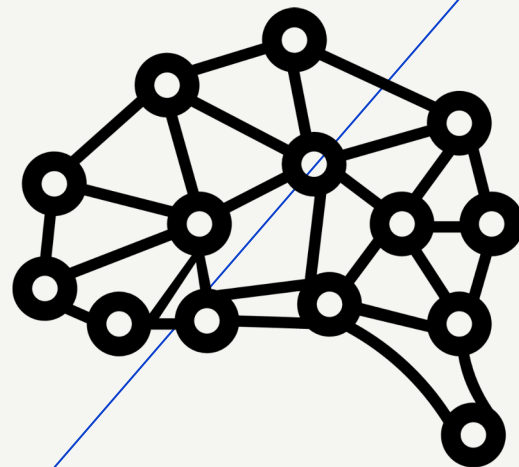


Navigating the Future of Data Science: Graphical Neural Networks

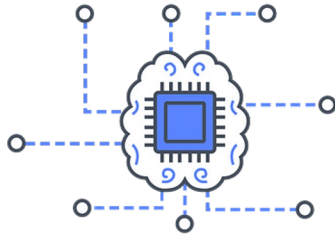
PHUSE UK SDE

Jun Se Kim,
Analytical Data Scientist,
Data & Statistical Sciences, PD Data
Sciences

