Abstract

We first consider a crowdsourcing platform where workers' responses to questions posed by a crowdsourcer are used to determine the hidden state of a multi-class labeling problem. As workers may be unreliable, we propose to perform sequential questioning in which the questions posed to the workers are designed based on previous questions and answers. We propose a Partially-Observable Markov Decision Process (POMDP) framework to determine the best questioning strategy, subject to the crowdsourcer's budget constraint. As this POMDP formulation is in general intractable, we develop a suboptimal approach based on a q-ary Ulam-Rényi game. We also propose a sampling heuristic, which can be used in tandem with standard POMDP solvers, using our Ulam-Rényi strategy. We demonstrate through simulations that our approaches outperform a non-sequential strategy based on error correction coding and which does not utilize workers' previous responses.

We next study how to design a dynamic selection strategy that poses different questions to different workers optimally at each step. Workers participating in a crowdsourcing platform can have a wide range of abilities and interests. An important problem in crowdsourcing is the task recommendation problem, in which tasks that best match a particular worker's preferences and reliabilities are recommended to that worker. A task recommendation scheme that assigns tasks more likely to be accepted by a worker who is more likely to complete it reliably results in better performance for the task requester. Without prior information about a worker, his preferences and reliabilities need to be learned over time. We propose a multi-armed bandit (MAB) framework to learn a worker's preferences and his reliabilities for different categories of tasks. However, unlike the classical MAB problem, the reward from the worker's completion of a task is unobservable. We therefore include the use of gold tasks (i.e., tasks whose solutions are known *a priori* and which do not produce any rewards) in our task recommendation procedure. Our model could be viewed as a new variant of MAB, in which the random rewards can only be observed at those time steps where gold tasks are used, and the accuracy of estimating the expected reward of recommending a task to a worker depends on the number of gold tasks used. We show that the optimal regret is $O(\sqrt{T})$, where T is the number of tasks recommended to the worker. We develop three task recommendation strategies to determine the number of gold tasks for different task categories, and show that they are order optimal. Simulations verify the efficiency of our approaches.

In the above work, stochastic independence of rewards is assumed for different arms (task categories), which enables the decision maker to consider each arm separately but leads to regret scaled linearly with the number of arms. Reward dependence between arms is therefore often assumed which enables the decision maker to gather information for more than one arm at each time. By doing so, a bandit algorithm can learn to choose good arms more quickly. One specific assumption of dependence is that arms are vectors containing numerical elements, and the expected reward of choosing one arm is an unknown linear function of the arm. Traditionally, the goal is still to maximize the cumulative reward. However, in some scenarios, the full reward can be decomposed to two components, and the objective of the learning process is to maximize the summation of only one component. One typical example is the discrimination prevent problem in recommendation systems. We study the orthogonal projection problem in linearly bandit, where the projection reward is defined as a linear function of the orthogonal projection of the arm into a subspace. We developed one algorithm which can achieve $O(\log T)$ regret for finite many arms set and $O(T^{2/3}(\log T)^{1/2})$ regret for the general compact set of infinite arms. Extensive experiments on both synthetic and real-world dataset verify the efficiency of our algorithm.

Publication List:

Q. Kang and W. P. Tay, "Task recommendation in crowdsourcing based on learning preferences and reliabilities." IEEE Trans. Services Comput., submitted.

Q. Kang and W. P. Tay, "Sequential multi-class labeling in crowdsourcing," IEEE Trans. Knowl. Data Eng., 2018, accepted.

Q. Kang and W. P. Tay, "Sequential multi-class labeling in crowdsourcing: A Ulam-Renyi game approach," in IEEE/WIC/ACM Int. Conf. on Web Intelligence, Leipzig, Germany, Aug. 2017