Developing a deep neural network model capable of conducting virtual gene manipulation experiment

Minh Nhat Nguyen, Son Le, David Tran
Dept of X, Bridge Institute, University of Southern California, Los Angeles, CA, USA

Abstract

The common approach of determining master regulator genes (MR) in disease has relied largely on trial-and-error experimental manipulation of individual genes, followed measuring the impact on disease phenotype. However this method is inefficient and incapable of identifying cooperative MR specific to individual patients. Therefore, there is a great need for a new algorithm that can identify cooperative MR specific to individual patients by virtual manipulation in computational simulation.

Auto-encoder model is an effective tool to model complex unknown relationships in biology due to their ability to learn and capture meaningful features and patterns. Utilizing transfer learning, we could train an auto-encoder network model based on large scale public data, and then fine tune it to a specific downstream task, identifying candidate therapeutic targets for specific disease. This approach can help accelerate discovery of key network regulators and candidate therapeutic targets.

Objectives

- Design an auto-encoder model capable of mapping genes-genes connection within specific human cell, with the model being trained on 237,824 training data points from RNA-seq dataset from Gene Expression Omnibus (GEO)
- Assess the performance of the model by judging its ability to model the gene gene relationships by comparing the model predicted values with the ground truth.

HPC’s Nodes Configurations

- Type of GPU: a100
- Total GPUs: 96
- Memory per CPUs: 32GB