Anu Sihra

San Francisco, California, USA | +1 3466382597 | anupreetsihra5@gmail.com | in/Anupreet-Sihra | GitHub

EXPERIENCE

Labelbox **San Francisco, CA**
Forward Deployed Engineer *May 2026 - Oct 2026*

- Built a computer-use data pipeline that generates tasks, runs agents, and captures full trajectories (actions, screenshots, step-level rewards), producing 100+ structured examples for post-training and evaluation.
- Served as the engineering point of contact for frontier-lab customers, translating model failure modes into human-in-the-loop annotation and preference-data programs that delivered 2,000+ expert-rated comparisons.
- Built internal tooling for AI evaluation workflows on computer-use tasks.

Paycom **Dallas, TX**
AI Software Developer *Jan 2025 - Feb 2026*

- Served as the engineering point of contact for enterprise clients, partnering with Product and Support to triage client-reported issues and ship tailored fixes, and turning recurring customer feedback into platform features across scheduling and payroll modules serving 100,000+ client accounts.
- Shipped AI-driven automation for enterprise labor-budget auditing - multi-step MCP tool calling over scheduling and payroll data to surface discrepancies - cutting manual auditing effort 30%.
- Built and scaled backend REST and GraphQL APIs for scheduling workflows - shift assignments, shift exchanges, and real-time labor cost tracking - with SQL query optimization that cut data-fetching overhead 40% under enterprise load.
- Re-architected the Scheduling Import Center with rebuilt React upload flows and server-side bulk validation, cutting import time for 50,000-row enterprise files from 3+ minutes to under 30 seconds.

Barclays **Pune, India**
Software Engineer (BA3) *Aug 2022 - Aug 2023*

- Built global tax applications (React, C#/.NET, SQL Server) computing charges across APAC, UK, and US regions; resolved 15 critical production defects with zero-downtime releases across year-end cycles.
- Automated the Year-End Tax Longview Reporting Pack using RESTful APIs and batch pipelines, cutting report generation from 3-4 hours to 40-50 minutes (80% faster).
- Executed multi-region tax-data migrations across transactional databases with automated validation, completing cutovers 87% faster than the prior process.

PROJECT

Avals - Local-First AI Agent Evaluation Platform | github.com/anupreeet/avals

- Built a local-first AI-agent evaluation platform (FastAPI, React 19) that launches any agent x model x task combination and live-scores trajectories on reward, cost, and correctness - 10+ concurrent eval jobs streamed over one multiplexed WebSocket.
- Engineered a durable sequence-numbered JSONL event log: every run survives server restarts with zero event loss, and clients replay full trajectories from any offset.
- Generated TypeScript clients from the OpenAPI spec - 100% of the API surface type-checked end-to-end - with a trajectory dashboard rolling up reward, cost, and token metrics per task.

HyperFrames Video Studio - Local AI Video Production Pipeline | github.com/anupreeet/hyperframes-video-studio

- Contributed to a fully local AI video studio (React/Vite, Express/TypeScript) that turns a script into a rendered 1080p MP4 - voiceover, captions, animated scenes - with 0 cloud API calls end-to-end.
- Implemented pipeline stages spanning Kokoro TTS narration, Whisper caption alignment, and 30 animated scene templates, with ffmpeg renders streamed live over SSE - script to final cut in under 10 minutes.

EDUCATION

Rice University **Houston, TX**
Master of Computer Science, Computer Science | GPA: 3.9/4 *Aug 2023 - Dec 2024*

Banasthali University **Jaipur, India**
Bachelor of Technology, Computer Science and Engineering | GPA: 8.97/10 *Jul 2018 - Jun 2022*

SKILLS

Languages & Frameworks: Python, TypeScript, JavaScript, SQL, C#/.NET, Java, React.js, Node.js, FastAPI, GraphQL, REST APIs
AI & Agents: AI Agents & Tool Calling, Model Context Protocol (MCP), RAG Pipelines
Evaluation & Post-Training: LLM Evaluation (LLM-as-Judge), RLHF & Preference Data, RL Environments & Gyms
Cloud, DevOps & Data: AWS (EC2, S3, Lambda), Docker & Kubernetes, CI/CD (Jenkins, GitHub Actions), PostgreSQL, Apache Trino