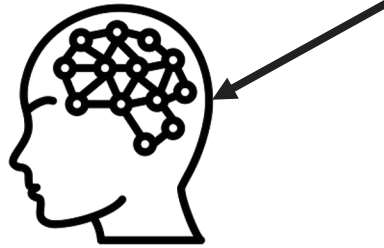


Apr 22, 2024 | Public Use

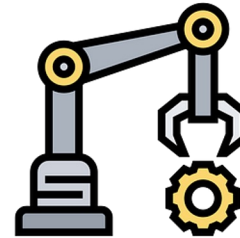
Word Association



Machine Learning



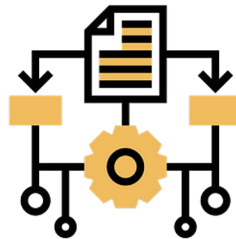
Neural Networks



Robotics



Expert Systems



Fuzzy Logic

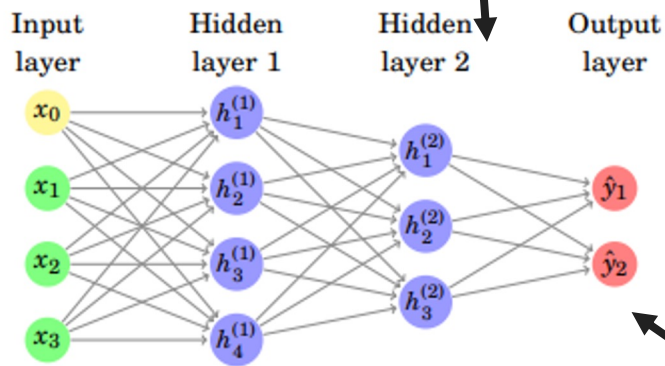


Natural Language
Processing

Neural Networks: The Basic Building Block of AI

Process input features to identify patterns and capture relevant features

features



Now train again but backpropagate to tune network

Contains neurons representing possible classes



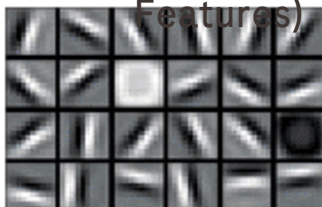
Data read in first layer

HOW A DEEP NEURAL NETWORK SEES

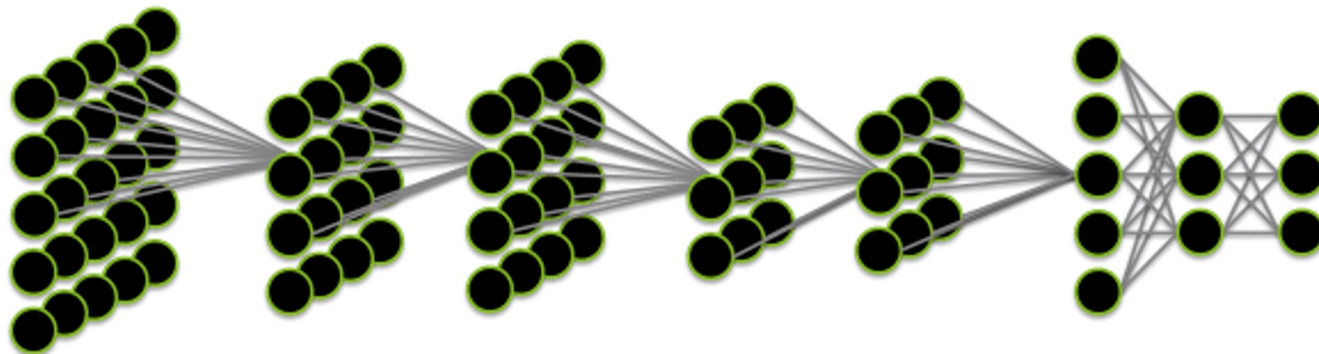
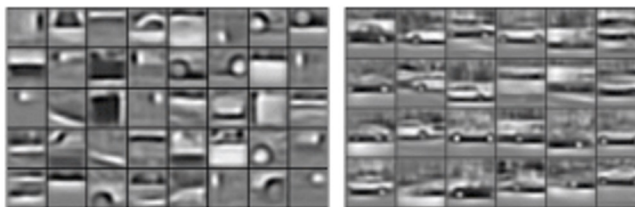
Raw Data



Edges (Low Level Features)

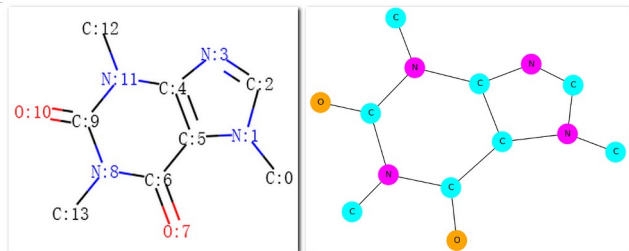
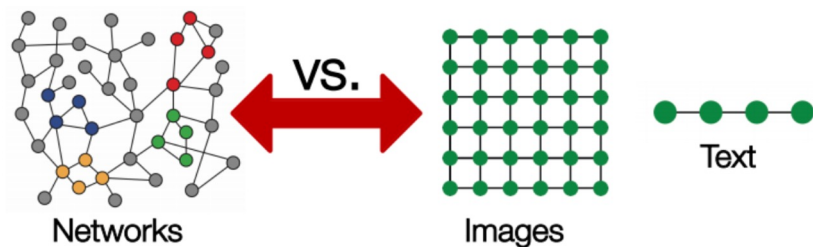


