Aryan Sharma

Fullerton, CA (Open to Relocation) | 714-213-4517

✉ aryan.sh.6302@gmail.com | 🔗 LinkedIn | 🐙 GitHub | 🌐 Portfolio

Skills

Technologies	Python, SQL, C++, NLP, Transformers (BERT, GPT, T5), Generative AI, LLMs, Prompt Engineering, Vector Databases, Chatbot Development, Web Scraping
Frameworks	TensorFlow, Scikit-Learn, NLTK, HuggingFace, Keras, LangChain, LangSmith, LangGraph, CrewAI, Ragas
Tools	Flask, Docker, Kubernetes, AWS, Azure App Services, MongoDB, Faiss, Selenium, ML Studio
Techniques	Exploratory Data Analysis (EDA), Data Mining, Neural Networks, Fine Tuning

Experience

ECS Dept., California State University, Fullerton *Jan 2024 – Present*
Teaching Assistant – Algorithm Engineering (Part-Time)

- Guided a cohort through dynamic programming, greedy methods, and NP-completeness via live demos, labs, and review sessions, so students could apply theory to real problems.

- Ran office hours with targeted code reviews and feedback, improving assignment quality and exam readiness across the class.

Pollack Library, California State University, Fullerton *Aug 2024 – Present*
LLM Research Engineer (Part-Time)

- Optimized and evaluated LLM pipelines used by students and faculty, focusing on faster responses and more relevant results for academic queries.

- Integrated AI tools (retrieval + summarization) into existing library systems to reduce manual searching and surface credible sources quickly.

Sabudh Foundation, Mohali, India *Jul 2021 – Jun 2023*
Data Science Intern

- Led a 7-member team delivering 5 projects; improved chatbot accuracy by **20%** and reduced misinterpretation by **15%** through better data curation and evaluation.

- Built clean documentation, flows, and handover guides that increased team efficiency by **30%**; completed **100+** hours of EDA and earned Dataiku/RASA certifications, cutting turnaround time by **35%**.

Projects

YOLOv8 Object Detector *Computer Vision*

- Developed and deployed object detector with transfer learning, achieving **92% accuracy** on custom dataset; optimized inference with TensorFlow and Ultralytics.

Mini Transformer from Scratch *NLP Core*

- Implemented encoder-decoder Transformer with self-attention and positional encoding to strengthen understanding of LLM internals.

Fine-tuned LLaMA2 with LoRA/QLoRA *LLM Customization*

- Fine-tuned LLaMA2 on domain dataset, reducing hallucinations by **18%**; deployed with HuggingFace PEFT for efficient scaling.

RAG-based Document Assistant *Enterprise GenAI*

- Built Retrieval-Augmented Generation (RAG) system with LangChain + Pinecone, enabling semantic search and context-aware Q&A across large document sets.

Multimodal AI Subtitle Generator *Speech + Language*

- Created subtitle generator using Whisper + LLMs for transcription, translation, and alignment; delivered **multi-language support** for video accessibility.

Education

California State University, Fullerton, CA *Expected Graduation: Dec 2025*
Master of Science, Computer Science — GPA: 3.8 / 4.0

I.K Gujral Punjab Technical University, Punjab, India *Aug 2019 – Jun 2023*
Bachelor of Technology, Computer Science & Engineering — GPA: 3.6 / 4.0