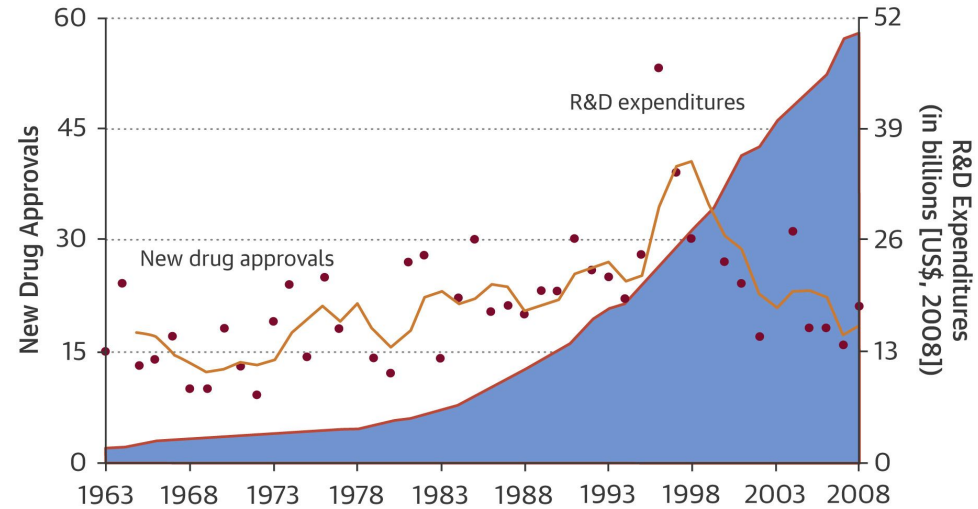
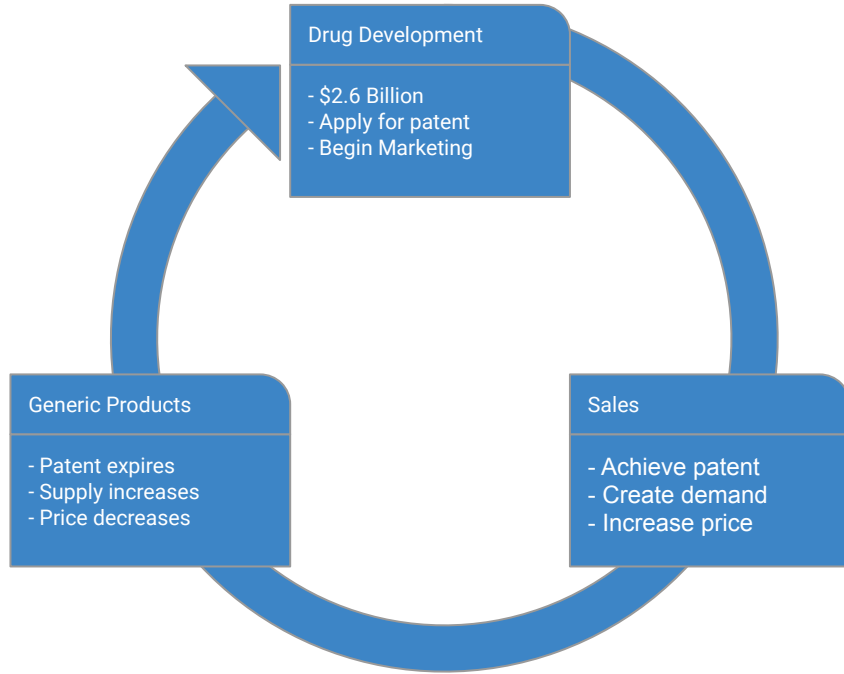




