Mohnish Bangaru
Machine Learning Engineer

Brooklyn, New York | (347) 856-8262 | mohnishbangaru@nyu.edu | LinkedIn

EDUCATION

New York University, Tandon School of Engineering   Sep 2023 - May 2025
Master of Science in Computer Engineering - GPA: 3.8/4.0   **New York, NY**

SRM Institute of Science and Technology   Jul 2017 – May 2021
Bachelor of Technology in Computer Science and Engineering- GPA: 3.8/4.0   **Chennai, IN**

PROFESSIONAL EXPERIENCE

New York University   **New York, NY**
Graduate Student Data Scientist   **Sep 2023 - Present**

- Built ETL pipelines for **30,000+ survey strengths**, reducing processing time by **50%**.
- Applied **NLP techniques** to analyze the open-ended text responses on surveys improving student feedback quality with **25% information gain**.
- Designed ML models and Tableau dashboards to generate visualizations and figures for what-if simulations related to institute performance,
- **Reduced deployment bottlenecks by 60%** by implementing **Dockerized ML pipelines**, leading to a faster turnaround for **A/B testing** of new models and data-driven improvements.
- Initiated **Onboarding** new hires in **machine learning best practices**, SOPs, and the tech stack.

Technology: Python, SQL, Microsoft Excel, Tableau, Qualtrics, Oracle DB, Pytorch, TensorFlow, Docker, Kubernetes.

KPMG   **Bangalore, IN**
Data Scientist   **Apr 2022 - May 2023**

- **Directed** automation of reporting workflows with **Python & ML workflows**, **saving 264+ hours annually** (12.5% efficiency gain).
- **Coordinated** a team of 3 data scientists to enhance anomaly detection in trading reports via machine learning, diminished false positives by 40% versus existing rule-based detection systems.

Analyst   **Jan 2021 - Apr 2022**

- **Revamped SQL queries** for real-time data access, cutting query execution time by **30%**.
- Implemented REST APIs achieving automated deployment and improving model deployment frequency by 15%.

Technology: Python, SQL, Microsoft Excel, Alteryx, JavaScript, Java, API Development.

PROJECTS

Efficient Deep Learning for Edge Deployment – New York University   **Jan 2025 - Present**

- Trimmed ResNet-50 (25.6M) with **quantization-aware training**, and **Residual Connections** which **reduced model size by 50%**, achieving a **memory footprint reduction of 65%**, and resulting in **10x faster** on-device **inference** speed.
- Deployed the model on a Raspberry pi 4, with an average **inference time** of **0.007 seconds** for image classification tasks.

Radiology Question-Answering Model – New York University   **Sep 2024 – Dec 2024**

- Leveraged the Vision-Language Transformer on 5,000 radiology scans from the MIMIC-CXR dataset alongside corresponding text reports, achieving a 96% accuracy in identifying conditions.

Fine tuning LLM's using Q-Lora – New York University   **Feb 2024 - May 2024**

- **Enhanced** the instruction-following capabilities of Pythia 6.9B and 12B models via targeted fine-tuning and quantized training on the Alpaca dataset, resulting in a 15% improvement in benchmark question-answering scores in a resource constrained setting.

SKILLS AND INTERESTS

- **Tools:** Hadoop, Spark, TensorFlow, PyTorch, JavaScript, Tableau, Power BI, Alteryx, Core ML.
- **Languages:** Python, Java, C++, React, R, SQL.
- **Relevant Coursework :** Advanced Computer Vision, Efficient ML and AI Accelerator, Deep-learning, Computing Systems Architecture, Computer Networks, Systems Engineering, Data Science for Business.
- **Certifications :** Deep Learning Specialization, Data Analytics with Python, Alteryx Core Developer.