Debarun Ghosh

Site Reliability Engineer | Generative AI Developer | Innovator in Research & Technology

Mobile: +91 9742085682 LinkedIn: Debarun - LinkedIn

Email: debarun.ghosh.2024@gmail.com GitHub: Debarun - GitHub Repository

As a current SRE Apprentice at ANZ in Bangalore, I am actively refining my skills in system reliability and AI-driven solutions. I hold a bachelor's degree in Electronics and Communications from REVA University and a Postgraduate Program in Applied Generative AI Specialization from Purdue University. My journey has equipped me with robust expertise in AI and machine learning, demonstrated across various projects and internships. I am adept at developing AI technologies that enhance operational efficiency and decision-making. Ambitious and forward-thinking, I am seeking to advance my career in technical roles within global tech giants, where I can contribute to and grow with innovative AI/ML initiatives.

Experience:

Professional:

Organization : **ANZ Operations and Technology Pvt. Ltd., Bengaluru**
Position : Site Reliability Engineer (SRE) - Apprentice
Duration : May 2024 – Present
Project : Architecture as Code Automation & AI PoC Use Cases
Description : Spearheading PoC development and production-grade deployment of AI-driven tools within the Payments Domain—focusing on automating Architecture as Code (AAC) practices and building innovative GenAI solutions. Core contributions include streamlining internal data workflows, engineering a document summarization bot, developing privacy-first local LLM chatbots, and enabling secure, compliant NLP use cases aligned with enterprise banking standards.
Tech Stack : Python, GCP, Terraform, Code Fresh, LLMOps, NLP, Full Stack Development, Docker

Role:

- I contribute to the development of AI-powered tools for Architecture as Code (AAC) automation and various chatbot-based use cases tailored for secure internal banking workflows within the LLMOps team.
- Engineered a Generative AI-based document summarization bot, significantly reducing manual data processing time by significant amount, and enabling rapid retrieval and decision-making across teams.
- Built and deployed full-stack applications leveraging React.js with Vue, Python and Docker to streamline ML operations and accelerate development cycles.
- Presented a proprietary offline LLaMA based chatbot tool for regulated environments at ANZ Unconference 2024, focusing on privacy-first NLP applications using local LLMs.
- Acted as a technical mentor within the SRE; delivered internal training sessions to upskill peers in Machine Learning Fundamentals, Generative AI integration, and LLM concepts.
- Collaborated cross-functionally with GitHub Copilot team to implement secure, scalable AI-first tooling that aligns with banking compliance standards.

Academic – Internship:

Organization : **DOM4U Pvt. Ltd., Bengaluru**
Position : AI/ML Integration Engineer - Internship
Duration : December 2023 – April 2024
Project : Generative AI and Applied Machine Learning utility
Description : Enhanced legacy industrial applications through AI and ML integration, automating and improving efficiency via deep learning and reverse engineering.
Tech Stack : Python, Machine Learning, CUDA, TensorFlow, OpenCV, YOLO, Keras, JIRA (Agile), Java, JS, MySQL, Deep Learning Framework, Business Static Analysis and Technical Documentation.

Role:

- Implemented AI models utilizing OpenCV, YOLO, and CUDA for real-time processing.
- Optimized predictive modelling techniques that reduced data analysis time by 40% and improved accuracy by 15%.
- Understanding Software Life Cycle and gained understanding as QA tester during Production.
- Addressing new ideology for new Gen AI and ML and how this can affect certain products overall usability.

Organization : **Mphasis Ltd., Bengaluru**
Position : Software Engineer Intern
Duration : May 2023 – July 2023
Project : Transformation of COBOL (Mainframe) to Active Programming Language
Description : Developed an automation tool to convert COBOL code to high-level programming languages using Machine Learning pre-trained models, enhancing mainframe application efficiency and modernization.
Platform : Anaconda (Jupyter), Machine Learning, TensorFlow, Keras, BERT, LLM, NLP, Transformers
Role:

- Utilized pretrained AI models – BERT, Transformers, boosting mainframe efficiency.
- Process automation benchmarks. Solved technical problems with innovative optimization methods.
- Honed AI and automation for enterprise applications.
- Directed project cycles, meeting strategic goals efficiently.

