

Machine Learning to Prioritize Drug Combinations in Cancers

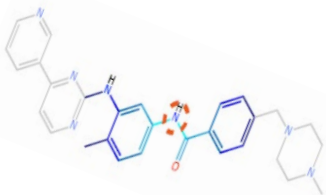
Yuanfang Guan

University of Michigan, Ann Arbor

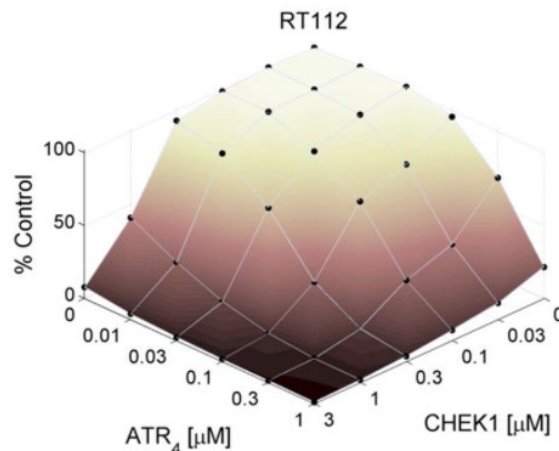
Mar 14, 2019

Predict which combination of small molecules is effective in treating cancers.

Chemical Structure

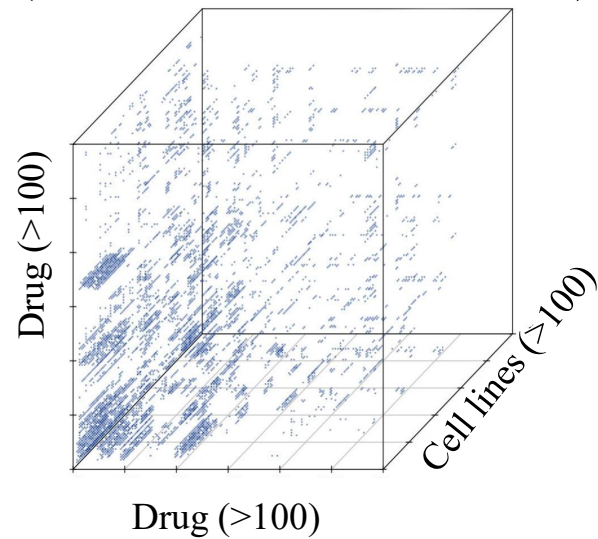


Drug Synergy

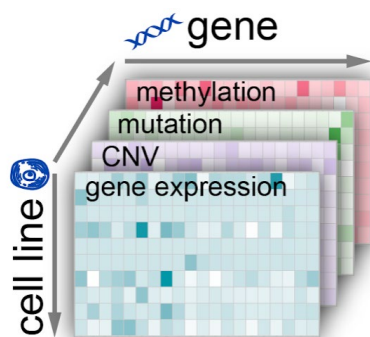


The challenge

Vast searching space
(>1 million combinations!)



Genomics



Benchmarked by the DREAM Challenge

The AstraZeneca-Sanger Drug Combination Prediction Challenge



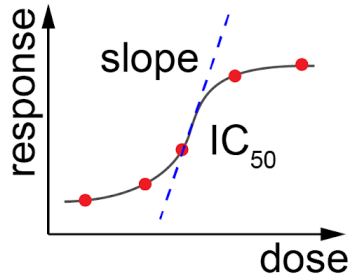
- 86 cancer cell lines
- 118 drugs
- 11,500 drug combinations
- **117 team submissions**
- Blind evaluation

We have approached accuracy of experimental replicates.

Li *et al.*, 2018 Cancer Research. Sep 15;78(18):5446-5457.

