**Datasets :**

1. **Reddit dataset:**

**Papers:** tgn , [shadow, cluster\_gcn(different samples)]

**Details**:

Column names: user\_id, item\_id, timestamp, state\_label, comma\_separated\_list\_of\_features

No of nodes: 11,000

No of edges: 157,474

Edge features: 172

Nodes with dynamic labels: 217

Availability: tgn format , graphsage format

Comments: in the tgn paper and the other 2 papers the Reddit dataset is a different dataset. in all the datasets used in shadow and cluster gcn, no timestamp information is given. moreover, the data is given in the raw format of json files.

1. **Wikipedia :**

**Papers:** tgn

**Details**:

Column names: user\_id, item\_id, timestamp, state\_label, comma\_separated\_list\_of\_features

No of nodes: 9227

No of edges: 672,447

Edge features: 172

Nodes with dynamic labels: 336

Availability : tgn format

1. **Twitter dataset**

Comments: In TGN paper the Twitter dataset instance used was from the recsys2020 competition. this dataset is not publicly available now.

1. **Bitcoin OTC**

**Details**:

Column names: SOURCE, TARGET, RATING, TIME

No of nodes: 5,881

No of edges: 35,592

Edge features: 1 (TARGET)

Comments: bitcoin otc is a dataset of bitcoin transactions. this dataset is having a similar structure to the datasets used in the tgn paper. This one is publicly available by Stanford University.

Also, there is another instance of a similar dataset like BitcoinOtc alpha.

1. **Flicker: sentence based image dataset . will not use .**

**Papers:** shadow

1. **Yelp: cancel no timestamp information**

**Papers:** shadow

1. **Ppi dataset: will not use. (whole protein structure is needed , no time variation)**

**Papers:** cluster\_gcn

**Details**:

Column names: user\_id, item\_id, timestamp, state\_label, comma\_separated\_list\_of\_features

No of nodes: 56,944

No of edges: 818,716

Edge features: 50

labels: 121

Availability : graphsage format