ANUBHAV ROY

Kolkata, WB 700054
(+91) 9874454959
roy.anubhav96@gmail.com

EXECUTIVE SUMMARY

Head of Technology at Nekko, driving the design and deployment of advanced AI/ML solutions that transform business processes and enable innovation at scale.

With deep expertise in Machine Learning, Deep Learning, and AI-powered systems, I specialize in building and fine-tuning models for real-world impact, including Large Language Models (LLMs), Vision Transformers (ViT), and predictive analytics.

Experience spans Generative AI (Fine-tuning, RAG, Multimodal), Computer Vision (object detection, instance segmentation), and NLP (LSTM, NER), as well as broader AI applications in automation, optimization, and decision-making.

Proficient in Big Data, IoT, and Cloud Computing, I lead technology strategies that redefine process efficiency, drive innovation, and create measurable value for clients

PROFESSIONAL EXPERIENCE:

Nekko   December 2024-Present
Head of AI

Accenture - LLM CoE   October 2021-December 2024
Machine Learning Engineer

Datycs   April 2020-October 2021
Data Scientist

Indian Oil Corporation Limited   May 2019-April 2020
Machine Learning Engineer

EnergyAustralia   January 2019-Apr 2019
Machine Learning Intern

Jaguar Land Rover India   August 2018-Nov 2018
Machine Learning Intern

For Nekko - from December 2024:
- **Achieved 80% cost reduction at 4x scale** by deploying fully containerized "LLM" solutions with model ownership, customization, and privacy controls.
- **Led a cross-functional team of 16** to design, develop, and deploy end-to-end AI systems across industries including legal tech, compliance, and retail intelligence.
- **Built a modular, open-source Data Management Platform** using Faiss, LangGraph, and LangChain — enabling scalable, vendor-agnostic RAG pipelines with seamless orchestration.
- **Developed a real-time Computer Vision anomaly detector** for PPE compliance and behavior deviation tracking — actively deployed in manufacturing and warehousing environments to monitor safety and catch policy violations in action.

For Accenture - Applied Intelligence from October 2021:
- Instruction Fine-tuning Base Models (Mistral, Llama) using DeepSpeed and FSDP.
- Creating Custom RAG Agents for problem-solving using Chain-of-Thought (CoT) and ReAct.
- Optimal Store Location Predictor: Identified prime areas for a top Health & Beauty firm.
- Priority-Based Ticket Generator: Achieved a 23% budget saving through delivery optimization.

For Datycs from April 2020:
(Contractor-Kavyant Technologies)
- Built Bi-LSTM models for multi-label classification of clinical data (ICD 10 Codes).
- Developed custom NER models to extract key information from patient charts.

- Implemented middleware for transfer-learning-based custom embeddings.

For Indian Oil Corporation Limited, Refineries Division from May 2019 to April 2020:
(Contractor-SMS Chemicals and Process Consultants Pvt Ltd)
- Implemented AI-based predictive maintenance for rotary equipment.
- Performed statistical and regression analysis on sensor data to forecast equipment failures.

For EnergyAustralia ML and Analytics Division from January 2019:
(Contractor: EnergyTech Ventures)
- Project: IaaS (Insight as a Service) Solution for Large Scale Commercial Real Estate
- Using ML on-trend patterns to maximise throughput and energy savings

For Jaguar Land Rover India from August 2018:
(Contractor: Xenon Automotive India Pvt Ltd)
- Developed Mask RCNN models to identify damages on vehicle panels.
- Estimated repair costs based on detected damage intensity.

EDUCATION
Jadavpur University (Kolkata, India), Master of Engineering., [August 2022]
- o Major in Software Engineering (I.T.) - 8.4 GPA

Government College of Engineering and Leather Technology (Kolkata, India),
B.Tech., [August 2019]
- o Major in Computer Science and Engineering (C.S.E) – 8.23 GPA

CERTIFICATIONS
- **AWS Certified Machine Learning – Specialty (2025–2028)**
- **AWS Certified Machine Learning Engineer – Associate (2024–2027)**
- **AWS Certified AI Fundamentals**
- **Azure AI Fundamentals**
- **GCP Professional Machine Learning Engineer**
- **Career Essentials in Generative AI by Microsoft and LinkedIn**
- **Green Software for Practitioners**

TECHNICAL SKILLS
- **Large Language Models (LLMs):** Fine-tuning, RAG, and Multimodal AI
- **Generative AI:** Chain-of-Thought (CoT), ReAct, Instruction Fine-tuning
- **Deep Learning:** Vision Transformers, LSTMs, Bi-LSTM
- **Cloud Platforms:** AWS, Azure, Google Cloud
- **Big Data & IoT Integration**
- **Computer Vision:** Object Detection, Instance Segmentation
- **Natural Language Processing:** Named Entity Recognition (NER), Semantic Search
- **Tools:** FSDP, DeepSpeed, Vertex AI Workbench

RESEARCH PUBLICATION:
Springer July 2023: A Guide to Applied Machine Learning for Biologists
A Study and Analysis of Image from Text Generation Models (Final Thesis) - **Jadavpur University 2022**
Bangla character recognition based on Mobilenet v1 and Inception v3 (Paper ID 23) – **ICETSD 2019**

AWARDS/ ACHIEVEMENTS:
- o Recommended Candidate – 1- Airforce Selection Board Dehradun 2019
- o DRDO Robotics and Unmanned Systems Exposition (DRUSE) top 200 across India
- o International Conference of Emerging Technologies for Sustainable Development (ICETSD) publication and presentation.