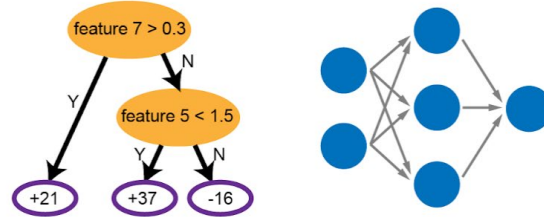
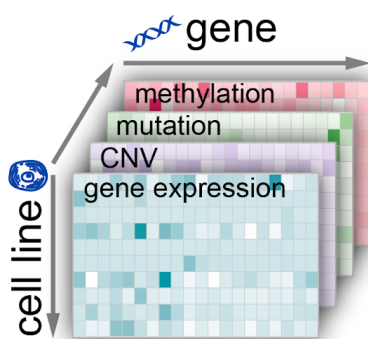
Machine Learning Predicts Drug Response & Synergy

Dose Response Curve



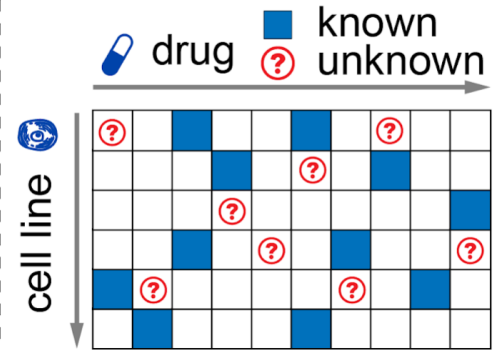
+

Genomic Profile

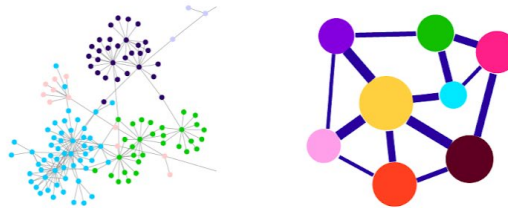


Machine Learning

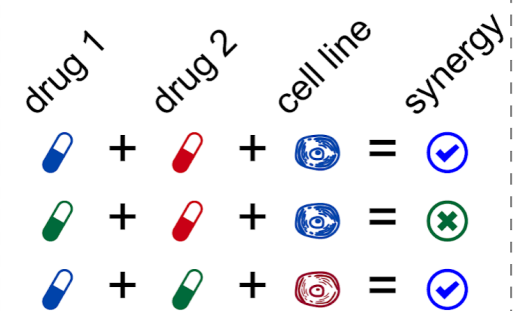
Single Drug Response



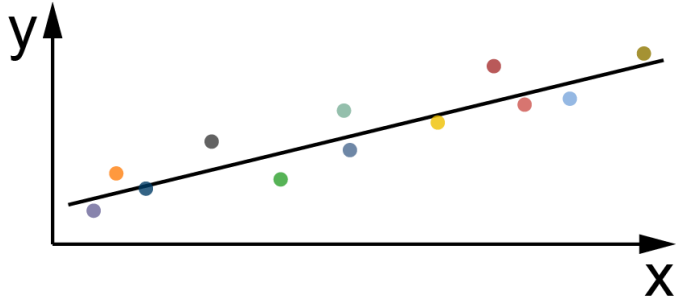
Domain Knowledge



Drug Synergy

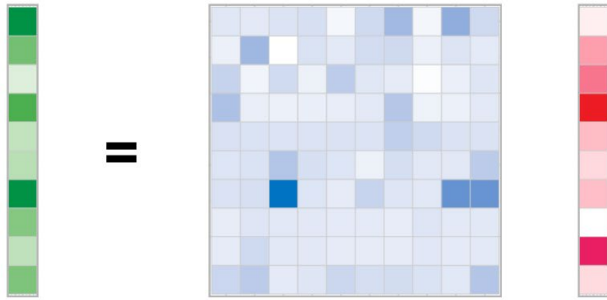


What are Machine Learning Problems



$$y = a(?) x$$

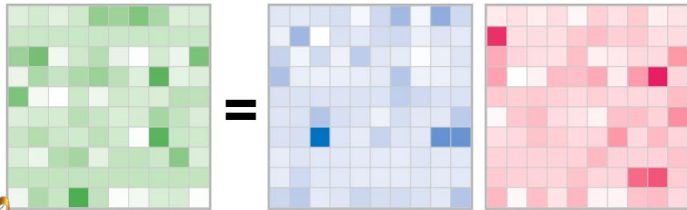
Knowing vector \mathbf{x} and \mathbf{y} ,
find the number \mathbf{a}



