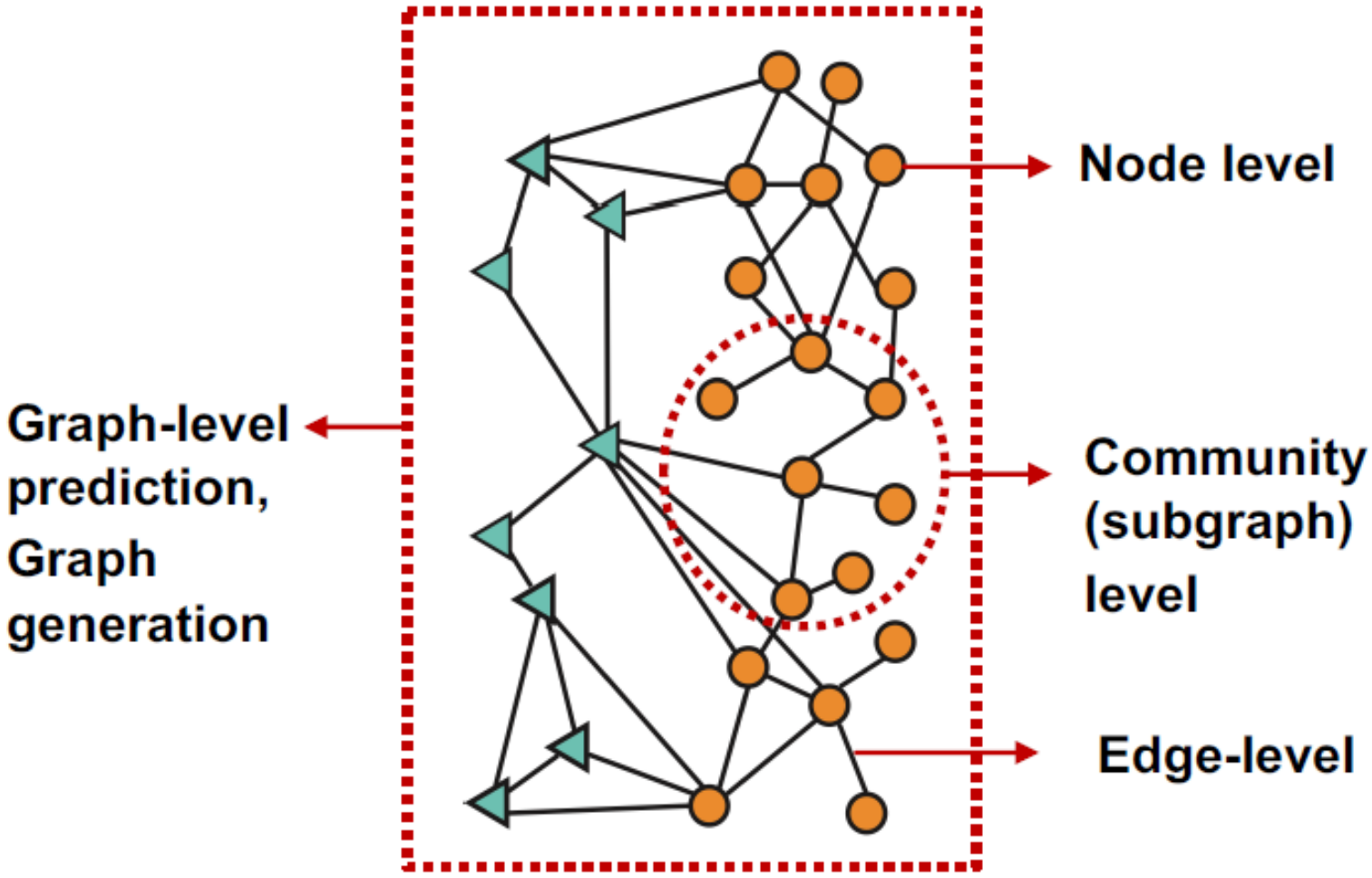
Graph Machine Learning

Applications of Graph ML

Slides adapted from Jure Leskovec's CS244
Graph Machine Learning

Different types of tasks



Classical Graph ML tasks

- **Node classification**: Predict a property of a node
  - **Example:** Categorize online users / items
- **Link prediction**: Predict whether there are missing links between two nodes
  - **Example:** Knowledge graph completion
- **Graph classification**: Categorize different graphs
  - **Example:** Molecule property prediction
- **Clustering**: Detect if nodes form a community
  - **Example:** Social circle detection
- **Other tasks**:
  - **Graph generation**: Drug discovery
  - **Graph evolution**: Physical simulation

