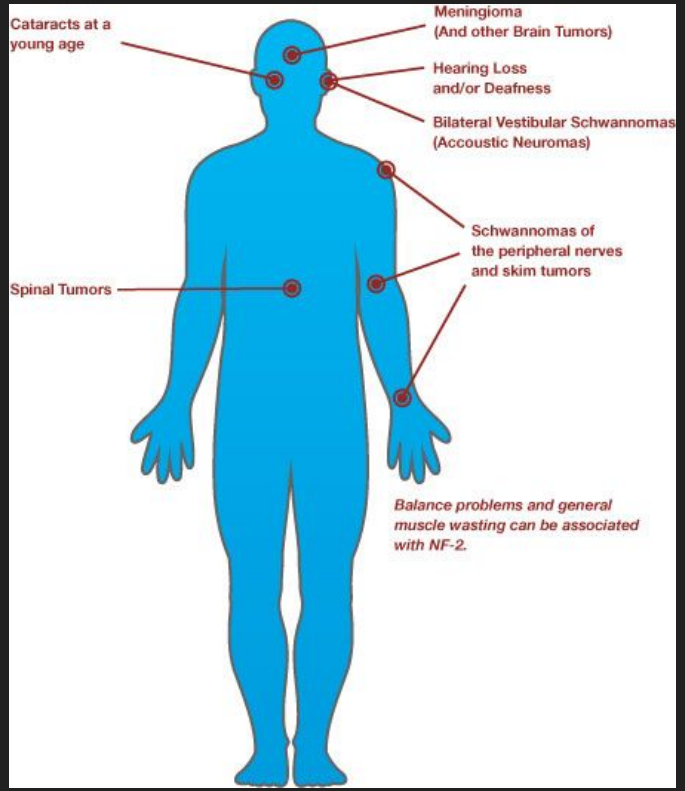
Applying an unsupervised transfer learning approach to identify novel targets

