

RIAKANTHA SHIVA KRISHNA

Houston, TX (Willing to Relocate) | +1(603) 438-8428

shivakrishna.a2609@gmail.com | [LinkedIn](#) | [Portfolio](#) | [GitHub](#) | OPT, No Sponsorship Needed

SUMMARY

Trained 124M-parameter transformer from scratch with zero GPU spend • Engineered LangChain RAG agents that eliminate 100% of manual data work for 7,008 companies • AWS Bedrock/SageMaker/Docker • Token efficiency & inference optimization specialist • Master's Data Science 4.0 GPA • Immediate contributor

SKILLS

Machine Learning & LLMs: LangChain · PyTorch · TensorFlow/Keras · Transformers · RAG Agents · LLM Optimization · Token Efficiency · Hugging Face · Groq API

MLOps & Cloud: AWS Bedrock · SageMaker · S3 · Docker · Git · Streamlit · CI/CD

Data Engineering: Python · Pandas · NumPy · SQL · pdfplumber · PyMuPDF · BeautifulSoup · yfinance · Regex · ETL Pipelines

Visualization & Tools: Tableau · Power BI · Seaborn · Matplotlib · Databricks (learning) · dbt (learning)

Other Skills: Data Structures

EXPERIENCE

Software Engineer (Volunteer)

July 2025 – Present

Guardian's Embrace

Remote

- Trained **124M-parameter transformer** from scratch on CPU-only hardware for ESG score prediction across 100 firms.
- Engineered **LangChain agent** that extracts structured JSON from 10-K/Q filings with 100% accuracy, eliminating all manual data entry.
- Built preprocessing pipeline that reduces 100-page PDFs to less than **2,400 tokens**, enabling viable CPU inference.
- Integrated **yfinance API** to compute **MA, RSI, MACD**, and volatility metrics at scale for 7,008 companies
- Key Tools: Python, LangChain, Pandas, BeautifulSoup, PyMuPDF, HuggingFace Embeddings

Freelance Data Scientist

Oct 2022 - May 2023

Independent Projects

Remote

- Developed **CNN** models in TensorFlow and Keras, achieving **92% accuracy** on 14,000 images through advanced data augmentation techniques
- Built **customer segmentation** system using **K-Means** that reduced high-churn segments by **20%**.
- Designed ensemble models (**Random Forest, Gradient Boosting, KNN**) with custom **feature engineering** that significantly improved marketing response prediction.
- Key Tools: Python, TensorFlow, Scikit-learn, Git, Tableau, Seaborn, Pandas, NumPy

PROJECTS

AutoPitcher | [GitHub](#)

Python · LangChain · Groq API · Streamlit · RAG

- Groq LLM-powered application that scrapes job postings, generates hyper-tailored outreach emails with dynamic project injection, and consistently achieves higher response rates.

Sales Analytics Project (Simulated Role: Junior Data Analyst at Atliq Hardware's) | [GitHub](#)

SQL · Power BI

- Implemented **SQL** queries to solve number of business use cases and delivered **end-to-end data pipeline** and executive dashboards highlighting **16.67% YoY** growth in unique products sold.

EDUCATION

Master's Degree, New England College, Data Science, (Grade: 4.0)

2023 - 2025

Bachelor of Technology, JNTUH, Electronics and Communication Engineering.

2018 – 2022