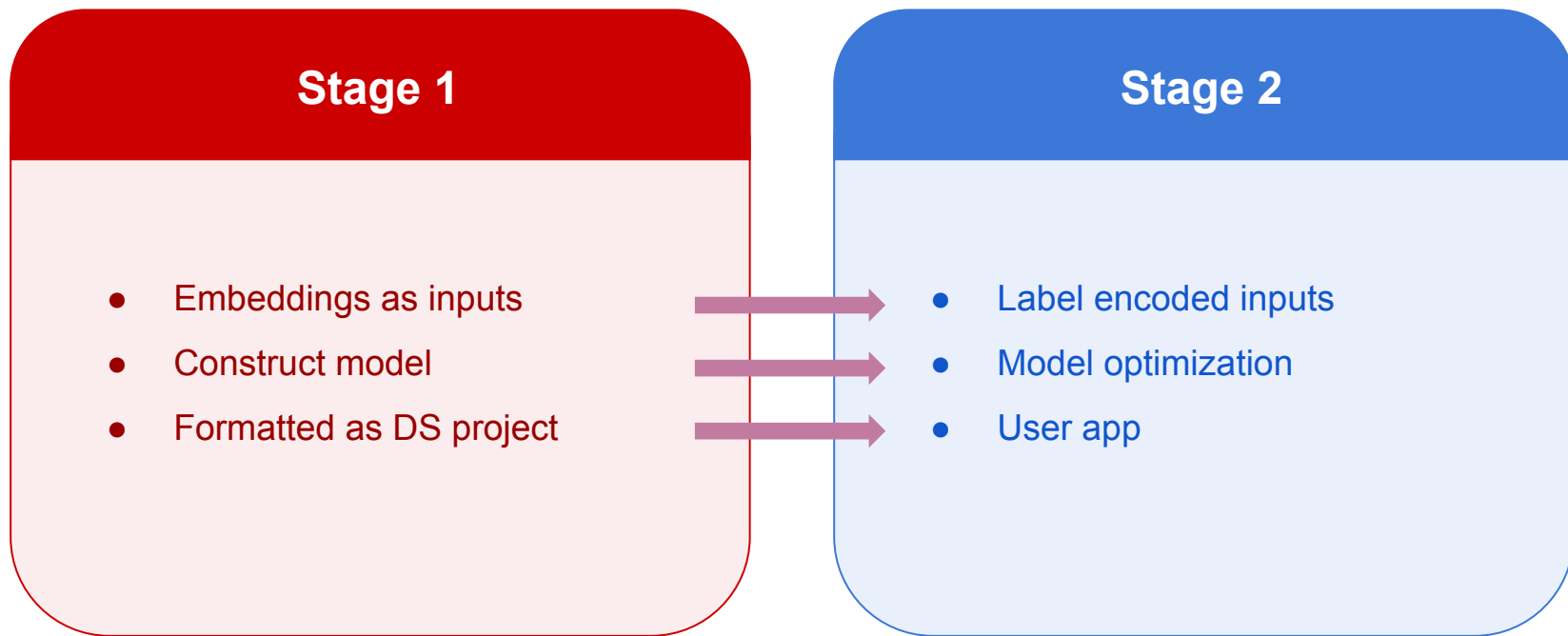
# Deep Drug Repurposer

By: Sarthak Kothari, Noah Demoës, Omair Shafi  
Ahmed, Forrest Hooton

# Recap: Motivation



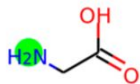
# Stage 2



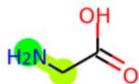
# Drug Fingerprints

- Fingerprints are low resolution molecular representations
- Multiple methods exist to create chemical fingerprints: ECFP, SMILES
- Objective: Create unique, machine readable representations of molecules

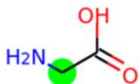
## Glycine:



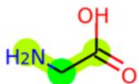
847957139



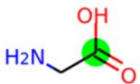
2592785365



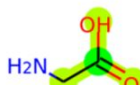
2245384272



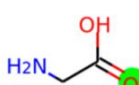
772927515



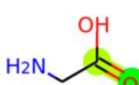
2246699815



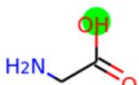
4278941385



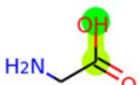
864942730



1510328189



864662311



1533864325

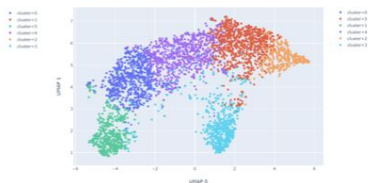
### Numerical Representation

1. Heavy atom connections
2. Non-hydrogen bonds
3. Atomic number
4. Sign of charge
5. Absolute charge
6. # attached hydrogens

