Atanu Maity

Email: amrkmrc@gmail.com
Mobile: +91-8016956715

Github: github.com/ytiam
Linkedin: https://www.linkedin.com/in/ytiam/

EDUCATION

- **Guru Ghasidas Central University** — CG, India
  Master of Science - Mathematics; Percentage: 48.2 — *May 2010 - April 2012*
  Courses: *Abstract Algebra, Linear Algebra, Fluid Dynamics, Mechanics, Calculus, Partial Differential Equation, Machine Learning, Deep Learning*

- **Calcutta University** — Kolkata, India
  Bachelor of Science - Mathematics; Percentage: 54 — *April 2006 - March 2009*
  Courses: *Linear Algebra, Numerical Analysis, Real Analysis, Vector Analysis, Functional Analysis, Complex Analysis*

SKILLS SUMMARY

- **Skills**: Data Science, Machine Learning, Deep Learning, NLP, Computer Vision, GenAI, Prompt Engineering, Issue-Fixer
- **Algorithms/Techniques**: LightGBM, Random-Forest, SVM, KNN, K-Means, Naive-Bayes, PCA, SVD, DBScan, t-SNE, Apriori, LDA, CNN, RNN, LSTM, Transformer, Auto-Encoder, Attention Mechanisms, ResNet, AlexNet, U-Net, GANs, RAG, PEFT, Few Shot Learning, In-Context Learning, Transfer Learning, Quantization
- **Languages**: Python, R, SQL, Bash, C, Cypher
- **Frameworks**: Numpy, Scikit, NLTK, SpaCy, TensorFlow, Keras, Pytorch, OpenCV, Flask, FastAPI, Py-Spark, Poetry, Pre-commit, Dataiku
- **Tools**: Docker, GIT, Airflow, MLFlow, MySQL, Neo4j, Github Actions
- **Platforms**: Linux, Web, Windows, Azure, AWS
- **Soft Skills**: Leadership, Mentoring, Continuous Learning, Writing, Time Management

EXPERIENCE

- **Accenture** — In-Office
  Senior Data Scientist - Associate Manager — *Sep 2024 - Present*
  - **Domain specific LLM Fine-tuning**: Finetuned GPT-2 small on domain restricted data using LoRA PEFT
  - **Improved Application Resiliency**: Developed and Integrated custom retry configuration and handler for LLM API endpoints
  - **LLM Payload Structure Unification**: Working on standardizing the different payload structure by designing an efficient API routing strategy tailored to diverse clients and their specific payload schema.

- **Capgemini** — In-Office
  Senior Data Scientist — *June 2021 - Sep 2024*
  - **Knowledge Graph Deployment**: Developed and delivered a comprehensive solution for a Knowledge Graph based question-answering system
  - **Custom Search Engine on Failure Analysis Report**: Developed and Deployed a tailor made customized Search Engine to Automate Information Retrieval task
  - **Process Automation**: Automated complex data processing pipelines for dashboard report generation using Selenium and Python.

- **Ayata** — In-Office
  Lead Data Scientist — *Apr 2016 - June 2021*
  - **Fraudulent Claim Estimation**: Developed and trained a custom embedding deep neural network on claim adjuster textual notes to predict the probability of fraudulent claims.
  - **Weather Parameters Forecasting**: Designed, trained, and deployed a multivariate forecasting model using LSTM, integrating radar data and satellite imagery to deliver predictive insights in the energy domain.
  - **Hail Damage Detection**: Trained a computer vision model to Identify and Measure the severity of Hail damage on Vehicle using CNN + YOLO - Insurance Domain.
  - **Predictive Model Building**: Worked on Different Predictive Modeling tasks for cross Insurance Clients using Random Forest, LightGBM, XGBoost, HMM

- **Karvy Analytics Ltd** — In-Office
  Data Scientist — *Apr 2015- March 2016*
  - **CLV Prediction**: Performed Unsupervised Learning using RFM Factors and Predicted CLV for effective marketing strategy.

PROJECTS

- **GENAI Hub (Generative AI, Prompt Engineering, NLP)**: Lead Developer - GenAI Based Use-case Development and Integrating those in a Flask based WebApp. Tech: Python, Langchain, RAG, Prompt Engineering, Flask, Jira, Azure Document Intelligence, Azure AI Search, Azure Web-App (Sept '23)
- **OCR - NLP search engine (NLP, Search Engine, Regex, Semantic Search, Vector Database, Indexing)**: Keyword and Semantic Mixed Hybrid Search Engine for 90 percent faster information processing and improved downstream tasks. Tech: Python, Dataiku, Sentence-Transformer, Flask, Faiss, LayoutParser (March '24)
- **Masked Language Model Training**: Language model to learn domain specific knowledge representation. Tech: Python, CuDA, PyTorch, Custom Tokenizer, Multi Headed Attention, RMSNorm (May '24)
- **GPT-2 Training from Scratch (NLP, Generative AI)**: Trained from scratch a Decoder only Transformer on a light weight GPU RTX-2090. Tech: Python, Pytorch, CuDA, Multi Headed Attention, LayerNorm, Residual Connection (Aug '22)

BLOGS

- **Blog: AirFlow and MLFlow Integration**: 300+ claps, Made development ready container. Tech: Airflow, MLFlow, Docker, Python (November '21)
- **Blog: Social Media Analytics in Pharma**: 150+ claps. Tech: Python, WebScraping, Bayesian Belief Network, Sentiment Analysis (November '21)

HONORS AND AWARDS

- Awarded title of Winnovation Award - Sept, 2023
- Capgemini Inspirer Award (3 times) for internal asset building and CoE (AI+Data) Contributions - Sept 2023, Sept 2022, April 2022

VOLUNTEER EXPERIENCE

- **Group Lead at Elite Developer Club Capgemini** India
  Conducted online technical Data Science sessions, reading research papers over 150+ developers.Jan 2022 - May 2024

- **SogetiLabs Fellow** Global
  Organized events, conducted workshops and delivered workshops reaching over 1000+ developers.Jan 2023 - Sept 2024