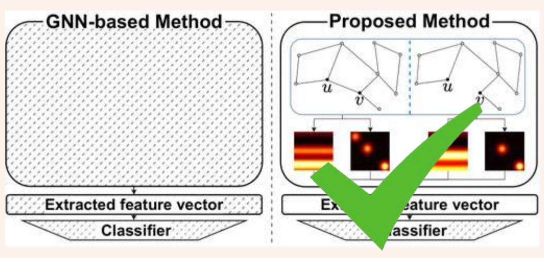
Interleaved summarization from **M-DocSum-7B:**

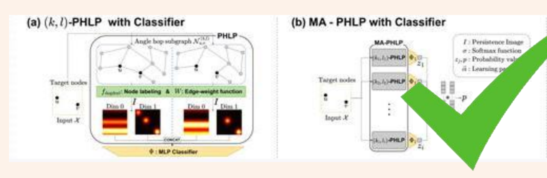
Research Background: Link prediction (LP) is a critical area in graph data analysis, focusing on inferring the connectivity between nodes. This task is essential for applications such as social network friend recommendations, knowledge graph completion, and drug-protein interaction prediction. While LP has been categorized into heuristic methods, embedding methods, and graph neural network (GNN)-based methods, GNN-based models have recently achieved significant improvements in capturing intricate relationships within graphs. However, the high performance of GNN-based models is challenging to interpret due to their complex neural network structures. Persistent homology (PH), a topological data analysis method, has been successfully applied to graph classification and node classification tasks but has seen limited use in LP. The TLC-GNN is a notable exception, integrating PH with GNNs to enhance LP performance, though it requires node attributes for its operation.

Com: 0.562 Acc: 0.5



Proposed Method: The proposed method, PHLP, leverages persistent homology to analyze the topological structure of graphs for LP. It employs angle hop subgraphs and a new node labeling technique called Degree DRNL to better distinguish graph information. PHLP calculates persistence images (PIs) for subgraphs with and without target links, transforming each target node into a vector comprising these PIs. The method uses a classifier to predict the existence of target links based on these vectors. Additionally, MA-PHLP is introduced to analyze data from multiple angles, enhancing the extraction of diverse topological information. The hybrid method integrates PHLP with existing subgraph methods, such as SEAL and WP, by concatenating PHLP vectors with the outputs of these models before classification.

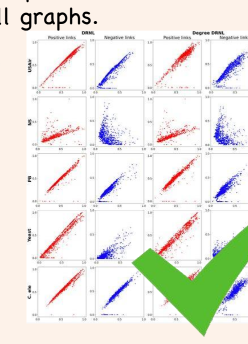
Com: 0.786 Acc: 1.0



Experimental Results: The performance of MA-PHLP was evaluated using the area under the curve (AUC) as the primary metric across eight benchmark datasets. MA-PHLP demonstrated competitive performance, achieving scores comparable to state-of-the-art models. Notably, it outperformed several baselines, including SEAL and WP, on most datasets. The Power dataset, which is known for its low performance in LP, showed significant improvement with MA-PHLP, achieving the highest AUC score among baseline models. The hybrid methods, which combined MA-PHLP with SEAL and WP, further enhanced performance, particularly on the Power dataset. Ablation studies revealed that Degree DRNL significantly improved performance, especially on the Power dataset, where it increased the AUC score by over 4 points. The analysis of subgraphs with node labeling showed that the number of nodes in positive links was smaller than in negative links, highlighting the importance of node selection in the analysis of small graphs.

Com: 0.429

Acc: 0.5



Conclusion: This study introduces PHLP, a novel method that applies persistent homology to analyze the topological structure of graphs for LP. By employing angle hop subgraphs and Degree DRNL, PHLP enhances the analysis of graph topology. Experimental results demonstrate that PHLP achieves competitive performance across benchmark datasets, including state-of-the-art performance on the Power dataset. The integration of PHLP with existing GNN-based methods further improves performance across all datasets. This approach provides explainable algorithms without relying on complex deep learning techniques, offering valuable insights into the factors that influence LP performance in graph data.

Com: 0.8 Acc: 1.0

Image Idx: None

Interleaved summarization from **Qwen2-VL-7B**

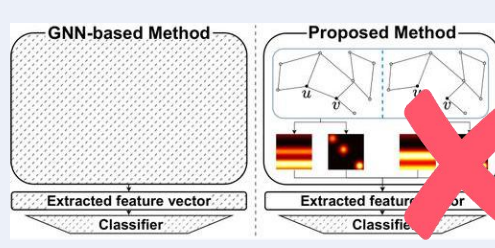
This paper proposes PHLP, an explainable method that applies PH to analyze the topological structure of graphs to overcome the limitations of GNN-based methods for LP. By employing the proposed methods, such as angle hop subgraphs and Degree DRNL, PHLP improves the analysis of the topological structure of graphs. The experimental results demonstrate that the proposed PHLP method achieves competitive performance across benchmark datasets, even SOTA performance, especially on the Power dataset. Additionally, when integrated with existing GNN-based methods, PHLP improves performance across all datasets. By analyzing the topological information of the given graphs, PHLP addresses the limitations of GNN-based methods and enhances overall performance. As demonstrated, PHLP provides explainable algorithms without relying on complex deep learning techniques, providing insight into the factors that significantly influence performance for the LP problem of graph data.

