数据说明

一共六个dat文件。每个文件中，每行数据用换行符隔开，每行中的项之间用逗号隔开。

●entity.dat (entity [str] , enetity\_id [int])

○1364行，网络中entity总数。

●tag.dat (tag [str] , tag\_id [int])

○50行，选取最流行的50个tag。

●edge\_weighted.dat (e1\_id [int] , e2\_id [int] , weight [int])

○173390行，entity间的带权边,其中e1\_id<e2\_id。平均边权重。

●mix\_edge\_weighted.dat (e\_id [int] , t\_id [int])

○4174行，tag与entity间的带权边，平均边权重。

●label.dat (e\_id [int] , label\_0 [bool] , label\_1 [bool] , label\_2 [bool] , label\_3 [bool] , label\_4 [bool] , label\_5 [bool] , label\_6 [bool] , label\_7 [bool] , label\_8 [bool] , label\_9 [bool] , label\_10 [bool] , label\_11 [bool] , label\_12 [bool] , label\_13 [bool] , label\_14 [bool] , label\_15 [bool] , label\_16 [bool] , label\_17 [bool])

○1364行，代表电影的题材类型。每部电影平均个label。

●label\_readme.dat (label [str] , label\_id [int])

○18行，总共18个题材类型的名称和id。