High Level Features

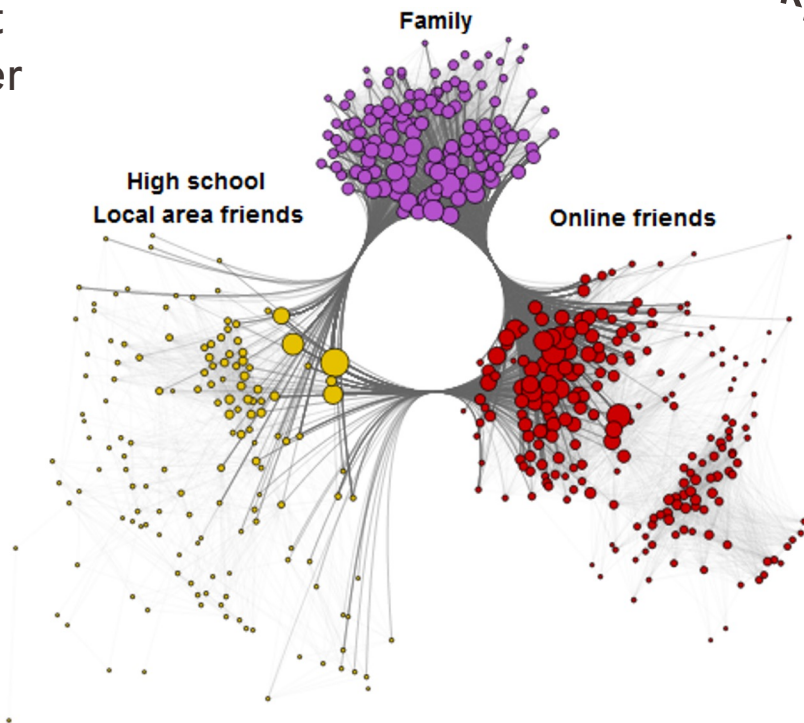


Graphs Are Everywhere

Nodes represent entities, edges represent connections or interactions. No spatial order

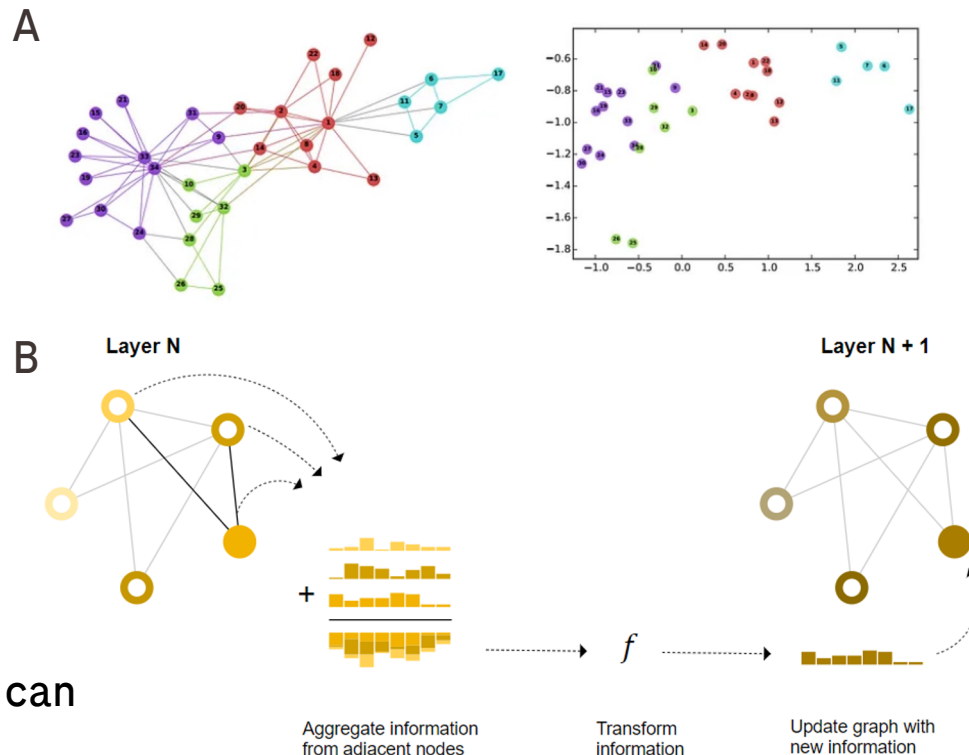


Relationships are Tricky :(



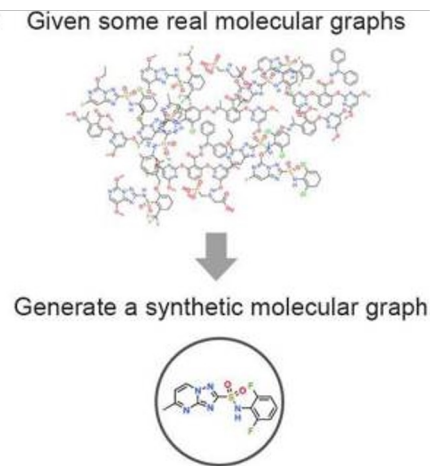
What are GNNs?