Com: 0.188 Acc: 0.75

Image Idx: None

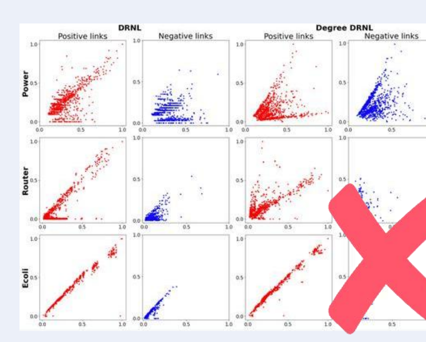
To address the link prediction problem, this paper proposes a novel method called PHLP, which employs persistent homology (PH) for link prediction. PHLP focuses on how the presence or absence of target links influences the overall topology. The PHLP utilizes the angle hop subgraph and new node labeling called Degree DRNL, distinguishing the information of graphs better than DRNL. Using only a classifier, PHLP performs similarly to state-of-the-art (SOTA) models on most benchmark datasets. Incorporating the outputs calculated using PHLP into the existing GNN-based SOTA models improves performance across all benchmark datasets. To the best of our knowledge, PHLP is the first method of applying PH to LP without GNNs. The proposed approach, employing PH while not relying on neural networks, enables the identification of crucial factors for improving performance.

Com: 0.429 Acc: 1.0



The authors compared the proposed method on the Power dataset. The advantages of the proposed method are ..., and the disadvantages are .... The experimental scheme/experimental setup is as follows: ... Experimental results and analysis: Compared with the ... model, the proposed model improved by ... on the ... dataset. This indicates ...

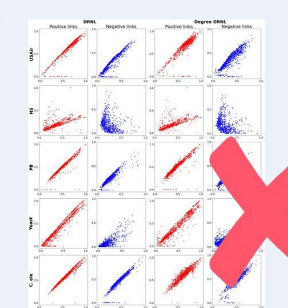
Com: 0.214 Acc: 0.0



This paper proposes PHLP, an explainable method that applies PH to analyze the topological structure of graphs to overcome the limitations of GNN-based methods for LP. By employing the proposed methods, such as angle hop subgraphs and Degree DRNL, PHLP improves the analysis of the topological structure of graphs. The experimental results demonstrate that the proposed PHLP method achieves competitive performance across benchmark datasets, even SOTA performance, especially on the Power dataset. Additionally, when integrated with existing GNN-based methods, PHLP improves performance across all datasets. By analyzing the topological information of the given graphs, PHLP addresses the limitations of GNN-based methods and enhances overall performance. As demonstrated, PHLP provides explainable algorithms without relying on complex deep learning techniques, providing insight into the factors that significantly influence performance for the LP problem of graph data.

Com: 0.8

Acc: 0.75



Evaluation Com and Acc of **Research Background:**

Partially Included +0.5

Completely Included +1

From M-DocSum:

Com = 4.5 / 8 = 0.563

From Qwen2-VL:

Com = 1.5 / 8 = 0.188

Key points from paragraph 1:

1. This paper addresses the link prediction (LP) problem, which involves inferring the connectivity between nodes in graph data.
2. LP is crucial for applications such as social networks, bioinformatics, recommendation systems, and supply chain optimization.
3. Existing LP methods are categorized into heuristic methods, embedding methods, and graph neural network (GNN)-based methods.
4. Heuristic methods rely on predefined structural features but struggle with capturing complex relationships.
5. Embedding methods map nodes into low-dimensional vector spaces but often require large dimensions and may fail to capture similar neighborhood structures.
6. GNN-based methods achieve superior performance by leveraging local and global graph information but are difficult to interpret due to their complexity.
7. Persistent homology (PH), a topological data analysis tool, has been successfully applied to graph and node classification tasks but is underexplored for LP tasks.
8. Previous works, such as TLC-GNN, have integrated PH with GNNs but require further research for datasets without node attributes.

From M-DocSum:

"accuracy": "Mostly Accurate",
"fluency": "Fluent",
"repetitions": 0,
"hallucinations": 0,
"distortions": 1

Acc = (3 - 0 - 1) / 4 = 0.5

From Qwen2-VL:

"accuracy": "Completely Accurate",
"fluency": "Fluent",
"repetitions": 1,
"hallucinations": 0,
"distortions": 0

Acc = (4 - 0 - 1) / 4 = 0.75

accuracy: "Completely Accurate" 4, "Mostly Accurate" 3, "Partially Accurate" 2, "Inaccurate" 1
fluency: "Fluent" 0, "Mostly Fluent" 1, "Not Fluent" 2
penalty_times = repetitions + hallucinations + distortions
Acc = (accuracy - fluency - penalty_times) / 4

Evaluation Criteria

Evaluation of **Reference Images:**

Reference images from M-DocSum: [1, 3, 4, 'None']

Ground truth reference images: [1, 3, 4, 'None']

Reference images from Qwen2-VL: ['None', 1, 5, 4]

ImgAcc = 3 / 3 = 1 NonAcc = 1 / 1 = 1 OMatch = 4 / 4 = 1
JacSim = Jaccard Similarity(set(1,3,4,'None'), set(1,3,4,'None')) = 4 / 4 = 1

ImgAcc = 0 / 3 = 0 NonAcc = 0 / 1 = 0 OMatch = 0 / 4 = 0
JacSim = Jaccard Similarity(set(1,3,4,'None'), set(1,4,5,'None')) = 2 / 4 = 0.5