

Marco Varrone

MSc student in CS - Supervisor: **Stefano Ceri**

Graduation: April 2020



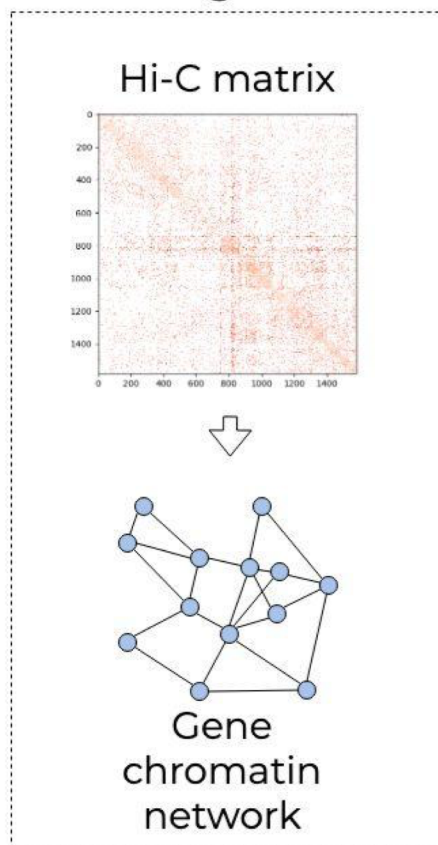
Co-expression network inference from chromatin conformation data through graph embedding

Relevant publications/submissions

- **Marco Varrone**, Luca Nanni (co-first), Giovanni Ciriello, Stefano Ceri, *Co-Expression Network Inference from Chromatin Conformation Data through Graph Embedding*. **Under review @ Intelligent Systems for Molecular Biology Conference (ISMB)**.

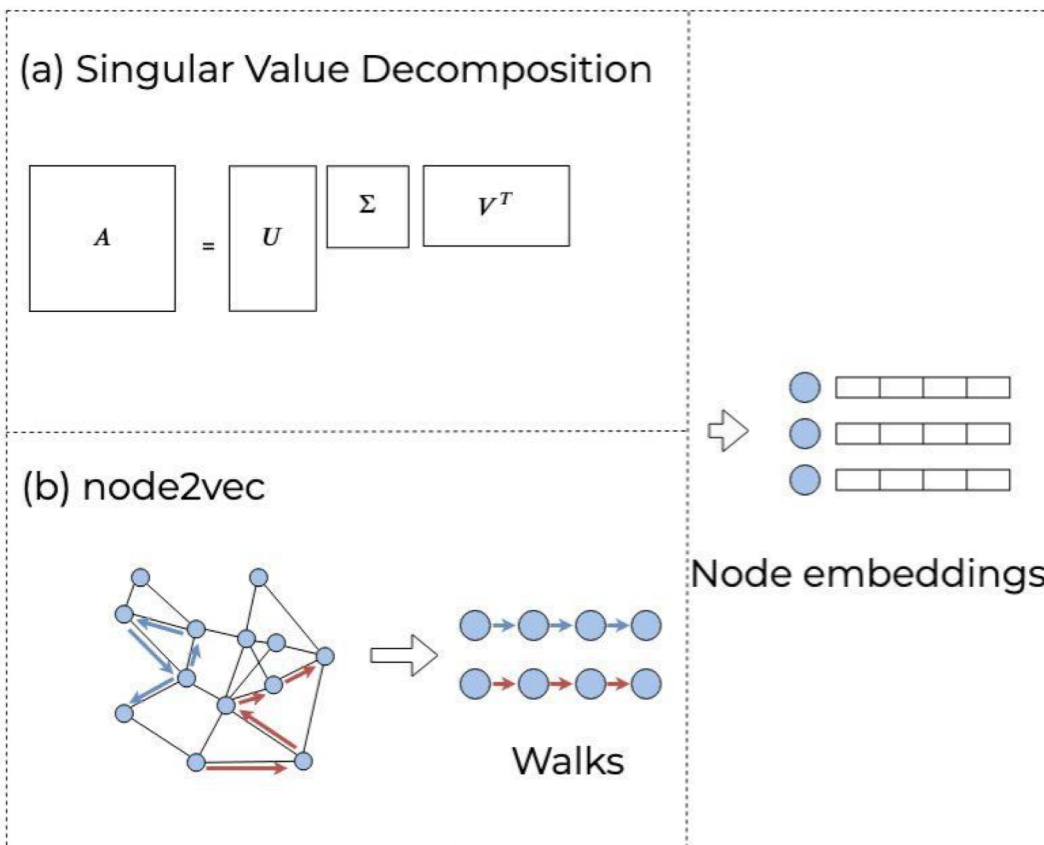
Can we predict co-expression from the chromatin interactions?