1. Node Embedding: Initial node representations.
2. Message passing: Nodes exchange information with neighbours.
3. Aggregation: Combining information from neighbours.
4. Pooling: Using the aggregated info we can perform tasks like node/edge prediction.



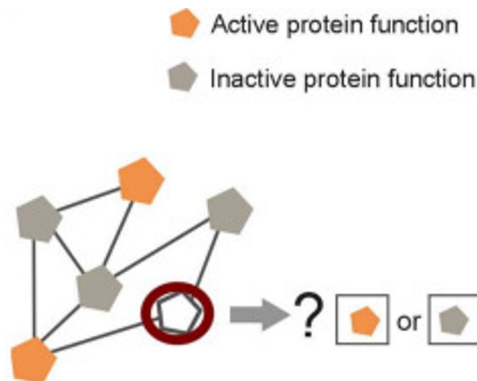
What Problems Can They Solve?

Graph Level Tasks



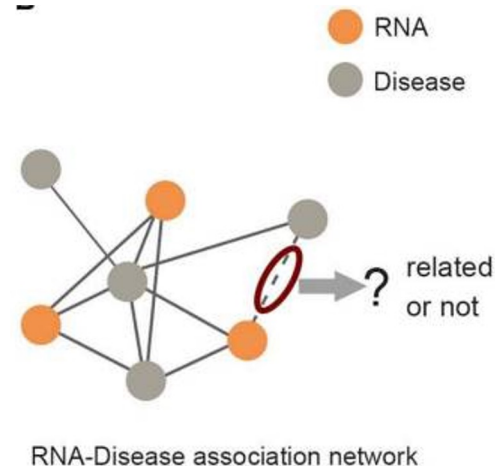
Predicting properties or labels for entire graphs.

Node Level Tasks



Predicting properties or labels for nodes.

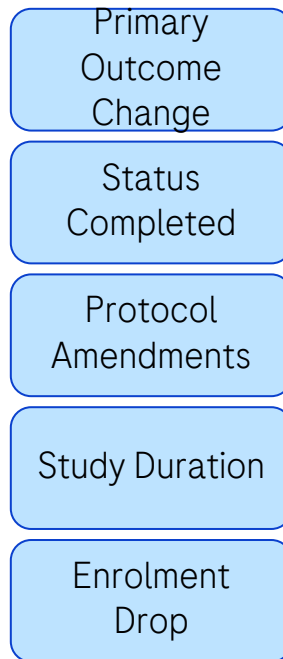
Edge Level Tasks



Predicting properties or labels for edges.

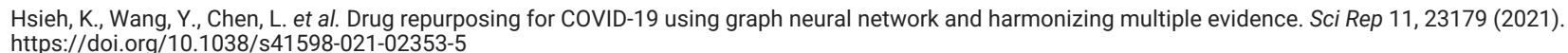
Risk Prediction For Clinical Trials Based On Protocol Design

- Clinical Trials have high failure rate
- GNN can predict outcome of CT from historic changes to Protocol.
- 135,940 unique CT protocols from ClinicalTrials.gov.
- CTs labelled as low, medium, or high risk based on Protocol features.
- AUROC of 0.8453 for 3 tertiary classification and 0.92 for binary.

