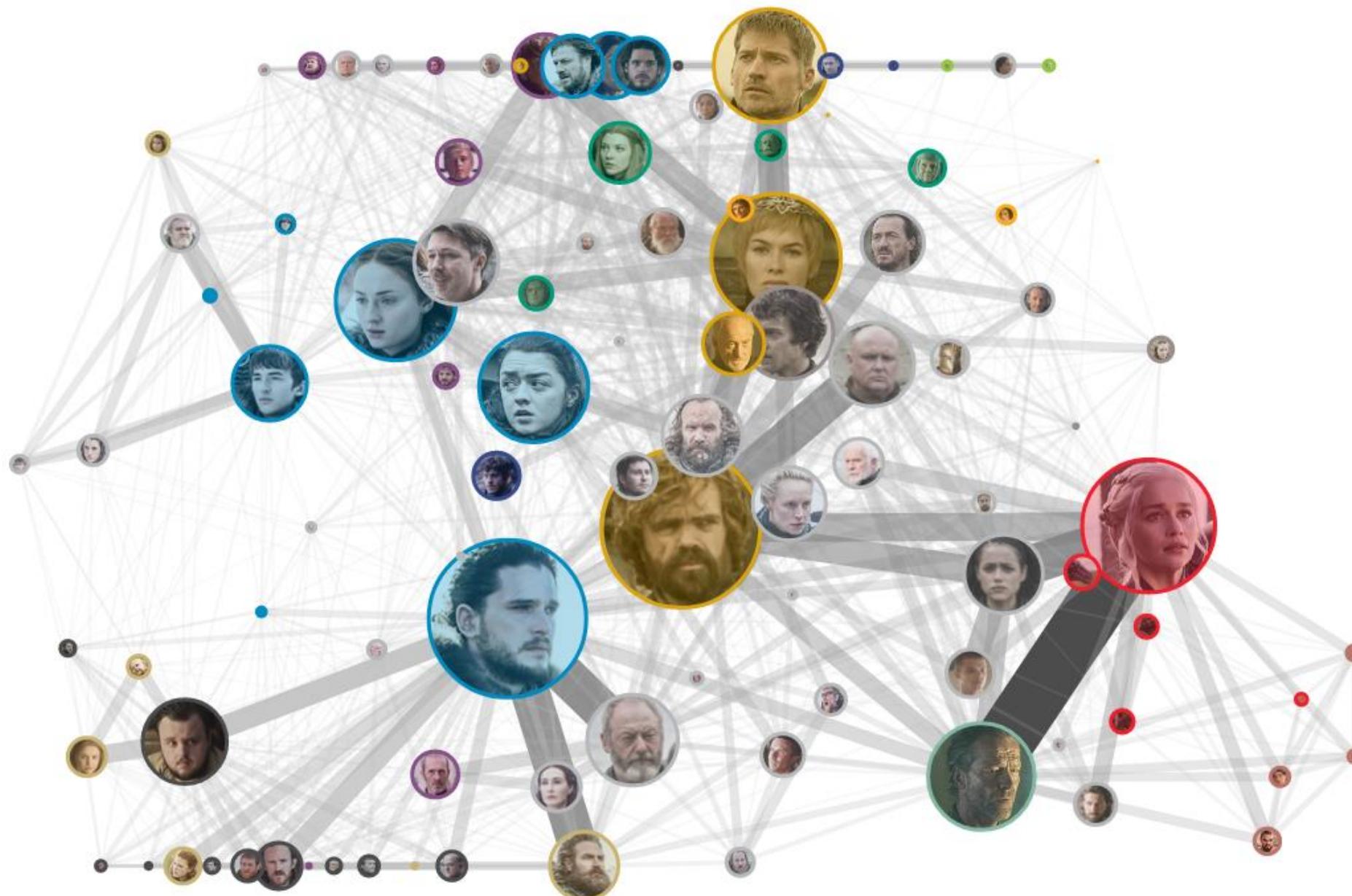


Graph Machine Learning: Foundations and Applications (AI60007)

Introduction to
Graph Machine
Learning

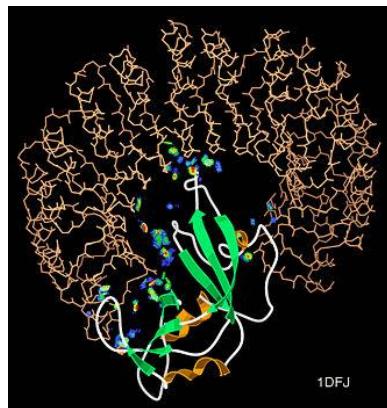
GML-FA, Autumn, 2022

Who is the real main character?

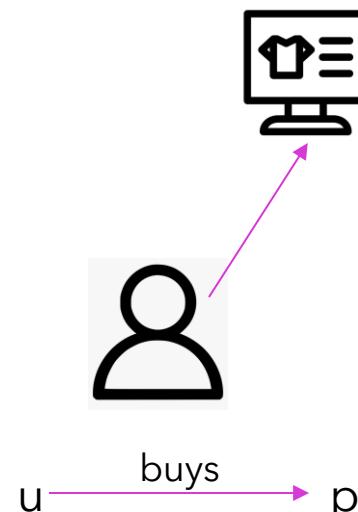
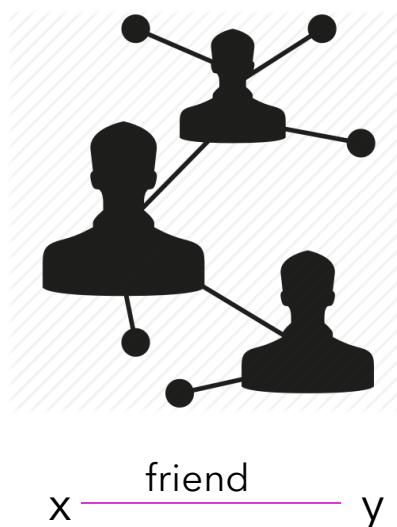


Graph: Universal Language for Data

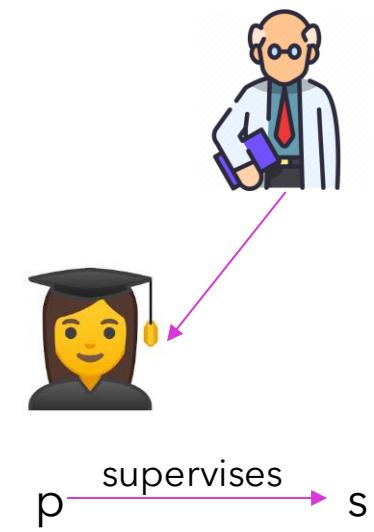
"In many ways, graphs are the main modality of data we receive from nature."



