# Fast and Sample Efficient Multi-Task Representation Learning in Stochastic Contextual Bandits

**Anonymous Authors**[1]

## Abstract

We study how representation learning can improve the learning efficiency of contextual bandit problems. We study the setting where we play $T$ contextual linear bandits with dimension $d$ simultaneously, and these $T$ bandit tasks collectively share a common linear representation with a dimensionality of $r \ll d$. We present a new algorithm based on alternating projected gradient descent (GD) and minimization estimator to recover a low-rank feature matrix. We obtain constructive provable guarantees for our estimator that provide a lower bound on the required sample complexity and an upper bound on the iteration complexity (total number of iterations needed to achieve a certain error level) of our proposed algorithm. We show that our algorithm achieves $\epsilon$-accurate recovery of the feature matrix with order $(d+T)r^3 \log(1/\epsilon)$ total samples and order $NTdr \log(1/\epsilon)$ time for any $\epsilon > 0$ that is lower bounded by the noise to signal ratio (NSR). Using the proposed estimator, we present a multi-task learning algorithm for linear contextual bandits and prove the regret bound of our algorithm. We presented experiments on synthetic and real-world MNIST data. We compared the performance of our algorithm against benchmark algorithms to illustrate our theoretical findings and demonstrate the effectiveness of our proposed algorithm.

## 1. Introduction

Contextual Bandits (CB) represent an online learning problem wherein sequential decisions are made based on observed contexts, aiming to optimize rewards in a dynamic environment with immediate feedback. In CBs, the environment presents a context in each round, and in response, the agent selects an action that yields a reward. The agent's objective is to choose actions to maximize cumulative reward over $N$ rounds. This introduces the exploration-exploitation dilemma, as the agent must balance exploratory actions to estimate the environment's reward function and exploitative actions that maximize the overall return (Bubeck & Cesa-Bianchi, 2012; Lattimore & Szepesvári, 2020). CB algorithms find applications in various fields, including robotics (Srivastava et al., 2014), clinical trials (Aziz et al., 2021), communications (Anandkumar et al., 2011), and recommender systems (Li et al., 2010).

Multi-task representation learning is the problem of learning a common low-dimensional representation among multiple related tasks (Caruana, 1997). Multi-task learning enables models to tackle multiple related tasks simultaneously, leveraging common patterns and improving overall performance (Zhang & Yang, 2018; Wang et al., 2016; Thekumparampil et al., 2021). By sharing knowledge across tasks, multi-task learning can lead to more efficient and effective models, especially when data is limited or expensive. Multi-task bandit learning has gained interest recently (Deshmukh et al., 2017; Fang & Tao, 2015; Cella et al., 2023; Hu et al., 2021; Yang et al., 2020). Many applications of CBs, such as recommending movies or TV shows to users and suggesting personalized treatment plans for patients with various medical conditions, involve related tasks. These applications can significantly benefit from this approach, as demonstrated in our empirical analysis in Section 6. This paper investigates the benefit of using representation learning in CBs theoretically and experimentally.

While representation learning has demonstrated remarkable success across various applications (Bengio et al., 2013), its theoretical understanding still remains underexplored. A prevalent assumption in the literature is the presence of a shared common representation among different tasks. (Maurer et al., 2016) introduced a general approach to learning data representation in both multi-task supervised learning and learning-to-learn scenarios. (Du et al., 2020) delved into few-shot learning through representation learning, making assumptions about a common representation shared between source and target tasks. (Tripuraneni et al., 2021) specifically addressed the challenge of multi-task linear regression with low-rank representation, presenting algorithms with robust statistical rates.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.