

Supplementary Material

XG-PUL: A Network-Based Framework for Disease Gene Prioritization Using Graph Representation Learning and Positive-Unlabeled Learning

Mahdiyeh Ayouman¹, Zahra Narimani¹

¹Department of Computer Science, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran

Contents

1	Section 1: Feature Engineering and Scenario Ablation Analysis	2
1.1	Ablation Study 1: Feature Space Hybridization	2
1.2	Ablation Study 2: Dimensionality Optimization and Computational Efficiency	2
2	Section 2: Comprehensive Performance Evaluation Tables	4
2.1	Evaluation Metrics and Methodological Summary	4
2.2	Aggregated Performance Profiles Across Disease Modules	4
3	Section 3: Top-K Threshold Comparative Analysis	8
3.1	Methodological Context and Robustness Evaluation	8
3.2	Sequential Threshold Performance Plots	9
4	Section 4: Continuous F_1-Score Trajectory Analysis across Top-K Thresholds	14
4.1	Methodological Context and Cross-Disease Trends	14
4.2	Disease-Specific F_1 -Score Comparison Curves	14
5	Section 5: Enrichment Analysis	19

1 Section 1: Feature Engineering and Scenario Ablation Analysis

1.1 Ablation Study 1: Feature Space Hybridization

A crucial phase in the development of the XG-PUL framework involved validating the selection of both the network embedding configuration and the final feature representation. To justify the embedding strategy, we first evaluated four Node2Vec hyperparameter configurations, including baseline, high-dimensional, deep-exploration, and ultra-high-dimensional settings. The corresponding hyperparameters, including embedding dimension, walk strategy (p and q), walk length, and sampling density, are summarized in Table 2.

The embedding configurations were assessed using an internal validation setting, in which held-out known disease genes (excluded from the seed set) were used for evaluation rather than a conventional train/test split. Representative results for seven complex diseases are presented in Table 3. Across these diseases, the HighDim configuration consistently achieved the strongest or near-best predictive performance in terms of both AUPR and F1@500. Based on these results, HighDim was selected as the default embedding configuration for all subsequent experiments.

After selecting the optimal embedding configuration, we evaluated three alternative feature representations:

- **Embedding Only (Top-150):** Node2Vec embeddings only after supervised feature selection.
- **Embedding + Topological Features (Top-150):** Node2Vec embeddings integrated with graph-derived topological features followed by supervised feature selection.
- **Embedding + Topological Features (All):** The complete feature set without feature selection.

Representative comparisons for these seven scenarios are presented in Table 1. Across all evaluated diseases, integrating graph-derived topological features consistently improved predictive performance compared with using embeddings alone. Although retaining all features produced a slight performance gain for prostate cancer, the improvement was marginal relative to the increased computational cost. These results justify the adoption of the combined feature representation with the top 150 selected features in the final XG-PUL framework.

1.2 Ablation Study 2: Dimensionality Optimization and Computational Efficiency

After confirming the benefit of integrating Node2Vec embeddings with graph-derived topological features, we next investigated the optimal feature dimensionality. Including all available topological descriptors introduces substantial redundancy and increases computational complexity, particularly for ensemble-based positive unlabeled learning.

Our evaluation demonstrated that the feature representation consisting of the top 150 selected features achieved performance comparable to the complete feature set across all representative diseases. Although the complete feature set yielded a slight improvement for prostate cancer, this gain was marginal relative to the increased computational cost. Consequently, selecting the top 150 features provides a better balance between predictive performance and computational efficiency and was therefore adopted throughout the proposed framework.

Table 1: Performance comparison of different feature representations for disease gene prioritization across seven complex diseases.

Disease	Scenario	AUPR	F1@500
Schizophrenia	Embed_only_top150	0.1767	0.1445
	Embed+Topological_Features_top150	0.1712	0.2069
	Embed+Topological_Features_all	0.1836	0.2241
Liver Cirrhosis	Embed_only_top150	0.0399	0.1086
	Embed+Topological_Features_top150	0.0320	0.0848
	Embed+Topological_Features_all	0.0410	0.0848
Colorectal Carcinoma	Embed_only_top150	0.2948	0.1828
	Embed+Topological_Features_top150	0.3063	0.1841
	Embed+Topological_Features_all	0.3139	0.1867
Intellectual Disability	Embed_only_top150	0.1194	0.1671
	Embed+Topological_Features_top150	0.3063	0.1841
	Embed+Topological_Features_all	0.1142	0.1718
Drug-Induced Liver Disease	Embed_only_top150	0.0151	0.0572
	Embed+Topological_Features_top150	0.0168	0.0604
	Embed+Topological_Features_all	0.0186	0.0731
Depressive Disorder	Embed_only_top150	0.0950	0.1751
	Embed+Topological_Features_top150	0.1007	0.1696
	Embed+Topological_Features_all	0.1116	0.1895
Chronic Alcoholic Intoxication	Embed_only_top150	0.0135	0.0692
	Embed+Topological_Features_top150	0.0142	0.0692
	Embed+Topological_Features_all	0.0151	0.0589

Table 2: Hyperparameter settings for Node2Vec embedding configurations used in the study.

Configuration	Dim.	Walk Len.	Num. Walks	Window	Min. Count	p	q
Baseline	128	80	10	10	3	1.0	1.0
High-Dim	256	80	10	10	3	1.0	1.0
Deep-Explore	128	80	20	10	3	0.5	2.0
Ultra-High-Dim	500	80	20	10	3	1.0	1.0

2 Section 2: Comprehensive Performance Evaluation Tables

2.1 Evaluation Metrics and Methodological Summary

To provide a rigorous numerical foundation for the evaluation curves presented in the subsequent sections, this section summarizes the empirical performance metrics across the investigated complex diseases. To prevent repetitive and oversized tabular layouts across all independent K thresholds, we report the macro-averaged performance metrics (Mean $\pm$ Standard Deviation) calculated across the entire progressive threshold spectrum, spanning $K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$.

The proposed XG-PUL framework is benchmarked against classical network propagation algorithms (DIAMOnD and RWR), community-detection baselines (MCL), and explainable graph neural network variations (XGDAG - GNNExplainer and XGDAG - GraphSVX). The evaluation tracks three standard metrics: Precision, Recall, and F_1 -score.

2.2 Aggregated Performance Profiles Across Disease Modules

Table 3: Performance comparison of node2vec embedding configurations across seven complex diseases.

Disease	Embedding	AUPR	F1@500
Chronic Alcoholic Intoxication	Baseline	0.0154	0.0769
	DeepExplore	0.0207	0.0846
	HighDim	0.0136	0.0538
	UltraHighDim	0.0152	0.0615
Colorectal Carcinoma	Baseline	0.2383	0.1732
	DeepExplore	0.2539	0.1640
	HighDim	0.2956	0.1820
	UltraHighDim	0.2608	0.1697
Depressive Disorder	Baseline	0.0534	0.1208
	DeepExplore	0.0824	0.1519
	HighDim	0.0965	0.1652
	UltraHighDim	0.0816	0.1475
Liver Cirrhosis	Baseline	0.0132	0.0543
	DeepExplore	0.0333	0.0849
	HighDim	0.0389	0.0985
	UltraHighDim	0.0262	0.0713
Schizophrenia	Baseline	0.1241	0.1780
	DeepExplore	0.1552	0.2015
	HighDim	0.1697	0.2042
	UltraHighDim	0.1541	0.1952
Drug-Induced Liver Disease	Baseline	0.0068	0.0413
	DeepExplore	0.0126	0.0509
	HighDim	0.0125	0.0541
	UltraHighDim	0.0203	0.0890
Intellectual Disability	Baseline	0.0807	0.1443
	DeepExplore	0.1028	0.1519
	HighDim	0.1180	0.1652
	UltraHighDim	0.1112	0.1652

Table 4: Macro-averaged performance evaluation across the complete K threshold spectrum ($K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$) for **Malignant Neoplasm of Breast (C0006142)**

Method	Mean Precision	Mean Recall	Mean F_1 -score
XG-PUL	0.908 ± 0.110	0.168 ± 0.151	0.248 ± 0.198
XGDAG - GraphSVX	0.626 ± 0.108	0.111 ± 0.101	0.164 ± 0.131
XGDAG - GNNExplainer	0.625 ± 0.108	0.111 ± 0.100	0.164 ± 0.130
DIAMOnD	0.621 ± 0.130	0.109 ± 0.102	0.161 ± 0.132
RWR	0.510 ± 0.099	0.095 ± 0.089	0.139 ± 0.116
MCL	0.434 ± 0.037	0.085 ± 0.079	0.125 ± 0.103

Table 5: Macro-averaged performance evaluation across the complete K threshold spectrum ($K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$) for **Schizophrenia (0036341)**.