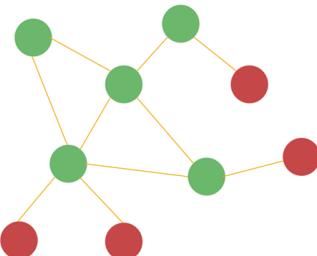
P1 $\xrightarrow{\text{interacts}}$ P2



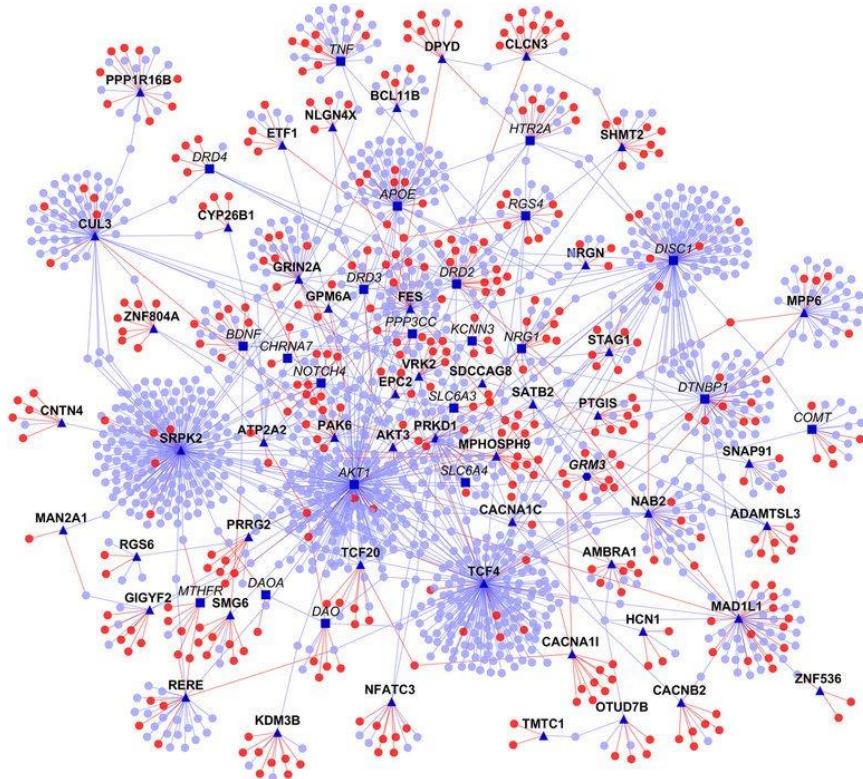
u $\xrightarrow{\text{buys}}$ p



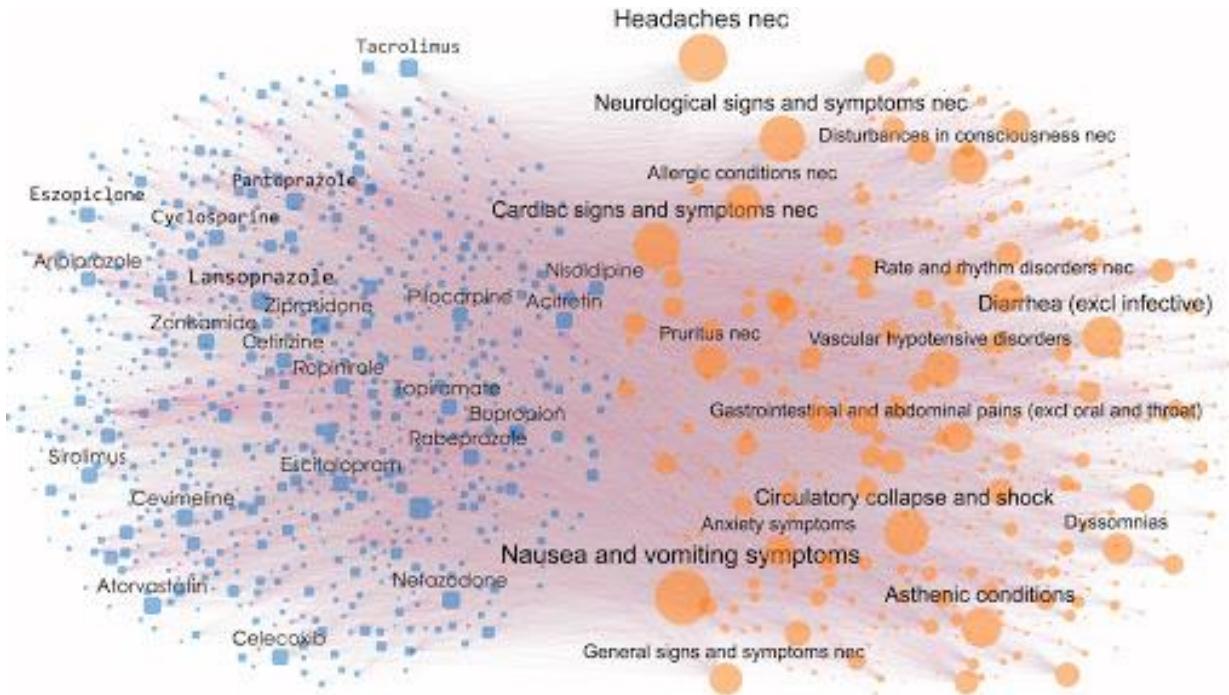
p $\xrightarrow{\text{supervises}}$ s



Real World Graphs



Protein-Protein Interaction



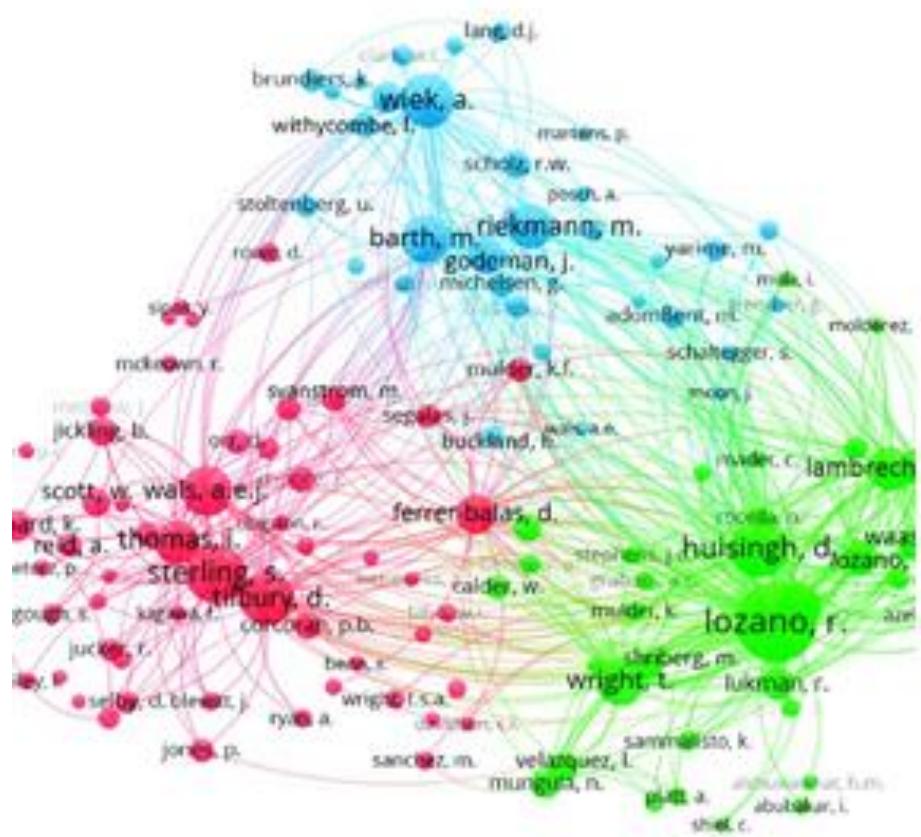
Drug Side effect Network

Real World Graphs

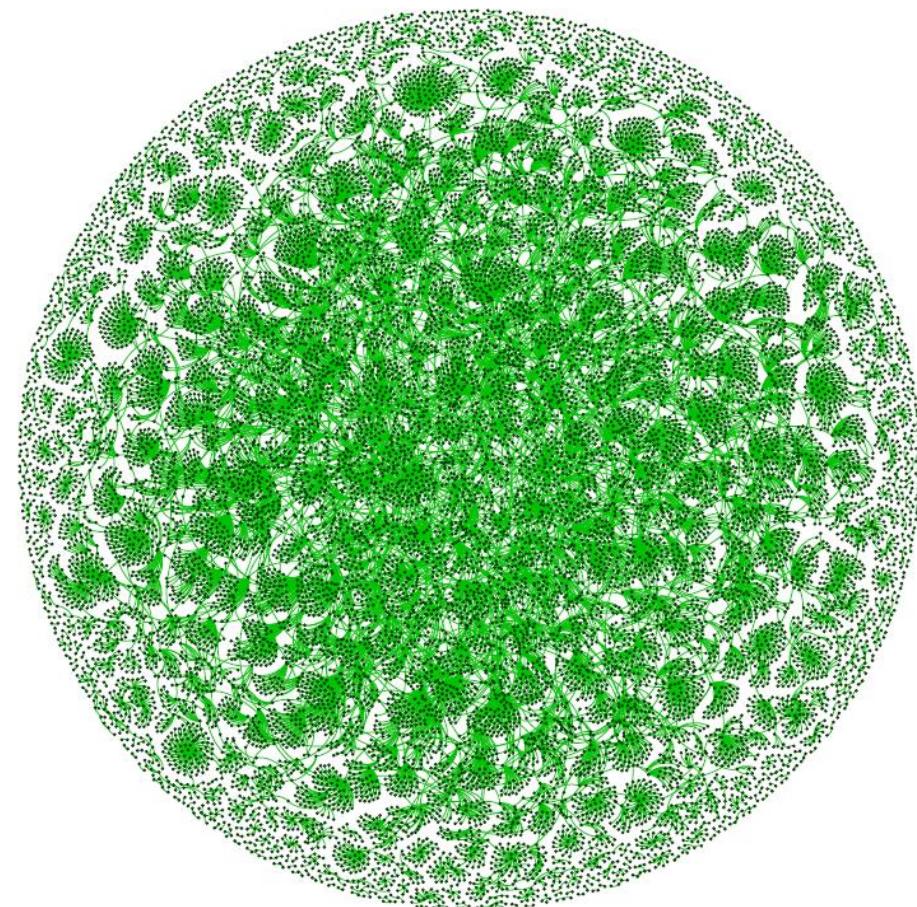


Transportation Network

Real World Graphs

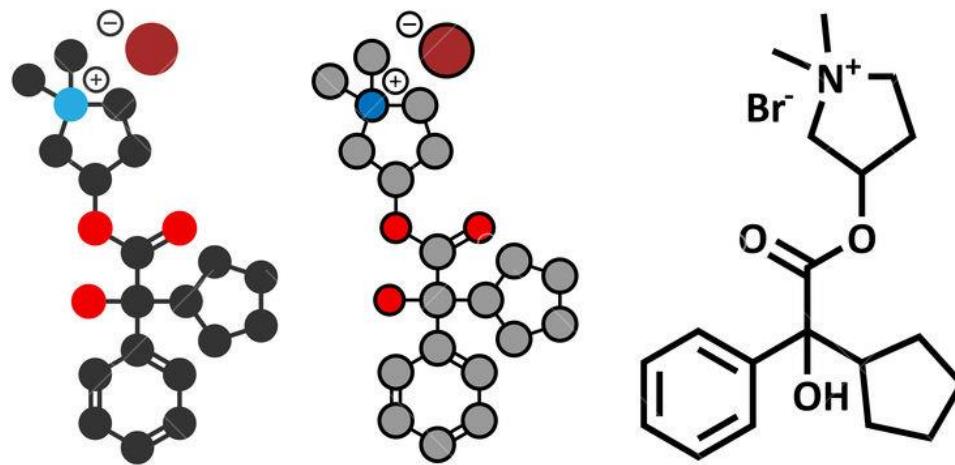


Citation Network



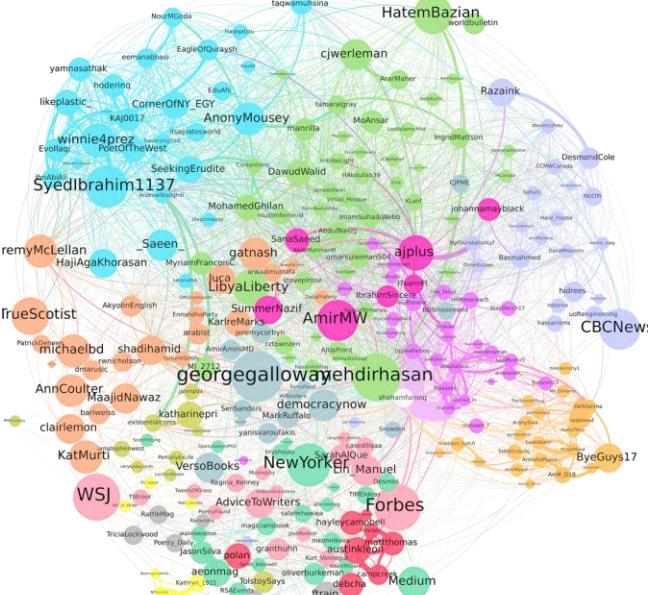
Academic Genealogy Network

Real World Graphs

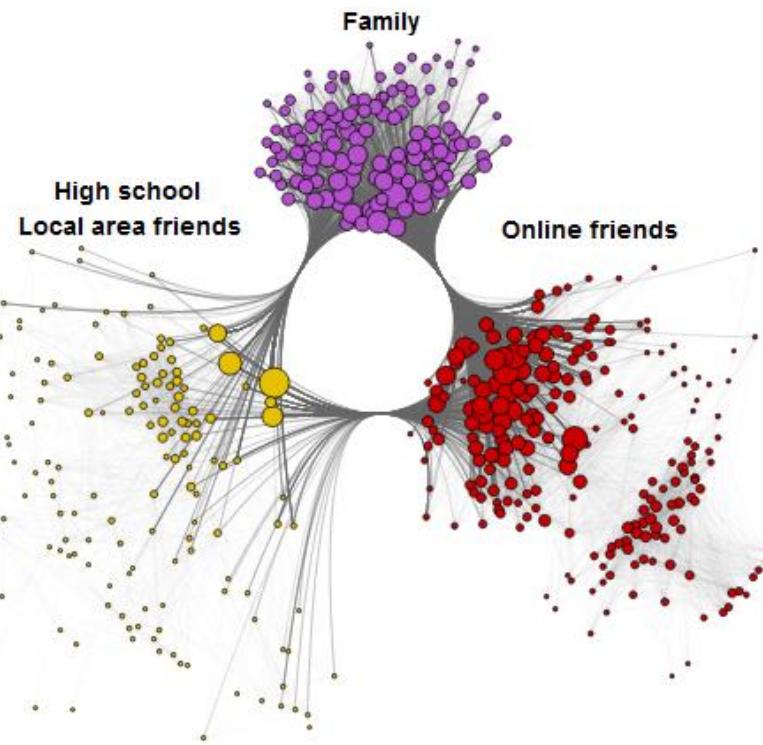


Drug molecule

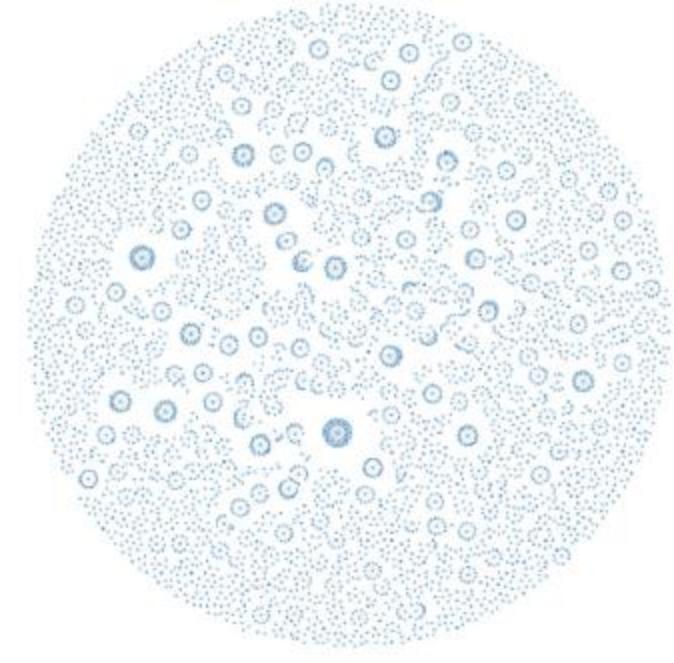
Real World Graphs



Twitter
Network

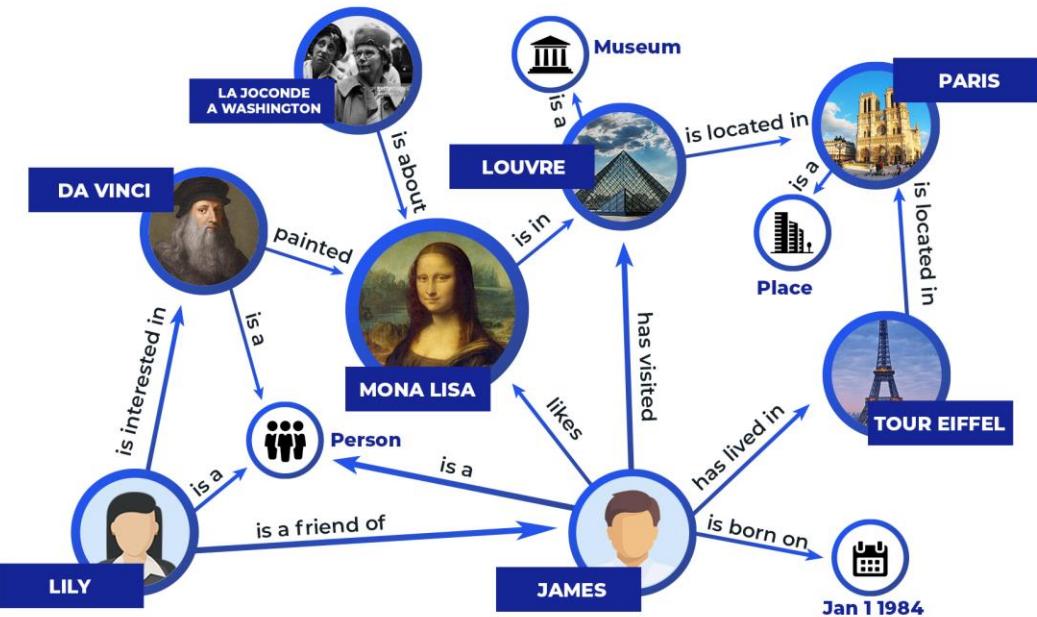


Communities in
Facebook
Network



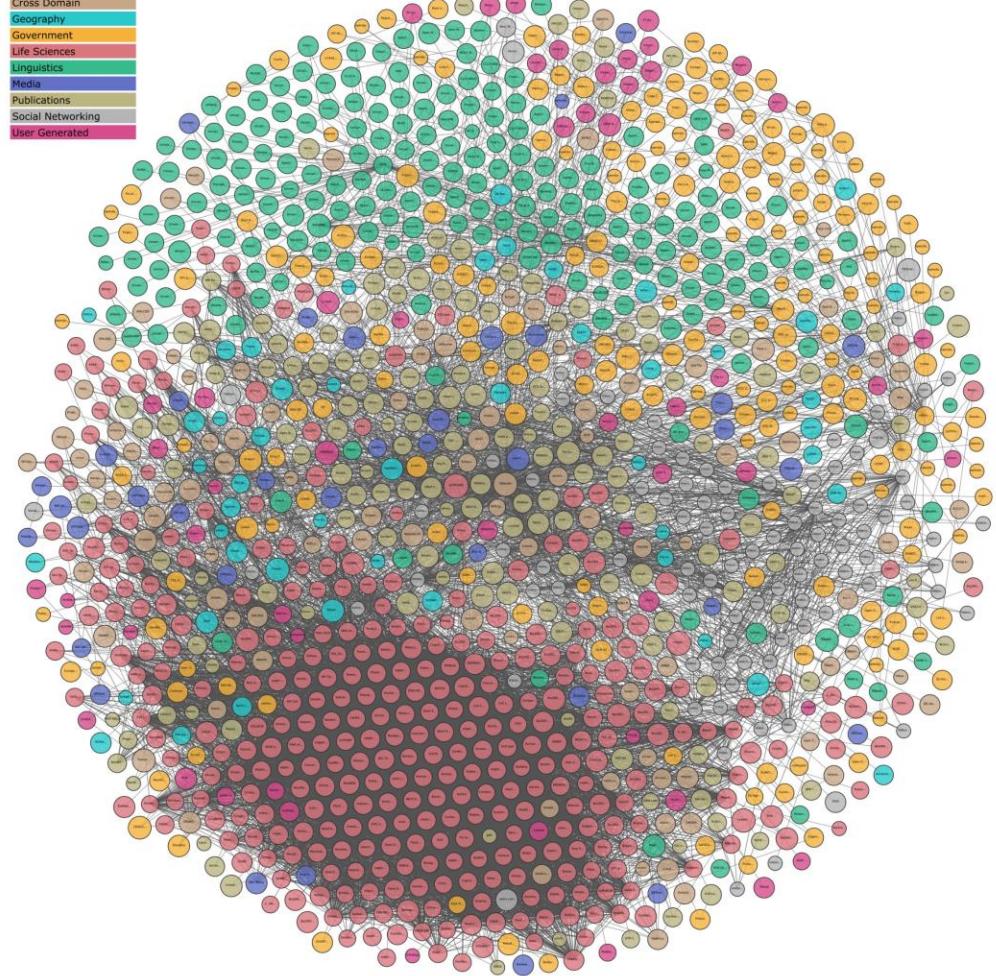
Product co-
purchasing
network

Real World Graphs



Legend

- Cross Domain
- Geography
- Government
- Life Sciences
- Linguistics
- Media
- Publications
- Social Networking
- User Generated