## Genes

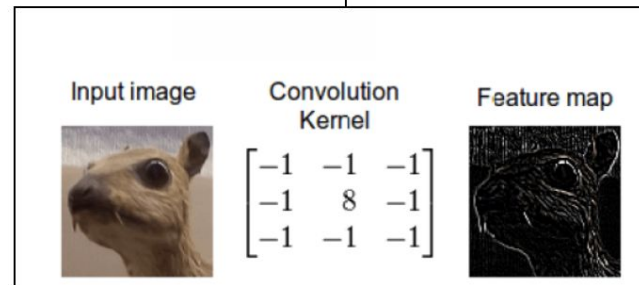
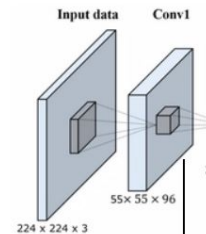


## 2) UMAP Reduction on Target Gene Embedding Vectors



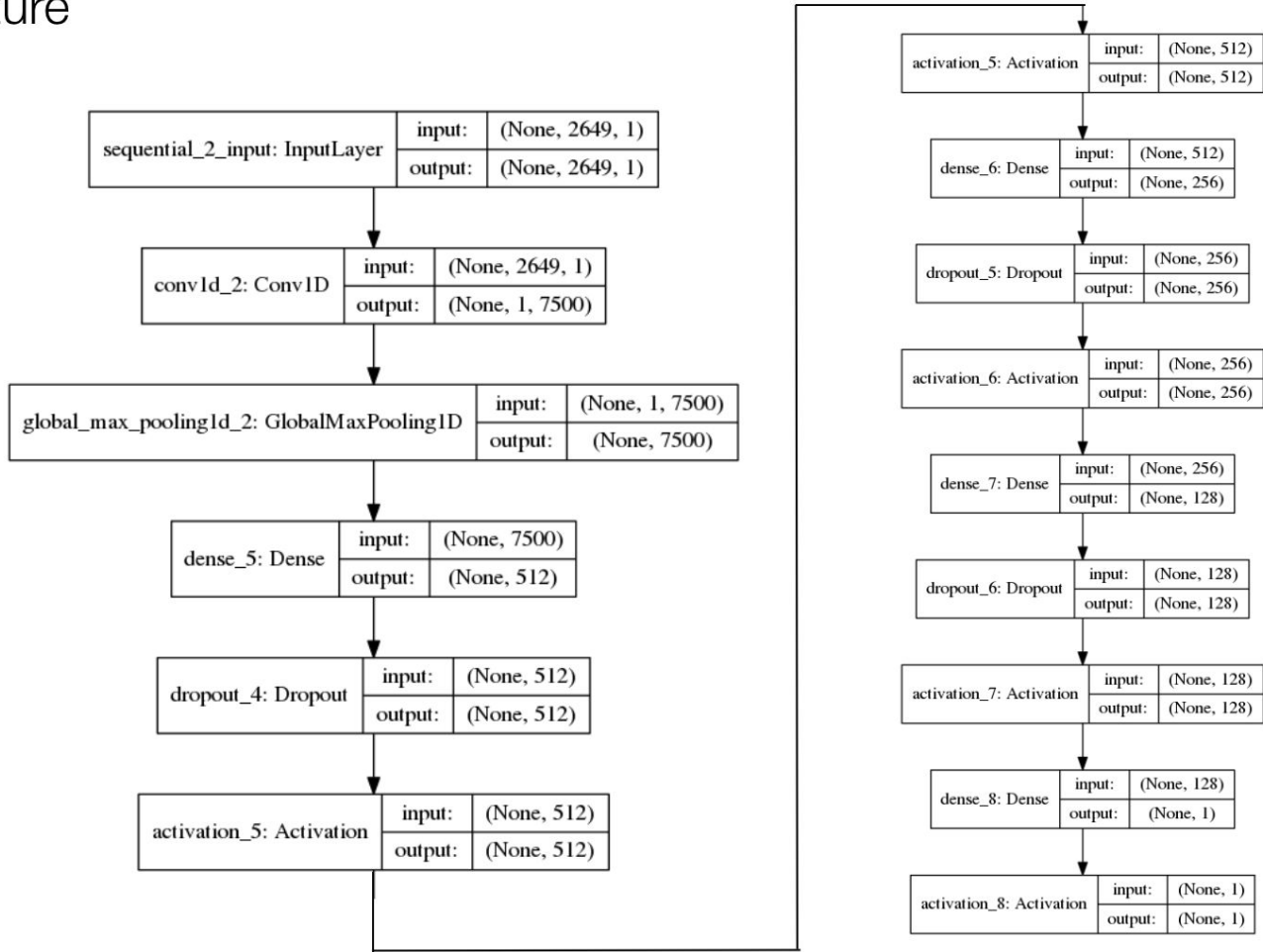
- Previous model used embeddings for feature reduction.
- Embeddings map the raw high dimensional data points to a lower dimensional space.
- While embeddings are a great way to reduce the feature space, they hide away the underlying structure (covariances).

