

Appendices

A. Additional Figures

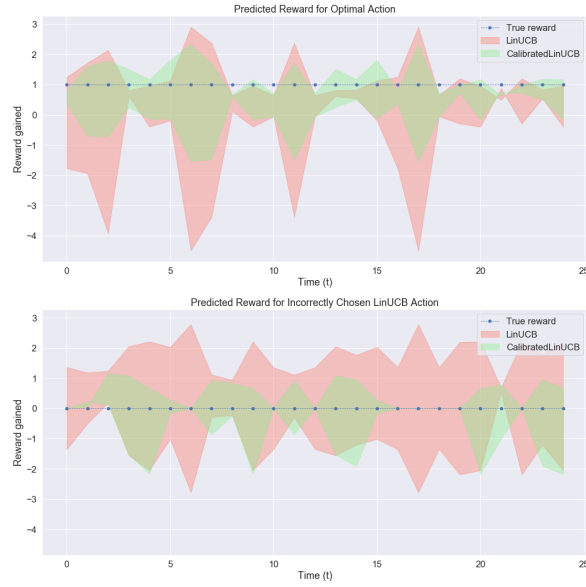


Figure 3. Predicted expected reward for both LinUCB and CalLinUCB algorithms on the covertype dataset. Figures show predictions at random timesteps where CalLinUCB chose the optimal action but LinUCB did not. Top: Predicted reward of both algorithms for the optimal action. Bottom: Predicted reward of both algorithms for the action which the algorithm chose to pick instead of the optimal action at that timestep. We can see LinUCB consistently underestimates reward from optimal action and overestimates reward from other actions. On the other hand, CalLinUCB is more accurate in its uncertainty predictions.

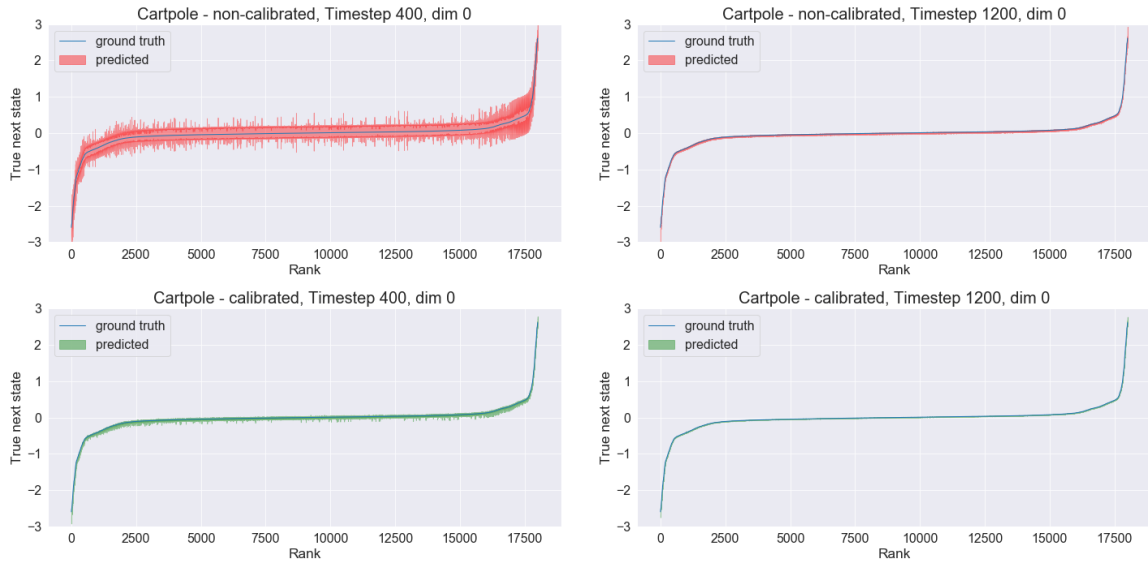


Figure 4. Cartpole future state predictions. The calibrated algorithm has much tighter uncertainties around the true next state in early training iterations. Later into training, their uncertainties are almost equivalent.