Volunteer:
Secretary IEEE Student Branch (REVA University)
Duration: March 2023 – October 2023

- Orchestrating technical events, elevating the professional growth of over 100 student members through tailored workshops and industry interface sessions.
- Executing data-driven approaches, yielding a 20% surge in branch membership by using IEEE vTools.
- Strengthening internal communications and project coordination, optimizing event planning and operational efficiency.

Education:

- **Purdue University** : **Postgraduate Program** in Applied Generative AI Specialization, August 2024.
- **REVA University** : **B. Tech** in Electronics, and Communications Engineering, June 2024 (Completed).

Projects:

- Nestle HR AI Assistant (Repository: http://bit.ly/4fR2pZ3) – **Purdue University**
  - Enhanced HR Efficiency: AI-driven assistant using LLaMA 3.1 8B for accurate, context-aware HR policy responses.
  - User-Friendly Interface: Gradio interface enables easy, natural language access to HR policy information.
- Monitoring Vision AI (Repository: https://bit.ly/3MgAG6m) – **REVA University**
  - It uses a TensorFlow/Keras model to monitor machine status, with a Streamlit interface for real-time predictions.
  - The system includes motion detection, status logging, and email alerts, enhancing monitoring capabilities.
  - The project captures video, preprocesses frames, and determines machine status for efficient management.
- 5G Network Energy and Channel Optimization (Repository: https://bit.ly/3FTE5Fd) – **IEEE NKCON 2023**
  - Engineered a Python-based optimization framework – Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) to reduce energy consumption in 5G networks.
  - Conducted iterative solution, fine-tuning parameters to minimize operational energy costs while optimizing channel allocation. Achieved optimal network results through rigorous algorithmic strategies.
- Neurological Disorder Classification & PoTS Diagnosis (Repository: http://bit.ly/3MAhSPZ) – **Self Explored**
  - Engineered a Python-based ML pipeline using logistic regression and CNN for accuracy.
  - Enhanced predictive models, achieving high accuracy validated through ROC analysis. Visualized results with confusion matrices.

Publication:

- Orchestrated the integration of Genetic Algorithms & particle swarm optimization in 5G networks, boosting energy efficiency by 30% and enhancing overall network performance (IEEE NKCON '23, January 2024 - https://bit.ly/3UAFrx5)
- Researched and implemented AI-driven strategies to reduce energy consumption in 5G networks, resulting in a 25% decrease in overall energy usage (IEEE ICONAT '23, April 2023 - https://bit.ly/47t1FUP)

Awards:

- First Runner-up in the National Level Project Exhibition at The National Institute of Engineering Mysuru, 2022
- National Creative Aptitude Exam 2022 – Holding All India Rank: 79
- Winner in Education Week event 'Write Way,' - IEEE Education Society Ad Hoc Young Professionals Committee, 2023

Certifications:

- AGS Advanced Generative AI: Models and Architecture – Purdue University – Simplilearn (Certification Id: 6961374)
- AGS Essentials of Generative AI, Prompt Engineering & ChatGPT – Purdue University – Simplilearn (Certification Id: 6835121)
- PwC Switzerland - Power BI Job Simulation – PwC – Forage (Certification Id: QsDKnZa6aJJmmBYKq)
- Virtual Experience Participate – Software Engineer – JPMC (Certification ID: brNC3Nvi6yTpA5Yh5)
- Virtual Experience Participate – Quantitative Research – JPMC (Certification ID: cntgfdtikim4jojJH)
- Python for Data Science, AI & Development – IBM (Certification ID: BN8J9CMANXQN)
- Virtual Experience Participate – Agile (Scrum) – JPMC (Certification ID: hAzifmFJwgwajgDPk)
- Microsoft Certified Systems Administrator: Machine Learning – Verzeo

Skills:

- GenAI – RAG, Prompt Engineering, GAN, VAEs,
- Deep Learning: PyTorch, Tensor flow, OpenCV, NNs, Keras, LLMOps, NLP
- Go-lang, Python, Machine Learning: Roc, Confusion Matrix, Analysis, Regression, Lambda
- C/C++, Linux, GCP, Terraform
- Oracle SQL/MySQL/PostgreSQL/SQLite, Big Data Analysis, Vector Database
- Microsoft Power Bi, Jira, Agile Methodology, LangChain for Workflow Design
- Basic Java/JS, HTML and CSS

Languages – English (Fluent), Hindi, Bengali, Kannada (Beginner), French (Beginner)