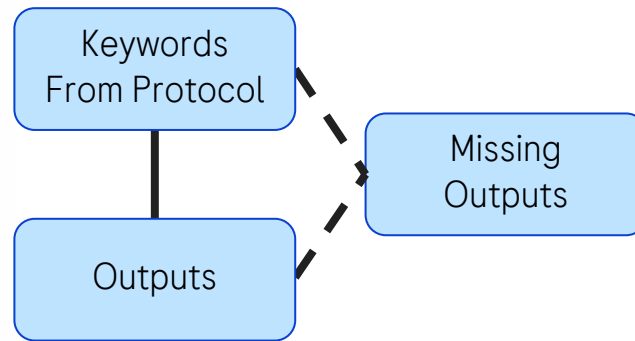
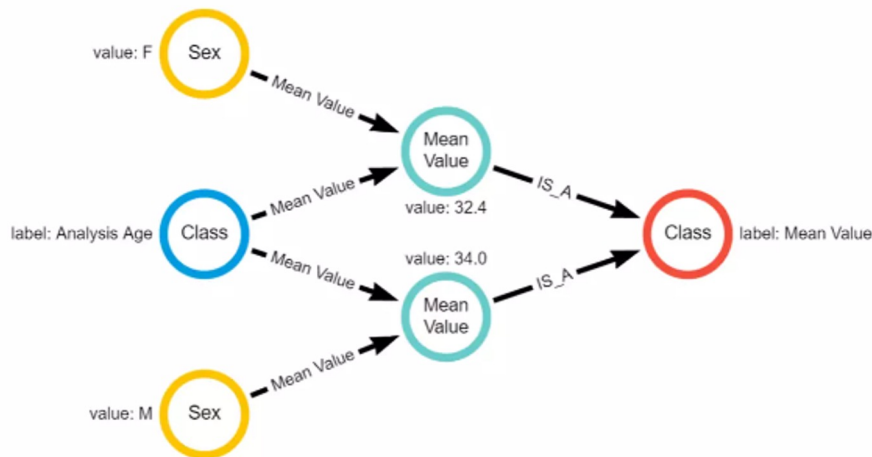


Risk criterion features

- Collected 3635 drugs related to COVID-19 treatment.
- Built COVID-19 knowledge graph with virus-host interactions.
- GNN used to analyze drug interactions.
- Prioritized drugs based on clinical trial history.
- Identified drug combinations with potential synergistic effects.



Thinking With Graphs



- ADAM and SDTM quality checking: Discover Duplication, Corrupted Nodes/Edges
- Output and Protocol Graph: Predict potentially missing outputs that would otherwise be adhoc

GSK

AstraZeneca

Hurdles To Overcome

- Practical implementation of GNNs still in its early stages in pharma.
- Data needs to be reprocessed as graphs.
- Ethical questions and security issues with GNNs?
- Value of AI solution outweigh the cost of production?



Key Takeaways

“The journey
of a thousand
miles begins
with a single
step.”

Lao Tzu



- To begin to apply AI innovatively, we need a basic knowledge of AI solutions available.
- GNNs are great predictive models.
- Can data used within your roles fit into a graph model?
- Could a GNN solution be useful in your role?

To Learn More..



- Stanford University Machine Learning with Graphs Course

Practical Deep Learning

A free course designed for people with some coding experience, who want to learn how to apply deep learning and machine learning to practical problems.

New!

We just launched a new >30 hour video course for more experienced students:

[Practical Deep Learning for Coders part 2: Deep Learning Foundations to Stable Diffusion](#)



This free course is designed for people (and bunnies!) with some coding experience who want to learn how to apply deep learning and machine learning to practical problems.

Deep learning can do all kinds of amazing things. For instance, all illustrations throughout this website are made with deep learning, using [DALL-E 2](#).

Welcome!

Practical Deep Learning for Coders 2022 part 1, recorded at the [University of Queensland](#), covers topics such as how to:



- Build and train deep learning models for computer vision, natural language processing, tabular analysis, and collaborative filtering problems
- Create random forests and regression models
- Deploy models
- Use PyTorch, the world's fastest growing deep learning software, plus popular libraries like fastai and Hugging Face

- Practical Deep Learning for Coders (Hands on Deep Learning Challenges)



- Open Graph Benchmark

Thank You For Listening!

Contact: `jun_se.kim@roche.com`

Doing now what patients need next