SHIVAM RATNANI

+1 (978) 799-7813 • shiv@ratnani.org • https://www.linkedin.com/in/shivamratnani/ • https://ratnani.org

---

EDUCATION

University of Wisconsin-Madison
Bachelor of Science (BS) - **Computer Science** & **Data Science** (Double Major) | Graduated: **May 2025**

WORK EXPERIENCE

Ember | Lead Software Engineer   **November 2025 – Present**
- IP protected by NDA

Expert AI | Lead Software Engineer   **May 2025 – November 2025**
- Rebuilt the platform across API, web, and mobile, delivering an agentic SOAP-note pipeline with parallelized sub-agents and a six-tier priority queue; achieved 92%+ structured-note acceptance with one-click Athena import.
- Drove revenue capture by cutting failed EHR uploads from 15% to <3% for 100+ physicians preventing lost charges.
- Migrated Flask to FastAPI & refactored to asyncio, doubling throughput, reducing p90 latency from 600ms to 200ms.
- Designed a central document aggregator to be injected for each SOAP subagent for better analysis and fewer retries.
- Launched a MedicalGPT product (Python API + Next.js UI) with 6 clinician modes including lab analysis, medical record analysis, differential diagnosis, treatment planning, guideline lookup, and general Q&A.
- Fine-tuned MedGemma 27B-it on Vertex AI using thousands of (no PHI) examples; achieved TTFT at 6 s or less.
- Built CI/CD on GCP Cloud Run with dev & prod containers reducing buggy deployments from 29% to 4%.
- Operated as de facto tech lead: managed contractors to finish the web revamp ensuring delivery against tight timelines.

Rightworks | Software Engineer Intern   **July 2024 – August 2024**
- Resolved 21 critical bugs in AccessHub's AngularJS web app, improving stability for 300,000+ active users
- Collaborated on development of new Windows AccessHub application written in C#
- Designed and implemented new state management system in AngularJS, significantly improving user experience by replacing properties ".pristine, .dirty, and .touched"
- Monitored and documented API performance with DataDog to optimize web application responsiveness
- Collaborated on Windows AccessHub C# application development and monitored API performance using DataDog

ActiveLinks | Co-Founder, Lead Developer   **December 2024 – November 2025**
- Developing iOS application in Swift/SwiftUI serving 500+ beta users with real-time data synchronization

EduTools | Co-Founder, Lead Developer   **May 2024 – January 2025**
- Founded and developed EduTools, a secure educational platform with integrated OCR analysis and personal GPT assistance, enabling 1000+ users to build collaborative knowledge bases from academic materials
- Architected dual-LLM system utilizing Mistral-7B for content analysis and specialized models (LLama3, Gemma2, Qwen2.5-coder, Claude 3.5 Sonnet) for targeted educational solutions
- Built cross-platform desktop application using Electron, React.js frontend and Node backend, deployed on Azure infrastructure with microservices architecture for scalable content processing
- Migrated from AWS to Azure, implementing cost-saving measures with Azure App Services, Blob Storage, CosmosDB, and Azure Functions

---

Projects:

Stock Volatility Analyzer: Python/React AWS application analyzing 19,000+ stocks with automated screening
Connect 4 AI: Python game using Minimax algorithm with alpha-beta pruning and Pygame GUI

Skills:

Languages: Python, JavaScript, TypeScript, Java, C#, C/C++, Swift, Kotlin, SQL, Go, R. Bash/Linux
Frameworks: FastAPI, Django, Flask, React.js, Node.js, Express.js, AngularJS, SwiftUI, React Native, .NET
Databases: PostgreSQL, MySQL, MongoDB, Redis, Firebase Firestore, Supabase
Cloud/DevOps: AWS (EC2, S3, Lambda, RDS, API Gateway, CloudWatch), Google Cloud Platform (Cloud Run, Cloud Build, Compute Engine, Cloud Storage, Cloud DNS), Microsoft Azure (AApp Services, Functions, Blob Storage, CosmosDB), Docker, Kubernetes, GitHub Actions, Expo EAS
AI/ML: LangChain, LangGraph, TensorFlow, Keras, scikit-learn, pandas, NumPy, Prompt & Context Engineering
Tools: Git, REST APIs, GraphQL, WebSockets, OAuth, JWT, Nginx, Gunicorn, pytest, Jest, SonarQube, DataDog