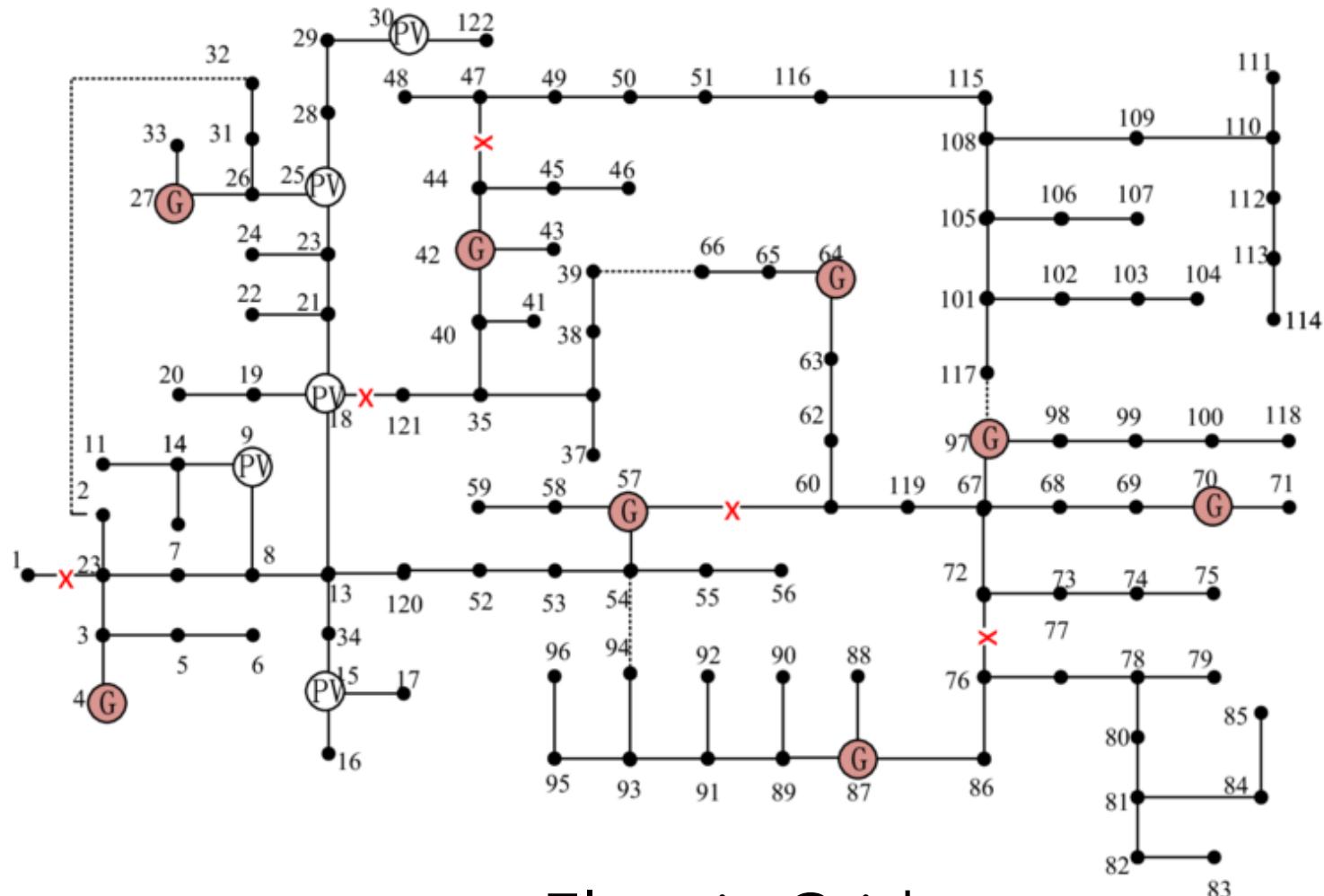


The Linked Open Data Cloud from lol-cloud.net



Knowledge Graph: The encyclopedia for the machines

Real World Graphs

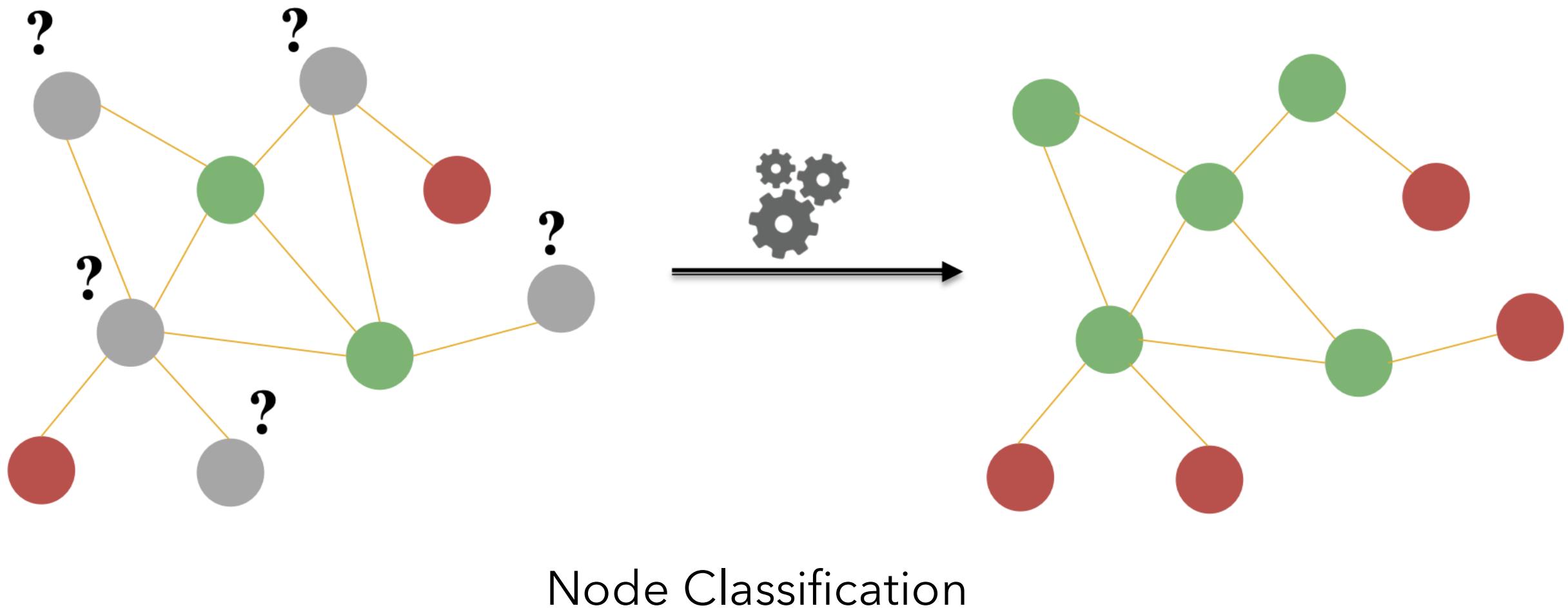


Graph Analytics

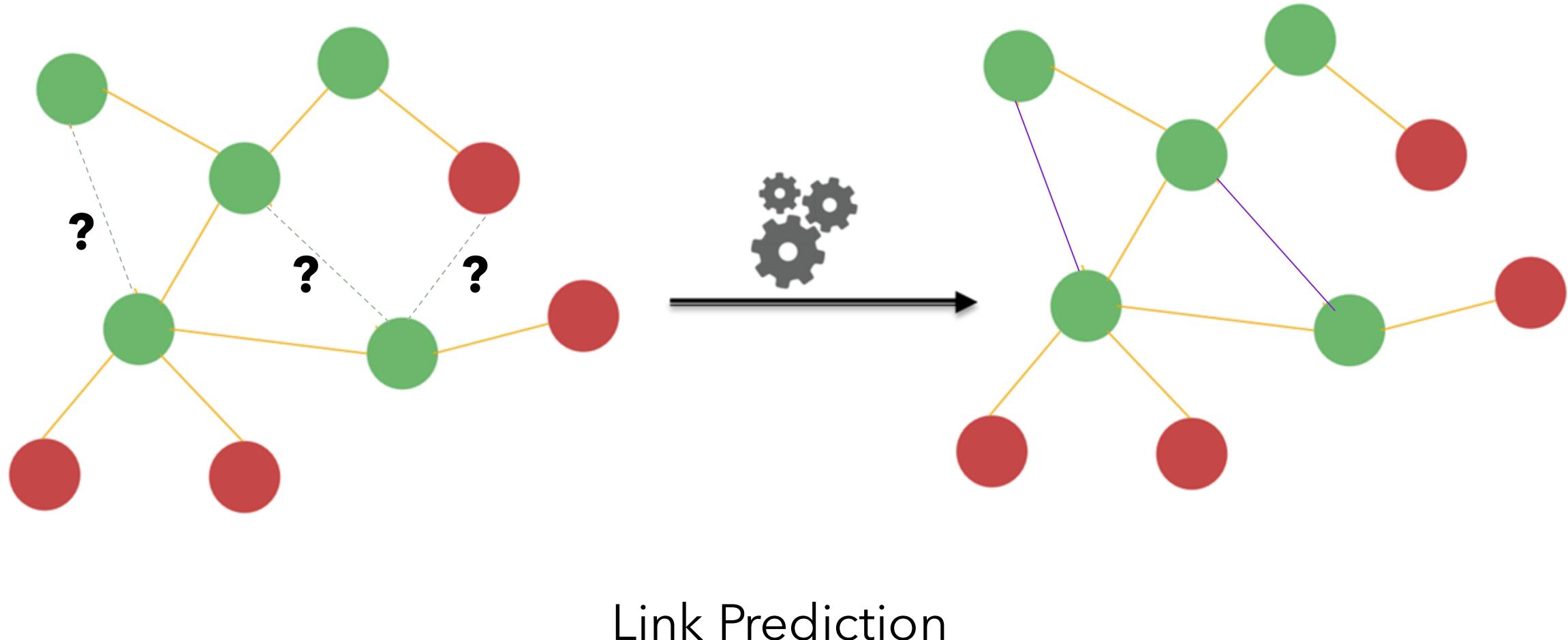
Structural
Analysis

Predictive
Modelling

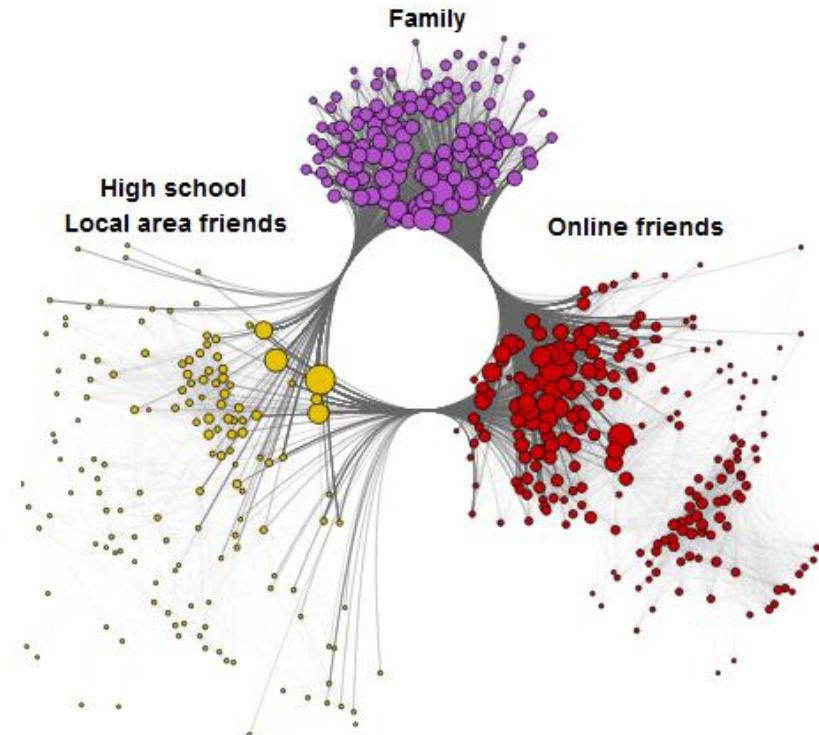
Machine Learning Tasks on Graph



Machine Learning Tasks on Graph

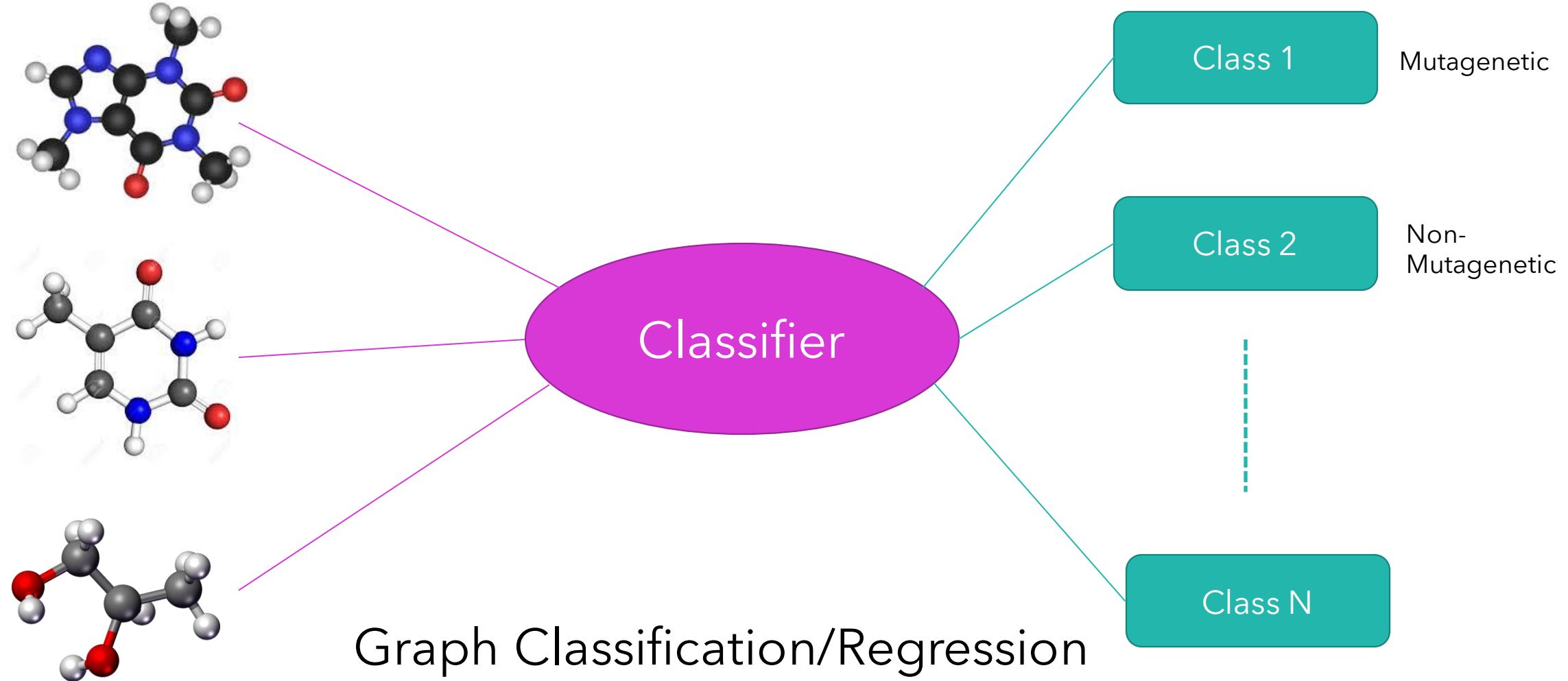


Machine Learning Tasks on Graph

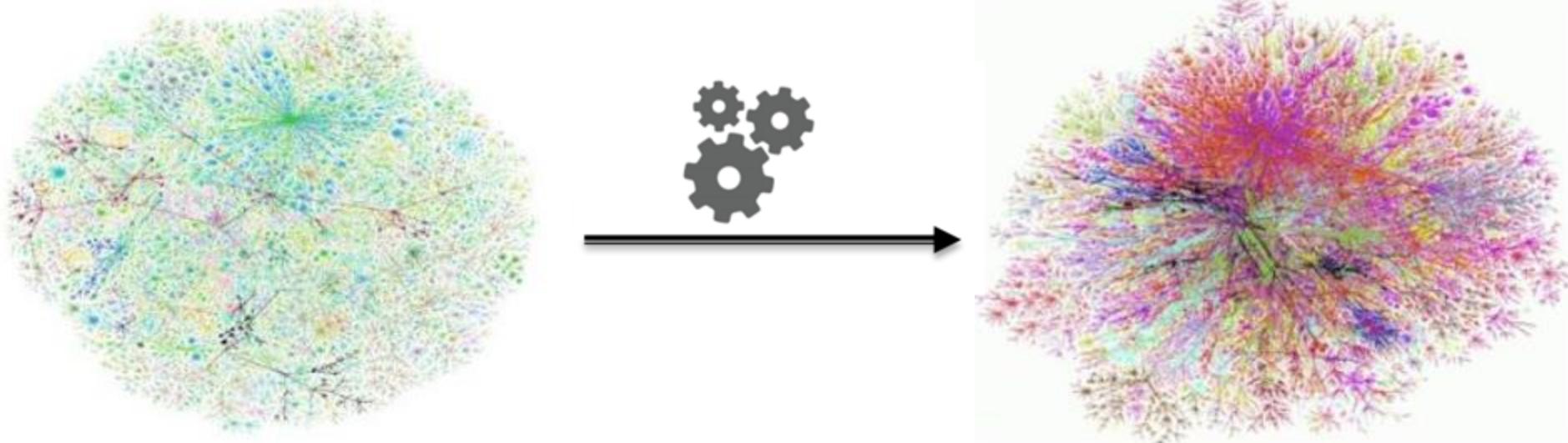


Community Detection or Graph Clustering

Machine Learning Tasks on Graph



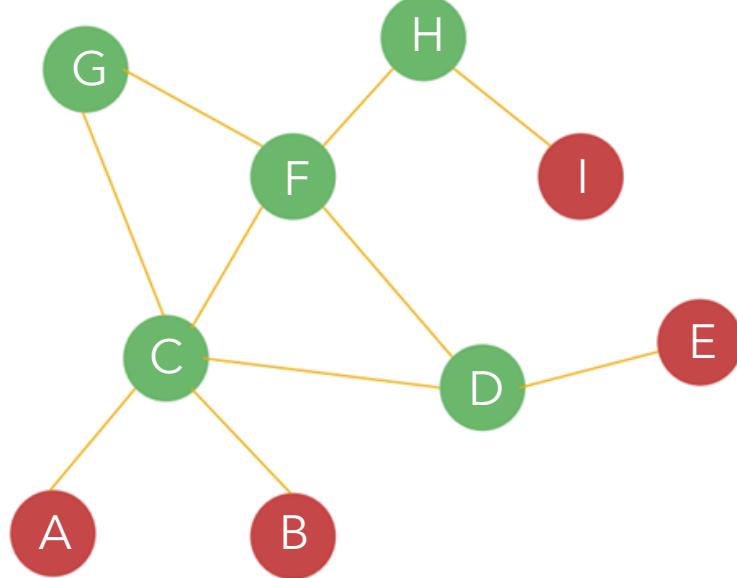
Machine Learning Tasks on Graph



Realistic graph generation: Given a set of graphs, generate similar graphs

Goal directed graph generation: Generate graph that satisfies constraints/optimizes objective

