

10월 30일,31일

자, Triple Passage와 Cycle Passage의 결과를 비교하자.

그러기 위해선, context score만 사용했을 때, context+graph score만 사용했을 때, graph score만 사용했을 때의 실험 결과를 비교해보고자 한다. 그리고 GSC, QA-GNN 두 가지 모델에 대해서 진행하고자 한다.

실험 세팅을 정리해보자면 다음과 같다. 3개의 seed에 대해서 진행하며, GSC, QA-GNN 두 가지 모델에 적용해서 실험을 진행한다.

노드 개수는 200개를 대상으로 진행한다.

GSC

1. Only context score(orig)

하이퍼 파라미터 세팅(기존 gsc와 동일)

- seed0(20231030_014245)
 - dev : 77.23, test : 71.80
- seed1(20231030_014529)
 - dev : 76.49, test : 70.83
- seed2(20231030_014618)
 - dev : 76.90, test : 71.23

2. Only context score(+Triple)

1. 하이퍼 파라미터 세팅(bs=64,mbs=4,ebs=8,top100)

- seed0(20231030_061344)
 - dev : 70.52, test : 70.59
- seed1(20231031_005921)
 - dev : 73.55, test : 70.19
- seed2(20231031_005933)

- dev : 69.29, test : 69.78

2. 하이퍼 파라미터 세팅(bs=32,mbs=4,ebs=8,top100)

- seed0(20231030_062845)
 - dev : 73.30, test : 69.70
- seed1(20231030_145056)
 - dev : 71.17, test : 70.19
- seed2(20231031_010132)
 - dev : 64.29, test : 64.79

3. 하이퍼 파라미터 세팅(bs=64,mbs=4,ebs=8,top50)

- seed0(20231030_152554)
 - dev : 71.33, test : 70.19
- seed1(20231030_154451)
 - dev : 72.07, test : 70.02
- seed2(20231031_040345)
 - dev : 70.93, test : 70.51

4. 하이퍼 파라미터 세팅(bs=32,mbs=4,ebs=8,top50)

- seed0(20231030_152451)
 - dev : 71.58, test : 70.75
- seed1(20231031_103549)
- seed2(20231031_103612)

3. Only context score(+Cycle Triple)

1. 하이퍼 파라미터 세팅(bs=64,mbs=4,ebs=8,top20)

- seed0(20231031_104906)
- seed1(20231031_123832)
- seed2

4. context score(orig) + graph score

5. context score(+ Triple) + graph score

1. 하이퍼 파라미터 세팅(CSQA-200(bs=64, mbs=4, ebs=8, epoch 30, top100)
max_len = 512(-gsc))
 - seed0(20231026_022502)
 - dev : 74.12, test : 74.21
 - seed1(20231026_022540)
 - dev : 73.05, test : 73.57
 - seed2(20231026_153619)
 - dev : 73.87, test : 73.33
2. 하이퍼 파라미터 세팅(CSQA-200(bs=64, mbs=4, ebs=8, epoch 30, top100)
max_len = 512(-gsc))
 - seed0
 - seed1
 - seed2
3. 하이퍼 파라미터 세팅(CSQA-200(bs=32, mbs=4, ebs=8, epoch 40, top50)
max_len = 512- GSC)
 - seed0
 - seed1
 - seed2
4. 하이퍼 파라미터 세팅(CSQA-200(bs=32, mbs=4, ebs=8, epoch 40, top50)
max_len = 512- GSC)
 - seed0(20231030_020451)
 - dev : 74.37, test : 73.65
 - seed1()
 - dev : , test :
 - seed2(20231028_152934)
 - dev : 75.76, test : 73.09

