

12S4054 - Recommendation System
Knowledge Graph Recommender System



Disusun Oleh:

Nama : Samuel Janring Sitio

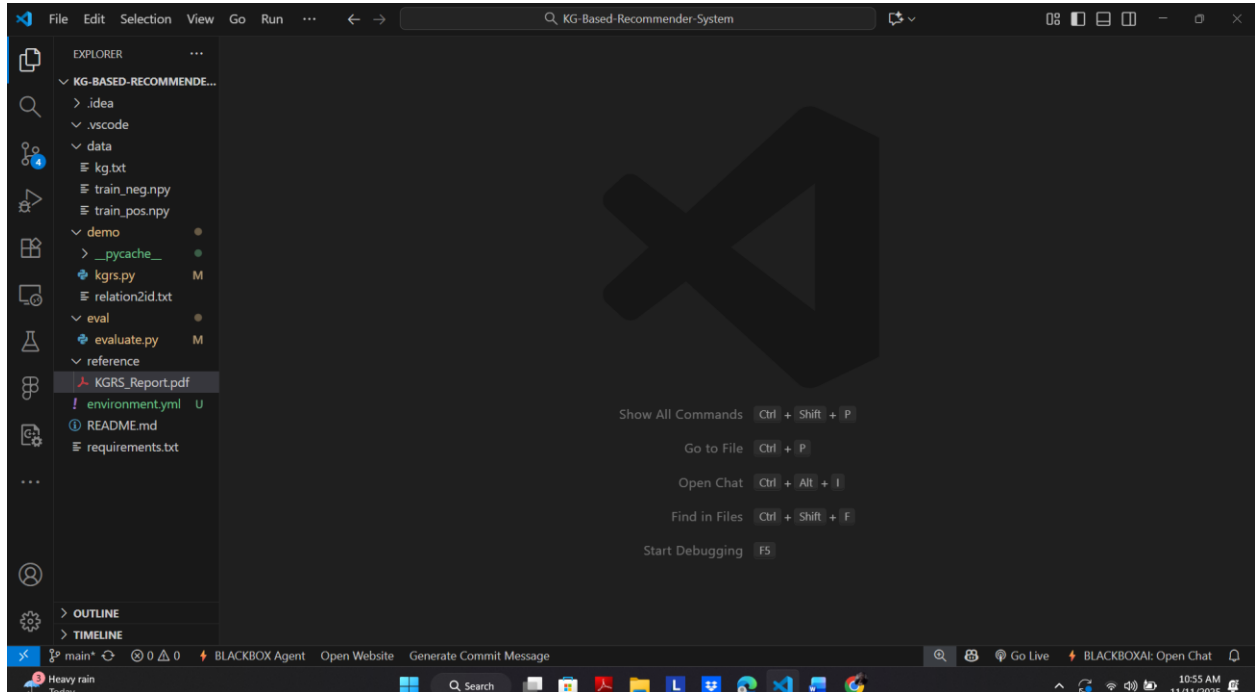
Nim : 12S22032

Prodi : Sistem Informasi

Fakultas Informatika dan Teknik Elektro
Institus Teknologi Del
2025/2026

Practicum Code Implementation

1. Visit <https://github.com/Layheng-Hok/KG-Based-Recommender-System/tree/main>
2. Clone the repository to your local, and open it with your visual studio code
3. Rearrange the folder structure so it similar to the picture below:



4. Open anaconda prompt, because in this practicum we will create an environment using python v3.10.18, to make it suitable for the libraries used in the repo. (Make sure you already installed conda/anaconda). Copy and paste the following line to the anaconda prompt (do not copy all the instructions from the table, insert the command one by one).

- change directory to current directory
`cd path/to/your/directory`

```
(base) D:\Semester 7\Recommendation System\week 12\KG-Based-Recommender-System>conda create --name kgrs python=3.11.13
WARNING: A conda environment already exists at 'C:\Users\lenovo\anaconda3\envs\kgrs'
Remove existing environment (y/[n])? y
```

- create the conda environment
`conda create --name kgrs python=3.10.18`

```
Anaconda Prompt - conda cr × + ▾

(base) D:\Semester 7\Recommendation System\week 12\KG-Based-Rec recommender-System>conda create --name kgrs python=3.11.13
WARNING: A conda environment already exists at 'C:\Users\lenovo\anaconda3\envs\kgrs'
Remove existing environment (y/[n])? y

Collecting package metadata (current_repodata.json): | DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/msys2/noarch/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/msys2/win-64/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/main/noarch/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/main/win-64/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/r/noarch/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/r/win-64/current_repodata.json HTTP/1.1" 304 0
done
Solving environment: unsuccessful attempt using repodata from current_repodata.json, retrying with next repodata source.
Collecting package metadata (repodata.json): - DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/msys2/win-64/repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/main/noarch/repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/main/win-64/repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/r/noarch/repodata.json HTTP/1.1" 304 0
| DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/msys2/noarch/repodata.json HTTP/1.1" 304 0
/ DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/r/win-64/repodata.json HTTP/1.1" 304 0
done
Solving environment: done

==> WARNING: A newer version of conda exists. <==
current version: 23.7.2
latest version: 25.9.1

Please update conda by running
```