$$\mathbf{y} = \mathbf{X} \mathbf{a}$$

Knowing matrix \mathbf{X} and vector \mathbf{Y} ,
find the vector \mathbf{a}

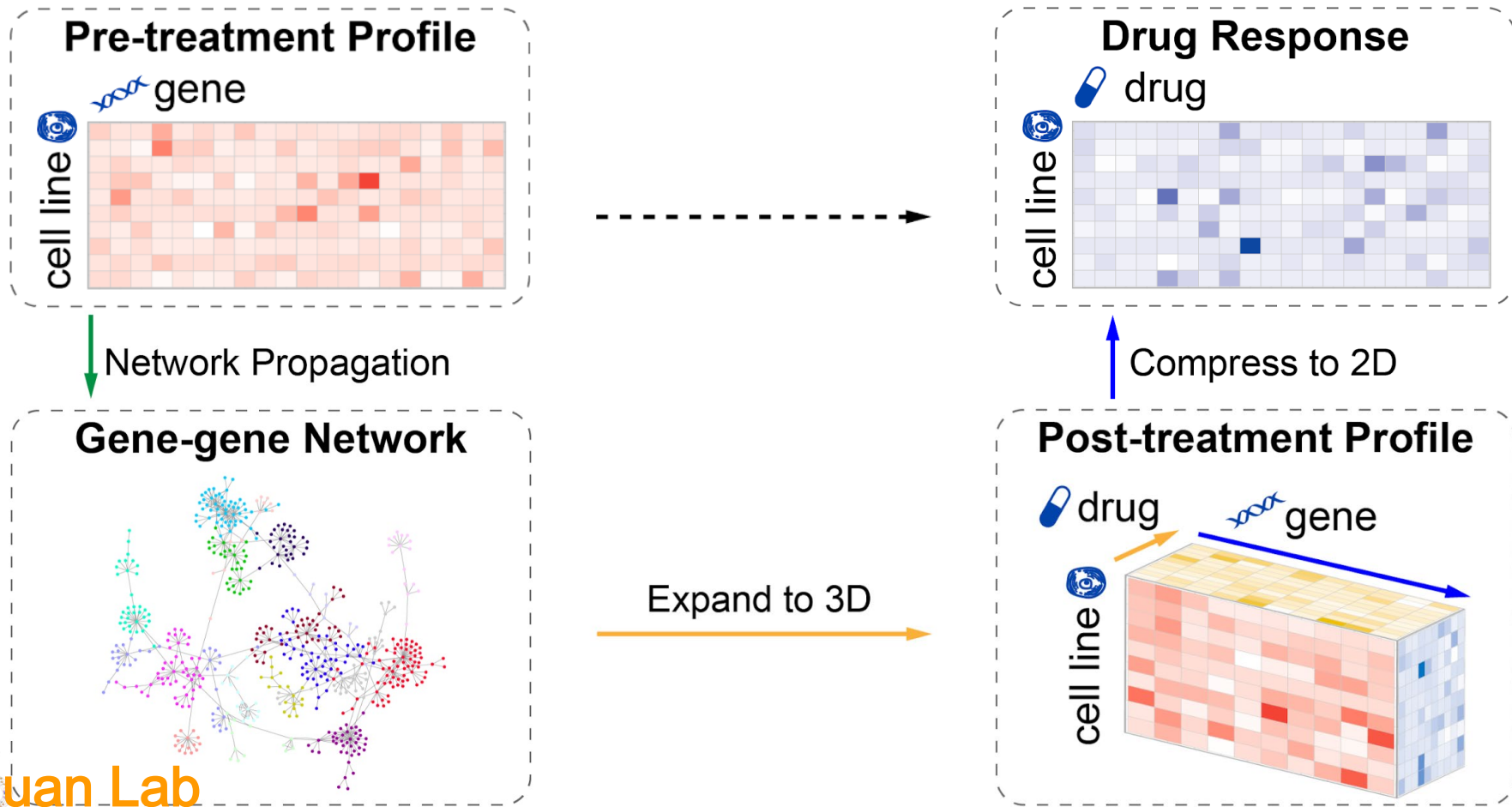
- Linear regression, Lasso, Ridge
- Support Vector Machine
- Gaussian Process Regression



