

# SOHAIL GIDWANI

Los Angeles, CA • [sohailgidwani15@gmail.com](mailto:sohailgidwani15@gmail.com) • [LinkedIn](#) • [Github](#) • [Portfolio](#) • +91 9503151319

Passionate software engineer focused on AI technologies and full-stack development, with problem-solving and collaborative skills.

## WORK EXPERIENCE

### Senior Software Engineer - I | Insaito, Inc.

May 2025 - Present

- **AI Agent Builder Platform:** Led the end-to-end architecture and development of a comprehensive AI agent builder platform. Built core infrastructure for open-source LLM deployment (Qwen 3, Mistral Small 24B 2), designed robust OAuth integration for 100+ third-party apps, and developed Model Context Protocol (MCP) servers to expose integrated app capabilities. Delivered seamless full-stack solutions, enabling users to create advanced AI agents with custom workflows and deep external integrations.

### Full Stack - Software Developer | IIFL Finance Ltd.

June 2023 - May 2025

- **Custom Data Chatbots (RAG) :** Built an internal employee support chatbot using NLP, Python, and Flask. Integrated with Qdrant vector database, Azure OpenAI service, and Zoho ticketing system. This AI-powered solution significantly reduced the number of support tickets raised by employees, streamlining internal processes. ([Certificate of Achievement](#))
- **Gold Loan Image Audit App:** Engineered AI-powered application using models like GroundingDino, Swin-Transformer, enhancing fraud detection and reducing potential loan fraud by 15%
- **CapitalGenie:** Designed and implemented an automated user support system that responded to queries via email and other channels. Leveraged IIFL's internal APIs to fetch user information, diagnose issues, and utilized GPT-4o to generate structured, personalized responses. This solution accelerated the resolution of the issue by 70% compared to manual processes.

### Web Developer (Intern) | Iremify.com

May 2021 - July 2021

- Developed and maintained web applications using NodeJS, Flask, and React; ensured cross-platform compatibility and delivered responsive designs.

## SKILLS

- **Programming:** Python, TypeScript, JavaScript, Java (intermediate), C/C++ (Basic), SQL
- **Full-Stack:** React, Next, Node.js, Express, Hono, FastAPI, Flask, RESTful APIs, HTML5, CSS3, Tailwind CSS, WebRTC
- **AI/ML:** TensorFlow, Keras, Scikit-Learn, NLP, Computer Vision, OpenCV Vector DBs, LLMs, CNN, LSTM, Transformer, MLOps, N8N, MCP
- **Databases & Cloud:** MySQL, PostgreSQL, MongoDB, Oracle, SQLAlchemy, Azure, AWS, Docker, Git, Linux, Serverless
- **Methodologies:** Agile, Test-Driven Development, Code Reviews

## EDUCATION

### B.E. Computer Engineering | University of Mumbai

Aug 2019 - May 2023

CGPA: 9.05/10

- **Relevant Coursework:** Artificial Intelligence, Machine Learning, Advanced DBMS, Data Structures & Algorithms, Operating Systems, Software Engineering, Cloud Computing, Object-Oriented Programming, Big Data Analytics, Computer Networks, Cryptography & System Security, Blockchain.

## PROJECTS

### Image Feature Detection & Captioning | Neural Networks, Streamlit ([Github](#))

Developed an AI app using CNN/VGG-16 for image features and LSTM/Transformer for captioning (BLEU: 0.80). Built a Streamlit web interface.

### ScribeGlobe (Blogging Site) | React.js, PostgreSQL, HONO, Cloudflare ([Github](#))

Built a responsive blogging platform with React.js and serverless backend on Cloudflare. Used PostgreSQL for data management.

### Project-TechUpdates | Python, TypeScript, JS ([Github](#))

Created a tool to scrape and categorize tech news, displaying ad-free updates on a minimal frontend.

## EXTRACURRICULAR ACTIVITIES

- **Event Leadership:** Organized "Blind Code," a sub-event of ASCENT 2022 (ISTE).
- **Competition Success:**
  - Won [Tech-A-Thon](#) at IIFL.
  - Runner-up at [Crack The Code](#) (Jai Hind, 2018).
  - Participated in [Feynwick](#) (KJSIEIT, 2021) and [Trident](#) (TSEC, 2019).
- **Courses:** 0-100 Cohort 2.0 (Harikart Singh), TensorFlow for deeplearning.