Graph Structure Learning via Lottery Hypothesis at Scale

Editors: Berrin Yanıkoğlu and Wray Buntine

1. Hyperparameters for training

We search for hyperparameters as in Table 1.

Hyperparameters	Searching fields
lr_adj	[0.00001,0.0001,0.001,0.01,0.1]
lr_weight	[0.00001,0.0001,0.001,0.01,0.1,1]
adj_sparsity	[-1,0.2,0.4,0.6,0.8]
weight_sparsity	[-1,0.2,0.4,0.6,0.8]
adj_optimizer	[SGD,Adam]

Table 1: Hyperparameters for training. lr adj and lr weight are learning rate for adjacency matrix and weight respectively in lottery searching. adj sparsity and weight sparsity are sparsity of edges and weight as mentioned in paper. -1 means that we don't prune the mask and maintain the trained scores in the model retrain step