Team AutoNF2

Challenge- NF2 tumors is a “Rare” disease

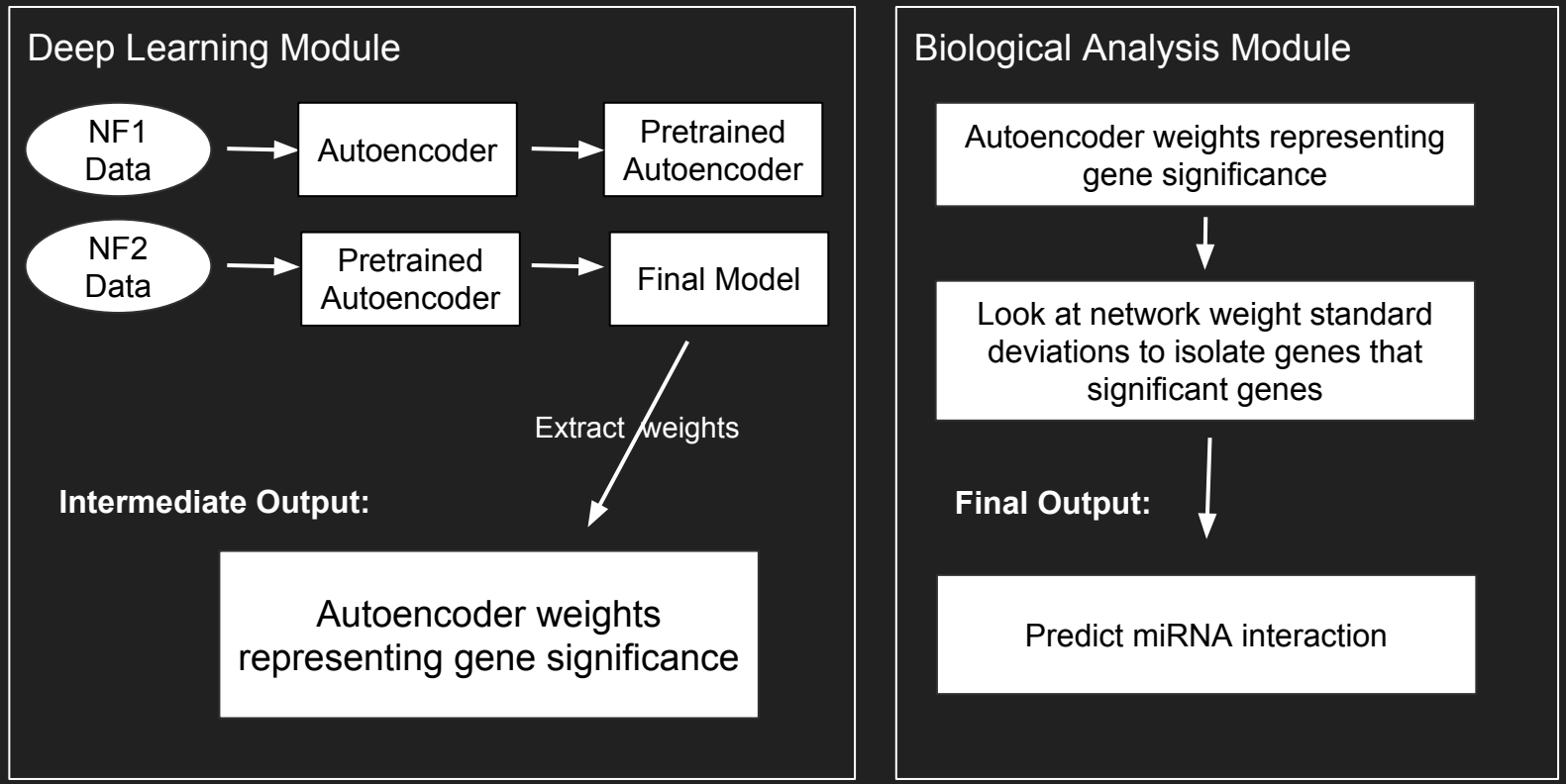
- NF2 tumors is rare genetic condition with an increased risk of other tumors in nervous system
- Limited resources makes a challenge to identify novel therapeutic targets



Need to identify therapeutic targets in NF2 tumor!

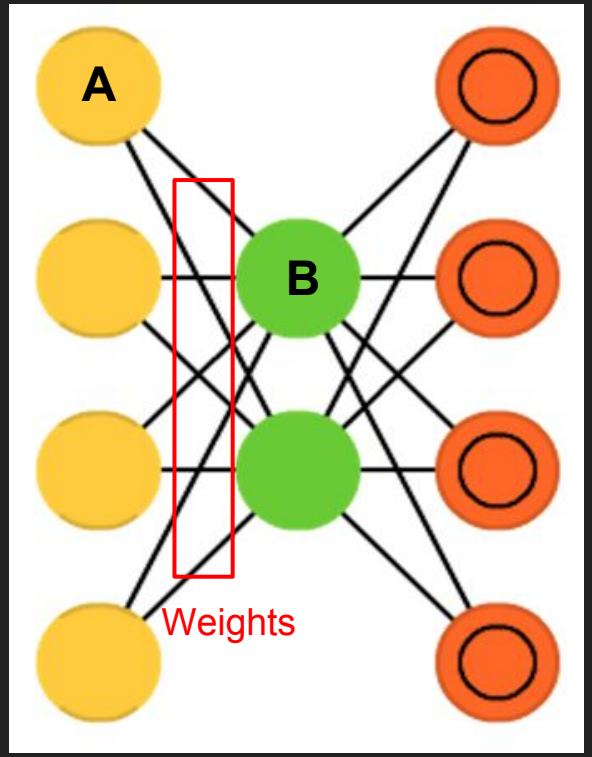
- Understanding of the genomic aberrations that drive these tumors remains incomplete
- Numerous clinical trials have failed to identify systemic medical therapies that can effectively control the relentless growth of these tumor
- Need for bigger sample size for more rigorous in silico techniques
- Potential solution: **Applying deep learning to identify potential novel therapeutic targets!**

