D.6. Experiments with GAT, SAGE, GIN types of GNN

We train GAT(Veličković et al., 2018), GraphSAGE(Hamilton et al., 2017), GIN (Xu et al., 2019) GNN model for a binary classification task, consisting of three convolutional layers, a max pooling layer, and a fully connected layer, following best practices from the literature (Vu & Thai, 2020). The model is trained with the Adam optimizer (Kingma & Ba, 2014) and a learning rate of 0.001 for 1000 epochs. The training/validation/testing split is 80%/10%/10%, and the corresponding accuracy measures are presented in Tables 9, 10 and 11.

Table 9. Accuracy of GAT on the graph binary classification task on NCI1, MUTAGENICITY, AIDS and PROTEINS.

	NCI1	MUTAGENICITY	AIDS	PROTEINS
Training	0.788	0.845	0.991	0.780
Validation	0.741	0.802	0.973	0.820
Testing	0.748	0.781	0.973	0.730

Table 10. Accuracy of GraphSAGE on the graph binary classification task on NCI1, MUTAGENICITY, AIDS and PROTEINS.

	NCI1	MUTAGENICITY	AIDS	PROTEINS
Training	0.854	0.896	0.992	0.825
Validation	0.783	0.856	0.978	0.847
Testing	0.809	0.795	0.940	0.748

Table 11. Accuracy of GIN on the graph binary classification task on NCI1, MUTAGENICITY, AIDS and PROTEINS.

	NCI1	MUTAGENICITY	AIDS	PROTEINS
Training	0.863	0.849	0.999	0.810
Validation	0.826	0.800	0.951	0.847
Testing	0.789	0.784	0.946	0.748

We experiment on ComRecGC and get the results for the FCR problem in Tables 12, 13 and 14. The parameters for ComRecGC are the same as for the experiments in Section 4 and Table 3, in particular $\Theta = 0.1$ and $\Delta = 0.02$.

Table 12. Results on the FCR problem for the ComRecGC method explaining the GAT trained model.

	NCI1		**Mutagenicity**		**AIDS**		**Proteins**	
	Coverage	Cost	Coverage	Cost	Coverage	Cost	Coverage	Cost
GCFExplainer	24.4%	5.26	47.3%	**5.82**	27.6%	7.12	42.6%	10.54
ComRecGC	**35.6%**	**5.02**	**55.7%**	6.05	**30.7%**	**6.89**	**42.9%**	**10.27**

Table 13. Results on the FCR problem for the ComRecGC method explaining the GraphSAGE trained model.

	NCI1		**Mutagenicity**		**AIDS**		**Proteins**	
	Coverage	Cost	Coverage	Cost	Coverage	Cost	Coverage	Cost
GCFExplainer	32.8%	4.86	46.5%	**5.46**	20.3%	7.38	68.6%	11.53
ComRecGC	**47.9%**	**4.76**	**50.9%**	5.90	**21.5%**	**7.16**	**69.4%**	**11.51**

Table 14. Results on the FCR problem for the ComRecGC method explaining the GIN trained model.

	NCI1		**Mutagenicity**		**AIDS**		**Proteins**	
	Coverage	Cost	Coverage	Cost	Coverage	Cost	Coverage	Cost
GCFExplainer	31.2%	5.13	30.4%	**6.05**	14.7%	7.68	47.3%	12.21
ComRecGC	**45.6%**	**4.58**	**33.7%**	6.41	**16.6%**	**7.34**	**48.6%**	**11.32**

D.7. Additional Dataset

We study the MDB-BINARY and MDB-MULTI datasets (Yanardag & Vishwanathan, 2015). The MDB-BINARY features movies from two genres (Action and Romance), where each graph represents a co-occurrence network of actors in a movie. A recourse represents a way to change the prediction from an action movie to a romance movie. The IMDB-MULTI includes movies from three genres (Comedy, Romance, and Sci-Fi). Since our current method is only interested in binary classification, we change the consider the following labels: Comedy and non-Comedy (i.e Romance and Scifi). A recourse represent a way to change the prediction from an comedy movie to a non-movie.

The parameters for ComRecGC are the same as for the experiments in Section 4 and Table 3, in particular $\Theta = 0.1$ and $\Delta = 0.02$. The results are presented in Table 15.

Table 15. Results on the FCR problem for the ComRecGC method explaining the GIN trained model.

	IMDB-BINARY		**IMDB-MULTI**	
	Coverage	Cost	Coverage	Cost
GCFExplainer	76.5%	8.33	19.9%	**7.65**
ComRecGC	**80.9%**	**8.10**	**21.9%**	7.70