SOHAM CHAKRABORTY

M.Tech in Computer Science

+91 8240125480 | sohamchakraborty10qw@gmail.com | Linkedin | Github | Portfolio | Kolkata, India

SC

SUMMARY

I am an M.Tech Computer Science student at ISI Kolkata, with hands-on experience in machine learning and data science, gained through an internship at Exposys Data Labs. I have constructed a Small Language Model from scratch and have applied my skills to deploy LLM applications using Hugging Face and Gradio. I am passionate about leveraging technology to address real-world challenges

EXPERIENCE

Data Science Intern

Exposys Data Labs

06/2023 - 07/2023 | Kolkata, India

A data analytics company focused on leveraging data for business solutions

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

EDUCATION

M. Tech in Computer Science

Indian Statistical Institute

07/2024 - 06/2026 | Kolkata, India

B. Tech in Computer Science & Engineering

Government College of Engineering and Textile Technology, Serampore

12/2020 - 06/2024 | Serampore, India

KEY ACHIEVEMENTS

Small Language Model Development

Built a Small Language Model (SLM) from scratch and deployed LLM apps using Hugging Face and Gradio

SKILLS

CMII | CSS | Data Structures | Deep Learning | Git | GitHub | HTML | Java | Linux | MatplotLib | NLP | Pandas | Python | PyTorch | Scikit-Learn | SQL | Gmail

PROJECTS

Small Language Model (SLM) from Scratch

01/2023 - 06/2023

A project to develop a Small Language Model from scratch using advanced machine learning techniques

- 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

01/2024 - 06/2024

Final year undergraduate project focusing on autonomous drone navigation through machine learning

- 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

INTERESTS

Technology Interests

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