Research Progress
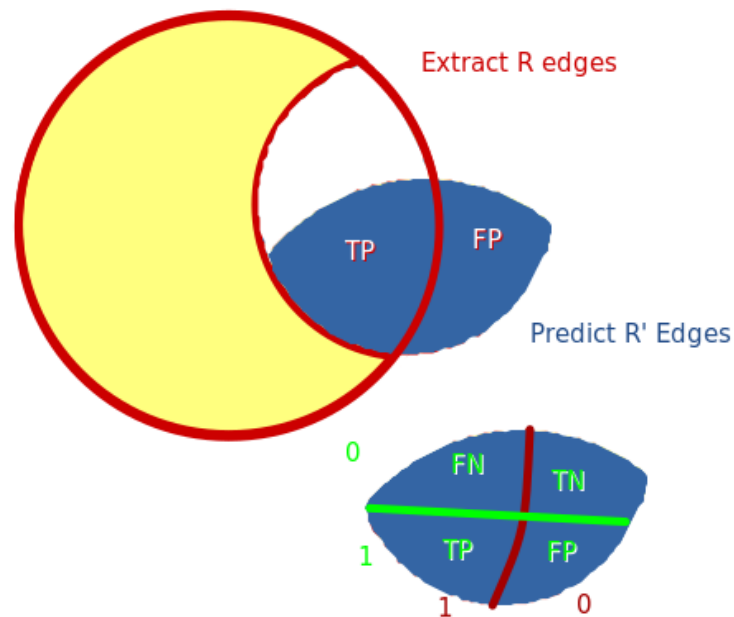
Nicolás López

July 3rd, 2020
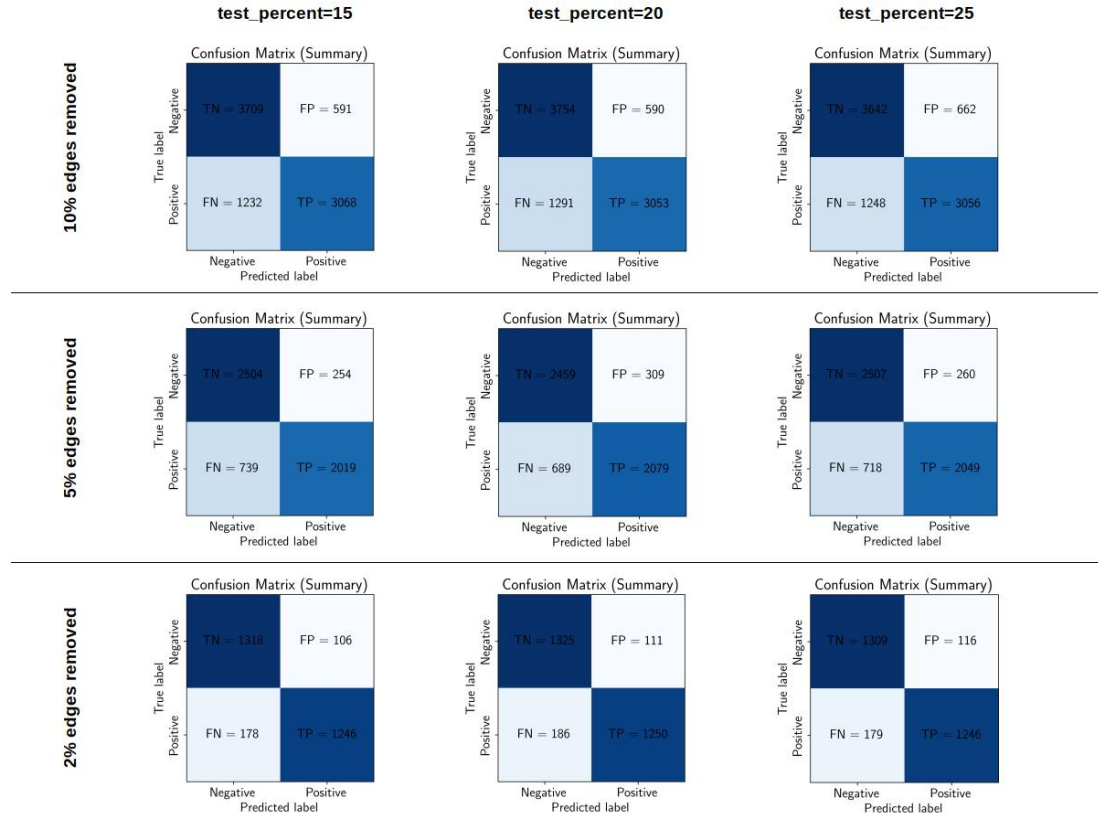
Protein-Protein Interactions

From the totam amount of edges of the network, the subset **R** is extracted and then predicted with the Length-3 degree-normalized Score (**L3**), obtaining the new set **R'**. From R', the set is balanced and after partitioning into 3 subsets (*training*, *test* and *validation*), the confusion matrices, feature importance plots and AUC-ROC plots are saved.

For the previous evaluation scheme, the following summary matrix is obtained.

test_percent=15   **test_percent=20**   **test_percent=25**

10% edges removed

Confusion Matrix (Summary)

	Predicted Negative	Predicted Positive
True Negative	TN = 3709	FP = 591
True Positive	FN = 1232	TP = 3068

Confusion Matrix (Summary)

	Predicted Negative	Predicted Positive
True Negative	TN = 3754	FP = 590
True Positive	FN = 1291	TP = 3053

Confusion Matrix (Summary)

	Predicted Negative	Predicted Positive
True Negative	TN = 3642	FP = 662
True Positive	FN = 1248	TP = 3056

5% edges removed

Confusion Matrix (Summary)

	Predicted Negative	Predicted Positive
True Negative	TN = 2504	FP = 254
True Positive	FN = 739	TP = 2019

Confusion Matrix (Summary)

	Predicted Negative	Predicted Positive
True Negative	TN = 2459	FP = 309
True Positive	FN = 689	TP = 2079

Confusion Matrix (Summary)

	Predicted Negative	Predicted Positive
True Negative	TN = 2507	FP = 260
True Positive	FN = 718	TP = 2049

2% edges removed

Confusion Matrix (Summary)

	Predicted Negative	Predicted Positive
True Negative	TN = 1318	FP = 106
True Positive	FN = 178	TP = 1246

Confusion Matrix (Summary)

	Predicted Negative	Predicted Positive
True Negative	TN = 1325	FP = 111
True Positive	FN = 186	TP = 1250

Confusion Matrix (Summary)

	Predicted Negative	Predicted Positive
True Negative	TN = 1309	FP = 116
True Positive	FN = 179	TP = 1246

After performing these steps, the original partition of the subset $\mathbf{R}$ was then evaluated. Results are shown below

L3 Confusion Matrix (Summary) | L3 Confusion Matrix (Summary) | L3 Confusion Matrix (Summary)

TN = 0 | FP = 42777 | FN = 0 | TP = 120543
TN = 0 | FP = 86209 | FN = 0 | TP = 240431
TN = 0 | FP = 129095 | FN = 0 | TP = 360865

TN = 0 | FP = 27488 | FN = 0 | TP = 54172
TN = 0 | FP = 55317 | FN = 0 | TP = 108003
TN = 0 | FP = 82955 | FN = 0 | TP = 162025

TN = 0 | FP = 14184 | FN = 0 | TP = 18486
TN = 0 | FP = 28341 | FN = 0 | TP = 36999
TN = 0 | FP = 42492 | FN = 0 | TP = 55518