Method	Mean Precision	Mean Recall	Mean F_1 -score
XG-PUL	0.493 $\pm$ 0.236	0.190 $\pm$ 0.146	0.204 $\pm$ 0.112
XGDAG - GraphSVX	0.220 $\pm$ 0.058	0.106 $\pm$ 0.092	0.110 $\pm$ 0.070
XGDAG - GNNExplainer	0.220 $\pm$ 0.058	0.107 $\pm$ 0.092	0.110 $\pm$ 0.070
DIAMOnD	0.213 $\pm$ 0.083	0.091 $\pm$ 0.074	0.095 $\pm$ 0.056
RWR	0.148 $\pm$ 0.027	0.084 $\pm$ 0.081	0.083 $\pm$ 0.062
MCL	0.119 $\pm$ 0.109	0.023 $\pm$ 0.010	0.029 $\pm$ 0.011

Table 6: Macro-averaged performance evaluation across the complete K threshold spectrum ($K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$) for **Liver Cirrhosis (C0023893)**.

Method	Mean Precision	Mean Recall	Mean F_1 -score
XG-PUL	0.080 $\pm$ 0.081	0.361 $\pm$ 0.217	0.082 $\pm$ 0.032
XGDAG - GraphSVX	0.020 $\pm$ 0.010	0.164 $\pm$ 0.132	0.026 $\pm$ 0.007
XGDAG - GNNExplainer	0.020 $\pm$ 0.010	0.157 $\pm$ 0.131	0.025 $\pm$ 0.005
DIAMOnD	0.001 $\pm$ 0.001	0.012 $\pm$ 0.012	0.001 $\pm$ 0.002
RWR	0.014 $\pm$ 0.007	0.137 $\pm$ 0.112	0.021 $\pm$ 0.009
MCL	0.008 $\pm$ 0.008	0.062 $\pm$ 0.044	0.011 $\pm$ 0.009

Table 7: Macro-averaged performance evaluation across the complete K threshold spectrum ($K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$) for **Colorectal Carcinoma (C0009402)**.

Method	Mean Precision	Mean Recall	Mean F_1 -score
XG-PUL	0.799 $\pm$ 0.164	0.170 $\pm$ 0.147	0.239 $\pm$ 0.178
XGDAG - GraphSVX	0.551 $\pm$ 0.112	0.118 $\pm$ 0.105	0.165 $\pm$ 0.125
XGDAG - GNNExplainer	0.552 $\pm$ 0.112	0.118 $\pm$ 0.105	0.165 $\pm$ 0.125
DIAMOnD	0.535 $\pm$ 0.168	0.104 $\pm$ 0.087	0.147 $\pm$ 0.103
RWR	0.411 $\pm$ 0.039	0.100 $\pm$ 0.094	0.138 $\pm$ 0.113
MCL	0.358 $\pm$ 0.039	0.086 $\pm$ 0.079	0.120 $\pm$ 0.095

Table 8: Macro-averaged performance evaluation across the complete K threshold spectrum ($K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$) for **Malignant Neoplasm of Prostate (C0376358)**.

Method	Mean Precision	Mean Recall	Mean F_1 -score
XG-PUL	0.780 $\pm$ 0.194	0.191 $\pm$ 0.160	0.255 $\pm$ 0.179
XGDAG - GraphSVX	0.523 $\pm$ 0.134	0.129 $\pm$ 0.111	0.172 $\pm$ 0.123
XGDAG - GNNExplainer	0.519 $\pm$ 0.128	0.129 $\pm$ 0.111	0.171 $\pm$ 0.123
DIAMOnD	0.537 $\pm$ 0.286	0.106 $\pm$ 0.089	0.141 $\pm$ 0.094
RWR	0.376 $\pm$ 0.085	0.103 $\pm$ 0.096	0.135 $\pm$ 0.107
MCL	0.253 $\pm$ 0.147	0.041 $\pm$ 0.024	0.058 $\pm$ 0.031

Table 9: Macro-averaged performance evaluation across the complete K threshold spectrum ($K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$) for **Intellectual Disability (C3714756)**.

Method	Mean Precision	Mean Recall	Mean F_1 -score
XG-PUL	0.380 $\pm$ 0.160	0.168 $\pm$ 0.135	0.172 $\pm$ 0.099
XGDAG - GraphSVX	0.204 $\pm$ 0.082	0.101 $\pm$ 0.090	0.100 $\pm$ 0.065
XGDAG - GNNExplainer	0.205 $\pm$ 0.083	0.100 $\pm$ 0.089	0.100 $\pm$ 0.065
DIAMOnD	0.185 $\pm$ 0.089	0.080 $\pm$ 0.065	0.082 $\pm$ 0.048
RWR	0.147 $\pm$ 0.028	0.087 $\pm$ 0.081	0.085 $\pm$ 0.061
MCL	0.138 $\pm$ 0.018	0.084 $\pm$ 0.077	0.083 $\pm$ 0.058

Table 10: Macro-averaged performance evaluation across the complete K threshold spectrum ($K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$) for **Bipolar Disorder (C0005586)**.

Method	Mean Precision	Mean Recall	Mean F_1 -score
XG-PUL	0.158 $\pm$ 0.079	0.174 $\pm$ 0.132	0.111 $\pm$ 0.052
XGDAG - GraphSVX	0.086 $\pm$ 0.051	0.093 $\pm$ 0.080	0.057 $\pm$ 0.023
XGDAG - GNNExplainer	0.086 $\pm$ 0.051	0.093 $\pm$ 0.080	0.057 $\pm$ 0.023
DIAMOnD	0.065 $\pm$ 0.045	0.052 $\pm$ 0.029	0.038 $\pm$ 0.016
RWR	0.053 $\pm$ 0.012	0.081 $\pm$ 0.073	0.047 $\pm$ 0.026
MCL	0.043 $\pm$ 0.045	0.019 $\pm$ 0.007	0.017 $\pm$ 0.009

Table 11: Macro-averaged performance evaluation across the complete K threshold spectrum ($K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$) for **Drug-Induced Liver Disease (C0860207)**.

Method	Mean Precision	Mean Recall	Mean F_1 -score
XG-PUL	0.032 $\pm$ 0.007	0.229 $\pm$ 0.197	0.044 $\pm$ 0.018
XGDAG - GraphSVX	0.034 $\pm$ 0.032	0.137 $\pm$ 0.107	0.033 $\pm$ 0.005
XGDAG - GNNExplainer	0.033 $\pm$ 0.032	0.137 $\pm$ 0.109	0.032 $\pm$ 0.005
DIAMOnD	0.004 $\pm$ 0.005	0.027 $\pm$ 0.018	0.007 $\pm$ 0.006
RWR	0.008 $\pm$ 0.006	0.085 $\pm$ 0.079	0.014 $\pm$ 0.010
MCL	0.007 $\pm$ 0.005	0.057 $\pm$ 0.043	0.012 $\pm$ 0.009

Table 12: Macro-averaged performance evaluation across the complete K threshold spectrum ($K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$) for **Depressive Disorder (C0011581)**.

Method	Mean Precision	Mean Recall	Mean F_1 -score
XG-PUL	0.334 $\pm$ 0.177	0.171 $\pm$ 0.137	0.159 $\pm$ 0.086
XGDAG - GraphSVX	0.216 $\pm$ 0.098	0.119 $\pm$ 0.098	0.109 $\pm$ 0.062
XGDAG - GNNExplainer	0.220 $\pm$ 0.106	0.119 $\pm$ 0.098	0.109 $\pm$ 0.062
DIAMOnD	0.185 $\pm$ 0.110	0.080 $\pm$ 0.054	0.079 $\pm$ 0.037
RWR	0.126 $\pm$ 0.029	0.091 $\pm$ 0.084	0.080 $\pm$ 0.056
MCL	0.059 $\pm$ 0.035	0.025 $\pm$ 0.015	0.026 $\pm$ 0.015

Table 13: Macro-averaged performance evaluation across the complete K threshold spectrum ($K \in [25, 50, 100, 200, 500, 750, 1000, 1500, 2000, 2500, 3000]$) for **Chronic Alcoholic Intoxication (C0001973)**.

Method	Mean Precision	Mean Recall	Mean F_1 -score
XG-PUL	0.061 ± 0.050	0.140 ± 0.117	0.052 ± 0.017
XGDAG - GraphSVX	0.042 ± 0.025	0.092 ± 0.073	0.036 ± 0.009
XGDAG - GNNExplainer	0.042 ± 0.025	0.093 ± 0.074	0.037 ± 0.010
DIAMOnD	0.022 ± 0.021	0.046 ± 0.026	0.023 ± 0.015
RWR	0.041 ± 0.030	0.091 ± 0.080	0.034 ± 0.008
MCL	0.012 ± 0.011	0.018 ± 0.010	0.009 ± 0.004

3 Section 3: Top- K Threshold Comparative Analysis

3.1 Methodological Context and Robustness Evaluation

To thoroughly investigate the stability and predictive robustness of the proposed XG-PUL framework, we analyze its performance against baseline and state-of-the-art models (including DIAMOnD, RWR, MCL, and XGDAG-based variations) across ten progressive top- K selection cutoffs ranging from $K = 25$ to $K = 3000$.

