

Sahil Shah

shahsahil8141@gmail.com | +91-8141312147 | [LinkedIn](#) | [GitHub](#)

SKILLS

- **Programming and Databases:** Python, SAS, Bash, SQL (MySQL, PostgreSQL), NoSQL (MongoDB), OOP
- **Machine Learning and Statistics:** Data Wrangling, Data Visualization, Feature Engineering, Model Development and Evaluation, Hypothesis Testing, A/B Testing, Time Series Forecasting, Statistical Analysis, Microsoft Excel
- **Generative AI and LLMs:** Retrieval-Augmented Generation (RAG), LangChain, Vector Databases (Pinecone, Chroma), OpenAI API, LLM Fine-Tuning (PEFT, LoRA, QLoRA), Prompt Engineering
- **Data Science Libraries:** NumPy, Pandas, Scikit-Learn, TensorFlow, Keras, NLTK, Matplotlib, Seaborn, Hugging Face Transformers
- **Cloud, DevOps and MLOps:** AWS (EC2, Lambda, S3, API Gateway, CodeBuild, CodePipeline, ECR, EMR, Batch, Cognito, CloudWatch, SES, SNS), Heroku, Docker, Git
- **Web Technologies and API Development:** Flask, FastAPI, HTML, CSS, AJAX

PROFESSIONAL EXPERIENCE

Tiger Analytics

Hyderabad, India

Senior Analyst – Data Science

April 2023 – Present

- Conceptualized and developed a fraud benchmarking module using transaction data for a leading US credit bureau; **transformed it into a standalone product**, driven by strong client acclaim and multi-stakeholder support.
- Executed a proactive fraud validation initiative using AWS services to enable pre-sales without client-provided data; accelerated sales cycles by delivering **900+ successful validations** with portfolio-level fraud insights.
- Developed custom first-party and third-party **fraud detection models** by retraining baseline models on client-specific data; increased **bad detection rates by up to 20%**, while collaborating with internal stakeholders to deliver client-ready model insights.
- Assessed fraud scoring models for First-party, Third-party, and Bust-out fraud types using SAS across diverse portfolios; enhanced fraud loss prevention — ranging from **\$1M–\$70M** — and ROI through precise model performance analysis.
- Built high-quality fraud datasets through rigorous data cleaning, feature engineering, and fraud tagging in Python and SAS; improved detection accuracy by continuously refining model features with domain-specific insights.
- Designed and built automated reporting solutions from scratch using Python and SAS to optimize validation processes and standardize EDA workflows; reduced manual effort by **89%**.

Softvan Pvt. Ltd

Ahmedabad, India

Senior Software Engineer – Data Science Focus

January 2022 – April 2023

- Spearheaded the development of a SaaS-based Quantitative and Qualitative **Data Analysis platform**, managing all stages from infrastructure design to deployment; optimized product and operational performance through **data-driven insights for clients including Amazon, Twitch, and Spotify**.
- Engineered an inference module utilizing **hypothesis testing, p-values, and advanced statistical methods** for client-specific reporting; reduced customer support workload by **80%**.
- Developed an NLP-based **semantic search** feature using **Sentence Transformers** for product review transcripts; improved video recommendations and user engagement accuracy.
- Implemented Google SSO authentication with AWS Cognito and built **70+ APIs** using Python, FastAPI, and MongoDB; enabled analytics delivery and data-driven platform capabilities.
- Designed and built a deep learning-based **video highlight generation** module with customizable subtitle and summary generation and automated server orchestration; reduced manual processing and enhanced content delivery efficiency.
- Implemented a **CI/CD pipeline** for production deployment and configured a custom Highcharts container for batch chart exports; improved deployment speed and reporting automation.

Softvan Pvt. Ltd

Ahmedabad, India

Software Engineer – Data Science Focus

July 2019 – December 2021

- Developed a regression model using social media and behavioral data to match student fraternities with brand campaigns; achieved **84% R² and increased revenue by 67%**.
- Implemented Instagram authentication and created **120+ APIs** using Python and AWS; enabled real-time data integration and scalable analytics across the platform.
- Designed workflows and integrated services like Qualtrics, TypeForm, Rev.AI, and Parallel Dots; enhanced platform functionality and facilitated actionable data insights.
- Built a data-driven social media platform for student fraternities at Stanford, UChicago, and others to promote global brands for rewards; oversaw the entire development lifecycle, resulting in **100+ campaigns and 6K+ student** registrations.

Softvan Pvt. Ltd

Ahmedabad, India

Software Engineer Intern

December 2018 – June 2019

- Developed backend APIs using Flask for a platform connecting entrepreneurs with resources and talent; enabled end-to-end startup and product development.
- Worked with a cross-functional team in an agile setup to gather requirements and implement features; ensured on-time and goal-aligned delivery.

EDUCATION

Gujarat Technological University, GIT College
Bachelor of Engineering in Information Technology

- Cumulative GPA: 8.69/10

Gandhinagar, India
July 2015 – May 2019

PROJECTS

Steel-Defect-Detection – [[GitHub](#)]

- Developed an automated defect detection system for steel sheets to boost industrial production efficiency and quality; achieved **94% F1 score**, significantly reducing manual inspections.
- Applied **transfer learning** with InceptionV4 to build an end-to-end web app with FastAPI backend, Git-LFS model management, and Docker; deployed on Heroku for scalable fault detection and reporting.

Ventilator-Pressure-Prediction – [[GitHub](#)]

- Automated ventilator control to support clinicians during pandemic-like scenarios; improved system precision in critical care with an **85% R² score**.
- Trained a Random Forest model and delivered a full-stack web app with FastAPI backend, Git-LFS, and Docker deployment; enhanced clinical data throughput and operational responsiveness.

CERTIFICATIONS

- Machine learning Certification (**Stanford University**)
- Deep learning Specialization Certification (DeepLearning.AI)
- Tools for Data Science Certification (**IBM**)
- Python Programming Masterclass Certification (Udemy)