

NEEL SACHIN SHAH

shahneelsachin@gmail.com | +1 930 215 9133 | LinkedIn: [neel-shah1901](#) | GitHub: [NeelShah09](#)

SUMMARY

Graduate Data Scientist with 2+ years of analytics experience and internships in *AI & Computer Vision*. Skilled in designing and deploying intelligent systems using LLMs, LangGraph, Hugging Face and Vector Stores. Experienced in multi-agent architectures, predictive analytics, and cloud-based automation that drive measurable business impact.

EDUCATION

Indiana University Bloomington, Luddy School of Informatics, Computing and Engineering **Indiana, USA**
Master of Science in Data Science (3.95 / 4.0) August 2024 - May 2026

University of Mumbai, Sardar Patel College of Engineering **Mumbai, India**
Bachelor of Technology (9.26 / 10.0) August 2018 - May 2022

TECHNICAL SKILLS

- Languages & Libraries:** Python, C++, R, NumPy, Scikit-learn, Tensorflow, PyTorch, Keras, HuggingFace Transformers
- LLM & GenAI:** LangChain, LangGraph, LlamaIndex, Groq API, Qdrant
- Machine Learning & Deep Learning:** Predictive Modelling, Computer Vision, Transfer Learning, NLP
- Data Engineering & MLOps:** Apache Airflow, Spark, AWS SageMaker, Glue, Lambda, Docker, FastAPI, CI/CD
- Database:** SQL(MySQL, PostgreSQL, Oracle SQL, Snowflake), NoSQL(MongoDB, Neo4j), Vector Stores(Qdrant)
- Visualization:** Matplotlib, Seaborn, Plotly, PowerBI, Quicksight, OpenCV
- Tools & Version Control:** Git, Github, Bitbucket, Vi/Vim

PROFESSIONAL EXPERIENCE

SproutsAI Inc. **California, USA**
AI & ML Intern **July 2025 – August 2025**

- Reduced time-to-insight by 80% by engineering a multi-agent AI platform translating natural language to MongoDB queries with automated Plotly visualizations enabling data-driven storytelling and actionable insights across 128+ collections.
- Improved query accuracy by 50% via a knowledge-graph based context retriever using **Neo4j** and **Qdrant** to map relationships across collections, combined with vector embeddings and community detection.
- Empowered non-technical users to perform complex data analysis through self-service analytics, significantly reducing reliance on data engineering teams.

ICICI Bank Limited **India**
Deputy Manager II (Focus: Data Engineering) **June 2022 – July 2024**

- Spearheaded the development and productionized an ETL pipelines including predictive load forecasting on 140M+ transactions using Airflow, SQL, and Excel, improving transaction processing efficiency by 5%.
- Saved INR 1M by improving core transaction processing logic recognized by analyzing 1TB+ of transaction log using bash scripting and command line.
- Led a project involving collaboration with 4+ cross-functional teams, driving smooth integration and on-time delivery of business requirements by facilitating communication, resolving roadblocks, and ensuring alignment across all teams.
- Directed the optimization of security audit process for the core banking application. Automated quarterly reports generation using Linux and Oracle SQL, that reduced manual audit time by 50% and confirmed compliance with RBI standards.

Wobot Intelligence Pvt. Ltd. (wobot.ai) **Remote, India**
Computer Vision Intern **September 2021 – January 2022**

- Trained and deployed a VGG19 model using TensorFlow, Keras, and Docker for safety gear detection, achieving 93% accuracy. Improved input data quality by 3% by implementing an OpenCV algorithm to filter anomalous CCTV frames.
- Converted models to TensorRT, reducing inference latency by 35% and enabling real-time video analytics.
- Developed a REST API in Go language to create RTSP streams for 100+ videos, enabling seamless video access and improving data streaming capabilities.

PROJECTS

AWS Predictive Maintenance Platform | *Python, SageMaker, Glue, Lambda, QuickSight*

- Architected an automated, end-to-end ML pipeline on AWS that ingests IoT sensor data via Kinesis, processes it with Glue ETL jobs for feature engineering, and stores it in an S3 data lake, enabling scalable predictive analytics.
- Productionized an XGBoost classifier (94% F1-score) to forecast equipment failure 72hr in advance, deploying via a SageMaker endpoint and automating retraining with MLOps pipelines.
- Operationalized predictions by implementing serverless Lambda/SNS alerts to trigger proactive maintenance workflows and building a QuickSight dashboard to demonstrate potential cost savings.

Smart Hospital Multi-Agent System (SHMAS) | *(2nd Runner-Up), Groq API, PostgreSQL, LangGraph, Streamlit*

- Led development of an AI-driven hospital management system using 5 collaborative agents (Mental Health Analyzer, Emergency Triage, Bed Allocator, Doctor Scheduler, Conflict Resolver) to automate patient workflows.
- Integrated Groq-powered deepseek-r1-distill-llama-70b model for real-time mood analysis and triage prioritization, reducing theoretical patient wait times by 40% through dynamic resource allocation.
- Recognized for innovation in Luddy Hackathon 2025 with 2nd Runner-up position among 50+ teams.