Node Level ML Tasks

Example 1 Protein folding

A protein chain acquires its native 3D structure

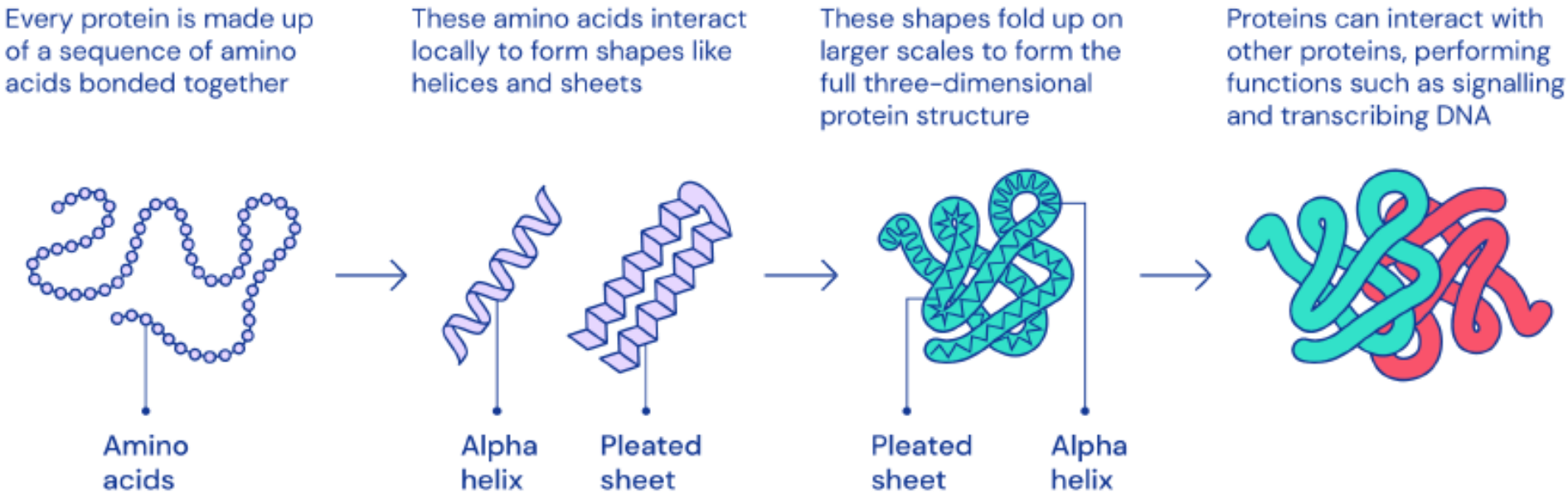
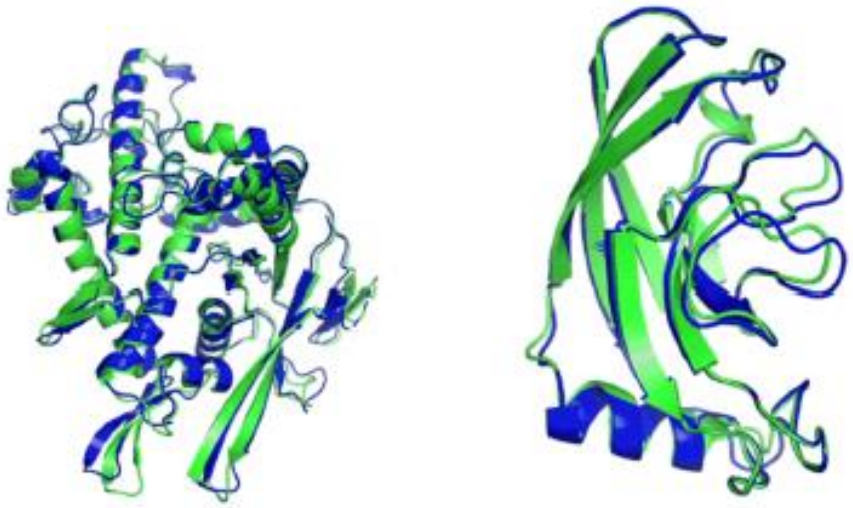


Image credit: DeepMind

The protein folding problem

Computationally predict a protein's 3D structure based solely on its amino acid sequence



T1037 / 6vr4
90.7 GDT
(RNA polymerase domain)

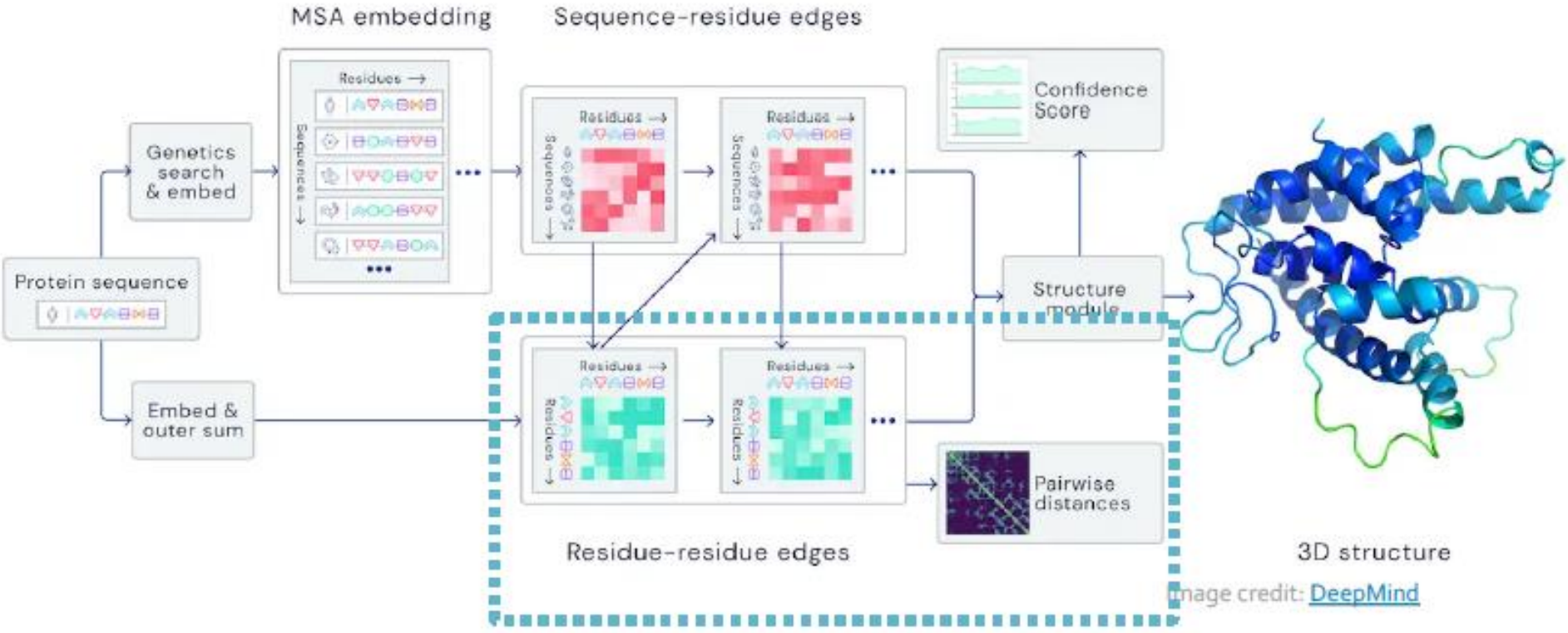
T1049 / 6y4f
93.3 GDT
(adhesin tip)

