

Safal Shrestha
safal.shrestha@nyu.edu
<https://safal312.github.io/>

Education

2020–24 **B.Sc.**, Computer Science, New York University Abu Dhabi
Minor: Applied Mathematics; Interactive Media
GPA: 3.929 – Latin Honors: magna cum laude

Research Interests

LLM Reasoning, Reinforcement Learning, Generalization, LLM Interpretability

Research Experience

- 2024– **Research Assistant**, New York University Abu Dhabi, (Advisor: [Dr. Keith Ross](#))
- Conducting research on reinforcement learning and reasoning in large language models (LLMs).
 - Collaborating with other research assistants and student teams on multiple projects; contributing to four papers.
 - Delivered guest sessions in Machine Learning course and performed extensive literature reviews to support ongoing research initiatives.
- 2024–24 **Post-Graduate Research Trainee**, New York University Abu Dhabi, (Advisor: [Dr. Mario D. Molina](#))
- Collected and organized large-scale interaction data from the StackExchange API to support research on online community behavior.
 - Investigated patterns of cooperative dynamics across stackexchange communities.
 - Reviewed relevant literature on cooperative behavior in online communities and network modeling.
- 2023–24 **Student Research Assistant**, New York University Abu Dhabi, (Advisor: [Dr. Minsu Park](#))
- Conducted research on trends in character inclusion and interaction dynamics in Hollywood films.
 - Developed a parallelized Selenium-wire scraper to collect 4,000+ data points from Amazon Prime Video.
 - Executed a comprehensive literature review on social dynamics in media and inclusion studies.

Publications

 [Google Scholar](#)

† → Equal contribution

Working papers (*Under Review*)

1. Aadim Nepal, **Safal Shrestha**, Anubhav Shrestha, Minwu Kim & Keith Ross. *Layer Importance for Mathematical Reasoning is Forged in Pre-Training and Invariant after Post-Training* 2025. arXiv: [2506.22638 \[cs.LG\]](https://arxiv.org/abs/2506.22638). <https://arxiv.org/abs/2506.22638>.
2. Minwu Kim[†], Anubhav Shrestha[†], **Safal Shrestha**, Aadim Nepal & Keith Ross. *Reinforcement Learning vs. Distillation: Understanding Accuracy and Capability in LLM Reasoning* 2025. arXiv: [2505.14216 \[cs.AI\]](https://arxiv.org/abs/2505.14216). <https://arxiv.org/abs/2505.14216>.
3. **Safal Shrestha**, Minwu Kim, Aadim Nepal, Anubhav Shrestha & Keith Ross. *Warm Up Before You Train: Unlocking General Reasoning in Resource-Constrained Settings* 2025. arXiv: [2505.13718 \[cs.AI\]](https://arxiv.org/abs/2505.13718). <https://arxiv.org/abs/2505.13718>.

Preprint

1. **Safal Shrestha**[†], Minwu Kim[†] & Keith Ross. *Mathematical Reasoning in Large Language Models: Assessing Logical and Arithmetic Errors across Wide Numerical Ranges* 2025. arXiv: [2502.08680 \[cs.LG\]](https://arxiv.org/abs/2502.08680). <https://arxiv.org/abs/2502.08680>.

Relevant Coursework

Reinforcement Learning, Machine Learning, Processing Big Data for Analytics, Agile Software Development, Computer Networks, Algorithms, Multivariable Calculus, Linear Algebra, Probability and Statistics, Computational Social Science, Research Design and Causality

Awards & Honors

2024	Founder's Day Award of Excellence
2023	Amazon MENA Transcend Challenge, Top-5
2019	Young Scientist Award, Nepal Youth Science Summit

Other Experience

Dec 2023 – Feb 2024	Data Engineer Intern , Liwa Capital Advisors, Abu Dhabi, UAE <ul style="list-style-type: none">– Created automated solutions using web scraping, Excel, and Power Automate, cutting report generation time by 98%.– Designed Power Automate workflow for news extraction from Bloomberg Terminal, enhancing decision-making.– Leveraged Excel automation with VBA macros to streamline data processing, resulting in improved efficiency.
May 2023 – Aug 2023	AI Engineer Intern , Digital Energy, Dubai, UAE <ul style="list-style-type: none">– Engineered an efficient Python web scraper to collect 10,000+ documents.– Processed documents to establish an effective domain-specific knowledge base.

– Built and deployed a modern chatbot UI and a FastAPI server to serve open-source Large Language Models (LLMs).

Activities

Feb 2021 – President, ADvocacy, Abu Dhabi, UAE
May 2024

– Led a team of 10+ people to start and manage initiatives for people of concern, migrant workers on campus and in the Filipino Embassy by organizing recreational events and language learning programs.

Feb 2021 – Social Media Coordinator, Nepali Student Association, Abu Dhabi, UAE
May 2023

– Organized Nepali cultural events and engaged the campus community with Nepali tradition and festivals, documented events, and published them on social media.

Skills

Coding	Python, JavaScript, R, C/C++
Software	L ^A T _E X, Git, RStudio, Figma, Adobe Creative Suite, Microsoft Office Suite
Languages	English (fluent), Hindi (fluent), Nepali (fluent)
