SANKET SHIGAONKAR

📞 (716)704-6664 ✉ sankets0930@gmail.com in sanketshigaonkar ⭕ sank3t9

EDUCATION

University at Buffalo, SUNY   **Buffalo, New York**
*Master of Science, **Computer Science** (AI/ML Specialization)*   *Aug 2024 – Dec 2025*

University of Mumbai   **Mumbai, India**
*Bachelor of Engineering, **Computer Science***   *Jun 2019 –May 2023*

TECHNICAL SKILLS

Languages: Python, SQL, C++
AI/ML: PyTorch, LangChain, Hugging Face, LLMs (Llama, Qwen), YOLOv8, OpenCV, FAISS, QLoRA, RAG Pipelines
Web & Backend: FastAPI, REST APIs, Streamlit
Cloud & DevOps: AWS (Lambda, DynamoDB), Docker, Apache Kafka, Redis, CI/CD, Git

EXPERIENCE

Dentite   **Jan 2025 – Jun 2025**
Software Engineer Intern   *Buffalo, NY*

- Engineered a synthetic data generation pipeline creating thousands of insurance card images via SVG/PNG templates with randomized metadata and realistic distortions to simulate real-world capture conditions.
- Optimized extraction accuracy by benchmarking and prompt-engineering multiple LLMs, validating performance against a custom-generated dataset across 6 major US insurance providers.
- Reduced model latency by 30% by architecting a secure REST API inference pipeline and optimizing data ingestion workflows while ensuring HIPAA compliance.
- Architected an automated claim reclamation feature that utilized extracted data to identify missed revenue, contributing to a 15% projected uplift for dental practices.

NetCore Solutions Pvt Ltd.   **Jan 2023 – Jul 2024**
Software Engineer   *Mumbai, India*

- Drove a 10% efficiency gain in collections by engineering a multilingual intent classification pipeline (5 Indic languages) using OpenAI Whisper and Llama, optimizing repayment forecasting.
- Automated credit risk assessment by configuring Experian PowerCurve strategies, eliminating manual underwriting for standard applications and ensuring real-time compliance.
- Streamlined customer onboarding by implementing Business Rules Engine (BRE) workflows, reducing processing overhead and accelerating loan disbursement cycles.
- Architected an internal Loan Management System (LMS), validating credit-bureau APIs with Postman/SoapUI for 100% integrity.

PROJECTS

Accio AI - An Aligned Conversational RAG Agent | *Qwen-4B, Hugging Face, FAISS, LangChain*   ⭕

- Engineered a Generative AI agent by fine-tuning Qwen-4B on 25,200+ samples via QLoRA-SFT for domain-specific reasoning.
- Architected a sub-100ms RAG pipeline with FAISS and Flash Attention 2, optimizing context retrieval to reduce hallucinations.
- Achieved 77% accuracy and 56% loss reduction via RLHF alignment, deploying the 4-bit quantized model on Streamlit.

Tickr - Live Crypto Price Monitor | *FastAPI, Apache Kafka, Redis, AWS Lambda, Docker, DynamoDB*

- Architected a real-time streaming platform with FastAPI and Kafka, ingesting high-frequency market data via Docker.
- Engineered a sub-millisecond alert engine using Redis for O(1) lookups, matching user thresholds against live streams.
- Deployed an event-driven AWS Lambda backend with DynamoDB, ensuring auto-scaling reliability during market volatility.

DeepDish - Deep Learning Calorie Estimator | *Python, PyTorch, YOLOv8, OpenCV, Streamlit, Pandas*   ⭕

- Engineered a multi-view food recognition pipeline using YOLOv8 on ECUST Dataset, achieving 0.975 mAP on 19 classes.
- Architected a volumetric engine using GrabCut and coin calibration to solve 2D-to-3D depth ambiguity.
- Deployed a real-time Streamlit calorie tracker, achieving <10% error on regular food shapes.

ACHIEVEMENTS

Patent Filed: Co-invented the "Smart YOLO Intersection Traffic System," a real-time congestion controller utilizing YOLOv8 for detection and Residual LSTMs for predictive signal timing.