● Experimental result
● Computational prediction

Image credit: DeepMind

AlphaFold: Solving Protein folding problem

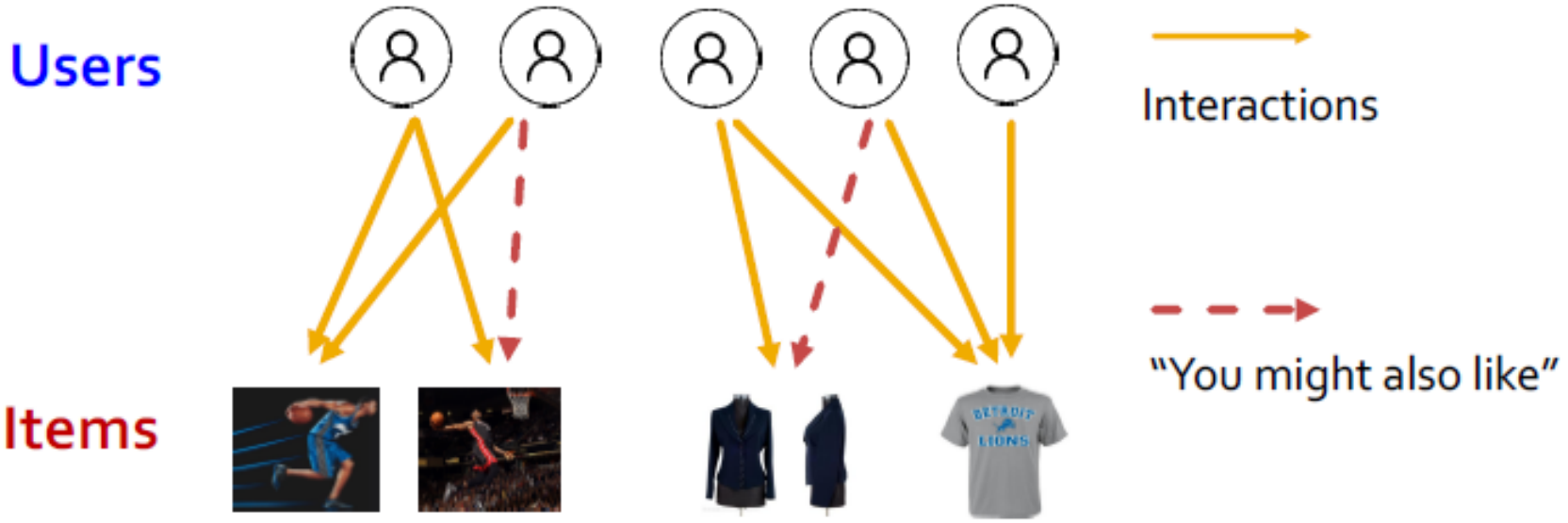
- **Key idea:** "Spatial graph"
  - **Nodes:** Amino acids in a protein sequence
  - **Edges:** Proximity between amino acids (residues)



Edge Level ML Tasks

Example – Recommender Systems

- **Users interacts with items**
  - Watch movies, buy merchandise, listen to music
  - **Nodes:** Users and items
  - **Edges:** User-item interactions
- **Goal: Recommend items users might like**

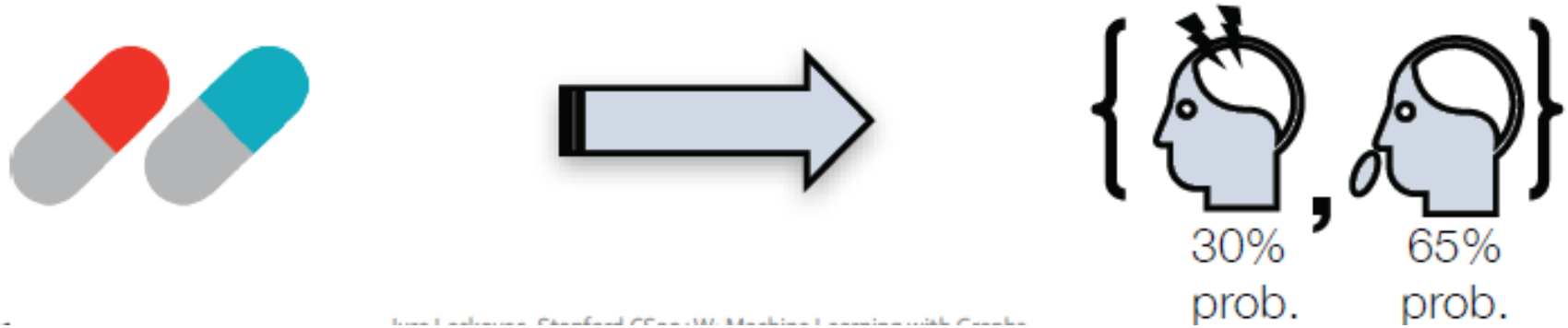


Example Drug Side Effects

Many patients take multiple drugs to treat complex or co-existing diseases:

