IEMS 490 Reinforcement Learning: UCB for Contextual Bandits

Yiming Peng

Department of IEMS

Northwestern University

December 13, 2018

1. I tune the parameter of UCB algorithm by a grid search and choose 0.5 in the end. Then I use it and the corresponding trained model on the test data and compute its regret against the optimal rewards (i.e. opt_rewards returned by sample_jester_data()5).