Soumen Kumar Mondal
Center for Machine Intelligence and Data Science (CMInDS)
Indian Institute of Technology Bombay, India
Specialization: **Data Science and Artificial Intelligence**
Email-ID : *mondalsoumen00@gmail.com* or *soumenkm@iitb.ac.in*

Male
M.S. by Research
DOB: 12/08/1996
Mobile: +91 8759240557
Website: https://soumenkm.github.io

Qualification	Specialization	Institute (all located in India)	Year	CPI/% (R)
MS by Research	Data Science & Artificial Intelligence	Indian Institute of Technology, Bombay	2023 - 2026	9.83 (1^{st})
MTech	Structural Engineering	Indian Institute of Science, Bangalore	2018 - 2020	9.30 (4^{th})
BTech	Civil Engineering	Jadavpur University, Kolkata	2014 - 2018	8.74 (6^{th})
12^{th} Standard	Science (PCM), Languages	Haldia High School (West Bengal Board)	2012 - 2014	89.60 (1^{st})
10^{th} Standard	Science, Arts, Languages	Haldia High School (West Bengal Board)	2011 - 2012	86.85 (2^{nd})

Technical Skills

- **Programming & Scripting Languages:** Python, C, C++, MATLAB
- **Tools and Technologies:** PyTorch, HuggingFace, LangChain, TensorFlow, Scikit-Learn, LaTeX, Git, Linux

List of Publications, Google Scholar ID: 154CUKcAAAAJ

1. **Mondal, Soumen Kumar**, Sen, S., Singhania, A., & Jyothi, P. (2025). **Language-Specific Neurons Do Not Facilitate Cross-Lingual Transfer**. *In Proceedings of the InsightsNLP in NAACL 2025 (oral)*. DOI: 10.18653/v1/2025.insights-1.6.
2. Sona, SE., **Mondal, Soumen Kumar**, Sen, S., Singhania, A., & Jyothi, P. (2025). **LoFTI: Localization and Factuality Transfer to Indian Locales**. *In Findings of the ACL 2025*. ArXiv: 2407.11833.
3. **Mondal, Soumen Kumar**, Varmora, A., Chanda, P., & Ramakrishnan, G. (2025). **FairPO: Robust Preference Optimization for Fair Multi-Label Learning**. *Under submission in NeurIPS 2025*. ArXiv: 2505.02433.

M.S. by Research in DS & AI at IIT Bombay

- **Improving Alignment and Control in Multilingual LLMs via Preference Optimization and Reinforcement Learning**
  (M.S. Thesis, Advisor: Prof. Preethi Jyothi, CSE, IIT Bombay) *(Spring 2025 - Present)*

  Developing methods to enhance alignment and steerability of multilingual LLMs, particularly for low-resource languages. Explores preference optimization (PPO, RLHF, DPO, SimPO, CPO) and reinforcement learning for advanced model editing and arithmetic capabilities. GitHub: soumenkm/ModelEditing (Currently Private)

- **Improving Downstream Task Performance in Multi-lingual LLMs by Intervening Language Specific Neurons**
  (M.S. Thesis, Advisor: Prof. Preethi Jyothi, CSE, IIT Bombay) *(Autumn 2024)*

  Investigated the functional role of language-specific neurons in multilingual LLMs concerning cross-lingual transfer. Employed LAPE and activation probability methods, leading to paper 1. GitHub: soumenkm/LangSpecificNeurons

- **IIT Bombay - Amazon Collaboration: Localizing Text Across Domains Using RARR Attribution Technique**
  (M.S. R&D Project, Advisor: Prof. Preethi Jyothi, CSE, IIT Bombay) *(Grade: 10, Spring 2024)*

  Explored techniques for generalizing text generation for localization across diverse domains using RARR attribution. Developed a benchmark for LLM localization evaluation, leading to paper 2. GitHub: soumenkm/RnD_Project

- **Cross-lingual Factual Knowledge Transfer in Multi-lingual Language Models**
  (M.S. Seminar, Advisor: Prof. Preethi Jyothi, CSE, IIT Bombay) *(Grade: 10, Spring 2024)*

  Analyzed factual knowledge representation and transfer in mBERT across languages using a probeless methodology to inspect internal model states. GitHub: soumenkm/TracingRootFacts

Machine Learning Course Projects at IIT Bombay

- **FairPO: Robust Preference Optimization for Fair Multi-Label Learning**
  (*Course Project, Optimisation for ML, Prof. Ganesh Ramakrishnan, CSE, IIT Bombay)* *(Spring 2025)*

  Proposed FairPO, a novel multi-label classification framework inspired by GRPO, to enhance fairness by partitioning labels and applying DPO/SimPO/CPO preference loss, leading to paper 3. GitHub: soumenkm/FairPO

- **Vision Transformer (ViT) Model Fine-Tuning with MillionAID Dataset using LoRA**
  (*Course Project, Advanced Deep Learning for CV, Prof. Biplab Banerjee, CSRE, IIT Bombay)* *(Autumn 2024)*

  Implemented Low-Rank Adaptation (LoRA) from scratch for efficient fine-tuning of a DINOv2-pretrained Vision Transformer (ViT) on the MillionAID remote sensing dataset. GitHub: soumenkm/IITB-GNR650-ADLCV/CodingProject

