CS6700 : Reinforcement Learning

Short Assignment on Bandits — Deadline: 09 Feb, 2019, 11:55 pm

- This is an individual assignment. Collaborations and discussions are strictly prohibited.
- Be precise with your explanations. Unnecessary verbosity will be penalized.
- Check the Moodle discussion forums regularly for updates regarding the assignment.
- **Please start early.**

1. You have come across Median Elimination as an algorithm to get (ϵ, δ)-PAC bounds on the best arm in a bandit problem. At every round, half of the arms are removed by removing arms with return estimates below the median of all estimates. How would this work if we removed only one-fourth of the worst estimated arms instead? Attempt a derivation of the new sample complexity. You can refer the paper on median elimination at this URL.