Soham Chakraborty

M.Tech in Computer Science, ISI Kolkata

+91 8240125480 | sohamchakraborty10qw@gmail.com |

Kolkata, India

-Linkedin | Github | Portfolio

Introduction

M.Tech Computer Science student at ISI Kolkata with hands-on experience in machine learning and a data science internship at Exposys Data Labs. Built a Small Language Model (SLM) from scratch and deployed LLM apps using Hugging Face and Gradio. Certified in Python by IIT Ropar and CMI-NPTEL.

Education

Indian Statistical Institute, M. Tech in Computer Science — July 2024 – June 2026

- **Percentage :** 81% (1st sem)
- **Relevant Coursework :** Machine Learning, Image Processing, Information Retrieval, DBMS, DSA, Statistical Methods, Computer Networks

Government College of Engineering and Textile Technology, Serampore, B. Tech in Computer Science & Engineering — Dec 2020 – June 2024

- **CGPA :** 9.44/10.0
- **Coursework :** NLP, Machine Learning, Operating Systems, Data Structures, Algorithms

Internship

Data Science Intern, Exposys Data Labs — June 2023 – July 2023

- Preprocessed customer datasets and engineered features using pandas and matplotlib.
- Applied clustering analysis and visualized patterns to aid customer segmentation.

Projects

Small Language Model (SLM) from Scratch
Python + Pytorch

- Implemented a transformer-based language model from scratch to understand key components: tokenization, multi-head self-attention, causal masking, and positional encoding
- Trained on TinyStories dataset using next-token prediction; built a custom training loop with cross-entropy loss
- Acquired hands-on understanding of transformer architecture and GPT-like models.
- Explored model inference and sampling; evaluated generated outputs to understand training limitations

Autonomous Drone Navigation — Report
B.Tech Final Year Project

- Developed and evaluated ML models, primarily RNNs, LSTMs, and GRUs, to navigate drones using sensor data and computer vision.
- Compared performance against LTC and CFC models; optimized hyperparameters via TPE sampling.
- Created offline datasets from rooftop environments and tested across conditions.

Skills

Technical Skills : Transformers | Hugging Face | PyTorch | Scikit-learn | Gradio | Pandas | Git | Python | C | SQL | Java | HTML | CSS | Matplotlib | Pyserini

Operating Systems: Windows, Linux

Soft Skills: Problem Solving, Self-Learning, Time Management

Interests: Large Language Models (LLMs), Deep Learning, Machine Learning, Information Retrieval, Artificial Intelligence