**Project #3**M1505.001600 정보모델링기법과 응용  
2019년도 봄학기

제출기한: **2019년 6월 17일 23:55까지**

\* (10점) 우선, 점수 예측 결과 top-n개 결과를 반환하는 get\_top\_n 함수를 작성하시오.  
(세부 배점은 project3.py 파일의 TODO 참고)

1. (10점) User-based Recommendation

주어진 user에 대하여, 다음의 알고리즘과 유사도 함수를 사용한 추천 결과 top-10 item을 텍스트 파일로 출력하시오.

* 알고리즘: KNNBasic, 유사도: cosine 파일명 1-1.txt
* 알고리즘: KNNWithMeans 유사도: pearson 파일명 1-2.txt

2. (10점) Item-based Recommendation

주어진 item에 대하여, 다음의 알고리즘과 유사도 함수를 사용한 추천 결과 top-10 user를 텍스트 파일로 출력하시오.

* 알고리즘: KNNBasic, 유사도: cosine 파일명 2-1.txt
* 알고리즘: KNNWithMeans 유사도: pearson 파일명 2-2.txt

3. (20점) Matrix-factorization Recommendation

주어진 user에 대하여, 다음의 알고리즘을 사용한 추천 결과 top-10 item을 텍스트 파일로 출력하시오.

* SVD(n\_factors=100, n\_epochs=50, biased=False) 파일명 3-1.txt
* SVD(n\_factors=200, n\_epochs=100, biased=True) 파일명 3-2.txt
* SVD++(n\_factors=100, n\_epochs=50) 파일명 3-3.txt
* SVD++(n\_factors=50, n\_epochs=100) 파일명 3-4.txt

\* 아래의 4 번은 베이스라인 모델 대비 성능 개선을 평가함. 평가지표는 RMSE를 사용함. 베이스라인 모델은 제공되지 않으며, 조교가 임의로 나눈 데이터에 대한 RMSE 값만 제시함.

4. (30점) 전체 데이터 셋에 대해, 1,3의 알고리즘의 파라미터를 조절하거나(유사도 선택 포함) 또는 새로운 알고리즘을 적용하여 가장 좋은 성능의 모델을 제출하시오. (Cross-validation 과정 및 파라미터 튜닝 과정 명시)

**Baseline Model RMSE : 0.6680**

\* 성능 평가 방법 \*

- 70,489쌍의 user/game/score 데이터를 조교가 임의로 training/test set으로 분리하여 training 및 test 진행

- 성능 평가 척도로 RMSE 사용

* 보고서(=발표 자료) 점수 10점
* 발표 점수 10점

***eTL에 업로드된 예시 코드를 응용하시오. 발표자료 및 보고서는 PPT나 PDF의 파일 형태로 python 코드, 텍스트 파일 출력 결과물과 함께 압축하여 eTL 사이트에 제출하시오.***