Debangshu Banerjee

Email: bdebangshu@iisc.ac.in
LinkedIn: debangshub
GitHub: Debangshu93
Google Scholar ID : FXkrG54AAAAJ

About Me

I am currently pursuing my doctoral dissertation under the guidance of Dr. Aditya Gopalan at the Indian Institute of Science. I am interested in the intersection of theory and practice of data-driven automated learning models in the field of AGI. Currently I am working on the allignment problem of LLMs with human preferences. At the begining of my dissertation program I had delved in depth in the field of structured and unstructured bandit problems and Markov Decision Processes. Before this I had done my Masters specializing in reinforcement learning and had done a stint at an Bengaluru based AI start-up.

Skills

Programming Languages : python, C

Deep Learning Frameworks : Pytorch, Tensorflow

LLM Frameworks : HuggingFace-Transformers

RL Frameworks : HuggingFace-TRL, RLLib

Distributed Systems : Accelerate

Education

PhD, Indian Institute of Science, Bangalore, India 2019,-Present
Advisor : Aditya Gopalan
- Department Of Electrical Communications and Engineering

M Tech, Indian Institute of Science, Bangalore, India 2016-2018
Advisor : Shalabh hatnagar
- Department Of Computer Science and Automation
- Thesis : Hex and Neurodynamic Programming *(https://arxiv.org/pdf/2008.06359.pdf)*

B Tech, Jadavpur University, Kolkata, India 2012-2016
Advisor : Ujjaini Sarkar
- Department Of Chemical Engineering

Industry Experience

Working Experience

AI Engineer, AlphaICs Corp., Bengaluru 2018-2019

- Inference on Proprietary Edge Device : Worked with the AI and Compiler teams to implement *native versions of deep learning models* related to vision, including **INCEPTION-V3**, **RESNET-50** and **VGG-NET** for *proprietary 16-bit integer edge* device processor (**RAP-E**). I was entrusted with training quantized versions of these networks to demonstrate inference for securing funding. Final versions had a test accuracy of 90%.
  Programming Platforms : **C**, **Python**, **TensorFlow**
- Tensorflow XLA-Framework library developement : Developed TeensorFlow XLA Libraries and Optimization for proprietary processor **RAP-C**. Software development and testing was done by working with AI and Compiler Teams. The results were improvements in speed and memory usage for most internal benchmarks. Developed Runtime Libraries and Optimization for RAP-Processor.
  Programming Platforms : **Assembly**, **C**, **TensorFlow**

Internship Experience

Research Intern, IBM Research Lab., Bengaluru Summer 2023

- Reinforcement Learning for Constrained Process Modelling : Use Reinforcement Learning Techniques to solve optimal control problems in the industrial production pipeline. Specifically, we explored novel algorithmic and theoretical ideas that could mitigate typical problems in such settings, such as the presence of limited real data and the absence of a reliable simulator (hence offline).
  Programming Platforms : **Pytorch**

Conference Publications

[1] **D. Banerjee**, A. Ghosh, S. Ray Chowdhury, and A. Gopalan, "Exploration in linear bandits with rich action sets and its implications for inference", in *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, F. Ruiz, J. Dy, and J.-W. van de Meent, Eds., ser. Proceedings of Machine Learning Research, vol. 206, PMLR, 25–27 Apr 2023, pp. 8233–8262.

Technical Reports

1. **D. Banerjee** and A. Gopalan, "Bad Values but Good Behavior: Learning Highly Misspecified Bandits and MDPs", 2023. (Currently under review)
   Summary Most practical problems which can be represented as a bandit learning problem inherently assume a model. Previous works have shown that when true bandits deviate from the assumed model structure, popular algorithms like LinUCB and ϵ-greedy can suffer a linear regret in the worst case. In this study, we systematically characterize a class of bandit problems which, even though misspecified, will suffer no extra linear regret. This also opens the possibility to designing intelligent features with low model complexity.
2. **D. Banerjee** and A. Gopalan, "On the minimax regret for linear bandits in a wide variety of action spaces", *arXiv preprint arXiv:2301.03597*, 2023.
   Summary Minimax regret in bandits is an important field of study in Bandit Literature. Most results focus on either the standard hypercube or the standard sphere as the decision set. In this study, we show a technique to expand the analysis to any L^p ball, where $1 \leq p < \infty$.
3. **D. Banerjee**, "Hex and Neurodynammic Programming", *arXiv preprint arXiv:2008.06359,*, 2017.
   Summary The projected Bellman error (PBE) has been shown to be a more stable objective function to optimize in reinforcement learning literature. This study extends the standard PBE error minimization algorithms like GTD and TDC to include multi-step look aheads. We compare and contrast various hyperparameters and standard RL algorithms in the game of HEX.
   Programming Platform : **Theano**