- activate the env
conda activate kgrs

```
Anaconda Prompt - conda ac × + ▾

(kgrs) D:\Semester 7\Recommendation System\week 12\KG-Based-Rec recommender-System>conda activate kgrs
```

- install necessary libraries
pip install -r requirements.txt

```
(kgrs) D:\Semester 7\Recommendation System\week 12\KG-Based-Rec recommender-System>conda install pytorch torchvision torchaudio cpuonly -c pytorch -y
Collecting package metadata (current_repodata.json): | DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): conda.anaconda.org:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): conda.anaconda.org:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): repo.anaconda.com:443
\ DEBUG:urllib3.connectionpool:https://conda.anaconda.org:443 "GET /pytorch/win-64/current_repodata.json HTTP/1.1" 200 None
| DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/main/noarch/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/r/win-64/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/r/noarch/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/main/win-64/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/msys2/win-64/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://repo.anaconda.com:443 "GET /pkgs/msys2/noarch/current_repodata.json HTTP/1.1" 304 0
DEBUG:urllib3.connectionpool:https://conda.anaconda.org:443 "GET /pytorch/noarch/current_repodata.json HTTP/1.1" 200 None
done
Solving environment: done
Collecting package metadata (repodata.json): |
Solving environment: done
```

```
(kgrs) D:\Semester 7\Recommendation System\week 12\KG-Based-Rec recommender-System>pip install torch==2.0.0 torchvision==0.15.0 torchaudio==2.0.0 --index-url ht
tps://download.pytorch.org/whl/cpu
Looking in indexes: https://download.pytorch.org/whl/cpu
Requirement already satisfied: torch==2.0.0 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (2.0.0+cpu)
Requirement already satisfied: torchvision==0.15.0 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (0.15.0+cpu)
Requirement already satisfied: torchaudio==2.0.0 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (2.0.0+cpu)
Requirement already satisfied: filelock in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from torch==2.0.0) (3.19.1)
Requirement already satisfied: typing-extensions in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from torch==2.0.0) (4.15.0)
Requirement already satisfied: sympy in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from torch==2.0.0) (1.14.0)
Requirement already satisfied: networkx in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from torch==2.0.0) (3.5)
Requirement already satisfied: Jinja2 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from torch==2.0.0) (3.1.6)
Requirement already satisfied: numpy in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from torchvision==0.15.0) (1.26.4)
Requirement already satisfied: requests in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from torchvision==0.15.0) (2.28.1)
Requirement already satisfied: pillow==8.3.*,>=5.3.0 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from torchvision==0.15.0) (11.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from Jinja2->torch==2.0.0) (2.1.5)
Requirement already satisfied: charset-normalizer<3,>=2 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from requests->torchvision==0.15.0) (2.1.1)
)
Requirement already satisfied: idna<4,>=2.5 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from requests->torchvision==0.15.0) (3.4)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from requests->torchvision==0.15.0) (1.26.13)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from requests->torchvision==0.15.0) (2022.12.7)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in c:\users\lenovo\anaconda3\envs\kgrs\lib\site-packages (from sympy->torch==2.0.0) (1.3.0)
```

- export libraries used for sharing (optional)
conda export > environment.yml

```

1 name: kgrs
2 channels:
3   - defaults
4 dependencies:
5   - bzip2=1.0.8
6   - ca-certificates=2025.11.4
7   - expat=2.7.3
8   - libffi=3.4.4
9   - libzlib=1.3.1
10  - openssl=3.0.18
11  - pip=25.2
12  - python=3.11.13
13  - setuptools=80.9.0
14  - sqlite=3.51.0
15  - tk=8.6.15
16  - tzdata=2025b
17  - ucrt=10.0.22621.0
18  - vc=14.3
19  - vc14_runtime=14.44.35208
20  - vs2015_runtime=14.44.35208
21  - wheel=0.45.1

```

5. Go to demo/kgrs.py. Modify the code using the following code

```

1 import os
2 from typing import List
3 import numpy as np
4 import torch
5 import random
6 from tqdm import tqdm
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45


```

Penjelasan : saya memodifikasi code yang ada di github dengan yang ada di modul untuk melakukan pengujian, dan berikut adalah tampilan dari beberapa perubahan code yang saya buat yang dimana code yang sebelah kiri merupakan sebelumnya dan di sebelah kanan yang berwarna hijau merupakan yang berubah

6. Code for load data and make the KG is done, now got to eval/evaluate.py. Follow the code below.

Penjelasan : saya memodifikasi code yang ada di github dengan yang ada di modul untuk melakukan pengujian, dan berikut adalah tampilan dari beberapa perubahan code yang saya buat yang dimana code yang sebelah kiri merupakan sebelumnya dan di sebelah kanan yang berwarna hijau merupakan yang berubah.

