Gaurav Surtani

Software Developer (AI/ML) | Full-Stack Engineer | San Jose, CA |

+1 408-690-5274 | gaurav.surtani.tech@gmail.com | linkedin.com/in/gaurav-surtani | github.com/gauravsurtani

Core Skills & Technologies

- **AI/ML:** Python, TensorFlow, PyTorch, Keras, Scikit-learn, NLP (BERT, GPT, LangChain), Knowledge Graphs, Agentic AI, Computer Vision (YOLO v8), GenAI, LLMs, ChromaDB, RAG
- **Data:** Pandas, NumPy, SQL, MongoDB, Neo4j, BigQuery, GeoPandas, ETL, Data Viz (Plotly, D3.js, Folium), Statistical Analysis
- **Web:** JavaScript, TypeScript, React, Next.js, Angular (15/16), Node.js, Flask, FastAPI, HTML/CSS, Tailwind, REST, GraphQL, Microservices
- **Tools:** Docker, Kubernetes, AWS, Azure, Git, CI/CD, Jira, Wix CMS, Strapi CMS
- **Soft Skills:** Collaboration, Rapid Prototyping, Project Leadership, Agile, Problem-Solving, Communication

Professional Experience

Full Stack AI Engineer| Bexorg, Inc. | *New Haven, CT* | **May 2024 – May 2025**

- Led cross-functional team initiatives, accelerating deployment cycles by 85% by streamlining CI/CD processes.
- Architected an async collaboration system, reducing meeting dependencies by 60% and cutting decision time from 7 days to under 4 hours.
- Designed a real-time data viewer, resulting in 45% faster data retrieval and 60% improvement in scientific analysis.
- Integrated ML-powered time series analysis and QC, reducing data issues by 70% and turnaround time by 80%.
- Developed a Gemini API-driven chatbot trained on experiment data using ChromaDB to utilize vector embeddings for RAG based chatbot.
- Prototyped a GenAI query canvas that translates natural language to visualization pipelines, making analysis 75% faster.

Machine Learning Developer | NASA Ames Research Center | *Mountain View, CA* | **Dec 2023 – Mar 2024**

- Developed a coordinate transformation pipeline (Pyproj) for flight path visualization, enabling accurate geospatial mapping in web dashboards (Folium).
- Containerized legacy systems, reducing researcher setup time from days to minutes and improving reproducibility.
- Launched a real-time time-series viewer for flight path adherence, supporting quick anomaly detection.
- Streamlined validation for 1000+ flight paths, cutting manual review from 4 hours to 15 minutes.

Research Assistant | San Jose State University | *San Jose, CA* | **Jul 2023 – Feb 2024**

- Created a real-time sensor fusion pipeline (LiDAR & KITTI) to enhance object detection and 3D bounding box accuracy using YOLO v8.
- Created OHANA, a web app for hurricane relief, integrating a Hugging Face LLM trained on multilingual mental health data—including Hawaiian—for crisis response and support center mapping.
- Prototyped multilingual TTS (Coqui, Firefox Voice) to deliver accessible support in 25+ languages.

Senior Software Engineer | Ascendion | *Basking Ridge, NJ* | **Nov 2022 – Aug 2023**

- Led a team of 15+ engineers to deliver enterprise-grade projects on time and within budget.
- Drove migration from AngularJS/13 to 15/16, reducing bundle size by 35% and improving app performance.
- Built a cloud economics dashboard, enabling teams to monitor VM usage/cost and achieve 30% cloud savings.
- Created an AI-powered VS Code extension for code quality using ChatGPT API, reducing documentation effort by 70%.
- Improved performance, cutting page load time from 23s to 2.5s and API response time by 75%.

Software Engineer | Collabera Inc. | *Basking Ridge, NJ* | **Dec 2021 – Dec 2022**

- Crafted an SRE & DevOps assessment tool, reducing operational overhead by 20% through scoring calculations.
- Designed a W3C-compliant Angular component library (40+ components) for rapid UI development.
- Implemented Azure ML workflows into a web app, enabling real-time data synchronization and automated processing.

Software Developer | LIBS | *Cotonou, Benin* | **Jan 2021 – Nov 2021**

- Led a team of 10+ engineers to deliver a warehouse and sales management system deployed across Africa.
- Orchestrated Tally cloud integration with warehouse management system using AWS Lambda for serverless data processing and EC2 for hosting core services, reducing manual intervention by 65%.

Software Engineer | Accenture | *Pune, India* | **Oct 2018 – Dec 2020**

- Optimized ETL pipelines for Barclays data migration, ensuring high data integrity and reliability.
- Implemented comprehensive data validation and QA checks, reducing errors and rework.
- Built automated ETL monitoring and reporting tools, improving operational oversight.

Frontend Developer Intern | Enthralltech Pvt Ltd | *Pune, India* | **May 2018 – Oct 2018**

- Led frontend development for an EdTech platform using Angular and MongoDB.
- Researched on a NoSQL/MongoDB chatbot for automated student support through Fuzzy Search, enhancing user engagement.

Education

Degree	Institution	Location	Year
M.S., Artificial Intelligence	San Jose State University	San Jose, CA	Aug 2023 – May 2025
B.E., Information Technology	Pune University (SAE)	Pune, India	Aug 2014 – May 2018