# Rohan Deb

+1 447-902-1605 | rd22@illinois.edu Homepage | Google Scholar | Linkedin

Interests: Reinforcement Learning, Bandits, Optimization, Active Learning, Deep Learning

## **EDUCATION**

Doctorate of Philosophy | Major: Computer Science Graduate Minor in Statistics (ongoing) University of Illinois, Urbana-Champaign Advisor: Arindam Banerjee GPA: 3.97/4

Master of Technology | *Major: Computer Science* Indian Institute of Science, Bangalore Advisor: Shalabh Bhatnagar GPA: 9.5/10

**Bachelor of Technology** | *Major: Computer Science* National Institute of Technology, Silchar GPA: 9.38/10

#### WORK EXPERIENCE

Applied Science Intern Amazon, SCOT-RL Advisor: Dean Foster

Applied Science Intern Amazon, AWS-AI Lab Advisor: Branislav Kveton

**Research Assistant** University of Illinois, Urbana-Champaign Advisor: Arindam Banerjee

**Project Associate** Indian Institute of Science, Bangalore Advisor: Gugan Thoppe

Research Intern Indian Institute of Technology, Madras Advisor: Kamakoti Veezhinathan

# PUBLICATIONS/PRE-PRINTS (\*EQUAL CONTRIBUTION)

- **Thompson Sampling for Constrained Bandits.** Rohan Deb, Mohammad Ghavamzadeh, Arindam Banerjee Accepted at The second Reinforcement Learning Conference (*RLC*), 2025 | openreview
- FisherSFT: Data-Efficient Supervised Fine-Tuning of Language Models Using Information Gain. Rohan Deb, Kiran Thekumparampil, Kousha Kalantari, Gaurush Hiranandani, Shoham Sabach, Branislav Kveton Accepted Submission at Forty-Second International Conference on Machine Learning (ICML), 2025 | arxiv
- Conservative Contextual Bandits: Beyond Linear Representations. Rohan Deb, Mohammad Ghavamzadeh, Arindam Banerjee Accepted Submission at 13th International Conference on Learning Representations (ICLR), 2025 | openreview
- **Contextual Bandits with Online Neural Regression.** Rohan Deb, Yikun Ban, Shiliang Zuo, Jingrui He, Arindam Banerjee Accepted at 12th International Conference on Learning Representations (ICLR), 2024 | arxiv | openreview
- Think Before You Duel: Understanding Complexities of Preference Learning under Constrained Resources. Rohan Deb, Aadirupa Saha, Arindam Banerjee Accepted at 27th International Conference on Artificial Intelligence and Statistics (AISTATS), 2024 | arxiv

Aug 2022 - now

Urbana-Champaign, IL, USA

Aug. 2019 – May 2021 Bangalore, KA, India

Aug. 2015 – May 2019 Silchar, India

May 2025 – Aug 2025 New York, USA

May 2024 – Aug 2024 San Jose, CA,USA

Aug 2022 – July 2023 Urbana-Champaign, IL, USA

> Aug 2021 – July 2022 Bangalore, KA, India

May 2017 – July 2017 Chennai, TN, India

You Duel: Understanding Complexities of Preference Learning under Constrained Resour

- Gradient Temporal Difference with Momentum: Stability and Convergence. Rohan Deb, Shalabh Bhatnagar Accepted at 36th AAAI Conference on Artificial Intelligence, 2022 | arxiv | AAAI
- Does Momentum Help in Stochastic Optimization? A sample complexity Analysis. Swetha Ganesh\*, Rohan Deb\*, Gugan Thoppe, Amarjit Buddhiraja Accepted at 39th Conference on Uncertainty in Artificial Intelligence (UAI), 2023 | UAI | arxiv
- Schedule Based Temporal Difference Algorithms. Rohan Deb\*, Meet Gandhi\*, Shalabh Bhatnagar Accepted at 58th Annual Allerton Conference on Communication, Control, and Computing, 2022 | IEEE | arxiv
- *N*-**Timescale Stochastic Approximation: Stability and Convergence.** Rohan Deb, Shalabh Bhatnagar *Under Submission at Stochastic Systems* | arxiv

# CURRENT RESEARCH PROJECTS

# **Optimal Design for Large Language Models**

Collaborators: Branislav Kveton, Kiran Koshy Thekumparampil

- We try to improve the statistical efficiency of fine tuning by selecting an informative subset of training examples. The key idea in our method is to select examples that maximize the Hessian of the log-likelihood of the LLM. We approximate it efficiently based on making a connection to uncertainty modeling in multinomial logistic regression models.
- Our approach is computationally efficient, analyzable, and performs well empirically. We demonstrate this on several problems, and back our claims with both quantitative results and an LLM evaluation.

