CSCI 4962 Discussion 7

Siddha Kilaru

Paper 1

“Finite-time Analysis of the Multiarmed Bandit Problem” by Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer:

The K-armed bandit problem is given a set of K random variables that are independent and not necessarily independent from one another, which is the best one to sample from over time to maximize utility. An algorithm called a policy chooses the following random variable to sample from given data from prior sampling. This paper is about empirically and analytically analyzing how the regret of certain policies changes over time. Moreover, the paper uses a proof that says the regret follows a logarithmic curve over time, and the research paper was able to recreate this using different policies such as $UCB1$, $UCB2$, and ϵ-greedy. One way to extend the paper is to drop the stationarity assumption of the K-armed bandit problem, which turns it into a stochastic K-armed bandit problem. Here, there needs to be more analysis done on how regret behaves over time because the problem is much more general. This analysis can be purely analytical, or maybe some conclusions can be drawn from using an algorithm (policy) designed with the stationarity assumption in mind.

Paper 2

Paper 3