

CS185 Homework 1 Report

Imitation Learning

Name Of Student: Chong Wei Kee

SID: 3041940996

Q2. Action Chunking with MSE Loss

Training Curves:

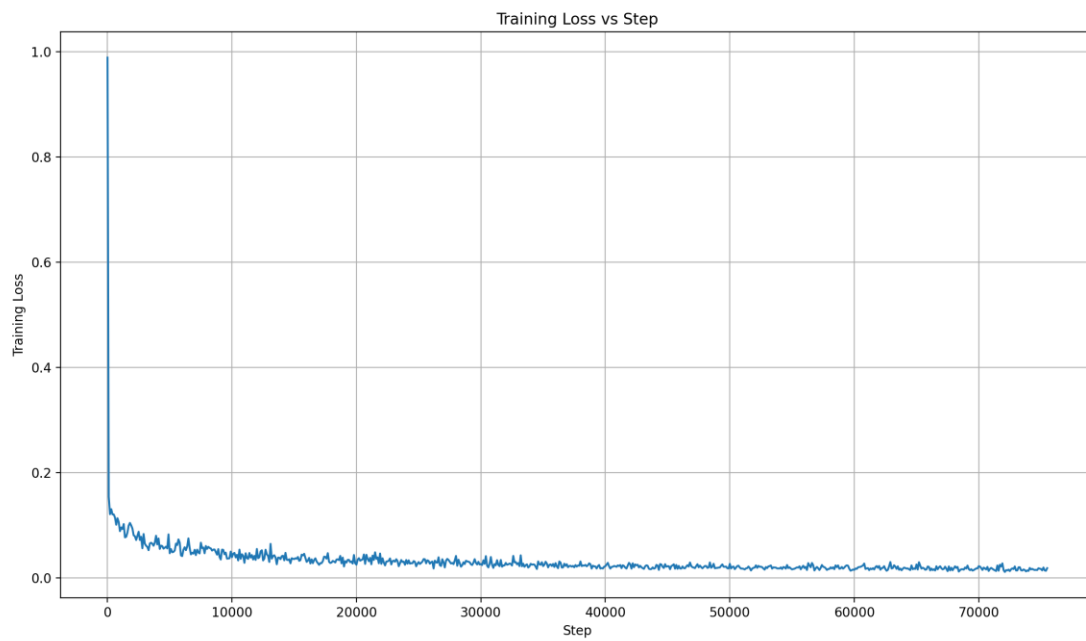


Fig 1: Plot of BasePolicy Training Loss Against Training Step

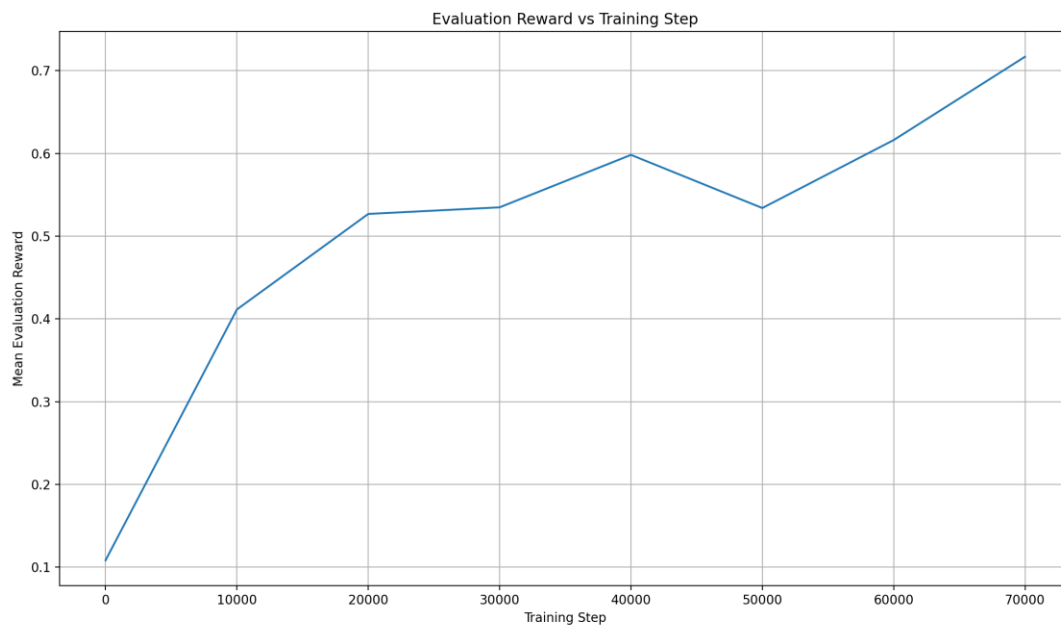


Fig 2: Plot of BasePolicy Reward Against Training Step

Description Of MLP Architecture:

I implemented a MLP architecture with hidden layers of sizes [128,256,128]. I added a ReLU activation function between each hidden layer. This is shown visually below. I achieved an eval/mean reward of 0.717.



