Suraj Jaiswal

ML Engineer, Neuroreef Labs
Masters in Computer Science and Engineering, IIT Gandhinagar

jaiswalsuraj@iitgn.ac.in
+91 7878189018, **LinkedIn**
jaiswalsuraj487.github.io

Education

Degree	Institution	CPI	Year
MTech (CSE)	Indian Institute of Technology Gandhinagar (IIT Gn)	8.5	2022 – 2024
BTech (CSE)	G.H. Patel College of Engineering and Technology, GTU	8.8	2018 – 2022

Experience

- **Machine Learning Engineer, NeuroReef Labs,** Banglore · Full-time **[Jun 2024 - Present]**
  - Developed product CareCortex (link: carecortex.ai): AI copilot for medical notetaking of Doctor-Patient conversation
  - **Fine-tuned large language models** (LLMs): on medical guidelines and medical codes to generate outputs for healthcare
  - **Medical Code Automation**: Developed a system that captures Doctor-Patient conversations to generate accurate ICD-10, CPT, SNOMED, and HCC codes using embedding-based similarity search and retrieval-augmented generation (RAG)
  - Interactive **Medical Chatbot**: Leveraged HyDE + RAG for real-time insights on patient chart details
  - Prompt Tuning: Aligned LLM outputs with medical guidelines using few-shot tuning for clinically accurate responses
  - Wrote SQL queries for PostgreSQL and managed security tasks using AWS services(S3, ECR, EC2)
  - Stack: Langchain, RAG, FastAPI, Docker, AWS, LLM finetuning, Asyncio
- **Machine Learning Intern, NeuroReef Labs** · Remote **[Dec 2023 - May 2024]**
  - Created, coded, and optimized **backend ML pipeline** for product **Medaura** (link:medaura.ai): Medical Q&A
  - Utilized HyDE + RAG using PubMed database to get answers using medical data
- **Teaching Assistant**, IIT Gandhinagar **[Jun 2022 - Jun 2024]**
  - **Natural Language Processing (**Prof. Mayank Singh**)** and **Machine Learning** (Prof. Nipun Batra): Assisted the professor in evaluating papers, assignments, viva, and quizzes
  - **Probability, Statistics, and Data Visualization** (Prof. Shanmuganathan) and **Computing** (Prof. Nipun Batra): Conducted hands-on lab sessions to instruct over 20+ students on probability distributions and fundamental ML concepts. Developed programming questions for over 300 students and provided Python tutorials to 20+ students

Research Publications

- **Eye in the Sky Detection and Compliance Monitoring of Brick** **[Jan 2024 - Jun 2024]**
  - Paper **accepted** in **ACM Compass** (Conference on Computing and Sustainable Societies) 2024
  - **Live demo:** brick-kilns-detector.streamlit.app of **Brick kiln detector application** on Streamlit enabling real-time brick kiln detection and compliance analysis for government and regulatory use
  - Developed a scalable deep learning model using **YOLOv8** for satellite-based detection of brick kilns **identifying 19,579 kilns** across nine states in the Indo-Gangetic plain
  - **Automated compliance monitoring of environmental policies** impacting public health, assessing the proximity of kilns to human habitats, rivers, and hospitals utilizing **Google Maps Static API and Sentinel API** for satellite data
- **Towards Scalable Identification of Brick Kilns from Satellite Imagery with Active Learning** **[July 2023 - Dec 2023]**
  GitHub repo**:** github.com/jaiswalsuraj487/brick-kiln-repo
  - **Paper** accepted and in nomination for **best paper award** in **NeurIPS 2023** Workshop on ReALML (Active learning and Machine learning in the real world)
  - Developed an efficient method for detecting brick kilns in satellite images using active learning techniques, **achieving 97% of oracle F1 score (0.976)** with a **70% reduction in manual annotation requirements**
  - Implemented in **Docker** container to ensure reproducibility, streamlined development and to use GPU on sever
  - **Identified over 700 new brick kilns** in the Indo-Gangetic region, showcasing the potential for global application in emissions monitoring and policy regulation

Achievements

- **1st Prize among 70+ teams - Third AI India Hackathon. Worked on *Neural dB* engine to search queries on *Google Drive* for** text and audio data
- Achieved all India 3000 rank in the Graduate Aptitude Test in Engineering in 2021

Projects

- **Meta-Learning: Hyper-Networks and Neural** **[Nov 2023]**
  Blog: jaiswalsuraj487.github.io/publications_and_projects/data/Hypernet_neural_process.html

  Meta-learning to learn task-specific networks to reconstruct whole given few context points of celebrity face image
- **Image to image for Climate Modelling using Auto-Encoder** **[Oct 2023]**
  Blog: jaiswalsuraj487.github.io/publications_and_projects/data/Autoencoder.html
  - Implemented Convolutional and UNet Auto-Encoder for multichannel input and output to predict pollution level in Delhi
- **The Third AI Engine Hackathon for Google Drive, ThirdAI Corp** **[Aug 2023]**
  GitHub repo: github.com/jaiswalsuraj487/TEGD_thirdai_hackathon
  - **Developed a localized drive search engine** for **retrieving** confidential file information within a person's Google Drive
  - Leveraged Third AI's extremely efficient NLP-based **NeuralDB** architecture that significantly **enhances user accessibility**

- **Enhancing Images with GAN-based Super Resolution** [May 2023]
  Blog link: jaiswalsuraj487.github.io/blogs/blogsData/Image_super_resolution.html
  - Implemented generative adversarial network to improve image quality using CNN with residual connections for generator and discriminator networks
- **Cryptocurrency Analysis & Trading Bot** [Jan 2022 – April 2022]
  GitHub repo: github.com/jaiswalsuraj487/Reddit-Cryptocurrency-Trading-Bot
  - Developed an AI **bot to fetch posts from subreddit** through the *Praw* library & the Reddit API
  - Implemented sentiment analysis on this collected data
  - **Enabled the bot to trade** specific **cryptocurrencies** using Binance API **based on technical indicators**, primarily using the RSI from the technical analysis library, and **integrated sentiment analysis into its trading decisions**.

Technical Skills

- **Languages:** Python, C, MySQL, Latex
- **Technologies:** Langchain, FastAPI, Tensorflow, Pytorch, JAX, Raytune, Sciki-learn, Numpy, Pandas, Matplotlib
- **Tools:** AWS, Hugging Face, Streamlit, Docker, Git, Visual Studio code, Excel, Notion, Jira