SAURABH BHAUSAHEB ZINJAD

480-913-5544   linkedin.com/in/saurabhzinjad   saurabhzinjad@gmail.com   github.com/Ztrimus

Versatile engineer with 6 years of experience developing production-ready ML solutions and full stack applications. Managing end-to-end pipelines, from design to deployment and optimization in cloud environments. Passionate about advancing AI infrastructure, fine-tuning foundation models, and delivering scalable, cost-effective solutions globally.

SKILLS

Languages: Python, JavaScript/TypeScript, C#, C++, SQL, Java, Bash, HTML, CSS
ML/AI: PyTorch, TensorFlow, scikit-learn, LangChain, LangGraph, OpenCV, Pandas, MLflow, Databricks
GenAI & Systems: RAG, LLM evaluation, fine-tuning, prompt engineering, agentic AI design, MLOps, observability, Kafka
Cloud/DevOps: AWS (Bedrock, Lambda, S3, CloudWatch API Gateway, Cognito), GCP, Azure, Docker, Kubernetes, Git
Full-Stack: React, Node.js, Flask, FastAPI, .NET, PostgreSQL, MySQL, MongoDB, Terraform

WORK EXPERIENCE

AI/ML Engineer II — *University of Phoenix, AZ, USA*   Feb 2025 – Present
· Built GenAI + explainability tools (LangGraph + OpenAI) to process unstructured data and auto-explain complex academic/financial calcs, enhance advisor support capacity by $\sim$ 6x
· Integrated organizational knowledge base with LLMs, building pipelines to extract code logic into step-by-step explanations, accelerating onboarding and improving developer productivity

AI Research Intern — *University of Phoenix, AZ, USA*   May 2024 – Aug 2024
· LLM summary dashboard with chain-of-thought reasoning integrated to Salesforce via Kafka; advisor efficiency ↑ 40%
· Text-to-Visual API service converting text to interactive diagrams; integrated with Slack/Teams/CRMs; documentation speed↑

Senior Machine Learning Engineer — *Tiger Analytics, Bangalore, India*   June 2022 – July 2023
· Trained & tuned predictive ML models and scalable Data pipelines (Python, PySpark, Databricks) achieving 18% client revenue
· Architected a multi-cloud model monitoring platform for lifecycle tracking/alerts; supported 10+ client adoptions

Software Engineer — *Winjit Technologies, Nashik, India*   January 2020 – June 2022
· Engineered 10+ RESTful APIs for real-time AI data processing; built 30+ low latency responsive UI React features
· Optimized large-scale MySQL, PostgreSQL databases, accelerating data access for AI model training and real-time inference
· Innovated data-driven UI framework with dynamic form reconfiguration, customizable CSS, reduced development time by 8x

Deep Learning Engineer — *Automation Teknix, Pune, India*   September 2019 – January 2020
· Developed lightweight object recognition engine with SSD & MobileNetV2, reducing computational cost by 30% for mobile use
· Streamlined CNN model using TensorFlow, OpenCV to improve image processing, reducing survey error in production by 22%

EDUCATION

Arizona State University, Tempe, USA   August 2023 – May 2025
Masters of Science in Computer Science - **Thesis**
Pune Institute of Computer Technology, Pune, India   July 2015 – June 2019
Bachelor of Engineering in Electronics and Telecommunications

PUBLICATION *first authored*

Can Typos Cause Harm? The Impact of Imperfect Input on LLM Safety   Sep 2024 – Aug 2025
· Systematic study of LLM safety under prompt perturbations; introduced safety flip rate and evaluated 5 models across 14 perturbation types & 11 harm categories—minor typos/paraphrases flip safety in 30% of cases
· Released open-source framework & augmentation pipeline for perturbation-aware safety evaluation and robust alignment tools

LLM-facilitated Pipeline for Personalized Resume Generation   Aug 2023 – July 2024
· Authored & published research paper at ACM's SIGIR conference on ML pipeline optimized for user personalization using LLM
· Formulated novel user personalization metrics, setting performance benchmarks for AI-enhanced resume evaluation
· Released open-source Python library integrating information retrieval leveraging web scraping, used by 1000+ users

PROJECTS

Multimodal AI System for Media Search — AWS + CLIP + vector DB   October 2024 – October 2024
· Won Big Data Prize for content-based search across video/audio/images; stored in Postgres delivered 92% faster media

Forest fire Detection using IoT Sensor Data   September 2021 – January 2022
· Implemented a TabNet classifier model with 98.7% accuracy on edge devices using AWS, TinyML, Docker, Redis, and Celery
· Utilized SMOTE and hyperparameter tuning to balance data and optimize recall, reducing errors detecting fire in 7 mins