To ensure maximum visual clarity and preserve the original high-resolution rendering of the evaluation metrics across all disease benchmarks, each threshold configuration is presented independently in a full-width sequential layout below.

As demonstrated in the following sequential plots, XG-PUL consistently maintains an edge or matches the most competitive configurations under tight thresholds (e.g., $K = 25, 50, 100$), which are crucial for immediate wet-lab experimental validation. Furthermore, as the threshold expands to wider bounds ($K \geq 500$), our model yields highly robust recall and stable F_1 -score trajectories, confirming that integrating graph representation learning with Positive-Unlabeled learning successfully circumvents the negative label noise inherent in classical network propagation techniques.

3.2 Sequential Threshold Performance Plots

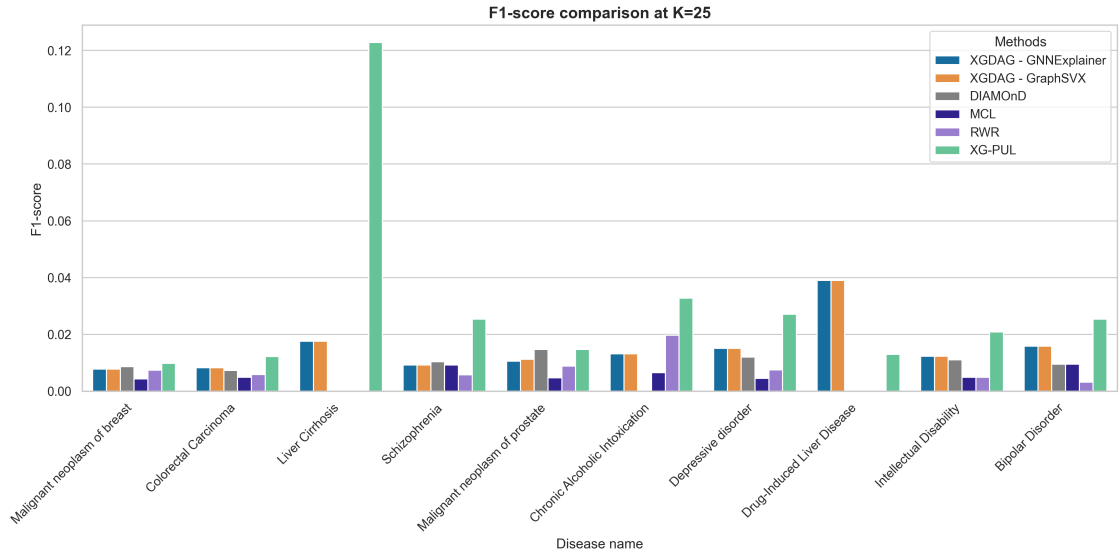


Figure 1: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 25$.

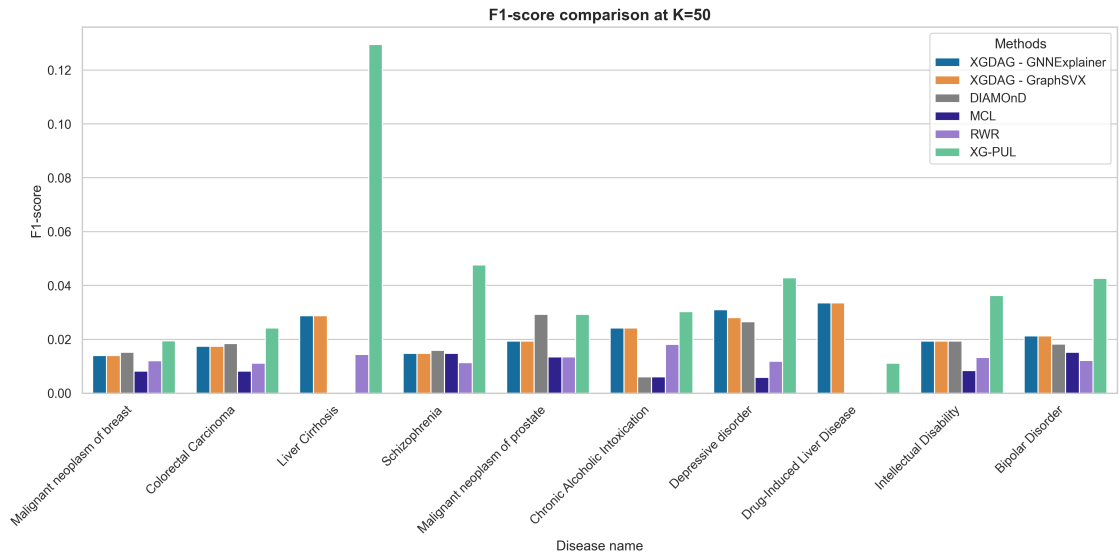


Figure 2: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 50$.

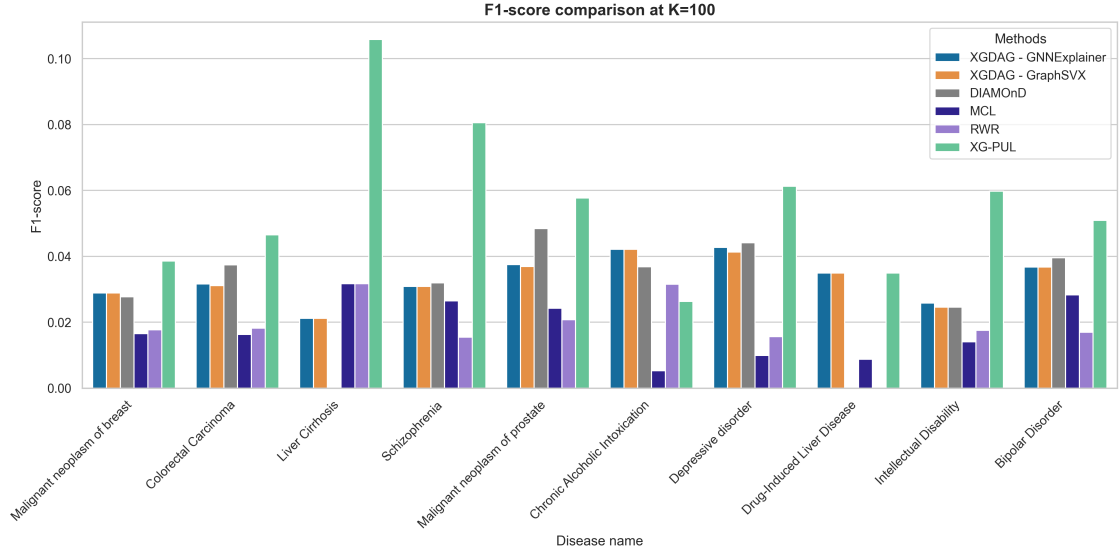


Figure 3: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 100$.

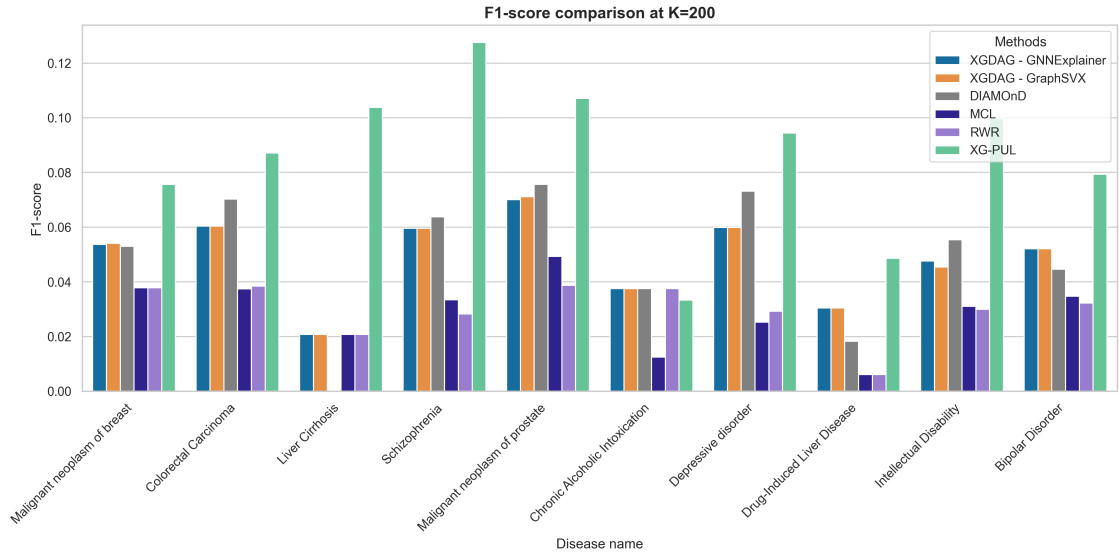


Figure 4: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 200$.