$$\mathbf{Y} = \mathbf{X} \mathbf{A}$$

Knowing matrices \mathbf{X} and \mathbf{Y} ,
find the matrix \mathbf{A}


The solution: network propagation predicts drug synergy

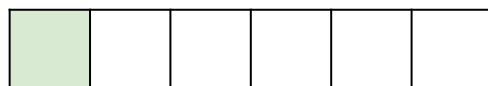


Molecular features of all genes

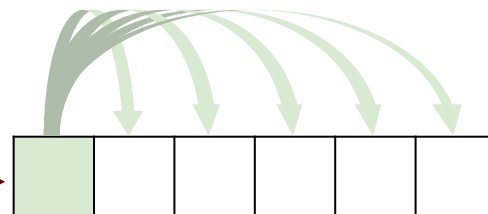


Molecular features of only drug target genes

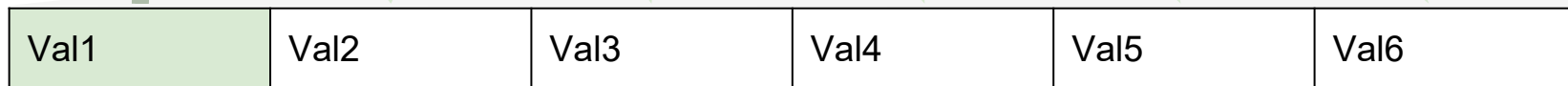
 Drug targets



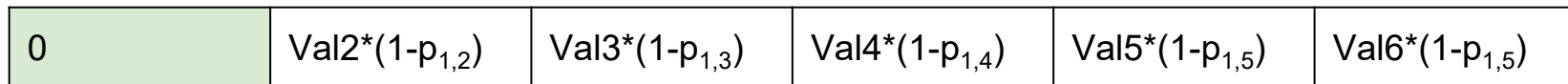
Map the target gene
to the gene network



Original expression patterns



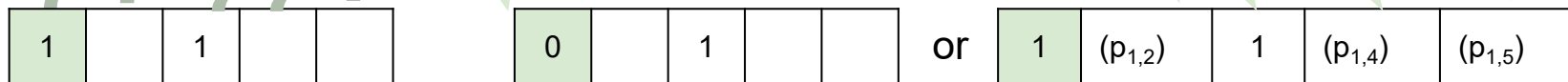
Simulate post-treatment expression patterns



Original mutation patterns



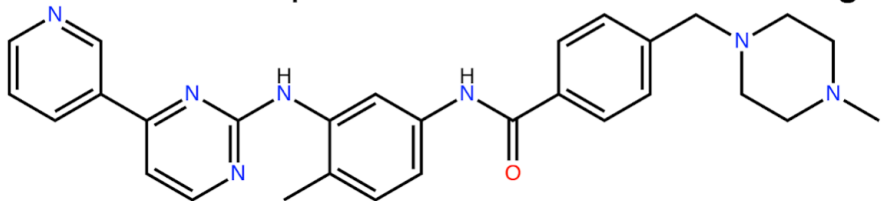
Simulated post-treatment mutation patterns



Future work: Graph Convolutional Neural Network (CNN)

