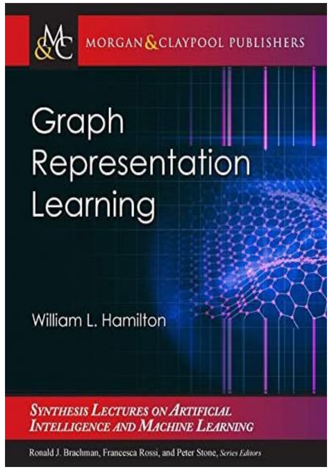


Graph Neural Networks Foundation, Frontier, and Applications

Lingfei Wu et. al.

Comprehensive, focus on applications and use cases

Free pre-print version is available



Graph Representation Learning

William L. Hamilton

Foundational, focus on building conceptual understanding

Free pre-print version is available

Thank You