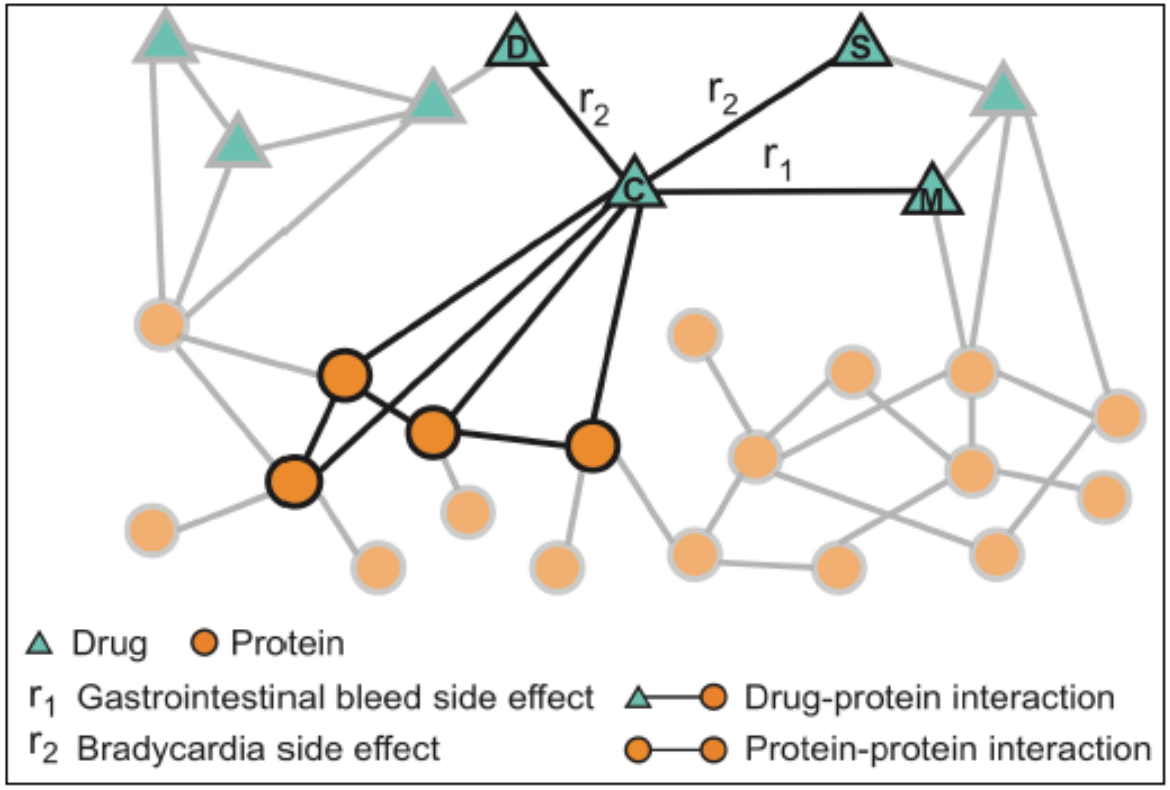
- 46% of people ages 70-79 take more than 5 drugs
- Many patients take more than 20 drugs to treat heart disease, depression, insomnia, etc.

Task: Given a pair of drugs predict adverse side effects

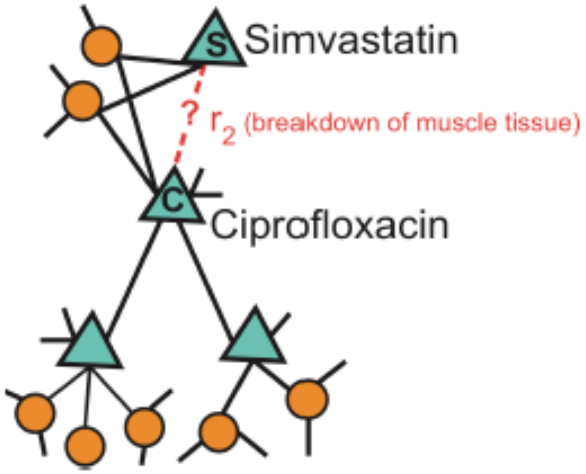


Biomedical Graph Link Prediction

- **Nodes**: Drugs & Proteins
- **Edges**: Interactions



Query: How likely will Simvastatin and Ciprofloxacin, when taken together, break down muscle tissue?