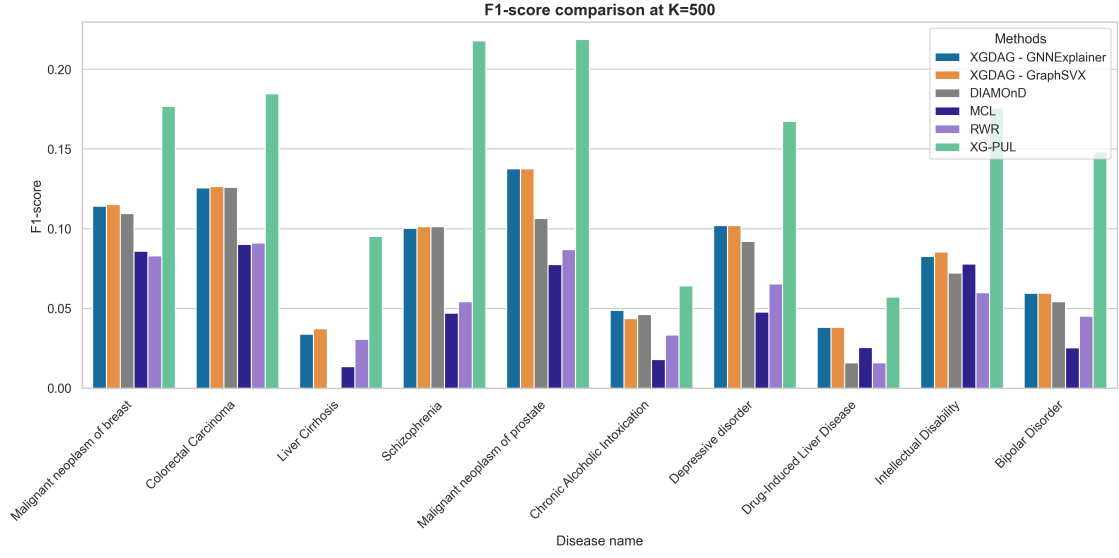


Figure 5: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 500$.

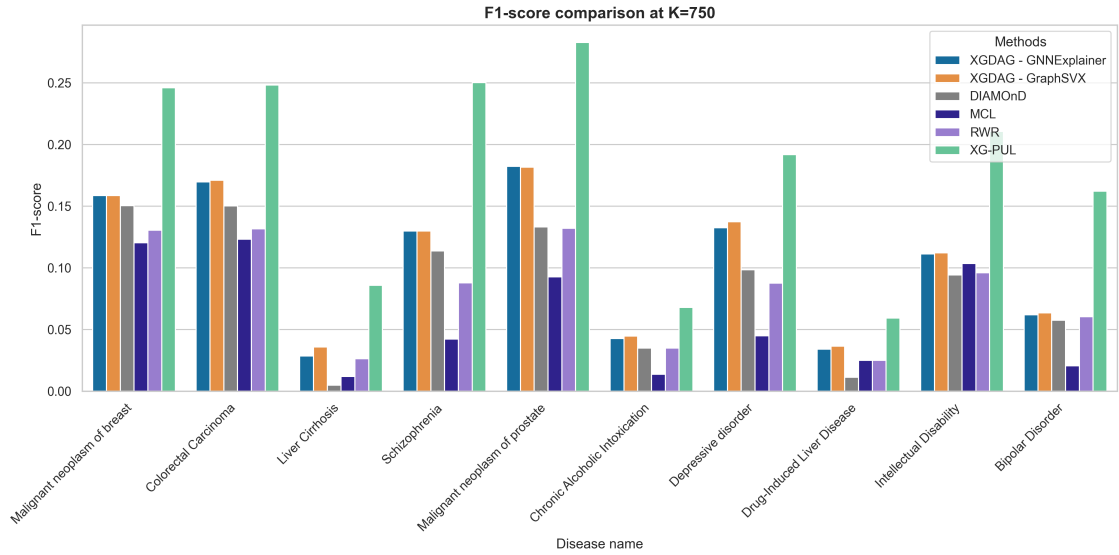


Figure 6: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 750$.

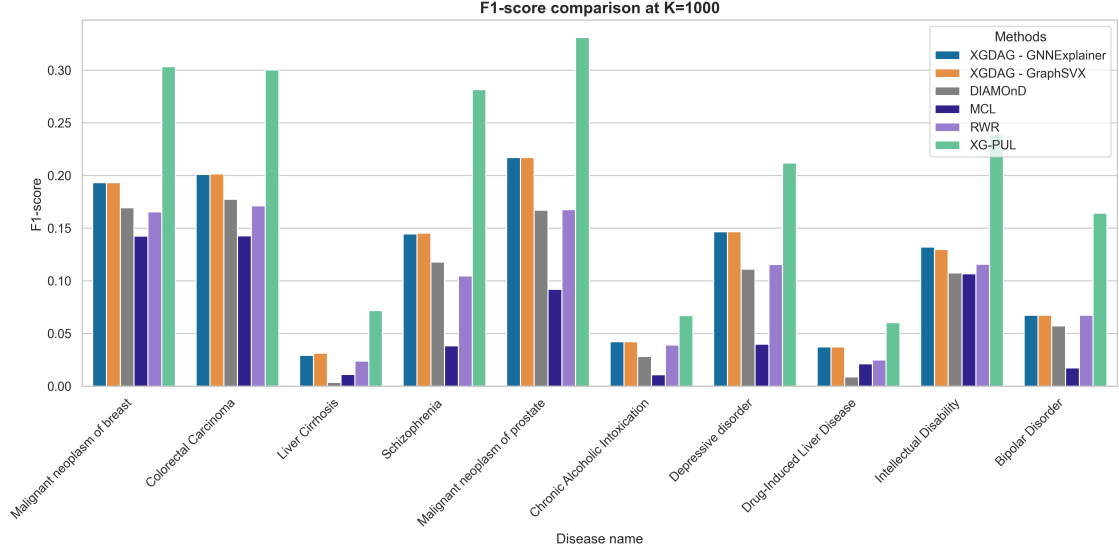


Figure 7: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 1000$.

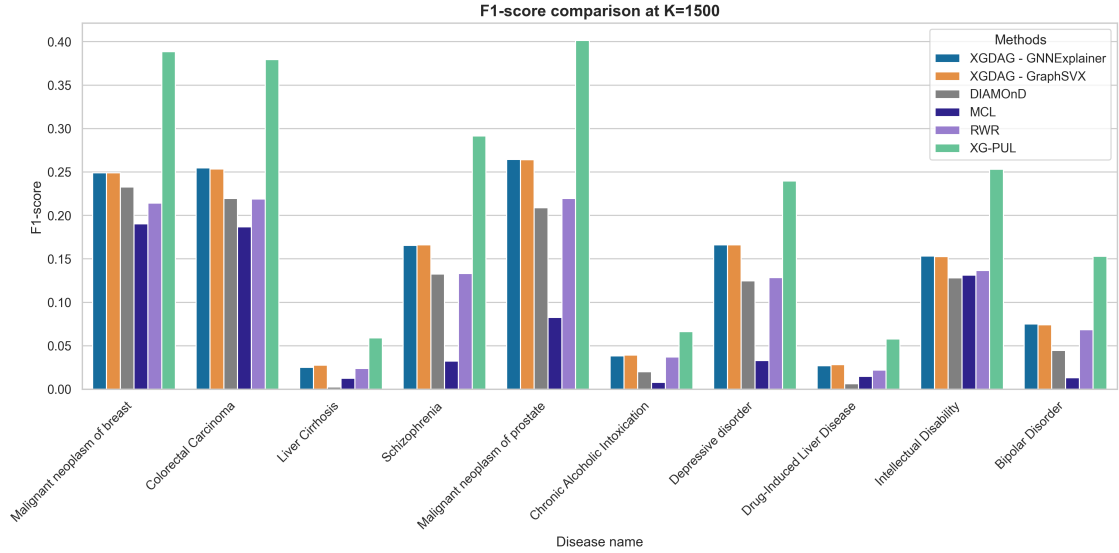


Figure 8: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 1500$.

