Devang Borkar

+1 (312) 358-5722 | devangborkar3@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Skills

- **Languages:** Python, C#, Java, TypeScript/JavaScript, SQL, Go, C++
- **AI/ML:** LLM Evaluations, multi-agent systems, LoRA fine-tuning, prompt engineering, ReACT agents, tool-calling, RAG, embeddings, adversarial test generation, AI red-teaming
- **Frameworks:** PyTorch, HuggingFace Transformers, FastAPI, LangChain, React, MongoDB, Redis
- **Cloud/APIs:** GCP, AWS, Docker, OpenAI/Anthropic/Gemini APIs, OpenRouter, CI/CD

Education

- M.S. Computer Science, University of California, Davis (GPA: 3.92)   *Expected Jun 2026*
- B.E. Computer Science, SPPU, Pune (GPA: 4.0)   *Graduated*

Experience

AI Engineer *PilotCrew AI*   Oct 2025 – Present

- Built an LLM evaluation platform (Python, FastAPI, MongoDB) to automate model benchmarking, AI red-teaming, and synthetic test generation for AI agents in production.
- Developed automated prompt optimization pipelines that analyze model outputs and iteratively improve prompts, increasing agent accuracy by 36% on adversarial test suites.
- Created a scalable benchmark infrastructure with 50+ Docker containers and 100+ Python tool integrations to evaluate LLM agents on hardware design automation tasks.

SWE Intern *LearnHaus AI*   Jun 2025 – Aug 2025

- Built a full-stack multimodal AI application (React, Python) from scratch, processing video, audio, and text inputs with optimized latency for real-time user feedback.
- Developed a LLM-as-Judge evaluation framework using GPT-5 and Gemini-3 to score model outputs across multiple quality dimensions; deployed to Google Cloud Platform.

Founding Engineer (Employee #2) *HammerTrade (Stealth Startup)*   Oct 2024 – Jun 2025

- Built scalable data pipelines and backend infrastructure (Python) from scratch to support ML model training and real-time trading simulations.
- Developed a configurable simulation environment for training reinforcement learning agents, enabling rapid experimentation with 10+ tunable hyperparameters.

Software Engineer *Hexaview Technologies*   Aug 2022 – Sep 2024

- Shipped 20+ production features for a Fortune 500 fintech platform; contributed to scalable microservices using C# and AWS Lambda.
- Optimized a legacy .NET backend servicing over 1 million monthly requests, applying key design patterns to reduce code complexity & successfully redesigning 50+ REST APIs

Projects

CausalFlow: AI Agent Debugging Framework   *[GitHub](#) | ICML '26 submission*

- Built an open-source Python framework to automatically detect and fix errors in LLM agent outputs using step-by-step analysis and automated repair generation at inference time.
- Tested on 3,000+ math and coding problems; successfully repaired 42.7% of failed AI outputs and improved agent performance by up to 31 percentage points without retraining.

Process Reward Model for On-Device LLMs   *[GitHub](#)*

- Fine-tuned an 8B parameter LLM using LoRA and PyTorch to verify reasoning steps; optimized for edge deployment with a lightweight 0.6B model on 1,319 math problems.
- Improved model accuracy by 21% using specualtive decoding over baseline methods; published model weights to HuggingFace.