Projects

1. *Towards Reliable Alignment: Uncertainty Aware RLHF*, (**ongoing**)
   Summary How reliable are learned reward models? As our experiments show, they might not be so reliable as to use them for aligning LLMs. We propose an uncertainty-aware methodology for fine-tuning LLMs, which enjoy *low-risk low-reward* properties compared to existing methods.
   Programming Platform : **Pytorch**, **HuggingFace-Transformers**, **HuggingFace-TRL**
2. *Deep Reinforcement Learning based on Human Feedback*, 2023
   Worked in a group to construct demos for the tutorial "Do you prefer learning from preferences" at Neurips 2023.
   Programming Platform : **Pytorch**
3. *Deep Reinforcement Learning based on Human Feedback*, 2023
   Worked in a group to construct demos for the tutorial "Do you prefer learning from preferences" at Neurips 2023.
   Programming Platform : **Pytorch**

4. *Deep Reinforcement Learning Tutorial*, 2022
   Conducted tutorials on concepts and implementation of deep RL techniques like DQN, TRPO, PPO, SAC, DDPG for students.
   Programming Platform : **Pytorch**

5. *On Randomized Least Squares Value Iteration for Misspecified MDPs*, 2021
   Summary Thompson Sampling is known for its ease of computability in the Bandit Literature. RLSVI is a natural extension of Thompson Sampling in MDPs. However, the study of RLSVI has mostly been restricted to finite state-action space domains in RL. Most attempts to analyze parameterized versions of RLSVI involves unrealistic model assumptions. In this work we analyze RLSVI without any model assumption. We recover the expected $\sqrt{T}$ term along with the usual linear term associated with model misspecification.

6. How Reliable are Test Numbers for Revealing the COVID-19 Ground Truth, 2021
   Data visualization for COVID project (Data Handling).See https://arxiv.org/pdf/2004.12782.pdf

Academic Experience

Research Experience

- **Stochastic Systems Lab, IISc** 2017-2018
  Advisor : Shalabh Bhatnagar
- **Learning Optimization Control and Automation (LOCA), IISc** 2019-present
  Advisor : Aditya Gopalan

Teaching

- **Teaching Assistant** at NPTEL Fall 2022, 2021
  Concentration Inequalities
- **Teaching Assistant** at Indian Institute of Science Fall 2023, 2022
  Random Processes (E2 202o)
- **Teaching Assistant** at Indian Instite of Science Fall 2020
  Online Prediction and Learning (E1 245)

Talks/Presentations

- **EECS Students Symposium, IISc, Bengaluru** 2024
  Student Speaker
- **AISTATS, Valencia, Spain** 2023
  Poster Presentation
- **IBM Day@IISC, Bengaluru, India** 2022
  Student Speaker
- **IISc WorkShop on Deep Reinforcement Learning, Bengaluru, India** 2021
  Student Speaker https://sml.csa.iisc.ac.in/drlworkshop/

Extracurricular Activities

- Technical Program Committee 2024
  SPCOM, 2024
- Web and Publication Team 2024
  SPCOM, 2024
- Student Volunteer 2023
  COLT, 2023
- Student Volunteer 2022

EECS Symposium 2022, 2023

- Webmaster at https://ai.iisc.ac.in/ 2019-2021
  General Website maintenace and content update
- Reviewer for various conferences including
  Neurips, ICML, ICLR, AAAI, AISTATS

Technical Courses

- **Analysis**
  - Real Analysis
  - Complex Analysis
  - Measure Theory
  - Functional Analysis
  - Distribution Theory
- **Probability**
  - Probability Theory
  - High Dimensional Statistics
  - Detection and Estimation
  - Concentration Inequalities
  - Stochastic Processes and Queuing Theory
- **Optimization**
  - Computational Methods of Optimization
  - Convex Optimization
  - Linear and Non Linear Optimization
- **Geometry**
  - Multivariable Calculus
  - Calculus on Manifolds
  - Differentiable Manifolds
- **Other Mathematical Topics**
  - Information Theory
  - Game Theory
  - Matrix Theory
  - Discrete Mathematics
- **Computer Science**
  - Data Structures
  - Algorithms
  - High Performance Computing
- **ML/AI**
  - Machine Learning
  - Reinforcement Learning
  - Topics in AI
  - Practical Data Science
  - Advanced Robotics Learnning and Control
  - Online Prediction and Learning