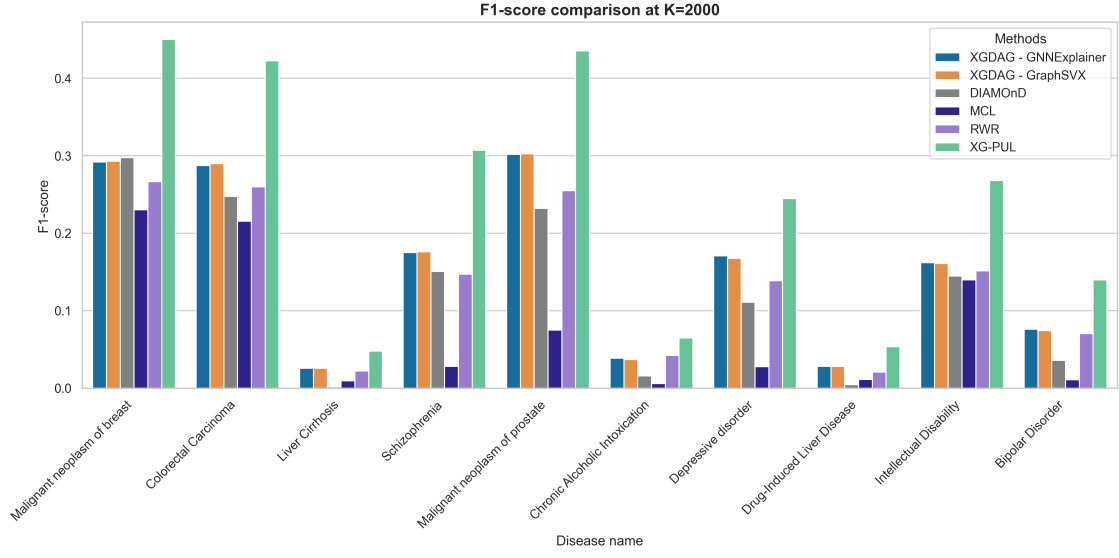


Figure 9: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 2000$.

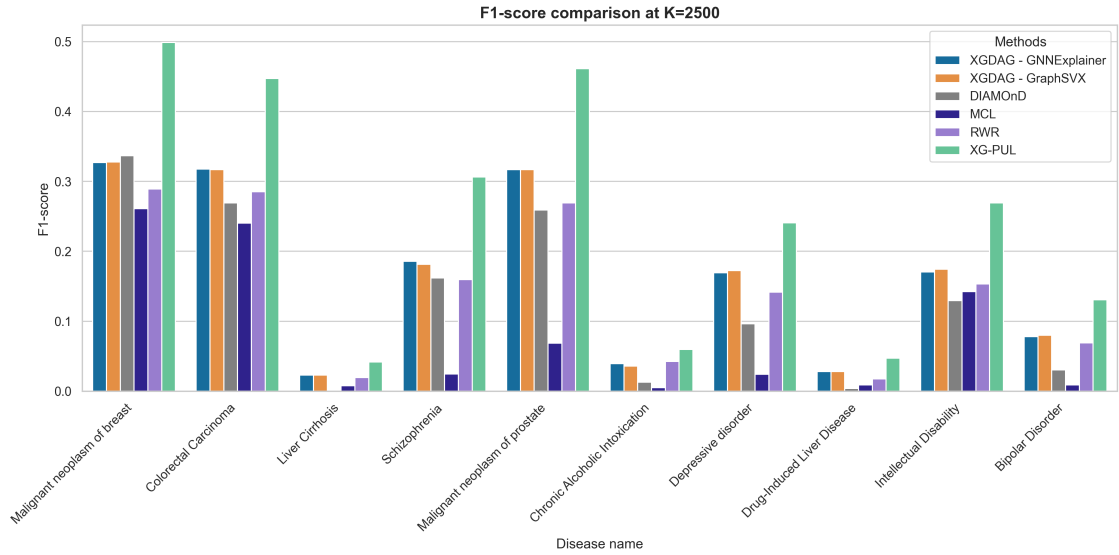


Figure 10: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 2500$.

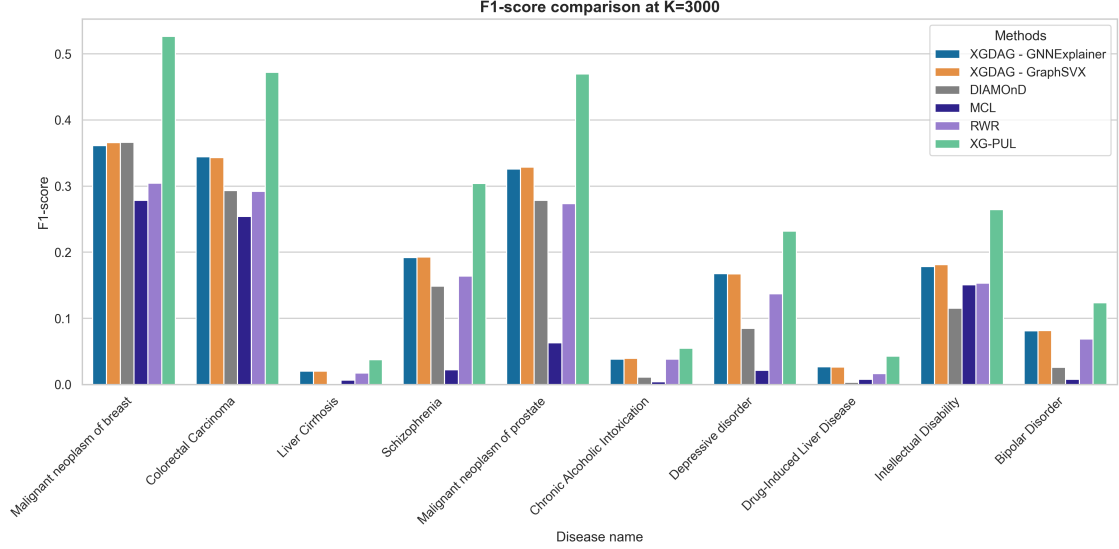


Figure 11: Comparative evaluation metrics across all ten complex diseases for top- K threshold $K = 3000$.

4 Section 4: Continuous F_1 -Score Trajectory Analysis across Top- K Thresholds

4.1 Methodological Context and Cross-Disease Trends

While the bar charts in Section S2 isolate individual evaluation thresholds, analyzing the continuous trajectory of the F_1 -score across a dense grid of K values provides deeper insights into model stability. In disease gene prioritization, a frequent challenge for network propagation algorithms (such as RWR and DIAMOnD) is the rapid decay of precision or early saturation of recall, which causes a sharp drop in the F_1 -score as the candidate list expands.

To demonstrate how the integration of embedding selections with bagging-based Positive-Unlabeled learning mitigates this limitation, we plot the F_1 -score as a continuous function of K (from $K = 25$ up to $K = 2500$).

Following the aggregate overview, Figures 12 through 20 disclose the independent, high-resolution F_1 -score comparison curves for each distinct disease in our benchmark dataset. These plots verify that the predictive superiority of XG-PUL is sustained across varying disease modules and topological contexts, rather than being skewed by a single highly-characterized disease.

4.2 Disease-Specific F_1 -Score Comparison Curves



Figure 12: Continuous F_1 -score trajectory comparison across varying top- K thresholds for Malignant Neoplasm of Breast.

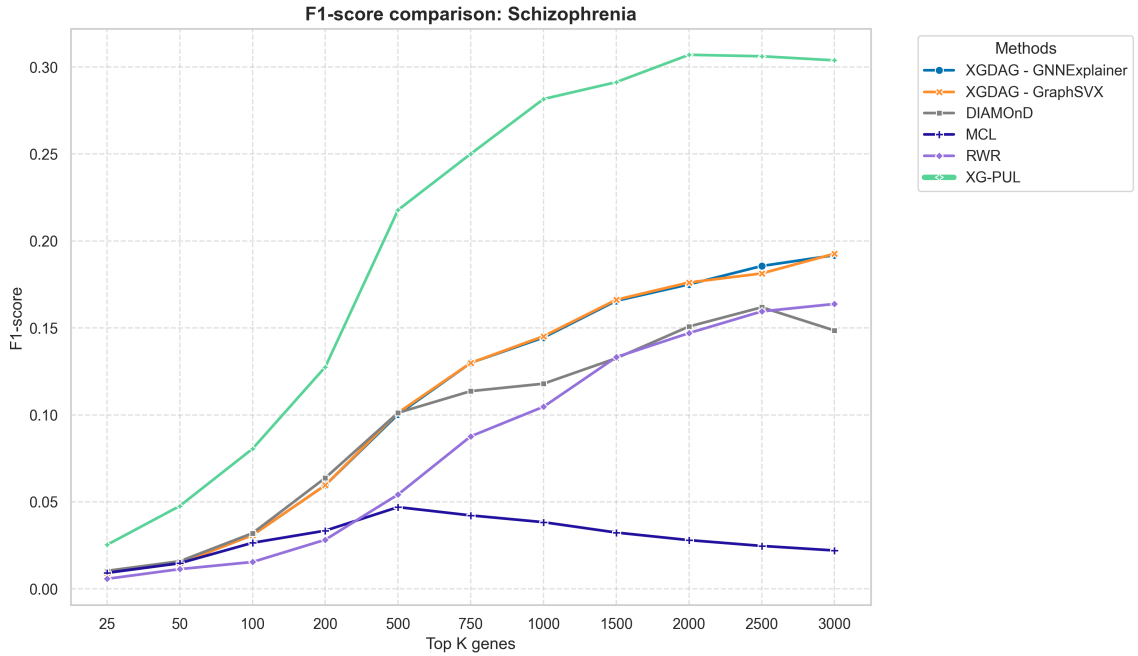


Figure 13: Continuous F_1 -score trajectory comparison across varying top- K thresholds for Schizophrenia.

