SANSHRIT SINGHAI

Data Scientist | ML Engineer | Optimization & GenAI
Atlanta, GA |+1(404) 903-4966 | singhai.sanshrit@live.com | LinkedIn | GitHub | Portfolio | H1B (Transfer Eligible)

SUMMARY

Data Scientist (3+ years, Georgia Tech MS) delivering ML, optimization, and GenAI systems at production scale. Strong in PyTorch, TensorFlow, OR-Tools, LangChain, BigQuery, and cloud-scale pipelines. Experienced in experimentation, scalable system design, and partnering with engineering/product teams to drive measurable business impact.

TECHNICAL SKILLS

- ML & GenAI: Python, PyTorch, TensorFlow, Scikit-learn, LLMs, LangChain, LangGraph, OpenCV, FAISS
- Data Engineering & Cloud: PySpark, Kafka, Azure (Event Hubs, Data Lake, DevOps), AWS (S3, General Cloud), GCP Big Query, Redis, Docker, Git, FastAPI
- Optimization, Graph & Analytics: OR-Tools, Gurobi, TigerGraph, GSQL, SQL, Tableau, Power BI, MATLAB

EXPERIENCE

ENRU – Logistics and Postal Optimization **Jun 2023 - Present**
Data Scientist *Atlanta, GA, US*

- Developed ML & DL models (TensorFlow/PyTorch, ensembles) for multi-label classification and accessorial cost prediction, improving cost estimation accuracy for new locations.
- Built a **RAG-based GenAI chatbot** (LangChain + FAISS) to enable natural-language planning queries, reducing analyst lookup time by **>60%**.
- Engineered a **profit-maximizing integer-programming optimizer** (OR-Tools, Gurobi) that improved routing decisions and increased revenue per truck by **25%**, while achieving a **10× speed improvement** over legacy heuristics.
- Designed a **real-time Azure EventHub pipeline** that streams 3PL freight data and supports downstream ML systems.
- Built and optimized **GSQL graph queries** for journey verification and truck assignment, enabling robust logistics decision-making.

Georgia Institute of Technology **Aug 2022 - May 2023**
Graduate Research Assistant (National Science Foundation Funded) *Atlanta, GA, US*

- Proposed and refined an LSTM-RNN framework for multivariate time series forecasting to predict mechanical properties in the excavating domain, achieving a **30% result improvement.**
- Developed an ANN surrogate for FEM simulations, reducing computation time **20x.**

Itasca Consulting Group **May 2022 - Aug 2022**
Machine Learning Software Intern *Minneapolis, MN, US*

- Developed a deep learning tool using neural networks to predict velocity fields, bearing capacity, and failure depths for 230K cases, achieving an **R2 score of 0.91**.
- Built an advanced mesh generation tool for Rhino Griddle based on computational geometry for efficient meshing.

PROJECTS

Document Chatbot with RAG and Llama

- Built a document-QA chatbot with RAG for semantic retrieval using FAISS, LangChain, and Ollama embeddings.
- Designed a Streamlit front-end with PDF parsing, semantic search, and context-aware responses.

Credit Card Fraud Detection (ML Pipeline)

- Trained XGBoost and Random Forest models on imbalanced transaction data for fraud classification.
- Engineered fraud-specific features (e.g., reversal detection, multi-swipe patterns) and handled class imbalance with SMOTE.

Unveiling Patterns and Insights in Chicago's Criminal Activity through Advanced Data Analysis

- Analyzed **7M+** crime records using PySpark on AWS Data Lake; performed large-scale ETL and feature engineering.
- Built and deployed an XGBoost model as a Flask API to predict crime risk by ZIP code.

EDUCATION

Georgia Institute of Technology **May 2023**
MS, Computational Science and Engineering (Machine Learning) - **GPA: 3.9/4.0** *Atlanta, GA, US*

- **Coursework:** CSE Algorithm, Computer Vision, Machine Learning, Data & Visual Analytics, AI

Shiv Nadar University **Jun 2020**
B.Tech, Civil Engineering (Computational Mechanics) – **GPA: 8.6/10** *Greater Noida,UP, India*

PUBLICATIONS

- **Singhai, S.**, C., Arson, Deep Learning Models for Subterranean Navigation and Soil Characterization, EMI 2023, Georgia Tech, Atlanta, GA
- **Singhai, S.**, Vikash, G., Uncertainties in Vertical Stress Distribution in Spatially Varying Random Elastic Half Space, International Journal of Geomechanics, ASCE