Soham Patel

Irvine, California | (929) 642-7699 | sohamhp@uci.edu | LinkedIn | GitHub | Portfolio

Education

University of California, Irvine — Irvine, CA
Master of Data Science — *Sep 2023 – Dec 2024*
Artificial Intelligence, Machine Learning, Statistics, Database Management, Bayesian Modeling, Big Data

SRM Institute of Science and Technology — Chennai, India
B.Tech in Computer Science Engineering — *Jun 2019 – May 2023*
Machine Learning, Statistics, DBMS, Artificial Intelligence, Data Structures, Operating Systems

Technical Skills

Programming: Python, C, R, C++, MATLAB, SQL
ML & DL: scikit-learn, Keras, PyTorch, TensorFlow, Neural Networks, Transformers, LLMs, XGBoost, NLP
Analytics/Visualization: Excel, Tableau, PowerBI, ArcGIS, Pandas, NumPy, Matplotlib, Seaborn
Data Engineering: ETL Pipelines, MLOps, MySQL, PostgreSQL, Apache Spark, Neo4j, MongoDB, Cassandra, Flink
Statistical Analysis: Regression, Bayesian Inference, A/B Testing, Time Series, Optimization
Cloud & Tools: SASS, SPSS, AWS, Azure, GCP, Airflow, OpenCV, Docker, Jenkins, GitHub, CI/CD, Kubernetes, Langchain

Experience

Machine Learning Developer — Dec 2024 – Present
The Blue Box Biomedical Solutions — *Remote*
- Developing a breast cancer detection system using a **CNN-LSTM model** in PyTorch, trained on biomedical sensor signals.
- Generated synthetic signals using **GAN framework** to augment data and improve model against class imbalance.
- Boosted model, achieving **ROC-AUC of 0.93** and **F1-score of 0.89** through feature engineering and architecture tuning.
- Containerized and deployed the model with **TorchServe**, building **REST APIs** for integration into diagnostic platforms.
- Integrated model into a **Flask dashboard**, enabling clinician interpretation and decisions in under 10 seconds.

Machine Learning Intern — Sep 2024 – Dec 2024
LiveGood Inc — *Irvine, CA*
- Applied **GLMs, PCA**, and **Cox models** using **Python (statsmodels, lifelines)** to model longevity predictors.
- Performed **cohort analysis** using **SQL, Python, and R** to identify **centenarian vs. non-centenarian** differences.
- Deployed statistical models on **AWS EC2** and integrated results into **Power BI dashboards** for marketing teams.
- Automated preprocessing workflows using **Airflow** and **Pandas**, **reducing analyst workload by 60%.**

Data Science Research Intern — Jun 2024 – Sep 2024
UCI Health — *Orange, CA*
- Built a robust **ETL pipeline on AWS** using Python, **S3**, and **Lambda**, enabling real-time medical record ingestion.
- Improved diagnostic model accuracy by **40%** using **TensorFlow and Keras** with optimized CNN layers.
- Developed a **GPT-3.5-based RAG chatbot** with Pinecone to assist doctors with instant access to burn injury literature.
- **Ran A/B tests** on model variants, **boosting F1-score by 10%** and validating robustness for clinical use.

Machine Learning Engineer — Dec 2021 – Sept 2023
Strategic Alliance — *India*
- Deployed a YOLO-based real-time passenger detection system using **TensorFlow**, increasing transport revenue by **30%**.
- Built an edge-deployed **social distancing tracker** (95% accuracy) using **OpenCV** and live camera feeds.
- Used **Docker and Kubernetes** for multi regional deployment, ensuring uptime, fault-tolerance, and horizontal scalability.
- Optimized large-scale ingestion pipelines using **Apache Spark**, reducing data processing latency by **20%**.

Projects

Healthcare Demand Forecasting — Jan 2024 – Jun 2024
- Developed a forecasting pipeline using **ARIMA** in Python to predict hospital bed demand, achieving **95% accuracy**.
- Integrated forecasts with **AWS dashboards** and delivered real-time insights for hospital admins to manage peak load.
- Improved operational efficiency by **20%** through optimized shift planning, reducing idle hours and patient wait time.
- Performed data cleaning and preprocessing using **Pandas, NumPy**, and automated updates with **Airflow**.

YOGDAAN: Emergency ML Dispatch System — Aug 2022 – Jul 2023
- Built a WhatsApp-integrated dispatch system using **Twilio API**, **YOLOv5**, and Python to detect emergencies.
- Deployed backend on **AWS EC2** with **S3-based image storage** and automated processing via **Lambda** functions.
- Automated alerts to authorities and NGOs using **reverse geolocation** logic and **Twilio's voice/SMS stack**.
- **Reduced manual filtering by 50%** and **improved emergency response time by 20%** even in underserved zones.