- **Learning to Classify Images under Noisy Labels using Turtle**
  (*Course Project, Advanced Deep Learning for CV, Prof. Biplab Banerjee, CSRE, IIT Bombay)* *(Autumn 2024)*

  Addressed label noise by combining a CLIP & DINO trained ViT ensemble with fine-tuning on denoised data. Achieved **88% accuracy** on CIFAR-100, despite 40% training label noise. GitHub: soumenkm/IITB-GNR650-ADLCV/Project1

- **Zero Shot Learning (ZSL) for Image Classification on AwA2 Dataset**
  (*Course Project, Advanced Deep Learning for CV, Prof. Biplab Banerjee, CSRE, IIT Bombay)* *(Autumn 2024)*

  Designed a ZSL pipeline (ViT, FastText, NN classifier with class normalization). Attained **40% test accuracy** on AwA2 with a challenging 50:50 train-test split. GitHub: soumenkm/IITB-GNR650-ADLCV/Project2

- **Fine Grained Image Classification on CUB Dataset using EfficientNet**
  (*Course Project, Deep Learning for CV, Prof. Biplab Banerjee, CSRE, IIT Bombay)* *(Spring 2024)*

  Optimized accuracy-parameter trade-off on the CUB dataset. Achieved **75% test accuracy** with an EfficientNet model (4.2M parameters). GitHub: soumenkm/GNR-638-Deep-CV/mini-project1

- **Deep Learning based System to Estimate the Calorie Content in Food from Images**
  (*Course Project, Foundations of Machine Learning, Prof. Sunita Sarawagi, CSE, IIT Bombay)* *(Autumn 2023)*

  Developed an automated calorie estimation system using YOLOv8 (detection) and GrabCut (segmentation), achieving **7.6% mean absolute error** across 19 food classes. GitHub: soumenkm/CS725-FML-Project

Machine Learning "From Scratch" Self Projects

- **Build GPT2 and BERT from Scratch**: Developed all core components of GPT-2 and BERT (multi-head self-attention, MLPs) from scratch. Developed a multi-GPU trainer for pre-training, LoRA (from scratch) fine-tuning, and instruction tuning. GitHub: soumenkm/Build-LLM-from-scratch and soumenkm/Build-BERT-from-scratch *(Autumn 2024)*
- **Build Diffusion Model (DDPM) from Scratch**: Developed a DDPM from scratch, implementing diffusion/reverse processes and sampling. Trained on CelebHQ. GitHub: soumenkm/Diffusion-Model-from-Scratch *(Autumn 2024)*
- **Build FFNN from Scratch**: Developed a Feedforward Neural Network from scratch using NumPy only, implementing forward and backpropagation and mini-batch SGD. GitHub: soumenkm/ML-Algorithms/FFNN *(Autumn 2023)*

Work Experience

- **Fujitsu Research, Bengaluru, India** *(May 2025 - July 2025)*
  AI Research Intern: Developed a proactive RCA system for hierarchical software log files using temporal & causal hypergraph and InstructRAG finetuning of State Space based Mamba model.
- **General Electric (GE Vernova), Bengaluru, India** *(Aug 2020 - July 2023)*
  System Value Optimisation Engineer: Designed several regression based ML systems to estimate the wind load in wind turbine. Developed Python modules to optimize fatigue simulation, enhancing system operational efficiency.

Courses at IIT Bombay

- CS 769: **Optimisation for Machine Learning**, Prof. Ganesh Ramakrishnan (*Grade: 10, Spring 2025*)
- GNR 602: **Advanced Satellite Image Processing**, Prof. BK Mohan (*Grade: 10, Spring 2025*)
- BB 610: **Biomedical Micro-systems**, Prof. Rohit Srivastava (Institute Elective) (*Grade: 10, Spring 2025*)
- GNR 650: **Advanced Deep Learning for Computer Vision**, Prof. Biplab Banerjee (*Grade: 10, Autumn 2024*)
- CS 601: **Algorithms and Complexity**, Prof. Akash Kumar (*Grade: 10, Autumn 2024*)
- SC 607: **Convex Optimisation**, Prof. Avishek Ghosh (*Grade: 10, Spring 2024*)
- GNR 638: **Deep Learning for Computer Vision**, Prof. Biplab Banerjee (*Grade: 10, Spring 2024*)
- CS 725: **Foundations of Machine Learning**, Prof. Sunita Sarawagi (*Grade: 10, Autumn 2023*)
- EE 635: **Applied Linear Algebra**, Prof. Dwaipayan Mukherjee (*Grade: 9, Autumn 2023*)
- IE 621: **Introduction to Probability & Stochastic Process**, Prof. KSM Rao (*Grade: 9, Autumn 2023*)

Teaching Assistant Positions at IIT Bombay

- IITB e-PG Diploma in AI: Mathematical Foundations of ML *(Spring 2025)*, CS 6106: Statistical Learning Theory *(Spring 2025)*, CS 725: Foundations of ML *(Autumn 2024)*, DS 303: Introduction to ML *(Spring 2024)*

Achievements

- Received the **Institute Academic Prize** for outstanding academic performance (Rank 1) at IIT Bombay. *(2024)*
- Received a **GE spotlight impact award** for contributions reducing business operational costs at GE Vernova. *(2022)*
- Won the **Innovate 2021 AI/ML challenge** organized by GE Vernova. *(2021)*

Hobbies

- Reading novels, Listening to music, Watching movies and TV series, Playing Cricket, Chess and Badminton.