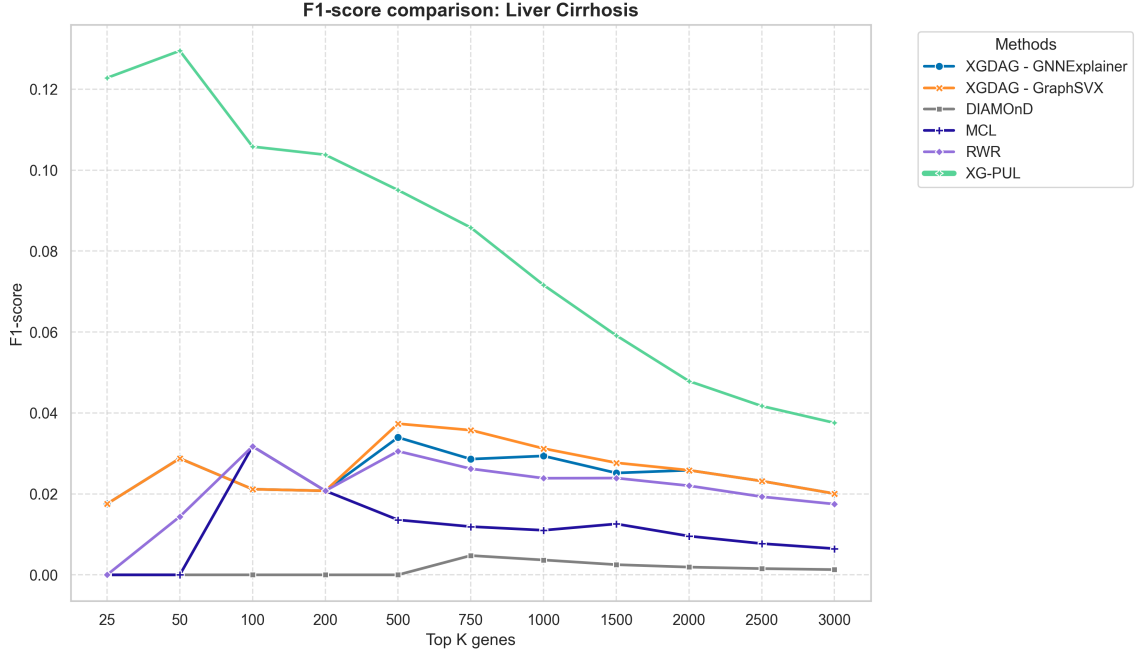


Figure 14: Continuous F_1 -score trajectory comparison across varying top- K thresholds for Liver Cirrhosis.

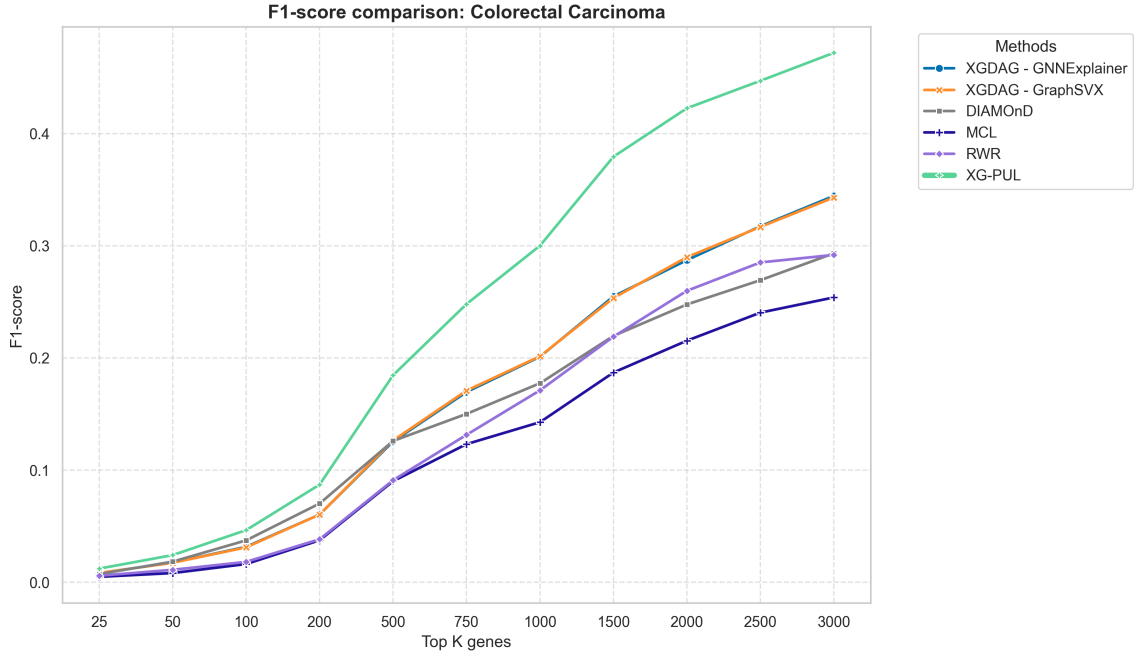


Figure 15: Continuous F_1 -score trajectory comparison across varying top- K thresholds for Colorectal Carcinoma.



Figure 16: Continuous F_1 -score trajectory comparison across varying top- K thresholds for Malignant Neoplasm of Prostate.

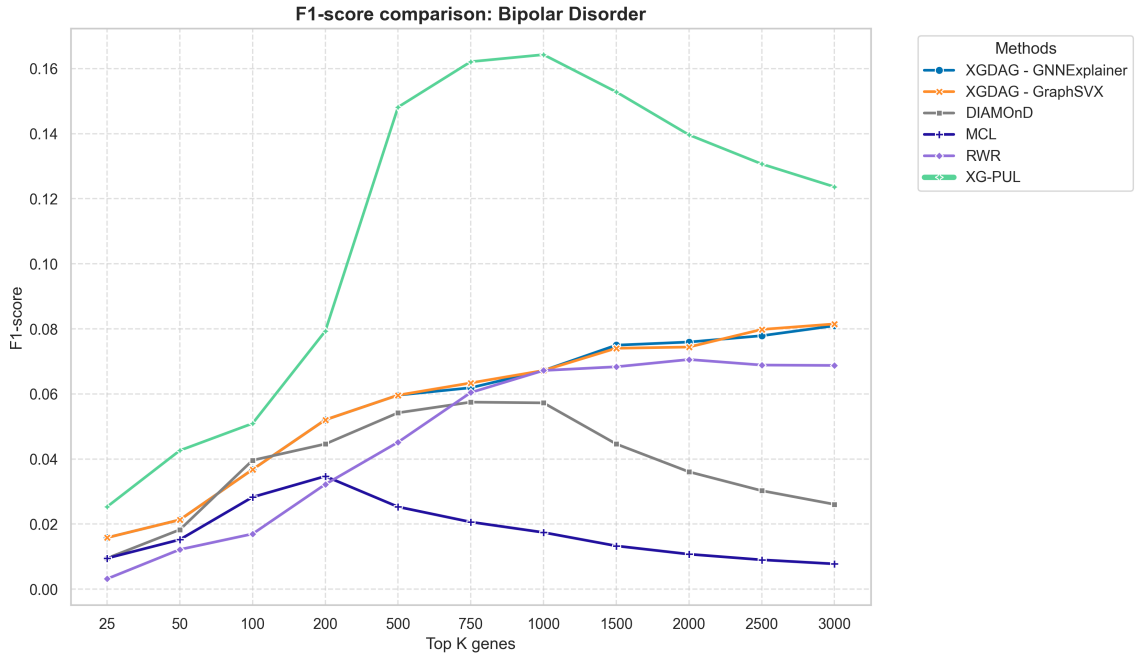


Figure 17: Continuous F_1 -score trajectory comparison across varying top- K thresholds for Bipolar Disorder.