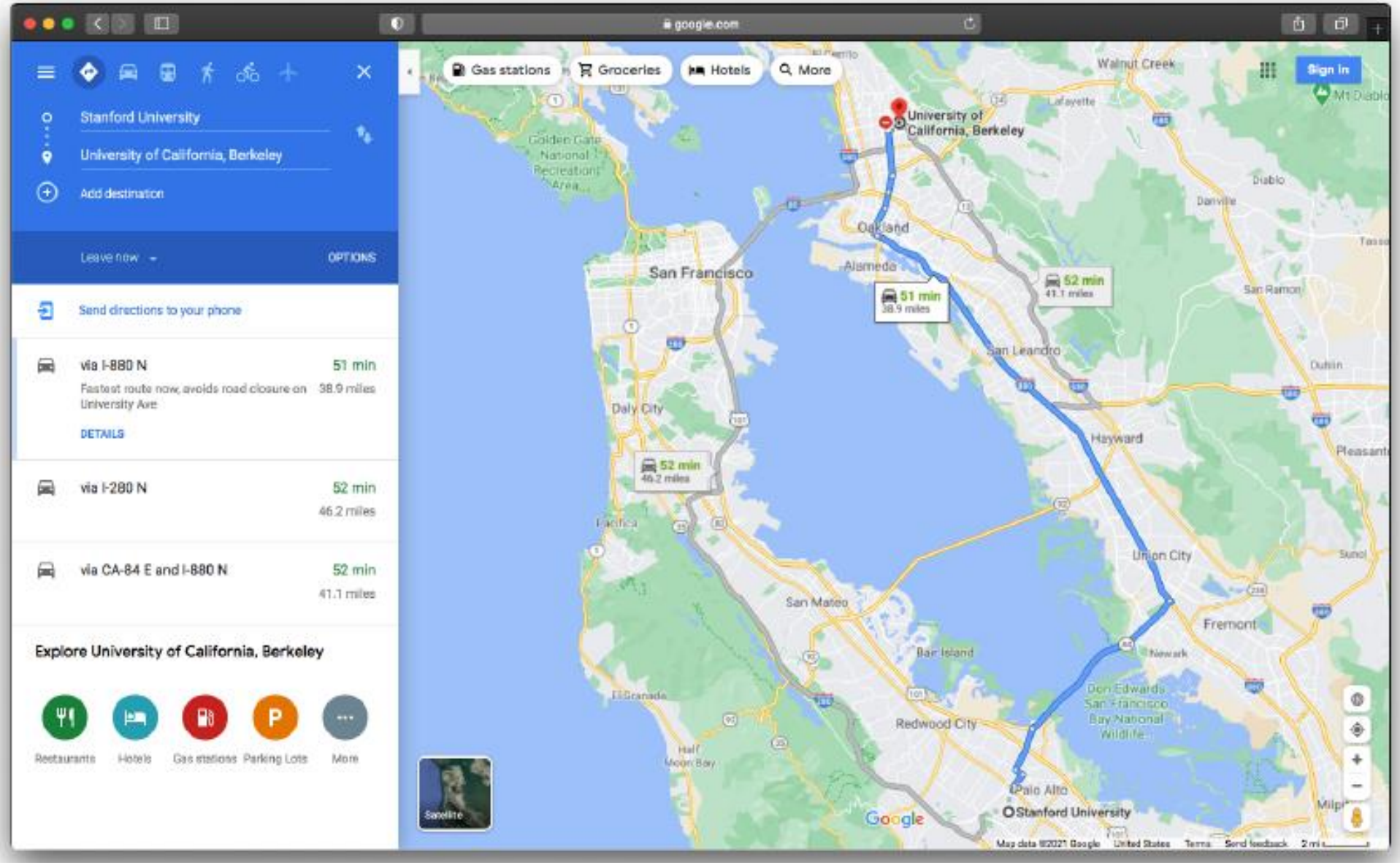


Results and Predictions

Rank	Drug c	Drug d	Side effect r	Evidence found
1	Pyrimethamine	Aliskiren	Sarcoma	Stage *et al.* 2015
2	Tigecycline	Bimatoprost	Autonomic neuropathy	
3	Omeprazole	Dacarbazine	Telangiectases	
4	Tolcapone	Pyrimethamine	Breast disorder	Bicker *et al.* 2017
5	Minoxidil	Paricalcitol	Cluster headache	
6	Omeprazole	Amoxicillin	Renal tubular acidosis	Russo *et al.* 2016
7	Anagrelide	Azelaic acid	Cerebral thrombosis	
8	Atorvastatin	Amlodipine	Muscle inflammation	Banakh *et al.* 2017
9	Aliskiren	Tioconazole	Breast inflammation	Parving *et al.* 2012
10	Estradiol	Nadolol	Endometriosis	

