





Rezwan Sheikh

 (917) 463-8110  rezwanswe23@gmail.com  LinkedIn  GitHub

Summary

Full-stack Software Engineer skilled in Next.js, React Native, Node, and TypeScript, with experience building secure and scalable web and mobile applications. Strong background in authentication, real-time features, API integrations, and system design. Experienced in developing AI-driven features using modern LLMs and RAG pipelines. Proven ability to ship production-quality projects, solve complex problems, and deliver end-to-end solutions.

Education

St. John's University <i>Bachelor of Science in Computer Science (GPA: 3.6 / 4.0)</i>	Expected Dec 2026 Queens, New York
---	---------------------------------------

Experience

Momen Technologies <i>Software Engineer</i> <ul style="list-style-type: none">Built an LLM-powered knowledge assistant with context-aware responses, cutting latency by 40% and scaling to 10K+ queries/dayDeveloped a React Native mobile app with Firebase Authentication, reducing login time by 30% and supporting 500+ concurrent usersImplemented a real-time AI chat interface with smooth animations and persistent sessions, enabling sub-200ms response times	May 2025 – Present Orlando, FL (Remote)
Momen Technologies <i>Software Engineer</i> <ul style="list-style-type: none">Fine-tuned DeepSeek-Chat 7B with LoRA, improving response relevance by 30% and reducing memory and training time by 40%Built a RAG pipeline with LangChain, MongoDB, and Pinecone, boosting factual accuracy by 20% and reducing hallucinationsDeployed a scalable NestJS backend on AWS EC2, reducing API response times by 35% through optimized load balancing and efficient request handling	Feb 2025 – May 2025 Orlando, FL (Remote)
NEW EIC <i>Software Engineering Intern</i> <ul style="list-style-type: none">Built a scalable interview management platform with cloud storage using MongoDB and AWS S3, achieving 3x faster retrieval and 70% lower storage usageArchitected a strategy management system with RBAC, real-time collaboration, and approval workflows, reducing approval time by 40%	Dec 2024 – Mar 2025 Boston, MA (Remote)
Institute for Machine Learning <i>Software Engineering Intern</i> <ul style="list-style-type: none">Developed AWS infrastructure (EC2, RDS, Lambda) to improve scalability and system performanceIncreased scalability by 40% and reduced data processing time by 25% through cloud infrastructure optimizationsBuilt time series models in Java, integrated econometric modules, and deployed real-time financial analytics	Sep 2024 – Dec 2024 Queens, NY

Projects

Horizon <i>React, Next.js, TailwindCSS, Plaid, Dwolla, Appwrite, Sentry</i> <ul style="list-style-type: none">Developed a financial web application using Plaid for bank account linking and Dwolla for secure fund transfersImplemented real-time error tracking with Sentry, reducing error resolution time by 50%
--

Technical Skills

Languages and Software: Java, Python, JavaScript/TypeScript, HTML/CSS, Node, Express, Nest.js, React, React Native, Next.js, Firebase, MongoDB, PostgreSQL, Supabase, Appwrite, TailwindCSS
AI/ML: NLP, Prompt Engineering, Multimodal AI, LLM Fine-Tuning (LoRA/QLoRA), RAG, Agentic AI
Tools: IntelliJ, Eclipse, AWS, Docker, GitHub, Jira, VS Code, Postman, Sentry, OpenAI, Hugging Face, LangChain, Pinecone
Certifications: AWS Certified Solutions Architect - Associate | Amazon Web Services (AWS) | Feb 2025