Semi-supervised learning

Known double knock out data provide hit and non-hit gene pair

Network Graphical Properties

- The degree: \( f_1 \)
- Clustering Coefficient: \( f_2 \)
- Stress Centrality: \( f_3 \)
- Shortest distance: \( f_4 \)
- Mutual neighbor #: \( f_5 \)
- Step depth length: \( f_6 \)

Samples: \( [S_1, S_2, \ldots, S_M] \)

Each sample: \( [f_{1g_1}^1, f_{1g_2}^2, \ldots, f_{Ng_2}^N] \)

Features of each gene or gene pair

Each gene pair \( g_1, g_2 \) as a sample:

Non-hit

Hit

Microarray data

Protein Complex

PPI data (BioGrid)