Subgraph Level ML Tasks

Example Traffic Prediction



Road Network as a Graph

- **Nodes:** Road segments
- **Edges:** Connectivity between road segments

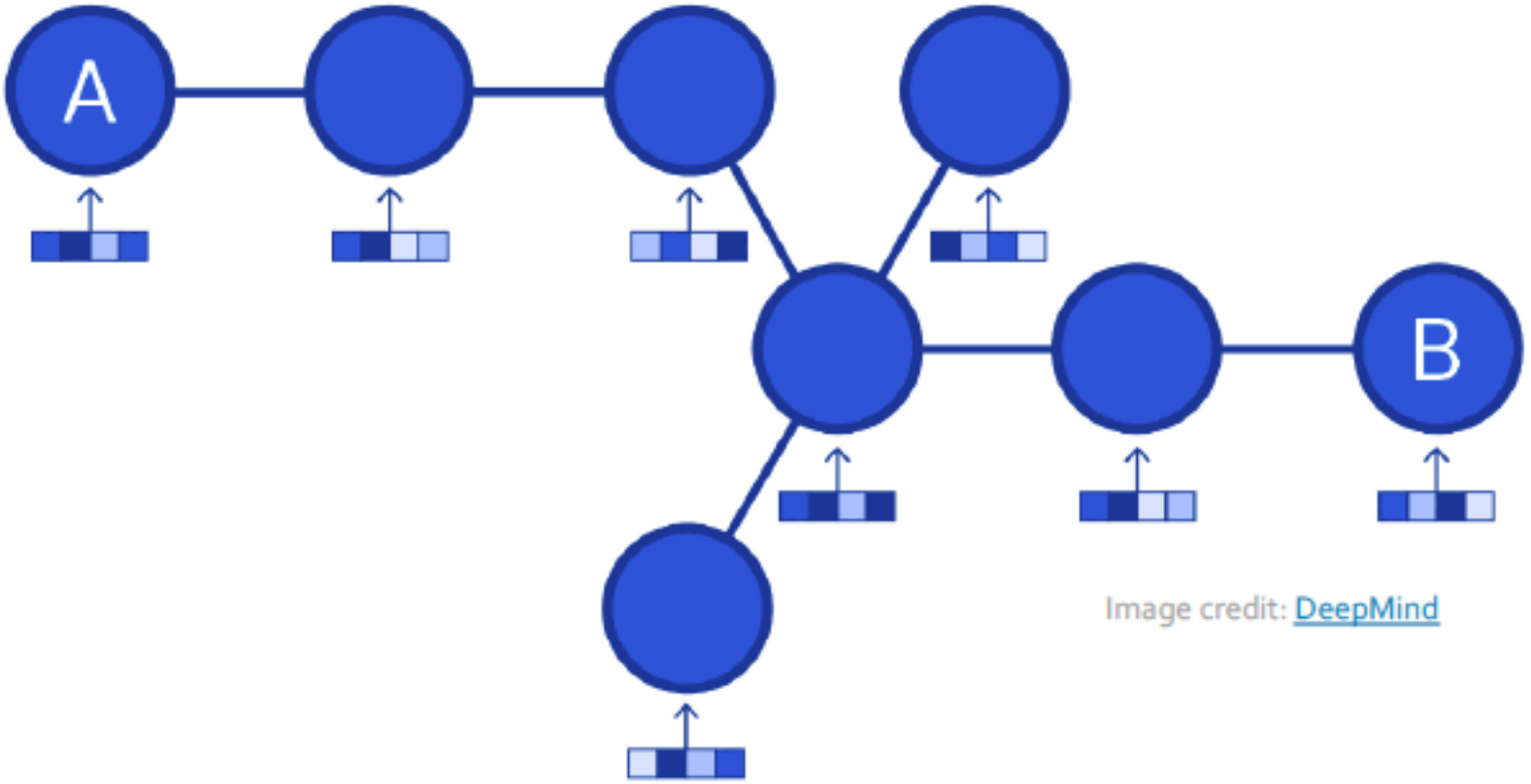
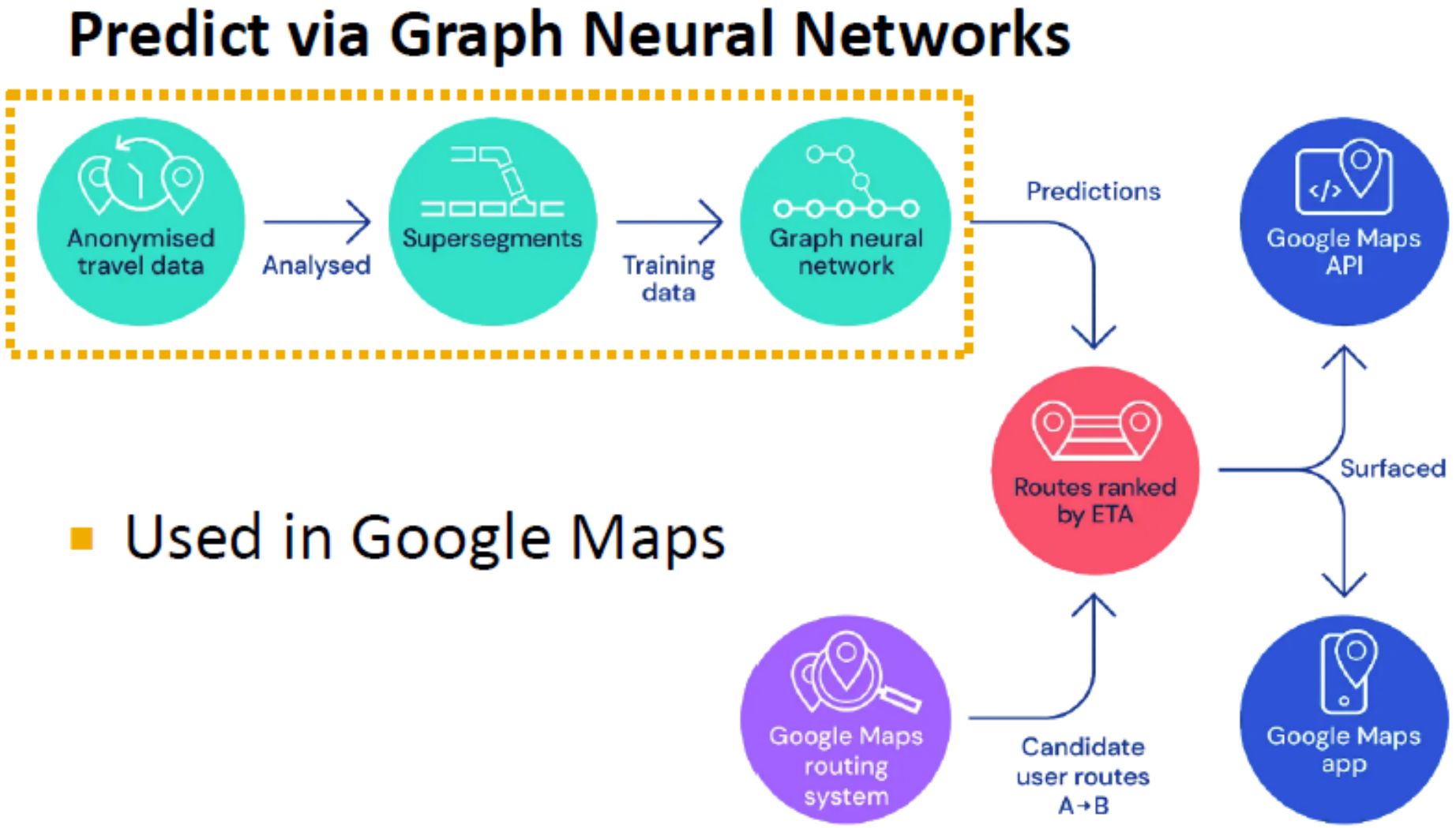