Graph Data Representation: Sparse

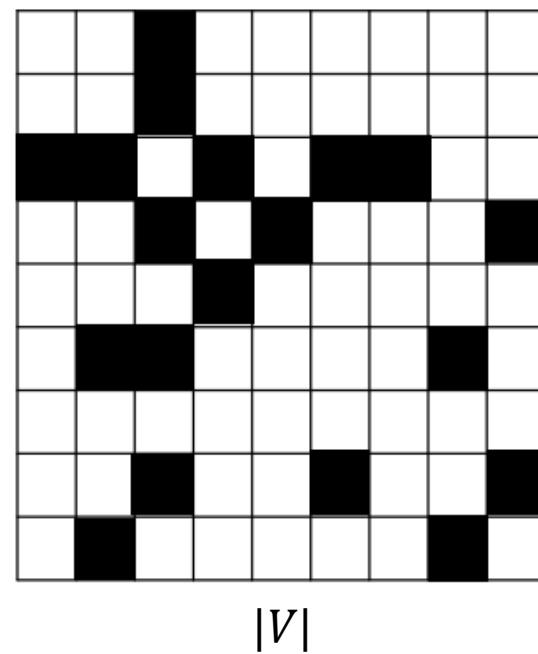
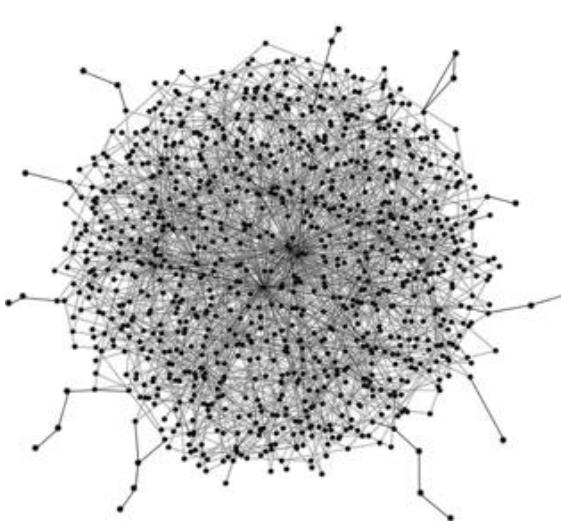


	A	B	C	D	E	F	G	H	I
A			1						
B			1						
C	1	1		1		1	1		
D			1	1					
E				1					
F		1	1					1	
G									
H					1			1	
I							1		

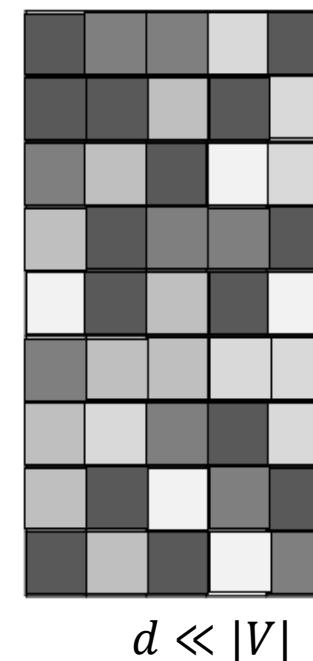
$\text{Sim}(\vec{C}, \vec{D})$

Sparse vectors are not reliable in estimating the similarity/distance
Inefficiency in processing large graphs

Graph Data Representation: Dense



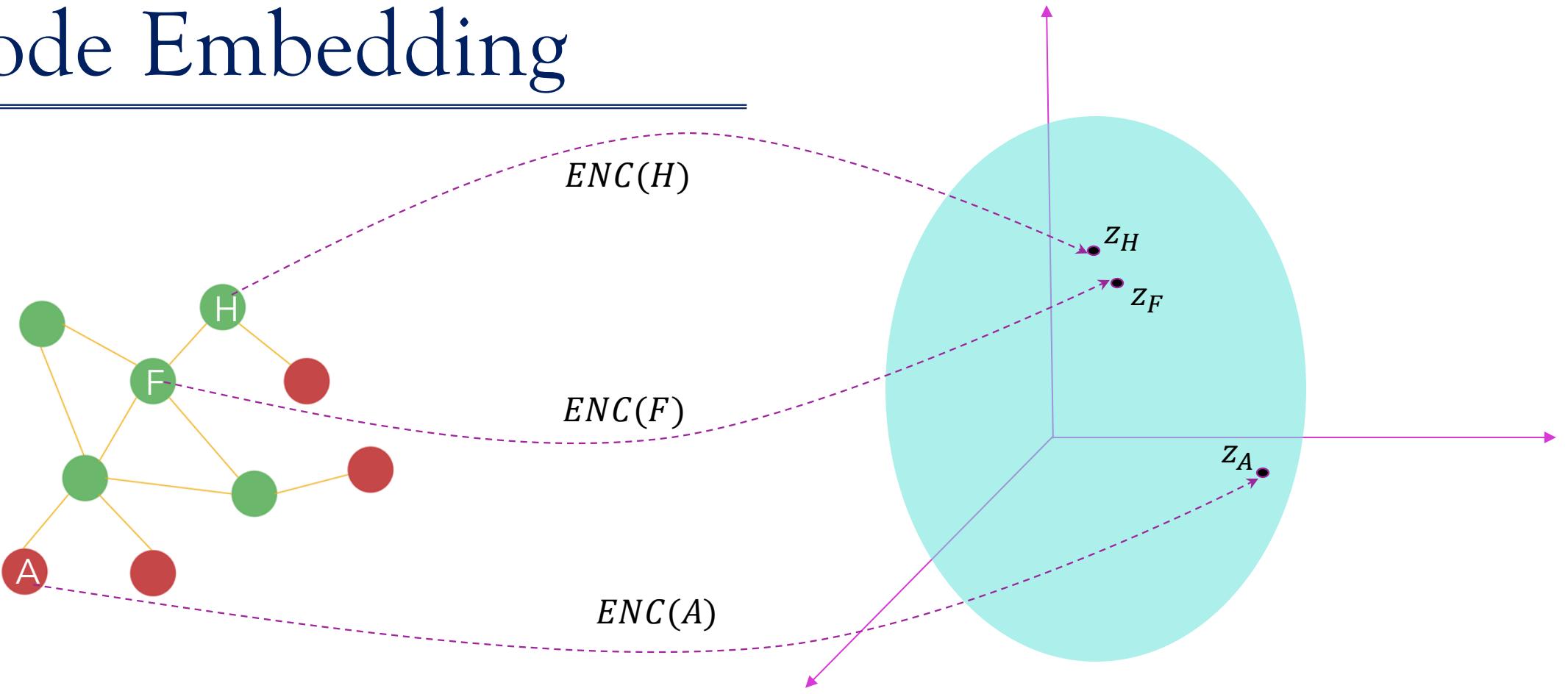
Latent
Dimensions



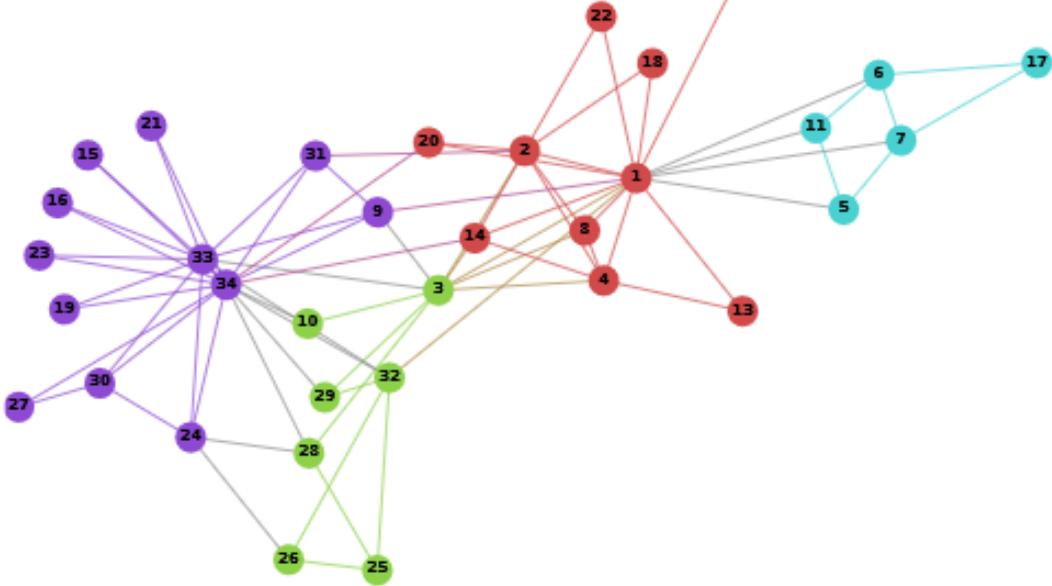
Downstream
Tasks

- Node classification
- Link prediction
- Graph clustering

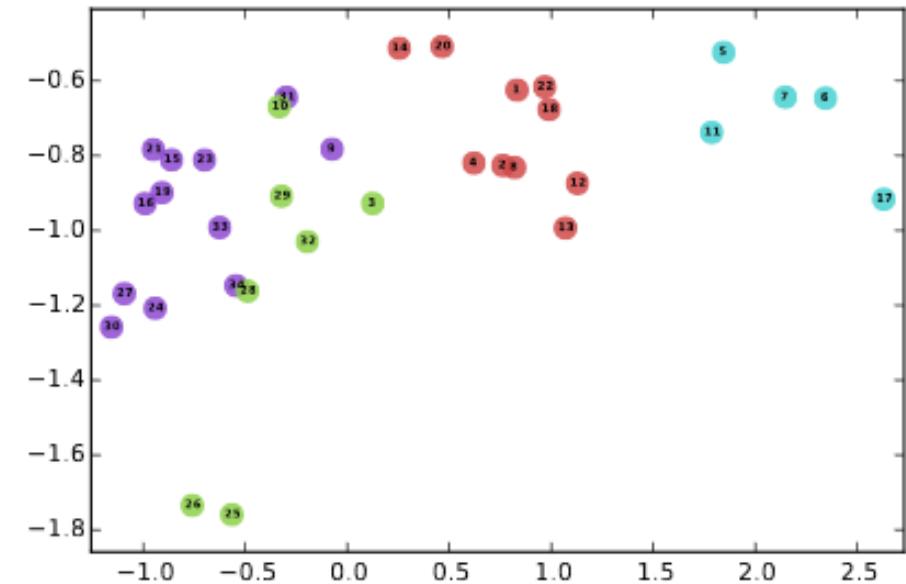
Node Embedding



Node Embedding



DeepWalk,
2013



Node2Vec,
2016

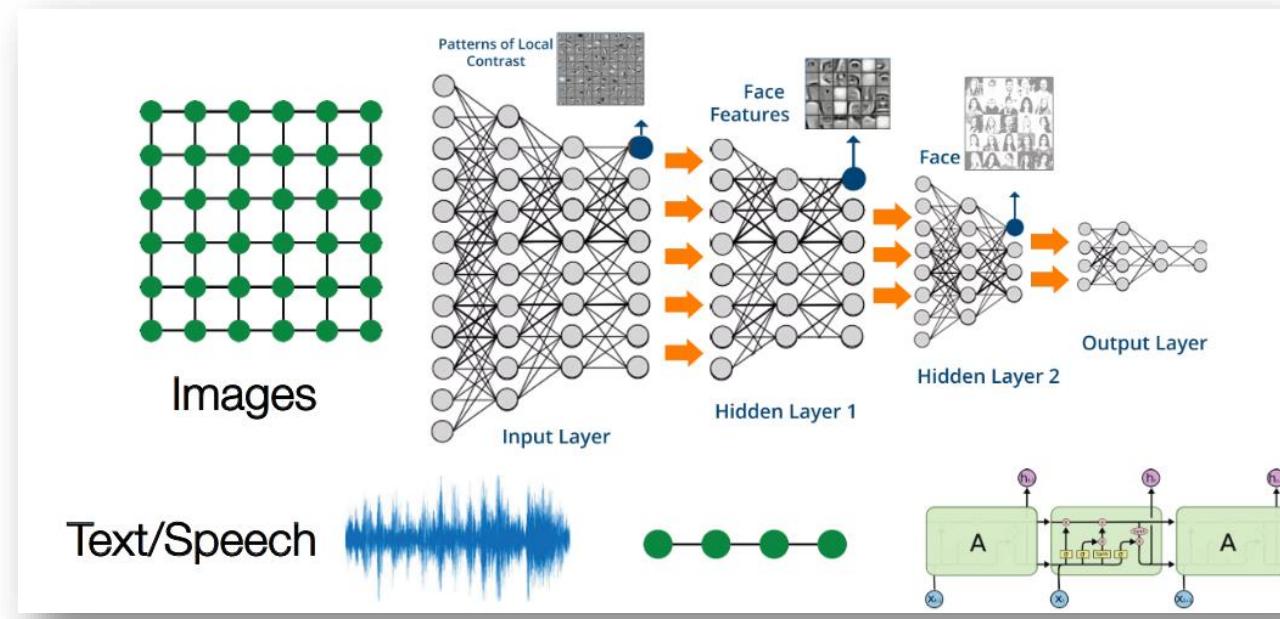
Struc2Vec,
2017

TransE,
2013

TransR,
2013

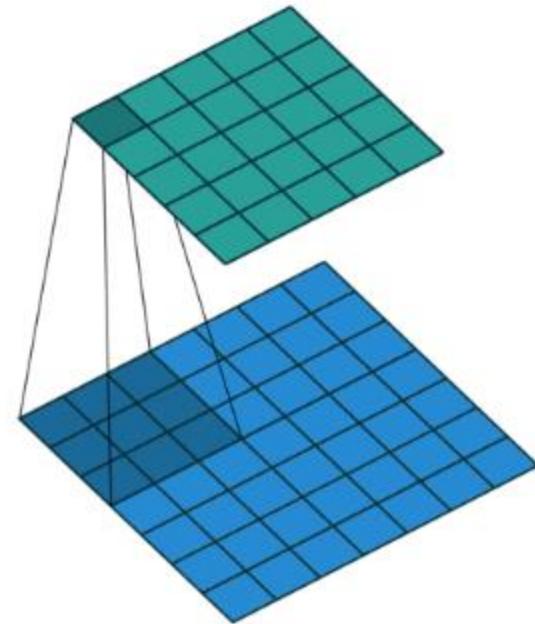
Deep Network for Graphs

$ENC(v)$ = Multiple layers of non-linear transformations over the graph structure

