Threshold자동계산 코드추가

# 목표

* 현재 모든 모델이 학습 후 deafault값(1)으로 설정됨
* 학습 완료 후 자동으로 Threshold를 구하는 기능을 추가하고자 함

# 기존 방식 확인, 변경 및 추가사항

* 기존 방식, 변경사항

Model\_training (api\_server/apis/multiproc/proc/proc\_annoy.py)  
influx에서 데이터 조회 -> 전체 데이터 concat후 scaler fit -> 각 pkl파일 scaler적용  
-> 조정된 데이터로 모델 학습 -> 모델 및 스케일러 저장

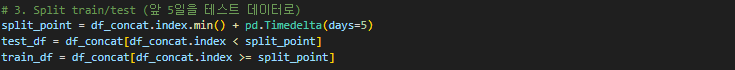
influx에서 데이터 조회 -> train\_test데이터 split -> train데이터로만 scaler fit ->   
-> train, test 데이터 scaler적용 -> train데이터로 모델 학습 -> 모델 및 스케일러 저장  
-> 학습된 모델로 test진행 -> thrshold계산 -> threshold csv파일로 저장

* 추가사항

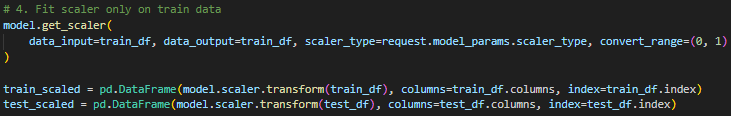
threshold계산함수가 들어있는 .py생성 (src\tag\_setting\threshold\_calculator.py)

# 변경 코드 확인

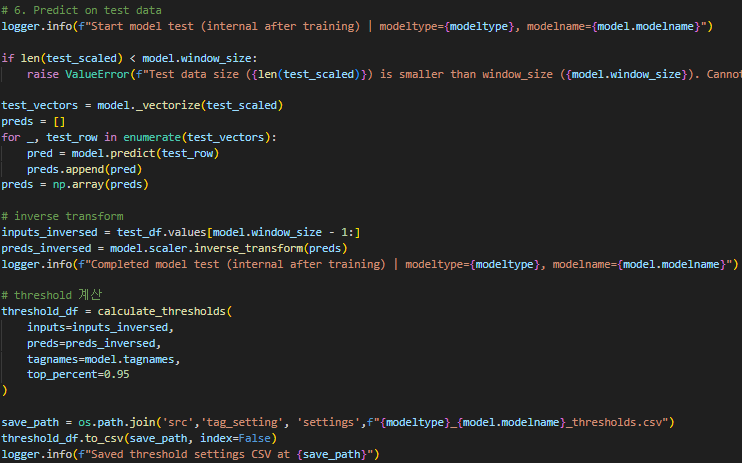
1. 데이터 조회 후 train, test데이터 split(test구간을 전체 데이터 앞 5일로 설정)



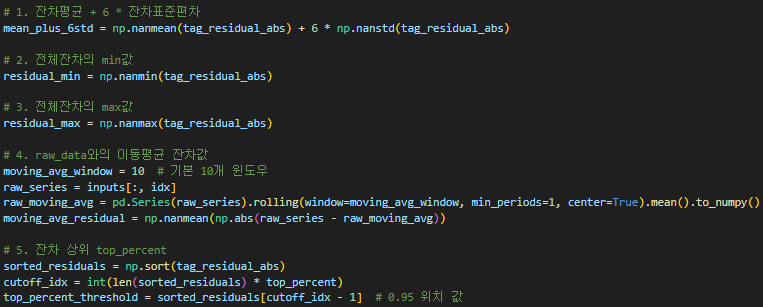
2. train데이터로만 scaler fit 한 후 scaler적용



3. 학습 완료 후 test진행(기존 test에서 진행했던 방식 사용)



4. 사용 계산식(threshold 계산식 4종류)



5. csv파일 결과 확인

training-server\src\tag\_setting\settings 폴더 안에 modeltype\_modelname으로 저장됨

# 추후 계획

현재 ANNOY모델에만 적용된 상태임으로 다른 모델에도 적용시킬 예정