Image credit: DeepMind

Traffic Prediction via GNN



- Used in Google Maps

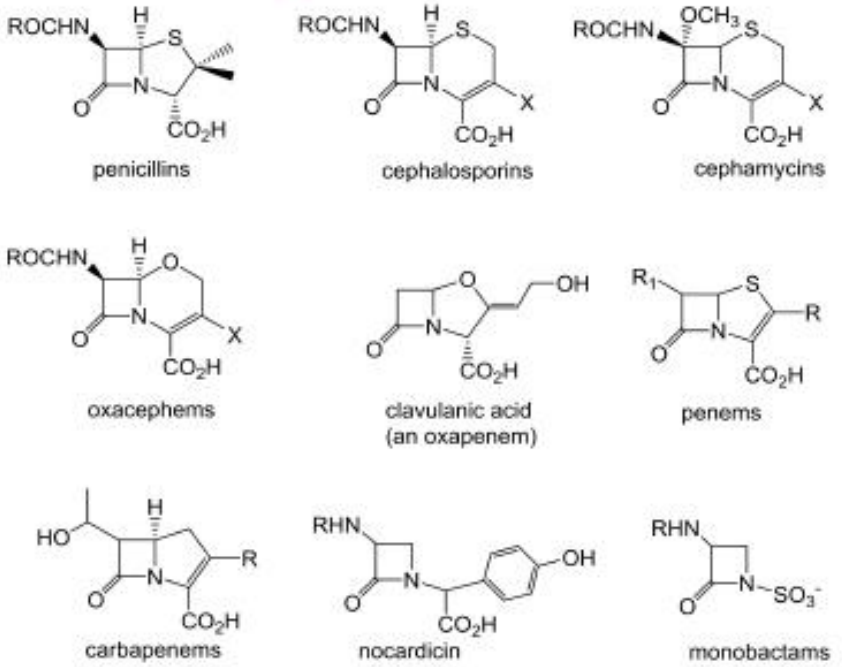
THE MODEL ARCHITECTURE FOR DETERMINING OPTIMAL ROUTES AND THEIR TRAVEL TIME.

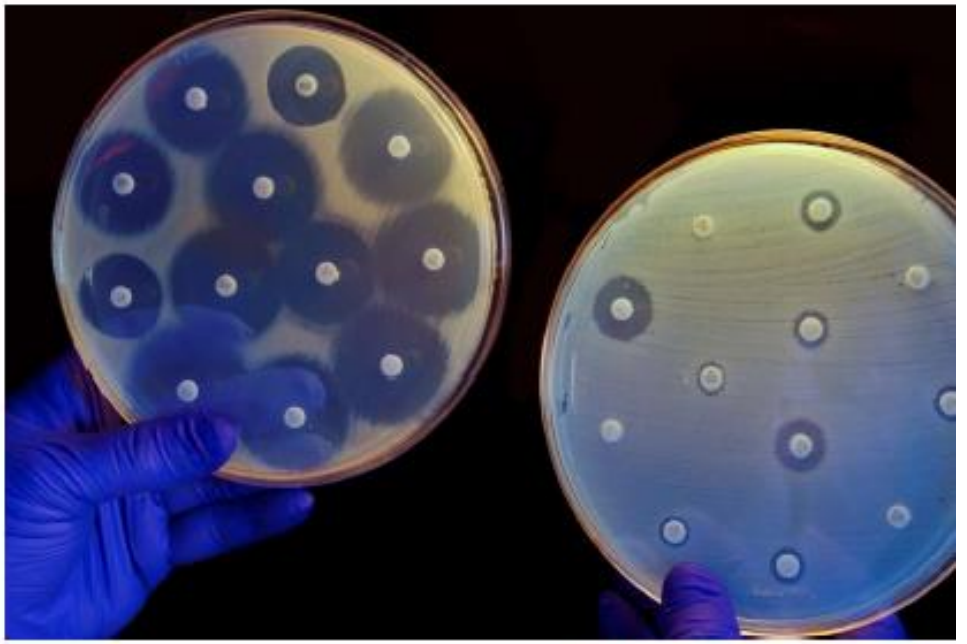
Image credit: DeepMind

Graph Level ML Tasks

Example Drug Discovery

- **Antibiotics are small molecular graphs**
  - **Nodes:** Atoms
  - **Edges:** Chemical bonds