**VS**

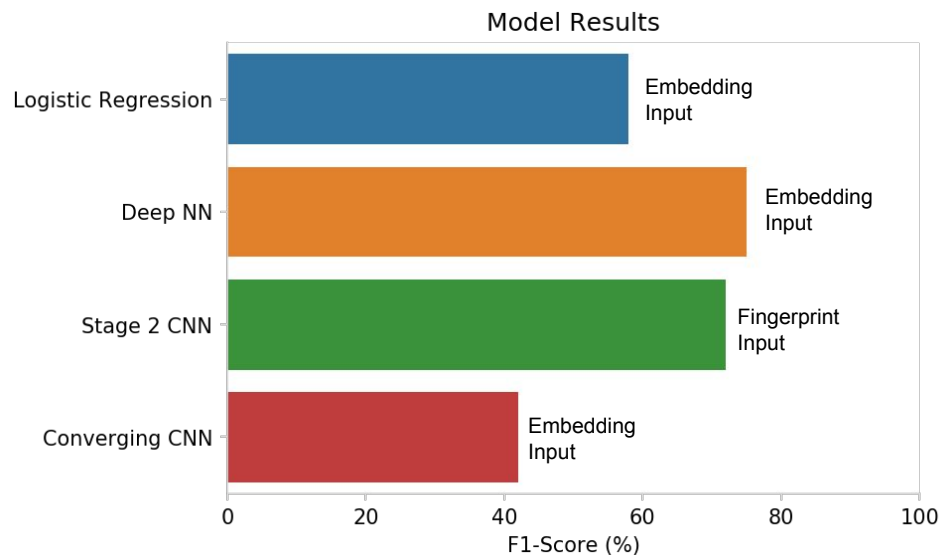


- Convolutional neural networks are a great way to learn higher order covariances and features, and leverage them for downstream learning tasks.
- However, a convolutional filter cannot learn from an embedding space that has hid away the structural covariances.

