Debarshi Chakraborty

in 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*

RELEVANT COURSEWORK

- **Data Science & Machine Learning**: Machine Learning, Deep Learning, Natural Language Processing, Computer Vision, Time Series and Survival Analysis
- **Big Data & Distributed Systems**: Java and Hadoop, Distributed Computing, Graph Databases
- **Mathematics & Optimization**: Probability and Stochastic Processes, Statistics, Linear Algebra , Convex Optimization

SKILLS SUMMARY

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

PROJECTS

- **Mood-Based Music Recommendation System (Machine Learning, Recommendation Systems)**: Developed a mood-based classification system using Spotify data to categorize songs into 7 mood categories (Calm, Dancing, Energetic, Happy, Neutral, Relaxing, Sad) with 92% accuracy. Designed a recommendation engine based on cosine similarity to suggest songs aligned with user mood preferences. Addressed class imbalance using SMOTE and implemented ensemble learning techniques (Random Forest, XGBoost, LightGBM).
  Tech Used: Python, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, XGBoost, LightGBM — December 2024

- **Custom Deep Learning-Based Document Summarization and Question Answering System (NLP, Deep Learning)**: Built custom deep learning models for document summarization and extractive question answering using LSTM/GRU-based Seq2Seq architectures with attention mechanisms. Fine-tuned pre-trained models (BART, T5, DistilBERT) for summarization and QA tasks, achieving competitive results on benchmark datasets. Evaluated models using ROUGE, BLEU, BERTScore, and Exact Match metrics. Deployed the QA system as an API with FastAPI for real-world usability.
  Tech Used: Python, PyTorch, Hugging Face Transformers, FastAPI — Ongoing

- **Distributed Large Language Models (LLMs) for Efficient Inference (Distributed Systems, NLP)**: Set up a distributed computing cluster to reduce inference time for large language models (LLMs) by leveraging tensor parallelism and parallelizing workloads across multiple nodes. Conducted a comprehensive literature review on distributed training and inference techniques such as model parallelism, data parallelism, and gradient compression. Implemented and evaluated distributed inference pipelines, significantly reducing latency while maintaining model accuracy.
  Tech Used: Python, C++, PyTorch, Hugging Face Transformers, Ray — Ongoing

CERTIFICATIONS

- **Introduction to Optimization Theory and Algorithms (NPTEL)** — *December 2024*
- **Building Applications with ChatGPT API (DeepLearning.AI)** — *December 2024*

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.

VOLUNTEER EXPERIENCE

- **Volunteer at Perceptron - Annual Tech Fest, RKMVERI** — *January 2025*
  Assisted in organizing and managing events, workshops, and technical activities for the university's annual tech fest, attended by more than 300 students and professionals. Played a key role in coordination and event execution, ensuring a smooth and impactful experience for all participants.