CSE 541 Course Project Guidelines

Spring 2025 – Instructor: Kevin Jamieson

Objective

The purpose of the course project is to deepen your understanding of multi-armed bandit algorithms by either:

1. Applying techniques you've learned to a real dataset or problem relevant to your interests (ideally aligned with your research), or
2. Exploring a theoretical area not covered in the course through a thorough literature review, culminating in identifying open questions or proposing new directions.

This is an opportunity to be creative. Bandits are a powerful abstraction for sequential decision-making under uncertainty. Whether your goal is scientific, practical, or theoretical, the project should demonstrate depth of thought and technical engagement.

Project Requirements

You may work with a group of size two (i.e., you and a partner) or solo. If you work with a partner the project is expected to be more ambitious.

- **Proposal (due Sunday, May 25):** <u>One page max</u> (including references–reports longer than 1 page will be severely penalized). Clearly state your goal, relevant literature or datasets, and a rough plan. The purpose here is to get feedback from Kevin about your ideas, so feel free to ask for advice if you're mulling a few different ideas.
- **Final Report (due Monday, June 10):** <u>Three pages max</u> (including references–reports longer than 3 pages will be severely penalized). The final report should demonstrate a deep understanding of the chosen topic and clearly articulate your reasoning and findings. For theoretical projects, you are not expected to derive new results. Instead, focus on carefully surveying the literature in a focused area we did not cover in class (some examples are below). Your report should clearly explain the setting, summarize what is known (with references to key papers), and highlight open problems or what remains under-explored. It is expected that you will review at least several papers in your research. For empirical projects, it is crucial to use a real dataset (not simulated data!) and be explicit about your modeling assumptions and how you justified them. Discuss how well these assumptions align with the data, and importantly, what you would have done differently with more time, better data, or more computational resources. What additional data would you collect to improve your approach? Are there alternative modeling choices you considered? Your report should reflect thoughtful reflection on both the strengths and limitations of your approach.

I don't care what Latex template you use. But please keep it to 11 size font and 1 inch margins (i.e, `\documentclass[11pt]{article} \usepackage[margin=1in]{geometry}`)

Suggested Project Directions

Here are example categories and project prompts to help you get started. You're welcome and encouraged to propose your own topic.

1. Empirical Studies on Real Data

Bandit algorithms are inherently about collecting data in an online fashion, which makes it difficult to benchmark them on static datasets. We recommend you get creative with making realistic environments (using synthetic data is discouraged). Sometimes enormous bandit-collected datasets are available that you can "replay" by fitting distributions to the data and assume those distributions are accurate (for example, https://nextml.github.io/caption-contest-data/). If you want to test contextual bandits algorithms, one popular thing to do is convert a multi-class classification dataset into a fully-observed contextual bandit dataset (see https://arxiv.org/abs/1802.04064 for an example). Alternatively, you can take a partially collected dataset like the Movielens dataset (https://grouplens.org/datasets/movielens/100k/) of users rating movies, use matrix completion to decompose the partially completed matrix into M = UV^T where each row of U represents a d-dimensional embedding of a user, and each row of V represents a d-dimensional embedding of a movie. Then use V as the feature vectors that you can recommend, and U as the unknown user vectors of users. If you're really ambitious, you could even try finetuning a weak LLM using evaluations from a strong open-source LLM (this is a contextual bandit problem). Some additional project ideas are below.

- Apply contextual bandit methods to personalized recommendation or ranking tasks. For example, use matrix factorization on Movielens data to simulate user preferences.
- Analyze medical trial data or A/B testing logs with bandit algorithms. Evaluate the potential impact of adaptive sampling.
- Implement off-policy evaluation methods and test how reliably different estimators perform under distributional shift.
- Use contextual bandits to fine-tune prompts or sampling strategies for LLMs (e.g., using GPT-generated data as feedback).
- There are a variety of ways to model hyperparameter tuning as a bandit problem. Review the paradigms and try some out.

2. Literature Reviews of Open Problems

- **Best of Both Worlds:** Survey algorithms that unify stochastic and adversarial bandit settings. What theoretical and practical challenges remain?
- **Bandits with Constraints:** Look at bandits with budget, fairness, or safety constraints. Which models are tractable? Where is the theory lacking?
- **Multi-agent Bandits:** Explore how bandit ideas extend to competitive or cooperative multi-agent settings. Can simple dynamics lead to equilibrium?

- **Causal Bandits:** What happens when action-outcome relationships follow a known (or learnable) causal graph? Summarize progress and identify unexplored frontiers.
- **Semi-bandit feedback:** Suppose a linear bandit problem where the arms $\mathcal{X} \subset \{0,1\}^d$. If $x \in \mathcal{X}$ is player, semi-bandit feedback gets reward $\langle \theta_t, x \rangle$ and observes $[\theta_t]_i$ for all i such that $x_i = 1$.
- **Infinite armed bandits:** Suppose instead of having a finite number of arms, the means of arms are drawn IID from some distribution, and you can draw as many arms as you want.
- **Options pricing:** In class we studied universal portfolio optimization for stocks. How do we model options pricing as an online learning problem? Some inspiration: `https://dl.acm.org/doi/10.1145/1132516.1132586`
- **Logistic bandits:** In this class we exclusively looked at additive noise. What if you observe binary values drawn according to a logistic model?
- **Best arm identification:** We mostly focused on regret minimization in this class, but just as important is optimization, or best arm identification. What's known for the fixed confidence and fixed budget settings?
- Some additional ideas from a previous offering of this course `https://courses.cs.washington.edu/courses/cse541/24sp/resources/project.pdf`.

Final Notes

- The best projects often connect to something you're already interested in. Don't hesitate to explore intersections with your research.
- All projects should involve some amount of implementation, theoretical analysis review—avoid purely conceptual proposals.
- You're encouraged to discuss your ideas with the instructor through the proposal for feedback and feasibility.