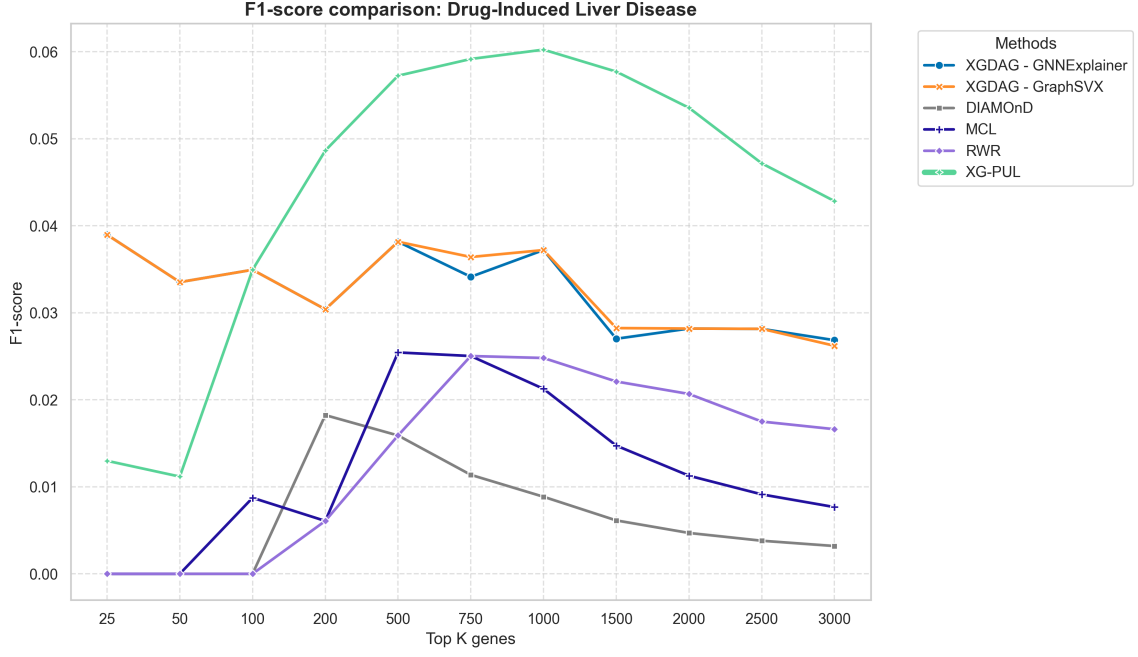


Figure 18: Continuous F_1 -score trajectory comparison across varying top- K thresholds for Drug-Induced Liver Disease.

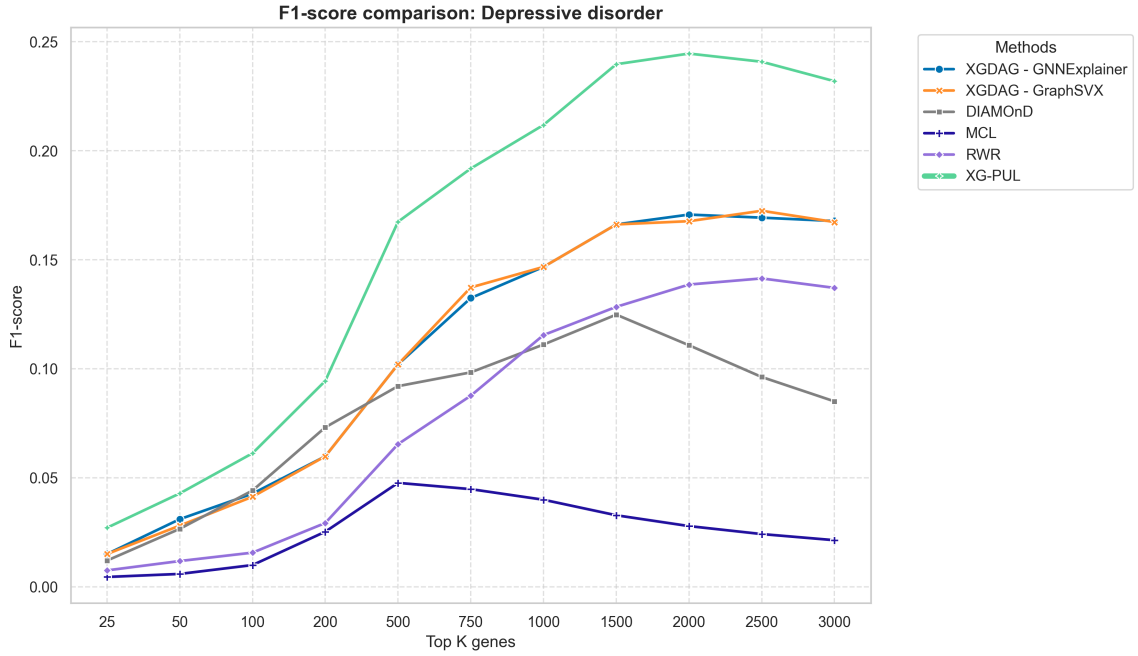


Figure 19: Continuous F_1 -score trajectory comparison across varying top- K thresholds for Depressive Disorder.

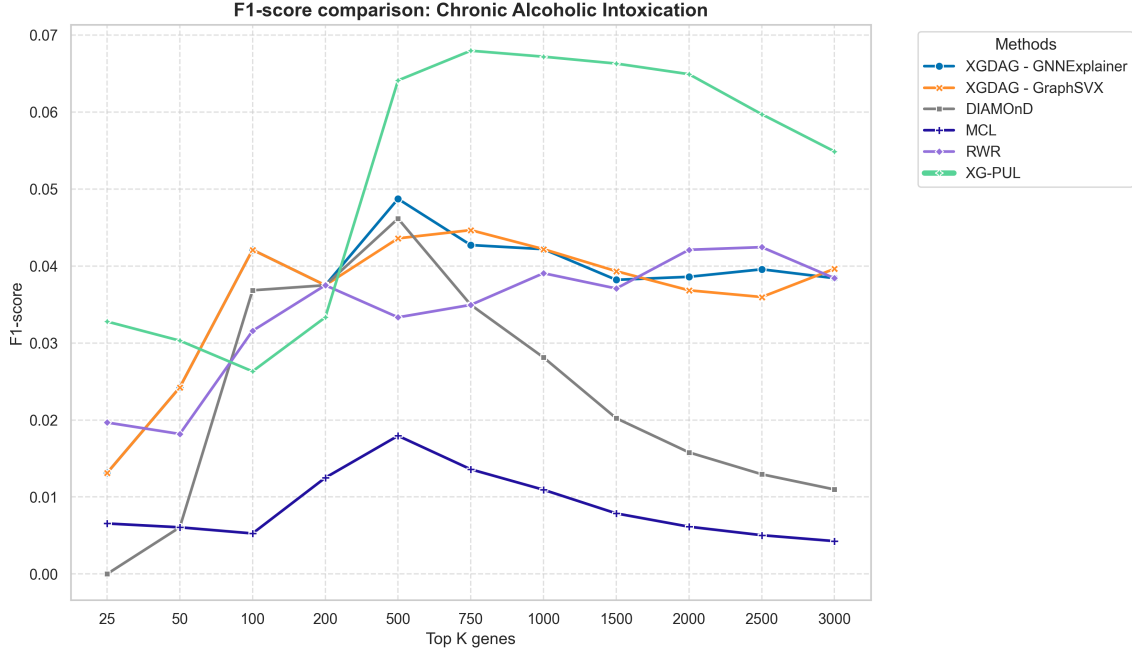


Figure 20: Continuous F_1 -score trajectory comparison across varying top- K thresholds for Chronic Alcoholic Intoxication.

5 Section 5: Enrichment Analysis

Table 11 summarizes the enrichment analysis results for the ten diseases investigated in this study. For each disease, the most statistically significant pathway, ontology, or related diseases identified by our proposed XG-PUL framework is reported. Model-prioritized genes were systematically cross-referenced with curated gene set databases using the Enrichr platform, including DisGeNET, GO Biological Process (2026), KEGG (2026), Reactome Pathways (2024), and WikiPathways (2024 Human). These cross-referenced results, aggregated in the corresponding tables (DisGeNET_table, GO_Biological_Process_2026_table, KEGG_2026_table, Reactome_Pathways_2024_table, and WikiPathways_2024_Human_table), highlight the most relevant pathophysiological modules underlying each phenotype. This comprehensive approach provides an objective biological validation of the candidate targets predicted by our approach.

Table 14: Enrichment analysis for the considered diseases.
We report the most enriched GO, pathway, or disease.