Graph Theory: $G = (V, E)$

G : Molecule Graph V : Atom Vertex E : Bond Edge

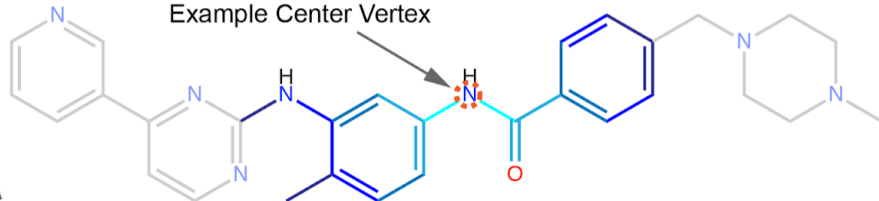


III

Convolution along Edges

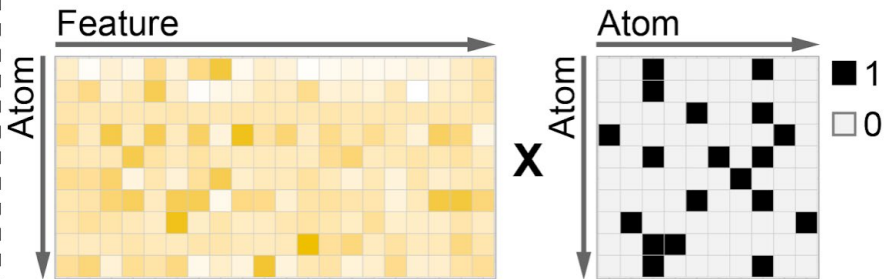
layer 1 layer 2 layer 3 layer 4 layer 5

Example Center Vertex

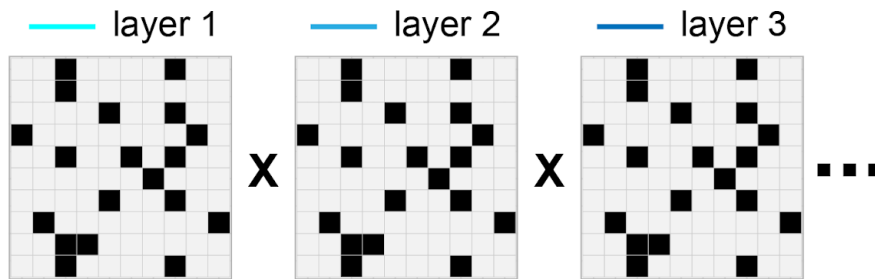


III

Input Feature and Adjacency Matrices



Aggregating Neighborhood Information

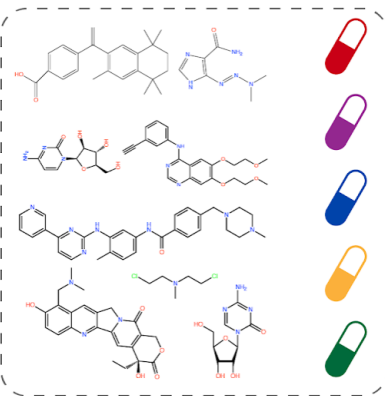


The Gene-Gene Network is also a Graph!

