Prakash Saini

Open to Relocate | +1 (737) 296-8315 | sainiprakash525@gmail.com
LinkedIn: https://www.linkedin.com/in/sainiprakash525 | GitHub: https://github.com/Prakash-sa

Software Engineer with 2+ years of experience delivering complex, high-performance software across the full stack. Skilled in Python, C++, Go, and JavaScript, with expertise in cloud-native services, container orchestration (Docker, Kubernetes), and API development. Recognized with HPE's Relentless Curious Award and runner-up in Global HPE Innovation Days.

Education

University of Texas at Austin | *MS in Information Science (Machine Learning)*   Austin, USA | *Aug 2022 – May 2024*
Courses: Machine Learning, Deep Learning, Natural Language Processing, AI in Health, Software Design & Impl

National Institute of Technology Surat | *BE in ECE*   Surat, India | *Aug 2017 – May 2021*

Technical Skills

Languages: Go, Python, Java, JavaScript/TypeScript, C/C++, SQL
Frontend: React, Node.js, Angular, Vite, Tailwind CSS, Cypress, Jest, HTML5/CSS3
Backend & APIs: Flask, FastAPI, Golang, Node.js, Spring Boot, GraphQL, OAuth2, JWT, Swagger/OpenAPI
Cloud & Containerization: AWS (S3, EC2, IAM, VPC, EKS, Lambda, RDS), Azure (Storage, Virtual Machines, Active Directory, VNet, AKS, Functions, SQL), GCP (Storage, Engine, IAM, VPC, GKE, Functions, SQL), Docker, Kubernetes
AI/ML & AI-Assisted Development: PyTorch, TensorFlow, LangChain, Hugging Face, NLP, GenAI, Agentic AI, n8n, LLM APIs (OpenAI, Claude, internal agents), AI-assisted coding tools (Cursor, Claude Code, GitHub Copilot)
DevOps & CI/CD: Git, GitHub Actions, Jenkins, Gradle, Infrastructure-as-Code (Terraform, Helm)
Observability & Monitoring: Prometheus, Grafana, New Relic, Splunk, OpenTelemetry, distributed tracing
Databases & Messaging: PostgreSQL, MySQL, MongoDB, Redis, Kafka
Other: Agile/Scrum, Jira, Confluence, VS Code, Code Reviews, Load Testing

Work Experience

Hewlett Packard Enterprise | *Software Developer*   Fort Collins, CO | *Jul 2024 – May 2025*
- Led the design and implementation of scalable, high-availability storage microservices supporting HPE Storage using React, Python, Go, Kubernetes, Docker, and AWS, following Agile Software Development methodologies.
- Pioneered an AI-powered automation chatbot using open-source tools like n8n, Qdrant, S3, and Ollama, reducing manual analysis time **by 90%** and delivering insights from architecture documents and sales reports.
- Collaborated with PMs, designers, and QA to deliver features with fast iteration cycles, ensuring alignment across UI/UX and backend deliverables, reducing release time **by 1 month**.
- Strengthened infrastructure security by enforcing FIPS compliance, integrating OAuth2/JWT authentication mechanisms, and resolving vulnerabilities flagged by **SonarQube**, reducing security risk **by 75%**.
- Successfully **led a team of 4 developers** for 3 release cycles following Agile methodologies.
- Integrated audit logging using event-driven patterns with Kafka with a new system design **within 1 week**.

Hewlett Packard Enterprise | *Software Engineer Intern*   Fort Collins, CO | *May 2023 – Aug 2023*
- Collaborated in developing the internationalization (I18N) and localization (L10N) feature for other services using React, Go, AWS, and GitHub, ensuring seamless integration into the entire codebase **within 7 days**.
- Built CI/CD pipelines using GitHub Actions to automate tests and deployment, reducing manual overhead **by 80%**.

Nokia Solutions and Networks | *Software Engineer*   Bangalore, India | *Jul 2021 – Jun 2022*
- Implemented automated system health checks and network monitoring tools using Python and Bash, reducing network setup time from **6-7 days to under 5 minutes** (99% reduction).
- Proactively identified and resolved UI/API performance bottlenecks, achieving **a 40% improvement** in load times through strategic optimizations.
- Participated in 24x7 on-call support rotation, authored technical documentation (design docs, runbooks, integration guides), resulting in resolving incidents related to microservice crashes.
- Created monitoring dashboards in Prometheus and Grafana, decreasing incident detection time **by 90%**.

Projects

- **News APP**: Adapted Mistral-7B-Instruct-GPTQ on the SAMSum dataset using QLoRA with 4-bit quantization, cutting GPU memory usage by 70%. Achieved ROUGE-L F1 of 0.162, BLEU score of 0.0198, and BERTScore F1 of 0.860, demonstrating high semantic alignment in generated summaries. **Website**, **GitHub**
- **Bio GPT**: Trained BioGPT with Haystack FARMReader on a domain-specific biomedical dataset, achieving 99.55% Exact Match and 99.61% F1 across 2,434 answerable questions, and 100% top-N accuracy for unanswerable cases. Deployed the model as a Flask-based API with Ngrok and frontend on Vercel. **Website**, **GitHub**

Awards & Recognitions

Relentless Curious Award by the Vice President at HPE
Runner-up in Global HPE Innovation Days
Winner of 4 National Level Hackathons

Leadership

- **Core Team Member in Developer Student Club (by Google Developers)**: Mentored 150+ students on Android Development and organized the HackCOVID-19 Hackathon with a team that was held on Devfolio.
- **Core Team Member in Cutting Edge Visionary**: Mentored 50+ students on Android Development, IoT, and organized events and debates.