

Suraj Jaiswal

Data Analyst & NLP Engineer
linkedin.com/in/surajjaiswal007

Email : jaiswalsuraj487@gmail.com

Mobile : +91 7878189018

Portfolio: jaiswalsuraj487.github.io

SUMMARY

Data Analyst at **Tiger Analytics** with **3+ years of experience** delivering impactful **ML** solutions. Alongside applications, I actively contribute to **ML research and publications** in top conferences (**NeurIPS** and **ACM**). Beyond work, I pursue a healthy lifestyle through yoga, surfing, traveling to new places, and badminton.

EDUCATION

• Indian Institute of Technology Gandhinagar (IIT)	Masters in Computer Science Engg	2022 – 2024
• Gujarat Technological University	Bachelors in Computer Science Engg	2018 – 2022

EXPERIENCE

• Data Analyst — Tiger Analytics	Dec 2024 – Present
○ Architected a Multi-agent AI interview orchestrator using LangChain + LangGraph to execute routing, phase-specific question generation, and response evaluation dynamically to the job description for the role	
○ Built an async FastAPI + WebSocket backend supporting real-time multi-phase interviews with guardrail	
○ Designed and deployed GenAI Bill-of-Materials automation for lighting parts pipeline using AWS Bedrock, MLflow, and Docker, reducing manual generation time by 80% through structured extraction	
○ Built LLM-based code evaluator using OpenAI API, prompt engineering, and multi-threaded Streamlit UI for parallel execution, improving throughput by 3× and reducing manual review time by 90%	
○ Engineered an end-to-end Discrete choice modeling framework using xLogit and Pandas to analyze 1.8K+ traveler preferences via Multinomial Logit (MNL) models and behavioral analytics	
○ Created interpretable marginals, utility, and demand estimation workflows to measure choice sensitivities and substitution probabilities under different scenario conditions on wide-to-long data	
• NLP Engineer — NeuroReef Labs	Oct 2023 – Oct 2024
○ Developed an AI copilot for medical note-taking using OpenAI LLMs fine-tuned on clinical guidelines	
○ Built RAG system to autogenerate ICD-10, CPT, SNOMED, and HCC codes for medical billing of patient visits	
○ Implemented a Medical chatbot using HyDE + RAG over Athena EHR data to deliver insights	
○ Applied few-shot prompt tuning for medical guideline-consistent and domain-aligned LLM outputs	

RESEARCH PUBLICATIONS

- **Space to Policy: Scalable Brick Kiln Detection and Automatic Compliance Monitoring with Geospatial Data** (link) Accepted at **ACM Journal on Computing and Sustainable Societies**
- **Eye in Sky: Detection and Compliance Monitoring of Brick Kilns using Satellite Imagery** (link) Accepted at **ACM Compass on Computing and Sustainable Societies poster track**
- **Towards Scalable Identification of Brick Kilns from Satellite Imagery with Active Learning** (link) Accepted at **NeurIPS Active Learning in the Real World**
Developed a scalable **YOLO-OBB-based deep learning pipeline** to detect and classify **30,000+ brick kilns** across 5 Indian states using satellite imagery to achieve **82% precision** over 520,000 km² of coverage

SKILLS

- **Programming Languages:** Python, SQL, Java, C
- **Libraries:** LangChain, Langgraph, FastAPI, YOLO, PyTorch, Tensorflow, Scikit, NumPy, Pandas, Matplotlib
- **Tools & Cloud:** AWS (S3, EC2, Bedrock, ECR), Boto3, Docker, Hugging Face, Streamlit, Git, VS Code, Excel

ACHIEVEMENTS

- Won "AI Engine for Google Drive" Hackathon, ThirdAI Corp
- GATE Computer Science – Secured **97th percentile** in the Graduate Aptitude Test in Engineering