Disease	Enriched Pathway / GO / Disease	Details	Reference
Malignant neoplasm of breast (C0006142)	SUMOylation of SUMOylation Proteins (Reactome)	The top predicted genes are most highly enriched for "SUMOylation of SUMOylation Proteins", highlighting the prominent role of SUMO modification processes in the pathogenesis of breast cancer. SUMOylation regulates DNA repair, transcription, and protein stability, all of which are critically implicated in tumor development.	Seeler & Dejean, 2017; Hendriks & Vertegaal, 2016

Table 14: Enrichment analysis for the considered diseases.
We report the most enriched GO, pathway, or disease.

Disease		Enriched Pathway / GO / Disease	Details	Reference
Schizophrenia (C0036341)		Nervous System Development (Reactome)	The prioritized genes showed significant enrichment for <i>Nervous System Development</i> , a term that is highly relevant to schizophrenia and consistent with the established neurodevelopmental contribution to its pathophysiology, even though signal transduction- and GPCR-related pathways ranked higher by statistical significance.	Owen et al., 2016; Kahn et al., 2015
Liver cirrhosis (C0023893)		Integrin Cell Surface Interactions (Reactome)	Genes prioritized by the model showed highest enrichment for "Integrin Cell Surface Interactions". This emphasizes the central role of integrin-mediated signaling and extracellular matrix (ECM) reorganization in liver fibrogenesis and the pathogenesis of cirrhosis, as consistently reported in recent research.	Bataller & Brenner, 2005; Kisseleva & Brenner, 2008
Colorectal carcinoma (C0009402)		Regulation of TP53 Activity (Reactome)	Genes prioritized by the model show the highest enrichment in Regulation of TP53 Activity, reflecting TP53s pivotal role in cell cycle regulation, DNA repair, and apoptosis in colorectal cancer pathogenesis. Dysfunction in this pathway is a hallmark of tumorigenesis.	Fearon & Vogelstein, 1990; Liebl & Hofmann, 2021
Malignant neoplasm of prostate (C0376358)		Immune System (Reactome)	The top enrichment in <i>Immune System</i> pathways is consistent with the known contribution of immune and inflammatory processes to prostate cancer development and progression, supporting current insights into tumor immunobiology.	Drake et al., 2014; Sfanos & De Marzo, 2012
Intellectual disability (C3714756)		Gene Expression (Transcription) (Reactome)	The top enrichment in gene expression/transcription-related pathways is consistent with the established role of transcriptional and epigenetic dysregulation in the etiology of intellectual disability, supporting evidence that chromatin modification and gene regulation defects contribute to neurodevelopmental disorders.	Mossink et al., 2021; Ciptasari & van Bokhoven, 2020
Bipolar disorder (C0005586)		Neuronal System (Reactome)	Model-predicted genes are enriched for <i>Neuronal System</i> processes, consistent with the role of synaptic, neuronal, and neurochemical dysregulation in bipolar disorder pathophysiology.	Grande et al., 2016; Fornito et al., 2015

Table 14: Enrichment analysis for the considered diseases.
We report the most enriched GO, pathway, or disease.

Disease	Enriched Pathway / GO / Disease	Details	Reference
Drug-induced liver disease (C0860207)	Liver carcinoma (DisGeNET)	The prioritized gene set showed significant enrichment for <i>Liver carcinoma</i> , suggesting that the model captures hepatic disease-related gene modules and shared molecular mechanisms linking drug-induced liver injury and hepatocarcinogenesis.	Zimmerman, 2000; Stravitz & Lee, 2019
Depressive disorder (C0011581)	GPCR Downstream Signalling (Reactome)	The prioritized genes are significantly enriched for GPCR downstream signalling, consistent with the central role of GPCR-mediated neurotransmission, including GABAergic signaling, in depression pathophysiology. These pathways are implicated in mood regulation, synaptic plasticity, and treatment response.	Krystal et al., 2019; Luscher et al., 2011
Chronic alcoholic intoxication (C0001973)	Signaling by GPCR (Reactome)	The model-predicted genes were significantly enriched for <i>Signaling by GPCR</i> , consistent with the central role of GPCR- and neurotransmitter-mediated signaling in alcohol-related neurobiological and systemic damage. This is in line with studies highlighting the impact of alcohol on neuronal signaling pathways and reward-related circuitry.	Goaszewski et al., 2026; Spanagel, 2009

References

- Seeler, J. S., & Dejean, A. (2017). SUMO and the robustness of cancer. *Nature Reviews Cancer*, 17(3), 184–197. doi:10.1038/nrc.2016.143
- Hendriks, I. A., & Vertegaal, A. C. O. (2016). A comprehensive compilation of SUMO proteomics. *Nature Reviews Molecular Cell Biology*, 17(9), 581–595. doi:10.1038/nrm.2016.81
- Owen, M. J., Sawa, A., & Mortensen, P. B. (2016). Schizophrenia. *The Lancet*, 388(10039), 86–97. doi:10.1016/S0140-6736(15)01121-6
- Kahn, R. S., Sommer, I. E., Murray, R. M., Meyer-Lindenberg, A., Weinberger, D. R., Cannon, T. D., ... & Insel, T. R. (2015). Schizophrenia. *Nature Reviews Disease Primers*, 1, 15067. doi:10.1038/nrdp.2015.67
- Bataller, R., & Brenner, S. L. (2005). Liver fibrosis. *The Journal of Clinical Investigation*, 115(2), 209–218. doi:10.1172/JCI24282
- Kisseleva, T., & Brenner, D. A. (2008). Fibrogenesis of parenchymal organs. *The FEBS Journal*, 288(5), 1197–1212. doi:10.1513/pats.200711-168DR
- Fearon, E. R., & Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell*, 61(5), 759–767. doi:10.1016/0092-8674(90)90186-I
- Liebl, M. C., & Hofmann, T. G. (2021). The Role of p53 Signaling in Colorectal Cancer. *Cancers*, 13(9), 2125. doi:10.3390/cancers13092125
- Drake, C. G., Lipson, E. J., & Brahmer, J. R. (2014). Breathing new life into immunotherapy: review of melanoma, lung and kidney cancer. *Nature Reviews Clinical Oncology*, 11(1), 24–37. doi:10.1038/nrclinonc.2013.208
- Sfanos, K. S., & De Marzo, A. M. (2012). Prostate cancer and inflammation: the evidence. *Histopathology*, 60(1), 199–215. doi:10.1111/j.1365-2559.2011.04033.x
- Mossink, B., van Rhijn, J. R., van de, S., Schubert, D., & Schenck, A. (2021). The emerging role of chromatin remodelers in neurodevelopmental disorders: a developmental perspective. *Cellular and Molecular Life Sciences*, 78(6), 2517–2563. doi:10.1007/s00018-020-03714-5
- Ciptasari, U., & van Bokhoven, H. (2020). The phenomenal epigenome in neurodevelopmental disorders. *Human Molecular Genetics*, 29(R1), R42–R50. doi:10.1093/hmg/ddaa175
- Grande, I., Berk, M., Birmaher, B., & Vieta, E. (2016). Bipolar disorder. *The Lancet*, 387(10027), 1561–1572. doi:10.1016/S0140-6736(15)00241-X
- Fornito, A., Zalesky, A., & Breakspear, M. (2015). The connectomics of brain disorders. *Nature Reviews Neuroscience*, 16(3), 159–172. doi:10.1038/nrn3901
- Zimmerman, H. J. (2000). Drug-induced liver disease. *Clinical Liver Disease*, 2(2), 39–41. doi:10.1016/s1089-3261(05)70097-0
- Stravitz, R. T., & Lee, W. M. (2019). Acute liver failure. *The Lancet*, 394(10201), 869–881. doi:10.1016/S0140-6736(19)31894-X
- Krystal, J. H., Abdallah, C. G., Sanacora, G., Charney, D. S., Duman, R. S., & Gordon, J. (2020). Ketamine: a paradigm shift for depression research and treatment. *Neuron*, 101(5), 774–778. doi:10.1016/j.neuron.2019.02.005

Luscher, B., Shen, Q., & Sahir, N. (2011). The GABAergic deficit hypothesis of major depressive disorder. *Molecular Psychiatry*, 16(4), 383–406. doi:10.1038/mp.2010.120

Gołaszewski, P., Wawrzyniak, A., Kłosowicz, M., Burbelka, A., & Balawender, K. (2026). Neurotoxicity of Chronic Alcohol Exposure: Mechanistic Insights, Cellular Disruption, and Emerging Therapeutic Strategies. *International Journal of Molecular Sciences*, 27(1), 299. doi:10.3390/ijms27010299

Spanagel, R. (2009). Alcoholism: a systems approach from molecular physiology to addictive behavior. *Physiological Reviews*, 89(2), 649–705. doi:10.1152/physrev.00013.2008