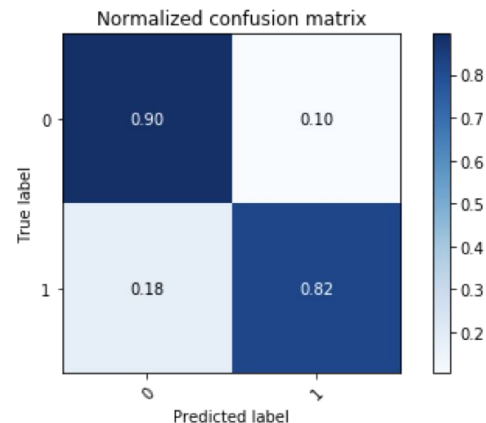
# Model: Architecture



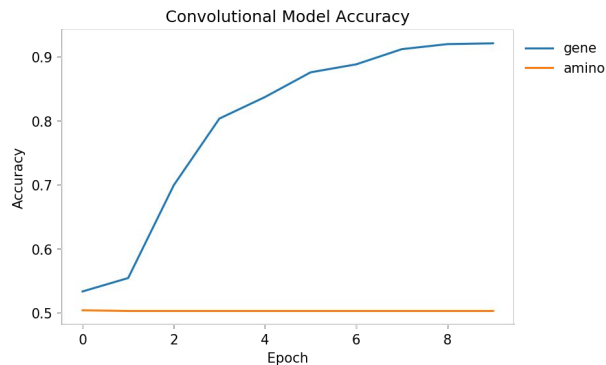
# Results



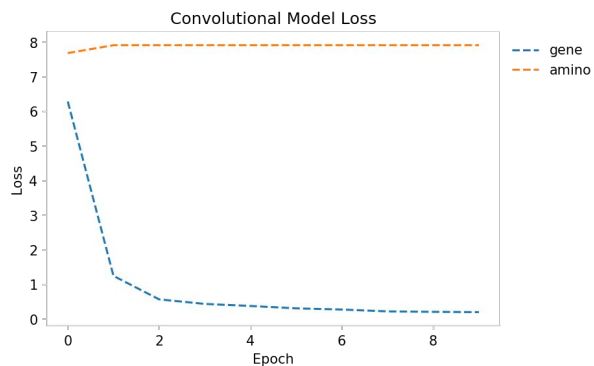
**Raw fingerprint input  
had noticeably better  
results**



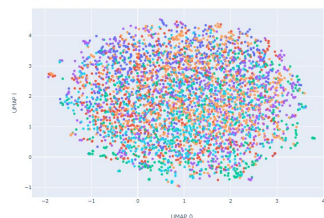
# Gene vs. Amino Acid Results (!)



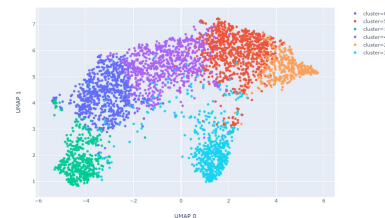
Model with amino  
acid representation  
is not learning



UMAP Reduction on Target Amino Embedding Vectors



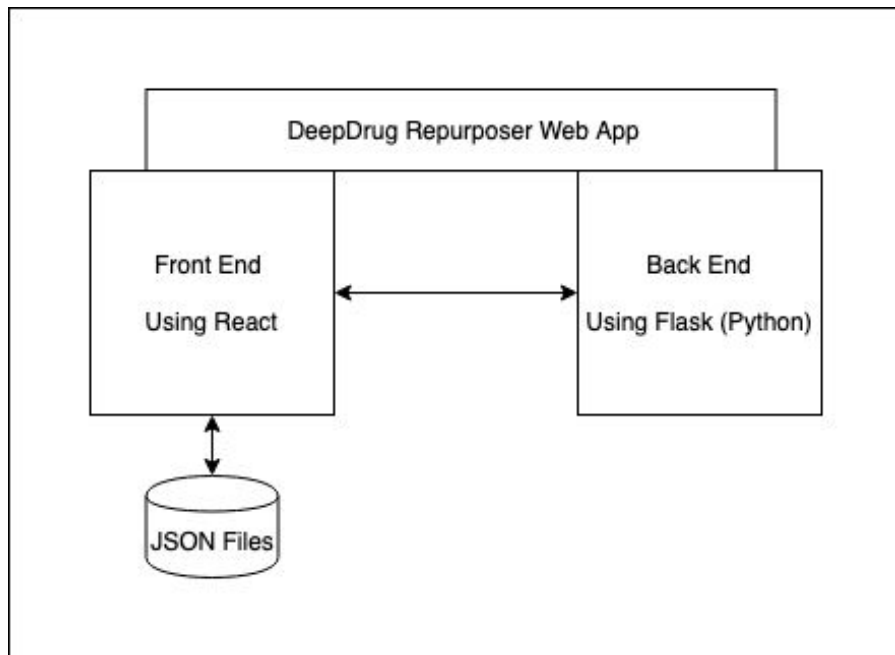
UMAP Reduction on Target Gene Embedding Vectors





# Technical Specifics of App

- Built using React for front end and Python for backend.
- App sends a GET request to the backend which runs the predictions against all the targets and returns the output to front end.
- Bank of input json files which can be used for getting the predictions.
- GCP for hosting this service





## Deep Drug Repurposer

Enter a drug or target name

A glass bottle is tipped over on a wooden surface, spilling various pills and capsules. The pills are in various colors including white, yellow, orange, green, red, and blue. The word "Questions" is centered over the image.

# Questions

Normalized confusion matrix

