Statement of Purpose for Ph.D. application for Fall 2018

Department: Computer Science   *Applicant: Subhojyoti Mukherjee*

I want to pursue Ph.D. in Computer Science and I aspire to become a professor in this field. My research interests span the areas of *Machine Learning, Reinforcement Learning, Online Optimization, and Recommender systems*. In my pursuit of these areas, I have explicitly focused in the area of Multi-armed Bandit (MAB) in my past works. In recent years, MABs have been increasingly gaining attention in a number of inter-disciplinary areas such as online education, online medical/health recommendation, online advertising, worker productivity management, and several other interesting industrial applications.

During my M.S at Indian Institute of Technology Madras, I collaborated with my advisors, *Dr. Balaraman Ravindran* and *Dr. Nandan Sudarsanam* as well as *Dr. K.P. Naveen* from Indian Institute of Technology Tirupati. A few of my past research works that resulted in publications in premier conferences from these collaborations (see Mukherjee et al. (2018), Mukherjee et al. (2017)) and a short description of them are listed below:-

1. **Efficient-UCBV: An Almost Optimal Algorithm using Variance Estimates:** In this work, we focus on the simple stochastic bandit model. We propose a novel variant of the UCB algorithm (referred to as Efficient-UCB-Variance (EUCBV)) for minimizing cumulative regret in the stochastic multi-armed bandit (MAB) setting. EUCBV incorporates the arm elimination strategy proposed in UCB-Improved, while taking into account the variance estimates to compute the arms' confidence bounds, similar to UCBV. Through a theoretical analysis we establish that EUCBV incurs a *gap-dependent* regret bound of $O\left(\frac{K\sigma_{\max}^2 \log(T\Delta^2/K)}{\Delta}\right)$ after T trials, where Δ is the minimal gap between optimal and sub-optimal arms; the above bound is an improvement over that of existing state-of-the-art UCB algorithms (such as UCB1, UCB-Improved, UCBV, MOSS). Further, EUCBV incurs a *gap-independent* regret bound of $O\left(\sqrt{KT}\right)$ which is an improvement over that of UCB1, UCBV and UCB-Improved, while being comparable with that of MOSS and OCUCB. Through an extensive numerical study, we show that EUCBV significantly outperforms the popular UCB variants (like MOSS, OCUCB, etc.) as well as Thompson sampling and Bayes-UCB algorithms. This work has been accepted for publication in *Proceedings of the Thirty-Second Association for the Advancement of Artificial Intelligence (AAAI-18)*.

2. **Thresholding Bandits with Augmented UCB:** In this work, we focus on a variant of the stochastic bandit model called the Thresholding Bandit Problem. We propose the Augmented-UCB (AugUCB) algorithm for a fixed-budget version of the thresholding bandit problem (TBP), where the objective is to identify a set of arms whose quality is above a threshold. A key feature of AugUCB is that it uses both mean and variance estimates to eliminate arms that have been sufficiently explored; to the best of our knowledge, this is the first algorithm to employ such an approach for the considered TBP. Theoretically, we obtain an upper bound on the loss (probability of misclassification) incurred by AugUCB. Although UCBEV in literature provides a better guarantee, it is important to emphasize that UCBEV has access to problem complexity (whose computation requires arms' mean and variances), and hence is not realistic in practice; this is in contrast to AugUCB whose implementation does not require any such complexity inputs. We conduct extensive simulation experiments to validate the performance of AugUCB. Through our simulation work, we establish that AugUCB, owing to its utilization of variance estimates, performs significantly better than the state-of-the-art APT, CSAR and other non variance-based algorithms. This work has been published in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)*.

Another important collaboration was with *Dr. Odalric Maillard* during a three-month internship at INRIA, SequeL Lab, Lille, France from September, 2017 to November, 2017. We worked in the area of piecewise stochastic MAB where there are changepoints when the distribution associated with each arm/action changes abruptly. We devised

actively adaptive algorithms to work in this environment, and we are soon planning to publish our research output. This internship has been an eye-opening experience for me where I worked with researchers explicitly on open problems and it was very exciting to get accustomed to the environment of a full-fledged research lab.

Further, I am going for a three-month internship in Adobe Research, San Jose, USA from January 2018 - April 2018 under the supervision of *Dr. Branislav Kveton* to work in the area of conservative contextual bandits. I have planned this internship so that I can actively engage with industry researchers and gain a full-some experience on how research is conducted in industrial labs and how it affects the day-to-day life of users interacting with industrial products.

Although I have mainly focused on immediate Reinforcement Learning settings and mostly theoretical works, I am open to a variety of research; there are several professors at Carnegie Mellon University such as *Dr. Pradeep Ravikumar* and *Dr. Barnabás Póczos* whose projects are especially appealing to me. Dr. Pradeep Ravikumar works in theoretical machine learning and its applications to high dimensional real-life data-sets which is an area I am interested to work in. Specifically, I want to work with him in the theoretical areas of learning theory and find common solutions where bandit theory can be applied. Also, I am interested to work with Dr. Barnabás Póczos in the area of online learning and optimization. I want to collaborate with him in several of his areas of interest such as automated discovery and efficient data processing in applied sciences including health-sciences and robotics, where there have been several bandit variants which can be applied.

References

Mukherjee, S.; Kolar Purushothama, N.; Sudarsanam, N.; and Ravindran, B. 2017. Thresholding bandits with augmented ucb. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, 2515–2521.

Mukherjee, S.; Kolar Purushothama, N.; Sudarsanam, N.; and Ravindran, B. 2018. Efficient-UCBV: An almost optimal algorithm using variance estimates. In *Proceedings of the 32nd Association for the Advancement of Artificial Intelligence (to appear)*.