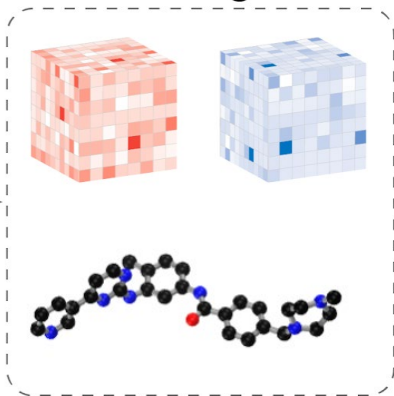


Integrating Gene-Gene Network into Deep Learning

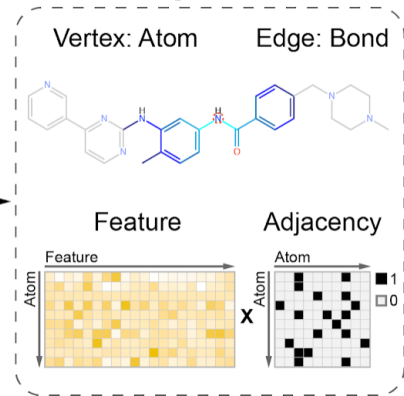
Small Molecules



3D Images



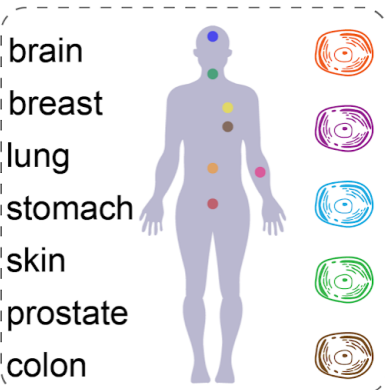
Graph CNN



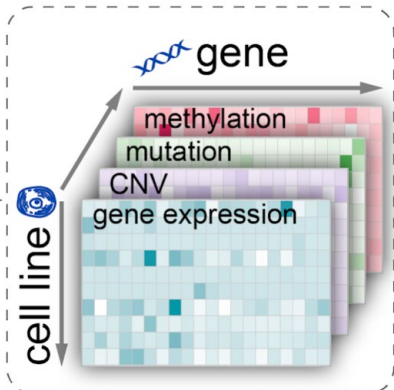
Principal Components



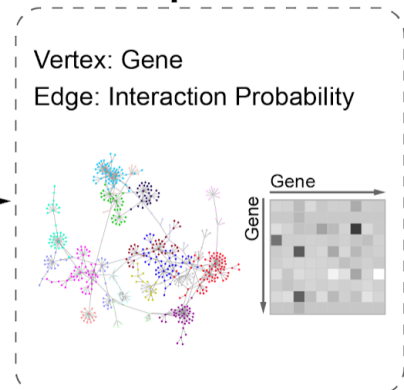
Cancer Cell Lines



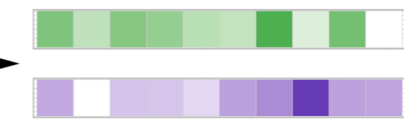
Genomic Profiles



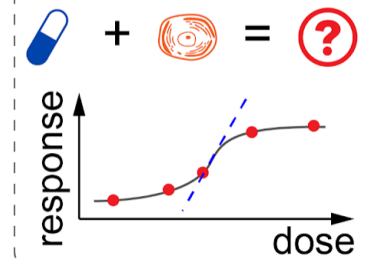
Graph CNN



Principal Components

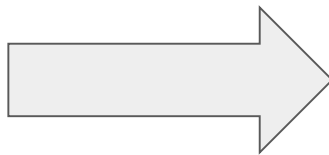
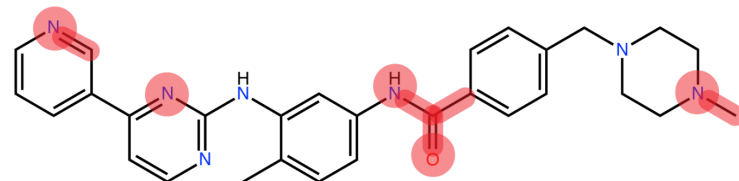
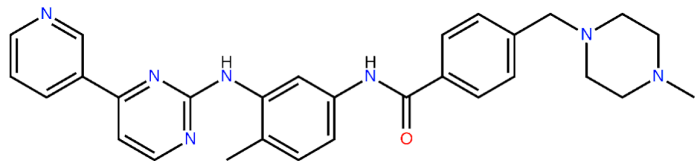
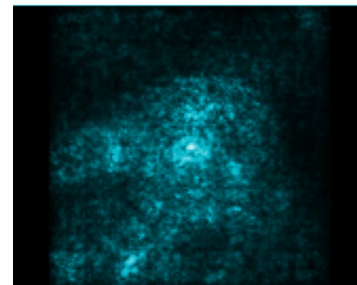


Prediction



Identifying Drug Synergy Biomarkers through Saliency Maps

Saliency maps visualize important features (e.g. pixels, genes, functional groups)



Thank you and questions