A multi-step deep learning pipeline



Autoencoder for dimensionality reduction

Input data is a vector of gene expressions (FPKM)



- An Autoencoder is trained to regenerate the input.
- Forces network to generate very efficient encoding of features
- Thus, weights contain information about what features are important.
- Large weight from A → B signifies that gene A could potentially influence the biological property underlying node B

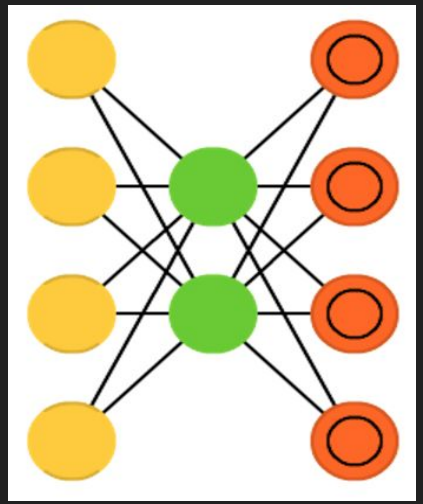
http://www.worldscientific.com/doi/pdf/10.1142/9789814644730_0014

Transfer Learning - train with limited data

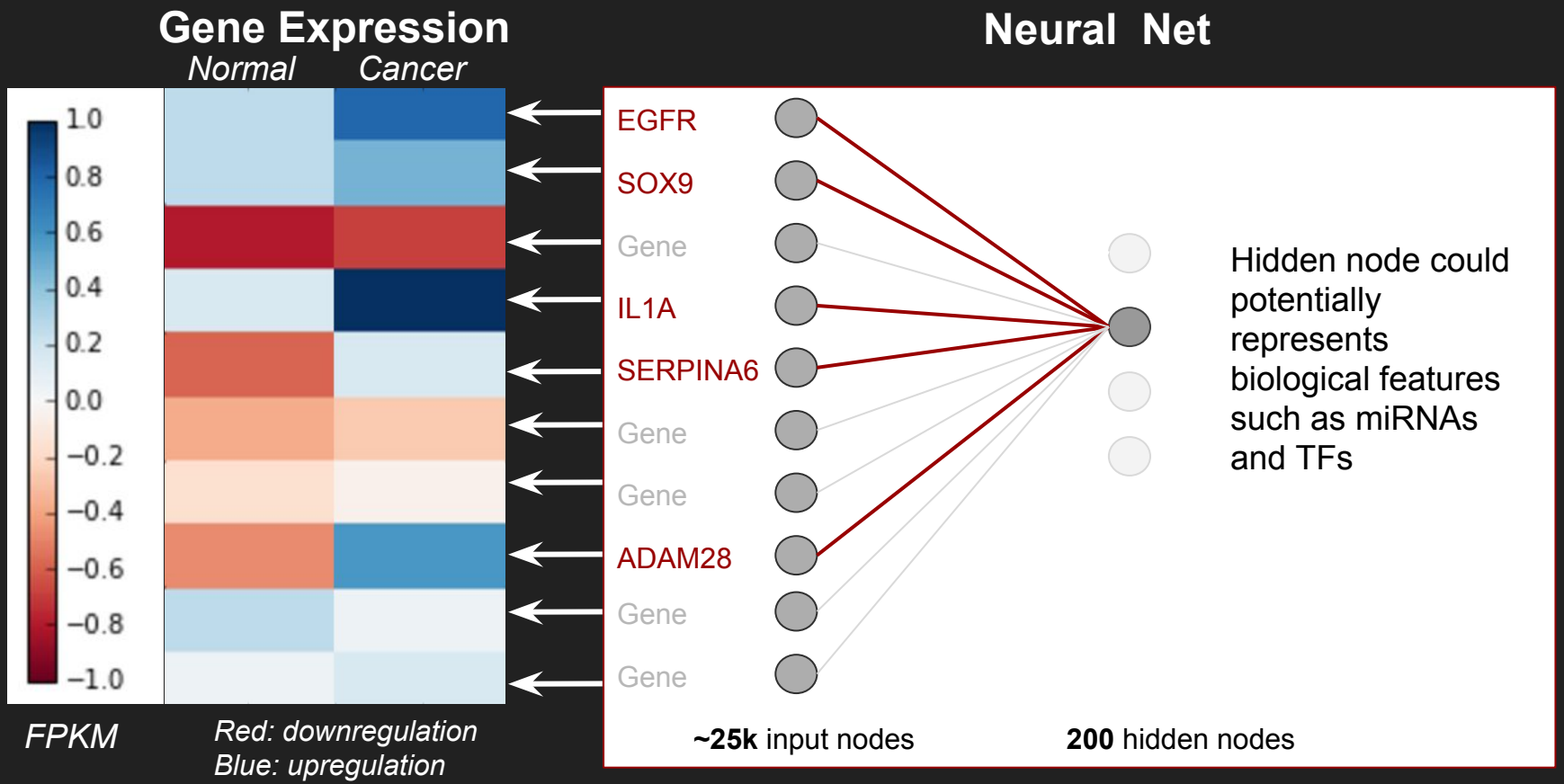
1. Train with larger NF1 dataset

2. Train with small NF2 dataset

Intuition: NF1 and NF2 are similar, means network pretraining is viable

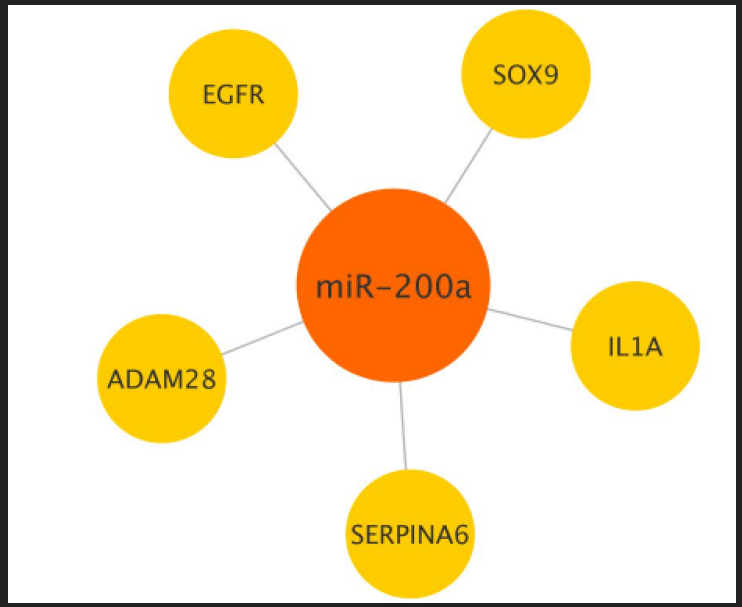


Interpreting constructed features reveals cancer related gene interactions



miR-200a as a potential circulating marker for NF2 tumor!

- miRNAs are a part of non-coding RNAs that comprises the 98% of the human genome!
- miR-200a is associated with poor survival in various tumor types (Nakamura et al. 2016)
- Neural net predicted miR-200a as a potential player in association with cancer causing genes



Source: http://mirdb.org/index.html

Future Directions

- First of a kind!
- We leverage a larger related data to effectively study a small dataset for NF2 tumors
- Perform more exhaustive search of hyperparameters for model
- Identify potential therapeutic targets e.g. EGFR inhibitor- Erlotinib
- Potential circulating miRNAs for early diagnostics e.g. miR-200a

A truly interdisciplinary team!!!!

Jo Varshney, DVM/PhD, Cancer Researcher (UCSF)
jyotika71@gmail.com

Nandita Damaraju, Software Engineer (Thermo Fisher)
nandita.damaraju@gmail.com

Ben Hsu, CS & Business
benhsu75@gmail.com

Teng Gao, CS & Bio student
gaoteng@wustl.edu