RAHUL GUPTA

notrahulgupta@gmail.com
rahulg.info

+619-764-8386

OpenAI Tier 3 Builder

ANTHROP\C Tier 2 Builder

SKILLS

- **Machine Learning & AI**: Deep Learning, NLP, **Transformer Models**, Computer Vision, Speech Recognition
- Optim Stack: Custom Activation Functions (SwiGLU, GELU), Attention Mechanisms (Causal, Long-form)
- Model Architecture: Transformer Variants, GPT-2, BERT, T5
- Model Optimization: **Pre-training, Fine-tuning, Hyperparameter Tuning, Transfer Learning**
- MLOps: Model Deployment, Monitoring, Versioning, A/B Testing
- Programming: Python, Java, TypeScript, C++
- Frameworks & Tools: PyTorch, TensorFlow, HuggingFace, OpenAI-SDK, Lang-chain, MLflow, Optuna
- Cloud & DevOps: **AWS SageMaker**, Azure ML, Docker, Jenkins, Kubernetes, Vercel
- Data Analysis: SQL, Jupyter, Pandas, NumPy, Pinecone DB

WORK EXPERIENCE

→ CTO & FOUNDER

FX **Fleuxlabs.ai**

Aug, 2022 - Present
San Diego, CA

- Built **NextJS LLM platform** with custom APIs: 900k+ daily tokens, **GPT-4/Anthropic integration**, Pinecone DB for shared context.
- Implemented shared buffer: Auth, file support, TTS. Exploring vectorized data for enhanced AI reasoning.
- Creating Deux: **LLM for web automation**: Causal attention, **HuggingFace transformers,** Dockerized Puppeteer for dynamic data extraction.

→ SDE-2 (RELEASE ENGINEERING)

Walmart Global Tech

June, 2021 - Aug, 2022
Sunnyvale, CA

- Engineered solutions to **reduce CI time** and increase developer productivity, collaborating with cross-functional teams (Platform, Infrastructure, InfoSec).
- Developed and maintained custom tools, **CI/CD pipelines for automating build, test, and release** of Walmart eCommerce (Glass) and business apps **(iOS, Android, React Native).**
- Built self-service CI/CD tools for code signing and distributing apps via App Center, Google Play Store, and Apple App Store Connect.
- **Owned entire CI/CD infrastructure and mobile release processes**, serving as the go-to person for all mobile CI/CD operations.
- Tools: Xcode, Python, **Jenkins, Fastlane, Nexus**, Azure DevOps, SonarQube, TestFlight

→ SDE-1 (TEST)

Walmart Global Tech

June, 2019 - May, 2021
Sunnyvale, CA

- Developed and owned **React.js application for end-to-end order processing,** reducing manual effort by 60% across multiple testing teams.
- Built and maintained full-stack automated testing framework for regression, functional, and end-to-end testing in Order Management.
- Implemented **continuous integration** using Jenkins, setting up servers and jobs for periodic test execution.
- Tools : React, Java, Node.js, MySQL, **Cassandra**, JIRA, **JavaScript, Selenium, Cucumber**

PROJECTS

- **Speaker Recognition**: Speaker Recognition: Engineered scalable GRU/LSTM-based system **processing 2TB of audio data,** achieving 77% (SOTA with LSTMs) accuracy & **sub-100ms inference time**, with potential applications in real-time fraud detection and customer authentication.

- Verte: **Fine-tuned Mistral-7B for spine-related queries using QLoRA,** integrating OpenAI Whisper and GPT-4 Turbo for 50% accuracy improvement.

EDUCATION

MS, COMPUTER SCIENCE

San Diego State University (2022-2024)

BS, COMPUTER SCIENCE

San Diego State University (2014-2018)