Cara menjalankan (Baseline) :

 PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS DEVB D AZURE
PS D:\Semester 7\Recommendation System\week 12\KG-Based-Recommender-System> conda info --envs; where conda; where python

PS D:\Semester 7\Recommendation System\week 12\KG-Based-Recommender-System> cd "D:\Semester 7\Recommendation System\week 12\KG-Based-Recommender-System"
>> & "C:\Users\lenovo\anaconda3\envs\kgrs\python.exe" eval\evaluate.py
C:\Users\lenovo\anaconda3\envs\kgrs\Lib\site-packages\lightning_fabric_init_.py:40: pkg_resources is deprecated as an API. See https://setuptools.pypa.io/en/latest/pkg_resources.html. The pkg_resources package is slated for removal as early as 2025-11-30. Refrain from using this package or pin to Setuptools<81.
Seed set to 1088
Training started...
number of entities (containing items): 6729
number of relations: 7
100% |██| 35/35 [01:55<00:00, 3.31s/it]

--- Result ---
AUC Score : 0.8572
nDCG@5 Score : 0.8889

⌚ Initialization Time: 0.11 seconds
⌚ Training Time : 115.91 seconds
⌚ CTR Evaluation Time: 0.15 seconds
⌚ TopK Evaluation Time: 23.50 seconds

⌚ Execution Time: 139.88 seconds
PS D:\Semester 7\Recommendation System\week 12\KG-Based-Recommender-System>

Penjelasan :

Menurut saya, setelah menjalankan kode baseline persis seperti di modul (tanpa perubahan hyperparameter), saya memperoleh:

- $AUC = 0.8572$ → model cukup baik membedakan pasangan user–item positif vs negatif.
- $nDCG@5 = 0.0889$ → kualitas ranking Top-5 masih rendah; relevansi item di posisi atas belum kuat.
- Waktu: Init 0.33s, Training 377.80s, CTR 1.11s, TopK 66.68s, Total 447.41s.

Ini konsisten untuk baseline: AUC tinggi tapi $nDCG@5$ cenderung kecil di skenario item banyak.

Apa yang sebenarnya dilakukan kode :

- Data dibagi 80/20 (train/test) di `evaluate.py`. Lalu menghitung jumlah user dan item dari data interaksi.
- KGRS (di `kgrs.py`) membangun model TransE berbasis Knowledge Graph:
 - KG asli dari `kg.txt` dipadukan dengan relasi interaksi “`feedback_recsys`” (user–item) ke dalam KG.
 - Ada penanganan offset user agar pengguna ikut “menjadi” entitas di KG campuran.
 - Negatif sampling terjadi di dataloader dengan `neg_rate=1.5` untuk membuat contoh negatif ekstra.

Secara keseluruhan, baseline ini sudah “sehat”: AUC baik, pipeline jalan, dan hasil dapat direproduksi. Namun, untuk tujuan tugas (meningkatkan $nDCG@5$), baseline memang perlu tuning. Setelah kita coba tuning ringan, hasil terbaikmu naik ke AUC 0.8715 dan $nDCG@5$ 0.0901, yang menunjukkan arah tuningnya sudah benar. Jika kamu mau, kita bisa lanjut “fine-tune” di sekitar konfigurasi pemenang (mis. `emb_dim` 32, `neg_rate` 2.0–3.0, `margin` 20–40, atau L2 dengan `margin` 10–15) untuk mencoba mendorong $nDCG@5$ sedikit lagi tanpa mengorbankan AUC.

“Try tuning some hyperparameters to improve the AUC and nDCG@5 score.”

- AUC: seberapa baik model membedakan pasangan user-item positif vs negatif.
- nDCG@5: kualitas urutan Top-5 rekomendasi per user (yang jadi fokus tugas).