I initially considered a MLP using the default hidden layers [128,128] but this achieved an eval/mean-reward of about 0.446, below the threshold of 0.5. I then went on to include the additional hidden layer of size 256. My intuition was increasing the dimensionality of the hidden layer from 128 to 256 might allow the model to capture more relationships regarding the observations, signals and rewards. Then condensing it from 256 to 128 might condense these relationships into more complex relationships which would compute the actions more accurately.

I further experimented with cosine scheduling but discovered it dropped my eval/mean-reward to 0.628. Hence, I reverted to my original implementation without scheduling.

Q3. Action Chunking with Flow Matching

Training Curves:

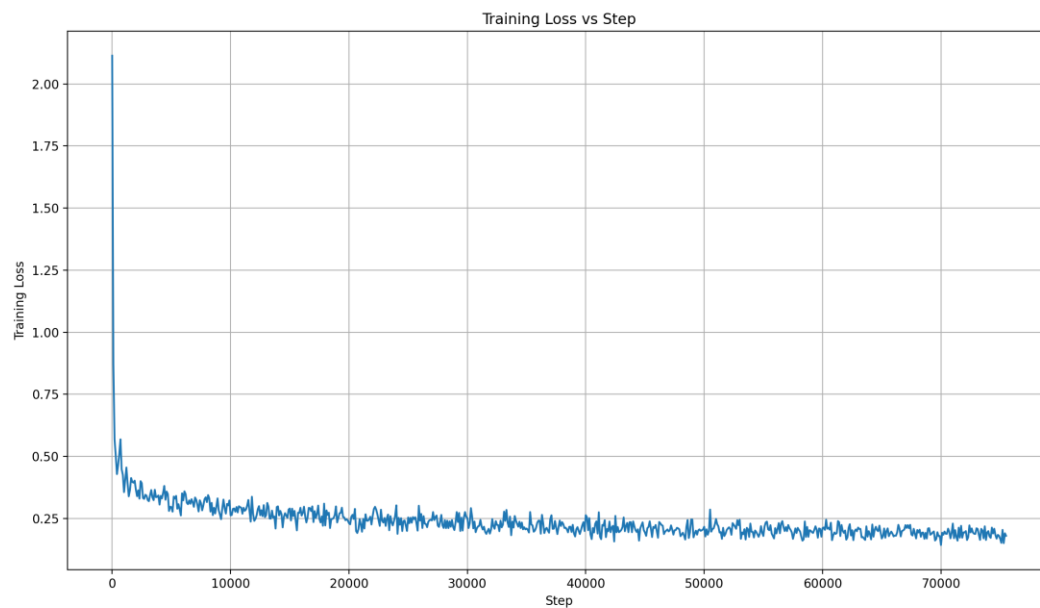


Fig 3: Plot of FlowMatchingPolicy Training Loss Against Training Step

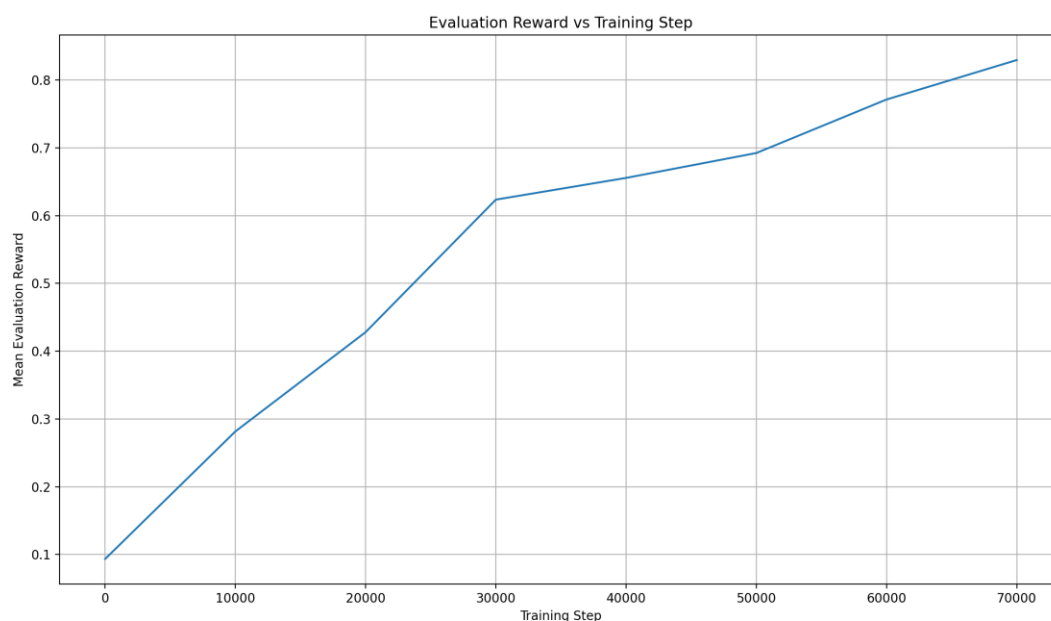


Fig 4: Plot of FlowMatchingPolicy Reward Against Training Step

Description Of MLP Architecture:

Only slightly edited such that the first layer inputs **(state_dim + (chunk_size * action_dim) + 1)** to account for the additional dimensionality due to the inputs of action and tau, since the model is now learning the vector field and not the action. I achieved a mean reward of 0.829.