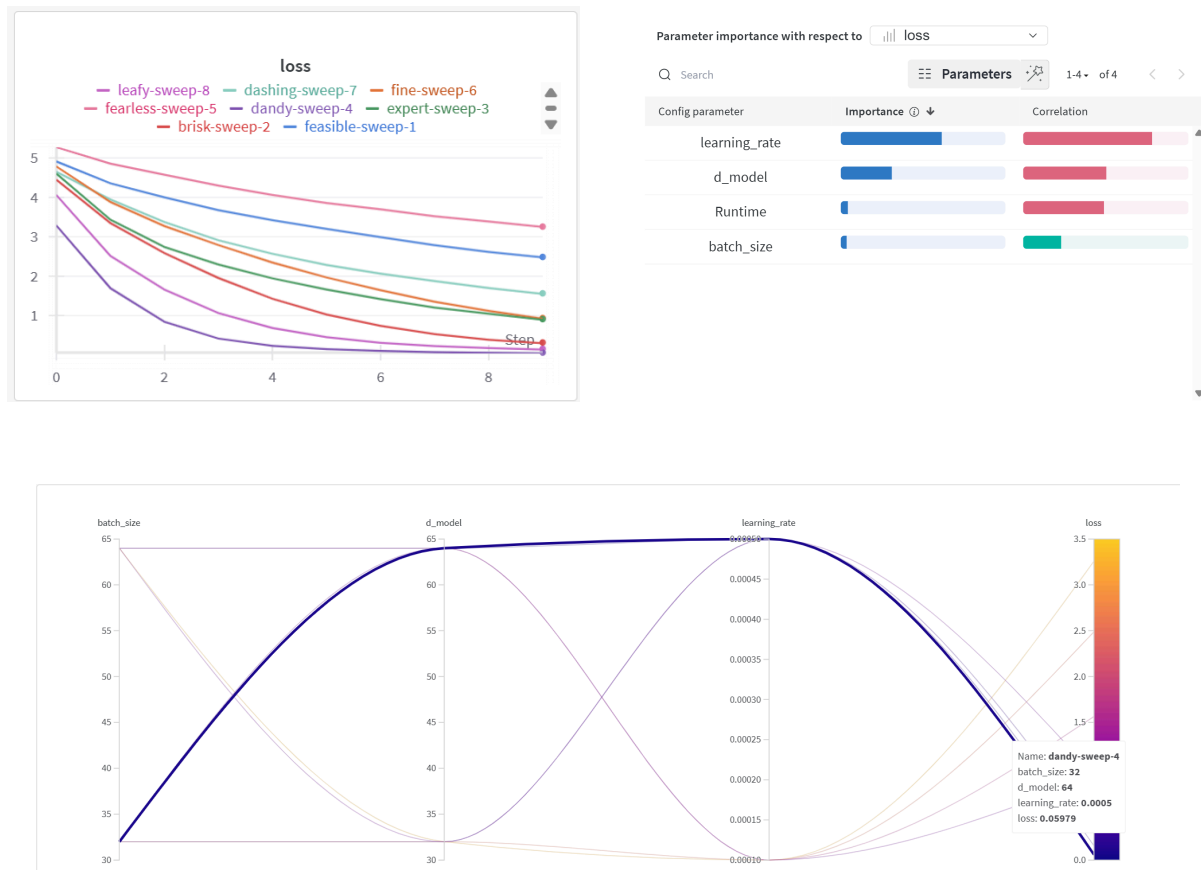


# Week 1 Report

@ DS 25 김소윤

## Sweeping



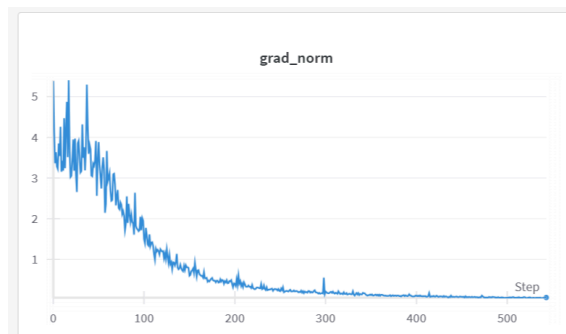
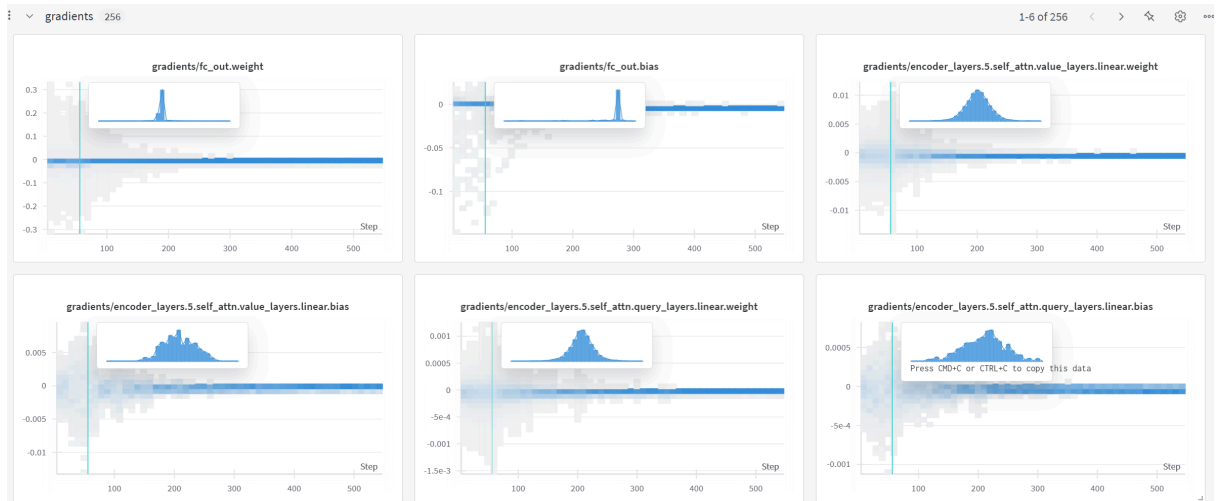
다른 파라미터들은 데이터셋 크기 기준으로 우선 결정했고, lr, batch size, d model 에 대해서만 sweeping 을 진행해 10 epoch동안 학습 양상 살핌.

여러가지 지표를 같이 보긴 했지만 어쨌든 regression 문제니 loss 위주로 봤음.

사진상 dandy-sweep-4가 빠르게 수렴해 가장 작은 loss를 보이므로 최종적으로 해당 파라미터 셋을 선정 해서 최종학습을 돌렸습니다.

파라미터 셋이 작아서 정확하진 않지만 lr, d\_model이 loss 에 대한 importance가 크다고 하니 그쪽으로 실험을 좀 더 돌려보면 좋을 것 같습니다.

## Final training



최종적으로 loss를 낮춰야겠지만,  
과정상에서 step에 따른 파라미터나 그라디언트 변  
화를 살펴보면 학습이 잘 되고 있는지 확인하기가  
 좋을 것 같아요