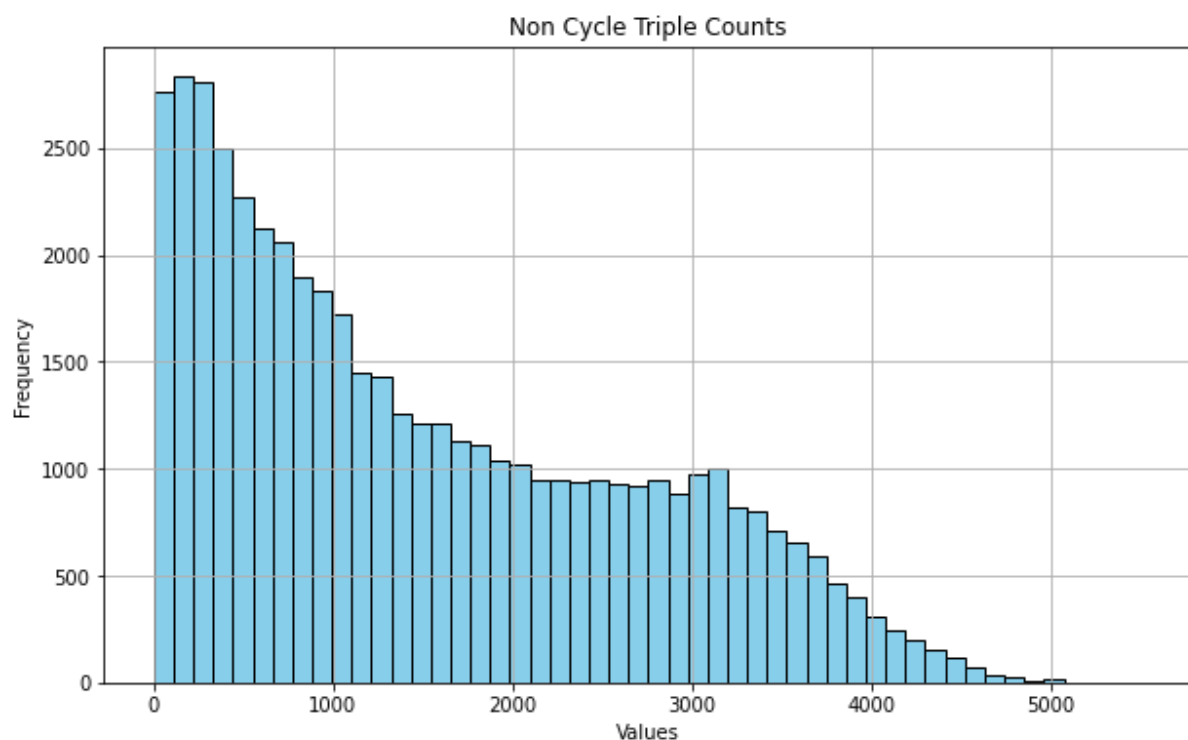
6. context score(+ Cycle Triple) + graph score

7. Only graph score

그래프에서 모든 triple 중 cycle에 들어가지 않는 triple들만을 사용한 triple context score 을 만드는 과정에서

cycle에 들어가지 않는 triple의 분포를 확인해보았다.

Train



Statistics for Your List:

Min: 2

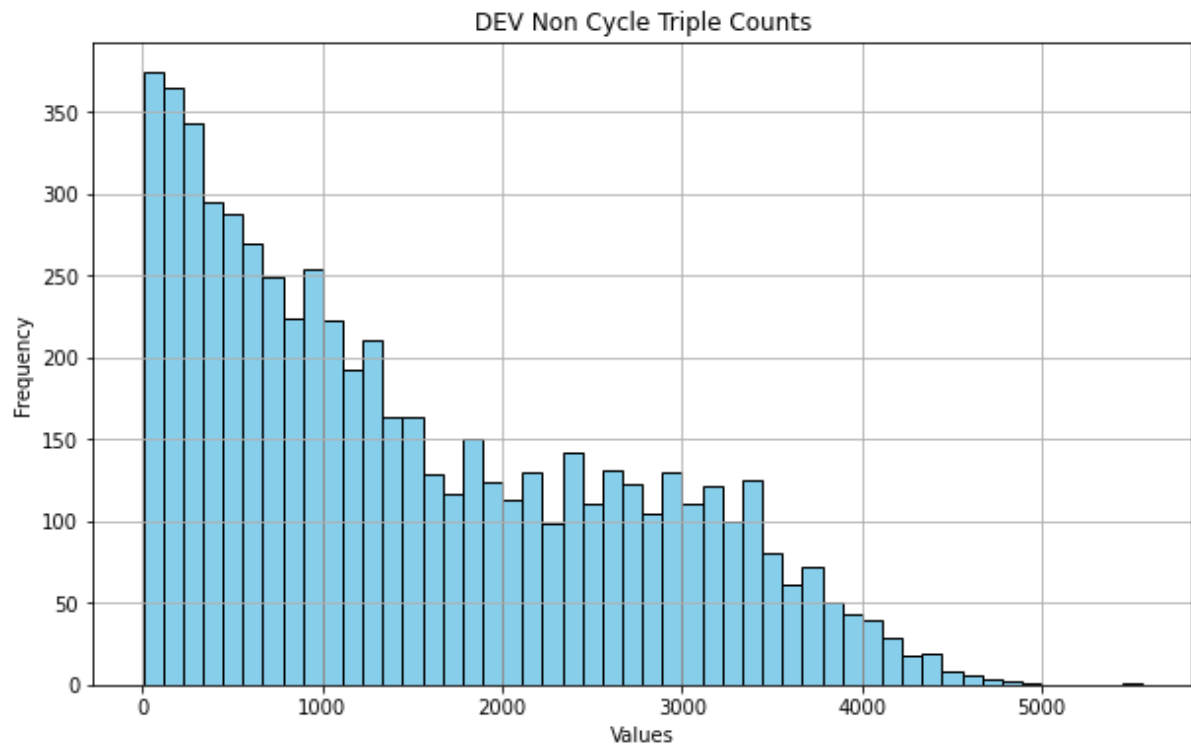
Max: 5514

Average: 1520.607863669028

Median: 1224.0

Standard Deviation: 1172.8462232070328

Valid



Statistics for Your List:

Min: 4

Max: 5554

Average: 1506.3469287469288

Median: 1214.0

Standard Deviation: 1162.458304322471

→ 위 두 가지 그림은 잘못됨, 왜냐하면 context node를 제외하지 않았기 때문