Aryan Patodiya

925-918-1058 | **Fresno,CA** | **aryanpatodiya018@gmail.com** | **linkedin.com/in/aryanpatodiya08/** | **github.com/Aryanp018** | aryanp-portfolio.netlify.app

EDUCATION

California State University — Fresno, CA
Master of Science in Computer Science — *Jan 2024 - Jan 2026*
Coursework: Reinforcement Learning, Deep Learning, Advanced Software Engineering, Applied Biometrics Security, Artificial Intelligence

Charotar University of Science and Technology — Anand,India
Bachelor of Technology in Computer Engineering — *Jul 2019 - Apr 2023*
GPA: 3.82/4.0
Coursework: Artificial Intelligence, Information Security, Cloud Computing, Big Data Analytics, Cryptography

TECHNICAL SKILLS

- **Programming & Core:** Python, Java, C++, SQL, TypeScript
- **ML/AI Frameworks:** PyTorch, TensorFlow, Transformers, OpenCV, Scikit-learn, LLMs, RL
- **Model Serving & Optimization:** Triton Inference Server, vLLM, TensorRT, ONNX, TorchScript, CUDA
- **Cloud & MLOps:** AWS (EC2, Lambda, S3, DynamoDB), Azure, Docker, Kubernetes, Terraform, MLflow, CI/CD (GitHub Actions, Jenkins)
- **Data & Databases:** Apache Kafka, Hadoop, MongoDB, Redis, PostgreSQL, MySQL
- **Software Development:** REST/GraphQL APIs, OOP, Multithreading, Design Patterns, Agile
- **Testing & Tools:** Unit Testing (JUnit, PyTest, Google Test), Git, JIRA, Automation Testing

WORK EXPERIENCE

SAC-Indian Space Research Organization — Ahmedabad, India
Machine Learning Intern – Hydrological Modeling & Forecasting — *Dec 2022 - Apr 2023*
- Developed a **probabilistic forecasting pipeline** using HMMs and Markov Chains, improving rainfall prediction accuracy by **20%**.
- Processed **hundreds of GBs of satellite data** via Pandas + AWS S3, reducing data prep time by **30%**.
- Integrated outputs into **flood-risk decision tools**, strengthening **early warning systems** across pilot regions.

Raven Technolabs — *Rajkot, India*
Machine Learning Engineering Intern – Cloud Infrastructure & Model Deployment — *May 2022 - July 2022*
- Built **microservices** (Spring Boot, Node.js) for **real-time ML inference**, reducing latency by **25%** under load.
- Deployed models via **REST/GraphQL APIs**, handling thousands of requests/day in staging.
- Automated **CI/CD pipelines** (AWS, GitHub Actions), cutting release cycles by **50%** and ensuring reproducibility.

Nanotech Technologies — *Ahmedabad,India*
Cofounder & Lead Software Engineer – Scalable Systems & Data Engineering — *Mar 2019 - Nov 2021*
- Directed a 1**4-member engineering** team to design edge-to-cloud **data pipelines** for industrial automation.
- Optimized backend systems via **low-level code refactoring + DB tuning**, boosting system performance by **20%**.
- Deployed **real-time monitoring infrastructure**, enabling predictive maintenance and future ML integration.

California State University — *Fresno, California*
*Machine Learning Research Assistant - Brain Computer Interfaces **(NIH-R15 Funded Research)*** — *May 2025 - Aug 2025*
- Designed a **scalable EEG preprocessing pipeline** (MATLAB, Python) processing **64k+ multi-channel trials**, cutting preprocessing time by **40%**.
- Built **optimized CNN architectures** with temporal–spatial convolutions, SE blocks, and GELU activations, achieving **98.7% balanced accuracy** and **AUC 0.9997**.
- Implemented a **robust evaluation framework** (test-time voting, stratified CV, augmentation), improving model robustness by **25%** and leading to **international conference acceptance**.

PROJECT EXPERIENCE

MarketPulse: Real-Time Stock Trend Predictor Github
- Designed **LSTM/GRU forecasting models** achieving **80% directional accuracy** on OHLCV data..
- Deployed production pipeline with **vector DB + optimized embeddings**, reducing query latency by **20%.**
- Deployed with **TensorFlow Serving + AWS**, delivering a **production-ready inference pipeline**.

DocuQuery: Fullstack Semantic Search Engine
- Built a **context-aware document retrieval system** using **LLaMA-2, LangChain, and FAISS**, improving search relevance by **30%**.
- Integrated **OpenAI embeddings** to improve retrieval; increased relevance of top results by **~30%**.
- Packaged as a **Dockerized microservice + Streamlit UI**, ensuring scalability and real-world usability.

TimeNet: Sequence Forecasting with RNNs
- Developed **RNN models** for rainfall & energy forecasting, outperforming ARIMA by **15–20%**.
- Streamlined data pipelines with **TensorFlow,** reducing training time by **20%**.
- Validated models across **multi-domain datasets**, ensuring adaptability.

SkyDefender: 3D Shooter & Duck Hunt Game
- Engineered a **full 3D game engine in C++ with OpenGL**, supporting **multi-level progression & dynamic environments**.
- Optimized graphics/audio subsystems with SOIL + irrKlang, sustaining **60 FPS** across complex scenes.

Optimizing Retrieval-Augmented Generation (RAG) with Reinforcement Learning *(Ongoing Research)*
- Developing a **GRPO-based reinforcement learning framework** to jointly optimize retrieval and generation components in RAG systems.
- Implementing a **hybrid reward mechanism** to balance factual accuracy and response coherence.
- Conducting **experiments on SQuAD and Natural Questions datasets** to evaluate hallucination reduction and performance gains.
- Designing a **comprehensive evaluation pipeline** using BLEU, ROUGE-L, and BERTScore for robust performance measurement.