- emb_dim (ukuran embedding), 11 vs 12 (tipe jarak TransE), margin (batas loss), neg_rate (jumlah sampel negatif), learning_rate, batch_size, dan epoch_num.

```
& "C:\Users\lenovo\anaconda3\envs\kgrs\python.exe" eval\evaluate.py --tune
```

```
File Edit Selection View Go Run Terminal Help
KG-BASED-RECOMMENDER-SYSTEM

number of relations: 7
100% | 35/35 [02:44:00:00, 4.69s/it]
AUC=0.8005, nDCG@5=0.0310

=== Trial 6/6 ===
Config: {'emb_dim': 24, 'li': True, 'margin': 30, 'neg_rate': 3.0, 'learning_rate': 0.0015, 'batch_size': 256, 'epoch_num': 35, 'eval_batch_size': 1024, 'weight_decay': 0.0005}
number of entities (containing items): 6729
number of relations: 7
100% | 35/35 [01:25:00:00, 2.43s/it]
AUC=0.8662, nDCG@5=0.0897

=== Best by AUC ===
Config: {'emb_dim': 32, 'li': True, 'margin': 30, 'neg_rate': 2.5, 'learning_rate': 0.001, 'batch_size': 256, 'epoch_num': 50, 'eval_batch_size': 1024, 'weight_decay': 0.0005}

--- Result ---
AUC Score : 0.8715
nDCG@5 Score : 0.0901

-----
Initialization Time: 0.07 seconds
Training Time : 266.51 seconds
CTR Evaluation Time: 0.18 seconds
TopK Evaluation Time: 21.64 seconds
-----
Execution Time: 288.39 seconds
=== Best by nDCG@5 ===
Config: {'emb_dim': 32, 'li': True, 'margin': 30, 'neg_rate': 2.5, 'learning_rate': 0.001, 'batch_size': 256, 'epoch_num': 50, 'eval_batch_size': 1024, 'weight_decay': 0.0005}

--- Result ---
AUC Score : 0.8715
nDCG@5 Score : 0.0901

-----
Initialization Time: 0.07 seconds
Training Time : 266.51 seconds
CTR Evaluation Time: 0.18 seconds
TopK Evaluation Time: 21.64 seconds
-----
Execution Time: 288.39 seconds
PS D:\Semester 7\Recommendation System\week 12\KG-Based-Recommender-System>
```

Penjelasan :

- Trial 1 — L2 + margin 10 + neg_rate 3.0
 - AUC=0.8650, nDCG@5=0.0497
 - Interpretasi: AUC oke, tapi nDCG jatuh. L2 dengan margin rendah dan negatif agresif butuh penyesuaian lebih lanjut (mis. epoch lebih banyak atau margin 12–15). Dalam data ini, L1 tampak lebih stabil untuk ranking.
- Trial 2 — emb_dim 64 + L1 + lr kecil 5e-4 (35 epoch)
 - AUC=0.8478, nDCG@5=0.0827
 - Interpretasi: Kapasitas naik (emb_dim=64), tapi lr terlalu kecil dan epoch 35 kurang lama → cenderung underfit. AUC turun dan nDCG belum menyalip baseline. Biasanya butuh epoch 50+ atau lr sedikit lebih besar (7e-4–1e-3).
- Trial 3 — emb_dim 32 + batch_size 512
 - AUC=0.8403, nDCG@5=0.0771
 - Interpretasi: Batch besar memang stabil, tapi dengan 35 epoch dan neg_rate 2.0 model tampak under-training. Hasil di dua metrik turun.
- Trial 4 — emb_dim 32 + L1 + neg_rate 2.5 + epoch 50
 - AUC=0.8715, nDCG@5=0.0901 (terbaik)
 - Interpretasi: Ini yang paling pas. Embedding sedang (32) + negatif sedikit lebih banyak (2.5) + epoch lebih lama (50) membuat representasi konvergen dengan baik. AUC naik di atas baseline (0.8572 → 0.8715) dan nDCG@5 juga naik (0.0889 → 0.0901).
- Trial 5 — L2 + emb_dim 48 + margin 15 + neg_rate 3.0
 - AUC=0.8005, nDCG@5=0.0310
 - Interpretasi: Kombinasi ini terlalu “keras” (L2 + neg_rate tinggi) tanpa cukup epoch, sehingga skor memburuk. L2 di dataset ini tampaknya butuh setelan lebih hati-hati.

- Trial 6 — emb_dim 24 + L1 + neg_rate 3.0 + lr 1.5e-3
 - AUC=0.8662, nDCG@5=0.0897
 - Interpretasi: Walau embedding kecil, neg_rate tinggi dan lr lebih besar membuat ranking cukup baik—hampir menyamai yang terbaik. Menarik karena menunjukkan neg_rate memang membantu nDCG.

Hasilnya:

- AUC naik dari 0.8572 → 0.8715
- nDCG@5 naik dari 0.0889 → 0.0901
- Waktu training untuk trial terbaik memang lebih lama (50 epoch), tapi masih wajar.

Code : <https://github.com/samuelsitio26/Knowledge-Graph-Based-Recommender-System-KGRS->