Supplementary file: Co-Training based prediction of multi-label Protein-Protein Interactions

January 12, 2024

Table S1: Accuracy comparison of models on SHS27K, SYS30K and SYS60K

Dataset	Partition scheme	Method				
		CoMPPI	AFTGAN	GNN-PPI (KEAP)	PIPR (optimized)	PIPR
SHS27K	Random	0.9359	0.9112	0.9121	0.8064	0.7090
	BFS	0.9021	0.8678	0.8700	0.8308	0.6966
	DFS	0.9243	0.8965	0.8961	0.8723	0.7087
SYS30K	Random	0.7988	0.7679	0.7508	0.6583	0.6722
	BFS	0.7904	0.7546	0.7546	0.7008	0.6354
	DFS	0.8585	0.8067	0.7943	0.7626	0.6728
SYS60K	Random	0.8268	0.8135	0.7484	0.7074	0.6923
	BFS	0.8217	0.7566	0.7524	0.7482	0.6347
	DFS	0.8813	0.8476	0.8358	0.7994	0.6671

Table S2: AU-PRC comparison of models on SHS27K, SYS30K and SYS60K

Dataset	Partition scheme	Method				
		CoMPPI	AFTGAN	GNN-PPI (KEAP)	PIPR (optimized)	PIPR
SHS27K	Random	0.9611	0.9382	0.9396	0.7715	0.6647
	BFS	0.9429	0.9196	0.9210	0.8781	0.8321
	DFS	0.958	0.9393	0.9400	0.9107	0.8545
SYS30K	Random	0.7381	0.7058	0.6321	0.4587	0.4792
	BFS	0.8391	0.814	0.814	0.768	0.75
	DFS	0.911	0.8562	0.8441	0.7963	0.7644
SYS60K	Random	0.7653	0.6735	0.623	0.5126	0.484
	BFS	0.8867	0.8396	0.8318	0.8195	0.7802
	DFS	0.9399	0.9158	0.9091	0.8744	0.8425

Table S3: AU-ROC comparison of models on SHS27K, SYS30K and SYS60K

Dataset	Partition scheme	Method				
		CoMPPI	AFTGAN	GNN-PPI (KEAP)	PIPR (optimized)	PIPR
SHS27K	Random	0.9802	0.9643	0.9655	0.8795	0.8382
	BFS	0.9648	0.9417	0.9440	0.9112	0.8919
	DFS	0.9762	0.9600	0.9595	0.9414	0.9169
SYS30K	Random	0.8762	0.8462	0.8210	0.6349	0.6704
	BFS	0.8778	0.8392	0.8392	0.7997	0.7835
	DFS	0.9390	0.8913	0.8767	0.8406	0.8266
SYS60K	Random	0.8982	0.8652	0.8231	0.7283	0.7128
	BFS	0.9099	0.8578	0.8488	0.8359	0.8125
	DFS	0.9572	0.9322	0.9246	0.8947	0.8891