# Achintya Kundu

achi.kundu@gmail.com | +91-9343653615 | Website | LinkedIn

## Current Position

Principal Machine Learning Engineer, HP Inc.

## Areas of Expertise

Large Language Models (LLMs), Machine Learning, Optimization, Deep Learning, Speech Processing

## Work Experience

- **Principal ML Engineer** [January 2025 – till date]
  HP AI Lab, Bangalore, India.
  Research Project: Faster On-device LLM Inference for Agentic AI
- **Staff Research Scientist** [December 2021 – October 2024]
  IBM Research, Singapore.
  Research Projects: Enhancing Training Efficiency of LLMs
  Efficiently Distilling LLMs for Edge Applications
  AI Model Optimization for Edge
- **Research Scientist** [August 2021 – December 2021]
  IBM Research, Bangalore, India.
  Research Project: Robust and Personalized Federated Learning
- **Senior Research Fellow** [March 2017 – July 2019]
  DST/INRIA sponsored project, Indian Institute of Science, Bangalore.
  Research Project: First order methods for Kernels and Submodular functions
- **Software Design Engineer** [August 2008 – July 2009]
  Texas Instruments, Bangalore, India.
  Project: HD Video Compression on Smartphone

## Education

- **Doctor of Philosophy**, 2022 **[CGPA: 7.5/8.0]**
  Department of Computer Science and Automation (CSA),
  Indian Institute of Science (IISc), Bangalore, India.
  Thesis: Novel First-order Algorithms for Non-smooth Optimization Problems in Machine Learning [Thesis Link]
  Research Supervisor: Prof. Chiranjib Bhattacharyya
- **Master of Engineering**, 2008 **[CGPA: 7.8/8.0]**
  Department of Electrical Communication Engineering (ECE),
  Indian Institute of Science (IISc), Bangalore, India.
  Project: Speech Enhancement - A Bayesian Estimation Approach using GMM
- **Bachelor of Engineering**, 2002 **[CGPA: 9.4/10]**
  Department of Electronics & Tele-communication Engineering (ETCE),
  Jadavpur University (JU), Kolkata, India.

## Awards

- Received First Patent Application Achievement Award from IBM in 2023.
- Received the IBM IRL Distinguished Paper Award 2022.
- Won the **Best Paper Award** at IEEE EDGE 2022.
- Received Honorable Mention Award in Yahoo! Key Scientific Challenges, 2011.
- Won the Best Perspective Seminar Award 2010-11 in the Dept. of CSA, IISc.

## Research Publications [Google Scholar]

- **A. Kundu**, R. D. Lee, *et*.al., "Enhancing Training Efficiency Using Packing with Flash Attention", *arXiv:2407.09105*, 2024. [Hugging Face Link]
- **A. Kundu**, F. Lim, *et*.al., "Efficiently Distilling LLMs for Edge Applications", *Conference of the North American Chapter of the Association for Computational Linguistics* **(NAACL)** , 2024, Vol. 6: Industry Track.
- **A. Kundu**, L. Wynter, *et*.al., "Transfer-Once-For-All: AI Model Optimization for Edge", *IEEE Intl. Conf. on Edge Computing and Communications* **(IEEE EDGE)**, 2023.
- **A. Kundu**, P. Yu, *et*.al., "Robustness and Personalization in Federated Learning: A Unified Approach via Regularization", *IEEE Intl. Conf. on Edge Computing and Communications* **(IEEE EDGE)**, 2022. **Best Paper Award**
- **A. Kundu**, F. Bach, *et*.al., "Convex Optimization over Intersection of Simple Sets: improved Convergence Rate Guarantees via an Exact Penalty Approach", *Intl. Conf. on Artificial Intelligence and Statistics* **(AISTATS)**, 2018.
- P. Pipada, **A. Kundu**, *et*.al., "LoadIQ: Learning to Identify Workload Phases from a Live Storage Trace", *USENIX Workshop on Hot Topics in Storage and File Systems* **(HotStorage)**, 2012.
- **A. Kundu**, V. Tankasali, *et*.al., "Efficient algorithms for learning kernels from multiple similarity matrices with general convex loss functions", *Neural Information Processing Systems* **(NeurIPS)**, 2010.
- **A. Kundu**, S. Chatterjee, *et*.al., "GMM Based Bayesian Approach to Speech Enhancement in Signal / Transform Domain", *Intl. Conf. on Acoustic, Speech, and Signal Processing* **(ICASSP)**, 2008.
- **A. Kundu**, S. Chatterjee, *et*.al., "Speech Enhancement Using Intra-frame Dependency in DCT Domain", *European Signal Processing Conference* **(EUSIPCO)**, 2008.
- **A. Kundu**, S. Chatterjee, *et*.al., "Subspace Based Speech Enhancement Using Gaussian Mixture Model", **InterSpeech**, 2008.

## Academic Achievements

- Received the **Gold Medal** for best Master of Engineering student (2007-08) in the Department of ECE, Indian Institute of Science, Bangalore, India.
- Received **Senate Commendation** at Indian Institute of Science, Bangalore for outstanding academic performance during Master of Engineering, 2006-08.
- Secured **6**$^{\text{th}}$ rank (Electronics & Communication branch) in Graduate Aptitude Test in Engineering (GATE), 2006.
- Achieved **17**$^{\text{th}}$ rank in West Bengal Joint Entrance Examination (an entrance test for undergraduate study in Engineering), 2002.

## Research Internships

- IBM Research, Bangalore, February – May 2021
  Research Project: Robust Federated Learning
- INRIA, Paris, June – August 2018
  Research Project: Non-smooth Optimization with Linear Minimization Oracle
- INRIA, Paris, September – December 2017
  Research Project: Convex Optimization over Intersection of Simple Sets
- Xerox Research Centre, Bangalore, July – October 2015
  Research Project: Learning Robust Predictors under Uncertainty using Copulas
- Amazon (Machine Learning Team), Bangalore, May – July 2014
  Research Project: Nonlinear Feature Transformation for Large-scale Classification
- Yahoo! Labs, Bangalore, May 2010 – July 2010
  Research Project: Forecasting Unique User Counts (Reach) for Advertisements

## Software Skills

- Deep Learning Framework: PyTorch
- Programming Language: Python, C
- Operating System: Linux