Debarshi Chakraborty

LinkedIn
GitHub

Email: debarshi2933@gmail.com
Mobile: +91-903-8564-470

Education

- **Ramakrishna Mission Vivekananda Educational and Research Institute**
  Master of Science - Big Data Analytics *July 2024 - Present*
- **Ramakrishna Mission Vivekananda Centenary College**
  Bachelor of Science - Mathematics Honors; CGPA: 7.63 *September 2021 - June 2024*

Experience

- **Research Intern** May 2025 - Present
  Indian Institute of Technology, Guwahati *Guwahati, India*
  - Collaborated with PhD scholars to design algorithms for the **Popular Matching Problem** in subcubic graphs and graphs of maximum degree 4.
  - Developed and analyzed **combinatorial and reduction-based algorithms** to determine the existence of popular matchings under bounded-degree constraints.
  - Implemented prototype solutions in `Python`, and conducted theoretical performance evaluations.
  - Gained hands-on experience in graph theory, complexity analysis, and algorithm design in a research setting.

Projects

- **Mood-Based Music Recommendation System — Machine Learning, Recommendation Systems (Dec 2024)**:
  - Built a classification model using Spotify data to categorize songs into 7 mood categories with 92% accuracy.
  - Developed a content-based recommendation engine using cosine similarity.
  - Resolved class imbalance with SMOTE and improved accuracy using ensemble methods.
  - **Technologies:** Python, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, XGBoost, LightGBM
- **Deep Learning Document Summarization and QA System — NLP, Deep Learning (May 2025)**:
  - Implemented a GRU-based Seq2Seq model with attention for document summarization.
  - Fine-tuned DistilBERT for extractive question answering on benchmark datasets.
  - Evaluated using ROUGE, BLEU, BERTScore, and Exact Match metrics.
  - **Technologies:** Python, PyTorch, Hugging Face Transformers
- **Distributed Inference for Large Language Models — Distributed Systems, NLP (May 2025)**:
  - Set up a distributed computing cluster to deploy LLMs such as DeepSeek R1 Distill and LLaMA 3.2B Instruct, achieving a 2–4× increase in tokens/sec throughput.
  - Conducted an extensive literature review on LLM inference strategies including tensor, model, and data parallelism.
  - Analyzed GitHub repositories of distributed LLM systems; documented architecture, scaling behavior, and memory layout.
  - **Technologies:** Python, C++
- **Real-Time (T+0) Trade Settlement System for the US Market — Blockchain, AI, Finance (Ongoing)**:
  - Designed a permissioned blockchain system using Hyperledger Besu and IBFT-2 consensus for atomic delivery-versus-payment (DvP) settlement of tokenized securities and cash.
  - Integrated Agentic AI for real-time trade validation, liquidity checks, and exception handling.
  - Utilized large language models (LLMs) for compliance reporting and anomaly detection.
  - Planned stress testing for 10,000+ trades/day and Monte Carlo simulations to quantify counterparty risk reduction.
  - **Technologies:** Blockchain, Smart Contracts, Python, Hyperledger Besu, AI Agents, LLMs

Skills Summary

- **Languages**: Python, R, JAVA
- **Frameworks**: Scikit-learn, NLTK, PyTorch
- **Libraries**: NumPy, Pandas, Matplotlib
- **Tools**: Git, GitHub, VS Code, Jupyter Notebooks, Neo4j
- **Platforms**: Linux, Web, Windows
- **Soft Skills**: Communication, Leadership, Event Management, Time Management

Achievements

- **Cleared IIT JAM 2024** *February 2024*
  Successfully cleared the IIT JAM 2024 examination, demonstrating proficiency in the subject and securing a place for higher studies in a prestigious institution.