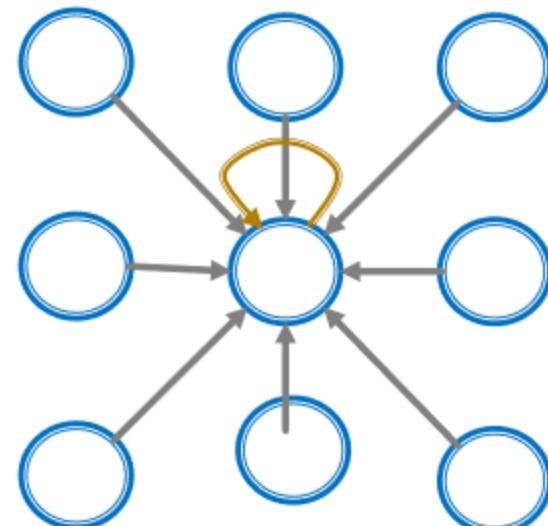


Graphs do not have regular structures like image, speech or language

Deep Network for Graphs: The Basic Idea



Image

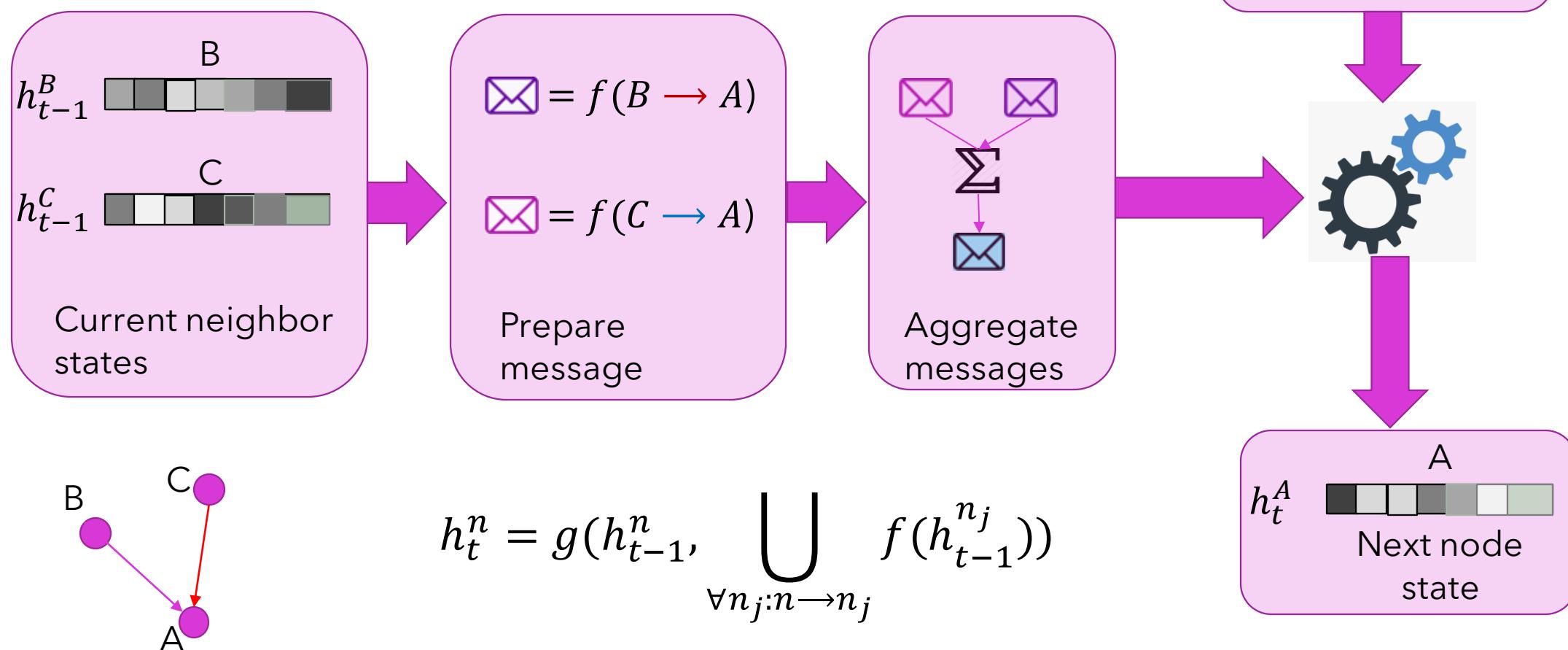


Graph

Transform messages ($W_i h_i$) from the neighbors and then combine ($\Sigma_i W_i h_i$)

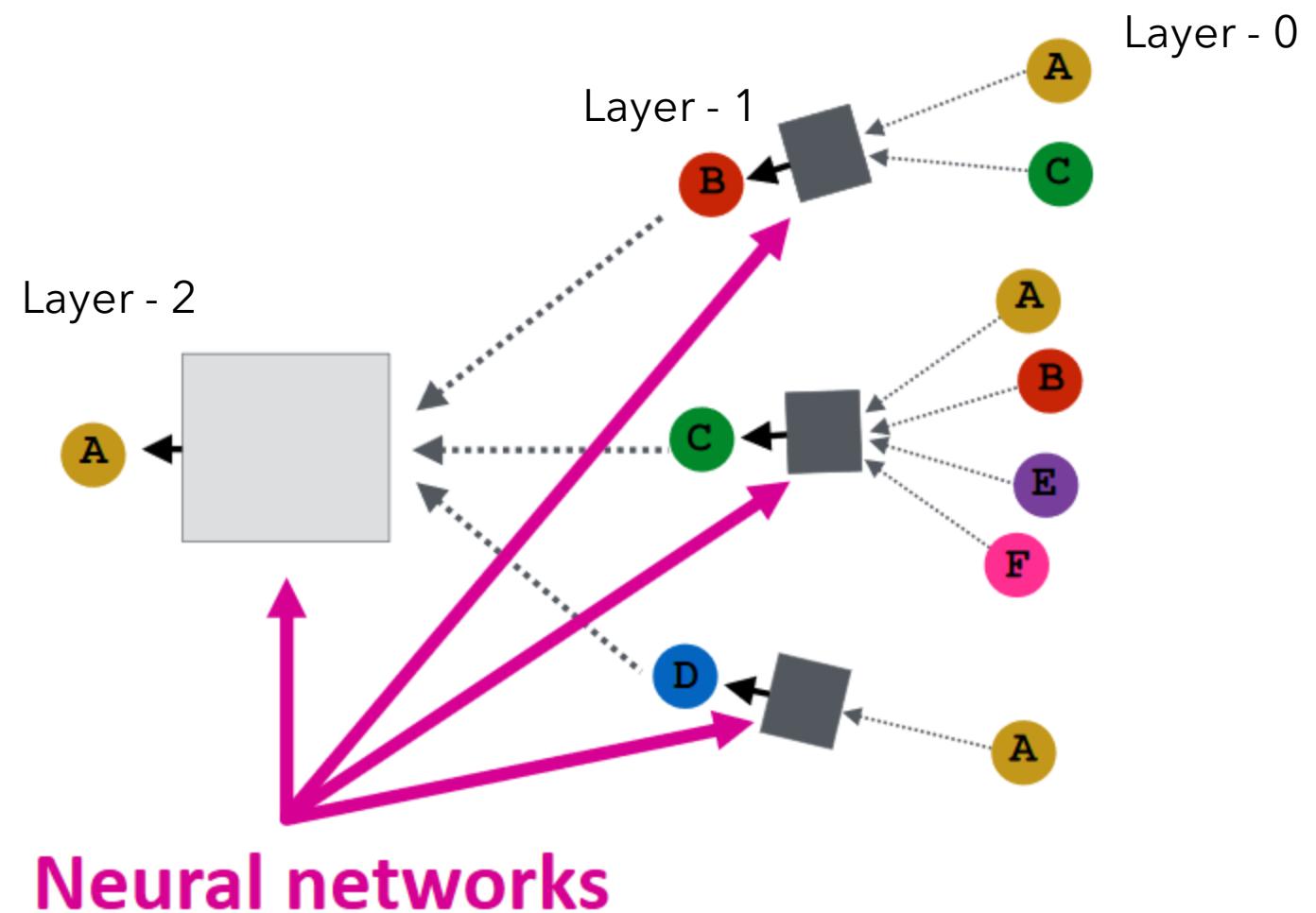
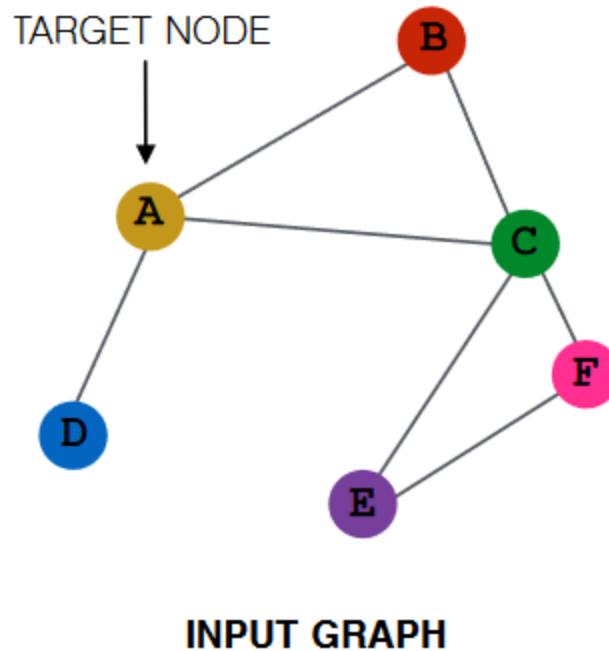
Deep Network for Graphs: The Basic Idea

Message Passing Neural Networks

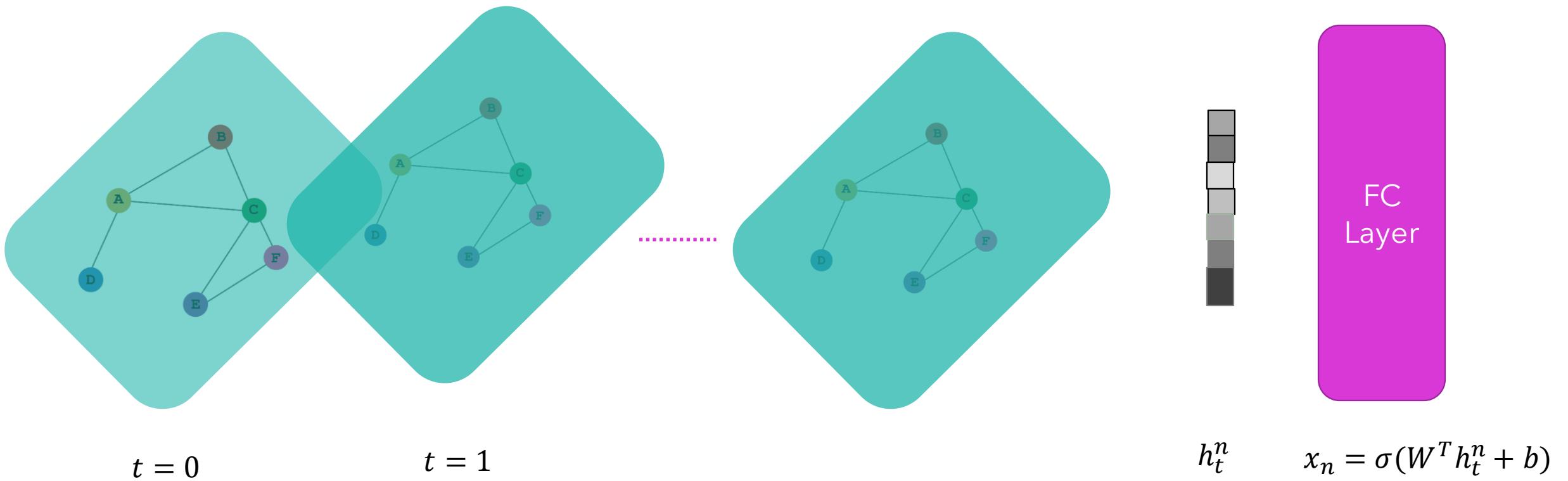


Deep Network for Graphs

Key Idea: Nodes learns aggregate information from their neighbors using neural network



End-to-End Learning



$$L = y_n \cdot \log x_n + (1 - y_n) \log(1 - x_n)$$

Different Graph Neural Network Models

Graph Convolution
Network (GCN)

GraphSAGE

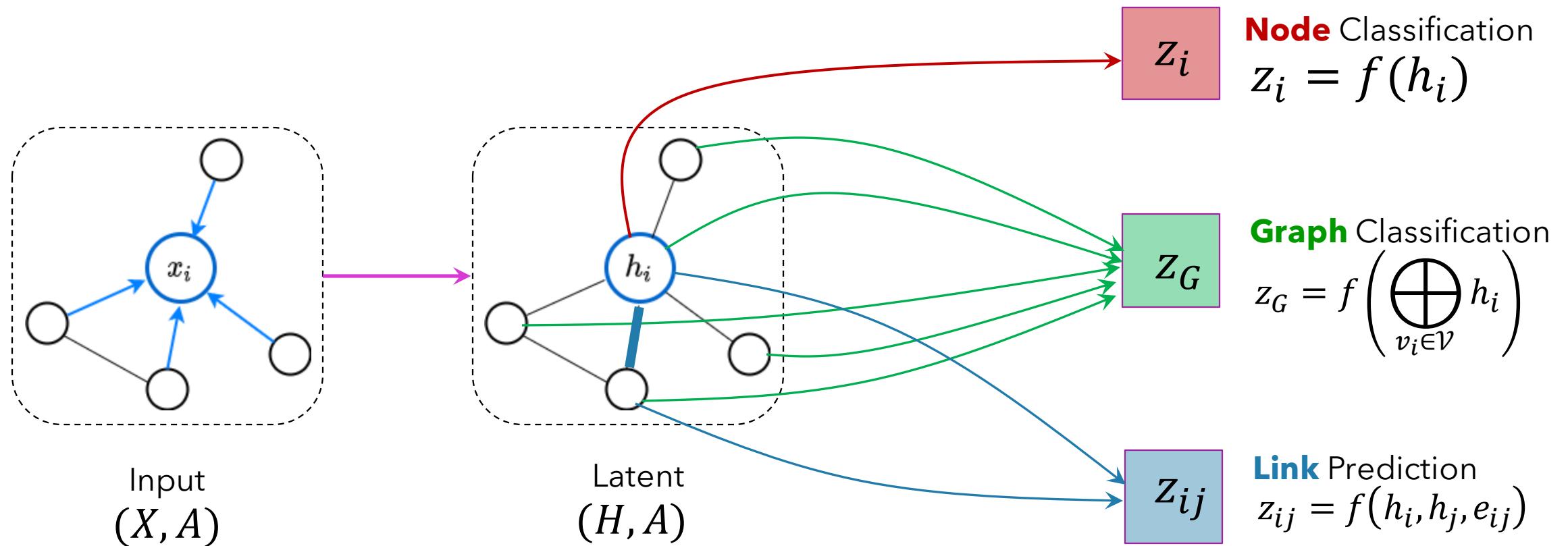
Gated Graph Neural
Network (GGNN)

