

Dataset	MBOP	TT	Ours
HalfCheetah-m	44.6	46.9 ±0.4	57.4±1.3
Hopper-m	48.8	61.1±3.6	75.9±12.3
Walker2d-m	41.0	79.0±2.8	84.7±4.5
HalfCheetah-m-r	42.3	41.9±2.5	53.4±1.3
Hopper-m-r	12.4	91.5±3.6	99.5±8.6
Walker2d-m-r	9.7	82.6±6.9	82.3±11.7
HalfCheetah-m-e	105.9	95.0±0.2	92.7±1.6
Hopper-m-e	55.1	110.0±2.7	91.3±18.4
Walker2d-m-e	70.2	101.9±6.8	110.2±0.7
MuJoCo total	430	709.9±29.5	747.4±60.4

Table 1: Average normalized score and the standard deviation of all algorithms over five seeds in the Gym. For every seed, we sample four trajectories and calculate the average return of these trajectories during the evaluation period. The highest-performing scores are highlighted. The score of TT and MBOP is the reported scores in Table 1 of TT.

1 Compare With Transformer And Model-based

We compared Trajectory Transformer and Model-based Offline Planning on MuJoCo tasks. By this Table, we find our algorithm better than TT on MuJoCo tasks without optimal dataset and competitive on other datasets.