

# Phaneendra Gavara

Data Scientist · ML Engineer · AI/LLM Engineer · Data Engineer

■ phaneendragavara436@gmail.com | ■ +1 623 320 6354

■ Tempe, AZ | ■ Portfolio | LinkedIn | GitHub

■ OPT (F-1) · Valid until Jun 29, 2029

## PROFESSIONAL SUMMARY

MS Data Science graduate from Arizona State University (GPA 3.90/4.0, Top 5%) specializing in AI/ML systems, LLM engineering, and production-grade data pipelines. Experienced in building end-to-end ML solutions — from RAG architectures and multi-agent frameworks to computer vision and time-series forecasting. Passionate about bridging research and deployment to deliver measurable business impact.

## EDUCATION

### M.S. Data Science, Analytics and Engineering

Aug 2023 – May 2026

Arizona State University — Ira A. Fulton Schools of Engineering — Tempe, AZ

- GPA: 3.90/4.0 | Concentration: Computing and Decision Analytics
- Top 5% of program — confirmed by three independent faculty letters of recommendation
- Relevant Courses: Deep Learning (A), Statistical ML (A), Distributed Computing (A+), Data Processing at Scale (A), Natural Language Processing (A), Data Visualization (A+), Knowledge Representation (A), Applied Regression (A), Topics in ML (A), Software Agility (A+)

### B.Tech. Electronics & Communication Engineering

2019 – 2023

IIIT Bhubaneswar — Bhubaneswar, India

- Final Year Thesis: Machine Learning for ECG Signal Anomaly Detection

## EXPERIENCE

### AI Research Intern

May 2024 – May 2025

Arizona State University — NVIDIA-funded project — Tempe, AZ

- Built a 3D Gaussian Splatting reconstruction pipeline; reduced inference time by 40% via CUDA optimizations.
- Developed novel rendering techniques improving visual fidelity 25% on standard benchmarks.
- Deployed an automated testing framework cutting validation time from weeks to hours.

### Teaching Assistant & Grader

Aug 2023 – May 2024

Arizona State University — Tempe, AZ

- TA for Data Processing at Scale (Spark, Hadoop) and EEE 120 (Digital Circuits) — 120+ students each.
- Held weekly office hours; automated grading scripts saving 8 hrs/week per course.

### Sports Society Manager

2021 – 2023

IIIT Bhubaneswar — Bhubaneswar, India

- Led college sports society; coordinated inter-university events with 500+ participants.
- Managed scheduling, budgeting, and volunteer coordination across 12+ sports disciplines.

## PROJECTS

### NeuroVault — AI Clinical Intelligence Platform

3rd Place, Sun Devil AI Hackathon 2024 (50+ teams)

- Production RAG pipeline (GPT-4o + Pinecone) processing 10,000+ medical documents with 94% retrieval accuracy.
- Achieved 89% diagnostic precision on complex multi-condition cases; 3x faster reporting than baseline.

### Republic of Bean — Multi-Agent AI Parliament Simulation

LangChain · GPT-4 · Python · FastAPI

- GPT-4 agents autonomously debate and vote on satirical policy motions using LangChain orchestration.
- Simulates democratic processes with role specialization, coalition dynamics, and filibuster mechanics.

### Real-Time Sign Language Translator

MediaPipe · LSTM · TensorFlow · React

- 96.2% accuracy on ASL alphabet recognition using hand-landmark sequences fed into LSTM classifier.
- Real-time browser interface achieving <50ms latency via TensorFlow.js + WebSocket streaming.

### Dine-Cast — Restaurant Demand Forecasting

SARIMA · Prophet · XGBoost · LightGBM

- Ensemble forecasting model achieving 12% MAE on hourly demand; optimized staffing and waste.
- Automated ingestion pipeline processing 3 years of POS data with drift detection.

## ARIA — AI Research Intelligence Assistant

GPT-4o · Pinecone · LangChain · Streamlit

- Multi-source RAG system (arXiv, Semantic Scholar, PubMed) with citation-aware retrieval.
- Reduces literature review time ~70%; supports gap analysis and hypothesis generation.

## SecureML — Privacy-Preserving Federated Learning

PySyft · PyTorch · Differential Privacy

- Federated CNN with differential privacy ( $\epsilon=1.2$ ) achieving 97.3% accuracy — matches centralized baseline.
- Demonstrated on medical imaging; privacy budget overhead <5% vs non-private training.

## TECHNICAL SKILLS

---

<b>Languages:</b>	Python, SQL, JavaScript/TypeScript, R, CUDA (basics), Bash
<b>ML / AI:</b>	PyTorch, TensorFlow, scikit-learn, XGBoost, LightGBM, Hugging Face, LangChain, LlamaIndex
<b>LLM / RAG:</b>	OpenAI GPT-4o, Pinecone, FAISS, Chroma, Retrieval-Augmented Generation, Prompt Engineering
<b>Data Engineering:</b>	Spark, Hadoop, Kafka, Airflow, dbt, PostgreSQL, MongoDB, Redis
<b>Cloud / MLOps:</b>	AWS (SageMaker, S3, Lambda), GCP (Vertex AI), Docker, Kubernetes, MLflow, FastAPI
<b>Visualization:</b>	Tableau, Plotly, D3.js, Matplotlib, Seaborn, Next.js/React

## ACHIEVEMENTS & ACTIVITIES

---

- 3rd Place — Sun Devil AI Hackathon 2024 (NeuroVault, 50+ competing teams)
- Top 5% of M.S. cohort at ASU — independently confirmed by three faculty LORs
- Teaching Assistant × 2 courses simultaneously; served 240+ students
- Sports Society Manager, IIT Bhubaneswar — led college-wide inter-university events
- GPA: 3.90/4.0 — straight A/A+ in all 10 graduate courses