

RUSHANK MIHIR SHAH

✉ rushankshah65@gmail.com | ☎ +1 213-234-8568 | 🌐 Portfolio | 📄 GitHub | 🔗 LinkedIn

PROFESSIONAL EXPERIENCE

- **LinkedIn Inc, Mountain View, CA — Software Engineer** Feb 2026–Present
 - Built agentic features for LinkedIn Hiring Assistant, serving 1.5M+ users, using OpenAI and open-source LLMs.
 - Led migration of two critical resume-processing pipelines (350K+ resumes/day, 350M total) from Apache Samza to Flink, saving ~\$91K/year, cutting memory by 45% and 11% (~8 TB total), and retiring 60 TiB of legacy Kafka capacity.
 - Built LLM evaluation pipelines and ran SFT and RL fine-tuning on Qwen models serving millions of requests daily.
- **C3.AI, Redwood City, CA — Associate Solution Engineer** July 2024–Jan 2026
 - Shipped 8+ enterprise GenAI apps (RAG with PGVector, Gemini/OpenAI) for government, manufacturing, and steel clients, driving \$4M+ revenue and \$10M+ in customer savings.
 - Built ingestion pipelines for structured (MSSQL) and unstructured (PDF, Docx) data with multimodal chunking and indexing for accurate retrieval.
 - Built a predictive maintenance platform for 1,500+ sensors and an ML model management system for 800+ models (Python, React).
- **ClaimBrite Inc, Remote (CO) — Software Engineer Intern** Feb 2024–May 2024
 - Built and launched a responsive, customer-facing website using Next.js and Tailwind CSS, improving user experience through optimized performance, modern design, and mobile-first responsiveness.
 - Collaborated cross-functionally with designers, product managers, and stakeholders in an agile development environment to collect user feedback, iterate on features, and implement design improvements, resulting in a more intuitive and user-centric website.
- **Harvard Medical School (McLean Hospital), Boston, MA — Research Developer** Jul 2021–Jul 2022
 - Owned the full-stack development of a HIPAA-compliant Mental Health iOS app in Flutter, engineering a secure and scalable backend on AWS to protect sensitive user data.
 - Leveraged AWS services including DynamoDB, Elastic Beanstalk, Route 53, Cognito, and S3 for backend development, routing, authentication, and video hosting, ensuring scalable and secure application infrastructure.
 - Acted as the project's only POI (Person of Interest) by contacting the cloud security team and leading the team with the security and hosting services.

EDUCATION

- **University of Southern California, Los Angeles, CA** Aug 2022–May 2024
M.S. in Computer Science. (CGPA 3.81/4)
- **Dwarkadas J. Sanghvi College of Engineering, Mumbai, India** Aug 2018–May 2022
B.E. in Information Technology. (CGPA 9.55/10)

ACADEMIC PROJECTS AND RESEARCH PAPERS

- **Virtual Labs**
 - Developed and deployed virtual lab simulations for schools across India, enabling over 2 million students to perform lab experiments remotely, bridging the gap where physical labs are unavailable.
 - Designed and implemented the front-end of the application using ReactJS, enhancing user engagement with interactive animations powered by Rive.
 - The project is included as part of CDAC and is hosted under the Vidyakash Portal, URL for open access.
- **Automated API and Database Management**
 - Devised a new system where developers can build an automated backend system, using ReactJS and NodeJS.
 - Ensured the security of user data using tools and techniques such as JWT-based authentication, format string vulnerability checking at every point, and API Rate Limiters to avoid DDoS attacks.
 - Authored and published a paper on the same system in International Journal for Research In Applied Science and Engineering Technology, DOI: 10.22214/ijraset.2022.41827.
- **Dynamic Word Embeddings for compression of deep learning models**
 - Designed a novel rolling hash algorithm to compress word embeddings by 94% with no loss in model accuracy, optimizing deep learning model performance on resource-constrained mobile devices.
 - Leverages Python, Pandas, NumPy, and PyTorch for implementation.

TECHNICAL SKILLS

- **Languages and Databases:** Java, Python, TypeScript, JavaScript, SQL, PostgreSQL, PGVector, NoSQL.
- **Data Streaming and Cloud:** Apache Flink, Apache Kafka, AWS (DynamoDB, EC2, S3, Elastic Beanstalk, Cognito).
- **AI/ML:** PyTorch, Generative AI, LLMs (OpenAI, Gemini), RAG, Vector Databases.
- **Frameworks and Libraries:** React, Next.js, Node.js, Express.js, Flutter, Tailwind CSS.
- **Tools and Platforms:** Git, GitHub, VSCode, REST APIs.