Table 5. Datasets of GNN explanation tasks.

Class	Dataset	Domain	Paper
Synthetic Datasets	BA-Shapes	Artificially generated	[60], [62], [66], [64], [71], [44], [85], [84], [92], [83], [63], [57]
	BA-Community		[60], [62], [63], [71], [14], [83]
	BA-2motifs		[62], [64], [71], [79]
	BA-3motif		[72]
	Tree-Cycles		[60], [62], [63], [66], [71], [44], [85], [92], [83]
	Tree-Grids		[60], [62], [63], [66], [71], [46], [83], [57]
	Is_Acyclic		[91]
	Infection		[57]
	Tree-BA		[46]
Real-world Datasets	MUTAG	Chemistry	[60], [62], [91], [64], [71], [72], [44], [85], [47], [92], [46], [83], [79], [57]
	ClinTox		[84]
	Tox21		[84]
	BBBP		[64], [84], [79]
	BACE		[84], [79]
	NCI1		[44], [85], [92]
	REDDIT-BINARY	Social network	[60]
	REDDIT-MULTI-5K		[47]
	Bitcoin-Alpha	Bitcoin exchange network	[63]
	Bitcoin-OTC		[63]
	Graph-SST2	Text classification	[64], [84], [83], [79]
	Twitter		[79]
	Cora	Citation network of machine learning	[14], [81]
	CiteSeer		[14], [44]
	PubMed	Citation network of biomedicine	[14], [81]
	VG-5	Computer vision	[72]
	MNIST		[63], [72]

A Datasets

We summarize synthetic and real-world datasets that are commonly used for evaluating the effectiveness of GNN explanation methods. The details of these datasets are presented in Table 5.

Synthetic Datasets. BA-Shapes [60] dataset consists of a base Barabási-Albert (BA) graph with 300 nodes and 80 "house"-structured network motifs, each consisting of five nodes. These motifs are randomly attached to nodes in the base graph. The nodes in the dataset are classified into four classes based on their structural roles: top, middle, and bottom nodes of the house motifs, and nodes that do not belong to any house.

BA-Community [60] dataset is created by combining two BA-Shapes graphs. The nodes in this dataset have feature vectors that follow a normal distribution. Each node is assigned to one of eight classes based on their structural roles within the house motifs and their community memberships.

BA-2motifs [62] dataset uses BA graphs as the basis. It includes graphs with two types of motifs: "house" motifs and five-node cycle motifs. The dataset is divided into two classes based on the type of motif attached to the graphs. Similarly, BA-3motif [72] dataset utilizes BA graphs as the basis. Each base graph is augmented with one of three motifs: house, cycle, or grid.

Tree-Cycles [60] dataset begins with a basic graph that is an 8-level balanced binary tree. Additionally, the dataset includes 80 six-node cycle motifs, which are randomly attached to nodes in the base graph.

Tree-Grids [60] dataset shares similarities with the Tree-Cycles dataset. However, instead of attaching cycle motifs to the base tree graph, 3-by-3 grid motifs are attached in the Tree-Grids dataset.

Other synthetic datasets include Is_Acyclic [110], BA-3motif [72], Infection [41], and Tree-BA [46].

Real-world Datasets. Depending on the specific domain, real-world datasets can be divided into the following categories:

- **Chemistry**: MUTAG [111] consists of 4,337 molecule graphs that have been labeled based on their mutagenic effect on a specific type of bacteria. Each graph in the dataset represents a molecule, and its label indicates whether it is mutagenic or non-mutagenic. ClinTox, Tox21, BBBP, and BACE [6] are publicly available and widely used in drug discovery and chemical informatics. These datasets consist of chemical molecule graphs that are labeled based on their specific chemical properties. NCI1 [33] dataset consists of 4,110 chemical compounds that are labeled as positive or negative in relation to cell lung cancer.
- **Social Networks**: REDDIT-BINARY [30] is a dataset consisting of 2,000 graphs, where each graph represents an online discussion thread from Reddit. Nodes in the graphs represent participating users, and edges indicate replies between users' comments. REDDIT-MULTI-5K [30] is a dataset that includes 4,999 social networks labeled with five different classes, representing the topics of question/answer communities.
- **Bitcoin Exchange Network**: Bitcoin-Alpha and Bitcoin-OTC datasets [112] are networks comprising 3,783 and 5881 accounts involved in Bitcoin trading on platforms. Each account is rated by other members on a scale of -10 (total distrust) to +10 (total trust).
- **Text Classification**: Graph-SST2 [113] and Twitter [114] are sentiment graph datasets for graph classification. Nodes denote words and edges represent the relationships between words.
- **Citation Network of Machine Learning**: Cora [115] and CiteSeer [116] consist of machine learning papers. Nodes in these datasets represent individual papers, while edges represent the citations between papers. Cora dataset is labeled with seven class labels, whereas CiteSeer dataset has six class labels.
- **Citation Network of Biomedicine**: PubMed [31] is a citation network in the field of biomedicine, comprising papers from the biomedical domain. In this dataset, papers are represented as nodes, and edges indicate that one paper has been cited by another paper.
- **Computer Vision**: VG-5 [72] is constructed using 4,443 (*images*, *scene graphs*) pairs from the Visual Genome dataset. Wherein, the graphs are labeled with five classes: stadium, street, farm, surfing, and forest. Each graph contains regions of the objects as nodes, while edges indicate the relationships between object nodes. MNIST [35] dataset consists of 70,000 images that are converted into graphs based on superpixel adjacency. Each graph in the MNIST dataset is labeled with one of ten digit classes.