SAMRIDDHA MAJUMDAR

+91 8910169688 ⋄ 5 Aswini Dutta Road, Kolkata, West Bengal, India ⋄ samriddha.majumdar706@gmail.com
Linkedin: samriddha-majumdar-7ba7b321b ⋄ ORCiD id: 0000-0002-8081-4607⋄ Github: sammajum706

EDUCATION

Jadavpur University — Kolkata , India
B.E Metallurgical and Material Engineering. — 2020 – 2024
CGPA: 4.0/4.0

Delhi Public School, Newtown — Kolkata , India
ISC (**Score:** 93.25 %) — 2020
Senior Secondary Education: 12th Standard (+2 Course)

Delhi Public School, Newtown — Kolkata , India
ICSE (**Score:** 93.8 %) — 2018
Secondary Education: 10th Standard

EXPERIENCE

ZS Associates — June 2024 - Present
Business Technology Solution Associate — *Pune, India*
Project: Analytics Program Optimization (APO)
Objective: In this internal project, the focus is on extracting insights from client biomedical data by primarily leveraging SQL and Generative AI. English instructions are automatically converted into SQL queries using Large Language Models (LLMs).

- Worked on data engineering module orchestration which includes services such as Crawler, Profiler, Interconnectedness, and Data Description.
- Hosted and tested service APIs using Docker and Putty.
- Deployed services into development environments using ArgoCD and TeamCity.

ZS Associates — Dec 2023 - Feb 2024
Business Technology Solution Associate Intern — *Pune, India*

Project: Trials.AI
Objective: Designed the Data Quality Management (DQM) strategy checks which plays a critical role in ensuring that the data being used to build or refine the ontology is accurate, consistent, and semantically meaningful.

- Designed and implemented a Data Quality Management framework for dataset integrity and reliability.
- Conducted critical checks such as file type verification, file size and duplicate file name validation, hierarchy and syntax checks, and predicate and Null ID assessments.
- Performed non-critical checks such as spell verification, ontology term ID and predictive validation, and duplicate record identification.
- Automated DQM reporting for summarizing outcomes of all checks for streamlined review and action.

Project: Knowledge Management Ops (KMOps)
Objective: Developed a system to create summaries of unstructured documents (pdf, ppt, docx) based on user queries.

- Developed a Retrieval-Augmented Generation (RAG) approach for efficient summarization of large docx documents.
- Utilized OCR techniques for information extraction from various formats, with data stored in JSON and managed in a Neo4j graph database.
- Tokenized and chunked sentences using NLTK to handle extensive information in docx files.
- Created vector embeddings with the BAAI/bge-base-en-v1.5 model and stored them in a FAISS database.

- Extracted top sentences based on user queries and summarized them within the LLM's context size using FAISS.
- Significantly reduced OpenAI API calls, improving processing time and performance compared to Elastic Search and Cosine Search methods.

CMATER Laboratory, Jadavpur University Dec 2021 - Jul 2023
Undergraduate Researcher, Supervised by Dr. Ram Sarkar *Kolkata, India*

- Developed a Feature Pyramid based Recurrent Residual U-Net model for phase segmentation of steel microstructure images.
- Worked on **MetalDAM** dataset which is a metallography dataset from additive manufacturing of steel.
- The proposed UNet model outperformed several state of art methods in MetalDAM dataset.
- Worked on a novel rank based ensemble method where Gamma Function is being proposed which combines the outcome of three standard CNN based transfer learning models for solving a two class classification problem of breast histopathological images, which outperformed the state of art by 4%.

KLIV Laboratory, Indian Institute of Technology, Kharagpur Jun 2022 - Nov 2022
Summer Research Intern *Kharagpur, India*

- Worked on Federated Deep Learning for biomedical image classification.
- Using medical image datasets for four diseases : COVID 19, Breast Cancer, Cervical Cancer, Skin Cancer
- Creating a benchmark between centralized and federated training algorithm.
- The implementation of FL approach can be seen here : Federated Deep Learning

T-M-S Laboratory, Indian Institute of Technology, Kanpur May 2023 - Jul 2023
Summer Research Intern , Supervised by Dr. NP Gurao *Kanpur, India*

- Worked in Crystal Plasticity Simulation for incorporating transformation induced plasticity (TRIP), twinning and slip effects in FCC single phase synthetic microstructures using a Fast Fourier based simulation software known as DAMASK.
- Mainly used two models in DAMASK : Phenopower model for slip and Dislotwin model for twinning and transformation.
- Utilised the Stacking Fault Energy parameter at a constant temperature to study the effects of twinning and transformation in FCC single phase synthetic microstructures.
- Utilised Atex, TSLOIM, and Paraview softwares for the post processing operations of the ouput provided by the DAMASK simulation software mainly for visualising the final deformed microstructure, IPF map, KAM map.
- Link for the poster presentation of the SURGE internship project: **Poster**.
- Pls refer to the **internship certificate**.

Amazon Machine Learning Summer School July, 2022

Intensive training sessions conducted by Amazon Applied Scientists on key Machine Learning and Deep Learning topics giving an emphasis on their industry level applications.
The certificate of completion can be found **here**

PUBLICATIONS

Gamma Function based Ensemble of CNN Models for Breast Cancer Detection in Histopathology Images

- Submitted at: **Expert Systems with Applications**
- Code Link: Gamma Function Ensemble

Metallographic Image Segmentation using Feature Pyramid based Recurrent Residual U-Net

- Submitted at: **Computational Materials Science**
- Code Link: Feature Pyramid RRUNet with squeeze and chanel excitation

PROJECTS

- Gamma function based ensemble of CNN models for breast cancer detection in histopathology images
- Federated Deep Learning Classification of Histopathological Medical Images
- Customer Sentiment Analysis Prediction

ONLINE COURSES

- Neural Networks and Deep Learning- **COURSERA**. The certficate can be found here
- Machine Learning- **COURSERA**. The certificate can be found here

SKILLS

Languages	Python, Java , C, C++, SQL
Technologies	Machine Learning, Computer Vision, Simulation, Image Processing, Natural Language Processing, Generative AI
Tools	Pytorch, Keras, TensorFlow, OpenCV, ScikitLearn, Langchain, Huggingface, AWS Sagemaker, AWS S3, Databricks, TeamCity, ArgoCD

ACHIEVEMENTS

Organising Secretary of Metallix,2024 Jadavpur University
Led the orchestration of the 15th edition of the annual techno-metal fest. Secured sponsorships from leading companies and startups, managed multiple internal teams, and introduced Hack-Met, an AI-based hackathon, as a pioneering addition to the event.

7th Position in Intel oneAPI Code Maven Hackathon
A technical hackathon organised by Intel and Techgig on the following topic:
"Predicting Custom Sentiment using Deep Learning" which consists of 10000 teams

Second Runner's Up in IEEE Pitch-O-Innovate
An idea presentation event on the particular topic:
" Application of AI in prevention and awareness of Domestic Violence"

Second Runner's Up in Gnosis(Quiz) : Metallix'23
A quiz event organized by Dept of Metallurgical and Material Engineering, JU in collaboration with Indian Institute of Metals, Kolkata Chapter

EXTRA-CURRICULAR

- Vocalist
- Acoustic Guitar Player
- Table Tennis
- Poster Designing