Graph Attention Network
(GAT)

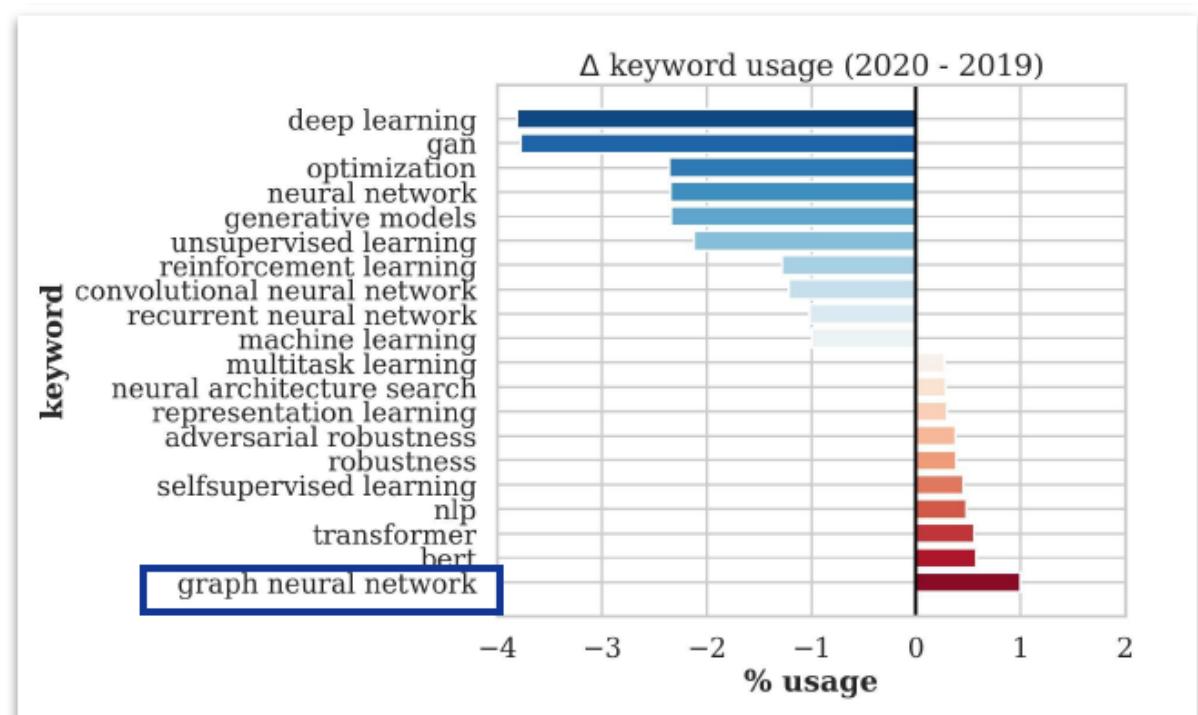
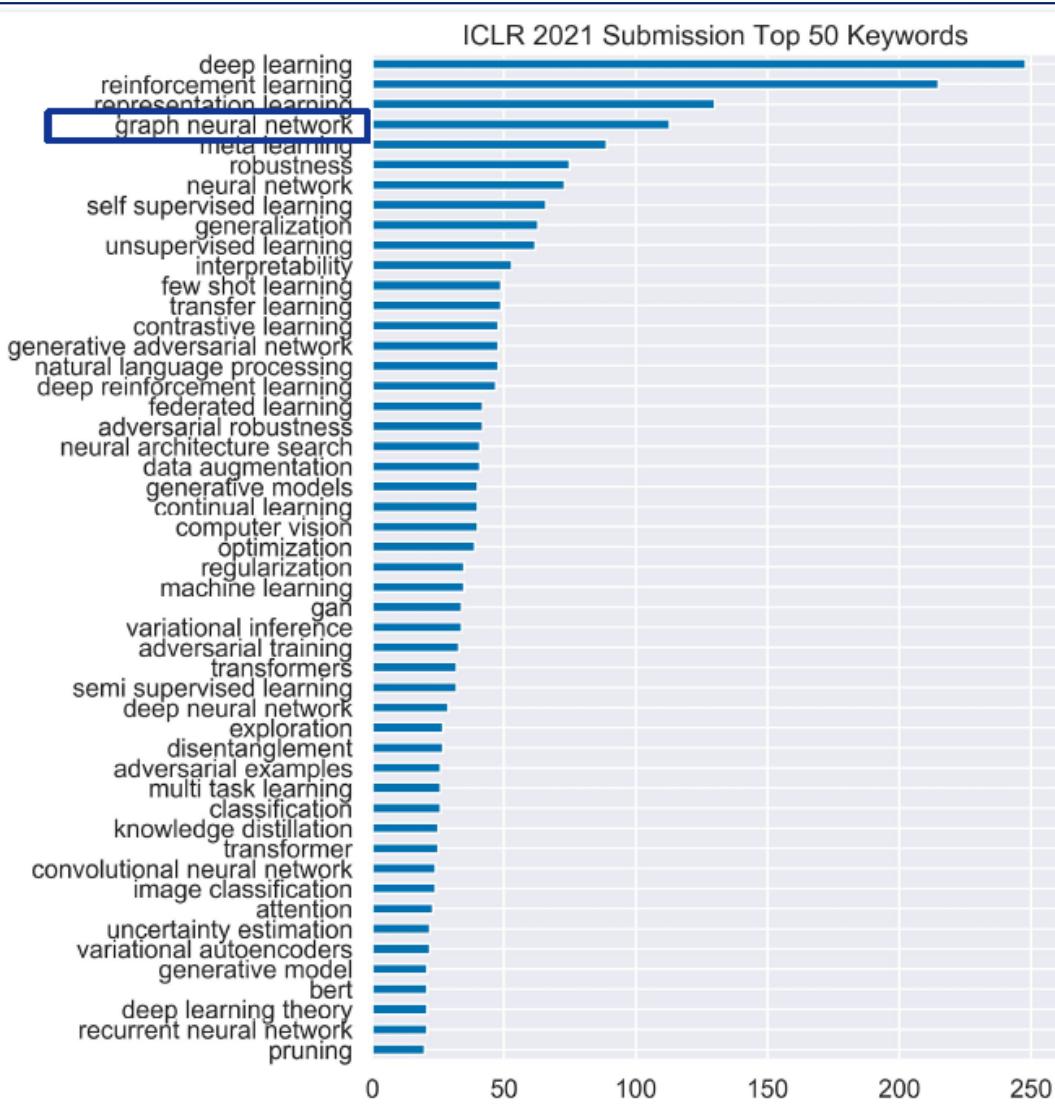
Generative Adversarial
Network for Graphs
(NetGAN)

GraphRNN

General Recipe of Graph ML

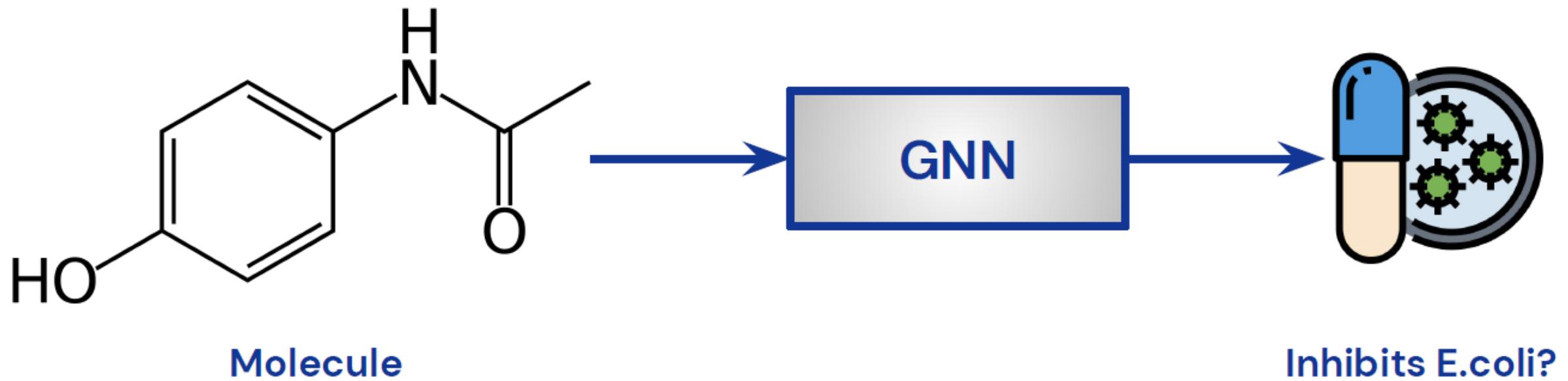


Recent Trend in ML Community



Applications: Science

Molecule Classification & Drug Discovery



Whether a molecule is a potent drug?

Molecule Classification & Drug Discovery

nature

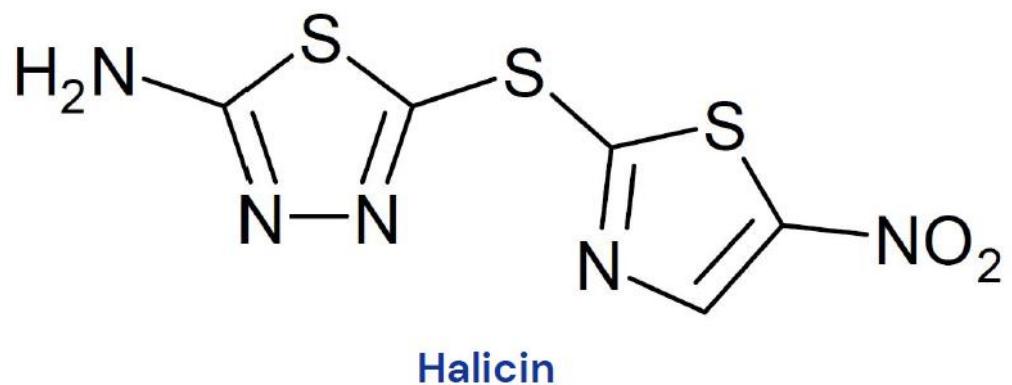
Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

nature > news > article

NEWS | 20 February 2020

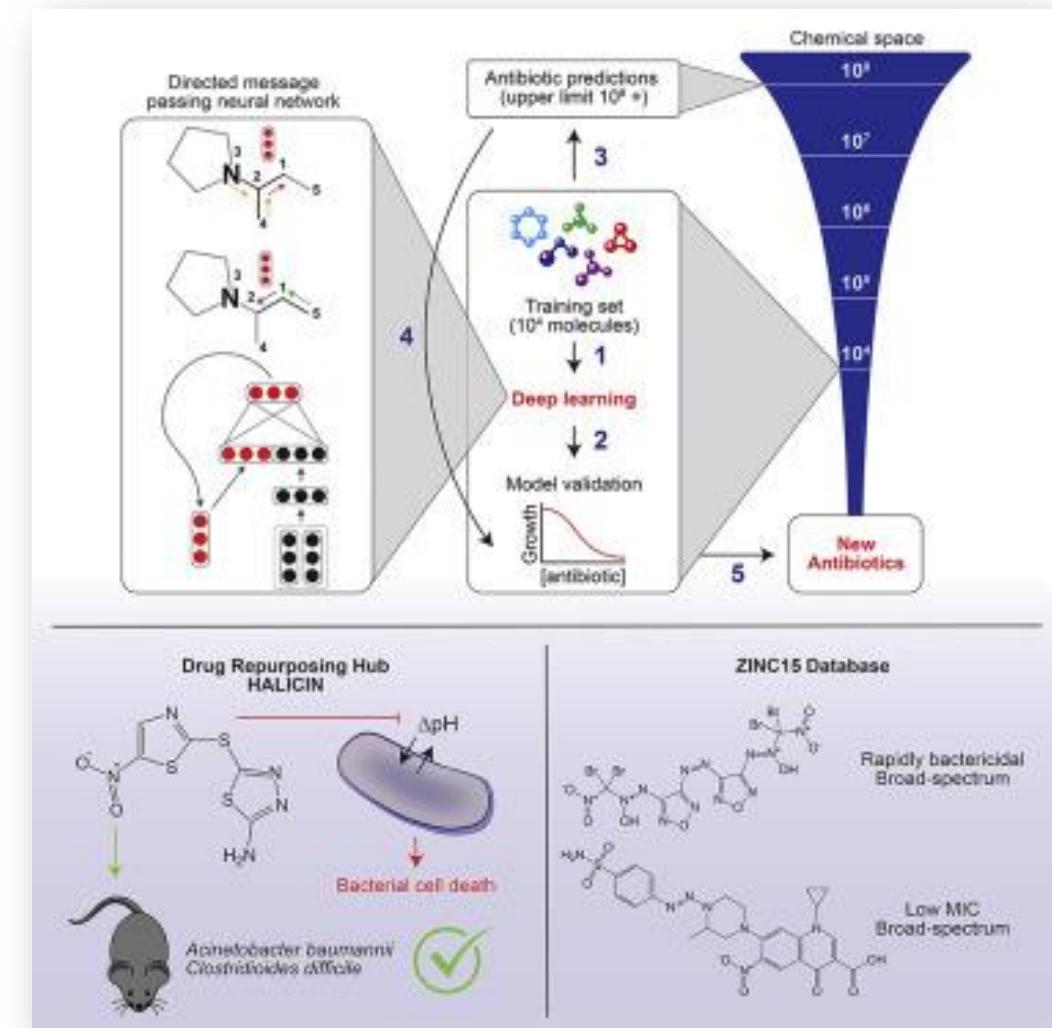
Powerful antibiotics discovered using AI

Machine learning spots molecules that work even against 'untreatable' strains of bacteria.



The chemical structure of Halicin is shown, featuring a central thiazine ring system substituted with a nitro group (NO_2) at position 2 and a 2-aminoethyl side chain at position 4. The side chain is further substituted with a thiomethyl group ($\text{CH}_2\text{S}-$).

