

Raajesh N

✉ greatraajesh10@gmail.com ☎ +91 95668 74940 in LinkedIn 🌐 Portfolio

Profile Summary

Machine Learning AI Engineer skilled in Python, Java, SQL, and neural network fine-tuning (e.g., Tiny-llama). Hands-on experience with LoRA, model optimization, and building real-world AI apps like chat bots and LLM tools. Strong in combining technical skills with a commerce background to create practical, data-driven solutions. Capable of handling end-to-end ML workflows, from training to deployment on platforms like AWS and Colab.

Education

Master of Computer Application (M.C.A.)

June 2023 – Dec 2025

SRM University

- GPA: 3.9/4.0
- **Specialization:** Data Science, Data Engineering, Data Warehousing, Machine Learning

Bachelor of Commerce (B.Com.,)

June 2020 – May 2023

SASTRA Deemed University

- GPA: 7.5/10.0
- **Specialization:** Direct Taxation, Insurance Law and Business Law

Experience

Intermediate Representative

Chennai, IN

BNY (Bank of New York)

July 2023 – Present

- Reduced processing time from 3 hours to 12 minutes for Contract Reviewer Process using Python.
- Cleaned massive data using SQLite for Competitor Analysis.
- Redesigned report structure (STP) for faster, atomic operations and better isolation.
- Mastered team processes and contributed 13+ improvement ideas for Client Experience and Work Quality.
- Worked with IDP (Intelligent Document Processing) to extract information from documents.
- Working with AI Agents to automate Client On boarding Processes.

Projects

Dino AI

GitHub [↗](#)

- In this project, hundreds of dinos learn to avoid obstacles (cacti and birds) by evolving their neural networks over generations. The AI uses a simple feedforward neural network as its "brain," and the population improves through genetic algorithms—mutation, crossover, and selection based on performance.
- Tools Used: Feedforward Neural Network, Reinforcement Learning, Deep Learning

TinyLLaMA Fine-Tuning for Python Code Generation

Colab [↗](#)

- Fine-tuned the TinyLLaMA-1.1B-Chat model using a subset of the codeparrot-clean dataset to improve performance on Python code generation tasks. Implemented LoRA (Low-Rank Adaptation) for efficient parameter tuning on resource-constrained GPUs using peft.
- Tools Used: PyTorch, PEFT, BitsAndBytes, Accelerate, Colab

GPT-2 Architecture Replicate

Colab [↗](#)

- Replicated GPT-2 architecture from scratch with custom layers, embeddings, and inference.
- Tools Used: Transformers, Machine Learning, GPT, LLM

Certifications

- **SIE(Securities Industries Essential)** – FINRA
- **Data Engineering Professional Certificate** — IBM (Coursera)
- **Agile Scrum Master Certification** — Udemy
- **Google Data Analytics Professional Certificate** — Google (Coursera)

Technologies

Languages: Python, Java, SQL, JavaScript

Technologies: PEFT, Hugging Face Transformers, TensorFlow, PyTorch, Scikit-learn

Skills

Skills: Python, SQL, Java, JavaScript, Data Lakes, Data Modeling, AWS (S3, Redshift), Scikit-learn, XGBoost, LightGBM, TensorFlow, PyTorch, Docker, FastAPI, Flask, Hugging Face Transformers, LLMs (GPT, BERT, TinyLLaMA), PEFT (LoRA), BitsAndBytes, OpenCV, CNNs, Transformers.