(1) Gene chromatin network generation



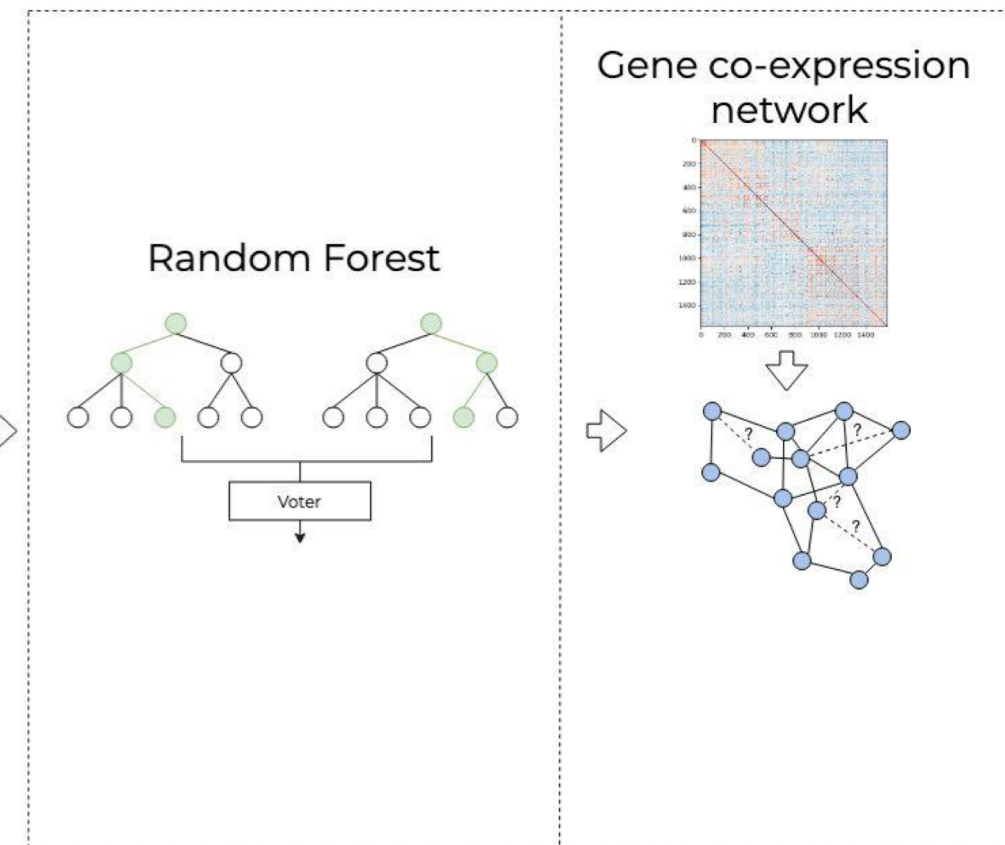
Network building

(2) Gene-node embedding generation



Embedding learning

(3) Link prediction



Model training

Evaluation

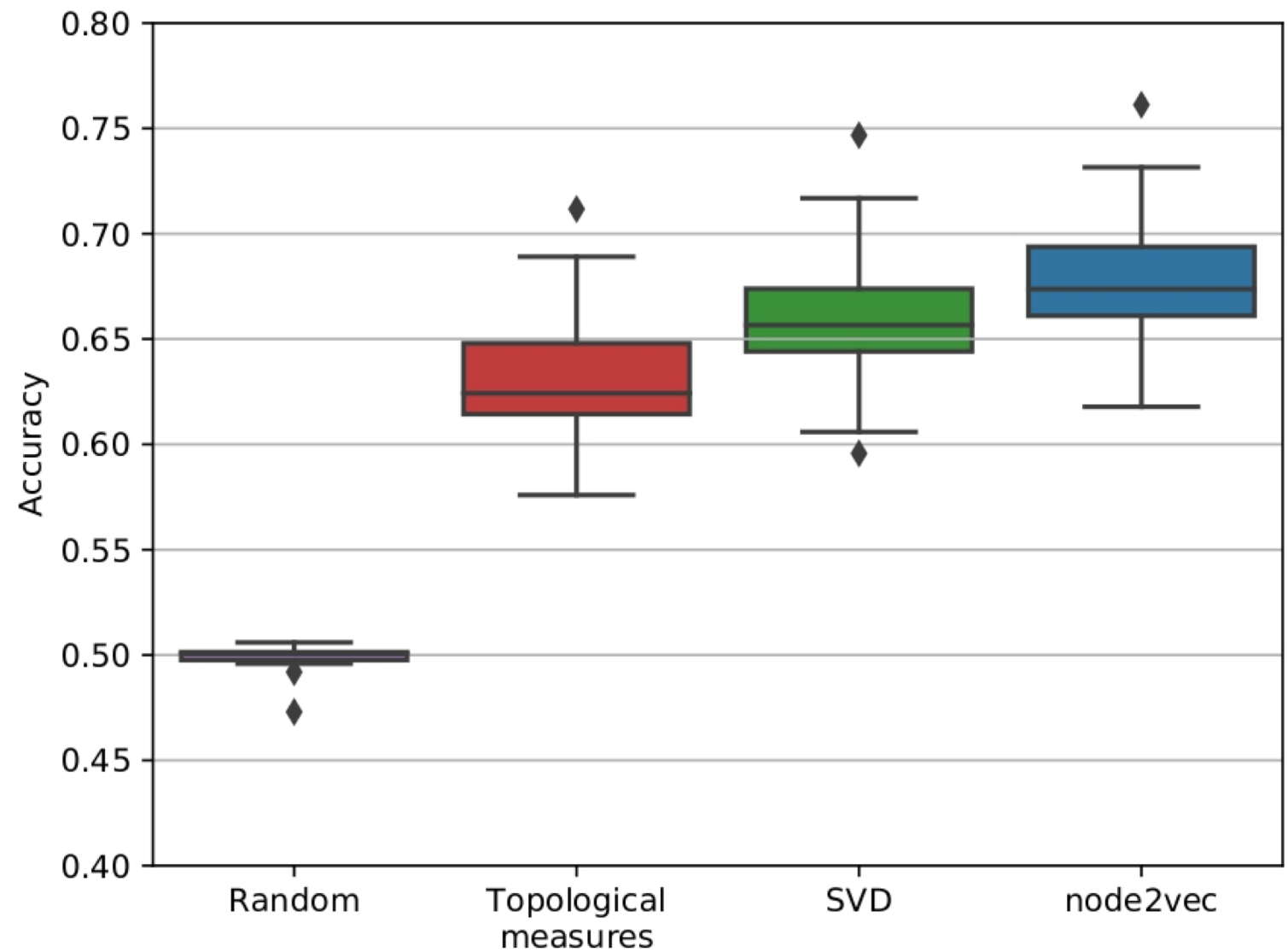
Results on 22 chromosomes

Gene expression data

TCGA Prostate cancer

Hi-C data

Prostate cancer cell-line
(22Rv1)



Future steps - Biological side

- **Interpretation of gene embeddings**
 - *What information do gene embeddings encode?*
 - *Is each dimension independent?*
- **Gene modules identification**
 - *Based on gene expression and chromatin structure information.*
 - *Are there tissue-specific clusters?*

Future steps - Computational side

- Use of **multi-layer** embedding:
 - *Combine Hi-C at different resolutions*
 - *Combine multiple Hi-C networks*
 - *Other networks (e.g. regulatory networks)*
- Use nodes and edge **attributes**:
Encode genomic annotations (e.g. mutations and histone modifications)
- Improve **explainability**:
Identify which nodes contributed the most to a certain prediction