Halicin



Pileup Mitigation in Large Hadron Collider



nature

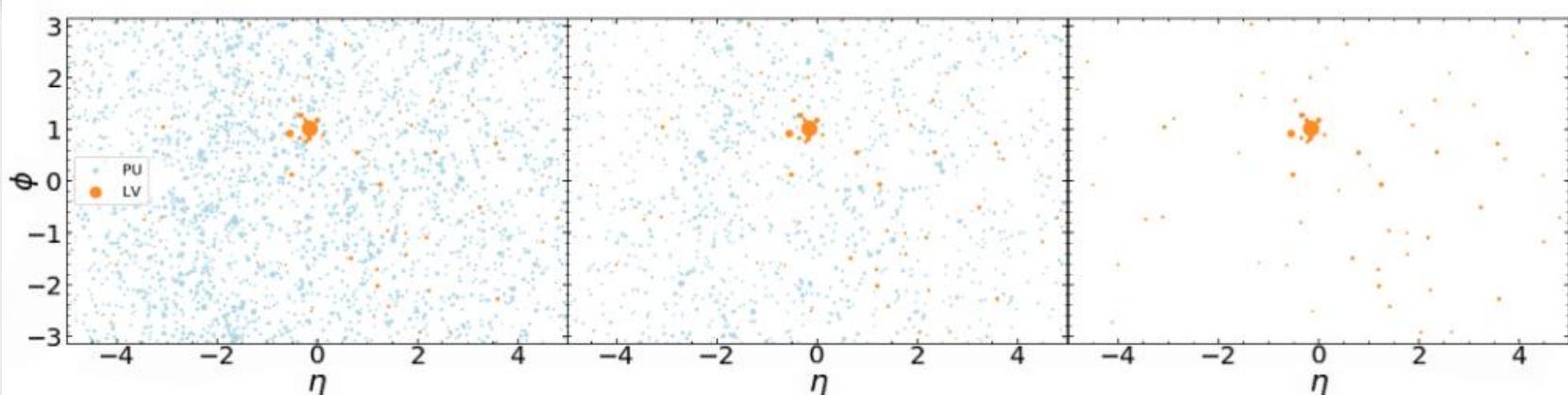
Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [news](#) > [article](#)

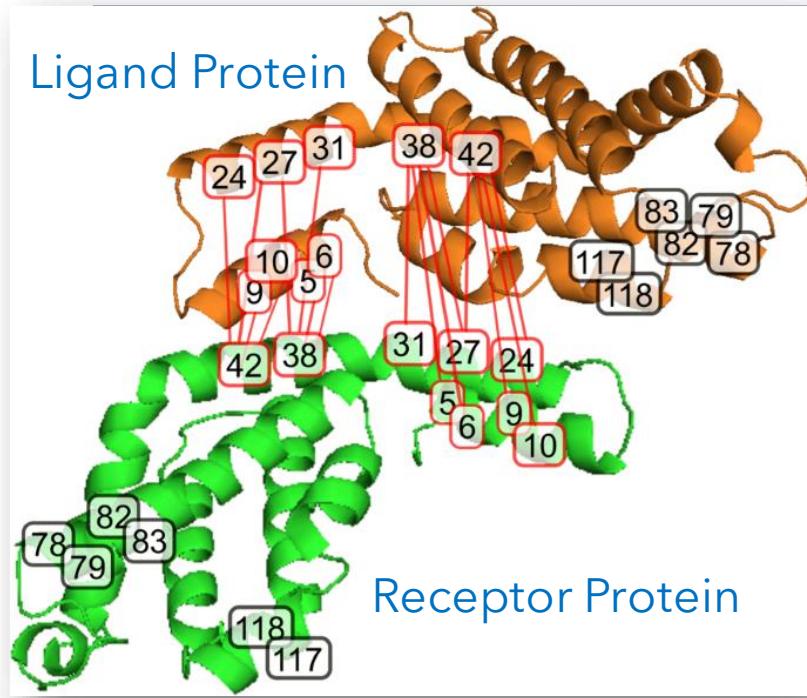
Published: 01 December 2015

Artificial intelligence called in to tackle LHC data deluge

[Davide Castelvecchi](#)



Protein Interface Prediction

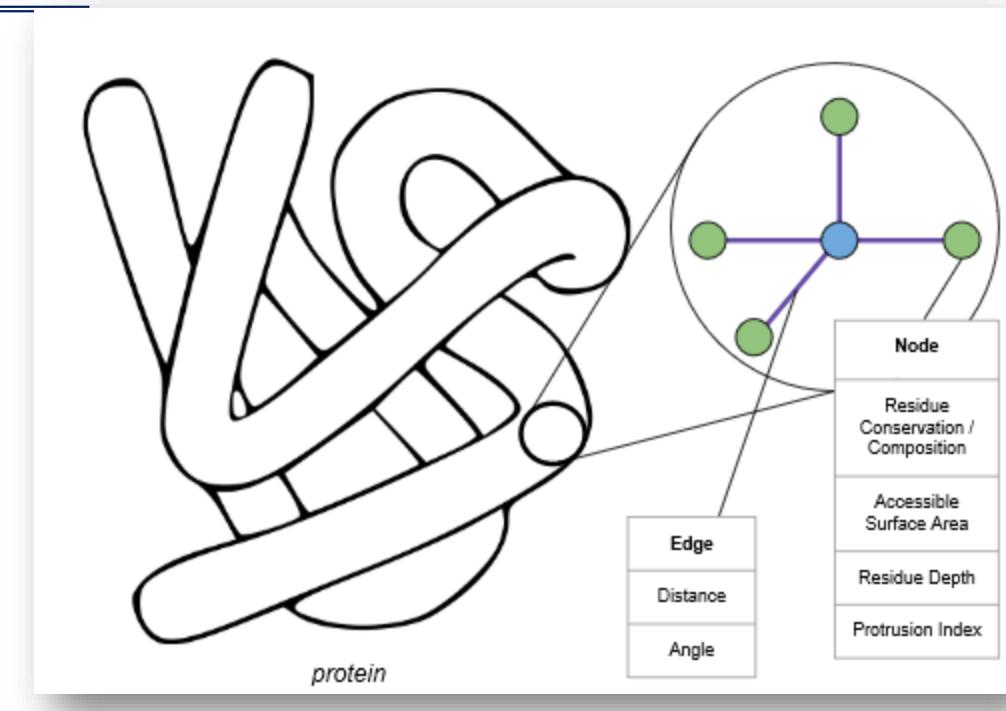


The Problem

Protein-Protein Docking

Important for identifying drug targets

X-ray crystallography is expensive



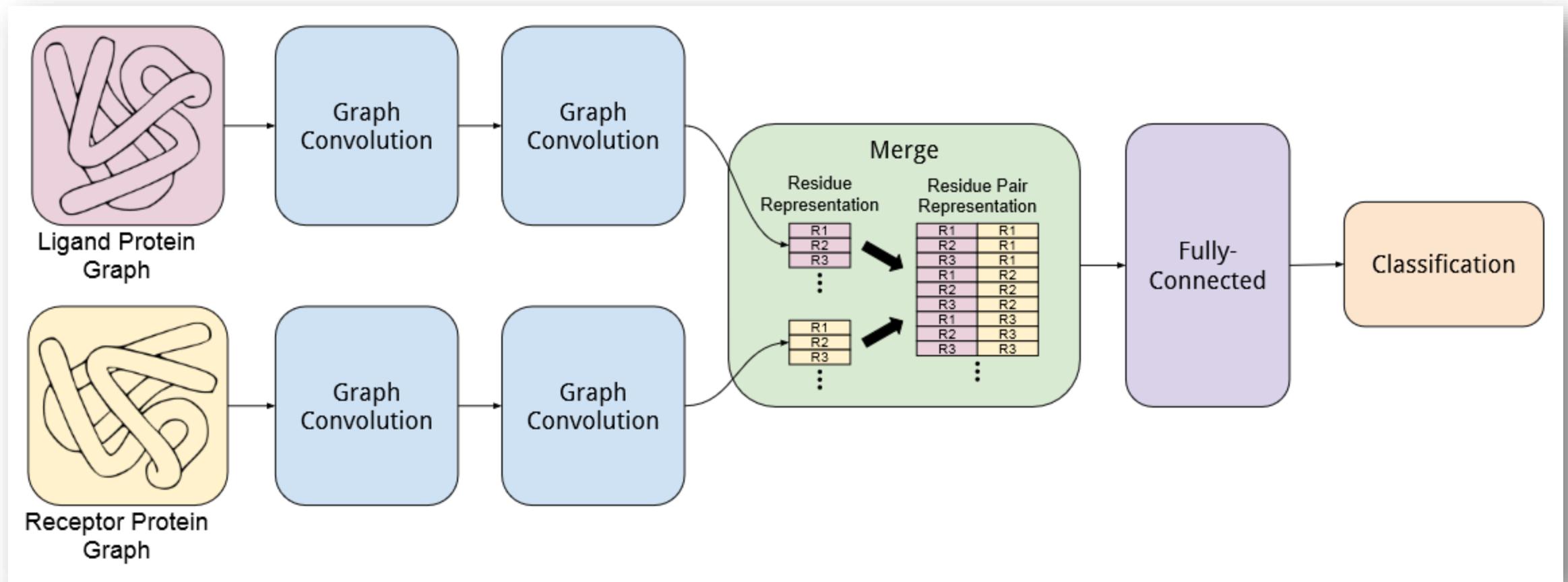
Protein Graph Representation

Amino acid residues are the binding sites

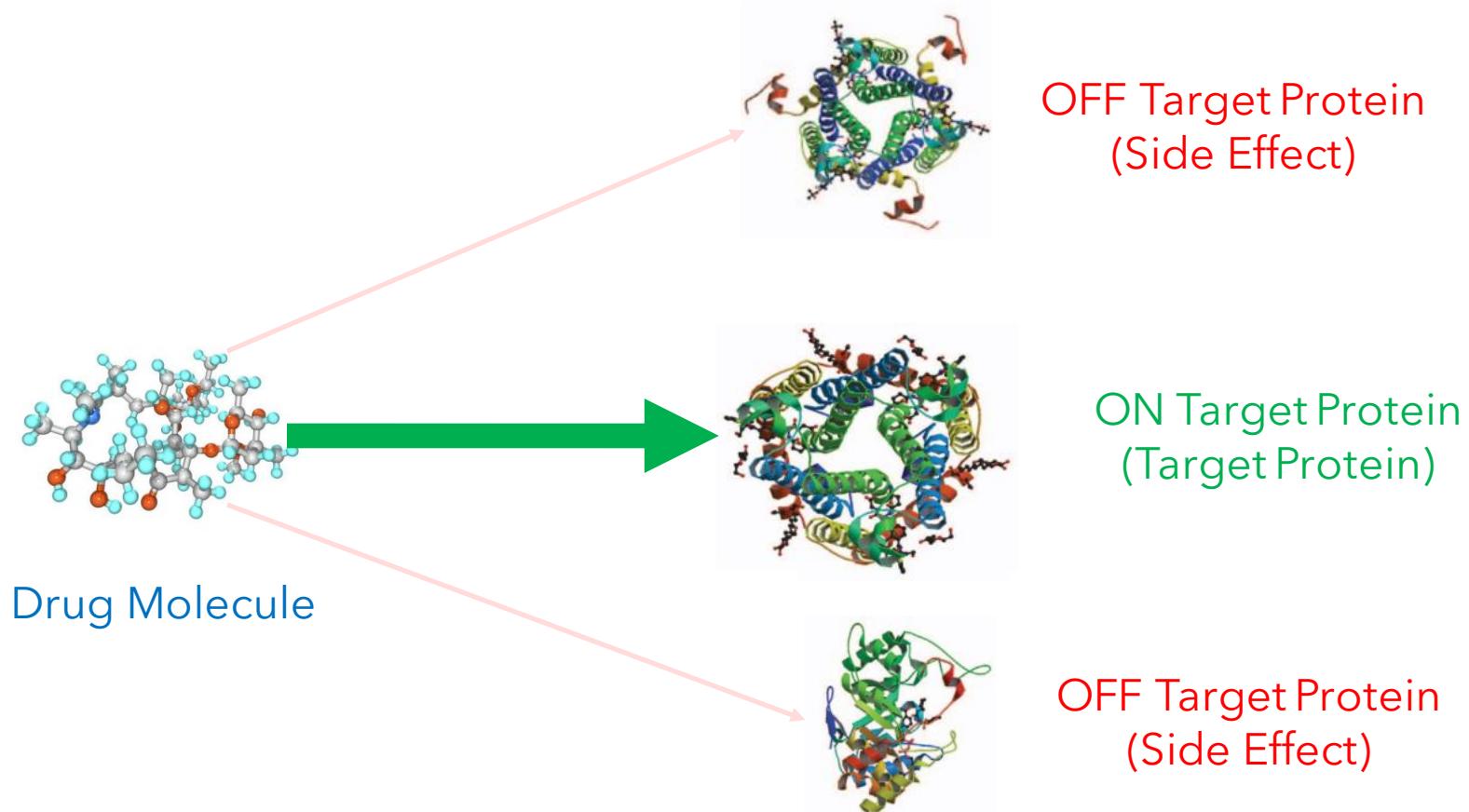
Amino acid residues are the nodes with features

Edges represent the spatial proximity

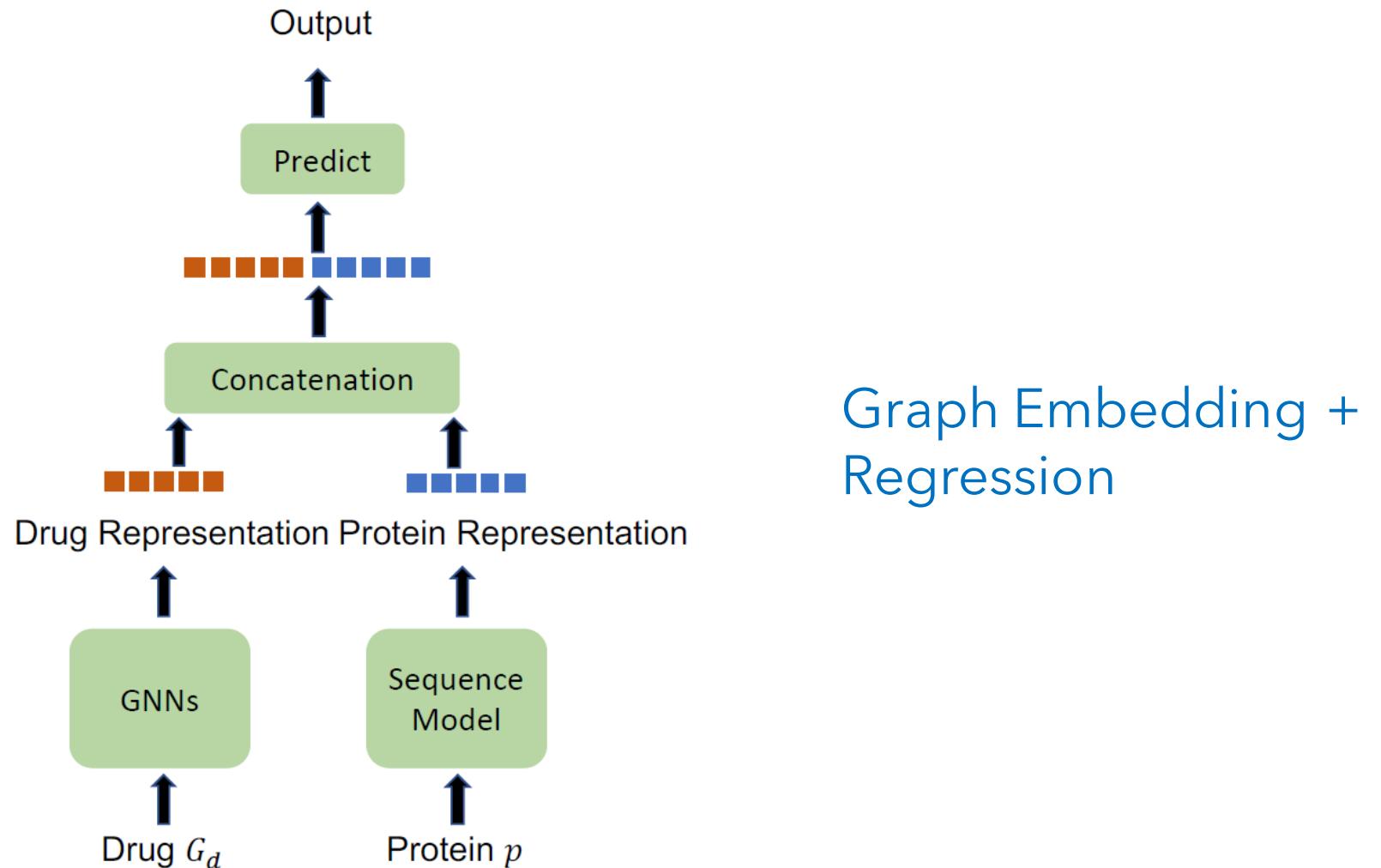
Applications: Protein Interface Prediction



