

# KESHAV SHARMA

[keshavprabhakar449@gmail.com](mailto:keshavprabhakar449@gmail.com) | +91-9877162427 | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

AI/ML engineer shipping 0-to-1 production systems end-to-end: multi-agent orchestration (LangGraph), open-weights fine-tuning (QLoRA), and stateful backends (FastAPI, PostgreSQL, vector stores). Currently the sole engineer on a cloud-native customs-compliance platform for cross-border manufacturers.

## Experience

- **Founding Engineer — Importex, Miami, US (Remote)** July 2026 – present
  - Sole engineer on Trade360, a cloud-native IMMEX / Annex 24-30 customs-compliance platform: data model, backend, API surface, and web client.
  - Designed a transactionally consistent FIFO/PEPS discharge engine reconciling pedimentos against SAT Data Stage/COVE feeds.
  - Built append-only audit logging with correction workflows, eliminating destructive edits on compliance records.
  - Automated ingestion across SAT web services and client ERPs (SAP, Oracle, Baan); tariff catalogues modelled as configuration, so rule changes ship without redeployment.
  - Ran client technical discovery; converted competitive analysis of 5 incumbent vendors into the scoped requirements baseline.
  - Shipped compliance modules to 5 client pilots across Mexican manufacturing operations.
- **AI / Data Engineer Intern — Tritone Analytics, Denver, US (Remote)** Jan 2026 – Mar 2026
  - Engineered a 0-to-1 LLM-driven ETL pipeline structuring 100K+ unstructured records, cutting manual effort by 85%.
  - Architected multi-agent LangGraph validation workflows with schema-level checks, eliminating record loss.
  - Shipped FastAPI REST services backed by ChromaDB for sub-second semantic search.
  - Debugged production infrastructure within a remote Git-based deployment workflow.
- **Data Science Intern — Sabudh Foundation** Jan 2025 – Apr 2025
  - Built TensorFlow training pipelines across 100K+ record production datasets.
  - Ran feature engineering and EDA to isolate the predictive signals driving model selection.

## Projects

- **Veritas Engine - Dual-Mode Claims Audit Swarm | LangGraph, Llama-3.3-70B, FastAPI, Next.js** | [Project Link](#)
  - Built a dual-mode 4-node LangGraph swarm auditing multi-modal document claims for VCs and researchers.
  - Evaluated each claim in an isolated Python sandbox; a compression node cut token payload by 90%.
- **Cerebrospatial Quantum Platform (CSQP) | PyTorch, LangGraph, OpenCV, PennyLane** | [Project Link](#)
  - Developed an autonomous CV pipeline tracking real-time intraoperative brain-tissue shift in pixels/mm.
  - Deployed a hybrid Quantum CNN (QCNN) with a LangGraph supervisor driving safety alerts.
- **SeismoSearch & Martian-Llama-3.1-8B | Python, Unsloth, QLoRA, ObsPy** | [Project Link](#)
  - Engineered an STA/LTA signal-processing pipeline isolating seismic anomalies from planetary telemetry.
  - Fine-tuned an 8B model via 4-bit QLoRA (rank 16) across 5 frequency profiles on a single consumer GPU.

## Skills

- **Languages:** Python, SQL, Java, C/C++, JavaScript/TypeScript, Bash
- **AI/ML:** LangGraph, LangChain, Multi-Agent Systems, RAG, QLoRA / Unsloth Fine-Tuning, 4-bit Quantization, PyTorch, TensorFlow, Hugging Face, OpenCV, PennyLane, Scikit-learn
- **Backend & Data:** FastAPI, Flask, Next.js, REST, WebSockets, ETL Pipelines, PostgreSQL, MongoDB, Redis, ChromaDB, Neo4j, Pandas, NumPy
- **Infrastructure:** Docker, AWS, CI/CD, Git, Linux, System Design, Sandboxed Execution, ERP & Third-Party API Integration

## Education

- **VIT-AP University, India** Sept 2023 – Aug 2027  
Branch – Computer Science and Engineering, Specialized in Data Analytics | CGPA: **8.32**

## Achievements and Certificates

- Winner, NASA Space Apps Challenge - Best Use of Science, Andhra Pradesh State Level Nov 2024
- President, SEDS Aurora Space Club - VIT-AP University Jun 2024 – Aug 2026
- Certifications: IBM Data Science Professional | DeepLearning.AI Machine Learning Specialization Dec 2025
- AWS Cloud Foundations and Architecture Oct 2025