# Black-box Metric Optimization for Pre-trained Models

Collaborators: Gaurush Hiranandani

- Given a pre-trained model and access to a black-box metric, our objective is to post-shift the model to optimize the given metric. Such models find applications in domain adaptation, fairness, distribution shift etc.
- Unlike previous work, we focus on cases where the metric could be an arbitrary function of the entries of the confusion matrix. Further we also plan on extending the setup to auto-regressive tasks with a focus on language models.

#### Neural Contextual Bandits with Random Sketching

Collaborators: Arindam Banerjee

• Existing bandit algorithms with neural networks are computationally inefficient and incur large regret owing to a huge number of parameters. We are focusing on using random sketching along with properties of hessian of the neural networks to obtain tighter regret guarantees along with faster algorithms.

#### Safety in Reinforcement Learning and Contextual Bandits

Collaborators: Mohammad Ghavamzadeh, Arindam Banerjee

- We study bandit and reinforcement learning problems under different safety models. For the conservative bandit setup where the agent needs to maintain performance with respect to a baseline we prove sub-linear regret for general reward functions. We also provide first order regret bounds that significantly improve the performance.
- Subsequently we are focusing on other safety setups such as stage-wise constraints and knapsack budgeted constraints for general reinforcement learning problems.

## TEACHING EXPERIENCE

Deep Generative Models, Teaching AssistantJan 2025 – May 2025University of Illinois, Urbana-ChampaignUrbana-Champaign, IL, USAReinforcement Learning, Teaching AssistantJan 2024 – May 2024University of Illinois, Urbana-ChampaignUrbana-Champaign, IL, USAIntroduction to Data Mining, Teaching AssistantAug 2023 – Dec 2023University of Illinois, Urbana-ChampaignUrbana-Champaign, IL, USAReinforcement Learning, Teaching AssistantAug 2023 – Dec 2023University of Illinois, Urbana-ChampaignUrbana-Champaign, IL, USAReinforcement Learning, Teaching AssistantIan 2022 – Apr 2022

**Reinforcement Learning**, Teaching Assistant Indian Institute of Science, Bangalore Jan 2022 – Apr 2022 Bangalore, KA, India

<b>Measure Theoretic Probability</b> , Teaching Assistant	Jan 2022 – Apr 2022
Indian Institute of Science, Bangalore	Bangalore, KA, India
<b>Topics in Stochastic Approximation Algorithms</b> , Teaching Assistant	Aug 2021 – Dec 2021
Indian Institute of Science, Bangalore	Bangalore, KA, India
Linear Algebra and Probability, Teaching Assistant	Aug 2021 – Dec 2021
Indian Institute of Science, Bangalore	Bangalore, KA, India
Machine Learning, Instructor	Feb 2022 - July 2022
Innomatics Research Labs	Hyderabad, TL, India
Introduction to Data Science, Instructor	May 2021 - Aug2021
Technology for all	Hyderabad, TL, India
HONORS AND AWARDS	

Computer Society of India Medal for Best Masters Student in Computer Science Indian Institute of Science, Bangalore	2022
<b>Undergraduate Medal for highest GPA in Computer Science</b> National Institute of Technology, Silchar	2020
<b>All India Computer Science rank 52</b> GATE (Graduate Aptitude Test in Engineering)	2019
Summer Research Fellowship Programme Indian Academy of Sciences	2017
Letter of appreciation for outstanding performance in High School exam. Ministry of Education	2015

## SELECTED COURSE WORK

Deep Generative and Dynamic models, Statistical Reinforcement Learning, Online Learning and Bandits, Deep Learning Theory, Stochastic Processes, Queuing Theory, Stochastic Approximation Algorithms, Machine Learning, Statistical Learning Theory, Pattern Recognition, Introduction to Robotics, Stochastic Calculus, High Dimensional Probability, Game Theory, Optimal Control

#### **PROFESSIONAL SERVICE**

- Co-Organizer, UIUC Machine Learning Seminar (CS 591 MLR)
- Organizer, Reading Group: Optimal Transport, Spring 2024, UIUC (Link)
- Organizer, Reading Group: Reinforcement Learning Theory, Winter 2023, UIUC (Link)
- Organizer, Reading Group: High Dimensional Probability, Fall 2023, UIUC (Link)
- Program Committee Member, 13th International Conference on Learning Representations, 2025
- Reviewer, 27th International Conference on Artificial Intelligence and Statistics (AISTATS), 2024, 2025
- Program Committee Member, 37th AAAI Conference on Artificial Intelligence, 2023, 2024, 2025
- Reviewer, European Control Conference (ECC), 2024
- Reviewer, IEEE Transactions on Automatic Control
- Reviewer, IEEE Control Systems Letters