Jaiydev Gupta

San Francisco, CA • (440) 990 5492 • jaiydev799@hotmail.com • LinkedIn • GitHub

Results-oriented Sr. Software Engineer with 3.5 years of experience at VISA Inc. specializing in developing scalable AI solutions

PROFESSIONAL EXPERIENCE

VISA Inc - VAS Innovation Labs, San Francisco, CA
Senior Software Engineer | 2025 - Present

- Led the development of an end-to-end automation framework by wrapping Langraph around Claude Code SDK, enabling the creation of highly specialized agents empowered MCP tools which could deliver features Autonomously.
- Designed and implemented s Neo4j Graph Database Solution that mapped out the dependencies for the 10 million line Issuing Solutions codebase, accelerating code conversion to Java at the rate of 10k-20k lines of code a day.
- Enhanced the pipeline by fine-tuning Mistral-Nemo-12B for code conversion (using a novel progressive prompt masking technique) for Tandem Application Language (TAL) to business logic improving accuracy.
- Served the finetuned LLM though dynamic Lora Adapter Switching via VLLM reducing memory footprint during inferencing from an additional 50-80GB vram to 150MB

VISA Inc – VAS Issuing Solutions, Denver, CO
Software Engineer | 2022 - 2025

- Continuous Pretraining on Mistral-Nemo-12B on 20,000 internal VISA documents using Fully Sharded Data Parallel (FSDP) to optimize resource utilization. enhancing the Visa-Assist application by improving prompt clarification and reducing hallucinations. Further quantized the model using HQQ Quantization to 4 bits for minimal accuracy loss and faster inference.
- Spearheaded the construction of a RAG pipeline for our micro-frontend chatbot, Utilizing OpenAI embeddings and Redis Vector Store, along with a Data Pipeline to sanitize incoming documents, ensuring data integrity and reliability.
- Implemented a message-level encryption strategy using dynamically generated public and private certificates at runtime, reducing fraudulent incidents by 20%.

PERSONAL PROJECTS

Knowledge Distillation of BERT on SST-2 sentiment analysis task - Finetuning

- Used a standard cross entropy loss for the teacher model and a KL divergence loss for the student model to match the soft probability distributions produced by the teacher.
- Achieved a minimal loss for 1000 samples for 10 epochs.

SKILLS

- Agents: Langraph, Langhain, Claude Code SDK, Langsmith, MCP, Transformers, VLLM, PyTorch
- Frontend: React, Bl@zor , Nextjs, Redux, Micro-frontends
- Backend: DotNet Core, Express, Node.js, Golang, Flask, FastApi
- DevOps: Docker, RedHat Kubernetes, AWS ECS
- Programming Languages: C, C#, Python, Java, JavaScript, Golang
- Database: Redis, Mongo, Neo4j