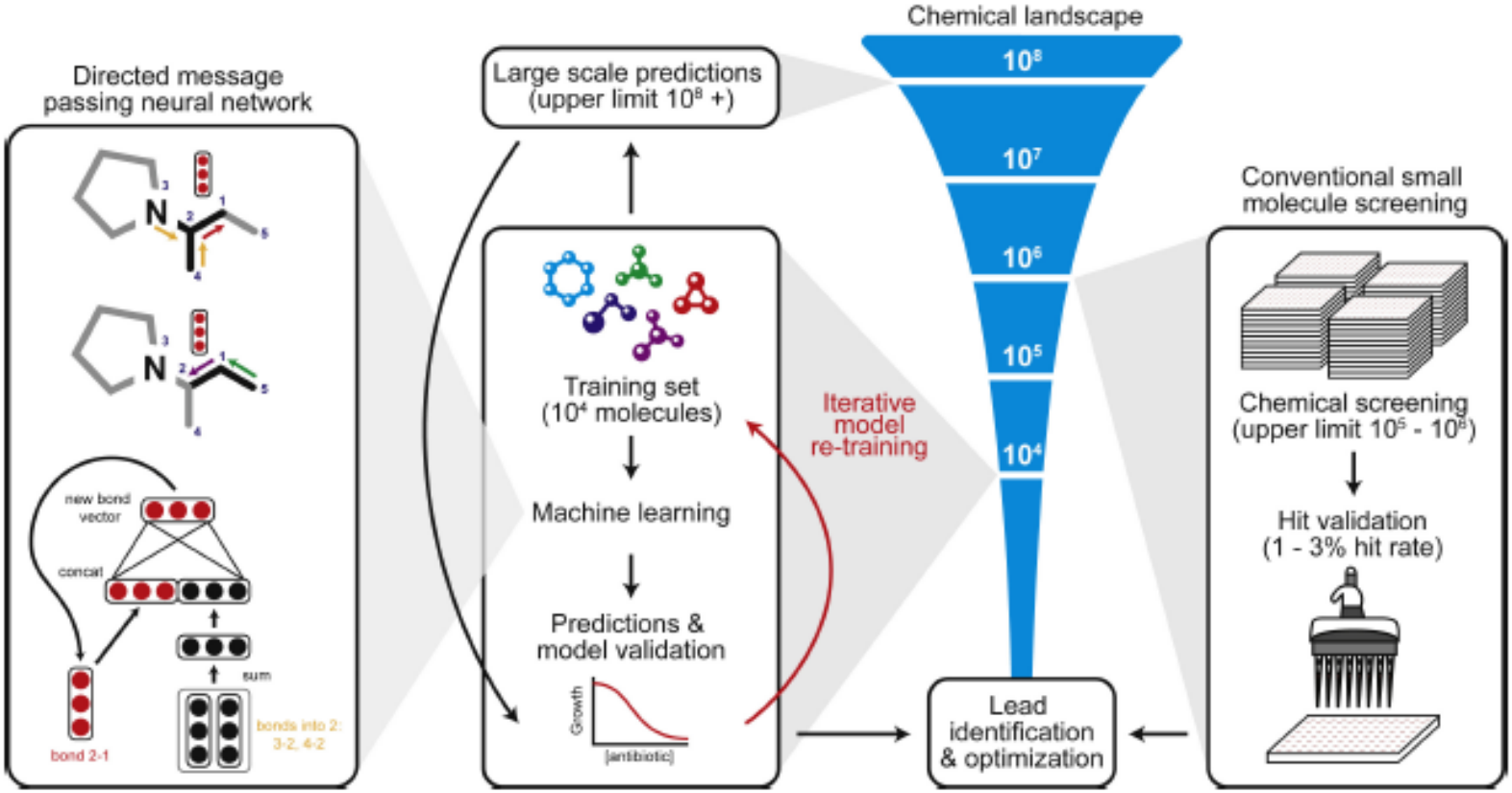


Konaklieva, Monika I. "Molecular targets of β-lactam-based antimicrobials: beyond the usual suspects." Antibiotics 3.2 (2014): 128-142.

Image credit: CNN

Models for Antibiotic Discovery

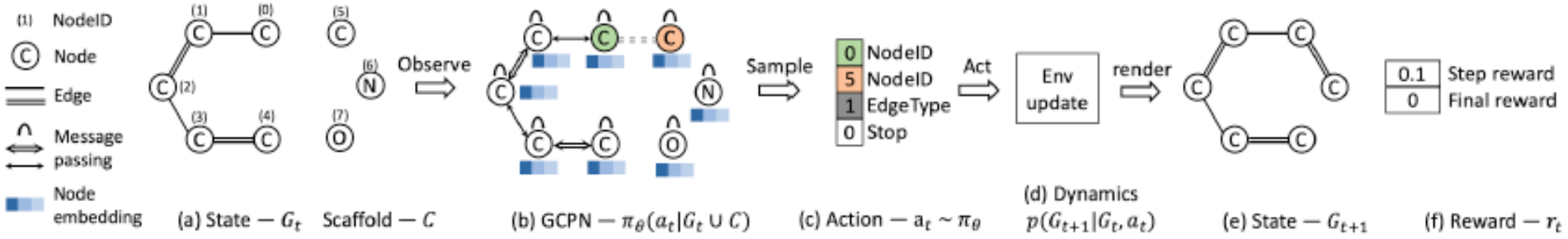
- A Graph Neural Network **graph classification model**
- Predict promising molecules from a pool of candidates



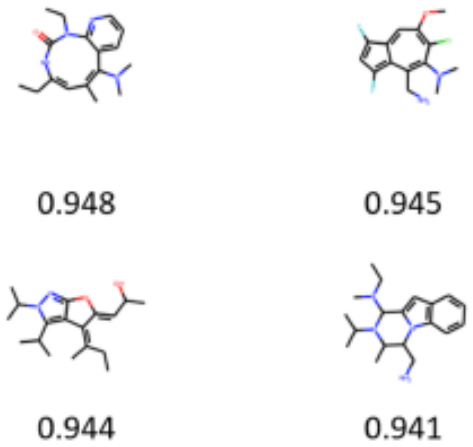
Stokes, Jonathan M., et al. "A deep learning approach to antibiotic discovery." Cell 180.4 (2020): 688-702.

Molecule Generation/Optimization

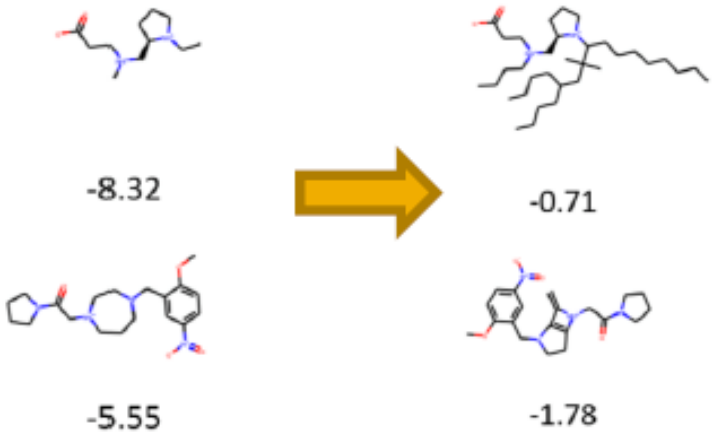
Graph generation: Generating novel molecules



Use case 1: Generate novel molecules with high drug likeness



Use case 2: Optimize existing molecules to have desirable properties



Physics Simulation

Physical simulation as a graph:

- **Nodes**: Particles
- **Edges**: Interaction between particles

Simulation Learning Framework

A graph evolution task:

- **Goal**: Predict how a graph will evolve over