



Homework #3: Twenty Newsgroup Classification

Sunglok Choi, Assistant Professor, Ph.D.
Computer Science and Engineering Department, SeoulTech
sunglok@seoultech.ac.kr | <https://mint-lab.github.io/>

Twenty Newsgroups Classification

- Task
 - Twenty Newsgroup 데이터셋의 multiclass classification을 수행 (# of classes: 20)
 - 주어진 skeleton code (hw03_skeleton.py) 를 수정
 - scikit-learn 외에 PyTorch 등 다른 도구 사용 가능
 - train()과 predict()의 인터페이스는 유지하며 내부를 수정
 - train()과 predict()에서는 주어진 dataset만 이용 (추가로 test data를 로드하여 사용하는 것 금지)
- Guideline
 - 제출 파일 (hw03_YourStudentID.py)
 - 제출 장소: e-Class > 숙제
- Point: 20
 - 점수는 balanced accuracy에 따라 자동으로 계산됨
 - 20 (≥ 0.70), 18 (≥ 0.67), 16 (≥ 0.64), 14 (≥ 0.61), 12 (≥ 0.58), 10 (≥ 0.55), ...
- Deadline: **2022/04/27 23:59**