

Azal Ahmad Khan

✉ khan1069@umn.edu  [azalahmadkhan.github.io](https://github.com/azalahmadkhan)  [Google Scholar](#)  +1(763)485-1561

Education

University of Minnesota, Twin-Cities 09/2024 – Present

Ph.D., Computer Science and Engineering (GPA: 3.95/4.0)

‣ Advisor: [Ali Anwar](#)

‣ Amazon ML Systems Fellow (2025) & GAGE Fellow (2024)

Indian Institute of Technology (IIT), Guwahati

11/2020 – 05/2024

B.Tech., Chemical Science and Technology

Research Experience

NEC Laboratories America, San Jose, CA

05/2025 – 08/2025

Research Intern with [Vijay Kumar](#) & [Manmohan Chandrakar](#)

University of Minnesota, Twin-Cities, MN

09/2024 – 12/2024

Graduate Assistant with [Ali Anwar](#)

University of Minnesota.

01/2022 – 05/2024

Research Intern with [Ali Anwar](#)

University of New South Wales.

01/2022 – 12/2023

Research Intern with [Rohitash Chandra](#)

Awards & Grants

2025 [Amazon Machine Learning System Fellowship.](#)

2024 [University of Minnesota GAGE Fellowship.](#)

2024 Travel Grant, [Google Research Week.](#)

2024 Highly Commended (**Top 10%**), [Global Undergraduate Research Award 2024.](#)

Publications & Pre-prints

* Equal Contribution

arXiv [Retrieval-of-Thought: Efficient Reasoning via Reusing Thoughts](#)

Ammar Ahmed*, **Azal Ahmad Khan***, Ayaan Ahmad, Sheng Di, Zirui Liu, Ali Anwar.
Pre-print.

NeurIPS '25 [Beyond Expectations: Quantile-Guided Alignment for Risk-Calibrated Language Models](#)

Xinran Wang, Jin Du, **Azal Ahmad Khan**, Qi Le, Enmao Diao, Jiawei Zhou, Jie Ding, Ali Anwar.

*The 39th Conference on Neural Information Processing Systems, 2025. **Spotlight (Top 3.1%)***

EMNLP '25 [Accelerating LLM Reasoning via Early Rejection with Partial Reward Modeling](#)

Seyyed Saeid Cheshmi*, **Azal Ahmad Khan***, Xinran Wang, Zirui Liu, Ali Anwar.
In Proceedings of Findings of EMNLP 2025.

arXiv [Sem-DPO: Mitigating Semantic Inconsistency in Preference Optimization for Prompt Engineering](#)

Anas Mohamed*, **Azal Ahmad Khan***, Xinran Wang, Ahmad Faraz Khan, Shuwen Ge, Saman Bahzad Khan, Ayaan Ahmad, Ali Anwar.
Pre-print.

- IROS '25 **Safety Aware Task Planning via Large Language Models in Robotics**
Azal Ahmad Khan*, Michael Andrev*, Muhammad Ali Murtaza, Sergio Aguilera, Rui Zhang, Jie Ding, Seth Hutchinson, Ali Anwar.
In the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2025.
- arXiv **LADs: Leveraging LLMs for AI-Driven DevOps**
 Ahmad Faraz Khan*, **Azal Ahmad Khan***, Anas Mohamed, Haider Ali, Suchithra Moolinti, Sabaat Haroon, Usman Tahir, Mattia Fazzini, Ali R. Butt, Ali Anwar.
Pre-print.
- IPDPS '25 **IP-FL: Incentivized and Personalized Federated Learning**
 Ahmad Faraz Khan, Xinran Wang, Qi Le, Zain ul Abdeen, **Azal Ahmad Khan**, Haider Ali, Ming Jin, Jie Ding, Ali Butt, Ali Anwar.
In Proceedings of the 39th IEEE International Parallel & Distributed Processing Symposium.
- EuroSys '24 **FLOAT: Federated Learning Optimizations with Automated Tuning**
 Ahmad Faraz Khan, **Azal Ahmad Khan**, Ahmed M Abdelmoniem, Samuel Fountain, Ali Butt, Ali Anwar.
In Proceedings of the Nineteenth European Conference on Computer Systems.
- BigData '24 **Personalized Federated Learning Techniques: Empirical Analysis**
Azal Ahmad Khan, Ahmad Faraz Khan, Haider Ali, Ali Anwar.
In IEEE International Conference on Big Data 2024 (Short Paper).
- BigData '24 **Mitigating Sycophancy in Large Language Models via Direct Preference Optimization**
Azal Ahmad Khan, Sayan Alam, Xinran Wang, Ahmad Faraz Khan, Debanga Raj Neog, and Ali Anwar.
In IEEE International Conference on Big Data 2024 (Short Paper).
- Elsevier **A review of ensemble learning and data augmentation models for class imbalanced problems: combination, implementation and evaluation**
Azal Ahmad Khan, Omkar Chaudhari, Rohitash Chandra.
Expert Systems with Applications, Volume 244, 122778.
- MDPI **A Quantum-Inspired Predator-Prey Algorithm for Real-Parameter Optimization**
Azal Ahmad Khan, Salman Hussain, Rohitash Chandra.
Algorithms 2024, 17(1), 33.

Technical Skills

Programming Python, C/C++ **Web Technologies** FastAPI
LLM Libraries and Frameworks LangChain, Ollama, OpenAI, AutoGen, Hugging Face, Agno
ML Tools Pandas, Numpy, Scipy, Matplotlib, Scikit-learn, Tensorflow, Pytorch, OpenCV
Miscellaneous Git, GitHub, SQL, L^AT_EX