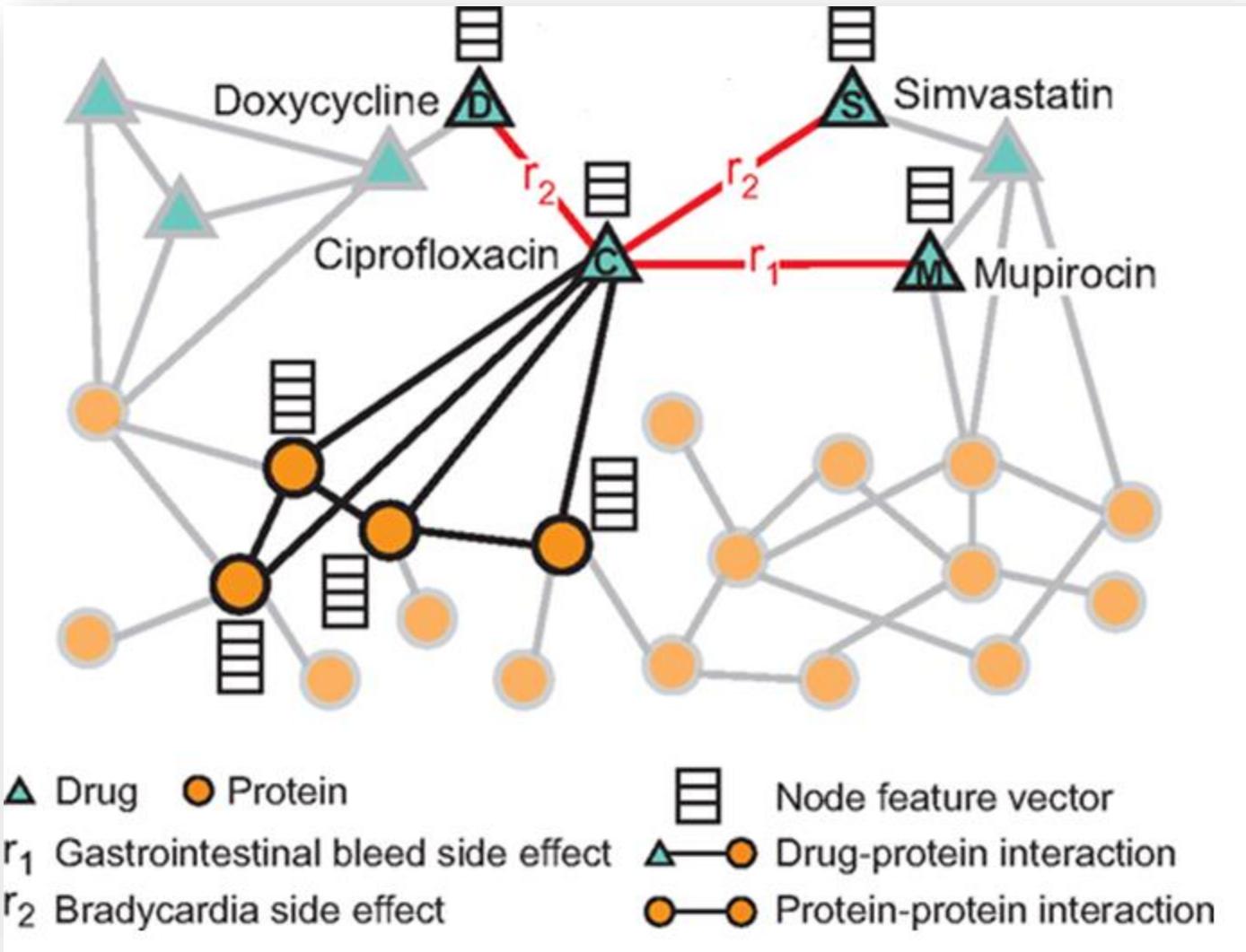
Drug-Target Affinity Binding Prediction



Drug-Target Affinity Binding Prediction



Polypharmacy Side Effect Prediction



Multi-modal Graph
Link Prediction Problem

Applications: Industry

Recommender Systems

Pinterest Engineering
Aug 15, 2018 · 8 min read · [Listen](#)

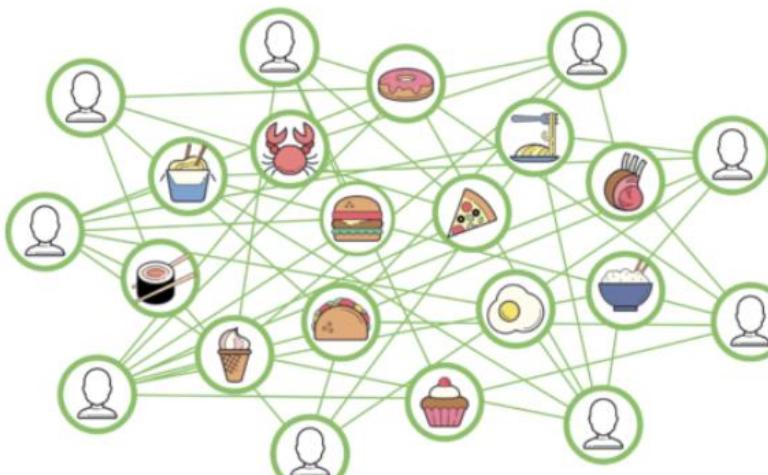
[Twitter](#) [Facebook](#) [LinkedIn](#) [Link](#) [Share](#) [...](#)

PinSage: A new graph convolutional neural network for food discovery

Food Discovery with Uber Eats: Using Graph Learning to Power Recommendations

Ankit Jain, Isaac Liu, Ankur Sarda, and Piero Molino

0



December 4, 2019

amazon | science

P-Companion: A principled framework for diversified complementary product recommendation

By Junheng Hao, Tong Zhao, Jin Li, Xin Luna Dong, Christos Faloutsos, Yizhou Sun, Wei Wang

2020

Query Item	Co-Purchase	Top-3 P-Companion Recommendations
		 
		  
	None	  
	None	  

Fast Chip Design (TPUv5)

nature

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [articles](#) > [article](#)

Article | Published: 09 June 2021

NATURE PODCAST | 09 June 2021

Google AI beats humans at designing computer chips

An AI that designs computer chips in hours, and zooming in on DNA's complex 3D structures.

A graph placer

Azalia Mirhoseini✉, Anna G Young-Joon Lee, Eric Johnson, Hang, Emre Tuncer, Quoc V. Benjamin Thompson & Noah Baker

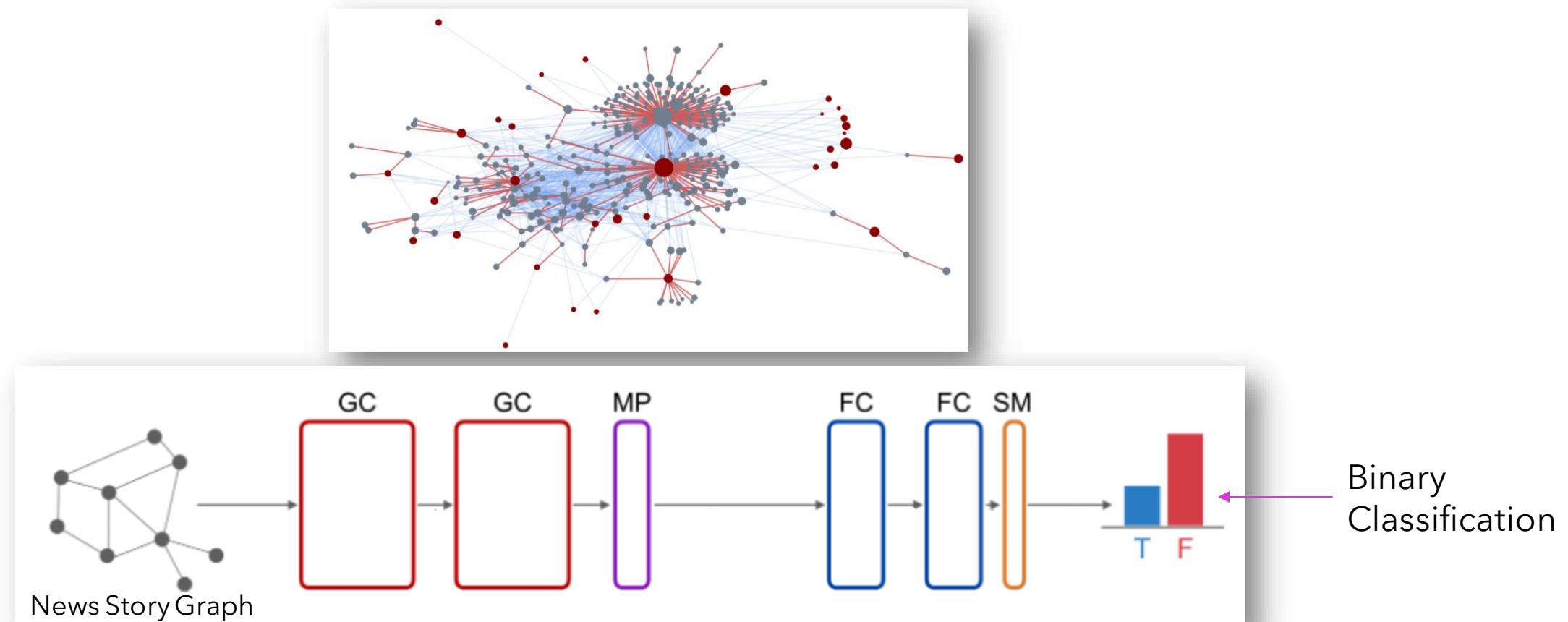
Graph ML in Google Maps



Data Mining & Social Networks

Applications: Fake News Detection

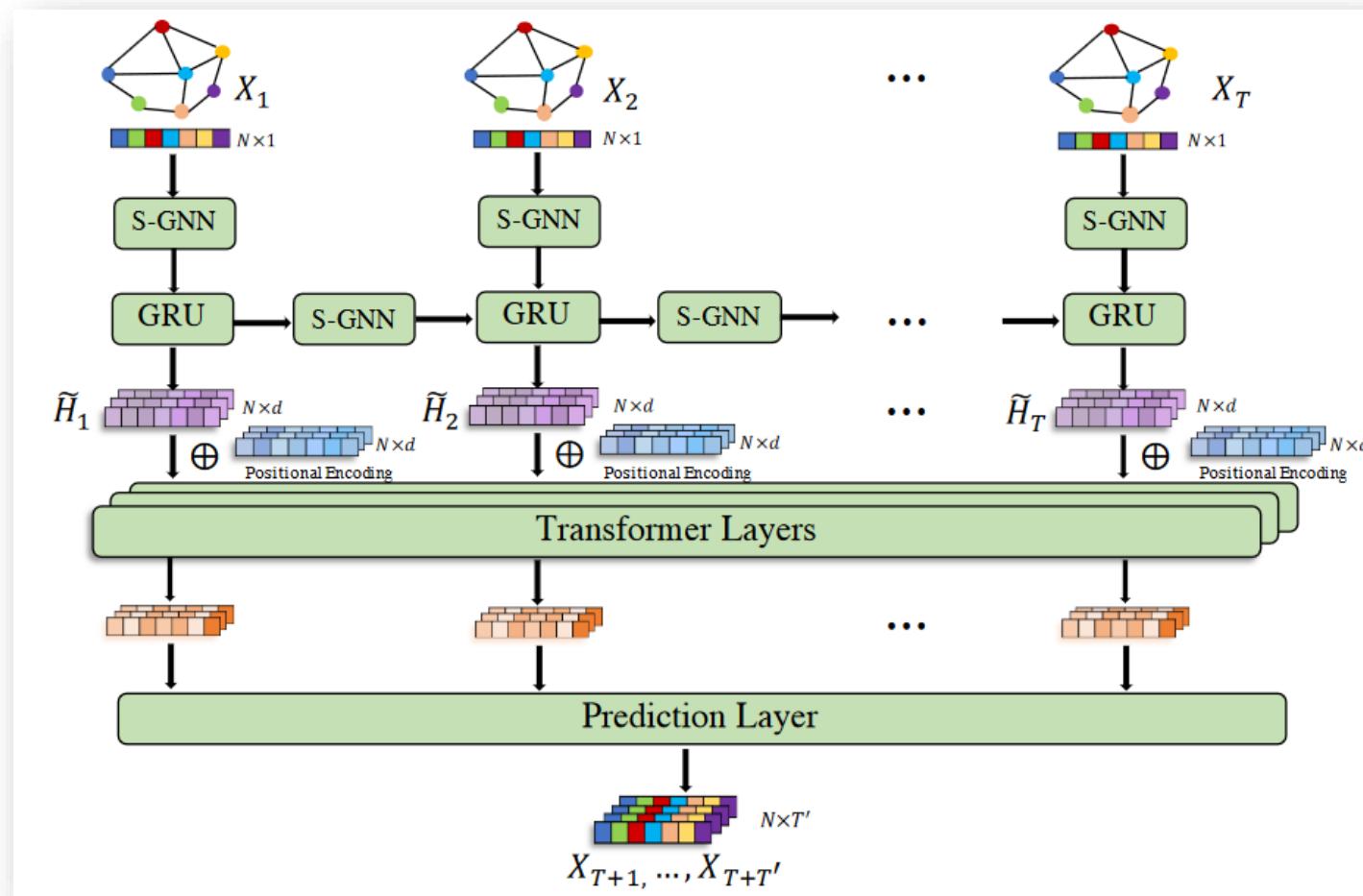
Fake News have different propagation pattern as compared to normal news



Graph Classification Problem

Traffic Prediction

$$(\hat{X}_{T+1}, \dots, \hat{X}_{T+T'}) = f(X_1, \dots, X_T)$$



Traffic flow prediction has been modelled as Time Series

- Captures temporality
- Ignores spatial dependency

S-GNN captures the spatial relationships + GRU captures the temporality

Aggregation of S-GNN and GRU outputs are used to predict the future states

Can be adopted for **air quality forecasting**

Graph DL Community

Graph Geometric Learning

Graph Representation Learning

Graph Neural Networks

ICML

NeurIPS

AAAI

WWW

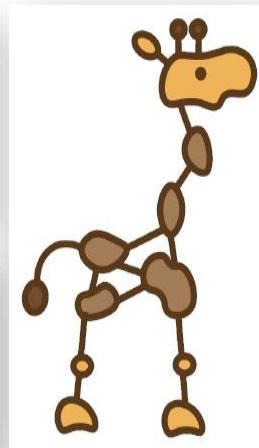
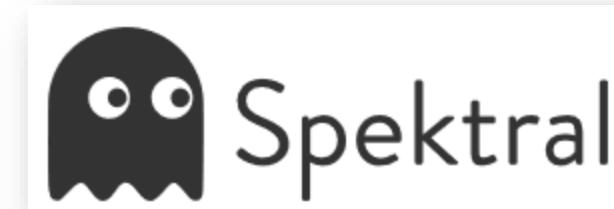
ICLR

KDD



[LoGaG Slack Channel](#)

Tools & Datasets



[Network Repository](#)

[Open Graph Benchmark](#)

Course Webpage

<https://ainimesh.github.io/GMLFA-AI60007/>