

Inference of gene regulatory networks from scRNA-Seq data benefits from prior knowledge.

A graph structured prior can be incorporated using graph autoencoders.

HMGU



Scan to download a digital copy of the poster.

Institute of
Epigenetics and Stem Cells

Institute of
Computational Biology

Institute of
Functional Epigenetics

Introduction

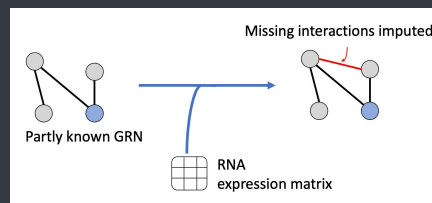
Genes are interacting in complex network structures. The computational reconstruction of such Gene Regulatory Networks (GRN) in the representation of graphs from sequencing data, remains a very important and challenging task. While individual gene interactions may be experimentally detected and validated by, e.g., knockout experiments, this approach lacks the scalability to infer large interaction networks.

References

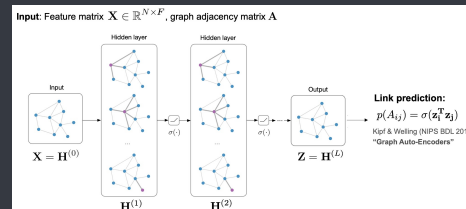
- [1] T. N. Kipf, M. Welling, Variational Graph Auto-Encoders, NIPS Workshop on Bayesian Deep Learning (2016)
- [2] M. Stone et al, Identifying strengths and weaknesses of methods for computational network inference from single cell RNA-seq data (2021)
- [3] J. Wang et al, Inductive inference of gene regulatory network using supervised and semi-supervised graph neural networks, Computational and structural biotechnology journal (2020)

Methodology

Here we use a Variational Graph Autoencoder (VGAE) to complete partly known GRNs borrowing information from RNA sequencing data sets. This self-supervised machine learning method uses a given incomplete GRN as prior knowledge to better predict missing gene interactions.



The model is based on the graph autoencoder proposed by T. Kipf 2016:



- **Encoder:** two layers of graph convolutional layers (GCN)
- **Decoder:** simple dot product
- **Loss function:** crossentropy loss of the resulting adjacency matrix to the prior knowledge graph

Results

The model outperforms methods that don't use prior knowledge. For the benchmarked DREAM5 data sets it achieved comparable performance to the GRGNN method by Wang et al 2020.

Dataset	Species	Cell type	Benchmark Algorithm	Prior knowledge	Test set	Benchmark AUROC	AUROC [0-1]
Gasch et al 2017	Yeast		SCODE	GAE: 99%		0.06	0.2
Jackson et al 2020	Yeast		Pearson	Benchmark: 0%	Imbalanced	0.04	0.32
Tran et al 2019 (A25)	Mouse	ESC	SILGGM			0.1	0.45
Tran et al 2019 (FBS)	Mouse	ESC	SILGGM			0.1	0.42
Dataset	Species	Cell type	Benchmark Algorithm	Prior knowledge	Test set	Benchmark AUROC	AUROC [0-1]
DREAM5 challenge	E. coli		GRGNN	67%	balanced	0.9	0.88
Marbach et. al. 2012	Yeast		Wang et al 2020			0.88	0.85

Acknowledgement

Marco Stock is supported by the Helmholtz Association under the joint research school "Munich School for Data Science - MUDS".

V1.0

Self-supervised completion of gene regulatory networks using graph autoencoders

Marco Stock, Antonio Scialdone

MUDS TUM