This tool is to process and collect the results of the experiments. When running our model, there will be a .xlsx file generated to store the results.

data_collection	2020/6/22 11:19	文件夹	
data_process	2020/6/27 11:51	文件夹	
Movielens_LCFN_1592661763571.xlsx	2020/6/20 23:14	Microsoft Excel ⊥	39 KB
Movielens_LCFN_1592711290100.xlsx	2020/6/21 13:01	Microsoft Excel ⊥	39 KB
Movielens_LCFN_1592728762480.xlsx	2020/6/21 17:51	Microsoft Excel ⊥	39 KB
Movielens_LCFN_1592683858826.xlsx	2020/6/21 5:23	Microsoft Excel <u>⊥</u>	39 KB
Movielens_LCFN_1592715678771.xlsx	2020/6/21 14:14	Microsoft Excel ⊥	39 KB
Movielens_LCFN_1592670693727.xlsx	2020/6/21 1:44	Microsoft Excel ⊥	39 KB

In each .xlsx file, there are three sheet. In the first sheet, we record the dataset, model name, and hyperparameters.

1	A	В	С	D	Е	F	G
1	DATASET	Movielens					
2	MODEL	LCFN					
3	LR	0.002					
4	LAMDA	0.02					
5	LAYER	1					
6	EMB_DIM	64					
7	FREQUENCY_USER	300					
8	FREQUENCY_ITEM	200					
9	BATCH_SIZE	10000					
10	SAMPLE_RATE	1					
11	IF_PRETRAIN	0					
12	N_EPOCH	200					
13	TEST_VALIDATION	Validation					
14	TOP_K	2	5	10	20	50	100
15	OPTIMIZATION	Adam					
16							

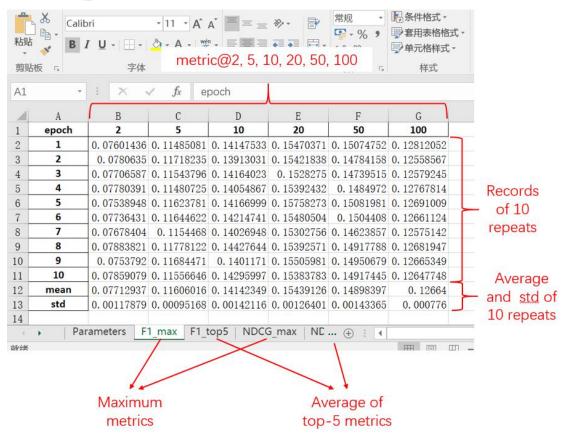
In the second and the third sheet, we record the F_1 -score and NDCG@{2, 5, 10, 20, 50, 100} during the 200 iterations, respectively.

4	A	В	C	D	Е	F	G	Н
1		2	5	10	20	50	100	
2	1	0.025491	0.037529	0.053227	0.065424	0.06836	0.063647	
3	2	0.022988	0.042395	0.056327	0.069336	0.070616	0.064615	
4	3	0.02905	0.048214	0.062033	0.070125	0.071661	0.06667	
5	4	0.027098	0.04641	0.060538	0.071113	0.071862	0.066837	
6	5	0.031212	0.049515	0.065905	0.077754	0.07928	0.073874	
7	6	0.031966	0.054503	0.072498	0.086336	0.08797	0.080106	
8	7	0.045252	0.069414	0.086018	0.098286	0.099943	0.08955	
9	8	0.044415	0.072093	0.090693	0.104253	0.104442	0.093117	
10	9	0.046861	0.079009	0.104209	0.115448	0.110983	0.097785	
11	10	0.052871	0.079945	0.098632	0.109409	0.110955	0.097611	
12	11	0.049676	0.08206	0.102947	0.116912	0.117828	0.102806	
13	12	0.047067	0.077682	0.100658	0.117632	0.119047	0.10567	
14	13	0.05586	0.089469	0.110189	0.12489	0.120897	0.105698	
15	14	0.058745	0.090191	0.112779	0.131653	0.129121	0.112494	
16	15	0.055181	0.086415	0.10719	0.123347	0.12354	0. 107585	
17	16	0.053729	0.085096	0.1104	0.125996	0.12523	0. 110235	
18	17	0.051863	0.085601	0.1093	0.127073	0.124483	0.108638	

When tuning models, for example, with respect to learning rate η and regularization coefficient λ , we run the model:

There will be $10 \times 5 \times 5$ files.

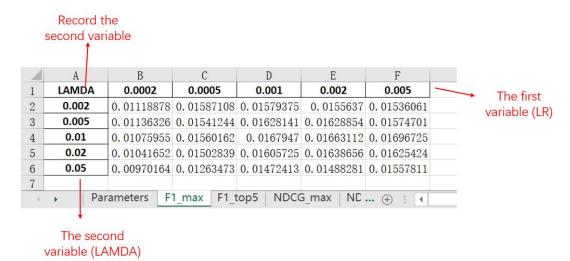
data_process.py can count the maximum and the average of top-5 metrics (F_1 -score and NDCG) during 200 iterations, and calculate the average of 10 repeats. data_process.py generates 5 x 5 files in folder *data process*.



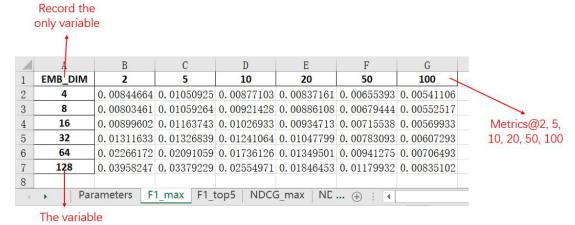
data_collection.py can count the metrics with varying hyperparameters from the files in folder data_process. data_collection.py generates a new files in the folder data_collection and removes the files in the folder data_process. There are 5 sheets in the newly generated file. The first sheet records the settings and hyperparameters. We can see which hyperparameters vary in this sheet.

A	A	В	С	D	Е	F		
1		1	2	3	4	5		
2	DATASET	Amazon						
3	MODEL	LCFN						
4	LR	0. 001	0.002	0. 0005	0. 005	0.0002		
5	LAMDA	0.02	0.002	0.05	0. 01	0.005		
6	LAYER	1						
7	EMB_DIM	64						
8	EQUENCY_US	100						
9	EQUENCY_ITE				Two varia	blos		
10	BATCH_SIZE	10000			I WO Valla	ibles		
	AMPLE_RATI	1						
12	IF_PRETRAIN	0						
13	N_EPOCH	200						
14	ST_VALIDATION	Validation			成图(Alt + A)			
15 TOP_K [2, 5.0, 10.0, 20.0, 50.0, 100.0],								
16 PTIMIZATIO Adam								
17	Parameters F1_max F1_top5 NDCG_max NE +							

In the 2nd-5th sheets, we can see the performances with different parameters.



If there is only one variable:



Please note that data_collection.py only supports at most 2 varying hyperparameters. DATASET and MODEL are not hyperparameters, thus 2 varying hyperparameters with varying DATASET and MODEL are allowed.