2. **기존에 발생하던 문제인 px4 환경에서의 reset 문제 해결**  
   - px4의 경우 기본적으로 extended kalman filter를 사용하기 때문에 이전 state 값들을 저장해놓고 사용 -> 따라서 시뮬레이터상에서 인위적으로 드론의 위치를 순간이동 시키게 되면 state에 들어가는 값이 비정상적으로 크게 변해 reset할 때 이상행동을 보임  
   - 해결 방법 : drone을 episode마다 reset할 때 위치를 순간이동시키는 것이 아닌 moveToPosition api로 이동  
   or episode 종료 후 환경 reset할 때 px4 재부팅

1. **각속도로 드론을 학습시키는 경우 rlpd 적용 불가능**  
   - offline data 획득이 어려움

- pwm (모터 출력)



- angular rate (각속도)



1. RLPD 적용하지 않은 기존 코드에서 random 위치 시작하는 task training중.

잘 되면 imu에 왜곡을 준 state로 test예정,

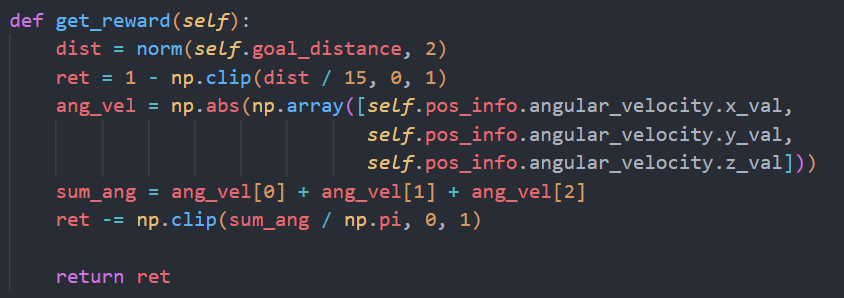
state stacking 구현중.

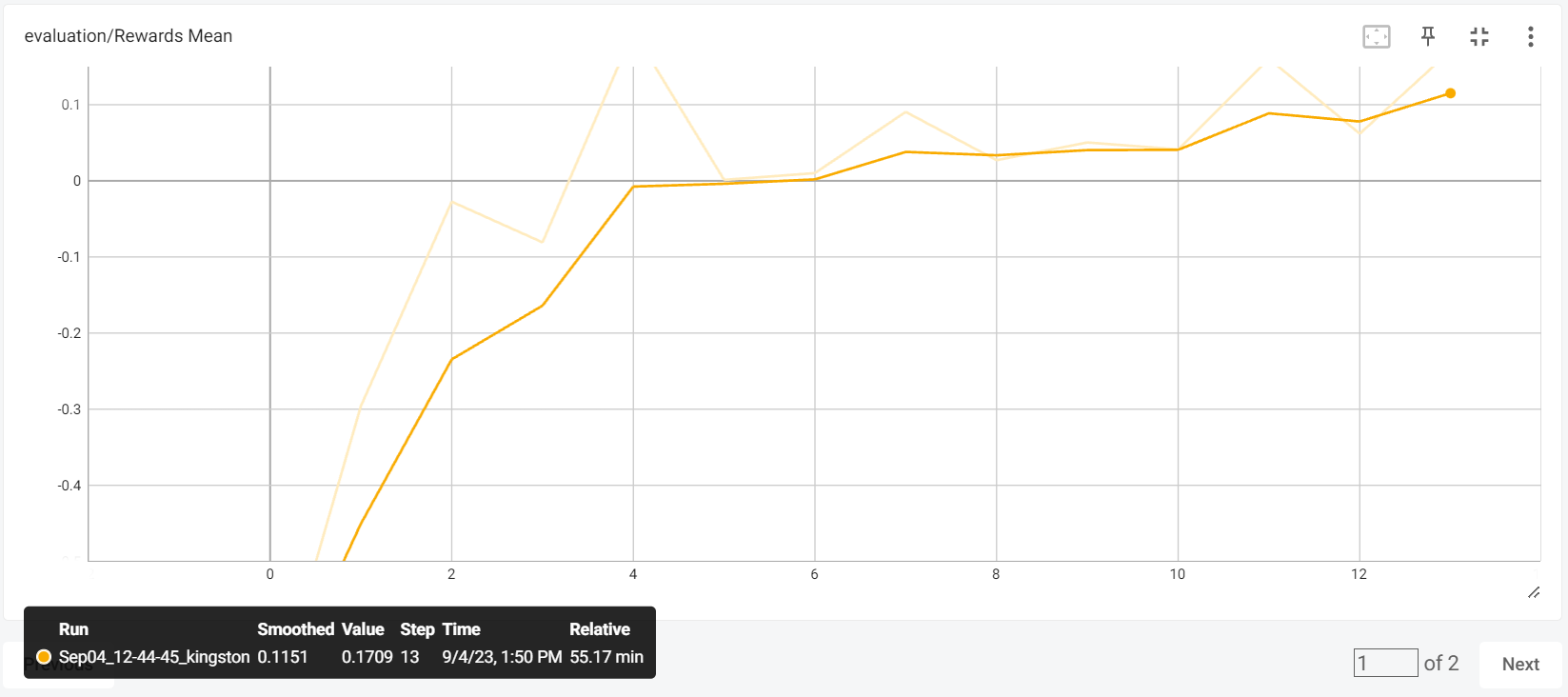
1. Reward shaping

목표 지점 13미터 이내의 랜덤 위치에서 생성해서 목표 지점(0,0,-20)을 찾아가는 task

기존에는 100만 step까지 학습을 시켜도 비행도 하지 못하는 현상이 계속해서 발생

Reward shaping으로 해결 (1만 step도 안되서 균형을 잘 잡고 목적지 근처까지 비행)





1. rlpd 코드 완성 및 수정  
     
   1000 step 즉 1 epoch 진행 후 기존 코드에서는 1번 update하던 방식을 1000번 update로 변경

tensorboard 추가

offline data 10만 스텝 사이즈로 변경 (기존 100만스텝으로 학습이 오래걸림) & offline data를 json 파일로 저장해 다음에 학습 진행할 때 빠르게 load가능하게 변경